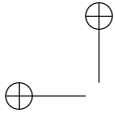
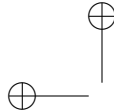


**FILTRAGGIO STATISTICO
(WIENER, LEVINSON,
KALMAN)
E APPLICAZIONI**

Giorgio Picci†



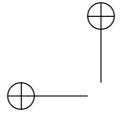
‡ Dipartimento di Ingegneria dell' Informazione, Università' di Padova, 35131
Padova, Italy.



Indice

PREFAZIONE	vii
PREFAZIONE	vii
Contenuto del Libro	viii
Capitolo 1. MODELLI DI SEGNALI ALEATORI	1
1. Segnali aleatori e processi stocastici	1
2. Modelli stocastici e identificazione	7
Capitolo 2. IL PROBLEMA DELLA STIMA	11
1. Introduzione alla Stima Bayesiana	11
2. Stimatori Lineari a Minima Varianza d'Errore	20
3. Esempi	28
4. Innovazioni e fattorizzazione di Cholesky	33
5. Componenti Principali e sviluppo di Karhunen-Loève	45
6. Stima di Direzioni	51
Esercizi	59
Capitolo 3. STIMATORI LINEARI A MINIMA VARIANZA PER PROCESSI STOCASTICI	61
1. Introduzione	61
2. Predizione, Filtraggio, Interpolazione	62
3. Euristiche del filtro di Wiener	67
4. Processi Ortonormalizzabili e Fattorizzazione Spettrale	72
Esercizi	88
Capitolo 4. FATTORIZZAZIONE SPETTRALE RAZIONALE E MODELLI ARMA	91
1. Introduzione	91
2. Fattorizzazione spettrale razionale: Il caso scalare	92
3. Fattorizzazione Spettrale di Matrici Razionali e Modelli ARMA Vettoriali	99
Esercizi	110
Capitolo 5. STRUTTURA DEI PROCESSI STAZIONARI	113
1. Processi Puramente Deterministici e Puramente non Deterministici	113
2. Convergenza di Martingale	122
Capitolo 6. FILTRI DI WIENER-KOLMOGOROV	125
1. Soluzione mediante la rappresentazione d'innovazione	125
2. Predittori di Wiener-Kolmogorov	131

3. L'Approccio polinomiale al filtraggio di Wiener	141
4. Interpolazione	145
5. Il processo coniugato e gli zeri della densità spettrale	148
Esercizi	153
Capitolo 7. VARIABILI ESOGENE, CAUSALITÀ E RETROAZIONE FRA PROCESSI STOCASTICI	157
1. Introduzione	157
2. Modelli a Retroazione	160
3. Reazione fra Processi Stocastici e Causalità	173
4. Modelli ARMAX e predizione in presenza di ingressi osservabili	177
5. Controllo a Minima Varianza	184
Esercizi	191
Capitolo 8. ALGORITMO DI LEVINSON	195
1. Introduzione	195
2. Innovazioni a memoria finita	195
3. Il caso scalare	203
4. Filtri di Levinson	206
5. Fattorizzazione veloce di matrici di Toeplitz	222
Esercizi	229
Capitolo 9. IL FILTRO DI KALMAN	231
1. Modelli di stato di segnali di dimensione finita	232
2. Alcune tecniche di Realizzazione	240
3. Modelli di stato non stazionari	252
4. Le equazioni del Filtro di Kalman	260
Esercizi	280
Capitolo 10. IL FILTRO DI KALMAN A REGIME	283
1. Teoria Asintotica dell'Equazione di Riccati	283
2. Fattorizzazione spettrale e ARE	303
3. Analisi del filtro a regime nel dominio della frequenza	308
4. Risoluzione dell'Equazione Algebrica di Riccati	309
Esercizi	323
Capitolo 11. IL FILTRO DI KALMAN MESSO IN PRATICA	327
1. Il Filtro di Kalman Esteso	327
2. Filtraggio Bayesiano non lineare a tempo discreto	336
3. Accordo ("tuning") del filtro di Kalman	341
Appendice A. RICHIAMI DI ALGEBRA DELLE MATRICI	349
Appendice B. RICHIAMI SUGLI SPAZI DI HILBERT	353
Appendice C. TRASFORMATA Z E SPAZI DI HARDY	359
Trasformata Z e Spazi di Hardy	360
Appendice D. ANALISI ARMONICA DEI PROCESSI STAZIONARI	365
Rappresentazione spettrale	369
Generalizzazione della trasformata di Fourier	370

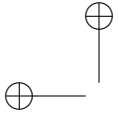


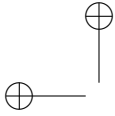
INDICE

v

Bibliografia

375





PREFAZIONE

La trasmissione, l'elaborazione, la gestione e l'organizzazione di segnali e di flussi di segnali ¹ sta diventando una parte sempre più importante dell'ingegneria moderna. Questi problemi caratterizzano un'area dell'ingegneria nota come *ingegneria dell'informazione*, e all'interno di quest'area, in particolare discipline come l'informatica, le telecomunicazioni e i controlli automatici.

Le trasformazioni epocali della società a cui stiamo assistendo (la diffusione di sistemi di comunicazione globale, l'automazione totale delle catene di produzione, lo sviluppo della robotica autonoma, ecc.) avvengono come conseguenza di tecnologie che sono imperniate sempre più sullo sviluppo di software e l'elaborazione di dati e di segnali piuttosto che sulla invenzione di nuove macchine e dispositivi di nuovo tipo (come viceversa avvenne per esempio con l'invenzione del transistor o all'epoca della rivoluzione industriale del diciottesimo secolo). In campo industriale si assiste al fatto sconcertante che aziende produttrici di software hanno fatturato e tassi di crescita molte volte superiori a grosse aziende che producono energia e macchinari. Entità fisicamente impalpabili come il "segnale" e l'"informazione" stanno caratterizzando lo sviluppo tecnologico della nostra era.

È una constatazione abbastanza ovvia che le problematiche fondamentali delle comunicazioni e del controllo hanno ragione d'essere perchè i segnali in gioco sono *quantità variabili nel tempo*, in altri termini, sono grandezze rappresentabili matematicamente come funzioni del tempo. Naturalmente nei sistemi di comunicazione la variabilità temporale è manifestazione del fatto che il segnale "contiene" (o trasporta) l'informazione utile, ad esempio una ripresa televisiva o una conversazione telefonica. Nei sistemi di controllo, viceversa, la variabilità temporale può essere una caratteristica indesiderata dovuta all'azione di disturbi; uno degli scopi dei sistemi di controllo automatico degli impianti industriali è di imporre un andamento temporale desiderato a certe variabili regolate o asservite, nonostante l'azione dei disturbi. Certe caratteristiche dell'andamento temporale dei segnali in gioco (limitatezza, assenza di oscillazioni, ecc.) sono comunque parametri fondamentali nell'analisi e nel progetto dei sistemi di controllo automatico.

Si comprende che una trattazione razionale e quantitativa di queste problematiche deve essere basata sull'uso di adeguati *modelli matematici*. In effetti, le metodologie moderne di progetto dei sistemi di controllo e di acquisizione, trattamento ed elaborazione di segnali, di codifica e compressione dati etc. sono sempre più basate sull'uso di modelli matematici dei segnali in gioco e, di conseguenza, sull'impiego estensivo del calcolatore. Il calcolatore, e più in generale l'informatica, intervengono a vari livelli, non solo come supporto fisico per l'implementazione concreta

¹La parola *segnale* è intesa come veicolo fisico per trasportare *informazione*. La natura fisica del segnale (tensione, intensità luminosa, magnetizzazione, ecc.) in quanto tale ha importanza secondaria.

di algoritmi, ma anche e soprattutto come strumento preliminare di *modellistica, analisi e progettazione* di sistemi.

Gli argomenti e i problemi che si trattano in questo testo hanno genericamente a che fare con l'analisi e il progetto di sistemi di elaborazione di segnali, basati su modelli matematici dinamici di tipo statistico. Il perché sia opportuno e naturale impiegare modelli statistici per descrivere le variabili e usare tecniche statistiche per la soluzione dei relativi problemi, è un fatto che scende dalla natura stessa dei segnali in gioco, che sono necessariamente *incerti*, cioè mai completamente noti a priori. Questo punto è fondamentale e verrà ripreso in dettaglio più avanti.

Contenuto del Libro

Nella soluzione di molti problemi dell'ingegneria del segnale ci si può ricondurre spesso a tre grandi categorie di problemi:

- *L'identificazione del modello*, che è il problema di ottenere una descrizione matematica del sistema e dei segnali che ne definiscono il comportamento. Questo è il primo passo, spesso il più oneroso, nella soluzione di molti problemi di filtraggio (misura), codifica e controllo. Il modello matematico è indispensabile ai fini di analisi e per la simulazione del sistema, e soprattutto, ai fini dell'impostazione e della soluzione dei problemi di misura e sintesi che sono descritti più sotto.
- *La misura*, che è il problema della progettazione di algoritmi per la ricostruzione di segnali non direttamente accessibili, ad esempio segnali che sono stati trasmessi da sorgenti remote o descrivono variabili "interne" essenziali del sistema, non "visibili" direttamente. La ricostruzione opera a partire da altre variabili osservabili che sono accessibili direttamente e che, tipicamente, rappresentano versioni distorte e rumorose (affette da errori di misura) delle prime.
- *La decisione e il controllo*, ovvero il progetto di algoritmi e di dispositivi (regolatori, controllori etc.), che hanno lo scopo di imporre a certe variabili "controllate" del sistema un andamento temporale desiderato.

La tematica affrontata in questo libro è soprattutto quella della "misura", ovvero della ricostruzione di variabili e segnali aleatori in presenza di distorsione e di mascheramento dovuto a rumore. La "misura" in questa accezione generale comprende classi di problemi molto generali che vengono chiamati *problemi inversi* in matematica applicata [18, 19]. Noi studieremo solo la sottoclasse che è normalmente chiamata *stima e filtraggio statistico*. Per ragioni di spazio, al problema della decisione e del controllo vero e proprio abbiamo potuto riservare uno spazio abbastanza modesto. Questa scelta è parzialmente giustificata dal fatto che le problematiche del controllo di sistemi incerti sono trattate diffusamente nella letteratura specializzata. Esse possono anche sembrare più particolari e tipiche del settore dell'Automatica, mentre questo libro ha² un'ambizione di interdisciplinarietà.

Ci sono molti problemi di filtraggio statistico e identificazione che pur essendo concettualmente simili o addirittura identici, nascono in contesti fisici o applicativi diversi (dalla bioingegneria alla telefonia, dalla elaborazione di immagini alla guida inerziale e alla navigazione automatica, dall'analisi e predizione di serie storiche

²... apparentemente in controtendenza con gli indirizzi attuali dell'istruzione universitaria in Italia.

in econometria, alla diagnostica di strutture e di edifici in ingegneria civile ...) e vengono spesso affrontati e risolti nella letteratura specializzata con tecniche ad hoc e col linguaggio specifico del settore applicativo da cui originano. Per esempio, la codifica di segnali audio e video basata sull'identificazione in linea e sull'uso di predittori lineari viene talvolta presentata come una tecnica inventata e sviluppata all'interno di un ristretto circolo di cultori di *signal processing* e sembra che alcuni non sospettino nemmeno che problematiche analoghe siano state discusse e risolte anche in statistica, nei controlli automatici, nell'econometria etc..

L'ambizione di questo libro è di offrire una serie di idee unificanti per l'analisi e la soluzione di problemi di stima e modellistica di segnali che si possono incontrare nelle comunicazioni e nel controllo ma anche in altri settori applicativi in cui si ha a che fare con segnali aleatori, senza necessariamente fare una trattazione indirizzata specificamente ad una certa comunità di ingegneri piuttosto che ad un'altra.

Per questo motivo, si è scelto di evitare la trattazione troppo specifica di alcuni argomenti orientati ad una particolare applicazione. Giusto per continuare con lo stesso esempio, la teoria della predizione lineare viene esposta, in certi testi indirizzati alla codifica e all'elaborazione numerica di segnali (ad esempio [136]), in modo così specialistico da rischiare di fornire strumenti *ad hoc* buoni solo per risolvere una certa classe di problemi (o un "problemmini") specifici del settore, senza fornire il minimo di comprensione generale che permetterebbe di affrontare problemi, anche leggermente diversi, che si presentano in altri settori dell'ingegneria. In conclusione, ci sembra che solo con una trattazione sufficientemente generale si possano far comprendere allo studente le grandi potenzialità applicative di questa disciplina. La generalità ha lo scopo di mettere l'utilizzatore in grado di trovare autonomamente la strada per modificare secondo necessità le soluzioni e gli algoritmi standard visti a lezione e arrivare a trovare le modifiche alla soluzione che si adattano meglio ai casi specifici che si possono incontrare nella carriera professionale. Sfortunatamente questi ultimi non capitano quasi mai nella forma standard trattata nei libri di testo o in ambiente accademico.

Le teorie orecchiate e mal digerite difficilmente si sanno poi mettere in pratica. L'applicabilità e il successo applicativo di molti strumenti metodologici dipendono essenzialmente da quanto profonda è la comprensione che se ne ha. Per favorire questa "comprensione orientata all'applicazione", ci sembra sia necessario curare tre aspetti:

- (1) La chiarezza nella formulazione dei problemi, nella descrizione delle ipotesi semplificatrici che stanno sempre alla base della loro formulazione matematica e nella derivazione formale delle soluzioni. Quest'ultima dovrebbe poi essere *ragionevolmente* rigorosa e completa.

Il rischio che alcuni vedono, nello scegliere questa strada, di tendere a motivare lo studente soprattutto con l'aspetto matematico della trattazione, vale la pena di essere corso ³.

- (2) L'attenzione che si dà agli aspetti *numerici* dei problemi e alle proprietà numeriche degli algoritmi che si usano per risolverli. L'utilizzatore deve sempre sapere quali sono le ragioni per cui un metodo di calcolo può dare luogo a sorprese (i migliori pacchetti software si "rifiutano" di attaccare

³C'è da dire che l'eleganza formale di una teoria è una caratteristica estremamente utile, proprio ai fini della sua applicabilità, perché ne rende semplici e naturali le idee e facilita la loro comprensione e memorizzazione.

problemi mal condizionati) e deve avere un'idea di quali azioni intraprendere in caso di malcondizionamento numerico. Deve anche sapere su che principi si basa e come reperire il software appropriato.

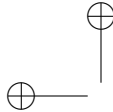
- (3) La capacità di formalizzare adeguatamente un problema ingegneristico in termini di modelli matematici che siano il più possibile semplici, ma che allo stesso tempo colgano le caratteristiche salienti del fenomeno fisico che si sta studiando e siano anche adatti alla classe e complessità degli algoritmi risolutivi che l'applicazione richiede.

Quest'ultima è in realtà un'"arte" che è l'essenza stessa dell'ingegneria moderna e che, purtroppo, è quasi impossibile comunicare con un manuale. Da un punto di vista più modesto, qui si è cercato di dare, attraverso esercizi e esempi inseriti nel testo, i rudimenti dell'arte di "porre in equazioni" alcuni problemi che si sono ritenuti "tipici" dell'area.

Questo libro è destinato principalmente a studenti di ingegneria e a ricercatori con interessi professionali nel campo dei controlli automatici, dell'elaborazione di segnali e delle telecomunicazioni. Dovrebbe essere anche di interesse allo statistico economico che lavora nel campo dei modelli econometrici dinamici e dell'analisi di serie storiche. In varie versioni preliminari è stato usato come testo per un corso universitario di teoria dell'Identificazione (ed elaborazione statistica di segnali) oppure di identificazione e analisi statistica dei dati, tenuto da diversi anni dall'autore presso la facoltà di ingegneria dell'Università di Padova.

In un corso semestrale di una settantina di ore, vengono normalmente svolti gli argomenti indicati nell'indice con eccezione talvolta del capitolo 8. La parte relativa all'identificazione, che è ovviamente indispensabile per l'uso pratico della teoria del filtraggio ma che, a sua volta, ne usa molti concetti, deve essere svolta alla fine del corso e fa riferimento ad un testo diverso.

I prerequisiti necessari per leggere questo libro sono sostanzialmente alcuni rudimenti di calcolo delle probabilità e della teoria dei processi stocastici stazionari (che vengono del resto richiamate in dettaglio in una appendice), e alcuni concetti fondamentali della teoria dei sistemi. Il linguaggio della teoria dei sistemi si usa specialmente nella seconda parte del testo quando si introducono i modelli di stato e il filtro di Kalman. La teoria dei sistemi è uno strumento naturale e necessario perchè il fine ultimo che si ha in mente qui è di produrre algoritmi risolutivi adatti all'uso in linea di sistemi di elaborazione digitale del segnale. Dopotutto la teoria dei sistemi è nata proprio allo scopo di fornire metodologie per manipolare modelli e segnali in modo congeniale all'uso in linea del calcolatore.



CAPITOLO 1

MODELLI DI SEGNALI ALEATORI

1. Segnali aleatori e processi stocastici

L'idea intuitiva di *segnale aleatorio* è associata a quella di imprevedibilità temporale. Le fluttuazioni di tensione in un circuito dovute al rumore termico, le fluttuazioni di ampiezza di un'onda radio dovute a variazioni accidentali delle costanti di propagazione del mezzo, o all'influenza di fenomeni di scarica elettrica atmosferica, l'andamento temporale del segnale elettrico microfonico durante una conversazione telefonica ecc., sono esempi di segnali aleatori familiari all'ingegnere delle comunicazioni. Altri esempi sono i cosiddetti "disturbi", di varia natura (prese di carico ecc.) con cui si devono fare i conti nella progettazione dei sistemi di regolazione automatica degli impianti industriali, l'andamento temporale di una stragrande maggioranza di variabili economiche (prezzi e domande di beni), di certe variabili biologiche ecc.

Nell'analisi dei sistemi di comunicazione e di controllo automatico i segnali aleatori giocano un ruolo particolarmente importante. Questo accade per diversi motivi:

- Perché lo scopo di un sistema di comunicazione è quello di trasmettere (nuova) informazione. Se l'andamento temporale del segnale fosse noto a priori non avrebbe senso trasmetterlo.
- Dualmente, perché lo scopo di un sistema di regolazione o controllo è, almeno in linea di principio, quello di annullare l'effetto di variabili incerte o di parametri variabili nel tempo in modo imprevedibile. Se l'evoluzione temporale delle variabili in gioco in un sistema da controllare, in particolare i disturbi, fosse nota a priori, non sarebbe necessario costruire regolatori o, più in generale, sistemi di controllo automatico. In tal caso si potrebbero infatti sintetizzare in anticipo le variabili di comando del sistema in modo da compensare esattamente i disturbi e ottenere l'andamento desiderato delle variabili controllate. In particolare, la retroazione non sarebbe necessaria.
- In entrambi i casi, si manipolano grandezze fisiche (tipicamente segnali elettrici o ottici variabili nel tempo) solo allo scopo di trasmettere informazione. Per ovvie ragioni, la trasmissione si fa alla potenza più bassa possibile e i segnali in gioco, specie in ricezione, sono di piccola ampiezza. Essi sono quindi facilmente mascherabili dal rumore e da disturbi dovuti ad interazioni non volute con l'ambiente fisico e artefatti di varia natura.

Finora abbiamo parlato di imprevedibilità temporale e di segnali aleatori in termini empirici, un po' vaghi e intuitivi. Ci proponiamo ora di elaborare la nozione empirica di segnale aleatorio per arrivare ad individuare una descrizione matematica naturale di queste caratteristiche allo scopo di proporre una definizione precisa di

modello (probabilistico) del segnale con cui lavorare in seguito. Quel che ci preme è di far chiarezza fin dall'inizio su una distinzione di carattere concettuale che è di fondamentale importanza anche da un punto di vista strettamente operativo. Come in ogni campo delle scienze sperimentali, si deve distinguere fra entità empiriche e modelli o rappresentazioni matematiche che se ne fanno. Le prime sono purtroppo sempre entità conoscibili e definibili solo in modo parziale ed approssimato. È bene che il lettore impari a distinguere in modo netto fra l'entità empirica “segnale aleatorio” e un certo ente matematico astratto, il *processo stocastico* che sceglieremo, fra altri modelli a priori possibili, allo scopo di catturarne in modo quantitativo la caratteristica di imprevedibilità temporale. Ciò che vorremmo evitare è di generare nel lettore la convinzione (errata ma assai diffusa) che le metodologie probabilistiche di stima e identificazione che svilupperemo più avanti siano strettamente applicabili solo a situazioni che in un qualche senso meritino la classificazione di “aleatorie” o “stocastiche” per l'esistenza di una struttura probabilistica evidente *a priori* in quel contesto. Di fatto *ogni* segnale (aleatorio o no) è a rigore rappresentabile come un processo stocastico e per usare questo modello non occorre affatto verificare che siano presenti nel problema in analisi meccanismi *fisici* palesi di tipo aleatorio, qualunque sia il significato che si voglia dare a questa frase. Ricordiamo a questo punto che la teoria della probabilità è solo un *modello matematico* astratto dalla realtà e come tale va riguardata: come un comodo strumento per descrivere in modo semplice e sintetico situazioni fisiche che potrebbero anche essere rappresentate mediante modelli “deterministici” (in genere molto complicati). Un esempio molto comune, come vedremo, è la descrizione “aggregata” mediante processi stocastici, degli errori che si commettono usando modelli approssimati per descrivere sistemi complessi. Un altro esempio è la descrizione “aggregata” degli errori numerici di arrotondamento in certi algoritmi di calcolo ecc.

Innanzitutto sembra a priori del tutto ovvio che il termine “segnale” (aleatorio o no) debba comunque essere sinonimo di *funzione*, $s : t \rightarrow s(t)$, t essendo l'argomento temporale (continuo o discreto). Si potrebbe poi dire che la caratteristica che contraddistingue l'“aleatorietà” o “incertezza” di un segnale s è una certa ignoranza a priori sull'andamento temporale futuro di s , anche quando ne è stata osservata la storia passata, diciamo fino ad un generico istante T . Insomma, si potrebbe dire che s è “aleatorio” quando la conoscenza di $s(t)$ per $t \leq T$ non determina univocamente l'andamento futuro $\{s(t) ; t > T\}$.

In realtà, l'analisi di questo tentativo di definizione porta ad una immediata contraddizione logica. In effetti, se l'assunto è che il segnale s debba essere una *funzione* (definita su tutto l'asse dei tempi), non può avere senso parlare di diversi prolungamenti futuri $\{s(t) ; t > T\}$, di s , giacché il “futuro” di s è, ovviamente, per definizione, parte di s stessa, in quanto funzione definita su tutto l'asse dei tempi. In altri termini, se s è una funzione, il suo andamento “futuro” $\{s(t) ; t > T\}$ è univocamente individuato da s .

Come si vede è necessario modificare la nostra idea iniziale sull'entità matematica che si adatta meglio a descrivere l'idea di “segnale aleatorio”. In effetti, per ottenere una descrizione plausibile e matematicamente sensata dell'idea intuitiva di aleatorietà bisogna considerare una *famiglia* di funzioni del tempo.

DEFINIZIONE 1.1. Un segnale¹ s è una terna $\{\mathbf{T}, V, \mathbf{S}\}$ dove

¹L'attributo “aleatorio” viene sottinteso.

- \mathbf{T} è l'insieme temporale, usualmente la retta reale \mathbb{R} o un sottointervallo del tipo $\{t \geq t_0\}$, nel caso di segnali a tempo continuo; \mathbb{Z} (gli interi) o un sottoinsieme del tipo $\{t \geq t_0, t_0 \in \mathbb{Z}\}$, nel caso di segnali a tempo discreto.
- V è l'alfabeto o spazio dei valori del segnale. In questo libro di norma $V = \mathbb{R}$ oppure \mathbb{R}^m .
- \mathbf{S} è lo spazio campionario. Ciascun elemento $s \in \mathbf{S}$ è una funzione $s : t \rightarrow s(t)$, dove $t \in \mathbf{T}$ e $s(t) \in V$, che chiameremo traiettoria del segnale.

Interpretazione. Lo spazio campionario S specifica quali sono le traiettorie ammissibili, cioè a priori possibili, del segnale. Per esempio, lo spazio campionario di un segnale telefonico potrebbe essere descritto come l'insieme delle funzioni continue del tempo con trasformata di Fourier a banda limitata compresa tra i 50 e i 3000 Hz, quello del segnale di scarica di un circuito RC (senza generatori) come una famiglia di funzioni esponenziali del tipo $\{A e^{-t/RC}; A \in \mathbb{R}\}$ ecc.

Diremo che un segnale è *deterministico* (o *determinato*) se il suo spazio campionario si riduce ad una sola traiettoria.

I segnali di cui ci interesseremo saranno quasi sempre segnali a *tempo discreto*, ottenuti ad esempio per campionamento temporale di variabili continue. La variabile temporale, che designeremo di norma con il simbolo t , può essere pensata semplicemente con un indice adimensionale che individua l'ordinamento dei campioni del segnale. I simboli comunemente usati per designare l'asse dei tempi sono \mathbb{Z} (gli interi), \mathbb{Z}_+ (gli interi non negativi) ed \mathbb{N} (i naturali).

Supponiamo di avere a che fare con un certo segnale aleatorio s e di disporre di una misura, diciamola $\{\bar{s}(t)\}_{t=0,1,\dots,T}$ del segnale nell'intervallo temporale $0 \leq t \leq T$. (Questo insieme di dati viene chiamato *serie temporale* o *serie storica* nella letteratura statistico-economica). Immaginiamo di riuscire a tracciare tutti i possibili "prolungamenti" di $\{\bar{s}(t)\}$ per $t > T$, ottenendo così un "universo" di possibili traiettorie future che il segnale potrebbe descrivere. Possiamo allora pensare di raffigurarci il meccanismo responsabile dell'aleatorietà del segnale come un processo di "estrazione" della traiettoria futura $\{s(t); t > T\}$ da una ipotetica "urna" contenente tutti i prolungamenti costruiti più sopra; si potrebbe dire che la "natura" sceglie la traiettoria effettiva dall'universo delle possibili. Lo schema richiama ovviamente il modello concettuale probabilistico. Per renderlo concreto esso andrà poi completato assegnando una misura quantitativa della *probabilità* con cui viene scelta una certa particolare traiettoria (o una certa famiglia di traiettorie) dall'universo dei possibili andamenti futuri. In realtà il ruolo dell'istante T è solo di ausilio concettuale. In questo schema "a urna" si può altrettanto bene pensare che la "natura" (cioè le condizioni sperimentali) sceglie l'*intera* traiettoria del segnale. Da notare che questa descrizione non ha nessuna pretesa di descrivere fisicamente il meccanismo di effettiva generazione del segnale² e serve solo per motivare la descrizione matematica di tipo probabilistico che daremo qui di seguito.

Questa consta di tre ingredienti fondamentali:

- 1) L'insieme di tutte le possibili traiettorie (o andamenti temporali) del segnale, diciamolo S . Questo insieme, lo *spazio campionario* del segnale, potrebbe semplicemente essere la collezione di tutte le possibili funzioni definite sull'asse temporale

²Per esempio lo "stato di natura" che al suo verificarsi sceglie una particolare traiettoria del segnale, dovrebbe essere qualcosa che ha capacità di prevedere il futuro, dato che dovrebbe determinare tutto l'andamento futuro della traiettoria prescelta.

(\mathbb{Z} , \mathbb{Z}_+ o \mathbb{N}), a valori reali. S è l'“urna” dalla quale viene estratta “a caso” la traiettoria del segnale.

2) Una famiglia di questioni (o proposizioni logiche) che ha interesse porsi riguardo all'andamento temporale del segnale. Ad esempio: “all'istante t l'ampiezza del segnale, $s(t)$, è compresa nell'intervallo $[a, b]$ ”, “esiste il limite per $t \rightarrow \infty$ di $s(t)$ ”, “per tutti i t pari il segnale prende valori negativi” ecc...

Queste sono ovviamente affermazioni che possono essere verificate (sono vere o false) solamente *dopo* che è avvenuta l'“estrazione” della traiettoria del segnale (cioè quando si dispone già di una *misura* del segnale). Quel che si vorrebbe, viceversa, è riuscire ad assegnare loro, *a priori*, un grado di verosimiglianza (una “probabilità”).

È ovvio che una qualunque affermazione relativa all'andamento temporale del segnale può essere rappresentata come *un sottoinsieme dello spazio campionario* S e che gli operatori logici del calcolo proposizionale (“e”, “oppure”, “non”) corrispondono in questa rappresentazione alle operazioni insiemistiche di intersezione, unione e complementazione³. Le proposizioni logiche, rappresentate come sottoinsiemi di S si chiamano *Eventi*.

Esempi importanti di eventi sono i sottoinsiemi di S corrispondenti ad affermazioni del tipo: “all'istante t_1 il segnale ha ampiezza nell'intervallo $I_1 = [a_1, b_1]$ e, ..., e, all'istante t_n il segnale ha ampiezza compresa nell'intervallo $I_n = [a_n, b_n]$ ”. Questi eventi si denotano comunemente mediante una scrittura del tipo

$$C_{t_1 \dots t_n}(I_1 \dots I_n) := \left\{ s \in S ; s(t_1) \in I_1, \dots, s(t_n) \in I_n \right\} \quad , \quad (1.1)$$

che si legge: “l'insieme degli elementi s di S definito dalle condizioni che appaiono dopo il punto e virgola, ovvero: $s(t_1)$ appartenga a I_1 ecc...”. Combinando fra di loro per mezzo di operatori insiemistici gli eventi del tipo (1.1) si ottengono rappresentazioni di proposizioni logiche via via più complesse. Senza entrare nel dettaglio della costruzione, diremo qui che si può arrivare ad individuare in questo modo una famiglia “massimale” di eventi, \mathcal{S} (che ha la struttura di σ -algebra *Booleana*) la quale racchiude tutte le affermazioni logiche di un qualche interesse pratico relative all'andamento del segnale.

3) Il terzo e più importante ingrediente è una *misura di probabilità* su \mathcal{S} , ovvero una funzione d'insieme $P : \mathcal{S} \rightarrow [0, 1]$ che permetta di assegnare a ciascun evento A di \mathcal{S} un numero compreso fra zero e uno, $P(A)$. Questo numero si interpreta come misura quantitativa della “fiducia” con cui ci si può aspettare a priori che l'evento A si verifichi effettivamente.

Naturalmente, in accordo con i dettami della teoria della probabilità, richiederemo che P sia una funzione lineare (e continua) rispetto alle operazioni d'insieme. Più precisamente, per ogni successione di eventi disgiunti $\{A_n\}$ di \mathcal{S} dovrà valere la

$$P \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_1^{\infty} P(A_n) \quad . \quad (1.2)$$

Questa proprietà (che si chiama “ σ -additività”) ha, com'è noto, numerose conseguenze importanti dal punto di vista del calcolo concreto delle probabilità di eventi. Una che ha un particolare significato in questo contesto è la seguente.

³L'insieme corrispondente a una certa proprietà matematica, \mathcal{P} , delle traiettorie è proprio l'insieme delle traiettorie di S che hanno la proprietà \mathcal{P} .

Siano t_1, \dots, t_n , n istanti arbitrari e consideriamo eventi del tipo (1.1) in cui gli intervalli I_1, \dots, I_n sono semirette del tipo $I_k = (-\infty, x_k)$, $k = 1 \dots n$, il cui estremo superiore si interpreta come una variabile corrente. La funzione di x_1, \dots, x_n , t_1, \dots, t_n , definita ponendo

$$F_n(x_1, \dots, x_n, t_1, \dots, t_n) := P(\{s \in S; s(t_1) \leq x_1, \dots, s(t_n) \leq x_n\}) \quad , \quad (1.3)$$

si chiama *funzione di distribuzione di probabilità (d.d.p) congiunta di ordine n* (o funzione di ripartizione congiunta di ordine n) relativa a P . È chiaro che si può definire un'intera famiglia di d.d.p. $\{F_n\}$ relative a P , indicata da n e che, ad esempio, tra F_{n-1} ed F_n sussiste una relazione, detta *di consistenza*, del tipo

$$\begin{aligned} F_n(x_1, \dots, x_{i-1}, +\infty, x_{i+1}, \dots, x_n, t_1 \dots t_n) = \\ = F_{n-1}(x_1, \dots, x_{i-1}, x_{i+1} \dots x_n, t_1 \dots t_{i-1}, t_{i+1}, \dots, t_n) \quad , \end{aligned}$$

che vale identicamente, qualunque sia $1 \leq i \leq n$.

Orbene, Kolmogorov [80] ha dimostrato che se sono note le (infinite) distribuzioni congiunte della famiglia $\{F_n\}$ allora *la probabilità P su \mathcal{S} è specificata completamente*, nel senso che esiste un procedimento che permette di calcolare $P(A)$ per ogni evento A di \mathcal{S} , a partire dalle $\{F_n\}$. Per assegnare la P occorre quindi in generale assegnare la famiglia (numerabile) dalle sue funzioni d.d.p. di ordine finito.

La descrizione matematica di un segnale aleatorio cui d'ora in avanti faremo riferimento sarà lo schema composto dei tre ingredienti $\{S, \mathcal{S}, P\}$ che abbiamo appena definito. Questa terna si chiama *processo stocastico* (a tempo discreto).

Per non confondere il lettore conviene a questo punto confrontare questa definizione con quella, un tantino più generale ma forse più astratta, che viene comunemente presentata nei testi di teoria della probabilità.

Secondo la definizione usuale, un processo stocastico (reale) è *una collezione di variabili aleatorie*, $\{y(t)\}$, *indicata dal parametro $t \in \mathbb{Z}$ (tempo discreto) oppure $t \in \mathbb{R}$ (tempo continuo)*, tutte definite su un medesimo spazio di probabilità astratto $\{\Omega, \mathcal{A}, P\}$.

L'insieme Ω è qualche volta chiamato spazio degli "eventi elementari" ω (o anche spazio degli "stati di natura"), \mathcal{A} è una σ -algebra di sottoinsiemi di Ω (gli "eventi") e P una misura di probabilità su \mathcal{A} . Ciascuna variabile aleatoria, $y(t)$, è una funzione (\mathcal{A} -misurabile) che ad ogni scelta (casuale) di uno "stato di natura" ω fa corrispondere il numero reale $y(t, \omega)$; in simboli

$$y(t) : \omega \rightarrow y(t, \omega) \quad . \quad (1.4)$$

Le *traiettorie* (o realizzazioni) del processo sono le successioni (reali) $\{y(\cdot, \bar{\omega})\}$ che si ottengono elencando in sequenza le determinazioni di ciascuna variabile aleatoria $y(t)$ corrispondenti alla selezione di un particolare evento elementare $\bar{\omega} \in \Omega$.

È facile verificare che la nostra definizione si può ricondurre a quella più generale appena riportata. Basta in effetti identificare lo spazio campionario S del processo con Ω ed \mathcal{S} con \mathcal{A} , col che $\{S, \mathcal{S}, P\}$ diventa una versione "concreta" dello spazio di probabilità astratto $\{\Omega, \mathcal{A}, P\}$. Se si definisce inoltre una variabile aleatoria $y(t)$ su $\{S, \mathcal{S}, P\}$ ponendo⁴

$$y(t, s) := s(t) \quad , \quad (1.5)$$

⁴Notare che la funzione è \mathcal{S} -misurabile per ogni t per costruzione.

ovvero, $y(t)$ associa a ciascuna traiettoria $s \in S$ il suo valore all'istante t , si vede che al variare di t , le determinazioni delle variabili $\{y(t)\}$ corrispondenti ad un elemento fissato $\bar{s} \in S$ descrivono, ovviamente, la traiettoria $\{\bar{s}(t)\}$ del processo.

Inversamente, si può sempre ridefinire un processo $\{y(t)\}$ dato su un generico spazio di probabilità astratto $\{\Omega, \mathcal{A}, P\}$, sul suo *spazio di probabilità campionario* $\{S, \mathcal{S}, \hat{P}\}$, dove S e \mathcal{S} sono lo spazio campionario e la relativa algebra di eventi del processo definite in precedenza, e \hat{P} è la misura di probabilità su \mathcal{S} *indotta dal processo*, che assegna all'insieme $\hat{A} \in \mathcal{S}$ di traiettorie del processo che corrispondono ad un evento $A \in \mathcal{A}$ la stessa probabilità

$$\hat{P}(\hat{A}) := P(A).$$

Questa rappresentazione equivalente del processo viene chiamata *rappresentazione sullo spazio campionario*.

La rappresentazione sullo spazio campionario è, univocamente individuata dal processo, una volta noto l'insieme in cui le $y(t)$ prendono valori ed è dato uno spazio di funzioni contenente le sue possibili traiettorie. Essa è particolarmente comoda perchè permette di non preoccuparsi affatto di descrivere con precisione lo spazio di probabilità sul quale il processo è definito e in moltissimi casi permette di definire un processo stocastico assegnandone semplicemente la famiglia delle distribuzioni di probabilità finito-dimensionali. Per esempio si può definire un processo *Gaussiano*, semplicemente assegnando la famiglia di densità di probabilità

$$p_n(x_1, \dots, x_n, t_1, \dots, t_n) = [(2\pi)^n \det \Sigma(t_1, \dots, t_n)]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[x - \mu]' \Sigma(t_1, \dots, t_n)^{-1}[x - \mu]\right\} \quad n = 1, 2, \dots$$

in cui

$$[x - \mu] := \begin{bmatrix} x_1 - \mu(t_1) \\ \vdots \\ x_n - \mu(t_n) \end{bmatrix} \quad \Sigma(t_1, \dots, t_n) := \begin{bmatrix} \sigma(t_1, t_1) & \sigma(t_1, t_2) & \dots & \sigma(t_1, t_n) \\ \sigma(t_2, t_1) & \sigma(t_2, t_2) & \dots & \sigma(t_2, t_n) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(t_n, t_1) & \sigma(t_n, t_2) & \dots & \sigma(t_n, t_n) \end{bmatrix}.$$

che sono a loro volta completamente note una volta assegnate la funzione *media* $\mu(t)$ e la *funzione di covarianza* $\sigma(t, s)$.

Finora abbiamo parlato solamente di segnali (e processi) a valori reali. Nelle applicazioni si ha spesso a che fare con misure simultanee di più segnali mutuamente correlati ed è necessario sviluppare algoritmi che elaborino le misure "in parallelo". Occorre allora generalizzare la nozione di processo (reale) considerata fin qui e introdurre quella di *processo stocastico a valori vettoriali*. La generalizzazione è di per sè abbastanza ovvia e vi accenneremo molto rapidamente. Un *processo stocastico vettoriale* $\{y(t)\}$ (a valori, ad esempio, in \mathbb{R}^m) sarà caratterizzato da uno spazio campionario, S , formato da successioni s a valori vettoriali (in \mathbb{R}^m).

Scriveremo il *vettore aleatorio* $y(t)$ come vettore colonna

$$y(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{bmatrix}$$

e $y(t)$ inteso come applicazione: $s \rightarrow s(t)$ sarà allora ancora rappresentabile come un vettore colonna di m variabili aleatorie scalari $y_k(t) : s \rightarrow s_k(t)$, $k = 1 \dots m$, dove

$s_k(t)$ è la k -sima componente di $s(t)$. Se $x = [x_1, \dots, x_m]$ denota una variabile corrente su \mathbb{R}^m e si usa la notazione $s(t) \leq x$ per intendere che $s_1(t) \leq x_1, \dots, s_m(t) \leq x_m$, si può allora ancora scrivere la funzione d.d.p. (del primo ordine) del processo come

$$F(x, t) = P\{s; s(t) \leq x\} = P\{s; y(t, s) \leq x\} \quad .$$

Espressioni analoghe possono ovviamente essere scritte anche per le d.d.p. di ordine superiore.

2. Modelli stocastici e identificazione

Fin qui abbiamo solo definito una classe molto generale di descrizioni matematiche di segnali aleatori. Una volta accettato il principio della descrizione probabilistica di un segnale aleatorio come processo stocastico, è però assolutamente necessario saper dare dei procedimenti che portino a costruire in concreto questo tipo di modelli. In termini generali si potrebbe dire che questo è lo scopo della teoria dell'*Identificazione*. Per comprendere meglio i termini in cui questo problema va posto è necessario discutere prima della differenza e del diverso ruolo giocati dai modelli stocastici che tratteremo in questo libro e i modelli deterministici usati tradizionalmente in ingegneria.

Tradizionalmente anche segnali intrinsecamente imprevedibili, come ad esempio il segnale vocale, vengono trattati e modellati in modo deterministico, ad esempio come quantità caratterizzate semplicemente da un certo contenuto armonico (banda di frequenze). In molte applicazioni, ad esempio la telefonia tradizionale, questo modo abbastanza primitivo di descrivere i dati è più che sufficiente. In sistemi sofisticati di comunicazione e controllo in cui si chiedono prestazioni spinte, si richiedono di norma anche operazioni sofisticate di trattamento del segnale, ad esempio una codifica efficiente, l'estrazione di segnale da rumore in condizioni molto sfavorevoli ecc. e per queste operazioni la sola conoscenza della banda può essere insufficiente.

La tendenza attuale è quella di eseguire la maggior parte delle operazioni di elaborazione del segnale in modo digitale e di vedere quindi un filtro o un regolatore come un *algoritmo* che verrà "programmato" su un certo supporto di calcolo. La versatilità e la capacità di calcolo dei moderni sistemi di elaborazione sono enormemente superiori a quelle dei dispositivi analogici o numerici del passato e, per questo motivo, è oggi possibile eseguire "in linea" (in tempo reale) operazioni molto più complesse di quelle che si ritenevano possibili solo pochi anni fa. La chiave per sfruttare appieno le potenzialità di questi sistemi di elaborazione è ovviamente la disponibilità di modelli matematici accurati dei segnali e sistemi in gioco. L'uso di modelli rozzi e semplicistici, anche quando il problema fornisce tutta l'informazione necessaria per descrizioni più verosimili del segnale, perde molte delle giustificazioni tradizionali.

L'uso di modelli matematici, in particolare di modelli probabilistici, del segnale permette di costruire algoritmi efficienti di controllo e filtraggio basati sull'estrazione "ottima" del segnale dal rumore, di definire in termini precisi la parte non ridondante di informazione contenuta in un segnale (l'innovazione) ai fini di una sua codifica efficiente ecc...

Naturalmente le ragioni che giustificano metodologie di progetto basate sull'uso di *modelli matematici probabilistici* dei segnali in gioco devono, in ultima analisi, venire dal successo pratico di queste tecniche e dai risultati che questo approccio

permette di conseguire. Citeremo a questo proposito solo un paio di esempi e tra questi innanzitutto la grande diffusione delle tecniche di *filtraggio alla Kalman* [2] in questi ultimi decenni. Queste tecniche hanno avuto un ruolo molto importante nel campo aerospaziale ed aeronautico negli anni 1960–70, ma hanno successivamente avuto impiego nei settori applicativi più disparati, dal controllo di processi, all'idrologia (predizione di piene di fiumi, ecc.), alla bioingegneria, all'economia ecc...

La ragione del successo del filtro di Kalman viene talvolta indicata nella particolare struttura algoritmica dello schema, ma questo è in realtà un aspetto (non trascurabile, ma) di secondo piano rispetto al fatto di essere un algoritmo costruito *su misura* di un particolare modello matematico *dei segnali in gioco*. Il filtro di Kalman è “cucito addosso” ad un modello probabilistico del segnale. Funziona bene per *quel* segnale, naturalmente se il modello è ragionevolmente accurato; può però funzionare malissimo per segnali che hanno ad esempio la stessa “banda” ma caratteristiche dinamiche e statistiche diverse. Si intravede come la filosofia di progetto tradizionale dei filtri sia profondamente modificata in questo contesto. Se si parte da modelli matematici dei segnali in gioco di struttura opportuna (in questo caso si tratta di modelli di stato lineari di tipo Gauss-Markov) è finalmente possibile porre il problema del filtraggio e del controllo in presenza di disturbi aleatori in termini matematicamente precisi ed arrivare a soluzioni le quali, una volta specificata la classe degli algoritmi ammissibili, sono le *migliori possibili*, tipicamente quelle che forniscono la minima varianza dell'errore di ricostruzione del segnale utile.

Il punto fondamentale nel filtraggio di Kalman è dunque la specifica di un *modello matematico dei segnali in gioco*. Il filtro ha una struttura fissa che risulta in buona parte predeterminata *dal modello*. Questa è in un certo senso l'esatto contrario della filosofia di progetto dei filtri digitali tradizionali che sono basati sul concetto di “banda passante” e hanno in linea di principio la pretesa di essere “universali”.

L'obiezione più comune al tipo di argomentazioni appena esposte è un certo scetticismo sull'effettiva possibilità di disporre (o saper costruire) modelli matematici sufficientemente accurati dei segnali in gioco. Questa obiezione è certamente fondata se si pretende che il modello abbia significato “fisico”, ovvero sia una schematizzazione del meccanismo fisico che genera il segnale, costruita in base a delle leggi fisiche elementari che governano questo meccanismo. In verità, anche nei pochissimi casi in cui la “fisica” delle variabili in gioco è nota, i modelli a cui conduce questa filosofia risultano spesso di impiego poco pratico, sia per la complicazione dovuta alla parametrizzazione “fisica” in cui compaiono molti parametri incogniti o poco noti, sia per le approssimazioni poco controllabili che inevitabilmente si introducono combinando descrizioni “fisiche” di molti sottosistemi elementari.

In realtà, nella stragrande maggioranza delle applicazioni un modello matematico deve solamente assolvere allo scopo di fornire uno schema sintetico che compendia certe caratteristiche dinamiche e statistiche con cui si vuole descrivere il segnale. Ad esempio un modello statistico di segnali ha lo scopo di produrre processi con una certa funzione densità spettrale di potenza (che può spesso essere stimata empiricamente) o una certa funzione covarianza ecc...

Questi modelli sono nella maggior parte dei casi solo descrizioni matematiche di traiettorie del segnale osservato e il fatto che ad esse si possa attribuire una qualche interpretazione fisica più o meno mediata può essere in qualche caso utile, ma non certamente indispensabile.

In moltissimi casi la questione della costruzione di un modello di segnale aleatorio può essere affrontata quindi con un approccio essenzialmente “fenomenologico”: quello di *descrivere matematicamente i dati osservati, non di scoprire le leggi fisiche che spiegano il meccanismo della loro generazione*. In questo contesto il problema diviene allora quello di sviluppare delle adeguate metodologie *statistiche* le quali, basandosi sull’osservazione di un “campione” del segnale, conducano appunto alla selezione di un’adeguata descrizione matematica (un modello probabilistico) del segnale osservato. Come si vede, l’obiettivo qui è assai meno ambizioso di quello della Fisica. Invece di ricercare le leggi matematiche che governano delle *classi* di fenomeni (individuate da certe caratteristiche comuni), si ricercano leggi che descrivano solo il particolare fenomeno osservato. La “generalizzazione”, tipica delle scienze sperimentali non riguarda la Statistica.

La teoria dell’*identificazione* tratta per l’appunto di queste metodologie di modellizzazione di segnali a partire da dati osservati. È un campo che è stato oggetto di ricerca molto attiva in questi ultimi decenni e entro certi limiti, è una teoria già ben collaudata. Si può affermare che oggi esistono algoritmi semplici, veloci e numericamente stabili, che risolvono il problema della modellizzazione di segnali e di sistemi dinamici in modo soddisfacente.

Concluderemo questo paragrafo accennando alla applicazione forse più nota di queste idee nel campo della codifica del segnale vocale. Questo segnale è l’esempio più classico di segnale aleatorio che veniva tradizionalmente descritto solo “per banda”.

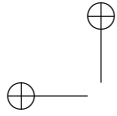
Attraverso una dettagliata analisi del processo di formazione e trasmissione dei suoni nella cavità orale (per i dettagli si può consultare il Rabiner-Schafer [38], pag. 105) si arriva a formulare una classe di modelli matematici di tipo ‘fisico’ del segnale vocale (campionato) che nei casi più semplici sono del tipo

$$y(t) = \sum_1^n a_i y(t-i) + e(t) + u(t) \quad , \quad (2.1)$$

dove $\{y(t)\}$ è il segnale vocale, $\{e(t)\}$ è un processo scorrelato a media zero (rumore bianco) e $\{u(t)\}$ è una sequenza (periodica) di impulsi formanti. Il segnale può essere considerato lentamente variabile rispetto al passo di campionamento e questo in concreto significa che i coefficienti del modello, $\{a_1, \dots, a_n\}$, e il periodo del treno di impulsi si possono considerare praticamente costanti per una durata di $10 \div 20$ msec (corrispondenti a $100 \div 200$ campioni ad una frequenza di campionamento di 10 kHz).

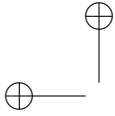
È abbastanza curioso il fatto che (come vedremo in seguito) modelli matematici del tipo (2.1) siano in realtà una descrizione standard di una classe molto vasta di segnale stazionari. Ciò che vale la pena di mettere in evidenza qui è che il grosso lavoro di modellizzazione “fisica” del segnale vocale serve a dare informazioni sostanziali solo sulla struttura del “segnale formante” $\{u(t)\}$. Per il resto (2.1) è un classico modello “autoregressivo” (AR) comunemente usato per rappresentare serie temporali stazionarie in svariati campi applicativi.

Usando le tecniche standard di identificazione parametrica si possono ottenere dalle stime “in linea” dei parametri $\{a_1, \dots, a_n\}$ del modello (2.1) e, impiegando degli algoritmi veloci di identificazione, queste stime possono essere ricalcolate per ciascun intervallo “di stazionarietà” (i $100 \div 200$ campioni di cui si diceva più sopra). In pratica i parametri stimati vengono man mano (codificati e) trasmessi



al posto del segnale vocale originale. È chiaro che, una volta noti i parametri del modello, in ricezione sarà sufficiente disporre di una sorgente di rumore bianco e di un sintetizzatore di impulsi per poter ricostruire il segnale.

Seguendo sostanzialmente questo schema di principio si può eliminare una quantità notevole di ridondanza nel segnale vocale. Con alcuni accorgimenti si può passare dalle velocità di trasmissione dei sistemi tradizionali di 64 kbit/s a $1 \div 2$ kbit/s!



CAPITOLO 2

IL PROBLEMA DELLA STIMA

1. Introduzione alla Stima Bayesiana

Il problema statistico della *stima* può essere visto come una formalizzazione matematica del *problema della misura*.

Lo scopo di un esperimento di misura è quello di ottenere informazioni su una variabile x , che non è osservabile direttamente, basandosi su dei valori numerici osservati, y , di certe altre variabili direttamente osservabili prodotte dallo “strumento” di misura. Quindi la misura di una certa variabile fisica x , o più in generale di una n -pla di variabili fisiche $x = [x_1, \dots, x_n]'$ si può vedere come un procedimento attraverso il quale ad x si associano dei dati numerici $y = [y_1, \dots, y_m]'$ che si suppongono “influenzati” in modo noto o parzialmente noto da x .

In condizioni ideali la conoscenza del dispositivo di accoppiamento tra x e le variabili osservabili (che chiameremo anche *osservazioni* o *misure*) e il valore numerico y di queste ultime determinatosi all'atto della misura dovrebbero permettere di risalire in modo univoco a un valore numerico per x . Di fatto, in tutte le situazioni di interesse pratico non è possibile isolare completamente il sistema fisico o la variabile fisica che si vuole misurare e il valore numerico y della variabile di osservazione viene determinato anche da un complesso di altre variabili non osservabili dovute all'ineliminabile interazione del sistema di misura con l'“ambiente” fisico circostante. Questa influenza, come comunemente si dice, della “condizione sperimentale”, ha l'effetto di rendere *incerto* il risultato della misura. In effetti le “condizioni sperimentali” variano nel tempo con la dinamica propria dell'ambiente fisico e, anche in misure ripetute della stessa variabile, effettuate con lo stesso strumento, le diverse condizioni sperimentali causano in generale diverse determinazioni numeriche del risultato della misura.

Non spenderemo troppe parole per cercare di convincere il lettore del fatto che questo tipo di situazione, descritta per concretezza con riferimento alle misure della Fisica sperimentale o dell'Elettrotecnica e in un contesto statico, è in realtà suscettibile di interpretazioni molto più generali. Per esempio la descrizione precedente si applica altrettanto bene, anziché alla misura di lunghezze, tensioni o masse, basata sull'elongazione dell'indice di uno strumento, alla ricostruzione di un messaggio x trasmesso attraverso un canale distorto e rumoroso, a partire dal corrispondente segnale y acquisito dal dispositivo di ricezione.

L'approccio Probabilistico. La presenza di incertezza sul risultato della misura porta in modo naturale a introdurre modelli di accoppiamento tra x e y di tipo probabilistico. In effetti, una delle motivazioni “storiche” della teoria della probabilità (Gauss [G]) è stata proprio quella di fornire uno schema matematico che permettesse la descrizione (anche se “descrivere” in questo caso non significa

necessariamente *rendere ragione*) del fatto che più misure successive fatte su una stessa grandezza fisica potessero produrre risultati numerici *diversi*.

Nel modello costruito da Gauss (che sta alla base della cosiddetta *teoria degli errori* della fisica sperimentale) i risultati di m misure successive, $\{y_k ; k = 1, \dots, m\}$, vengono modellati secondo la relazione

$$y_k = h(x) + \varepsilon_k \quad , \quad (1.1)$$

dove $h(x)$ è la caratteristica ideale dello strumento, una funzione nota di x , e gli ε_k , $k = 1, \dots, m$, sono ciascuno la somma di un gran numero di “errori accidentali”, termini additivi individualmente piccoli e a media nulla, dovuti all’interazione con l’ambiente. Tra questi si può anche includere la deviazione (ignota) della caratteristica “reale” dello strumento da quella ideale $h(x)$. Per descrivere questa situazione si può immaginare di introdurre uno spazio di probabilità in cui gli eventi elementari sono tutte le possibili diverse condizioni sperimentali che con il loro verificarsi possono mutare la determinazione numerica delle variabili osservabili. I termini d’errore ε_k che intervengono nella (1.1) diventano così *variabili aleatorie*. Dalla definizione qualitativa di errore accidentale si può poi arrivare ad una caratterizzazione matematica degli ε_k come variabili indipendenti a distribuzione Gaussiana (per il teorema del limite centrale) a media zero. La varianza dei termini d’errore può essere incognita ma è spesso determinabile in base alle specifiche di precisione dello strumento.

La (1.1) diventa così un semplice modello probabilistico che specifica la distribuzione di probabilità delle misure y , nota a meno del “parametro” incognito x . Quest’ultimo è appunto il valore numerico della grandezza fisica da determinarsi.

Già in questo semplicissimo esempio non è immediatamente evidente quale possa essere un modo ragionevole per ricavare il valore di x a partire dai dati osservati. In effetti, dato che la possibilità di ricostruire univocamente x è in pratica da escludersi, si deve impostare il problema come un *problema di approssimazione*: quello della ricerca della *migliore ricostruzione possibile* del valore di x , in base ai dati misurati e alla conoscenza del modello probabilistico che descrive le misure.

Questo è un tipico *problema di stima* statistica.

In questo capitolo tratteremo del problema della stima seguendo il cosiddetto *approccio Bayesiano*. Questo approccio, a differenza di quello “classico” o “Fisheriano”, fa riferimento a situazioni in cui esiste una informazione a priori di tipo probabilistico anche sulla variabile da stimare, x . Si parte quindi dall’assunto che x non sia una grandezza “certa” descrivibile per mezzo di un “valore vero” deterministico ma incognito, ma piuttosto essa stessa sia una quantità *aleatoria* che non è suscettibile, per sua stessa natura, di essere conosciuta esattamente.

I possibili valori di x sono descritti come determinazioni di una variabile aleatoria (o di un vettore aleatorio n -dimensionale) che si distribuiscono in modo casuale su \mathbb{R} (o su \mathbb{R}^n) secondo una certa legge di probabilità.

Da un punto di vista pratico si può ben dire che molto spesso si hanno a disposizione informazioni a priori su x sufficienti a giustificare l’adozione dell’approccio Bayesiano. Per esempio sono in generale noti un “valore nominale” di x e dei parametri che ne quantificano la dispersione attorno al valore nominale (“tolleranza”, “classe di precisione” ecc.). Queste informazioni possono quasi sempre essere tradotte in parametri di tipo probabilistico come media e varianza e talvolta anche in una vera e propria distribuzione di probabilità. Ne viene che, al di là dell’assunto ideologico sulla natura di x (quantità certa ma ignota, oppure quantità aleatoria)

che può sembrare più o meno plausibile secondo le circostanze, la reale motivazione dell'uso dell'approccio Bayesiano risiede nell'effettiva disponibilità di dati probabilistici "a priori" su x . Se questi dati sono effettivamente disponibili sarebbe errato non tenerne conto.

Procediamo ora a descrivere formalmente che cosa si intende per un problema di stima Bayesiana. In questo capitolo sarà necessario distinguere accuratamente variabili aleatorie da valori campionari. Per le prime useremo perciò **caratteri in grassetto minuscoli** e per i secondi i corrispondenti caratteri tipografici in corpo normale.

Sia \mathbf{x} un vettore aleatorio n -dimensionale la cui distribuzione di probabilità supporremo per il momento completamente nota; \mathbf{x} non è accessibile, il che significa che non si possono osservare i valori campionari x di \mathbf{x} .

Di norma assumeremo tutte le variabili aleatorie in gioco di tipo continuo e dotate di densità. Indicheremo con

$$p_{\mathbf{x}}(x) = p_{\mathbf{x}}(x_1, \dots, x_n) \quad (1.2)$$

la densità di probabilità di \mathbf{x} , detta anche densità *a priori*.

Sia poi \mathbf{y} il vettore aleatorio m -dimensionale che rappresenta le misure. Supporremo nota la *densità condizionata* di \mathbf{y} dato il valore x assunto da \mathbf{x} , $f(y | \mathbf{x} = x)$, che denoteremo anche con il simbolo $f(y | x)$,

$$f(y | x) = P(y \leq \mathbf{y} \leq y + dy | \mathbf{x} = x) / dy \quad (1.3)$$

Questa funzione si può considerare come la descrizione matematica dello strumento di misura o del canale di trasmissione.

Il *problema di stima* (o *inferenza statistica Bayesiana*) su \mathbf{x} è di ricostruire il vettore aleatorio \mathbf{x} sulla base del modello probabilistico (1.2) e (1.3) e dell'osservazione $\mathbf{y} = y$ della variabile di misura. Questo può essere inteso semplicemente come il problema di calcolare la nuova distribuzione di probabilità di \mathbf{x} determinata dall'osservazione del valore campionario $\mathbf{y} = y$, oppure come la ricostruzione del *valore campionario* $x = \mathbf{x}(\omega)$ che si è determinato all'atto dell'esperimento di misura.

Se $p_{\mathbf{x}}$ è nota completamente, il primo problema ha la soluzione più ovvia e più esauriente che ci si possa aspettare. In effetti dalla (1.2) e dalla (1.3) si può ricavare la densità condizionata di \mathbf{x} data l'osservazione $\mathbf{y} = y$,

$$f(x | \mathbf{y} = y) = \frac{f(y | x) p_{\mathbf{x}}(x)}{\int_{\mathbb{R}^n} f(y | x) p_{\mathbf{x}}(x) dx} = \frac{p_{\mathbf{y}\mathbf{x}}(y, x)}{p_{\mathbf{y}}(y)} \quad (1.4)$$

e questa è la descrizione più esauriente possibile di quanto l'osservazione $\mathbf{y} = y$ modifica la nostra conoscenza a priori di \mathbf{x} , data da $p_{\mathbf{x}}$. La (1.4) viene comunemente chiamata *densità di probabilità a posteriori* di \mathbf{x} .

In realtà la (1.4) è una descrizione in un certo senso "troppo completa" della ricostruzione di \mathbf{x} e non è di pratico impiego. In pratica si ha bisogno di valori numerici con cui portare avanti i calcoli, cioè di una *stima puntuale* di \mathbf{x} che dica qual è il *valore campionario* (non osservabile) $\mathbf{x} = x$ che ha dato luogo all'osservazione $\mathbf{y} = y$.

Il problema della stima puntuale si può visualizzare meglio assumendo che il meccanismo di accoppiamento fra la variabile misurata non accessibile \mathbf{x} e quella misurabile \mathbf{y} sia descrivibile analiticamente mediante una relazione del tipo

$$\mathbf{y} = h(\mathbf{x}, \mathbf{w}) \quad (1.5)$$

in cui il vettore aleatorio \mathbf{w} (il “rumore di misura”) rappresenta le variabili di interazione dell’ambiente fisico col dispositivo di misura. Queste sono le variabili che causano incertezza nella misura, per cui il valore assunto da \mathbf{x} durante l’esperimento non si riflette nell’osservazione $\mathbf{y} = y$ in modo certo e prevedibile.

Al momento dell’esperimento le condizioni sperimentali danno dunque luogo ad una particolare determinazione delle variabili di interazione $\mathbf{w}(\omega) = w$ e queste a loro volta ad un’osservazione y legata al valore incognito x assunto da \mathbf{x} , dalla relazione

$$y = h(x, w) \quad , \quad (1.6)$$

In questo schema, il *problema della stima puntuale Bayesiana* è quello di ricostruire la determinazione x del vettore aleatorio \mathbf{x} che ha “provocato” all’atto della misura l’osservazione $\mathbf{y} = y$.

Dato che w è intrinsecamente difficile (in effetti impossibile) da conoscere, x non può mai essere ricavata esattamente dalla (1.6) e si vede bene che questo problema va posto come un *problema di approssimazione*. In tutti i problemi di questa natura, si richiede la definizione preliminare di un criterio con cui giudicare la bontà dell’approssimazione. Una classe “ragionevole” di criteri di approssimazione è definita qui sotto.

DEFINIZIONE 2.1. *Chiameremo “funzione costo” (o “perdita”) una funzione scalare $c : \mathbb{R}^n \rightarrow \mathbb{R}$ della variabile $\xi = z - x$, $z \in \mathbb{R}^n$, che è strettamente convessa, non-negativa e nulla nell’origine. Una funzione costo è simmetrica, se $c(-\xi) = c(\xi)$.*

Un esempio semplice di funzione costo simmetrica è la forma quadratica

$$c(z - x) = \|z - x\|_Q^2 \quad , \quad (1.7)$$

dove $\|x\|_Q^2 := x'Qx$ e Q è una matrice simmetrica e definita positiva.

Il vettore z gioca qui il ruolo di *approssimante* di x . Dato che x è incognito, sembra necessario cercare un vettore z che minimizzi c per tutti i possibili valori del parametro incognito x .

In realtà l’informazione probabilistica presente nel problema rende possibile un approccio meno brutale. Sappiamo infatti che certi valori di x sono a priori più probabili di altri; meglio ancora, la misura $\mathbf{y} = y$ ci dice addirittura quanto più probabili sono certi valori di x per effetto dell’osservazione $\mathbf{y} = y$. Bisognerà quindi “pesare” la variabile x mediante la densità a posteriori, $f(x | y)$.

È naturale allora introdurre il *rischio atteso condizionato*,

$$R(z, y) := E [c(z - \mathbf{x}) | \mathbf{y} = y] = \int_{\mathbb{R}^n} c(z - x) f(x | y) dx \quad , \quad (1.8)$$

e definire la *stima (puntuale) Bayesiana* di x , quel vettore $z = \hat{x}$ che minimizza, rispetto a z , il rischio atteso $R(z, y)$,

$$\hat{x} = \underset{z}{\text{Arg min}} R(z, y) \quad . \quad (1.9)$$

Abbiamo data per scontata l’esistenza e l’unicità del minimo, che dimostreremo in un caso particolare qui sotto. Naturalmente \hat{x} dipende dalla scelta della funzione costo c . Questa dipendenza è però assai poco critica almeno in una larga classe di problemi di stima.

TEOREMA 2.1. *Se la funzione costo c è simmetrica e la densità a posteriori $f(\cdot | y)$ è anch'essa simmetrica attorno alla sua media $\mu(y) := E(\mathbf{x} | \mathbf{y} = y)$, allora la stima puntuale \hat{x} definita dalla (1.9) non dipende dalla funzione costo e coincide con la media condizionata del vettore \mathbf{x} dato $\mathbf{y} = y$.*

DIMOSTRAZIONE. Supponiamo inizialmente che $\mu(y) = E(\mathbf{x} | \mathbf{y} = y) = 0$, per cui la simmetria di f si scrive $f(x | y) = f(-x | y)$. Da questa relazione segue che

$$R(z, y) = E[c(z - \mathbf{x}) | \mathbf{y} = y] = E[c(z + \mathbf{x}) | \mathbf{y} = y] = E[c(\mathbf{x} - z) | \mathbf{y} = y]$$

l'ultima delle quali segue dalla simmetria di c . Quindi vale anche la

$$R(z, y) = E\left[\frac{c(\mathbf{x} - z) + c(\mathbf{x} + z)}{2} | \mathbf{y} = y\right] > E[c(\mathbf{x}) | \mathbf{y} = y] = R(0, y), \quad z \neq 0$$

per la convessità stretta. Eppertanto $\mu(y) = 0 = \text{Arg min}_z R(z, y)$.

Nel caso generale $\mu(y) \neq 0$, posto $\Delta \mathbf{x} := \mathbf{x} - \mu(y)$, $\Delta z := z - \mu(y)$ si ha $E[\Delta \mathbf{x} | \mathbf{y} = y] = 0$ e $\Delta z - \Delta \mathbf{x} = z - \mathbf{x}$ per cui

$$R(z, y) = E[c(\Delta z - \Delta \mathbf{x}) | \mathbf{y} = y] > E[c(\Delta \mathbf{x}) | \mathbf{y} = y] = R(\mu(y), y), \quad z \neq \mu(y)$$

ovvero $\mu(y) = \text{Arg min}_z R(z, y)$. \square

Notiamo che, una volta appurato il fatto che si possa usare per c una qualunque funzione costo simmetrica, ad esempio la distanza Euclidea

$$c(z, x) = \|z - x\|^2, \quad (1.10)$$

l'enunciato del teorema segue facilmente in base a una ben nota caratterizzazione variazionale della media di una distribuzione di probabilità che è richiamata nel problema seguente.

PROBLEMA 2.1. *Dimostrare che per una funzione costo quadratica come la (1.10) e, più in generale, come la (1.7), si ha $\hat{x} = E(\mathbf{x} | \mathbf{y} = y)$, qualunque sia la densità condizionata $f(x | y)$ purché, ovviamente, la media condizionata abbia senso. Studiare prima il caso in cui \mathbf{x} e \mathbf{y} sono scalari. Si veda a questo proposito anche il teorema 2.2 che segue.* \diamond

La stima \hat{x} che minimizza il rischio atteso corrispondente al costo quadratico (1.10) viene spesso chiamata anche *stima (Bayesiana) ai minimi quadrati* di x . Questa terminologia è un poco ambigua e in letteratura si tendono a denominare troppe cose col termine “minimi quadrati”. Noi non la useremo mai in questo contesto probabilistico.

Dato che si ha sempre unicità del minimo in (1.9) al variare del campione osservato della variabile di misura, y , la stima puntuale \hat{x} di x si può interpretare come il valore assunto da una *funzione* $\hat{\mathbf{x}} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ che viene chiamata *stimatore Bayesiano (a rischio minimo) di \mathbf{x} in base a \mathbf{y}* .

Uno *stimatore* è in generale una *funzione dei dati* che può evidentemente essere interpretato come un procedimento di calcolo (algoritmo) che trasforma i dati di misura in stime puntuali secondo uno schema a blocchi del tipo

In questo spirito, denoteremo sempre con il simbolo $E(\mathbf{x} | \mathbf{y})$, da leggersi “la media condizionata di \mathbf{x} dato \mathbf{y} ”, la *funzione* definita su \mathbb{R}^m dalla assegnazione

$$E(\mathbf{x} | \mathbf{y}) : y \rightarrow E(\mathbf{x} | \mathbf{y} = y) \quad ;$$

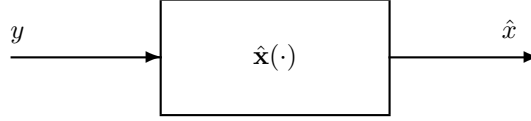


FIGURA 1.1. Stimatore.

Per funzioni costo e densità a posteriori simmetriche (Teorema 2.1) questa funzione dei dati è lo stimatore Bayesiano a rischio minimo di \mathbf{x} in base a \mathbf{y} . In realtà le proprietà della media condizionata possono essere caratterizzate meglio nel modo seguente.

TEOREMA 2.2. *La media condizionata $E(\mathbf{x} | \mathbf{y})$ è, tra tutte le funzioni misurabili $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$, quella che minimizza la media del quadrato della distanza $\|\mathbf{x} - g(\mathbf{y})\|_Q$ fra $g(\mathbf{y})$ e il vettore aleatorio \mathbf{x} ; in altre parole,*

$$E(\mathbf{x} | \mathbf{y}) = \text{Arg} \min_{g(\cdot)} E\|\mathbf{x} - g(\mathbf{y})\|_Q^2 \quad (1.11)$$

qualunque sia la matrice simmetrica definita positiva $Q \in \mathbb{R}^{n \times n}$.

DIMOSTRAZIONE. Notiamo innanzitutto che si può scrivere

$$\begin{aligned} E\|\mathbf{x} - g(\mathbf{y})\|_Q^2 &= \int_{\mathbb{R}^m} p_{\mathbf{y}}(y) E\|\mathbf{x} - g(\mathbf{y})\|_Q^2 | \mathbf{y} = y \, dy \\ &= E\{E\|\mathbf{x} - g(\mathbf{y})\|_Q^2 | \mathbf{y}\} \end{aligned} \quad (1.12)$$

e l'ultimo membro di questa uguaglianza si può scrivere in una forma particolarmente significativa se si usa l'identità

$$\begin{aligned} &E\{\|\mathbf{x} - E(\mathbf{x} | \mathbf{y}) + E(\mathbf{x} | \mathbf{y}) - g(\mathbf{y})\|_Q^2 | \mathbf{y}\} \\ &= E\{[\mathbf{x} - E(\mathbf{x} | \mathbf{y})]' Q[\mathbf{x} - E(\mathbf{x} | \mathbf{y})] | \mathbf{y}\} \\ &+ 2E\{[\mathbf{x} - E(\mathbf{x} | \mathbf{y})]' Q[E(\mathbf{x} | \mathbf{y}) - g(\mathbf{y})] | \mathbf{y}\} \\ &+ E\{[E(\mathbf{x} | \mathbf{y}) - g(\mathbf{y})]' Q[E(\mathbf{x} | \mathbf{y}) - g(\mathbf{y})] | \mathbf{y}\} \quad . \end{aligned} \quad (1.13)$$

Il secondo addendo a secondo membro di questa identità è zero giacché

$$\begin{aligned} &E\{[\mathbf{x} - E(\mathbf{x} | \mathbf{y})]' Q[E(\mathbf{x} | \mathbf{y}) - g(\mathbf{y})] | \mathbf{y}\} \\ &= \{[E(\mathbf{x} | \mathbf{y}) - E(\mathbf{x} | \mathbf{y})]' Q[E(\mathbf{x} | \mathbf{y}) - g(\mathbf{y})]\} = 0 \quad . \end{aligned}$$

Prendendo allora la media a primo e secondo membro nella (1.13) si trova

$$E\|\mathbf{x} - g(\mathbf{y})\|_Q^2 = E\|\mathbf{x} - E(\mathbf{x} | \mathbf{y})\|_Q^2 + E\|E(\mathbf{x} | \mathbf{y}) - g(\mathbf{y})\|_Q^2 \quad ,$$

dove i due addendi a destra sono entrambi non negativi e il primo non dipende da g . Ne segue che il minimo si ottiene per $g(\mathbf{y}) = E(\mathbf{x} | \mathbf{y})$. \square

Come appare dalla dimostrazione, se Q è solamente semidefinita positiva la proprietà di minimo sussiste ancora ma il minimo non è più unico.

Se restringiamo a priori la classe degli stimatori g a quelli cosiddetti *corretti*¹, aventi media $Eg(\mathbf{y})$ uguale a $E\mathbf{x}$, e che garantiscono quindi un errore di stima $\mathbf{x} - g(\mathbf{y})$ a media nulla, l'errore quadratico medio (1.12) si può interpretare, per $Q = I$, come la *varianza scalare dell' "errore di stima"* $\mathbf{x} - g(\mathbf{y})$. È ovvio che la media condizionata è uno stimatore corretto. D' altra parte la correttezza di g è un'ovvia condizione necessaria per la minimizzazione dell'errore quadratico medio, come mostrato nel seguente esercizio.

PROBLEMA 2.2. *Si ponga $\bar{\mathbf{x}} = \mathbf{x} - E\mathbf{x}$ e $\bar{g}(\mathbf{y}) := g(\mathbf{y}) - Eg(\mathbf{y})$. Giustificare l'affermazione che la correttezza di g è condizione necessaria per l'ottimalità, usando l'identità*

$$E\|\mathbf{x} - g(\mathbf{y})\|_Q^2 = E\|\bar{\mathbf{x}} - \bar{g}(\mathbf{y})\|_Q^2 + \|E\mathbf{x} - Eg(\mathbf{y})\|_Q^2 \quad .$$

\diamond

Per questo motivo, diremo in seguito che la media condizionata di \mathbf{x} dato \mathbf{y} è lo stimatore che ha la *minima varianza d'errore nella classe delle funzioni misurabili di \mathbf{y}* che soddisfano alla condizione di correttezza $Eg(\mathbf{y}) = E\mathbf{x}$.

Lo stimatore M.A.P. È abbastanza scontato che la definizione di stimatore puntuale come la funzione dei dati osservati che *minimizza* una funzione rischio possa essere sostituita da una, del tutto equivalente, di funzione dei dati che *massimizza* una funzione *guadagno atteso* del tipo

$$G(z, y) := E[\gamma(z - \mathbf{x}) | \mathbf{y} = y] = \int_{\mathbb{R}^n} \gamma(z - x) f(x | y) dx \quad , \quad (1.14)$$

dove adesso $\gamma : \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione della variabile $\xi = z - x$, $z \in \mathbb{R}^n$, che è *concava*, e massima nell'origine. D'altra parte, se si prende γ non negativa a forma di picco centrato nell'origine e nulla al di fuori di un piccolo intorno di $\xi = 0$, si può approssimare arbitrariamente bene una funzione δ di Dirac e ottenere una funzione guadagno atteso

$$G(z, y) \sim f(z | y)$$

uguale alla densità a posteriori corrispondente all'osservazione $\mathbf{y} = y$. Il corrispondente stimatore a guadagno atteso massimo

$$\hat{x}_{MAP}(y) := \underset{z}{\text{Arg max}} f(z | y) \quad (1.15)$$

è ispirato dal principio di rendere massima in $x = \hat{x}$ la probabilità a posteriori di \mathbf{x} data l'osservazione $\mathbf{y} = y$. Esso si chiama stimatore (del) *Massimo a Posteriori* (MAP) ed è molto usato in statistica Bayesiana.

Come si vede lo stimatore MAP è la *moda condizionata* della densità di probabilità di \mathbf{x} data l'osservazione $\mathbf{y} = y$. Naturalmente nel caso di densità unimodali e simmetriche la moda e la media coincidono e non si trova nulla di nuovo.

Da questi primi risultati si vede bene che la statistica Bayesiana, e in ispecie la *stima* Bayesiana, può essere vista come un *problema di calcolo delle probabilità*

¹Unbiased in inglese.

senza fare appello alcuno a ragionamenti di tipo induttivo. In effetti lo stimatore Bayesiano è (almeno nella grande maggioranza dei casi interessanti) una media condizionata e il calcolo della media condizionata è una questione che ha assai poco di “statistico” nel senso comune del termine.

Purtroppo però la media condizionata $E(\mathbf{x} \mid \mathbf{y})$ si riesce a calcolare esplicitamente solo in pochissimi casi. Uno particolarmente importante è quello in cui la distribuzione congiunta di \mathbf{x} e \mathbf{y} è *Gaussiana*.

TEOREMA 2.3. *Se \mathbf{x} e \mathbf{y} sono congiuntamente Gaussiani, ovvero se il vettore $n + m$ -dimensionale $\mathbf{z} = [\mathbf{x}', \mathbf{y}']'$ è Gaussiano, di media*

$$\mu_{\mathbf{z}} = \begin{bmatrix} \mu_{\mathbf{x}} \\ \mu_{\mathbf{y}} \end{bmatrix} \quad (1.16)$$

e varianza

$$\Sigma_{\mathbf{z}} = \begin{bmatrix} \Sigma_{\mathbf{x}} & \Sigma_{\mathbf{xy}} \\ \Sigma_{\mathbf{yx}} & \Sigma_{\mathbf{y}} \end{bmatrix}, \quad (1.17)$$

la densità condizionata di \mathbf{x} dato \mathbf{y} è ancora *Gaussiana*. Se $\Sigma_{\mathbf{y}} > 0$ la sua media e varianza (condizionate) sono date dalle espressioni

$$E(\mathbf{x} \mid \mathbf{y}) = \mu_{\mathbf{x}} + \Sigma_{\mathbf{xy}} \Sigma_{\mathbf{y}}^{-1} (\mathbf{y} - \mu_{\mathbf{y}}) \quad (1.18)$$

$$\text{Var}(\mathbf{x} \mid \mathbf{y}) = \Sigma_{\mathbf{x}} - \Sigma_{\mathbf{xy}} \Sigma_{\mathbf{y}}^{-1} \Sigma_{\mathbf{yx}}. \quad (1.19)$$

DIMOSTRAZIONE. Per semplicità di notazioni ragioneremo sul vettore ($n + m$ -dimensionale) $\bar{\mathbf{z}} := \mathbf{z} - \mu_{\mathbf{z}}$, che riscriveremo nella forma

$$\bar{\mathbf{z}} = \begin{bmatrix} \bar{\mathbf{x}} \\ \bar{\mathbf{y}} \end{bmatrix}.$$

Ovviamente $\bar{\mathbf{z}} \sim N(0, \Sigma_{\mathbf{z}})$ con $\Sigma_{\mathbf{z}}$ data dalla (1.17).

Prendiamo una trasformazione lineare delle variabili $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ definita dalla:

$$\begin{cases} \tilde{\mathbf{x}} = \bar{\mathbf{x}} + A\bar{\mathbf{y}} \\ \tilde{\mathbf{y}} = \bar{\mathbf{y}} \end{cases} \quad (1.20)$$

e scegliamo la matrice $n \times m$, A , in modo tale che $\tilde{\mathbf{x}}$ e $\tilde{\mathbf{y}}$ siano *scorrelati*. Imponendo che $E \tilde{\mathbf{x}}\tilde{\mathbf{y}}' = 0$, si trova

$$0 = \Sigma_{\mathbf{xy}} + A \Sigma_{\mathbf{y}}$$

e se $\Sigma_{\mathbf{y}} > 0$, questa equazione ha l'unica soluzione

$$A = -\Sigma_{\mathbf{xy}} \Sigma_{\mathbf{y}}^{-1}. \quad (1.21)$$

Chiaramente il vettore $\tilde{\mathbf{z}} := [\tilde{\mathbf{x}}' \tilde{\mathbf{y}}']'$ è ancora Gaussiano a media zero e varianza

$$\Sigma_{\tilde{\mathbf{z}}} = \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} \Sigma_{\mathbf{z}} \begin{bmatrix} I & 0 \\ A' & I \end{bmatrix} = \begin{bmatrix} \Sigma_{\tilde{\mathbf{x}}} & 0 \\ 0 & \Sigma_{\mathbf{y}} \end{bmatrix}.$$

Inoltre, per una nota proprietà della distribuzione Gaussiana, le componenti $\tilde{\mathbf{x}}$ e $\tilde{\mathbf{y}}$ ($= \bar{\mathbf{y}}$), essendo scorrelate, sono anche indipendenti, per cui

$$p_{\tilde{\mathbf{z}}}(\tilde{\mathbf{z}}) = p_{\tilde{\mathbf{x}}}(\tilde{\mathbf{x}}) p_{\tilde{\mathbf{y}}}(\tilde{\mathbf{y}}) = p_{\tilde{\mathbf{x}}}(\tilde{\mathbf{x}}) p_{\bar{\mathbf{y}}}(\tilde{\mathbf{y}}).$$

Dalla prima delle (1.20) riscritta come $\bar{\mathbf{x}} = \tilde{\mathbf{x}} - A\bar{\mathbf{y}}$ e dall'indipendenza di $\tilde{\mathbf{x}}$ da $\bar{\mathbf{y}}$ segue allora che la media condizionata di $\bar{\mathbf{x}}$ dato $\bar{\mathbf{y}}$ è

$$E(\bar{\mathbf{x}} | \bar{\mathbf{y}}) = -A\bar{\mathbf{y}} = \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \bar{\mathbf{y}} \quad (1.22)$$

A questo risultato si può ovviamente arrivare anche calcolando la densità condizionata di $\bar{\mathbf{x}}$ dato $\bar{\mathbf{y}} = y$ mediante la classica formula di Bayes. Per fare questo notiamo innanzitutto che lo Jacobiano della trasformazione (1.20) è una matrice triangolare superiormente con diagonale l'identità e pertanto il suo determinante è unitario. Applicando le note regole per il calcolo della densità di una funzione di variabili casuali si trova allora

$$p_{\bar{\mathbf{z}}}(z) = p_{\tilde{\mathbf{z}}}(\tilde{z}) \Big|_{\substack{\tilde{x} = x + Ay \\ \tilde{y} = y}} = p_{\tilde{\mathbf{x}}}(x + Ay) p_{\tilde{\mathbf{y}}}(y)$$

e la densità condizionata di $\bar{\mathbf{x}}$ dato $\bar{\mathbf{y}} = y$ è quindi $p_{\tilde{\mathbf{x}}}(x + Ay)$. Questa densità è Gaussiana, di media $-Ay$ e varianza $\Sigma_{\tilde{\mathbf{x}}}$. Si ritrova così la formula (1.22) e l'espressione per la varianza condizionata

$$\begin{aligned} \text{Var}(\bar{\mathbf{x}} | \bar{\mathbf{y}}) = \Sigma_{\tilde{\mathbf{x}}} &= E(\bar{\mathbf{x}} + A\bar{\mathbf{y}})(\bar{\mathbf{x}} + A\bar{\mathbf{y}})' \\ &= E\tilde{\mathbf{x}}\tilde{\mathbf{x}}' + EA\bar{\mathbf{y}}\bar{\mathbf{y}}' \\ &= \Sigma_{\tilde{\mathbf{x}}} - \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \Sigma_{\mathbf{y}\mathbf{x}} \quad . \end{aligned} \quad (1.23)$$

Introducendo i valori medi in (1.22) si trova infine la (1.18). \square

Se $\Sigma_{\mathbf{y}}$ è singolare si possono ancora dare delle espressioni esplicite per la media e la varianza condizionata di \mathbf{x} , ma in questo caso occorre usare la "pseudoinversa" di $\Sigma_{\mathbf{y}}$. Questo argomento verrà affrontato più avanti; si veda la Proposizione 2.6.

Se si interpreta $E(\mathbf{x} | \mathbf{y})$ come lo stimatore Bayesiano di \mathbf{x} in base a \mathbf{y} , allora il vettore $\tilde{\mathbf{x}}$ introdotto nel corso della dimostrazione è l'*errore residuo di stima*

$$\tilde{\mathbf{x}} := \bar{\mathbf{x}} - \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \bar{\mathbf{y}} = \mathbf{x} - E(\mathbf{x} | \mathbf{y}) \quad . \quad (1.24)$$

Questa quantità gode allora della proprietà notevole di essere *indipendente dai dati* \mathbf{y} . Intuitivamente, $\tilde{\mathbf{x}}$ è quanto resta di \mathbf{x} dopo averne sottratto la migliore approssimazione possibile basata sulla conoscenza di \mathbf{y} . L'indipendenza esprime il fatto che i dati non contengono più alcuna informazione utile agli effetti della ricostruzione del "residuo di stima" $\tilde{\mathbf{x}} = \mathbf{x} - E(\mathbf{x} | \mathbf{y})$. In altri termini, *i dati sono stati sfruttati nel modo migliore possibile*.

L'indipendenza spiega inoltre come mai la varianza condizionata di \mathbf{x} dato \mathbf{y} non dipenda dai dati. Questa varianza condizionata coincide infatti con quella (incondizionata) dell'errore di stima, $\Sigma_{\tilde{\mathbf{x}}}$, proprio grazie all'indipendenza di $\tilde{\mathbf{x}}$ e \mathbf{y} .

La *varianza d'errore*, espressa dalla formula (1.19), ovvero

$$\Sigma_{\tilde{\mathbf{x}}} = \Sigma_{\mathbf{x}} - \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \Sigma_{\mathbf{y}\mathbf{x}} \quad , \quad (1.25)$$

può essere interpretata come la *differenza tra la varianza a priori*, $\Sigma_{\mathbf{x}}$ di \mathbf{x} , e *quella dello stimatore Bayesiano* $E(\mathbf{x} | \mathbf{y})$. Dalla (1.18) (o dalla (1.22)) segue infatti agevolmente che

$$\text{Var} E(\mathbf{x} | \mathbf{y}) = \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \Sigma_{\mathbf{y}\mathbf{x}} \quad . \quad (1.26)$$

La (1.26) rappresenta dunque la *riduzione nell'incertezza a priori* sul valore assunto da \mathbf{x} introdotta dallo stimatore. Ne segue che uno stimatore è tanto più efficiente quanto più la varianza (1.26) è prossima a $\Sigma_{\mathbf{x}}$.

2. Stimatori Lineari a Minima Varianza d'Errore

Dati due vettori aleatori \mathbf{x} e \mathbf{y} , lo stimatore (corretto) a minima varianza (d'errore) di \mathbf{x} in base a \mathbf{y} ha, come abbiamo visto, una forma particolarmente semplice se \mathbf{x} e \mathbf{y} sono Gaussiani. In questo caso infatti $E(\mathbf{x} | \mathbf{y})$ è *lineare* nelle osservazioni e si può calcolare a partire dai momenti congiunti del primo e secondo ordine di \mathbf{x} e \mathbf{y} .

Questa semplicità scompare se l'ipotesi di Gaussianità viene meno. In questo caso $E(\mathbf{x} | \mathbf{y})$ è in generale una funzione *non lineare* delle osservazioni che si sa calcolare esplicitamente solo in pochissimi casi.

Dato che lo stimatore a minima varianza d'errore è la funzione $g(\cdot)$ che minimizza $E \|\mathbf{x} - g(\mathbf{y})\|^2$ nella classe (estremamente vasta!) di tutte le funzioni (misurabili) dei dati (che soddisfano la condizione di correttezza), è naturale chiedersi se restringendo a priori la classe delle funzioni g ammissibili non si ottengano degli stimatori più semplici da calcolare.

DEFINIZIONE 2.2. Si chiama *miglior stimatore lineare, o stimatore lineare a minima varianza (d'errore), la funzione lineare delle osservazioni*

$$g(\mathbf{y}) = A\mathbf{y} + b \quad , \quad A \in \mathbb{R}^{n \times m} \quad , \quad b \in \mathbb{R}^n$$

che minimizza la media del quadrato dell'errore di stima $E \|\mathbf{x} - g(\mathbf{y})\|^2$.

Lo stimatore lineare a m.v.² di \mathbf{x} si ottiene dunque risolvendo il problema

$$\min \begin{cases} A, b \\ A \in \mathbb{R}^{n \times m}, b \in \mathbb{R}^n \end{cases} E \|A\mathbf{y} + b - \mathbf{x}\|^2. \quad (2.1)$$

Se \mathbf{x} e \mathbf{y} hanno media nulla, certamente il valore di b per cui si ha il minimo è zero. In effetti se

$$g_*(\mathbf{y}) = A_*\mathbf{y} + b_*$$

fosse lo stimatore lineare a m.v. ed $E\mathbf{x} = 0$, $E\mathbf{y} = 0$, si avrebbe:

$$\begin{aligned} E \|A_*\mathbf{y} + b_* - \mathbf{x}\|^2 &= E \|A_*\mathbf{y} - \mathbf{x}\|^2 + 2E(A_*\mathbf{y} - \mathbf{x})' b_* + \|b_*\|^2 \\ &= E \|A_*\mathbf{y} - \mathbf{x}\|^2 + \|b_*\|^2 \geq E \|A_*\mathbf{y} - \mathbf{x}\|^2 \end{aligned}$$

e quindi $b_* = 0$ perché altrimenti $A_*\mathbf{y}$ sarebbe uno stimatore *migliore* di $A_*\mathbf{y} + b_*$.

Cerchiamo allora lo stimatore di $\bar{\mathbf{x}} = \mathbf{x} - \mu_{\mathbf{x}}$ in funzione di $\bar{\mathbf{y}} = \mathbf{y} - \mu_{\mathbf{y}}$ della forma $g(\bar{\mathbf{y}}) = A\bar{\mathbf{y}}$. Allo scopo dobbiamo considerare il problema di minimo

$$\min_A E \|A\bar{\mathbf{y}} - \bar{\mathbf{x}}\|^2 \quad . \quad (2.2)$$

Una volta risolto questo problema, il caso generale si riporta facilmente al caso di media nulla notando che

$$\min_{A,b} E \|A\mathbf{y} + b - \mathbf{x}\|^2 = \min_{A,b} E \|A\bar{\mathbf{y}} - \bar{\mathbf{x}} + c\|^2$$

dove $c = A\mu_{\mathbf{y}} + b - \mu_{\mathbf{x}}$. Dato che $E(A\bar{\mathbf{y}} - \bar{\mathbf{x}}) = 0$ il problema di minimizzazione si riduce a:

$$\min_{A,b} \left(E \|A\bar{\mathbf{y}} - \bar{\mathbf{x}}\|^2 + \|A\mu_{\mathbf{y}} + b - \mu_{\mathbf{x}}\|^2 \right) \quad .$$

²L'abbreviazione m.v. sta per "minima varianza".

Quindi se A_* minimizza il primo termine della somma, il vettore

$$b_* := \mu_{\mathbf{x}} - A_* \mu_{\mathbf{y}} \quad (2.3)$$

minimizza il secondo.

Formulazione geometrica e Teorema delle Proiezioni. Consideriamo la famiglia delle variabili aleatorie *scalari* (reali) definite su uno stesso spazio di probabilità $\{\Omega, \mathcal{A}, P\}$ (che si può pensare come lo spazio di probabilità relativo all'esperimento di misura che stiamo descrivendo) aventi

- (1) Media nulla,
- (2) Varianza finita.

Denotiamo col simbolo \mathbf{H} questo spazio e notiamo che \mathbf{H} è uno *spazio vettoriale* reale sul quale può essere definito un *prodotto interno* $\langle \cdot, \cdot \rangle_H$ tramite la

$$\langle \mathbf{x}, \mathbf{y} \rangle_H := E\mathbf{x}\mathbf{y} = \int_{\Omega} \mathbf{x}(\omega) \mathbf{y}(\omega) P(d\omega) = \int_{\mathbb{R}^2} xy p_{\mathbf{x}\mathbf{y}}(x, y) dx dy \quad (2.4)$$

il quale a sua volta definisce una *norma*

$$\|\mathbf{x}\|_H^2 = E\mathbf{x}^2 = \text{var}(\mathbf{x}) \quad , \quad \mathbf{x} \in \mathbf{H} \quad . \quad (2.5)$$

Se conveniamo di dichiarare uguale a zero una variabile aleatoria che è zero con probabilità uno, allora questa norma soddisfa le ben note proprietà che qualificano una norma in uno spazio metrico, in particolare la disuguaglianza triangolare e l'equivalenza tra $\|\mathbf{x}\|_H = 0$ e $\mathbf{x} = 0$, (quest'ultima eguaglianza va interpretata nel senso che \mathbf{x} è una variabile aleatoria “quasi certamente” nulla).

La convergenza di variabili aleatorie di \mathbf{H} rispetto alla norma $\|\cdot\|_H$ si chiama convergenza in *media quadratica*. Si può mostrare che \mathbf{H} è uno spazio *completo* rispetto alla convergenza in media quadratica ed è quindi uno *spazio di Hilbert* reale che verrà chiamato lo *spazio di Hilbert delle variabili del secondo ordine relativo all'esperimento* $\{\Omega, \mathcal{A}, P\}$. (Le definizioni fondamentali sugli spazi di Hilbert sono brevemente riportate all'inizio dell'Appendice B). L'introduzione di questa idea, dovuta essenzialmente a Kolmogorov [82], permette di impostare tutta la teoria della stima in termini geometrici, con enorme guadagno in chiarezza e semplicità concettuale.

Denoteremo la relazione di ortogonalità in \mathbf{H} , che evidentemente corrisponde a incorrelazione, col simbolo \perp .

Dato un vettore aleatorio \mathbf{y} a m componenti, il sottospazio vettoriale di \mathbf{H} generato dalle componenti $\mathbf{y}_i, i = 1, \dots, m$, cioè l'insieme di tutte le combinazioni lineari delle componenti scalari $\mathbf{y}_i, i = 1 \dots m$, di \mathbf{y} si denota col simbolo $\mathbf{H}(\mathbf{y})$,

$$\mathbf{H}(\mathbf{y}) := \text{span} \{\mathbf{y}_1 \dots \mathbf{y}_m\} = \left\{ \sum_{i=1}^m a_i \mathbf{y}_i ; a_i \in \mathbb{R} \right\} \quad . \quad (2.6)$$

Se le $\{\mathbf{y}_i\}$, pensate come funzioni di ω , sono linearmente indipendenti, $\mathbf{H}(\mathbf{y})$ ha dimensione m . Com'è noto dall'algebra lineare, questo accade allora e solo allora che la matrice gramiana,

$$G(\mathbf{y}) := [\langle \mathbf{y}_i, \mathbf{y}_j \rangle_H]_{ij=1 \dots m}$$

costruita con le componenti del vettore aleatorio \mathbf{y} , ha rango m . Equivalentemente, dato che $G(\mathbf{y})$ coincide con la *matrice varianza* $\Sigma_{\mathbf{y}}$, possiamo affermare che

LEMMA 2.1. *Le componenti scalari del vettore aleatorio \mathbf{y} sono linearmente indipendenti (equivalentemente, il sottospazio $\mathbf{H}(\mathbf{y})$ ha dimensione m) se e solo se $\Sigma_{\mathbf{y}}$ è una matrice definita positiva.*

Dunque l'ipotesi $\Sigma_{\mathbf{y}} > 0$ è equivalente al fatto che l'informazione contenuta in ciascuna componente del vettore \mathbf{y} non sia ricavabile dalle rimanenti, cioè che non vi sia ridondanza nei dati.

Notiamo che se \mathbf{y} è un vettore di dimensione finita, $\mathbf{H}(\mathbf{y})$ è ovviamente un sottospazio *chiuso* dello spazio di Hilbert \mathbf{H} .

Motivati dal problema (2.2) consideriamo ora la seguente questione.

PROBLEMA 2.3. *Data una variabile casuale scalare \mathbf{x} , trovare la funzione lineare $a'\mathbf{y}$, del vettore aleatorio \mathbf{y} , che ha minima distanza, in \mathbf{H} , da \mathbf{x} . Equivalentemente, trovare l'elemento $\mathbf{z} \in \mathbf{H}(\mathbf{y})$ per cui la norma*

$$\|\mathbf{x} - \mathbf{z}\|_H^2 = E[\mathbf{x} - \mathbf{z}]^2$$

è minima.

La soluzione di questo problema è fornita dal seguente, fondamentale, *Teorema delle proiezioni*.

TEOREMA 2.4. *Sia $\mathbf{H}(\mathbf{y})$ un sottospazio chiuso di \mathbf{H} . La variabile casuale $\mathbf{z}_* \in \mathbf{H}(\mathbf{y})$ che minimizza $\|\mathbf{x} - \mathbf{z}\|_H^2$ è unica ed è la proiezione ortogonale di \mathbf{x} su $\mathbf{H}(\mathbf{y})$.*

Condizione necessaria e sufficiente perché \mathbf{z}_ sia la proiezione ortogonale di \mathbf{x} su $\mathbf{H}(\mathbf{y})$ è che sia*

$$\mathbf{x} - \mathbf{z}_* \perp \mathbf{H}(\mathbf{y}) \quad (2.7)$$

(“principio di ortogonalità”), che equivale alla condizione

$$E(\mathbf{x} - \mathbf{z}_*) \mathbf{y}_i = 0 \quad \text{per} \quad i = 1, 2, \dots \quad (2.8)$$

dove $\mathbf{y}_i, i = 1, 2, \dots$ sono generatori di $\mathbf{H}(\mathbf{y})$.

DIMOSTRAZIONE. Mostriamo che $\mathbf{z}_* \in \mathbf{H}(\mathbf{y})$ minimizza $\|\mathbf{z} - \mathbf{x}\|_H^2$ se e solo se vale la (2.7). Sia \mathbf{z} un qualunque elemento di $\mathbf{H}(\mathbf{y})$; chiaramente $\mathbf{z}_* - \mathbf{z}$ appartiene ancora ad $\mathbf{H}(\mathbf{y})$, perciò, se vale la (2.7), si ha:

$$\begin{aligned} \|\mathbf{x} - \mathbf{z}\|_H^2 &= \|\mathbf{x} - \mathbf{z}_* + \mathbf{z}_* - \mathbf{z}\|_H^2 = \|\mathbf{x} - \mathbf{z}_*\|_H^2 \\ &\quad + 2\langle \mathbf{x} - \mathbf{z}_*, \mathbf{z}_* - \mathbf{z} \rangle_H + \|\mathbf{z}_* - \mathbf{z}\|_H^2 \\ &= \|\mathbf{x} - \mathbf{z}_*\|_H^2 + \|\mathbf{z}_* - \mathbf{z}\|_H^2 \end{aligned} \quad (2.9)$$

e il minimo di questa espressione si raggiunge ovviamente per $\mathbf{z} = \mathbf{z}_*$. Supponiamo viceversa che $\mathbf{z}_* \in \mathbf{H}(\mathbf{y})$ minimizzi $\|\mathbf{x} - \mathbf{z}\|_H^2$. Definiamo la variabile

$$\mathbf{z}_i := \mathbf{z}_* + \frac{\langle \mathbf{x} - \mathbf{z}_*, \mathbf{y}_i \rangle_H}{\|\mathbf{y}_i\|_H^2} \mathbf{y}_i \quad .$$

Si noti che $\mathbf{z} \in \mathbf{H}(\mathbf{y})$ e pertanto $\|\mathbf{x} - \mathbf{z}_i\|_H^2 \geq \|\mathbf{x} - \mathbf{z}_*\|_H^2$. Calcolando $\|\mathbf{x} - \mathbf{z}_i\|_H^2$ si trova

$$\begin{aligned} \|\mathbf{x} - \mathbf{z}_i\|_H^2 &= \left\| (\mathbf{x} - \mathbf{z}_*) - \frac{\langle \mathbf{x} - \mathbf{z}_*, \mathbf{y}_i \rangle_H}{\|\mathbf{y}_i\|_H^2} \mathbf{y}_i \right\|_H^2 \\ &= \|\mathbf{x} - \mathbf{z}_*\|_H^2 - \frac{\langle \mathbf{x} - \mathbf{z}_*, \mathbf{y}_i \rangle_H^2}{\|\mathbf{y}_i\|_H^2} \end{aligned}$$

e per l'osservazione precedente si deve necessariamente avere

$$\langle \mathbf{x} - \mathbf{z}_*, \mathbf{y}_i \rangle_H = 0 \quad .$$

Notiamo che dalla (2.9) segue immediatamente anche l'unicità di \mathbf{z}_* . Se infatti $\mathbf{z}_1 \in \mathbf{H}(\mathbf{y})$ minimizza $\|\mathbf{x} - \mathbf{z}\|_H^2$, ne segue che $\|\mathbf{x} - \mathbf{z}_1\|_H^2 = \|\mathbf{x} - \mathbf{z}_*\|_H^2$ e la (2.9) porge (per $\mathbf{z} = \mathbf{z}_1$)

$$\|\mathbf{z}_* - \mathbf{z}_1\|_H^2 = 0$$

ovvero $\mathbf{z}_* = \mathbf{z}_1$. □

OSSERVAZIONE 2.1. È importante, in vista di future estensioni, notare che gli argomenti usati nella prova sono indipendenti dalla dimensione di $\mathbf{H}(\mathbf{y})$. Il principio di ortogonalità continua in particolare a valere anche nel caso in cui $\dim \mathbf{H}(\mathbf{y}) = \infty$, cioè nel caso in cui $\mathbf{H}(\mathbf{y})$ è generato da una arbitraria famiglia (anziché un numero finito) di variabili casuali $\{\mathbf{y}_\alpha\}$. In questo caso però è bene puntualizzare che l'esistenza dell'elemento $\mathbf{z}_* \in \mathbf{H}(\mathbf{y})$ a distanza minima da \mathbf{x} si può garantire solo se il sottospazio $\mathbf{H}(\mathbf{y})$ è un sottospazio *chiuso* (completo) di \mathbf{H} . Si veda ad esempio il Teorema 1, a pag. 23, in Halmos [61].

La proprietà di completezza, che è automatica nel caso di dimensioni finite, per $m = \infty$ richiede che il sottospazio $\mathbf{H}(\mathbf{y})$ venga definito in modo appropriato ed è inevitabile l'introduzione di un certo apparato matematico (dalla teoria degli spazi di Hilbert) che nel presente contesto ($m < \infty$) può essere evitata. Nel caso che ci interessa qui, l'esistenza della variabile \mathbf{z}_* a distanza minima da \mathbf{x} verrà provata costruttivamente fra poco. ◇

Il teorema ha una intuitiva interpretazione geometrica che è illustrata nella figura 2.1.

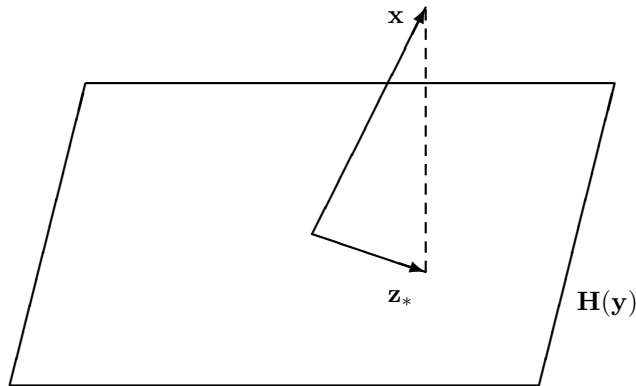


FIGURA 2.1. Proiezione ortogonale.

Imponendo la condizione di ortogonalità (2.8) del Teorema 2.4 si vede subito che la variabile di $\mathbf{H}(\mathbf{y})$,

$$\mathbf{z} = \sum_1^m a_i \mathbf{y}_i$$

è la proiezione di \mathbf{x} su $\mathbf{H}(\mathbf{y})$ se e solo se il vettore $a = [a_1 \dots a_m]'$ è soluzione del sistema di equazioni lineari

$$E\mathbf{x}\mathbf{y}_i = E a' \mathbf{y}_i \mathbf{y}_i \quad , \quad i = 1 \dots m$$

ovvero

$$a' E\mathbf{y}\mathbf{y}' = E\mathbf{x}\mathbf{y}'$$

che si può riscrivere

$$a' \Sigma_{\mathbf{y}} = \sigma_{\mathbf{x}\mathbf{y}} \quad (2.10)$$

dove $\sigma_{\mathbf{x}\mathbf{y}} = [\sigma_{\mathbf{x}\mathbf{y}_1}, \dots, \sigma_{\mathbf{x}\mathbf{y}_m}]$ è il vettore riga delle covarianze di \mathbf{x} e \mathbf{y} .

Se $\Sigma_{\mathbf{y}} > 0$ la (2.10) può essere risolta ottenendo

$$a'_* = \sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \quad (2.11)$$

per cui lo stimatore a m.v. di \mathbf{x} in base al vettore aleatorio \mathbf{y} è dato dalla formula

$$\mathbf{z}_* = a'_* \mathbf{y} = \sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \mathbf{y} \quad . \quad (2.12)$$

Se \mathbf{x} invece di essere una variabile scalare è un *vettore* aleatorio n -dimensionale, il problema diventa quello di trovare n elementi $\mathbf{z}_i \in \mathbf{H}(\mathbf{y})$, $i = 1 \dots n$ (ovvero un vettore n -dimensionale \mathbf{z} a componenti in $\mathbf{H}(\mathbf{y})$) tali che

$$\sum_{i=1}^n \|\mathbf{z}_i - \mathbf{x}_i\|_H^2 = E \left(\sum_1^n (\mathbf{z}_i - \mathbf{x}_i)^2 \right) = E \|\mathbf{z} - \mathbf{x}\|^2 \quad (2.13)$$

sia minima. Infatti \mathbf{z} minimizza l'ultimo membro se e solo se le componenti \mathbf{z}_i minimizzano separatamente le distanze $\|\mathbf{z}_i - \mathbf{x}_i\|_H^2$, per cui, se si usa una notazione matriciale, del tipo

$$\mathbf{z} = A\mathbf{y} \quad , \quad A \in \mathbb{R}^{n \times m}$$

le righe a'_{i*} di A_* sono ottenibili immediatamente dalla (2.11) come

$$a'_{i*} = \sigma_{\mathbf{x}_i\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \quad , \quad i = 1 \dots n$$

che si può in definitiva riscrivere nella forma

$$A_* = \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1}$$

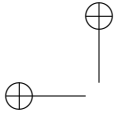
ovvero

$$\mathbf{z}_* = \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \mathbf{y} \quad . \quad (2.14)$$

A posteriori, notiamo la seguente importante proprietà.

PROPOSIZIONE 2.1. *Il miglior stimatore lineare di \mathbf{x} in base a \mathbf{y} dipende solo dai momenti (congiunti) del primo e del secondo ordine di \mathbf{x} e \mathbf{y} . La sua espressione generale è identica a quella del caso Gaussiano (1.18), ovvero*

$$\hat{\mathbf{x}}(\mathbf{y}) = \mu_{\mathbf{x}} + \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} (\mathbf{y} - \mu_{\mathbf{y}}) \quad . \quad (2.15)$$



Questo risultato era prevedibile perché nel caso in cui \mathbf{x}, \mathbf{y} sono congiuntamente Gaussiani il *miglior stimatore in senso assoluto è lineare* e deve pertanto coincidere con quello che si ottiene restringendo la classe delle funzioni $g(\cdot)$ a quelle lineari.

Ne scende che anche la varianza dell'errore di stima $\tilde{\mathbf{x}} := \mathbf{x} - \hat{\mathbf{x}}(\mathbf{y})$ è data dalla stessa espressione del caso Gaussiano,

$$\text{Var}(\tilde{\mathbf{x}}) = \Sigma_{\mathbf{x}} - \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \Sigma_{\mathbf{y}\mathbf{x}} \quad (2.16)$$

e si può calcolare a partire dai soli momenti del secondo ordine (congiunti) di \mathbf{x} e \mathbf{y} .

Nel seguito ci interesseremo quasi esclusivamente di *stimatori lineari* e pertanto faremo di norma riferimento a problemi in cui *l'informazione probabilistica* sul legame tra le osservazioni \mathbf{y} e l'incognita \mathbf{x} , così come la descrizione a priori di \mathbf{x} , siano specificate *unicamente tramite i momenti del primo e del secondo ordine*. È evidente che aggiungendo l'informazione addizionale di Gaussianità si ha *completa* specificazione della struttura probabilistica (dato che le distribuzioni Gaussiane sono completamente individuate dai momenti del primo e secondo ordine). In questo caso lo stimatore non è solo il migliore nella classe degli stimatori lineari, ma addirittura il migliore nella classe di tutte le possibili funzioni (non lineari) delle misure \mathbf{y} .

Il fatto che lo stimatore lineare a m.v. coincida con la media condizionata $E(\mathbf{x} | \mathbf{y})$ nel caso Gaussiano, rende particolarmente suggestiva una notazione del tipo

$$\hat{E}(\mathbf{x} | \mathbf{y}) \quad \text{oppure} \quad \hat{E}[\mathbf{x} | H(\mathbf{y})] \quad . \quad (2.17)$$

In effetti lo stimatore lineare a m.v. viene spesso chiamato *media condizionata in senso debole* di \mathbf{x} dato \mathbf{y} . L'analogia è suffragata dal fatto che la vera media condizionata $E(\mathbf{x} | \mathbf{y})$ può, anche nel caso in cui le distribuzioni congiunte non sono Gaussiane, essere definita come una proiezione su un opportuno spazio vettoriale costruito sui dati. Ne segue che le proprietà dei due operatori sono molto simili, almeno sul piano formale. Le proprietà di $\hat{E}(\mathbf{x} | \mathbf{y})$ sono in effetti le stesse di $E(\mathbf{x} | \mathbf{y})$ pur di sostituire alla parola "indipendenza" e "funzione misurabile dei dati" le rispettive versioni *deboli* e cioè "ortogonalità" (o incorrelazione) e "funzione lineare dei dati". È opportuno però mettere in guardia il lettore dal pericolo di confondere i due concetti e di arrivare a convincersi (almeno inconsciamente) che i due stimatori possano più o meno considerarsi equivalenti. In casi in cui le distribuzioni in gioco sono molto lontane dalla Gaussianità c'è viceversa da aspettarsi che "l'approssimazione lineare" $\hat{E}(\mathbf{x} | \mathbf{y})$ di $E(\mathbf{x} | \mathbf{y})$ sia di scarso significato. Si veda a questo proposito l'Esempio 2.4 riportato più sotto.

NOTAZIONI . A rigore lo spazio di Hilbert \mathbf{H} e ogni suo sottospazio $\mathbf{H}(\mathbf{y})$ sono composti di variabili aleatorie *scalari*. Onde evitare inutili giri di parole sarà opportuno stabilire delle convenzioni e notazioni da usare nella manipolazione di quantità vettoriali. D'ora innanzi converremo che notazioni del tipo $\mathbf{z} \in \mathbf{H}(\mathbf{y})$, $\mathbf{z} \perp \mathbf{H}(\mathbf{y})$ ecc., in cui \mathbf{z} è un *vettore* aleatorio, significhino che tutte le componenti $\{\mathbf{z}_k\}$ appartengono o sono ortogonali al sottospazio in questione ecc... Significato ovvio avrà allora anche il simbolo $\hat{E}(\mathbf{x} | \mathbf{y})$ nel caso di \mathbf{x} vettoriale, che interpreteremo come il vettore a componenti uguali alle proiezioni ortogonali $\hat{E}(\mathbf{x}_k | \mathbf{y})$ delle componenti scalari \mathbf{x}_k , $k = 1, \dots, n$.

Nell'enunciato che segue riformuleremo le proprietà dello stimatore lineare a m.v. in una forma a cui faremo spesso riferimento in seguito.

COROLLARIO 2.1. *Se \mathbf{x} e \mathbf{y} sono due vettori aleatori di dimensioni rispettive n ed m (a media nulla e varianza finita), lo stimatore lineare a m.v. $\hat{E}(\mathbf{x} | \mathbf{y})$ di \mathbf{x} in base a \mathbf{y} è quel vettore \mathbf{z} (a componenti) in $\mathbf{H}(\mathbf{y})$ tale per cui*

$$E(\mathbf{x} - \mathbf{z}) \mathbf{y}' = 0 \quad . \quad (2.18)$$

Inoltre lo stimatore $\hat{E}(\mathbf{x} | \mathbf{y})$ ha la minima matrice di varianza d'errore tra tutte le funzioni lineari dei dati, ovvero

$$E\{[\mathbf{x} - \hat{E}(\mathbf{x} | \mathbf{y})] [\mathbf{x} - \hat{E}(\mathbf{x} | \mathbf{y})]'\} \leq E[(\mathbf{x} - \mathbf{z}) (\mathbf{x} - \mathbf{z})'] \quad (2.19)$$

per ogni vettore aleatorio n -dimensionale $\mathbf{z} \in \mathbf{H}(\mathbf{y})$. La disuguaglianza $A \leq B \dots$ tra matrici simmetriche è da intendersi nel senso che $B - A$ è semidefinita positiva.

DIMOSTRAZIONE. La (2.18) è il principio di ortogonalità (2.8), ovvero la

$$E(\mathbf{x}_i - \mathbf{z}_i) \mathbf{y}_j = 0 \quad , \quad i = 1, \dots, n \quad j = 1, \dots, m$$

scritta in forma matriciale.

Per provare la (2.19) si scriva $\mathbf{z} = A\mathbf{y}$. Calcolando la varianza matriciale di $\mathbf{x} - \mathbf{z}$ si trova

$$\begin{aligned} \text{Var}(\mathbf{x} - \mathbf{z}) &= \text{Var}(\mathbf{x} - \hat{E}(\mathbf{x} | \mathbf{y}) + \hat{E}(\mathbf{x} | \mathbf{y}) - A\mathbf{y}) \\ &= (\Sigma_{\mathbf{x}} - \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \Sigma_{\mathbf{y}\mathbf{x}}) + (A - \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1}) \Sigma_{\mathbf{y}} (A' - \Sigma_{\mathbf{y}}^{-1} \Sigma_{\mathbf{y}\mathbf{x}}) \end{aligned}$$

ed evidentemente il minimo si raggiunge per $A = \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1}$. \square

Il modello lineare. In molte situazioni concrete il meccanismo fisico che lega le osservazioni \mathbf{y} al vettore da stimare \mathbf{x} può essere approssimato mediante una relazione lineare del tipo

$$\mathbf{y} = S\mathbf{x} + \mathbf{w} \quad (2.20)$$

dove S è una matrice nota, \mathbf{x} e \mathbf{w} sono vettori aleatori *scorrelati*, di varianze $P := \text{Var}(\mathbf{x})$ ed $R := \text{Var}(\mathbf{w})$ supposte note. Per semplicità di notazioni supporremo \mathbf{x} e \mathbf{w} a media nulla. Lo schema (2.20) rappresenta evidentemente un prototipo di misure prodotte da un sensore lineare in presenza di "rumore additivo" (\mathbf{w}).

Le matrici $\Sigma_{\mathbf{x}\mathbf{y}}$ e $\Sigma_{\mathbf{y}}$ che servono per calcolare lo stimatore lineare a minima varianza $\hat{E}(\mathbf{x} | \mathbf{y})$ sono calcolabili facilmente a partire dai parametri del modello (2.20). Innanzitutto, per l'ortogonalità di \mathbf{x} e \mathbf{w} , si trova che $\Sigma_{\mathbf{y}\mathbf{x}}$ è data dalla relazione

$$\Sigma_{\mathbf{y}\mathbf{x}} = S \Sigma_{\mathbf{x}} = SP$$

per cui

$$\Sigma_{\mathbf{x}\mathbf{y}} = PS' \quad . \quad (2.21)$$

Inoltre,

$$\Sigma_{\mathbf{y}} = SPS' + R \quad (2.22)$$

e se $\Sigma_{\mathbf{y}} > 0$, cosa che si verifica ad esempio se $R > 0$, lo stimatore di \mathbf{x} in base a \mathbf{y} è

$$\hat{E}(\mathbf{x} | \mathbf{y}) = PS'(SPS' + R)^{-1} \mathbf{y} \quad . \quad (2.23)$$

La varianza dell'errore di stima, Λ , si calcola poi a partire dall'espressione (2.16) e risulta essere

$$\Lambda = P - PS'(SPS' + R)^{-1} SP \quad . \quad (2.24)$$

OSSERVAZIONE 2.2. Ci si può domandare quanto rappresentativa sia la relazione lineare (2.20) della situazione generale a cui la teoria fa riferimento. In altri termini quale sottoclasse di problemi di stima lineare sia descrivibile in termini di un modello lineare del tipo (2.20).

È facile convincersi che un *qualunque* problema posto a partire da una *arbitraria* matrice di varianze e covarianze dei vettori \mathbf{x} e \mathbf{y} può essere riformulato come un problema di stima su un modello lineare.

Si può infatti sempre pensare di rappresentare \mathbf{y} come somma della sua proiezione ortogonale su $\mathbf{H}(\mathbf{x})$ e del termine d'errore $\tilde{\mathbf{y}} = \mathbf{y} - \hat{E}(\mathbf{y} | \mathbf{x})$,

$$\mathbf{y} = \hat{E}(\mathbf{y} | \mathbf{x}) + \tilde{\mathbf{y}} \quad (2.25)$$

dove $\hat{E}(\mathbf{y} | \mathbf{x})$ è una funzione lineare di \mathbf{x} . Se $\Sigma_{\mathbf{x}} > 0$ questa funzione può essere scritta nella forma $\hat{E}(\mathbf{y} | \mathbf{x}) = S\mathbf{x}$, con S data da:

$$S = \Sigma_{\mathbf{y}\mathbf{x}} \Sigma_{\mathbf{x}}^{-1} \quad .$$

Inoltre, identificando \mathbf{w} con $\tilde{\mathbf{y}}$ (che è scorrelato da \mathbf{x} per il principio di ortogonalità), la (2.25) è formalmente identica alla (2.20). La varianza, R , del termine di rumore è data dalla formula $R = \Sigma_{\mathbf{y}} - \Sigma_{\mathbf{y}\mathbf{x}} \Sigma_{\mathbf{x}}^{-1} \Sigma_{\mathbf{x}\mathbf{y}}$. \diamond

Torniamo ora alle espressioni (2.23) e (2.24). Torna comodo dare a queste formule un aspetto diverso, sia a scopo di analizzare l'influenza dei vari parametri in gioco sulla stima, sia a scopo di facilitare il calcolo dello stimatore in certe situazioni comuni nella pratica, ad esempio quando R è una matrice diagonale.

In seguito incontreremo spesso manipolazioni di questo tipo e tutte, come vedremo, saranno riconducibili a una fondamentale identità del calcolo matriciale nota sotto il nome di *lemma d'inversione di matrice*, il cui enunciato è riportato in Appendice A.

TEOREMA 2.5. *Si assuma che la varianza a priori del vettore \mathbf{x} nel modello lineare (2.20), sia invertibile. Allora lo stimatore lineare a varianza minima del vettore aleatorio \mathbf{x} , può esprimersi nella forma*

$$\hat{E}(\mathbf{x} | \mathbf{y}) = \mu_{\mathbf{x}} + (P^{-1} + S'R^{-1}S)^{-1} S'R^{-1}(\mathbf{y} - \mu_{\mathbf{y}}) \quad (2.26)$$

e la relativa matrice varianza d'errore, Λ , è data anche dalla formula

$$\Lambda = (P^{-1} + S'R^{-1}S)^{-1} \quad . \quad (2.27)$$

DIMOSTRAZIONE. La (2.27) si ottiene dalla (2.24) usando il lemma d'inversione di matrice (0.1) in appendice A con $A = P^{-1}$, $B = S'$, $C = R^{-1}$.

Per quanto riguarda la (2.26), usando ancora il lemma A.1 in appendice A, si ottiene

$$\begin{aligned}\hat{E}(\mathbf{x} | \mathbf{y}) &= \mu_{\mathbf{x}} + PS' [R^{-1} - R^{-1} S(S'R^{-1}S + P^{-1})^{-1} S'R^{-1}] (\mathbf{y} - \mu_{\mathbf{y}}) \\ &= \mu_{\mathbf{x}} + [P - PS'R^{-1} S(S'R^{-1}S + P^{-1})^{-1}] S'R^{-1} (\mathbf{y} - \mu_{\mathbf{y}}) \\ &= \mu_{\mathbf{x}} + [P(S'R^{-1}S + P^{-1}) - PS'R^{-1}S] (S'R^{-1}S + P^{-1})^{-1} S'R^{-1} (\mathbf{y} - \mu_{\mathbf{y}})\end{aligned}$$

e svolgendo i calcoli si vede facilmente che il termine fra parentesi quadre è l'identità. \square

Nelle formule (2.26) e (2.27) è molto evidente l'influenza della *varianza a priori di \mathbf{x}* . Se la conoscenza a priori su \mathbf{x} è molto incerta, P è grande e verosimilmente P^{-1} può essere trascurata rispetto all'addendo $S'R^{-1}S$. Se ne deduce che *in assenza di informazioni a priori su \mathbf{x}* lo stimatore a minima varianza si riduce a

$$\hat{E}(\mathbf{x} | \mathbf{y}) = \mu_{\mathbf{x}} + (S'R^{-1}S)^{-1} S'R^{-1} (\mathbf{y} - \mu_{\mathbf{y}}) \quad (2.28)$$

e la sua varianza a

$$\Lambda = (S'R^{-1}S)^{-1} \quad (2.29)$$

Queste formule sono in effetti esattamente quelle che danno la stima (di Markov) del "parametro" \mathbf{x} usando l'approccio Fisheriano. Si veda [116].

3. Esempi

ESEMPIO 2.1 (Trattamento statistico di misure ripetute). Supponiamo di eseguire n misure di una variabile fisica, ad esempio massa, lunghezza, tensione etc. supponendo che la variabile sia suscettibile di misurazione diretta. Supponiamo ad esempio di effettuare misure di resistenza su un resistore di valore nominale x_0 e tolleranza α (espressa in % del valore nominale), non necessariamente usando lo stesso metodo. Supponiamo che le misure siano eseguite in modo indipendente l'una dall'altra, ad esempio in diverse condizioni sperimentali o cambiando l'apparato di misura ad ogni prova.

Se indichiamo con \mathbf{y}_i la variabile che rappresenta l' i -sima misura e con \mathbf{x} il valore della resistenza, si può scrivere

$$\mathbf{y}_i = \mathbf{x} + \mathbf{w}_i \quad , \quad i = 1, \dots, n \quad (3.1)$$

dove \mathbf{w}_i sono variabili casuali indipendenti, di media zero (errori "accidentali") e varianza calcolabile in base a un'analisi di precisione del metodo. In genere questa precisione viene espressa in valore percentuale riferito ad un valore (nominale) di "fondo scala", x_M , dello strumento. Si individua in questo modo una fascia di variazione i cui limiti superiore e inferiore, $\pm\beta x_M$, (se β è la precisione dello strumento) danno un campo dentro il quale sta, con probabilità molto alta, il valore campionario dell'errore accidentale \mathbf{w} . Ipotizzando una distribuzione normale per \mathbf{w}_i e una probabilità del 95%, si può così ricavare la varianza, σ_i^2 , di \mathbf{w}_i usando la relazione

$$2\sigma_i = \beta_i x_M$$

dove β_i è la precisione dell' i -simo strumento di misura.

Con un ragionamento analogo si può poi ricavare la varianza *a priori* σ_x^2 di \mathbf{x} . Basta allo scopo ricavare σ_x dalla

$$2\sigma_x = \alpha x_0$$

dove α è la tolleranza.

In questo modo sono noti tutti i dati a priori del modello, sia quelli relativi alla variabile \mathbf{x} (che avrà media a priori $\mu_{\mathbf{x}} = x_0$ e varianza a priori $(\alpha x_0/2)^2$) sia quelli relativi alle variabili di rumore (di media zero, varianze σ_i^2 e scorrelate).

Il modello (3.1) si può dunque riscrivere in forma vettoriale come

$$\mathbf{y} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \mathbf{x} + \mathbf{w}$$

con $\mu_x = x_0$, $P = \sigma_x^2$ ed R diagonale,

$$R = \text{diag} \{ \sigma_1^2, \dots, \sigma_n^2 \} \quad .$$

Ovviamente supporremo (per ragioni fisiche) \mathbf{x} e \mathbf{w} scorrelate.

Applicando la (2.26) si trova

$$\hat{E}(\mathbf{x} | \mathbf{y}) = x_0 + \left(\frac{1}{\sigma_x^2} + \sum_1^n \frac{1}{\sigma_i^2} \right)^{-1} \sum_1^n \frac{1}{\sigma_i^2} (\mathbf{y}_i - x_0) \quad .$$

Se $\sigma_i^2 = \sigma^2$, $i = 1 \dots n$, cioè le n misure hanno tutte la stessa precisione, questa espressione si riduce alla

$$\hat{E}(\mathbf{x} | \mathbf{y}) = x_0 + \frac{1}{\frac{\sigma_x^2}{\sigma^2} + n} \sum_1^n (\mathbf{y}_i - x_0)$$

che è circa uguale alla media campionaria $\bar{\mathbf{y}}_n$ se $\sigma^2/\sigma_x^2 \ll n$.

Usando la (2.27) la varianza dell'errore di stima vale

$$\lambda^2 = \sigma_x^2 \left(1 + \sum_1^n \frac{\sigma_x^2}{\sigma_i^2} \right)^{-1} \quad .$$

Si vede che se σ_x^2 è grande, λ^2 è all'incirca uguale a

$$\lambda^2 \cong \left(\sum_1^n \frac{1}{\sigma_i^2} \right)^{-1}$$

ovvero a σ^2/n se le misure hanno tutte la stessa precisione.

Notiamo che in pratica il valore calcolato di λ^2 serve alla fine per esprimere l'incertezza nella stima, ad esempio presentando il risultato delle misure nella forma $x = \hat{x} \pm 2\lambda$, dove $\hat{x} = \hat{E}(\mathbf{x} | \mathbf{y} = y)$ è il valore stimato in base ai dati y . Questo modo di scrivere ha una chiara interpretazione nel caso in cui le variabili in gioco sono Gaussiane. Il valore della resistenza x sta allora con probabilità del 95% nell'intervallo $[\hat{x} - 2\lambda, \hat{x} + 2\lambda]$. (Il lettore è invitato a giustificare questa affermazione partendo dalla relazione $\mathbf{x} = \hat{E}(\mathbf{x} | \mathbf{y}) + \tilde{\mathbf{x}}$. \diamond)

ESEMPIO 2.2 (Stima di un ritardo). Tratteremo molto succintamente un problema di stima che nasce da applicazioni a sistemi radar/sonar. La variabile da stimare è il ritardo con cui si riceve l' "eco" di un segnale sinusoidale riflesso da un "target" a distanza incognita r dalla sorgente.

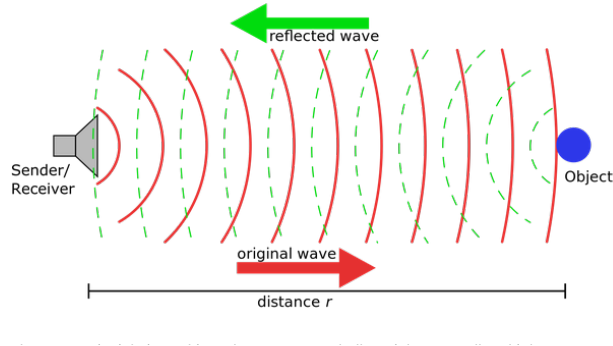


FIGURA 3.1. Onde Radar

La forma tipica di un segnale radar o sonar in ricezione, dopo rigenerazione e campionamento, è

$$\mathbf{y}(t) = A \sin(\omega_0 t + \boldsymbol{\theta}) + \mathbf{w}(t) \quad , \quad t = 1, \dots, N \quad (3.2)$$

dove

- A è un'ampiezza normalizzata nota;
- $\boldsymbol{\theta}$ è il ritardo di fase dell'eco (proporzionale alla distanza dell'ostacolo), che supporremo una variabile aleatoria uniformemente distribuita sull'intervallo $[0, 2\pi]$;
- $\mathbf{w}(t)$ è rumore additivo che supporremo bianco a media zero, indipendente da $\boldsymbol{\theta}$ e di varianza nota σ^2 .

Il problema è di stimare il ritardo $\boldsymbol{\theta}$. Questo problema non lineare può essere trattato con la teoria degli stimatori lineari a m.v. introducendo una variabile fittizia

$$\mathbf{x} = [\cos \boldsymbol{\theta} \quad \sin \boldsymbol{\theta}]' \quad (3.3)$$

e riscrivendo la (3.2) usando l'identità $\sin(\omega_0 t + \boldsymbol{\theta}) = \sin \omega_0 t \cos \boldsymbol{\theta} + \cos \omega_0 t \sin \boldsymbol{\theta}$, come un modello lineare $\mathbf{y} = S\mathbf{x} + \mathbf{w}$ in cui

$$\begin{aligned} \mathbf{y} &= [y(1), \dots, y(N)]' \\ S &= A \begin{bmatrix} \sin \omega_0 & \cos \omega_0 \\ \vdots & \vdots \\ \sin \omega_0 N & \cos \omega_0 N \end{bmatrix} \\ \mathbf{w} &= [w(1), \dots, w(N)]' \quad . \end{aligned} \quad (3.4)$$

Dalla distribuzione di $\boldsymbol{\theta}$ si ricavano facilmente i momenti

$$\mu_{\mathbf{x}} = 0 \quad , \quad P := \text{Var}(\mathbf{x}) = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \quad ; \quad (3.5)$$

inoltre $R = \sigma^2 I_N$, dove I_N è l'identità $N \times N$.

Nelle formule (2.26) e (2.27) servono le espressioni seguenti

$$S' R^{-1} S = \frac{A^2}{\sigma^2} \begin{bmatrix} \sum_1^N \sin^2 \omega_0 t & \sum_1^N \sin \omega_0 t \cos \omega_0 t \\ \sum_1^N \sin \omega_0 t \cos \omega_0 t & \sum_1^N \cos^2 \omega_0 t \end{bmatrix}$$

$$S' R^{-1} \mathbf{y} = \frac{A}{\sigma^2} \begin{bmatrix} \sum_1^N y(t) \sin \omega_0 t \\ \sum_1^N y(t) \cos \omega_0 t \end{bmatrix}$$

e queste formule suggeriscono una semplice espressione asintotica per lo stimatore. Scrivendo

$$\hat{E}(\mathbf{x} | \mathbf{y}) = \left[\frac{1}{N} (P^{-1} + S' R^{-1} S) \right]^{-1} \frac{1}{N} S' R^{-1} \mathbf{y} \quad (3.6)$$

e notando che, per $N \rightarrow \infty$, assumendo $\omega_0 \neq h\pi$ con h intero,

$$\frac{1}{N} \sum_1^N \sin^2 \omega_0 t \rightarrow \frac{1}{2}, \quad \frac{1}{N} \sum_1^N \sin \omega_0 t \cos \omega_0 t \rightarrow 0, \quad \frac{1}{N} \sum_1^N \cos^2 \omega_0 t \rightarrow \frac{1}{2},$$

si conclude che

$$\hat{E}(\mathbf{x} | \mathbf{y}) \cong \frac{2}{A} \begin{bmatrix} \frac{1}{N} \sum_1^N y(t) \sin \omega_0 t \\ \frac{1}{N} \sum_1^N y(t) \cos \omega_0 t \end{bmatrix} \quad (3.7)$$

con matrice varianza d'errore

$$\Lambda \cong \frac{2\sigma^2}{A^2 N} I_2 \quad . \quad (3.8)$$

La formula (3.7) ha un significato estremamente intuitivo e si sarebbe potuta “indovinare” ignorando il termine di rumore nella (3.2). \diamond

L'importanza della teoria della stima lineare a m.v. va al di là della soluzione di semplici problemi di misura in un contesto strettamente “statico”, come quello appena visto. In realtà le idee chiave sulle quali è fondato il calcolo dello stimatore (2.15) e in particolare il *principio di ortogonalità* formano la base metodologica per affrontare problemi nei quali si vuole ricostruire non già una variabile casuale fissata, ma un *segnale aleatorio* variabile nel tempo.

ESEMPIO 2.3 (Stimatori a memoria finita). Questo esempio serve a mostrare come la teoria statica sia in grado di fornire la soluzione di alcuni semplici problemi di stima su processi stocastici, nell'ipotesi che le osservazioni siano in *numero finito e costante nel tempo* (non si abbia cioè “accumulo di informazione” col passare del tempo).

Supponiamo che $\{\mathbf{x}(t)\}$ sia un processo stocastico scalare la cui funzione di covarianza è nota. Per semplicità di notazioni supporremo che $\sigma(t, s) = E[\mathbf{x}(t) \mathbf{x}(s)]$ dipenda solo dalla differenza dei due argomenti temporali t ed s (il processo è così stazionario “del secondo ordine”) e che $E\mathbf{x}(t) = 0$.

Siano t_1, \dots, t_n , n istanti temporali arbitrari ma distinti e si supponga di voler calcolare il *miglior stimatore lineare di $\mathbf{x}(t)$ in base alle n osservazioni $\mathbf{x}(t + t_1), \dots, \mathbf{x}(t + t_n)$* .

Identificando il vettore \mathbf{y} con $[\mathbf{x}(t+t_1), \dots, \mathbf{x}(t+t_n)]'$ e \mathbf{x} con $x(t)$ si può usare immediatamente la formula (2.14) con

$$\begin{aligned}\Sigma_{\mathbf{xy}} &= [\sigma(t_1), \dots, \sigma(t_n)] \\ \Sigma_{\mathbf{y}} &= \begin{bmatrix} \sigma(0) & \sigma(t_1-t_2) & \dots & \sigma(t_1-t_n) \\ \sigma(t_1-t_2) & \sigma(0) & & \dots \\ \dots & & \ddots & \\ \sigma(t_1-t_n) & & & \sigma(0) \end{bmatrix}.\end{aligned}$$

Se $\Sigma_{\mathbf{y}}$ è invertibile, lo stimatore $\hat{\mathbf{x}}(t | t_1, \dots, t_n)$ di $x(t)$ è dato dalla

$$\hat{\mathbf{x}}(t | t_1 \dots t_n) = \Sigma_{\mathbf{xy}} \Sigma_{\mathbf{y}}^{-1} \begin{bmatrix} x(t+t_1) \\ \vdots \\ x(t+t_n) \end{bmatrix}. \quad (3.9)$$

Questo stimatore è un “predittore” del processo $\{\mathbf{x}(t)\}$ se t_1, \dots, t_n sono negativi, un “ricostruttore del passato” se t_1, \dots, t_n sono positivi e un “interpolatore” nel caso generale. Da notare che (per l'ipotesi di stazionarietà) la trasformazione che definisce la funzione lineare delle osservazioni nella (3.9) non dipende da t (è tempo-invariante).

Lo stimatore “degenera” (cioè ha varianza d'errore nulla) se e solo se $\mathbf{x}(0) \in \text{span}\{\mathbf{x}(t_1), \dots, \mathbf{x}(t_n)\}$. La condizione di invertibilità di $\Sigma_{\mathbf{y}}$ dipende dalla struttura di $\sigma(t)$. Si può per esempio verificare che prendendo la funzione covarianza più semplice possibile, $\sigma(t) = p\lambda_0^{-|t|}$, dove $p = \sigma(0)$ è un numero positivo e $0 < \lambda_0 < 1$, la matrice

$$\Sigma_{\mathbf{y}} = \sigma(0) \begin{bmatrix} 1 & \lambda_0^{t_2-t_1} & \dots & \lambda_0^{t_n-t_1} \\ \lambda_0^{t_2-t_1} & 1 & & \dots \\ \dots & & \ddots & \\ \lambda_0^{t_n-t_1} & & & 1 \end{bmatrix} \quad (3.10)$$

è sempre non singolare qualunque siano $t_1 \neq \dots \neq t_n$.

Consideriamo il caso in cui $t_1 = -1, \dots, t_n = -n$ e lo stimatore (3.9) è il *predittore di un passo* di $\mathbf{x}(t)$ basato sulle n variabili aleatorie $\mathbf{x}(t-1), \dots, \mathbf{x}(t-n)$ immediatamente precedenti. In questo caso denoteremo $\Sigma_{\mathbf{xy}}$ e $\Sigma_{\mathbf{y}}$ con i simboli σ_n e Σ_n rispettivamente. Notiamo che Σ_{n+1} si ottiene “bordando” Σ_n secondo lo schema

$$\Sigma_{n+1} = \begin{bmatrix} \sigma(0) & \sigma_n \\ \sigma_n' & \Sigma_n \end{bmatrix}, \quad n = 1, 2, \dots \quad (3.11)$$

dove $\sigma_n = [\sigma(1), \dots, \sigma(n)]$. Consideriamo dunque il *predittore a memoria n*,

$$\hat{\mathbf{x}}_n(t) := \hat{E}[\mathbf{x}(t) | \mathbf{x}(t-1), \dots, \mathbf{x}(t-n)]$$

e il relativo errore di predizione $\mathbf{e}_n(t) := \mathbf{x}(t) - \hat{\mathbf{x}}_n(t)$.

Usando il principio di ortogonalità $\mathbf{e}_n(t) \perp \{\mathbf{x}(t-1), \dots, \mathbf{x}(t-n)\}$, si trova che la varianza del vettore $[\mathbf{e}_n(t) \mathbf{x}(t-1), \dots, \mathbf{x}(t-n)]$ è *diagonale a blocchi* e, detta $\lambda_n^2 := \text{var } \mathbf{e}_n(t)$, si ha

$$\begin{bmatrix} \lambda_n^2 & 0 \\ 0 & \Sigma_n \end{bmatrix} = T_n \Sigma_{n+1} T_n' \quad (3.12)$$

dove T_n è la matrice che opera la trasformazione $[\mathbf{x}(t) \mathbf{x}(t-1), \dots, \mathbf{x}(t-n)] \rightarrow [\mathbf{e}_n(t) \mathbf{x}(t-1), \dots, \mathbf{x}(t-n)]$, ovvero

$$T_n = \begin{bmatrix} 1 & -\sigma_n \Sigma_n^{-1} \\ 0 & I \end{bmatrix} . \quad (3.13)$$

Si noti che $\det T_n = 1$. La diagonalizzazione a blocchi (3.12) può così servire per dimostrare una formula notevole in teoria della predizione, ovvero la

$$\text{var } \mathbf{e}_n(t) = \lambda_n^2 = \frac{\det \Sigma_{n+1}}{\det \Sigma_n} . \quad (3.14)$$

Il processo $\{\mathbf{e}_n(t)\}$ si chiama *innovazione a memoria n* del processo $\{\mathbf{x}(t)\}$ e verrà studiato e utilizzato diffusamente nel capitolo 8. \diamond

ESEMPIO 2.4 (Un caso in cui lo stimatore lineare non ha significato). Consideriamo il processo $\{\mathbf{x}(t)\} := \{\sin \omega t, t \in \mathbb{Z}\}$ in cui ω è una variabile casuale uniformemente distribuita sull'intervallo $[-\pi, \pi]$ e supponiamo di voler costruire uno stimatore di $\mathbf{x}(t)$ (per un t fissato) in base a un numero finito ma peraltro qualunque di osservazioni $\{\mathbf{x}(s), s < t\}$.

Dato che per $s \neq t$

$$E[\mathbf{x}(t) \mathbf{x}(s)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin \omega t \sin \omega s d\omega = 0$$

le variabili $\mathbf{x}(t)$ e $\{\mathbf{x}(s), s < t\}$ sono scorrelate. Lo stimatore lineare a minima varianza di $\mathbf{x}(t)$ è perciò la costante (deterministica) $E\mathbf{x}(t) = 0$ e la varianza d'errore è quindi uguale all'intera varianza a priori di $\mathbf{x}(t)$.

Notando viceversa che $\{\mathbf{x}(t)\}$ soddisfa un'equazione alle differenze del tipo

$$\mathbf{x}(t) - 2 \cos \omega \mathbf{x}(t-1) + \mathbf{x}(t-2) = 0 \quad t \in \mathbb{Z},$$

si può ricavare il valore campionario di ω corrispondente all'osservazione di una traiettoria attraverso la

$$\omega = \text{Arccos} \left[\frac{\mathbf{x}(t) + \mathbf{x}(t-2)}{2\mathbf{x}(t-1)} \right]$$

e quindi costruire uno *stimatore non lineare di $\mathbf{x}(t)$* basato sui dati passati $\{\mathbf{x}(t-1), \mathbf{x}(t-2), \mathbf{x}(t-3)\}$,

$$\hat{\mathbf{x}}(t) = \sin \left\{ \text{Arccos} \left[\frac{\mathbf{x}(t-1) + \mathbf{x}(t-3)}{2\mathbf{x}(t-2)} \right] t \right\}$$

che fornisce *esattamente* il valore di $\mathbf{x}(t)$ (e ha ovviamente varianza d'errore uguale a zero). \diamond

4. Innovazioni e fattorizzazione di Cholesky

Cambiamento di base. Supponiamo che le m componenti di $\mathbf{y} = [\mathbf{y}_1, \dots, \mathbf{y}_m]'$ siano linearmente indipendenti. Consideriamo altre m variabili casuali di \mathbf{H} , linearmente indipendenti,

$$\mathbf{z}_1, \dots, \mathbf{z}_m$$

tali che

$$\mathbf{H}(\mathbf{z}) = \mathbf{H}(\mathbf{y}) . \quad (4.1)$$

È chiaro che il vettore aleatorio \mathbf{z} dev'essere legato a \mathbf{y} da una trasformazione lineare del tipo

$$\mathbf{z} = T\mathbf{y} \quad (4.2)$$

con $T \in \mathbb{R}^{m \times m}$ non singolare (dato che per la (4.1) si deve anche poter scrivere $\mathbf{y} = S\mathbf{z}$, ovvero $\mathbf{y} = ST\mathbf{y}$, $S \in \mathbb{R}^{m \times m}$ da cui $\Sigma_{\mathbf{y}} = ST\Sigma_{\mathbf{y}}$ ed $ST = I$). Le componenti di \mathbf{z} sono una nuova base in $\mathbf{H}(\mathbf{y})$. Dalla (4.2) scende che T si può calcolare dai momenti congiunti di \mathbf{z} e \mathbf{y} ,

$$\Sigma_{\mathbf{zy}} = T\Sigma_{\mathbf{y}} \quad , \quad T = \Sigma_{\mathbf{zy}} \Sigma_{\mathbf{y}}^{-1} \quad (4.3)$$

per cui ovviamente $\mathbf{z} = \hat{E}(\mathbf{z} | \mathbf{y})$ (\mathbf{z} coincide con la sua proiezione su $\mathbf{H}(\mathbf{y})$). Inoltre

$$\Sigma_{\mathbf{z}} = T\Sigma_{\mathbf{y}}T' \quad . \quad (4.4)$$

Sia ora \mathbf{x} un arbitrario vettore aleatorio di \mathbf{H} e consideriamo lo stimatore $\hat{E}(\mathbf{x} | \mathbf{y}) = A\mathbf{y}$ di \mathbf{x} in base ai dati $(\mathbf{y}_1 \dots \mathbf{y}_m)$. Vogliamo vedere come cambia A per effetto del cambiamento di base (4.2).

Sia allora $\hat{E}(\mathbf{x} | \mathbf{z}) = \hat{A}\mathbf{z}$ con

$$\hat{A} = \Sigma_{\mathbf{xz}} \Sigma_{\mathbf{z}}^{-1} \quad . \quad (4.5)$$

Usando la (4.2) si ha $\Sigma_{\mathbf{xz}} = \Sigma_{\mathbf{xy}}T'$. Dalla (4.4) scende allora immediatamente che

$$\hat{A} = \Sigma_{\mathbf{xy}} \Sigma_{\mathbf{y}}^{-1} T^{-1} = AT^{-1} \quad . \quad (4.6)$$

Quindi, dato $\hat{E}(\mathbf{x} | \mathbf{y}) = A\mathbf{y}$, per ottenere lo stimatore $\hat{E}(\mathbf{x} | \mathbf{z})$ basta semplicemente operare la trasformazione (4.2) sulla variabile \mathbf{y} . In altre parole la matrice A cambia solo per effetto del cambiamento di base in $\mathbf{H}(\mathbf{y})$.

Questo risultato è in realtà la versione “debole” di una proprietà molto generale dell’operatore media condizionata $E(\mathbf{x} | \mathbf{y})$, che si dimostra in teoria della probabilità: $E(\mathbf{x} | \mathbf{y})$ dipende solo dalla σ -algebra indotta dal vettore aleatorio \mathbf{y} . In parole più semplici, se $\mathbf{z} = \varphi(\mathbf{y})$, dove φ è una qualunque funzione invertibile di \mathbf{y} , si ha

$$E(\mathbf{x} | \mathbf{z}) = [E(\mathbf{x} | \mathbf{y})]_{\mathbf{y}=\varphi^{-1}(\mathbf{z})} \quad .$$

PROBLEMA 2.4. Siano \mathbf{x} e \mathbf{y} vettori aleatori di \mathbf{H} (a media nulla!). Sia $\hat{E}(\mathbf{x} | \mathbf{y}) = A\mathbf{y}$ e supponiamo che $\mathbf{z} = T\mathbf{y} + b$, $b \in \mathbb{R}^m$, con T invertibile. Mostrare che lo stimatore lineare a minima varianza di \mathbf{x} in base a \mathbf{z} si ottiene mediante la formula

$$\hat{E}(\mathbf{x} | \mathbf{z}) = AT^{-1}(\mathbf{z} - b) \quad .$$

In particolare notare che lo stimatore $\hat{E}(\mathbf{x} | \mathbf{y} + b)$ non dipende da b . Questo è il motivo per cui si può sempre interpretare, anche se \mathbf{y} ha media diversa da zero, lo stimatore lineare $\hat{E}(\mathbf{x} | \mathbf{y})$ basato su \mathbf{y} come una proiezione ortogonale, anche se quest’ultima è relativa allo spazio $\mathbf{H}(\bar{\mathbf{y}})$, dove $\bar{\mathbf{y}} := \mathbf{y} - \mu_{\mathbf{y}}$. \diamond

Stima con dati linearmente dipendenti. In questa sezione ricaveremo una formula che generalizza la (2.12) al caso in cui $\Sigma_{\mathbf{y}}$ è singolare. Naturalmente continueremo a ipotizzare che \mathbf{x} e \mathbf{y} abbiano media nulla (come abbiamo già osservato, una volta risolto il problema in queste ipotesi, l’estensione al caso di medie non nulle si ottiene facilmente).

Supponiamo quindi che $\Sigma_{\mathbf{y}}$ abbia rango $r \leq m$, ovvero possa avere uno spazio nullo non banale e quindi che le componenti del vettore \mathbf{y} non siano variabili aleatorie linearmente indipendenti. Sia $\mathbf{z} = T\mathbf{y}$, $T \in \mathbb{R}^{r \times m}$ una base r -dimensionale di $\mathbf{H}(\mathbf{y})$ (ad esempio un sottoinsieme linearmente indipendente delle m componenti di \mathbf{y}) per cui

$$\mathbf{H}(\mathbf{z}) = \mathbf{H}(\mathbf{y}) \quad (4.7)$$

Come abbiamo visto, dato che \mathbf{z} e \mathbf{y} generano il medesimo spazio di Hilbert, la stima lineare a minima varianza di \mathbf{x} dato \mathbf{y} coincide con la stima lineare a minima varianza di \mathbf{x} dato \mathbf{z} . Dato che $\Sigma_{\mathbf{z}}$ è non singolare, per calcolare $\hat{E}[\mathbf{x} | \mathbf{y}]$ è sufficiente determinare il vettore \mathbf{z} e applicare la formula consueta:

$$\hat{E}[\mathbf{x} | \mathbf{y}] = \hat{E}[\mathbf{x} | \mathbf{z}] = \Sigma_{\mathbf{xz}} \Sigma_{\mathbf{z}}^{-1} \mathbf{z}. \quad (4.8)$$

Per implementare la soluzione appena trovata partendo dai dati del problema, $\Sigma_{\mathbf{xy}}$ e $\Sigma_{\mathbf{y}}$, abbiamo bisogno di una procedura per calcolare la matrice di trasformazione \bar{T} .

Allo scopo sia $V \in \mathbb{R}^{m \times (m-r)}$ una matrice le cui colonne sono una base per $\ker \Sigma_{\mathbf{y}}$ (lo spazio nullo di $\Sigma_{\mathbf{y}}$). In altri termini, $\Sigma_{\mathbf{y}} v = 0$, $v \in \mathbb{R}^m$, se e solo se v è combinazione lineare delle colonne di V . Sia inoltre $U \in \mathbb{R}^{m \times r}$ una matrice le cui colonne sono una base per lo spazio immagine $\text{Im } \Sigma_{\mathbf{y}} = [\ker \Sigma_{\mathbf{y}}]^\perp$. Si osservi che la matrice $m \times m$

$$T := [U | V]$$

è non singolare per costruzione. Dimostriamo ora il seguente risultato

LEMMA 2.2. *I vettori \mathbf{z} che formano una base per $\mathbf{H}(\mathbf{y})$ sono tutti e soli quelli del tipo*

$$\mathbf{z} = U' \mathbf{y} \quad (4.9)$$

dove U è una matrice le cui colonne formano una base per $\text{Im } \Sigma_{\mathbf{y}}$. Conseguentemente, lo stimatore lineare a minima varianza di \mathbf{x} si può scrivere nella forma

$$\hat{E}[\mathbf{x} | \mathbf{y}] = \Sigma_{\mathbf{xz}} \Sigma_{\mathbf{z}}^{-1} \mathbf{z} = E[\mathbf{xy}'U] (E[U' \mathbf{yy}'U])^{-1} U' \mathbf{y} = \Sigma_{\mathbf{xy}} U (U' \Sigma_{\mathbf{y}} U)^{-1} U' \mathbf{y}. \quad (4.10)$$

la matrice $U(U' \Sigma_{\mathbf{y}} U)^{-1} U'$ non dipende dalla particolare scelta di U ma solo dallo spazio immagine $\text{Im } \Sigma_{\mathbf{y}}$.

DIMOSTRAZIONE. Siano $\mathbf{y}_U := U' \mathbf{y}$ e $\mathbf{y}_V := V' \mathbf{y}$ con U e V matrici le cui colonne formano una base per $\text{Im } \Sigma_{\mathbf{y}}$ e $\ker \Sigma_{\mathbf{y}}$, rispettivamente.

Osserviamo per prima cosa che $\mathbf{y}_V := V' \mathbf{y} = 0$. Infatti, si ha

$$\Sigma_{\mathbf{y}_V} = E[V' \mathbf{yy}' V] = V' E[\mathbf{yy}' V] = V' \Sigma_{\mathbf{y}} V = 0$$

ossia la varianza di \mathbf{y}_V è nulla o, equivalentemente, il vettore aleatorio \mathbf{y}_V è nullo con probabilità 1. In particolare, quindi, si ha

$$T' \mathbf{y} = \begin{bmatrix} U' \\ V' \end{bmatrix} \mathbf{y} = \begin{bmatrix} \mathbf{y}_U \\ 0 \end{bmatrix} \quad (4.11)$$

Poiché T' è non singolare, $\mathbf{H}(\mathbf{y}) = \mathbf{H}(T' \mathbf{y})$. D'altronde, poiché le ultime componenti di $T' \mathbf{y}$ sono nulle, si ha $\mathbf{H}(T' \mathbf{y}) = \mathbf{H}(\mathbf{y}_U)$ e quindi

$$\mathbf{H}(\mathbf{y}) = \mathbf{H}(\mathbf{y}_U)$$

Inoltre, $\Sigma_{\mathbf{y}_U} = U' \Sigma_{\mathbf{y}} U$ è definita positiva. Infatti se $w \in \ker \Sigma_{\mathbf{y}_U}$ allora $\Sigma_{\mathbf{y}} U w = 0$. Ma poiché le colonne di U sono nel complemento ortogonale di $\ker \Sigma_{\mathbf{y}}$, $U w = 0$ e, poiché le colonne di U sono linearmente indipendenti (sono una base di $[\ker \Sigma_{\mathbf{y}}]^\perp$), $w = 0$.

Scrivendo infine $U = \bar{U} S$ con \bar{U} a colonne ortonormali e $S \in \mathbb{R}^{r \times r}$ non singolare e sostituendo all'ultimo membro di (4.10) si trova che quest'ultimo termine dipende in effetti solo da \bar{U} . \square

Osserviamo che la matrice U non è univocamente determinata, anzi vi sono infiniti modi per sceglierla. A ciascuna scelta corrisponde un diverso vettore \mathbf{z} ma l'ultimo membro della (4.10) che fornisce una formula per il calcolo dello stimatore ottimo, è indipendente da quale scelta si è operata. In particolare U può essere scelta a colonne ortogonali.

Calcolo dello stimatore ottimo mediante la pseudoinversa. Per poter utilizzare la formula (4.10) occorre calcolare preliminarmente una base per lo spazio vettoriale $\text{Im } \Sigma_{\mathbf{y}}$. Esiste in realtà un modo molto più diretto di riscrivere la (4.10) e quindi di calcolare lo stimatore lineare ottimo.

Ricordiamo a questo proposito un classico risultato di algebra lineare.

LEMMA 2.3. *Data una matrice $A \in \mathbb{R}^{n \times m}$ esiste un'unica matrice $A^+ \in \mathbb{R}^{m \times n}$ che soddisfa le seguenti quattro proprietà.*

$$A^+AA^+ = A^+, \quad AA^+A = A, \quad (AA^+)' = AA^+, \quad (A^+A)' = A^+A. \quad (4.12)$$

DEFINIZIONE 2.3. *La matrice $A^+ \in \mathbb{R}^{m \times n}$ che soddisfa le quattro proprietà (4.12) si dice pseudoinversa di Moore-Penrose di A .*

Si osservi che quando A è non singolare $A^+ = A^{-1}$. Per una trattazione completa delle proprietà della pseudoinversa si rimanda a [16]. Qui ci limitiamo a richiamare un paio di proprietà che saranno utili nel seguito e che possono essere verificate facilmente basandosi sulle proprietà (4.12).

- (1) Siano $A \in \mathbb{R}^{n \times m}$, $T_1 \in \mathbb{R}^{n \times n}$ e $T_2 \in \mathbb{R}^{m \times m}$, con T_1 e T_2 matrici ortogonali. Allora si ha

$$(T_1AT_2)^+ = T_2^{-1}A^+T_1^{-1} = T_2'A^+T_1'. \quad (4.13)$$

- (2) Sia

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times m}$$

con $A_1 \in \mathbb{R}^{p \times p}$ invertibile. Allora si ha,

$$A^+ = \begin{bmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{bmatrix}. \quad (4.14)$$

Il risultato che segue consente di esprimere in modo molto semplice lo stimatore lineare ottimo attraverso l'uso della pseudoinversa.

TEOREMA 2.6. *Lo stimatore lineare a minima varianza è dato dalla formula*

$$\hat{E}[\mathbf{x} | \mathbf{y}] = \Sigma_{\mathbf{xy}} \Sigma_{\mathbf{y}}^+ \mathbf{y}. \quad (4.15)$$

DIMOSTRAZIONE. Alla luce del lemma 2.2, è sufficiente dimostrare che vale l'identità $U(U'\Sigma_{\mathbf{y}}U)^{-1}U' = \Sigma_{\mathbf{y}}^+$, avendo scelto U in modo che le sue colonne formino una base ortonormale di $\text{Im } \Sigma_{\mathbf{y}}$. Prendendo anche V a colonne ortonormali e

ricordando che $V'\Sigma_{\mathbf{y}} = 0$ e $\Sigma_{\mathbf{y}}V = 0$, si ha

$$\begin{aligned}\Sigma_{\mathbf{y}}^+ &= \left[T^{-\top} \underbrace{\begin{bmatrix} U' \\ V' \end{bmatrix}}_{T'} \Sigma_{\mathbf{y}} \underbrace{[U \mid V]}_T T^{-1} \right]^+ = T \left[\begin{bmatrix} U' \\ V' \end{bmatrix} \Sigma_{\mathbf{y}} [U \mid V] \right]^+ T' \\ &= T \begin{bmatrix} U'\Sigma_{\mathbf{y}}U & 0 \\ 0 & 0 \end{bmatrix}^+ T' = [U \mid V] \begin{bmatrix} (U'\Sigma_{\mathbf{y}}U)^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U' \\ V' \end{bmatrix} \\ &= U(U'\Sigma_{\mathbf{y}}U)^{-1}U' \end{aligned} \quad (4.16)$$

che è quanto volevamo dimostrare. \square

In conclusione osserviamo che, vista l'unicità della pseudoinversa, la (4.15) dimostra direttamente ciò che avevamo già notato, ossia che la stima ottima non dipende dalla matrice U scelta.

Il calcolo della pseudoinversa può essere fatto in modo numericamente efficiente usando delle routines disponibili in quasi tutti i pacchetti *software* scientifici. Quindi la (4.15) è effettivamente il modo più semplice ed efficiente per calcolare lo stimatore nel caso di misure dipendenti.

Innovazioni. C'è un caso in cui il calcolo dello stimatore a m.v. diventa assolutamente banale e precisamente quando le componenti $(\mathbf{y}_1, \dots, \mathbf{y}_m)$ del vettore aleatorio \mathbf{y} sono *ortonormali* (cioè scorrelate e di varianza unitaria). La matrice $\Sigma_{\mathbf{y}}$ diventa allora *l'identità* e la formula per lo stimatore lineare di \mathbf{x} diventa

$$\hat{E}(\mathbf{x} \mid \mathbf{y}) = \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1} \mathbf{y} = \Sigma_{\mathbf{x}\mathbf{y}} \mathbf{y} \quad (4.17)$$

cioè la matrice A è *semplicemente la matrice di covarianza, $\Sigma_{\mathbf{x}\mathbf{y}}$, tra \mathbf{x} e \mathbf{y} .*

Vediamo allora come trasformare, mediante un cambio in base in $\mathbf{H}(\mathbf{y})$, \mathbf{y} in un vettore aleatorio a componenti ortonormali. In vista della generalizzazione ai processi stocastici del procedimento, sarà conveniente considerare le componenti $\mathbf{y}_1, \dots, \mathbf{y}_m$ come osservazioni che si fanno *sequenzialmente* nel tempo (\mathbf{y}_1 è la prima e \mathbf{y}_m è l'ultima). Assumeremo $\{\mathbf{y}_1, \dots, \mathbf{y}_m\}$ *linearmente indipendenti*, cioè supporremo la matrice

$$\Sigma_{\mathbf{y}} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{21} & & & \\ \sigma_{m1} & \dots & \dots & \sigma_{mm} \end{bmatrix}, \quad \sigma_{ij} = E(\mathbf{y}_i \mathbf{y}_j)$$

(simmetrica e) definita positiva. Useremo inoltre estensivamente la notazione $\mathbf{y}^t := [\mathbf{y}_1, \dots, \mathbf{y}_t]'$ ($t \leq m$). Evidentemente $\mathbf{y}^m \equiv \mathbf{y}$.

Il procedimento di ortonormalizzazione dei "vettori" $(\mathbf{y}_1, \dots, \mathbf{y}_m)$ di $\mathbf{H}(\mathbf{y})$, che illustreremo fra poco, si chiama procedimento di Gram-Schmidt e ha validità molto generale. Sulla sua importanza concreta dal punto di vista della soluzione numerica di problemi di stima lineare torneremo più avanti.

Il procedimento lavora sequenzialmente sugli elementi della successione $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m$. Si pone

$$\begin{aligned} \mathbf{e}_1 &:= \mathbf{y}_1 \quad ; & \boldsymbol{\varepsilon}_1 &:= \mathbf{e}_1 / \|\mathbf{e}_1\| \\ \mathbf{e}_2 &:= \mathbf{y}_2 - \langle \mathbf{y}_2, \boldsymbol{\varepsilon}_1 \rangle \boldsymbol{\varepsilon}_1 \quad ; & \boldsymbol{\varepsilon}_2 &:= \mathbf{e}_2 / \|\mathbf{e}_2\| \\ & \dots\dots\dots & & \\ \mathbf{e}_t &:= \mathbf{y}_t - \sum_{k=1}^{t-1} \langle \mathbf{y}_t, \boldsymbol{\varepsilon}_k \rangle \boldsymbol{\varepsilon}_k \quad ; & \boldsymbol{\varepsilon}_t &:= \mathbf{e}_t / \|\mathbf{e}_t\| \\ & t = 2, 3, \dots, m \quad . \end{aligned} \quad (4.18)$$

Nelle formule (4.18) abbiamo sottaciuto il pedice H nei prodotti scalari e nei simboli che denotano la norma.

TEOREMA 2.7. *Le variabili casuali $\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots, \boldsymbol{\varepsilon}_m\}$ definite attraverso il procedimento di Gram-Schmidt (4.18) sono ortonormali.*

Per ogni t , $\boldsymbol{\varepsilon}^t$ è legato a \mathbf{y}^t da una trasformazione non singolare L_t , di struttura triangolare inferiore,

$$\mathbf{y}^t = L_t \boldsymbol{\varepsilon}^t \quad (4.19)$$

di modo che

$$H(\boldsymbol{\varepsilon}^t) = H(\mathbf{y}^t) \quad , \quad t = 1, 2, \dots, m \quad . \quad (4.20)$$

Le variabili $\{\mathbf{e}_t, t = 1, 2, \dots, m\}$ sono gli errori di predizione di un passo di \mathbf{y}_t in base a \mathbf{y}^{t-1} , nel senso che

$$\mathbf{e}_t = \mathbf{y}_t - \hat{E}(\mathbf{y}_t | \mathbf{y}^{t-1}) \quad (4.21)$$

ed $\{\boldsymbol{\varepsilon}_t, t = 1, 2, \dots, m\}$ sono i corrispondenti errori di predizione normalizzati (scalati a varianza unitaria). La matrice $L \equiv L_m$ è un fattore a struttura triangolare inferiore della matrice $\Sigma_{\mathbf{y}}$, ovvero

$$\Sigma_{\mathbf{y}} = LL' \quad . \quad (4.22)$$

DIMOSTRAZIONE. L'ortonormalità degli $\{\boldsymbol{\varepsilon}_t\}$ si dimostra facilmente per induzione a partire dall'ultima formula in (4.18). Se $\{\boldsymbol{\varepsilon}^{t-1}\}$ ha componenti ortonormali, si ha infatti

$$\langle \mathbf{e}_t, \boldsymbol{\varepsilon}_k \rangle = \langle \mathbf{y}_t, \boldsymbol{\varepsilon}_k \rangle - \langle \mathbf{y}_t, \boldsymbol{\varepsilon}_k \rangle = 0 \quad , \quad k < t$$

e quindi $\langle \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_k \rangle = 0$ per $k = 1, 2, \dots, t-1$. Questa proprietà è del resto evidente se si riflette sul significato geometrico del procedimento.

Esplicitando rispetto a \mathbf{y}^t nelle (4.18) si trova poi

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_t \end{bmatrix} = \begin{bmatrix} \|\mathbf{e}_1\| & 0 & \dots & 0 \\ \langle \mathbf{y}_2, \boldsymbol{\varepsilon}_1 \rangle & \|\mathbf{e}_2\| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{y}_t, \boldsymbol{\varepsilon}_1 \rangle & \langle \mathbf{y}_t, \boldsymbol{\varepsilon}_2 \rangle & \dots & \|\mathbf{e}_t\| \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_t \end{bmatrix} \quad (4.23)$$

che è una trasformazione a struttura triangolare inferiore come la (4.19). Notiamo che i termini diagonali sono tutti non nulli, ovvero $\|\mathbf{e}_t\| > 0$ per ogni t , giacché $\mathbf{e}_t = 0$ implicherebbe $\mathbf{y}_t \in \mathbf{H}(\boldsymbol{\varepsilon}^{t-1}) \subset \mathbf{H}(\mathbf{y}^{t-1})$ per la (4.18) e questo è manifestamente impossibile data l'indipendenza lineare delle $\{\mathbf{y}_t\}$. In conclusione, la trasformazione è invertibile e questo permette immediatamente di concludere che vale anche la proprietà (4.20).

Dato che gli $\{\boldsymbol{\varepsilon}_k, k \leq t\}$ sono ortonormali, la (4.17) porge facilmente

$$\hat{E}[\mathbf{y}_t | \boldsymbol{\varepsilon}^{t-1}] = \sum_{k=1}^{t-1} \langle \mathbf{y}_t, \boldsymbol{\varepsilon}_k \rangle \boldsymbol{\varepsilon}_k$$

e quest'ultima espressione coincide, per la formula dei cambiamenti di base (4.6), con $\hat{E}[\mathbf{y}_t | \mathbf{y}^{t-1}]$.

La (4.22) è infine un'immediata conseguenza della (4.19) (scritta per $t = m$) e dell'ortonormalità di $\boldsymbol{\varepsilon}^m$. \square

Il vettore $\boldsymbol{\varepsilon}$ si chiama *innovazione normalizzata* di \mathbf{y} (\mathbf{e} è l'innovazione non normalizzata). Questa terminologia, introdotta da Wiener e Masani [152], è in sostanza un modo pittoresco di ribadire il fatto che $\boldsymbol{\varepsilon}_t$ è un errore di predizione. La variabile $\boldsymbol{\varepsilon}_t$ contiene in effetti la “nuova informazione” che è portata dalla t -esima misura \mathbf{y}_t , quando si hanno già a disposizione le $t - 1$ precedenti. La parte “prevedibile” (in base a \mathbf{y}^{t-1}) di \mathbf{y}_t è infatti proprio $\hat{E}[\mathbf{y}_t | \mathbf{y}^{t-1}]$.

Notiamo che $\boldsymbol{\varepsilon}_t$ si può in linea di principio ricavare esplicitamente come funzione lineare di \mathbf{y}^t risolvendo sequenzialmente le (4.18), ma in concreto questo si fa invertendo la matrice L ed estraendone la t -esima riga. Di fatto l'inversa di L è ancora inferiormente triangolare per cui, detta g_t la t -esima riga di L^{-1} , si vede che $\boldsymbol{\varepsilon}_t = g_t \mathbf{y} = g_t \mathbf{y}^t$ dipende in effetti solo da \mathbf{y}^t . Questa proprietà di *equivalenza causale* tra $\boldsymbol{\varepsilon}$ e \mathbf{y} , che è sintetizzata nella relazione (4.20), è di importanza fondamentale.

Prima di occuparci del calcolo dell'innovazione conviene esaminare brevemente la questione della sua *unicità*.

Sia $\bar{\boldsymbol{\varepsilon}} = \{\bar{\boldsymbol{\varepsilon}}_t, t = 1, \dots, m\}$ un'altra base ortonormale in $\mathbf{H}(\mathbf{y})$. Dato che $\mathbf{H}(\mathbf{y}) = \mathbf{H}(\boldsymbol{\varepsilon}) = \mathbf{H}(\bar{\boldsymbol{\varepsilon}})$, i vettori $\boldsymbol{\varepsilon}$ ed $\bar{\boldsymbol{\varepsilon}}$ sono legati da una trasformazione (non singolare) $T \in \mathbb{R}^{m \times m}$,

$$\bar{\boldsymbol{\varepsilon}} = T\boldsymbol{\varepsilon} \quad .$$

Siccome $\Sigma_{\bar{\boldsymbol{\varepsilon}}} = I$, $\Sigma_{\boldsymbol{\varepsilon}} = I$, la matrice T deve essere ortogonale, ovvero:

$$TT' = I \quad . \quad (4.24)$$

Ora, se si impone anche a $\bar{\boldsymbol{\varepsilon}}$ di soddisfare la condizione di equivalenza causale (4.20) è facile riconoscere che T deve essere di struttura triangolare inferiore. Questa condizione implica automaticamente che T debba essere una *matrice segno*: infatti, dalla (4.24) $T^{-1} = T'$; ma allora T^{-1} deve essere allo stesso tempo triangolare inferiormente e superiormente e pertanto T è diagonale. Dalla (4.24) segue poi che gli elementi di T sono radici quadrate dell'unità. Abbiamo così dimostrato che l'innovazione è essenzialmente unica.

PROPOSIZIONE 2.2. *Esiste un'unica (a meno del segno) m -pla di variabili ortonormali $\{\boldsymbol{\varepsilon}_t\}$ che soddisfa la condizione di equivalenza causale (4.20).*

L'unicità del vettore delle innovazioni ha una controparte algebrica nell'unicità della fattorizzazione (4.22) della matrice simmetrica e definita positiva $\Sigma_{\mathbf{y}}$. Come abbiamo già notato, per calcolare concretamente l'innovazione di \mathbf{y} occorre proprio fattorizzare $\Sigma_{\mathbf{y}}$ nella forma (4.22). Il vettore ortonormale $\boldsymbol{\varepsilon}$ si ottiene poi trasformando \mathbf{y} attraverso la $\boldsymbol{\varepsilon} = L^{-1}\mathbf{y}$.

In algebra lineare si dimostra che ogni matrice Q simmetrica e definita positiva può essere fattorizzata nella forma $Q = LL'$ dove L è triangolare inferiormente,

$$\begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & & \\ \vdots & & & \\ q_{n1} & \cdots & \cdots & q_{nn} \end{bmatrix} = \begin{bmatrix} \ell_{11} & 0 & & 0 \\ \ell_{21} & \ell_{22} & & \\ \vdots & & & 0 \\ \ell_{n1} & \cdots & \cdots & \ell_{nn} \end{bmatrix} \begin{bmatrix} \ell_{11} & \ell_{21} & \cdots & \ell_{n1} \\ 0 & \ell_{22} & & \\ \vdots & & & \\ 0 & & 0 & \ell_{nn} \end{bmatrix} \quad (4.25)$$

dove i termini diagonali ℓ_{ii} (che sono radici quadrate) sono determinati a meno del segno e possono per esempio essere presi tutti positivi. Questa decomposizione si chiama *fattorizzazione di Cholesky* di Q ed è uno degli strumenti di calcolo fondamentali in Statistica.

Il fattore $L = [\ell_{ij}]$ si calcola con un semplice algoritmo che deriveremo nel corso della dimostrazione del teorema seguente.

TEOREMA 2.8. *Se la matrice simmetrica Q è definita positiva, allora esiste un'unica matrice L , triangolare inferiormente e ad elementi diagonali positivi, tale che $Q = LL'$.*

DIMOSTRAZIONE. Se $n = 1$ il risultato è ovvio ($L = +\sqrt{q_{11}}$) e si può così procedere per induzione sulla dimensione k di Q .

Sia allora

$$Q_{k+1} = \begin{bmatrix} Q_k & r \\ r' & q \end{bmatrix}$$

di dimensione $(k+1) \times (k+1)$, simmetrica e definita positiva. Perché Q_{k+1} possa esprimersi come prodotto dei due fattori *triangolari*:

$$L_{k+1} = \begin{bmatrix} L_k & 0 \\ \ell' & \lambda \end{bmatrix}, \quad L'_{k+1} = \begin{bmatrix} L'_k & \ell \\ 0 & \lambda \end{bmatrix}$$

occorre e basta che valgano le

$$\begin{aligned} L_k L'_k &= Q_k \\ L_k \ell &= r \\ \ell' \ell + \lambda^2 &= q \end{aligned}$$

La prima di queste relazioni è verificata in base all'ipotesi induttiva giacché Q_k è simmetrica e *definita positiva* (ricordare il criterio di Sylvester sui determinanti dei minori principali).

La seconda fornisce

$$\ell = L_k^{-1} r$$

e la terza relazione ha *un'unica* soluzione positiva, λ , giacché la differenza

$$q - \ell' \ell = q - r' (L'_k)^{-1} L_k^{-1} r = q - r' Q_k^{-1} r$$

è *positiva*. Risulta infatti:

$$q - r' Q_k^{-1} r = [r' Q_k^{-1} \quad -1] \begin{bmatrix} Q_k & r \\ r' & q \end{bmatrix} \begin{bmatrix} Q_k^{-1} r \\ -1 \end{bmatrix}$$

e il secondo membro è manifestamente positivo. Alternativamente, se si interpreta Q come matrice varianza di un vettore aleatorio \mathbf{y} (cosa che è evidentemente sempre possibile) la quantità $\lambda^2 = q - r' Q_k^{-1} r$ altri non è se non la *varianza dell'errore di predizione* di \mathbf{y}_{k+1} in base a \mathbf{y}^k (vedi la formula (2.16)). \square

Dalla dimostrazione scende subito il celebre algoritmo per il calcolo del fattore di Cholesky, L .

Questo algoritmo ha struttura ricorsiva e trasforma successivamente le righe della parte triangolare inferiore di Q (che è simmetrica) in quelle del fattore L .

Notiamo fin d'ora che questa proprietà permette facilmente di aggiornare il calcolo del fattore L in corrispondenza all'introduzione di ulteriori righe (e colonne) nella matrice Q .

Algoritmo di Cholesky.

ALGORITMO 2.1. *Il seguente algoritmo esegue la fattorizzazione di Cholesky di una matrice $Q = Q' > 0$.*

1. *Gli elementi diagonali ℓ_{ii} si calcolano con la*

$$\ell_{ii} = \sqrt{q_{ii} - \sum_{j=1}^{i-1} \ell_{ij}^2} \quad , \quad i = 1, \dots, n \quad . \quad (4.26)$$

2. *Supponendo di aver calcolato le prime $i - 1$ righe di L , l'elemento di posto (ij) nella i -sima riga è dato dalla*

$$\ell_{ij} = \frac{1}{\ell_{jj}} \left(q_{ij} - \sum_{k=1}^{j-1} \ell_{ik} \ell_{jk} \right) \quad , \quad j = 1, \dots, i - 1 \quad (j < i). \quad (4.27)$$

Per usare la (4.27) bisogna aver già calcolato i precedenti elementi della riga i -sima, $\ell_{i1}, \dots, \ell_{i,j-1}$. I numeri $\ell_{j1}, \dots, \ell_{j,j-1}$ sono i corrispondenti elementi della riga j -sima e sono tutti noti perché la riga j -sima sta "sopra" la riga i -sima ($j < i$). L'algoritmo lavora dall'alto in basso e, per ogni riga, da sinistra a destra.

Estensione ai processi stocastici. È importante rendersi conto del fatto che il vincolo che ci siamo posti a priori sulla finitezza della dimensione m di \mathbf{y} non gioca alcun ruolo nella costruzione dell'innovazione.

Il procedimento di Gram-Schmidt può infatti essere definito, ancora attraverso la (4.18), per una successione *infinita* di misure, ovvero per un *processo stocastico* $\{\mathbf{y}(t), t = 1, 2, \dots\}$ costituito da misure scalari (ma più in generale anche vettoriali) eseguite sequenzialmente nel tempo. Il Teorema 2.7 e il relativo corollario continuano a valere anche in questo caso purché, beninteso, la matrice varianza $\Sigma_{\mathbf{y}}(t) := E[\mathbf{y}^t(\mathbf{y}^t)']$ sia *definita positiva per ogni t* . Con questa puntualizzazione anche l'enunciato del Teorema 2.8 (e la relativa dimostrazione) e l'algoritmo di fattorizzazione che ne segue rimangono validi anche per $m = \infty$.

Se ci si pone in questo contesto, si può dare una suggestiva interpretazione "sistemistica" alla relazione (4.19). Isolando l'equazione corrispondente alla t -sima riga si può infatti scrivere

$$\mathbf{y}(t) = \ell_t \boldsymbol{\varepsilon}^t = \sum_1^t \ell(t, s) \boldsymbol{\varepsilon}(s) \quad (4.28)$$

dove $\ell_t = [\ell(t, 1), \dots, \ell(t, t)]$ è la t -sima riga di L_t . Questa scrittura esprime $\mathbf{y}(t)$ come l'uscita di un sistema dinamico lineare, di risposta impulsiva $\ell(t, \cdot)$, corrispondente al processo di ingresso $\{\boldsymbol{\varepsilon}(t)\}$ che è rumore bianco (normalizzato). L'innovazione è allora la risposta del *sistema inverso*, di risposta impulsiva $g_t = [g(t, 1), \dots, g(t, t)]$ uguale alla t -sima riga di L_t^{-1} , in corrispondenza all'ingresso $\{\mathbf{y}(t)\}$.

ESEMPIO 2.5 (Innovazione di un processo stazionario a media mobile). Consideriamo il processo $\{\mathbf{y}(t)\}$ descritto dal modello a “media mobile”

$$\mathbf{y}(t) = \mathbf{w}(t) + a \mathbf{w}(t-1) \quad , \quad t = 1, 2, \dots$$

dove $\{\mathbf{w}(t)\}$ è rumore bianco a media zero e varianza σ^2 . È immediato calcolare la funzione di covarianza di $\{\mathbf{y}(t)\}$,

$$\sigma(\tau) = E \mathbf{y}(t+\tau) \mathbf{y}(t)' = \begin{cases} \sigma^2(1+a^2) & \tau = 0 \\ \sigma^2 a & \tau = \pm 1 \\ 0 & |\tau| > 1 \end{cases}$$

per cui $\Sigma_t := E \mathbf{y}^t (\mathbf{y}^t)'$, $t = 1, 2$, ha la struttura tridiagonale

$$\Sigma_t = \sigma^2 \begin{bmatrix} 1+a^2 & a & 0 & \dots & 0 \\ a & 1+a^2 & \ddots & & \\ 0 & & \ddots & \ddots & \\ \vdots & & & \ddots & a \\ 0 & & a & & 1+a^2 \end{bmatrix}$$

e il determinante (normalizzato) $d_t := \frac{1}{\sigma^{2t}} \det \Sigma_t$ soddisfa l'equazione alle differenze lineare

$$d_{t+1} = (1+a^2) d_t - a^2 d_{t-1} \quad , \quad t \geq 1$$

dalla quale si ricava facilmente l'espressione

$$d_t = \sum_0^t (a^2)^k \quad .$$

Come si vede, d_t è diverso da zero per tutti i t . Possiamo allora applicare a Σ_t l'algoritmo di Cholesky per calcolare il fattore triangolare $L_t = [\ell(i, j)]$ che ci fornirà la rappresentazione d'innovazione dal processo $\{\mathbf{y}(t)\}$.

Ragionando per induzione si mostra facilmente che L_t deve essere bidiagonale

$$L_t = \begin{bmatrix} \ell(1, 1) & 0 & \dots & 0 \\ \ell(2, 1) & \ell(2, 2) & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & \ell(t, t-1) & \ell(t, t) \end{bmatrix}$$

e i termini $\ell(t, t-1)$ ed $\ell(t, t)$ si ricavano dalle formule (4.26) e (4.27) ponendo $\ell_{t,s} := \ell(t, s) = 0$ per $s < t-1$. Si trovano così le formule ricorsive

$$\begin{aligned}\ell(t, t) &= [\sigma(0) - \ell(t, t-1)^2]^{\frac{1}{2}} \\ \ell(t, t-1) &= \frac{1}{\ell(t-1, t-1)} \sigma(1)\end{aligned}$$

che possono essere risolte iterando su $t = 2, 3, \dots$, partendo dalla condizione iniziale

$$\ell(1, 1) = +\sqrt{\sigma^2(1+a^2)} \quad .$$

È però più semplice rifarsi all'osservazione riportata alla fine della dimostrazione del Teorema 2.8, in cui si identifica $\ell^2(t, t)$ con la varianza dell'errore di predizione di un passo di $\mathbf{y}(t)$ in base alle $t-1$ osservazioni precedenti, \mathbf{y}^{t-1} , del processo $\{\mathbf{y}(t)\}$. Usando la formula (3.14) derivata nell'Esempio 2.3, si ha

$$\ell^2(t, t) = \frac{\det \Sigma_t}{\det \Sigma_{t-1}} = \sigma^2 \frac{d_t}{d_{t-1}}$$

e

$$\ell(t, t-1) = \sigma a \sqrt{\frac{d_{t-2}}{d_{t-1}}} \quad .$$

Come si vede, queste espressioni dipendono effettivamente dal tempo, per cui la rappresentazione d'innovazione di $\{\mathbf{y}(t)\}$

$$\mathbf{y}(t) = \ell(t, t) \boldsymbol{\varepsilon}(t) + \ell(t, t-1) \boldsymbol{\varepsilon}(t-1) \quad (4.29)$$

è tempo-variante. Il fatto notevole è però che i coefficienti del modello (4.29) *tendono ad assestarsi per* $t \rightarrow \infty$. Se $|a| < 1$, infatti, il rapporto d_t/d_{t-1} converge a 1; se $|a| > 1$, si ha

$$\frac{d_t}{d_{t-1}} = \frac{a^{2t}}{a^{2(t-1)}} \frac{\sum_0^t a^{-2k}}{\sum_0^{t-1} a^{-2k}} \rightarrow a^2$$

mentre, nel caso $|a| = 1$, $d_t/d_{t-1} = (t+1)/t = 1 + 1/t$ che converge ancora a 1. In conclusione, per $t \rightarrow \infty$ i parametri del modello (4.29) convergono rispettivamente a

$$\begin{aligned}\ell(t, t) \rightarrow \ell_0 &= \begin{cases} \sigma & \text{se } |a| \leq 1 \\ |a| \sigma & \text{se } |a| > 1 \end{cases} \\ \ell(t, t-1) \rightarrow \ell_1 &= \begin{cases} a\sigma & \text{se } |a| \leq 1 \\ \frac{a}{|a|} \sigma & \text{se } |a| > 1 \end{cases}\end{aligned}$$

e il modello d'innovazione diventa “asintoticamente stazionario”,

$$\mathbf{y}(t) = \ell_0 \boldsymbol{\varepsilon}(t) + \ell_1 \boldsymbol{\varepsilon}(t-1) \quad .$$

In realtà questo tipo di comportamento asintotico della risposta impulsiva $\ell(t, \cdot)$ si verifica per una classe molto ampia di processi stazionari. Su di esso è fondato un famoso algoritmo di fattorizzazione spettrale di Bauer [14] e un approccio al filtraggio ricorsivo dovuto a Rissanen e Barbosa [123]. Si veda il Teorema 4.3 del capitolo V più avanti. \diamond

Risoluzione di equazioni algebriche con la fattorizzazione di Cholesky. La fattorizzazione di Cholesky viene correntemente impiegata per risolvere numericamente sistemi di equazioni algebriche del tipo:

$$Q\mathbf{x} = b \quad (4.30)$$

dove Q è *simmetrica e definita positiva*. Equazioni di questo genere sono molto comuni in statistica e teoria della stima, dove tipicamente Q è una matrice di varianza.

Per risolvere la (4.30) si può usare il seguente algoritmo.

- (1) Si calcola la fattorizzazione di Cholesky

$$Q = LL' \quad .$$

- (2) Posto $L'\mathbf{x} = z$ (z è una nuova incognita) si risolve il sistema lineare

$$Lz = b \quad .$$

Siccome L è triangolare inferiormente, questo calcolo è immediato: la soluzione si ottiene per sostituzioni successive dall'alto.

- (3) Determinato z , si risolve la

$$L'\mathbf{x} = z \quad .$$

Ancora, la soluzione è immediata: dato che L' è triangolare superiormente il calcolo si fa per sostituzioni successive dal basso.

Questo algoritmo può in particolare essere usato per calcolare l'inversa di Q ; basta allo scopo risolvere colonna per colonna la

$$Lz = I$$

che fornisca $\mathbf{z} = L^{-1}$ (ancora di struttura triangolare inferiore) e pertanto

$$Q^{-1} = \mathbf{z}'\mathbf{z} \quad .$$

Un'altra applicazione dell'algoritmo è nel calcolo della matrice $A = \Sigma_{\mathbf{x}\mathbf{y}} \Sigma_{\mathbf{y}}^{-1}$ che individua lo stimatore lineare a m.v. di \mathbf{x} in base a \mathbf{y} . Dal principio di ortogonalità scende infatti che A è individuata dalla

$$\Sigma_{\mathbf{x}\mathbf{y}} = A \Sigma_{\mathbf{y}} \quad (4.31)$$

e la trasposta di questa relazione è appunto di struttura identica alla (4.30) (ovviamente considerata colonna per colonna).

È istruttivo scrivere la "soluzione" della (4.31) corrispondente all'introduzione della fattorizzazione di Cholesky $\Sigma_{\mathbf{y}} = LL'$. Si trova

$$A = (\Sigma_{\mathbf{x}\mathbf{y}} L^{-\top}) L^{-1} \quad . \quad (4.32)$$

Questa formula verrà generalizzata al caso in cui \mathbf{x} e \mathbf{y} sono due processi stocastici (stazionari) nel Capitolo VII. Il lettore è invitato a notare che lo stimatore (4.32) esegue preliminarmente uno "sbiancamento" delle osservazioni \mathbf{y} attraverso il fattore L^{-1} e calcola successivamente lo stimatore di \mathbf{x} in base al vettore delle innovazioni $\boldsymbol{\varepsilon}$ (infatti $\Sigma_{\mathbf{x}\mathbf{y}} L^{-1} = \Sigma_{\mathbf{x}\boldsymbol{\varepsilon}}$).

L'idea della ortonormalizzazione (o sbiancamento) preliminare delle osservazioni sta, come vedremo presto, alla base di tutta la teoria "dinamica" della stima.

5. Componenti Principali e sviluppo di Karhunen-Loève

Il problema che vogliamo studiare in questo capitolo è quello dell'approssimazione di segnali aleatori mediante combinazioni lineari finite di andamenti temporali deterministici, che chiameremo “modi”, e della scelta ottima di questi modi ai fini di una rappresentazione il più possibile parsimoniosa del segnale. Le applicazioni di queste tecniche riguardano la compressione di dati e il riconoscimento di configurazioni.

Sviluppo in Componenti Principali. Inizieremo studiando il problema per un tratto finito di un segnale scalare a tempo discreto, $\{\mathbf{y}(1), \dots, \mathbf{y}(t), \dots, \mathbf{y}(T)\}$, i cui campioni penseremo organizzati in un vettore aleatorio T -dimensionale \mathbf{y} , a media nulla di matrice covarianza nota Σ , definita positiva.

Consideriamo una qualunque base ortonormale $\{\mathbf{x}_k; k = 1, \dots, T\}$ di variabili aleatorie scalari in $\mathbf{H}(\mathbf{y})$ e uno sviluppo del tipo

$$\mathbf{y} = \sum_{k=1}^T \alpha_k \mathbf{x}_k u_k, \quad u_k \in \mathbb{R}^T \quad (5.1)$$

dove α_k sono numeri reali e gli u_k dei vettori deterministici che debbono evidentemente essere vincolati dalla condizione

$$\alpha_k u_k := E\{\mathbf{y}\mathbf{x}_k\}, \quad k = 1, \dots, T.$$

PROPOSIZIONE 2.3. *Il vettore aleatorio \mathbf{y} è rappresentabile mediante uno sviluppo del tipo (5.1) in cui i vettori deterministici u_k sono ortonormali, i.e. $u'_k u_j = \delta_{k,j}$, se e solo se essi formano un sistema di autovettori normalizzati per la matrice di covarianza, Σ , di \mathbf{y} . In tal caso, detto $\lambda_k > 0$ l'autovalore di Σ corrispondente all'autovettore u_k , si ha $\alpha_k = \sqrt{\lambda_k}$ e le variabili aleatorie \mathbf{x}_k sono date dalla formula*

$$\mathbf{x}_k = \frac{1}{\sqrt{\lambda_k}} u'_k \mathbf{y} = \frac{1}{\sqrt{\lambda_k}} \sum_{t=1}^T u_k(t) \mathbf{y}(t), \quad k = 1, \dots, T. \quad (5.2)$$

DIMOSTRAZIONE. Sufficienza: se gli u_k sono gli autovettori normalizzati di Σ , dalla

$$\Sigma u_k = \lambda_k u_k, \quad k = 1, \dots, T$$

si vede facilmente che le variabili \mathbf{x}_k definite in (5.2) sono ortonormali. Infatti

$$E\mathbf{x}_k \mathbf{x}_j = \frac{1}{\sqrt{\lambda_k \lambda_j}} u'_k \Sigma u_j = \delta_{k,j}$$

Dato che le T variabili \mathbf{x}_k appartengono a $\mathbf{H}(\mathbf{y})$, esse formano una base ortonormale per questo spazio. Scrivendo $\mathbf{x} := [\mathbf{x}_1, \dots, \mathbf{x}_T]'$, si ha pertanto $\mathbf{H}(\mathbf{y}) = \mathbf{H}(\mathbf{x})$, cosicché

$$\mathbf{y} = \hat{E}[\mathbf{y} | \mathbf{x}] = \sum_{k=1}^T E\{\mathbf{y}\mathbf{x}_k\} \mathbf{x}_k$$

e siccome $E\{\mathbf{y}\mathbf{x}_k\} = \frac{1}{\sqrt{\lambda_k}} E\{\mathbf{y}\mathbf{y}'\} u_k = \sqrt{\lambda_k} u_k$ si ritrova una rappresentazione del tipo (5.1) in cui $\alpha_k = \sqrt{\lambda_k}$.

Necessità: Calcolando la matrice di covarianza di \mathbf{y} in base alla (5.1) si trova

$$\Sigma = \sum_{k=1}^T u_k E[\mathbf{x}_k]^2 u'_k := U \text{diag} \{\lambda_1, \dots, \lambda_T\} U'$$

dove $U := [u_1, \dots, u_T]$ è una matrice ortogonale e pertanto deve necessariamente coincidere con la matrice degli autovettori normalizzati di Σ . \square

I vettori u_k vengono chiamati *componenti principali* o, più descrittivamente, *modi propri* del segnale. In seguito supporremo sempre di organizzare gli autovalori di Σ (e quindi anche i rispettivi coefficienti aleatori \mathbf{x}_k dei modi u_k) in ordine *decrescente*, i.e.

$$\lambda_1 \geq \dots \geq \lambda_T > 0$$

Con questa convenzione, lo sviluppo “biortonormale” (5.1), dove naturalmente $\alpha_k = \sqrt{\lambda_k}$ e le \mathbf{x}_k sono ottenute mediante le (5.2), si chiama *Sviluppo in Componenti Principali* o anche *Sviluppo di Karhunen-Loève* del segnale \mathbf{y} .

In molte situazioni accade che l’ “energia statistica” del segnale $E \|\mathbf{y}\|^2 = \sum_{k=1}^T \lambda_k E\{\mathbf{x}_k\}^2 = \sum_{k=1}^T \lambda_k$ è concentrata su pochi modi propri. In altri termini c’è un indice n per cui risulta

$$\lambda_1 + \dots + \lambda_n \gg \lambda_{n+1} + \dots + \lambda_T$$

e gli autovalori di indice maggiore di n si possono ritenere all’incirca uguali a zero, di modo che, per gli scopi che si hanno in vista,

$$\mathbf{y} \simeq \hat{\mathbf{y}}_n := \sum_{k=1}^n \sqrt{\lambda_k} \mathbf{x}_k u_k = \sum_{k=1}^n (u_k' \mathbf{y}) u_k \quad (5.3)$$

Questa formula si usa per eseguire compressione di dati ai fini della trasmissione, dell’immagazzinamento (*data storage*) o della classificazione di segnali. In pratica la matrice di covarianza Σ non è nota a priori. Però, se si dispone di un certo numero di misure dello stesso segnale, diciamole $\{y_1, \dots, y_N\}$, si può usare una stima della covarianza, ottenuta ad esempio per mezzo di una formula del tipo

$$\hat{\Sigma} = \frac{1}{N} \sum_{k=1}^N y_k y_k'$$

e poi, una volta calcolati gli autovalori dominanti e i corrispondenti autovettori (normalizzati), si effettua la compressione di una arbitraria traiettoria osservata, y , del segnale che si vuole codificare o classificare, calcolandone i coefficienti $\{x_k; k = 1, \dots, n\}$ che la rappresentano nella decomposizione (5.3).

Notiamo che l’ errore di approssimazione $\tilde{\mathbf{y}}_n := \mathbf{y} - \hat{\mathbf{y}}_n$ è ortogonale a \mathbf{y} , per cui

$$\Sigma = \text{Var } \mathbf{y} = \text{Var } \hat{\mathbf{y}}_n + \text{Var } \tilde{\mathbf{y}}_n := \hat{\Sigma}_n + \tilde{\Sigma}_n$$

e quindi la varianza dell’approssimazione si può esprimere come

$$\hat{\Sigma}_n = \sum_{k=1}^n \lambda_k u_k E\{\mathbf{x}_k^2\} u_k' = \sum_{k=1}^n \lambda_k u_k u_k'$$

Da questa espressione risulta immediatamente che

PROPOSIZIONE 2.4. *La matrice varianza di $\hat{\mathbf{y}}_n$ è la miglior approssimazione (in norma ℓ^2 o in norma di Frobenius) di rango n , della matrice varianza del vettore \mathbf{y} .*

DIMOSTRAZIONE. L’affermazione segue facilmente dalla proprietà di approssimazione ottima della SVD troncata di una matrice qualunque, enunciata nel corollario A.1 dell’ Appendice A. Basta riconoscere che per una matrice simmetrica la SVD coincide con l’ordinaria decomposizione in autovettori e autovalori. \square

In realtà non è evidente se il procedimento di approssimazione suggerito più sopra goda di analoghe proprietà di ottimalità dal punto di vista della rappresentazione del segnale stesso. Cerchiamo di capire innanzitutto quale sarebbe un criterio di approssimazione naturale e ragionevole da usare. Intanto, è evidente che il secondo membro della (5.3) può essere visto come una trasformazione lineare rappresentata da una matrice (deterministica) M , simmetrica, semidefinita positiva e di rango n , che opera sul vettore aleatorio \mathbf{y} . Una M di questo genere si può sempre scrivere nella forma $M = WW'$ con W di dimensione $T \times n$ di rango colonne pieno. Il fatto che la matrice M abbia rango n ($\leq T$), implica che l'approssimazione $\hat{\mathbf{y}} := M\mathbf{y}$, genera un sottospazio n -dimensionale di $\mathbf{H}(\mathbf{y})$. Possiamo così dire che la (5.3) fornisce un'approssimazione $\hat{\mathbf{y}} := M\mathbf{y}$ di rango n , di \mathbf{y} .

Motivati da quanto appena visto, consideriamo il problema di approssimazione ottima di rango n del vettore aleatorio \mathbf{y} , che ha la naturale formulazione seguente. Trovare una matrice $M \in \mathbb{R}^{T \times T}$ di rango n , che minimizza l'errore quadratico medio di approssimazione, ovvero risolvere il problema di minimo

$$\min_{\text{rango}(M) = n} E\{\|\mathbf{y} - M\mathbf{y}\|^2\} \quad (5.4)$$

Da notare che in questo problema si ricerca in sostanza il sottospazio “ottimo” di $\mathbf{H}(\mathbf{y})$ su cui “proiettare”, per cui si ha a che fare con un problema di approssimazione diverso da quello che abbiamo considerato all'inizio del capitolo, quando abbiamo introdotto gli stimatori lineari a minima varianza.

Come accade per la (5.3), è naturale richiedere che l'approssimazione $M\mathbf{y}$ sia scorrelata dal residuo di stima, ovvero

$$\mathbf{y} - M\mathbf{y} \perp M\mathbf{y}. \quad (5.5)$$

Questa condizione è equivalente alla

$$M\Sigma - M\Sigma M' = 0$$

ovvero, introducendo una qualunque radice quadrata $\Sigma^{1/2}$ di Σ e definendo $\hat{M} := \Sigma^{-1/2}M\Sigma^{1/2}$, alla

$$\hat{M} = \hat{M}\hat{M}'$$

la quale implica che \hat{M} dev'essere simmetrica e *idempotente* (i.e. $\hat{M} = \hat{M}^2$) ovvero una *proiezione ortogonale* da \mathbb{R}^T su un suo sottospazio n -dimensionale. Quindi M deve avere la struttura

$$M = \Sigma^{1/2}\Pi\Sigma^{-1/2}, \quad \Pi = \Pi^2 \quad \Pi = \Pi' \quad (5.6)$$

dove $\Sigma^{1/2}$ è una qualunque radice quadrata di Σ e Π è una matrice proiezione ortogonale di rango n .

TEOREMA 2.9. *Le soluzioni del problema di minimo (5.4) che soddisfano alla condizione di ortogonalità (5.5) sono del tipo*

$$M = WW', \quad W = U_nQ_n$$

dove U_n è la matrice $T \times n$ le cui colonne sono i primi n autovettori normalizzati di Σ , ordinati rispetto all'ordinamento in senso decrescente dei corrispondenti autovalori e Q_n è una arbitraria matrice ortogonale $n \times n$.

DIMOSTRAZIONE. Sia $\Lambda := \text{diag}\{\lambda_1, \dots, \lambda_T\}$ e $\Sigma = U\Lambda U'$ la decomposizione spettrale di Σ in cui U è la matrice ortogonale degli autovettori. Possiamo scegliere come radice quadrata di Σ la matrice $\Sigma^{1/2} := U\Lambda^{1/2}$.

Notiamo che il vettore aleatorio $\mathbf{e} := \Sigma^{-1/2} \mathbf{y}$ ha (comunque si scelga $\Sigma^{1/2}$) componenti ortonormali, per cui, usando la struttura (5.6) la funzione costo del nostro problema di minimo si può riscrivere

$$\begin{aligned} E\{\|\mathbf{y} - M\mathbf{y}\|^2\} &= E\{\|\Sigma^{1/2}\mathbf{e} - \Sigma^{1/2}\Pi\Sigma^{-1/2}\mathbf{y}\|^2\} = E\{\|\Sigma^{1/2}(\mathbf{e} - \Pi\mathbf{e})\|^2\} = \\ &E\{\|\Lambda^{1/2}(\mathbf{e} - \Pi\mathbf{e})\|^2\} = E(\mathbf{e} - \Pi\mathbf{e})'\Lambda(\mathbf{e} - \Pi\mathbf{e}) = \text{Tr}[\Lambda E(\mathbf{e} - \Pi\mathbf{e})(\mathbf{e} - \Pi\mathbf{e})'] \end{aligned}$$

dove $\text{Tr} A := \sum a_{kk}$ sta per traccia della matrice A . Ne segue così che il nostro problema di minimo si può riscrivere equivalentemente come

$$\min_{\text{rango}(\Pi) = n} \text{Tr}\{\Lambda\Pi^\perp\}$$

dove $\Pi^\perp := I - \Pi$ è la proiezione ortogonale sul sottospazio complemento ortogonale di $\text{Im } \Pi$.

Assumendo sempre che gli autovalori siano ordinati in senso decrescente, i.e. $\{\lambda_1 \geq \dots \geq \lambda_T\}$, si vede bene che il minimo di questa funzione di Π si raggiunge quando Π proietta sullo spazio generato dai primi n assi coordinati. Insomma $\Pi_{\text{ottimo}} = \text{diag}\{I_n, 0\}$ e il minimo vale $\lambda_{n+1} + \dots + \lambda_T$. È evidente allora che

$$M = U\Lambda^{1/2}\Pi_{\text{ottimo}}\Lambda^{-1/2}U' = U_nU_n'$$

Naturalmente la moltiplicazione di U_n per una qualunque matrice ortogonale $n \times n$ non cambia il risultato. \square

Questo risultato conferma che lo sviluppo troncato di Karhunen-Loève (5.3) è ottimo anche dal punto di vista della rappresentazione del segnale. Da notare che la rappresentazione richiede solo la conoscenza dei primi n autovettori (o modi) principali di Σ e non dei corrispondenti autovalori. La caratterizzazione della rappresentazione come soluzione di un problema di ottimizzazione, può essere sfruttata vantaggiosamente dal punto di vista computazionale. Si veda ad esempio [159].

Approssimazione di segnali a tempo continuo*. Per tradizione lo sviluppo di Karhunen-Loève è normalmente introdotto in letteratura nel caso di segnali a tempo continuo. Sebbene da un punto di vista concettuale non ci siano novità, in questo caso la decomposizione spettrale della matrice covarianza Σ deve essere rimpiazzata da uno sviluppo in autofunzioni e la trattazione diviene necessariamente un pò più tecnica.

Consideriamo allora un processo scalare a tempo continuo $\mathbf{y} := \{\mathbf{y}(t); t \in T\}$, dove ora la variabile t è una variabile reale a valori in un certo intervallo temporale T . Supponiamo che il processo abbia media nulla e che la sua funzione di covarianza $R(t, s) := E\{\mathbf{y}(t)\mathbf{y}(s)\}$ sia *continua in entrambi gli argomenti*. Questa condizione è equivalente alla continuità in media quadratica del processo. Supponiamo che valga la condizione

$$\int_T \int_T R(t, s)^2 dt ds < \infty \quad (5.7)$$

(che, se T è un intervallo finito è automaticamente soddisfatta). Consideriamo, sullo spazio dei segnali a energia finita, diciamolo $C^2[T]$, che sono funzioni continue in T , dotato del prodotto interno $\langle f, g \rangle := \int_T f(t)g(t) dt$, l'operatore lineare Σ_R definito dalla posizione

$$[\Sigma_R f](t) := \int_T R(t, s) f(s) ds \quad (5.8)$$

Si può verificare che questo operatore mappa segnali di $C^2[T]$ in segnali dello stesso tipo. Gli *autovalori* e le corrispondenti *autofunzioni* di un operatore integrale di questo tipo, sono le coppie λ, φ con $0 < \|\varphi\|_{L^2[T]} < \infty$, che soddisfano alla

$$[\Sigma_R \varphi](t) = \int_T R(t, s) \varphi(s) ds = \lambda \varphi(t) \quad t \in T \quad (5.9)$$

La novità rispetto al caso a dimensioni finite è che non è affatto pacifico che gli autovalori esistano. In generale lo spettro di un operatore lineare potrebbe anche essere di tipo continuo e le autofunzioni mancare del tutto. Sotto la condizione (5.7), l'operatore Σ_R ha però la proprietà notevole di essere un *operatore compatto* (o *completamente continuo* come si dice nella letteratura russa, vedere ad esempio [2]) e si mostra che vale il seguente notevole risultato, che generalizza la decomposizione spettrale delle matrici simmetriche e semidefinite positive.

TEOREMA 2.10 (Mercèr). *Nelle ipotesi poste, valgono le seguenti affermazioni:*

- (1) *Il problema agli autovalori (5.9) ha soluzioni (esiste almeno un autovalore). L'autovalore massimo λ_0 è dato dalla formula*

$$\lambda_0 = \max_{\|\varphi\|_{L^2[T]}=1} \langle \Sigma_R \varphi, \varphi \rangle = \max_{\|\varphi\|_{L^2[T]}=1} \int_T \int_T R(t, s) \varphi(t) \varphi(s) dt ds \quad (5.10)$$

L'autofunzione corrispondente, $\varphi_0(t)$, è una funzione continua.

- (2) *La funzione $R_1(t, s) := R(t, s) - \lambda_0 \varphi_0(t) \varphi_0(s)$ è ancora una funzione covarianza (di tipo positivo) che soddisfa la condizione (5.7). Quindi il problema agli autovalori*

$$[\Sigma_{R_1} \varphi](t) := \int_T R_1(t, s) \varphi(s) ds = \lambda \varphi(t) \quad (5.11)$$

ha ancora un autovalore massimo, λ_1 , dato dalla formula

$$\lambda_1 = \max_{\|\varphi\|_{L^2[T]}=1} \langle \Sigma_{R_1} \varphi, \varphi \rangle = \max_{\|\varphi\|_{L^2[T]}=1} \int_T \int_T R_1(t, s) \varphi(t) \varphi(s) dt ds \quad (5.12)$$

e risulta $\lambda_1 \leq \lambda_0$.

- (3) *Il procedimento può essere iterato. Gli autovalori del problema (5.9) formano una sequenza di numeri positivi (non necessariamente distinti), monotona decrescente, il cui unico punto di accumulazione può essere 0. Le corrispondenti autofunzioni sono tutte continue a energia finita e possono essere rese ortonormali, i.e.*

$$\int_T \varphi_k(t) \varphi_j(t) dt = \delta_{k,j}$$

- (4) *La funzione covarianza ammette la rappresentazione*

$$R(t, s) = \sum_{k=0}^{\infty} \lambda_k \varphi_k(t) \varphi_k(s), \quad t, s \in T \times T \quad (5.13)$$

dove la serie converge puntualmente, uniformemente in $T \times T$.

Si può anche dimostrare che l'approssimazione di rango $n + 1$ della funzione covarianza ottenuta troncando lo sviluppo (5.13) ai primi $n + 1$ termini,

$$R(t, s) \simeq R_n(t, s) := \sum_{k=0}^n \lambda_k \varphi_k(t) \varphi_k(s), \quad (5.14)$$

è la migliore possibile in una varietà di sensi. Per esempio, l'operatore lineare Σ_{R_n} definito dalla

$$[\Sigma_{R_n} f](t) := \int_T R_n(t, s) f(s) ds$$

è il miglior approssimante di rango $n + 1$ di Σ_R , nel senso che risolve il problema

$$\min_{\text{rango}(\Sigma) = n+1} \|\Sigma_R - \Sigma\| \quad (5.15)$$

e il minimo è uguale a λ_{n+1} , il primo autovalore trascurato. La norma di un operatore lineare è definita in (0.2), (0.3) dell' Appendice B.

Il seguente risultato è l'analogo a tempo continuo della Proposizione 2.3

PROPOSIZIONE 2.5. *Se la funzione covarianza del segnale aleatorio \mathbf{y} è continua e soddisfa la (5.7), allora \mathbf{y} è rappresentabile mediante lo sviluppo*

$$\mathbf{y}(t) = \sum_{k=0}^{+\infty} \sqrt{\lambda_k} \varphi_k(t) \mathbf{x}_k \quad (5.16)$$

convergente in media quadratica, uniformemente rispetto a $t \in T$, dove λ_k ; $k = 0, 1, \dots$ sono gli autovalori dell'operatore Σ_R , ordinati in senso decrescente, φ_k ; $k = 0, 1, \dots$ le corrispondenti autofunzioni normalizzate e le variabili aleatorie \mathbf{x}_k ; $k = 0, 1, \dots$ sono definite dalla

$$\mathbf{x}_k = \frac{1}{\sqrt{\lambda_k}} \int_T \varphi_k(t) \mathbf{y}(t) dt \quad (5.17)$$

Le variabili \mathbf{x}_k ; $k = 0, 1, \dots$ formano una base ortonormale per lo spazio di Hilbert $\mathbf{H}(\mathbf{y})$, generato dal processo \mathbf{y} .

Come nel caso di segnali discreti, la rappresentazione (5.16) viene normalmente troncata ad un numero finito di termini. In effetti, si può dimostrare, in modo analogo a quanto fatto per il teorema 2.9, che lo sviluppo finito

$$\mathbf{y}_n(t) = \sum_{k=0}^n \sqrt{\lambda_k} \varphi_k(t) \mathbf{x}_k \quad (5.18)$$

fornisce, per ogni $t \in T$, la migliore approssimazione di rango $n + 1$ di $\mathbf{y}(t)$ nel senso del minimo errore quadratico medio. Questo significa in particolare che, ad esempio, l'analisi (in serie) di Fourier del segnale ha in generale peggiori proprietà di approssimazione dello sviluppo (5.18) a parità di numero di termini.

La bontà dell'approssimazione si può misurare in termini di energia statistica. Dato che

$$E \int_T \mathbf{y}(t)^2 dt = \int_T E \mathbf{y}(t)^2 dt = \sum_{k=0}^{+\infty} \lambda_k$$

e l'energia di \mathbf{y}_n è data da un'analoga espressione in cui la somma ha solo i primi $n + 1$ termini, si vede bene che l'errore che si commette approssimando \mathbf{y} con \mathbf{y}_n dipende da quanto velocemente tendono a zero gli autovalori di Σ_R . In termini relativi l'energia dell'errore può così essere espressa come il rapporto

$$\frac{\sum_{k=n+1}^{+\infty} \lambda_k}{\sum_{k=0}^{+\infty} \lambda_k} = 1 - \frac{\sum_{k=0}^n \lambda_k}{\sum_{k=0}^{+\infty} \lambda_k}$$

Sviluppo di K-L di processi stazionari. Se \mathbf{y} è un processo stazionario si ha $R(t, s) = R(t - s)$ ed è facile convincersi che se T è un intervallo illimitato, ad esempio $T = [0, +\infty)$, la condizione (5.7) non può essere verificata. Quindi *lo sviluppo di K-L non esiste se il processo è stazionario e l'intervallo T è illimitato*. Se ci si restringe ad un intervallo finito, ad esempio $T = [-a, a]$, si può sviluppare la funzione covarianza in serie di Fourier,

$$\Sigma(\tau) = \sum_{k=0}^{+\infty} \sigma_k \cos \frac{k\pi\tau}{a}$$

Sostituendo nell'equazione integrale (5.9) e tenendo conto dell'ortogonalità delle funzioni coseno, si vede immediatamente che gli autovalori sono semplicemente i coefficienti di Fourier, ovvero $\lambda_k = \sigma_k$ e le autofunzioni normalizzate sono

$$\varphi_k(t) = \frac{1}{\sqrt{a}} \cos \frac{k\pi t}{a}$$

Quindi le variabili aleatorie \mathbf{x}_k sono i coefficienti (aleatori) di Fourier del segnale \mathbf{y} e si ritrova in sostanza la rappresentazione del processo mediante il suo sviluppo in serie di Fourier. Lo sviluppo di Karhunen-Loève può in questo caso essere visto come una generalizzazione della trasformata di Fourier.

6. Stima di Direzioni

In questa ultima sezione discuteremo una classe di problemi di stima Bayesiana che hanno una certa importanza nelle applicazioni e che non si possono trattare efficacemente mediante la teoria della stima lineare a minima varianza. Si tratta di problemi in cui la variabile da stimare è una *direzione*.

Nelle applicazioni più semplici si tratta di puntare ad un bersaglio (“target”), ad esempio un punto P , che può anche muoversi aleatoriamente in \mathbb{R}^3 . Per ipotesi il sensore a disposizione non può misurare la distanza del punto dal centro di osservazione O (convenzionalmente coincidente con l'origine del sistema di coordinate) e si può al più pensare di ricostruire il versore del raggio OP che congiunge il centro ottico al target. Possiamo in sostanza pensare di descrivere sia la variabile osservata che le misure fornite dallo strumento come *direzioni* e rappresentarle mediante vettori \mathbf{x} e \mathbf{y}_k appartenenti alla sfera unitaria centrata in O . In questa formulazione \mathbf{x} è la direzione incognita del punto P in \mathbb{R}^3 e \mathbf{y}_k , $k = 1, \dots, m$ (sempre vettori della sfera unitaria) sono misure rumorose della direzione “vera” \mathbf{x} . Si potrebbe dire che il problema è formulato come un problema di *stima sulla sfera unitaria*. In generale la sfera unitaria in \mathbb{R}^n si denota col simbolo \mathbb{S}^{n-1} . Essa è definita come l'insieme dei vettori di \mathbb{R}^n di norma Euclidea unitaria ed è quindi una superficie di dimensione $n - 1$.

Può capitare di avere a che fare con problemi di stima di direzioni quando si impiegano antenne a schiere oppure quando il sensore è di tipo ottico, e in particolare in problemi di ricostruzione di scene e/o del moto a partire da immagini acquisite mediante telecamere (visione computazionale). Normalmente in questi problemi le osservazioni sono dati angolari rumorosi, relativi alla direzione di un certo “target” che si cerca di inseguire. In questo capitolo ci limiteremo però a trattare solamente il problema di stima della direzione di un target fisso.

È evidente che in un problema di questa natura il modo in cui il rumore di misura entra nelle osservazioni per distorcerne il valore “ideale” \mathbf{x} , non può essere

additivo e una formulazione realistica di questo problema dovrà necessariamente scostarsi dal classico paradigma di “rumore additivo a media nulla”.

La Distribuzione di Langevin. Una famiglia molto semplice di distribuzioni di probabilità sulla sfera \mathbb{S}^2 è definita dalla *densità di Langevin*

$$p(x) = \frac{\kappa}{4\pi \sinh \kappa} \exp \kappa \mu' x, \quad \|x\| = 1 \quad (6.1)$$

rispetto alla misura d'area elementare $d\sigma_x$, della superficie sferica. In coordinate sferiche $x = (\theta, \phi)$, $d\sigma = \sin \theta d\theta d\phi$. Il parametro vettoriale $\mu \in \mathbb{S}^2$ (μ è convenzionalmente normalizzato a lunghezza unitaria) è la *moda* della distribuzione, mentre il numero positivo $\kappa > 0$ si chiama la *concentrazione* della distribuzione. Per $\kappa \rightarrow 0$ la densità tende alla distribuzione uniforme, mentre per $\kappa \rightarrow \infty$, p tende alla distribuzione di Dirac concentrata in $x = \mu$. La funzione densità (6.1), che si denoterà col simbolo $L(\mu, \kappa)$, fu introdotta da Langevin (1905) nel suo modello statistico del magnetismo [91]. Dopo Langevin, questa distribuzione è stata riscoperta numerose volte e ri-attribuita ad esempio a von Mises e ad altri autori [146]. Ciò che rende la distribuzione di Langevin utile in statistica Bayesiana è il fatto che si tratta di una distribuzione di classe esponenziale, la cui forma funzionale si preserva per moltiplicazione, similmente a quanto accade per la Gaussiana. Notiamo anche che, in analogia alla proprietà della distribuzione Gaussiana, di preservare la forma funzionale rispetto ad operazioni lineari (la somma di variabili Gaussiane è Gaussiana), la distribuzione di Langevin si preserva per *rotazioni*, nel senso che, se $\mathbf{x} \sim L(\mu, \kappa)$ e $\mathbf{y} = R\mathbf{x}$ con R matrice di rotazione ($RR' = I$) allora $\mathbf{y} \sim L(R'\mu, \kappa)$, come si verifica facilmente in base alle note regole del calcolo delle probabilità.

Introducendo un sistema di coordinate sferiche con polo la direzione della moda, (6.1) può essere riscritta nella forma

$$p(\theta, \phi) = \frac{\kappa}{4\pi \sinh \kappa} \exp \kappa \cos \theta \quad 0 \leq \theta \leq \pi.$$

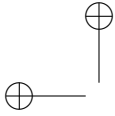
che rende evidente la simmetria di rotazione di $L(\mu, \kappa)$ rispetto alla direzione della moda μ .

La (6.1) è l'espressione della distribuzione di Langevin sulla sfera unitaria di \mathbb{R}^3 . Per dimensioni superiori la costante di normalizzazione ha un'espressione più complicata. La distribuzione su \mathbb{S}^{n-1} , $n \geq 3$, ha densità

$$p(x) = \frac{\kappa^{(n/2-1)}}{(2\pi)^{n/2} I_{n/2-1}(\kappa)} \exp \kappa \mu' x, \quad \|x\| = 1 \quad (6.2)$$

rispetto alla misura elementare di superficie sferica, dove $I_{n/2-1}(x)$ è la funzione di Bessel modificata del primo tipo. Più in generale, un'arbitraria densità su \mathbb{S}^{n-1} può essere espressa come l'esponenziale di uno sviluppo in armoniche sferiche, che sono funzioni ortonormali sulla sfera unitaria trattate in molti testi, ad esempio in [146, p. 80-88]. Dato che nel sistema di coordinate sferiche la prima armonica sferica è proprio $\cos \theta$, la distribuzione di Langevin può essere pensata come un' approssimazione del primo ordine di una distribuzione sferica generale.

La Distribuzione Gaussiana Angolare. Come abbiamo già accennato, alcune proprietà della distribuzione di Langevin sono l'analogo sulla sfera delle proprietà delle distribuzioni Gaussiane su spazi lineari Euclidei. Ci sono vari tentativi in letteratura di derivare assiomaticamente la distribuzione di Langevin dalla Gaussiana. Probabilmente il risultato più semplice di questo tipo è l'osservazione,



apparentemente dovuta a Fisher [47], che la distribuzione di un vettore aleatorio Gaussiano \mathbf{x} con distribuzione *isotropa* $\mathcal{N}(m, \sigma^2 I)$, condizionata dall'evento $\{\|\mathbf{x}\| = 1\}$, è Langevin con moda $\mu = m/\|m\|$ e concentrazione $\kappa = \|m\|/\sigma^2$.

Un risultato molto più utile, che discuteremo qui sotto, è la sostanziale identità della cosiddetta distribuzione *Gaussiana angolare* alla Langevin.

La distribuzione Gaussiana angolare (che si denota col simbolo Ag) è la densità di probabilità del versore di direzione $\mathbf{x} := \boldsymbol{\xi}/\|\boldsymbol{\xi}\|$ di un vettore $\boldsymbol{\xi}$ che ha distribuzione Gaussiana *isotropa*, i.e. $\boldsymbol{\xi} \sim \mathcal{N}(m, \sigma^2 I)$. La distribuzione si ottiene calcolando la marginale di $\mathcal{N}(m, \sigma^2 I)$ sulla sfera unitaria $\|x\| = 1$. È mostrato nell'Appendice C di [146] che la Gaussiana angolare è una combinazione convessa di distribuzioni di Langevin con parametro di concentrazione κ , variabile

$$Ag(x) = N \int_0^{+\infty} \kappa^{n-1} e^{-\frac{1}{2} \frac{\kappa^2}{\alpha^2}} e^{\kappa \mu' x} d\kappa,$$

dove $\mu = \frac{m}{\|m\|}$ $\alpha = \frac{\|m\|}{\sigma}$. Si vede da questa formula che Ag dipende dai parametri della Gaussiana m, σ^2 solamente attraverso μ e α^2 . Si denota perciò col simbolo $Ag(\mu, \alpha^2)$.

Questa notazione è conveniente perchè per valori moderati o grandi del parametro α , la densità $Ag(\mu, \alpha^2)$ è, per tutti gli scopi pratici, *coincidente* con la $L(\mu, \alpha^2)$ ³.

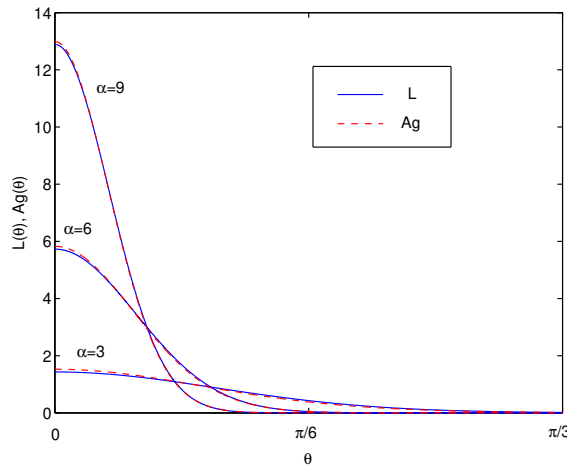


FIGURA 6.1. Distribuzioni Gaussiana angolare e di Langevin

In effetti, la distribuzione Gaussiana angolare approssima la distribuzione di Langevin anche per α piccoli, quando entrambe sono prossime alla distribuzione uniforme, ma in questo caso (del resto poco interessante) la relazione tra α e κ è diversa. In seguito noi riterremo generalmente che valga la $Ag(\mu, \alpha^2) = L(\mu, \alpha^2)$.

³“Moderati o grandi” qui significa che α^2 dev’essere maggiore di qualche decina, per avere un’approssimazione delle funzioni densità dell’ordine di qualche percento nelle rispettive ampezze.

Da notare che tutte le distribuzioni $\mathcal{N}(\rho m, \rho^2 \sigma^2 I)$, $\rho > 0$, danno origine alla stessa Gaussiana angolare. Questa è in effetti la famiglia di tutte le Gaussiane isotrope che generano la stessa Gaussiana angolare.

La Gaussiana angolare (e quindi la distribuzione di Langevin) hanno un ruolo naturale nel descrivere osservazioni direzionali, come è illustrato nell'esempio seguente.

Siano ξ, ζ due vettori aleatori Gaussiani isotropi con $\xi \sim \mathcal{N}(m, \sigma^2 I)$, $\zeta \sim \mathcal{N}(0, \sigma_z^2 I)$. Supponiamo di osservare la direzione del vettore

$$\eta = C\xi + \zeta \quad (6.3)$$

dove C è una matrice ortogonale ($CC' = I$). Se ξ, ζ fossero indipendenti la distribuzione di η sarebbe Gaussiana isotropa e la direzione $\mathbf{y} := \eta/\|\eta\|$ avrebbe una distribuzione Gaussiana angolare (i.e. Langevin) $\mathbf{y} \approx L(m/\|m\|, \frac{\|m\|^2}{\sigma^2 + \sigma_z^2})$.

In realtà, a prescindere da come ξ, ζ sono correlate, la densità condizionata $p(\mathbf{y} | \xi)$ è Gaussiana angolare. Questo accade perchè la distribuzione condizionata di η dato $\xi = \xi$ è Gaussiana di media $C\xi$ e varianza σ_z^2 . Quindi

$$p(\mathbf{y} | \xi = \xi) = Ag(C\xi/\|\xi\|, \|C\xi\|^2/\sigma_z^2) = Ag(Cx, \|\xi\|^2/\sigma_z^2)$$

dove x è il versore di ξ .

Quel che interessa in pratica è però la densità condizionata $p(\mathbf{y} | \mathbf{x})$. A questo proposito si ha il seguente risultato.

PROPOSIZIONE 2.6. *Se la perturbazione additiva ζ nel modello lineare (6.3) è Gaussiana isotropa e la sua varianza σ_z^2 , è proporzionale a $\|\xi\|^2$, i.e. $\sigma_z^2 = \sigma_0^2 \|\xi\|^2$, allora la densità condizionata $p(\mathbf{y} | \mathbf{x})$ del versore dell'osservazione (6.3) è Gaussiana Angolare.*

DIMOSTRAZIONE. Poniamo $r := \|\xi\|$ e notiamo che $p(\mathbf{y} | \xi = \xi) = p(\mathbf{y} | x, r)$. Dato che $p(\mathbf{y}, r | x) = p(\mathbf{y} | x, r)p(r | x)$, applicando una nota regola del calcolo delle probabilità si ha

$$p(\mathbf{y} | x) = \int_0^\infty r^2 p(\mathbf{y} | x, r) p(r | x) dr$$

e siccome nelle ipotesi poste, $p(\mathbf{y} | x, r)$ non dipende da r , si può portare fuori dall'integrale. \square

La dipendenza da ξ della varianza del rumore nel modello additivo (6.3), (che è uguale alla varianza condizionata di η dato $\xi = \xi$) si può definire *condizione di "rumore angolare"*. Da notare che questa condizione preclude l'indipendenza di ξ da ζ . Questo fatto, che a prima vista può sorprendere, è in realtà in accordo con la nozione intuitiva di rumore angolare. Ogni perturbazione infinitesima $d\xi$ di un vettore aleatorio ξ di norma fissata (deterministica) e direzione $\mathbf{x} \in \mathbb{S}^2$, che mantenga il vettore $\xi + d\xi$ della stessa norma (a meno di infinitesimi di ordine superiore), può essere rappresentata come l'effetto di una rotazione attorno ad un vettore polare infinitesimo $d\mathbf{p}$ ovvero

$$d\xi = d\mathbf{p} \wedge \xi = (d\mathbf{p} \wedge \mathbf{x}) \|\xi\| \quad (6.4)$$

dove il simbolo \wedge indica il prodotto vettoriale (o “esterno”)⁴. In questo modo il vettore $d\xi$ sta sul piano tangente alla sfera di raggio $r = \|\xi\|$, nel punto ξ ed ha ampiezza proporzionale alla norma $\|\xi\|$ (e quindi varianza proporzionale al suo quadrato). Come si vede, la perturbazione deve essere legata geometricamente alla direzione e al modulo del vettore ξ e non può quindi esserne indipendente.

Si può così affermare, sulla base della proposizione 2.6, che la distribuzione Gaussiana angolare (i.e. la Langevin) descrive statisticamente l’effetto di piccole perturbazioni angolari di una direzione fissa. In altre parole, per piccole perturbazioni angolari, la densità condizionata $p(y | \mathbf{x} = x)$ di osservare la direzione y sulla sfera unitaria, quando la “vera” direzione osservata è $\mathbf{x} = x$, è la densità di Langevin

$$p(y | \mathbf{x}) = \frac{\kappa}{4\pi \sinh \kappa} \exp \kappa \mathbf{x}' y \quad (6.5)$$

In questo contesto quindi si può sempre interpretare la distribuzione ordinaria $L(\mu, \kappa)$ come una distribuzione *condizionata* calcolata in base alla condizione $\mathbf{x} = \mu$. Il fatto che questa distribuzione sia unimodale e a simmetria di rotazione attorno alla direzione di partenza $\mathbf{x} = \mu$ si può interpretare dicendo che non c’è errore sistematico (*bias*) introdotto dallo strumento di misura.

Se la perturbazione angolare non è infinitesima (ci possono essere grandi errori angolari di osservazione) il ragionamento fatto più sopra non vale. Integrando la relazione (6.4) si trova che una perturbazione angolare *finita* di una direzione $\mathbf{x} \in \mathbb{S}^2$ può essere espressa dalla formula

$$\mathbf{y} := R(\mathbf{p}) \mathbf{x} \quad (6.6)$$

dove $R(\mathbf{p}) := \exp\{\mathbf{p}\wedge\}$ è una matrice ortogonale a determinante +1 e rappresenta quindi una rotazione (aleatoria) attorno al vettore polare \mathbf{p} . Le matrici di rotazione formano un gruppo che si denota col simbolo $SO(3)$.

L’uscita \mathbf{y} di un sensore di direzione rumoroso è quindi in generale rappresentabile da moltiplicazione per una matrice di rotazione aleatoria $R = R(\mathbf{p}) \in SO(3)$. In questo caso, la distribuzione condizionata può ovviamente essere un’arbitraria funzione non negativa sulla sfera unitaria.

I parametri (μ, κ) della distribuzione di Langevin possono essere espressi in funzione del vettore media m della distribuzione. Per la distribuzione su \mathbb{S}^2 valgono le

$$\mu = \frac{m}{\|m\|}, \quad \frac{\cosh \kappa}{\sinh \kappa} - \frac{1}{\kappa} = \|m\| \quad (6.7)$$

e si può verificare che queste formule definiscono una corrispondenza biunivoca tra m e (μ, κ) . In altre parole, in analogia a quanto accade con la Gaussiana, il parametro vettoriale m determina la distribuzione $L(\mu, \kappa)$ completamente. La proposizione seguente stabilisce una caratterizzazione della distribuzione di Langevin che è l’analogo su \mathbb{S}^{n-1} della ben nota caratterizzazione della Gaussiana, che si dimostra essere, tra tutte le distribuzioni sullo spazio Euclideo \mathbb{R}^n che hanno media e varianza fissate, quella a massima entropia.

⁴Se $p := [p_1 p_2 p_3]'$ il prodotto esterno $p\wedge$ è l’operatore lineare in \mathbb{R}^3 definito dalla matrice emisimmetrica

$$p\wedge = \begin{bmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{bmatrix}$$

PROPOSIZIONE 2.7. *Tra tutte le distribuzioni di probabilità sulla sfera unitaria che hanno vettore media fissato uguale ad m , la distribuzione di Langevin è quella di massima entropia.*

DIMOSTRAZIONE. L'entropia di una densità f (o di una distribuzione $dF(x) := f(x) d\sigma_x$) sulla sfera, è

$$H_f := - \int_{\mathbb{S}^{n-1}} \log f(x) f(x) d\sigma_x$$

Denotiamo per brevità la distribuzione di Langevin a media m col simbolo $l(x)$. Usando l'espressione (6.2) di l , si ricava

$$\begin{aligned} H_l &= - \int_{\mathbb{S}^{n-1}} \log l(x) l(x) d\sigma_x \\ &= - \log \frac{\kappa^{(n/2-1)}}{(2\pi)^{n/2} I_{n/2-1}(\kappa)} - \kappa \mu' m \\ &= - \int_{\mathbb{S}^{n-1}} \log l(x) f(x) d\sigma_x \end{aligned}$$

per un'arbitraria distribuzione f di media m . Segue che la differenza

$$\begin{aligned} H_l - H_f &= \\ &= - \int_{\mathbb{S}^{n-1}} \log l(x) f(x) d\sigma_x + \int_{\mathbb{S}^{n-1}} \log f(x) f(x) d\sigma_x \\ &= \int_{\mathbb{S}^{n-1}} \log \frac{f(x)}{l(x)} f(x) d\sigma_x \end{aligned}$$

è la cosiddetta *pseudo-distanza di Kullback- Leibler* di f da l (chiamata anche *divergenza* o *entropia relativa*, si veda [35] oppure l'articolo di Van Schuppen e Stoortvogel in [25, p. 314]. Si dimostra che questa pseudodistanza è positiva a meno che non sia $f = l$, nel qual caso è zero. Quindi $H_l \geq H_f$ per ogni f di media m . \square

Miglior approssimazione mediante una Distribuzione di Langevin. La distanza di Kullback- Leibler introdotta nel corso della dimostrazione precedente ha delle proprietà che la rendono una metrica naturale in problemi di approssimazione di densità di probabilità. Si veda ad es. [88].

Sia P una arbitraria misura di probabilità sulla sfera unitaria, assolutamente continua rispetto alla misura di superficie sferica elementare $d\sigma_x = \sin\theta d\theta d\varphi$. Vogliamo approssimare la densità $f(x) = dP/d\sigma_x$ con una densità del tipo Langevin, i.e. con una densità della classe parametrica

$$\mathcal{L} = \left\{ \ell(x) = \frac{\kappa}{4\pi \sinh(\kappa)} \exp\{\kappa \mu' x\} \quad , \quad \kappa \geq 0, \|\mu\| = 1 \right\}; \quad (6.8)$$

usando come criterio di approssimazione la (pseudo-)distanza di Kullback- Leibler. Dobbiamo quindi cercare il minimo di,

$$K(f, \ell_{(\mu, \kappa)}) = E_f \log \frac{f(x)}{\ell_{(\mu, \kappa)}(x)} = \int_{\mathbb{S}^2} \log \frac{f(x)}{\ell_{(\mu, \kappa)}(x)} f(x) d\sigma_x. \quad (6.9)$$

Il calcolo del minimo:

$$\min_{\{(\mu, \kappa) : \kappa \geq 0, \|\mu\|=1\}} K(f, \ell_{(\mu, \kappa)}) \quad (6.10)$$

può essere risolto usando i moltiplicatori di Lagrange e derivando rispetto ai parametri μ and κ . Dopo alcuni semplici passaggi si trova che il minimo si raggiunge per:

$$\begin{aligned} \frac{\cosh \kappa}{\sinh \kappa} - \frac{1}{\kappa} - \mu' m &= 0 \\ \kappa m - \lambda \mu &= 0 \end{aligned} \quad (6.11)$$

dove m è il vettore media della distribuzione P

$$m = \int_{\mathbb{S}^2} x f(x) d\sigma_x. \quad (6.12)$$

Si ha così il seguente risultato

PROPOSIZIONE 2.8. *La distribuzione di Langevin che meglio approssima una arbitraria distribuzione P sulla sfera unitaria, nel senso della (pseudo-)distanza di Kullback- Leibler, è quella che ha la stessa media m di P .*

Ne segue che il problema di approssimazione si risolve semplicemente prendendo la distribuzione di Langevin che ha la stessa media di P . In sostanza, l'unico parametro della distribuzione P che conta per l'approssimazione è il vettore media. Questo risultato è l'analogo sulla sfera di quanto succede con le distribuzioni Gaussiane in \mathbb{R}^n . La Gaussiana che meglio approssima una distribuzione arbitraria su \mathbb{R}^n , è quella che ha la stessa media e la stessa varianza.

Stima MAP di una Direzione. Supponiamo di misurare una direzione incognita $\mathbf{x} \in \mathbb{S}^2$ che è descritta da una densità a priori di tipo Langevin,

$$\mathbf{x} \sim L(x_0, \kappa_0)$$

mediante un sensore affetto da rumore angolare, descrivibile mediante una densità condizionata $p(y|x)$ che è ancora una distribuzione di Langevin. Per ragioni fisiche è ragionevole supporre che la direzione misurata \mathbf{x} e la rotazione aleatoria $d\mathbf{p} \wedge$ che corrompe le osservazioni di \mathbf{x} , siano quantità aleatorie indipendenti. Possiamo così formare la densità a posteriori $p(x|y)$ usando la regola di Bayes. La densità congiunta è

$$p(x, y) = p(y|x)p(x) = A(\kappa, \kappa_0) \exp \hat{\kappa} \hat{\mu}' x$$

dove

$$A(\kappa, \kappa_0) = \frac{\kappa}{4\pi \sinh \kappa} \frac{\kappa_0}{4\pi \sinh \kappa_0}$$

$$\hat{\kappa} \hat{\mu}' x := \kappa y' x + \kappa_0 x_0' x.$$

In questa formula, $\hat{\kappa} = \hat{\kappa}(y, x_0) > 0$ e $\hat{\mu} = \hat{\mu}(y, x_0)$ sono funzioni di y e della moda a priori x_0 , date esplicitamente dalle relazioni

$$\hat{\mu} := \frac{\kappa y + \kappa_0 x_0}{\hat{\kappa}} \quad \hat{\kappa} := \|\kappa y + \kappa_0 x_0\|. \quad (6.13)$$

Notiamo che $\|\hat{\mu}\| = 1$.

Dividendo per la marginale si ottiene così la densità a posteriori

$$p(x|\mathbf{y}) = \frac{\hat{\kappa}}{4\pi \sinh \hat{\kappa}} \exp \hat{\kappa}(\mathbf{y}) \hat{\mu}'(\mathbf{y}) x$$

che è ancora di tipo Langevin. La moda condizionata $\hat{\mu}(\mathbf{y})$, che è la *stima Bayesiana Massima a Posteriori* di \mathbf{x} , data l'osservazione \mathbf{y} e la concentrazione condizionata $\hat{\kappa}(\mathbf{y})$ si calcolano facilmente in questo semplice esempio. Esse sono date proprio dalla formula (6.13).

Queste formule possono essere generalizzate al caso di un numero arbitrario di osservazioni dello stesso target affette da rumori angolari indipendenti.

Supponiamo di avere una successione di osservazioni della direzione \mathbf{x} affette da rumore angolare

$$\mathbf{y}(t) := R(\mathbf{p}(t)) \mathbf{x} = \exp\{\mathbf{p}(t) \wedge\} \mathbf{x} \quad t = 1, 2, \dots \quad (6.14)$$

dove le \mathbf{p} sono variabili aleatorie indipendenti e identicamente distribuite, indipendenti da \mathbf{x} . Le $\mathbf{y}(t)$ sono quindi condizionatamente indipendenti data \mathbf{x} . Assumeremo inoltre che le rotazioni aleatorie siano “piccole”, di modo che $p(y(t) | \mathbf{x}) = L(\mathbf{x}, \kappa)$, dove κ è un parametro di concentrazione del rumore angolare che per semplicità supporremo essere lo stesso in tutte le misure. Introducendo la notazione $\mathbf{y}^t := [\mathbf{y}(1), \dots, \mathbf{y}(t)]'$, possiamo scrivere

$$p(\mathbf{y}^t | \mathbf{x}) = \frac{\kappa^t}{(4\pi \sinh \kappa)^t} \exp \kappa \langle \mathbf{x}, \sum_{s=1}^t \mathbf{y}(s) \rangle \quad (6.15)$$

dove $\langle \cdot, \cdot \rangle$ denota il prodotto scalare in \mathbb{R}^3 . Assumendo ancora una distribuzione a priori di $\mathbf{x} \sim L(x_0, \kappa_0)$, si calcola, con lo stesso procedimento visto sopra, la densità posteriori

$$p(x | \mathbf{y}^t) = \frac{\hat{\kappa}(t)}{(4\pi \sinh \hat{\kappa}(t))} \exp \hat{\kappa}(t) \langle \hat{\mu}(t), x \rangle \quad (6.16)$$

che è ancora di tipo Langevin con parametri

$$\hat{\mu}(t) = \frac{1}{\hat{\kappa}(t)} \left(\kappa \sum_{s=1}^t \mathbf{y}(s) + \kappa_0 x_0 \right) \quad (6.17)$$

$$\hat{\kappa}(t) = \left\| \kappa \sum_{s=1}^t \mathbf{y}(s) + \kappa_0 x_0 \right\| \quad (6.18)$$

Da notare che nel caso di distribuzione a priori uniforme ($\kappa_0 = 0$), la prima di queste formule diventa

$$\hat{\mu}(t) = \frac{\sum_{s=1}^t \mathbf{y}(s)}{\left\| \sum_{s=1}^t \mathbf{y}(s) \right\|} \quad (6.19)$$

È istruttivo confrontare questa espressione con lo stimatore a minima varianza nel caso Gaussiano, che è semplicemente la media aritmetica delle osservazioni, che, ovviamente, non preserva la norma unitaria degli addendi come invece fa la (6.19).

Lo stimatore (6.17) può essere scritto in forma ricorsiva, mediante uno schema di calcolo che aggiorna la stima calcolata al tempo t , in corrispondenza all'acquisizione della $t + 1$ -sima misura.

PROPOSIZIONE 2.9. *Lo stimatore MAP, $\hat{\mu}(t)$, della direzione aleatoria \mathbf{x} osservata in presenza di rumore angolare a distribuzione di tipo Langevin di concentrazione κ , si aggiorna nel tempo secondo le ricorsioni*

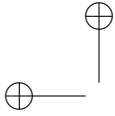
$$\hat{\mu}(t+1) = \frac{1}{\hat{\kappa}(t+1)} (\hat{\kappa}(t) \hat{\mu}(t) + \kappa y(t+1)) \quad (6.20)$$

$$\hat{\kappa}(t+1) = \left\| \hat{\kappa}(t) \hat{\mu}(t) + \kappa y(t+1) \right\| \quad (6.21)$$

con condizioni iniziali $\hat{\mu}(0) = x_0$ e $\hat{\kappa}(0) = \kappa_0$.

Queste formule assomigliano alle ricorsioni del Filtro di Kalman ⁵ e possono essere generalizzate al caso di inseguimento di direzioni variabili di moto aleatorio. L'argomento della stima dinamica di direzioni è trattato in [115, 34].

⁵In effetti ne rappresentano una semplice versione non lineare.

**Esercizi**

2-1 Calcolare l'interpolatore lineare a m.v.

$$\hat{\mathbf{x}}(t) = \hat{E}[\mathbf{x}(t) \mid \mathbf{x}(t-1), \mathbf{x}(t+1)]$$

per un processo stazionario a media nulla di covarianza $\sigma(\tau)$. Come dev'essere $\sigma(\tau)$ per avere un errore di stima $\tilde{\mathbf{x}}(t)$ uguale a zero?

2-2 Sia $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$ un vettore aleatorio a media nulla con $\text{var}(\mathbf{y}_1)$ e $\text{var}(\mathbf{y}_2)$ positive ma la cui matrice varianza è singolare. Dimostrare che $\mathbf{y}_1 = \alpha \mathbf{y}_2$ con $\alpha \neq 0$.

2-3 Sia $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$ un vettore aleatorio a media nulla la cui matrice varianza è non singolare. Qual è la condizione sulle due componenti $\mathbf{y}_1, \mathbf{y}_2$ necessaria e sufficiente per la validità della

$$\hat{E}[\mathbf{x} \mid \mathbf{y}] = \hat{E}[\mathbf{x} \mid \mathbf{y}_1] + \hat{E}[\mathbf{x} \mid \mathbf{y}_2]$$

per una arbitraria variabile scalare a media nulla \mathbf{x} ?

2-4 Dimostrare che la matrice varianza (3.10) nell'esempio 2.3 è invertibile se $0 < \lambda_0 < 1$ ma è singolare per $n \geq 2$ se $\lambda_0 = 1$, indipendentemente dalla scelta dei tempi $t_1 \leq \dots \leq t_n$.

2-5 Nel modello lineare a due blocchi

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}$$

in cui le osservazioni \mathbf{y}_i sono m_i -dimensionali, si assuma $\text{Var}(\mathbf{x}) := P > 0$, $\text{Var}(\mathbf{w}_i) := R_i > 0$, $i = 1, 2$, e $\mathbf{w}_1 \perp \mathbf{x} \perp \mathbf{w}_2$ (tutte le variabili sono a media nulla). Utilizzando il lemma di inversione di matrice, dimostrare che lo stimatore $\hat{\mathbf{x}} = \hat{E}[\mathbf{x} \mid \mathbf{y}_1, \mathbf{y}_2]$ si può scrivere come una funzione lineare degli stimatori "decentralizzati" $\hat{\mathbf{x}}_i = \hat{E}[\mathbf{x} \mid \mathbf{y}_i]$, $i = 1, 2$, della forma

$$\hat{\mathbf{x}} = Q^{-1}[A \hat{\mathbf{x}}_1 + B \hat{\mathbf{x}}_2] \quad .$$

Trovare l'espressione di Q , A , B .

2-6 (Stima distribuita). Si supponga che $\mathbf{y}_1, \dots, \mathbf{y}_N$ siano sottovettori che formano una partizione del vettore di osservazione \mathbf{y} , descritto da un modello lineare $\mathbf{y} = S\mathbf{x} + \mathbf{w}$ in cui valgano le ipotesi standard (medie nulle, $R > 0$ ecc.) Siano $\hat{\mathbf{x}}_i = \hat{E}[\mathbf{x} \mid \mathbf{y}_i]$ gli stimatori "locali" di \mathbf{x} basati sulla conoscenza dei parametri del modello globale ma solo sull'osservazione parziale \mathbf{y}_i , $i = 1, \dots, N$.

Si vogliono trovare condizioni sulla struttura del modello lineare sotto le quali vale l'uguaglianza ("fusione" delle stime)

$$\hat{E}[\mathbf{x} \mid \mathbf{y}_1, \dots, \mathbf{y}_N] = \hat{E}[\mathbf{x} \mid \hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_N]$$

in altri termini, si vuole che lo stimatore $\hat{\mathbf{x}} = \hat{E}[\mathbf{x} \mid \mathbf{y}]$ sia esprimibile come una funzione (lineare) degli stimatori locali $\{\hat{\mathbf{x}}_i\}$.

- Mostrare che $\mathbf{w}_i \perp \mathbf{x} \perp \mathbf{w}_j$, $\forall i \neq j$ (\mathbf{w}_i è il sottovettore di \mathbf{w} corrispondente a \mathbf{y}_i) è una condizione sufficiente per la fusione e determinare la relazione lineare tra $\hat{\mathbf{x}}$ e $\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_N$.
- La condizione precedente è anche necessaria?

2-7 Siano date tre variabili aleatorie $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$ a media nulla e matrice di covarianza Σ non singolare. Si considerino le tre proiezioni ortogonali $\hat{E}[\mathbf{y}_1 | \mathbf{y}_2 \mathbf{y}_3]$, $\hat{E}[\mathbf{y}_2 | \mathbf{y}_1 \mathbf{y}_3]$, $\hat{E}[\mathbf{y}_3 | \mathbf{y}_1 \mathbf{y}_2]$, ciascuna delle quali dipende da due parametri. Studiare sotto quali condizioni queste proiezioni determinano univocamente Σ . Notare che Σ , essendo simmetrica, dipende da sei parametri.

È possibile assegnarle *arbitrariamente* tutte e tre come funzioni lineari dei rispettivi argomenti e determinare così Σ ?

2-8 Sia $\mathbf{y} := \{\mathbf{y}(\tau) | \tau \in \mathbb{R}\}$ un processo scalare a tempo continuo di media nulla e funzione covarianza $\sigma(t, s)$ finita in ogni punto del piano $\{t, s\}$. Dimostrare che

- (1) \mathbf{y} è continuo in media quadratica ($\lim_{t \rightarrow s} \|\mathbf{y}(t) - \mathbf{y}(s)\| = 0$ per ogni $t, s \in \mathbb{R}$), se e solo se σ è continua su tutti i punti della diagonale $t = s$ del piano $\{t, s\}$.
- (2) \mathbf{y} è derivabile in media quadratica se e solo se esiste la derivata seconda mista $\partial^2 \sigma / \partial t \partial s$ in tutti i punti della diagonale $t = s$ del piano $\{t, s\}$.
- (3) L'integrale in media quadratica $\mathbf{x} := \int_0^T \mathbf{y}(s) ds$ ha varianza finita se e solo se

$$\int_0^T \int_0^T \sigma(t, s) dt ds < \infty$$

Si veda ad esempio il testo di Wong, [156, p. 77-80].

2-9 Sia ancora $\mathbf{y} := \{\mathbf{y}(\tau) | \tau \in \mathbb{R}\}$ un processo scalare a tempo continuo di media nulla e funzione covarianza $\sigma(t, s)$ finita in ogni punto del piano $\{t, s\}$. Sia $h > 0$ un passo di campionamento fissato. Discutere i seguenti problemi

- (1) (Discretizzazione della derivata). Trovare l'approssimazione a minima varianza d'errore della derivata in media quadratica $\frac{d\mathbf{y}}{dt}(t_0)$, basata sui valori discreti $\{\mathbf{y}(t_0 - nh), \dots, \mathbf{y}(t_0 - h), \mathbf{y}(t_0), \mathbf{y}(t_0 + h), \dots, \mathbf{y}(t_0 + nh)\}$.
- (2) (Formula di quadratura a minima varianza). Trovare l'approssimazione a minima varianza d'errore dell'integrale in media quadratica \mathbf{x} definito nel problema precedente, basata sui valori discreti $\{\mathbf{y}(0), \mathbf{y}(h), \dots, \mathbf{y}(nh)\}$, $nh = T$. In altri termini determinare i coefficienti $c(k)$ della formula di quadratura

$$\hat{\mathbf{x}} = \sum_{k=0}^n c(k) \mathbf{y}(kh)$$

in modo tale che la varianza di $\mathbf{x} - \hat{\mathbf{x}}$ sia minima.



CAPITOLO 3

STIMATORI LINEARI A MINIMA VARIANZA PER PROCESSI STOCASTICI

1. Introduzione

Nel capitolo precedente ci siamo occupati del problema della stima di un certo vettore aleatorio \mathbf{x} di dimensione finita, basata sull'osservazione di un secondo vettore m -dimensionale \mathbf{y} . I dati di misura erano quindi i valori campionari di una m -pla *fissata* di variabili casuali $(\mathbf{y}_1, \dots, \mathbf{y}_m)$ e in tutte le considerazioni fatte m è stato trattato come un dato del problema fissato a priori.

In realtà, molto spesso nell'elaborazione di segnali di interesse in ingegneria, non si ha a che fare con una m -pla di valori osservati data una volta per tutte, ma si ha piuttosto un *flusso di misure*. Normalmente i dispositivi di acquisizione dei dati forniscono valori campionati a cadenza fissa di segnali che si ricevono per periodi molto lunghi di tempo e quindi si hanno continuamente a disposizione nuove misure che si aggiungono a quelle precedenti. La quantità che interessa stimare è poi generalmente essa stessa variabile nel tempo; tipicamente si tratta di un segnale che contiene l'informazione utile e di cui le osservazioni rappresentano una versione distorta (dal mezzo fisico di misura e trasmissione del segnale) e “rumorosa”.

In questi casi occorre riformulare il problema di stima considerando sia i dati che la grandezza da stimare come due *processi stocastici* (in generale a valori vettoriali) $\{\mathbf{y}(t)\}$ e $\{\mathbf{x}(t)\}$. Sorge così il problema della modellizzazione matematica dei processi $\{\mathbf{y}(t)\}$ e $\{\mathbf{x}(t)\}$.

Diciamo subito che per ovvi motivi di semplicità di trattazione matematica, il nostro interesse sarà rivolto principalmente alla classe degli *stimatori lineari*, cosicché nella maggior parte dei casi sarà sufficiente disporre di descrizioni che specificano soltanto i *momenti congiunti del primo e del secondo ordine* di $\{\mathbf{y}(t)\}$ e $\{\mathbf{x}(t)\}$. Questi dati sono sufficienti a individuare lo stimatore se si usano criteri di valutazione della bontà di stima di tipo quadratico, in particolare con quello della minima varianza d'errore.

In un contesto dinamico le questioni di modellistica di segnali aleatori assumono importanza primaria. Per l'applicabilità pratica della teoria è necessario infatti che le descrizioni matematiche dei processi in gioco siano semplici e agevolmente interpretabili ai fini della descrizione e della manipolazione dei segnali che si incontrano in pratica nelle comunicazioni e nei controlli. In queste applicazioni i segnali di solito sono generati e combinati tra di loro per mezzo di “filtri” o più in generale di sistemi dinamici ed è molto più naturale per l'ingegnere descrivere i segnali in gioco in un sistema di comunicazione o controllo in termini di trasformazioni dinamiche e di connessioni “ingresso-uscita”, piuttosto che pensare di avere direttamente disponibili le auto e mutue covarianze. Per questa ragione il primo obiettivo che ci dovremo porre sarà quello di studiare in modo approfondito la *modellistica* dei

processi del secondo ordine da questo punto di vista. Come vedremo nel seguito, la soluzione dei problemi di stima dinamica su processi stocastici è basata in modo essenziale sulle proprietà *sistemistiche* del modello dinamico con cui si rappresentano i segnali. Per esempio, il cosiddetto approccio di *Wiener-Kolmogorov* (che esporremo nei capitoli seguenti) dipende in modo essenziale da un importante teorema di rappresentazione (di H. Wold) il quale stabilisce che una vasta classe di processi stocastici stazionari può essere rappresentata come l'uscita di un sistema dinamico lineare con ingresso un opportuno rumore bianco. Vedremo nel seguito che la soluzione del problema di stima nell'approccio di Wiener-Kolmogorov si riduce sostanzialmente al calcolo di rappresentazioni dinamiche di questo tipo dei processi $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$.

Un altro elemento nuovo che sorge nel contesto dinamico è che i dati si accumulano nel tempo: la *base temporale*, ovvero l'intervallo di tempo I che comprende i dati $\{\mathbf{y}(t)\}$ su cui basare la stima, in generale si espande nel tempo. In queste condizioni l'algoritmo di calcolo delle stime (lo stimatore) deve necessariamente essere *tempo-variante*. Dato che in molte applicazioni lo stimatore deve, come si suol dire, lavorare in *tempo reale*, cioè fornire le stime del segnale messaggio sequenzialmente nel tempo man mano che arrivano i dati di misura, si intravede fin d'ora l'importanza di una adeguata struttura algoritmica della soluzione.

2. Predizione, Filtraggio, Interpolazione

Ci sono tre prototipi fondamentali di problema di stima su processi stocastici, la *predizione*, il *filtraggio* e l'*interpolazione*. Questi tre problemi, pur non esaurendo la totalità dei casi interessanti per le applicazioni, costituiscono delle utili schematizzazioni alle quali è spesso possibile ricondursi.

Nella descrizione che segue supporremo $\{\mathbf{y}(t)\}$ e $\{\mathbf{x}(t)\}$ processi del secondo ordine di dimensioni rispettive m ed n . Useremo il simbolo $\hat{\mathbf{x}}(t | I)$ per indicare lo stimatore (lineare a minima varianza) della variabile casuale $\mathbf{x}(t)$ (o del vettore casuale $\mathbf{x}(t)$) basato sui dati $\{\mathbf{y}(t) ; t \in I\}$.

1) Problema della predizione

Si vuole stimare il processo

$$\mathbf{x}(t) = \mathbf{y}(t + h) \quad , \quad (2.1)$$

dove h è un intero positivo. La base temporale per la stima della variabile $\mathbf{y}(t + h)$ è l'intervallo di tempo

$$I = [t_0, t] \quad . \quad (2.2)$$

In questo problema si cerca quindi, per ogni t , la miglior predizione del processo $\{\mathbf{y}(t)\}$, “ h passi in avanti” basata sull'informazione “corrente”, cioè su tutte misure disponibili all'istante t , $\{\mathbf{y}(s) ; t_0 \leq s \leq t\}$. Il predittore si denota col simbolo

$$\hat{\mathbf{y}}(t + h | t) \quad (2.3)$$

Il problema della predizione fu storicamente il primo problema di stima dinamica formulato e studiato in letteratura. Fu risolto indipendentemente da Kolmogorov [81, 83] e da Wiener [150] attorno al 1940. Il lavoro di Wiener era motivato soprattutto dall'applicazione a sistemi di puntamento automatico per batterie antiaeree e venne raccolto in un famoso rapporto [149] a circolazione limitata datato Febbraio 1942, noto come “the yellow peril” tra gli ingegneri dell'epoca per la sua non facilissima comprensione. Il rapporto fu pubblicato solo dopo la guerra e diede

origine a una vastissima letteratura. Chi spiegò come le idee di Wiener si potessero effettivamente mettere in pratica furono Bode e Shannon [26] nel 1949.

2) Problema del filtraggio

Questo è il tipico problema delle comunicazioni. Le misure (il segnale in ricezione), $\{\mathbf{y}(t)\}$, sono elaborate per ricostruire un segnale “messaggio” $\{\mathbf{x}(t)\}$ che in generale è distorto da un canale o da un sensore (che noi supporremo lineare) di caratteristiche note e corrotto da rumore $\{\mathbf{n}(t)\}$ di caratteristiche statistiche note. Gli schemi che vengono impiegati comunemente per descrivere questa situazione sono, nei casi più semplici del tipo “segnale con sovrapposto rumore *additivo*”, cioè

$$\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{n}(t) \quad (2.4)$$

ovvero, più in generale, da

$$\mathbf{y}(t) = H \circ \mathbf{x}(t) + \mathbf{n}(t) \quad , \quad (2.5)$$

dove di solito (ma non necessariamente) $\{\mathbf{x}(t)\}$ ed $\{\mathbf{n}(t)\}$ sono ipotizzati scorrelati (o indipendenti). Nello schema (2.5) il simbolo $H \circ$ denota una trasformazione nota, ad esempio un operatore di convoluzione, oppure, nei casi più semplici, una trasformazione istantanea che si riduce alla moltiplicazione per una matrice $m \times n$.

La ricostruzione del messaggio, cioè la stima di $\mathbf{x}(t)$, deve essere fatta ad ogni istante, usando tutta l'informazione disponibile *fino a quell'istante* ($I = [t_0, t]$). In altri termini la stima di $\mathbf{x}(t)$, che denotiamo col simbolo

$$\hat{\mathbf{x}}(t | t) \quad , \quad (2.6)$$

deve (al pari della predizione (2.3)) *essere ricalcolata “in tempo reale”* ad ogni istante t .

3) Problema dell'interpolazione

La caratteristica dei problemi di interpolazione è la possibilità di usare anche dati “futuri” per ottenere stime di $\mathbf{x}(t)$. Tipicamente l'interpolazione si usa quando la stima del segnale non serve per prendere delle decisioni in tempo reale o è comunque indifferente se la sua ricostruzione viene effettuata con un certo ritardo. In questi casi interessa piuttosto la bontà della ricostruzione del messaggio $\{\mathbf{x}(t)\}$.

Nell'*interpolazione a intervallo fisso* la base temporale è fissata

$$I = [t_0, t_1] \quad (2.7)$$

e si chiede di ricostruire il segnale $\{\mathbf{x}(t)\}$ (legato alle osservazioni mediante, ad esempio, relazioni del tipo (2.4) o (2.5)) per ogni t in I . L'interpolatore a intervallo fisso

$$\hat{\mathbf{x}}(t | I) \quad (2.8)$$

è generalmente inteso come un algoritmo che lavora “fuori linea”. Si noti comunque che col progredire dell'hardware, l'elaborazione di segnali in “tempo reale” diventa praticamente possibile anche usando algoritmi a struttura non causale. In applicazioni in cui la velocità e il tempo di calcolo non sono una limitazione essenziale, si possono immagazzinare successivamente tratti di segnale di durata finita (*frames*) e ricostruire il segnale messaggio $\{\mathbf{x}(t)\}$ mediante interpolatori a intervallo fisso basati sui dati di frames acquisiti in successione.

In letteratura vengono classificate come problemi di interpolazione anche alcune varianti del problema del filtraggio. In questi casi la base temporale è sempre

relativa alle misure “correnti”,

$$I = [t_0, t] \quad .$$

Nell’*interpolazione a punto fisso* si chiede di ricostruire, per ogni t , la variabile $\mathbf{x}(\tau)$, dove ora τ è un istante temporale *fissato* (tipicamente l’istante iniziale), precedente t . Lo stimatore relativo si denota col simbolo

$$\hat{\mathbf{x}}(\tau | t) \quad , \quad (\tau < t) \quad . \quad (2.9)$$

Nell’*interpolazione a ritardo fisso* si ha a disposizione un certo ritardo, $h > 0$, che permette di utilizzare anche le misure $\mathbf{y}(t-h+1), \dots, \mathbf{y}(t)$ per ricostruire il segnale $\mathbf{x}(t-h)$ all’istante $t-h$.

Lo scopo è ovviamente quello di sfruttare il ritardo per migliorare la qualità della stima. L’interpolatore a ritardo fisso viene indicato mediante il simbolo

$$\hat{\mathbf{x}}(t-h | t) \quad . \quad (2.10)$$

Si noti che anche in questi ultimi due esempi gli interpolatori debbono in generale essere in grado di lavorare in “tempo reale”, più o meno analogamente al filtro che fornisce $\hat{\mathbf{x}}(t | t)$.

Sul concetto di “soluzione” di un problema di stima. La descrizione che abbiamo dato di questi problemi tocca solo superficialmente la questione, oltremodo importante, di ciò che si debba esattamente intendere per *soluzione* di un problema di predizione, filtraggio o interpolazione.

Allo scopo di rendersi conto della problematica connessa all’implementazione pratica delle formule risolutive, analizzeremo sommariamente un problema di stima in tempo reale della variabile corrente di un processo messaggio, $\mathbf{x}(t)$, basata su osservazioni $\{\mathbf{y}(t)\}$ che si suppone vengano acquisite per un tempo indefinito. Supponiamo che l’istante iniziale t_0 sia finito. Uno *stimatore lineare* di $\mathbf{x}(t)$ in base ai dati $\{\mathbf{y}(s) ; t_0 \leq s \leq t\}$ avrà allora la struttura

$$\hat{\mathbf{x}}(t | t) = \sum_{t_0}^t H(t, k) \mathbf{y}(k) \quad , \quad (2.11)$$

dove le $\{H(t, k) ; k = t_0, t_0+1, \dots, t\}$ sono matrici $n \times m$. $H(t, \cdot)$ può essere pensata come la risposta impulsiva di un certo sistema lineare che accetta in ingresso i dati di misura e restituisce in uscita le corrispondenti stime $\hat{\mathbf{x}}(t | t)$ di $\mathbf{x}(t)$. Questo sistema lineare (lo stimatore) deve in linea di principio essere determinato “fuori linea”, cioè prima di iniziare la fase di elaborazione dei dati vera e propria, in base alla specifica delle statistiche congiunte dei segnali in gioco.

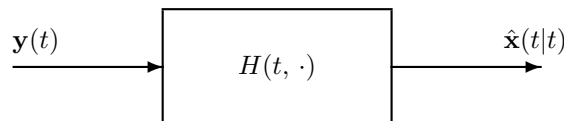


FIGURA 2.1. Stimatore dinamico.

In effetti, il filtro lineare a minima varianza che calcola $\hat{\mathbf{x}}(t | t)$ si determina in base al principio di ortogonalità,

$$\mathbf{x}(t) - \hat{\mathbf{x}}(t | t) \perp \mathbf{y}(s) \quad , \quad t_0 \leq s \leq t \quad ,$$

il quale fornisce le relazioni

$$\Sigma_{\mathbf{x}\mathbf{y}}(t, s) = \sum_{t_0}^t H(t, k) \Sigma_{\mathbf{y}}(k, s) \quad , \quad t_0 \leq s \leq t \quad , \quad (2.12)$$

che costituiscono un sistema di $(t - t_0 + 1)$ equazioni lineari nelle $t - t_0 + 1$ incognite

$$H(t, t_0) , H(t, t_0 + 1) , \dots , H(t, t) \quad .$$

Questo sistema dovrà essere risolto (sotto opportune condizioni di non singolarità delle $\Sigma_{\mathbf{y}}(t, s)$) *per ogni* istante finito t . Sembra fino a questo punto che non vi sia alcun aspetto sostanzialmente nuovo rispetto alla teoria statica della stima lineare sviluppata nel Cap. 2.

In realtà la “soluzione” del problema rappresentata dalla (2.12) è del tutto insoddisfacente e per almeno due importanti ragioni.

- (1) Dato che si ha a che fare con un “flusso” di misure occorre risolvere la (2.12) per tutti i $t > t_0$ e ricavare la corrispondente risposta impulsiva $H(t, \cdot)$ (fuori linea) per tutti i $t > t_0$. Ovviamente, al crescere di t la soluzione del sistema può diventare molto onerosa e, al limite, per t molto grandi, praticamente impossibile.
- (2) In fase di elaborazione dei dati di misura $\{\mathbf{y}(t)\}$, per il calcolo della stima $\hat{\mathbf{x}}(t | t)$ all'istante t , occorre avere *immediatamente* disponibile l'intera risposta impulsiva $H(t, \cdot)$. Ciò significa che il dispositivo di elaborazione deve memorizzare in realtà *tutte* le risposte impulsive $H(t, \cdot)$ necessarie per il calcolo delle stime “future” in $t > t_0$. Questo è ovviamente un altro vincolo che preclude l'utilizzazione pratica della “soluzione” (2.12) nella maggioranza dei casi concreti.

Un terzo motivo che si potrebbe aggiungere ai due precedenti è che lo schema proposto ha delle forti carenze dal punto di vista della “economia di calcolo”: tutti i calcoli fatti relativi a dati acquisiti fino all'istante t e in particolare $H(t, \cdot)$ e $\hat{\mathbf{x}}(t | t)$, devono essere “buttati via” quando arriva la $t + 1$ -esima misura, $\mathbf{y}(t + 1)$. In effetti, per produrre la nuova stima occorre ricalcolare completamente $H(t + 1, \cdot)$ e $\hat{\mathbf{x}}(t + 1 | t + 1)$ e non è affatto chiaro se e come si possano utilizzare allo scopo i risultati delle elaborazioni passate.

Le teorie del filtraggio statistico. La possibilità di ottenere soluzioni praticamente utilizzabili in un problema di stima su processi stocastici è legata all'introduzione di certe ipotesi sui segnali aleatori in gioco. In questo capitolo elencheremo alcune classiche “teorie del filtraggio statistico”, come vengono comunemente denominate, che sono state sviluppate in letteratura e testate con successo in applicazioni di svariata natura nel corso degli ultimi decenni. Queste teorie sono basate su una serie assunzioni sui segnali in gioco che possono sembrare a prima vista notevolmente diverse. Le ipotesi che si fanno (notare che a stretto rigore nessuna ipotesi è mai verificata “esattamente” nei problemi reali) possono sembrare più o meno limitative per certe applicazioni specifiche. Anche se le assunzioni sembrano coprire situazioni rigidamente delimitate, in realtà è spesso possibile usare teorie

diverse per studiare un medesimo problema. Le buone teorie non sono mai troppo sensibili al tipo di ipotesi sui dati a cui fanno riferimento e devono permettere un certo grado di flessibilità sulle ipotesi in cui si opera la formalizzazione senza portare a conseguenze disastrose. In effetti, spesso basta una lieve “perturbazione” nella descrizione matematica dei dati disponibili, corrispondente magari a una formalizzazione meno rigida della realtà, per poter arrivare a un risultato utilizzabile. Tanto per fare un esempio, in pratica si può qualche volta ipotizzare la stazionarietà del segnale pur di lavorare con scale dei tempi opportune.

Ricordiamo in ogni caso che le “soluzioni” che una teoria fornisce sono sempre da riguardarsi come soluzioni di un problema astratto ottenuto schematizzando e semplificando la realtà. Per questo motivo è importante che esse siano il più possibile “robuste” rispetto a piccole perturbazioni sulla struttura dei segnali e diano garanzia che l’errore che si commette usandole in situazioni a stretto rigore non contemplata nelle ipotesi non porti a risultati catastrofici (ad esempio l’instabilità del filtro, ecc.).

Fatte queste premesse, descriviamo ora sommariamente i tre approcci principali al problema della stima (lineare) su processi stocastici.

Il primo e storicamente il più antico, è quello sviluppato da N. Wiener e A.N. Kolmogorov all’incirca intorno agli anni 1940-1945. Questo approccio fa riferimento alle seguenti ipotesi.

- (1) L’intervallo d’osservazione I è *illimitato inferiormente*,

$$I = (-\infty, t] \quad . \quad (2.13)$$

- (2) *I processi $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ sono congiuntamente stazionari (in senso debole).*

È possibile mostrare che in queste ipotesi la risposta impulsiva dello stimatore $H(t, k)$ dipende solo dalla differenza dei due argomenti temporali t e k . Lo stimatore diventa insomma un sistema lineare *invariante nel tempo* che può quindi essere calcolato una volta per tutte, note le statistiche dei segnali. Originariamente, il problema fondamentale della teoria era quello di determinarne esplicitamente la risposta impulsiva $\{H(\tau), \tau \geq 0\}$ oppure la corrispondente matrice di trasferimento $\hat{H}(z)$, in forma chiusa, cioè si cercava di *calcolare analiticamente* $H(\tau)$ o $\hat{H}(z)$, supponendo che le covarianze $\Sigma_{\mathbf{xy}}(\tau)$ e $\Sigma_{\mathbf{y}}(\tau)$, ovvero (equivalentemente) gli spettri corrispondenti, fossero noti analiticamente.

Il calcolo di H si fa risolvendo l’equazione (2.12) (equazione che nelle ipotesi attuali viene chiamata *equazione di Wiener-Hopf*) mediante tecniche di “fattorizzazione spettrale”.

Il secondo approccio trae origine dall’articolo di N. Levinson [92] pubblicato come Appendice B alla monografia di Wiener [150]. C’è in questo articolo una formulazione del problema della predizione che *non richiede intervallo d’osservazione infinito*. Un ruolo essenziale è però giocato dalla *stazionarietà* e in particolare dalla particolare struttura (di “Toeplitz”) che assume la matrice varianza di un processo stazionario.

In questo contesto la risposta impulsiva dello stimatore è comunque non stazionaria. In altre parole, la risposta impulsiva H varia al variare di t .

Levinson ha però mostrato che $H(t, \cdot)$ si può aggiornare mediante uno schema di calcolo “ricorsivo”, che sfrutta in modo molto intelligente i calcoli eseguiti per determinare $H(t-1, \cdot)$ al passo precedente.

La soluzione di Levinson ha quindi una struttura completamente diversa da quella di Wiener. L'idea del calcolo esplicito (dell'espressione analitica) di H viene completamente abbandonata. La soluzione consiste invece in un algoritmo che aggiorna $H(t, \cdot)$ (e la stima) in modo efficiente e veloce, sorprendentemente adatto, per i tempi in cui fu ideato, all'uso del calcolatore.

Il terzo approccio, dovuto a R.E. Kalman [74], è strettamente indirizzato verso un tipo di soluzione ricorsiva che è idealmente adatta all'implementazione su sistemi di elaborazione digitali. Si cercano, nella teoria del "filtro di Kalman", soluzioni che hanno la forma di equazioni di evoluzione dello stato di un sistema dinamico lineare di dimensione finita del tipo:

$$\hat{\mathbf{x}}(t+1) = A \hat{\mathbf{x}}(t) + K \mathbf{y}(t+1) \quad . \quad (2.14)$$

La (2.14) è il prototipo di un algoritmo a memoria fissa (uguale alla dimensione di $\hat{\mathbf{x}}(t)$) che aggiorna ad ogni passo la "vecchia" stima $\hat{\mathbf{x}}(t)$ in conseguenza dell'acquisizione della nuova misura $\mathbf{y}(t+1)$.

Per ottenere strutture del tipo (2.14) occorre fare l'ipotesi che anche i segnali in gioco siano descrivibili (o modellabili) mediante sistemi dinamici stocastici di dimensione finita. Questa ipotesi non ha nulla a che vedere con la stazionarietà, né tantomeno con lo spostamento dell'istante iniziale t_0 a $-\infty$. Essa enfatizza invece l'opportunità di descrivere i processi stocastici in gioco mediante *modelli dinamici* piuttosto che mediante dati probabilistici "astratti" come le auto e mutue covarianze.

Negli ultimi decenni le tre teorie sono state sottoposte a un'analisi molto dettagliata e a vari tipi di confronti e generalizzazioni. Si può dire che si tratta di tre punti di vista più o meno intercambiabili, almeno in un campo abbastanza esteso di problemi. L'enfasi originaria sul calcolo analitico della funzione di trasferimento del filtro dell'approccio di Wiener-Kolmogorov è stata abbandonata, ma ugualmente, come vedremo in seguito, le idee essenziali della teoria sono utili e di pratico impiego in molte situazioni concrete. Ugualmente le idee di Levinson sono state generalizzate e applicate a problemi in cui i segnali sono descritti mediante modelli dinamici a dimensioni finite. Si sono così messe a punto versioni cosiddette "veloci" del filtro di Kalman che sono riportate in letteratura sotto nomi svariatisimi (Chandrasekar ecc.). Lo studio intensivo del problema di fattorizzazione spettrale, essenziale per calcolare il "filtro di Wiener", ha poi permesso di chiarire e analizzare in dettaglio il comportamento asintotico del filtro di Kalman e anche di arrivare a proporre in tempi recenti algoritmi numerici efficienti per il calcolo dei relativi parametri.

3. Euristica del filtro di Wiener

A parte alcune difficoltà tecniche connesse con la posizione $t_0 = -\infty$ (e dovute al fatto che bisogna manipolare funzioni lineari di infinite variabili aleatorie), la teoria di Wiener-Kolmogorov appare come la più naturale estensione della teoria statica della stima a m.v., sviluppata nel Cap. 2.

In effetti se si ammette che lo stimatore lineare di $\mathbf{x}(t)$ in base ai dati

$$\mathbf{y}^t := \left\{ \mathbf{y}(s) ; s \leq t \right\}$$

abbia la forma

$$\hat{\mathbf{x}}(t | t) = \sum_{-\infty}^t H(t, k) \mathbf{y}(k) \quad , \quad (3.1)$$

dove la serie a secondo membro si suppone convergente in media quadratica, il principio di ortogonalità e la stazionarietà (debole) dei processi in gioco forniscono per $H(t, \cdot)$ l'equazione

$$\Sigma_{\mathbf{xy}}(t-s) = \sum_{-\infty}^t H(t, k) \Sigma_{\mathbf{y}}(k-s) \quad , \quad s \leq t \quad , \quad (3.2)$$

che è l'ovvia generalizzazione della (2.12).

Supponiamo ora che l'equazione (3.2), che è in realtà un sistema di infinite equazioni lineari nell'incognita $H(t, \cdot)$, abbia *un'unica soluzione*¹

LEMMA 3.1. *Nelle ipotesi di stazionarietà congiunta e intervallo di osservazione illimitato inferiormente, $H(t, k)$ dipende solo dalla differenza, $t - k$, dei due argomenti temporali.*

DIMOSTRAZIONE. Scrivendo il principio di ortogonalità relativo a $\hat{\mathbf{x}}(t+h | t+h)$, con $h \in \mathbb{Z}$ arbitrario, si trova

$$\Sigma_{\mathbf{xy}}(t+h-s) = \sum_{-\infty}^{t+h} H(t+h, k) \Sigma_{\mathbf{y}}(k-s) \quad , \quad s \leq t+h \quad ;$$

ponendo $\tau = k - h$ e $\sigma = s - h$, questa espressione può poi essere riscritta nella forma

$$\Sigma_{\mathbf{xy}}(t-\sigma) = \sum_{-\infty}^t H(t+h, \tau+h) \Sigma_{\mathbf{y}}(\tau-\sigma) \quad , \quad \sigma \leq t \quad ,$$

che coincide con la (3.2). Per l'unicità della soluzione si deve allora avere

$$H(t, \tau) = H(t+h, \tau+h)$$

per ogni τ e h e questa condizione ovviamente implica l'asserto. \square

Notiamo che l'ipotesi $t_0 = -\infty$ è essenziale per provare l'invarianza temporale del filtro (3.1). In queste condizioni lo stimatore lavora "a regime" e ci si può aspettare che $\hat{\mathbf{x}}(t | t)$ sia esso stesso un processo (debolmente) *stazionario*. Per giustificare questa affermazione bisognerebbe usare una serie di concetti che verranno introdotti in seguito e noi torneremo su questa questione a tempo debito. Notiamo però che la stazionarietà è intuitivamente legata alla "stabilità" dello stimatore, ovvero al comportamento della sua risposta impulsiva $H(\cdot)$. È chiaro che la stabilità del filtro $H(\cdot)$ è una proprietà importante che sarebbe opportuno richiedere in ogni caso, per ovvi motivi di utilizzazione pratica.

Il calcolo della risposta impulsiva del filtro di Wiener si riduce, come abbiamo visto, alla soluzione dell'equazione (3.2). Tenendo conto della invarianza temporale

¹L'esistenza e unicità dello stimatore a minima varianza è garantita dal teorema delle proiezioni. Qui però si assumono come candidati per lo stimatore dei funzionali dei dati passati della forma (3.1) e noi non abbiamo ancora appurato se i funzionali lineari in $\mathbf{H}(\mathbf{y}^t)$ hanno effettivamente questa forma.

di H , questa equazione, dopo alcuni semplici cambiamenti di variabile, può essere posta nella forma standard

$$\Sigma_{\mathbf{xy}}(t) = \sum_0^{\infty} H(s) \Sigma_{\mathbf{y}}(t-s) \quad , \quad t \geq 0 \quad , \quad (3.3)$$

che è l'equazione di Wiener-Hopf studiata dai due matematici di cui porta il nome negli anni precedenti la seconda guerra mondiale. In realtà Wiener e Hopf si interessarono dell'analoga a tempo continuo

$$\Sigma_{\mathbf{xy}}(t) = \int_0^{\infty} H(s) \Sigma_{\mathbf{y}}(t-s) ds \quad , \quad t \geq 0 \quad .$$

In questa equazione e nella (3.3), le funzioni (note e incognite) della variabile t sono definite naturalmente su tutto l'asse dei tempi. In particolare le covarianze (note) $\Sigma_{\mathbf{xy}}$ e $\Sigma_{\mathbf{y}}$ sono quantità intrinsecamente definite su tutto l'asse dei tempi. Noi cerchiamo soluzioni $H(\cdot)$ che siano funzioni *causali* di t , ovvero, per le quali valga la condizione

$$H(t) = 0 \quad , \quad t < 0 \quad . \quad (3.4)$$

La condizione (3.4) è ovviamente legata al fatto che $\hat{\mathbf{x}}(t | t)$ deve dipendere solo dalla storia presente e passata di $\{\mathbf{y}(t)\}$ ma non da quella futura.

La (3.3) sembra molto simile a un'ordinaria equazione di convoluzione e si è tentati di risolverla usando l'analisi armonica (trasformata di Fourier). Ad esempio, usando la trasformata z (bilatera) e senza preoccuparci troppo di questioni di esistenza degli spettri $S_{\mathbf{xy}}$ e $S_{\mathbf{y}}$, corrispondenti alle covarianze $\Sigma_{\mathbf{xy}}$ e $\Sigma_{\mathbf{y}}$, sembrerebbe di poter trasformare la (3.3) nella

$$S_{\mathbf{xy}}(z) = \hat{H}(z) S_{\mathbf{y}}(z) \quad ,$$

la quale darebbe immediatamente

$$\hat{H}(z) = S_{\mathbf{xy}}(z) S_{\mathbf{y}}(z)^{-1} \quad . \quad (3.5)$$

In realtà la condizione “ $t \geq 0$ ” (che abbiamo ignorato nel ricavare la (3.5)) è critica e la rende considerevolmente diversa e più difficile da trattare di un'equazione di convoluzione. Ne dà testimonianza il fatto che la funzione di trasferimento $\hat{H}(z)$ nella (3.5) non è in generale trasformata di una funzione *causale* nel senso (3.4) e quindi non può rappresentare la risposta impulsiva del filtro.

PROBLEMA 3.1. *Si può verificare facilmente (e lo lasciamo come esercizio per il lettore) che la (3.5) fornisce la funzione di trasferimento dell'interpolatore a intervallo fisso*

$$\hat{\mathbf{x}}(t | \mathbb{Z}) = \hat{E} [\mathbf{x}(t) | \mathbf{y}(s) ; -\infty < s < +\infty] \quad . \quad (3.6)$$

Mostrare che lo spettro del processo d'errore $\tilde{\mathbf{x}}(t | \mathbb{Z}) = \mathbf{x}(t) - \hat{\mathbf{x}}(t | \mathbb{Z})$ dell'interpolatore di Wiener (3.6) è dato dalla formula

$$S_{\tilde{\mathbf{x}}}(e^{j\omega}) = S_{\mathbf{x}}(e^{j\omega}) - S_{\mathbf{xy}}(e^{j\omega}) S_{\mathbf{y}}^{-1}(e^{j\omega}) S_{\mathbf{yx}}(e^{j\omega}) \quad , \quad (3.7)$$

dove

$$S_{\mathbf{yx}}(e^{j\omega}) = S'_{\mathbf{xy}}(e^{-j\omega}) = S_{\mathbf{xy}}(e^{j\omega})^* \quad .$$

◇

La condizione $t \geq 0$ nell'equazione di Wiener-Hopf (3.3) è legata proprio alla struttura causale del filtro. Per comprenderne il significato prendiamo in considerazione la situazione in cui $\{\mathbf{y}(t)\}$ è *rumore bianco* (debolmente) stazionario, ovvero

$$\Sigma_{\mathbf{y}}(t-s) = I \delta(t-s) \quad ,$$

dove δ è il simbolo di Kronecker. La (3.3) si risolve allora immediatamente e tenendo anche conto della condizione (3.4) si ottiene

$$H(t) = \begin{cases} \Sigma_{\mathbf{xy}}(t) & \text{per } t \geq 0 \\ 0 & \text{per } t < 0 \end{cases} \quad (3.8)$$

cioè $H(\cdot)$ coincide con il tratto causale della covarianza $\Sigma_{\mathbf{xy}}(\cdot)$. Questo modo di risolvere la (3.3) è in un certo senso assolutamente generale. Per estenderlo ai casi che ci interessano è però necessario un riesame approfondito dell'idea di *innovazione* del processo $\{\mathbf{y}(t)\}$ nel presente contesto.

Formulazione geometrica. Procediamo ora a generalizzare ai processi stocastici la formulazione geometrica del problema della stima a minima varianza d'errore introdotta al Capitolo II, sezione 2. Considereremo il caso (rilevante per l'impostazione di Wiener-Kolmogorov) in cui la base temporale delle misure è infinita. Per il momento non sarà necessario assumere stazionarietà del processo di misura.

Nello spazio di Hilbert \mathbf{H} delle variabili del second'ordine relative all'esperimento in esame, consideriamo l'insieme delle combinazioni lineari *finite*, a coefficienti reali, costruite a partire dalla famiglia di variabili aleatorie $\{\mathbf{y}_k(t) ; k = 1, \dots, m ; t \in \mathbb{Z}\}$ del processo delle osservazioni,

$$\tilde{\mathbf{H}}(\mathbf{y}) := \left\{ \sum a_{ki} \mathbf{y}_k(t_i) ; a_{ki} \in \mathbf{R} , t_i \in \mathbb{Z} , k = 1, \dots, m \right\} \quad . \quad (3.9)$$

Chiaramente $\tilde{\mathbf{H}}(\mathbf{y})$ ha struttura di spazio vettoriale reale.

Ora però si ha a che fare con una famiglia infinita di generatori e non è più vero in generale che i limiti di successioni di Cauchy (in m.q.) di variabili di $\tilde{\mathbf{H}}(\mathbf{y})$ appartengano ancora a $\tilde{\mathbf{H}}(\mathbf{y})$. Questa proprietà di chiusura dello spazio rispetto all'operazione di passaggio al limite, che si chiama "completezza", è però essenziale per l'applicabilità del teorema delle proiezioni (si ricordi l'osservazione fatta al Capitolo 2 a proposito della validità del teorema delle proiezioni in spazi di dimensione infinita). Dato che uno degli obiettivi principali in questo corso è proprio quello di calcolare stimatori lineari per segnali aleatori, ci dovremo preoccupare della questione dell'esistenza e del calcolo di proiezioni ortogonali su sottospazi generati da sequenze di dati

La difficoltà si risolve in modo formale: si "completa" $\tilde{\mathbf{H}}(\mathbf{y})$ aggiungendovi tutti i limiti delle successioni di Cauchy. Si ottiene così un sottospazio vettoriale chiuso di \mathbf{H} , al quale si può estendere la nozione di prodotto scalare di $\tilde{\mathbf{H}}(\mathbf{y})$. In effetti se $u = \lim u_n$ e $v = \lim v_k$ in m.q., si può porre:

$$\langle u, v \rangle := \lim_{n,k} \langle u_n, v_k \rangle \quad .$$

L'operazione di completamento produce così un sottospazio *di Hilbert*, $\mathbf{H}(\mathbf{y})$, di \mathbf{H} che si chiama "spazio dei funzionali lineari" del processo del secondo ordine $\{\mathbf{y}(t)\}$. Per indicare questa operazione di completamento useremo una notazione sintetica del tipo

$$\mathbf{H}(\mathbf{y}) = \overline{\text{span}} \{ \mathbf{y}(t) ; t \in \mathbb{Z} \} \quad . \quad (3.10)$$

Un elemento di $\mathbf{H}(\mathbf{y})$ dipende (linearmente) da tutta la “storia” del processo $\{\mathbf{y}(t)\}$. Notiamo però che la forma esplicita di questa dipendenza è adesso incognita. Che tipo di funzionali lineari stiano in $\mathbf{H}(\mathbf{y})$ è in generale difficile da stabilirsi. Ad esempio è in generale difficile stabilire se, e per quali classi di coefficienti $\{a(t)\}$, somme del tipo $\sum_{-\infty}^{+\infty} a'(t) \mathbf{y}(t)$ (convergenti in m.q.) esauriscono $\mathbf{H}(\mathbf{y})$ a meno che non valgano, come vedremo più avanti, delle condizioni particolari sul processo. Un caso particolarmente importante, che vedremo tra poco, in cui si conosce la forma dei funzionali lineari di $\mathbf{H}(\mathbf{y})$ è quando $\{\mathbf{y}(t)\}$ è un *processo bianco*.

Come abbiamo avuto modo di vedere nella sezione precedente, nei problemi di stima dinamica lo spazio delle osservazioni è raramente quello generato dall'intera “storia” del processo $\{\mathbf{y}(t)\}$. In problemi di stima in linea si ha tipicamente disponibile la storia passata e presente del processo, generata dalle variabili $\{\mathbf{y}(s) ; -\infty \leq s \leq t\}$.

Prendendo nella definizione (3.9) le variabili temporali t_i nell'intervallo $(-\infty, t]$ anziché in tutto \mathbb{Z} si arriva, attraverso la medesima costruzione, a definire una famiglia di *sottospazi* di $\mathbf{H}(\mathbf{y})$

$$\mathbf{H}_t^-(\mathbf{y}) := \overline{\text{span}} \left\{ \mathbf{y}(s) ; s \leq t \right\} , \quad (3.11)$$

che contengono i funzionali lineari della “storia passata e presente” del processo fino all'istante t . Analogamente, per simmetria si possono definire i sottospazi dei funzionali lineari della storia futura e presente,

$$\mathbf{H}_t^+(\mathbf{y}) := \overline{\text{span}} \left\{ \mathbf{y}(s) ; s \geq t \right\} . \quad (3.12)$$

La famiglia di sottospazi $\{\mathbf{H}_t^-(\mathbf{y})\}$ è evidentemente *crescente* con t , ovvero

$$\mathbf{H}_s^-(\mathbf{y}) \subseteq \mathbf{H}_t^-(\mathbf{y}) \quad , \quad s \leq t \quad , \quad (3.13)$$

mentre $\{\mathbf{H}_t^+(\mathbf{y})\}$ è *decrescente*,

$$\mathbf{H}_s^+(\mathbf{y}) \supseteq \mathbf{H}_t^+(\mathbf{y}) \quad , \quad s \leq t \quad , \quad (3.14)$$

e con notazione suggestiva si può scrivere

$$\begin{aligned} \lim_{t \rightarrow +\infty} \mathbf{H}_t^-(\mathbf{y}) &= \mathbf{H}(\mathbf{y}) \\ \lim_{t \rightarrow -\infty} \mathbf{H}_t^+(\mathbf{y}) &= \mathbf{H}(\mathbf{y}) \end{aligned} \quad (3.15)$$

dove i limiti stanno a indicare lo spazio di Hilbert generato dalle rispettive successioni (monotone) di sottospazi.

Dato che per la maggior parte del tempo ci occuperemo di problemi con struttura di osservazioni causale (= passato più presente), useremo la notazione semplificata $\mathbf{H}_t(\mathbf{y})$ (o anche $\mathbf{H}(\mathbf{y}^t)$) per indicare lo spazio (dei funzionali lineari) della storia passata e presente $\mathbf{H}_t^-(\mathbf{y})$.

Proiezioni Ortogonali. Ricordiamo che, dato un sottospazio chiuso \mathbf{Y} di \mathbf{H} , ogni elemento $\mathbf{x} \in \mathbf{H}$ si può decomporre in modo unico nella somma

$$\mathbf{x} = \hat{\mathbf{x}} + \tilde{\mathbf{x}} \quad ,$$

dove $\hat{\mathbf{x}} \in \mathbf{Y}$ e $\tilde{\mathbf{x}} \perp \mathbf{Y}$.

La variabile casuale $\hat{\mathbf{x}}$ è allora la *proiezione ortogonale di \mathbf{x} su \mathbf{Y}* . In accordo con quanto stabilito al Capitolo 2, denoteremo di norma l'operatore di proiezione

ortogonale da \mathbf{H} su \mathbf{Y} col simbolo $\hat{E}[\cdot | \mathbf{Y}]$. Due proprietà ovvie dell'operatore di proiezione ortogonale sono

- $\hat{E}[\cdot | \mathbf{Y}]$ è *idempotente*, cioè

$$\hat{E} \left[\hat{E}[\cdot | \mathbf{Y}] | \mathbf{Y} \right] = \hat{E}[\cdot | \mathbf{Y}] \quad .$$

- Se $\mathbf{Y}_1 \subset \mathbf{Y}_2$,

$$\hat{E} \left[\hat{E}[\cdot | \mathbf{Y}_2] | \mathbf{Y}_1 \right] = \hat{E}[\cdot | \mathbf{Y}_1] \quad .$$

In particolare, se $s < t$

$$\hat{E} \left[\hat{E}[\mathbf{x} | \mathbf{H}_t(\mathbf{y})] | \mathbf{H}_s(\mathbf{y}) \right] = \hat{E}[\mathbf{x} | \mathbf{H}_s(\mathbf{y})]$$

per una qualunque variabile $\mathbf{x} \in \mathbf{H}$.

Nel caso di proiezioni di quantità vettoriali ricorreremo sistematicamente all'abuso di notazioni che abbiamo introdotto (e già usato) nel capitolo precedente. \diamond

Dato un vettore n -dimensionale $\mathbf{x} \in \mathbf{H}$, lo stimatore lineare a m.v. di \mathbf{x} in base alle misure $\{\mathbf{y}(s) ; s \in I\}$ è, per il teorema 2.4, proprio la proiezione ortogonale (componente per componente!) di \mathbf{x} sul sottospazio chiuso di \mathbf{H} generato dai dati, diciamolo

$$\mathbf{Y} := \overline{\text{span}}\{\mathbf{y}(s) ; s \in I\} \quad .$$

Questa proiezione, $\hat{\mathbf{x}} = \hat{E}[\mathbf{x} | \mathbf{Y}]$, si può in linea di principio trovare come soluzione dell'equazione

$$E(\mathbf{x} - \hat{\mathbf{x}}) \mathbf{y}(s)' = 0 \quad , \quad s \in I \quad ,$$

(principio di ortogonalità) la quale, dato che \mathbf{Y} è un sottospazio chiuso, ammette sempre una ed una sola soluzione $\hat{\mathbf{x}}$.

Purtroppo però, per tradurre questa formulazione astratta in un metodo di calcolo concreto dello stimatore, occorre sapere qual è la forma generale di un funzionale lineare dei dati, ovvero come $\hat{\mathbf{x}}$ dipende esplicitamente da $\{\mathbf{y}(s) ; s \in I\}$. Come abbiamo già accennato, questa è una questione abbastanza delicata alla quale non è possibile dare una risposta generale. Anche per questo motivo si ricorre all'ortonormalizzazione ("sbiancamento") preliminare del processo di misura. Questo argomento verrà trattato in dettaglio nella rimanente parte del capitolo.

4. Processi Ortonormalizzabili e Fattorizzazione Spettrale

In questa sezione mostreremo che un'ampia classe di processi stocastici ammette una rappresentazione come uscita di un filtro lineare tempo-invariante con in ingresso rumore bianco. La rappresentazione d'innovazione (di un processo stazionario) è in sostanza un caso particolare di questa rappresentazione. Il calcolo della rappresentazione verrà ricondotto ad un problema di *fattorizzazione spettrale*.

Questa rappresentazione fondamentale vale in condizioni molto generali, in particolare senza alcuna ipotesi di *razionalità degli spettri*. Per evitare questa ipotesi, che è più naturale introdurre in seguito, in vista delle applicazioni che tratteremo, è necessaria una trattazione matematica che purtroppo richiede conoscenze di analisi complessa che possono esulare dal bagaglio matematico di uno studente di ingegneria. In particolare le dimostrazioni di alcuni enunciati possono risultare di non facilissima lettura. Queste dimostrazioni verranno contrassegnate mediante un

asterisco; esse possono essere tralasciate senza una sostanziale perdita di continuità logica per il seguito.

D' ora in avanti in questo capitolo considereremo solo processi (del secondo ordine, a media nulla, debolmente) *stazionari*, definiti su tutto l'asse dei tempi \mathbb{Z} . Useremo diffusamente le notazioni e i risultati dell'Appendice B.

Diamo innanzitutto la definizione di processo stazionario *ortonormalizzabile*.²

DEFINIZIONE 3.1. *Il processo m -dimensionale $\{\mathbf{y}(t)\}$ è ortonormalizzabile se esiste un rumore bianco normalizzato p -dimensionale $\{\boldsymbol{\varepsilon}(t)\}$ tale che*

- (1) *I processi $\{\mathbf{y}(t)\}$ ed $\{\boldsymbol{\varepsilon}(t)\}$ sono congiuntamente stazionari.*
- (2) *Si ha*

$$\mathbf{H}(\mathbf{y}) = \mathbf{H}(\boldsymbol{\varepsilon}) \quad . \quad (4.1)$$

Il processo $\{\mathbf{y}(t)\}$ è causalmente ortonormalizzabile o puramente non deterministico (abbreviato a p.n.d. nel seguito) se esiste un processo bianco normalizzato $\{\boldsymbol{\varepsilon}(t)\}$ che, oltre alle condizioni 1. e 2. precedenti, soddisfa anche alla

$$(3) \quad \mathbf{H}_t(\mathbf{y}) = \mathbf{H}_t(\boldsymbol{\varepsilon}) \quad , \quad t \in \mathbb{Z} \quad , \quad (4.2)$$

detta condizione di equivalenza causale. \diamond

Dalla definizione si vede che $\{\boldsymbol{\varepsilon}(t)\}$ è una “speciale” base ortonormale per $\mathbf{H}(\mathbf{y})$. L'attributo “speciale” si riferisce alla proprietà di *invarianza temporale*, implicita nella condizione di stazionarietà 1), che non è in generale posseduta da una arbitraria base ortonormale di $\mathbf{H}(\mathbf{y})$.

ESEMPIO 3.1 (Un processo stazionario non ortonormalizzabile). Un esempio banale di processo non ortonormalizzabile è il seguente. Siano \mathbf{x} e \mathbf{y} due variabili aleatorie a media nulla con

$$E\mathbf{x}^2 = E\mathbf{y}^2 = \sigma^2 \quad , \quad E\mathbf{x}\mathbf{y} = 0 \quad , \quad (4.3)$$

e ω_0 una pulsazione fissa (deterministica). Il processo stazionario $\{\mathbf{y}(t)\}$ definito dalla

$$\mathbf{y}(t) = \mathbf{x} \cos \omega_0 t + \mathbf{y} \sin \omega_0 t \quad , \quad t \in \mathbb{Z} \quad , \quad (4.4)$$

è ovviamente non ortonormalizzabile. Infatti $\mathbf{H}(\mathbf{y}) = \text{span}\{\mathbf{x}, \mathbf{y}\}$ che ha dimensione due, mentre ogni spazio di Hilbert $\mathbf{H}(\boldsymbol{\varepsilon})$ generato da un rumore bianco $\boldsymbol{\varepsilon}$ ha sempre dimensione infinita.

Il lettore può facilmente generalizzare questo esempio alla somma di un numero qualunque di oscillazioni armoniche elementari tra loro scorrelate del tipo (4.4). \diamond

I processi ortonormalizzabili sono esattamente quelli che si possono generare “filtrando” un processo bianco ortonormale mediante un filtro lineare (non necessariamente *causale*) tempo-invariante ed ℓ^2 -stabile, dove “ ℓ^2 -stabile” significa che la risposta impulsiva del filtro è un segnale a quadrato sommabile (o a energia finita). Questo fondamentale teorema di rappresentazione fu sostanzialmente scoperto da H. Wold [155] nel 1938.

²Alcuni autori, ad es. Rozanov [126], chiamano i processi ortonormalizzabili *processi a rango costante*. La dizione, come vedremo più avanti, si riferisce ad una proprietà della densità spettrale di questi processi.

TEOREMA 3.1. *Ogni processo ortonormalizzabile ammette rappresentazioni del tipo*

$$\mathbf{y}(t) = \sum_{-\infty}^{+\infty} W(t-s) \boldsymbol{\varepsilon}(s) \quad , \quad (4.5)$$

dove $\{\boldsymbol{\varepsilon}(t)\}$ è rumore bianco normalizzato e $W(\cdot)$ è una funzione matriciale del tempo $\mathbb{Z} \rightarrow \mathbb{R}^{m \times p}$ le cui righe $w_k(\cdot)$, $k = 1, \dots, m$, sono a quadrato sommabile, i.e.³

$$\sum_{t=-\infty}^{+\infty} |w_k(t)|^2 < \infty \quad , \quad k = 1, \dots, m \quad . \quad (4.6)$$

Se il processo è p.n.d. esistono rappresentazioni (4.5) in cui $W(\cdot)$ è una funzione causale, ovvero

$$W(t) = 0 \quad , \quad t < 0 \quad . \quad (4.7)$$

DIMOSTRAZIONE. Se scriviamo il processo bianco $\{\boldsymbol{\varepsilon}(t)\}$ componente per componente

$$\{\boldsymbol{\varepsilon}(t)\} = \left\{ \dots, \boldsymbol{\varepsilon}_1(0), \dots, \boldsymbol{\varepsilon}_p(0), \dots, \boldsymbol{\varepsilon}_1(t), \dots, \boldsymbol{\varepsilon}_p(t), \dots \right\},$$

otteniamo un'ordinaria base ortonormale per $\mathbf{H}(\mathbf{y})$. Un'arbitraria variabile scalare $\boldsymbol{\eta} \in \mathbf{H}(\mathbf{y})$ ha quindi un'unica rappresentazione rispetto alla base

$$\boldsymbol{\eta} = \sum_{t=-\infty}^{+\infty} c(-t) \boldsymbol{\varepsilon}(t) = \sum_{t=-\infty}^{+\infty} \sum_{k=1}^p c_k(-t) \boldsymbol{\varepsilon}_k(t) \quad , \quad (4.8)$$

dove $c(t)$ è il vettore riga definito dalla

$$c(t) = [c_1(t), \dots, c_p(t)] \quad , \quad c_k(-t) = E \boldsymbol{\eta} \boldsymbol{\varepsilon}_k(t) \quad , \quad (4.9)$$

e il segno “-” nell'argomento è posto per uniformarsi alla convenzione usata per le funzioni di covarianza, pensando $\boldsymbol{\eta}$ come variabile assegnata all'istante $t = 0$. La serie (4.8) converge in media quadratica e vale l'uguaglianza

$$\sum_{-\infty}^{+\infty} |c(-t)|^2 = \sum_{-\infty}^{+\infty} |c(t)|^2 = \|\boldsymbol{\eta}\|^2 < \infty \quad (4.10)$$

che scende immediatamente dal corollario B.1 dell'Appendice B. La rappresentazione (4.5) si ottiene in particolare prendendo per $\boldsymbol{\eta}$ la variabile aleatoria $\mathbf{y}_k(t)$ e notando che per la stazionarietà congiunta di $\{\boldsymbol{\varepsilon}(t)\}$ e $\{\mathbf{y}(t)\}$ i coefficienti dello sviluppo

$$w_{kj}(t-s) := E \mathbf{y}_k(t) \boldsymbol{\varepsilon}_j(s) \quad (4.11)$$

dipendono solo dalla differenza $t-s$.

Se $\{\boldsymbol{\varepsilon}(t)\}$ soddisfa la condizione di equivalenza causale (4.2) si ha $\boldsymbol{\varepsilon}(t+h) \perp \mathbf{H}_t(\mathbf{y})$ per ogni $h > 0$ e quindi dalla (4.11) scende

$$w_{kj}(-h) = E \mathbf{y}_k(t) \boldsymbol{\varepsilon}_j(t+h) = 0 \quad , \quad k, j = 1, \dots, m.$$

che è la condizione di causalità (4.7). \square

³ $|\cdot|$ denota qui la norma Euclidea in \mathbb{R}^p .

I processi p.n.d. sono quelli che corrispondono alla nozione intuitiva di fenomeni caratterizzati da “imprevedibilità temporale”. Infatti si vede bene, imponendo la condizione di causalità (4.7) alla rappresentazione (4.5), che passando dall’istante t a $t + 1$ si somma alla combinazione lineare di valori passati $\{\varepsilon(s); s \leq t\}$ che determinano $y(t)$, un nuovo contributo “imprevedibile” $W(0)\varepsilon(t + 1)$, che è completamente scorrelato dalla storia passata, in formule,

$$\mathbf{y}(t + 1) = W(0)\varepsilon(t + 1) + \sum_{s=-\infty}^t W(t + 1 - s) \varepsilon(s)$$

L’esempio di processo non ortonormalizzabile dato in precedenza è viceversa quello di un processo le cui traiettorie si potrebbero chiamare “prevedibili” e che non corrisponde affatto alla nozione di aleatorietà come “imprevedibilità temporale”. In effetti i valori futuri di questo processo possono essere predetti *esattamente* (cioè con errore di predizione zero) se se ne conoscono i valori campionari in un numero finito di istanti (nel caso in esame bastano due istanti diversi).

Lo studio e la classificazione dei processi stazionari da questo punto di vista verrà affrontato nel capitolo 5.

Sulla corrispondenza isometrica tra variabili aleatorie di un processo stazionario e segnali a energia finita*. Le formule (4.8) e (4.9) forniscono la rappresentazione di una arbitraria variabile aleatoria $\boldsymbol{\eta} \in \mathbf{H}(\boldsymbol{\varepsilon})$ come “funzionale lineare” del processo bianco generatore $\{\varepsilon(t)\}$. Questa rappresentazione si può pensare come una corrispondenza biunivoca (una volta *fissato* il processo bianco $\{\varepsilon(t)\}$ che in generale *non è unico*) tra variabili aleatorie $\boldsymbol{\eta} \in \mathbf{H}(\boldsymbol{\varepsilon})$ e i relativi coefficienti p -dimensionali $\{c(-t), t \in \mathbb{Z}\}$, che formano un segnale *a energia finita*. Ricordiamo (vedere l’Appendice B) che l’insieme dei segnali p -dimensionali a energia finita diventa uno spazio di Hilbert reale denominato $\ell_p^2(\mathbb{Z})$ (notazione abbreviata a ℓ_p^2 quando non c’è pericolo di confusione) se è dotato del prodotto scalare

$$\langle f, g \rangle_{\ell_p^2(\mathbb{Z})} := \sum_{-\infty}^{+\infty} f(t) g(t)' \quad .$$

Possiamo a questo punto introdurre una notazione un pò astratta ma molto conveniente. Introduciamo l’applicazione (o operatore) lineare $I_{\boldsymbol{\varepsilon}} : \ell_p^2(\mathbb{Z}) \rightarrow \mathbf{H}(\mathbf{y})$ che fa corrispondere alla risposta impulsiva (non necessariamente causale) $c(\cdot) \in \ell_p^2$ la variabile aleatoria $\boldsymbol{\eta}$ definita dalla somma di convoluzione (4.8). Introduciamo anche l’operatore di *traslazione temporale*⁴ sul processo $\{\mathbf{y}(t)\}$, cioè l’applicazione U_t che per $t, s \in \mathbb{Z}$ associa a $\mathbf{y}_k(s)$ la variabile traslata $\mathbf{y}_k(t + s)$ di modo che in particolare si può scrivere formalmente

$$\mathbf{y}_k(t) = U_t \mathbf{y}_k(0), \quad t \in \mathbb{Z} \quad k = 1, \dots, m.$$

Dato che la somma di convoluzione (4.8) è convergente in media quadratica, l’identità di Parseval (4.10) permette di affermare che l’applicazione $I_{\boldsymbol{\varepsilon}} : c(\cdot) \mapsto \boldsymbol{\eta}$ è una *isometria* da $\ell_p^2(\mathbb{Z})$ su $\mathbf{H}(\mathbf{y})$, cioè un’applicazione lineare che preserva la norma.

Notiamo in particolare che l’isometria $I_{\boldsymbol{\varepsilon}}$ associa le variabili del processo $\mathbf{y}_k(0)$ relative all’istante $t = 0$, alle funzioni riga $w_k(\cdot)$ della matrice $W(\cdot)$ in (4.5) e le variabili all’istante t , $\mathbf{y}_k(t)$, alle funzioni *traslate* $w_k(\cdot + t)$, per $k = 1, \dots, m$.

⁴L’operatore di traslazione temporale è stato introdotto nell’Appendice B. Esso verrà studiato più diffusamente nel Capitolo 5.

Quest'ultima osservazione conduce al seguente ovvio (ma utile) risultato.

COROLLARIO 3.1. *La traslazione temporale U_t sul processo $\{\mathbf{y}(t)\}$, cioè l'applicazione che per $t \in \mathbb{Z}$ trasforma $\mathbf{y}_k(0)$ nella variabile traslata $\mathbf{y}_k(t)$ per $k = 1, \dots, m$, corrisponde alla traslazione temporale a sinistra di t unità, σ_t , delle risposte impulsive $w_k(\cdot)$. In formule, se $\mathbf{y}_k(0) = I_\varepsilon[w_k(\cdot)]$, allora $\mathbf{y}_k(t) = I_\varepsilon[w_k(\cdot + t)]$, per $t \in \mathbb{Z}$, $k = 1, \dots, m$ e quindi,*

$$I_\varepsilon \sigma_t = U_t I_\varepsilon$$

di modo che U_t è unitariamente equivalente all'operatore di traslazione σ_t in ℓ_p^2 .

Fattorizzazione Spettrale. Sia \hat{W} la trasformata di Fourier di W . Come spiegato nell'Appendice B, considereremo formalmente la z -trasformata di W come la funzione analitica $\hat{W}(z)$ definita sulla corona circolare contenente la circonferenza unitaria $\{|z| = 1\}$ in cui la serie di Laurent

$$\hat{W}(z) = \sum_{-\infty}^{+\infty} W(t) z^{-t} \quad (4.12)$$

è convergente. Se la serie (4.12) non converge puntualmente in un intorno della circonferenza unità, nondimeno la trasformata di Fourier di W esiste (come limite in L^2) e penseremo allora $\hat{W}(z) := \hat{W}(e^{j\omega})|_{e^{j\omega}=z}$ definita (quasi ovunque) solo su $\{|z| = 1\}$. Da notare che con questa convenzione la coniugata trasposta $\hat{W}(e^{j\omega})^*$ di $\hat{W}(e^{j\omega})$ si trasforma, attraverso la posizione $z = e^{j\omega}$, in $\hat{W}(1/z)^\top$. Per questo motivo scriveremo spesso (con un certo abuso di notazioni) $\hat{W}(z)^*$ per intendere $\hat{W}(1/z)^\top$, anche quando z è una variabile complessa che non sta sulla circonferenza unità.

Nelle manipolazioni di formule è molto comodo esprimere le somme di convoluzione come la (4.5) mediante prodotti di z -trasformate. Nel nostro caso però i segnali di “ingresso” e “uscita” nella (4.5) sono processi stocastici e per giustificare rigorosamente il formalismo sarebbe necessario definire la *trasformata di Fourier di un processo stazionario*, in particolare di un rumore bianco $\{\varepsilon(t)\}$. Questo non è immediato (si può mostrare che le traiettorie di un rumore bianco sono segnali a energia infinita con probabilità 1) e necessiterebbe in ogni caso di un lungo “detour” matematico. Noi qui ci accontenteremo di introdurre una notazione simbolica del tipo

$$\mathbf{y}(t) = \hat{W}(z) \varepsilon(t) \quad (4.13)$$

e opereremo formalmente tenendo in mente il significato sistemistico della relazione (4.5), in cui $\{\mathbf{y}(t)\}$ si interpreta come l'*uscita di un filtro lineare* non necessariamente causale, di funzione di trasferimento $\hat{W}(z)$, corrispondente all'applicazione del segnale di ingresso $\{\varepsilon(t)\}$.

In seguito applicheremo le note regole di composizione dell'algebra degli schemi a blocchi dei controlli automatici, come se i processi in gioco fossero segnali deterministici. La trattazione rigorosa dal punto di vista formale, che si può ad esempio trovare nel testo di Rozanov [126], giustifica in generale questo modo di procedere. Sull'interpretazione sistemistica della somma di convoluzione ritorneremo comunque alla fine del capitolo.

Il risultato fondamentale dell'intero capitolo è il seguente.

TEOREMA 3.2. *Il processo m -dimensionale $\{\mathbf{y}(t)\}$ è ortonormalizzabile se e solo se ha distribuzione spettrale assolutamente continua e la relativa matrice densità*

$\omega \rightarrow S(e^{j\omega})$ è fattorizzabile, ovvero esistono funzioni matriciali $m \times p$, $\omega \rightarrow W(e^{j\omega})$, tali che per tutti gli $\omega \in [-\pi, \pi]$,

$$S(e^{j\omega}) = W(e^{j\omega}) W(e^{j\omega})^* \quad . \quad (4.14)$$

Inoltre $\{\mathbf{y}(t)\}$ è p.n.d. se e solo se esistono “fattori spettrali” $W(e^{j\omega})$, che sono i valori al contorno sulla circonferenza unitaria $\{|z| = 1\}$ di matrici $W(z)$ analitiche in $\{|z| > 1\}$.

DIMOSTRAZIONE. (Necessità) Se $\{\mathbf{y}(t)\}$ è ortonormalizzabile allora esiste un rumore bianco normalizzato $\{\boldsymbol{\varepsilon}(t)\}$ per cui vale una rappresentazione del tipo (4.5). Per l’isometria fra $\mathbf{H}(\mathbf{y})$ ed $\ell_p^2(\mathbb{Z})$ cui si è accennato più sopra, si può scrivere la covarianza del processo $\{\mathbf{y}(t)\}$ come

$$\begin{aligned} \sigma_{kj}(\tau) &= \left\langle \mathbf{y}_k(t + \tau), \mathbf{y}_j(t) \right\rangle_{\mathbf{H}(\mathbf{y})} = \left\langle w_k(\cdot + \tau), w_j(\cdot) \right\rangle_{\ell_p^2} \\ &= \int_{-\pi}^{\pi} e^{j\omega\tau} \hat{w}_k(e^{j\omega}) \hat{w}_j(e^{j\omega})^* \frac{d\omega}{2\pi} \quad , \quad k, j = 1, \dots, m \quad , \quad (4.15) \end{aligned}$$

dove nell’ultimo passaggio si è usata l’isometria fra lo spazio dei segnali a energia finita $\ell_p^2(\mathbb{Z})$ e quello delle rispettive trasformate di Fourier, $L_p^2(\mathbf{C})$ (ovviamente tenendo conto del fatto che la trasformata di $w_k(\cdot + \tau)$ è $\hat{w}_k(e^{j\omega}) e^{j\omega\tau}$).

Confrontando la (4.15) con la rappresentazione di $\Sigma_{\mathbf{y}}$ nel Teorema di Herglotz D.1, Appendice B, si conclude, data l’unicità della distribuzione spettrale, che effettivamente dF è assolutamente continua e la sua densità è fattorizzabile nella forma (4.14). Ovviamente se $\{\mathbf{y}(t)\}$ è p.n.d. $\hat{W}(e^{j\omega})$ è la trasformata di una risposta impulsiva causale a righe in ℓ_p^2 . Ne segue che le \hat{w}_k appartengono allo spazio di Hardy H_p^2 (si veda l’Appendice B) e quindi sono estendibili a funzioni analitiche in $\{|z| > 1\}$.

La prova della sufficienza risulterà dalla discussione che segue. \square

LEMMA 3.2. *Tutti i rumori bianchi normalizzati $\{\boldsymbol{\varepsilon}(t)\}$ tali per cui $\mathbf{H}(\mathbf{y}) = \mathbf{H}(\boldsymbol{\varepsilon})$ hanno la stessa dimensione p . Questo numero è un invariante del processo che si chiama rango o molteplicità di $\{\mathbf{y}(t)\}$.*

Per ogni fattore spettrale W corrispondente ad una rappresentazione di $\{\mathbf{y}(t)\}$ del tipo (4.13) si ha

$$\text{rango } W(e^{j\omega}) = p \quad (4.16)$$

per quasi tutti gli $\omega \in [-\pi, \pi]$. Il rango della matrice densità $S(e^{j\omega})$ di un processo ortonormalizzabile è quindi costante (quasi ovunque) in $[-\pi, \pi]$ e uguale a p . Il rango di $\{\mathbf{y}(t)\}$ coincide pertanto con quello della sua matrice densità spettrale.

DIMOSTRAZIONE. Siano $\{\boldsymbol{\varepsilon}_1(t)\}$ e $\{\boldsymbol{\varepsilon}_2(t)\}$ due rumori bianchi normalizzati per cui $\mathbf{H}(\boldsymbol{\varepsilon}_1) = \mathbf{H}(\boldsymbol{\varepsilon}_2)$. Mostriamo che $p_1 = \dim \boldsymbol{\varepsilon}_1(t)$ e $p_2 = \dim \boldsymbol{\varepsilon}_2(t)$ sono uguali. Possiamo allo scopo usare le rappresentazioni

$$\boldsymbol{\varepsilon}_1(t) = Q_1(z) \boldsymbol{\varepsilon}_2(t) \quad , \quad \boldsymbol{\varepsilon}_2(t) = Q_2(z) \boldsymbol{\varepsilon}_1(t) \quad , \quad (4.17)$$

garantite dal Teorema 3.1 applicato ai processi $\{\boldsymbol{\varepsilon}_1(t)\}$ ed $\{\boldsymbol{\varepsilon}_2(t)\}$. La parte già dimostrata del Teorema 3.2 implica che Q_1 e Q_2 debbono essere fattori spettrali

delle relative densità, ovvero

$$\begin{aligned} Q_1(e^{j\omega}) Q_1(e^{j\omega})^* &= I_{p_1} \\ Q_2(e^{j\omega}) Q_2(e^{j\omega})^* &= I_{p_2} \end{aligned} \quad (4.18)$$

quasi ovunque per $\omega \in [-\pi, \pi]$. Queste relazioni implicano che $Q_1(e^{j\omega})$ e $Q_2(e^{j\omega})$ hanno le righe (quasi ovunque) linearmente indipendenti e pertanto

$$\text{rango } Q_1(e^{j\omega}) = p_1 \leq p_2, \quad \text{rango } Q_2(e^{j\omega}) = p_2 \leq p_1$$

da cui segue facilmente che dev'essere $p_1 = p_2 = p$.

Per dimostrare la (4.16) occorre mostrare che non esiste un intervallo aperto in $[-\pi, \pi]$ sul quale il rango di $W(e^{j\omega})$ (che ha dimensione $m \times p$) è strettamente minore di p . Supponiamo allora per assurdo che l'equazione $W(e^{j\omega}) x_\omega = 0$ abbia soluzioni x_ω non banali per tutti i punti ω di un intervallo aperto $\Delta \subset [-\pi, \pi]$. Si può allora dimostrare (ma noi qui lo daremo per scontato) che è possibile scegliere i vettori p -dimensionali x_ω nello spazio nullo di $W(e^{j\omega})$ in modo tale da formare una funzione f (che scriveremo come un vettore *riga*) definita da $f(e^{j\omega}) := x_\omega^*$ per $\omega \in \Delta$ e zero per $\omega \notin \Delta$, appartenente allo spazio $L_p^2(\mathbf{C})$ e di norma $\|f\|_{L_p^2}$ positiva. Formiamo allora la variabile aleatoria $\boldsymbol{\eta} := f(z) \boldsymbol{\varepsilon}(t)$, che è non nulla e appartiene per costruzione ad $\mathbf{H}(\boldsymbol{\varepsilon}) = \mathbf{H}(\mathbf{y})$. Mostriamo che $\boldsymbol{\eta} \perp \mathbf{H}(\mathbf{y})$. In effetti si trova

$$E \mathbf{y}(t) \boldsymbol{\eta} = \int_{-\pi}^{\pi} e^{j\omega t} W(e^{j\omega}) f(e^{j\omega})^* \frac{d\omega}{2\pi} = 0$$

per ogni $t \in \mathbb{Z}$, il che è palesemente assurdo. Ne segue che $W(e^{j\omega}) x_\omega = 0$ non può avere soluzioni non banali su un intervallo aperto e la condizione (4.16) è così dimostrata. L'ultima affermazione scende allora dalla disuguaglianza di Sylvester [52, pag. 66]. \square

ESEMPIO 3.2 (Altri processi stazionari non ortonormalizzabili). Dal risultato appena dimostrato scende che un'altra classe non banale di processi non ortonormalizzabili è costituita dai cosiddetti processi (scalari) a *banda limitata*, quelli cioè che hanno spettro nullo per tutte le frequenze al di fuori di una banda di misura finita, ad esempio $S(e^{j\omega}) = 0$; $|\omega| > \omega_0$ dove $0 < \omega_0 < \pi$. In questo caso infatti il rango di $S(e^{j\omega})$ varia in un intervallo di misura positiva. \diamond

Il Teorema di fattorizzazione spettrale 3.2 è di importanza fondamentale. Esso dà infatti origine ad un procedimento mediante il quale in linea di principio si possono costruire tutte le rappresentazioni del tipo (4.5) di un processo (ortonormalizzabile) $\{\mathbf{y}(t)\}$. Il procedimento può essere suddiviso in tre fasi:

1 - Fattorizzazione spettrale. Trovare i fattori spettrali $W(e^{j\omega})$, di dimensione $m \times p$, della matrice densità spettrale di potenza $S(e^{j\omega})$, dove p è uguale al rango quasi ovunque di $S(e^{j\omega})$ su $[-\pi, \pi]$. Questi fattori spettrali hanno (quasi ovunque) rango p uguale al numero di colonne e possiedono quindi un'inversa sinistra $W(e^{j\omega})^{-L}$.

2 - Sbiancamento. Costruire il processo p -dimensionale

$$\boldsymbol{\varepsilon}(t) := W(e^{j\omega})^{-L} \mathbf{y}(t) \quad . \quad (4.19)$$

Occorre dimostrare che la (4.19) definisce effettivamente $\{\varepsilon(t)\}$ come processo stazionario del secondo ordine. Notiamo allo scopo che la funzione matriciale S_ε definita dalla

$$\begin{aligned} S_\varepsilon(e^{j\omega}) &:= W(e^{j\omega})^{-L} S(e^{j\omega}) [W(e^{j\omega})^{-L}]^* \\ &= W(e^{j\omega})^{-L} W(e^{j\omega}) W(e^{j\omega})^* [W(e^{j\omega})^*]^{-R} = I_p \end{aligned} \quad (4.20)$$

è ovviamente integrabile sulla circonferenza unitaria. Ne segue che le righe di W^{-L} appartengono allo spazio $L_p^2(\mathbf{C}, S \frac{d\omega}{2\pi})$ e quindi per il Corollario D.2, Appendice B, la somma di convoluzione (4.19) è convergente in media quadratica e definisce un processo stazionario $\{\varepsilon(t)\}$ il cui spettro (calcolato mediante le note regole) è proprio S_ε in (4.20). Pertanto $\{\varepsilon(t)\}$ è *rumore bianco normalizzato p-dimensionale*.

Occorre a questo punto premettere un lemma che evidenzia il senso in cui c'è "dipendenza lineare" tra le m componenti di un processo di rango p .

LEMMA 3.3. *Se $\{\mathbf{y}(t)\}$ è ortonormalizzabile di rango p , allora per ogni fattore spettrale W di dimensione $m \times p$ si ha $(I - WW^{-L})\mathbf{y}(t) = 0$, per ogni t , con probabilità uno.*

DIMOSTRAZIONE. Basta calcolare lo spettro di $(I - WW^{-L})\mathbf{y}(t)$.

Si trova

$$\left(I - W(e^{j\omega}) W(e^{j\omega})^{-L} \right) W(e^{j\omega}) W(e^{j\omega})^* \left(I - W(e^{j\omega}) W(e^{j\omega})^{-L} \right)^*$$

che è zero, dato che $W^{-L}W = I$. □

3 - Filtro formante. Il terzo passo del procedimento consiste nel moltiplicare ambo i membri di (4.19) a sinistra per $W(e^{j\omega})$.⁵ Per il lemma appena enunciato $\mathbf{y}(t) = W(e^{j\omega}) W(e^{j\omega})^{-L} \mathbf{y}(t)$, per cui si trova

$$\mathbf{y}(t) = W(e^{j\omega}) \varepsilon(t) \quad ,$$

che è la rappresentazione di tipo (4.5) (o, equivalentemente, (4.13)) del processo $\{\mathbf{y}(t)\}$ come uscita di un filtro ℓ^2 -stabile con ingresso rumore bianco normalizzato. \diamond

Il procedimento in tre passi appena illustrato può essere visto come una giustificazione della prima affermazione nell'enunciato del Teorema 3.2 (sufficienza). La parte più interessante del teorema è però quella riguardante la caratterizzazione dei processi p.n.d.

Processi p.n.d. e fattori spettrali analitici. Supponiamo che esistano fattori spettrali $W(z)$ $m \times p$, analitici in $\{|z| > 1\}$. Ad ognuno di questi fattori la costruzione precedente permette di associare ancora un rumore bianco normalizzato p -dimensionale $\{\varepsilon(t)\}$ e una corrispondente rappresentazione ingresso-uscita di $\{\mathbf{y}(t)\}$

$$\mathbf{y}(t) = W(z) \varepsilon(t) \quad ,$$

in cui ora si ha però una relazione *causale* tra $\{\varepsilon(t)\}$ e $\{\mathbf{y}(t)\}$, dato che, essendo $W(z)$, analitica in $\{|z| > 1\}$, la sua trasformata inversa è una risposta impulsiva *causale*. Questo implica $\mathbf{y}(t) \in \mathbf{H}_t(\varepsilon)$ per ogni $t \in \mathbb{Z}$ e quindi

$$\mathbf{H}_t(\mathbf{y}) \subset \mathbf{H}_t(\varepsilon) \quad .$$

⁵Cioè applicare l'operatore di convoluzione con nucleo \tilde{W} uguale alla trasformata inversa di Fourier di W al processo $\{\varepsilon(t)\}$ definito in (4.19).

Dovremo ora mostrare che se esistono fattori spettrali analitici ne esiste sempre almeno uno per cui anche l'operazione di "sbiancamento" descritta dalla formula (4.19) è un'operazione *causale*, ovvero produce un processo bianco per cui

$$\varepsilon(t) \in \mathbf{H}_t(\mathbf{y}) \quad , \quad t \in \mathbb{Z} \quad .$$

Questa è una faccenda molto più delicata che verrà qui discussa in generale e risolta a livello astratto, basandosi su alcuni risultati della teoria degli spazi di Hardy. Questi sono spazi di Hilbert di funzioni analitiche che hanno molte proprietà utili e permettono di risolvere in modo elegante e compatto i problemi di filtraggio e stima su processi stazionari. In appendice B si può trovare un sunto delle definizioni fondamentali e delle proprietà più importanti di questi spazi.

Innanzitutto cerchiamo di capire *quanti* fattori spettrali analitici ci possono essere. Bastano dei semplici esempi per convincersi che la condizione di analiticità in $\{|z| > 1\}$ del fattore spettrale non è certamente sufficiente a garantirne l'unicità. Per esempio, per la densità spettrale scalare razionale

$$S(z) = \frac{5/4 + 1/2 (z + z^{-1})}{10/9 + 1/3 (z + z^{-1})} = \frac{(z + 1/2) (z^{-1} + 1/2)}{(z + 1/3) (z^{-1} + 1/3)}$$

(in cui si è posto $e^{j\omega} = z$), si possono prendere come fattori analitici in $\{|z| > 1\}$ le funzioni

$$W_1(z) = \frac{z + 1/2}{z + 1/3} \quad , \quad W_2(z) = \frac{z + 2}{z + 1/3}$$

ma, più in generale, è evidente che anche tutte le funzioni del tipo

$$W(z) = W_1(z) U(z) \quad ,$$

in cui $U(z)$ è analitica in $\{|z| > 1\}$ e soddisfa alla relazione $U(z)U(1/z) = 1$, sono fattori spettrali analitici.

Come richiamato nell'appendice B, le funzioni matriciali $U(z)$ di dimensione $p \times p$ che sono analitiche e uniformemente limitate in $\{|z| > 1\}$ e soddisfano sul cerchio unitario la condizione

$$U(e^{j\omega}) U(e^{j\omega})^* = I_p \quad (4.21)$$

(cioè sono matrici *unitarie* sulla circonferenza unità) vengono in letteratura chiamate funzioni *Inner* oppure *passa-tutto stabili*⁶.

I più semplici esempi di funzioni inner scalari sono del tipo $U(z) = z^k$, dove k è un intero negativo. Più in generale, le funzioni inner razionali sono i cosiddetti *prodotti di Blasckhe* finiti,

$$U(z) = \prod_k \left[\frac{1 + \alpha_k z}{z + \bar{\alpha}_k} \right]^{n_k} \quad , \quad |\alpha_k| < 1 \quad . \quad (4.22)$$

che hanno zeri in posizione reciproca rispetto ai poli $\{-\alpha_k, -\bar{\alpha}_k\}$.

Dato che manifestamente ci sono infinite funzioni inner, ci sono anche infiniti fattori spettrali analitici. Come vedremo però, soltanto *uno* degli infiniti fattori spettrali analitici soddisfa alla condizione di equivalenza causale (4.2).

⁶Nella letteratura ingegneristica il termine "passa-tutto" viene usato per funzioni (che possono essere anche matriciali) a valori unitari sulla circonferenza unità (cioè funzioni per cui vale la condizione (4.21)) ma che non sono necessariamente analitiche in $\{|z| > 1\}$.

DEFINIZIONE 3.2. Sia $\hat{W}(z)$ una matrice $m \times p$ le cui righe, $\hat{w}_k(z)$, $k = 1, \dots, m$, sono funzioni analitiche in $\{|z| > 1\}$ appartenenti allo spazio H_p^2 . Siano $w_k(\cdot) \in \ell_p^{2+}$ le antitrasformate (necessariamente causali) delle \hat{w}_k . La funzione \hat{W} si dice a fase minima o esterna in H_p^2 se

$$\overline{\text{span}}_{\ell_p^2(\mathbb{Z})} \{w_k(\cdot - t); k = 1, \dots, m, t \geq 0\} = \ell_p^{2+} \quad (4.23)$$

o, equivalentemente, se

$$\overline{\text{span}}_{H_p^2} \{\hat{w}_k(z) z^{-t}; k = 1, \dots, m, t \geq 0\} = H_p^2 \quad (4.24)$$

Ovvero, se il sottospazio chiuso di ℓ_p^2 , generato dalla famiglia delle risposte impulsive "traslate" $\{w_k(\cdot - t); k = 1, \dots, m, t \geq 0\}$, coincide con lo spazio ℓ_p^{2+} di tutti i segnali causali p -dimensionali a energia finita.

Nel caso scalare ($m = p = 1$) la condizione (4.23) o (4.24) si può interpretare dicendo che $w(\cdot)$ o, rispettivamente, \hat{w} , è un vettore ciclico per l'operatore di ritardo σ^{-1} (o, rispettivamente, l'operatore di moltiplicazione per z^{-1} in H^2).

LEMMA 3.4. Una matrice $m \times p$, ($m \geq p$), a fase minima in H_p^2 , ha rango uguale a p su tutta la regione $\{|z| > 1\}$ del piano complesso, incluso il punto all'infinito.

DIMOSTRAZIONE. La (4.24) dice che l'operatore di moltiplicazione per $\hat{W}(z)$, $M_W : f(z) \mapsto f(z)\hat{W}(z)$ (ricordare che le funzioni di H_p^2 sono vettori riga), mappa la sottovarietà lineare di H_p^2 dei polinomi analitici (i.e. le serie di Laurent finite del tipo $f(z) = \sum_{k \geq 0} a_k z^{-k}$; $a_k \in \mathbb{R}^m$) in un sottospazio denso in H_p^2 . Dato che i polinomi analitici sono certamente densi in H_p^2 , si può ben dire che l'operatore di moltiplicazione per una matrice a fase minima è suriettivo. In particolare è possibile trovare successioni di polinomi analitici φ_k^n , $n = 1, 2, \dots$ tali che $\varphi_k^n \hat{W}(z) \rightarrow e_k$, $k = 1, \dots, p$, e_k essendo il vettore costante $[0, \dots, 1, \dots, 0]$ con l'1 al k -simo posto (che è una funzione di H_p^2 !). In altri termini, è possibile trovare una successione di matrici analitiche $\{\Phi^n\}$ a righe in H_m^2 tale che

$$\lim_{n \rightarrow \infty} \Phi^n(z) \hat{W}(z) = I_p$$

per tutti gli $z \in \{|z| > 1\}$.

Se esistesse un vettore non nullo $a \in \mathbb{R}^p$ tale che $\hat{W}(z_0)a \equiv 0$ per qualche z_0 nel complemento del cerchio unitario, questo non potrebbe verificarsi. \square

Se il rango di una matrice W di dimensione $m \times p$, ($m \geq p$) che è funzione analitica della variabile complessa z , è minore di p , significa che esiste un numero complesso z_0 ed un vettore $a \in \mathbb{C}^p$ per cui $W(z_0)a = 0$. In questo caso z_0 si chiama uno zero di $W(z)$ e a è la direzione di zero corrispondente. Evidentemente, se W è quadrata, z_0 è uno zero di W se e solo se è uno zero del suo determinante. Il lemma 3.4 si può allora enunciare anche dicendo che una funzione matriciale a fase minima non può avere zeri in $\{|z| > 1\}$, incluso il punto all'infinito.

TEOREMA 3.3. Si ha equivalenza causale fra $\{\mathbf{y}(t)\}$ e il processo $\boldsymbol{\varepsilon}(t) = W(z)^{-L}\mathbf{y}(t)$ se e solo se $W(z)$ è un fattore spettrale a fase minima.

DIMOSTRAZIONE. Dalle relazioni $\mathbf{y}_k(0) = I_\varepsilon[w_k(\cdot)]$, $k = 1, \dots, m$ o, equivalentemente, dalle

$$\mathbf{y}_k(-t) = I_\varepsilon[w_k(\cdot - t)] \quad , \quad k = 1, \dots, m, t \geq 0 \quad ,$$

segue

$$\begin{aligned} \mathbf{H}_0(\mathbf{y}) &= \overline{\text{span}} \left\{ \mathbf{y}_k(-t); k = 1, \dots, m, t \geq 0 \right\} \\ &= I_\varepsilon [\overline{\text{span}} w_k(\cdot - t); k = 1, \dots, m, t \geq 0] \quad . \end{aligned}$$

Se $\hat{W}(z)$ è a fase minima, l'ultimo termine è uguale a $I_\varepsilon[\ell_p^{2+}]$ che è proprio l'insieme $\mathbf{H}_0(\varepsilon)$ di *tutti* i funzionali lineari della storia passata del processo bianco $\{\varepsilon(t)\}$ fino all'istante $t = 0$. Se $\tilde{W}(z)$ è a fase minima si ha dunque $\mathbf{H}_0(\mathbf{y}) = \mathbf{H}_0(\varepsilon)$. Di fatto, se \hat{W} è a fase minima si può sostituire nella (4.23) a $t = 0$ un arbitrario istante t_0 e a ℓ_p^{2+} lo spazio dei segnali a energia finita nulli per $t < t_0$ (la verifica è lasciata al lettore) per cui con lo stesso ragionamento si perviene in realtà all'uguaglianza $\mathbf{H}_{t_0}(\mathbf{y}) = \mathbf{H}_{t_0}(\varepsilon)$ per ogni t_0 .

Viceversa, dall'equivalenza causale si deduce in particolare $\mathbf{H}_0(\mathbf{y}) = \mathbf{H}_0(\varepsilon)$ e ripercorrendo a ritroso il ragionamento fatto, si arriva a concludere che deve valere la (4.23). \square

Come si riconoscono le funzioni a fase minima? Le funzioni a fase minima scalari ammettono una rappresentazione integrale esplicita che si può trovare ad esempio nel testo di Hoffmann [65]. La loro caratterizzazione analitica in generale è un pò complicata e noi non la riporteremo. Qui ci accontenteremo di illustrare una condizione necessaria che tornerà utile per caratterizzare il caso razionale.

LEMMA 3.5. *Una matrice $W(z)$ a righe in H_p^2 è a fase minima se e solo se non ammette fattori (destri) inner non banali. In altri termini, in ogni eventuale fattorizzazione del tipo*

$$W(z) = \tilde{W}(z) Q(z)$$

con $\tilde{W}(z)$ analitica in $\{|z| > 1\}$ e $Q(z)$ $p \times p$, matrice inner, $Q(z)$ deve essere una matrice ortogonale costante.

DIMOSTRAZIONE. Questa proposizione è sostanzialmente un corollario di un risultato "avanzato" della teoria delle funzioni analitiche, il teorema di Beurling che è enunciato in appendice B che non vale qui la pena di richiamare esplicitamente. Cercheremo qui di darne una dimostrazione elementare, accontentandoci di discutere il caso scalare.

Sia $\hat{w}(z)$ una funzione a fase minima scalare e supponiamo che essa ammetta uno zero di molteplicità k in $z = a$ con $|a| > 1$. Se a è reale si può scrivere

$$\hat{w}(z) = w_0(z) \left(\frac{1 - z/a}{z - 1/a} \right)^k \quad (4.25)$$

con $w_0(z)$ analitica in $\{|z| > 1\}$, $w_0(a) \neq 0$, e dotata di uno zero "riflesso" (dentro il cerchio unitario) in $z = 1/a$ di uguale molteplicità k . Nel caso in cui a è complesso c'è ovviamente anche lo zero coniugato e bisognerà moltiplicare il secondo membro in (4.25) per lo stesso fattore in cui al posto di a compare il coniugato \bar{a} . Per non appesantire la trattazione lasceremo la discussione di questo caso al lettore.

Il fattore $Q(z) = \left(\frac{1 - z/a}{z - 1/a} \right)^k$ è inner. La dimostrazione si basa sul seguente fatto: Per ogni funzione inner non costante, $Q(z)$, esistono funzioni $\hat{g}(z) \in H^2$ tali che $\bar{g}(z) := Q(1/z) \hat{g}(z)$ è la trasformata zeta di una successione (strettamente)

anticausale appartenente al complemento ortogonale, $(H^2)^\perp$, dello spazio H^2 in $L^2(\mathbf{C})$.

Nel nostro caso basta ad esempio prendere

$$\hat{g}(z) = \frac{z}{(z - 1/a)^k} \in H^2$$

per ottenere

$$\bar{g}(z) := Q(1/z) \hat{g}(z) = \frac{z}{(1 - z/a)^k}$$

che è la z -trasformata di un segnale a energia finita anticausale, nullo in $t = 0$. Mostriamo ora che *tutte* le successioni (causali) $f(\cdot)$ del sottospazio $\overline{\text{span}}\{w(\cdot - t); t \geq 0\} \subseteq \ell^{2+}$ sono ortogonali all'antitrasformata, $g(\cdot)$, di $\hat{g}(z) \in H^2$. In questo modo si mostra che $\hat{w}(z)$ non può essere a fase minima giacché $\overline{\text{span}}\{w(\cdot - t); t \geq 0\}$ non può ovviamente esaurire tutto ℓ^{2+} , essendo ortogonale a $g(\cdot)$.

Per mostrare l'ortogonalità si ragiona prima sulle funzioni $\hat{f}(z)$ della varietà lineare, diciamola W , delle combinazioni lineari finite $\text{span}\{z^{-t} \hat{w}(z); t \geq 0\}$. Dalla fattorizzazione (4.25) segue che ognuna di queste funzioni si può scrivere come prodotto $\varphi(z) Q(z)$ con $\varphi(z) \in H^2$. Calcolando il prodotto scalare in $L^2(\mathbf{C})$ e ricordando che $Q^*(e^{j\omega}) = Q(e^{-j\omega})$ si trova

$$\langle f(\cdot), g(\cdot) \rangle_{\ell^2} = \langle \hat{f}, \hat{g} \rangle_{L^2} = \langle \varphi Q, \hat{g} \rangle_{L^2} = \langle \varphi, Q^* \hat{g} \rangle_{L^2} = \langle \varphi, \bar{g} \rangle_{L^2} = 0 \quad .$$

Ora, la trasformata, \hat{f} , di un qualunque elemento $f \in \overline{\text{span}}\{w(\cdot - t); t \geq 0\}$ è un limite in H^2 di una successione $\{\hat{f}_n\} \subset W$. Per la continuità del prodotto scalare in $L^2(\mathbf{C})$ si ha così

$$\langle f, g \rangle_{\ell^2} = \lim_{n \rightarrow \infty} \langle \hat{f}_n, \hat{g} \rangle_{L^2} = 0$$

per ogni f , il che è quanto si voleva dimostrare. \square

Il teorema seguente descrive la famiglia dei fattori spettrali analitici, cosiddetti di rango "pieno" (p), di un processo p.n.d.

TEOREMA 3.4. *Se la matrice densità spettrale S ammette fattori analitici in $\{|z| > 1\}$, essa ne ammette anche uno a fase minima, \hat{W} . Questo fattore è unico a meno di moltiplicazione a destra per un'arbitraria matrice ortogonale costante di dimensione $p \times p$.*

Ogni altro fattore analitico di uguale dimensione $m \times p$, $G(z)$, è legato al fattore a fase minima, $\hat{W}(z)$, dalla relazione

$$G(z) = \hat{W}(z) Q(z) \quad , \quad (4.26)$$

dove $Q(z)$ è una funzione passa-tutto $p \times p$.

DIMOSTRAZIONE. Le righe di ogni fattore $m \times p$, diciamolo $G(z)$, analitico in $\{|z| > 1\}$, appartengono necessariamente ad H_p^2 . Se $G(z)$ non ammette fattori destri passa-tutto (non banali), è già a fase minima per il lemma 3.5. D'altro canto, se $G(z)$ ammette fattorizzazioni del tipo $G(z) = W(z) Q(z)$ con W analitica e Q passa-tutto $p \times p$, è ovvio che anche $W(z)$ deve essere un fattore spettrale analitico.

Ripetendo il ragionamento su $W(z)$, si conclude che ogni fattore analitico si può scrivere nella forma (4.26) con $\hat{W}(z)$ fattore spettrale a fase minima.

Rimane da dimostrare che il fattore a fase minima è unico. Supponiamo che ve ne sia più di uno. Per fissare le idee, supponiamo che si possa scrivere

$$\mathbf{y}(t) = \sum_{-\infty}^t W_1(t-s) \boldsymbol{\varepsilon}_1(s) = \sum_{-\infty}^t W_2(t-s) \boldsymbol{\varepsilon}_2(s) \quad (4.27)$$

con $\{\boldsymbol{\varepsilon}_i(t)\}$ processi m -dimensionali ortonormali e $\{W_i(t)\}$ successioni causali a quadrato sommabile, $i = 1, 2$.

Per l'equivalenza causale si avrà allora

$$\mathbf{H}_t(\boldsymbol{\varepsilon}_1) = \mathbf{H}_t(\boldsymbol{\varepsilon}_2) \quad , \quad \forall t \in \mathbb{Z} \quad .$$

Quindi $\boldsymbol{\varepsilon}_1(t)$ si può ottenere come uscita di un filtro *causale* ed ℓ^2 -stabile pilotato dal rumore bianco $\{\boldsymbol{\varepsilon}_2(t)\}$

$$\boldsymbol{\varepsilon}_1(t) = \sum_{-\infty}^t Q_1(t-s) \boldsymbol{\varepsilon}_2(s) \quad . \quad (4.28)$$

Viceversa, si potrà analogamente esprimere $\boldsymbol{\varepsilon}_2(t)$ come

$$\boldsymbol{\varepsilon}_2(t) = \sum_{-\infty}^t Q_2(t-s) \boldsymbol{\varepsilon}_1(s) \quad , \quad (4.29)$$

in cui $\{Q_2(t)\}$ è ancora una successione causale ed ℓ^2 -stabile. Ora, la matrice densità spettrale di potenza dei due processi bianchi $\{\boldsymbol{\varepsilon}_i(t)\}$, $i = 1, 2$, è l'identità per cui

$$I_p = \hat{Q}_1(e^{j\omega}) \hat{Q}_1(e^{j\omega})^* = \hat{Q}_2(e^{j\omega}) \hat{Q}_2(e^{j\omega})^* \quad . \quad (4.30)$$

Quindi le matrici di trasferimento dei due sistemi lineari (4.28) e (4.29) sono passa-tutto.

Inoltre, imponendo le (4.28) e (4.29) si ricava che $\hat{Q}_1(z)$ e $\hat{Q}_2(z)$ sono ciascuna l'inversa dell'altra in $\{|z| \geq 1\}$.

Ora, ogni funzione passa-tutto $Q(z)$ è (definita e) analitica in $\{|z| > 1\}$ e limitata sulla circonferenza unità (dato che è unitaria in $\{|z| = 1\}$, cfr. la (4.21)). Dalla teoria delle funzioni analitiche ([51, pag. 189, Th. 12.27]) è noto che una funzione passa-tutto Q si può prolungare all'interno del cerchio $\{|z| < 1\}$, attraversando la circonferenza unitaria, in tutti (e soli) i punti $z = e^{j\omega_0}$ per cui $|Q(e^{j\omega_0})^{-1}|$ è limitata e in un intorno di questi punti il prolungamento analitico è dato dalla formula

$$Q(z) = \left(Q(1/z)'\right)^{-1} \quad \text{per} \quad \{|z| < 1\} \quad . \quad (4.31)$$

Ovviamente $Q(z)$ ha dei poli in $\{|z| < 1\}$ (esattamente in coincidenza degli zeri di $Q(1/z)'$ in $|1/z| > 1$) e quindi l'estensione è in generale solo locale.

Nel caso di $\hat{Q}_1(z)$ si ha però $\hat{Q}_1(e^{j\omega})^{-1} = \hat{Q}_2(e^{j\omega})$ che è uniformemente limitata sulla circonferenza e inoltre l'estensione all'interno del cerchio unitario,

$$\hat{Q}_1(z) = \left(\hat{Q}_1(1/z)'\right)^{-1} = \hat{Q}_2(1/z)' \quad , \quad \{|z| < 1\} \quad ,$$

è definita per ogni z del cerchio unitario $\{|z| < 1\}$, giacché $\hat{Q}_2(1/z)$ è analitica (e uniformemente limitata) in $|1/z| > 1$. Questo implica che $\hat{Q}_1(z)$ è analitica su *tutto il piano complesso* e quindi deve essere una *matrice unitaria costante*. Il ragionamento si applica per simmetria anche a $\hat{Q}_2(z)$. Nella (4.27) W_1 e W_2 quindi differiscono per la moltiplicazione a destra per una matrice ortogonale costante $p \times p$. \square

Il fattore spettrale canonico e la rappresentazione d'innovazione. Sarà utile in seguito sfruttare l'arbitrarietà di scelta della matrice ortogonale costante per dare al fattore spettrale a fase minima del teorema precedente una struttura speciale particolarmente semplice, che chiameremo *canonica*. Faremo questo solo per *processi di rango pieno* ($p = m$), che è la situazione che si incontra più di frequente nei problemi di stima.

Se il rango (q.o.) di $S(e^{j\omega})$ è m , è ovvio che anche il fattore spettrale a fase minima, $\hat{W}(z)$, dovrà essere quadrato di rango m , cioè non singolare, in $\{|z| > 1\}$. Dallo sviluppo di $\hat{W}(z)$,

$$\hat{W}(z) = \sum_0^{+\infty} W(t) z^{-t} \quad , \quad \{|z| > 1\} \quad , \quad (4.32)$$

si vede che $\hat{W}(\infty) = W(0)$. Scende dal lemma 3.4 che anche questa matrice deve essere *non singolare*. Questo fatto è così importante che merita di essere enunciato separatamente.

PROPOSIZIONE 3.1. *Se il processo y è a rango pieno e $\{W(t)\}$ è l'antitrasformata zeta del fattore spettrale analitico a fase minima $\hat{W}(z)$, il termine di grado zero, $W(0)$, nello sviluppo (4.32) è una matrice non singolare.*

DIMOSTRAZIONE. Daremo una dimostrazione “probabilistica” che non fa uso del lemma 3.4. Detto $\{\hat{\varepsilon}(t)\}$ il rumore bianco corrispondente a $\hat{W}(z)$, si ha

$$\mathbf{y}(t) = W(0) \hat{\varepsilon}(t) + W(1) \hat{\varepsilon}(t-1) + \dots$$

Se $a \in \mathbb{R}^m$ è nello spazio nullo di $W(0)$ (cioè $W(0)a = 0$), la variabile casuale $\boldsymbol{\eta} := a' \hat{\varepsilon}(t) \in H_t(\hat{\varepsilon})$ è ortogonale ad $H_t(\mathbf{y})$. Infatti si ha $\boldsymbol{\eta} \perp H_{t-1}(\hat{\varepsilon}) = H_{t-1}(\mathbf{y})$ ed $\boldsymbol{\eta} \perp \mathbf{y}(t)$ giacché

$$E \mathbf{y}(t) \boldsymbol{\eta} = W(0) a = 0 \quad .$$

Dato che per ipotesi $H_t(\hat{\varepsilon}) = H_t(\mathbf{y})$, l'unica possibilità è che sia $a = 0$. \square

PROBLEMA 3.2. *Dimostrare che vale anche il viceversa. Pi'ù in generale, dimostrare che un processo stazionario p.n.d. ha rango p se e solo se $W(0)$ ha rango p . Giustificare in questo modo l'affermazione che si trova qualche volta in letteratura, che il rango di \mathbf{y} è uguale al rango della matrice di covarianza dell'innovazione non normalizzata Λ definita più sotto (vedere la (4.35)).* \diamond

Si può così usare la libertà di scegliere il fattore ortogonale costante U in modo tale da rendere $\hat{W}(\infty)U$ triangolare inferiore con elementi positivi sulla diagonale.

Questa particolare normalizzazione di $\hat{W}(z)$ verrà denotata⁷ con il simbolo $\hat{L}(z)$ (“L” sta per “Lower triangular”). Chiameremo $\hat{L}(z)$ *fattore spettrale canonico*.

È possibile viceversa normalizzare $\hat{W}(\infty)$ all'identità. Scegliendo il fattore a fase minima come

$$\hat{W}(z) := \hat{L}(z) \hat{L}(\infty)^{-1} \quad , \quad (4.33)$$

si ottiene la rappresentazione di $\mathbf{y}(t)$ in termini di un processo bianco non normalizzato

$$\mathbf{e}(t) := \hat{L}(\infty) \boldsymbol{\varepsilon}(t) \quad (4.34)$$

⁷Spesso il circonflesso “^” verrà omissso.

di varianza

$$E \mathbf{e}(t) \mathbf{e}(t)' = \Lambda = \hat{L}(\infty) \hat{L}(\infty)' > 0 \quad . \quad (4.35)$$

L'equazione di fattorizzazione spettrale corrispondente al rumore bianco non normalizzato $\{\mathbf{e}(t)\}$ si modifica corrispondentemente in

$$S(e^{j\omega}) = \hat{W}(e^{j\omega}) \Lambda \hat{W}(e^{j\omega})^* \quad , \quad \omega \in [-\pi, \pi] \quad . \quad (4.36)$$

DEFINIZIONE 3.3. *Chiameremo le rappresentazioni del processo p.n.d. $\{\mathbf{y}(t)\}$,*

$$\mathbf{y}(t) = \hat{L}(z) \boldsymbol{\varepsilon}(t) \quad , \quad \mathbf{y}(t) = \hat{W}(z) \mathbf{e}(t) \quad , \quad (4.37)$$

dove $\hat{L}(z)$ è il fattore spettrale canonico e $\hat{W}(z)$ il fattore spettrale a fase minima normalizzato all'identità all'infinito, la rappresentazione d'innovazione normalizzata e non normalizzata, rispettivamente. I processi bianchi $\{\boldsymbol{\varepsilon}(t)\}$ ed $\{\mathbf{e}(t)\}$ corrispondenti a $\hat{L}(z)$ e $\hat{W}(z)$ saranno denominati innovazione normalizzata e innovazione di $\{\mathbf{y}(t)\}$.

Naturalmente a questo punto è d'obbligo, per giustificare la terminologia, dare un'interpretazione dei processi $\{\boldsymbol{\varepsilon}(t)\}$ ed $\{\mathbf{e}(t)\}$ in termini di ortonormalizzazione "alla Gram-Schmidt" del processo $\{\mathbf{y}(t)\}$.

Definiamo

$$\hat{\mathbf{y}}(t) := \sum_1^{\infty} W(k) \mathbf{e}(t-k)$$

e riscriviamo la seconda rappresentazione in (4.37) nel dominio del tempo come

$$\mathbf{y}(t) = \mathbf{e}(t) + \hat{\mathbf{y}}(t) \quad .$$

È ovvio che la somma a secondo membro è ortogonale. In effetti, dalla rappresentazione (4.37), $\hat{\mathbf{y}}(t) \in \mathbf{H}_{t-1}(\mathbf{e}) = \mathbf{H}_{t-1}(\mathbf{y})$ per cui $\hat{\mathbf{y}}(t)$ è evidentemente la proiezione ortogonale di $\mathbf{y}(t)$ su $\mathbf{H}_{t-1}(\mathbf{y})$, in altre parole $\hat{\mathbf{y}}(t)$ è il predittore a minima varianza di $\mathbf{y}(t)$ basato su tutta la storia passata, $\mathbf{H}_{t-1}(\mathbf{y})$, di $\{\mathbf{y}(t)\}$

$$\hat{\mathbf{y}}(t) = \hat{\mathbf{y}}(t | t-1) = \hat{E} [\mathbf{y}(t) | \mathbf{H}_{t-1}(\mathbf{y})]$$

e il termine $\mathbf{e}(t) \perp \mathbf{H}_{t-1}(\mathbf{e}) = \mathbf{H}_{t-1}(\mathbf{y})$, il relativo errore di predizione di un passo

$$\mathbf{e}(t) = \mathbf{y}(t) - \hat{\mathbf{y}}(t | t-1) \quad . \quad (4.38)$$

Ovviamente $\boldsymbol{\varepsilon}(t)$ ha il significato corrispondente, di errore di predizione normalizzato.

Fattorizzazione spettrale e fattorizzazione di Cholesky. Sia $\{\mathbf{y}(t)\}$ un processo p.n.d. di rango p di covarianza $\Sigma(\tau) := E\mathbf{y}(t+\tau) \mathbf{y}(t)'$, $\tau \in \mathbb{Z}$ e $\hat{W}(z) = \sum_0^{\infty} W(t) z^{-t}$ un fattore spettrale analitico di dimensione $m \times p$. Riscrivendo in notazione matriciale la seconda uguaglianza della formula (4.15) si trova una relazione notevole che lega $\Sigma(\cdot)$ alla risposta impulsiva $W(\cdot)$ del corrispondente "filtro formante",

$$\Sigma(\tau) = \sum_{t=0}^{\infty} W(t+\tau) W(t)' \quad \tau \in \mathbb{Z}. \quad (4.39)$$

Notare che questa formula è la relazione di fattorizzazione spettrale scritta nel dominio del tempo.

PROBLEMA 3.3. *Dimostrare direttamente la relazione $\Sigma(-\tau) = \Sigma(\tau)'$ mediante il cambio di variabile $s = t + \tau$ nella somma (4.39). Notare che il limite inferiore si può anche porre uguale a $-\infty$.* \diamond

Facciamo vedere che la (4.39) si può interpretare anche come fattorizzazione di Cholesky a blocchi di una opportuna matrice varianza (infinita) del processo $\{\mathbf{y}(t)\}$. Definiamo allo scopo le matrici

$$\Sigma_0 = \Sigma(0), \Sigma_1 = \begin{bmatrix} \Sigma(0) & \Sigma(-1) \\ \Sigma(1) & \Sigma(0) \end{bmatrix}, \dots, \Sigma_n = \begin{bmatrix} \Sigma(0) & \dots & \Sigma(-n) \\ \vdots & \ddots & \\ \Sigma(n) & & \Sigma(0) \end{bmatrix}, \dots$$

(Σ_n ha ovviamente il significato di matrice varianza del vettore $\mathbf{y}_n := [\mathbf{y}(t)', \dots, \mathbf{y}(t+n)']'$) e sia Σ_∞ la matrice infinita ottenuta “espandendo” indefinitamente Σ_n (per $n \rightarrow \infty$). Definendo la matrice doppiamente infinita W_∞ a struttura triangolare inferiore,

$$W_\infty := \begin{bmatrix} \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & W(0) & 0 & 0 & 0 & \dots \\ \dots & W(1) & W(0) & 0 & 0 & \dots \\ \dots & W(2) & W(1) & W(0) & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (4.40)$$

si verifica facilmente che la relazione (4.39) si può formalmente scrivere come fattorizzazione “alla Cholesky” di Σ_∞ ,

$$\Sigma_\infty = W_\infty W_\infty' \quad (4.41)$$

In particolare, ciascuna sottomatrice Σ_n fattorizza come prodotto di una sottomatrice, $W_{n,\infty}$, costituita da n blocchi riga contigui di W_∞ , per la relativa trasposta

$$\Sigma_n = W_{n,\infty} W_{n,\infty}' \quad (4.42)$$

PROBLEMA 3.4. *Usando questa fattorizzazione, dimostrare che per un processo p.n.d. di rango pieno tutte le sottomatrici Σ_n di Σ_∞ sono definite positive.* \diamond

Il filtro sbiancante. Concludiamo questo capitolo cercando di ottenere una interpretazione un po' più concreta del concetto di equivalenza causale tra $\{\mathbf{y}(t)\}$ ed $\{\boldsymbol{\varepsilon}(t)\}$.

La condizione (4.2) implica che, per ogni t , $\boldsymbol{\varepsilon}(t) \in \mathbf{H}_t(\mathbf{y})$, che si interpreta dicendo che $\boldsymbol{\varepsilon}(t)$ è una funzione lineare dei valori passati e presenti di $\{\mathbf{y}(t)\}$, all'istante t . Ciò suggerisce che $\boldsymbol{\varepsilon}(t)$ si possa ottenere come uscita di un sistema dinamico lineare causale e tempo-invariante il cui ingresso è proprio il processo originario, $\{\mathbf{y}(t)\}$.

In effetti, detta $\Gamma(z)$ la matrice inversa del fattore spettrale canonico

$$\Gamma(z) := L(z)^{-1}$$

si ha:

$$\Gamma(e^{j\omega}) S(e^{j\omega}) \Gamma'(e^{-j\omega}) = I \quad (4.43)$$

per cui le righe di Γ sono funzioni di $L_m^2(\mathbf{C}, Sd\omega/2\pi)$. Detta allora $G(\cdot)$ l'antitrasformata generalizzata di Γ in $\ell_m^2(\mathbb{Z}, \Sigma)$ (si veda l'Appendice B) si può scrivere

$$\boldsymbol{\varepsilon}(t) = \sum_{-\infty}^t G(t-s) \mathbf{y}(s) \quad (4.44)$$

e la serie a secondo membro è convergente in media quadratica (Corollario D.2 dell'Appendice B). Si può quindi interpretare il primo membro di (4.43) come la matrice densità spettrale del processo $\{\boldsymbol{\varepsilon}(t)\}$ definito dalla (4.44). Se ne ricava che $\{\boldsymbol{\varepsilon}(t)\}$ è effettivamente rumore bianco normalizzato. Il filtro lineare (4.44) viene chiamato *filtro sbiancante*.⁸

Ricordiamo però che il significato operativo della (4.44) sta nella possibilità di interpretare la somma di convoluzione come *operazione di filtraggio, applicabile alle singole traiettorie di $\{\mathbf{y}(t)\}$* . Ora $\Gamma(z)$, seppure analitica in $\{|z| > 1\}$, può avere singolarità, ad esempio poli, sulla circonferenza unitaria $\{|z| = 1\}$ e in questo caso la somma di convoluzione, benché convergente in media quadratica, potrebbe benissimo *non convergere nel senso ordinario* anche per un insieme di traiettorie di $\{\mathbf{y}(t)\}$ di probabilità uno.

Nel prossimo capitolo dimostreremo che l'interpretazione sistemistica di (4.44) è perfettamente legittima se S è razionale e se vale una maggiorazione del tipo

$$S(e^{j\omega}) > cI \quad , \quad c > 0 \quad , \quad (4.45)$$

la cui validità è chiaramente equivalente all'assenza di zeri di $S(z)$ sulla circonferenza unita.

Se, viceversa, la condizione (4.45) non vale, sorgono dei problemi che non sono di carattere totalmente accademico. Infatti in questo caso la $\Gamma(z)$ è la funzione di trasferimento di un sistema generalmente *instabile* e il significato pratico della formula (4.44) diventa questionabile.

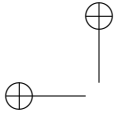
Commenti conclusivi. In questo Capitolo abbiamo generalizzato al caso dei processi stazionari del secondo ordine la nozione di rappresentazione d'innovazione introdotta al Capitolo 2. La difficoltà principale di questa generalizzazione sta nel fatto che l'intervallo temporale in cui si vogliono ortonormalizzare i dati è inferiormente illimitato ($t_0 = -\infty$). In questo caso l'applicazione "ingenua" dell'algoritmo di Gram-Schmidt non porta a procedure costruttive. Occorre prima individuare la classe dei processi che possono essere descritti tramite innovazioni (ad esempio, per il processo dell'esempio 3.1 l'algoritmo di Gram-Schmidt fornisce $e(t) = 0$ ad ogni passo).

Per i processi ortonormalizzabili p.n.d., l'analogo della fattorizzazione di Cholesky della matrice di varianza è dato dalla fattorizzazione spettrale (4.14). Si noti che in questo contesto la "bi-causalità" della rappresentazione è in un caso equivalente alla struttura triangolare inferiore di L (e della sua inversa) e nell'altro alla condizione di *fase minima* (o canonicità) di $L(z)$. Questa analogia tra le due fattorizzazioni non è soltanto formale e verrà ripresa e utilizzata più avanti.

Esercizi

3-1 Usando l'equivalenza causale e basandosi sul fatto che $\hat{E}\{\mathbf{y}(t) \mid \mathbf{e}(s); s \in \mathbb{Z}\} = \mathbf{y}(t)$, verificare che vale la relazione $S(z) = S_{ye}(z)\Lambda^{-1}S_{ey}(z)$ dove $\Lambda = \text{Var}\{\mathbf{e}(t)\} > 0$. Notare che $S_{ye}(z)$ è la trasformata zeta di una successione causale e pertanto la matrice $W(z) := S_{ye}(z)\Lambda^{-1/2}$, è analitica in $\{|z| > 1\}$. Usando queste considerazioni, dimostrare direttamente che l'equivalenza causale implica l'esistenza di fattori spettrali analitici fuori dal cerchio.

⁸“Whitening filter” in letteratura.



3-2 (Generalizzazione del Teorema di Rappresentazione 3.1).

Sia $\{\mathbf{w}(t)\}$ un processo bianco normalizzato, congiuntamente stazionario ad un processo ortonormalizzabile $\{\mathbf{y}(t)\}$. Supponiamo che $\dim \mathbf{w}(t) = r \geq p = \text{rango} \{\mathbf{y}(t)\}$ e che $\mathbf{H}(\mathbf{w}) \supset \mathbf{H}(\mathbf{y})$.

Mostrare che $\{\mathbf{y}(t)\}$ ammette una rappresentazione del tipo (4.5) con W di dimensione $m \times r$ e “ingresso bianco” $\{\mathbf{w}(t)\}$.

Si può avere uguaglianza tra i due spazi di Hilbert $\mathbf{H}(\mathbf{w})$, $\mathbf{H}(\mathbf{y})$ se $r > p$?

3-3 Sia ε un processo ortonormale che soddisfa alla condizione (4.1). Supponiamo di usare $m - p$ generatori di rumore bianco e di aggiungere alle p componenti di ε altre $r - p$ componenti scorrelate, ε_{p+k} , $k = 1, \dots, r - p$ in modo tale che il processo “aumentato” r -dimensionale così ottenuto sia ancora bianco. Usando il risultato dell’esercizio precedente, generalizzare il Teorema di fattorizzazione spettrale 3.2 riscrivendo la condizione di fattorizzazione dello spettro in termini di fattori spettrali $W(e^{j\omega})$ di dimensione $m \times r$, con r qualunque, purchè $r \geq p$. Dimostrare che la condizione (4.16) continua a valere anche per questi fattori spettrali “maggiorati”.

3-4 Considerare la rappresentazione definita nel Problema 3-2. Mostrare (confrontando gli spazi di Hilbert generati dai processi in questione) che non è possibile ricostruire il processo bianco $\{\mathbf{w}(t)\}$ per mezzo di un filtro lineare con ingresso il solo processo $\{\mathbf{y}(t)\}$ e che per fare questo, in ingresso al filtro sbiancante bisognerebbe aggiungere un altro processo stazionario completamente scorrelato da $\{\mathbf{y}(t)\}$. Dare una spiegazione intuitiva della risposta.

3-5 Sia $G(z)$ un fattore spettrale analitico in $\{|z| > 1\}$ di dimensione $m \times r$, $r \geq p$. Mostrare che ogni $G(z)$ di questo tipo si può fattorizzare come $G(z) = \hat{W}(z) Q(z)$, dove $\hat{W}(z)$ è il fattore a fase minima e $Q(z)$ è una matrice $p \times r$ analitica in $\{|z| > 1\}$ che soddisfa alla condizione

$$Q(z) Q \left(\frac{1}{z} \right)' = I_p \quad .$$

3-6 Dimostrare che la funzione $w(z) = 1 - z^{-1}$ è a fase minima in H^2 . (Mostrare che con opportune combinazioni lineari delle $\{w(z) z^{-t}, t = 0, 1, \dots\}$ si ottengono tutti i monomi $\{z^{-t}, t \geq 0\}$ e poi applicare la definizione).

3-7 Nella fattorizzazione (4.42) la matrice infinita $W_{n,\infty}$ ha la struttura

$$\begin{bmatrix} \dots & W(1) & W(0) & 0 & \dots & 0 \\ \dots & W(2) & W(1) & W(0) & \dots & 0 \\ \dots & & & \ddots & \ddots & \\ \dots & W(n) & \dots & \dots & W(1) & W(0) \end{bmatrix} \quad .$$

Mostrare che, se $p = \text{rango} \{\mathbf{y}(t)\} = m$, le righe di questa matrice sono tutte linearmente indipendenti.

(Si proceda nel modo seguente.

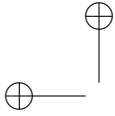
Sia $a' = [a'_1, \dots, a'_n]'$, $a_k \in \mathbb{R}^m$, un vettore nello spazio nullo sinistro di $W_{n,\infty}$, i.e.,

$a' W_{n,\infty} = 0$. Dalle

$$a'_n W(0) = 0, a'_{n-1} W(0) + a'_n W(1) = 0, \dots, \sum_1^n a'_k W(t+k) = 0, \\ t \geq 0,$$

ricavare una relazione nel dominio della frequenza, che contraddice la rango $W(e^{j\omega}) = m$.

Che cosa si può dire nel caso in cui $p < m$?



CAPITOLO 4

FATTORIZZAZIONE SPETTRALE RAZIONALE E MODELLI ARMA

1. Introduzione

Da quanto abbiamo esposto finora, è evidente che il passo fondamentale per la costruzione esplicita di rappresentazioni di un processo stazionario come uscita di un filtro lineare pilotato da rumore bianco consiste sostanzialmente nella soluzione di un problema di *fattorizzazione spettrale*. Come vedremo nel seguito, questo problema è di fatto il problema fondamentale cui si riduce tutta la teoria del filtraggio statistico “alla Wiener-Kolmogorov”.

È bene dire subito che in pratica si sa come eseguire la fattorizzazione spettrale solo nel caso di densità spettrali di potenza che sono *funzioni razionali di $e^{j\omega}$* . In questo capitolo studieremo appunto la fattorizzazione spettrale di spettri razionali e analizzeremo la struttura delle rappresentazioni di un processo stazionario che corrispondono a fattori spettrali razionali. C'è da dire che la restrizione a spettri razionali è sufficiente per moltissime applicazioni; in linea di principio si può poi sempre approssimare lo spettro di un arbitrario segnale stazionario ortonormalizzabile mediante uno spettro razionale di “complessità” (ordine) sufficientemente elevata.

Dato che le funzioni razionali hanno automaticamente estensione dalla circonferenza unità $\{z = e^{j\omega}\}$ a tutto il piano complesso (eccettuato al più un numero finito di poli) scriveremo sempre gli spettri come funzioni della variabile complessa z , definite in tutto il piano complesso. Ricordiamo anche che i processi stazionari di cui ci occupiamo sono *reali* e hanno quindi densità spettrali che sono funzioni (in generale matriciali $m \times m$) di tipo *reale*. In particolare l'aggiunta $S(e^{j\omega})^*$ della matrice $S(e^{j\omega})$ è uguale a $S(e^{-j\omega})'$.

Inizieremo studiando il caso dei processi scalari.

2. Fattorizzazione spettrale razionale: Il caso scalare

Sia $S(z)$ la densità spettrale di un processo *scalare* stazionario del secondo ordine $\{\mathbf{y}(t)\}$. Come è noto, $S(z)$ ha le seguenti proprietà:

- 1) $S(z) = S(z^{-1})$
- 2) $S(e^{j\omega}) \geq 0$, $\omega \in [-\pi, \pi]$
- 3) $\int_{-\pi}^{\pi} S(e^{j\omega}) \frac{d\omega}{2\pi} = \sigma_{\mathbf{y}}(0) = E \mathbf{y}(t)^2 < \infty$.

Supponiamo che $S(z)$ sia *una funzione razionale di z* .

Per la proprietà 1), se S si annulla in z_0 , certamente anche $1/z_0$ è uno zero di S . Analogamente, se p_0 è un polo di S anche $1/p_0$ dovrà esserlo. In entrambi i casi le molteplicità dovranno essere le stesse. Ne deriva che $S(z)$ si può scrivere nella forma

$$S(z) = cz^r \frac{\prod_1^m (z - \alpha_i) (z - 1/\alpha_i)}{\prod_1^n (z - \beta_i) (z - 1/\beta_i)} , \quad (2.1)$$

dove α_i e β_i sono reali (e diversi da zero) oppure, se complessi, appaiono assieme ai loro coniugati (questo perchè S è una funzione *reale* della variabile complessa z). In ogni caso non c'è perdita di generalità se prendiamo α_i e β_i tali che $|\alpha_i| \leq 1$, $|\beta_i| \leq 1$. Notiamo che per l'integrabilità di $S(z)$ sulla circonferenza unità (condizione 3) *non ci possono essere poli a modulo unitario*, per cui in realtà si avrà $|\beta_i| < 1$.

Dalla $(z - 1/\alpha_i) = -z/\alpha_i(z^{-1} - \alpha_i)$ scende poi che la (2.1) si può scrivere:

$$S(z) = \lambda^2 z^k \frac{\prod_1^m (z - \alpha_i) (z^{-1} - \alpha_i)}{\prod_1^n (z - \beta_i) (z^{-1} - \beta_i)} , \quad (2.2)$$

dove, per la condizione di positività 2), λ^2 dev'essere reale e non negativo. Dalla 1) segue allora che $z^k = z^{-k}$, per cui k *deve essere zero*. In conclusione, si può scrivere:

$$S(z) = G(z) G(z^{-1}) \quad ; \quad G(z) := \lambda \frac{\prod_1^m (z - \alpha_i)}{\prod_1^n (z - \beta_i)} . \quad (2.3)$$

Abbiamo così costruito un fattore spettrale, $G(z)$, di $S(z)$ che ha le seguenti proprietà.

- a) I *poli* (al finito) di $G(z)$ stanno tutti dentro il cerchio unitario, dato che $|\beta_i| < 1$. Quindi $G(z)$ è un fattore spettrale analitico (in $\{|z| > 1\}$) a meno di possibili poli all'infinito. Interpretando $G(z)$ come la funzione di trasferimento di un sistema lineare (non necessariamente causale), diremo che essa è *strettamente stabile*¹.
- b) Gli *zeri* (al finito) di $G(z)$ stanno tutti nel cerchio unitario chiuso (dato che $|\alpha_i| \leq 1$). Notiamo che gli eventuali zeri di $S(z)$ sulla circonferenza unità, $\{z \mid |z| = 1\}$, debbono avere molteplicità *pari* altrimenti sarebbe $S(e^{j\omega}) < 0$ in un loro intorno. Essi possono così venire assegnati per metà al fattore $G(z)$ e per metà al coniugato $G(z^{-1})$.

Notiamo che gli zeri al finito di $G(z)$ sono all'interno del cerchio unitario, cioè in $\{|z| < 1\}$, se e solo se la condizione 2) vale in senso stretto, cioè la densità spettrale $S(e^{j\omega})$ è strettamente positiva.

Da questo fattore spettrale, con un piccolo ritocco che sistema il comportamento all'infinito, è immediato ottenere quello a fase minima.

¹Useremo spesso questa terminologia "controllistica" nel caso di funzioni di trasferimento razionali.

PROPOSIZIONE 4.1. *Ogni densità spettrale razionale scalare $S(z)$ ammette un fattore spettrale, $L(z)$, che gode delle proprietà a) e b) sopra elencate e in più ha numeratore e denominatore di uguale grado, pari al massimo tra m ed n . Questo fattore si ottiene moltiplicando $G(z)$, definita in (2.3), per un opportuno fattore monomio z^k in modo da ottenere l'uguaglianza $n = m$.*

Il fattore spettrale $L(z)$ è a fase minima. Esso coincide con il fattore spettrale canonico definito nel Capitolo 4 ed è unico a meno della moltiplicazione per il fattore scalare ± 1 . Esso può essere interpretato come la funzione di trasferimento di un filtro lineare causale, strettamente stabile e causalmente invertibile.

DIMOSTRAZIONE. Come è noto dal capitolo precedente, una funzione razionale analitica e priva di zeri in $\{|z| > 1\}$ (incluso il punto all'infinito) è a fase minima. Quindi $L(z)$ è proprio il fattore canonico a fase minima. L'unicità, a meno di una costante moltiplicativa di modulo 1, scende dal procedimento di fattorizzazione e del resto è cosa nota per quanto visto nel Capitolo 4. \square

Come si vede, la funzione di trasferimento inversa di $L(z)$

$$\Gamma(z) := 1/L(z) \quad (2.4)$$

è per costruzione ancora analitica, cioè priva di poli in $\{|z| > 1\}$ (incluso il punto all'infinito), però non più in generale strettamente stabile, dato che, se $S(e^{j\omega})$ si annulla in qualche punto della circonferenza unità possono esservi poli di $\Gamma(z)$ a modulo uguale a uno. Si può comunque espandere Γ in una serie di Laurent convergente in $\{|z| > 1\}$ e definire direttamente la risposta impulsiva del filtro sbiancante, g , come la successione dei coefficienti dello sviluppo (l'antitrasformata Z usuale della teoria dei segnali e dei controlli). Quindi *la risposta impulsiva g del filtro sbiancante è un segnale causale*, anche se non a energia finita in generale. Ricordiamo comunque che nonostante la possibile instabilità di $\Gamma(z)$, la somma

$$\varepsilon(t) = \sum_{k=-\infty}^t g(t-k) \mathbf{y}(k) \quad (2.5)$$

converge egualmente in media quadratica e definisce un processo a varianza finita $\{\varepsilon(t)\}$ che è bianco ortonormale (si confronti la discussione a proposito della formula (4.44) al Capitolo 4).

Da questa osservazione e dal fatto che $L(z)$ è una funzione di H^2 (per cui l'antitrasformata $\ell(\cdot)$ è un segnale a energia finita nullo per $t < 0$) seguono allora le inclusioni

$$\varepsilon(t) \in \mathbf{H}_t(\mathbf{y}), \quad \mathbf{y}(t) \in \mathbf{H}_t(\varepsilon) \quad \forall t \quad .$$

Dalla prima di queste relazioni scende che $\varepsilon(s) \in \mathbf{H}_s(\mathbf{y})$, $\forall s \leq t$, per cui $\mathbf{H}_t(\varepsilon) \subseteq \mathbf{H}_t(\mathbf{y})$. Dalla seconda, analogamente, $\mathbf{H}_t(\mathbf{y}) \subseteq \mathbf{H}_t(\varepsilon)$ e quindi

$$\mathbf{H}_t(\varepsilon) = \mathbf{H}_t(\mathbf{y}) \quad , \quad \forall t \quad ,$$

il che prova, anche in questo caso particolare, che

PROPOSIZIONE 4.2. *Il processo $\{\varepsilon(t)\}$ definito dalla (2.5) è l'innovazione normalizzata di $\{\mathbf{y}(t)\}$.*

Sull'interpretazione sistemistica del filtro sbiancante. Nel caso razionale le operazioni di sbiancamento e di ricostruzione del processo tramite il filtro formante si possono generalmente interpretare come vere e proprie operazioni di convoluzione *sulle singole traiettorie dei processi* in gioco (e non solamente come somme convergenti in media quadratica).

PROPOSIZIONE 4.3. *Se la densità spettrale S è priva di zeri sulla circonferenza unità, $\{|z| = 1\}$, la somma di convoluzione che definisce $\{\varepsilon(t)\}$ come uscita del sistema lineare di funzione di trasferimento $\Gamma(z)$ con ingresso $\{\mathbf{y}(t)\}$,*

$$\varepsilon(t) = \sum_{k=-\infty}^t g(t-k) \mathbf{y}(k) \quad , \quad (2.6)$$

è convergente (in media quadratica e) con probabilità uno e definisce un processo stazionario $\{\varepsilon(t)\}$ bianco ortonormale. Similmente, le traiettorie di $\{\mathbf{y}(t)\}$ si ottengono come l'uscita del sistema lineare di funzione di trasferimento $L(z)$ corrispondenti alle traiettorie del rumore bianco d'ingresso $\{\varepsilon(t)\}$ definito in (2.5), ovvero

$$\mathbf{y}(t) = \sum_{k=-\infty}^t \ell(t-k) \varepsilon(k) \quad , \quad (2.7)$$

dove la somma è ancora convergente con probabilità uno. Nelle formule (2.5) e (2.7) $g(\cdot)$ ed $\ell(\cdot)$ sono le risposte impulsive corrispondenti, rispettivamente, alle funzioni di trasferimento $\Gamma(z)$ ed $L(z)$.

DIMOSTRAZIONE. La dimostrazione di questo enunciato consiste sostanzialmente nel mostrare che l'espressione (2.6) è una somma convergente con probabilità uno. Premettiamo allo scopo il seguente lemma.

LEMMA 4.1. *Se σ è la funzione di covarianza di un processo p.n.d. a spettro razionale, esiste un numero λ , con $0 < \lambda < 1$ e una costante positiva A^2 , tali che*

$$\sigma(\tau) \leq A^2 \lambda^{|\tau|} \quad \tau \in \mathbb{Z} \quad . \quad (2.8)$$

DIMOSTRAZIONE. La dimostrazione di questo lemma è basata sull'osservazione che per $\tau \geq 0$ la funzione $\sigma(\cdot)$ è una combinazione lineare di esponenziali che tendono a zero per $\tau \rightarrow \infty$. Inoltre $\sigma(-\tau) = \sigma(\tau)$. I dettagli sono lasciati come esercizio per il lettore. \square

Per ogni t fissato, la varianza della ridotta n -esima

$$\varepsilon_n(t) := \sum_{k=0}^n g(k) \mathbf{y}(t-k)$$

della serie (2.6) soddisfa la disuguaglianza

$$E[\varepsilon_n(t)]^2 = \sum_{hk=0}^n g(h)\sigma(h-k)g(k) \leq A^2 \left[\sum_{h=0}^n g(h)^2 \right] \left[\sum_{h=0}^n \lambda^{2h} \right]$$

dove l'ultimo passaggio segue dal lemma 4.1 e dalla disuguaglianza di Schwartz. Nello stesso modo si ottiene ancora

$$E[\varepsilon_n(t) - \varepsilon_m(t)]^2 = E \left[\sum_{k=m}^n g(k) \mathbf{y}(t-k) \right]^2 \leq A^2 \sum_{k=m}^n \lambda^{2k} \sum_{k=m}^n g(k)^2.$$

Ora, se $S(e^{j\omega}) \neq 0$, per tutti gli $\omega \in [-\pi, \pi]$, la funzione di trasferimento $\Gamma(z)$ è strettamente stabile e tutti i modi della corrispondente risposta impulsiva $g(\cdot)$ decadono a zero esponenzialmente per cui $\sum_{k=m}^n g(k)^2 \leq \|g\|_{\ell^2}^2 < \infty$. Usando la diseuguaglianza di Chebichev, si trova

$$P\left\{|\varepsilon_n(t) - \varepsilon_m(t)| \geq c\right\} \leq \frac{A^2 \|g\|_{\ell^2}^2}{c^2} \lambda^{2m} \sum_{k=0}^{n-m} \lambda^{2k}$$

per cui

$$\sum_{m=0}^{+\infty} \sup_{n \geq m} P\left\{|\varepsilon_n(t) - \varepsilon_m(t)| \geq c\right\} < \infty \quad .$$

Questo implica, in base a noti risultati di teoria della probabilità, che la successione $\{\varepsilon_n(t)\}$ converge con probabilità uno, uniformemente in t . (Vedere ad esempio il testo di Wong [156, Proposizione 6.4, p. 22]). La stessa dimostrazione si può adattare per mostrare che (2.7) è convergente con probabilità uno. \square

Da osservare che questo risultato dipende in modo cruciale dalla razionalità dello spettro, ma in un certo senso è anche molto generale, perchè non dipende dalla distribuzione di probabilità di \mathbf{y} .

La possibile instabilità di $\Gamma(z)$ è un fatto indesiderabile: in pratica una modellizzazione approssimata o qualunque perturbazione non voluta sovrapposta a $\{\mathbf{y}(t)\}$ possono eccitare i modi oscillatori del filtro e causare divergenza della risposta. In effetti la presenza di poli in $\{|z| = 1\}$ di $\Gamma(z)$ è una conseguenza ovvia dell'annullarsi di $S(e^{j\omega})$. Il rumore bianco $\{\varepsilon(t)\}$ ha uno spettro piatto che contiene tutte le frequenze. Se $S(e^{\pm j\omega_0}) = 0$ significa che al processo $\{\mathbf{y}(t)\}$ “manca la componente armonica di frequenza ω_0 ”. È chiaro allora che per ottenere uno spettro piatto a partire da $\{\mathbf{y}(t)\}$, il filtro $\Gamma(z)$ dovrà introdurre le frequenze mancanti, cioè appunto avere poli in $e^{\pm j\omega_0}$.

In pratica, se si può pensare sovrapposta a $\{\mathbf{y}(t)\}$ una piccola componente di rumore bianco scorrelato, $\{\mathbf{n}(t)\}$ di varianza σ_n^2 ,

$$\mathbf{z}(t) := \mathbf{y}(t) + \mathbf{n}(t)$$

lo spettro di $\{\mathbf{z}(t)\}$,

$$S_z(z) = S_y(z) + \sigma_n^2 \quad (2.9)$$

diventa strettamente positivo sulla circonferenza unità e si evitano le degenerazioni prima accennate.

Modelli ARMA. Il fattore spettrale canonico si può scrivere come rapporto di polinomi in z ,

$$L(z) = \frac{C(z)}{A(z)} = \frac{c_0 z^n + c_1 z^{n-1} + \dots + c_n}{z^n + a_1 z^{n-1} + \dots + a_n} \quad , \quad (2.10)$$

con $A(z)$ e $C(z)$ polinomi ² relativamente primi di grado n in z . L'indeterminazione nei coefficienti è stata eliminata prendendo $a_0 = 1$. Possiamo moltiplicare il

²Riserveremo la lettera B che potrebbe sembrare piu' naturale da usare al numeratore, per i modelli ARMA con ingressi che introdurremo al Capitolo 8.

numeratore e il denominatore nella (2.10) per z^{-n} , ottenendo una rappresentazione della funzione razionale $L(z)$ come rapporto di polinomi nell'indeterminata z^{-1} :

$$L(z) = \frac{c_0 + c_1 z^{-1} + \dots + c_n z^{-n}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} := \frac{C^*(z^{-1})}{A^*(z^{-1})} \quad , \quad (2.11)$$

dove l'asterisco denota il *polinomio reciproco*: $P^*(\lambda) := \lambda^n P(\lambda^{-1})$, essendo n il grado di $P(\lambda)$. L'espressione (2.11) sostituita formalmente nella rappresentazione $\mathbf{y}(t) = L(z) \boldsymbol{\varepsilon}(t)$ fornisce l'uguaglianza simbolica $A^*(z^{-1}) \mathbf{y}(t) = C^*(z^{-1}) \boldsymbol{\varepsilon}(t)$, che lega il processo $\{\mathbf{y}(t)\}$ alle sue innovazioni $\{\boldsymbol{\varepsilon}(t)\}$. Se si interpreta la moltiplicazione per z^{-1} come operatore di ritardo unitario, questa uguaglianza si può riscrivere nel dominio del tempo come un'equazione alle differenze di ordine n ,

$$\mathbf{y}(t) + \sum_{i=1}^n a_i \mathbf{y}(t-i) = c_0 \boldsymbol{\varepsilon}(t) + \sum_{i=1}^n c_i \boldsymbol{\varepsilon}(t-i) \quad , \quad t \in \mathbb{Z} \quad . \quad (2.12)$$

Equazioni di questo tipo vengono chiamate nella letteratura statistica *Modelli AR-MA*. L'attributo "ARMA" sta per "Auto-Regressive-Moving-Average". Il modello si dice *puramente autoregressivo* se $c_i = 0$, $i = 1, \dots, n$, e a *pura media mobile* se $a_i = 0$, $i = 1, \dots, n$. Si noti che in ogni caso deve essere $c_0 \neq 0$.

Il modello (2.12) può anche essere scritto in una forma leggermente diversa, usando al posto di $\boldsymbol{\varepsilon}(t)$ l'innovazione *non normalizzata* $\{\mathbf{e}(t)\}$ che è legata alla prima da una relazione di proporzionalità (cfr. la (4.34))

$$\mathbf{e}(t) = c_0 \boldsymbol{\varepsilon}(t) \quad (2.13)$$

I coefficienti del numeratore vengono così normalizzati a

$$\bar{c}_0 = 1 \quad , \quad \bar{c}_i = c_i/c_0 \quad i = 1, \dots, n \quad .$$

La (2.12) viene allora sostituita dall'equazione

$$\mathbf{y}(t) + \sum_1^n a_i \mathbf{y}(t-i) = \mathbf{e}(t) + \sum_1^n \bar{c}_i \mathbf{e}(t-i) \quad , \quad \lambda^2 = \text{var}[e(t)] = c_0^2 \quad , \quad (2.14)$$

che è evidentemente il modello ARMA corrispondente al fattore spettrale a fase minima, $W(z)$, normalizzato all'infinito ($W(\infty) = 1$).

Raccogliamo queste ultime informazioni nell'enunciato seguente.

TEOREMA 4.1. *Ogni processo stazionario puramente non deterministico il cui spettro di potenza sia una funzione razionale di $z = e^{j\omega}$ è rappresentabile come soluzione di un'equazione alle differenze del tipo (2.12) (oppure (2.14)) in cui $\{\boldsymbol{\varepsilon}(t)\}$ è un processo ortonormale (risp. bianco) e il polinomio*

$$A(z) = z^n + a_1 z^{n-1} + \dots + a_n$$

ha tutti gli zeri all'interno del cerchio unitario, $\{|z| < 1\}$, del piano complesso. Il processo ortonormale $\{\boldsymbol{\varepsilon}(t)\}$ è l'innovazione normalizzata di $\{\mathbf{y}(t)\}$ (rispettivamente, $\{\mathbf{e}(t)\}$ è l'innovazione non normalizzata) se e solo se gli zeri del polinomio

$$C(z) = c_0 z^n + c_1 z^{n-1} + \dots + c_n$$

stanno nel cerchio unitario $\{|z| \leq 1\}$. \diamond

ESEMPIO 4.1 (Esempio elementare di fattorizzazione spettrale). Dato che la covarianza di un processo scalare (reale) è una funzione pari: $\sigma(\tau) = \sigma(-\tau)$, lo spettro $S(e^{j\omega})$ di un processo reale (trasformata di Fourier di $\sigma(\tau)$) è sempre funzione di

$\cos \omega$. Dato poi che $(z+a)(z^{-1}+a) = 1+a^2+a(z+z^{-1})$, uno spettro razionale è sempre un rapporto di prodotti di termini del tipo $1+a^2+2a\cos\omega$, $|a| \leq 1$. Il lettore è invitato a tracciare il grafico di queste componenti spettrali elementari per vari valori del parametro reale a .

Sia

$$S(e^{j\omega}) = \frac{1.04 + 0.4 \cos \omega}{1.25 + \cos \omega} .$$

Per calcolarne la fattorizzazione spettrale scriviamo allora $S(z)$ come

$$S(z) = \frac{1.04 + 0.2(z+z^{-1})}{1.25 + 0.5(z+z^{-1})} = \frac{0.2(z+0.2)(z+5)}{0.5(z+0.5)(z+2)} = \frac{z+0.2}{z+0.5} \frac{z^{-1}+0.2}{z^{-1}+0.5}$$

da cui si vede immediatamente che

$$L(z) = \frac{z+0.2}{z+0.5}$$

è il fattore spettrale canonico. Come visto si può così scrivere $\mathbf{y}(t) = L(z)\boldsymbol{\varepsilon}(t)$ e questa relazione simbolica corrisponde ad un'equazione alle differenze che lega $\{\mathbf{y}(t)\}$ all'innovazione $\{\boldsymbol{\varepsilon}(t)\}$ del tipo

$$\mathbf{y}(t+1) + 0.5\mathbf{y}(t) = \boldsymbol{\varepsilon}(t+1) + 0.2\boldsymbol{\varepsilon}(t) .$$

Per mettere in evidenza la causalità del legame ingresso-uscita si preferisce scrivere questa equazione mediante *differenze all'indietro* (modello ARMA)

$$\mathbf{y}(t) + 0.5\mathbf{y}(t-1) = \boldsymbol{\varepsilon}(t) + 0.2\boldsymbol{\varepsilon}(t-1) , \quad t \in \mathbb{Z} . \quad (2.15)$$

Si può guardare a questa relazione pensando $\{\boldsymbol{\varepsilon}(t)\}$ indifferentemente come l'ingresso o l'uscita. Nel primo caso essa esprime $\mathbf{y}(t)$ come funzione lineare delle variabili passate e presente del processo $\{\boldsymbol{\varepsilon}(t)\}$, nel secondo caso si pensa $\{\mathbf{y}(t)\}$ come l'ingresso e la relazione (2.15) definisce il filtro sbiancante. Essa fornisce quindi un algoritmo sequenziale per calcolare i campioni $\{\dots\boldsymbol{\varepsilon}(t-1), \boldsymbol{\varepsilon}(t), \dots\}$ del processo di innovazione corrispondenti ad una specifica traiettoria osservata $\{\dots\mathbf{y}(t-1), \mathbf{y}(t), \dots\}$ del processo $\{\mathbf{y}(t)\}$. \diamond

Qualche volta in letteratura si parla di “processi ARMA” per riferirsi a processi stazionari descrivibili mediante equazioni alle differenze lineari con ingresso rumore bianco, del tipo (2.12) o (2.14). Questa denominazione non significa altro se non che il processo ha spettro razionale.

Equazioni di Yule-Walker. La funzione di covarianza di un processo a spettro razionale si può esprimere come soluzione di un'equazione alle differenze lineare che è sostanzialmente la stessa equazione del modello ARMA. Questa osservazione è importante perchè, come vedremo, sta alla base di alcuni metodi elementari di identificazione. Per verificare questa affermazione si può semplicemente moltiplicare la (2.14) per $\mathbf{y}(s)$ e poi prendere l'aspettazione dei due membri. Un metodo di procedere alternativo è il seguente.

Si nota innanzitutto che la funzione di covarianza incrociata $\sigma_{ye}(t, s)$ è esprimibile mediante la formula

$$\sigma_{ye}(t-s) = \sum_{k=0}^{+\infty} w(k) E\{\mathbf{e}(t-k)\mathbf{e}(s)\} = \sum_{k=0}^{+\infty} w(k) \lambda^2 \delta(t-s-k) \quad t \geq s ;$$

mentre $\sigma_{ye}(t-s) = 0$ per $t < s$. Ne deriva che σ_{ye} è la risposta del sistema lineare di funzione di trasferimento $W(z)$ (il fattore spettrale a fase minima normalizzato) ad

un impulso di Kronecker $\delta(t)$ moltiplicato per la varianza dell'innovazione. Questo implica che $\sigma_{ey}(t)$ ($= \sigma_{ye}(-t)$) è una funzione anticausale la cui Z-trasformata è $\lambda^2 W(1/z)$. Dato che $S_e(z) = \lambda^2$, la relazione di fattorizzazione spettrale si può così riscrivere come

$$S(z) = S_{ye}(z)S_e^{-1}(z)S_{ey}(z) = W(z)S_{ey}(z).$$

Interpretando questa relazione nel dominio del tempo si arriva al seguente risultato.

PROPOSIZIONE 4.4 (Equazioni di Yule-Walker). *La funzione di covarianza σ di un processo ARMA stazionario soddisfa il seguente sistema di equazioni alle differenze di ordine $2n$*

$$\sigma(t) + \sum_{i=1}^n a_i \sigma(t-i) = \sigma_{ey}(t) + \sum_{i=1}^n \bar{c}_i \sigma_{ey}(t-i) \quad (2.16)$$

$$\sigma_{ey}(t) = \sigma_{ye}(-t) \quad (2.17)$$

$$\sigma_{ye}(t) + \sum_{i=1}^n a_i \sigma_{ye}(t-i) = \lambda^2 \left[\delta(t) + \sum_{i=1}^n \bar{c}_i \delta(t-i) \right] \quad (2.18)$$

Se la funzione di trasferimento $W(z)$ è di tipo AR la (2.16) si semplifica nella

$$\sigma(t) + \sum_{i=1}^n a_i \sigma(t-i) = \sigma_{ey}(t)$$

Scrivendo questa equazione per $t = 0, 1, \dots, n$ in forma matriciale e ricordando che per processi scalari si ha $\sigma(-k) = \sigma(k)$, si trova

$$\begin{bmatrix} 1 & a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} \sigma(0) & \sigma(1) & \dots & \sigma(n) \\ \sigma(1) & \sigma(0) & \dots & \sigma(n-1) \\ \vdots & & \ddots & \vdots \\ \sigma(n) & \sigma(n-1) & \dots & \sigma(0) \end{bmatrix} = \begin{bmatrix} \lambda^2 & 0 & 0 & \dots & 0 \end{bmatrix} \quad (2.19)$$

Più in generale, dato che per $t > 0$ la covarianza $\sigma_{ey}(t)$ a secondo membro della (2.16) è nulla, la funzione σ soddisfa, per $t \geq 0$, l'equazione alle differenze di ordine n

$$\sigma(t+n) + \sum_{i=1}^n a_i \sigma(t+n-i) = \lambda^2 \delta(t) \quad t \geq 0 \quad (2.20)$$

La formula (2.20) e l'analogha relazione che si può scrivere partendo dalla (2.20) possono essere interpretate come relazioni che esprimono in forma implicita i parametri $\{a_k\}$ di $W(z)$ in funzione della covarianza σ del processo \mathbf{y} e stanno alla base di alcuni metodi elementari di identificazione, si veda ad esempio [132].

Naturalmente la (2.16) o la (2.20) possono invece essere interpretate come equazioni nell'incognita σ . Notare che per risolvere la (2.20) (che può sembrar più semplice) servono n condizioni iniziali che non sono date. Queste possono ovviamente essere trovate risolvendo la (2.16) per tempi negativi.

Il risultato ha una versione matriciale sostanzialmente analoga la cui formulazione verrà lasciata per esercizio al lettore.

3. Fattorizzazione Spettrale di Matrici Razionali e Modelli ARMA Vettoriali

Nel caso di processi vettoriali il calcolo della rappresentazione d'innovazione richiede la fattorizzazione di una *matrice* densità spettrale $S(z)$ di dimensione $m \times m$. Noi qui supporremo che gli elementi di $S(z)$ siano tutti *funzioni razionali* di z , fatto che per brevità si esprime anche dicendo che $S(z)$ è una *funzione razionale* di z . Ricordiamo che le tre proprietà fondamentali di uno spettro scalare nel caso matriciale si generalizzano a

- 1') $S(z) = S(z^{-1})'$
- 2') $S(e^{j\omega}) \geq 0$, $\omega \in [-\pi, \pi]$
- 3') $\int_{-\pi}^{\pi} S_{kh}(e^{j\omega}) \frac{d\omega}{2\pi} = E \mathbf{y}_k(t) \mathbf{y}_h(t) = \sigma_{k,h}(0) < \infty$ $k, h = 1, \dots, m$.

La condizione 1') si può esprimere esplicitamente come

$$S_{kh}(z) = S_{hk}(z^{-1}) \quad , \quad h, k = 1, \dots, m \quad . \quad (3.1)$$

La 2') significa che S è una matrice *semidefinita positiva* sulla circonferenza unità. Questa condizione è più informativa se posta in relazione al *rango* di S . Supponiamo che $S(z)$ abbia rango m quasi ovunque per $z \in \mathbb{C}$ (cioè sia invertibile eccettuato al più un numero finito di punti del piano complesso). In questo caso c'è una nozione univoca di *zeri* di una matrice razionale come zeri del suo determinante. La 2') implica allora che $S(e^{j\omega})$ può diventare singolare al più in un numero finito di frequenze nell'intervallo $[-\pi, \pi]$. La 2') può in particolare valere con il segno di disuguaglianza stretta, $S(e^{j\omega}) > 0$, (il che significa che $S(e^{j\omega})$ è *strettamente* definita positiva). In questo caso diremo che lo spettro *non ha zeri sulla circonferenza unità*.

Se il rango di S è invece $p < m$, la 2') non può ovviamente valere mai in senso stretto. In questo caso, l'assenza di zeri sulla circonferenza unità corrisponde alla condizione

$$\text{rango } S(e^{j\omega}) = p \quad , \quad \forall \omega \in [-\pi, \pi] \quad .$$

Per giustificare questa affermazione occorrerebbe una definizione precisa del concetto di *zero* di una matrice razionale su cui non possiamo dilungarci. Si vedano a questo proposito i testi [72, 86]. Notiamo infine che la 3'), analogamente al caso scalare, *esclude la presenza di poli* (degli elementi) di $S(z)$ sulla circonferenza unità.

TEOREMA 4.2. *Ogni matrice razionale di rango p , soddisfacente le condizioni 1'), 2'), 3') ammette fattori spettrali razionali $m \times p$, analitici in $\{|z| \geq 1\}$.*

Questo fondamentale risultato, dovuto a Youla [161], dice in sostanza che tutti i processi a distribuzione spettrale assolutamente continua e *spettro razionale non solo sono ortonormalizzabili, ma anche automaticamente p.n.d.* Le due nozioni introdotte nella Definizione 3.1 in questo caso coincidono. Non è così però per uno spettro non razionale.

L'articolo di Youla [161] presenta anche una procedura per il calcolo del fattore a fase minima, basato sulla riduzione di $S(z)$ alla cosiddetta *forma canonica di Smith-McMillan* (si veda ad esempio [72, pp. 443–451]). Questo “algoritmo” di fattorizzazione spettrale ha oggi solo interesse storico ed è servito a suo tempo soprattutto a dimostrare che nel caso matriciale esistevano procedure con cui i fattori spettrali potevano effettivamente essere calcolati.

Esistono sostanzialmente due gruppi di tecniche di calcolo per risolvere il problema della fattorizzazione spettrale. Il primo fa uso della teoria delle *rappresentazioni polinomiali fratte* (MFD) delle matrici razionali e riduce il problema a quello della fattorizzazione di una matrice *polinomiale*. Il secondo fa uso di modelli di stato e parte dalla conoscenza di una realizzazione minima della matrice razionale $S(z)$. Il problema del calcolo dei fattori spettrali si riduce in questo contesto alla risoluzione di un'equazione algebrica di Riccati. Questo secondo approccio costituisce una naturale appendice alla teoria del filtro di Kalman e verrà studiato nel Cap. 10.

In questo capitolo ci sembra più naturale trattare tecniche basate su descrizioni ingresso-uscita e modelli polinomiali di funzioni di trasferimento. Per fare questo sarà necessario introdurre qualche definizione e un pò di apparato algebrico. Il lettore interessato solo ai risultati principali può passare direttamente alla sezione successiva.

Rappresentazioni polinomiali fratte di matrici razionali*. Converrà generalizzare leggermente la definizione usuale e definire *polinomio (a coefficienti reali) nell'indeterminata z* una qualunque somma *finita* del tipo

$$p(z) := \sum_m^n p_k z^k \quad , \quad p_k \in \mathbb{R} \quad , \quad m, n \in \mathbb{Z} \quad , \quad m \leq n \quad . \quad (3.2)$$

Secondo questa definizione, un polinomio può quindi contenere sia potenze positive che negative dell'indeterminata z (questi oggetti sono qualche volta chiamati “dipolinomi” o “polinomi di Laurent” in letteratura), ad esempio: $p(z) = 3z^2 + z^{-1} + 5z^{-7}$. Denoteremo l'insieme dei polinomi a coefficienti in \mathbb{R} con il simbolo $\mathbb{R}[z]$. In algebra non è necessario dare alcun significato matematico preciso all'indeterminata z ; l'espressione z^k serve solo da “segnaposto” per il relativo coefficiente. Questo ci lascerà piena libertà in seguito di interpretare all'occorrenza z come variabile complessa, come variabile sulla circonferenza unità ($z = e^{j\omega}$), oppure come operatore di anticipo di un passo su successioni (segnali temporali).

Se si definiscono la somma e il prodotto di polinomi secondo le usuali regole di somma e moltiplicazione delle potenze dell'indeterminata z ,

$$p(z) + q(z) = \sum_k (p_k + q_k) z^k, \quad p(z)q(z) = \sum_k \sum_h (p_{k-h} q_h) z^k$$

si verifica facilmente che $\mathbb{R}[z]$ è un *anello commutativo* con unità il polinomio identicamente uguale a 1. I suoi elementi *invertibili*, $u(z)$, tali per cui esiste una soluzione polinomiale $x(z)$ dell'equazione $x(z)u(z) - 1 = 0$, vengono chiamati *unimodulari* e giocano lo stesso ruolo delle “costanti” non nulle nella teoria dei polinomi “ordinari” (a potenze non negative). Gli elementi unimodulari di $\mathbb{R}[z]$ sono tutti i monomi az^k con $a \neq 0$.

Useremo sempre la convenzione che nella rappresentazione (3.2) i termini p_m e p_n siano entrambi diversi da zero. Il *grado*, $\deg(p)$, di $p(z)$ si definisce allora come la differenza $\deg(p) := n - m$. Da notare che con questa definizione i monomi az^k hanno grado zero³.

³Ovviamente, secondo questa definizione, il grado di un polinomio ordinario a potenze non negative di z (che denoteremo all'occorrenza col simbolo $\deg_+(p)$) è in generale un numero diverso. Ad esempio il grado ordinario di z^5 è uguale a 5 mentre per noi $\deg(z^5) = 0$.

In letteratura si usa talvolta denotare i polinomi di Laurent col simbolo $p(z, z^{-1})$, il che sembra, a nostro avviso, una scelta particolarmente infelice, dato che insinua l'idea che un polinomio di Laurent sia la stessa cosa di un polinomio ordinario in due indeterminate z e z^{-1} . Assumendo

In realtà $\mathbb{R}[z]$ è anche un *dominio euclideo*, cosa sostanzialmente esprimibile mediante il seguente enunciato che daremo per noto.

LEMMA 4.2. *Per ogni a, b in $\mathbb{R}[z]$ con $b \neq 0$ esistono polinomi q ed r siffatti che*

- i) $\deg(r) < \deg(b)$ oppure $r = 0$,
- ii) $a = qb + r$.

I polinomi q ed r , detti, rispettivamente, quoziente e resto della divisione di a per b , non sono necessariamente unici.

Da notare che l'unicità del quoziente si ha viceversa nel caso di polinomi “ordinari” in z che formano un cosiddetto dominio euclideo *proprio*. (Informazioni ulteriori su questi concetti si possono trovare nei testi di algebra commutativa. Un'esposizione ridotta all'essenziale si può trovare ad esempio nell'Appendice A e B del Vidyasagar [143]). Se $r \neq 0$, in effetti il procedimento di divisione si può iterare ottenendo $r = q_1 b + r_1$ ancora con $\deg(r_1) < \deg(b)$. Sostituendo in ii), si trova così

$$a = (q + q_1)b + r_1 \quad .$$

Chiaramente ora $r_1 \neq 0$ e il procedimento si può reiterare su r_1 , ecc... Questo *algoritmo euclideo di divisione successiva* permette di ottenere successivamente i termini dello sviluppo di Laurent attorno a $z = \infty$ di una funzione razionale

$$\frac{a(z)}{b(z)} = q_m z^m + \dots + q_0 + q_{-1} z^{-1} + q_{-2} z^{-2} + \dots$$

ed è particolarmente utile in molte questioni connesse al calcolo di z -antitrasformate.

L'insieme delle *funzioni razionali reali* si denota con il simbolo $\mathbb{R}(z)$. Ogni $f \in \mathbb{R}(z)$ si può rappresentare come frazione $f(z) = q(z)/p(z)$ con $q, p \in \mathbb{R}[z]$. Due rappresentazioni si dicono *equivalenti* se p e q differiscono per uno stesso fattore polinomiale. La rappresentazione è *copriva* (o relativamente prima) se gli unici fattori comuni a q e p sono *polinomi unimodulari*. Da notare che con questa definizione

$$1 + \alpha z^{-1}, \quad \frac{z + \alpha}{z}$$

sono, come ci si aspetta debba essere, rappresentazioni coprime della stessa funzione razionale.

Come abbiamo già visto nella (2.11), il fattore comune unimodulare si può scegliere in modo tale da rendere q e p polinomi a potenze solo positive o solo negative di z e da normalizzare un coefficiente (ad esempio nel denominatore) all'unità.

È immediato verificare che $\mathbb{R}(z)$ è un *corpo* (ogni elemento non nullo ha un inverso), quindi l'insieme, $\mathbb{R}^n(z)$, delle n -ple $f(z) = \{f_1(z), \dots, f_n(z)\}$, $f_k(z) \in \mathbb{R}(z)$, che scriveremo per il momento come vettori colonna⁴ è uno *spazio vettoriale* su $\mathbb{R}(z)$. Notare che $\mathbb{R}^n(z)$ ha esattamente dimensione n , una base *canonica* essendo data dalla n -pla di funzioni razionali $\{e_1, \dots, e_n\}$, con $e_k(z)$ il vettore razionale a componenti nulle eccetto la k -sima uguale all'unità moltiplicativa, 1, di $\mathbb{R}(z)$ (la funzione razionale “identicamente uguale a 1”).

$m < 0$, questo oggetto dovrebbe allora naturalmente avere grado uguale a $\nu = \max\{n, -m\}$ che in generale è diverso da quello che si vorrebbe

⁴Per i nostri scopi vedremo poi che sarà più conveniente usare invece vettori riga.

Le applicazioni $\mathbb{R}(z)$ -lineari

$$M(z) : \mathbb{R}^m(z) \rightarrow \mathbb{R}^n(z) \quad ,$$

hanno sempre una rappresentazione come moltiplicazione (a destra, usando la convenzione descritta) per matrici $n \times m$ ad elementi in $\mathbb{R}(z)$. Queste matrici formano esse stesse uno spazio vettoriale che si denota $\mathbb{R}^{n \times m}(z)$. In particolare le matrici polinomiali $n \times m$ formano un sottospazio di $\mathbb{R}^{n \times m}(z)$ che si denota con $\mathbb{R}^{n \times m}[z]$. Si definisce così il *rango* di una matrice razionale $M(z) \in \mathbb{R}^{n \times m}(z)$ come la dimensione dello spazio immagine in $\mathbb{R}^n(z)$. Per non confonderlo con il rango della matrice $M(z) \in \mathbb{C}^{n \times m}$, quando a z si dà il significato di variabile complessa, si usano nei due casi le scritture $\text{rank}_{\mathbb{R}(z)} M(z)$ e $\text{rank}_{\mathbb{C}} M(z)$, rispettivamente. Il primo viene anche chiamato *rango normale* in letteratura. La notazione standard è un po' complicata e noi useremo $\text{rank } M$ al posto di $\text{rank}_{\mathbb{R}(z)} M(z)$, almeno quando non c'è pericolo di confusione. Se $M_1(z)$ ed $M_2(z) \in \mathbb{R}^{n \times m}(z)$, pensate come funzioni di $z \in \mathbb{C}$, coincidono su un insieme aperto del piano complesso, ovvero su un tratto di curva regolare di lunghezza positiva, o più semplicemente su un insieme *numerabile* di punti distinti, esse coincidono su tutto \mathbb{C} . Da questo fatto scende che il $\text{rank}_{\mathbb{C}} M(z)$ è “genericamente” costante al variare di $z \in \mathbb{C}$ eccettuato al più un insieme finito di punti. Questo valore (che è ovviamente $\leq \min(n, m)$) si chiama *rango generico* o *rango quasi ovunque* (q.o.) di $M(z)$ su \mathbb{C} . È abbastanza ovvio che debba essere

$$\text{rank}_{\mathbb{R}(z)} M(z) = \text{rank}_{\mathbb{C}} M(z) \quad \text{q.o.}$$

Ai fini della rappresentazione di funzioni di trasferimento matriciali, gli oggetti più interessanti sono le *matrici polinomiali*. In particolare, le matrici polinomiali *quadrate* $\mathbb{R}^{m \times m}[z]$ formano un anello (non commutativo) con identità, rispetto all'usuale operazione di somma e prodotto fra matrici. Gli elementi invertibili (a inversa polinomiale) di questo anello si chiamano *matrici unimodulari*. È facile convincersi che vale il seguente risultato.

LEMMA 4.3. *Una matrice polinomiale $A(z) \in \mathbb{R}^{m \times m}[z]$ è unimodulare se e solo se il suo determinante, $\det A(z)$, è un polinomio unimodulare. Equivalentemente, $A(z)$ è unimodulare se e solo se*

$$\text{rank}_{\mathbb{C}} A(z) = m, \quad z \neq 0 \quad (3.3)$$

Date due matrici polinomiali $A(z) \in \mathbb{R}^{m \times m}[z]$ e $B(z) \in \mathbb{R}^{m \times p}[z]$ con $\text{rank } A = m$ (cioè $A(z)$ non singolare q.o.), l'espressione

$$M(z) = A(z)^{-1} B(z) \quad (3.4)$$

definisce una matrice razionale in $\mathbb{R}^{m \times p}(z)$. Ogni coppia di matrici polinomiali (A, B) per cui vale la rappresentazione (3.4) si chiama una *rappresentazione polinomiale fratta (sinistra)*, brevemente, una r.p.f. (sinistra), di $M(z)$.

LEMMA 4.4. *Ogni matrice razionale di $\mathbb{R}^{m \times p}(z)$ ammette rappresentazioni polinomiali fratte.*⁵

DIMOSTRAZIONE. In effetti, detto $m(z)$ il minimo comune multiplo dei denominatori degli elementi $\{m_{ij}(z)\}$ di una arbitraria matrice razionale $M(z)$, si può scrivere $M(z) = 1/m(z) N(z)$ con $N(z)$ polinomiale. Detta $D(z) := I_m m(z)$ con I_m l'identità $m \times m$, si ha per $M(z)$ la r.p.f. $M(z) = D(z)^{-1} N(z)$. \square

⁵Nel seguito la dizione “rappresentazione polinomiale fratta” verrà abbreviata a r.p.f.

L'utilità delle r.p.f. risiede nel fatto che se $M(z)$ è la matrice di trasferimento di un sistema lineare, una r.p.f. del tipo (3.4) permette di realizzare immediatamente il legame ingresso-uscita del sistema nella forma di un sistema di equazioni alle differenze, che in notazione simbolica vettoriale ha la forma

$$A(z) \mathbf{y}(t) = B(z) \mathbf{u}(t) \quad , \quad (3.5)$$

nella quale adesso l'indeterminata z va interpretata come operatore di anticipo unitario su sequenze (e.g., $(z\mathbf{y})(t) = \mathbf{y}(t+1)$).

Supponiamo per un momento che anche $B(z)$ sia quadrata $m \times m$. Allora, una misura naturale della "complessità" od "ordine" del modello (3.5) è data dal massimo fra i gradi dei determinanti dei due polinomi a denominatore, $\det A(z)$, e a numeratore, $\det B(z)$. Modelli (che rappresentano la stessa matrice di trasferimento) di "ordine" minimo saranno così descritti da equazioni alle differenze "il meno complicate possibile" e per mezzo di un numero generalmente basso di parametri.

Un primo requisito che sembra ragionevole perchè una r.p.f. sia "a minima complessità" nel senso intuitivo appena discusso, è che non vi siano fattori (matriciali sinistri) polinomiali comuni ad $A(z)$ e $B(z)$. Una matrice $Q(z) \in \mathbb{R}^{m \times m}[z]$ tale che

$$A(z) = Q(z) A_1(z) \quad , \quad B(z) = Q(z) B_1(z) \quad , \quad (3.6)$$

con $A_1(z) \in \mathbb{R}^{m \times m}[z]$ e $B_1(z) \in \mathbb{R}^{m \times p}[z]$, si chiama *divisore polinomiale comune sinistro* della coppia (A, B) . Vorremmo quindi eliminare tutti i divisori sinistri "non banali". A differenza del caso scalare però i divisori non banali risultano essere molti più di quanto ci si sarebbe potuto aspettare. Notiamo allo scopo che se $Q(z)$ è unimodulare, anche $\det Q(z)$ è unimodulare e si ha $\deg[\det A(z)] = \deg[\det A_1(z)]$ (vale anche l'analoga per $B(z)$ se B è quadrata invertibile). Le due r.p.f. corrispondenti a (3.6), ovvero

$$M(z) = A(z)^{-1} B(z) = A_1(z)^{-1} B_1(z)$$

hanno perciò in questo caso la stessa "complessità".

DEFINIZIONE 4.1. *La r.p.f. (3.4) si dice coprime a sinistra se gli unici divisori polinomiali comuni sinistri di $(A(z), B(z))$ sono unimodulari.*

Una r.p.f. *coprime* (a sinistra o a destra) viene anche detta *irriducibile*. Dalla definizione segue che tutte le r.p.f. *irriducibili* di una stessa matrice razionale formano una classe di equivalenza rispetto alla moltiplicazione per matrici unimodulari (che formano un gruppo rispetto all'usuale prodotto tra matrici). La libertà di scelta del fattore unimodulare si può sfruttare per dare ad $A(z)$ e $B(z)$ una struttura speciale. Ad esempio moltiplicando per

$$Q(z) = Q_0 \operatorname{diag}\{z^{k_1}, \dots, z^{k_m}\} \quad , \quad (3.7)$$

dove Q_0 è una matrice costante nonsingolare e (k_1, \dots, k_m) sono degli interi opportuni, si possono trasformare $A(z)$ e $B(z)$ in *polinomi nel solo operatore di ritardo* z^{-1} con il coefficiente del monomio di grado zero, A_0 , del polinomio $A(z)$ normalizzato all'identità. Per fare questo bisogna introdurre la nozione di matrice polinomiale *ridotta*.

Una matrice polinomiale quadrata $m \times m$, $A(z)$, si dice *ridotta*, (o *propria*) *per righe* se la matrice A_h costruita prendendo solo i coefficienti dei monomi di grado ordinario⁶ più elevato in ciascuna *riga* di $A(z)$ è non singolare.

ESEMPIO 4.2 (Sulle matrici ridotte). Alla matrice polinomiale

$$A(z) = \begin{bmatrix} z^2 + 2z + 1 + z^{-1} & z + 1 + z^{-1} \\ z + 2 & 1 \end{bmatrix}$$

corrisponde una matrice dei coefficienti di grado massimo per righe $A_h = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ che è singolare e quindi $A(z)$ non è ridotta per righe. \diamond

Una qualunque matrice polinomiale può ovviamente essere scritta nella forma

$$A(z) = \text{diag}\{z^{n_1}, \dots, z^{n_m}\} A_h + A_1(z) \quad , \quad (3.8)$$

dove $A_1(z)$ è ancora polinomiale ma con i gradi di tutti i monomi di ciascuna riga rispettivamente minori di n_1, \dots, n_m . Questi interi non negativi prendono il nome di *gradi riga* della matrice. Si può verificare che il grado del determinante di $A(z)$ soddisfa la disuguaglianza

$$\text{deg det}(A(z)) \leq \sum_{k=1}^m n_k .$$

Le matrici ridotte per riga sono proprio quelle per cui la disuguaglianza vale con il segno di uguale. Quindi *in una matrice ridotta il grado del determinante è uguale alla somma dei gradi riga*. Si dimostra che *ogni matrice polinomiale non singolare può essere trasformata in una ridotta per righe* per mezzo di premoltiplicazione per una opportuna matrice unimodulare. Si veda ad es. [72, p. 386]. L'idea è di ridurre il più possibile il grado di ciascuna riga di $A(z)$ per mezzo di opportune sequenze di operazioni elementari sulle righe. Queste operazioni elementari consistono in:

- (1) scambio di righe,
- (2) somma di un multiplo polinomiale di una riga ad un'altra riga,
- (3) moltiplicazione di righe per monomi αz^k , con $\alpha \neq 0$.

È facile convincersi che ciascuna di queste operazioni corrisponde a premoltiplicazione per una matrice unimodulare. Illustriamo il procedimento sull'esempio precedente.

ESEMPIO 4.3 (Seguito). Moltiplicando la seconda riga di $A(z)$ per z e sottraendola dalla prima si ottiene

$$\begin{bmatrix} 1 & -z \\ 0 & 1 \end{bmatrix} A(z) = \begin{bmatrix} 1 + z^{-1} & 1 + z^{-1} \\ z + 2 & 1 \end{bmatrix} := \tilde{A}(z)$$

per cui

$$\tilde{A}_h = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad ,$$

che è nonsingolare. Ne segue che $\tilde{A}(z)$ si può scrivere

$$\begin{bmatrix} 1 & 0 \\ 0 & z \end{bmatrix} \tilde{A}_h + \begin{bmatrix} z^{-1} & z^{-1} \\ 2 & 1 \end{bmatrix} .$$

⁶In letteratura questa definizione si riferisce a matrici polinomiali nella sola indeterminata z . Qui il termine "grado" è inteso nel senso ordinario.

che è propria per righe. \diamond

Una volta resa $A(z)$ propria per righe è evidente che prendendo nell'espressione (3.7) $Q_0 = A_h^{-1}$ e $k_1 = -n_1, \dots, k_m = -n_m$ si trasforma $A(z)$ in un polinomio matriciale in z^{-1} col coefficiente di grado zero normalizzato all'identità

$$A(z) = I + A_1 z^{-1} + \dots + A_\nu z^{-\nu} \quad (3.9)$$

che chiameremo *monico* nel seguito.

ESEMPIO 4.4 (Seguito). Prendendo $Q(z) = \tilde{A}_h^{-1} \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}$ e moltiplicando per $\tilde{A}(z)$ si ottiene in effetti

$$Q(z) \tilde{A}(z) = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \tilde{A}(z) = I + \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix} z^{-1} \quad .$$

che è monico. \diamond

Come è noto, una matrice razionale $M(z) \in \mathbb{R}^{m \times n}(z)$ si dice *propria* se non ha poli all'infinito, i.e., se $M_{kj}(\infty) < \infty$, $k = 1 \dots m$, $j = 1 \dots n$; in particolare, $M(z)$ si dice *strettamente propria* se $M_{kj}(\infty) = 0$ per ogni k, j . È immediato verificare che se la

$$M(z) = A(z)^{-1} B(z)$$

è una r.p.f. di una matrice propria in cui il polinomio “denominatore” $A(z)$ è un polinomio monico nell'operatore di ritardo z^{-1} , $B(z)$ deve necessariamente essere ancora un polinomio in z^{-1} , ovvero

$$B(z) = B_0 + B_1 z^{-1} + \dots + B_\nu z^{-\nu} \quad ,$$

in cui $B_0 = M(\infty)$.

Come accennato poco fa, per lavorare con modelli “non ridondanti” è necessario usare r.p.f. *coprime*. Servono allo scopo

- tests per verificare se $A(z)$ e $B(z)$ ammettono divisori sinistri comuni non unimodulari;
- algoritmi per estrarre gli eventuali divisori comuni da $A(z)$ e $B(z)$, e in specie un procedimento per estrarre il *massimo divisore comune* (sinistro) $m \times m$ dei due polinomi.

È impossibile descrivere anche succintamente la batteria di tests e procedure numeriche che sono state sviluppate allo scopo. Rimandiamo per questo ai testi di Kailath [72, pagg. 377 e seguenti], Kucera [86] etc., accontentandoci di illustrare solo un procedimento elementare per estrarre il massimo divisore comune.

LEMMA 4.5. (Algoritmo per estrarre il M.C.D. sinistro di due matrici polinomiali)

Date $A(z)$ e $B(z)$ di dimensioni $m \times m$ ed $m \times p$ con $\det A(z)$ non identicamente nullo (cioè $\text{rank } A = m$) esiste una matrice unimodulare $(m+p) \times (m+p)$, $U(z)$, che annulla identicamente le ultime p colonne di $[A(z) B(z)]$, ovvero

$$[A(z) B(z)] U(z) = [R(z) 0] \quad (3.10)$$

La matrice $R(z) \in \mathbb{R}^{m \times m}[z]$ è non singolare ed è un M.C.D. sinistro di $A(z)$, $B(z)$.

$U(z)$ si può costruire per composizione successiva di matrici elementari che eseguono le seguenti operazioni.

- (1) Scambio di colonne.

- (2) *Somma di un multiplo polinomiale di una colonna ad un'altra colonna.*
 (3) *Moltiplicazione di colonne per monomi αz^k con $\alpha \neq 0$.*

DIMOSTRAZIONE. Moltiplicando a destra (3.10) per l'inversa $Q(z) = U(z)^{-1}$ e partizionando, si ottiene

$$[A(z) \ B(z)] = [R(z) \ 0] \begin{bmatrix} Q_{11}(z) & Q_{12}(z) \\ Q_{21}(z) & Q_{22}(z) \end{bmatrix} = [R(z) \ Q_{11}(z) \ R(z) \ Q_{12}(z)], \quad (3.11)$$

dove Q_{11} e Q_{12} sono polinomiali. Ne segue che R è un divisore sinistro. D'altro canto, dalla (3.10) segue anche che

$$R(z) = A(z) U_{11}(z) + B(z) U_{21}(z)$$

(con ovvio significato dei simboli), per cui ogni divisore sinistro di A e B divide anche R . Per definizione R è il M.C.D. \square

Ovviamente il M.C.D. è unico solo a meno di moltiplicazione a destra per matrici unimodulari. Se $R(z)$ è essa stessa unimodulare, $A(z)$ e $B(z)$ sono coprime.

L'*algoritmo euclideo delle divisioni successive* permette di organizzare il calcolo di $R(z)$ in un numero finito (anche se in genere molto elevato) di operazioni elementari sulla matrice polinomiale $m \times (m+p)$ $P(z) := [A(z) \ B(z)]$. Si procede nel modo seguente.

- (1) Si scambiano le colonne di $P(z)$ finché l'elemento di posto (1,1), $p_{11}(z)$, ha grado minimo nella corrispondente riga.
- (2) Si fanno le divisioni $p_{1j}(z) = q_{1j}(z) p_{11}(z) + r_{1j}(z)$, $j = 2, \dots, m+p$, e si somma alla colonna j -sima la prima moltiplicata per $-q_{1j}(z)$. Il nuovo elemento nella prima riga diventa $r_{1j}(z)$ che ha grado inferiore a $p_{11}(z)$.
- (3) Si ripete il procedimento finché si arriva ad ottenere $p_{1j}(z) = 0$, $j = 2, \dots, m+p$.
- (4) Si passa all'elemento (2,2) della seconda colonna e, lasciando invariata la prima colonna di $P(z)$, si ripete il procedimento dei punti 1, 2, 3 precedenti fino ad ottenere $p_{2j}(z) = 0$, $j = 3, \dots, m+p$, etc...

In questo modo si ottiene un M.C.D. $R(z)$ a struttura triangolare inferiore.

Dall'idea su cui è basata la dimostrazione del lemma 4.5 si ricava un test semplicissimo che serve per verificare se due matrici polinomiali sono coprime.

PROPOSIZIONE 4.5. *Le matrici polinomiali $A(z) \in \mathbb{R}^{m \times m}[z]$ e $B(z) \in \mathbb{R}^{m \times p}[z]$ sono coprime a sinistra se e solo se*

$$\text{rank}_{\mathbb{C}} [A(z) \ B(z)] = m \quad z \neq 0 \quad (3.12)$$

DIMOSTRAZIONE. Chiaramente $A(z)$ e $B(z)$ sono coprime se e solo se $R(z)$ è unimodulare. La dimostrazione scende allora dal criterio di unimodularità del lemma 4.3, applicato alla (3.11). \square

ESEMPIO 4.5 (Un caso degenere). Anche le matrici polinomiali si possono formalmente rappresentare mediante rpf. Ad esempio si può rappresentare $G(z) =$

$$\begin{bmatrix} z^{-2} \\ z^{-1} + z^{-2} \end{bmatrix} \text{ come } (z^2 I)^{-1} \begin{bmatrix} 1 \\ z + 1 \end{bmatrix}. \text{ Applicando l'algoritmo di triangolarizzazione}$$

inferiore ai due fattori polinomiali,

$$\begin{aligned} [A(z) \ B(z)] &= \begin{bmatrix} z^2 & 0 & 1 \\ 0 & z^2 & z+1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & z^2 \\ z+1 & z^2 & 0 \end{bmatrix} \rightarrow \\ &\rightarrow \begin{bmatrix} 1 & 0 & 0 \\ z+1 & z^2 & -z^2(z+1) \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ z+1 & z^2 & 0 \end{bmatrix} . \end{aligned}$$

si trova che il M.C.D. è $\begin{bmatrix} 1 & 0 \\ z+2 & z^2 \end{bmatrix}$, che è unimodulare e quindi $A(z)$ e $B(z)$ sono coprime a sinistra. Questo è in accordo col fatto (di immedita verifica) che il rango di $[A(z) \ B(z)]$ per $z \neq 0$ è sempre uguale a 2. \diamond

Fattorizzazione di spettri polinomiali. Uno *spettro polinomiale* (o di tipo MA) è una matrice densità spettrale *polinomiale* in z . Dalla condizione di simmetria “Hermitiana” ^{1'}, segue facilmente che ogni spettro polinomiale si può scrivere nella forma

$$S(z) = \sum_{-N}^N S_k z^{-k} , \quad (3.13)$$

dove i coefficienti $\{S_k\}$ sono matrici reali che rispettano le regole di simmetria della corrispondente funzione covarianza, ovvero

$$S'_0 = S_0 \quad , \quad S'_k = S_{-k} \quad , \quad k = 1, \dots, N \quad ,$$

e la condizione di *positività* che si può esprimere dicendo che tutte le sottomatrici quadrate principali, $\Sigma_0, \Sigma_1, \Sigma_2, \dots, \Sigma_n$, della matrice infinita

$$\Sigma := \begin{bmatrix} S_0 & S_{-1} & \dots & S_{-N} & 0 & \dots \\ S_1 & S_0 & & & & \ddots \\ \vdots & & \ddots & & & \ddots \\ S_N & & & S_0 & & \\ 0 & \ddots & & & \ddots & \end{bmatrix} \quad (3.14)$$

sono definite positive. La matrice Σ ha la cosiddetta *struttura di Toeplitz* ($\Sigma_{ij} = S_{i-j}$) *a banda*.

LEMMA 4.6. *Ogni densità spettrale razionale $S(z)$ ammette una fattorizzazione del tipo:*⁷

$$S(z) = D(z^{-1})^{-1} M(z) D(z)^{-\top} , \quad (3.15)$$

dove $D(z)$ è una matrice polinomiale nonsingolare ed $M(z)$ è uno spettro polinomiale delle stesse dimensioni di $S(z)$. In particolare, esistono fattorizzazioni in cui $D(z^{-1})$ è un polinomio monico ($D_0 = I$) nella sola indeterminata z^{-1} con $\det D(\lambda)$ privo di zeri in $\{|\lambda| \geq 1\}$.

⁷ $A^{-\top}$ indica l'inversa della trasposta di A .

DIMOSTRAZIONE. Sia $m(z)$ il minimo comune multiplo dei denominatori, $m_{ij}(z)$, degli elementi di $S(z)$. Supponiamo, senza perdita di generalità, che le rappresentazioni polinomiali fratte $s_{ij} = n_{ij}/m_{ij}$ di questi elementi siano tutte prive di fattori comuni. Si ha allora

$$\begin{aligned} m(z) &= \text{m.c.m.} \{ m_{ij}(z) ; i, j = 1 \dots m \} \\ &= \text{m.c.m.} \{ m_{ji}(z^{-1}) ; i, j = 1 \dots m \} = m(z^{-1}) \quad , \end{aligned}$$

per cui $m(z)$ è un polinomio simmetrico, univocamente determinato a meno di una costante moltiplicativa. Per la condizione 3') $m(z)$ non può avere zeri sulla circonferenza unità $\{z = e^{j\omega}\}$ e quindi $m(e^{j\omega}) = \overline{m(e^{j\omega})}$ è reale per tutti gli $\omega \in [-\pi, \pi]$ e sempre dello stesso segno. Scegliendo opportunamente il segno della costante moltiplicativa si ottiene così $m(e^{j\omega}) > 0$ su tutta la circonferenza unità. Segue che $m(z)$ si può fattorizzare nella forma $m(z) = d(z) d(1/z) = d^*(z^{-1}) d^*(z)$ con $d(z) = \alpha(z^n + d_1 z^{n-1} + d_n)$ polinomio reale con tutti gli zeri in $\{|z| < 1\}$. Normalizzando opportunamente $m(z)$ si può prendere $\alpha = 1$. A questo punto è ovvio che, ponendo

$$D(z^{-1}) := I_m d^*(z^{-1}) \quad , \quad D(z) = I_m d^*(z) \quad ,$$

la matrice $M(z)$ definita dalla

$$D(z^{-1}) S(z) D(z)' := M(z)$$

è uno spettro polinomiale ed $S(z)$ è quindi fattorizzabile nella forma descritta nell'enunciato. \square

Una volta nota una rappresentazione polinomiale del tipo (3.15), la fattorizzazione spettrale di $S(z)$ si riduce a quella dello *spettro polinomiale* $M(z)$. Per matrici densità spettrale polinomiale d'altra parte esistono algoritmi efficienti che permettono di calcolare direttamente il *fattore polinomiale a fase minima*, privo di zeri in $\{|z| > 1\}$. Questi algoritmi sono essenzialmente di due tipi:

- (1) algoritmi basati sulla fattorizzazione di Cholesky della matrice di covarianza Σ , [14] [121] o su tecniche analoghe per rappresentazioni di stato equivalenti [5];
- (2) algoritmi iterativi di tipo Newton [154] [69].

La prima categoria è basata sul legame che esiste tra fattori spettrali analitici e fattorizzazioni triangolari inferiori della matrice infinita Σ discusso alla fine del secondo capitolo.

TEOREMA 4.3 (Bauer-Rissanen). *La matrice di Toeplitz a banda, Σ , associata ad uno spettro polinomiale $S(z)$, ammette una fattorizzazione di Cholesky a blocchi*

$$\Sigma = LL' \quad ,$$

dove L è una matrice a struttura triangolare inferiore a banda,

$$L = \begin{bmatrix} L_{00} & & 0 & & 0 & & 0 \\ \vdots & \ddots & & & & & \\ L_{N0} & & L_{NN} & & & & \\ 0 & \ddots & & \ddots & & & \\ \vdots & & L_{k,k-N} & \dots & L_{kk} & & \\ \dots & & \dots & & & & \ddots \end{bmatrix} \quad (3.16)$$

in cui gli elementi della riga k -sima convergono per $k \rightarrow \infty$

$$\lim_{k \rightarrow \infty} L_{k,k-r} = \bar{L}_r \quad . \quad (3.17)$$

Il polinomio matriciale $\bar{L}(z^{-1})$ costruito con i coefficienti limite \bar{L}_r

$$\bar{L}(z^{-1}) := \bar{L}_0 + \bar{L}_1 z^{-1} + \dots + \bar{L}_N z^{-N} \quad (3.18)$$

è un fattore spettrale a fase minima di $S(z)$.

La dimostrazione di questo enunciato è basata sulla cosiddetta “decomposizione di Wold” e verrà data nel prossimo capitolo.

Il calcolo dei coefficienti L_{kj} , $j = k - N, \dots, k$ della riga k -sima può essere fatto mediante una facile generalizzazione della fattorizzazione di Cholesky al caso di matrici positive e simmetriche a blocchi. Moltiplicando il blocco riga k -simo di L per il blocco colonna j -simo di L' e imponendo l'uguaglianza con il blocco di posto (k, j) di Σ (che è S_{k-j} se $|k - j| \leq N$ e zero altrimenti), si ottiene il seguente algoritmo.

Algoritmo di Cholesky a blocchi. L'algoritmo lavora sui blocchi della banda inferiore di Σ , dall'alto al basso e da sinistra verso destra.

- *Inizializzazione* (fattorizzazione di ggS_0)

$$L_{00} L'_{00} = S_0 \quad .$$

- Per $j < k$ (ovvero per $j = k - N, k - N + 1, \dots, k - 1$, se k è abbastanza grande)

$$L_{kj} = \left(S_{k-j} - \sum_{r=k-N}^{j-1} L_{kr} L'_{jr} \right) L_{jj}^{-\top} \quad . \quad (3.19)$$

- I termini diagonali L_{kk} si calcolano come fattori di Cholesky (triangolari inferiormente a elementi positivi sulla diagonale) nella fattorizzazione

$$L_{kk} L'_{kk} = S_0 - \sum_{r=k-N}^{k-1} L_{kr} L'_{kr} \quad . \quad (3.20)$$

Generalizzazione a spettri di rango $p < m$. Così come è scritto, l'algoritmo richiede che Σ sia definita positiva, intendendo con questo che devono essere definite positive tutte le sottomatrici quadrate principali $\Sigma_0, \Sigma_1, \dots$ di Σ . Se Σ_k diventa singolare per qualche k , il secondo membro in (3.20) cessa di essere (strettamente) definito positivo e la fattorizzazione di Cholesky non si può più eseguire. La condizione di positività stretta, che come abbiamo già visto corrisponde a rango $S(z) = m$, può però essere rilassata. Se rango $S(z) = p < m$, basta prendere

i blocchi L_{kj} di dimensione $m \times p$ (anziché quadrati) e per il calcolo degli L_{kk} sostituire alle fattorizzazioni di Cholesky delle arbitrarie fattorizzazioni “a rango pieno” $m \times p$ della matrice a secondo membro in (3.20), che ha (quindi) rango p . \diamond

L’algoritmo di Bauer-Rissanen richiede un numero di operazioni per ciclo (calcolo di un nuovo blocco riga di L) proporzionale a $m^3 N^2$ e può diventare molto inefficiente in problemi di dimensione elevata. Inoltre, se ci sono zeri di $S(z)$ sulla circonferenza unità (o molto prossimi alla circonferenza unità) la convergenza diventa molto lenta (cfr. l’esempio scalare 2.5 alla fine del Cap. 2). Per questo motivo è consigliabile da un lato cercare di lavorare con rappresentazioni irriducibili dello spettro (N piccolo) e dall’altro usare tecniche più veloci nei casi critici. La *fattorizzazione veloce* di matrici di Toeplitz verrà trattata al Cap. 8 (Algoritmo di Levinson). I *metodi di tipo Newton* hanno convergenza quadratica e possono essere una valida alternativa alla fattorizzazione di Cholesky. La loro descrizione però è un po’ complicata e necessita l’introduzione di argomenti specialistici sulla soluzione di equazioni polinomiali. Per questo motivo rimandiamo il lettore agli articoli originali [154, 69].

Modelli ARMA vettoriali. È evidente che se si dispone di una rappresentazione fratta del tipo (3.15) di uno spettro razionale $S(z)$, il calcolo del fattore a fase minima $\bar{L}(z)$ del polinomio numeratore $M(z)$ di $S(z)$, effettuabile ad esempio mediante l’algoritmo di Bauer-Rissanen descritto più sopra, fornisce direttamente una r.p.f. del fattore spettrale canonico,

$$L(z) = D(z^{-1})^{-1} \bar{L}(z^{-1}) \quad .$$

in cui i polinomi matriciali D e \bar{L} sono polinomi nell’ indeterminata z^{-1} . Se lo spettro è a rango pieno, $\bar{L}(z^{-1})$ è quadrata $m \times m$ con \bar{L}_0 triangolare inferiore ad elementi positivi sulla diagonale. Normalizzando il coefficiente di grado zero all’identità, ovvero definendo

$$N(z^{-1}) := \bar{L}(z^{-1}) \bar{L}_0^{-1} \quad ,$$

si ottiene così una r.p.f. del fattore a fase minima normalizzato $W(z) = D(z^{-1})^{-1} N(z^{-1})$ che fornisce il modello ARMA di innovazione

$$D(z^{-1})\mathbf{y}(t) = N(z^{-1})\mathbf{e}(t) \quad .$$

Da notare che

$$\Lambda := \bar{L}_0 \bar{L}_0'$$

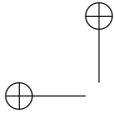
è la varianza dell’innovazione (non normalizzata) $\mathbf{e}(t)$ di $\{\mathbf{y}(t)\}$.

Esercizi

4-1 Talvolta si conviene di dire che un processo stazionario ha “densità spettrale” anche se la distribuzione relativa ha salti. Si parla in questi casi di “righe spettrali” (impulsi di Dirac) nella densità. Sia φ una variabile aleatoria a distribuzione uniforme su $[-\pi, \pi]$ e A una costante deterministica. Mostrare che il processo

$$\mathbf{x}(t) = A \sin(\omega_0 t + \varphi) \quad , \quad t \in \mathbb{Z} \quad ,$$

è (debolmente) stazionario e calcolarne la funzione di covarianza $\sigma_x(\tau)$ e lo spettro. Mostrare che un processo di questo tipo ha una “densità spettrale” che consiste di due righe spettrali. Provare a eseguire la fattorizzazione spettrale di questa densità.



4-2 Calcolare la densità spettrale S_0 corrispondente alla funzione di covarianza $\sigma_0(\tau) = A^2 \lambda^{-|\tau|}$, $|\lambda| < 1$ e trovare il relativo modello ARMA di innovazione.

4-3 Partendo dall'osservazione che una densità spettrale razionale $S(e^{j\omega})$ di un processo reale, dev'essere una funzione limitata, mostrare che esistono sempre costanti A^2 e $0 < \lambda < 1$ tali che

$$S(e^{j\omega}) \leq \frac{A^2}{1 + \lambda^2 - 2\lambda \cos \omega} \quad \omega \in [-\pi, \pi].$$

4-3 Detta $\Sigma_t := \mathbb{E} \mathbf{y}^t (\mathbf{y}^t)'$ la $t + 1$ -sima matrice di Toeplitz di covarianza di un processo \mathbf{y} di densità spettrale S , mostrare che la disuguaglianza $S(e^{j\omega}) \leq S_0(e^{j\omega})$, $\omega \in [-\pi, \pi]$, è equivalente alla

$$\Sigma_t \leq \Sigma_t^0 \quad \text{per } t \geq 0$$

dove la disuguaglianza è intesa nel senso dell'ordinamento semidefinito positivo. (Mostrare che la disuguaglianza tra gli spettri è equivalente alla disuguaglianza

$$\sum_{t,s} f(t) \sigma(t-s) f(s) \leq \sum_{t,s} f(t) \sigma_0(t-s) f(s)$$

dove f è una qualunque successione a supporto compatto).

4-4 Verificare se la r.p.f.

$$\frac{1}{(z+1)^2(z+2)} \begin{bmatrix} z(z+1)(z+2) & z^2(z+1) \\ (z+1)^2 & z^2(z+2) \end{bmatrix}$$

è coprime. Calcolare il M.C.D. sinistro dei polinomi a denominatore e numeratore.

4-5 Scrivere la matrice razionale

$$\begin{bmatrix} \frac{z}{z+1} & \frac{z^2}{(z+1)(z+2)} \\ \frac{1}{z+2} & \frac{z^2}{(z+1)^2} \end{bmatrix}$$

come r.p.f. coprime e monica nell'indeterminata z^{-1} .

4-6 Verificare usando il criterio del rango che la r.p.f

$$W(z) = \frac{1}{z+0.8} \begin{bmatrix} z-0.7 & 1 \\ 1.95 & z-0.5 \end{bmatrix} = \begin{bmatrix} z+0.8 & 0 \\ 0 & z+0.8 \end{bmatrix}^{-1} \begin{bmatrix} z-0.7 & 1 \\ 1.95 & z-0.5 \end{bmatrix},$$

non è coprime. Mostrare che il numeratore ha due zeri uno dei quali è in $z = 0.8$. Calcolare il M.C.D. sinistro dei due polinomi matriciali e verificare che è un multiplo del polinomio

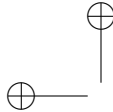
$$\begin{bmatrix} 1 & 0 \\ z-0.5 & z+0.8 \end{bmatrix}$$

Ricavare infine la rappresentazione coprime

$$W(z) = \begin{bmatrix} z+0.8 & 0 \\ -(z-0.5) & 1 \end{bmatrix}^{-1} \begin{bmatrix} z-0.7 & 1 \\ -(z-2) & 0 \end{bmatrix},$$

da cui si vede che $W(z)$ ha un solo zero in $z = 2$.

4-7 Scrivere l'analogia dell'equazione matriciale di Yule-Walker (2.19), partendo dalla relazione generale (2.20). Notare che nella matrice (di Toeplitz) del corrispondente sistema compaiono ora le covarianze $\{\sigma(n), \sigma(n+1), \dots\}$ relative a tempi $\geq n$.



CAPITOLO 5

STRUTTURA DEI PROCESSI STAZIONARI

Come abbiamo avuto modo di notare, esistono processi stazionari le cui traiettorie sono funzioni deterministiche e assolutamente “predicibili” del tempo e viceversa altri, tipicamente i processi ortonormalizzabili e i processi p.n.d., le cui traiettorie sono molto più “irregolari” e corrispondono molto più dei primi alla nozione intuitiva di “imprevedibilità” connessa ai fenomeni aleatori. Come si è visto, tutti i processi p.n.d. si possono costruire, o, se si preferisce “simulare”, inviando un rumore bianco, cioè una successione completamente imprevedibile di campioni aleatori, in ingresso ad un filtro lineare stabile e tempo-invariante.

In questo capitolo caratterizzeremo questi due tipi di “comportamento dinamico” e classificheremo i processi stazionari del secondo ordine in due classi. Per comprendere l’essenza di questa diversità di comportamento è purtroppo necessaria una trattazione matematica un pò astratta e occorre far riferimento a nozioni che in generale non fanno parte del bagaglio culturale di uno studente di ingegneria. Il lettore interessato solo ai risultati principali può tranquillamente leggersi le definizioni nella prossima sezione, passare immediatamente al Teorema 5.2 e saltare il resto di questo capitolo per passare ai successivi, senza sostanziale perdita di continuità.

1. Processi Puramente Deterministici e Puramente non Deterministici

Per studiare la struttura di un processo stazionario in questo capitolo ci rifaremo soprattutto all’ *operatore di traslazione temporale* del processo, introdotto brevemente nel corollario 3.1 e in Appendice B. Ricordiamo che l’operatore di traslazione di un passo, U , è definito inizialmente sulle variabili del processo $\{\mathbf{y}(t)\}$ dalla posizione, $U : \mathbf{y}_k(t) \rightarrow \mathbf{y}_k(t+1)$, $k = 1 \dots m$, $t \in \mathbb{Z}$. Si può poi estendere U per linearità allo spazio vettoriale lineare $\tilde{\mathbf{H}}(\mathbf{y})$ generata dal processo ponendo

$$U \sum a_{k_i} \mathbf{y}_k(t_i) := \sum a_{k_i} \mathbf{y}_k(t_i + 1) \quad (1.1)$$

per ogni combinazione lineare finita di componenti scalari $\{\mathbf{y}_k(t_i)\}$ del processo. Dalla stazionarietà (del secondo ordine) scende che U è un operatore *isometrico* ovvero

$$\|U\mathbf{z}\| = \|\mathbf{z}\| \quad \forall \mathbf{z} \in \tilde{\mathbf{H}}(\mathbf{y}) \quad ,$$

come si può facilmente verificare (basta in effetti ricordare che $\langle \mathbf{y}_k(t_i), \mathbf{y}_h(t_j) \rangle = \sigma_{kh}(t_i - t_j) = \langle \mathbf{y}_k(t_i + 1), \mathbf{y}_h(t_j + 1) \rangle$).

L’operatore di traslazione può poi essere esteso per continuità¹ a tutto $\mathbf{H}(\mathbf{y})$ preservandone il carattere di isometria (si veda il teorema B.1). In effetti, è facile convincersi che l’operatore esteso a $\mathbf{H}(\mathbf{y})$ è suriettivo e quindi *unitario*. Questo implica che U^{-1} esiste e coincide con l’operatore “aggiunto”, U^* . Quest’ultimo,

¹Dato che $\|U(\mathbf{x}_n - \mathbf{x})\| = \|\mathbf{x}_n - \mathbf{x}\|$, U è in effetti uniformemente continuo.

come è noto, è definito dalla relazione $\langle U\mathbf{x}, \mathbf{z} \rangle = \langle \mathbf{x}, U^*\mathbf{z} \rangle$, $\forall \mathbf{x}, \mathbf{z} \in \mathbf{H}(\mathbf{y})$. Si possono così descrivere le traslazioni temporali all'indietro di durata t arbitraria, tramite la formula

$$U_{-t} := U^{-t} = \underbrace{U^{-1} \dots U^{-1}}_{t \text{ volte}} = (U^*)^t \quad , \quad t \geq 0 \quad . \quad (1.2)$$

Riassumendo, si può caratterizzare la famiglia $\{U_t\}_{t \in \mathbb{Z}}$ definita dalle relazioni (1.1), (1.2) dicendo che è *un gruppo ad un parametro di operatori unitari che applicano $\mathbf{H}(\mathbf{y})$ su sè stesso*.

Qualche volta, per brevità, il gruppo delle traslazioni temporali del processo, $\{U_t\}$, verrà chiamato lo *shift temporale* di $\{\mathbf{y}(t)\}$.

Se si ritorna per un attimo alla definizione di $\mathbf{H}_t(\mathbf{y})$ si vede immediatamente che per un processo stazionario si può scrivere

$$\mathbf{H}_{t+s}(\mathbf{y}) = U_s \mathbf{H}_t(\mathbf{y}) \quad \forall t, s \in \mathbb{Z} \quad . \quad (1.3)$$

A questo punto, ricordiamo che lo stimatore lineare a minima varianza del vettore casuale $\mathbf{y}(t+1)$ sulla base delle osservazioni passate e presenti

$$\mathbf{y}^t := \{\mathbf{y}(s), s \leq t\},$$

a suo tempo indicata con i simboli $\hat{\mathbf{y}}(t+1 | t)$, $\hat{E}[\mathbf{y}(t+1) | \mathbf{y}^t]$, e $\hat{E}[\mathbf{y}(t+1) | \mathbf{H}_t(\mathbf{y})]$, si può interpretare come (il vettore delle) proiezione(i) ortogonale(i) di $\mathbf{y}(t+1)$ su $\mathbf{H}_t(\mathbf{y})$. Ricordiamo ancora che la varianza scalare dell'errore di predizione di u n passo, λ^2 , è definita dalla formula

$$\lambda^2 = E \sum_{k=1}^m [\mathbf{y}_k(t+1) - \hat{\mathbf{y}}_k(t+1 | t)]^2 := \|\mathbf{y}(t+1) - \hat{\mathbf{y}}(t+1 | t)\|^2. \quad (1.4)$$

Dimostriamo ora formalmente una proprietà di invarianza temporale della proiezione ortogonale che intuitivamente è pressochè ovvia.

LEMMA 5.1. *Sia $\mathbf{z} \in \mathbf{H}(\mathbf{y})$, \mathbf{K} un qualunque sottospazio di $\mathbf{H}(\mathbf{y})$ e U_s l'operatore di traslazione del processo $\{\mathbf{y}(t)\}$. Vale allora la relazione*

$$\hat{E}[U_s \mathbf{z} | U_s \mathbf{K}] = U_s \hat{E}[\mathbf{z} | \mathbf{K}] \quad , \quad s \in \mathbb{Z} \quad . \quad (1.5)$$

In particolare, il "processo predittore" $\hat{\mathbf{y}}(t+1 | t)$ è stazionario e la varianza dell'errore di predizione di un passo è costante nel tempo.

DIMOSTRAZIONE. L'ultima affermazione scende dalla relazione

$$\hat{\mathbf{y}}(t+s+1 | t+s) = \hat{E}[U_s \mathbf{y}(t+1) | U_s \mathbf{H}_t(\mathbf{y})] = U_s \hat{E}[\mathbf{y}(t+1) | \mathbf{H}_t(\mathbf{y})], \quad (1.6)$$

per cui è sufficiente mostrare la validità della relazione generale (1.5).

La proiezione $\hat{E}[U_s \mathbf{z} | U_s \mathbf{K}]$ è l'unico vettore di $U_s \mathbf{K}$ che soddisfa alla

$$\langle U_s \mathbf{z} - \hat{E}[U_s \mathbf{z} | U_s \mathbf{K}], U_s \mathbf{k} \rangle = 0 \quad , \quad (1.7)$$

per ogni $\mathbf{k} \in \mathbf{K}$. Usando la definizione di operatore aggiunto U_s^* e la $U_s^* U_s = I$ la (1.7) si può riscrivere

$$\langle \mathbf{z} - U_s^* \hat{E}[U_s \mathbf{z} | U_s \mathbf{K}], \mathbf{k} \rangle = 0 \quad , \quad \forall \mathbf{k} \in \mathbf{K} \quad ,$$

per cui

$$U_s^* \hat{E}[U_s \mathbf{z} | U_s \mathbf{K}] = \hat{E}[\mathbf{z} | \mathbf{K}] \quad ;$$

moltiplicando a sinistra per U_s si trova così

$$\hat{E}[U_s \mathbf{z} \mid U_s \mathbf{K}] = U_s \hat{E}[\mathbf{z} \mid \mathbf{K}] \quad , \quad (1.8)$$

che è la relazione cercata. \square

Calcolando la varianza dell'errore di predizione all'istante $t + s$, si trova allora

$$\begin{aligned} & \|\mathbf{y}(t + s + 1) - \hat{E}[\mathbf{y}(t + s + 1) \mid \mathbf{H}_{t+s}(\mathbf{y})]\|^2 = \\ & = \|U_s[\mathbf{y}(t + 1) - \hat{E}[\mathbf{y}(t + 1) \mid \mathbf{H}_t(\mathbf{y})]]\|^2 = \lambda^2 \end{aligned}$$

per qualunque s .

Vi sono casi degeneri in cui l'errore di predizione è nullo. I processi per cui questo accade sono evidentemente caratterizzati dalla proprietà che $\mathbf{y}(t+1) \in \mathbf{H}_t(\mathbf{y})$ e, per il lemma appena dimostrato, se questa circostanza si verifica per qualche t essa vale per *tutti* i $t \in \mathbb{Z}$.

DEFINIZIONE 5.1. *Un processo del secondo ordine $\{\mathbf{y}(t)\}$ si dice puramente deterministico se $\lambda^2 = 0$, ovvero se*

$$\mathbf{H}_t(\mathbf{y}) = \mathbf{H}_{t+1}(\mathbf{y}) \quad , \quad \forall t \in \mathbb{Z} \quad , \quad (1.9)$$

Notiamo subito che se vale la (1.9), tutti i sottospazi “passato”, $\mathbf{H}_t(\mathbf{y})$, $t \in \mathbb{Z}$, di $\{\mathbf{y}(t)\}$ coincidono e si ha

$$\mathbf{H}_t(\mathbf{y}) = \mathbf{H}(\mathbf{y}) \quad , \quad \forall t \in \mathbb{Z} \quad . \quad (1.10)$$

Un processo puramente deterministico (abbreviato a p.d. nel seguito) è per definizione *esattamente predicibile* in base all'osservazione del suo passato.²

Per caratterizzare la classe di processi per i quali questa degenerazione non si verifica conviene introdurre alcune considerazioni di carattere geometrico.

DEFINIZIONE 5.2. *Il passato remoto di un processo del secondo ordine $\{\mathbf{y}(t)\}$ è il sottospazio $\mathbf{H}_{-\infty}(\mathbf{y})$ definito dalla*

$$\mathbf{H}_{-\infty}(\mathbf{y}) := \bigcap_{t \leq k} \mathbf{H}_t(\mathbf{y}) \quad . \quad (1.11)$$

Dato che gli $\mathbf{H}_t(\mathbf{y})$ formano una famiglia non decrescente, l'intersezione nella (1.11) è indipendente dall'indice (finito) k che si considera. Notiamo allora che $\{\mathbf{y}(t)\}$ è p.d. se e solo se

$$\mathbf{H}_{-\infty}(\mathbf{y}) = \mathbf{H}_t(\mathbf{y}) = \mathbf{H}(\mathbf{y}) \quad , \quad \forall t \in \mathbb{Z} \quad , \quad (1.12)$$

come si vede immediatamente confrontando le (1.10) e (1.11). In particolare tutta la storia di un processo p.d. è contenuta nel suo passato remoto, nel senso che $\mathbf{H}_{-\infty}(\mathbf{y}) = \mathbf{H}(\mathbf{y})$.

Come spiegheremo meglio più avanti, tutta “l'aleatorietà” di un processo p.d. si può pensare concentrata nelle “condizioni iniziali” a $t = -\infty$. Intuitivamente, ci si aspetta che, a causa dell'uguaglianza (1.12), le traiettorie di un processo p.d. siano funzioni il cui andamento temporale è completamente “prevedibile”: in effetti ciascuna traiettoria deve essere univocamente determinata dalla propria restrizione ad un intorno di $t = -\infty$ (ovvero ad intervalli del tipo $(-\infty, t]$ con t arbitrario).

²Dovremmo in realtà dire che ogni $\mathbf{y}(t+k)$, $k \geq 0$, è esattamente predicibile mediante uno stimatore lineare basato sulla storia passata $\mathbf{H}_t(\mathbf{y})$ di $\{\mathbf{y}(t)\}$.

La situazione opposta si ha per processi il cui passato remoto è triviale (costituito solo dalla variabile aleatoria zero). Di che tipo di processi si tratti è chiarito dal seguente fondamentale teorema di Wold.

TEOREMA 5.1 (H. Wold). *I processi del secondo ordine il cui passato remoto è triviale, ovvero*

$$\mathbf{H}_{-\infty}(\mathbf{y}) = \{0\} \quad , \quad (1.13)$$

sono processi puramente non deterministici.

DIMOSTRAZIONE. Costruiamo, per ogni t , il complemento ortogonale, \mathbf{E}_t , di $\mathbf{H}_{t-1}(\mathbf{y})$ in $\mathbf{H}_t(\mathbf{y})$

$$\mathbf{E}_t := \mathbf{H}_t(\mathbf{y}) \ominus \mathbf{H}_{t-1}(\mathbf{y}) \quad . \quad (1.14)$$

e notiamo subito che \mathbf{E}_t è lo spazio vettoriale generato dalle componenti dell'errore di predizione di un passo $\tilde{\mathbf{y}}_k(t | t-1) := \mathbf{y}_k(t) - \hat{\mathbf{y}}_k(t | t-1)$, $k = 1 \dots m$, per cui

$$\mathbf{E}_t = \mathbf{H}(\tilde{\mathbf{y}}(t | t-1)) \quad (1.15)$$

e come tale ha al più dimensione m .

Dato inoltre che $\mathbf{E}_t \perp \mathbf{H}_{t-1}(\mathbf{y})$ e $\mathbf{H}_{t-1}(t) \supset \mathbf{H}_s(\mathbf{y}) \supset \mathbf{E}_s$, $s \leq t-1$, si ha

$$\mathbf{E}_t \perp \mathbf{E}_s \quad , \quad \forall s < t \quad , \quad (1.16)$$

per cui la famiglia $\{\mathbf{E}_t\}_{t \in \mathbb{Z}}$ è formata di sottospazi ortogonali a due a due. Iterando la decomposizione ortogonale $\mathbf{H}_t(\mathbf{y}) = \mathbf{E}_t \oplus \mathbf{H}_{t-1}(\mathbf{y})$ si ottiene

$$\mathbf{H}_t(\mathbf{y}) = \mathbf{E}_t \oplus \mathbf{E}_{t-1} \oplus \dots \oplus \mathbf{H}_s(\mathbf{y}) \quad , \quad s < t \quad . \quad (1.17)$$

Sia ora \mathbf{z} un qualunque elemento di $\mathbf{H}_t(\mathbf{y})$. Proiettando \mathbf{z} sulla somma diretta ortogonale al secondo membro della (1.17) si ottiene:

$$\mathbf{z} = \hat{\mathbf{z}}(t) + \hat{\mathbf{z}}(t-1) + \dots + \hat{\mathbf{z}}(s+1) + \hat{\mathbf{z}}_s \quad , \quad (1.18)$$

dove i termini

$$\hat{\mathbf{z}}(i) := \hat{E}[\mathbf{z} | \mathbf{E}_i] \quad , \quad i = s+1, \dots, t$$

$$\hat{\mathbf{z}}_s := \hat{E}[\mathbf{z} | \mathbf{H}_s(\mathbf{y})]$$

sono mutuamente ortogonali. Notiamo ora che la somma a termini ortogonali

$$\sum_{s+1}^t \hat{\mathbf{z}}(i) \quad (1.19)$$

è convergente (in norma) per $s \rightarrow -\infty$, dato che dalla (1.18) segue che

$$\left\| \sum_{s+1}^t \hat{\mathbf{z}}(i) \right\|^2 = \sum_{s+1}^t \|\hat{\mathbf{z}}(i)\|^2 \leq \|\mathbf{z}\|^2 < \infty$$

per ogni $s < t$ (si veda l'Appendice B). Ne segue allora che esiste anche il $\lim \hat{\mathbf{z}}_s$ per $s \rightarrow -\infty$, perchè nella (1.18) gli addendi (1.19) e $\hat{\mathbf{z}}_s$ hanno somma costante, \mathbf{z} , indipendente da s .

Ora, dato che $\hat{\mathbf{z}}_s \in \mathbf{H}_s(\mathbf{y})$, $\forall s$, si deve avere

$$\lim_{s \rightarrow -\infty} \hat{\mathbf{z}}_s \in \bigcap_s \mathbf{H}_s(\mathbf{y}) = \mathbf{H}_{-\infty}(\mathbf{y}) \quad (1.20)$$

e quindi il limite è zero se il passato remoto del processo $\{\mathbf{y}(t)\}$ è triviale.

Ne deriva che in questo caso (e solo in questo caso) ogni $\mathbf{z} \in \mathbf{H}_t(\mathbf{y})$ si può esprimere come somma infinita ortogonale di proiezioni sui sottospazi \mathbf{E}_t

$$\mathbf{z} = \sum_{-\infty}^t \hat{\mathbf{z}}(s)$$

ovvero, equivalentemente, $\mathbf{H}_t(\mathbf{y})$ si può decomporre in somma infinita ortogonale

$$\mathbf{H}_t(\mathbf{y}) = \bigoplus_{-\infty}^t \mathbf{E}_s \quad . \quad (1.21)$$

Notiamo adesso che per la (1.15) e la (1.6) si ha:

$$\mathbf{E}_{t+\Delta} = U_{\Delta} \mathbf{E}_t \quad . \quad (1.22)$$

Fissiamo allora una base (p -dimensionale), $\mathbf{e}(s) = [\mathbf{e}_1(s), \dots, \mathbf{e}_p(s)]'$ in \mathbf{E}_s .

Chiaramente si ha $p \leq m$ e l'uguaglianza $p = m$ vale solo quando le componenti di $\tilde{\mathbf{y}}(s | s-1)$ sono linearmente indipendenti (ovvero la matrice varianza d'errore di predizione $E[\tilde{\mathbf{y}}(s | s-1) \tilde{\mathbf{y}}(s | s-1)']$ è definita positiva). In questo caso si può prendere $\mathbf{e}(s) = \tilde{\mathbf{y}}(s | s-1)$.

Definiamo in ogni caso il processo $\{\mathbf{e}(t)\}$ ponendo

$$\mathbf{e}(t) := U_{t-s} \mathbf{e}(s) \quad \forall t \quad . \quad (1.23)$$

In altre parole $\{\mathbf{y}(t)\}$ ed $\{\mathbf{e}(t)\}$ si propagano nel tempo con lo stesso operatore di traslazione e sono così processi *congiuntamente stazionari*. È chiaro dalla relazione (1.22) che con questa definizione si ottiene, per ciascun t , ancora una base in \mathbf{E}_t . Il processo $\{\mathbf{e}(t)\}$ è inoltre a variabili ortogonali, per la (1.16). Da questa scelta di base segue allora

$$\mathbf{H}_t(\mathbf{e}) = \bigoplus_{-\infty}^t \mathbf{H}(\mathbf{e}(s)) = \bigoplus_{-\infty}^t \mathbf{E}_s = \mathbf{H}_t(\mathbf{y}) \quad ,$$

in cui l'ultima uguaglianza scende dalla (1.21). Abbiamo così costruito un processo bianco stazionario, causalmente equivalente a $\{\mathbf{y}(t)\}$.

La proiezione di $\mathbf{y}_k(t)$ su \mathbf{E}_s ha la forma

$$\hat{E}[\mathbf{y}_k(t) | \mathbf{e}(s)] = w_k(t, s) \mathbf{e}(s) \quad ,$$

dove $w_k(t, s)$ è il vettore riga (deterministico) a p componenti delle covarianze di $\mathbf{y}_k(t)$ ed $\mathbf{e}(s)$ diviso per la varianza (costante) di $\mathbf{e}(s)$. Per la stazionarietà congiunta w_k dipende solo dalla differenza $t - s$ e, col solito abuso di notazioni scriveremo $w_k(t, s) = w_k(t - s)$. Compattando in notazione matriciale, si ha così

$$\hat{E}[\mathbf{y}(t) | \mathbf{e}(s)] = W(t - s) \mathbf{e}(s) \quad , \quad (1.24)$$

dove $W(t - s)$ è una matrice $m \times p$ univocamente individuata dai momenti congiunti di $\mathbf{y}(t)$ ed $\mathbf{e}(s)$ tramite la nota formula

$$W(t - s) = E[\mathbf{y}(t) \mathbf{e}(s)'] \Lambda^{-1} \quad , \quad (1.25)$$

dove Λ è la matrice varianza di $\mathbf{e}(s)$. Dalla (1.25) si vede in particolare che $W(t) = 0$ per $t < 0$.

La decomposizione (1.21) fornisce così l'espressione di $\mathbf{y}(t)$ come risposta di un sistema lineare causale di risposta impulsiva $W(\cdot)$, all'ingresso bianco $\{\mathbf{e}(t)\}$

$$\mathbf{y}(t) = \sum_{-\infty}^t W(t-s) \mathbf{e}(s) \quad . \quad (1.26)$$

Questo mostra che $\{\mathbf{y}(t)\}$ è p.n.d.

Per mostrare la necessità della condizione (1.13) basta notare che se vale una rappresentazione *causale* come la (1.26) si ha $\mathbf{y}(t) \in \mathbf{H}_t(\mathbf{e})$ per ogni t e quindi $\mathbf{H}_t(\mathbf{y}) \subset \mathbf{H}_t(\mathbf{e})$, $t \in \mathbb{Z}$. La conclusione si basa allora sul lemma seguente. \square

LEMMA 5.2. *Il passato remoto di un processo a variabili ortogonali è triviale.*

DIMOSTRAZIONE. Sia $\{\mathbf{u}(t)\}$ un processo (non necessariamente stazionario) a media nulla con $E|\mathbf{u}(t)|^2 < \infty$, $\forall t$, e $\mathbf{u}(t) \perp \mathbf{u}(s)$, $\forall t, s \in \mathbb{Z}$. Per l'ortogonalità di $\{\mathbf{u}(t)\}$ ogni variabile $\mathbf{h} \in \mathbf{H}_t(\mathbf{u})$ è ortogonale a tutti i valori futuri $\mathbf{u}(k)$, $k > t$. In particolare, se $\mathbf{h} \in \mathbf{H}_{-\infty}(\mathbf{u})$, \mathbf{h} deve essere ortogonale a tutti gli $\mathbf{u}(t)$, $t \in \mathbb{Z}$. Segue che $\langle \mathbf{h}, \mathbf{z}_k \rangle = 0$ per ogni $\mathbf{z}_k \in \tilde{\mathbf{H}}(\mathbf{u})$. Approssimando un'arbitraria variabile $\mathbf{z} \in \mathbf{H}(\mathbf{u})$ con una successione $\{\mathbf{z}_k\} \subset \tilde{\mathbf{H}}(\mathbf{u})$ e passando al limite, si ha $\langle \mathbf{h}, \mathbf{z} \rangle = 0$, $\forall \mathbf{z} \in \mathbf{H}(\mathbf{u})$. Ma l'unica variabile di $\mathbf{H}(\mathbf{u})$ ortogonale a tutto $\mathbf{H}(\mathbf{u})$ è la variabile zero. \square

OSSERVAZIONE 5.1. È evidente che la (1.26) è proprio una *rappresentazione di innovazione* del processo p.n.d. $\{\mathbf{y}(t)\}$ e la trasformata zeta, $\hat{W}(z)$, della risposta impulsiva W , definita in (1.25), un fattore spettrale a fase minima. Se $\{\mathbf{y}(t)\}$ è a rango pieno (e quindi $\hat{W}(z)$ è quadrata $m \times m$) si vede, usando il principio di ortogonalità, che

$$\hat{W}(\infty) = W(0) = E[\mathbf{y}(t) \tilde{\mathbf{y}}(t | t-1)'] \Lambda^{-1} = E[\tilde{\mathbf{y}}(t | t-1) \tilde{\mathbf{y}}(t | t-1)'] \Lambda^{-1} = I,$$

cosicché $\hat{W}(z)$ è precisamente il fattore a fase minima *normalizzato* della Definizione 3.3. \diamond

Decomposizione di Wold. Le due definizioni di processo p.d. e p.n.d. appena date descrivono due casi estremi. Come vadano le cose in generale è chiarito dal seguente teorema (ancora dovuto a Wold [155]).

TEOREMA 5.2. *Ogni processo del secondo ordine ammette una decomposizione*

$$\mathbf{y}(t) = \mathbf{u}(t) + \mathbf{v}(t) \quad , \quad (1.27)$$

dove $\{\mathbf{u}(t)\}$ è p.n.d., $\{\mathbf{v}(t)\}$ è p.d. e $\{\mathbf{u}(t)\}$ e $\{\mathbf{v}(t)\}$ sono scorrelati, ovvero $\langle \mathbf{u}_k(t), \mathbf{v}_j(s) \rangle = 0$, $\forall t, s \in \mathbb{Z}$, $\forall k, j = 1, \dots, m$. La decomposizione (1.27) è unica se $\mathbf{H}_t(\mathbf{u}) \subset \mathbf{H}_t(\mathbf{y})$ e $\mathbf{H}_t(\mathbf{v}) \subset \mathbf{H}_t(\mathbf{y})$.

DIMOSTRAZIONE. Sia $\mathbf{H}_{-\infty}(\mathbf{y})$ il passato remoto del processo. Faremo vedere che la proiezione di $\mathbf{y}(t)$ su $\mathbf{H}_{-\infty}(\mathbf{y})$,

$$\mathbf{v}(t) := \hat{E}[\mathbf{y}(t) | \mathbf{H}_{-\infty}(\mathbf{y})] \quad (1.28)$$

e il relativo "errore di stima"

$$\mathbf{u}(t) := \mathbf{y}(t) - \hat{E}[\mathbf{y}(t) | \mathbf{H}_{-\infty}(\mathbf{y})] = \mathbf{y}(t) - \mathbf{v}(t) \quad (1.29)$$

sono appunto la parte p.d. e (rispettivamente) p.n.d. di $\{\mathbf{y}(t)\}$. Il nucleo della dimostrazione consiste nel far vedere che

$$\mathbf{H}_t(\mathbf{y}) = \mathbf{H}_t(\mathbf{u}) \oplus \mathbf{H}_{-\infty}(\mathbf{y}) \quad , \quad \forall t \in \mathbb{Z} \quad , \quad (1.30)$$

dove la somma diretta è ortogonale.

Dalla (1.29) segue che, anche per $s \leq t$, $\mathbf{y}(s)$ è somma ortogonale di un elemento di $\mathbf{H}_t(\mathbf{u})$, $\mathbf{u}(s)$, e di uno in $\mathbf{H}_{-\infty}(\mathbf{y})$, $\hat{E}[\mathbf{y}(s) | \mathbf{H}_{-\infty}]$.

Ovviamente, per il principio di ortogonalità, una qualunque combinazione lineare finita di variabili $\{\mathbf{u}(s) ; s \leq t\}$ è ortogonale ad $\mathbf{H}_{-\infty}(\mathbf{y})$.

Ne deriva che

$$\mathbf{H}_t(\mathbf{u}) \oplus \mathbf{H}_{-\infty}(\mathbf{y}) \supset \mathbf{H}_t(\mathbf{y}) \quad .$$

D'altra parte $\mathbf{H}_{-\infty}(\mathbf{y}) \subset \mathbf{H}_t(\mathbf{y})$ e $\mathbf{H}_t(\mathbf{u}) \subset \mathbf{H}_t(\mathbf{y})$. L'ultima inclusione segue dalla $\mathbf{u}(s) = \mathbf{y}(s) - \hat{E}[\mathbf{y}(s) | \mathbf{H}_{-\infty}(\mathbf{y})]$. Se ne conclude che vale la (1.30).

Ora, $\{\mathbf{u}(t)\}$ è p.n.d. dato che $\bigcap_t \mathbf{H}_t(\mathbf{u}) \subset \bigcap_t \mathbf{H}_t(\mathbf{y}) = \mathbf{H}_{-\infty}(\mathbf{y})$ e ogni elemento nella prima intersezione è ortogonale ad $\mathbf{H}_{-\infty}(\mathbf{y})$. Per quanto riguarda $\{\mathbf{v}(t)\}$ notiamo che dalla $\mathbf{y}(t) = \mathbf{u}(t) + \mathbf{v}(t)$ segue

$$\mathbf{H}_t(\mathbf{y}) = \mathbf{H}_t(\mathbf{u}) \oplus \mathbf{H}_t(\mathbf{v})$$

e dato che il sottospazio $\mathbf{H}_t(\mathbf{u})$ può avere un unico complemento ortogonale in $\mathbf{H}_t(\mathbf{y})$, dalla (1.30) si trae che $\mathbf{H}_t(\mathbf{v}) = \mathbf{H}_{-\infty}(\mathbf{y})$, $\forall t$. \square

Cercheremo ora di ricavare qualche informazione sulla struttura della componente p.d. di un processo del secondo ordine $\{\mathbf{y}(t)\}$.

Notiamo a questo proposito che se $\mathbf{z} \in \mathbf{H}_{-\infty}(\mathbf{y})$ allora, qualunque sia s , la variabile traslata $U_s \mathbf{z}$ sta ancora nel passato remoto, dato che $\mathbf{z} \in \mathbf{H}_t(\mathbf{y}) \quad \forall t \Rightarrow U_s \mathbf{z} \in \mathbf{H}_{t+s}(\mathbf{y})$, $\forall t$ e quindi

$$U_s \mathbf{z} \in \bigcap_t \mathbf{H}_{t+s}(\mathbf{y}) = \bigcap_t \mathbf{H}_t(\mathbf{y}) = \mathbf{H}_{-\infty}(\mathbf{y}) \quad .$$

In particolare valgono le $U \mathbf{H}_{-\infty}(\mathbf{y}) \subset \mathbf{H}_{-\infty}(\mathbf{y})$ e $U^{-1} \mathbf{H}_{-\infty}(\mathbf{y}) \subset \mathbf{H}_{-\infty}(\mathbf{y})$ (notare infatti che s può essere anche negativo). Un sottospazio come $\mathbf{H}_{-\infty}(\mathbf{y})$ invariante sia per U che per U^* ($= U^{-1}$), si dice *doppiamente invariante*.

Si ha quindi la seguente caratterizzazione di $\mathbf{H}_{-\infty}$.

PROPOSIZIONE 5.1. *Il passato remoto è un sottospazio doppiamente invariante per l'operatore di traslazione U , ovvero*

$$U \mathbf{H}_{-\infty}(\mathbf{y}) = \mathbf{H}_{-\infty}(\mathbf{y}) \quad . \quad (1.31)$$

Processi Puramente Deterministici Elementari. Notiamo che i più semplici sottospazi doppiamente invarianti per una trasformazione lineare sono generati da *autovettori* ed è perciò naturale cercare di studiare $\mathbf{H}_{-\infty}$ ricercando le variabili aleatorie \mathbf{z} che soddisfano la

$$U \mathbf{z} = \mu \mathbf{z} \quad , \quad \mathbf{z} \neq 0 \quad . \quad (1.32)$$

In questo problema, l'isometria di U implica $\|\mathbf{z}\|^2 = \langle U \mathbf{z}, U \mathbf{z} \rangle = |\mu|^2 \|\mathbf{z}\|^2$ e perciò $|\mu|^2 = 1$, ovvero

$$\mu = e^{j\lambda} \quad , \quad \lambda \in [-\pi, \pi] \quad . \quad (1.33)$$

Inoltre, dato che U trasforma variabili reali in variabili reali, gli autovalori di U sono in coppie complesse coniugate, $e^{\pm j\lambda}$, e associati rispettivamente ad autovettori

$\mathbf{z}_\pm = \mathbf{x} \pm j\mathbf{y}$ tra loro ortogonali (ricordare infatti che $UU^* = U^*U = I$, e pertanto U è una trasformazione lineare “normale” che ha sempre autovettori ortogonali.).

Supponiamo dunque che l’operatore di traslazione associato ad un processo del secondo ordine $\{\mathbf{y}(t)\}$ abbia due autovalori (semplici) in $e^{\pm j\lambda_0}$ con associati gli autovettori \mathbf{z}_+ e \mathbf{z}_- . Sia \mathbf{H} il sottospazio bidimensionale di $\mathbf{H}(\mathbf{y})$ generato da questi due autovettori. Chiaramente \mathbf{H} ammette una base reale, ovvero \mathbf{H} si può costruire anche come

$$\mathbf{H} = \text{span} \{\mathbf{x}, \mathbf{y}\} \quad ,$$

dove le variabili casuali \mathbf{x} e \mathbf{y} si ricavano da \mathbf{z}_+ e \mathbf{z}_- attraverso le relazioni $\mathbf{x} = \frac{1}{2}(\mathbf{z}_+ + \mathbf{z}_-)$, $\mathbf{y} = \frac{1}{2j}(\mathbf{z}_+ - \mathbf{z}_-)$. Da queste formule è immediato verificare che anche \mathbf{x} e \mathbf{y} sono tra loro ortogonali.

Vediamo come U_t opera su \mathbf{H} . Con calcoli ovvi si ricavano dalla $U^t \mathbf{z}_\pm = e^{\pm j\lambda_0 t} \mathbf{z}_\pm$ le formule

$$\begin{aligned} U^t \mathbf{x} &= \mathbf{x} \cos \lambda_0 t - \mathbf{y} \sin \lambda_0 t \\ U^t \mathbf{y} &= \mathbf{x} \sin \lambda_0 t + \mathbf{y} \cos \lambda_0 t \end{aligned}$$

che possono essere riscritte nella forma

$$\begin{bmatrix} U^t \mathbf{x} \\ U^t \mathbf{y} \end{bmatrix} = \begin{bmatrix} \cos \lambda_0 & -\sin \lambda_0 \\ \sin \lambda_0 & \cos \lambda_0 \end{bmatrix}^t \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} := F^t \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \quad , \quad (1.34)$$

dove la matrice (deterministica)

$$F := \begin{bmatrix} \cos \lambda_0 & -\sin \lambda_0 \\ \sin \lambda_0 & \cos \lambda_0 \end{bmatrix}$$

è ortogonale (e quindi non singolare). Si vede immediatamente dalla (1.34) che \mathbf{H} è doppiamente invariante per U nel senso della Proposizione 5.1.

Consideriamo ora la proiezione, $\mathbf{v}(t)$, di $\mathbf{y}(t)$ su \mathbf{H} . Supporremo per semplicità che il processo $\{\mathbf{y}(t)\}$ sia scalare.

All’istante $t = 0$ questa proiezione è una certa combinazione lineare delle due variabili di base, del tipo $a\mathbf{x} + b\mathbf{y}$.

In virtù della formula precedente si vede che $\mathbf{v}(t)$ si può esprimere come

$$\begin{aligned} \mathbf{v}(t) &:= \hat{E}[\mathbf{y}(t) | \mathbf{H}] = U_t \hat{E}[\mathbf{y}(0) | \mathbf{H}] = U_t(a\mathbf{x} + b\mathbf{y}) \\ &= [a, b] [U^t \mathbf{x} \ U^t \mathbf{y}]' \\ &= [a, b] F^t \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \\ &= (a\mathbf{x} + b\mathbf{y}) \cos \lambda_0 t + (b\mathbf{x} - a\mathbf{y}) \sin \lambda_0 t \\ &:= \boldsymbol{\xi} \cos \lambda_0 t + \boldsymbol{\eta} \sin \lambda_0 t \end{aligned} \quad (1.35)$$

dove $\boldsymbol{\xi}$ ed $\boldsymbol{\eta}$ sono due variabili di \mathbf{H} . Dato che $\{\mathbf{v}(t)\}$ deve essere stazionario e pertanto $\|\mathbf{v}(t)\|^2 = \|\mathbf{v}(0)\|^2$, $\forall t$, usando l’ortogonalità di \mathbf{x} e \mathbf{y} , si ricava che $\boldsymbol{\xi}$ ed $\boldsymbol{\eta}$ devono sottostare alle condizioni

$$E(\boldsymbol{\xi}\boldsymbol{\eta}) = 0 \quad , \quad E\boldsymbol{\xi}^2 = E\boldsymbol{\eta}^2 = A^2 \quad ,$$

le quali implicano che la covarianza di $\{\mathbf{v}(t)\}$ ha la forma

$$\sigma_v(\tau) = A^2 \cos \lambda_0 \tau \quad . \quad (1.36)$$

Notiamo che il processo $\{\mathbf{v}(t)\}$ è puramente deterministico, dato che $\mathbf{H}_{-\infty}(\mathbf{v}) = \mathbf{H}_t(\mathbf{v}) = \mathbf{H}$, $\forall t$. Inoltre il suo spettro consiste in due “righe” (impulsi di Dirac nella

densità spettrale) di ampiezza $A^2/2$ alle frequenze $\pm\lambda_0/2\pi$. Chiaramente queste righe si ritroveranno nello spettro (complessivo) del processo $\{\mathbf{y}(t)\}$.

PROBLEMA 5.1. *Dimostrare che le righe spettrali di $\{\sigma_v(t)\}$ si ritrovano nella distribuzione spettrale del processo $\{\mathbf{y}(t)\}$.*

Soluzione: Ovviamente si può scrivere $\mathbf{y}(t) = \mathbf{v}(t) + \tilde{\mathbf{y}}(t)$ con \mathbf{v} e $\tilde{\mathbf{y}}$ processi completamente scorrelati, quindi la funzione covarianza di \mathbf{y} è somma della covarianza di \mathbf{v} e di quella di $\tilde{\mathbf{y}}$. \diamond

Come vedremo meglio in seguito, l'argomento appena esposto si può invertire, nel senso che ad ogni coppia di righe spettrali del processo $\{\mathbf{y}(t)\}$ corrisponde una decomposizione ortogonale $\mathbf{y}(t) = \mathbf{v}(t) + \tilde{\mathbf{y}}(t)$, dove $\{\mathbf{v}(t)\}$ è una componente p.d. elementare del processo del tipo (1.35) la cui funzione covarianza $\sigma_v(\tau)$ è del tipo (1.36). La componente $\{\mathbf{v}(t)\}$ genera un sottospazio bidimensionale doppiamente invariante di $\mathbf{H}(\mathbf{y})$ a cui corrisponde una coppia di autovalori complessi coniugati del relativo operatore di traslazione U .

Più in generale, ad una serie di righe spettrali alle frequenze $\lambda_k/2\pi$, $k = 1, 2, \dots$, corrisponde una componente p.d. quasi periodica, somma di oscillazioni aleatorie elementari $\{\mathbf{v}_k(t)\}$, $k = 1, 2, \dots$, fra loro scorrelate, di frequenze $\{\lambda_k/2\pi\}$.

Abbiamo così stabilito un'importante corrispondenza tra righe spettrali di $\{\mathbf{y}(t)\}$, autovalori del relativo operatore di traslazione e componenti p.d. del processo.

OSSERVAZIONE 5.2. Che una oscillazione aleatoria elementare come $\{\mathbf{v}(t)\}$ sia p.d., si può vedere anche direttamente, usando la formula (1.35) e il teorema di Cayley-Hamilton,

$$F^2 + a_1 F + a_2 I = 0 \quad ,$$

dove a_1 e a_2 sono i coefficienti del polinomio caratteristico di F . Infatti dalla (1.35) risulta che

$$\mathbf{v}(t+1) + a_1 \mathbf{v}(t) + a_2 \mathbf{v}(t-1) = \begin{bmatrix} a & b \end{bmatrix} (F^2 + a_1 F + a_2 I) F^{t-1} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = 0$$

e perciò $\mathbf{v}(t+1)$ è esattamente prevedibile se sono noti $\mathbf{v}(t)$ e $\mathbf{v}(t-1)$, in quanto

$$\mathbf{v}(t+1) = -a_1 \mathbf{v}(t) - a_2 \mathbf{v}(t-1) \quad .$$

\diamond

La Condizione di Szegő. È bene chiarire che la discussione precedente sulle righe spettrali descrive solo una classe di possibili componenti p.d. di un processo stazionario $\{\mathbf{y}(t)\}$. In generale possono essere presenti componenti p.d. di natura diversa. In particolare, un processo stazionario può essere p.d. anche quando la sua distribuzione spettrale ammette densità (è assolutamente continua). In questi casi però la struttura temporale del processo non è una semplice somma di oscillazioni armoniche elementari. In effetti, descrivere le traiettorie di un processo di questo tipo è a tutt'oggi uno dei classici problemi aperti in teoria della probabilità.

Siano \mathbf{y}_a e \mathbf{y}_s le componenti assolutamente continua e singolare del processo \mathbf{y} , introdotte in Appendice D, all'Osservazione D.1. È evidente che \mathbf{y}_s , che non ha densità spettrale, dovrà necessariamente essere puramente deterministico.

Si può dare una condizione necessaria e sufficiente affinché \mathbf{y}_a sia p.n.d. Un classico criterio di Szegő-Kolmogorov [82], [126], afferma che

TEOREMA 5.3 (Szegö-Kolmogorov). *Un processo stazionario del secondo ordine di rango pieno, a spettro assolutamente continuo, è p.n.d. se e solo se la sua densità spettrale $S(e^{j\omega})$ soddisfa alla condizione³*

$$\int_{-\pi}^{\pi} \log \det S(e^{j\omega}) \frac{d\omega}{2\pi} > -\infty \quad . \quad (1.37)$$

Viceversa, se l'integrale (1.37) diverge, il processo è puramente deterministico.

La condizione esprime in sostanza il fatto che la densità spettrale di un processo p.n.d. non può annullarsi per “troppi” valori di ω e nel caso di zeri puntuali dà implicitamente un limite sulla “velocità” con cui uno spettro può annullarsi. Per processi scalari, la (1.37) dice in particolare che lo spettro di un processo p.n.d. non può essere nullo su un intervallo di frequenze di misura positiva.

Si può verificare agevolmente che per spettri razionali (di rango pieno) la (1.37) è automaticamente soddisfatta.

PROBLEMA 5.2. *Mostrare che i cosiddetti segnali a banda limitata non solo non sono ortonormalizzabili, ma, per la seconda parte del teorema di Szegö, sono puramente deterministici!* \diamond

2. Convergenza di Martingale

La costruzione dei sottospazi \mathbf{E}_t nella prova del Teorema di Wold è un'idea suscettibile di importanti generalizzazioni e applicazioni.

Supponiamo di avere a disposizione una successione *crescente* di sottospazi (chiusi) $\{\mathbf{H}_t ; t \in \mathbb{Z}\}$ di un certo spazio \mathbf{H} di variabili aleatorie del secondo ordine. Definiamo

$$\mathbf{H}_{-\infty} := \bigcap_t \mathbf{H}_t \quad , \quad \mathbf{H}_{\infty} := \overline{\text{span}} \{\mathbf{H}_t ; t \in \mathbb{Z}\} \quad .$$

(\mathbf{H}_{∞} è la somma vettoriale degli $\{\mathbf{H}_t\}$, chiusa rispetto alla convergenza in media quadratica) e

$$\begin{aligned} \bar{\mathbf{H}}_t &:= \mathbf{H}_t \ominus \mathbf{H}_{-\infty} \\ \mathbf{E}_t &:= \bar{\mathbf{H}}_t \ominus \bar{\mathbf{H}}_{t-1} \quad , \quad t \in \mathbb{Z} \quad . \end{aligned} \quad (2.1)$$

Esattamente come nel caso particolare $\mathbf{H}_t = \mathbf{H}_t(\mathbf{y})$, i sottospazi $\{\mathbf{E}_t\}$ sono a due a due ortogonali e forniscono uno sviluppo in somma diretta ortogonale di \mathbf{H}_t

$$\mathbf{H}_t = \left(\bigoplus_{k=-\infty}^t \mathbf{E}_k \right) \oplus \mathbf{H}_{-\infty} \quad , \quad (2.2)$$

con significato già chiarito a suo tempo quando fu ricavata la (1.21). Uno sviluppo analogo vale anche per \mathbf{H}_{∞} .

Sia ora \mathbf{z} una variabile (scalare) di \mathbf{H} . Il processo

$$\mathbf{y}(t) := \hat{E}(\mathbf{z} | \mathbf{H}_t) \quad , \quad t \in \mathbb{Z} \quad , \quad (2.3)$$

ha la proprietà di essere “condizionatamente costante” nel tempo

$$\hat{E}[\mathbf{y}(t) | \mathbf{H}_s] = \mathbf{y}(s) \quad , \quad s \leq t \quad .$$

³Si può mostrare che l'integrale è sempre limitato superiormente se $S(e^{j\omega})$ è una densità spettrale.

Questa proprietà in letteratura è chiamata *proprietà di Martingala* (in senso debole) *rispetto alla famiglia crescente di sottospazi* $\{\mathbf{H}_t\}$. Il risultato seguente è la versione debole di un teorema di convergenza (“delle martingale”) molto utile e importante in teoria della probabilità.

TEOREMA 5.4 (Doob). *Ogni martingala a varianza finita ammette limite per $t \rightarrow \pm\infty$. Si ha in particolare*

$$\begin{aligned}\lim_{t \rightarrow -\infty} \hat{E}[\mathbf{z} | \mathbf{H}_t] &= \hat{E}[\mathbf{z} | \mathbf{H}_{-\infty}] \\ \lim_{t \rightarrow +\infty} \hat{E}[\mathbf{z} | \mathbf{H}_t] &= \hat{E}[\mathbf{z} | \mathbf{H}_{\infty}]\end{aligned}\quad (2.4)$$

dove i limiti sono in media quadratica.

DIMOSTRAZIONE. Dimostreremo solo la seconda relazione limite in (2.4). La prima si ottiene con un ragionamento molto simile a quello fatto per trovare la decomposizione di Wold.

Proiettando \mathbf{z} sui termini dello sviluppo in somma diretta ortogonale (2.2) si ottiene

$$\mathbf{y}(t) = \sum_{-\infty}^t \mathbf{x}_n + \mathbf{x}_{-\infty} \quad , \quad (2.5)$$

dove $\mathbf{x}_n := \hat{E}[\mathbf{z} | \mathbf{E}_n]$, $\mathbf{x}_{-\infty} := \hat{E}[\mathbf{z} | \mathbf{H}_{-\infty}]$ e la serie è convergente in media quadratica per ogni t fissato. Per mostrare che $\mathbf{y}(t)$ converge, è necessario mostrare che converge la serie $\sum_{-\infty}^{+\infty} \mathbf{x}_n$, i cui addendi, $\{\mathbf{x}_n\}$, sono variabili ortogonali.

Come dimostrato in Appendice B, una serie di questo tipo converge se e solo se la somma delle varianze $\sum_{-\infty}^{+\infty} \|\mathbf{x}_n\|^2$ converge. Questa condizione è d'altra parte garantita dalla

$$\sum_{-\infty}^t \|\mathbf{x}_n\|^2 = \left\| \sum_{-\infty}^t \mathbf{x}_n \right\|^2 \leq \|\mathbf{z}\|^2 \quad , \quad \forall t \quad .$$

Il limite, $\mathbf{y}(\infty)$, che appartiene ad \mathbf{H}_{∞} per costruzione è tale per cui $\mathbf{z} - \mathbf{y}(\infty) \perp \mathbf{H}_{\infty}$, come si vede passando al limite per $t \rightarrow \infty$ sulla relazione $\langle \mathbf{z} - \mathbf{y}(t), \mathbf{h}(t) \rangle = 0$ valida $\forall \mathbf{h}(t) \in \mathbf{H}_t$ e per ogni t . Pertanto $\mathbf{y}(\infty) = \hat{E}[\mathbf{z} | \mathbf{H}_{\infty}]$. \square

Come applicazione di questo risultato dimostreremo la convergenza dell'algoritmo di Bauer illustrato alla fine del paragrafo precedente.

Dimostrazione del Teorema 4.3. Supporremo $\{\mathbf{y}(t)\}$ di rango pieno. Si usa il procedimento di ortogonalizzazione di Gram-Schmidt sulla successione vettoriale $\mathbf{y}(t-n), \dots, \mathbf{y}(t)$, partendo da $\mathbf{y}(t-n)$, generando ricorsivamente dei vettori ortogonali m -dimensionali che denoteremo $\boldsymbol{\varepsilon}_0(t-n), \dots, \boldsymbol{\varepsilon}_n(t)$ rispettivamente. Com'è ovvio, il procedimento corrisponde alla fattorizzazione di Cholesky (a blocchi) della sottomatrice principale di dimensione $(n+1)m \times (n+1)m$, Σ_n , di Σ e conduce ad una rappresentazione in cui interviene la corrispondente sottomatrice triangolare del fattore di Cholesky L , di Σ ,

$$\begin{bmatrix} \mathbf{y}(t-n) \\ \vdots \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} L_{00} & & 0 \\ & \ddots & \\ & & L_{nn} \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}_0(t-n) \\ \vdots \\ \boldsymbol{\varepsilon}_n(t) \end{bmatrix} . \quad (2.6)$$

Ciascun $\varepsilon_n(t)$ si ottiene normalizzando a varianza identità l'innovazione a memoria n di $\mathbf{y}(t)$, definita come

$$\mathbf{e}_n(t) := \mathbf{y}(t) - \hat{E}[\mathbf{y}(t) | \mathbf{H}_{[t-n,t]}] \quad , \quad (2.7)$$

che è calcolata rispetto ad un passato di n campioni precedenti $\mathbf{y}(t)$:

$$\mathbf{H}_{[t-n,t]} := \text{span} \{ \mathbf{y}(t-n), \dots, \mathbf{y}(t-1) \} \quad . \quad (2.8)$$

Al crescere di n i sottospazi (2.8) formano una successione crescente che genera, per $n \rightarrow \infty$, proprio l'intero passato $\mathbf{H}_t(\mathbf{y})$ di $\{\mathbf{y}(t)\}$.

Dal teorema di convergenza delle martingale si ricava allora

$$\lim_{n \rightarrow \infty} \mathbf{e}_n(t) = \mathbf{e}(t) \quad ,$$

dove $\mathbf{e}(t) = \tilde{\mathbf{y}}(t | t-1)$ è l'innovazione non normalizzata del processo. Svolgendo l'ultimo blocco riga nella (2.6), è immediato verificare che il coefficiente matriciale $L_{n,n-j}$ che moltiplica $\mathbf{e}_n(t-j)$, $j = 0, 1, \dots, n$, è dato dalla classica formula

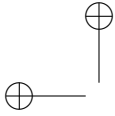
$$L_{n,n-j} = E[\mathbf{y}(t) | \mathbf{e}_{n-j}(t-j)'] (\Lambda'_{n-j})^{-1/2} \quad , \quad (2.9)$$

dove $\Lambda_{n-j}^{1/2} := L_{n-j,n-j}$ è il fattore di Cholesky della varianza, Λ_{n-j} , di $e_{n-j}(t-j)$. Mantenendo t e j fissi e facendo tendere n all'infinito, il secondo membro in (2.9) converge a $E[\mathbf{y}(t) \mathbf{e}(t-j)'] (\Lambda')^{-1/2}$ e pertanto

$$\lim_{n \rightarrow \infty} L_{n,n-j} = W(j) (\Lambda')^{-1/2} \quad ,$$

in base all'espressione (1.25) della risposta impulsiva, W , del fattore a fase minima.

□



CAPITOLO 6

FILTRI DI WIENER-KOLMOGOROV

Riprendiamo in questo capitolo il problema di calcolare lo stimatore lineare a minima varianza di un processo messaggio $\{\mathbf{x}(t)\}$ sulla base dell'osservazione di una traiettoria passata del segnale in ricezione $\{\mathbf{y}(t)\}$.

1. Soluzione mediante la rappresentazione d'innovazione

Supponiamo per il momento, per chiarezza di esposizione, che i segnali in gioco siano scalari. Come abbiamo già visto, la soluzione dell'equazione di Wiener-Hopf (2.12) è semplicissima se $\{\mathbf{y}(t)\}$ è un processo ortonormale. Infatti in questo caso si ha $\sigma_{\mathbf{y}}(t-s) = \delta(t-s)$ e la risposta impulsiva del filtro è

$$h(t) = \begin{cases} \sigma_{\mathbf{xy}}(t) & \text{per } t \geq 0 \\ 0 & \text{per } t < 0 \end{cases} . \quad (1.1)$$

Dalla (1.1) scende allora che la funzione di trasferimento del filtro di Wiener (che deve essere *causale* dato che $\hat{\mathbf{x}}(t|t)$ è funzione solo del passato \mathbf{y}^t) si ottiene come z -trasformata del “tratto causale” di $\sigma_{\mathbf{xy}}(t)$, ovvero

$$H(z) = \sum_0^{+\infty} \sigma_{\mathbf{xy}}(t) z^{-t} := [S_{\mathbf{xy}}(z)]_+ . \quad (1.2)$$

dove l'espressione $[F(z)]_+$ indica la *parte causale* di $F(z)$. (Per maggiori dettagli sulla definizione di parte causale si consulti l'Appendice C.)

Supponiamo ora che $\{\mathbf{y}(t)\}$ sia un processo *puramente non deterministico*. Sappiamo allora che $\{\mathbf{y}(t)\}$ può essere ortonormalizzato per mezzo di una operazione di “sbiancamento” lineare e *causale* che fornisce il suo processo delle innovazioni $\{\boldsymbol{\varepsilon}(t)\}$. L'uguaglianza dei due spazi $\mathbf{H}_t(\mathbf{y})$ e $\mathbf{H}_t(\boldsymbol{\varepsilon})$ suggerisce che $\boldsymbol{\varepsilon}^t$ (che contiene la “stessa informazione” di \mathbf{y}^t) possa essere usata al posto di \mathbf{y}^t nel calcolo del filtro. Procediamo così in due tempi:

- Determiniamo lo stimatore come se le osservazioni fossero effettivamente il processo d'innovazione $\{\boldsymbol{\varepsilon}(t)\}$ anziché il processo $\{\mathbf{y}(t)\}$.

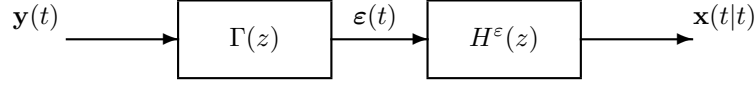


FIGURA 1.1. Struttura in cascata del filtro di Wiener.

- Combiniamo lo stimatore così ottenuto con il filtro sbiancante che dà le innovazioni $\{\varepsilon(t)\}$ in funzione del passato di $\{\mathbf{y}(t)\}$, ovvero

$$\varepsilon(t) = \sum_{-\infty}^t g(t-k) \mathbf{y}(k) \quad . \quad (1.3)$$

Detta $H^\varepsilon(z)$ la funzione di trasferimento del filtro di Wiener basato sulle innovazioni, la funzione di trasferimento effettiva, $H(z)$, si troverà allora combinando in cascata $H^\varepsilon(z)$ e il filtro sbiancante (1.3), come suggerito dallo schema a blocchi seguente

Questo schema a blocchi permette di scrivere immediatamente la funzione di trasferimento dello stimatore come

$$H(z) = H^\varepsilon(z) \Gamma(z) = [S_{\mathbf{x}\varepsilon}(z)]_+ 1/L(z) \quad , \quad (1.4)$$

dato che per la (1.2) la funzione di trasferimento $H^\varepsilon(z)$ coincide con la parte causale di $S_{\mathbf{x}\varepsilon}(z)$.

Rimane quindi solo da determinare $S_{\mathbf{x}\varepsilon}(z)$, note che siano $S_{\mathbf{y}}(z)$ ed $S_{\mathbf{x}\mathbf{y}}(z)$. Per fare questo notiamo che

$$\sigma_{\mathbf{x}\mathbf{y}}(t) = E[\mathbf{x}(t) \mathbf{y}(0)] = E\left[\mathbf{x}(t) \sum_{-\infty}^0 \ell(-k) \varepsilon(k)\right] = \sum_0^{\infty} \ell(k) \sigma_{\mathbf{x}\varepsilon}(t+k)$$

e perciò, prendendo la trasformata zeta di $\sigma_{\mathbf{x}\mathbf{y}}$, si trova

$$\begin{aligned} S_{\mathbf{x}\mathbf{y}}(z) &= \sum_{-\infty}^{+\infty} \sigma_{\mathbf{x}\mathbf{y}}(t) z^{-t} \\ &= \sum_{t=-\infty}^{+\infty} \sum_{i=-\infty}^{+\infty} \ell(i) z^i \sigma_{\mathbf{x}\varepsilon}(t+i) z^{-(t+i)} = L(z^{-1}) S_{\mathbf{x}\varepsilon}(z) \quad . \end{aligned}$$

Si ottiene quindi

$$S_{\mathbf{x}\varepsilon}(z) = S_{\mathbf{x}\mathbf{y}}(z)/L(z^{-1}) \quad . \quad (1.5)$$

Sostituendo (1.5) in (1.4) si ricava la celebre espressione della funzione di trasferimento del *filtro di Wiener*

$$H(z) = \left[\frac{S_{\mathbf{x}\mathbf{y}}(z)}{L(z^{-1})} \right]_+ \frac{1}{L(z)} \quad , \quad (1.6)$$

La formula (1.6) richiede la *fattorizzazione spettrale* di $S_{\mathbf{y}}(z)$ (per procurarsi $L(z)$) e il calcolo della parte causale di $S_{\mathbf{x}\varepsilon}(z) = S_{\mathbf{x}\mathbf{y}}(z)/L(z^{-1})$.

Calcolo della Parte Causale. Sia $F(z)$ la trasformata zeta di un segnale f scalare a energia finita. Supporremo che $F(z)$ ammetta sviluppo in serie di Laurent $F(z) = \sum_{-\infty}^{+\infty} f(k) z^{-k}$, convergente in una corona circolare $r < |z| < R$ contenente la circonferenza unità $\{|z| = 1\}$.

Ricordiamo che la *parte causale*, $[F(z)]_+$, di $F(z)$, è semplicemente la parte dello sviluppo di Laurent a potenze di z *negative o nulle*. Propriamente bisognerebbe dire che $[F(z)]_+$ è la trasformata zeta della parte causale di f , cioè della proiezione ortogonale di f sul sottospazio ℓ^{2+} dei segnali causali a energia finita. Mostriamo qui che la parte causale di una *funzione razionale*, trasformata zeta di un segnale a energia finita, si può calcolare prendendo lo sviluppo in frazioni parziali relativo ai poli di $F(z)$ *all'interno del cerchio unità*.

Ogni funzione razionale di questo tipo si può infatti sviluppare in frazioni parziali,

$$F(z) = \sum_{|p_k| < 1} F_{p_k}(z) + \sum_{|p_k| > 1} F_{p_k}(z) + F_\infty(z) + C \quad , \quad (1.7)$$

dove ciascun addendo $F_{p_k}(z)$ è la cosiddetta *parte principale* (o *caratteristica*) di $F(z)$ relativa al polo p_k , di molteplicità ν_k , data da un'espressione del tipo

$$F_{p_k}(z) = \sum_{j=1}^{\nu_k} \frac{R_{k,j}}{(z - p_k)^j} \quad (1.8)$$

che si riduce a $R_k/(z - p_k)$ se p_k è un polo semplice al finito. Il termine $F_\infty(z)$ è un polinomio in z (la parte principale relativa a $p = \infty$) e C una costante.

PROBLEMA 6.1. *Si consideri lo sviluppo di Laurent della caratteristica (1.8) con $|p_k| \neq 1$, relativo ad un intorno anulare del cerchio unità $\{|z| = 1\}$. Dimostrare che lo sviluppo*

- *contiene solo potenze negative o nulle di z se e solo se $|p_k| < 1$;*
- *contiene solo potenze positive o nulle di z se e solo se $|p_k| > 1$ (incluso $p_k = \infty$).*

Soluzione nel caso di poli semplici: *Ad esempio, se $|a| > 1$ lo sviluppo di Laurent di $1/(z - a)$ attorno a $\{|z| = 1\}$ è $-1/a \sum_0^\infty a^{-t} z^t$, che è convergente in $|z| < |a|$. Viceversa, se $|a| < 1$, lo sviluppo di Laurent di $1/(z - a)$ attorno a $\{|z| = 1\}$ è $z^{-1} \sum_{t=0}^\infty a^t z^{-t}$, convergente in $|z| > |a|$. \diamond*

Ne deriva che la parte causale di $F(z)$ è, *a meno di una costante additiva*¹, data dalla somma delle parti principali di $F(z)$ relative ai poli in $\{|z| < 1\}$.

Un metodo per determinare la costante additiva, diciamola C_0 , è ad esempio quello di imporre che la parte *strettamente anticausale* di $F(z)$

$$[[F(z)]]_- := F(z) - [F(z)]_+ = F(z) - \left[\sum_{|p_k| < 1} F_{p_k}(z) + C_0 \right]$$

sia nulla nell'origine, $z = 0$.

Nella sezione 3 daremo un metodo generale per il calcolo della parte causale, basato sulla soluzione di equazioni polinomiali.

Vediamo come si usano le formule per il calcolo del filtro in un semplice esempio.

¹Che ovviamente non è necessariamente uguale a C in (1.7).

ESEMPIO 6.1 (Filtro di Wiener con rumore bianco additivo). Vogliamo calcolare lo stimatore lineare a minima varianza di un processo messaggio $\{\mathbf{x}(t)\}$ sulla base dell'osservazione di una traiettoria passata del segnale in ricezione $\{\mathbf{y}(t)\}$. Tutti i processi sono stazionari, p.n.d. e a media nulla.

Supponiamo che le osservazioni siano affette da rumore bianco additivo

$$\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{n}(t) \quad (1.9)$$

e che $\{\mathbf{x}(t)\}$ ed $\{\mathbf{n}(t)\}$ siano *scorrelati*. Si ha allora

$$\begin{aligned} S_{\mathbf{y}}(z) &= S_{\mathbf{x}}(z) + S_{\mathbf{n}}(z) \\ S_{\mathbf{xy}}(z) &= S_{\mathbf{x}}(z) \end{aligned}$$

dove $S_{\mathbf{n}}(z) = \sigma^2$. Il problema è di costruire un filtro lineare che separi il “messaggio” $\{\mathbf{x}(t)\}$ dal rumore, in modo tale che la varianza dell'errore di ricostruzione sia la minima possibile.

Introduciamo il fattore spettrale canonico, $L(z)$, di $S_{\mathbf{y}}(z)$, che deve soddisfare alla condizione

$$L(z) L(z^{-1}) = S_{\mathbf{x}}(z) + \sigma^2 \quad ,$$

dalla quale si ricava

$$S_{\mathbf{x}\epsilon}(z) = \frac{S_{\mathbf{x}}(z)}{L(z^{-1})} = L(z) - \frac{\sigma^2}{L(z^{-1})} \quad ,$$

per cui la funzione di trasferimento del filtro è

$$H(z) = \frac{1}{L(z)} \left[L(z) - \frac{\sigma^2}{L(z^{-1})} \right]_+ = 1 - \frac{1}{L(z)} \left[\frac{\sigma^2}{L(z^{-1})} \right]_+ \quad .$$

Chiaramente la parte causale di $1/L(z^{-1})$ è semplicemente il coefficiente di grado zero nello sviluppo in serie di Laurent

$$\left[1/L(z^{-1}) \right]_+ = \left[\Gamma(z^{-1}) \right]_+ = g(0) = 1/\ell(0)$$

(dove i simboli $\Gamma(z)$ e $g(t)$ si riferiscono come al solito al filtro sbiancante) e pertanto si ha

$$H(z) = 1 - \Gamma(z) \frac{\sigma^2}{\ell(0)} \quad . \quad (1.10)$$

Il filtro toglie quindi alle osservazioni $\{\mathbf{y}(t)\}$ il rumore bianco

$$\hat{\mathbf{n}}(t) := \frac{\sigma^2}{\ell(0)} \boldsymbol{\varepsilon}(t) = \frac{\sigma^2}{\lambda^2} \mathbf{e}(t)$$

($\mathbf{e}(t)$ è l'innovazione non normalizzata di varianza $\lambda^2 = \ell^2(0)$), che coincide con la stima di $\mathbf{n}(t)$ in base ai dati \mathbf{y}^t (si veda la Figura 1.2 qui sotto).

Notiamo anche che la stabilità del filtro (1.10) è assicurata dalla stabilità del filtro sbiancante $\Gamma(z)$. In effetti $\Gamma(z)$ non può avere poli in $\{|z| = 1\}$, essendo $S_{\mathbf{y}}(e^{j\theta}) \geq \sigma^2 > 0$ per tutti i θ . \diamond

Quanto detto in questo paragrafo è manifestamente di validità assai generale, in particolare assolutamente indipendente dall'ipotesi di *razionalità* degli spettri.

Notiamo però che, dal punto di vista del calcolo pratico, esistono metodi generali per eseguire le operazioni di fattorizzazione spettrale di $S_{\mathbf{y}}(z)$ e il calcolo della parte causale di $S_{\mathbf{x}\epsilon}(z) = S_{\mathbf{xy}}(z)/L(z^{-1})$, solo se si ha a che fare con funzioni

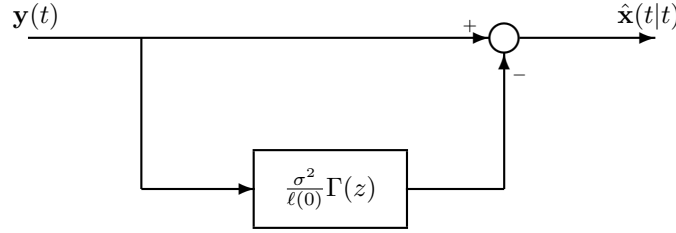
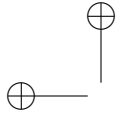


FIGURA 1.2. Schema a blocchi del filtro.

razionali. Se ne deduce (ed è bene tener presente questo fatto fin d'ora) che tutta la teoria di Wiener-Kolmogorov è di uso concreto solo se si manipolano *segnali a spettro razionale*.

Caso di processi vettoriali. La formula (1.6) è stata derivata con riferimento a processi scalari. Se i segnali in gioco sono vettoriali (e il processo $\{\mathbf{y}(t)\}$ è di rango pieno) la (1.6) va sostituita dalla

$$H(z) = [S_{\mathbf{xy}}(z) L^{-\top}(z^{-1})]_+ L^{-1}(z) \quad . \quad (1.11)$$

Per ricavare la (1.11) basta ripercorrere il ragionamento fatto nel caso scalare e ricavare $S_{\mathbf{x}\epsilon}$ dalla relazione

$$S_{\mathbf{xy}}(z) = \sum_{t=-\infty}^{+\infty} \left(\sum_{k=0}^{+\infty} \Sigma_{\mathbf{x}\epsilon}(t+k) L'(k) \right) z^{-t} = S_{\mathbf{x}\epsilon}(z) L'(z^{-1}) \quad .$$

PROBLEMA 6.2. *Dimostrare che $F(z) := S_{\mathbf{xy}}(z) L^{-\top}(z^{-1})$ è la trasformata zeta di una successione $\{F(t)\}_{t \in \mathbb{Z}}$ a quadrato sommabile.*

Far uso della formula che esprime la densità spettrale dell'errore di interpolazione $\tilde{\mathbf{x}}(t | \mathbb{Z})$, riscritta come $S_{\mathbf{x}}(z) - F(z) F(z^{-1})^T = S_{\tilde{\mathbf{x}}}(z)$, che è ovviamente nonnegativa e integrabile sulla circonferenza unità.

Se ne concluda che la risposta impulsiva $H^\epsilon(t) = F(t)$ per $t \geq 0$ (e 0 per $t < 0$) è a quadrato sommabile, indipendentemente da ogni ipotesi di razionalità sugli spettri. Che cosa si può invece dire della risposta impulsiva $H(t)$, del filtro di Wiener (1.11)? \diamond

Calcolo della varianza dell'errore di stima. L'espressione dello spettro dell'errore di stima si ricava notando che per $t = s + \tau, \tau \geq 0$, la covarianza dell'errore di stima si può esprimere come

$$\begin{aligned} \Sigma_{\tilde{\mathbf{x}}}(\tau) &= E [\mathbf{x}(t) - \hat{\mathbf{x}}(t | t)] [\mathbf{x}(s) - \hat{\mathbf{x}}(s | s)]' = E [\mathbf{x}(t) - \hat{\mathbf{x}}(t | t)] \mathbf{x}(s)' \\ &= \Sigma_{\mathbf{x}}(\tau) - \Sigma_{\tilde{\mathbf{x}}\mathbf{x}}(\tau) \end{aligned}$$

dove $\Sigma_{\tilde{\mathbf{x}}\mathbf{x}}(\tau)$ è la covarianza incrociata dello stimatore e del processo messaggio. Ne segue che

$$S_{\tilde{\mathbf{x}}}(z) = S_{\mathbf{x}}(z) - S_{\tilde{\mathbf{x}}\mathbf{x}}(z)$$

dove lo spettro incrociato al secondo membro può essere calcolato pensando $\hat{\mathbf{x}}(t | t)$ descritto dall'espressione simbolica $\hat{\mathbf{x}}(t | t) = H^\epsilon(z) \boldsymbol{\epsilon}(t)$ e introducendo lo spettro incrociato $S_{\boldsymbol{\epsilon}\mathbf{x}}(z) = L(z)^{-1} S_{\mathbf{xy}}(z^{-1})'$ (cf. la derivazione della (1.5) per il caso scalare). Si trova così

$$S_{\tilde{\mathbf{x}}}(z) = S_{\mathbf{x}}(z) - H^\epsilon(z) S_{\boldsymbol{\epsilon}\mathbf{x}}(z) = S_{\mathbf{x}}(z) - H^\epsilon(z) S_{\mathbf{x}\epsilon}(z^{-1})' \quad (1.12)$$

Da questa espressione scende una formula che (almeno in certi casi) permette di calcolare la varianza dell'errore di stima.

PROPOSIZIONE 6.1. *La varianza d'errore di stima $\text{Var } \tilde{\mathbf{x}}(t | t)$ è calcolabile mediante la formula*

$$\text{Var } \tilde{\mathbf{x}}(t | t) = \int_{-\pi}^{\pi} [S_{\mathbf{x}}(e^{j\omega}) - H^{\varepsilon}(e^{j\omega})H^{\varepsilon}(e^{j\omega})^*] \frac{d\omega}{2\pi}, \quad (1.13)$$

dove l'asterisco denota, come d'uso, il coniugato trasposto.

DIMOSTRAZIONE. Ricordando che

$$\text{Var } \tilde{\mathbf{x}}(t | t) = \int_{-\pi}^{\pi} S_{\tilde{\mathbf{x}}}(e^{j\omega}) \frac{d\omega}{2\pi},$$

si ottiene l'espressione

$$\text{Var } \tilde{\mathbf{x}}(t | t) = \int_{-\pi}^{\pi} [S_{\mathbf{x}}(e^{j\omega}) - H^{\varepsilon}(e^{j\omega})S_{\mathbf{x}\varepsilon}(e^{j\omega})^*] \frac{d\omega}{2\pi},$$

Ora, per due arbitrarie funzioni $\hat{f}, \hat{g} \in L_p^2(\mathbf{C}, \frac{d\omega}{2\pi})$, il teorema delle proiezioni fornisce per la proiezione ortogonale $P^{H_p^2} \hat{f}$ di \hat{f} su H_p^2 , l'identità

$$\int_{-\pi}^{\pi} P^{H_p^2} \hat{f}(e^{j\omega}) \hat{g}(e^{j\omega})^* \frac{d\omega}{2\pi} = \langle P^{H_p^2} \hat{f}, \hat{g} \rangle_{L^2(\mathbf{C}, \frac{d\omega}{2\pi})} = \langle P^{H_p^2} \hat{f}, P^{H_p^2} \hat{g} \rangle_{L^2(\mathbf{C}, \frac{d\omega}{2\pi})}.$$

Basta a questo punto ricordare che $H^{\varepsilon} = [S_{\mathbf{x}\varepsilon}]_+$ è la proiezione ortogonale, riga per riga, di $S_{\mathbf{x}\varepsilon}$ sul sottospazio (di Hardy) H_p^2 delle Fourier-trasformate delle funzioni causali a quadrato sommabile. \square

PROBLEMA 6.3. Si scomponga $\mathbf{x}(t)$ nella somma ortogonale

$$\tilde{\mathbf{x}}(t | t) = [\mathbf{x}(t) - \hat{\mathbf{x}}(t | \mathbb{Z})] + [\hat{\mathbf{x}}(t | \mathbb{Z}) - \hat{\mathbf{x}}(t | t)]$$

dove $\hat{\mathbf{x}}(t | \mathbb{Z})$ è la stima non causale (3.6), Cap. 3, basata su tutta la storia del processo $\{\mathbf{y}(t)\}$ e $\mathbf{x}(t) - \hat{\mathbf{x}}(t | \mathbb{Z}) := \tilde{\mathbf{x}}(t | \mathbb{Z})$ il relativo errore.

Questo termine di errore è ortogonale a tutto lo spazio $\mathbf{H}(\mathbf{y})$ e contribuisce in toto all'errore di stima causale $\tilde{\mathbf{x}}(t | t)$. Il termine rimanente, $\hat{\mathbf{x}}(t | \mathbb{Z}) - \hat{\mathbf{x}}(t | t)$ è la parte di $\hat{\mathbf{x}}(t | \mathbb{Z})$ che è ortogonale al passato $\mathbf{H}_t(\mathbf{y})$.

Il termine d'errore $\hat{\mathbf{x}}(t | \mathbb{Z}) - \hat{\mathbf{x}}(t | t)$ è la proiezione ortogonale di $\mathbf{x}(t)$ sullo spazio delle innovazioni future $\{\varepsilon(s), s > t\}$ ed è esprimibile come risposta del filtro strettamente anticausale

$$H_-^{\varepsilon}(z) := \left[\left[\frac{S_{\mathbf{x}\mathbf{y}}(z)}{L(z^{-1})} \right] \right]_- = \frac{S_{\mathbf{x}\mathbf{y}}(z)}{L(z^{-1})} - H^{\varepsilon}(z)$$

all'ingresso $\{\varepsilon(t)\}$ (supponiamo che i segnali in gioco siano scalari).

Difatti $\hat{\mathbf{x}}(t | \mathbb{Z}) - \hat{\mathbf{x}}(t | t) \in \mathbf{H}(\mathbf{y})$, è ortogonale a $\mathbf{H}_t(\mathbf{y}) = \mathbf{H}_t(\varepsilon)$ e quindi appartiene al futuro "stretto" $\mathbf{H}_{t+1}^+(\varepsilon) := \overline{\text{span}}\{\varepsilon(s), s > t\}$. D'altro canto, la proiezione ortogonale di $\hat{\mathbf{x}}(t | t)$ su questo sottospazio è zero e quella di $\hat{\mathbf{x}}(t | \mathbb{Z})$ coincide con la proiezione di $\mathbf{x}(t)$.

Ricavare per questa via l'espressione della varianza dell'errore di filtraggio,

$$\text{var } \tilde{\mathbf{x}}(t | t) = \int_{-\pi}^{\pi} \left[S_{\tilde{\mathbf{x}}(t|\mathbb{Z})}(e^{j\omega}) + |H_-^{\varepsilon}(e^{j\omega})|^2 \right] \frac{d\omega}{2\pi}, \quad (1.14)$$

dove $S_{\bar{\mathbf{x}}(t|\mathbb{Z})}$ è data dalla formula (3.7) del Capitolo 3 e dimostrare che questa espressione è equivalente alla (1.13). \diamond

PROBLEMA 6.4. *Mostrare che la funzione di trasferimento del predittore di Wiener a k passi in avanti, $\hat{\mathbf{y}}(t+k | t)$, del processo $\{\mathbf{y}(t)\}$ è data dalla formula*

$$H_k(z) = [z^k L(z)]_+ L^{-1}(z) \quad , \quad (1.15)$$

dove $L(z)$ è il fattore spettrale canonico di $S_{\mathbf{y}}(z)$. \diamond

2. Predittori di Wiener-Kolmogorov

L'esempio più semplice e forse più importante di filtro statistico è il *predittore*. Predittori lineari a minima varianza giocano un ruolo fondamentale in *identificazione* e in generale nella modellizzazione e codifica di segnali aleatori [67], nel controllo stocastico, oltre che naturalmente come strumento di previsione di serie temporali in econometria, idrologia ecc.

Sebbene la formula per la funzione di trasferimento del predittore di un processo stazionario p.n.d. si ricavi facilmente da quella generale (1.6) del filtro di Wiener (si veda il problema 6.4), riporteremo qui una derivazione diretta che permette di comprenderne a fondo la struttura dinamica e di ricavare anche una semplice formula per l'errore di stima. Per semplicità di notazioni tratteremo il caso scalare. Il lettore è invitato a ripercorrere il ragionamento nel caso di processi a valori vettoriali.

Supponiamo di voler costruire il predittore a minima varianza a k passi in avanti di \mathbf{y} , data la storia passata e presente del processo, \mathbf{y}^t . Usando la rappresentazione d'innovazione $\mathbf{y}(t) = L(z) \boldsymbol{\varepsilon}(t)$ si può scrivere

$$\begin{aligned} \mathbf{y}(t+k) &= \sum_{s=t+1}^{t+k} \ell(t+k-s) \boldsymbol{\varepsilon}(s) + \sum_{s=-\infty}^t \ell(t+k-s) \boldsymbol{\varepsilon}(s) \quad (2.1) \\ &= \ell(0) \boldsymbol{\varepsilon}(t+k) + \dots + \ell(k-1) \boldsymbol{\varepsilon}(t+1) + \sum_{s=-\infty}^t \ell_k(t-s) \boldsymbol{\varepsilon}(s) \end{aligned}$$

dove abbiamo posto

$$\ell_k = \begin{cases} \ell(t+k) & t \geq 0 \\ 0 & t < 0 \end{cases} \quad . \quad (2.2)$$

L'ultimo termine che compare nella (2.1) è una funzione lineare di $\boldsymbol{\varepsilon}^t$ e dato che $\{\boldsymbol{\varepsilon}(t)\}$ è l'innovazione di $\{\mathbf{y}(t)\}$ si ha

$$\sum_{s=-\infty}^t \ell_k(t-s) \boldsymbol{\varepsilon}(s) \in \mathbf{H}_t(\mathbf{y}) \quad ; \quad (2.3)$$

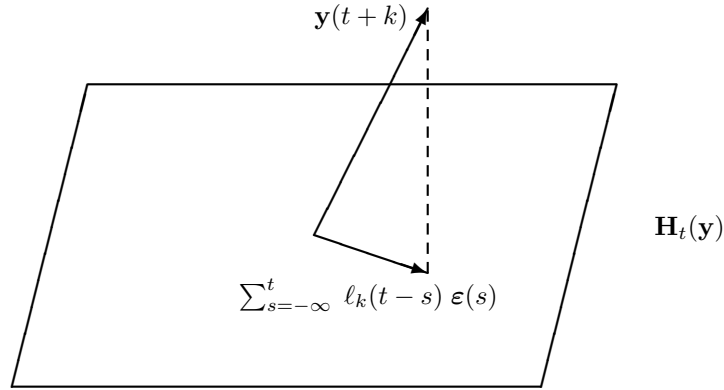
inoltre $\boldsymbol{\varepsilon}(t+i)$ è ortogonale ad $\mathbf{H}_t(\mathbf{y})$, $\forall i \geq 1$, per cui la somma

$$\tilde{\mathbf{y}}(t+k | t) := \ell(0) \boldsymbol{\varepsilon}(t+k) + \dots + \ell(k-1) \boldsymbol{\varepsilon}(t+1)$$

è ortogonale ad $\mathbf{H}_t(\mathbf{y})$. Si ha così la situazione di figura.

Segue immediatamente dal teorema delle proiezioni che *lo stimatore lineare ottimo* di $\mathbf{y}(t+k)$ in base a \mathbf{y}^t è

$$\hat{\mathbf{y}}(t+k | t) = \sum_{s=-\infty}^t \ell_k(t-s) \boldsymbol{\varepsilon}(s) \quad (2.4)$$

FIGURA 2.1. Predittore a k passi.

mentre l'errore di predizione è dato dall'espressione

$$\tilde{\mathbf{y}}(t+k|t) = \sum_0^{k-1} \ell(s) \boldsymbol{\varepsilon}(t+k-s) \quad . \quad (2.5)$$

Nell'equazione (2.4), $\hat{\mathbf{y}}(t+k|t)$ è espresso nella base ortonormale $\boldsymbol{\varepsilon}^t$ di $\mathbf{H}_t(\mathbf{y})$. Per avere un'espressione del predittore in funzione dei dati effettivamente osservati, \mathbf{y}^t , cioè un'espressione del tipo

$$\hat{\mathbf{y}}(t+k|t) = \sum_{s=-\infty}^t h(t-s) \mathbf{y}(s) \quad (2.6)$$

basta usare l'espressione di $\{\boldsymbol{\varepsilon}(t)\}$ come uscita del filtro sbiancante. Conviene, per semplificare le notazioni, usare la trasformata zeta. Introducendo la trasformata di ℓ_k ,

$$L_k(z) = \sum_0^{\infty} \ell_k(t) z^{-t} \quad , \quad (2.7)$$

la funzione di trasferimento $H(z)$ si ottiene combinando in cascata $L(z)^{-1}$ ed $L_k(z)$ come suggerito dallo schema a blocchi di figura 2.2.

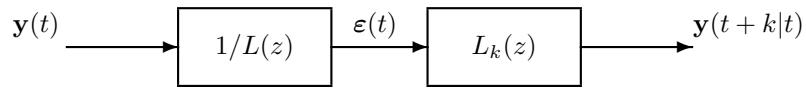


FIGURA 2.2. Schema a blocchi del Predittore di Wiener.

Si ottiene quindi il seguente risultato.

PROPOSIZIONE 6.2. *La funzione di trasferimento del predittore lineare a minima varianza a k passi in avanti del processo p.n.d. $\{\mathbf{y}(t)\}$ è data dalla formula*

$$H(z) = L_k(z) L(z)^{-1} \quad , \quad (2.8)$$

dove $L_k(z)$ è la zeta trasformata della parte causale della risposta impulsiva $\ell(\cdot)$ del fattore spettrale canonico $L(z)$ traslata a sinistra di k passi.

Il ragionamento che ci ha condotto a questo risultato serve da un lato a chiarire meglio la struttura del filtro predittore e dall'altro a trovare direttamente l'espressione esplicita (2.5) dell'errore di predizione.

Predittore di Wiener per modelli ARMA scalari. Fin qui non abbiamo fatto ipotesi di razionalità. Se si assume spettro razionale, ovvero si assume di disporre del fattore spettrale canonico nella forma (cfr. le (2.10) e (2.11))

$$L(z) = \frac{C(z)}{A(z)} = \frac{z^{-n} C(z)}{z^{-n} A(z)} = \frac{C^*(z^{-1})}{A^*(z^{-1})}$$

si possono calcolare esplicitamente $L_k(z)$ ed $H(z)$ per mezzo di semplici manipolazioni algebriche dei polinomi $A(z)$ e $C(z)$.

Usiamo la definizione (2.2) di $\ell_k(t)$ per scrivere l'ovvia relazione

$$L(z) = \ell(0) + \ell(1) z^{-1} + \dots + \ell(k-1) z^{-k+1} + z^{-k} L_k(z) \quad (2.9)$$

Dalla (2.9) segue che se $L(z)$ è razionale lo è anche $L_k(z)$ e le due funzioni hanno lo stesso denominatore. Quindi $L_k(z)$ si può scrivere come

$$L_k(z) = \frac{C_k(z)}{A(z)} \quad (2.10)$$

dove il polinomio $C_k(z)$ soddisfa la relazione

$$C(z) = \left[\sum_{i=0}^{k-1} \ell(i) z^{-i} \right] A(z) + z^{-k} C_k(z) \quad (2.11)$$

A questa relazione si può dare un aspetto familiare moltiplicando tutti i membri per z^k , ottenendo

$$z^k C(z) = \left[\sum_{i=0}^{k-1} \ell(i) z^{k-i} \right] A(z) + C_k(z)$$

da cui si vede che $C_k(z)$ (e quindi anche la parte causale $L_k(z)$) si può calcolare mediante una divisione di polinomi. Più precisamente, $C_k(z)$ si può calcolare applicando k volte l'algoritmo Euclideo delle divisioni successive alla divisione di $z^k C(z)$ per $A(z)$ e prendendo il resto, $C_k(z)$, al k -simo passo.

Alternativamente, moltiplicando la (2.11) per z^{-n} e ponendo $C_k^*(z^{-1}) := z^{-n} C_k(z)$ si ottiene una relazione che può essere interpretata come una divisione di polinomi nella variabile z^{-1} . Il termine tra parentesi quadre in (2.11) appare come il quoziente e $z^{-k} C_k(z)$ come il resto della divisione di $C^*(z^{-1})$ per $A^*(z^{-1})$.

PROPOSIZIONE 6.3. *La funzione di trasferimento del predittore a k passi di un processo a spettro razionale è data dalla formula*

$$H(z) = \frac{C_k(z)}{C(z)} = \frac{C_k^*(z^{-1})}{C^*(z^{-1})} \quad , \quad (2.12)$$

dove il polinomio $z^{-k} C_k^*(z^{-1})$ è il resto dopo k divisioni successive del polinomio $C^*(z^{-1})$ per $A^*(z^{-1})$, rispettivamente numeratore e denominatore del fattore spettrale canonico $L(z)$.

La varianza dell'errore di predizione è data dalla formula

$$\text{var } \hat{\mathbf{y}}(t+k | t) = \ell^2(0) + \dots + \ell^2(k-1) = \sum_0^{k-1} \ell^2(i) \quad . \quad (2.13)$$

Come si vede, nel caso di spettro razionale il calcolo del predittore si risolve in una operazione di *fattorizzazione spettrale* (per ottenere $L(z)$) e in una *divisione tra polinomi*.

Notiamo dalla (2.13) che l'errore di predizione aumenta al crescere di k , per cui per k grandi il predittore ha in genere scarso significato.

L'*efficienza* (o "percentuale di varianza spiegata", come viene spesso denominata) è un parametro importante per caratterizzare il comportamento del predittore. Essa è misurata dal rapporto

$$\eta = \frac{\text{var } [\hat{\mathbf{y}}(t+k | t)]}{\text{var } [\mathbf{y}(t+k)]} = 1 - \frac{\sum_0^{k-1} \ell^2(i)}{\sum_0^\infty \ell^2(i)} \quad . \quad (2.14)$$

Il termine a denominatore è ovviamente uguale a $\text{var } [\mathbf{y}(t)] = \sigma_{\mathbf{y}}(0)$ e può essere calcolato mediante la

$$\sigma_{\mathbf{y}}(0) = \int_{-\pi}^{\pi} S_{\mathbf{y}}(e^{j\omega}) \frac{d\omega}{2\pi} = \int_{-\pi}^{\pi} |L(e^{j\omega})|^2 \frac{d\omega}{2\pi} \quad . \quad (2.15)$$

Notiamo che η è tanto più vicina a 1 quanto più la "coda" $\{\ell_k(t), t \geq 0\}$ della risposta impulsiva del fattore spettrale si mantiene sensibilmente diversa da zero. Ci si aspetta quindi che il predittore lavori tanto meglio quanto più "lenta" è la dinamica di $L(z)$, e in definitiva quanto più lenta è la dinamica del processo $\{\mathbf{y}(t)\}$.

L'efficienza (2.14) può ovviamente essere riscritta anche in funzione dei campioni della risposta impulsiva, $w(t)$, della rappresentazione d'innovazione *non normalizzata* di $\{\mathbf{y}(t)\}$.

Detta λ^2 la varianza di $\{\mathbf{e}(t)\}$, si ha (cfr. ad esempio le (4.34), Cap IV e (2.13) Cap. V) $\ell^2(0) = \lambda^2 (= c_0^2$ nella rappresentazione (2.12)). Dividendo numeratore e denominatore per $\ell^2(0)$ in (2.14) si ottiene

$$\eta = 1 - \sum_0^{k-1} w(i)^2 / \sum_0^\infty w(i)^2 \quad , \quad (2.16)$$

con $w(0) = 1$. Questa espressione mette in luce il fatto che η *non dipende dalla varianza dell'ingresso* $\{\mathbf{e}(t)\}$.

Un'osservazione molto importante è che il predittore può essere realizzato *in forma ricorsiva*. Usando la notazione

$$\hat{\mathbf{y}}_k(t) := \hat{\mathbf{y}}(t+k | t)$$

ed esprimendo numeratore e denominatore della (2.12) nella forma

$$\begin{aligned} C_k^*(z^{-1}) &= \gamma_0 + \gamma_1 z^{-1} + \gamma_{n-1} z^{-n+1} \\ C^*(z^{-1}) &= c_0 + c_1 z^{-1} + \dots + c_n z^{-n} \end{aligned} \quad (2.17)$$

si ricava l'equazione alle differenze,

$$c_0 \hat{\mathbf{y}}_k(t) + c_1 \hat{\mathbf{y}}_k(t-1) + \dots + c_n \hat{\mathbf{y}}_k(t-n) = \gamma_0 \mathbf{y}(t) + \dots + \gamma_{n-1} \mathbf{y}(t-n+1), \quad (2.18)$$

che permette di calcolare ricorsivamente $\hat{\mathbf{y}}_k(t)$ una volta note le stime $\hat{\mathbf{y}}_k(t-1), \dots, \hat{\mathbf{y}}_k(t-n)$ in n istanti precedenti e le osservazioni precedenti, \mathbf{y}^t .

La (2.18) si può riscrivere in forma compatta come

$$C^*(z^{-1}) \hat{\mathbf{y}}_k(t+k | t) = C_k^*(z^{-1}) \mathbf{y}(t) \quad . \quad (2.19)$$

Notiamo che, se $S_{\mathbf{y}}(e^{j\omega}) > 0$, il polinomio $C(z)$ è (strettamente) *stabile* e i modi dell'equazione alle differenze (2.19) tendono a zero quando $t \rightarrow \infty$.

In pratica però i dati cominciano a fluire a partire da un certo istante *finito* t_0 (che supporremo nel seguito essere zero) e ci sono delle difficoltà per inizializzare l'algoritmo (2.18), dato che le condizioni iniziali $\hat{\mathbf{y}}_k(-1), \dots, \hat{\mathbf{y}}_k(-n)$, effetto dell'ipotetica "evoluzione passata in $(-\infty, -1]$ " del processo, sono *incognite*.

Usualmente si inizializza l'algoritmo prendendo *condizioni iniziali nulle*, $\hat{\mathbf{y}}_k(i) = 0$, $i = -1, \dots, -n$.

In questo modo si introduce però un *errore*. Notiamo infatti che l'espressione "esatta" del predittore che si ottiene risolvendo la (2.18), *partendo da* $t = -\infty$, è

$$\hat{\mathbf{y}}_k(t) = \sum_{-\infty}^t h(t-s) \mathbf{y}(s) = \sum_0^t h(t-s) \mathbf{y}(s) + \sum_{-\infty}^{-1} h(t-s) \mathbf{y}(s) \quad (2.20)$$

($h(\cdot)$ è la risposta impulsiva del predittore). Il secondo addendo è l'effetto dell'informazione acquisita in passato, $\{\mathbf{y}(s) ; s < 0\}$, sulla previsione all'istante t di $\mathbf{y}(t+k)$, e ovviamente è ignoto e non calcolabile dato che si osserva (una traiettoria di) $\{\mathbf{y}(t)\}$ solo da $t = 0$ in poi. Il primo addendo è la risposta del predittore a $\{\mathbf{y}(t)\}_{t \geq 0}$ partendo da condizioni iniziali nulle (è la *risposta forzata* del predittore) ed è la parte calcolabile in pratica.

Notiamo che il termine dipendente dalla storia passata di $\{\mathbf{y}(t)\}$ si può sempre scrivere come

$$\sum_{-\infty}^{-1} h(t-s) \mathbf{y}(s) = \sum_1^n \varphi_i(t) \hat{\mathbf{x}}_i \quad , \quad (2.21)$$

dove $\{\hat{\mathbf{x}}_i, i = 1, \dots, n\}$ sono condizioni iniziali opportune e le funzioni $\varphi_i(t)$ sono combinazioni lineari *dei modi* dell'equazione (2.18) (o (2.19)). Per la (stretta) stabilità $\lim_{t \rightarrow \infty} \varphi_i(t) = 0$ e il termine (2.21) diventa trascurabile per t abbastanza grandi.

ESEMPIO 6.2 (Calcolo del predittore di un processo ARMA). Si voglia calcolare il predittore di un passo del processo $\{\mathbf{y}(t)\}$ rappresentato dal modello ARMA

$$\mathbf{y}(t) + 0.7 \mathbf{y}(t-1) = \mathbf{n}(t) + 2 \mathbf{n}(t-1) \quad , \quad (2.22)$$

dove $\{\mathbf{n}(t)\}$ è rumore bianco di varianza σ^2 .

Riscrivendo il modello (2.22), in notazioni simboliche,

$$\mathbf{y}(t) = \frac{B^*(z^{-1})}{A^*(z^{-1})} \mathbf{n}(t) = \frac{1 + 2z^{-1}}{1 + 0.7z^{-1}} \mathbf{n}(t) \quad ,$$

si vede che lo zero di $B(z)$ è in $z = -2$ che *non sta dentro il cerchio unità*, quindi la (2.22) *non è una rappresentazione d'innovazione*. Notiamo ancora che

$$S_{\mathbf{y}}(z) = \sigma^2 \frac{(z+2)(z^{-1}+2)}{(z+0.7)(z^{-1}+0.7)} = 4\sigma^2 \frac{(1+1/2z)(1+1/2z^{-1})}{(1+0.7z)(1+0.7z^{-1})} \quad ,$$

per cui $L(z) = 2\sigma(1 + 1/2 z^{-1})/(1 + 0.7 z^{-1})$ e quindi la

$$\mathbf{y}(t) + 0.7 \mathbf{y}(t-1) = 2\sigma(\boldsymbol{\varepsilon}(t) + 1/2 \boldsymbol{\varepsilon}(t-1)) \quad (2.23)$$

è la rappresentazione d'innovazione di $\{\mathbf{y}(t)\}$. Indicheremo con $C(z)$ il polinomio $2\sigma(z + 1/2)$.

Per calcolare il predittore di un passo, $\hat{\mathbf{y}}(t+1 | t)$ di $\{\mathbf{y}(t)\}$ occorre procurarsi $C_1^*(z^{-1})$ partendo dalla relazione

$$C^*(z^{-1}) = \ell(0) A^*(z^{-1}) + z^{-1} C_1^*(z^{-1}) \quad ,$$

ovvero

$$2\sigma(1 + 1/2 z^{-1}) = 2\sigma(1 + 0.7 z^{-1}) + z^{-1} C_1^*(z^{-1}) \quad ,$$

per cui

$$C_1^*(z^{-1}) = -0.4 \sigma = \gamma_0 \quad .$$

La funzione di trasferimento del predittore è così

$$H(z) = \frac{-0.4}{2(1 + 1/2 z^{-1})} = \frac{-0.2}{1 + 0.5 z^{-1}}$$

e la relativa equazione alle differenze

$$\hat{\mathbf{y}}(t | t-1) + 0.5 \hat{\mathbf{y}}(t-1 | t-2) = -0.2 \mathbf{y}(t) \quad .$$

La varianza dell'errore di predizione è $\ell^2(0) = 4\sigma^2$.

Volendo calcolare l'efficienza occorre procurarsi la varianza $\sigma_{\mathbf{y}}(0)$ di $\{\mathbf{y}(t)\}$. Per fare questo si può integrare la densità di potenza del processo, ricavabile immediatamente dal modello ARMA,

$$\sigma_{\mathbf{y}}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{\mathbf{y}}(e^{j\omega}) d\omega = \frac{\sigma^2}{2\pi j} \oint_{|z|=1} \frac{B(z)}{A(z)} \frac{B(z^{-1})}{A(z^{-1})} \frac{dz}{z} \quad .$$

Qui di seguito verrà descritto un algoritmo che calcola integrali (del modulo quadrato) di funzioni razionali di questo tipo. Vedere il Teorema 2.3 e la bibliografia relativa a pag.157 del classico testo [10]. \diamond

Algoritmo di Åstrom per il calcolo di integrali quadratici di funzioni razionali.

ALGORITMO 6.1. *Si voglia calcolare l'integrale*

$$\sigma^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(e^{j\omega}) d\omega = \frac{1}{2\pi j} \oint_{|z|=1} \frac{B(z)}{A(z)} \frac{B(z^{-1})}{A(z^{-1})} \frac{dz}{z} \quad .$$

L'algoritmo per il calcolo di σ^2 è definito dalle formule ricorrenti

$$I_k - \beta_k^2 = I_{k-1}(1 - \alpha_k^2) \quad , \quad I_0 = \beta_0^2 \quad , \quad k = 1, \dots, n \quad , \quad (2.24)$$

dove I_k è l'integrale,

$$I_k = \frac{1}{2\pi j} \oint_{|z|=1} \frac{B_k(z)}{A_k(z)} \frac{B_k(z^{-1})}{A_k(z^{-1})} \frac{dz}{z} \quad (2.25)$$

e i polinomi $A_k(z)$ e $B_k(z)$ di grado k

$$A_k(z) = a_0^k z^k + a_1^k z^{k-1} + \dots + a_k^k$$

$$B_k(z) = b_0^k z^k + b_1^k z^{k-1} + \dots + b_k^k$$

si calcolano ricorsivamente attraverso le formule

$$\begin{aligned} A_{k-1}(z) &= z^{-1} [A_k(z) - \alpha_k A_k^*(z)] \quad , \quad A_n(z) = A(z) \quad , \\ B_{k-1}(z) &= z^{-1} [B_k(z) - \beta_k A_k^*(z)] \quad , \quad B_n(z) = B(z) \quad . \end{aligned} \quad (2.26)$$

($P^*(z)$ è il polinomio reciproco di $P(z)$, $P^*(z) := z^n P(z^{-1})$, dove n è il grado di $P(z)$).

I coefficienti α_k e β_k nella (2.26) (e nella (2.24)) sono dati dalle relazioni

$$\begin{aligned} \alpha_k &= a_k^k / a_0^k \\ \beta_k &= b_k^k / a_0^k \quad , \quad k = n, n-1, \dots \end{aligned} \quad (2.27)$$

e sono noti non appena sono disponibili $A_k(z)$ e $B_k(z)$.

Questo algoritmo è strettamente imparentato all'*algoritmo di Levinson* (si veda il successivo Cap. 8) e può essere generalizzato al calcolo di integrali di spettri incrociati razionali [10].

ESEMPIO 6.3 (Continuazione). Nel nostro esempio si può prendere

$$\begin{aligned} A(z) &= z + 0.7 & A^*(z) &= 1 + 0.7z \\ B(z) &= 2z + 1 & B^*(z) &= 2 + z \end{aligned} \quad .$$

Dato che il grado n è 1, si pone

$$A(z) = A_1(z) \quad , \quad B(z) = B_1(z) \quad ;$$

con $k = n$ la (2.26) diventa

$$\begin{aligned} A_0(z) &= z^{-1} [z + 0.7 - \alpha_1(1 + 0.7z)] \\ B_0(z) &= z^{-1} [2z + 1 - \beta_1(1 + 0.7z)] \quad , \end{aligned}$$

con

$$\begin{aligned} \alpha_1 &= a_1^1 / a_0^1 = 0.7 / 1 = 0.7 \\ \beta_1 &= b_1^1 / a_0^1 = 1 / 1 = 1 \end{aligned} \quad .$$

Quindi si trova

$$\begin{aligned} A_0(z) &= z^{-1} [z + 0.7 - 0.7 - 0.49z] = 0.51 \\ B_0(z) &= z^{-1} [2z + 1 - 1 - 0.7z] = 1.3 \end{aligned} \quad ,$$

per cui

$$a_0^0 = 0.51 \quad , \quad b_0^0 = 1.3$$

e

$$\beta_0 = \frac{1.3}{0.51} = 2.55 \quad .$$

Dalla (2.24) si ha così

$$I_1 - \beta_1^2 = \beta_0^2 [1 - \alpha_1^2] \quad ,$$

ovvero

$$I_1 = 1 + (2.55)^2 [1 - (0.7)^2] = 1 + 6.5 [1 - 0.49] = 1 + 3.32 = 4.32 \quad ,$$

quindi

$$\sigma_{\mathbf{y}}(0) = 4.32 \sigma^2 \quad .$$

La varianza dell'errore è $4\sigma^2$ per cui si ha una riduzione di incertezza (relativa) $4.32 - 4/4.32 = 0.32/4.32 \simeq 0.075$, del 7.5%, che è molto scarsa. \diamond

Un modo rapido per calcolare approssimativamente $\sigma_{\mathbf{y}}(0)$ è mediante la formula

$$\sigma_{\mathbf{y}}(0) = \sum_0^{\infty} \ell^2(i) \quad ,$$

troncando ad un numero finito di termini. I coefficienti $\{\ell(i)\}$ si trovano dividendo numeratore e denominatore di $L(z)$ mediante l'algoritmo euclideo.

Notiamo che questo procedimento si può applicare ad una qualunque espressione di $\{\mathbf{y}(t)\}$ del tipo

$$\mathbf{y}(t) = \sum_0^{\infty} w(k) \mathbf{n}(t-k) \quad ,$$

dove $\{\mathbf{n}(t)\}$ è rumore bianco (non necessariamente l'innovazione).

PROBLEMA 6.5. Ripercorrere la derivazione del predittore di Wiener nel caso di spettro razionale (Proposizione 6.3) usando la rappresentazione d'innovazione *non normalizzata* $\mathbf{y}(t) = W(z) \mathbf{e}(t)$, dove $W(z)$ è il fattore a fase minima normalizzato ad 1 all'infinito. Ricordare che anche il numeratore di $W(z)$ è *monico*: lo si denoti con $C^*(z^{-1}) = 1 + c_1 z^{-1} + \dots + c_n z^{-n}$. \diamond

PROBLEMA 6.6. Mostrare direttamente (cioè senza applicare la formula (1.15) ma usando solo la definizione di innovazione) che il predittore di un passo $\hat{\mathbf{y}}(t+1 | t)$ di un processo vettoriale descritto da un modello d'innovazione "puramente autoregressivo" (A.R.)

$$\mathbf{y}(t) + \sum_1^{\nu} A_k \mathbf{y}(t-k) = \mathbf{e}(t) \quad , \quad t \in \mathbb{Z} \quad ,$$

è dato dalla formula

$$\hat{\mathbf{y}}(t | t-1) = - \sum_1^{\nu} A_k \mathbf{y}(t-k) \quad . \quad (2.28)$$

\diamond

PROBLEMA 6.7. Analogo dell'esercizio precedente per il processo vettoriale $\{\mathbf{y}(t)\}$ descritto dal modello d'innovazione ARMA matriciale

$$\mathbf{y}(t) + \sum_1^{\nu} A_k \mathbf{y}(t-k) = \mathbf{e}(t) + \sum_1^{\nu} C_k \mathbf{e}(t-k) \quad , \quad t \in \mathbb{Z} \quad ,$$

nel quale, senza perdita di generalità l'indice ν è stato preso uguale per i due polinomi. Se il modello è d'innovazione il predittore è dato dalla formula ricorsiva

$$\hat{\mathbf{y}}(t | t-1) + \sum_1^{\nu} C_k \hat{\mathbf{y}}(t-k | t-k-1) = \sum_1^{\nu} (C_k - A_k) \mathbf{y}(t-k) \quad . \quad (2.29)$$

In particolare, per un processo vettoriale $\{\mathbf{y}(t)\}$ descritto dal modello d'innovazione a *pura media mobile* (M.A.)

$$\mathbf{y}(t) = \mathbf{e}(t) + \sum_1^{\nu} C_k \mathbf{e}(t-k) \quad , \quad t \in \mathbb{Z} \quad ,$$

il predittore è dato dalla formula

$$\hat{\mathbf{y}}(t | t-1) + \sum_1^{\nu} C_k \hat{\mathbf{y}}(t-k | t-k-1) = \sum_1^{\nu} C_k \mathbf{y}(t-k) \quad . \quad (2.30)$$

◇

Nel caso di processi vettoriali, le formule per il calcolo del predittore a più passi partendo da un modello ARMA a coefficienti matriciali, non sono così semplici. Come c'è da aspettarsi, la difficoltà è sostanzialmente causata dalla non commutatività del prodotto di matrici polinomiali che impedisce di eseguire la semplificazione dei denominatori di $L_k(z)$ e $L(z)$ nella formula (2.8).

Un'applicazione alle comunicazioni: il PCM differenziale. I segnali audio e video sono esempi di segnali altamente “ridondanti” nel senso che è possibile predirne abbastanza bene l'andamento futuro con predittori basati sulle statistiche “recenti” della forma d'onda osservata. In pratica essi possono venire trattati statisticamente (identificati) nel breve periodo come se fossero stazionari. Ad esempio l'identificazione della voce può essere basata su finestre temporali di 10÷20 ms e ripetuta periodicamente per riaggiustare il modello.

Per questa ragione è plausibile e naturale ricercare schemi di codifica che permettano di trasmettere solo la parte *non ridondante* del segnale, riducendo in questo modo la velocità di trasmissione, senza perdere in qualità (rapporto segnale/disturbo) in ricezione.

È intuitivamente evidente che un codificatore ideale per un segnale stazionario di spettro (razionale) noto dovrebbe trasmettere solo l'*innovazione* (ed eventualmente, “una tantum”, i coefficienti del modello ARMA d'innovazione o del relativo predittore di un passo). La ragione per cui in questo modo si ottiene un effettivo guadagno in termini di velocità di trasmissione può essere spiegata in termini generali ricordando che a parità di potenza (varianza) il segnale stazionario che ha massima entropia media per campione è il rumore bianco (Gaussiano) [35]. Per avere una spiegazione in termini più concreti occorre esaminare il sistema di codifica tenendo conto della *quantizzazione* (conversione A/D) che inevitabilmente deve precedere la fase di (codifica vera e propria e) trasmissione del segnale analogico digitalizzato. Il numero di livelli discreti di quantizzazione (cioè il numero di bits/campione) deve essere scelto in base a specifiche sul rapporto segnale/disturbo che si desidera avere in trasmissione. Quest'ultimo si può esprimere come rapporto fra la varianza del segnale $\sigma_{\mathbf{y}}^2$ e quella dell'errore di quantizzazione σ_Q^2 . Supponendo per semplicità la quantizzazione uniforme, si vede che σ_Q^2 è proporzionale al quadrato dell'ampiezza del quanto Δ

$$\sigma_Q^2 = k^2 \Delta^2 \quad .$$

(Ad esempio per segnali a distribuzione di probabilità uniforme in ampiezza si ha $k^2 = 1/12$, [67]). Inoltre l'ampiezza del quanto è a sua volta proporzionale alla deviazione standard (o valore efficace), $\sigma_{\mathbf{y}}$, del segnale quantizzato, ovvero

$$\Delta = \frac{c_{\mathbf{y}} \sigma_{\mathbf{y}}}{2^R} \quad ,$$

dove $c_{\mathbf{y}}$ è una costante che dipende dalla distribuzione di probabilità di ampiezza del segnale e 2^R è il numero di livelli di quantizzazione. Si trova così la ben nota

relazione

$$\frac{\sigma_{\mathbf{y}}^2}{\sigma_Q^2} = \frac{1}{(kc_{\mathbf{y}})^2} (2^R)^2 \quad , \quad (2.31)$$

che lega il rapporto segnale/disturbo alla velocità di trasmissione R (la cosiddetta *bit-rate*).

Questa relazione è valida nell'ipotesi di quantizzare direttamente il segnale $\{\mathbf{y}(t)\}$.

Se invece si premette un filtro sbiancante e si invia in ingresso al quantizzatore l'innovazione $\{\mathbf{e}(t)\}$ anziché il segnale, bisogna tenere conto che il rapporto segnale/disturbo per $\{\mathbf{e}(t)\}$ che corrisponde a quello desiderato (e prefissato) per il segnale di interesse $\{\mathbf{y}(t)\}$ è *in genere molto più basso*.

Questa affermazione si può giustificare partendo dalla formula

$$\sigma_{\mathbf{y}}^2 = \sigma_e^2 \left(\sum_0^{\infty} w(t)^2 \right) := \sigma_e^2 \|w\|^2 \quad , \quad (2.32)$$

dove $\sigma_e^2 = \lambda^2$ è la varianza dell'innovazione e $w(\cdot)$ è la risposta impulsiva del fattore a fase minima normalizzato. Da notare che $\|w\|^2$ è sempre maggiore di uno, dato che $w(0) = 1$. Per il segnale vocale tipicamente $\|w\|^2$ è dell'ordine delle decine, ma può arrivare in certe situazioni anche alle centinaia. Se si trasmette $\{\mathbf{e}(t)\}$, la velocità di trasmissione R_e andrà determinata dalla formula (2.31) in cui però al primo membro va il rapporto segnale/disturbo relativo all'innovazione, *che si ricava da quello prefissato per il segnale $\{\mathbf{y}(t)\}$* , ovvero

$$\frac{\sigma_e^2}{\sigma_Q^2} = \frac{1}{\|w\|^2} \frac{\sigma_{\mathbf{y}}^2}{\sigma_Q^2} \quad .$$

Ne risulta che si può codificare (quantizzare) il segnale usando solamente

$$2^{R_e} = \frac{2^R}{\|w\|}$$

livelli discreti per campione, anziché i 2^R necessari se si fosse quantizzato direttamente $\{\mathbf{y}(t)\}$.

Tentando una spiegazione intuitiva, si potrebbe dire che gli errori di predizione di un segnale "lentamente variabile", con $\|w\|^2$ grande, possono essere quantizzati con un numero inferiore di bits (o di livelli 2^R), perché hanno ampiezze significative molto più piccole dell'ampiezza del segnale originale. Gli errori di predizione di un passo possono essere visti come *variazioni* del segnale rispetto al suo andamento medio (condizionato); per questo motivo questo schema di codifica si chiama *PCM differenziale*.

PROBLEMA 6.8. Dimostrare che tra tutte le rappresentazioni mediante un filtro formante stabile di un processo scalare, del tipo $\mathbf{y}(t) = W(z) \mathbf{w}(t)$, dove $W(\infty) = 1$, quella d'innovazione non normalizzata $\mathbf{y}(t) = \hat{W}(z) \mathbf{e}(t)$ ha la risposta impulsiva a massima norma ℓ^2 .

Suggerimento: Notare che si ha sempre $\mathbf{H}_t(\mathbf{e}) \subset \mathbf{H}_t(\mathbf{w})$. Confrontare le varianze dell'errore di predizione di $\mathbf{y}(t+1)$ basato sul passato $\mathbf{H}_t(\mathbf{e})$ e quello basato sul passato del processo \mathbf{w} , $\mathbf{H}_t(\mathbf{w})$. Evidentemente la prima sarà sempre più grande (o al più uguale, ma solo in un caso) a quella del secondo. \diamond

3. L'Approccio polinomiale al filtraggio di Wiener

Come abbiamo visto, il calcolo della funzione di trasferimento del filtro di Wiener-Kolmogorov nel caso di processi a spettro razionale si riduce ad una fattorizzazione spettrale dello spettro del processo di misura e ad una operazione di calcolo di una parte causale, che nel caso del predittore si riduce ad una operazione algebrica di divisione fra polinomi. In questo paragrafo mostreremo che l'operazione di calcolo della parte causale può essere ridotta alla soluzione di un'equazione polinomiale lineare (un'equazione cosiddetta "Diofantea") che interviene in molte questioni di sintesi di filtri e di sintesi di leggi di controllo. Tratteremo problemi di filtraggio un poco più generali del tipo seguente. Sia

$$\mathbf{y}(t) = \mathbf{s}(t) + \mathbf{d}(t) \quad , \quad (3.1)$$

dove $\{\mathbf{s}(t)\}$, il segnale messaggio, e $\{\mathbf{d}(t)\}$, il disturbo, sono due processi p.n.d. a spettro razionale che supporremo scorrelati e (ancora per semplicità di trattazione) *scalari*. Si desidera trovare lo stimatore lineare a minima varianza

$$\hat{\mathbf{s}}(t+k | t) = \hat{E}[\mathbf{s}(t+k) | \mathbf{H}_t(\mathbf{y})] \quad ,$$

dove k è un intero fissato, non necessariamente positivo.

Si assumono note le rappresentazioni ARMA *d'innovazione* del segnale e del disturbo

$$\mathbf{s}(t) = \frac{N_1(z)}{D_1(z)} \boldsymbol{\varepsilon}_1(t) \quad , \quad \mathbf{d}(t) = \frac{N_2(z)}{D_2(z)} \boldsymbol{\varepsilon}_2(t) \quad , \quad (3.2)$$

in cui i polinomi D_1 e D_2 di gradi rispettivi n_1 ed n_2 sono *monici* e le coppie (N_k, D_k) , $k = 1, 2$, sono prive di fattori comuni. Tranne avviso contrario, tutti i polinomi che considereremo saranno polinomi a potenze non negative in z .

Usando la formula generale (1.6), il calcolo del filtro si può fare in due passi.

1. Fattorizzazione spettrale di $S_y(z)$

Il fattore spettrale canonico $L(z) = N(z)/D(z)$ di $S_y(z)$ si calcola partendo dall'espressione $S_y = S_s + S_d$, ovvero

$$\frac{N(z)N^*(z)}{D(z)D^*(z)} = \frac{N_1(z)N_1^*(z)}{D_1(z)D_1^*(z)} + \frac{N_2(z)N_2^*(z)}{D_2(z)D_2^*(z)} \quad , \quad (3.3)$$

dove l'asterisco indica, come al solito, il polinomio reciproco.

Da questa relazione si vede immediatamente che il denominatore di $L(z)$ deve essere

$$D(z) = D_1(z)D_2(z) \quad , \quad (3.4)$$

mentre il numeratore N si trova risolvendo il problema di fattorizzazione spettrale *polinomiale*

$$NN^* = N_1N_1^*D_2D_2^* + N_2N_2^*D_1D_1^* \quad , \quad (3.5)$$

in cui i polinomi indicati senza l'asterisco possono essere, se lo si desidera, facilmente trasformati in polinomi nella sola variabile z^{-1} . In effetti, dato che $\deg(N_1D_2) = \deg(N_2D_1) = n_1 + n_2 := n$, basta all'uopo moltiplicare ambo i membri di (3.5) per z^{-n} . Il calcolo di N può ad esempio essere fatto mediante l'algoritmo di Bauer.

2. Calcolo della parte causale di $S_{\mathbf{x}\mathbf{y}}(z)/L(z^{-1})$

Nel nostro problema $S_{\mathbf{x}\mathbf{y}}(z) = z^k S_s(z)$ ed $L(z^{-1})$ si può scrivere come N^*/D^* , per

cui

$$S_{\mathbf{xy}}(z) \frac{1}{L(z^{-1})} = z^k \frac{N_1 N_1^*}{D_1 D_1^*} \frac{D_1^* D_2^*}{N^*} = \frac{z^k N_1 N_1^* D_2^*}{D_1 N^*} . \quad (3.6)$$

La parte causale di questa funzione è una funzione *razionale* analitica in $\{|z| \geq 1\}$ (incluso il punto all'infinito) e si può quindi rappresentare come rapporto di polinomi $X(z)/A(z)$, in cui $A(z)$ è *stabile*, *monico* e $\deg X(z) \leq \deg A(z)$. In effetti, per quanto detto all'inizio del capitolo a proposito del calcolo della parte causale di funzioni razionali, si vede che:

- Se $k \geq 0$, deve essere necessariamente $A(z) = D_1(z)$, dato che gli zeri di $D_1(z)$ sono tutti e soli i poli della (3.6) situati in $\{|z| < 1\}$.

Se ne conclude che il polinomio a numeratore X dovrà essere tale da rendere la differenza

$$F_-(z) := \frac{z^k N_1 N_1^* D_2^*}{D_1 N^*} - \frac{X}{D_1}$$

una funzione razionale appartenente al complemento ortogonale, \bar{H}_0^2 , di H^2 in $L^2(\mathbb{C})$ e quindi necessariamente della forma $Y(z)/N^*(z)$ con $Y(z)$ polinomio a potenze non negative in z e almeno uno zero nell'origine. Si trova così l'*equazione Diofantea*²

$$N^* X + Y D_1 = z^k N_1 N_1^* D_2 \quad , \quad (3.7)$$

la quale è una equazione lineare nelle due incognite polinomiali $X(z)$ e $Y(z)$. La soluzione che interessa dovrà avere il grado di $X(z)$ minore o uguale a quello di D_1 , i.e., $n_X \leq n_1 = \deg D_1$. Ovviamente il grado $n_Y + n_1$ di $Y D_1$ per ogni soluzione $Y(z)$ dovrà essere uguale al massimo tra i gradi degli altri due addendi polinomiali nell'equazione (3.7), ovvero $n_Y + n_1 = \max\{n + n_X, n + n_1 + k\}$. Da questa relazione si vede che $n_Y = n + k$.

- Se $k < 0$, oltre ai poli corrispondenti agli zeri di D_1 , la (3.6) ha in $\{|z| < 1\}$ anche un polo nell'origine di molteplicità pari a $k - k_1$, dove k_1 è la molteplicità di un eventuale zero nell'origine di $N_1(z)$. In ogni caso si può prendere il denominatore della parte causale uguale ad $A(z) = z^{|k|} D(z)$, per cui la parte anticausale,

$$F_-(z) := \frac{N_1 N_1^* D_2^*}{z^{|k|} D_1 N^*} - \frac{X}{z^{|k|} D_1} \quad ,$$

si potrà rappresentare come una funzione razionale della stessa forma, $Y(z)/N^*(z)$, del caso precedente. I polinomi X e Y sono ora soluzioni dell'equazione

$$N^* X + Y z^{|k|} D_1 = N_1 N_1^* D_2 \quad , \quad (3.8)$$

di grado rispettivo $n_X \leq n_1 + k$ ed $n_Y = n - |k|$, con $Y(z)$ ancora soggetto alla condizione $Y(0) = 0$.

Per quanto riguarda la risoluzione delle equazioni Diofantee (3.7) e (3.8), riportiamo dalla letteratura (e.g. [86], [143, p. 11]) il seguente enunciato.

²Questa è una denominazione spesso usata in letteratura per equazioni polinomiali del tipo $AX + YB = C$. Alcuni autori preferiscono chiamarle *equazioni di Bezout* e c'è stata una certa diatriba tra i sostenitori delle due denominazioni. Sembra alla fine che abbiano prevalso gli estimatori di Diofanto, anche se è certo che le equazioni "Diofantee", studiate nel terzo secolo d.C. dal celebre matematico di Alessandria, hanno poco a che fare con le equazioni polinomiali che ci interessano qui ([131])

TEOREMA 6.1. *Siano A, B, C polinomi assegnati in $\mathbb{R}[z]$. L'equazione Diofantea,*

$$AX + YB = C \quad , \quad (3.9)$$

- (1) *ammette soluzioni (X, Y) in $\mathbb{R}[z]$ se e solo se il M.C.D. di (A, B) è anche un divisore di C ;*
- (2) *ammette soluzioni, qualunque sia C , se e solo se (A, B) sono coprimi.*
- (3) *Una soluzione particolare della (3.9) è data dalla*

$$X_0 = PC_0 \quad , \quad Y_0 = QC_0 \quad , \quad (3.10)$$

dove P e Q (al pari di S e T) sono polinomi in $\mathbb{R}[z]$ tali che

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} P & S \\ Q & T \end{bmatrix} = \begin{bmatrix} R & 0 \end{bmatrix} \quad ,$$

$R \in \mathbb{R}[z]$ è il M.C.D. di (A, B) (lemma 4.4) e C_0 è definito dalla fattorizzazione $C = RC_0$.

- (4) *Tutte e sole le soluzioni dell'equazione Diofantea (3.9) si ottengono sommando ad una soluzione particolare della (3.9) una qualunque soluzione dell'equazione omogenea $AX + YB = 0$. Queste ultime hanno la forma*

$$X = HB \quad , \quad Y = -AH \quad , \quad (3.11)$$

dove $H \in \mathbb{R}[z]$ è un polinomio arbitrario.

- (5) *Si assumano A e B coprimi e di gradi rispettivamente n_A ed n_B . Esiste allora una soluzione (\hat{X}, \hat{Y}) tale che*

$$\deg(\hat{X}) < n_B \quad (3.12)$$

(e, dualmente, una (\tilde{X}, \tilde{Y}) tale che $\deg \tilde{Y} < n_A$).

- (6) *Se si considera l'equazione (3.9) nel dominio Euclideo proprio dei polinomi a potenze non negative di z , la soluzione (\hat{X}, \hat{Y}) soddisfacente (3.12) è unica. Essa viene denominata soluzione di grado minimo in X . La soluzione di grado minimo \hat{X} si ottiene da una soluzione particolare (ad esempio X_0 nella (3.10)) come resto della divisione di X per B ,*

$$X = \hat{Q}B + \hat{X}$$

e scegliendo $H = -\hat{Q}$ nell'addendo soluzione dell'omogenea (3.11) (in questo modo anche $\hat{Y} = Y - A\hat{Q}$ rimane univocamente determinato).

COROLLARIO 6.1. *Se (A, B) sono coprimi e se $A(0) \neq 0$, allora esiste un'unica soluzione di (3.9) tale che*

$$\deg X \leq n_B \quad , \quad Y(0) = 0$$

nel dominio dei polinomi a potenze non negative di z .

DIMOSTRAZIONE. Tutte le soluzioni si ottengono da quella di grado minimo mediante la

$$X = \hat{X} + HB \quad , \quad Y = \hat{Y} - AH \quad .$$

Tutte e sole quelle con $\deg X \leq n_B$ corrispondono quindi ad $H = \text{costante}$ (polinomio di grado zero). Ma se $A(0) \neq 0$ esiste una sola costante H tale per cui $Y(0) = \hat{Y}(0) - A(0)H = 0$. \square

Scende allora dal Corollario 6.1 che entrambe le equazioni Diofantee (3.7) e (3.8) hanno un'unica soluzione (X, Y) tale che, rispettivamente, $\deg X \leq n_1$, $\deg X \leq n_1 + k$ e $Y(0) = 0$.

Il calcolo della parte causale e anticausale della (3.6) si riduce quindi alla soluzione di un'equazione polinomiale lineare.

Nel caso, già visto al paragrafo precedente, del calcolo del *predittore* $\hat{y}(t+k | t)$, si possono particolarizzare le formule supponendo il disturbo $d = 0$ in (3.1). Il caso di disturbo nullo si può far corrispondere ad $N_2 = 0$, $D_2 = 1$, quindi ad $N = N_1$ e in definitiva si perviene alla seguente particolarizzazione della (3.7)

$$1 \cdot X + Y D_1 = z^k N_1 \quad ,$$

la quale è null'altro se non la (2.21) in notazioni differenti. In particolare la soluzione (X, Y) con $\deg X \leq \deg D_1$ e $Y(0) = 0$ è

$$X(z) = C_k(z) \quad , \quad Y(z) = z^k \sum_{i=0}^{k-1} \ell(i) z^{-i} \quad .$$

Usando il principio di identità dei polinomi, un'equazione Diofantea può essere convertita in un ordinario sistema di equazioni lineari. In particolare, le soluzioni (X, Y) delle (3.7) e (3.8), soggette ai vincoli discussi in precedenza sul grado di $X(z)$ e sull'annullarsi del coefficiente di grado zero in $Y(z)$, possono essere calcolate risolvendo un opportuno sistema di equazioni lineari nei relativi coefficienti. Per scrivere questi sistemi in forma compatta introduciamo l'intero non negativo m uguale ad n_1 se $k \geq 0$ ed $n_1 + |k|$ se $k < 0$. Denotiamo il polinomio N^* (che gioca il ruolo di A nella forma standard (3.9)) come

$$N^*(z) = a_n z^n + \dots + a_1 z + a_0$$

e definiamo $B(z) := D_1(z)$ se $k \geq 0$ e $B(z) := z^{|k|} D_1(z)$ se $k < 0$. Formalmente $B(z)$ ha grado m e si può scrivere

$$B(z) = b_m z^m + \dots + b_1 z + b_0 \quad .$$

Il termine noto $C(z)$ è uguale, rispettivamente, a $z^k N_1 N_1^* D_2$ se $k \geq 0$ e ad $N_1 N_1^* D_2$ se $k < 0$ (per cui $\deg C(z) = n + m + k$ in entrambi i casi). Scrivendo

$$C(z) = c_{n+m+k} z^{n+m+k} + \dots + c_1 z + c_0$$

e

$$\begin{aligned} X(z) &= x_m z^m + \dots + x_1 z + x_0 \\ Y(z) &= y_{n+k} z^{n+k} + \dots + y_1 z \end{aligned}$$

si ricava dalle (3.7) o (3.8) un sistema lineare del tipo

dove $N_0(z)$ è il polinomio a zeri in $\{|z| < 1\}$ soluzione dell'equazione di fattorizzazione spettrale polinomiale,

$$N_0(z) N_0^*(z) = \sigma_x^2 N(z) N^*(z) + \sigma^2 D(z) D^*(z) \quad .$$

Usando la (4.1), la formula che definisce l'interpolatore a intervallo fisso $\hat{\mathbf{x}}(t) := \hat{\mathbf{x}}(t | \mathbb{Z})$ si può riscrivere nella forma

$$\begin{aligned} \hat{\mathbf{x}}(t) &= S_x(z) S_y(z)^{-1} \mathbf{y}(t) \\ &= \frac{N(z) N^*(z)}{D(z) D^*(z)} \frac{D(z) D^*(z)}{N_0(z) N_0^*(z)} \mathbf{y}(t) \\ &= \frac{N(z)}{N_0(z)} \frac{N^*(z)}{N_0^*(z)} \mathbf{y}(t) \quad . \end{aligned}$$

Da questa formula si vede che la stima $\hat{\mathbf{x}}(t | \mathbb{Z})$ si può calcolare in due “passate” successive, una che elabora i dati originari $\{\mathbf{y}(t)\}$ mediante il filtro *anticausale* di funzione di trasferimento $\bar{H}(z) := N^*(z)/N_0^*(z)$ partendo dall'istante finale (in teoria $t_1 = +\infty$) e progredendo *all'indietro* nel tempo, e l'altra che elabora il segnale in uscita del filtro anticausale, diciamolo $\{\mathbf{z}(t)\}$, mediante il filtro *causale* di funzione di trasferimento $H(z) = N(z)/N_0(z) = N^*(z^{-1})/N_0^*(z^{-1})$.

Scrivendo i due polinomi $N^*(z)$ ed $N_0^*(z)$ esplicitamente,

$$\begin{aligned} N^*(z) &= p_0 + p_1 z + \dots + p_n z^n \\ N_0^*(z) &= q_0 + q_1 z + \dots + q_n z^n \end{aligned}$$

e normalizzando q_0 a uno (giacché si può supporre senza perdita di generalità che q_0 sia diverso da zero), l'algoritmo di interpolazione si può così riscrivere mediante le due ricorsioni in cascata,

$$\mathbf{z}(t) + \sum_1^n q_k \mathbf{z}(t+k) = \sum_0^n p_k \mathbf{y}(t+k) \quad (4.2a)$$

$$\hat{\mathbf{x}}(t) + \sum_1^n q_k \hat{\mathbf{x}}(t-k) = \sum_0^n p_k \mathbf{z}(t-k) \quad . \quad (4.2b)$$

Si noti che la (4.2) in pratica va inizializzata ad un tempo finale finito t_1 assegnando n valori terminali per la variabile ausiliaria $\mathbf{z}(t)$. Questi valori terminali possono essere presi arbitrariamente (ad esempio tutti uguali a zero) perché i modi del sistema dinamico (4.2) hanno un andamento esponenziale *che tende a zero per* $t \rightarrow -\infty$ giacché gli zeri di $N_0(z)$ hanno tutti modulo strettamente maggiore di uno. Quando l'effetto del “transitorio” dovuto alle condizioni terminali “sbagliate” si è esaurito, la soluzione di (4.2) è un processo stazionario.

Ovviamente un ragionamento perfettamente duale vale per il filtro causale (4.2) che in pratica deve essere inizializzato all'istante finito t_0 , fornendo n valori iniziali arbitrari per $\hat{\mathbf{x}}(t)$. In (4.2) i modi del sistema tendono a zero esponenzialmente quando $t \rightarrow +\infty$ e anche in questo caso la stima a varianza minima $\hat{\mathbf{x}}(t)$ si ottiene solo dopo qualche decina di campioni, a transitorio esaurito. È interessante qui mettere in evidenza che la stabilità esponenziale dei modi dei due schemi ricorsivi (4.2) e (4.2) è proprio la caratteristica fondamentale che permette l'implementazione pratica dell'algoritmo. Se non si avesse tendenza a zero dei modi, l'effetto degli inevitabili errori di arrotondamento numerico ecc. sovrapposti alle variabili in gioco

tenderebbe viceversa ad *esaltarsi nel tempo*. Questo fatto renderebbe lo schema numericamente inutilizzabile.

Un'alternativa del filtro interpolatore a “due passate” che abbiamo appena illustrato è lo schema additivo “alla Mayne-Fraser” [102] che si ottiene decomponendo la funzione di trasferimento dell'interpolatore in una parte causale e una anticausale.

Predizione all'indietro. Supponiamo che $\{\mathbf{y}(t)\}$ sia un processo m -dimensionale a rango pieno e a spettro razionale $S(z)$. L'*innovazione (normalizzata) all'indietro* di $\{\mathbf{y}(t)\}$ è un processo bianco ortonormale $\{\bar{\boldsymbol{\varepsilon}}(t)\}$ tale per cui

$$\mathbf{H}_t^+(\bar{\boldsymbol{\varepsilon}}) = \mathbf{H}_t^+(\mathbf{y}) \quad , \quad \forall t \quad , \quad (4.3)$$

in altre parole, $\{\bar{\boldsymbol{\varepsilon}}(t)\}$ è *anticausalmente equivalente* a $\{\mathbf{y}(t)\}$. La *rappresentazione d'innovazione (normalizzata) all'indietro* di $\mathbf{y}(t)$ che si ottiene dalla (4.3),⁴

$$\mathbf{y}(t) = \sum_t^{+\infty} \bar{L}(t-s) \bar{\boldsymbol{\varepsilon}}(s) := \bar{L}(z) \boldsymbol{\varepsilon}(t) \quad ,$$

si trova calcolando il fattore spettrale $\bar{L}(z)$ di dimensione $m \times m$ di $S(z)$, le cui righe appartengono allo spazio di Hardy coniugato \bar{H}_m^2 (e quindi sono funzioni di z analitiche in $\{|z| \leq 1\}$) ed è tale che $\det(\bar{L}(z))$ non si annulla mai all'interno del cerchio unita, $\{|z| < 1\}$. Questo fattore è essenzialmente unico (a meno di moltiplicazione a destra per una matrice ortogonale costante) e si potrebbe chiamare fattore a fase minima *coniugato* dello spettro $S(z)$.

Da notare che $\bar{L}(0) = [\bar{L}(z)]_{z=0}$ è non singolare e può essere reso a struttura triangolare inferiore per mezzo di una opportuna normalizzazione ortogonale. Il fattore a fase minima coniugato, normalizzato in modo tale che $\bar{L}(0)$ sia di struttura triangolare inferiore nell'origine verrà chiamato *fattore anticanonico*.

PROPOSIZIONE 6.4. Sia $\tilde{L}(z)$ il fattore spettrale canonico dello spettro trasposto, ovvero

$$S^T(z) = \tilde{L}(z) \tilde{L}(1/z)^T \quad ;$$

allora il fattore spettrale anticanonico di $S(z)$ è dato dalla formula

$$\bar{L}(z) = \tilde{L}(1/z) \quad .$$

DIMOSTRAZIONE. La proposizione scende immediatamente dall'unicità del fattore spettrale canonico. Infatti $S^T(z) = \bar{L}(1/z) \bar{L}(z)^T$ e $\bar{L}(1/z)$ è un fattore spettrale canonico di $S^T(z)$. \square

L'enunciato della Proposizione 6.4 vale in condizioni più generali di quelle ipotizzate qui. In particolare, è vero per processi di rango arbitrario p . Naturalmente nel caso scalare si ha semplicemente $\bar{L}(z) = \tilde{L}(1/z)$.

PROBLEMA 6.9. Dimostrare che per processi p.n.d. di rango pieno la funzione di trasferimento $\bar{H}(z)$ dello stimatore anticausale $\hat{\mathbf{x}}(t) := \hat{E}[\mathbf{x}(t) | \mathbf{H}_t^+(\mathbf{y})]$ è data dalla formula

$$\bar{H}(z) = [S_{\mathbf{xy}}(z) \bar{L}(z^{-1})^{-\top}]_- \bar{L}(z)^{-1} \quad ,$$

⁴Da notare il solito abuso di notazioni: dovremmo scrivere $\hat{\bar{L}}$ per indicare la zeta trasformata di \bar{L} .

dove $[F(z)]_-$ indica la *parte anticausale* di F , ovvero la parte dello sviluppo di Laurent di $F(z)$ relativa ad un intorno di $\{|z| = 1\}$, a potenze *positive o nulle* in z .
 \diamond

PROBLEMA 6.10. Mostrare che per un processo scalare gli errori di predizione $\tilde{\mathbf{y}}(t+k | t)$ e $\bar{\mathbf{y}}(t-k | t) := \mathbf{y}(t-k) - \hat{E}[\mathbf{y}(t-k) | \mathbf{H}_t^+(\mathbf{y})]$ ($k > 0$) hanno la stessa varianza. In particolare, le innovazioni (non normalizzate) in avanti e all'indietro, $\{\mathbf{e}(t)\}$, $\{\bar{\mathbf{e}}(t)\}$, hanno uguale varianza, $\lambda^2 = \bar{\lambda}^2$.

Questa simmetria si può far risalire al fatto che un processo scalare è *reversibile*, nel senso che $E\mathbf{y}(t+\tau)\mathbf{y}(t) = \sigma(\tau) = \sigma(-\tau) = E\mathbf{y}(t-\tau)\mathbf{y}(t)$.
 \diamond

5. Il processo coniugato e gli zeri della densità spettrale

Sia $\{\mathbf{y}(t)\}$ un processo scalare di densità spettrale $S(z)$ con cui si modellano certe osservazioni fatte sull'intervallo infinito $(-\infty, +\infty)$. Supponiamo che, per un qualunque motivo, il dato $\mathbf{y}(t)$ relativo ad un certo istante fissato, t , sia mancante. Ci si richiede di *calcolare lo stimatore lineare a minima varianza, $\hat{\mathbf{y}}(t)$ di $\mathbf{y}(t)$, basato sulle rimanenti misure $\{\mathbf{y}(s) ; s \neq t\}$.*

Introduciamo allo scopo lo spazio dei dati⁵

$$\tilde{\mathbf{H}}_t(\mathbf{y}) := \overline{\text{span}}\{\mathbf{y}(s) ; s \neq t\}$$

e procediamo euristicamente, supponendo che lo stimatore abbia la forma

$$\hat{\mathbf{y}}(t) = \sum_{s=-\infty}^{+\infty} h(t-s) \mathbf{y}(s) \quad , \quad h(0) = 0 \quad ,$$

dove la somma è convergente in media quadratica. In questa ipotesi, l'errore di stima si scrive

$$\begin{aligned} \tilde{\mathbf{y}}(t) = \mathbf{y}(t) - \hat{\mathbf{y}}(t) &= \sum_{-\infty}^{+\infty} [\delta(t-s) - h(t-s)] \mathbf{y}(s) \\ &:= \sum_{-\infty}^{+\infty} g(t-s) \mathbf{y}(s) \quad , \end{aligned}$$

con $g(\cdot) := \delta(\cdot) - h(\cdot)$, $\delta(\cdot)$ essendo il delta di Kronecker centrato in $t = 0$. Imponendo il principio di ortogonalità $E\tilde{\mathbf{y}}(t)\mathbf{y}(s) = 0$ per tutti gli $s \neq t$, e ponendo $a^2 := E\tilde{\mathbf{y}}(t)\mathbf{y}(t) = E(\tilde{\mathbf{y}}(t))^2$ (la varianza dell'errore di interpolazione), si trova dopo alcuni semplici passaggi l'equazione di convoluzione

$$\sum_{-\infty}^{+\infty} g(t-\tau) \sigma_y(\tau) = a^2 \delta(t) \quad , \quad (5.1)$$

dalla quale, passando alle zeta-trasformate, si ottiene formalmente la funzione di trasferimento da \mathbf{y} a $\tilde{\mathbf{y}}(t)$,

$$G(z) = \frac{a^2}{S(z)} \quad . \quad (5.2)$$

⁵E' ovvio che per la stazionarietà del processo \mathbf{y} ci si potrebbe riferire all'istante fissato $t = 0$ in tutti i passaggi di questa sezione. Continueremo a fare riferimento ad un istante t generico solo per chiarezza di notazioni.

A questo punto si può cercare di calcolare la varianza di $\tilde{\mathbf{y}}(t)$. Si trova

$$\begin{aligned} a^2 = E\left(\tilde{\mathbf{y}}(t)\right)^2 &= \int_{-\pi}^{\pi} G(e^{j\omega}) G(e^{-j\omega}) S(e^{j\omega}) \frac{d\omega}{2\pi} \\ &= a^4 \int_{-\pi}^{\pi} \frac{1}{S(e^{j\omega})} \frac{d\omega}{2\pi} . \end{aligned}$$

Da notare che l'ultimo integrale esiste sempre (può eventualmente essere uguale a $+\infty$) perché S è una funzione non negativa. *Supponendo* $a^2 \neq 0$, si trova la formula

$$\text{var } \tilde{\mathbf{y}}(t) = a^2 = \frac{1}{\int_{-\pi}^{\pi} \frac{1}{S(e^{j\omega})} \frac{d\omega}{2\pi}} . \quad (5.3)$$

Se la varianza d'errore è uguale a zero si ha $\mathbf{y}(t) \in \dot{\mathbf{H}}_t(\mathbf{y})$ e in un certo senso si può dire che le variabili del processo $\{\mathbf{y}(t)\}$ sono una famiglia *ridondante* di generatori per lo spazio $\mathbf{H}(\mathbf{y})$. Conveniamo allora di dire che le variabili aleatorie $\{\mathbf{y}(t)\}_{t \in \mathbb{Z}}$ sono una famiglia *minima* o *linearmente indipendente* di generatori dello spazio $\mathbf{H}(\mathbf{y})$ se $\text{var } \tilde{\mathbf{y}}(t) \neq 0$ e quindi *nessuna variabile del processo, $\mathbf{y}(t)$, può essere espressa come combinazione lineare, finita o infinita, delle rimanenti*. In [126, p. 99] un processo che ha questa proprietà viene chiamato *minimo*.

Kolmogorov [82, 83], [126, p. 98-99] ha caratterizzato esattamente i processi minimi.

TEOREMA 6.2 (Kolmogorov). *Si ha $\text{var } \tilde{\mathbf{y}}(t) \neq 0$ se e solo se il processo \mathbf{y} ammette densità spettrale e*

$$\int_{-\pi}^{\pi} \frac{1}{S(e^{j\omega})} \frac{d\omega}{2\pi} < +\infty . \quad (5.4)$$

In questo caso, $\text{var } \tilde{\mathbf{y}}(t)$ è fornita dalla formula (5.3).

Da notare che se l'integrale è finito, anche il reciproco, $1/S(z)$, è uno spettro di potenza. È in effetti immediato verificare che $1/S(z)$ è proprio lo spettro del processo errore $\{\tilde{\mathbf{y}}(t)\}$. In letteratura questo processo è chiamato *processo coniugato* di $\{\mathbf{y}(t)\}$ [101].

Nel caso di *spettri razionali*, è ovvio che la condizione (5.4) si verifica se e solo se non vi sono zeri di $S(z)$ sulla circonferenza unità $\{|z| = 1\}$. Il Teorema 6.2 fornisce in questo caso una interessante interpretazione probabilistica degli zeri dello spettro $S(z)$ sulla circonferenza unità.

COROLLARIO 6.2. *Le variabili di un processo a spettro razionale, \mathbf{y} formano una famiglia minima di generatori per $\mathbf{H}(\mathbf{y})$ se e solo se $S(z)$ non ha zeri in $\{|z| = 1\}$.*

Una dimostrazione diretta del corollario si può ottenere basandosi sull'osservazione seguente: *Se ci sono zeri di $S(z)$ a modulo unitario, esistono successioni $g(\cdot)$ tali per cui*

$$\sum_{-\infty}^{+\infty} \sigma_y(t-s) g(s) = 0 , \quad \forall t \in \mathbb{Z} , \quad (5.5)$$

e quindi, in accordo con l'equazione di convoluzione (5.1), ci si trova nel caso limite in cui $a^2 = 0$. *Viceversa se una $g(\cdot)$ non identicamente nulla soddisfa la condizione (5.5), ci debbono essere zeri dello spettro sulla circonferenza unità.* Verifichiamo quest'ultimo fatto in un caso semplice.

Supponiamo che $S(z)$ sia uno spettro polinomiale (di un processo scalare a media mobile) del tipo

$$S(z) = \sum_{-N}^N \sigma(k) z^k \quad .$$

Ogni successione $\{g(t)\}$ che soddisfa la (5.5) soddisfa allora l'equazione alle differenze

$$\sigma(-N) g(t) + \dots + \sigma(0) g(t+N) + \dots + \sigma(N) g(t+2N) = 0 \quad , \quad \forall t \in \mathbb{Z} \quad , \quad (5.6)$$

che ha equazione caratteristica

$$\sigma(N) z^{2N} + \dots + \sigma(0) z^N + \dots + \sigma(-N) = 0 \quad . \quad (5.7)$$

le cui soluzioni sono le N coppie reciproche di zeri dello spettro $S(z)$. Nello spazio vettoriale $2N$ -dimensionale dalle soluzioni dell'equazione alle differenze omogenea (5.6) possiamo in realtà prendere solo le successioni $g(\cdot)$ tali per cui la somma $\sum_{-\infty}^{+\infty} g(t-s) \mathbf{y}(s)$ è convergente in media quadratica. Come mostrato nell'Appendice D questa condizione è equivalente alla

$$\sum_{t,s=-\infty}^{+\infty} g(t) \sigma(t-s) g(s) < \infty \quad . \quad (5.8)$$

Le soluzioni ammissibili g , dell'equazione (5.6) debbono quindi essere successioni (definite su \mathbb{Z}) che soddisfano alla condizione di sommabilità (5.8).

La dimostrazione del corollario 6.2 poggia allora sulla seguente proposizione.

PROPOSIZIONE 6.5. *Le soluzioni ammissibili (e non identicamente nulle) dell'equazione (5.6) possono solo essere combinazioni lineari di modi che corrispondono a zeri di $S(z)$ sulla circonferenza unità.*

DIMOSTRAZIONE. Mostriamo che se $S(z)$ non ha zeri sulla circonferenza unità, l'unica soluzione ammissibile di (5.6) è $g = 0$. In effetti, se $S(e^{j\omega}) > 0$ su $[-\pi, \pi]$, vale una disuguaglianza del tipo

$$S(e^{j\omega}) > \sigma^2, \quad \omega \in [-\pi, \pi]$$

per un opportuno $\sigma^2 > 0$. Introducendo, per $T \geq N$, la matrice di covarianza $(2T+1) \times (2T+1)$

$$\Sigma_T := \begin{bmatrix} \sigma(0) & \sigma(-1) & \dots & \sigma(-N) & 0 & \dots \\ \sigma(1) & \sigma(0) & \dots & & \sigma(-N) & 0 \\ \vdots & & & \ddots & & \sigma(-N) \\ \sigma(N) & \dots & & & & \\ 0 & \sigma(N) & & & & \\ \dots & & \sigma(N) & & \dots & \sigma(0) \end{bmatrix} := E \begin{bmatrix} \mathbf{y}(-T) \\ \vdots \\ \mathbf{y}(T) \end{bmatrix} \begin{bmatrix} \mathbf{y}(-T) \\ \vdots \\ \mathbf{y}(T) \end{bmatrix}'$$

è facile vedere (cf. l'esercizio 5.3) che questa disuguaglianza è equivalente alla

$$\Sigma_T \geq \sigma^2 I_T$$

e pertanto il troncamento all'intervallo $[-T, T]$ di una qualunque soluzione g di (5.6), diciamolo g_T , soddisfa alla disuguaglianza

$$g_T \Sigma_T g_T' \geq \|g_T\|^2$$

dove $\|\cdot\|$ è la norma Euclidea. Se g contiene modi corrispondenti a zeri che non stanno sul cerchio unitario $\|g_T\|^2 \rightarrow \infty$ per $T \rightarrow \infty$ e la condizione (5.8) non può essere soddisfatta. \square

PROBLEMA 6.11. Il lettore completi l'argomento e lo estenda al caso di un processo scalare a spettro razionale arbitrario. *Suggerimento:* si esprima $S(z)$ come rapporto di polinomi, $S(z) = \Delta(z)/D(z)$ dove $\Delta(z)$ e $D(z)$ sono spettri polinomiali. Dimostrare che la (5.5) è equivalente all'equazione simbolica $\Delta(z)g = 0$. \diamond

Le idee appena esposte si estendono al caso di processi vettoriali. In questo contesto la condizione di minimalità (5.4) si generalizza alla

$$\int_{-\pi}^{+\pi} \text{Tr} [S(e^{j\omega})^{-1}] \frac{d\omega}{2\pi} < \infty \tag{5.9}$$

che richiede che la matrice densità spettrale di \mathbf{y} debba essere genericamente invertibile. In altri termini, un'ovvia condizione necessaria affinché le variabili di un processo vettoriale \mathbf{y} siano linearmente indipendenti è che \mathbf{y} sia di rango pieno. Questo è in accordo con l'intuizione, dato che altrimenti una qualche componente \mathbf{y}_k sarebbe combinazione lineare delle rimanenti $\{\mathbf{y}_j, j \neq k\}$. Da notare che per la seconda relazione in (0.6), Appendice A, la norma di Frobenius di $S(e^{j\omega})^{-1}$ è uguale alla sua traccia, per cui la condizione (5.9) è equivalente all'integrabilità di tutte le componenti scalari di $S(e^{j\omega})^{-1}$ sulla circonferenza unità.

PROPOSIZIONE 6.6. *Ogni processo minimo è puramente non deterministico.*

DIMOSTRAZIONE. La dimostrazione è basata sulla disuguaglianza, valida per matrici semidefinite positive,

$$(\det A)^{1/n} \leq \frac{1}{n} \text{Tr} A$$

che si può ad esempio trovare in [66, p. 477], la quale porge

$$\int_{-\pi}^{+\pi} \log \det [S(e^{j\omega})^{-1}] \frac{d\omega}{2\pi} \leq \int_{-\pi}^{+\pi} \log \text{Tr} [S(e^{j\omega})^{-1}] \frac{d\omega}{2\pi} \leq \int_{-\pi}^{+\pi} \text{Tr} [S(e^{j\omega})^{-1}] \frac{d\omega}{2\pi}$$

dove l'ultimo passaggio vale in virtù della disuguaglianza di Jensen. Dato che $\log \det [S(e^{j\omega})^{-1}] = -\log \det [S(e^{j\omega})]$, e l'ultimo integrale è minore di $+\infty$ si trova

$$\int_{-\pi}^{+\pi} \log \det [S(e^{j\omega})] \frac{d\omega}{2\pi} > -\infty$$

che è la condizione di Szegö-Kolmogorov (1.37). \square

Il concetto di processo minimo richiama ovviamente l'idea di *base* (non necessariamente ortogonale) per lo spazio di Hilbert $\mathbf{H}(\mathbf{y})$. A questo proposito bisogna mettere in guardia il lettore sul fatto che l'estensione dell'idea di base dal caso finito-dimensionale (dove una base è esattamente un sistema minimo di generatori) a quello infinito-dimensionale, richiede in più una condizione di convergenza. La definizione di base che noi adotteremo in questo contesto è la seguente (cfr. [126, p. 104] per una definizione leggermente diversa).

DEFINIZIONE 6.1. Sia \mathbf{y} un processo aleatorio m -dimensionale a spettro assolutamente continuo. Le variabili di \mathbf{y} formano una base per lo spazio $\mathbf{H}(\mathbf{y})$ se, ogni variabile $\boldsymbol{\xi} \in \mathbf{H}(\mathbf{y})$ può esprimersi in modo unico come

$$\boldsymbol{\xi} = \sum_{t=-\infty}^{+\infty} c(t)\mathbf{y}(t)$$

dove la somma è convergente in media quadratica.

È naturale allora chiedersi se la minimalità di \mathbf{y} sia equivalente alla proprietà di essere una base. La risposta a questa domanda è contenuta nel seguente enunciato (che noi non dimostreremo).

TEOREMA 6.3 (Rozanov, Masani). *Le variabili del processo \mathbf{y} formano una base per lo spazio $\mathbf{H}(\mathbf{y})$ se e solo se il processo è minimo, ovvero se e solo se vale la condizione (5.9).*

Sia allora $\hat{W}(z)$ il fattore spettrale a fase minima (normalizzato) dello spettro di \mathbf{y} , e il processo di innovazione e $\hat{G}(z) = \hat{W}(z)^{-1}$ la funzione di trasferimento del relativo filtro sbiancante. Segue da quanto abbiamo detto finora che,

COROLLARIO 6.3. *Se il processo \mathbf{y} è minimo, allora \hat{G} è una matrice a righe in H_m^2 e quindi ammette uno sviluppo di Laurent*

$$\hat{G}(z) = \sum_0^{+\infty} G(k) z^{-k}$$

in cui i $G(k)$ sono proprio i coefficienti di Fourier (ordinari) di \hat{G} . In questo caso il processo d'innovazione \mathbf{e} ha la rappresentazione, convergente in media quadratica,

$$\mathbf{e}(t) = \sum_{-\infty}^t G(t-k)\mathbf{y}(k)$$

in cui $G(\cdot)$ è una successione a quadrato sommabile.

PROBLEMA 6.12. Il lettore dimostri il corollario usando il fatto che $S(z)^{-1}$ è la densità spettrale di un processo puramente non deterministico. \diamond

Nel caso razionale, è facile convincersi che la condizione di minimalità (5.9), che è in sostanza equivalente all'integrabilità dell'inversa della matrice $S(z)$ sulla circonferenza unitaria, è ancora equivalente all'assenza di zeri di $S(z)$ (ovvero del suo determinante) sulla circonferenza unità. Riscopriamo così il fatto (già noto) che in questo caso il filtro sbiancante è strettamente stabile.

In questioni legate all'identificazione ed in ispecie all'identificazione cosiddetta "a sottospazi" [141] tornano utili delle condizioni di tipo geometrico per la minimalità. Il seguente risultato descrive una condizione intuitiva che è formalmente identica a quella dell'algebra lineare.

TEOREMA 6.4. *Se le variabili $\{\mathbf{y}(t)\}$, formano una base per lo spazio $\mathbf{H}(\mathbf{y})$, allora, per un qualunque istante $t \in \mathbb{Z}$, passato e futuro (stretto) sono in somma diretta, ovvero:*

$$\mathbf{H}_t^-(\mathbf{y}) \cap \mathbf{H}_{t+1}^+(\mathbf{y}) = \{0\} \quad (5.10)$$

Viceversa, se vale la (5.10) e se la varianza $\Sigma(0) := \text{Var}\{\mathbf{y}(t)\}$ è non singolare (in particolare se il processo è a rango pieno), le variabili $\{\mathbf{y}(t)\}$, formano una base per lo spazio $\mathbf{H}(\mathbf{y})$.

DIMOSTRAZIONE. La condizione è necessaria. Infatti, se l'intersezione (5.10) fosse non banale ci sarebbe (almeno) una variabile aleatoria, diciamola $\boldsymbol{\eta}$ esprimibile come combinazione lineare, convergente in media quadratica, sia delle variabili del passato che del futuro stretto

$$\boldsymbol{\eta} = \sum_{-\infty}^t c_-(k)\mathbf{y}(k) = \sum_{t+1}^{+\infty} c_+(k)\mathbf{y}(k)$$

per cui, definendo $c(k) := c_-(k)$ per $k \leq t$ e $c(k) := -c_+(k)$ per $k > t$, si avrebbe

$$\sum_{-\infty}^{+\infty} c(k)\mathbf{y}(k) = 0$$

dove la serie è ovviamente convergente in media quadratica. Questo è in contrasto con la minimalità del processo.

La condizione è sufficiente. Se per assurdo il processo non fosse minimo, ci sarebbe una sequenza $\{c(k)\}$ non identicamente nulla per cui la serie $\sum_{-\infty}^{+\infty} c(k)\mathbf{y}(k)$ converge (in media quadratica) alla variabile aleatoria nulla. Si avrebbe allora, per qualche t_0 ,

$$\sum_{-\infty}^{t_0} c(k)\mathbf{y}(k) = \sum_{t_0+1}^{+\infty} c(k)\mathbf{y}(k).$$

dove la serie a sinistra (e quindi anche quella a destra) converge, in virtù del fatto che la covarianza del processo è una funzione di tipo positivo e quindi vale una disuguaglianza del tipo

$$\sum_{t_0-N}^{t_0} \sum_{t_0-N}^{t_0} c(t)\Sigma(t-s)c(s)^\top \leq \sum_{t_0-N}^{t_0+N} \sum_{t_0-N}^{t_0+N} c(t)\Sigma(t-s)c(s)^\top.$$

Ora, per qualche t_0 la variabile aleatoria

$$\boldsymbol{\eta}(t_0) := \sum_{-\infty}^{t_0} c(k)\mathbf{y}(k)$$

dev'essere non nulla, perchè altrimenti sarebbe nulla anche $\boldsymbol{\eta}(t_0-1)$ e quindi avremmo anche $c(t_0)\mathbf{y}(t_0) = 0$ per ogni t_0 il che è in contrasto con l'ipotesi che $\Sigma(0) > 0$. Siccome la prima somma appartiene a $\mathbf{H}_{t_0}^-(\mathbf{y})$ e la seconda a $\mathbf{H}_{t_0}^+(\mathbf{y})$, ne segue che $\mathbf{H}_{t_0}^-(\mathbf{y}) \cap \mathbf{H}_{t_0}^+(\mathbf{y}) \neq \{0\}$ per qualche t_0 e, per la stazionarietà, l'intersezione sarebbe quindi non zero per ogni t_0 . \square

Se lo spettro di \mathbf{y} è razionale, la condizione (5.10) è equivalente all'assenza di zeri dello spettro sulla circonferenza unità. Questa condizione è sostanzialmente stata ricavata, in modo molto macchinoso, in [63].

Esercizi

6-1 Siano \mathbf{y} e \mathbf{u} due processi stazionari p.n.d. tali per cui

$$\mathbf{H}_t(\mathbf{y}) \subset \mathbf{H}_t(\mathbf{u}) \quad t \in \mathbb{Z}$$

Mostrare che esiste una funzione di trasferimento $\hat{H}(z)$ analitica in $\{|z| > 1\}$, tale che

$$\mathbf{y}(t) = \sum_{-\infty}^t H(t-s)\mathbf{u}(s) \quad (*)$$

dove H è l'antitrasformata zeta di \hat{H} e la somma è convergente in media quadratica.

6-2 Si assuma, viceversa, che \mathbf{y} e \mathbf{u} (stazionari e p.n.d.) siano legati da una relazione del tipo (*) (con la somma convergente in media quadratica) in cui $H(t) = 0$; $t < 0$. È vero che allora $\mathbf{H}_t(\mathbf{y}) \subset \mathbf{H}_t(\mathbf{u})$?

In caso contrario, dare delle condizioni necessarie e sufficienti perchè questo avvenga.

6-3 Siano \mathbf{y} e \mathbf{u} due processi stazionari p.n.d. tali per cui

$$\mathbf{H}_t(\mathbf{y}) \subset \mathbf{H}_t(\mathbf{u}) \quad t \in \mathbb{Z}$$

Usando la relazione $\hat{E}\{\mathbf{y}(t) | \mathbf{u}(s); s \leq t\} = \mathbf{y}(t)$ dimostrare che gli spettri di \mathbf{y} e \mathbf{u} sono legati dalla

$$S_y(e^{j\omega}) = S_{yu}(e^{j\omega})S_u^{-1}(e^{j\omega})S_{uy}(e^{j\omega}) = \hat{H}(e^{j\omega})S_u(e^{j\omega})\hat{H}(e^{-j\omega})'$$

(Teorema di Wiener-Kintchine).

6-4 Dato un processo descritto dal modello M.A.

$$\mathbf{y}(t) = (1 - 9/2z^{-1} + 2z^{-2}) \mathbf{v}(t)$$

calcolare lo stimatore lineare a minima varianza $\hat{\mathbf{v}}(t | t) = \hat{E}[\mathbf{v}(t) | \mathbf{y}^t]$, dato il passato infinito di \mathbf{y} .

Mostrare che il processo $\{\hat{\mathbf{v}}(t | t)\}$ è un multiplo costante dell'innovazione di \mathbf{y} .

6-5 (Generalizzazione del Problema 6-4) Sia dato un processo descritto dal modello M.A.

$$\mathbf{y}(t) = \mathbf{v}(t) + \sum_1^n c_k \mathbf{v}(t - k) := C^*(z^{-1}) \mathbf{v}(t),$$

in cui $\{\mathbf{v}(t)\}$ è rumore bianco di varianza σ^2 , si assuma il polinomio $C(z) := z^n + c_1 z^{n-1} + \dots + c_n$ privo di zeri sulla circonferenza unita $\{|z| = 1\}$ ma del resto arbitrario.

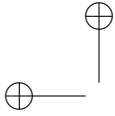
- Si calcoli lo spettro incrociato $S_{vy}(z)$ e si verifichi che la funzione $S_{v\varepsilon}(z) := S_{vy}(z)/L(z^{-1})$ è multipla scalare di una funzione passa-tutto. In particolare i suoi poli e zeri sono a coppie reciproche.
- Lo stimatore lineare a minima varianza $\hat{\mathbf{v}}(t | t) = \hat{E}[\mathbf{v}(t) | \mathbf{y}^t]$, dato il passato infinito di \mathbf{y} , si può scrivere anche come

$$\hat{\mathbf{v}}(t | t) = \hat{E}[\mathbf{v}(t) | \varepsilon^t]$$

- Si mostri che il processo $\{\hat{\mathbf{v}}(t | t)\}$ è sempre un multiplo costante dell'innovazione $\varepsilon(t)$, di \mathbf{y} .
- Verificare che $\mathbf{v}(t) \perp \mathbf{y}(s)$ per tutti gli $s < t$. È corretto dedurne che $\hat{E}[\mathbf{v}(t) | \mathbf{H}_t(\mathbf{y})] = \hat{E}[\mathbf{v}(t) | \mathbf{y}(t)]$?
- Per l'osservazione fatta al punto precedente, si ha $\mathbf{v}(t) \perp \varepsilon^{t-1}$ per tutti gli $s < t$ e in questo caso è *corretto* affermare che $\hat{E}[\mathbf{v}(t) | \mathbf{H}_t(\varepsilon)] = \hat{E}[\mathbf{v}(t) | \varepsilon(t)]$!

6-6 Calcolare lo stimatore lineare a minima varianza $\hat{\mathbf{y}}(t-1 | t) := \hat{E}[\mathbf{y}(t-1) | \mathbf{y}(t)]$ per il processo stazionario descritto dal modello AR

$$\mathbf{y}(t) + a\mathbf{y}(t-1) = \mathbf{e}(t)$$



dove $|a| < 1$ e $\{\mathbf{e}(t)\}$ è rumore bianco di varianza σ^2 . Verificare che

$$\hat{E}[\mathbf{y}(t-1) | \mathbf{y}(t)] = \hat{E}[\mathbf{y}(t-1) | \mathbf{H}_t^+(\mathbf{y})]$$

dove $\mathbf{H}_t^+(\mathbf{y})$ è lo spazio generato dalla storia futura di \mathbf{y} dall'istante t (incluso) in poi.

6-7 Generalizzare l'algoritmo di interpolazione a "due passate" (formule (4.2)) al caso di rumore additivo di spettro (razionale) arbitrario.

6-8 Sia φ una variabile aleatoria a distribuzione uniforme su $[-\pi, \pi]$ e A una costante deterministica. Mostrare che il processo

$$\mathbf{x}(t) = A \sin(\omega_0 t + \varphi) \quad , \quad t \in \mathbb{Z} \quad ,$$

è (debolmente) stazionario e calcolarne la funzione di covarianza $\sigma_{\mathbf{x}}(\tau)$. Discutere il calcolo dello stimatore $\hat{\mathbf{x}}(t | t) = \hat{E}[\mathbf{x}(t) | \mathbf{H}_t(\mathbf{y})]$ a partire da misure rumorose di \mathbf{x}

$$\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{w}(t) \quad ,$$

dove $\{\mathbf{w}(t)\}$ è rumore bianco scorrelato da $\{\mathbf{x}(t)\}$.

6-9 Dato il processo stazionario $\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{v}(t)$

$$\mathbf{x}(t) = (1 + 1/3z^{-1})\mathbf{n}(t) \quad \mathbf{v}(t) = \frac{\sqrt{5}}{2}(1 + 2/3z^{-1})\mathbf{e}(t)$$

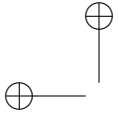
dove $\mathbf{n}(t)$ ed $\mathbf{e}(t)$ sono bianchi normalizzati e tra loro scorrelati, calcolare

- (1) lo stimatore a M.V., $\hat{\mathbf{x}}(t)$, di $\mathbf{x}(t)$ basato su tutta la storia di \mathbf{y}
- (2) lo stimatore a M.V., $\hat{\mathbf{x}}_-(t)$, di $\mathbf{x}(t)$ basato sul passato $\{\mathbf{y}(s) | s < t\}$ (predittore di un passo)
- (3) lo stimatore a M.V., $\hat{\mathbf{x}}_+(t)$, di $\mathbf{x}(t)$ basato sul futuro $\{\mathbf{y}(s) | s \geq t\}$

Verificare che si può esprimere $\hat{\mathbf{x}}(t)$ come combinazione lineare

$$\hat{\mathbf{x}}(t) = A\hat{\mathbf{x}}_-(t) + B\hat{\mathbf{x}}_+(t)$$

e calcolare i coefficienti A, B . Confrontare le varianze d'errore rispettive nei tre casi.



CAPITOLO 7

VARIABILI ESOGENE, CAUSALITÀ E RETROAZIONE FRA PROCESSI STOCASTICI

1. Introduzione

In molte applicazioni, ad esempio nel campo dei controlli automatici ed in econometria, si ha a che fare con segnali la cui evoluzione temporale è influenzata da variabili misurabili che vengono chiamate “esogene” o “esterne” o di “ingresso”. Di solito queste variabili hanno una connotazione speciale nel problema. Esse giocano il ruolo di dati che servono a “spiegare” l’andamento temporale della variabile che interessa, $\{\mathbf{y}(t)\}$, ma che di norma non si ha interesse a descrivere di per sé.

L’introduzione di variabili esogene si rende spesso necessaria per descrivere la dinamica di $\{\mathbf{y}(t)\}$, per esempio ai fini della predizione, perché il passato proprio \mathbf{y}^t , da solo, riesce a spiegare solo parzialmente il futuro di $\{\mathbf{y}(t)\}$ e conduce a predittori di efficienza molto bassa. Si pensa in questo caso ad un nesso funzionale di tipo “causa-effetto” tra l’andamento temporale delle variabili di ingresso, diciamo $\{\mathbf{u}(t)\}$, e quella “di uscita” $\{\mathbf{y}(t)\}$, per esempio come quello tradizionalmente adottato nel campo dei controlli automatici, corrispondente allo schema a blocchi in cui l’uscita effettiva $\mathbf{y}(t)$ è formata da una parte perfettamente prevedibile nota

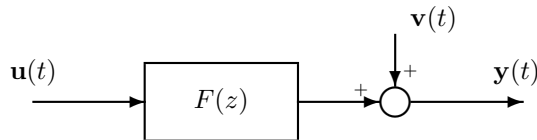


FIGURA 1.1. Modello del segnale \mathbf{y} con disturbo additivo.

l’evoluzione passata dell’ingresso (il termine $F(z)\mathbf{u}(t)$, in cui si assume che $F(z)$ sia una funzione di trasferimento causale) e da una componente imprevedibile di “disturbo”, $\{\mathbf{v}(t)\}$. Naturalmente ci siamo posti in un ambito lineare tempo-invariante e abbiamo riportato, come si suole dire, tutti i disturbi all’uscita.

Nello spirito della teoria che stiamo esponendo, supponendo che sia ragionevole modellare i disturbi come processi *stazionari del secondo ordine* (che spesso

supporremo a spettro razionale) e, pensando di descrivere $\{\mathbf{v}(t)\}$, supposto p.n.d., mediante la sua rappresentazione d'innovazione

$$\mathbf{v}(t) = G(z) \mathbf{e}(t) \quad ,$$

si può quindi pensare di modellare il processo d'uscita mediante una equazione simbolica del tipo

$$\mathbf{y}(t) = F(z) \mathbf{u}(t) + G(z) \mathbf{e}(t) \quad . \quad (1.1)$$

Modelli di questo genere possono essere un punto di partenza naturale per affrontare problemi di stima lineare sulla variabile $\{\mathbf{y}(t)\}$ quando è necessario tener conto, oltre che della storia passata di \mathbf{y} , anche di quella delle variabili di "ingresso" \mathbf{u} .

Occorre a questo punto fare una breve riflessione sul modo in cui si descrive matematicamente il segnale di ingresso.

Dato che non interessa (o non si hanno informazioni su) la dinamica di $\{\mathbf{u}(t)\}$, la scelta più semplice è quella di pensare $\{\mathbf{u}(t)\}$ un segnale *deterministico*, ignoto a priori ma misurabile, nel senso che ad ogni istante t si conoscono tutti i campioni passati $\mathbf{u}^t := \{\mathbf{u}(s) ; s \leq t\}$. Con questa scelta l'"aleatorietà" di $\{\mathbf{y}(t)\}$ risulta concentrata tutta nel segnale di disturbo e l'evoluzione deterministica $F(z) \mathbf{u}(t)$ in (1.1) appare semplicemente come un termine di *media* nel processo $\{\mathbf{y}(t)\}$. Sfortunatamente però questa scelta non è sempre adeguata e può condurre a risultati errati.

La descrizione "deterministica" ¹ di $\{\mathbf{u}(t)\}$ può risultare in particolare inadeguata quando $\{\mathbf{u}(t)\}$ e $\{\mathbf{y}(t)\}$ descrivono segnali che intervengono in sistemi fisici, biologici o economici in cui si può essere in presenza di *retroazione* da $\{\mathbf{y}(t)\}$ a $\{\mathbf{u}(t)\}$.

Similmente, le variabili accessibili di un impianto industriale (ad esempio una cartiera o un impianto chimico) oppure le variabili misurabili in un sistema economico interconnesso, possono facilmente essere influenzate, talvolta in modo poco visibile e difficilmente rivelabile a priori, dall'andamento temporale di altre variabili dell'impianto o del sistema, anche se non esistono veri e propri meccanismi fisici di retroazione (regolatori PID, ecc.). In pratica può essere molto difficile escludere a priori che alcune componenti di $\{\mathbf{u}(t)\}$ non vengano scelte o almeno modificate in relazione all'andamento dell'uscita $\{\mathbf{y}(t)\}$.

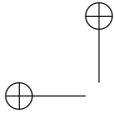
Se questo è il caso, il segnale $\{\mathbf{u}(t)\}$ non può essere trattato come un parametro certo e deve essere modellato anch'esso come un processo stocastico.

Supponiamo allora di descrivere anche $\{\mathbf{u}(t)\}$ come un processo stazionario. Esso può, in esatta analogia a quanto fatto per $\{\mathbf{y}(t)\}$, essere decomposto a sua volta in una parte che congloba l'influenza dell'andamento passato, \mathbf{y}^t , dell'uscita, più un termine "autonomo" che in qualche modo descrive la parte della dinamica dell'ingresso non influenzata dall'uscita.

Si dovrebbero ora scrivere, in perfetta simmetria a quanto fatto in (1.1), *due* relazioni "ingresso-uscita più disturbo" del tipo

$$\begin{aligned} \mathbf{y}(t) &= F(z) \mathbf{u}(t) + G(z) \mathbf{e}_1(t) \quad , \\ \mathbf{u}(t) &= H(z) \mathbf{y}(t) + K(z) \mathbf{e}_2(t) \quad . \end{aligned} \quad (1.2)$$

¹In realtà dire che si descrive $\{\mathbf{u}(t)\}$ come segnale "deterministico ma incognito" significa che non si vuol ipotizzare e usare alcun modello statistico del segnale. Si tratta in realtà di ipotizzare implicitamente una situazione in cui la soluzione del problema è *indipendente da quale possa essere il modello matematico di $\{\mathbf{u}(t)\}$* .



Nelle ipotesi fatte è in effetti possibile dare anche una giustificazione formale dell'esistenza di rappresentazioni lineari congiunte dei due processi del tipo (1.2). Si possono ad esempio identificare i termini causali $F(z) \mathbf{u}(t)$ e $H(z) \mathbf{y}(t)$ con le proiezioni (stimatori di Wiener) $\hat{E}[\mathbf{y}(t) | \mathbf{u}^t]$ e $\hat{E}[\mathbf{u}(t) | \mathbf{y}^t]$, e $G(z) \mathbf{e}_1(t)$ e $K(z) \mathbf{e}_2(t)$ come le rappresentazioni d'innovazione dei due errori di stima $\{\tilde{\mathbf{y}}(t)\}$ e $\{\tilde{\mathbf{u}}(t)\}$. In realtà, come vedremo più avanti, una descrizione congiunta del tipo (1.2) è altamente *non unica* e la decomposizione ortogonale ottenuta proiettando sul passato di $\{\mathbf{u}(t)\}$ e $\{\mathbf{y}(t)\}$ è solo *una* delle molte possibili (in effetti non è nemmeno una delle più convenienti).

È immediato verificare che il modello congiunto (1.2) dei due segnali corrisponde ad uno schema a blocchi del tipo

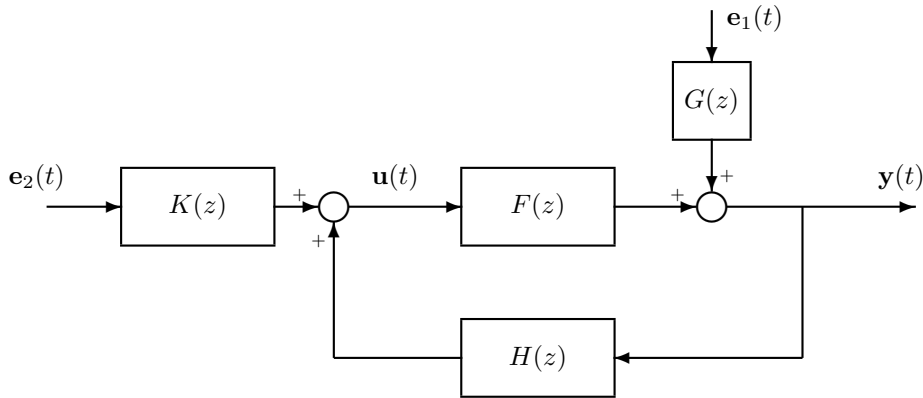


FIGURA 1.2. Modello congiunto dei segnali \mathbf{y} e \mathbf{u} .

che evidenzia in modo del tutto generale proprio la struttura a *controreazione*, la cui possibile presenza nel contesto *fisico* (o economico) era stata l'origine delle nostre preoccupazioni.

A rischio di disorientare il lettore, è bene a questo punto osservare che non esistono più ragioni "sistemistiche" per distinguere le variabili "esogene" $\{\mathbf{u}(t)\}$ da quelle "di uscita" $\{\mathbf{y}(t)\}$. Il ruolo delle due è in effetti completamente simmetrico e non c'è alcuna ragione di affermare che $\{\mathbf{u}(t)\}$ *causa* $\{\mathbf{y}(t)\}$ (è un "ingresso") più di quanto non sia vero il viceversa. L'uso della parola "ingresso", riferendosi al significato deterministico del termine ($\mathbf{y}(t)$ è funzione *solo* del passato di $\{\mathbf{u}(t)\}$, $\mathbf{y}(t) = f(\mathbf{u}^t)$), in questo contesto può dare luogo a seri errori.

PROBLEMA 7.1. Un modo ingenuo di definire l'analogo stocastico della nozione deterministica di causalità potrebbe essere il seguente: $\{\mathbf{u}(t)\}$ *causa* $\{\mathbf{y}(t)\}$ (cioè \mathbf{u} è un *ingresso* e \mathbf{y} la corrispondente uscita) se, per ogni t , $\mathbf{y}(t)$ è scorrelato (o indipendente) dagli ingressi futuri, i.e.,

$$E \mathbf{y}(t) \mathbf{u}(t+k) = 0 \quad , \quad \forall k > 0 \quad . \quad (1.3)$$

Sia $\mathbf{y}(t) = F(z) \mathbf{u}(t)$ con $F(z)$ strettamente stabile e causale e $\{\mathbf{u}(t)\}$ p.n.d.. Mostrare che la (1.3) può essere vera solo se $\{\mathbf{u}(t)\}$ è *rumore bianco*. \diamond

A causa della retroazione, la possibilità, che era stata data per scontata all'inizio della nostra discussione, di usare il passato delle variabili "esogene" per costruire

predittori di $\mathbf{y}(t)$, trattando queste variabili come ingressi deterministici, non è quindi affatto scontata. Comprendere *quando* e *come* questo può essere possibile sarà in sostanza il tema principale di questo capitolo.

Diciamo fin d'ora che la descrizione deterministica di $\{\mathbf{u}(t)\}$ può essere “corretta” nell'ipotesi di *assenza di reazione* (definiremo tra breve in modo preciso questo concetto). Il senso intuitivo di questa affermazione è il seguente:

Se non c'è reazione un modello probabilistico congiunto di $\{\mathbf{y}(t)\}$ e $\{\mathbf{u}(t)\}$ che descrive i due processi accoppiati (mediante i momenti congiunti del secondo ordine) conduce, ad esempio nella soluzione del problema di predizione di $\{\mathbf{y}(t)\}$ in base ai dati $(\mathbf{y}^t, \mathbf{u}^t)$, a formule che non dipendono dalla descrizione statistica congiunta di $\{\mathbf{u}(t)\}$ e $\{\mathbf{y}(t)\}$ e sono di struttura analoga a quelle che si ottengono supponendo che $\{\mathbf{u}(t)\}$ sia un segnale deterministico.

Questa affermazione è da intendere come un enunciato di carattere generale e verrà meglio precisata nel corso dell'esposizione che segue.

2. Modelli a Retroazione

In questa sezione studieremo una classe di modelli lineari del tipo (1.2) che potrebbe ben servire per descrivere un sistema di controllo con retroazione dall'uscita, con ingresso di riferimento $\{\mathbf{r}(t)\}$ e soggetto ad un segnale di disturbo agente sull'uscita $\{\mathbf{v}(t)\}$, entrambi modellabili come processi aleatori stazionari p.n.d.. Questa classe di modelli, oltre ad avere interesse come descrizione naturale da usare per un approccio statistico all'analisi e al progetto di sistemi di controllo, sarà utile anche per discutere la soluzione di problemi di predizione e filtraggio in condizioni di presenza di reazione.

Consideriamo due processi congiuntamente stazionari del secondo ordine $\{\mathbf{y}(t)\}$ e $\{\mathbf{u}(t)\}$ di dimensioni m e p rispettivamente.

Supponiamo che il cosiddetto *processo congiunto*,

$$\mathbf{z}(t) := \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{bmatrix} , \quad (2.1)$$

sia p.n.d., di rango pieno $r = m + p$ e (per concretezza) a *spettro razionale* strettamente definito positivo sulla circonferenza unità

$$S(e^{j\omega}) > 0 \quad , \quad \omega \in [-\pi, \pi] \quad . \quad (2.2)$$

DEFINIZIONE 7.1. *Chiameremo Modello a retroazione del processo stocastico $\{\mathbf{z}(t)\}$, una rappresentazione mediante uno schema del tipo descritto in Figura 8.2, ovvero una descrizione simbolica del tipo*

$$\begin{aligned} \mathbf{y}(t) &= F(z) \mathbf{u}(t) + \mathbf{v}(t) \quad , \\ \mathbf{u}(t) &= H(z) \mathbf{y}(t) + \mathbf{r}(t) \quad , \end{aligned} \quad (2.3)$$

soggetta alle seguenti condizioni:

- (1) *Le matrici di trasferimento $F(z)$ e $H(z)$ sono funzioni razionali di z causali (regolari all'infinito), ma non necessariamente analitiche fuori dal cerchio unità. Inoltre esiste almeno un ritardo in catena di azione diretta, i.e.,*

$$F(\infty) = 0 \quad . \quad (2.4)$$

- (2) I processi \mathbf{v} e \mathbf{r} hanno spettro razionale, strettamente positivo sulla circonferenza unità. Senza perdita di generalità, si possono pensare assegnati tramite le rispettive rappresentazioni d'innovazione

$$\mathbf{v}(t) := G(z) \mathbf{w}_1(t) \quad , \quad \mathbf{r}(t) := K(z) \mathbf{w}_2(t) \quad (2.5)$$

dove $\{\mathbf{w}_1(t)\}$ e $\{\mathbf{w}_2(t)\}$ sono bianchi e scorrelati, ovvero

$$E \left\{ \begin{bmatrix} \mathbf{w}_1(t) \\ \mathbf{w}_2(t) \end{bmatrix} \begin{bmatrix} \mathbf{w}_1(s)' \\ \mathbf{w}_2(s)' \end{bmatrix} \right\} = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \delta(t-s) \quad , \quad (2.6)$$

dove Q_1 e Q_2 sono (strettamente) definite positive. Assumeremo in genere che (2.5) siano rappresentazioni d'innovazione non normalizzata con G e K a fase minima e $G(\infty) = I_m$, $K(\infty) = I_p$.

- (3) L'interconnessione a retroazione (Figura 8.2) corrispondente al modello (2.3) è internamente stabile. In altri termini la funzione di trasferimento in catena chiusa, $T(z)$, che lega gli ingressi $\{\mathbf{v}(t)\}$ ed $\{\mathbf{r}(t)\}$ alle variabili $\{\mathbf{y}(t)\}$ e $\{\mathbf{u}(t)\}$,

$$T(z) = \begin{bmatrix} (I - FH)^{-1} & (I - FH)^{-1} F \\ (I - HF)^{-1} H & (I - HF)^{-1} \end{bmatrix} \quad , \quad (2.7)$$

è priva di poli in $\{|z| \geq 1\}$. \diamond

È immediato verificare che il processo congiunto $\left\{ \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{bmatrix} \right\}$ descritto da un modello a retroazione è stazionario, di rango pieno e a spettro razionale, strettamente definito positivo sulla circonferenza unità. In effetti, risolvendo le (2.3) rispetto alle variabili $\mathbf{y}(t)$ e $\mathbf{u}(t)$ si trova

$$\begin{bmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{bmatrix} = \begin{bmatrix} (I - FH)^{-1} G & (I - FH)^{-1} FK \\ (I - HF)^{-1} HG & (I - HF)^{-1} K \end{bmatrix} \begin{bmatrix} \mathbf{w}_1(t) \\ \mathbf{w}_2(t) \end{bmatrix} := W(z) \mathbf{w}(t) \quad (2.8)$$

dove la matrice $W(z) = T(z) \text{diag}\{G(z), K(z)\}$ è razionale e analitica in $\{|z| \geq 1\}$, incluso il punto all'infinito, giacché²

$$W(\infty) = \begin{bmatrix} I_m & 0 \\ H(\infty) & I_p \end{bmatrix} \quad (2.9)$$

la quale dimostra fra l'altro che $W(z)$ è di rango pieno in un intorno di $z = \infty$ e quindi quasi ovunque sul piano complesso. Dalla rappresentazione (2.8) deriva la fattorizzazione "diagonale a blocchi" dello spettro congiunto $S(z)$

$$S(z) = W(z) \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} W(z^{-1})^T \quad , \quad (2.10)$$

con $W(z)$ quadrata e analitica in $\{|z| \geq 1\}$.

Un modello a retroazione individua quindi un particolare fattore spettrale quadrato, analitico in $\{|z| \geq 1\}$ e triangolare inferiormente a blocchi all'infinito, dello spettro congiunto dei due processi stazionari $\{\mathbf{y}(t)\}$ e $\{\mathbf{u}(t)\}$ che rappresenta.

DEFINIZIONE 7.2. Fissati m e p con $m + p = r$, diremo che $W(z)$ è un fattore spettrale normalizzato a blocchi di $S(z)$ se

²Notare che la presenza del ritardo in catena di azione diretta fa sì che l'interconnessione a retroazione di Figura 8.2 sia *ben posta*, cioè che le equazioni implicite (2.3) siano risolvibili e la funzione di trasferimento in catena chiusa $W(z)$ sia causale.

- (1) $W(z)$ è una matrice analitica all'infinito, triangolare inferiormente a blocchi con

$$W(\infty) = \begin{bmatrix} I_m & 0 \\ * & I_p \end{bmatrix} \begin{matrix} m \\ p \end{matrix}, \quad (2.11)$$

ovvero $W_{11}(\infty) = I_m$, $W_{12}(\infty) = 0$ e $W_{22}(\infty) = I_p$.

- (2) $W(z)$ soddisfa una relazione di fattorizzazione spettrale del tipo (2.10) con Q_1 e Q_2 simmetriche, definite positive e di dimensioni $m \times m$ e $p \times p$, rispettivamente.

Da quanto abbiamo appena visto, scende che il fattore spettrale corrispondente ad un modello a retroazione è sempre normalizzato a blocchi. La proposizione seguente stabilisce che un qualunque fattore spettrale $W(z)$ regolare all'infinito, cioè tale che $W(\infty)$ è finita e non singolare, può sempre essere normalizzato a blocchi mediante una semplice moltiplicazione a destra per una opportuna matrice costante.

PROPOSIZIONE 7.1. *Un qualunque fattore spettrale di $S(z)$, quadrato e regolare all'infinito, può essere normalizzato a blocchi moltiplicandolo a destra per una opportuna matrice costante. Le varianze Q_1 e Q_2 sono univocamente determinate da $W(z)$.*

DIMOSTRAZIONE. Senza perdita di generalità, possiamo assumere che il fattore spettrale, diciamolo $\bar{W}(z)$, sia normalizzato all'infinito ($\bar{W}(\infty) = I$) e soddisfi una relazione di fattorizzazione spettrale del tipo

$$S(z) = \bar{W}(z) Q \bar{W}'(z^{-1}),$$

dove $Q = Q' > 0$. Detto $\{\mathbf{n}(t)\}$ il processo di rumore bianco r -dimensionale di varianza Q associato al fattore spettrale $\bar{W}(z)$, siano $\{\mathbf{n}_1(t)\}$ e $\{\mathbf{n}_2(t)\}$ i processi costruiti prendendo le sue prime m e le ultime p componenti. Definendo

$$\mathbf{w}_1(t) := \mathbf{n}_1(t), \quad \mathbf{w}_2(t) := \mathbf{n}_2(t) - \hat{E}[\mathbf{n}_2(t) | \mathbf{n}_1(t)],$$

evidentemente si ottengono due rumori bianchi $\{\mathbf{w}_1(t)\}$ e $\{\mathbf{w}_2(t)\}$ scorrelati, la cui covarianza ha la struttura (2.6). Inoltre $\mathbf{w}(t) := [\mathbf{w}_1'(t) \mathbf{w}_2'(t)]'$ si ottiene da $\mathbf{n}(t)$ attraverso la trasformazione di struttura triangolare inferiore

$$\begin{bmatrix} \mathbf{w}_1(t) \\ \mathbf{w}_2(t) \end{bmatrix} = \begin{bmatrix} I & 0 \\ -Q_{21}Q_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} \mathbf{n}_1(t) \\ \mathbf{n}_2(t) \end{bmatrix} := T \mathbf{n}(t), \quad (2.12)$$

dove $Q_{ij} := E \mathbf{n}_i(t) \mathbf{n}_j'(t)$, $ij = 1, 2$. Ne segue che, esiste una trasformazione costante T triangolare inferiormente che "ortogonalizza" a blocchi il rumore bianco associato a \bar{W} . La nuova matrice $W(z) := \bar{W}(z)T^{-1}$ (T^{-1} è ancora triangolare inferiormente) soddisfa alle condizioni 1. e 2. con Q_1 e Q_2 opportune. \square

Quanto "particolare" è la rappresentazione a retroazione di due processi stazionari? Dato uno spettro congiunto $S(z)$ definito positivo sulla circonferenza unità, ma del resto arbitrario, è possibile riconoscere se esso può o meno essere lo spettro congiunto di due processi stazionari (\mathbf{y}, \mathbf{u}) generati da un modello a retroazione?

Dimostreremo in questa sezione che uno spettro razionale definito positivo sulla circonferenza unità, può sempre essere "realizzato" da uno schema a retroazione.

Questo significa che preso un fattore spettrale quadrato normalizzato a blocchi dello spettro (dalla Proposizione 7.1 segue che è sempre possibile costruirne), è generalmente possibile fargli corrispondere un modello a retroazione che rappresenta i due processi componenti, $\{\mathbf{y}(t)\}$ e $\{\mathbf{u}(t)\}$, come le variabili di uscita e di

controllo di un sistema a retroazione ben posto, internamente stabile e con processi di riferimento e disturbo scorrelati.

Supponiamo che $W(z)$ sia un fattore spettrale quadrato normalizzato a blocchi di $S(z)$. Imponendo la struttura “a retroazione” (2.8) si ricavano facilmente le relazioni

$$\begin{aligned} F &= W_{12} W_{22}^{-1} & G &= W_{11} - W_{12} W_{22}^{-1} W_{21} \\ H &= W_{21} W_{11}^{-1} & K &= W_{22} - W_{21} W_{11}^{-1} W_{12} \end{aligned} \quad (2.13)$$

che forniscono quattro funzioni razionali causali mediante le quali si possono formalmente rappresentare $\{\mathbf{y}(t)\}$ e $\{\mathbf{u}(t)\}$ tramite un modello a retroazione del tipo (2.3). In effetti $F(\infty) = 0$ e $F(z)$ è *strettamente causale*, $H(\infty) = W_{21}(\infty)$ è finita, $G(\infty) = I_m$ e $K(\infty) = I_p$. Questa corrispondenza è per ora meramente un fatto formale perché, anche se si parte da un fattore spettrale $W(z)$ analitico in $\{|z| \geq 1\}$, non c'è garanzia a priori che il modello a retroazione definito dalle quattro funzioni di trasferimento (2.13) sia *internamente stabile*.

ESEMPIO 7.1 (Un fattore spettrale analitico con modello a retroazione instabile). Il fattore spettrale 2×2

$$W(z) = \frac{1}{z + 0.8} \begin{bmatrix} z - 0.7 & 1 \\ 1.95 & z - 0.5 \end{bmatrix}, \quad (2.14)$$

cui corrispondono

$$\begin{aligned} F(z) &= \frac{1}{z - 0.5} & G(z) &= \frac{z - 2}{z - 0.5} \\ H(z) &= \frac{1.95}{z - 0.7} & K(z) &= \frac{z - 2}{z - 0.7} \end{aligned}$$

e

$$\left(1 - F(z)H(z)\right)^{-1} = \frac{(z - 0.5)(z - 0.7)}{(z - 2)(z + 0.8)},$$

pur essendo analitico, genera un modello a retroazione *instabile*. Notiamo che nel formare il prodotto $T(z) \text{diag}\{G(z), K(z)\} = W(z)$, il polo in catena chiusa in $z = 2$ viene cancellato. Notiamo anche che la cancellazione di poli instabili è possibile proprio perché, anche se $W(z)$ è analitica in $\{|z| \geq 1\}$, $G(z)$ e $K(z)$ non vengono necessariamente generate a fase minima (e nemmeno stabili) dalla formula (2.13). \diamond

Per studiare il problema di rappresentazione (o realizzazione) che abbiamo posto più sopra, possiamo innanzitutto rimarcare che un modello a retroazione è completamente definito dalle funzioni di trasferimento d'anello $F(z)$ e $H(z)$ e dai due spettri $S_v(z)$ e $S_r(z)$ dei due processi di disturbo, $\mathbf{v}(t)$, e di riferimento $\mathbf{r}(t)$. Quale fattore spettrale di $S_v(z)$ e $S_r(z)$ venga scelto per rappresentarli è ininfluente. Quindi, d'ora in avanti, con la dizione “modello a retroazione” intenderemo la *classe d'equivalenza* $\{F(z), H(z), G(z), K(z)\}$ dove $G(z)$ e $K(z)$ variano nella famiglia dei fattori spettrali quadrati (non necessariamente analitici) normalizzati all'infinito ($G(\infty) = I_m$, $K(\infty) = I_p$) dei rispettivi spettri.

Per descrivere la corrispondenza tra fattori spettrali quadrati normalizzati e modelli a retroazione, questi ultimi intesi nel senso più generale che abbiamo appena precisato, è necessario introdurre una relazione di equivalenza anche tra fattori spettrali normalizzati a blocchi.

DEFINIZIONE 7.3. Due fattori spettrali quadrati normalizzati a blocchi $W^1(z)$ e $W^2(z)$ (non necessariamente analitici) dello spettro, $S(z)$, del processo congiunto (2.1) si dicono diagonalmente equivalenti (a blocchi) se

$$W^2(z) = W^1(z) \operatorname{diag} \{V_1(z), V_2(z)\} \quad , \quad (2.15)$$

dove $V_1(z)$ e $V_2(z)$, di dimensioni $m \times m$ e $p \times p$, sono matrici “paraunitarie”, ovvero tali che

$$V_k(z) Q_k^1 V_k(z^{-1})^T = Q_k^2 \quad k = 1, 2. \quad (2.16)$$

le $Q^i = \operatorname{diag} \{Q_1^i, Q_2^i\}$, $i = 1, 2$ essendo le rispettive matrici varianza dei fattori $W^i(z)$, i.e. $S(z) = W^i(z) Q^i W^i(z^{-1})'$.

Notiamo che una matrice paraunitaria razionale è limitata e quindi analitica in un intorno della circonferenza unitaria. Una matrice paraunitaria *passa-tutto* è una matrice paraunitaria analitica in $\{|z| \geq 1\}$ che quindi preserva l'analiticità dei fattori spettrali.

Osserviamo ora che il modello a retroazione corrispondente al fattore diagonalmente equivalente ad un dato $W(z)$, $\hat{W} = W \operatorname{diag} \{V_1, V_2\}$, è definito da

$$\hat{F} = \hat{W}_{12} \hat{W}_{22}^{-1} = F \quad , \quad \hat{H} = \hat{W}_{21} \hat{W}_{11}^{-1} = H \quad (2.17)$$

per cui F ed H non cambiano. In altri termini:

PROPOSIZIONE 7.2. L'equivalenza diagonale tra fattori spettrali è una relazione di equivalenza che preserva le funzioni di trasferimento F ed H (e quindi la matrice di trasferimento in catena chiusa $T(z)$) e trasforma solo le matrici G e K in altri fattori spettrali normalizzati degli stessi spettri $S_v(z)$ e $S_r(z)$.

In particolare G e K possono essere rese a fase minima mediante equivalenza diagonale senza cambiare le funzioni di trasferimento d'anello.

Questo perché $G(\infty) = I_m$ e $K(\infty) = I_p$ ed esistono certamente (molte) matrici paraunitarie $V_1(z)$ e $V_2(z)$ tali da rendere

$$\begin{aligned} \hat{G} &:= G V_1 = (W_{11} - W_{12} W_{22}^{-1} W_{21}) V_1 \\ \hat{K} &:= K V_2 = (W_{22} - W_{21} W_{11}^{-1} W_{12}) V_2 \end{aligned} \quad (2.18)$$

fattori spettrali a fase minima di $S_v(z)$ e $S_r(z)$. Un arbitrario fattore razionale quadrato $W(z)$ normalizzato a blocchi è quindi sempre diagonalmente equivalente ad un fattore $\hat{W}(z)$ in cui le rispettive \hat{G} e \hat{K} sono a fase minima. Purtroppo però, come dimostra l'esempio precedente, non è sempre vero che il modello a retroazione costruito mediante la trasformazione (2.13) a partire da un fattore spettrale *analitico* $W(z)$, ha la proprietà fondamentale di *stabilità interna*, ovvero che la funzione di trasferimento a catena chiusa $T(z)$ è analitica fuori dal cerchio unitario.

Studieremo ora un particolare fattore spettrale per cui questa circostanza si verifica. Il caso generale verrà affrontato più avanti.

Il fattore spettrale normalizzato a fase minima e il modello a retroazione d'innovazione. Per quanto ne sappiamo finora è possibile modellare lo stesso processo congiunto mediante una moltitudine di fattori spettrali normalizzati a blocchi ma non necessariamente mediante altrettanti modelli a retroazione (internamente stabili). Faremo vedere in questa sezione che il modello a retroazione corrispondente al fattore spettrale normalizzato a blocchi *a fase minima*, esiste ed é

ben definito. Questo fattore a fase minima é unico (giacché la normalizzazione all'infinito lo fissa univocamente assieme alle matrici di covarianza dei rumori bianchi di ingresso) e fa intervenire nel corrispondente modello una versione opportunamente ortogonalizzata a blocchi del processo d'innovazione.

Per costruire il modello a retroazione d'innovazione (non normalizzata) occorre costruire il *fattore a fase minima normalizzato a blocchi*, diciamolo $\hat{W}(z)$, per il quale $\hat{W}_{11}(\infty) = I_m$, $\hat{W}_{22}(\infty) = I_p$. Per ottenere la particolare struttura all'infinito che permette di soddisfare le condizioni 1. e 2. della definizione 7.2 si può per esempio partire dal fattore a fase minima normalizzato ($W(\infty) = I$) e, seguendo il procedimento di ortogonalizzazione indicato nella prova della proposizione 7.1, associargli un rumore bianco, $\{\hat{\mathbf{e}}(t)\}$, le cui componenti sono

$$\hat{\mathbf{e}}_1(t) = \mathbf{e}_1(t) \quad , \quad \hat{\mathbf{e}}_2(t) = \mathbf{e}_2(t) - \hat{E}[\mathbf{e}_2(t) | \mathbf{e}_1(t)] \quad , \quad (2.19)$$

che sono scorrelate e di varianze rispettive

$$\begin{aligned} \hat{Q}_1 &= \Lambda_{11} = \text{Var}(\mathbf{e}_1(t)) \\ \hat{Q}_2 &:= \Lambda_{22} - \Lambda_{21} \Lambda_{11}^{-1} \Lambda_{12} \end{aligned} \quad (2.20)$$

Λ_{ij} , $ij = 1, 2$, essendo i corrispondenti blocchi della matrice Λ , varianza dell'innovazione *congiunta* non normalizzata, $\{\mathbf{e}(t)\}$, di $\{\{\mathbf{y}'(t), \mathbf{u}'(t)\}'\}$. Il significato della seconda componente, $\{\hat{\mathbf{e}}_2(t)\}$, dell'innovazione ortogonalizzata a blocchi, è spiegato nell'enunciato seguente.

PROPOSIZIONE 7.3. *Le componenti dell'innovazione ortogonalizzata a blocchi $\{\hat{\mathbf{e}}(t)\}$, definita dalla formula (2.19), sono ancora errori di predizione di un passo; si ha infatti*

$$\begin{aligned} \hat{\mathbf{e}}_1(t) = \mathbf{e}_1(t) &= \mathbf{y}(t) - \hat{E}[\mathbf{y}(t) | \mathbf{y}^{t-1} \mathbf{u}^{t-1}] \\ \hat{\mathbf{e}}_2(t) &= \mathbf{u}(t) - \hat{E}[\mathbf{u}(t) | \mathbf{y}^t \mathbf{u}^{t-1}] \quad . \end{aligned} \quad (2.21)$$

Le due componenti di $\{\hat{\mathbf{e}}(t)\}$ in (2.21) sono i rumori bianchi che intervengono nel modello a retroazione d'innovazione non normalizzata, corrispondente al fattore $\hat{W}(z)$,

$$\begin{aligned} \mathbf{y}(t) &= \hat{F}(z) \mathbf{u}(t) + \hat{G}(z) \hat{\mathbf{e}}_1(t) \\ \mathbf{u}(t) &= \hat{H}(z) \mathbf{y}(t) + \hat{K}(z) \hat{\mathbf{e}}_2(t) \quad . \end{aligned} \quad (2.22)$$

DIMOSTRAZIONE. La rappresentazione di $\hat{\mathbf{e}}_2(t)$ nella (2.21) si ottiene notando che dalla $\mathbf{H}(\mathbf{y}^t, \mathbf{u}^{t-1}) = \mathbf{H}(\mathbf{y}^{t-1}, \mathbf{u}^{t-1}) \oplus \mathbf{H}(\mathbf{e}_1(t))$, dove la somma diretta è ortogonale, si ricava

$$\hat{E}[\mathbf{u}(t) | \mathbf{y}^t, \mathbf{u}^{t-1}] = \hat{E}[\mathbf{u}(t) | \mathbf{y}^{t-1}, \mathbf{u}^{t-1}] + \hat{E}[\mathbf{u}(t) | \mathbf{e}_1(t)] \quad . \quad (2.23)$$

Ricordando la definizione di $\mathbf{e}_2(t) := \mathbf{u}(t) - \hat{E}[\mathbf{u}(t) | \mathbf{y}^{t-1} \mathbf{u}^{t-1}]$, e sostituendo nell'ultimo addendo, si trova $\hat{E}[\mathbf{u}(t) | \mathbf{e}_1(t)] = \hat{E}[\mathbf{e}_2(t) | \mathbf{e}_1(t)]$. Basta allora sottrarre i due membri della (2.23) da $\mathbf{u}(t)$ per ottenere l'espressione voluta per $\hat{\mathbf{e}}_2(t)$. \square

Da notare che se si parte dal fattore spettrale canonico, $L(z)$, di $S(z)$, la relazione di fattorizzazione spettrale si può scrivere formalmente come

$$S(z) = L(z) \begin{bmatrix} I_m & 0 \\ 0 & I_p \end{bmatrix} L'(1/z)$$

con I_m e I_p identità di dimensioni m e p . La condizione 1. della definizione 7.2 non è però in generale soddisfatta perché $L(\infty) = \Lambda^{1/2}$, (il fattore di Cholesky della varianza d'innovazione Λ) seppur triangolare inferiormente, non può in generale essere normalizzato in modo da far diventare i blocchi $L_{11}(\infty), L_{22}(\infty)$ uguali all'identità.

Pertanto il fattore canonico *non* è, secondo la nostra definizione, un fattore a fase minima normalizzato a blocchi. Esso ammette comunque un corrispondente modello a retroazione

$$\begin{aligned} \mathbf{y}(t) &= \bar{F}(z) \mathbf{u}(t) + \bar{G}(z) \boldsymbol{\varepsilon}_1(t) \\ \mathbf{u}(t) &= \bar{H}(z) \mathbf{y}(t) + \bar{K}(z) \boldsymbol{\varepsilon}_2(t) \end{aligned} \quad (2.24)$$

in cui i rumori bianchi normalizzati $\{\boldsymbol{\varepsilon}_1(t)\}$ ed $\{\boldsymbol{\varepsilon}_2(t)\}$ sono semplicemente le prime m e le ultime p componenti dell'*innovazione normalizzata del processo congiunto* $\{[\mathbf{y}(t), \mathbf{u}(t)']'\}$.

È importante notare che anche se $W(z)$ è a fase minima, nel modello a retroazione corrispondente $G(z)$ e $K(z)$ non risultano necessariamente a fase minima. L'esempio seguente dovrebbe servire a chiarire questo punto.

ESEMPIO 7.2. La matrice

$$W(z) = \frac{1}{z - 1/2} \begin{bmatrix} z - 1/2 & 1 \\ -3/2(z - 1/2) & z - 2 \end{bmatrix}$$

è analitica in $\{|z| \geq 1\}$, con $W(\infty) = \begin{bmatrix} 1 & 0 \\ -3/2 & 1 \end{bmatrix}$ e la sua inversa,

$$W^{-1}(z) = \frac{1}{z - 1/2} \begin{bmatrix} z - 2 & -1 \\ 3/2(z - 1/2) & z - 1/2 \end{bmatrix},$$

è manifestamente ancora analitica in $\{|z| \geq 1\}$ e quindi $W(z)$ è a fase minima. Dalla normalizzazione all'infinito, si può pensare che $W(z) \equiv L(z)$ sia il fattore spettrale canonico di un processo congiunto bidimensionale. Dalla formula (2.25) si ottengono le funzioni di trasferimento del modello a retroazione d'innovazione normalizzata,

$$\begin{aligned} \bar{G} &= \frac{z - 1/2}{z - 2} & \bar{K} &= 1 \\ \bar{F} &= \frac{1}{z - 2} & \bar{H} &= -3/2 \end{aligned}$$

in cui \bar{G} e \bar{F} hanno in comune un polo instabile in $z = 2$. È facile vedere che $W(z)$ è diagonalmente equivalente al fattore spettrale

$$\tilde{W}(z) = \frac{1}{z - 1/2} \begin{bmatrix} z - 2 & 1 \\ -3/2(z - 2) & z - 2 \end{bmatrix}$$

a cui corrispondono le funzioni a fase minima $G(z) = 1$ e $H(z) = 1$, ma, come si vede, $\tilde{W}(z)$ non è più a fase minima, dato che ha uno zero in $z = 2$. \diamond

L'enunciato che segue caratterizza il *modello a retroazione d'innovazione* che corrisponde al fattore spettrale normalizzato a blocchi e a fase minima dello spettro congiunto $S(z)$.

TEOREMA 7.1. *Se in un modello a retroazione le funzioni di trasferimento F ed H sono analitiche in $\{|z| \geq 1\}$ e G e K sono a fase minima, allora il corrispondente fattore spettrale $W(z)$ in (2.8) è strettamente a fase minima³. Viceversa, se il fattore spettrale analitico e normalizzato a blocchi dello spettro congiunto, diciamolo $\hat{W}(z)$, è a fase minima, allora le matrici razionali \hat{F} , \hat{H} , \hat{G} , \hat{K} definite a partire da $\hat{W}(z)$, dalle relazioni inverse (2.13), sono tali che:*

- (1) *gli zeri di $\hat{G}(z)$ e $\hat{K}(z)$ sono in $\{|z| < 1\}$;*
- (2) *la funzione di trasferimento $\hat{T}(z)$, corrispondente ad \hat{F} e \hat{H} (2.7), è analitica in $\{|z| \geq 1\}$;*
- (3) *le matrici $\hat{F}(z)$, $\hat{H}(z)$, $\hat{G}(z)$ e $\hat{K}(z)$, sono analitiche in $\{|z| \geq 1\}$, a meno di eventuali poli instabili che sono comuni a (\hat{G}, \hat{F}) ed a (\hat{K}, \hat{H}) e si cancellano formando i prodotti $\hat{G}^{-1}\hat{F}$ e $\hat{K}^{-1}\hat{H}$.*

In particolare, il modello a retroazione corrispondente al fattore a fase minima normalizzato è sempre ben posto (si ha stabilità interna dell'interconnessione a catena chiusa).

DIMOSTRAZIONE. Per quanto riguarda la prima affermazione, notiamo che la funzione di trasferimento in catena chiusa $T(z)$ è per definizione analitica in $\{|z| \geq 1\}$; se G e K sono a fase minima, ovviamente $W(z)$ è analitica in $\{|z| \geq 1\}$. Dall'espressione per l'inversa di $W(z)$, che riportiamo qui di seguito,

$$W(z)^{-1} = \begin{bmatrix} G^{-1} & -G^{-1}F \\ -K^{-1}H & K^{-1} \end{bmatrix}, \quad (2.25)$$

si trae poi che gli zeri di $W(z)$ sono esattamente i poli delle matrici razionali che compaiono nei quattro blocchi di $W(z)^{-1}$. Notiamo che G e K non hanno zeri in $\{|z| = 1\}$ per definizione ed F e H sono per ipotesi prive di poli sulla circonferenza unità, per cui tutti gli zeri di $W(z)$ stanno strettamente all'interno del cerchio unitario.

Viceversa, essendo $\hat{W}(z)$ un fattore a fase minima di uno spettro privo di zeri in $\{|z| = 1\}$ scende, ancora dalla formula per l'inversa \hat{W}^{-1} , che deve valere la condizione 1. Inoltre $\hat{G}^{-1}\hat{F}$ e $\hat{K}^{-1}\hat{H}$ devono essere analitiche in $\{|z| \geq 1\}$. Questo implica che, a meno di cancellazioni di poli instabili comuni a (\hat{G}, \hat{F}) ed a (\hat{K}, \hat{H}) , le matrici $\hat{F}(z)$ e $\hat{H}(z)$ devono essere analitiche in $\{|z| \geq 1\}$. A questo punto riscriviamo la (2.25) nella forma

$$\hat{W}^{-1} = \text{diag}\{\hat{G}^{-1}, \hat{K}^{-1}\} \begin{bmatrix} I & -\hat{F} \\ -\hat{H} & I \end{bmatrix}.$$

Da questa uguaglianza si trae

$$\begin{bmatrix} \hat{G} & 0 \\ 0 & \hat{K} \end{bmatrix} = \begin{bmatrix} I & -\hat{F} \\ -\hat{H} & I \end{bmatrix} \hat{W},$$

³Questo significa che anche gli zeri di $W(z)$ sono in $\{|z| < 1\}$.

la quale implica che gli *unici* poli instabili di \hat{G} e \hat{K} sono quelli eventuali in comune con \hat{F} e (rispettivamente) \hat{H} . L'analiticità di $\hat{T}(z)$ scende infine dalla formula

$$\hat{T}(z) = \hat{W}(z) \operatorname{diag}\{\hat{G}(z)^{-1}, \hat{K}(z)^{-1}\}, \quad (2.26)$$

stante che $\hat{G}(z)^{-1}$, $\hat{K}(z)^{-1}$ sono analitici. \square

Nel modello a retroazione di innovazione, generato dal fattore a fase minima \hat{W} , ci possono quindi essere poli instabili nelle funzioni di trasferimento \hat{F} , \hat{H} , \hat{G} e \hat{K} . Il fatto che $\hat{G}^{-1}\hat{F}$, $\hat{K}^{-1}\hat{H}$ debbano essere comunque analitiche, può essere interpretato come una condizione di dipendenza causale dell'innovazione dal processo congiunto. Infatti dalla prima delle (2.22) si ha ad esempio

$$\hat{\mathbf{e}}_1(t) = \hat{G}(z)^{-1} [\mathbf{y}(t) - \hat{F}(z) \mathbf{u}(t)]$$

dalla quale si vede bene come le condizioni di analiticità 3. del teorema 7.1 servano a garantire che $\hat{\mathbf{e}}_1(t) \in \mathbf{H}_t(\mathbf{y}, \mathbf{u})$.

Anche per il modello a innovazioni normalizzate (2.24), si può dare una caratterizzazione analoga al teorema 7.1 che stabilisce in particolare che \bar{G} e \bar{K} debbono essere fattori spettrali canonici (a fase minima) dei rispettivi spettri $S_{\mathbf{v}}$ ed $S_{\mathbf{r}}$, "a meno di cancellazioni con eventuali fattori a poli instabili" presenti, rispettivamente, anche in \bar{F} e \bar{H} .

Invarianza del Modello a Reazione. Inizieremo questa sezione mostrando che un fattore spettrale analitico normalizzato a blocchi genera "quasi sempre" un modello a retroazione (internamente stabile).

C'è una difficoltà tecnica che si incontra nello stabilire questo risultato, che è dovuta alla possibile cancellazioni di poli instabili e che implica, come abbiamo già avuto modo di constatare, che fattori spettrali analitici possono generare modelli a retroazione instabili. Per affrontare queste questioni conviene fare uso di opportune rappresentazioni polinomiali fratte.

Siano

$$[F(z) \ G(z)] = A(z^{-1})^{-1} [B(z^{-1}) \ C(z^{-1})], \quad (2.27)$$

$$[H(z) \ K(z)] = D(z^{-1})^{-1} [M(z^{-1}) \ N(z^{-1})] \quad (2.28)$$

rappresentazioni polinomiali fratte coprime a sinistra in cui $A(z^{-1})$, $C(z^{-1})$, $D(z^{-1})$, $N(z^{-1})$, sono polinomi matriciali *monici* nell'indeterminata z^{-1} . Usando direttamente le espressioni simboliche corrispondenti nel modello a retroazione (2.3),

$$\begin{aligned} A(z^{-1})\mathbf{y}(t) - B(z^{-1})\mathbf{u}(t) &= C(z^{-1})\mathbf{w}_1(t) \\ D(z^{-1})\mathbf{u}(t) - M(z^{-1})\mathbf{y}(t) &= N(z^{-1})\mathbf{w}_2(t) \end{aligned}$$

si vede immediatamente che il fattore spettrale $W(z)$ corrispondente (2.8), si può rappresentare attraverso la r.p.f.

$$W(z) = \begin{bmatrix} A(z^{-1}) & -B(z^{-1}) \\ -M(z^{-1}) & D(z^{-1}) \end{bmatrix}^{-1} \begin{bmatrix} C(z^{-1}) & 0 \\ 0 & N(z^{-1}) \end{bmatrix}, \quad (2.29)$$

mentre, usando le

$$\begin{aligned} A(z^{-1})\mathbf{y}(t) - B(z^{-1})\mathbf{u}(t) &= A(z^{-1})\mathbf{v}(t) \\ D(z^{-1})\mathbf{u}(t) - M(z^{-1})\mathbf{y}(t) &= D(z^{-1})\mathbf{r}(t) \end{aligned}$$

si vede che la funzioni di trasferimento in catena chiusa $T(z)$ di (2.7) ha la r.p.f.

$$T(z) = \begin{bmatrix} A(z^{-1}) & -B(z^{-1}) \\ -M(z^{-1}) & D(z^{-1}) \end{bmatrix}^{-1} \begin{bmatrix} A(z^{-1}) & 0 \\ 0 & D(z^{-1}) \end{bmatrix}, \quad (2.30)$$

in cui il polinomio matriciale a denominatore ha esattamente la stessa struttura, triangolare inferiore all'infinito con blocchi unità sulla diagonale, del fattore W . Si potrebbe dire che le due r.p.f. sono *normalizzate a blocchi*.

Usando la caratterizzazione della coprimalità data al Captolo V, Proposizione 4.5, è possibile mostrare facilmente che la r.p.f. (2.30) è coprima a sinistra *se e solo se tali sono le r.p.f.*

$$F(z) = A(z^{-1})^{-1} B(z^{-1}) \quad H(z) = D(z^{-1})^{-1} M(z^{-1}) \quad (2.31)$$

(notare che questo non è necessariamente il caso anche se le rappresentazioni (2.27) e (2.28) sono coprime). In questo caso $T(z)$ è analitica in $\{|z| \geq 1\}$, se e solo se

$$\det \begin{bmatrix} A(z^{-1}) & -B(z^{-1}) \\ -M(z^{-1}) & D(z^{-1}) \end{bmatrix} \neq 0 \quad \text{per } |z| \geq 1 \quad (2.32)$$

Sfortunatamente la coprimalità delle r.p.f. (2.27) (2.28), anche assieme a quella delle (2.31), non garantisce affatto quella della r.p.f. (2.29) del fattore spettrale $W(z)$.

PROBLEMA 7.2. Verificare che nell'esempio (2.14) le (2.27) (2.28), (2.31) sono tutte coprime, c'è uno zero del determinante del denominatore di (2.30) in $z = 2$ che rende il modello a retroazione instabile ma la rpf (2.29) non è coprima e nel formare $W(z)$ questo zero viene cancellato. \diamond

Come visto nell'esempio, (2.29) non è in generale coprima e gli zeri del determinante del polinomio matriciale a denominatore in (2.29) non sono necessariamente tutti poli di $W(z)$. Questo evidentemente dipende dalla possibile cancellazione di questi zeri con zeri del numeratore (cioè con zeri comuni a G e/o a K) per cui $W(z)$ potrebbe risultare analitica anche se $T(z)$ non lo è, si veda a questo proposito l'esempio (2.14). Questa eventualità è relativamente poco importante per le applicazioni e conviene introdurre una condizione tecnica che la escluda.

DEFINIZIONE 7.4. Siano F , H , G , K , espresse mediante le r.p.f. coprime (2.27) (2.28), le funzioni di trasferimento dell'interconnessione a retroazione corrispondente attraverso la (2.13), al fattore spettrale $W(z)$. Diremo che la r.p.f. (2.29), ha denominatore stabile, se vale la condizione (2.32), ovvero, il determinante del polinomio denominatore non ha zeri in $\{|z| \geq 1\}$. Ovviamente un fattore $W(z)$ a denominatore stabile è analitico, indipendentemente dalla coprimalità.

Sotto questa condizione, l'analiticità di $W(z)$ e di $T(z)$ sono essenzialmente equivalenti, come stabilito dalla proposizione che segue.

PROPOSIZIONE 7.4 (Anderson, Gevers). *Se le r.p.f. (2.27) (2.28) sono coprime, le due affermazioni seguenti sono equivalenti:*

- (1) *Il fattore spettrale $W(z)$ ha denominatore stabile (in particolare $W(z)$ è analitico in $\{|z| \geq 1\}$).*
- (2) *La funzione di trasferimento corrispondente, $T(z)$, è analitica in $\{|z| \geq 1\}$ e gli eventuali poli instabili di G e di K sono in comune rispettivamente a F e a H ; più precisamente, si cancellano formando, rispettivamente, i prodotti $G^{-1}F$ e $K^{-1}H$.*

Inoltre $W(z)$ è a fase minima se e solo se valgono la (2.32) e la

$$\det C(z^{-1}) \det N(z^{-1}) \neq 0 \quad \text{per } |z| \geq 1 \quad (2.33)$$

DIMOSTRAZIONE. (2 \Rightarrow 1) Siano P e Q i polinomi M.C.D. sinistri di (A, B) e di (D, M) , ovvero sia

$$A = PA_1, \quad B = PB_1, \quad D = QD_1, \quad M = QM_1$$

e $F = A_1^{-1}B_1$, $H = D_1^{-1}M_1$ le relative fattorizzazioni coprime. Nella discussione del Problema 7.2 si è mostrato che l'analiticità di $T(z)$ è equivalente a

$$\det \begin{bmatrix} A_1(z^{-1}) & -B_1(z^{-1}) \\ -M_1(z^{-1}) & D_1(z^{-1}) \end{bmatrix} \neq 0 \quad \text{per } |z| \geq 1$$

Pertanto,

$$\det \begin{bmatrix} A(z^{-1}) & -B(z^{-1}) \\ -M(z^{-1}) & D(z^{-1}) \end{bmatrix} = \det \begin{bmatrix} P(z^{-1}) & 0 \\ 0 & Q(z^{-1}) \end{bmatrix} \det \begin{bmatrix} A_1(z^{-1}) & -B_1(z^{-1}) \\ -M_1(z^{-1}) & D_1(z^{-1}) \end{bmatrix},$$

e gli eventuali zeri in $\{|z| \geq 1\}$ del primo membro debbono necessariamente essere zeri di $\det P(z^{-1}) \det Q(z^{-1})$. Ora, dato che la rappresentazione (2.27) è coprima, P e C non possono avere fattori polinomiali sinistri comuni non banali e pertanto P è un divisore "effettivo" del denominatore di G , nel senso che gli zeri di $\det P(z^{-1})$ sono tutti poli di $G(z)$ e quindi gli eventuali zeri instabili di $\det P(z^{-1})$ sono tutti poli instabili di $G(z)$. Per la condizione (2) dell'enunciato questi poli debbono essere in comune ad F e quindi debbono essere cancellati formando il prodotto $G(z)^{-1}F(z) = C(z^{-1})^{-1}P(z^{-1})B_1(z^{-1})$, ma questo è impossibile perchè C e P sono coprime a sinistra. Si vede così che $\det P(z^{-1})$ non può avere zeri instabili.

Con un ragionamento perfettamente analogo, si mostra che nemmeno $\det Q(z^{-1})$ può avere zeri instabili.

(1 \Rightarrow 2): Dalla fattorizzazione introdotta più sopra, segue immediatamente che $\det P$ e $\det Q$ possono avere zeri solo in $\{|z| < 1\}$ e $\det \begin{bmatrix} A_1(z^{-1}) & -B_1(z^{-1}) \\ -M_1(z^{-1}) & D_1(z^{-1}) \end{bmatrix}$ può avere zeri solo in $\{|z| < 1\}$.

Infine, se $W(z)$ è a fase minima, segue dalle affermazioni (1) e (3) del Teorema 7.1 (che è un caso particolare del risultato che stiamo dimostrando) che i M.C.D. sinistri di (A, C) e di (D, N) , diciamoli R ed S

$$A = RA_1, \quad C = RC_1, \quad D = SD_1, \quad N = SN_1$$

debbono essere polinomi stabili perché ogni eventuale zero in $\{|z| \geq 1\}$ di $\det R$ o $\det S$ dovrebbe essere anche zero del denominatore di F o H e questa possibilità è esclusa dalla coprimalità delle r.p.f. (2.27), (2.28). La condizione (2.33) è allora implicata dal fatto che C_1 e N_1 sono stabili (infatti gli zeri di G e K sono tutti in $\{|z| < 1\}$). \square

Discende dal risultato appena dimostrato che ogni fattore spettrale analitico normalizzato a blocchi che ha denominatore stabile⁴, si può realizzare mediante un modello a retroazione (internamente stabile). Il teorema afferma anche che,

⁴Come si è visto, questa è una condizione tecnica che serve a escludere la possibilità di cancellazioni polo-zero instabili nella r.p.f. (2.29) di $W(z)$. La sua presenza è responsabile della qualificazione "quasi sempre", usata nell'introduzione di questa sezione.

viceversa, i modelli a retroazione che hanno F e G fissate e G e K fattori spettrali normalizzati arbitrari di $S_v(z)$ e $S_r(z)$, ma tali che gli eventuali poli instabili di G e di K siano in comune rispettivamente a F e ad H , generano fattori spettrali normalizzati a blocchi a denominatore stabile.

La condizione di poli instabili in comune si può esprimere anche nella forma⁵

$$\delta \left[F_+ : G_+ \right] = \delta [F_+] \quad (2.34)$$

$$\delta \left[H_+ : K_+ \right] = \delta [H_+] \quad (2.35)$$

dove $\delta[\cdot]$ è il grado di MacMillan e F_+ è la *parte instabile* di F (cioè lo sviluppo in frazioni parziali relativo ai poli in $\{|z| \geq 1\}$).

Presenteremo ora un risultato notevole che particolarizza ai fattori spettrali normalizzati a blocchi a denominatore stabile, la parametrizzazione dei fattori analitici (formula (4.26), Teorema 3.4) data al capitolo 4. Da questo enunciato scenderà anche la conclusione, a priori abbastanza sorprendente, che “genericamente” una coppia di processi stocastici stazionari ammette un *unico modello a retroazione*. Questa proprietà di *invarianza* del modello a retroazione è una proprietà fondamentale che torna utile in molte applicazioni.

TEOREMA 7.2 (Ng, Goodwin, Anderson, Gevers). *Ogni fattore spettrale razionale analitico, normalizzato a blocchi, $W(z)$, di $S(z)$, a denominatore stabile, è diagonalmente equivalente a quello a fase minima.*

DIMOSTRAZIONE. Supponiamo di esprimere le funzioni di trasferimento del modello a reazione derivato da $W(z)$ mediante r.p.f. coprime come in (2.27) e (2.28). Per ipotesi $W(z)$ ha una rappresentazione polinomiale del tipo (2.29) in cui il denominatore è un polinomio matriciale stabile (cioè il suo determinante soddisfa la (2.32)). Gli spettri polinomiali

$$\Psi_A(z) := C(z^{-1})Q_1C'(z)' \quad \Psi_D(z) := N(z^{-1})Q_2N'(z)'$$

hanno entrambi rango pieno e possono essere fattorizzati mediante fattori polinomiali a fase minima triangolari inferiormente all'infinito (vedere l'algoritmo di Bauer-Rissanen, Teorema 4.3, Cap. V). Estraendo i fattori triangolari, si può coscrivere

$$\Psi_A(z) := \bar{C}(z^{-1})L_C L'_C \bar{C}'(z)' \quad \Psi_D(z) := \bar{N}(z^{-1})L_N L'_N \bar{N}'(z)'$$

dove i polinomi matriciali \bar{C} e \bar{N} sono monici, $\det \bar{C}(z^{-1}), \det \bar{N}(z^{-1}) \neq 0$ per $|z| \geq 1$, e

$$\hat{Q}_1 := L_C L'_C > 0 \quad \hat{Q}_2 := L_N L'_N > 0.$$

Queste due matrici sono le varianze dei processi d'innovazione. In effetti il fattore spettrale normalizzato a blocchi

$$\hat{W}(z) = \begin{bmatrix} A(z^{-1}) & -B(z^{-1}) \\ -M(z^{-1}) & D(z^{-1}) \end{bmatrix}^{-1} \begin{bmatrix} \bar{C}(z^{-1}) & 0 \\ 0 & \bar{N}(z^{-1}) \end{bmatrix}$$

è a fase minima, dato che

$$\hat{W}(z) = W(z) \text{diag}\{\hat{V}_1(z) \hat{V}_2(z)\}$$

⁵L'equivalenza di queste due condizioni é dimostrata in dettaglio nel lemma 8 dell'articolo [54].

dove,

$$\begin{aligned}\hat{V}_1(z) &:= C(z^{-1})^{-1}\bar{C}(z^{-1}) \\ \hat{V}_2(z) &:= N(z^{-1})^{-1}\bar{N}(z^{-1})\end{aligned}$$

ed è diagonalmente equivalente a $W(z)$. \square

PROBLEMA 7.3. Il fattore spettrale 2×2 dell'esempio (2.14)

$$W(z) = \frac{1}{z+0.8} \begin{bmatrix} z-0.7 & 1 \\ 1.95 & z-0.5 \end{bmatrix},$$

è analitico ma non a denominatore stabile. Infatti alle funzioni di trasferimento

$$\begin{aligned}F(z) &= \frac{1}{z-0.5} & G(z) &= \frac{z-2}{z-0.5} \\ H(z) &= \frac{1.95}{z-0.7} & K(z) &= \frac{z-2}{z-0.7}\end{aligned}$$

corrisponde il polinomio denominatore

$$\begin{bmatrix} A(z^{-1}) & -B(z^{-1}) \\ -M(z^{-1}) & D(z^{-1}) \end{bmatrix} = \begin{bmatrix} 1-0.5z^{-1} & -1z^{-1} \\ -1.95z^{-1} & 1-0.7z^{-1} \end{bmatrix},$$

il cui determinante è $(1-2z^{-1})(1+0.8z^{-1})$. Dalla rappresentazione coprime (vedere l'esercizio 5-6, Cap V)

$$W(z) = \begin{bmatrix} z+0.8 & 0 \\ -(z-0.5) & 1 \end{bmatrix}^{-1} \begin{bmatrix} z-0.7 & 1 \\ -(z-2) & 0 \end{bmatrix},$$

si vede che $W(z)$ ha uno zero instabile in $z=2$ e pertanto non è a fase minima.

Dimostrare che $W(z)$ non può essere diagonalmente equivalente al fattore a fase minima. \diamond

Il risultato fondamentale di questa sezione, sull'invarianza del modello a reazione, si può allora formulare nel modo seguente.

TEOREMA 7.3. *Tutti i fattori spettrali razionali quadrati dello spettro congiunto $S(z)$, normalizzati a blocchi e a denominatore stabile, sono tra loro diagonalmente equivalenti. Ne segue che tutte le rappresentazioni a retroazione dei due processi stazionari di spettro congiunto $S(z)$ hanno la stessa funzione di trasferimento in catena di azione diretta $F(z)$, la stessa funzione di trasferimento del canale di retroazione $H(z)$ e gli stessi spettri dei segnali di riferimento $S_{\mathbf{r}}(z)$ e di disturbo $S_{\mathbf{v}}(z)$.*

In altri termini, nonostante i fattori spettrali quadrati stabili e normalizzati di $S(z)$ possano essere infiniti, il modello a retroazione corrispondente è sostanzialmente *unico*.

DIMOSTRAZIONE. Per la dimostrazione del teorema basta notare che un qualunque fattore quadrato normalizzato a blocchi a denominatore stabile, è diagonalmente equivalente al fattore normalizzato a fase minima (Teorema 7.2). Dato che l'equivalenza diagonale è una relazione di equivalenza (quindi in particolare transitiva), questo fatto implica immediatamente la prima affermazione del teorema. L'affermazione sull'invarianza di F e H e degli spettri $S_{\mathbf{r}}(z)$ e $S_{\mathbf{v}}(z)$ è stata già giustificata nella proposizione 7.2. \square

La proprietà di unicità del modello a reazione corrispondente ad un dato spettro congiunto $S(z)$, è un fenomeno particolarmente importante nell'*identificazione* di sistemi operanti in condizioni di “normale funzionamento” quando, di norma, è assai difficile escludere che vi sia influenza mutua (reazione) fra i segnali misurati in ingresso e in uscita al sistema. La presenza di reazione potrebbe far sorgere un problema di *identificabilità* in catena chiusa. In termini astratti, il problema a cui ci riferiamo è il seguente:

PROBLEMA 7.4 (Identificabilità in catena chiusa). *Supponiamo che il processo congiunto $\{\mathbf{y}(t), \mathbf{u}(t)\}$ sia effettivamente costituito dall'uscita e dall'ingresso di un sistema di controllo lineare retroazionato con funzioni di trasferimento d'anello “vere” $F_0(z)$, $H_0(z)$ e soggetto a segnali stocastici di riferimento e di disturbo di spettri rispettivi $S_r(z)$ e $S_v(z)$. Supponiamo che sia possibile misurare i segnali $\{\mathbf{y}(t), \mathbf{u}(t)\}$ per un intervallo infinitamente lungo di tempo e quindi di poter risalire esattamente allo spettro congiunto $S(z)$ dei due segnali. Ci si chiede se sia possibile ricavare univocamente $F_0(z)$ (ed eventualmente anche $H_0(z)$ e gli spettri $S_r(z)$, $S_v(z)$), dalle misure effettuate, ovvero dalla conoscenza dello spettro congiunto $S(z)$. \square*

Questo problema e alcune sue varianti sono stati studiati lungamente in letteratura [4, 54]. Il teorema appena enunciato stabilisce che “genericamente” il problema dell'identificabilità in catena chiusa ha risposta affermativa. Con questo intendiamo che esiste una applicazione “genericamente” biiettiva che associa ad uno spettro congiunto (a rango pieno e senza zeri sulla circonferenza unità) $S(z)$ un unico modello a retroazione ($F(z)$, $H(z)$, $S_r(z)$, $S_v(z)$) (si potrebbe dire la sua “realizzazione” mediante uno schema a retroazione) e viceversa.

3. Reazione fra Processi Stocastici e Causalità

Sembra intuitivamente chiaro che la *reazione* fra i due processi $\{\mathbf{y}(t)\}$ e $\{\mathbf{u}(t)\}$ sia dovuta alla presenza del termine $H(z)$ nelle equazioni del modello a retroazione (2.3). Si potrebbe così definire “senza reazione” la coppia $(\{\mathbf{y}(t)\}, \{\mathbf{u}(t)\})$, quando $H(z) \equiv 0$. Purtroppo questa condizione dipende da una particolare normalizzazione a blocchi dei fattori spettrali e possono darsi dei casi di modelli a retroazione in cui $H(z) \equiv 0$ in una particolare rappresentazione, mentre $H(z) \neq 0$ per altre (si veda l'esempio alla fine di questa sezione). Questo fatto è da certi punti di vista insoddisfacente e si desidera una definizione “intrinseca” di assenza di reazione che sia indipendente dal particolare modello che si sta usando.

Una definizione abbastanza intuitiva proposta da Granger⁶ [59] richiede che ad ogni istante t l'evoluzione futura dell'ingresso sia condizionatamente scorrelata (indipendente, nel caso gaussiano) dal passato di $\{\mathbf{y}(t)\}$, se è nota la storia passata, \mathbf{u}^{t-1} , del processo di ingresso.

DEFINIZIONE 7.5. *La coppia di processi del secondo ordine $(\{\mathbf{y}(t)\}, \{\mathbf{u}(t)\})$ è “senza reazione” se*

$$\hat{E} [\mathbf{u}(t+k) | \mathbf{y}^t \mathbf{u}^{t-1}] = \hat{E} [\mathbf{u}(t+k) | \mathbf{u}^{t-1}] \quad (3.1)$$

per ogni $t \in \mathbb{Z}$ e per ogni $k \geq 0$.

⁶La prima discussione del concetto di causalità nel contesto dei processi stocastici, che come vedremo è equivalente all'assenza di reazione, risale a N. Wiener [151] nel 1956.

Agli effetti del calcolo del predittore lineare a minima varianza di $\mathbf{u}(t+k)$ in base ad $(\mathbf{y}^t, \mathbf{u}^{t-1})$ la storia passata di $\{\mathbf{y}(t)\}$ non aggiunge quindi alcuna informazione utile che non sia già contenuta in quella $(\mathbf{H}_{t-1}(u))$ del processo di ingresso.

Analizzando la (3.1), si vede che questa relazione può essere interpretata geometricamente come una relazione di *ortogonalità (o incorrelazione) condizionata* fra le variabili aleatorie del sottospazio $\mathbf{H}_t^+(u) = \overline{\text{span}}\{\mathbf{u}(s) ; s \geq t\}$ (il futuro all'istante t di $\{\mathbf{u}(t)\}$), e quella del *passato*, $\mathbf{H}_t(\mathbf{y})$, di $\{\mathbf{y}(t)\}$, dato il passato “stretto”, $\mathbf{H}_{t-1}(u)$, di $\{\mathbf{u}(t)\}$ allo stesso istante.

Questa condizione si scrive simbolicamente come

$$\mathbf{H}_t^+(\mathbf{u}) \perp \mathbf{H}_t(\mathbf{y}) \mid \mathbf{H}_{t-1}(\mathbf{u})$$

e si legge: “ $\mathbf{H}_t^+(\mathbf{u})$ e $\mathbf{H}_t(\mathbf{y})$ sono condizionatamente ortogonali dato $\mathbf{H}_{t-1}(\mathbf{u})$ ”. La condizione, per sua stessa natura, è *simmetrica* rispetto ad $\mathbf{H}_t^+(\mathbf{u})$ e $\mathbf{H}_t(\mathbf{y})$.

Notiamo che, a rigore, nella definizione si sarebbe dovuto parlare di assenza di reazione *in senso lato* giacché la condizione (3.1) riguarda solo i momenti congiunti del secondo ordine dei due processi. Si potrebbe in effetti introdurre anche una nozione di assenza di reazione *in senso stretto*, esprimibile in termini di *indipendenza condizionata* del futuro di $\{\mathbf{u}(t)\}$ dal passato di $\{\mathbf{y}(t)\}$, nota la storia passata di $\{\mathbf{u}(t)\}$, (equivalente, nel caso di processi Gaussiani a quella data). Questa nozione non interverrà mai direttamente in questo libro e noi non ce ne occuperemo più. Invece riserveremo la denominazione di “assenza di reazione in senso lato” a una condizione un pò più debole della (3.1) che introdurremo più sotto.

Ortogonalità Condizionata. Ricordiamo, per comodità del lettore, che due sottospazi di variabili aleatorie del secondo ordine \mathbf{X}, \mathbf{Y} , si definiscono *condizionatamente correlati* o, meglio, *condizionatamente ortogonali*, dato un terzo sottospazio di variabili del secondo ordine \mathbf{Z} , se risulta

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \hat{E}(\mathbf{x} \mid \mathbf{Z}), \hat{E}(\mathbf{y} \mid \mathbf{Z}) \rangle \quad (3.2)$$

per ogni $\mathbf{x} \in \mathbf{X}$, $\mathbf{y} \in \mathbf{Y}$. Che questa sia la naturale versione “debole” dell’indipendenza condizionata si può vedere supponendo che \mathbf{X} , \mathbf{Y} e \mathbf{Z} siano popolati da variabili congiuntamente Gaussiane. Notando che la (3.2) è equivalente alla

$$\langle \mathbf{x} - E(\mathbf{x} \mid \mathbf{Z}), \mathbf{y} - E(\mathbf{y} \mid \mathbf{Z}) \rangle = \text{cov}(\mathbf{x}, \mathbf{y} \mid \mathbf{Z}) = 0 \quad ,$$

si vede che la *covarianza incrociata condizionata di \mathbf{x} e \mathbf{y} , date le variabili del sottospazio \mathbf{Z}* , è zero e la densità congiunta condizionata $p_{\mathbf{xy}}(\cdot \mid \mathbf{Z})$ fattorizza quindi nel prodotto delle due densità $p_{\mathbf{x}}(\cdot \mid \mathbf{Z}) p_{\mathbf{y}}(\cdot \mid \mathbf{Z})$ (si può inizialmente supporre \mathbf{Z} generato da un vettore z di dimensione finita e notare che il ragionamento vale, indipendentemente dalla dimensione di \mathbf{Z}). La nozione intuitiva di incorrelazione condizionata è poi catturata in modo esplicito dal seguente enunciato (la cui dimostrazione si può ad esempio trovare in [94]).

LEMMA 7.1. *Si ha $\mathbf{X} \perp \mathbf{Y} \mid \mathbf{Z}$ se e solo se vale una delle due condizioni equivalenti*

- (i) $\hat{E}(\mathbf{x} \mid \mathbf{Y} \vee \mathbf{Z}) = \hat{E}(\mathbf{x} \mid \mathbf{Z}) \quad \mathbf{x} \in \mathbf{X}$
- (ii) $\hat{E}(\mathbf{y} \mid \mathbf{X} \vee \mathbf{Z}) = \hat{E}(\mathbf{y} \mid \mathbf{Z}) \quad \mathbf{y} \in \mathbf{Y}$

dove $\mathbf{X} \vee \mathbf{Y}$ denota lo spazio di Hilbert generato dalle variabili aleatorie dei sottospazi \mathbf{X} e \mathbf{Y} (ovvero la chiusura della somma vettoriale $\mathbf{X} + \mathbf{Y}$).

Come si vede, le (i) e (ii) sono condizioni di “non dipendenza” di un predittore da una parte dei dati, esattamente dello stesso tipo della (3.1).

Il risultato seguente fornisce una caratterizzazione dei processi senza reazione e chiarisce in modo definitivo il legame tra assenza di reazione e causalità.

TEOREMA 7.4 (Caines, Chan, Gevers, Anderson). *La coppia $(\{\mathbf{y}(t)\}, \{\mathbf{u}(t)\})$ è senza reazione se e solo se vale una qualunque delle seguenti condizioni.*

- (1) Per ogni $s \leq t$ e per ogni $t \in \mathbb{Z}$, risulta

$$\hat{E}[\mathbf{y}(s) | \mathbf{H}(\mathbf{u})] = \hat{E}[\mathbf{y}(s) | \mathbf{H}_{t-1}(\mathbf{u})] = \hat{E}[\mathbf{y}(s) | \mathbf{u}^{t-1}] \quad ; \quad (3.3)$$

in particolare per $s = t$, lo stimatore lineare $\hat{E}[\mathbf{y}(t) | \mathbf{u}^{t-1}]$ coincide con quello, $\hat{E}[\mathbf{y}(t) | \mathbf{H}(\mathbf{u})]$, basato su tutta la storia di $\{\mathbf{u}(t)\}$.

- (2) Il fattore spettrale a fase minima, $\hat{W}(z)$, normalizzato all'infinito ($\hat{W}(\infty) = I$), dello spettro congiunto dei due processi ha il blocco di indice 21 identicamente nullo,

$$\hat{W}_{21}(z) \equiv 0 \quad (3.4)$$

e la varianza dell'innovazione, Λ , è diagonale a blocchi. Quest'ultima condizione è equivalente a dire che le due componenti dell'innovazione $\mathbf{e}(t) = [\mathbf{e}_1(t)', \mathbf{e}_2(t)']$ di dimensione m e p sono completamente scorrelate e che $\mathbf{e}_2(t) = \hat{\mathbf{e}}_2(t)$.

- (3) Nel modello a retroazione d'innovazione (2.22), i processi $\{\mathbf{u}(t)\}$ ed $\{\mathbf{e}_1(t)\}$ sono scorrelati

$$E \mathbf{u}(t) \mathbf{e}_1(s)' = 0 \quad , \quad \forall t, s \in \mathbb{Z} \quad . \quad (3.5)$$

- (4) Nei modelli a retroazione d'innovazione (2.24) e (2.22) il blocco di retroazione (dall'uscita) è nullo, i.e.,

$$\bar{H}(z) \equiv 0 \quad , \quad \hat{H}(z) \equiv 0 \quad . \quad (3.6)$$

Osservazione. La (3.3) può essere interpretata come la definizione di *causalità stocastica*.

DIMOSTRAZIONE. L'equivalenza fra assenza di reazione, (3.1), e causalità, (3.3), si riconduce al fatto che le due condizioni sono due modi diversi di esprimere l'ortogonalità condizionata di $\mathbf{H}_t^+(\mathbf{u})$ e $\mathbf{H}_t(\mathbf{y})$ dato $\mathbf{H}_{t-1}(\mathbf{u})$ (lemma 7.1).

Mostriamo che l'assenza di reazione implica la (3.5) (condizione (3)).

Dalla (2.21) si ha $\hat{\mathbf{e}}_2(t) = \mathbf{u}(t) - \hat{E}[\mathbf{u}(t) | \mathbf{y}^t \mathbf{u}^{t-1}] = \mathbf{u}(t) - \hat{E}[\mathbf{u}(t) | \mathbf{u}^{t-1}] =$ all'innovazione, $\mathbf{e}_u(t)$, del processo $\{\mathbf{u}(t)\}$. Dato che $\{\mathbf{e}_1(t)\} \perp \{\hat{\mathbf{e}}_2(t)\}$ e, come si è visto, $\mathbf{H}(\hat{\mathbf{e}}_2) = \mathbf{H}(\mathbf{u})$ in assenza di reazione, segue la (3).

È immediato che (3) \Rightarrow (4). Difatti, moltiplicando la

$$\mathbf{u}(t) = (I - \hat{H}\hat{F})^{-1} \hat{H}\hat{G} \mathbf{e}_1(t) + (I - \hat{H}\hat{F})^{-1} \hat{K} \hat{\mathbf{e}}_2(t)$$

per $\mathbf{e}_1(s)^T$ e calcolando l'aspettazione, si trova

$$\left(I - \hat{H}(z) \hat{F}(z) \right)^{-1} \hat{H}(z) \hat{G}(z) \Lambda_1 \equiv 0 \quad ,$$

identicamente. Dato che $\Lambda_1 > 0$, $\hat{G}(\infty) = I_m$ e $(I - \hat{H}(\infty) \hat{F}(\infty))^{-1} = I_p$, l'unica possibilità è che sia $\hat{H}(z) \equiv 0$. Un ragionamento analogo può essere fatto sul modello d'innovazione (2.24).

Che (4) \Rightarrow (3) segue immediatamente dal fatto che se $\hat{H} \equiv 0$ nel modello d'innovazione (2.22), necessariamente deve essere $\hat{\mathbf{e}}_2(t) = \mathbf{e}_u(t)$ e quindi $\{\hat{\mathbf{e}}_1(t)\} \perp \{\mathbf{e}_u(t)\}$ il che implica $\{\mathbf{e}_1(t)\} \perp \{\mathbf{u}(t)\}$.

Che ((3) + (4)) \Rightarrow (2) è conseguenza immediata della formula (2.8) per $\hat{W}(z)$ e della (2.23) la quale, con $\{\mathbf{u}(t)\} \perp \{\mathbf{e}_1(t)\}$ (condizione (3)), dice proprio che $\hat{\mathbf{e}}_2(t) = \mathbf{e}_2(t)$.

Infine, dalla (3.4) si trae, usando la formula di inversione (2.13), che $\hat{H} \equiv 0$, \hat{K} e \hat{G} sono a fase minima e \hat{F} è *analitica in* $\{|z| \geq 1\}$. Questo, oltre a giustificare l'implicazione (2) \Rightarrow (4), serve a dimostrare anche l'implicazione mancante (2) \Rightarrow (1). Infatti, calcolando la proiezione $\hat{E}[\mathbf{y}(t) | \mathbf{H}(\mathbf{u})]$, si trova

$$\hat{E}[\mathbf{y}(t) | \mathbf{H}(\mathbf{u})] = \hat{F}(z) \mathbf{u}(t) = \hat{E} \left[\hat{G}(z) \mathbf{e}_1(t) | \mathbf{H}(\mathbf{u}) \right] = \hat{F}(z) \mathbf{u}(t)$$

e il secondo membro è una funzione *strettamente causale* di $\{\mathbf{u}(t)\}$ dato che $\hat{F}(\infty) = 0$ e $\hat{F}(z)$ è analitica in $\{|z| \geq 1\}$. Pertanto $\hat{E}[\mathbf{y}(t) | \mathbf{H}(\mathbf{u})] = \hat{E}[\mathbf{y}(t) | \mathbf{u}^{t-1}]$, $\forall t$, e questa uguaglianza è equivalente alla (3.3).

Con questo ultimo passaggio abbiamo chiuso il ciclo

$$(1) \iff (2.21) \implies (3) \iff (4) \iff (2) \implies (1)$$

e completata la dimostrazione. \square

Come abbiamo visto anche nel corso della dimostrazione, se non c'è reazione non possono esservi cancellazioni di poli instabili e l'enunciato del Teorema 7.1 vale senza condizioni. In particolare *nei modelli a retroazione d'innovazione \hat{F} ed \hat{F} sono analitiche in $\{|z| \geq 1\}$ e \hat{G} e \hat{G} strettamente a fase minima.*

Le condizioni (2), (3), (4) nell'enunciato si riferiscono al fattore spettrale canonico (o a fase minima) e al relativo modello a retroazione *d'innovazione*. È abbastanza naturale chiedersi se esse non debbano valere in generale, per un *qualunque* fattore spettrale e per il modello a retroazione corrispondente. La risposta è in generale negativa. Come abbiamo già avuto modo di vedere, l'ortogonalizzazione dei segnali di riferimento e disturbo e la normalizzazione a blocchi del fattore spettrale giocano un ruolo fondamentale nell'interpretazione dei modelli a retroazione.

ESEMPIO 7.3 (Gevers-Anderson). Consideriamo il modello a retroazione

$$\begin{aligned} \mathbf{y}(t) &= \frac{1/5}{z + 1/2} \mathbf{u}(t) + \frac{z + 3/5}{z + 4/5} \mathbf{e}_1(t) \\ \mathbf{u}(t) &= \frac{z + 1/2}{z + 3/4} \mathbf{e}_2(t) \end{aligned}$$

dove $\{\mathbf{e}_1(t)\} \perp \{\mathbf{e}_2(t)\}$ e $\text{var } \mathbf{e}_1(t) = \text{var } \mathbf{e}_2(t) = 1$.

È facile verificare che questo modello è d'innovazione e, dato che $H(z) = 0$, non c'è reazione da y a u . Il relativo fattore a fase minima ha la r.p.f. coprima

$$\hat{W}(z) = D(z)^{-1} \hat{N}(z) = \begin{bmatrix} z + 4/5 & -1/5 \\ 0 & z + 3/4 \end{bmatrix}^{-1} \begin{bmatrix} z + 3/5 & 6/5 \\ 0 & z + 1/2 \end{bmatrix} .$$

Partendo da questa rappresentazione si può calcolare una fattorizzazione spettrale polinomiale $\hat{N}(z) \hat{N}(1/z)^T = \bar{N}(z) Q \bar{N}(1/z)^T$ dove

$$\bar{N}(z) = \begin{bmatrix} z + 2.7826 & 0.8068 \\ -1.085 & z + 0.882 \end{bmatrix} , \quad Q = \begin{bmatrix} 0.159 & 0.197 \\ 0.197 & 0.813 \end{bmatrix} ,$$

e costruire la matrice razionale analitica $\bar{W}(z) := D(z)^{-1} \bar{N}(z)$ che soddisfa la relazione di fattorizzazione spettrale $\bar{W}(z) Q \bar{W}(1/z)' = \hat{W}(z) \hat{W}(1/z)'$. Da notare che in questo esempio $\bar{W}(z)$ non è normalizzato a blocchi.

Si vede facilmente che $\bar{W}_{21} \neq 0$ e il modello a retroazione corrispondente ha $H(z) \neq 0$. \diamond

Assenza di reazione “in senso lato”. L'assenza di reazione potrebbe essere definita da una condizione meno stringente della (3.1), richiedendo che il futuro, $\mathbf{H}_t^+(\mathbf{u})$, e il passato di $\{\mathbf{y}(t)\}$, $\mathbf{H}_t(\mathbf{y})$, siano condizionatamente ortogonali dato il passato *all'istante* t , $\mathbf{H}_t(\mathbf{u})$, di $\{\mathbf{u}(t)\}$,

$$\mathbf{H}_t^+(\mathbf{u}) \perp \mathbf{H}_t(\mathbf{y}) \mid \mathbf{H}_t(\mathbf{u}) \quad (3.7)$$

Questa condizione ha un aspetto più simmetrico, ma non si lega in modo naturale alla struttura dei modelli a retroazione, nel senso della Definizione 7.5, per i quali si richiede l'esistenza di almeno un ritardo (causalità stretta) in $F(z)$.

PROBLEMA 7.5. Dimostrare che l'assenza di reazione implica l'assenza di reazione in senso lato, i.e. $\mathbf{H}_t^+(\mathbf{u}) \perp \mathbf{H}_t(\mathbf{y}) \mid \mathbf{H}_{t-1}(\mathbf{u})$ implica $\mathbf{H}_t^+(\mathbf{u}) \perp \mathbf{H}_t(\mathbf{y}) \mid \mathbf{H}_t(\mathbf{u})$, ma non viceversa.

Suggerimento: Per il Teorema 7.4, (1), l'assenza di reazione è equivalente alla causalità “stretta” $\hat{E}[\mathbf{y}(t) \mid \mathbf{H}(\mathbf{u})] = \hat{E}[\mathbf{y}(t) \mid \mathbf{u}^{t-1}]$, ($F(z)$ strettamente propria), che implica la causalità “in senso lato”, i.e.

$$\hat{E}[\mathbf{y}(t) \mid \mathbf{H}(\mathbf{u})] = \hat{E}[\mathbf{y}(t) \mid \mathbf{u}^t]$$

che si ha se $F(z)$ è propria ma non strettamente propria, in generale. Quindi il viceversa in generale non vale. \diamond

4. Modelli ARMAX e predizione in presenza di ingressi osservabili

Come accennato nell'introduzione di questo capitolo, in molti problemi pratici il modello del canale di reazione è incognito o mal noto ed è necessario cercare di risolvere i problemi di stima su \mathbf{y} facendo riferimento solo al modello della catena di azione diretta che noi qui riscriveremo nella forma

$$\mathbf{y}(t) = F(z) \mathbf{u}(t) + G(z) \mathbf{w}(t) \quad , \quad (4.1)$$

dove abbiamo eliminato il pedice 1, perché normalmente ci riferiremo solo a rappresentazioni di questo tipo. Oltre a supporre naturalmente che \mathbf{w} sia un processo bianco di varianza Q definita positiva, supporremo qui che le funzioni di trasferimento F e G soddisfino alle seguente ipotesi.

IPOTESI 1. *Nel modello (4.1) $F(z)$ e $G(z)$ sono matrici razionali che soddisfano alle seguenti condizioni*

- (1) *C'è almeno un ritardo in F e G è normalizzata all'infinito, i.e. $F(\infty) = 0$ e $G(\infty) = I$.*
- (2) *$G(z)^{-1}$ e $G(z)^{-1}F(z)$ sono analitiche in $\{|z| \geq 1\}$.*

Come vedremo presto, queste ipotesi sono sostanzialmente equivalenti a dire che (4.1) si può interpretare come la “prima metà” di un modello a retroazione, in accordo con quanto abbiamo imparato finora sulla struttura generale di questi modelli.

Inizialmente supporremo per semplicità che $\{\mathbf{y}(t)\}$ sia un processo *scalare* ($m = 1$). La razionalità permette di scrivere le funzioni di trasferimento $F(z)$ e $G(z)$ nella forma

$$F(z) = N_F(z)/D_F(z) \quad , \quad G(z) = N_G(z)/D_G(z) \quad ,$$

dove N_F è un polinomio vettoriale p -dimensionale e D_F, N_G, D_G sono polinomi scalari in z .

Riducendo a denominatore comune, la (4.1) può evidentemente risciversi come

$$A^*(z^{-1}) \mathbf{y}(t) = B^*(z^{-1}) \mathbf{u}(t) + C^*(z^{-1}) \mathbf{w}(t) \quad , \quad (4.2)$$

dove

$$\begin{aligned} A(z) &= D_F(z) D_G(z) \\ B(z) &= N_F(z) D_G(z) \\ C(z) &= N_G(z) D_F(z) \quad . \end{aligned}$$

Supponiamo di avere eliminato tutti i fattori comuni ai tre polinomi $A(z), B(z), C(z)$. Per le ipotesi fatte è evidente che $C(z)$ ha allora tutti gli zeri interni al cerchio unità. Notiamo che $A(z)$ ha gli zeri all'interno del cerchio solo se non ci sono cancellazioni di poli instabili, nel senso spiegato nelle sezioni precedenti. Per la causalità possiamo supporre che A, B, C abbiano uguale grado come polinomi in z . Potremo quindi scrivere

$$\begin{aligned} A^*(z^{-1}) &= 1 + a_1 z^{-1} + \dots + a_n z^{-n} \\ B^*(z^{-1}) &= b_1 z^{-1} + \dots + b_n z^{-n} \\ C^*(z^{-1}) &= 1 + c_1 z^{-1} + \dots + c_n z^{-n} \end{aligned}$$

dove eventualmente qualche coefficiente (b_k) e/o (c_k) può essere nullo. Se si fosse partiti da un modello con rumore bianco normalizzato, nel polinomio $C^*(z^{-1})$ si avrebbe un termine di grado zero, c_0 , che sarebbe comunque sempre *diverso da zero* per ipotesi. Il suo quadrato è evidentemente uguale alla varianza di $\mathbf{w}(t)$.

Modelli del tipo (4.2) si chiamano in letteratura *modelli ARMAX* o anche “modelli ARMA con ingressi”. La generalizzazione al caso vettoriale dei modelli ARMAX si ottiene immediatamente partendo dalla r.p.f. (2.27). A questo proposito è utile notare che,

PROPOSIZIONE 7.5. *Se la r.p.f. (2.27) è coprime e se vale l'ipotesi 1, il polinomio matriciale $C(z)$ ha tutti gli zeri interni al cerchio unità.*

DIMOSTRAZIONE. Infatti $G(z)^{-1}F(z) = C(z^{-1})^{-1}B(z^{-1})$ può essere analitica fuori del cerchio solo se gli eventuali zeri instabili di $C(z)$ si cancellano con quelli di $B(z)$ ed esiste quindi un divisore polinomiale sinistro $Q(z^{-1})$ comune a $C(z^{-1})$ e a $B(z^{-1})$ a zeri instabili. D'altra parte $G(z)^{-1} = C(z^{-1})^{-1}A(z^{-1})$ può essere analitica fuori del cerchio solo se $Q(z^{-1})$ è un divisore sinistro anche di $A(z^{-1})$. Questo significa che $Q(z^{-1})$ sarebbe un divisore sinistro di tutti i polinomi $A(z^{-1}), B(z^{-1}), C(z^{-1})$, il che è impossibile vista l'ipotesi di coprimalità. \square

Vogliamo ora occuparci del *problema della predizione del processo* $\{\mathbf{y}(t)\}$ *in base alla storia passata* $\mathbf{z}^t := (\mathbf{y}^t, \mathbf{u}^t)$ *congiunta di* $\{\mathbf{y}(t)\}$ *e dell'ingresso* $\{\mathbf{u}(t)\}$. Cerchiamo di procedere senza fare ipotesi sulla struttura di \mathbf{u} (o del canale di retroazione).

Ricordiamo a questo proposito che il predittore di un passo, $\hat{\mathbf{y}}(t+1 | t)$, di $\mathbf{y}(t+1)$ in base a \mathbf{z}^t è univocamente individuato in base al principio di ortogonalità come l'elemento, diciamolo $\hat{\mathbf{y}}(\mathbf{y}^t, \mathbf{u}^t)$, di $\mathbf{H}_t(\mathbf{y}, \mathbf{u})$ ($= \mathbf{H}_t(\mathbf{z})$) per cui vale una decomposizione del tipo

$$\mathbf{y}(t+1) = \hat{\mathbf{y}}(\mathbf{y}^t, \mathbf{u}^t) + \mathbf{n}(t+1) \quad , \quad (4.3)$$

dove la variabile $\mathbf{n}(t+1)$ è ortogonale ad $\mathbf{H}_t(\mathbf{y}, \mathbf{u})$.

Dalla (4.1) scritta per $t = t+1$ si ricava

$$\mathbf{y}(t+1) = \frac{B^*(z^{-1})}{A^*(z^{-1})} \mathbf{u}(t+1) + \frac{C^*(z^{-1})}{A^*(z^{-1})} \mathbf{w}(t+1) \quad . \quad (4.4)$$

Come già fatto a suo tempo per il calcolo del predittore di Wiener senza ingressi, definiamo $C_1^*(z^{-1})$ (o $C_1(z)$) attraverso la

$$C^*(z^{-1}) = A^*(z^{-1}) + z^{-1} C_1^*(z^{-1}) \quad . \quad (4.5)$$

Questa relazione è analoga alla (2.11) (con un rumore bianco normalizzato avremmo avuto $c_0 A^*(z^{-1})$ al posto di $A^*(z^{-1})$).

Usando (4.5) nella (4.4) si trova così

$$\mathbf{y}(t+1) = \frac{B^*(z^{-1})}{A^*(z^{-1})} \mathbf{u}(t+1) + \frac{C_1^*(z^{-1})}{A^*(z^{-1})} \mathbf{w}(t) + \mathbf{w}(t+1) \quad . \quad (4.6)$$

Il nostro obiettivo è quello di porre questa espressione in una forma analoga alla (4.3).

Le cose qui sono semplici perché nel modello (4.1) esiste almeno un ritardo nella catena di azione diretta e quindi $b_0 = 0$ in $B^*(z^{-1})$. In questo caso la (4.6) si può riscrivere come

$$\mathbf{y}(t+1) = \frac{B_1^*(z^{-1})}{A^*(z^{-1})} \mathbf{u}(t) + \frac{C_1^*(z^{-1})}{A^*(z^{-1})} \mathbf{w}(t) + \mathbf{w}(t+1) \quad , \quad (4.7)$$

dove B_1^* è definito dalla relazione

$$B^*(z^{-1}) = z^{-1} B_1^*(z^{-1}) = z^{-1}(b_1 + b_2 z^{-1} + \dots + b_n z^{-n+1}) \quad .$$

I primi due termini a secondo membro possono poi essere scritti in modo tale da evidenziare la loro dipendenza causale da $(\mathbf{y}^t, \mathbf{u}^t)$. Per fare questo basta ricavare $\mathbf{w}(t)$ dalla (4.2)

$$\mathbf{w}(t) = \frac{1}{C^*(z^{-1})} [A^*(z^{-1}) \mathbf{y}(t) - B^*(z^{-1}) \mathbf{u}(t)]$$

e sostituire in (4.7), ottenendo

$$\mathbf{y}(t+1) = \frac{C_1^*(z^{-1})}{C^*(z^{-1})} \mathbf{y}(t) + \left[\frac{B_1^*(z^{-1})}{A^*(z^{-1})} - \frac{z^{-1} C_1^*(z^{-1}) B_1^*(z^{-1})}{A^*(z^{-1}) C^*(z^{-1})} \right] \mathbf{u}(t) + \mathbf{w}(t+1) .$$

Usando la (4.5) il termine fra parentesi quadre si riduce a $B_1^*(z^{-1})/C^*(z^{-1})$, per cui l'espressione cercata ha la forma

$$\mathbf{y}(t+1) = \frac{C_1^*(z^{-1})}{C^*(z^{-1})} \mathbf{y}(t) + \frac{B_1^*(z^{-1})}{C^*(z^{-1})} \mathbf{u}(t) + \mathbf{w}(t+1)$$

e per la stabilità stretta di $C(z)$ i primi due addendi stanno certamente in $\mathbf{H}_t(\mathbf{y}, \mathbf{u})$. Il problema si riduce a vedere se vale l'ortogonalità

$$\mathbf{w}(t+1) \perp \mathbf{H}_t(\mathbf{y}, \mathbf{u}) \quad (4.8)$$

nel qual caso si può concludere che *il predittore di un passo è dato dalla formula*

$$\hat{\mathbf{y}}(t+1 | t) = \frac{C_1^*(z^{-1})}{C^*(z^{-1})} \mathbf{y}(t) + \frac{B_1^*(z^{-1})}{C^*(z^{-1})} \mathbf{u}(t) \quad , \quad (4.9)$$

dove il polinomio $C_1^*(z^{-1})$ è definito dalla (4.5) e $B_1^*(z^{-1})$ è uguale a $zB^*(z^{-1})$.

OSSERVAZIONE 7.1. La derivazione dell'espressione (4.9) per il predittore di un passo è ovviamente corretta nell'ipotesi che il modello ARMAX disponibile sia *d'innovazione* (sia cioè la prima equazione di un modello a retroazione d'innovazione per il processo congiunto). In questo caso $\mathbf{w}(t+1) = \mathbf{e}(t+1) = \mathbf{y}(t+1) - \hat{E}[\mathbf{y}(t+1) | \mathbf{y}^t, \mathbf{u}^t]$ è ovviamente ortogonale a $\mathbf{H}_t(\mathbf{y}, \mathbf{u})$. Supporre che il modello (4.1) sia di innovazione non sposta di molto i termini della questione perchè il calcolo del predittore e quello del modello d'innovazione sono, in fondo, questioni equivalenti e il problema si riproporrebbe come quello di decidere quando (4.1) è un modello di innovazione. \diamond

PROPOSIZIONE 7.6. *Se il modello (4.1) soddisfa alle ipotesi poste e il processo \mathbf{u} è generato da una retroazione lineare causale, cioè una retroazione del tipo*

$$\mathbf{u}(t) = H(z) \mathbf{y}(t) + \mathbf{r}(t) \quad ,$$

dove $H(z)$ è una matrice di trasferimento non necessariamente razionale, regolare all'infinito, \mathbf{r} un processo stazionario p.n.d. scorrelato da \mathbf{w} e la corrispondente funzione di trasferimento in catena chiusa, $T(z)$, è analitica in $\{|z| \geq 1\}$, allora vale la condizione di ortogonalità (4.8) e il predittore è dato dalla formula (4.9).

DIMOSTRAZIONE. Sia $\mathbf{r}(t) = K(z) \mathbf{w}_2(t)$ la rappresentazione d'innovazione del processo $\mathbf{r}(t)$ mediante il rumore bianco $\mathbf{w}_2(t)$. Evidentemente mettendo assieme la legge di retroazione e il modello (4.1) si trova un modello a retroazione (internamente stabile) che è del tipo standard (2.3) studiato nelle sezioni precedenti, eccettuata la possibile non razionalità delle funzioni che compaiono nel canale di retroazione (e della quale non ci preoccuperemo troppo). In ogni caso, la stabilità di $T(z)$ e l'analiticità del fattore spettrale $W(z)$ corrispondente (garantita dalle ipotesi fatte) permette di dire che sia $\mathbf{y}(t)$ che $\mathbf{u}(t)$ sono funzionali del passato dei rumori bianchi di ingresso, i.e. $\mathbf{H}_t(\mathbf{y}, \mathbf{u}) \subset \mathbf{H}_t(\mathbf{w}, \mathbf{w}_2)$ e pertanto $\mathbf{w}(t+1) \perp \mathbf{H}_t(\mathbf{y}, \mathbf{u})$. \square

OSSERVAZIONE 7.2. Da notare che in particolare potrebbe essere $H(z) \equiv 0$ e $\mathbf{u} \equiv \mathbf{r}$. In questo caso (ovviamente non c'è reazione e) $F(z)$ dev'essere analitica in $\{|z| \geq 1\}$. Un ovvio corollario della Proposizione 7.6 è che $\mathbf{w}(t+1)$ è l'errore di predizione (innovazione) di $\mathbf{y}(t+1)$ in base ai dati congiunti $\{\mathbf{y}^t, \mathbf{u}^t\}$ e quindi il modello (4.1) è di innovazione.

Notiamo che il predittore di un passo (4.9) sembra non dipendere affatto dalla dinamica del "canale di reazione" $\mathbf{y}(t) \rightarrow \mathbf{u}(t)$. Sembra così lecito (e talvolta si trovano in letteratura affermazioni di questo tipo) usare la formula prescindendo completamente da una qualunque forma di informazione a priori su \mathbf{u} e trattare, insomma, i dati \mathbf{u}^t come se fossero *deterministici*. In realtà l'indipendenza del predittore dalla struttura probabilistica del segnale $\{\mathbf{u}(t)\}$ si verifica solo nel caso di retroazione causale. \diamond

PROBLEMA 7.6. Discutere il calcolo del predittore di un semplice modello ARMAX in cui $\mathbf{u}(t) = \mathbf{y}(t+k) + \mathbf{w}_2(t)$, $k > 0$. \diamond

Una difficoltà che si incontra nell'applicare la formula ai casi che si incontrano in pratica è che guardando *solo* alla descrizione della “catena diretta” (4.1), non si può verificare se il modello è d'innovazione perché per fare questo, come abbiamo appena visto (vedi anche il Teorema 7.1), si richiede anche una certa conoscenza dell' “altra metà” del modello che descrive $\{\mathbf{u}(t)\}$.

In effetti un caso in cui il problema si semplifica molto è quando c'è *assenza di reazione*. L' enunciato che segue tratta appunto questa situazione ed è particolarmente utile nelle applicazioni.

PROPOSIZIONE 7.7. *Se il processo $\{\mathbf{y}(t)\}$ è descritto da un modello razionale del tipo*

$$\mathbf{y}(t) = F(z) \mathbf{u}(t) + G(z) \mathbf{w}(t) \quad , \quad (4.10)$$

in cui $\{\mathbf{u}(t)\}$ ed $\{\mathbf{w}(t)\}$ sono scorrelati ed $F(z)$ è analitica in $\{|z| \geq 1\}$ e nulla all'infinito, la coppia $\{\mathbf{y}(t)\}, \{\mathbf{u}(t)\}$ è senza reazione. Se inoltre $G(z)$ è a fase minima, il modello (4.10) è necessariamente d'innovazione. In queste ipotesi il predittore di un passo $\hat{\mathbf{y}}(t+1 | t)$ si può calcolare con la formula (4.9).

DIMOSTRAZIONE. Si può parafrasare l'ultima parte ((2) \Rightarrow (1)) della prova del Teorema 7.4 e mostrare che $\hat{E}[\mathbf{y}(t) | \mathbf{H}(\mathbf{u})] = F(z) \mathbf{u}(t) = \hat{E}[\mathbf{y}(t) | \mathbf{u}^{t-1}]$. Accoppiando alla (4.10) la rappresentazione d'innovazione, $\mathbf{u}(t) = K(z) \mathbf{e}_u(t)$, di $\{\mathbf{u}(t)\}$, si ottiene allora, se $G(z)$ è presa a fase minima, il modello a retroazione d'innovazione del processo congiunto $[\{\mathbf{y}(t)\}, \{\mathbf{u}(t)\}]$ (teorema 7.1). \square

PROBLEMA 7.7. *Dimostrare direttamente che, nelle ipotesi della proposizione precedente, il modello è di innovazione (e quindi $\mathbf{w}(t) = \mathbf{e}(t)$).*

Soluzione. Dalla causalità di $F(z)$ e $G(z)$ segue infatti che $\mathbf{y}(s) \in \mathbf{H}_s(\mathbf{u}, \mathbf{w}) \forall s$ e quindi $\mathbf{H}_t(\mathbf{y}, \mathbf{u}) \subset \mathbf{H}_t(\mathbf{u}, \mathbf{w})$ cosicchè $\mathbf{w}(t+1) \perp \mathbf{H}_t(\mathbf{y}, \mathbf{u})$. Ne deriva che il predittore di un passo $\hat{E}[\mathbf{y}(t+1) | \mathbf{y}^t, \mathbf{u}^t]$ è ancora dato dall'espressione (4.9) e quindi $\mathbf{w}(t+1)$ è proprio il relativo errore di predizione. \diamond

Dalla (4.9) si vede agevolmente che la stima $\hat{\mathbf{y}}(t) := \hat{\mathbf{y}}(t+1 | t)$ soddisfa l'equazione stocastica alle differenze

$$C^*(z^{-1}) \hat{\mathbf{y}}(t) = B_1^*(z^{-1}) \mathbf{u}(t) + C_1^*(z^{-1}) \mathbf{y}(t) \quad , \quad (4.11)$$

che permette di calcolare ricorsivamente la predizione $\hat{\mathbf{y}}(t)$ come funzione dei dati osservati \mathbf{u}^t e \mathbf{y}^t , note n stime precedenti $\hat{\mathbf{y}}(t_0-1) \dots \hat{\mathbf{y}}(t_0-n)$ all'istante iniziale t_0 .

A proposito dell'inizializzazione di questo schema bisogna poi ripetere pari pari le considerazioni fatte, relative alla formula di predizione senza ingressi del Cap. VII. Anche in questo caso in pratica bisogna inizializzare il predittore con condizioni iniziali arbitrarie, ad esempio $\hat{\mathbf{y}}(i) = 0, i = 1 \dots n$. Se $C(z)$ ha tutti gli zeri in $\{|z| < 1\}$, l'effetto delle condizioni iniziali sbagliate si estingue nel tempo e a transitorio esaurito si ottiene la stima ottimale.

PROBLEMA 7.8. Derivare la formula (4.11) nel caso di processi vettoriali. \diamond

Notiamo che la presenza del ritardo nella catena di azione diretta ha permesso di semplificare notevolmente i calcoli. Se si parte da un modello del tipo (4.10) in cui non c'è il ritardo in $F(z)$, il problema di calcolare $\hat{\mathbf{y}}(t+1 | t)$ (senza ipotesi addizionali) diventa equivalente a quello del calcolo del predittore del processo congiunto, $\hat{\mathbf{z}}(t+1 | t)$ e non si può risolvere senza conoscere la dinamica del canale di reazione.

Caso di accoppiamento istantaneo tra \mathbf{u} e \mathbf{y} . Se non c'è il ritardo in $F(z)$, ma però si suppone che *non ci sia reazione*⁷ usando la scomposizione $B^*(z^{-1}) = b_0 A^*(z^{-1}) + z^{-1} B_1^*(z^{-1})$ e usando al posto della (4.7) l'espressione

$$\mathbf{y}(t+1) - b_0 \mathbf{u}(t+1) = \frac{B_1^*(z^{-1})}{A^*(z^{-1})} \mathbf{u}(t) + \frac{C_1^*(z^{-1})}{A^*(z^{-1})} \mathbf{w}(t) + \mathbf{w}(t+1) \quad .$$

si può concludere facilmente che il predittore di un passo di $\mathbf{y}(t+1) - b_0 \mathbf{u}(t+1)$ in base a $(\mathbf{y}^t, \mathbf{u}^t)$ è

$$\frac{B_1^*(z^{-1})}{A^*(z^{-1})} \mathbf{u}(t) + \frac{C_1^*(z^{-1})}{A^*(z^{-1})} \mathbf{w}(t)$$

che si può calcolare mediante la formula ricorsiva

$$C^*(z^{-1}) [\hat{\mathbf{y}}(t+1|t) - b_0 \hat{\mathbf{u}}(t+1|t)] = B_1^*(z^{-1}) \mathbf{u}(t) + C_1^*(z^{-1}) [\mathbf{y}(t) - b_0 \mathbf{u}(t)] \quad .$$

In questa formula il predittore $\hat{\mathbf{u}}(t+1|t) = \hat{E}[\mathbf{u}(t+1) | \mathbf{y}^t, \mathbf{u}^t]$ non dipende da \mathbf{y}^t per l'assenza di reazione e coincide con il predittore in base alla storia propria di \mathbf{u} , $\hat{E}[\mathbf{u}(t+1) | \mathbf{u}^t]$. Questo predittore deve naturalmente essere costruito in base alla dinamica propria del segnale di ingresso, mediante uno schema del tipo (2.12) e aggiornato "in parallelo".

L'errore di predizione dell'uscita,

$$\tilde{\mathbf{y}}(t+1|t) = \tilde{\mathbf{y}}(t+1|t) - b_0 \tilde{\mathbf{u}}(t+1|t) + b_0 \tilde{\mathbf{u}}(t+1|t) = \mathbf{w}(t+1) + b_0 \tilde{\mathbf{u}}(t+1|t)$$

ha varianza $\sigma^2 + b_0^2 \lambda_{\mathbf{u}}^2$ dove $\sigma^2 = \text{var } \mathbf{w}(t)$ e $\lambda_{\mathbf{u}}^2$ è la varianza dell'innovazione di \mathbf{u} .

Notare che il processo bianco $\{\mathbf{w}(t)\}$ non è l'innovazione di $\{\mathbf{y}(t)\}$ (naturalmente qui si intende l'innovazione rispetto al passato del processo congiunto $(\mathbf{y}^{t-1}, \mathbf{u}^{t-1})$). \diamond

Calcolo del predittore a k passi. Se oltre ad essere $b_0 = 0$ si ha anche $b_1 = \dots = b_{k-1} = 0$, ovvero

$$B^*(z^{-1}) = z^{-k} B_k^*(z^{-1}) \quad ,$$

e ci si trova nella fortunata circostanza di voler calcolare il predittore a k passi in avanti quando *c'è un ritardo di almeno k passi tra ingresso e uscita*, l'espressione (4.9) si può facilmente generalizzare e il calcolo di $\hat{\mathbf{y}}(t+k|t)$ è altrettanto semplice. Posto infatti

$$C^*(z^{-1}) = Q^*(z^{-1}) A^*(z^{-1}) + z^{-k} C_k^*(z^{-1}) \quad , \quad (4.12)$$

dove Q^* è il quoziente della divisione di C^* per A^*

$$Q^*(z^{-1}) = 1 + q_1 z^{-1} + \dots + q_{k-1} z^{-k-1} \quad (4.13)$$

($Q^*(z^{-1})$ è di grado $k-1$ in z^{-1}), si ha

$$\mathbf{y}(t+k) = Q^*(z^{-1}) \frac{B_k^*(z^{-1})}{C^*(z^{-1})} \mathbf{u}(t) + \frac{C_k^*(z^{-1})}{C^*(z^{-1})} \mathbf{y}(t) + \mathbf{e}(t+k) + \dots + q_{k-1} \mathbf{e}(t+1) \quad ,$$

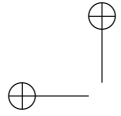
dalla quale

$$\hat{\mathbf{y}}(t+k|t) = Q^*(z^{-1}) \frac{B_k^*(z^{-1})}{C^*(z^{-1})} \mathbf{u}(t) + \frac{C_k^*(z^{-1})}{C^*(z^{-1})} \mathbf{y}(t) \quad (4.14)$$

e la varianza dell'errore di predizione è in questo caso

$$\text{var } \tilde{\mathbf{y}}(t+k|t) = \text{var} [Q^*(z^{-1}) \mathbf{e}(t+k)] = \lambda^2 (1 + q_1^2 + \dots + q_{k-1}^2) \quad ,$$

⁷Naturalmente qui l'assenza di reazione è da intendersi *in senso lato*, cf. la definizione 3.7, dato che non può evidentemente esserci causalità stretta da \mathbf{u} a \mathbf{y} .



dove $\lambda^2 = \text{var } \mathbf{e}(t)$ è la varianza dell'innovazione (non normalizzata).

5. Controllo a Minima Varianza

In questa sezione discuteremo brevemente un'applicazione estremamente importante delle formule di predizione che abbiamo appena visto. L'argomento meriterebbe ben altro spazio di quello che possiamo effettivamente dedicargli.

Supponiamo di avere un modello dinamico del tipo (4.1) in cui $F(z)$ si interpreta come la funzione di trasferimento di un impianto da controllare in cui la variabile manipolabile è $\mathbf{u}(t)$ e quella da regolare è $\mathbf{y}(t)$. In realtà si tratterà in genere di *variazioni* delle variabili attorno a valori nominali, dato che i modelli lineari che adottiamo qui descrivono di solito la dinamica linearizzata del sistema da controllare. Nello schema additivo (4.1) il termine $G(z)\mathbf{e}(t)$ è il modello probabilistico dei disturbi riportati all'uscita, che si suppongono descrivibili mediante un processo stazionario a spettro razionale e a media zero.

Supponiamo di voler progettare un compensatore che renda più piccole possibile le fluttuazioni dell'uscita (dovute ai disturbi) attorno ad un certo valore nominale che si vuole imporre alla variabile. Nella discussione che faremo più avanti questo valore nominale è assunto uguale a zero, ma potrebbe più in generale trattarsi di un valore di "set point" aggiustabile ma mantenuto costante nel tempo. Chiaramente di questi valori noti della variabile di riferimento, così come della presenza di eventuali componenti medie del disturbo, si può tenere conto in fase preliminare di modellizzazione, e si possono sempre ridefinire le variabili in gioco in modo che risultino a media zero. Questo problema è in sostanza il classico *problema della regolazione* della teoria dei Controlli Automatici. Più in generale si possono voler minimizzare le fluttuazioni dell'uscita attorno ad una certa traiettoria "nominale" imposta al sistema. In questo caso si parla di *asservimento* ("tracking" in inglese). La traiettoria da inseguire può essere nota a priori oppure aleatoria e in quest'ultimo caso la si suppone assimilabile alla traiettoria di un processo stocastico stazionario di cui si suppongono note le statistiche del second'ordine.

Dato che la variabile da regolare è soggetta a disturbi aleatori e quindi è descritta come una grandezza stocastica, l'obiettivo del progetto del sistema di regolazione si deve evidentemente porre come quello della minimizzazione in senso statistico delle fluttuazioni attorno al valore nominale. In modo naturale si arriva così all'idea di *minimizzare la varianza di $\{\mathbf{y}(t)\}$* .

Sarà opportuno far riferimento ad una classe di *leggi di controllo* (o *compensatori*) "ammissibili". La classe che ci interessa è quella delle leggi di controllo a controreazione, *causali* (o "non anticipative") che rendono *stabile*, in un senso che preciseremo tra poco, il sistema in catena chiusa.

DEFINIZIONE 7.6. *Sia dato il sistema di controllo lineare con disturbi sull'uscita*

$$\mathbf{y}(t) = F(z)\mathbf{u}(t) + \mathbf{d}(t) \quad , \quad (5.1)$$

in cui $F(z)$ è una funzione razionale di z , causale ma non necessariamente stabile, e $\{\mathbf{d}(t)\}$ un processo stazionario a media nulla e spettro razionale (rappresentabile senza perdita di generalità mediante la rappresentazione d'innovazione $\mathbf{d}(t) = G(z)\mathbf{e}(t)$, con $\{\mathbf{e}(t)\}$ un processo bianco di varianza λ^2 e $G(z)$ a fase minima). Una Legge di controllo ammissibile per il sistema (5.1) è una retroazione lineare, $\mathbf{u}(t) = H(z)\mathbf{y}(t)$, in cui $H(z)$ è una funzione di trasferimento causale (ovvero analitica in un intorno di $z = \infty$), siffatta da rendere le due funzioni di

trasferimento in catena chiusa

$$\left(1 - F(z) H(z)\right)^{-1}, \quad \left(1 - F(z) H(z)\right)^{-1} H(z) \quad (5.2)$$

strettamente stabili e causali (cioè analitiche in $\{|z| \geq 1\}$).

Notiamo subito che una legge di controllo ammissibile assicura che, per ogni t , l'ingresso $\mathbf{u}(t) = H(z) \mathbf{y}(t)$ è un funzionale lineare a varianza finita della storia passata di $\{\mathbf{y}(t)\}$ e che l'uscita, a sua volta, è un processo stazionario del secondo ordine, funzionale causale della storia passata del disturbo $\{\mathbf{d}(t)\}$.

In problemi di *asservimento* (o di regolazione con riferimento aggiustabile) la definizione si può ampliare leggermente definendo *ammissibili* leggi di controllo del tipo

$$\mathbf{u}(t) = H_1(z) \mathbf{y}(t) + H_2(z) \mathbf{r}(t) \quad , \quad (5.3)$$

dove $\mathbf{r}(t)$ è un segnale di riferimento (scorrelato da $\{\mathbf{d}(t)\}$) che si vuole inseguire. In questo caso si chiede che $H_1(z)$ sia una funzione di trasferimento causale mentre $H_2(z)$ potrebbe anche essere anticipativa, dato che in molte applicazioni alcuni campioni futuri del segnale di riferimento possono essere noti all'istante t . Si deve comunque chiedere che la legge di controllo (5.3) renda *internamente stabile* l'interconnessione a controreazione, ovvero renda la matrice di trasferimento $T(z)$ del sistema in catena chiusa

$$\begin{bmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{bmatrix} = \begin{bmatrix} \frac{FH_2}{1 - FH_1} & \frac{1}{1 - FH_1} \\ \frac{H_2}{1 - FH_1} & \frac{H_1}{1 - FH_1} \end{bmatrix} \begin{bmatrix} \mathbf{r}(t) \\ \mathbf{d}(t) \end{bmatrix}$$

analitica in $\{|z| \geq 1\}$. Questo problema è un poco più generale del problema del regolatore e verrà discusso a parte più avanti.

PROBLEMA 7.9. *Il problema del regolatore a minima varianza è quello di trovare una legge di controllo ammissibile che minimizzi la varianza del processo di uscita del sistema (5.1).*

L'idea che permette di risolvere questo problema è di separare la dinamica del processo di uscita "influenzabile" dalla variabile di controllo dalla parte "completamente aleatoria" sulla quale il controllo non ha influenza. La legge di controllo ottima (se ne esiste una) si trova minimizzando la varianza della componente "influenzabile" dal controllo.

Ricordiamo che la funzione di trasferimento disturbo-uscita, $(1 - FH)^{-1}$, corrispondente ad una qualunque legge di controllo ammissibile, è analitica in $\{|z| \geq 1\}$. Introducendo allora la rappresentazione d'innovazione (non normalizzata) $\mathbf{d}(t) = G(z) \mathbf{e}(t)$ del disturbo, si vede che la relazione $\mathbf{e}(t) \rightarrow \mathbf{y}(t)$ espressa dalla

$$\mathbf{y}(t) = \frac{G}{1 - FH} \mathbf{e}(t) \quad (5.4)$$

è *causale*, per cui $\mathbf{y}(t) \in \mathbf{H}_t(e)$ per ogni legge di controllo H ammissibile. Inoltre, per la normalizzazione di G ($G(\infty) = 1$) e la causalità stretta di F , ($F(\infty) = 0$), si può scrivere

$$G(z) = 1 + z^{-1} G_1(z) \quad , \quad F(z) = z^{-1} F_1(z) \quad , \quad (5.5)$$

con G_1 ed F_1 funzioni razionali proprie. Ne scende che $G(\infty)[1 - F(\infty)H(\infty)]^{-1} = 1$ e si può decomporre $\mathbf{y}(t+1)$ nella *somma ortogonale*

$$\mathbf{y}(t+1) = \left(\frac{G}{1 - FH} - 1 \right) \mathbf{e}(t+1) + \mathbf{e}(t+1) \quad . \quad (5.6)$$

Ricavando $\mathbf{e}(t+1)$ dalla (5.4) e sostituendo nel primo addendo in (5.6), si trova, dopo un paio di semplici passaggi, che il primo termine della somma (5.6) si può esprimere come

$$\left(1 - \frac{1}{G} + \frac{FH}{G} \right) \mathbf{y}(t+1) = \left(\frac{G_1}{G} + \frac{F_1H}{G} \right) \mathbf{y}(t) \quad , \quad (5.7)$$

dove $G_1(z)$ ed $F_1(z)$ sono le parti strettamente causali di $G(z)$ ed $F(z)$. Notiamo ora che per l'ortogonalità della decomposizione (5.6) e la stazionarietà di $\{\mathbf{y}(t)\}$ (garantita dall'uso di una legge di controllo ammissibile) si può scrivere

$$\text{var}(\mathbf{y}(t)) = \text{var}(\mathbf{y}(t+1)) = \text{var} \left[\left(\frac{G_1}{G} + \frac{F_1}{G} H \right) \mathbf{y}(t) \right] + \text{var}(\mathbf{e}(t+1)) \quad . \quad (5.8)$$

Evidentemente, $\text{var}(\mathbf{e}(t+1)) = \lambda^2$ *non dipende da* $H(z)$, per cui il termine "influenzabile" dalla legge di controllo è quello fra parentesi quadre. Minimizzare rispetto ad H la varianza di $\mathbf{y}(t)$ è pertanto equivalente a minimizzare quella di $(G_1 + F_1H) \frac{1}{G} \mathbf{y}(t)$.

Proviamo ad effettuare la minimizzazione dimenticando per un attimo che si tratta di una minimizzazione *vincolata* (H deve essere *ammissibile*). Sembra di poter rendere la varianza del termine in questione addirittura *uguale a zero*, prendendo $H(z)$ uguale a

$$H(z) = - \frac{G_1(z)}{F_1(z)} \quad . \quad (5.9)$$

Notiamo che questa legge di controllo *cancella tutta la parte prevedibile* (in base alle uscite passate) *del disturbo*. La legge di controllo (5.9) è però ammissibile solo in certe circostanze, che sono discusse in dettaglio nell'enunciato del teorema seguente.

TEOREMA 7.5. *Si assuma lo spettro del disturbo $S_d(z)$ privo di zeri sulla circonferenza unita. Se gli zeri di $F(z)$ sono tutti all'interno del cerchio unitario, esiste una legge di controllo ammissibile che minimizza la varianza dell'uscita del sistema (5.1). Sotto questa condizione, ammesso che vi siano esattamente $k \geq 1$ ritardi nella funzione di trasferimento del processo, ovvero, ammesso che*

$$F(z) = z^{-k} F_k(z) \quad , \quad F_k(\infty) \neq 0 \quad , \quad (5.10)$$

la legge di controllo (ammissibile) a minima varianza è data dalla formula

$$H(z) = - \frac{1}{Q_k^*(z^{-1})} \frac{G_k(z)}{F_k(z)} \quad , \quad (5.11)$$

dove $G_k(z)$ e $Q_k^*(z^{-1}) = 1 + q_1 z^{-1} + \dots + q_{k-1} z^{-k+1}$ sono legati dalla

$$G(z) = Q_k^*(z^{-1}) + z^{-k} G_k(z) \quad (5.12)$$

(che è analoga alle (2.11) e (4.12)). Con la legge di controllo (5.11) la funzione di trasferimento disturbo-uscita in catena chiusa è

$$\frac{1}{1 - F(z)H(z)} = Q_k^*(z^{-1}) \frac{1}{G(z)} \quad , \quad (5.13)$$

per cui il processo di uscita a varianza minima è una pura media mobile di ordine $k - 1$

$$\mathbf{y}(t) = \mathbf{e}(t) + q_1 \mathbf{e}(t-1) + \dots + q_{k-1} \mathbf{e}(t-k+1) \quad , \quad (5.14)$$

di varianza $\lambda^2(1 + q_1^2 + \dots + q_{k-1}^2)$.

DIMOSTRAZIONE. Seguendo la traccia del ragionamento fatto in precedenza per $k = 1$, notiamo che per ogni controllo ammissibile la relazione $\mathbf{y}(t) = [1 - F(z)H(z)]^{-1} G(z)\mathbf{e}(t)$ è una relazione *causale* e si ha pertanto $\mathbf{y}(t) \in \mathbf{H}_t(\mathbf{e})$, $\forall t$. Dall'identità

$$\mathbf{y}(t+k) = \left[\left(\frac{G}{1-FH} - Q_k^* \right) \mathbf{e}(t+k) \right] + Q_k^* \mathbf{e}(t+k) \quad ,$$

riscrivendo il termine fra parentesi quadre nella forma

$$\frac{G - Q_k^* + Q_k^* FH}{1-FH} \mathbf{e}(t+k) = \frac{G_k + Q_k^* F_k H}{1-FH} \mathbf{e}(t)$$

(cfr. le (5.10) e (5.12)) ed esprimendo $\mathbf{e}(t)$ tramite la (5.4), si ottiene la decomposizione *ortogonale*

$$\mathbf{y}(t+k) = \left[(G_k + Q_k^* F_k H) \frac{1}{G} \mathbf{y}(t) \right] + Q_k^* \mathbf{e}(t+k) \quad . \quad (5.15)$$

Dato che $\text{var } Q_k^*(z^{-1})\mathbf{e}(t+k)$ non dipende da $H(z)$, la varianza di $\mathbf{y}(t)$ si minimizza solo se la varianza del termine fra parentesi quadre è minima. Con la legge di controllo (5.11) questo minimo può essere reso addirittura uguale a zero. Si tratta ora di verificare che la legge (5.11) è effettivamente ammissibile. Dato che $Q_k^*(z^{-1})$ è monico ($G(\infty) = 1$), si vede subito che $H(\infty) = -G_k(\infty)/F_k(\infty) < \infty$, per cui H è *causale*.

Nelle ipotesi fatte la funzione di trasferimento $\mathbf{d}(t) \rightarrow \mathbf{y}(t)$, riportata nella formula (5.13), è analitica in $\{|z| \geq 1\}$. Inoltre la funzione di trasferimento $\mathbf{d}(t) \rightarrow \mathbf{u}(t)$,

$$W_{\mathbf{u}}(z) = -\frac{1}{F_k(z)} \frac{G_k(z)}{G(z)} \quad , \quad (5.16)$$

è analitica in $\{|z| \geq 1\}$ perché $F_k(z)$ non ha zeri instabili e $G(z)$ non si annulla in $\{|z| \geq 1\}$. Ne scende che (5.11) è effettivamente ammissibile. \square

OSSERVAZIONE 7.3. La condizione sugli *zeri di $F(z)$* (che in letteratura si trova anche espressa dicendo che $F(z)$ deve essere “a fase minima”) è essenziale per la stabilità interna dell'interconnessione a controreazione, che risulta applicando la legge di controllo (5.11). Infatti eventuali zeri instabili di F (o di F_k), a meno di improbabili cancellazioni con zeri di G_k , portano ad instabilità e a divergenza della variabile $\{\mathbf{u}(t)\}$. Da notare inoltre che la definizione di legge di controllo ammissibile che abbiamo dato si riferisce sostanzialmente ad un concetto di stabilità a retroazione di un sistema in cui l'*unico ingresso agente* è il disturbo $\{\mathbf{d}(t)\}$. In realtà, in sistemi di asservimento o di regolazione con riferimento variabile è importante considerare anche il comportamento (in catena chiusa) $\mathbf{r}(t) \rightarrow \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{bmatrix}$ ed è perciò più naturale riferirsi ad un concetto di ammissibilità, definito con riferimento alla stabilità interna del sistema con due ingressi (\mathbf{r}, \mathbf{d}) e due uscite, (\mathbf{y}, \mathbf{u}). Calcolando la matrice di trasferimento in catena chiusa, si vede che si richiede allora

anche la stabilità (analiticità in $\{|z| \geq 1\}$) della funzione di trasferimento $r \rightarrow y$

$$\frac{F(z)}{1 - F(z)H(z)} = F(z) \frac{Q_k^*(z^{-1})}{G(z)} \quad (5.17)$$

e questo significherebbe in sostanza richiedere anche la *stabilità di $F(z)$* il che è poco allettante per le applicazioni dove di norma $F(z)$ deve avere poli in $z = 1$ per garantire la precisione a regime. In questo caso è necessario riimpostare il problema usando controllori a due gradi di libertà del tipo (5.3). \diamond

ESEMPIO 7.4 (Calcolo del regolatore a minima varianza). Supponiamo di voler trovare il regolatore a minima varianza per un sistema in cui

$$\begin{aligned} A^*(z^{-1}) &= 1 - 1.7z^{-1} + 0.7z^{-2} \\ B^*(z^{-1}) &= z^{-1}(1 + 0.5z^{-1}) \\ C^*(z^{-1}) &= 1 + 0.5z^{-1} - 0.6z^{-2} - 0.9z^{-3} \end{aligned} .$$

Notiamo che $A(z)$ e $C(z)$ hanno in comune uno zero in $z = 1$. Si verifica facilmente che gli altri zeri sono dentro la circonferenza unità per cui $G(z) = C(z)/A(z)$ è a fase minima.

Per calcolare il controllore notiamo innanzitutto che

$$(1 + 0.5z^{-1} - 0.6z^{-2} - 0.9z^{-3}) = (1 - 1.7z^{-1} + 0.7z^{-2}) + z^{-1}(2.2 - 1.3z^{-1} - 0.9z^{-2}) ,$$

per cui $C_1^*(z^{-1}) = 2.2 - 1.3z^{-1} - 0.9z^{-2}$. La legge di controllo a varianza minima è allora

$$\mathbf{u}(t) = - \frac{2.2 - 1.3z^{-1} - 0.9z^{-2}}{1 + 0.5z^{-1}} \mathbf{y}(t) ,$$

che può essere scritta in forma ricorsiva come

$$\mathbf{u}(t) = -0.5\mathbf{u}(t-1) - 2.2\mathbf{y}(t) + 1.3\mathbf{y}(t-1) + 0.9\mathbf{y}(t-2) .$$

Con questo compensatore si ha $\mathbf{y}(t) = \mathbf{e}(t)$ e quindi

$$\text{var } \mathbf{y}(t) = \lambda^2 = \text{var } \mathbf{e}(t) .$$

Se invece di $C^*(z^{-1})$ si usa $\lambda C^*(z^{-1})$ (ovvero la varianza del rumore è normalizzata ad 1), dalla

$$\lambda C^*(z^{-1}) = \lambda A^*(z^{-1}) + z^{-1} \lambda C_1^*(z^{-1})$$

si vede che il termine $C_1^*(z^{-1})$ va ora moltiplicato per λ .

Tenendo conto del fatto che ora $c_0 = \lambda$, la formula (5.11) dà ancora la *stessa* legge di controllo. L'uscita in catena chiusa ovviamente si scrive $\mathbf{y}(t) = \lambda \mathbf{e}(t)$.

Notiamo infine che se avessimo introdotto esplicitamente la variabile di riferimento nella legge di controllo, nella forma $\mathbf{u}(t) = -C_1^*(z)/B_1^*(z) \mathbf{y}(t) + \mathbf{r}(t)$, l'uscita in catena chiusa sarebbe risultata

$$\mathbf{y}(t) = \frac{z^{-1}B_1^*(z^{-1})}{C^*(z^{-1})} \mathbf{r}(t) + \mathbf{e}(t)$$

e il polo instabile in $z = 1$ dovuto a $C(z)$ sarebbe apparso esplicitamente nella funzione di trasferimento $\mathbf{r} \rightarrow \mathbf{y}$. \diamond

Il fatto che la legge di controllo ottima *non dipende dalla varianza dell'innovazione* è assolutamente generale. Questo comportamento è conseguenza del fatto che il predittore è esso stesso invariante rispetto alla trasformazione

$$\mathbf{e}(t) \rightarrow \lambda \mathbf{e}(t) ,$$

cioè è lo stesso indipendentemente dalla varianza del rumore d'innovazione.

ESEMPIO 7.5 (Un caso degenere). Si vuol studiare il problema del regolatore a minima varianza per un sistema descritto dal modello

$$(1 - 5/6 z^{-1} + 1/6 z^{-2}) \mathbf{y}(t) = z^{-2}(1 + 1/2 z^{-1}) \mathbf{u}(t) + (1 - 7/3 z^{-1} + 2/3 z^{-2}) \mathbf{w}(t)$$

dove \mathbf{w} è rumore bianco di varianza σ^2 . Questo è in un certo senso un caso degenere perchè dalle fattorizzazioni

$$\begin{aligned} z^2 - 7/3 z + 2/3 &= (z - 2)(z - 1/3) \\ z^2 - 5/6 z + 1/6 &= (z - 1/2)(z - 1/3) \end{aligned}$$

segue che il modello si può riscrivere

$$\mathbf{y}(t) = z^{-2} \frac{1 + 1/2 z^{-1}}{(1 - 1/2 z^{-1})(1 - 1/3 z^{-1})} \mathbf{u}(t) + \frac{1 - 2 z^{-1}}{(1 - 1/2 z^{-1})} \mathbf{w}(t)$$

ed è evidente che la rappresentazione del disturbo non è di innovazione. Si vede subito che il disturbo additivo sull'uscita, $\mathbf{d}(t)$, è in effetti rumore bianco di varianza $4\sigma^2$ (e la $G(z)$ è evidentemente uguale a 1 in questo caso). Ne segue che nella decomposizione (5.12),

$$G(z) = Q_2^*(z^{-1}) + z^{-2} G_2(z)$$

si ha $Q_2^*(z^{-1}) = 1$ e $G_2(z) = 0$. La legge di controllo candidata per l'ottimalità è quindi

$$H(z) = -\frac{1}{Q_2^*(z^{-1})} \frac{G_2(z)}{F_2(z)} = 0$$

il che significa $\mathbf{u}(t) \equiv 0$, in sostanza una legge di controllo *in catena aperta!* Con questo regolatore si ha $\mathbf{y}(t) = \mathbf{d}(t)$ e anche se formalmente l'uscita è ancora rumore bianco avente la minima varianza possibile, la soluzione in questo caso è priva di valore pratico. Questo fenomeno è dovuto al fatto che il regolatore a minima varianza che abbiamo studiato in questo capitolo annulla la parte prevedibile dell'uscita dovuta al disturbo, il che equivale in sostanza ad annullare l'uscita del predittore $\hat{\mathbf{y}}(t+k|t)$. In questo esempio il predittore è funzione solo dell'ingresso passato,

$$\hat{\mathbf{y}}(t+k|t) = \frac{B_k^*(z^{-1})}{A^*(z^{-1})} \mathbf{u}(t)$$

e se si insiste nel pensare che u sia un processo stocastico generato da una legge di controllo in retroazione, la minimizzazione della varianza porta alla soluzione $\mathbf{u}(t) = 0$. \diamond

Asservimento a minima varianza. Considereremo qui solo il caso in cui ad ogni istante t sono noti i valori

$$\mathbf{r}(t+1), \dots, \mathbf{r}(t+k)$$

della variabile di riferimento \mathbf{r} e k è *esattamente* il numero di ritardi presente nella funzione di trasferimento $F(z)$ del sistema. In pratica i valori del segnale di riferimento da inseguire possono essere noti (ad esempio assegnata a priori) anche per un orizzonte temporale più lungo dei k istanti futuri successivi, però la legge di controllo che calcoleremo non userà questa eventuale informazione in più.

Per trovare la legge di controllo a due gradi di libertà (5.3), si procede in due passi. Prima si determina $H_2(z)$ imponendo che $E \mathbf{y}(t) = \mathbf{r}(t)$ per ogni t . Dato

che i valori di \mathbf{r} sono noti, si può trattare il segnale di riferimento come un segnale deterministico e quindi

$$E \mathbf{y}(t) = \frac{F(z) H_2(z)}{1 - F(z) H_1(z)} \mathbf{r}(t) \quad ,$$

dalla quale, imponendo $E \mathbf{y}(t) = \mathbf{r}(t)$, si ricavare la condizione $F(H_2 + H_1) = 1$ ovvero,

$$H_2(z) = -H_1(z) + \frac{1}{F(z)} \quad (5.18)$$

Il secondo passo consiste nel progettare un regolatore a minima varianza che regoli le fluttuazioni della variabile $\tilde{\mathbf{y}}(t) := \mathbf{y}(t) - \mathbf{r}(t)$, la quale, come si vede immediatamente, è descritta dalla relazione

$$\tilde{\mathbf{y}}(t) = \frac{1}{1 - F(z) H_1(z)} \mathbf{d}(t).$$

che corrisponde formalmente ad una legge di controllo del tipo $\mathbf{u}(t) = H_1(z) \tilde{\mathbf{y}}(t)$ applicata al sistema $\tilde{\mathbf{y}}(t) = F(z) \mathbf{u}(t) + \mathbf{d}(t)$.

Questo problema di regolazione a minima varianza è dello stesso tipo discusso e risolto all'inizio di questo capitolo. Usando la struttura (5.10) di F si ottiene ancora la decomposizione ortogonale

$$\tilde{\mathbf{y}}(t+k) = \left[(G_k + Q_k^* F_k H_1) \frac{1}{G} \tilde{\mathbf{y}}(t) \right] + Q_k^* \mathbf{e}(t+k) \quad ,$$

dalla quale si ricava la soluzione ottima che ha la nota espressione

$$\tilde{H}_1(z) = -\frac{1}{Q_k^*(z^{-1})} \frac{G_k(z)}{F_k(z)} \quad .$$

C'è da notare che questo regolatore si applica agli scostamenti $\mathbf{y}(t) - \mathbf{r}(t)$ e conviene perciò riscrivere la (5.3) nella forma

$$\mathbf{u}(t) = \tilde{H}_1(z) \tilde{\mathbf{y}}(t) + \tilde{H}_2(z) \mathbf{r}(t) \quad , \quad (5.19)$$

la quale corrisponde al classico schema che combina un compensatore in "feedforward" ($\tilde{H}_2(z)$) e una retroazione basata sull'errore di asservimento ($\tilde{H}_1(z) \tilde{\mathbf{y}}(t)$) di figura

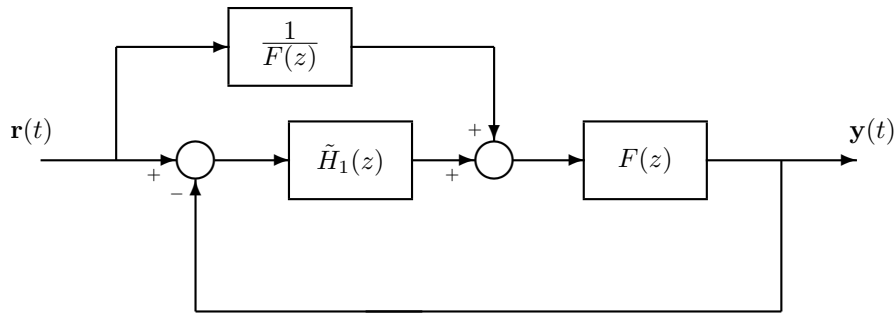
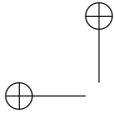


FIGURA 5.1. Schema a catena chiusa con compensatore *feedforward*.



Confrontando con (5.3) si trova

$$\tilde{H}_1(z) = H_1(z) \quad \tilde{H}_2(z) = H_2(z) + H_1(z) = \frac{1}{F(z)}$$

che mostra l'azione elementare di feedforward esplicitata dal controllore $\tilde{H}_2(z)$.

Sostituendo poi in (5.18), si trova una formula esplicita per $H_2(z)$

$$\begin{aligned} H_2(z) &= \frac{G_k(z) + z^k Q_k^*(z^{-1})}{Q_k^*(z^{-1})F_k(z)} \\ &= \frac{z^k G(z)}{Q_k^*(z^{-1})F_k(z)} \end{aligned}$$

Come era da prevedere questa funzione di trasferimento anticipa di k passi. Usando le espressioni appena trovate si può mettere il controllore (5.3) in forma ricorsiva. Partendo dalla

$$Q_k^*(z^{-1})F_k(z) \mathbf{u}(t) = -G_k(z) \mathbf{y}(t) + G(z) \mathbf{r}(t+k)$$

ed esprimendo F_k e G come rapporto di polinomi in z^{-1} , si arriva alla

$$Q_k^*(z^{-1})B_k^*(z^{-1}) \mathbf{u}(t) = -C_k^*(z^{-1}) \mathbf{y}(t) + C^*(z^{-1}) \mathbf{r}(t+k) \quad .$$

La verifica che questa legge di controllo rende il sistema in catena chiusa internamente stabile *se gli zeri di $F_k(z)$ sono stabili*, è lasciata al lettore.

Come commento conclusivo menzioniamo il fatto che il regolatore (5.11) può anche dare luogo ad ampie escursioni della variabile di controllo e questo può essere inammissibile in certi tipi di applicazioni. Per ridurre queste escursioni occorrerebbe considerare indici di comportamento in cui entrasse una somma pesata delle varianze di uscita e dell'ingresso. Questo tipo di considerazioni ci porterebbe però troppo lontano. Rinviamo il lettore alla letteratura ed in particolare a [7, 22], per una trattazione del problema più esauriente di quanto possa essere fatto in questa sede.

Esercizi

7-1 Calcolare il fattore spettrale $W(z)$, della densità congiunta dei processi descritti dal modello a retroazione

$$\begin{aligned} \mathbf{y}(t) &= \frac{1}{z-2} \mathbf{u}(t) + \mathbf{w}_1(t) \\ \mathbf{u}(t) &= -1.5 \mathbf{y}(t) + \mathbf{w}_2(t) \end{aligned}$$

dove \mathbf{w}_1 , \mathbf{w}_2 sono rumori bianchi di varianze q_1 , q_2 positive. Verificare che $W(z)$ è analitico ma non a fase minima.

Determinare un opportuno fattore passa-tutto $V(z) = (1-\alpha z)/(z-\alpha)$ in modo tale che

$$\hat{W}(z) := W(z) \begin{bmatrix} V(z) & 0 \\ 0 & 1 \end{bmatrix}$$

sia un fattore spettrale privo di zeri e poli in $\{|z| \geq 1\}$. Verificare infine che il modello a retroazione (d'innovazione!) corrispondente è

$$\begin{aligned} \mathbf{y}(t) &= \frac{1}{z-2} \mathbf{u}(t) + \frac{1-2z}{z-2} \mathbf{e}_1(t) \\ \mathbf{u}(t) &= -1.5 \mathbf{y}(t) + \mathbf{e}_2(t) \end{aligned}$$

e discutere il risultato ottenuto alla luce del Teorema 7.2.

7-2 Trovare l'analogo del Teorema 7.4 per caratterizzare la condizione di assenza di reazione in senso debole descritta dalla (3.7).

7-3 Lo schema a blocchi seguente descrive un cosiddetto modello *a Errori nelle Variabili* (EIV)

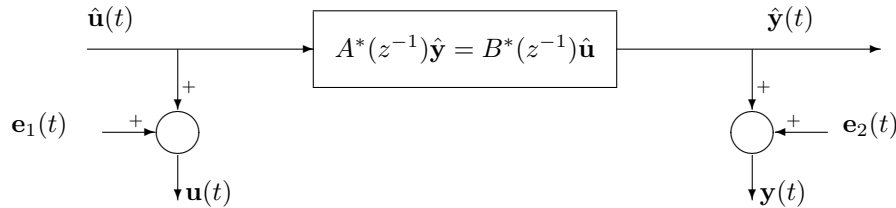


FIGURA 5.2. Schema a blocchi del modello EIV.

Le variabili cosiddette “vere”, $\hat{\mathbf{u}}(t)$ e $\hat{\mathbf{y}}(t)$ sono legate tra di loro da una equazione alle differenze del tipo

$$A^*(z^{-1})\hat{\mathbf{y}}(t) = B^*(z^{-1})\hat{\mathbf{u}}(t)$$

e sono osservate in presenza di rumori additivi $\mathbf{e}_1(t)$ e $\mathbf{e}_2(t)$ che si assumono bianchi, scorrelati tra di loro ($\mathbf{e}_1 \perp \mathbf{e}_2$) e con le variabili vere,

$$\begin{cases} \mathbf{u}(t) = \hat{\mathbf{u}}(t) + \mathbf{e}_1(t) & \mathbf{e}_1 \perp \hat{\mathbf{u}} \\ \mathbf{y}(t) = \hat{\mathbf{y}}(t) + \mathbf{e}_2(t) & \mathbf{e}_2 \perp \hat{\mathbf{y}} \end{cases}$$

Mostrare che $\mathbf{y}(t)$ e $\mathbf{u}(t)$ sono legate da un modello ARMAX. Calcolare i relativi polinomi e il rumore bianco in ingresso al modello.

Discutere se e sotto quali condizioni ci può essere assenza di reazione da \mathbf{y} ad \mathbf{u} .

7-4 Discutere il progetto di un regolatore a minima varianza per un sistema discreto descritto dal modello

$$\mathbf{y}(t) = \frac{z^{-2}(1 + 0.5z^{-1})}{(1 + 2z^{-1})(1 + z^{-1})}\mathbf{u}(t) + \frac{1 + 0.3z^{-1}}{1 + 0.8z^{-1}}\mathbf{w}(t)$$

dove \mathbf{w} è bianco di varianza σ^2 . Notare che l'impianto da regolare è instabile.

7-5 Un sistema di controllo a dati campionati soggetto a disturbi aleatori è descritto in forma simbolica dal seguente modello statistico lineare,

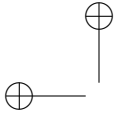
$$\mathbf{y}(t) = \frac{z + 0.2}{z(z - 1)(z + 0.8)}\mathbf{u}(t) + \frac{z + 0.5}{z + 0.8}\mathbf{w}(t)$$

dove $\mathbf{u}(t)$ è l'ingresso manipolabile e $\mathbf{w}(t)$ è un processo di rumore bianco stazionario di varianza nota σ^2 .

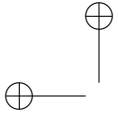
Si vuole progettare un controllore lineare che segua, con varianza d'errore minima, una traiettoria di riferimento $\mathbf{r}(t)$ della quale ad ogni istante t sono noti i valori futuri $\mathbf{r}(t + 1)$, $\mathbf{r}(t + 2)$ nei due successivi istanti di campionamento.

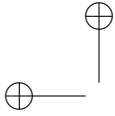
Calcolare la funzione di trasferimento in catena chiusa

$$\begin{bmatrix} \mathbf{r}(t) \\ \mathbf{d}(t) \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{bmatrix}$$



dove $\mathbf{d}(t)$ è il disturbo aleatorio additivo nel modello e verificare la stabilità interna del sistema in catena chiusa.





CAPITOLO 8

ALGORITMO DI LEVINSON

1. Introduzione

In questo capitolo considereremo problemi di stima su processi stazionari con *dati finiti* in cui le misure cominciano a fluire a partire da un certo istante iniziale finito, t_0 .

Il problema di cui ci occuperemo sarà ancora quello del calcolo dello stimatore lineare a minima varianza $\hat{E}[\mathbf{x}(t) | \mathbf{y}(t_0), \dots, \mathbf{y}(t)]$ di un processo messaggio $\{\mathbf{x}(t)\}$ in base alla storia passata finite del processo $\{\mathbf{y}(t)\}$, che denoteremo

$$\mathbf{y}_{[t_0, t]} := \begin{bmatrix} \mathbf{y}(t_0) \\ \vdots \\ \mathbf{y}(t) \end{bmatrix}, \quad t \geq t_0 \quad . \quad (1.1)$$

Il risultato fondamentale del capitolo sarà uno schema di calcolo “ricorsivo” (l’algoritmo di Levinson) che permette di costruire la stima basata sui dati $\mathbf{y}_{[t_0, t+1]}$ mediante un “aggiornamento” di quella precedente, basata su $\mathbf{y}_{[t_0, t]}$.

Faremo sempre l’ipotesi che $\{\mathbf{y}(t)\}$ sia un processo del secondo ordine, a media nulla, *puramente non deterministico e di rango pieno*. Notiamo che, nonostante la stazionarietà del processo di misura, il fatto che si usino stringhe finite di dati porta inevitabilmente a considerare stimatori che sono funzioni “non stazionarie” delle misure, descritti cioè da sistemi varianti nel tempo. Questa inevitabile complicazione potrebbe indurre a dubitare dell’utilità stessa di imporre al processo $\{\mathbf{y}(t)\}$ l’ipotesi di stazionarietà, che in pratica è verificabile solo in modo molto grossolano (tanto varrebbe partire in condizioni di assoluta generalità e considerare processi non-stazionari).

Mostreremo in questo capitolo che la stazionarietà è viceversa una condizione strutturale preziosa che può essere sfruttata in modo da ottenere proprio la *struttura ricorsiva* dell’algoritmo di Levinson. Come vedremo questa struttura ha proprietà numeriche estremamente interessanti.

2. Innovazioni a memoria finita

Come al solito, le operazioni sui dati di misura necessarie per costruire gli stimatori risultano grandemente facilitate se si introduce il processo di innovazione. In questo contesto definiremo *il processo di innovazione in avanti a memoria (finita)* n , $\{\mathbf{e}_n(t)\}$, di $\{\mathbf{y}(t)\}$, come il processo degli errori di predizione di un passo

$$\mathbf{e}_n(t) := \mathbf{y}(t) - \hat{E}[\mathbf{y}(t) | \mathbf{y}_{[t-n, t]}] \quad , \quad t \in \mathbb{Z} \quad , \quad (2.1)$$

relativi al predittore a memoria finita di $\mathbf{y}(t)$ in base alla storia passata $\mathbf{y}_{[t-n,t]} \equiv \mathbf{y}_{[t-n,t-1]}$. (La notazione $[t-n,t] \equiv [t-n,t-1]$ è introdotta per ragioni tipografiche). Denoteremo in seguito con il simbolo $\mathbf{H}(\mathbf{y}_{[t-n,t]})$ il sottospazio di $\mathbf{H}(\mathbf{y})$ generato da $\{\mathbf{y}(t-n), \dots, \mathbf{y}(t)\}$.

Chiaramente, per n fissato, $\{\mathbf{e}_n(t)\}$ è un processo stazionario giacché, in base alla nota formula (1.5) del Capitolo 6,

$$\mathbf{e}_n(t+s) = \mathbf{y}(t+s) - \hat{E}[U_s \mathbf{y}(t) | U_s \mathbf{H}(\mathbf{y}_{[t-n,t]})] = U_s \mathbf{e}_n(t) \quad . \quad (2.2)$$

Il tema centrale di questo paragrafo sarà la derivazione di formule ricorsive per l'aggiornamento dell'innovazione $\mathbf{e}_n(t)$, corrispondente all'aggiunta di un nuovo dato nella memoria del predittore. Per far questo fisseremo l'istante t e studieremo le modifiche di $\mathbf{e}_n(t)$ conseguenti all'ampliamento della base dei dati da $\mathbf{y}_{[t-n,t]}$ a $\mathbf{y}_{[t-n-1,t]}$. Vogliamo insomma riuscire a calcolare $\mathbf{e}_{n+1}(t)$ come funzione della (vecchia) innovazione $\mathbf{e}_n(t)$ e dell'ulteriore dato di misura $\mathbf{y}(t-n-1)$ che si aggiunge al vettore di n dati $\mathbf{y}_{[t-n,t]}$ che sono in memoria all'istante t . Le formule che otterremo saranno successivamente estese al calcolo di predittori e di stimatori più generali.

L'approccio seguito normalmente nei testi per la derivazione dell'algoritmo di Levinson è sostanzialmente algebrico e consiste nel manipolare in modo accorto un sistema di equazioni lineari (si veda la (2.21) più sotto) e un certo altro sistema "duale", in modo da esprimere la soluzione "a memoria $n+1$ " esplicitamente come funzione della soluzione "a memoria n ". Noi cercheremo di dare una veste un tantino più "probabilistica" a queste manipolazioni formali, nella speranza di facilitare la comprensione delle ragioni che stanno alla base del loro successo.

Come primo passo studiamo *l'ampliamento della memoria dell'innovazione a memoria finita* $\mathbf{e}_n(t)$.

È evidente che sarà conveniente aggiungere ai vecchi dati $\mathbf{y}_{[t-n,t]}$ solo quella parte di $\mathbf{y}(t-n-1)$ che contiene "nuova informazione". In altri termini si potrà benissimo togliere dal nuovo dato $\mathbf{y}(t-n-1)$ il suo "predittore" lineare (a minima varianza d'errore) basato sui dati $\mathbf{y}_{[t-n,t]}$ senza perdere alcunché di rilevante ai fini del calcolo di $\mathbf{e}_{n+1}(t)$, almeno nel contesto lineare in cui stiamo operando.

Questa semplice osservazione conduce all'idea di *innovazione all'indietro*.

DEFINIZIONE 8.1. *Si chiama "innovazione all'indietro a memoria n di $\{\mathbf{y}(t)\}$ " il processo (stazionario) $\{\bar{\mathbf{e}}_n(t)\}$ definito dalla formula*

$$\bar{\mathbf{e}}_n(t-n) = \mathbf{y}(t-n) - \hat{E}[\mathbf{y}(t-n) | \mathbf{y}_{(t-n,t)}] \quad , \quad (2.3)$$

dove $\mathbf{y}_{(t-n,t)} := \mathbf{y}_{[t-n+1,t]}$ è il vettore delle n osservazioni $\{\mathbf{y}(t-n+1), \dots, \mathbf{y}(t)\}$, che seguono $\mathbf{y}(t-n)$.

Come si vede, $\bar{\mathbf{e}}_n(t-n)$ è semplicemente l'errore di "predizione all'indietro" di $\mathbf{y}(t-n)$ in base ai dati $\mathbf{y}(t-n+1), \dots, \mathbf{y}(t)$.

Nonostante sia naturale indicare l'argomento temporale in $\bar{\mathbf{e}}_n$ con $t-n$ anziché t , la simbologia adottata è un pò pesante e cercheremo di semplificarla un poco introducendo la notazione:

$$\tilde{\mathbf{e}}_n(t) \equiv \bar{\mathbf{e}}_n(t-n)$$

che si rivelerà comoda nel seguito, quando si tratterà di manipolare formule .

Ovviamente $\{\tilde{\mathbf{e}}_n(t)\}$ è (per ogni n fissato) ancora un processo stazionario. La proprietà che ci interessa di più è però la seguente.

LEMMA 8.1. Per ogni $t \in \mathbb{Z}$ i vettori $\bar{\mathbf{e}}_0(t), \dots, \bar{\mathbf{e}}_k(t-k)$ formano una base ortogonale dello spazio $\mathbf{H}(\mathbf{y}_{[t-k,t]})$ generato da $\mathbf{y}(t-k), \dots, \mathbf{y}(t)$, ovvero

$$\text{span}\{\bar{\mathbf{e}}_0(t), \dots, \bar{\mathbf{e}}_k(t-k)\} = \text{span}\{\mathbf{y}(t-k), \dots, \mathbf{y}(t)\} = \mathbf{H}(\mathbf{y}_{[t-k,t]}) \quad (2.4)$$

$$\bar{\mathbf{e}}_{k+1}(t-k-1) \perp \mathbf{H}(\mathbf{y}_{[t-k,t]}) \quad (2.5)$$

per $k = 0, 1, \dots$

Dualmente, i vettori aleatori $\mathbf{e}_0(t), \mathbf{e}_1(t+1), \dots, \mathbf{e}_k(t+k), \dots$ formano una base ortogonale di $\mathbf{H}(\mathbf{y}_{[t,t+k]})$, ovvero

$$\text{span}\{\mathbf{e}_0(t), \mathbf{e}_1(t+1), \dots, \mathbf{e}_k(t+k)\} = \mathbf{H}(\mathbf{y}_{[t,t+k]}) \quad (2.6)$$

$$\mathbf{e}_{k+1}(t+k+1) \perp \mathbf{H}(\mathbf{y}_{[t,t+k]}) \quad (2.7)$$

per ogni t e $k \geq 0$.

DIMOSTRAZIONE. Notiamo in effetti che i vettori aleatori $\bar{\mathbf{e}}_0(t), \bar{\mathbf{e}}_1(t-1), \dots$ sono costruiti tramite il classico procedimento di Gram-Schmidt (senza normalizzazione) partendo da $\mathbf{y}(t)$ e procedendo all'indietro, cioè ortogonalizzando successivamente $\mathbf{y}(t), \mathbf{y}(t-1), \dots$. Si ha, ad esempio,

$$\begin{aligned} \bar{\mathbf{e}}_0(t) &= \mathbf{y}(t) \\ \bar{\mathbf{e}}_1(t-1) &= \mathbf{y}(t-1) - \hat{E}[\mathbf{y}(t-1) | \bar{\mathbf{e}}_0(t)] \\ \bar{\mathbf{e}}_2(t-2) &= \mathbf{y}(t-2) - \hat{E}[\mathbf{y}(t-2) | \bar{\mathbf{e}}_1(t-1), \bar{\mathbf{e}}_0(t)] \end{aligned}$$

ecc... Dualmente, $\mathbf{e}_k(t)$ è costruito attraverso un procedimento di ortogonalizzazione in avanti delle variabili $\mathbf{y}(t-k), \mathbf{y}(t-k+1), \dots, \mathbf{y}(t)$ e da questo seguono immediatamente le proprietà (2.6) e (2.7). \square

Ricordiamo che la definizione di $\tilde{\mathbf{e}}_n(t)$ fa riferimento a un problema di predizione (all'indietro) in cui i dati sono relativi all'intervallo temporale $(t-n, t]$. Il problema originario (in avanti) è, viceversa, relativo a dati disponibili in $[t-n, t)$. Mantenendo t costante e allargando la memoria di un dato (ovvero introducendo l'ulteriore osservazione $\mathbf{y}(t-n-1)$) si acquisisce pertanto una "nuova informazione" rappresentata dall'innovazione all'indietro,

$$\tilde{\mathbf{e}}_n(t-1) = \mathbf{y}(t-n-1) - \hat{E}[\mathbf{y}(t-n-1) | \mathbf{y}_{[t-n,t]}] \quad (2.8)$$

TEOREMA 8.1. I processi di innovazione in avanti e all'indietro con memoria di n dati, definiti dalle formule (2.1) e (2.3) soddisfano il sistema di equazioni alle differenze (nella variabile n)

$$\mathbf{e}_{n+1}(t) = \mathbf{e}_n(t) - \bar{\Gamma}_n \tilde{\mathbf{e}}_n(t-1) \quad , \quad \mathbf{e}_0(t) = \mathbf{y}(t) \quad (2.9a)$$

$$\tilde{\mathbf{e}}_{n+1}(t) = \tilde{\mathbf{e}}_n(t-1) - \Gamma_n \mathbf{e}_n(t) \quad , \quad \tilde{\mathbf{e}}_0(t) = \mathbf{y}(t) \quad , \quad (2.9b)$$

dove le matrici $m \times m$, Γ_n e $\bar{\Gamma}_n$, dette "coefficienti di riflessione", sono date dalle formule

$$\bar{\Gamma}_n = E[\mathbf{e}_n(t) \tilde{\mathbf{e}}_n(t-1)'] \bar{\Lambda}_n^{-1} \quad , \quad (2.10a)$$

$$\Gamma_n = E[\tilde{\mathbf{e}}_n(t-1) \mathbf{e}_n(t)'] \Lambda_n^{-1} \quad . \quad (2.10b)$$

Le matrici Λ_n e $\bar{\Lambda}_n$ sono le varianze degli errori di predizione a memoria n ,

$$\Lambda_n = E \mathbf{e}_n(t) \mathbf{e}_n(t)' \quad , \quad \bar{\Lambda}_n = E \tilde{\mathbf{e}}_n(t) \tilde{\mathbf{e}}_n(t)' \quad (2.11)$$

e, al pari di Γ_n e $\bar{\Gamma}_n$, non dipendono da t per la stazionarietà dei processi di innovazione. Esse sono legate ai coefficienti di riflessione dalle formule ricorsive

$$\Lambda_{n+1} = (I - \bar{\Gamma}_n \Gamma_n) \Lambda_n \quad , \quad \Lambda_0 = \Sigma(0) \quad , \quad (2.12a)$$

$$\bar{\Lambda}_{n+1} = (I - \Gamma_n \bar{\Gamma}_n) \bar{\Lambda}_n \quad , \quad \bar{\Lambda}_0 = \Sigma(0) \quad . \quad (2.12b)$$

DIMOSTRAZIONE. Il vettore $\tilde{\mathbf{e}}_n(t-1)$ è per costruzione ortogonale a $\mathbf{y}_{[t-n,t]}$, per cui

$$\mathbf{H}(\mathbf{y}_{[t-n-1,t]}) = \mathbf{H}(\mathbf{y}_{[t-n,t]}) \oplus \mathbf{H}(\tilde{\mathbf{e}}_n(t-1)) \quad . \quad (2.13)$$

Questa decomposizione ortogonale permette allora di esprimere l'errore di predizione in avanti a memoria $n+1$, come

$$\mathbf{e}_{n+1}(t) = \mathbf{y}(t) - \hat{E}[\mathbf{y}(t) | \mathbf{y}_{[t-n-1,t]}] = \mathbf{y}(t) - \hat{E}[\mathbf{y}(t) | \mathbf{y}_{[t-n,t]}] - \hat{E}[\mathbf{y}(t) | \tilde{\mathbf{e}}_n(t-1)] \quad .$$

Notiamo allora che $\mathbf{y}(t) = \mathbf{e}_n(t) + \hat{E}[\mathbf{y}(t) | \mathbf{y}_{[t-n,t]}]$ (vedi la (2.1)), per cui l'ortogonalità di $\mathbf{y}_{[t-n,t]}$ a $\tilde{\mathbf{e}}_n(t-1)$ porta a concludere che l'ultimo addendo si può scrivere

$$\hat{E}[\mathbf{y}(t) | \tilde{\mathbf{e}}_n(t-1)] = \hat{E}[\mathbf{e}_n(t) | \tilde{\mathbf{e}}_n(t-1)] \quad .$$

Sostituendo questa relazione nella precedente si ottiene

$$\mathbf{e}_{n+1}(t) = \mathbf{e}_n(t) - \hat{E}[\mathbf{e}_n(t) | \tilde{\mathbf{e}}_n(t-1)] \quad ,$$

che è la prima delle (2.9).

La simmetria di questa espressione fa nascere il sospetto che in questo problema $\{\mathbf{e}_n(t)\}$ ed $\{\tilde{\mathbf{e}}_n(t)\}$ giochino un ruolo completamente simmetrico.

In effetti, *partendo* ora da un problema di *predizione all'indietro* in cui si desidera predire il vettore fisso $\mathbf{y}(t-n-1)$ con memoria crescente da n ad $n+1$ dati e ortogonalizzando gli $n+1$ dati $\mathbf{y}_{(t-n-1,t]}$ secondo la

$$\begin{aligned} \mathbf{H}(\mathbf{y}_{(t-n-1,t]}) &= \mathbf{H}(\mathbf{y}_{(t-n-1,t-1]}) \oplus \mathbf{H}(\mathbf{e}_n(t)) \\ &= \mathbf{H}(\mathbf{y}_{[t-n,t]}) \oplus \mathbf{H}(\mathbf{e}_n(t)) \end{aligned} \quad (2.14)$$

si ottiene una formula per l'aggiornamento di $\tilde{\mathbf{e}}_n(t-1)$, perfettamente analoga a quella trovata per $\{\mathbf{e}_n(t)\}$, ovvero

$$\begin{aligned} \tilde{\mathbf{e}}_{n+1}(t) &= \mathbf{y}(t-n-1) - \hat{E}[\mathbf{y}(t-n-1) | \mathbf{H}(\mathbf{y}_{[t-n,t]})] - \hat{E}[\mathbf{y}(t-n-1) | \mathbf{H}(\mathbf{e}_n(t))] \\ &= \tilde{\mathbf{e}}_n(t-1) - E[\tilde{\mathbf{e}}_n(t-1) | \mathbf{e}_n(t)] \quad . \end{aligned}$$

A questo punto l'unica cosa che rimane da provare sono le formule (2.12) giacché le (2.9) sono semplicemente le espressioni esplicite degli operatori di proiezione.

Calcolando la varianza dei due membri nella (2.9) si ha

$$\Lambda_{n+1} = \Lambda_n - E[\mathbf{e}_n(t) \tilde{\mathbf{e}}_n(t-1)'] \bar{\Gamma}'_n - \bar{\Gamma}_n E[\tilde{\mathbf{e}}_n(t-1) \mathbf{e}_n(t)'] + \bar{\Gamma}_n \bar{\Lambda}_n \bar{\Gamma}'_n$$

e, usando le (2.10),

$$\Lambda_{n+1} = \Lambda_n - \bar{\Gamma}_n \Gamma_n \Lambda_n$$

che è appunto la (2.12). La (2.12) si ricava in modo analogo. \square

Abbiamo così dimostrato il risultato fondamentale della teoria, che potremmo chiamare "forma astratta" dell'algoritmo di Levinson. Le relazioni (2.9) compaiono nel lavoro di Benveniste e Chaure [17].

Forma concreta dell'algoritmo. In questo paragrafo studieremo un'implementazione concreta delle formule (2.9).

Dato che $\mathbf{e}_n(t)$ è una funzione lineare dei dati $\mathbf{y}_{[t-n,t]}$, si può rappresentare nella forma

$$\mathbf{e}_n(t) = \sum_{k=t-n}^t A_n(t, k) \mathbf{y}(k) \quad , \quad (2.15)$$

dove $A_n(t, k)$ sono matrici $m \times m$, univocamente determinate in base alle ipotesi poste sul processo (infatti $\mathbf{y}(t-n), \dots, \mathbf{y}(t)$ sono vettori linearmente indipendenti in $\mathbf{H}(\mathbf{y})$). Dalla stazionarietà di $\{\mathbf{e}_n(t)\}$ segue che

$$\mathbf{e}_n(t+s) = U_s \mathbf{e}_n(t) = \sum_{j=t+s-n}^{t+s} A_n(t+s, j) \mathbf{y}(j) = \sum_{k=t-n}^t A_n(t, k) \mathbf{y}(k+s)$$

e, ponendo $j = k + s$ nella prima sommatoria si trova

$$A_n(t+s, k+s) = A_n(t, k) \quad , \quad \forall s \in \mathbb{Z} \quad ,$$

per cui $A_n(t, k)$ dipende solo dalla differenza $t - k$ dei due argomenti temporali, t e k .

La (2.15) può così essere riscritta nella forma

$$\mathbf{e}_n(t) = \sum_{k=t-n}^t A_n(t-k) \mathbf{y}(k) = \sum_0^n A_n(k) \mathbf{y}(t-k) \quad . \quad (2.16)$$

Esaminando la formula (2.1) che definisce $\mathbf{e}_n(t)$ si vede poi immediatamente che deve essere

$$A_n(0) = I \quad (2.17)$$

e anche

$$\hat{E}[\mathbf{y}(t) \mid \mathbf{y}_{[t-n,t]}] = - \sum_1^n A_n(k) \mathbf{y}(t-k) \quad , \quad (2.18)$$

in altre parole, gli $\{-A_n(k) \ , \ k = 1, \dots, n\}$ sono i coefficienti del predittore di un passo a memoria n , del processo $\{\mathbf{y}(t)\}$.

Il calcolo dei coefficienti $A_n(k)$ del "filtro sbiancante" a memoria finita (2.16) è quindi esattamente lo stesso problema del calcolo esplicito del predittore di un passo. Sarà bene tener presente questa equivalenza in seguito.

Ovviamente gli $A_n(k)$ si possono determinare usando il principio di ortogonalità. Quest'ultimo, scritto nella forma

$$\bar{E} \mathbf{e}_n(t) \mathbf{y}(t-s)' = 0 \quad , \quad s = 1, \dots, n \quad , \quad (2.19)$$

fornisce

$$\sum_{k=0}^n A_n(k) \Sigma(s-k) = 0 \quad , \quad s = 1, \dots, n \quad , \quad (2.20)$$

dove $\Sigma(t)$ è la matrice di covarianza di $\{\mathbf{y}(t)\}$.

La (2.20) si può riscrivere in forma matriciale,

$$[A_n(0) \ A_n(1) \ \dots \ A_n(n)] \begin{bmatrix} \Sigma(1) & \dots & \Sigma(n) \\ \Sigma(0) & & \Sigma(n-1) \\ \dots & & \\ \Sigma(1-n) & \dots & \Sigma(0) \end{bmatrix} = \mathbf{0} \quad . \quad (2.21)$$

Questa equazione, tenuto conto della $A_n(0) = I$, può essere vista come caso particolare della relazione generale (2.12) che determina i coefficienti del predittore.

Anche l'innovazione all'indietro può, per ogni n fissato, essere espressa nella forma

$$\tilde{\mathbf{e}}_n(t) = \bar{\mathbf{e}}_n(t-n) = \sum_{k=t-n}^t B_n(t-k) \mathbf{y}(k) = \sum_0^n B_n(k) \mathbf{y}(t-k) \quad , \quad (2.22)$$

dove le matrici $B_n(k)$, $k = 0, 1, \dots, n-1$ sono (in accordo con la (2.3)) semplicemente i coefficienti, cambiati di segno, del predittore all'indietro, a memoria n di $\mathbf{y}(t-n)$,

$$\hat{E}[\mathbf{y}(t-n) | \mathbf{y}_{(t-n,t)}] = - \sum_0^{n-1} B_n(k) \mathbf{y}(t-k) \quad , \quad (2.23a)$$

$$B_n(n) = I \quad . \quad (2.23b)$$

In perfetta dualità a quanto visto per gli $\{A_n(k)\}$ questi coefficienti si possono, in linea di principio, determinare in base al principio di ortogonalità

$$E \tilde{\mathbf{e}}_n(t) \mathbf{y}(t-s)' = 0 \quad , \quad s = 0, 1, \dots, n-1 \quad ,$$

il quale fornisce

$$\sum_0^n B_n(k) \Sigma(s-k) = 0 \quad , \quad s = 0, 1, \dots, n-1 \quad . \quad (2.24)$$

Queste n equazioni matriciali possono ancora venire riscritte in modo compatto nella forma

$$[B_n(0) \dots B_n(n)] \begin{bmatrix} \Sigma(0) & \dots & \Sigma(n-1) \\ \Sigma(-1) & & \dots \\ \dots & & \Sigma(0) \\ \Sigma(-n) & \dots & \Sigma(-1) \end{bmatrix} = 0 \quad , \quad (2.25)$$

la quale è, in un senso che spiegheremo meglio tra poco, *duale* della (2.21).

Forma concreta dell'algoritmo di Levinson. Confrontando le formule astratte (2.9) e le rappresentazioni "concrete" delle innovazioni a memoria finita (2.16) e (2.22), si capisce la possibilità di riscriverle sotto forma di algoritmo per l'aggiornamento ricorsivo rispetto ad n , dei coefficienti $\{A_n(k)\}$ e $\{B_n(k)\}$.

TEOREMA 8.2 (Algoritmo di Levinson). *Le matrici*

$$A_n := [A_n(0), \dots, A_n(n)] \quad (2.26)$$

$$B_n := [B_n(0), \dots, B_n(n)] \quad (2.27)$$

dei coefficienti dei filtri sbiancanti che ortogonalizzano, rispettivamente, in avanti e all'indietro le variabili $\mathbf{y}(t-n), \mathbf{y}(t-n+1), \dots, \mathbf{y}(t)$ del processo stazionario $\{\mathbf{y}(t)\}$ si aggiornano, per $n = 0, 1, \dots$, mediante lo schema ricorsivo seguente.

Note A_n e B_n (di dimensione $m \times (n+1)m$), A_{n+1} e B_{n+1} si ottengono mediante le

$$\begin{aligned} A_{n+1} &= [A_n \ 0] - \bar{\Gamma}_n [0 \ B_n] \quad , \quad A_0 = I \quad , \\ B_{n+1} &= [0 \ B_n] - \Gamma_n [A_n \ 0] \quad , \quad B_0 = I \quad , \end{aligned} \quad (2.28)$$

dove i coefficienti di riflessione $\Gamma_n, \bar{\Gamma}_n$ si ottengono dalle

$$\bar{\Gamma}_n = P_n \bar{\Lambda}_n^{-1} \quad , \quad \Gamma_n = P_n' \Lambda_n^{-1} \quad (2.29)$$

la matrice P_n essendo definita dalla,

$$P_n := \sum_0^n A_n(k) \Sigma(n+1-k) = \sum_0^n \Sigma(k+1) B_n(k)' \quad . \quad (2.30)$$

Le matrici varianza dell'errore di predizione a memoria n , rispettivamente in avanti, Λ_n , e all'indietro, $\bar{\Lambda}_n$, si aggiornano mediante le equazioni

$$\begin{aligned} \Lambda_{n+1} &= (I - \bar{\Gamma}_n \Gamma_n) \Lambda_n \quad , \quad \Lambda_0 = \Sigma(0) \\ \bar{\Lambda}_{n+1} &= (I - \Gamma_n \bar{\Gamma}_n) \bar{\Lambda}_n \quad , \quad \bar{\Lambda}_0 = \Sigma(0) \quad . \end{aligned}$$

DIMOSTRAZIONE. Introducendo le espressioni (2.16) e (2.22) in (2.9) e tenendo conto dell'unicità delle rappresentazioni di $\mathbf{e}_{n+1}(t)$ ed $\tilde{\mathbf{e}}_{n+1}(t)$ come funzioni lineari delle variabili del processo, si ottiene

$$\sum_0^{n+1} A_{n+1}(k) \mathbf{y}(t-k) = \sum_0^n A_n(k) \mathbf{y}(t-k) - \bar{\Gamma}_n \sum_0^n B_n(k) \mathbf{y}(t-1-k)$$

e

$$\sum_0^{n+1} B_{n+1}(k) \mathbf{y}(t-k) = \sum_0^n B_n(k) \mathbf{y}(t-1-k) - \Gamma_n \sum_0^n A_n(k) \mathbf{y}(t-k) \quad .$$

Uguagliando i coefficienti dei termini in cui $\mathbf{y}(\cdot)$ appare con lo stesso argomento temporale (questo è una specie di "principio di identità" simile a quello che vale per i coefficienti dei polinomi), si arriva facilmente alle formule (2.28).

Per ricavare le espressioni (2.29) dei coefficienti di riflessione Γ_n e $\bar{\Gamma}_n$ in funzione dei coefficienti A_n e B_n , basta ricordare che dalle definizioni (2.1) e (2.8) di $\mathbf{e}_n(t)$ ed $\tilde{\mathbf{e}}_n(t-1)$, seguono le relazioni di ortogonalità

$$\begin{aligned} \mathbf{e}_n(t) &\perp \hat{E}[\mathbf{y}(t-n-1) \mid \mathbf{y}_{[t-n,t]}] \\ \tilde{\mathbf{e}}_n(t-1) &\perp \hat{E}[\mathbf{y}(t) \mid \mathbf{y}_{[t-n,t]}] \end{aligned}$$

per cui

$$E \mathbf{e}_n(t) \tilde{\mathbf{e}}_n(t-1)' = E \mathbf{e}_n(t) \mathbf{y}(t-n-1)' = E \mathbf{y}(t) \tilde{\mathbf{e}}_n(t-1)' \quad ,$$

e usando le espressioni (2.16) e (2.22) delle innovazioni a memoria finita, si vede che gli ultimi due termini forniscono proprio le espressioni in (2.30). Il resto è ormai ovvio. \square

COROLLARIO 8.1. *I coefficienti di riflessione a memoria n sono legati fra loro dalla relazione*

$$\bar{\Gamma}_n \bar{\Lambda}_n = [\Gamma_n \Lambda_n]' \quad (2.31)$$

DIMOSTRAZIONE. La (2.31) segue immediatamente dalla (2.29) oppure confrontando tra loro la (2.10a) e la (2.10b). \square

Le formule (2.28) sono state ricavate, per il caso di processi scalari (e non precisamente nella forma in cui le abbiamo presentate) da Levinson [92] e son state successivamente riscoperte e valorizzate da Durbin [40] nel 1960. Le formule di aggiornamento per il caso vettoriale sono state ottenute successivamente da Whittle [148] e da Wiggins e Robinson [153]. Come si vede esse hanno la struttura di un

algoritmo ricorsivo per l'aggiornamento dei coefficienti dei filtri sbiancanti (2.16) e (2.22), o equivalentemente, i coefficienti del predittore di un passo (a memoria finita) di un processo stazionario $\{\mathbf{y}(t)\}$ del quale sia nota la matrice di covarianza, all'aumentare della dimensione, n , della memoria.

Dalla struttura delle equazioni di aggiornamento (2.28) emerge il ruolo fondamentale giocato dai coefficienti di riflessione Γ_n e $\bar{\Gamma}_n$. In effetti *la conoscenza delle successioni* $\{\Gamma_n\}$ e $\{\bar{\Gamma}_n\}$ è equivalente a quella delle successioni dei coefficienti $\{A_n\}$ e $\{B_n\}$ nel seguente senso: assegnata la successione $\{\Gamma_k, \bar{\Gamma}_k; k = 0, 1, \dots, n\}$ le equazioni (2.28) determinano $\{A_k, B_k\}$ per $k = 1, 2, \dots, n+1$ e, viceversa, noti i coefficienti $\{A_k, B_k\}$ per $k = 1, 2, \dots, n+1$, i coefficienti di riflessione sono determinati dalla

$$\bar{\Gamma}_k = -A_{k+1}(k+1), \quad \Gamma_k = -B_{k+1}(0) \quad k = 0, 1, \dots, n. \quad (2.32)$$

Questa osservazione andrà, come vedremo più avanti, complementata con quella forse ancora più importante, che lega la *stabilità* dei filtri sbiancanti (2.16) e (2.22) a certe condizioni di carattere algebrico sui coefficienti di riflessione.

Usando la relazione (2.16) si possono esprimere le innovazioni a memoria finita $\mathbf{e}_n(t+n)$ e $\bar{\mathbf{e}}_n(t) \equiv \bar{\mathbf{e}}_n(t+n)$ nella forma simbolica

$$\mathbf{e}_n(t+n) = A_n(z) \mathbf{y}(t), \quad \bar{\mathbf{e}}_n(t) = B_n(z) \mathbf{y}(t) \quad n = 0, 1, 2, \dots$$

dove i polinomi matriciali $\{A_n(z)\}$ (e i duali $\{B_n(z)\}$) sono definiti mediante la posizione

$$\begin{aligned} A_n(z) &:= z^n \sum_0^n A_n(k) z^{-k} = \sum_0^n A_n(k) z^{n-k} \\ B_n(z) &:= z^n \sum_0^n B_n(k) z^{-k} = \sum_0^n B_n(k) z^{n-k} \end{aligned} \quad (2.33)$$

Notare che, in accordo con il significato attribuitogli a suo tempo, l'innovazione all'indietro $\bar{\mathbf{e}}_n(t)$ è espressa come una combinazione anticausale di dati futuri.

PROBLEMA 8.1. *Verificare che in forma "concreta", l'ortogonalità dei vettori innovazione a memoria finita $\mathbf{e}_0(t), \mathbf{e}_1(t+1), \dots, \mathbf{e}_k(t+k), \dots$, stabilita al lemma 8.1, si può scrivere come*

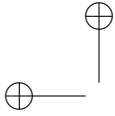
$$\oint_{|z|=1} A_k(z) S(z) A_p(z^{-1})' \frac{dz}{z} = \Lambda_k \delta(k-p) \quad (2.34)$$

La matrice peso S è la densità spettrale del processo \mathbf{y} . La (2.34) mostra che i polinomi $\{A_n(z)\}$ sono tra loro ortogonali (riga per riga) rispetto al prodotto scalare dello spazio $L_m^2([-\pi, \pi], S \frac{d\omega}{2\pi})$ (si veda l'Appendice B).

Scrivere la relazione di ortogonalità che vale per i polinomi duali $B_n(z)$.

Polinomi (scalari) ortogonali sulla circonferenza unità rispetto ad un prodotto scalare "pesato" di questo tipo, sono stati introdotti da G. Szegö [137] e studiati successivamente da vari autori. Il caso matriciale è trattato in profondità negli articoli [37, 162]. \diamond

Problemi con istante iniziale fissato. Notiamo che in un problema di predizione con istante iniziale (t_0) finito e fissato a priori, conviene usare l'algoritmo pensando l'istante più antico, $t-n$, fissato e uguale a t_0 e immaginare che la "base temporale" dello stimatore si allarghi incrementando t . Questo anche se il procedimento che abbiamo seguito nel derivare l'algoritmo presupponeva, viceversa, t costante. In effetti tutte le formule ricavate finora devono ovviamente essere invarianti rispetto alla "data di riferimento" che si considera, per la stazionarietà del processo.



Se si sceglie $t_0 = 0$ basta sostituire in tutte le formule il pedice n con t . In particolare, il problema di predizione all'indietro, con memoria di t dati, consiste ora nello stimare $\mathbf{y}(0)$ in base ai dati $\mathbf{y}(1), \dots, \mathbf{y}(t)$.

OSSERVAZIONE 8.1. *Notiamo che per calcolare i coefficienti che interessano in un problema di predizione, ovvero le matrici $-A_t(k)$, $k = 1, \dots, t$, e la varianza dell'errore, Λ_t , occorre comunque aggiornare anche le quantità "duali", $B_t(k)$, $k = 1, \dots, t$, e $\bar{\Lambda}_t$.* \diamond

Questa è in fondo una caratteristica comune a molti problemi di ottimizzazione (nei quali si può, in fin dei conti, far rientrare anche la teoria della stima a minima varianza)

3. Il caso scalare

Nel caso in cui il processo $\{\mathbf{y}(t)\}$ è scalare ($m = 1$) si ha una notevole semplificazione nelle formule. La ragione della semplificazione risiede nella proprietà di *reversibilità temporale* del processo, che può essere spiegata in due parole dicendo che il processo originale $\{\mathbf{y}(t)\}$ e il suo "reciproco temporale" $\{\mathbf{z}(t)\}$, definito ponendo

$$\mathbf{z}(t) = \mathbf{y}(-t) \quad , \quad \forall t \in \mathbb{Z} \quad , \quad (3.1)$$

hanno la stessa funzione di covarianza. In effetti

$$E\mathbf{z}(t + \tau) \mathbf{z}(t) = E\mathbf{y}(-t - \tau) \mathbf{y}(-t) = \sigma(-\tau) = \sigma(\tau)$$

Notiamo che questa proprietà, più che con la dimensione m , ha a che vedere con la parità della matrice di covarianza $\Sigma(t)$ del processo (cioè con la proprietà $\Sigma(-t) = \Sigma(t)$, equivalente, nel caso vettoriale, alla $\Sigma(t) = \Sigma'(t)$). Se $\Sigma(t)$ è una funzione pari di t , tutto quanto diremo più sotto è generalizzabile al caso vettoriale.

Riscriviamo la relazione (2.21) che determina i coefficienti del filtro sbiancante in avanti. Dato che siamo nel caso scalare useremo dappertutto lettere minuscole.

Tenendo conto che $\sigma(-t) = \sigma(t)$ e che $a_n(0) = 1$, l'equazione (2.21) si può riscrivere nella forma (non omogenea)

$$[a_n(1), \dots, a_n(n)] \begin{bmatrix} \sigma(0) & \dots & \sigma(n-1) \\ \dots & & \dots \\ \sigma(n-1) & \dots & \sigma(0) \end{bmatrix} = -[\sigma(1), \dots, \sigma(n)] \quad (3.2)$$

ovvero, in forma compatta,

$$a_n \Sigma_n = -\sigma_n \quad , \quad (3.3)$$

dove Σ_n è la matrice di *Toeplitz* delle covarianze che compare a primo membro in (3.2) e σ_n è il vettore riga $[\sigma(1), \dots, \sigma(n)]$.

Ricordiamo che una matrice quadrata T si dice "di Toeplitz" (o di Toeplitz a blocchi) se l'elemento (o il blocco) di posto (i, j) , t_{ij} dipende solo dalla differenza $i - j$. In altri termini, gli elementi che appartengono alla diagonale principale e alle sue parallele in T sono tutti uguali. Se T è anche simmetrica (com'è nel nostro caso) viene qualche volta chiamata "ipersimmetrica".

Riscrivendo in modo analogo la (2.25) che determina i coefficienti $b_n(k)$ del filtro sbiancante all'indietro, si trova

$$[b_n(0), \dots, b_n(n-1)] \begin{bmatrix} \sigma(0) & \dots & \sigma(n-1) \\ \dots & & \dots \\ \sigma(n-1) & \dots & \sigma(0) \end{bmatrix} = -[\sigma(n), \dots, \sigma(1)] \quad , \quad (3.4)$$

dove ora abbiamo usato il fatto che $b_n(n) = 1$. A secondo membro di questa equazione compare un vettore riga che ha *le stesse componenti di σ_n nella (3.3) ma in ordine opposto*. Definiamo in generale il *reciproco*, x^* , di un vettore n -dimensionale $x = [x_1, \dots, x_n]$ ponendo

$$x^* := [x_n, \dots, x_1] \quad , \quad (3.5)$$

ovvero

$$x_k^* = x_{n-k} \quad , \quad k = 1, \dots, n \quad .$$

Si vede allora che il vettore, b_n , dei coefficienti del predittore all'indietro soddisfa l'equazione:

$$b_n \Sigma_n = -\sigma_n^* \quad . \quad (3.6)$$

Ora, una proprietà generale delle matrici di Toeplitz simmetriche (ricordiamo che la simmetria corrisponde alla reversibilità temporale) è di essere invarianti rispetto alla “reciprocazione”, ovvero

$$(xT)^* = x^* T \quad , \quad \forall x \in \mathbb{R}^n \quad . \quad (3.7)$$

(Vale ovviamente anche l'analoga per vettori colonna e moltiplicazione a destra).

Ne scende allora immediatamente, data l'unicità della soluzione delle equazioni (3.2) e (3.4), *che b_n è necessariamente il reciproco di a_n* , ovvero

$$b_n(k) = a_n(n-k) \quad , \quad k = 0, 1, \dots, n-1 \quad . \quad (3.8)$$

Notiamo subito che la (3.8) vale anche per $k = n$ (dato che $a_n(0) = b_n(n) = 1$) e quindi

$$\begin{aligned} \tilde{\mathbf{e}}_n(t) &= \sum_0^n b_n(k) \mathbf{y}(t-k) = \sum_0^n a_n(n-k) \mathbf{y}(t-k) \\ &= \sum_{j=0}^n a_n(j) \mathbf{y}(t-n+j) = a_n \begin{bmatrix} \mathbf{y}(t-n) \\ \vdots \\ \mathbf{y}(t) \end{bmatrix} \quad . \end{aligned} \quad (3.9)$$

Da questa formula segue che $\bar{\lambda}_n^2 = \text{var } \tilde{\mathbf{e}}_n(t) = a_n \Sigma_n a_n' = \text{var } \mathbf{e}_n(t) = \lambda_n^2$. In altri termini *le varianze dei due processi di innovazione a memoria finita $\tilde{\mathbf{e}}_n(t)$ e $\mathbf{e}_n(t)$ coincidono*, ovvero

$$\bar{\lambda}_n^2 = \lambda_n^2 \quad . \quad (3.10)$$

Abbiamo così eliminato completamente le quantità duali b_n e $\bar{\lambda}_n$ che entravano nell'algoritmo generale del Teorema 8.2.

COROLLARIO 8.2 (Algoritmo di Levinson, caso scalare). *Il vettore dei coefficienti detto*

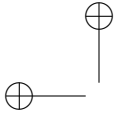
$$a_n := [1, a_n(1), \dots, a_n(n)] \quad ,$$

del filtro sbiancante a memoria n che ortogonalizza (in avanti) le variabili del processo scalare $\{\mathbf{y}(t)\}$ di covarianza $\sigma(t)$, si aggiorna ricorsivamente in n , secondo lo schema

$$a_{n+1} = [a_n \ 0] - \gamma_n [0 \ a_n^*] \quad , \quad a_0 = 1 \quad , \quad (3.11)$$

dove a_n^* è il vettore reciproco di a_n , di componenti

$$a_n^*(k) = a_n(n-k) \quad , \quad k = 0, 1, \dots, n \quad , \quad (3.12)$$



e γ_n è il coefficiente di riflessione

$$\gamma_n = \frac{1}{\lambda_n^2} \sum_0^n a_n(k) \sigma(n+1-k) \quad . \quad (3.13)$$

La varianza dell'errore di predizione a memoria n , λ_n^2 , che compare nella formula precedente, si ottiene come soluzione dell'equazione alle differenze

$$\lambda_{n+1}^2 = (1 - \gamma_n^2) \lambda_n^2 \quad , \quad \lambda_0^2 = \sigma(0) \quad . \quad (3.14)$$

A proposito dell'utilizzazione di questo schema in problemi di stima con istante iniziale fissato, si possono ripetere qui le considerazioni fatte prima. Assumendo ancora $t - n = t_0 = 0$, bisognerà sostituire dappertutto n con t . In particolare, il predittore di un passo $\hat{\mathbf{y}}(t | t - 1)$ basato sugli $n = t$ dati, $\mathbf{y}(0), \mathbf{y}(1), \dots, \mathbf{y}(t - 1)$, potrà scriversi

$$\hat{\mathbf{y}}(t | t - 1) = - \sum_1^t a_t(k) \mathbf{y}(t - k)$$

e il vettore dei coefficienti $a_t := [1, a_t(1), \dots, a_t(t)]$ potrà essere aggiornato usando lo schema ricorsivo del corollario in cui si è sostituita la variabile n con t .

Dato che nel caso scalare si ha l'uguaglianza delle varianze d'errore $\lambda_n^2 = \text{var } \mathbf{e}_n(t) = \text{var } \tilde{\mathbf{e}}_n(t - 1)$, la definizione (2.10) del coefficiente di riflessione, porge,

$$\gamma_n = \frac{E \mathbf{e}_n(t) \tilde{\mathbf{e}}_n(t - 1)}{\lambda_n^2} = \frac{E \mathbf{e}_n(t) \tilde{\mathbf{e}}_n(t - 1)}{(\text{var } \mathbf{e}_n(t))^{1/2} (\text{var } \tilde{\mathbf{e}}_n(t - 1))^{1/2}} \quad (3.15)$$

dalla quale scende che γ_n è un coefficiente di correlazione. In realtà è un coefficiente di correlazione tra le variabili $\mathbf{e}_n(t)$ e $\tilde{\mathbf{e}}_n(t - n - 1)$ che sono relative agli estremi dell'intervallo $(t - n, t - 1)$. In letteratura γ_n viene chiamato *coefficiente di correlazione parziale*. Ovviamente la (3.15) implica che $|\gamma_n| \leq 1$ per ogni n . Questa condizione, associata alla formula (3.14), che descrive il comportamento della varianza dell'errore di predizione al crescere del numero di dati disponibili, implica ovviamente che le varianze $\{\lambda_n^2\}$ formano una *successione monotona decrescente*

$$\lambda_{n+1}^2 \leq \lambda_n^2 \quad , \quad n = 0, 1, \dots \quad ,$$

il che (com'era da aspettarsi) significa che il predittore a memoria finita migliora le sue prestazioni al crescere di n . ◇

PROBLEMA 8.2. *Dimostrare che se il processo (scalare) \mathbf{y} è puramente non deterministico, tutti i coefficienti di riflessione $\{\gamma_n\}$ debbono essere strettamente minori di uno, i.e.*

$$\gamma_n < 1, \quad n = 1, 2, \dots \quad (3.16)$$

Viceversa, si assuma che valga la la (3.16). Dimostrare usando la relazione (3.14), che si ha $\det \Sigma_n > 0$ per ogni n . Questa condizione implica che il processo \mathbf{y} non può avere righe spettrali (si tratta di una condizione più debole del puro non determinismo).

Purtroppo non sembra ancora esistere una caratterizzazione così netta dei coefficienti di riflessione nel caso vettoriale. Si vedano a questo proposito [105] e [132, p. 318-319].

4. Filtri di Levinson

Studieremo ora le proprietà dinamiche del predittore di Levinson (o meglio, del filtro sbiancante) a memoria finita n , supponendo di aver fissato n .

Notiamo che la funzione di trasferimento (matriciale) del filtro sbiancante in avanti (2.16) si può esprimere come

$$\sum_0^n A_n(k) z^{-k} = z^{-n} \sum_0^n A_n(k) z^{n-k} := A_n^*(z^{-1}), \quad (4.1)$$

Analogamente quella del filtro sbiancante all'indietro (2.22) si può scrivere

$$\sum_0^n B_n(k) z^{-k} = z^{-n} \sum_0^n B_n(k) z^{n-k} := B_n^*(z^{-1}) \quad . \quad (4.2)$$

Aderendo ad una convenzione usata già in passato, abbiamo usato i simboli $A_n^*(z)$ e $B_n^*(z)$ per denotare i polinomi matriciali *reciproci*¹ dei polinomi matriciali $A_n(z)$ e $B_n(z)$, introdotti in (2.33).

Notiamo che nel caso scalare la condizione (3.8) fa sì che $B_n(z)$ sia esso stesso il reciproco di $A_n(z)$:

$$B_n(z) = z^n A_n(z^{-1}) := A_n^*(z) \quad . \quad (4.3)$$

Notiamo in particolare che gli zeri di $B_n(z)$ sono esattamente i reciproci (ora nel senso usuale del termine) di quelli di $A_n(z)$.

Dimostriamo ora un'importante proprietà del filtro sbiancante (a memoria n): la sua funzione di trasferimento è (strettamente stabile e causale e) *ad inversa strettamente stabile e causale*. In altri termini, $A_n^*(z^{-1})$ può essere davvero pensata come l'inversa di un opportuno fattore spettrale a fase minima.

TEOREMA 8.3 (Whittle). *Nelle ipotesi su $\{\mathbf{y}(t)\}$ fatte all'inizio di questo capitolo, il polinomio $A_n(z)$ ha tutti gli zeri dentro al cerchio unitario aperto. In altri termini l'equazione algebrica*

$$\det A_n(z) = 0$$

ha solo radici a modulo strettamente minore di uno. Dualmente, $\det B_n(z)$ può solamente annullarsi fuori del cerchio unità.

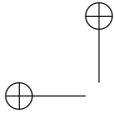
DIMOSTRAZIONE. Definiamo il vettore $m \times n$ -dimensionale $\mathbf{z}(t)$ ponendo

$$\mathbf{z}(t) := \begin{bmatrix} \mathbf{y}(t - n + 1) \\ \dots \\ \mathbf{y}(t) \end{bmatrix}$$

e consideriamo lo stimatore a minima varianza di $\mathbf{z}(t)$ in base a $\mathbf{z}(t-1) := U_{-1} \mathbf{z}(t)$. Chiaramente lo stimatore $\hat{\mathbf{z}}(t) = \hat{E}[\mathbf{z}(t) \mid \mathbf{z}(t-1)]$ si potrà scrivere come una funzione lineare dei dati

$$\hat{\mathbf{z}}(t) = F \mathbf{z}(t-1) \quad , \quad (4.4)$$

¹Il polinomio *reciproco* di un polinomio matriciale è definito, analogamente a quanto stipulato nel caso scalare, cf. la definizione (3.5), come quello che ha coefficienti (matriciali) che formano una sequenza reciproca, di quella del polinomio di partenza.



dove F (che è una matrice quadrata di dimensione $mn \times mn$) ha la struttura compagna a blocchi seguente,

$$F = \begin{bmatrix} 0 & I & 0 \\ & & I \\ -A_n(n) & \dots & -A_n(1) \end{bmatrix} . \tag{4.5}$$

Notiamo che l'ultima riga di F corrisponde alla relazione (2.18). L'errore di stima, $\mathbf{e}(t)$, (che per definizione è ortogonale a $\mathbf{z}(t-1)$) ha, corrispondentemente, le prime $m(n-1)$ componenti uguali a zero e le ultime m uguali ad $\mathbf{e}_n(t)$.

Partendo ora dalla decomposizione ortogonale

$$\mathbf{z}(t) = F \mathbf{z}(t-1) + \mathbf{e}(t) \tag{4.6}$$

e calcolando le matrici varianza a primo e secondo membro si trova

$$\Sigma = F \Sigma F' + \Lambda \quad , \tag{4.7}$$

dove Σ è la matrice varianza di $\mathbf{z}(t)$ (e di $\mathbf{z}(t-1)$) e Λ è la matrice diagonale a blocchi

$$\Lambda = E \mathbf{e}(t) \mathbf{e}'(t) = \text{diag} \{0, \dots, 0, \Lambda_n\} \quad . \tag{4.8}$$

Ricordiamo che, in virtù delle ipotesi poste (in particolare il fatto che $\{\mathbf{y}(t)\}$ è p.n.d.), Σ è *definita positiva*. D'altro canto la (4.7) è la ben nota *equazione di Lyapunov discreta* associata al sistema discreto (4.6). Si sa (si veda ad esempio il lemma 9.2, Cap. 10) che, detta $\Lambda^{1/2}$ una qualunque radice quadrata di Λ (ovvero $\Lambda^{1/2} \cdot (\Lambda^{1/2})' = \Lambda$), se $(F, \Lambda^{1/2})$ è completamente raggiungibile, allora l'esistenza di una soluzione definita positiva (Σ) della (2.16) *implica la stabilità asintotica di F* .

Dato che

$$\det(\lambda I - F) = \det A_n(\lambda) \quad ,$$

come si può facilmente verificare (ad esempio mediante operazioni elementari sui blocchi colonna di $\lambda I - F$), si vede che la prova del Teorema si riduce a dimostrare la raggiungibilità completa di $(F, \Lambda^{1/2})$.

Prendiamo $\Lambda^{1/2}$ uguale a $[0, \dots, (\Lambda_n^{1/2})']'$. Questa matrice ha rango m per l'ipotesi di rango pieno fatta sul processo $\{\mathbf{y}(t)\}$. Calcolando la matrice di raggiungibilità $R : [\Lambda^{1/2}, F\Lambda^{1/2}, \dots, F^{n-1}\Lambda^{1/2}]$ si trova (i termini indicati con * non interessano ai fini della dimostrazione),

$$R = \begin{bmatrix} 0 & \dots & 0 & \Lambda_n^{1/2} \\ & & & * \\ 0 & \Lambda_n^{1/2} & & \\ \Lambda_n^{1/2} & * & \dots & * \end{bmatrix} \quad ,$$

Questa espressione mostra chiaramente che R è non singolare prova così la raggiungibilità di $(F, \Lambda^{1/2})$.

Con un ragionamento duale si può provare l'affermazione relativa a $B_n(z)$. \square

Per quanto appena visto, la relazione (2.16) può essere invertita fornendo una rappresentazione vettoriale di tipo autoregressivo del processo $\{\mathbf{y}(t)\}$. Il processo stazionario $\{\mathbf{y}(t)\}$ può insomma essere rappresentato come l'uscita di un *filtro causale e strettamente stabile di tipo AR*, pilotato dal processo delle innovazioni

a memoria finita $\{\mathbf{e}_n(t)\}$. Denoteremo per brevità con $W_n(z) = [A_n^*(z^{-1})]^{-1}$ la matrice di trasferimento del filtro,

$$W_n(z) = \left[\sum_0^n A_n(k) z^{-k} \right]^{-1} \quad (4.9)$$

Dato che il processo $\{\mathbf{e}_n(t)\}$ non è in generale bianco, la terminologia non è a stretto rigore corretta; essa ha però un senso approssimato che preciseremo fra un attimo.

Notiamo per completezza che $\{\mathbf{y}(t)\}$ può, dualmente, essere rappresentato anche come l'uscita di un modello puramente autoregressivo di funzione di trasferimento

$$\bar{W}_n(z) = [B_n^*(z^{-1})]^{-1} = \left[\sum_0^n B_n(k) z^{-k} \right]^{-1} = z^n \left[\sum_0^n B_n(k) z^{n-k} \right]^{-1} \quad (4.10)$$

pilotato dal rumore (ancora in generale non bianco) $\{\bar{\mathbf{e}}_n(t)\}$.

Dato che la (4.10) è strettamente instabile (ha poli tutti in $\{|z| > 1\}$), è opportuno a questo punto reintrodurre l'innovazione all'indietro $\bar{\mathbf{e}}_n(t) = \bar{\mathbf{e}}_n(t+n)$ e riscrivere la (2.22) nella forma

$$\mathbf{y}(t) + \sum_1^n B_n(n-k) \mathbf{y}(t+k) = \bar{\mathbf{e}}_n(t)$$

che va interpretata come uno schema autoregressivo che genera $\{\mathbf{y}(t)\}$ ricorsivamente *all'indietro* nel tempo elaborando i dati futuri da t in poi, $\{\mathbf{y}(t+k) \mid k = 1, \dots, n\}$.

I filtri autoregressivi di Levinson $W_n(z)$ possono essere pensati come *fattori spettrali approssimati* della densità spettrale $S(z)$ del processo $\{\mathbf{y}(t)\}$, nel senso che

$$S(z) \simeq W_n(z) \Lambda_n W_n'(z^{-1}) \quad .$$

Per precisare meglio di che tipo di approssimazione si tratti, introduciamo la seguente definizione.

DEFINIZIONE 8.2. *La matrice densità spettrale*

$$S_n(z) = \sum_{-\infty}^{+\infty} \Sigma_n(\tau) z^{-\tau} \quad (4.11)$$

è un interpolante di ordine n di $S(z)$ se $\Sigma_n(\tau) = \Sigma(\tau)$ per $\tau = 0, \pm 1, \pm 2, \dots, \pm n$, ovvero, equivalentemente, se

$$\Sigma_n(\tau) = \int_{-\pi}^{\pi} S_n(e^{i\omega}) e^{i\omega\tau} \frac{d\omega}{2\pi} = \Sigma(\tau) \quad , \quad \tau = 0, 1, \dots, n \quad . \quad (4.12)$$

È evidente che se una matrice razionale $S_n(z)$ è un interpolante di $S(z)$ per ogni n allora, dall'unicità dei coefficienti di Fourier di $S(z)$, scende che necessariamente $S_n(z) = S(z)$.

Consideriamo allora il processo stazionario $\{\mathbf{y}_n(t)\}$ ottenuto filtrando un *rumore bianco* $\{\mathbf{u}(t)\}$ di varianza Λ_n mediante il filtro autoregressivo (4.9),

$$A_n^*(z^{-1}) \mathbf{y}_n(t) = \mathbf{u}(t) \quad , \quad E \mathbf{u}(t) \mathbf{u}'(s) = \Lambda_n \delta(t-s) \quad . \quad (4.13)$$

Verifichiamo che,

PROPOSIZIONE 8.1. *Lo spettro, $S_n(z)$, di $\{\mathbf{y}_n(t)\}$,*

$$S_n(z) = W_n(z) \Lambda_n W_n'(z^{-1}) \quad , \quad (4.14)$$

è un interpolante di ordine n di $S(z)$.

DIMOSTRAZIONE. Com'è noto, la funzione di covarianza $\Sigma_n(t)$ del processo $\{\mathbf{y}_n(t)\}$ si ottiene risolvendo il sistema di equazioni alle differenze (che si ricava dalla (4.13) tenendo conto del fatto che $\{\mathbf{u}(t)\}$ è necessariamente l'innovazione di $\{\mathbf{y}_n(t)\}$ e pertanto $E \mathbf{u}(t) \mathbf{y}(t)' = E \mathbf{u}(t) \mathbf{u}(t)' = \Lambda_n$)

$$\begin{aligned} \sum_0^n A_n(k) \Sigma_n(-k) &= \Lambda_n \\ \sum_0^n A_n(k) \Sigma_n(s-k) &= 0 \quad , \quad s = 1, 2, \dots \quad . \end{aligned} \quad (4.15)$$

Ricordando ora la definizione (2.16) di $\mathbf{e}_n(t)$ e calcolando esplicitamente Λ_n usando questa formula, si trova

$$\Lambda_n = E \mathbf{e}_n(t) \mathbf{e}_n(t)' = E \mathbf{e}_n(t) \mathbf{y}(t)' = \sum_0^n A_n(k) \Sigma(-k) \quad .$$

D'altro canto i coefficienti $A_n(k)$ soddisfano al sistema di equazioni (2.20) (oppure (2.21)). Aggiungendo a queste equazioni quella testé ricavata si trova:

$$\begin{aligned} \sum_0^n A_n(k) \Sigma(-k) &= \Lambda_n \\ \sum_0^n A_n(k) \Sigma(s-k) &= 0 \quad , \quad s = 1, \dots, n \end{aligned}$$

e confrontando con la (4.15) se ne deduce che

$$\Sigma_n(t) = \Sigma(t) \quad , \quad t = 0; \pm 1, \dots, \pm n \quad . \quad (4.16)$$

□

Dato che $\Sigma(t)$ non soddisfa necessariamente le (2.20) per $|t| > n$, l'evoluzione temporale delle due covarianze risulterà *diversa* fuori dall'intervallo $[-n, n]$, quella di $\Sigma_n(t)$ essendo univocamente determinata dall'equazione alle differenze (4.15) e quella di $\Sigma(t)$ potendo a priori essere sostanzialmente arbitraria (a parte naturalmente il vincolo di positività).

Chiaramente le considerazioni appena fatte possono essere ripetute per il processo "duale" $\{\bar{\mathbf{y}}_n(t)\}$ definito risolvendo all'indietro nel tempo l'equazione:

$$B_n^*(z^{-1}) \bar{\mathbf{y}}_n(t) = \bar{\mathbf{u}}(t) \quad , \quad E \bar{\mathbf{u}}(t) \bar{\mathbf{u}}(s)' = \bar{\Lambda}_n \delta(t-s) \quad . \quad (4.17)$$

Conviene anche in questo caso riscrivere la (4.17) moltiplicando a destra e a sinistra per z^n e definire il rumore bianco $\bar{\mathbf{u}}(t)$ come $\bar{\mathbf{u}}(t+n)$ nella precedente espressione. Con questa operazione si ha

$$B_n(z) \bar{\mathbf{y}}_n(t) = \bar{\mathbf{u}}(t) \quad , \quad E \bar{\mathbf{u}}(t) \bar{\mathbf{u}}(s)' = \bar{\Lambda}_n \delta(t-s) \quad . \quad (4.18)$$

La (4.18) definisce una relazione completamente anticausale (e anticausalmente invertibile) tra $\bar{\mathbf{y}}_n(t)$ e $\bar{\mathbf{u}}(t)$ nel senso che $\bar{\mathbf{u}}(t)$ è una funzione del futuro di $\{\bar{\mathbf{y}}_n(t)\}$ da t in poi e, viceversa, la (4.18) può essere risolta all'indietro (per l'instabilità

completa di $B_n(z)$) e fornire $\bar{\mathbf{y}}_n(t)$ come processo stazionario funzione del futuro di $\{\bar{\mathbf{u}}(t)\}$ dall'istante t in poi. Lo spettro di $\{\bar{\mathbf{y}}_n(t)\}$ può così scriversi ancora nella forma

$$\bar{S}_n(z) = \bar{W}_n(z) \bar{\Lambda}_n \bar{W}'(z^{-1}) \quad .$$

Moltiplicando a destra la (4.18) prima per $\bar{\mathbf{y}}_n(t)'$ e poi per $\bar{\mathbf{y}}_n(t+n-s)'$ (che è ortogonale ad $\bar{\mathbf{u}}(t)$ se $n-s > 0$) si ottengono le formule analoghe alle (4.15) per la covarianza $\bar{\Sigma}_n(t)$ di $\{\bar{\mathbf{y}}_n(t)\}$ e cioè le equazioni di evoluzione all'indietro,

$$\begin{aligned} \sum_0^n B_n(k) \bar{\Sigma}_n(n-k) &= \bar{\Lambda}_n \\ \sum_0^n B_n(k) \bar{\Sigma}_n(s-k) &= 0 \quad , \quad s = n-1, n-2, \dots \end{aligned} \quad (4.19)$$

Confrontando ora le (4.19) con le (2.24) (o la (2.25)) si ottiene un risultato analogo alle (4.16),

$$\bar{\Sigma}_n(t) = \Sigma(t) \quad , \quad |t| \leq n \quad ,$$

che ovviamente implica l'uguaglianza $\Sigma_n(t) = \bar{\Sigma}_n(t)$, $\forall t$ e, in definitiva, la

$$\bar{S}_n(z) = S_n(z) \quad . \quad (4.20)$$

Quanto sopra si può riassumere nel modo seguente.

PROPOSIZIONE 8.2. *I processi stocastici $\{\mathbf{y}_n(t)\}$, $\{\bar{\mathbf{y}}_n(t)\}$ hanno lo stesso spettro razionale $S_n(z)$ (sono processi statisticamente equivalenti). Le rappresentazioni definite dalle equazioni (4.13) e (4.18) sono rappresentazioni di innovazione non normalizzata rispettivamente in avanti e all'indietro associate a $S_n(z)$.*

In altre parole, come visto al capitolo 6 (Predizione all'indietro), $B_n(z)$ e $\bar{\Lambda}_n$ si possono in linea di principio calcolare partendo da una fattorizzazione a fase minima di $S_n(z)'$, necessariamente ancora di tipo autoregressivo,

$$S_n(z)' = \left[I + \sum_1^n \tilde{A}_k z^{-k} \right]^{-1} \bar{\Lambda}_n \left[I + \sum_1^n \tilde{A}_k z^k \right]^{-\top}$$

che fornisce

$$S_n(z) = \left[I + \sum_1^n \tilde{A}_k z^k \right]^{-1} \bar{\Lambda}_n \left[I + \sum_1^n \tilde{A}'_k z^{-k} \right]^{-1}$$

dove la funzione razionale $\left[I + \sum_1^n \tilde{A}_k z^k \right]^{-1}$ ha tutti i suoi poli e i suoi zeri fuori del cerchio unitario. Questo permette di identificare $B_n(z)$ e $\bar{\Lambda}_n$ come

$$B_n(z) = \tilde{A}_n^{-1} \left[I + \sum_1^n \tilde{A}_k z^k \right] \quad \bar{\Lambda}_n = \tilde{A}_n^{-1} \bar{\Lambda}_n \tilde{A}_n^{-\top} \quad (4.21)$$

Tralascieremo la dimostrazione che \tilde{A}_n è invertibile.

Estensione positiva di covarianze. Finora abbiamo ragionato (in linea con quanto siamo venuti esponendo finora) come se la matrice densità spettrale del processo di osservazione fosse un dato del problema. In realtà in pratica questo è assai raramente il caso. In molti problemi di elaborazione di segnali si può spesso pensare che siano invece note le n matrici di covarianza

$$\{\Sigma(\tau) \mid \tau = 0, \pm 1, \pm 2, \dots, \pm n\} \quad (4.22)$$

ottenute ad esempio per mezzo di un qualche procedimento di stima a partire dai dati osservati. Supporremo che le (4.22) siano il tratto iniziale di una funzione di covarianza “vera” $\Sigma(\tau)$ incognita, il che è equivalente a supporre che la matrice di Toeplitz

$$\Sigma_n := \begin{bmatrix} \Sigma(0) & \Sigma(1) & \dots & \Sigma(n) \\ \Sigma(-1) & \Sigma(0) & \dots & \\ \vdots & & & \vdots \\ \Sigma(-n) & \dots & & \Sigma(0) \end{bmatrix}, \quad \Sigma(-\tau) = \Sigma(\tau)'$$

sia *definite positiva*. Il problema che ci si pone è quello di costruire una densità spettrale $S_n(z)$ (descritta ad esempio dallo sviluppo (4.11)) che soddisfi alle condizioni di interpolazione (4.12) o, equivalentemente, un’ *estensione* della sequenza (4.22) per $|\tau| > n$, che, beninteso, sia ancora una funzione di covarianza. Questo è un classico problema, la cui versione scalare, nota come il *problema di estensione positiva di Caratheodory*, è stata molto studiata in analisi complessa all’inizio del novecento e, per la sua importanza in elaborazione di segnali, ha successivamente dato luogo ad una letteratura sterminata, si vedano ad es [78, 53, 29].

È chiaro che il filtro di Levinson $W_n(z)$ (e il suo duale), fornisce, attraverso la formula (4.14) un’estensione di Caratheodory *razionale* della sequenza (4.22).

Di estensioni positive definite, che come si dice comunemente, *interpolano* le $n + 1$ matrici date $\Sigma(\tau)$, $\tau = 0, 1, \dots, n$, se ne possono però costruire infinite [78, 53, 29]. Finora la (4.14) non sembra godere di meriti speciali.

In realtà l’estensione positiva basata sul filtro di Levinson ha una certa importanza pratica, specialmente nel campo della *stima di spettri*, perché, come faremo vedere tra breve, essa è l’*estensione positiva a massima entropia* dei dati (4.22). Per apprezzare il significato di questa affermazione conviene riferirsi al seguente problema, che in un certo senso idealizza quello della stima di spettri.

Sia $\{\mathbf{y}(t)\}$ un processo stazionario m -dimensionale, di media zero e distribuzione di probabilità incognita. L’*unica* informazione disponibile sulla legge di probabilità del processo sia costituita dalle $n + 1$ matrici di covarianza (4.22).

Si chiede di determinare una distribuzione di probabilità per il processo $\{\mathbf{y}(t)\}$ (e, nella fattispecie, il relativo spettro di potenza) che sia compatibile con i dati noti (4.22) ma nello stesso tempo non richieda assolutamente alcuna altra ipotesi addizionale (non verificabile) sul processo.

Si vuole, in altre parole, trovare una distribuzione di probabilità (stazionaria e a media zero) per $\{\mathbf{y}(t)\}$ determinata *unicamente* in base alla conoscenza dei momenti (4.22) e corrispondente, quindi, alla massima incertezza possibile su tutti gli altri parametri che possano individuarla.

Matematicamente questo problema viene formulato come un *problema di massimizzazione dell’entropia della legge di probabilità incognita* con vincoli corrispondenti alla conoscenza delle $n + 1$ covarianze (4.22).

Per una definizione precisa dell'entropia (della legge di probabilità) di un processo a tempo discreto rimandiamo il lettore (ad esempio) al testo [25, pag. 314-326] oppure a [35, pag. 63], dove sono raccolte anche le formule cui faremo riferimento in seguito.

Per un noto teorema di Shannon [128], la classe delle distribuzioni stazionarie con momenti del second'ordine finiti per cui si raggiunge il massimo dell'entropia è quella delle *distribuzioni Gaussiane*. Più precisamente, se l'informazione disponibile a priori è costituita solo dall'assegnazione di alcuni momenti del second'ordine (come nella (4.22)), la ricerca della distribuzione a massima entropia può essere ristretta alla classe delle distribuzioni gaussiane (naturalmente stazionarie e a media zero).

Ciascuna distribuzione in questa classe sarà allora caratterizzata univocamente assegnando una funzione di covarianza $R(t)$, $m \times m$, soddisfacente ai vincoli

$$R(t) = \Sigma(t) \quad , \quad t = 0, \pm 1, \dots, \pm n \quad . \quad (4.23)$$

Ricordiamo che l'entropia di una distribuzione Gaussianiana stazionaria di covarianza $R(t)$ ha la classica espressione

$$H(R) = \frac{1}{2} \log \det \Lambda + \frac{m}{2} \log 2\pi e \quad , \quad (4.24)$$

dove Λ è la *matrice varianza dell'errore di predizione di un passo* (in base a tutta la storia passata)

$$\mathbf{e}(t) = \mathbf{y}(t) - \hat{E}[\mathbf{y}(t) | H_{t-1}(\mathbf{y})]$$

relativo ad un qualunque processo stazionario $\{\mathbf{y}(t)\}$, di covarianza $R(t)$. Ora, dato che si ha in vista un problema di massimizzazione di $H(R)$, sarà evidentemente equivalente restringere la nostra attenzione a quelle funzioni $R(t)$ per cui $\det \Lambda > 0$, che caratterizzano processi p.n.d. e di rango pieno. In questo caso la (4.24) può essere espressa dalla celebre formula di Kolmogorov (si veda Rozanov [126, pag. 75])

$$H(R) = \frac{1}{2} \int_{-\pi}^{\pi} \log \det \Phi(e^{i\omega}) \frac{d\omega}{2\pi} + \frac{m}{2} \log 2\pi e \quad ,$$

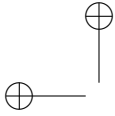
in cui appare la densità spettrale $\Phi(z)$ corrispondente alla funzione di covarianza $R(t)$.

Possiamo adesso enunciare la proprietà fondamentale dei filtri autoregressivi di Levinson $W_n(z)$.

TEOREMA 8.4. *La densità spettrale $S_n(z) = W_n(z) \Lambda_n W_n(z^{-1})'$ è l'interpolante delle $n+1$ covarianze assegnate $\Sigma(\tau)$, $\tau = 0, 1, \dots, n$, che ha massima entropia. Equivalentemente, l'uscita $\{\mathbf{y}_n(t)\}$ del filtro di Levinson $W_n(z)$ eccitato da rumore bianco di covarianza Λ_n è il processo stazionario che ha la più grande varianza d'errore di predizione di un passo nella classe dei processi stazionari la cui funzione di covarianza assume, per $\tau = 0, \pm 1, \dots, \pm n$, i valori assegnati (4.22).*

(In questo enunciato l'errore di predizione va inteso relativamente al passato infinito $\{\mathbf{y}_n(s) ; -\infty < s < t$ e "più grande" significa che $\det [\text{Var } \hat{\mathbf{y}}_n(t | t-1)]$ è il massimo possibile, compatibilmente con le condizioni (4.23)).

DIMOSTRAZIONE. Daremo la dimostrazione solo nel caso scalare ($m = 1$). Si vede in questo caso dalla formula (4.24) che la massimizzazione di $H(r)$ si consegue se e solo se si massimizza rispetto ad $r : t \rightarrow r(t)$ la varianza, che qui denoteremo col



simbolo q , dell'errore di predizione associata ad (un processo $\{\mathbf{y}(t)\}$ di covarianza) r . Quest'ultima deve soddisfare alle condizioni

$$r(h) = \sigma(h) \quad , \quad h = 0, 1, \dots, n \quad , \quad (4.25)$$

dove le $\{\sigma(h)\}$ sono assegnate a priori.

Sia allora

$$\mathbf{y}(t) = \sum_{-\infty}^t \ell(t-s) \boldsymbol{\varepsilon}(s) \quad , \quad \text{var } \boldsymbol{\varepsilon}(t) = 1 \quad , \quad (4.26)$$

la rappresentazione d'innovazione normalizzata, corrispondente alla funzione di covarianza $r(t)$. Si avrà, per definizione,

$$r(t) = \sum_0^{\infty} \ell(k) \ell(t+k) \quad , \quad t \geq 0 \quad , \quad (4.27)$$

e siccome (vedi il Capitolo 4)

$$q = \text{var } \mathbf{e}(t) = \ell^2(0) \quad , \quad (4.28)$$

si vede che conviene reimpostare il problema massimizzando rispetto ad $\ell := \{\ell(t) ; t \geq 0\}$ anziché rispetto alla covarianza.

Sia ℓ^{2+} lo spazio di Hilbert delle successioni $x = \{x(t) ; t \geq 0\}$ a quadrato sommabile. Questo spazio è dotato del prodotto interno

$$\langle x, y \rangle := \sum_0^{\infty} x(t) y(t) \quad .$$

Sappiamo che la risposta impulsiva $\ell = \{\ell(t)\}$ del fattore spettrale canonico sta in ℓ^{2+} . Definiamo allora l'operatore lineare

$$\delta_0 : \ell^{2+} \rightarrow \ell^{2+} \quad , \quad \delta_0(x) := \{x(0), 0, 0, \dots\} \quad .$$

Il nostro problema può così essere riformulato nel modo seguente.

Trovare l'elemento ℓ di ℓ^{2+} che massimizza la norma $\|\delta_0(\ell)\|^2 = \ell^2(0)$ (= var $\mathbf{e}(t)$) ed è soggetto alle condizioni (4.25), ovvero alle

$$\langle \ell, T_h \ell \rangle = \sigma(h) \quad , \quad h = 0, 1, \dots, n \quad , \quad (4.29)$$

dove T_h è l'operatore di traslazione a sinistra in ℓ^{2+} definito dalla

$$(T_h x)(k) = x(h+k) \quad , \quad k \geq 0 \quad .$$

Chiaramente T_h trasla a sinistra di h unità temporali la sequenza x e scarta i termini con indice negativo.

Introduciamo $n+1$ moltiplicatori di Lagrange, che denotiamo col simbolo $\lambda(0), \lambda(1), \dots, \lambda(n)$ e definiamo la funzione di Lagrange

$$L(\ell) := \|\delta_0(\ell)\|^2 + \sum_0^n \lambda(h) [\sigma(h) - \langle \ell, T_h \ell \rangle] \quad .$$

Ad una variazione, $d\ell$, di ℓ in ℓ^{2+} corrisponde una variazione del funzionale “aumentato” $L(\ell)$ pari a:

$$dL(\ell) = 2\langle \delta_0(\ell), d\ell \rangle - \sum_0^n \lambda(h) [\langle d\ell, T_h \ell \rangle + \langle \ell, T_h d\ell \rangle] \quad (4.30)$$

$$= 2\langle \delta_0(\ell), d\ell \rangle - \sum_0^n \lambda(h) [\langle T_h \ell, d\ell \rangle + \langle T_h^* \ell, d\ell \rangle] \quad (4.31)$$

dato che le $\{\sigma(h)\}$ non dipendono da ℓ . Notiamo ora che se $h \geq 0$ si ha

$$\langle x, T_h y \rangle = \sum_0^\infty x(t) y(t+h) = \sum_h^\infty x(t-h) y(t) = \langle T_{-h} x, y \rangle \quad , \quad t \geq 0 \quad .$$

Quindi $\langle T_h^* \ell, d\ell \rangle$ si può anche scrivere come $\langle T_{-h} \ell, d\ell \rangle$.

Imponendo ora che la variazione dL nella (4.31) sia zero per tutte le possibili variazioni $d\ell$ in ℓ^{2+} , si ottiene:

$$2\delta_0(\ell) - \sum_0^n \lambda(h) [T_h \ell + T_{-h} \ell] = 0 \quad ,$$

la sequenza nulla in ℓ^{2+} . Si riscrive questa relazione in modo compatto nella forma

$$2\delta_0(\ell) - \sum_{-n}^n b(h) T_h \ell = 0$$

dove $b(h) := \lambda(-h)$ per $h < 0$, $b(0) = 2\lambda(0)$, e $b(h) := \lambda(h)$ per $h > 0$. Calcolando il prodotto scalare del primo membro di questa espressione con $T_s \ell$ e ricordando che $\langle T_h \ell, T_s \ell \rangle = r(h-s)$, si trova

$$2\langle \delta_0(\ell), T_s \ell \rangle - \sum_{h=-n}^n b(h) r(h-s) = 0 \quad . \quad (4.32)$$

Il primo termine di questa relazione è particolarmente semplice da calcolare se $s \leq 0$. In questo caso infatti $\langle \delta_0(\ell), T_s \ell \rangle = \ell(0)^2 \delta(s)$ e pertanto la (4.32) si può riscrivere in forma esplicita come

$$2\ell(0)^2 \delta(s) = \sum_{h=-n}^n b(h) r(h-s) \quad s \leq 0 \quad . \quad (4.33)$$

In realtà la somma a secondo membro è una funzione pari di s perché, per $s > 0$, ponendo $\tau := -s \leq 0$, essa è uguale a

$$\begin{aligned} \sum_{h=-n}^n b(h) r(s-h) &= \sum_{h=-n}^n b(h) r(-\tau-h) = \\ \sum_{h=-n}^n b(-k) r(-\tau+k) &= \sum_{h=-n}^n b(k) r(k-\tau) \end{aligned}$$

e pertanto la (4.33) vale per ogni s . Passando alle Z-trasformate si vede che lo spettro dell'interpolante a massima entropia $S_n(z)$ (trasformata Z di r), dev'essere lo spettro di un modello AR di ordine n

$$S_n(z) = \frac{2\ell(0)^2}{B_n(z)}$$

dove il polinomio $B_n(z) := \sum_{-n}^n b(h) z^h$ dovrà evidentemente assumere valori positivi sulla circonferenza unitaria e ammettere quindi una fattorizzazione

$$B_n(z) = c^2 A^*(z^{-1}) A^*(z)$$

in cui $A^*(z^{-1}) := \sum_{h=0}^n a(h) z^{-h}$ è un polinomio monico a zeri in $\{|z| < 1\}$.

In altri termini, la (4.33) mostra come la soluzione estrema ℓ debba necessariamente soddisfare l'equazione alle differenze

$$\ell(t) + \sum_1^n a(h) \ell(t-h) = c \delta(t) \quad , \quad (4.34)$$

dove $c = \ell(0)$ è una costante che, al pari dei coefficienti $\{a(h)\}$ (funzioni dei moltiplicatori di Lagrange $\{\lambda(h)\}$), andrà determinata in base ai vincoli (4.25).

La (4.34) esprime il fatto che $\ell = \{\ell(t) ; t \geq 0\}$ è la risposta impulsiva del sistema puramente autoregressivo di funzione di trasferimento

$$L(z) = c \left[1 + \sum_1^n a(h) z^{-h} \right]^{-1} .$$

Ne segue che la funzione di covarianza r del processo $\{y(t)\}$ definito nella (4.26) è necessariamente legata ai coefficienti c e $\{a(h)\}$ dalle relazioni

$$\begin{aligned} r(0) + \sum_1^n a(h) r(-h) &= c^2 \\ r(s) + \sum_1^n a(h) r(s-h) &= 0 \quad , \quad s > 0 \quad . \end{aligned}$$

Scrivendo la seconda equazione per $s = 1, \dots, n$ e tenendo conto della condizione $r(t) = \sigma(t)$ per $|t| \leq n$, si trova l'esatto analogo dell'equazione (3.2) che determina i coefficienti $\{a_n(k)\}$ del filtro di Levinson $W_n(z)$. Dato che la matrice T_n che compare in questa equazione è stata assunta non singolare, ne segue che $a_n(h) = a(h)$, $h = 1, \dots, n$.

A questo punto è facile concludere che $c^2 = \lambda_n$. □

L'enunciazione di questo risultato sembra sia dovuta a J.P. Burg. Delle dimostrazioni alternative (non completamente convincenti) si possono trovare negli articoli di Van Der Bos [139] ed Edward e Fitelson [43].

Filtri a traliccio. Concluderemo questo paragrafo accennando ad una realizzazione dei filtri di Levinson di notevole importanza nelle applicazioni.

Siano $A_n^*(z^{-1})$ e $B_n^*(z^{-1})$ le funzioni di trasferimento dei filtri sbiancanti a memoria finita, n . In notazione vettoriale,

$$\begin{bmatrix} \mathbf{e}_n(t) \\ \hat{\mathbf{e}}_n(t) \end{bmatrix} = \begin{bmatrix} A_n^*(z^{-1}) \\ B_n^*(z^{-1}) \end{bmatrix} \mathbf{y}(t)$$

Usando le equazioni di aggiornamento dei coefficienti A_n e B_n dell'algoritmo di Levinson (riportate nel Teorema 8.2)) è facile ricavare dalle formule ricorsive di tipo analogo direttamente per i polinomi $A_n^*(z^{-1})$ e $B_n^*(z^{-1})$. Queste formule hanno l'aspetto seguente:

$$\begin{bmatrix} A_{n+1}^*(z^{-1}) \\ B_{n+1}^*(z^{-1}) \end{bmatrix} = \begin{bmatrix} I & -z^{-1} \bar{\Gamma}_n \\ -\Gamma_n & z^{-1} I \end{bmatrix} \begin{bmatrix} A_n^*(z^{-1}) \\ B_n^*(z^{-1}) \end{bmatrix} \quad , \quad \begin{matrix} A_0^*(z^{-1}) = I \\ B_0^*(z^{-1}) = I \end{matrix} \quad . \quad (4.35)$$

Come si vede, moltiplicando a destra queste relazioni per $\mathbf{y}(t)$ si riottengono le relazioni (2.9) di aggiornamento delle innovazioni a memoria finita, scritte in forma simbolica. Ora, la (4.35) esprime la *funzione di trasferimento dei filtri a memoria $n + 1$ come cascata della funzione di trasferimento dei filtri a memoria n e del blocco di struttura "a traliccio" tratteggiato nella figura seguente.*

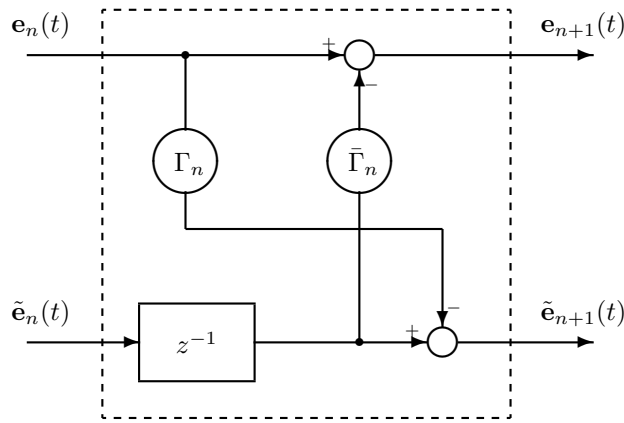


FIGURA 4.1. Blocco a traliccio.

Chiaramente, iterando il ragionamento, il filtro sbiancante può essere realizzato mediante la cascata di esattamente n blocchi a traliccio, ciascuno avente la struttura tratteggiata in figura 4.1. Notiamo che ciascun blocco è individuato assegnando i coefficienti di riflessione $\Gamma_n, \bar{\Gamma}_n$ al passo n -esimo.

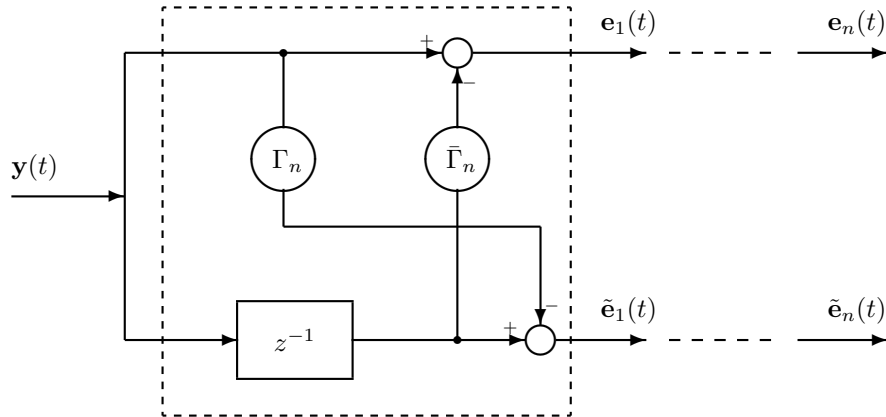
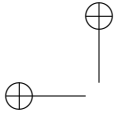


FIGURA 4.2. Struttura a traliccio.

Questi schemi a traliccio trovano applicazione in molte implementazioni circuitali di algoritmi di elaborazione del segnale. Uno particolarmente importante è l'analisi, la codifica e la ricostruzione in linea del segnale vocale. Il motivo è che in questi schemi il parametro di complessità n , può, in un certo senso, essere scelto



dall' utente, senza bisogno di modificare la struttura del filtro ma solo aggiungendo o escludendo un certo numero di sezioni elementari a traliccio della cascata. Si può anche mostrare [124] che in generale queste realizzazioni sono meno sensibili agli errori di quantizzazione e di arrotondamento sui coefficienti (Γ_n e $\bar{\Gamma}_n$) di quanto lo siano gli schemi di realizzazione classici del filtro (si veda ad es. la sezione 2 del prossimo capitolo) in cui si usano direttamente i coefficienti A_n e B_n . Questi ultimi non sono poi evidentemente *aggiornabili con n* dato che passando da memoria n ad $n + 1$ cambiano *tutti* i coefficienti del filtro.

Predizione e Filtraggio con i filtri a traliccio. Se si desidera costruire un *predittore a memoria n* basta derivare il segnale $\mathbf{e}_n(t)$ all'uscita ed eseguire l'operazione $\mathbf{y}(t) - \mathbf{e}_n(t)$ come nello schema di figura 4.3.

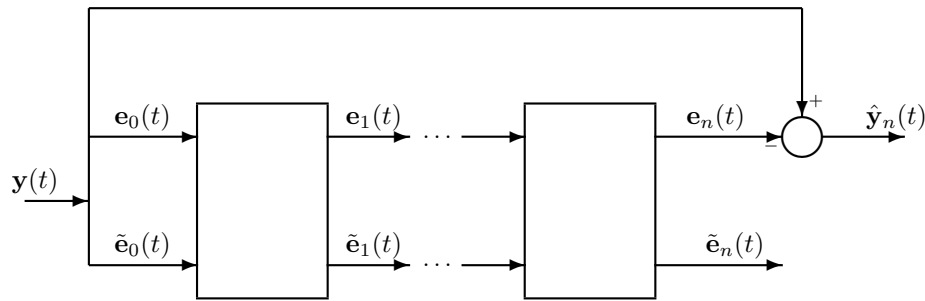


FIGURA 4.3. Predittore a memoria n .

Supponiamo invece di voler calcolare la stima filtrata a memoria n di un processo $\{\mathbf{x}(t)\}$ congiuntamente stazionario con $\{\mathbf{y}(t)\}$. Si può in questo caso procedere come segue.

Il filtro a memoria finita si può esprimere in funzione dell'innovazione all'indietro

$$\hat{E}[\mathbf{x}(t) | \mathbf{y}(t), \dots, \mathbf{y}(t - n)] = \hat{E}[\mathbf{x}(t) | \bar{\mathbf{e}}_0(t), \dots, \bar{\mathbf{e}}_n(t - n)] := \sum_{k=0}^n C_n(k) \bar{\mathbf{e}}_k(t), \tag{4.36}$$

dove i coefficienti $C_n(k)$ si possono calcolare usando l'espressione del filtro sbiancante (2.22),

$$C_n(k) = E[\mathbf{x}(t)\bar{\mathbf{e}}_k(t)] \bar{\Lambda}_k^{-1} = \sum_{i=0}^k E[\mathbf{x}(t)\mathbf{y}(t - i)] B_k(i)' \bar{\Lambda}_k^{-1} = \sum_{i=0}^k \Sigma_{\mathbf{xy}}(i) B_k(i)' \bar{\Lambda}_k^{-1}$$

e questa, definendo la matrice

$$\Sigma_k := [\Sigma_{\mathbf{xy}}(0) \dots \Sigma_{\mathbf{xy}}(n)]$$

si può riscrivere

$$C_n(k)\bar{\Lambda}_k = \Sigma_k B_k'. \tag{4.37}$$

Questa formula mostra ancora il fatto notevole che *anche i coefficienti $C_n(k)$ del filtro non dipendono dalla memoria n* . Quest'ultima può quindi essere variata come parametro di progetto nella realizzazione del filtro semplicemente aggiungendo o

escludendo sezioni elementari a traliccio, esattamente come succedeva per il filtro sbiancante. Un possibile schema realizzativo del filtro è rappresentato nella figura 4.4.

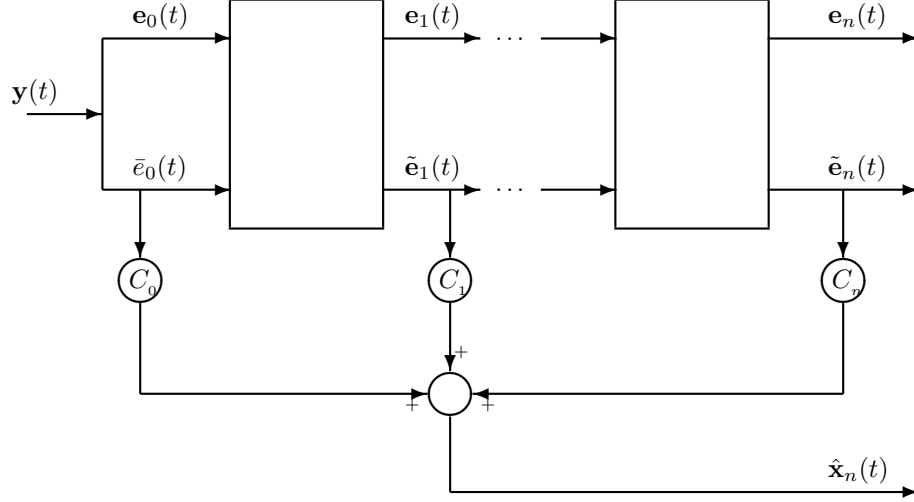


FIGURA 4.4. Stima filtrata a memoria n .

PROBLEMA 8.3. *Risolvere il problema del filtraggio a memoria n usando le innovazioni a memoria finita in avanti, $\{\mathbf{e}_k(t)\}$ e disegnare il relativo schema a blocchi.* \diamond

Esistono in letteratura numerose varianti della struttura di base di figura 4.4. Si vedano ad esempio il lavoro già citato di Benveniste e Chaure [17], i lavori di Friedlander [50] e la monografia [136].

Realizzazioni a traliccio di funzioni razionali. Come abbiamo visto i filtri sbiancanti a memoria finita, rappresentati come una cascata di sezioni a traliccio del tipo di figura 4.2 hanno notevoli proprietà “circuitali” (e in realtà anche notevoli proprietà numeriche [124]). Ci si può chiedere se la realizzabilità mediante schemi a traliccio sia una caratteristica speciale dei filtri sbiancanti di Levinson o se valga, più in generale. Mostriamo innanzitutto che la struttura a traliccio per il filtro sbiancante a memoria n si trasporta anche al filtro formante corrispondente. Per far questo dobbiamo sostanzialmente *invertire lo schema di figura 4.2*, o, equivalentemente, invertire le relazioni (4.35).

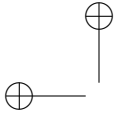
Allo scopo, possiamo equivalentemente partire dalle relazioni “astratte” (2.9) e ricavarne, dopo alcuni semplici passaggi le relazioni

$$\bar{\Delta}_n \mathbf{e}_n(t) = \mathbf{e}_{n+1}(t) + \bar{\Gamma}_n \tilde{\mathbf{e}}_{n+1}(t), \quad (4.38)$$

$$\Delta_n \tilde{\mathbf{e}}_n(t-1) = \Gamma_n \mathbf{e}_{n+1}(t) + \tilde{\mathbf{e}}_{n+1}(t) \quad (4.39)$$

dove, le matrici

$$\Delta_n = I - \Gamma_n \bar{\Gamma}_n, \quad \bar{\Delta}_n = I - \bar{\Gamma}_n \Gamma_n \quad (4.40)$$



sono per ipotesi invertibili. Applicando alla (4.39) l'operatore di anticipo di un passo e sostituendo le espressioni delle innovazioni a memoria finita,

$$\mathbf{e}_n(t) = A_n^*(z^{-1})\mathbf{y}(t), \quad \tilde{\mathbf{e}}_n(t) = B_n^*(z^{-1})\mathbf{y}(t)$$

si perviene così alle relazioni:

$$\begin{bmatrix} \bar{\Delta}_n A_n^*(z^{-1}) \\ \Delta_n B_n^*(z^{-1}) \end{bmatrix} = \begin{bmatrix} I & \bar{\Gamma}_n \\ z\Gamma_n & zI \end{bmatrix} \begin{bmatrix} A_{n+1}^*(z^{-1}) \\ B_{n+1}^*(z^{-1}) \end{bmatrix} \quad (4.41)$$

che vanno iterate per n decrescenti fino ad $n = 0$. Supponiamo allora di avere a disposizione, per un certo $N > 0$ fissato, i processi $\mathbf{e}_N, \tilde{\mathbf{e}}_N$, uscite del filtro sbiancante a memoria N e del suo duale diciamoli $A_N^*(z^{-1}), B_N^*(z^{-1})$. Iterando la (4.41) possiamo realizzare il filtro formante che trasforma gli ingressi $\mathbf{e}_N, \tilde{\mathbf{e}}_N$ nell'uscita $\mathbf{y}(t)$ (più precisamente nei processi $\mathbf{e}_0(t)$ e $\tilde{\mathbf{e}}_0(t)$, entrambi coincidenti con $\mathbf{y}(t)$), come cascata di sezioni a traliccio del tipo di figura 4.5,

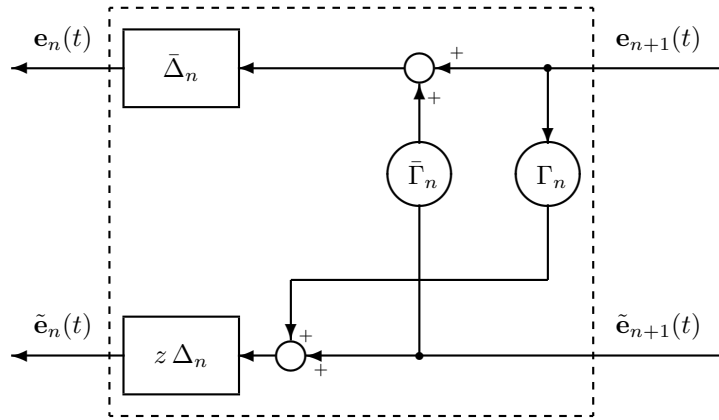


FIGURA 4.5. Blocco a traliccio per il filtro formante

Consideriamo adesso una qualunque funzione di trasferimento matriciale di tipo puramente autoregressivo (AR),

$$W(z) = \left[I + \sum_{k=1}^N A_k z^{-k} \right]^{-1} B_0 \quad A_k \in \mathbb{R}^{m \times m} \quad (4.42)$$

dove supponiamo che B_0 abbia rango m .

Mostriamo che ogni $W(z)$ di tipo AR che è stabile (non ha poli in $|z| \geq 1$) è un filtro formante di Levinson e può essere rappresentata come una cascata di blocchi a traliccio.

Allo scopo si può, ragionare come segue.

- (1) Utilizziamo la funzione di trasferimento (4.42) come filtro formante e generiamo un processo stazionario m -dimensionale \mathbf{y} partendo da un rumore bianco normalizzato \mathbf{w} (di dimensione pari al numero di colonne di B_0), i.e.

$$\mathbf{y}(t) + \sum_{k=1}^N A_k \mathbf{y}(t-k) = B_0 \mathbf{w}(t). \quad (4.43)$$

Dato che B_0 ha rango m il processo bianco $\mathbf{e}_N(t) := B_0 \mathbf{w}(t)$ ha matrice varianza $\Lambda_N := B_0 B_0'$ che è definita positiva.

- (2) Per ipotesi il polinomio $\Delta(z) := \det \left[I z^N + \sum_{k=1}^N A_k z^{N-k} \right]$ ha tutti i suoi zeri dentro la circonferenza unità e quindi \mathbf{e}_N è l'innovazione di \mathbf{y} . Di fatto, \mathbf{e}_N è anche l'innovazione a memoria N (vedere gli esercizi (8.3) e (8.4) alla fine del capitolo). Quindi il polinomio matriciale $I + \sum_{k=1}^N A_k z^{-k}$ è un filtro sbiancante a memoria N e, come ben sappiamo, può quindi essere realizzato con una struttura a traliccio.
- (3) Segue che la funzione di trasferimento $W(z)$ e il filtro formante di Levinson di ordine N del processo \mathbf{y} , opportunamente normalizzato: $W_N(z) \Lambda_N^{1/2}$, differiscono per la moltiplicazione a destra per una matrice ortogonale costante². Quindi, per quanto visto sopra, $W(z)$ è realizzabile come una cascata di blocchi a traliccio.

I coefficienti $\Gamma_k, \bar{\Gamma}_k, k = 1, \dots, N$ che definiscono la struttura a traliccio hanno quindi il significato statistico di coefficienti di riflessione (in particolare sono numeri a modulo strettamente minore di uno nel caso scalare) e si potrebbero in linea di principio, calcolare applicando l'algoritmo di Levinson "concreto" alla successione delle covarianze $\Sigma(0), \dots, \Sigma(N-1)$ del processo \mathbf{y} . Quest'ultima si può ottenere risolvendo l'equazione di Lyapunov

$$\Sigma = F \Sigma F' + \Lambda$$

dove F è la matrice compagna a blocchi

$$F = \begin{bmatrix} 0 & I & 0 \\ & & I \\ -A_N & \dots & -A_1 \end{bmatrix}.$$

e Λ è la matrice diagonale a blocchi data dalla formula (4.8). Come visto nella dimostrazione del Teorema 8.3, c'è un'unica soluzione Σ_N , definita positiva, di questa equazione, che è proprio la matrice di Toeplitz a blocchi delle varianze del processo \mathbf{y} . In formule,

$$\Sigma_N = E \left\{ \begin{bmatrix} \mathbf{y}(t-N+1) \\ \dots \\ \mathbf{y}(t) \end{bmatrix} \begin{bmatrix} \mathbf{y}(t-N+1) \\ \dots \\ \mathbf{y}(t) \end{bmatrix}' \right\}.$$

In realtà se fosse noto, oltre ad $A_N(z)$, anche il polinomio duale, $B_N(z)$, i coefficienti di riflessione $\Gamma_k, \bar{\Gamma}_k, k = N-1, \dots, 0$, si potrebbero ottenere direttamente dall'iterazione "all' indietro" (4.41) che possiamo inizializzare al passo N ponendo:

$$\begin{bmatrix} \bar{\Delta}_{N-1} A_{N-1}^*(z^{-1}) \\ \Delta_{N-1} B_{N-1}^*(z^{-1}) \end{bmatrix} = \begin{bmatrix} I & \bar{\Gamma}_{N-1} \\ z \Gamma_{N-1} & z I \end{bmatrix} \begin{bmatrix} A_N^*(z^{-1}) \\ B_N^*(z^{-1}) \end{bmatrix}. \quad (4.44)$$

Notiamo che se $A_N(z)$ e $B_N(z)$ sono dati, sono automaticamente noti anche i coefficienti di riflessione $\Gamma_{N-1}, \bar{\Gamma}_{N-1}$ che valgono (cfr. la (2.32))

$$\bar{\Gamma}_{N-1} = -A_N(N), \quad \Gamma_{N-1} = -B_N(0).$$

Di fatto questo modo di procedere non è completamente soddisfacente perchè richiede la conoscenza del polinomio duale $B_N(z)$, che va calcolato preliminarmente,

²In effetti, sia $W(z)$ che il filtro di Levinson di ordine N sono allora fattori spettrali a fase minima della densità spettrale razionale $S(z) = W(z)W(z^{-1})'$.

usando per esempio una fattorizzazione spettrale (duale) del trasposto dello spettro polinomiale $A^*(z^{-1})[A^*(z)]'$. Questa fattorizzazione spettrale si può ad esempio eseguire con l'algoritmo di Bauer-Rissanen del Capitolo 4. Si veda [8] per una rassegna di metodi alternativi.

L'iterazione (4.41) si può però impiegare direttamente nel caso scalare in cui si sa che $B_N(z)$ è semplicemente il polinomio reciproco di A (cf la (3.8)).

PROPOSIZIONE 8.3. *Ogni funzione razionale scalare $W(z) = b_0/A_N^*(z^{-1})$ con $A_N^*(z) = 1 + \sum_{k=1}^N a_N(k)z^{-k}$ polinomio (monico di grado N) a zeri strettamente dentro il cerchio unità, può essere realizzata mediante la cascata di N blocchi a traliccio del tipo di figura 4.5 dove la successione dei coefficienti di riflessione si calcola con il seguente algoritmo.*

For $n = N, N-1, \dots, 0$ do:

$$\begin{aligned} A_{n-1}^*(z^{-1}) &= \frac{1}{1 - \gamma_{n-1}^2} [A_n^*(z^{-1}) + \gamma_{n-1}A_n(z^{-1})] \\ \gamma_{n-1} &= -a_n(n) \quad n = N, N-1, \dots, 0. \end{aligned} \quad (4.45)$$

Notiamo che, una volta noti i coefficienti di riflessione, le varianze λ_n^2 si potrebbero calcolare separatamente iterando all'indietro la relazione (3.14) con condizione finale $\lambda_N^2 = b_0^2$. Esse comunque non intervengono nell'algoritmo.

Dalla proposizione 8.3 scende il celebre test di Schur-Cohn per verificare se un polinomio ha tutte le radici dentro il cerchio unità.

TEOREMA 8.5. *Condizione necessaria e sufficiente affinché le radici dell'equazione $A_N(z) = 0$ siano tutte interne alla circonferenza unità è che i $\{\gamma_n\}$ calcolati con l'algoritmo (4.45) abbiano tutti modulo strettamente minore di uno. Esiste in effetti una corrispondenza biunivoca tra il sottoinsieme di \mathbb{R}^N popolato dai coefficienti $\{a_n(1), a_n(2), \dots, a_n(N)\}$ dei polinomi (monici di grado N) a zeri strettamente dentro il cerchio unitario e l'ipercubo $(-1, 1)^N \subset \mathbb{R}^N$, popolato dai corrispondenti coefficienti di riflessione $\{\gamma_0, \gamma_1, \dots, \gamma_{N-1}\}$.*

DIMOSTRAZIONE. La necessità della condizione scende dalla discussione che abbiamo condotto fin qui; in particolare se $W(z)$ è (strettamente) stabile vale l'interpretazione statistica dei $\{\gamma_n\}$ come coefficienti di riflessione, che debbono essere a modulo strettamente minore di uno perchè il processo generato col filtro formante $W(z)$ è p.n.d..

La sufficienza si può provare riscrivendo la prima delle relazioni (4.35) dell'algoritmo di Levinson per il caso scalare, nella forma

$$a_{n+1}^*(z^{-1}) = a_n^*(z^{-1}) - z^{-1}\gamma_n a_n(z^{-1})$$

ovvero, moltiplicando per z^{n+1} ,

$$a_{n+1}(z) = za_n(z) - \gamma_n a_n^*(z), \quad n = 0, 1, \dots \quad (*)$$

Ricordiamo che tutti i polinomi "parziali" $a_n(z)$ sono monici. È ovvio che se $|\gamma_0| < 1$ il polinomio $a_1(z)$ ha il suo unico zero dentro il cerchio. Possiamo così ragionare per induzione. Assumiamo che $a_n(z)$ sia stabile e riscriviamo la relazione (*) nella forma

$$a_{n+1}(z) = a_n(z) \left[z - \gamma_n \frac{a_n^*(z)}{a_n(z)} \right],$$

da cui è evidente che gli zeri di $a_{n+1}(z)$ sono quelli di $a_n(z)$ e il valore di z che annulla l'espressione tra parentesi quadre, che deve esistere e essere unico, in quanto $n+1$ -simo zero di $a_{n+1}(z)$. Per l'ipotesi induttiva, $\frac{a_n^*(z)}{a_n(z)}$ è una funzione analitica fuori dal cerchio (incluso il punto all'infinito) e ha modulo uguale a uno sulla circonferenza unità. Dato che $|\gamma_n| < 1$, per il principio del massimo modulo si deve avere

$$\left| \gamma_n \frac{a_n^*(z)}{a_n(z)} \right| < 1, \quad |z| \geq 1$$

da cui segue che l'uguaglianza $z = \gamma_n \frac{a_n^*(z)}{a_n(z)}$ non può verificarsi per $|z| \geq 1$ e deve quindi necessariamente verificarsi per un valore di z interno al cerchio unità.

Lasceremo al lettore la dimostrazione dell'ultima affermazione. \square

Per l'applicazione del test occorre notare che il calcolo ricorsivo dei coefficienti $\{\gamma_n\}$ di una funzione di trasferimento scalare di tipo AR può essere portato a termine a partire da un polinomio a denominatore affatto generale (non necessariamente a zeri dentro il cerchio), purchè qualche γ_n non risulti uguale ad uno nel qual caso l'algoritmo si interrompe. In questo caso però c'è necessariamente qualche zero a modulo unitario. Vedere a questo proposito l'esercizio 8.1 alla fine del capitolo.

Una discussione molto dettagliata dell'algoritmo di Levinson nel caso scalare è contenuta nel capitolo 11.3 del libro di Roberts e Mullis [124].

5. Fattorizzazione veloce di matrici di Toeplitz

Un modo di interpretare l'algoritmo di Levinson, che ha di per sé poco a che vedere con la statistica e la teoria della probabilità, è di considerarlo semplicemente come uno strumento per risolvere sistemi di equazioni lineari del tipo

$$Tx = b, \quad (5.1)$$

dove T è una *matrice di Toeplitz*. Per esempio, l'equazione (3.3) che determina i coefficienti del predittore a memoria n è (modulo una trasposizione) del tipo (5.1).

Ora, se si esegue un conteggio del numero di operazioni elementari (moltiplicazioni e somme) necessario per arrivare alla soluzione a_n attraverso n applicazioni successive dell'algoritmo di Levinson, si trova che questo numero è all'incirca proporzionale ad n^2 . Usando invece uno degli algoritmi standard (ad esempio l'eliminazione di Gauss con riordino oppure l'algoritmo di Cholesky descritto nel Capitolo 2) si trova che questo numero è viceversa proporzionale ad n^3 . L'algoritmo di Levinson fornisce quindi la soluzione impiegando un ordine di grandezza di operazioni in meno. Chiaramente questo fatto è una conseguenza della speciale struttura della matrice T , che viene appunto sfruttata in modo intelligente nell'algoritmo.

Siccome moltissimi problemi di stima e identificazione si riducono in ultima analisi alla soluzione di equazioni lineari del tipo (5.1) (o successioni di equazioni di questo tipo) in cui T è di Toeplitz (o di Toeplitz a blocchi), diventa di estremo interesse generalizzare l'algoritmo di Levinson in modo da poter trattare situazioni in cui il secondo membro (b) nell'equazione (5.1) non è necessariamente della forma standard che si ritrova nel problema di predizione studiato al paragrafo precedente.

Una possibile strada per raggiungere lo scopo potrebbe essere quella di costruire una fattorizzazione triangolare (o di Cholesky) di T del tipo $T = LL'$, con L triangolare inferiore. Ovviamente se si assume T simmetrica e definita positiva

(e questo sarà sempre il caso nel seguito) L si può calcolare con l'algoritmo di fattorizzazione di Cholesky, con un costo computazionale proporzionale al cubo della dimensione di T . Vedremo che si può fare di meglio.

Facciamo riferimento a *matrici di Toeplitz a blocchi* ($m \times m$) che siano *simmetriche* (ovvero il blocco $(m \times m)$ T_{ij} è uguale al trasposto T'_{ji} del blocco di posto ji) e *definite positive*. Una matrice di questo tipo può sempre essere pensata come la covarianza di $n + 1$ campioni successivi estratti da un processo stocastico m -dimensionale, stazionario, p.n.d. e di rango pieno. In formule,

$$T := \begin{bmatrix} T_0 & T_{-1} & & T_{-n} \\ T_1 & \ddots & \ddots & \\ & \ddots & \ddots & T_{-1} \\ T_n & & T_1 & T_0 \end{bmatrix} = E \left\{ \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}(t+1) \\ \vdots \\ \mathbf{y}(t+n) \end{bmatrix} [\mathbf{y}'(t), \dots, \mathbf{y}'(t+n)] \right\} \quad (5.2)$$

col che, ovviamente, $T_k \equiv \Sigma(k) = E \mathbf{y}(t+k) \mathbf{y}'(t)$ per $k = 0, \pm 1, \dots, \pm n$ e quindi $T_{-k} = T'_k$.

Prendiamo in considerazione il processo delle innovazioni in avanti $\{\mathbf{e}_k(t)\}$ a memoria k , associato al processo stazionario $\{\mathbf{y}(t)\}$. Ricordiamo (lemma 8.1) che $\mathbf{e}_k(t)$ è costruito attraverso un procedimento di ortogonalizzazione in avanti delle variabili $\mathbf{y}(t-k), \mathbf{y}(t-k+1), \dots, \mathbf{y}(t)$, e che i vettori aleatori $\mathbf{e}_0(t), \mathbf{e}_1(t+1), \dots, \mathbf{e}_k(t+k), \dots$ sono quindi mutuamente ortogonali, ovvero

$$E \mathbf{e}_k(t+k) \mathbf{e}'_j(t+j) = \Lambda_k \delta(k-j) \quad (5.3)$$

e, inoltre, tali che $\text{span}\{\mathbf{e}_0(t), \mathbf{e}_1(t+1), \dots, \mathbf{e}_k(t+k)\} = \text{span}\{\mathbf{y}(t), \mathbf{y}(t+1), \dots, \mathbf{y}(t+k)\}$ per ogni $k \geq 0$.

Usando ora la rappresentazione (2.16) per $\mathbf{e}_k(t+k)$ si possono scrivere le relazioni

$$\begin{bmatrix} \mathbf{e}_0(t) \\ \mathbf{e}_1(t+1) \\ \vdots \\ \mathbf{e}_n(t+n) \end{bmatrix} = \begin{bmatrix} I & & & \\ A_1(1) & I & & \\ \dots & \dots & \dots & \\ A_n(n) & \dots & A_n(1) & I \end{bmatrix} \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}(t+1) \\ \vdots \\ \mathbf{y}(t-n) \end{bmatrix}, \quad (5.4)$$

dove la matrice, diciamola \tilde{A} , che compare a secondo membro è *triangolare inferiore con tutti uni sulla diagonale*. Calcolando le varianze a primo e secondo membro e tenendo conto dell'ortogonalità dei vettori $\{\mathbf{e}_k(t+k)\}$ dichiarata nel lemma 8.1, si trova immediatamente la relazione

$$\begin{bmatrix} \Lambda_0 & & \\ & \ddots & \\ & & \Lambda_n \end{bmatrix} = \tilde{A} T \tilde{A}' \quad (5.5)$$

la quale si può leggere dicendo che: *la trasformazione triangolare inferiore \tilde{A} , costruita con i coefficienti (matriciali) A_k come mostrato nella formula (5.4), diagonalizza (a blocchi) la matrice di Toeplitz T* .

Notiamo che se avessimo normalizzato a varianza unitaria le innovazioni $\{\mathbf{e}_k(t)\}$, a primo membro nella (5.5) ci sarebbe proprio l'identità. Evidentemente, la (5.5)

fornisce esplicitamente l'inversa di T in forma fattorizzata

$$T^{-1} = \tilde{A}' \begin{bmatrix} \Lambda_0^{-1} & & \\ & \ddots & \\ & & \Lambda_n^{-1} \end{bmatrix} \tilde{A} \quad , \quad (5.6)$$

in cui ora A' è triangolare superiormente con diagonale unitaria.

La conclusione di questo semplice calcolo è dunque la seguente: *senza alcun aggravio ulteriore dei calcoli, l'algoritmo di Levinson fornisce direttamente l'inversa della matrice T in forma fattorizzata.* Notiamo infatti che il calcolo delle inverse, Λ_k^{-1} , delle varianze d'innovazione deve comunque essere fatto per aggiornare i coefficienti A_k e B_k (vedi l'enunciato del Teorema 8.2).

A partire dalla (5.6) si possono anche dare delle formule ricorsive che calcolano il blocco $(T^{-1})_{i+1,j+1}$ in funzione di $(T^{-1})_{ij}$ e di certi elementi di A_i e B_j . Queste formule si possono derivare agevolmente usando l'espressione ricorsiva (2.28) per l'aggiornamento dei coefficienti A_k e B_k . Esse non sembrano essere di grande utilità e pertanto non le riporteremo qui. Il loro prototipo fu scoperto (per il caso scalare) da W. Trench [138] nel 1964.

È evidente che usando le innovazioni *all'indietro* $\{\tilde{\mathbf{e}}_0(t), \dots, \tilde{\mathbf{e}}_n(t)\}$, si può arrivare a formule simili alla (5.6). Scrivendo l'analogia delle (5.4) si trova infatti

$$\begin{bmatrix} \tilde{\mathbf{e}}_0(t) \\ \tilde{\mathbf{e}}_1(t) \\ \vdots \\ \tilde{\mathbf{e}}_n(t) \end{bmatrix} = \begin{bmatrix} I & & & \\ B_1(0) & I & & \\ \dots & & \dots & \\ B_n(0) & \dots & B_n(n-1) & I \end{bmatrix} \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}(t-1) \\ \vdots \\ \mathbf{y}(t-n) \end{bmatrix} \quad (5.7)$$

e calcolando le varianze si arriva alla

$$\begin{bmatrix} \bar{\Lambda}_0 & & \\ & \ddots & \\ & & \bar{\Lambda}_n \end{bmatrix} = \tilde{B} T^* \tilde{B}' \quad , \quad (5.8)$$

dove $T^* := \text{Var}[\mathbf{y}(t)' \mathbf{y}(t-1)', \dots, \mathbf{y}(t-n)']'$. Notiamo che $T^* = T$ solo nel caso in cui si ha reversibilità temporale (ad esempio nel caso scalare). In tal caso infatti $T_{-k} = \Sigma(-k) = \Sigma(k) = T_k$.

Esiste però una semplice relazione tra T e T^* . Definiamo la matrice

$$J := \begin{bmatrix} 0 & \dots & 0 & I \\ \vdots & & & \vdots \\ I & 0 & \dots & 0 \end{bmatrix} \quad ,$$

dove i blocchi identità sono di dimensione $m \times m$. Premoltiplicando per J si rovescia l'ordine dei blocchi riga e postmoltiplicando per J si rovescia l'ordine dei blocchi colonna. Si ha quindi

$$J T J = T^* \quad .$$

Definendo allora $\tilde{B}^* := \tilde{B} J$, la fattorizzazione (5.8) può essere scritta nella forma

$$\begin{bmatrix} \bar{\Lambda}_0 & & \\ & \ddots & \\ & & \bar{\Lambda}_n \end{bmatrix} = \tilde{B}^* T (\tilde{B}^*)' \quad , \quad (5.9)$$

dove adesso \tilde{B}^* è ancora triangolare inferiormente ma rispetto all'antidiagonale principale, ovvero

$$B^* = \begin{bmatrix} 0 & \dots & I & B_1(0) \\ I & B_n(n-1) & \dots & B_n(0) \end{bmatrix} .$$

Moltiplicando per J a destra e a sinistra nella (5.9) si arriva così alla

$$\begin{bmatrix} \bar{\Lambda}_n & & \\ & \ddots & \\ & & \bar{\Lambda}_0 \end{bmatrix} = \tilde{C} T \tilde{C}' , \quad (5.10)$$

dove

$$\tilde{C} := \begin{bmatrix} I & B_n(n-1) & \dots & B_n(0) \\ & \ddots & & \\ & & I & B_1(0) \\ & & & I \end{bmatrix} . \quad (5.11)$$

Questa formula conduce quindi ad una *fattorizzazione dell'inversa di T* ,

$$T^{-1} = \tilde{C}' \begin{bmatrix} \bar{\Lambda}_n^{-1} & & \\ & \ddots & \\ & & \bar{\Lambda}_0^{-1} \end{bmatrix} \tilde{C} , \quad (5.12)$$

in cui \tilde{C}' è triangolare inferiore con diagonale unitaria. Notiamo che in queste formule la matrice \tilde{C} e quella delle varianze d'errore si aggiornano con n , "dal basso verso l'alto".

In un certo senso questo corrisponde all'aver invertito la direzione dei tempi.

Le formule (5.6) e (5.12) risolvono in particolare il problema del calcolo della soluzione $x = T^{-1}b$ della (5.1) con b qualunque, con un costo di calcolo proporzionale ad n^2 . Occorre ancora dire che esistono numerose varianti dei risultati che abbiamo appena esposto e spesso in letteratura si trovano formule simili a quella ricavata qui, riportate sotto i nomi più esotici. Per una panoramica che discute anche il caso in cui T non è necessariamente simmetrica (o positiva) rimandiamo al lavoro di Kailath, Vieira e Morf [73] e alla bibliografia ivi citata.

L'ultima questione che vogliamo affrontare e che è un logico complemento a quella trattata più sopra è quella del *calcolo veloce* ("alla Levinson" per intenderci) *della fattorizzazione di Cholesky di T* .

Questo è un problema che si incontra assai spesso ed è per esempio intimamente legato al calcolo della fattorizzazione spettrale della matrice $S(z)$, densità spettrale di potenza associata al processo $\{\mathbf{y}(t)\}$.

Con le notazioni introdotte nella formula (5.2) si ha

$$S(z) = \sum_{-\infty}^{+\infty} T_k z^{-k}$$

e questa corrispondenza ci permetterà facilmente di associare fattorizzazioni triangolari della matrice a blocchi T a fattori spettrali approssimati di $S(z)$.

Come è già stato più volte messo in evidenza, la fattorizzazione di Cholesky della matrice di covarianza $[T_{kj} := E\mathbf{y}(t+k)\mathbf{y}'(t+j)]$, $k, j = 0, 1, \dots, n$, di un

processo stocastico $\{\mathbf{y}(t)\}$, può essere vista come un metodo per calcolare esplicitamente la *rappresentazione d'innovazione* (normalizzata) di $\{\mathbf{y}(t)\}$. Con questa terminologia intendiamo qui riferirci alla rappresentazione delle $n+1$ variabili (vettoriali di dimensione m), $\mathbf{y}(t), \mathbf{y}(t+1), \dots, \mathbf{y}(t+n)$ come funzioni lineari e causali delle innovazioni a memoria finita $\varepsilon_0(t), \varepsilon_1(t+1), \dots, \varepsilon_n(t+n)$ (normalizzate) ottenute ortonormalizzando le variabili del processo una dopo l'altra nella successione temporale assegnata.

La fattorizzazione di Cholesky della matrice di Toeplitz T definita nella (5.2) sarà quindi nota (modulo una banale operazione di normalizzazione) quando sia disponibile la matrice dei coefficienti, \tilde{W} , nella rappresentazione (d'innovazione non normalizzata)

$$\begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}(t+1) \\ \vdots \\ \mathbf{y}(t+n) \end{bmatrix} = \begin{bmatrix} I & & & \\ W_1(0) & I & & \\ \vdots & & \ddots & \\ W_n(0) & \dots & W_n(n-1) & I \end{bmatrix} \begin{bmatrix} \mathbf{e}_0(t) \\ \mathbf{e}_1(t+1) \\ \vdots \\ \mathbf{e}_n(t+n) \end{bmatrix}. \quad (5.13)$$

Prendendo in effetti le covarianze a primo e secondo membro si trova

$$T = \tilde{W} \begin{bmatrix} \Lambda_0 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \Lambda_n \end{bmatrix} \tilde{W}' \quad , \quad (5.14)$$

in cui \tilde{W} è manifestamente triangolare con diagonale unitaria. Volendo la fattorizzazione di Cholesky basta moltiplicare a destra ciascun blocco riga, W_k , in \tilde{W} per il fattore di Cholesky $\Lambda_k^{1/2}$ di Λ_k .

Confrontando la rappresentazione d'innovazione (5.13) con quella dei filtri sbiancanti a memoria finita (5.4) si vede che le due relazioni sono una l'inversa dell'altra. In particolare le matrici dei coefficienti \tilde{W} e \tilde{A} sono legate dalla relazione

$$\tilde{W} = \tilde{A}^{-1} \quad . \quad (5.15)$$

Sfruttando le equazioni di aggiornamento dei coefficienti A_k (i blocchi riga della matrice \tilde{A}) e la struttura triangolare inferiore di \tilde{W} e \tilde{A} , si può arrivare ad un algoritmo ricorsivo per l'aggiornamento dei blocchi riga,

$$W_k = [W_k(0), \dots, W_k(k-1)I] \quad ,$$

che ha struttura simile all'algoritmo di Levinson e richiede, per il calcolo di \tilde{W} , un numero di operazioni elementari proporzionale ad n^2 . Questo algoritmo di fattorizzazione "veloce" di T è stato derivato per la prima volta da J. Rissanen [121] mediante una serie di manipolazioni formali sull'equazione (5.15).

Ne presenteremo una derivazione "probabilistica" qui sotto.

L'osservazione da cui parte il procedimento è che

$$W_k(j) = E \mathbf{y}(t+k) \mathbf{e}'_j(t+j) \cdot \Lambda_j^{-1}$$

per $k = 0, 1, \dots, n$ e, per ogni k fissato, j variabile nell'intervallo $0 \leq j \leq k$. Definendo

$$L_k(j) := W_k(j) \Lambda_j \quad (5.16)$$

si vede che, usando le formule di aggiornamento per l'innovazione a memoria finita (2.9), valgono le relazioni

$$\begin{aligned} L_{k+1}(j+1) &= E \mathbf{y}(t+k+1) [\mathbf{e}_j(t+j+1) - \bar{\Gamma}_j \tilde{\mathbf{e}}_j(t+j)]' \\ &= E \mathbf{y}(t+k) \mathbf{e}'_j(t+j) - E \mathbf{y}(t+k) \tilde{\mathbf{e}}_j(t-1-j)' \bar{\Gamma}'_j \\ &:= L_k(j) - P_k(j) \bar{\Gamma}'_j \quad , \end{aligned} \quad (5.17)$$

dove $P_k(j)$ è definita dalla

$$P_k(j) := E \mathbf{y}(t+k) \tilde{\mathbf{e}}'_j(t-1-j) \quad .$$

Usando la (2.9) si trova analogamente

$$P_k(j+1) = E \mathbf{y}(t+k) [\tilde{\mathbf{e}}_j(t-1+j) - \Gamma_j \mathbf{e}_j(t+j)]' = P_k(j) - L_k(j) \Gamma'_j \quad . \quad (5.18)$$

Queste formule valgono per $0 \leq j \leq k$ e permettono il calcolo ricorsivo di $L_k := [L_k(0), \dots, L_k(k)]$ all'aumentare di k , naturalmente noti che siano i coefficienti di riflessione $\bar{\Gamma}_j$ e Γ_j .

Per trovare le *condizioni iniziali* con cui far partire l'algoritmo basta notare che

$$\begin{aligned} L_k(0) &= E \mathbf{y}(t+k) \mathbf{e}'_0(t) = E \mathbf{y}(t+k) \mathbf{y}'(t) = T_k \\ P_k(0) &= E \mathbf{y}(t+k) \tilde{\mathbf{e}}'_0(t-1) = E \mathbf{y}(t+k) \mathbf{y}'(t-1) = T_{k+1} \quad . \end{aligned} \quad (5.19)$$

I *coefficienti di riflessione* possono essere calcolati ricorsivamente a partire dalle $L_k(j)$ e $P_k(j)$. Le formule di aggiornamento si ricavano nel modo seguente.

Notiamo innanzitutto che le "condizioni finali" associate alle ricorsioni (5.17) e (5.18) sono, rispettivamente:

$$L_k(k) = E \mathbf{y}(t+k) \mathbf{e}'_k(t+k) = E \mathbf{e}_k(t+k) \mathbf{e}'_k(t+k) = \Lambda_k \quad (5.20)$$

e

$$\begin{aligned} P_k(k) &= E \mathbf{y}(t+k) \tilde{\mathbf{e}}'_k(t-1+k) = E \mathbf{y}(t) \tilde{\mathbf{e}}'_k(t-1) \\ &= E \mathbf{e}_k(t) \tilde{\mathbf{e}}'_k(t-1) = P_k \quad , \end{aligned} \quad (5.21)$$

dove P_k è la matrice ($m \times m$!) introdotta nell'enunciato del Teorema 8.2 allo scopo di calcolare i coefficienti di riflessione. Ricordiamo in effetti che $\bar{\Gamma}_k$ e Γ_k si calcolano a partire dalle formule

$$P_k = \bar{\Gamma}_k \bar{\Lambda}_k = \Lambda_k \Gamma'_k \quad ,$$

dalle quali risulta immediatamente che

$$\Gamma'_k = \Lambda_k^{-1} P_k = L_k(k)^{-1} P_k(k) \quad . \quad (5.22)$$

Notiamo di passaggio che combinando questa espressione con l'equazione di aggiornamento (5.18) per $P_k(j)$, scritta per $j = k$, si ottiene:

$$P_k(k+1) = P_k(k) - L_k(k) L_k(k)^{-1} P_k(k) = 0$$

e questa uguaglianza mostra che nella (5.18) basta in realtà far correre l'indice j da 0 a $k-1$.

Per calcolare $\bar{\Gamma}_k$ occorre procurarsi la varianza dell'innovazione all'indietro $\bar{\Lambda}_k$. Ricordando le equazioni di aggiornamento (2.12) si può scrivere:

$$\bar{\Lambda}_{k+1} = \bar{\Lambda}_k - \Gamma_k \bar{\Gamma}_k \bar{\Lambda}_k = \bar{\Lambda}_k - \Gamma_k P_k \quad ,$$

ovvero:

$$\begin{aligned}\bar{\Lambda}_{k+1} &= \bar{\Lambda}_k - P'_k(k) L_k(k)^{-1} P_k(k) \\ \bar{\Lambda}_0 &= T_0 \quad .\end{aligned}\tag{5.23}$$

Chiaramente basterà aggiornare $\bar{\Lambda}_k$ alla fine di ogni iterazione su j e calcolare quindi il successivo valore del coefficiente di riflessione $\bar{\Gamma}'_k$ mediante la

$$\bar{\Gamma}'_k = \bar{\Lambda}_k^{-1} P'_k(k) \quad .\tag{5.24}$$

Combinando infine le (5.17), (5.18), (5.22), (5.23) e (5.24) si arriva all'algoritmo cercato.

Fattorizzazione di Cholesky di una matrice di Toeplitz a blocchi T .

ALGORITMO 8.1. *Il seguente algoritmo calcola la fattorizzazione di Cholesky "veloce" di una matrice T di Toeplitz a blocchi di dimensione $m \times m$*

$$T = \begin{bmatrix} T_0 & T_{-1} & \dots & T_{-n} \\ T_1 & & & \\ \vdots & & & \\ T_n & & & T_0 \end{bmatrix}, \quad T_{-k} = T'_k$$

che si assume non singolare.

- (1) *Si calcola dapprima un fattore inferiormente triangolare (a blocchi), \tilde{L} di T , che soddisfa l'equazione*

$$T = \tilde{L} \tilde{\Lambda}^{-1} \tilde{L}' \quad ,$$

dove $\tilde{\Lambda}$ è una matrice diagonale a blocchi

$$\tilde{\Lambda} = \text{diag} \{ \Lambda_0, \dots, \Lambda_n \}$$

in cui $\Lambda_k = \Lambda'_k > 0$, $k = 0, 1, \dots, n$. I blocchi di posto $(k+1, j)$ di \tilde{L} , $j = 0, 1, \dots, k+1$ (cioè il blocco riga di indice $k+1$) si ottengono da quelli del blocco riga precedente attraverso le relazioni

$$L_{k+1}(j+1) = L_k(j) - P_k(j) \bar{\Lambda}_j^{-1} P'_j(j) \quad , \quad 0 \leq j \leq k \quad ,\tag{5.25}$$

dove $P_k(j)$ e $\bar{\Lambda}_k$ sono matrici $(m \times m)$ calcolabili tramite la

$$P_k(j+1) = P_k(j) - L_k(j) L_j^{-1}(j) P_j(j) \quad , \quad 0 \leq j \leq k\tag{5.26}$$

e la

$$\bar{\Lambda}_{k+1} = \bar{\Lambda}_k - P'_k(k) L_k^{-1}(k) P_k(k) \quad .\tag{5.27}$$

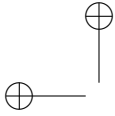
Le condizioni iniziali associate agli schemi ricorsivi (5.25), (5.26) e (5.27) sono, rispettivamente,

$$L_k(0) = T_k \quad , \quad P_k(0) = T_{k+1} \quad , \quad \bar{\Lambda}_0 = T_0 \quad .\tag{5.28}$$

Gli elementi della matrice diagonale $\tilde{\Lambda}$ che interviene nella fattorizzazione sono dati dalla

$$\Lambda_k = L_k(k) \quad , \quad k = 0, 1, \dots, n \quad .$$

- (2) *Il fattore di Cholesky, L , di T si costruisce moltiplicando a destra ciascun blocco riga, L_k , di \tilde{L} per $\Lambda_k^{-1/2}$, dove $\Lambda_k^{1/2}$ è a sua volta il fattore di Cholesky (triangolare inferiore) del blocco $(m \times m)$ di posto k in $\tilde{\Lambda}$.*



- (3) Le matrici $W_k(j)$ nella rappresentazione d'innovazione non normalizzata (5.13) si calcolano mediante la formula

$$W_k(j) = L_k(j) L_j(j)^{-1} \quad , \quad j = 0, 1, \dots, k \quad . \quad (5.29)$$

Si noti che il calcolo della $\{W_k(j) \ , \ 0 \leq j \leq k-1\}$ è esplicitamente richiesto per l'aggiornamento di $P_k(j)$ fino a $k = n-1$.

- (4) Nel caso scalare ($m = 1$), $\bar{\Lambda}_k = \Lambda_k = \ell_k(k)$ e l'algoritmo si riduce alla coppia di equazioni

$$\begin{aligned} \ell_{k+1}(j+1) &= \ell_k(j) - \frac{p_j(j)}{\ell_j(j)} p_k(j) \quad , \quad 0 \leq j \leq k \\ p_k(j+1) &= p_k(j) - \frac{p_j(j)}{\ell_j(j)} \ell_k(j) \quad , \quad 0 \leq j \leq k \quad , \end{aligned} \quad (5.30)$$

associata alle condizioni iniziali $\ell_k(0) = T_k = \sigma(k)$ e $p_k(0) = T_{k+1} = \sigma(k+1)$.

Ovviamente si può scrivere una versione duale all'algoritmo partendo dalla matrice $T^* = JTJ$ e arrivare in questo modo a costruire una fattorizzazione di T del tipo triangolare superiormente (a blocchi). Lascерemo i dettagli di questa costruzione al lettore.

Esercizi

8-1 (Varianti del problema 8.2) Usando la relazione (3.14) mostrare che per un processo scalare p.n.d. si ha sempre $|\gamma_n| < 1$. Viceversa, supponiamo che n sia il più piccolo intero per cui $|\gamma_n| = 1$. Mostrare che il processo \mathbf{y} è descritto da un'equazione alle differenze (deterministica !) di ordine n (e quindi è p.d.). Mostrare che se $|\gamma_n| = 1$ allora $|\gamma_{n+k}| = 1, \forall k \geq 0$.

8-2 Nella prima affermazione del problema 8.2, l'ipotesi che \mathbf{y} sia p.n.d. può essere indebolita. Per dimostrare che $|\gamma_n| < 1, \forall n$, è in realtà sufficiente assumere che $\det \Sigma_n > 0, \forall n$. Dimostrare questa affermazione usando la diagonalizzazione introdotta in (3.12).

8-3 Supponiamo che \mathbf{y} sia un processo scalare descritto da un modello autoregressivo (d'innovazione) di ordine n ,

$$A^*(z^{-1})\mathbf{y}(t) = \mathbf{e}(t) .$$

Dimostrare che l'innovazione a memoria n , $\mathbf{e}_n(t)$, coincide con $\mathbf{e}(t)$.

Dimostrare che per ogni $k \geq 0$ risulta

$$\lambda_{n+k}^2 = \lambda_n^2 = \lambda^2 = \mathbb{E} \mathbf{e}(t)^2$$

e che quindi tutti i coefficienti di riflessione $\{\gamma_{n+k}, k \geq 0\}$, sono nulli.

8-4 Generalizzare le affermazioni dell'esercizio precedente al caso di un processo vettoriale (di rango pieno).

8-5 Supponiamo che \mathbf{y} sia m -dimensionale puramente non deterministico e di rango pieno. Dimostrare che la varianza dell'innovazione a memoria finita Λ_n converge per $n \rightarrow \infty$ in modo monotono (i.e. $\Lambda_{n+1} \leq \Lambda_n$) alla varianza Λ , dell'innovazione stazionaria $\mathbf{e}(t)$. Usare la convergenza in media quadratica $\lim_{n \rightarrow \infty} \mathbf{e}_n(t) = \mathbf{e}(t)$ che è stata dimostrata nella prova del teorema 4.3 alla fine del capitolo 5.

Mostrare che vale anche la relazione limite duale, $\bar{\Lambda}_n \downarrow \bar{\Lambda}$, per la successione delle varianze dell'innovazione all'indietro.

8-6 Usando la definizione dei coefficienti di riflessione, si mostri che valgono le relazioni

$$\Lambda_{n+1} = \Lambda_n - P_n \bar{\Lambda}_n^{-1} P_n'$$

$$\bar{\Lambda}_{n+1} = \bar{\Lambda}_n - P_n' \Lambda_n^{-1} P_n.$$

Usando i risultati dell'esercizio precedente, passando al limite per $n \rightarrow \infty$, si dimostri che $\lim_{n \rightarrow \infty} P_n = 0$ e dedurre quindi che *i coefficienti di riflessione tendono a zero con n* , ovvero

$$\lim_{n \rightarrow \infty} \Gamma_n = 0 \quad \lim_{n \rightarrow \infty} \bar{\Gamma}_n = 0.$$

8-7 Ricavare in questo modo la disuguaglianza matriciale

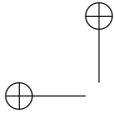
$$0 \leq (I - \bar{\Gamma}_n \Gamma_n) < I$$

e l'analoga per il prodotto $\Gamma_n \bar{\Gamma}_n$, basandosi sulla relazione duale $\bar{\Lambda}_{n+1} \leq \bar{\Lambda}_n$.

8-7 Calcolare i coefficienti di riflessione di un processo descritto dal modello MA,

$$\mathbf{y}(t) = (1 + bz^{-1})\mathbf{e}(t), \quad |b| < 1$$

dove \mathbf{e} è un processo bianco di varianza λ^2 .



CAPITOLO 9

IL FILTRO DI KALMAN

L'impostazione data nel 1960 da R.E. Kalman [74, 75] al problema del filtraggio statistico ha introdotto delle innovazioni profonde.

Nell'approccio di Wiener-Kolmogorov l'enfasi primaria è posta nel calcolo esplicito della funzione di trasferimento del filtro. In realtà, come si è visto nei capitoli precedenti, la formula risolutiva di Wiener (2.23) data al Cap. 8 è di uso pratico solo nel caso in cui i segnali in gioco hanno spettro razionale. Questo sia perchè sostanzialmente solo in questo caso si possono portare avanti i calcoli impliciti nella formula (specialmente la fattorizzazione spettrale), sia poi perchè l'implementazione del filtro mediante un algoritmo di complessità finita (fisicamente realizzabile) richiede che la sua funzione di trasferimento sia razionale. In effetti, lo schema numerico (o algoritmo) che *realizza* le operazioni sul segnale descritte dalla funzione di trasferimento $H(z)$ deve, si badi bene, richiedere sempre un numero *finito* di operazioni elementari sul segnale ad ogni passo di campionamento e, come è ben noto dalla teoria dei sistemi (teoria della realizzazione), la funzione di trasferimento $H(z)$ può essere realizzata mediante un algoritmo ricorsivo a memoria finita, ovvero un'equazione alle differenze, se e solo se $H(z)$ è una funzione *razionale* di z .

Il ruolo giocato dalla razionalità si è d'altronde visto assai chiaramente nella soluzione del problema della predizione.

È importante ricordare anche che la razionalità dello spettro di $\{\mathbf{y}(t)\}$ è equivalente alla rappresentabilità di $\{\mathbf{y}(t)\}$ come soluzione di un'equazione stocastica alle differenze di ordine finito pilotata da rumore bianco (ad esempio il modello ARMA di innovazione). La conoscenza di questo modello (che è in effetti equivalente a quella dello spettro) permette di risolvere il problema della predizione in modo molto più diretto e produce immediatamente una descrizione di tipo equazione alle differenze anche per il predittore. In altre parole porta direttamente ad una soluzione ideale dal punto di vista algoritmico.

Non è difficile convincersi che questo fenomeno di "preservazione della razionalità", quando si passa da un modello dei processi in gioco di tipo ARMA al corrispondente filtro, è un fatto assolutamente generale. Basta esaminare la formula generale (2.23) e si vede subito che per ottenere una funzione di trasferimento del filtro razionale e quindi algoritmi di stima a memoria finita (del tipo equazione lineare alle differenze, per intenderci) occorre supporre che i segnali in gioco abbiano spettri razionali e quindi siano essi stessi modellabili per mezzo di schemi di analoga struttura.

Se si esamina il problema in modo un po' più approfondito, ci si accorge anche che la razionalità in quanto tale è una proprietà contingente, legata all'uso delle tecniche dell'analisi armonica, che sono impiegabili solo per processi *stazionari*. Il fatto chiave è l'ipotesi che i segnali in gioco siano descrivibili per mezzo di *sistemi dinamici (lineari) di dimensione finita*. La fondamentale novità dell'approccio di

Kalman sta nell'aver introdotto questo tipo di modelli fin dall'inizio, all'atto stesso della posizione del problema.

Così com'è, questa ipotesi sembra ragionevole ma abbastanza innocua. Le sue conseguenze sono però profonde. Vedremo in seguito come in questo contesto il problema del calcolo esplicito della funzione di trasferimento del filtro scompaia completamente. Lo stimatore di Kalman ha una struttura algoritmica fissa e nota a priori a meno di un parametro. Si tratta di un sistema dinamico lineare a tempo discreto e a memoria finita, pilotato dalle osservazioni, che costituisce quindi uno schema numerico *ricorsivo* particolarmente adatto ad essere programmato e usato per l'elaborazione di segnali in linea. I suoi parametri sono noti a priori a partire dalle specifiche del modello, a meno di una matrice "guadagno" che può essere calcolata fuori linea risolvendo un'equazione alle differenze non lineare del tipo di Riccati.

Si può tranquillamente affermare che le tecniche statistiche di elaborazione del segnale diventano di impiego pratico e generalizzato nell'ingegneria solo a partire dagli anni sessanta e proprio in conseguenza dell'introduzione di modelli di stato e del filtro di Kalman. Questo del resto non sembra affatto essere un evento casuale. Gli anni sessanta corrispondono all'inizio dell'era in cui il calcolatore numerico cominciò a diffondersi in molti settori dell'ingegneria e l'idea di sistema dinamico (a tempo discreto) non è altro dopotutto, se non una naturale generalizzazione di quella di "algoritmo ricorsivo" o funzione "ricorsivamente computabile". La teoria dei sistemi è essa stessa, in un certo senso, la matematica che formalizza e risolve i problemi ingegneristici di analisi, controllo e misura (o stima) in termini di *algoritmi ricorsivi* (da notare che la nozione di sistema dinamico della teoria dei sistemi è sostanzialmente la stessa di funzione ricorsiva).

1. Modelli di stato di segnali di dimensione finita

Sia $\{\mathbf{y}(t)\}$ un processo stocastico m -dimensionale definito sull'intervallo temporale $[t_0, +\infty)$. Diremo che $\{\mathbf{y}(t)\}$ è un *processo di dimensione finita* se si può rappresentare nella forma

$$\begin{cases} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{n}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{n}(t) \end{cases}, \quad t \geq t_0, \quad (1.1)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0,$$

dove $\{\mathbf{n}(t)\}$ è un processo di rumore bianco p -dimensionale di varianza unitaria (processo ortonormale), i.e.

$$E \mathbf{n}(t) \mathbf{n}(s)' = I \delta(t-s), \quad (1.2)$$

$\{\mathbf{x}(t)\}$ è un processo di dimensione n , il cui valore iniziale \mathbf{x}_0 è un vettore aleatorio *scorrelato dalla storia presente e futura di $\{\mathbf{n}(t)\}$* ,

$$E \mathbf{x}_0 \mathbf{n}(t)' = 0, \quad \forall t \geq t_0, \quad (1.3)$$

di cui supponiamo noti i momenti del primo e secondo ordine

$$E \mathbf{x}_0 = \boldsymbol{\mu}_0, \quad \text{Var } \mathbf{x}_0 = \boldsymbol{\Sigma}_0 \quad (1.4)$$

Infine le matrici A, B, C, D sono quantità deterministiche, di dimensioni compatibili con quelle delle variabili in gioco.

Si vede che un processo di dimensione finita è sostanzialmente un processo che può essere rappresentato come l'uscita di un sistema dinamico lineare di dimensione finita pilotato da rumore bianco. Ogni rappresentazione di questo tipo verrà chiamata *realizzazione*¹ di $\{\mathbf{y}(t)\}$.

NOTA BENE: Buona parte delle considerazioni che verranno fatte in questo capitolo, compresa la derivazione delle equazioni dinamiche del filtro di Kalman, valgono nel caso generale in cui le matrici A, B, C, D che caratterizzano il modello (1.1) sono funzioni arbitrarie (ma note, ovvero “deterministiche”) del tempo. Per non appesantire la notazione e soprattutto per non oscurare la trattazione con formule complicate faremo però riferimento solo al caso in cui *i parametri* A, B, C, D *sono costanti*. Il lettore può facilmente ripercorrere l'analisi che segue eliminando questa ipotesi e introducendo, quando necessario, gli opportuni cambiamenti. \diamond

Il processo $\{\mathbf{x}(t)\}$ si chiama *processo di stato* del modello (1.1). Esso gode di una notevole proprietà.

LEMMA 9.1. *Il processo di stato $\{\mathbf{x}(t)\}$ è un processo di Markov in senso lato, nel senso che vale la*

$$\hat{E}[\mathbf{x}(t) | \mathbf{H}_s(\mathbf{x})] = \hat{E}[\mathbf{x}(t) | \mathbf{x}(s)] \quad , \quad \forall t \geq s \quad , \quad (1.5)$$

dove il primo membro nella (1.5) rappresenta la proiezione ortogonale di $\{\mathbf{x}(t)\}$ sul sottospazio di Hilbert $\mathbf{H}_s(\mathbf{x})$ generato dalle variabili aleatorie $\{\mathbf{x}(t_0), \mathbf{x}(t_0 + 1), \dots, \mathbf{x}(s)\}$.

Se $\{\mathbf{n}(t)\}$ e \mathbf{x}_0 sono congiuntamente Gaussiani, allora $\{\mathbf{x}(t)\}$ è Gaussiano e di Markov in senso stretto.

DIMOSTRAZIONE. Risolvendo la (1.1) si ottiene

$$\mathbf{x}(t) = A^{t-t_0} \mathbf{x}_0 + \sum_{k=t_0}^{t-1} A^{t-1-k} B \mathbf{n}(k) \quad ,$$

dalla quale si vede che $\mathbf{x}(t)$ è funzione (lineare) di \mathbf{x}_0 ed $\{\mathbf{n}(t_0), \dots, \mathbf{n}(t-1)\}$. Ne risulta, per l'ipotesi (1.3), che $\mathbf{x}(t)$ è scorrelato da $\{\mathbf{n}(t), \mathbf{n}(t+1), \dots\}$, ovvero

$$E[\mathbf{x}(s) \mathbf{n}(t)'] = 0 \quad , \quad \forall t \geq s \quad . \quad (1.6)$$

In particolare nell'espressione

$$\mathbf{x}(t) = A^{t-s} \mathbf{x}(s) + \sum_s^{t-1} A^{t-1-k} B \mathbf{n}(k)$$

i due addendi a secondo membro risultano ortogonali.

Si noti che per quanto appena detto si ha

$$\{\mathbf{n}(k) \ , \ k \geq s\} \perp \mathbf{H}_s(\mathbf{x}) \quad (1.7)$$

per ogni s e pertanto la proiezione di $\mathbf{x}(t)$ su $\mathbf{H}_s(\mathbf{x})$ coincide con $A^{t-s} \mathbf{x}(s)$, cioè

$$\hat{E}[\mathbf{x}(t) | \mathbf{H}_s(\mathbf{x})] = \hat{E}[\mathbf{x}(t) | \mathbf{x}(s)] = A^{t-s} \mathbf{x}(s) \quad .$$

¹In questo libro evitiamo di usare il termine “realizzazione” (come viene talvolta fatto) per intendere una “determinazione campionaria” o *traiettoria* del processo. La parola “realizzazione” verrà invece usata per denotare una *rappresentazione* del processo in termini di variabili di stato, similmente a quanto si fa in teoria dei sistemi.

Nel caso Gaussiano la (1.5) vale per le medie condizionate. D'altro canto la probabilità condizionata $P\{\mathbf{x}(t) \in A \mid \mathbf{x}(t_0), \dots, \mathbf{x}(s)\}$ in questo caso dipende dai dati $\{\mathbf{x}(t_0), \dots, \mathbf{x}(s)\}$ solo attraverso la media condizionata e quindi se vale la (1.5) essa è equivalente alla

$$P\{\mathbf{x}(t) \in A \mid \mathbf{x}(t_0), \dots, \mathbf{x}(s)\} = P\{\mathbf{x}(t) \in A \mid \mathbf{x}(s)\}$$

per tutti $t \geq s$, per ogni insieme di Borel $A \subset \mathbb{R}$. \square

È evidente che la proprietà di Markov costituisce la naturale generalizzazione probabilistica della proprietà di memoria dinamica del vettore di stato nei sistemi dinamici deterministici. Tutta l'informazione contenuta nella storia passata e presente, $\mathbf{H}_s(\mathbf{x})$, che serve per predire l'evoluzione futura del processo $\{\mathbf{x}(t)\}$ dall'istante s in poi, è concentrata nel solo vettore casuale $\mathbf{x}(s)$ che rappresenta l'informazione "presente". La conoscenza del passato prima di s , $\{\mathbf{x}(k), k < s\}$, non aggiunge nulla che non sia già contenuto in $\mathbf{x}(s)$.

La proprietà di "memoria dinamica" della variabile di stato vale anche nei confronti del processo di uscita $\{\mathbf{y}(t)\}$. Formalmente essa si può esprimere nel modo seguente.

PROPOSIZIONE 9.1. *I processi di stato e di uscita di un sistema stocastico lineare del tipo (1.1) soddisfano alle relazioni*

$$\hat{E}[\mathbf{x}(t) \mid \mathbf{H}(\mathbf{x}^s, \mathbf{y}^{s-1})] = \hat{E}[\mathbf{x}(t) \mid \mathbf{x}(s)] \quad , \quad t \geq s \quad , \quad (1.8)$$

$$\hat{E}[\mathbf{y}(t) \mid \mathbf{H}(\mathbf{x}^s, \mathbf{y}^{s-1})] = \hat{E}[\mathbf{y}(t) \mid \mathbf{x}(s)] \quad , \quad t \geq s \quad , \quad (1.9)$$

dove $\mathbf{H}(\mathbf{x}^s, \mathbf{y}^{s-1})$ è lo spazio di Hilbert generato da $\mathbf{H}_s(\mathbf{x})$ e dal passato "stretto" $\{\mathbf{y}(t_0), \mathbf{y}(t_0 + 1), \dots, \mathbf{y}(s - 1)\}$ del processo di uscita.

DIMOSTRAZIONE. Le (1.8) e (1.9) scendono immediatamente dalla condizione di ortogonalità

$$\mathbf{H}_s^+(\mathbf{n}) \perp \mathbf{H}(\mathbf{x}^s, \mathbf{y}^{s-1}) \quad , \quad (1.10)$$

che generalizza la (1.7) vista nella prova del lemma precedente. Qui $\mathbf{H}_s^+(\mathbf{n})$ è la storia futura, dal tempo s (incluso) in poi, del rumore bianco $\{\mathbf{n}(t)\}$. La (1.10) d'altro canto è ancora una conseguenza della dipendenza lineare e *causale* delle variabili $\{\mathbf{x}(\tau), \mathbf{y}(\tau - 1); t_0 \leq \tau \leq s\}$ dalla condizione iniziale $\mathbf{x}(t_0)$ e dal rumore passato $\{\mathbf{n}(\tau); t_0 \leq \tau < s\}$. \square

Dalla dimostrazione segue che la proprietà di Markov e quella, più generale, di memoria dinamica del processo di stato $\{\mathbf{x}(t)\}$ si possono scrivere come relazioni di *ortogonalità (incorrelazione) condizionata*. Per la definizione di ortogonalità condizionata e le proprietà principali si rimanda il lettore alla sezione 3.

A questo scopo, definiamo il vettore $n + m$ dimensionale

$$\mathbf{z}(t) := \begin{bmatrix} \mathbf{x}(t + 1) \\ \mathbf{y}(t) \end{bmatrix} \quad (1.11)$$

e compattiamo un pò le notazioni introducendo

$$\mathbf{H}_t^+(\mathbf{z}) := \mathbf{H}_{t+1}^+(\mathbf{x}) \vee \mathbf{H}_t^+(\mathbf{y}), \quad \mathbf{H}_t^-(\mathbf{z}) := \mathbf{H}_{t+1}^-(\mathbf{x}) \vee \mathbf{H}_t^-(\mathbf{y})$$

Usando lo stesso argomento della dimostrazione precedente, è immediato verificare che

PROPOSIZIONE 9.2. *Se i processi (\mathbf{x}, \mathbf{y}) sono legati da un modello lineare del tipo (1.1), detto $\mathbf{X}_t := \text{span}\{\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)\}$, lo spazio di stato del sistema all'istante t , vale la proprietà di ortogonalità condizionata*

$$\mathbf{H}_t^-(\mathbf{z}) \perp \mathbf{H}_t^+(\mathbf{z}) \mid \mathbf{X}_t \quad t_0 \leq t \quad (1.12)$$

Da cui, dato che $\mathbf{z}(t) \in \mathbf{H}_t^+(\mathbf{z})$, usando il lemma ?? si vede che le relazioni (1.8) e (1.8) sono in realtà una conseguenza della proprietà generale (1.12).
si può così scrivere

$$\hat{\mathbf{z}}(t \mid t-1) = \hat{E}[\mathbf{z}(t) \mid \mathbf{x}(t)] = \begin{bmatrix} A \\ C \end{bmatrix} \mathbf{x}(t).$$

notiamo che il modello (1.1) corrisponde alla decomposizione

$$\mathbf{z}(t) = \hat{\mathbf{z}}(t \mid t-1) + \mathbf{v}(t), \quad (1.13)$$

dove $\hat{\mathbf{z}}(t \mid t-1)$ è il predittore di un passo di $\mathbf{z}(t)$ basato sulla storia passata (a partire dall'istante t_0) e $\mathbf{v}(t) := \mathbf{z}(t) - \hat{\mathbf{z}}(t \mid t-1)$ è il corrispondente processo di innovazione (non normalizzata).

Calcolo dei momenti. È facile calcolare i momenti del primo e del second'ordine dei processi $\{\mathbf{y}(t)\}$ e $\{\mathbf{x}(t)\}$ a partire dalla descrizione (1.1). Per le medie si ha

$$\begin{aligned} \mu_{\mathbf{x}}(t+1) &= A \mu_{\mathbf{x}}(t) \quad , \quad \mu_{\mathbf{x}}(t_0) = \mu_0 \quad , \\ \mu_{\mathbf{y}}(t) &= C \mu_{\mathbf{x}}(t) \quad , \end{aligned} \quad (1.14)$$

dove $\mu_0 = E \mathbf{x}_0$.

La covarianza di $\{\mathbf{x}(t)\}$, diciamola $\Sigma_{\mathbf{x}}(t, s)$, può essere scritta, usando la (1.7), nella forma

$$\Sigma_{\mathbf{x}}(t, s) = A^{t-s} \Sigma(s) \quad , \quad t \geq s \quad ,$$

dove la matrice $\Sigma(s) := \text{Var}(\mathbf{x}(s))$ (detta *varianza di stato*) soddisfa l'equazione matriciale lineare

$$\Sigma(t+1) = A \Sigma(t) A' + B B' \quad , \quad \Sigma(t_0) = \Sigma_0 \quad .$$

Σ_0 essendo la varianza della condizione iniziale

$$\Sigma_0 = \text{Var} \{ \mathbf{x}_0 \}$$

Notando che $\mathbf{y}(s) \in \mathbf{H}_{s+1}(\mathbf{x}, \mathbf{y})$ e usando la (1.9), la covarianza di $\mathbf{y}(t)$ si può allora esprimere, per $t > s$,

$$\begin{aligned} \Sigma_{\mathbf{y}}(t, s) &= E \mathbf{y}(t) \mathbf{y}(s)' = E \left\{ E[\mathbf{y}(t) \mid \mathbf{H}_{s+1}(\mathbf{x}, \mathbf{y})] \mathbf{y}(s)' \right\} \\ &= C A^{t-s-1} E \left\{ \mathbf{x}(s+1) \mathbf{y}(s)' \right\} \\ &= C A^{t-s-1} [A \Sigma(s) C' + B D'] \quad , \end{aligned}$$

mentre per $t = s$ si ha

$$\Sigma_{\mathbf{y}}(t, t) = C \Sigma(t) C' + D D' \quad .$$

Ricordiamo che $\Sigma_{\mathbf{x}}(t, s)$ e $\Sigma_{\mathbf{y}}(t, s)$ sono determinate dalle relazioni appena viste, anche per $t < s$, dato che per una qualunque matrice di covarianza si ha $\Sigma(s, t) = \Sigma(t, s)'$.

Possiamo allora raccogliere queste osservazioni nella proposizione seguente.

PROPOSIZIONE 9.3. *L'evoluzione temporale della media e della covarianza dei processi $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ descritti dal sistema stocastico (1.1) è determinata, rispettivamente, dalle equazioni (1.14) e dalle*

$$\Sigma_{\mathbf{x}}(t, s) = \begin{cases} A^{t-s} \Sigma(s) & t \geq s \\ \Sigma(t) (A')^{s-t} & t \leq s \end{cases} \quad (1.15)$$

$$\Sigma_{\mathbf{y}}(t, s) = \begin{cases} CA^{t-s-1} G(s) & t > s \\ C \Sigma(t) C' + DD' & t = s \\ G(t)' (A')^{s-t-1} C' & t < s \end{cases} \quad (1.16)$$

dove $G(s) := A \Sigma(s) C' + BD'$ e la matrice varianza di stato $\Sigma(t)$ è soluzione dell'equazione lineare

$$\Sigma(t+1) = A \Sigma(t) A' + BB' \quad , \quad \Sigma(t_0) = \Sigma_0 \quad . \quad (1.17)$$

Stazionarietà asintotica. Per capire quali sono i campi di applicazione naturale delle due teorie di Wiener-Kolmogorov e di Kalman è utile mettere in chiaro le condizioni sotto le quali il modello (1.1) descrive dei *processi stazionari*.

Diremo a questo proposito che un processo del second'ordine $\{\mathbf{y}(t)\}$, definito per $t \geq t_0$, è *asintoticamente stazionario* (in senso lato) se per $t - t_0 \rightarrow \infty$, la matrice di covarianza $\Sigma_{\mathbf{y}}(t, s)$, $t, s \geq t_0$, tende ad una funzione di covarianza che dipende dalla sola differenza $t - s$.

Ovviamente per la stazionarietà asintotica occorrerà innanzitutto che il sistema dinamico (1.1) sia a *parametri costanti*, ovvero le matrici A, B, C, D siano indipendenti dal tempo, cosa che assumeremo d'ora in avanti.

È abbastanza intuitivo che la restrizione più importante dovrà poi riguardare la *stabilità asintotica della matrice*² A . Ricordiamo allo scopo il seguente utile risultato che generalizza il *lemma di Lyapunov*, noto dalla teoria dei sistemi.

LEMMA 9.2. *Sia Q una matrice simmetrica e semidefinita positiva e B una qualunque "radice quadrata" di Q , ovvero B soddisfi la*

$$Q = BB' \quad . \quad (1.18)$$

Allora due qualunque delle condizioni seguenti implicano la rimanente:

- i) *la coppia (A, B) è completamente raggiungibile;*
- ii) *la matrice A è asintoticamente stabile;*
- iii) *l'equazione matriciale di Lyapunov*

$$X = AXA' + Q \quad (1.19)$$

ammette una soluzione P che è simmetrica e definita positiva.

Se A è asintoticamente stabile, la soluzione della (1.19) è necessariamente unica ed è data dalla formula

$$P = \sum_0^{\infty} A^k Q (A')^k \quad . \quad (1.20)$$

²La stabilità asintotica a cui ci si riferisce è, ovviamente, quella del sistema lineare $x(t+1) = Ax(t)$. In questo testo, per evitare di ripetere continuamente, con un lungo giro di parole, che ci si riferisce alla stabilità del sistema dinamico descritto dalla matrice A , parleremo semplicemente di stabilità o stabilità asintotica di A .

DIMOSTRAZIONE. Cominciamo dall'ultima affermazione e dimostriamo un risultato un po' più generale di quello enunciato.

LEMMA 9.3. *L'equazione di Lyapunov (1.19) ha un'unica soluzione qualunque sia il termine noto $Q \in \mathbb{R}^{n \times n}$, se e solo se lo spettro di A non contiene elementi reciproci ($\lambda_k \in \sigma(A) \Rightarrow 1/\lambda_k \notin \sigma(A)$).*

DIMOSTRAZIONE. L'applicazione da $\mathbb{R}^{n \times n}$ in sè stesso

$$X \mapsto X - AXA' \quad (1.21)$$

è lineare e i suoi autovalori sono $1 - \lambda_k \lambda_j$ se $\{\lambda_k\}$ sono gli autovalori di A . In effetti, detto a_k l'autovettore di A corrispondente a λ_k , si vede immediatamente che

$$X_{kj} = a_k a_j'$$

è l'"automatrice" di (1.21) corrispondente all'autovalore $1 - \lambda_k \lambda_j$.

Come è ben noto, l'applicazione (1.21) è iniettiva (e quindi anche suriettiva) se e solo se non ha autovalori nulli, cioè $1 - \lambda_k \lambda_j \neq 0$. \square

Pertanto se A è asintoticamente stabile esiste sempre una e una sola soluzione dell'equazione di Lyapunov (1.19). È poi immediato verificare che la somma della serie (convergente) (1.20) è una soluzione, e quindi l'unica soluzione.

Dimostrare che i) + ii) \Rightarrow iii) diventa a questo punto banale, stante che P domina (è "maggiore" di) il gramiano di raggiungibilità di (A, B) .

Che i) + iii) \Rightarrow ii) si può dimostrare ragionando per assurdo.

Si assuma che A abbia un autovalore λ_0 di modulo maggiore o uguale a 1 con corrispondente autovettore (in genere complesso) a . Segue dall'ipotesi iii) che

$$a^* P a = |\lambda_0|^2 a^* P a + a^* B B' a \quad ,$$

ovvero

$$(1 - |\lambda_0|^2) a^* P a = a^* B B' a \quad , \quad (1.22)$$

dove il primo membro di questa uguaglianza è ≤ 0 ($a^* P a > 0$), mentre il secondo è ≥ 0 . Ne segue che entrambi debbono essere nulli. In particolare deve essere $a^* B = 0$ e pertanto c'è un autovettore (a) di A che è ortogonale alle colonne di B . Questo fatto contraddice la raggiungibilità completa di (A, B) . Ne consegue che $|\lambda_0|^2$ deve essere < 1 .

Con un ragionamento dello stesso tipo si mostra che ii) + iii) \Rightarrow i). Infatti supponendo questa volta che esista un vettore $a \neq 0$ ortogonale alla colonna di $[B \ AB \ \dots \ A^{n-1} B]$, usando Cayley-Hamilton si mostra facilmente che a deve anche essere ortogonale alle colonne di $A^k B$ qualunque sia k .

Pertanto

$$a^* \sum_0^{+\infty} A^k B B' (A')^k a = a^* P a = 0 \quad ,$$

il che contraddice la positività stretta di P . \square

Usando questo lemma si possono discutere le condizioni di stazionarietà asintotica dei processi \mathbf{x} e \mathbf{y} nel sistema (1.1).

PROPOSIZIONE 9.4. *Se la matrice A è asintoticamente stabile i processi $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ descritti dal modello lineare (1.1) tendono, quando $t - t_0 \rightarrow +\infty$, a diventare congiuntamente stazionari e le rispettive covarianze sono, per $t - t_0 \rightarrow +\infty$, descritte dalle relazioni*

$$\Sigma_{\mathbf{x}}(t - s) = A^{t-s} \bar{\Sigma} \quad , \quad t \geq s \quad , \quad (1.23a)$$

$$\Sigma_{\mathbf{y}}(t - s) = \begin{cases} CA^{t-s-1} \bar{G} & t > s \\ C \bar{\Sigma} C' + DD' & t = s \end{cases} \quad (1.23b)$$

dove $\bar{G} := A \bar{\Sigma} C' + BD'$ e $\bar{\Sigma} := \lim_{t-t_0 \rightarrow +\infty} \Sigma(t)$ è la matrice varianza asintotica dello stato, soluzione dell'equazione di Lyapunov,

$$\bar{\Sigma} = A \bar{\Sigma} A' + BB' \quad . \quad (1.24)$$

Inoltre, se $\mathbf{x}(t_0)$ ha covarianza (iniziale) $\Sigma_0 = \bar{\Sigma}$ allora $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ sono congiuntamente stazionari in $t \geq t_0$ e le loro covarianze sono date dalle relazioni (1.23) e (1.24) su tutto l'intervallo $t \geq t_0$.

DIMOSTRAZIONE. La dimostrazione si riconduce, vista la Proposizione 9.3, a far vedere che la soluzione, $\Sigma(t)$, dell'equazione (1.17) converge quando $t - t_0 \rightarrow +\infty$ e converge ad una matrice $\bar{\Sigma}$ che soddisfa l'equazione di Lyapunov (1.24). La cosa è in effetti un'ovvia conseguenza dell'asintotica stabilità di A , dato che risolvendo la (1.17) si ottiene

$$\Sigma(t) = A^{t-t_0} \Sigma_0 (A')^{t-t_0} + \sum_0^{t-t_0-1} A^k BB' (A')^k \quad , \quad (1.25)$$

dalla quale si vede che, nelle ipotesi poste,

$$\begin{aligned} \lim_{t-t_0 \rightarrow +\infty} \Sigma(t) &= \lim_{t-t_0 \rightarrow +\infty} \sum_0^{t-t_0-1} A^k BB' (A')^k \\ &= \sum_0^{\infty} A^k BB' (A')^k := \bar{\Sigma} < \infty \end{aligned}$$

ed è immediato verificare che $\bar{\Sigma}$ soddisfa la (1.24). \square

Notiamo che il risultato appena dimostrato può anche essere letto in un modo leggermente diverso e forse più significativo: se si fa tendere t (mantenendo fisso t_0) a $+\infty$ ed A è asintoticamente stabile, una volta esaurito il "transitorio" dovuto alle condizioni iniziali $\mathbf{x}(t_0)$ (che in generale non sono tali da garantire stazionarietà), i processi $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ diventano stazionari e le relative matrici di covarianza (asintotiche) sono date dalle (1.23a), (1.23b) e (1.24). Si vede in particolare che la varianza asintotica dello stato è una costante, $\bar{\Sigma}$, che *non dipende dalla condizione iniziale* Σ_0 .

Chiaramente si possono avere dei modelli (1.1) che descrivono processi asintoticamente stazionari anche se A non è asintoticamente stabile. Si pensi al caso banale in cui $\mathbf{n}(t) = 0$; ad esempio $\mathbf{x}(t+1) = \mathbf{x}(t)$. Se si suppone però che (A, B) sia *raggiungibile*, il che significa che il rumore bianco influenza (alla lunga) tutte le componenti del vettore di stato $x_i(t)$, $i = 1, \dots, n$, allora l'asintotica stabilità di A diventa anche una condizione necessaria per la stazionarietà.

PROPOSIZIONE 9.5. *Se (A, B) è raggiungibile e il processo di stato $\{\mathbf{x}(t)\}$ del modello (1.1) è asintoticamente stazionario, allora A è asintoticamente stabile. In questo caso $\bar{\Sigma}$ è definita positiva.*

DIMOSTRAZIONE. In effetti allora $\Sigma(t)$ deve convergere quando $t - t_0 \rightarrow +\infty$ e necessariamente verso una matrice simmetrica e semidefinita positiva che soddisfa la (1.24). Basta allora notare che la formula risolutiva (1.25) implica, per $t - t_0$ abbastanza grande, che

$$\Sigma(t) \geq \sum_n^{t-t_0-1} A^k B B' (A')^k + \sum_0^{n-1} A^k B B' (A')^k > 0 \quad ,$$

dove l'ultimo addendo a secondo membro è il gramiano di raggiungibilità di (A, B) , definito positivo per ipotesi.

Passando al limite si trova $\bar{\Sigma} > 0$ e in base all'ultima affermazione del lemma 9.2, A è asintoticamente stabile. \square

Se A è asintoticamente stabile e t è sufficientemente grande da ritenere che il "transitorio" del sistema (1.1) si sia esaurito, il processo $\{\mathbf{y}(t)\}$ è praticamente un processo stazionario. Per calcolare il suo spettro basta osservare che le equazioni (1.1) descrivono $\{\mathbf{y}(t)\}$ come risposta al rumore bianco ortonormale $\{\mathbf{n}(t)\}$ del sistema lineare e invariante di matrice di trasferimento

$$W(z) = C(zI - A)^{-1} B + D \quad . \quad (1.26)$$

L'espressione della matrice densità spettrale di potenza del processo di uscita è allora semplicemente

$$S(z) = W(z) W(z^{-1})' \quad . \quad (1.27)$$

Quindi $W(z)$ è un *fattore spettrale*. La (1.27) può essere riscritta nella forma più esplicita

$$S(z) = [C(zI - A)^{-1} I] \begin{bmatrix} B \\ D \end{bmatrix} [B' D'] \begin{bmatrix} (z^{-1}I - A')^{-1} C' \\ I \end{bmatrix} \quad . \quad (1.28)$$

Notiamo che la matrice

$$\begin{bmatrix} B \\ D \end{bmatrix} [B' D'] = \begin{bmatrix} BB' & BD' \\ DB' & DD' \end{bmatrix} := \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \quad (1.29)$$

è la matrice delle varianze e covarianze dei due rumori bianchi $B\mathbf{n}(t)$ e $D\mathbf{n}(t)$ che vengono spesso chiamati *rumore di modello* e di *osservazione* (o di *misura*). Introducendo i rumori di modello e di osservazione, il sistema lineare (1.1) può essere riscritto in modo più sintetico ponendo

$$\mathbf{v}(t) := B\mathbf{n}(t) \quad , \quad \mathbf{w}(t) := D\mathbf{n}(t) \quad , \quad (1.30)$$

e specificando ora, anziché le matrici B e D , la matrice di covarianza

$$E \left\{ \begin{bmatrix} \mathbf{v}(t) \\ \mathbf{w}(t) \end{bmatrix} [\mathbf{v}(t)' \ \mathbf{w}(t)'] \right\} = \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \quad . \quad (1.31)$$

In molte applicazioni pratiche $\{\mathbf{v}(t)\}$ e $\{\mathbf{w}(t)\}$ risultano *scorrelati* e in questi casi la notazione è effettivamente più conveniente. Questa situazione corrisponde, nelle notazioni originali, al verificarsi della condizione

$$S = BD' = 0 \quad ,$$

che si ha ad esempio quando $B = [\bar{B} \ 0]$ e $D = [0 \ \bar{D}]$ (le matrici essendo partizionate conformemente).

Nel caso di rumore di modello e di osservazione scorrelati, lo spettro di $\{\mathbf{y}(t)\}$ assume la forma

$$S(z) = C(zI - A^{-1}) Q(z^{-1} I - A')^{-1} C' + R \quad . \quad (1.32)$$

Notiamo che se i rumori di modello e di osservazione sono scorrelati ed R è *definita positiva*, certamente $S(e^{j\omega}) > 0$ per tutti gli ω in $[-\pi, \pi]$. Ne segue che *nessun fattore spettrale di $S(z)$ può avere zeri sulla circonferenza unitaria*. Questa condizione, la cui utilità si è già avuto modo di apprezzare nella teoria di Wiener-Kolmogorov, gioca un ruolo importante anche nella teoria del filtro di Kalman.

Notiamo infine che la funzione di trasferimento $W(z)$ è un fattore spettrale (in genere non quadrato) che non è necessariamente un fattore spettrale a fase minima (o canonico). Il calcolo esplicito del fattore spettrale canonico di $S(z)$ si otterrà come sottoprodotto della teoria asintotica del filtro di Kalman che esamineremo più avanti.

Ovviamente la matrice $W(z)$ è una matrice di funzioni razionali proprie della variabile z . Ogni matrice di questo tipo si può esprimere come rapporto di matrici polinomiali (vedere il Cap. 4)

$$W(z) = A(z)^{-1} C(z) = A^{*-1}(z^{-1}) C^*(z^{-1}) \quad ,$$

dove $A^*(z^{-1})$ e $C^*(z^{-1})$ sono matrici polinomiali di dimensioni $m \times m$ ed $m \times p$ rispettivamente.

Dalla teoria dei sistemi (si veda ad esempio [72]) si sa che l'asintotica stabilità di A implica che il polinomio $A(z)$ può essere scelto in modo tale da avere tutti i suoi zeri (le soluzioni di $\det A(z) = 0$) dentro il cerchio unitario. Inoltre il coefficiente di grado massimo in z può sempre essere normalizzato all'identità. In definitiva il processo $\{\mathbf{y}(t)\}$ può essere (asintoticamente) descritto mediante un modello ARMA multivariabile

$$\mathbf{y}(t) + \sum_1^{\nu} A_k \mathbf{y}(t-k) = C_0 \mathbf{n}(t) + \sum_1^{\nu} C_k \mathbf{n}(t-k) \quad . \quad (1.33)$$

Vale anche il ragionamento inverso. Nella prossima sezione riporteremo alcuni semplici procedimenti di *realizzazione* di modelli ARMA e ARMAX.

2. Alcune tecniche di Realizzazione

Per iniziare con il caso più semplice, supponiamo che il processo $\{\mathbf{y}(t)\}$ sia *scalare* e soddisfi un'equazione alle differenze del tipo

$$\mathbf{y}(t) + \sum_1^n a_k \mathbf{y}(t-k) = c_0 \mathbf{n}(t) + \sum_1^n c_k \mathbf{n}(t-k) \quad , \quad (2.1)$$

con $\{\mathbf{n}(t)\}$ rumore bianco normalizzato. Per quanto ci interessa in questa sezione non è necessario assumere alcunché sulla stazionarietà di $\{\mathbf{y}(t)\}$, né tantomeno sulla stabilità del sistema (2.1). In generale penseremo che l'equazione (2.1) descriva il processo $\{\mathbf{y}(t)\}$ a partire da un certo istante iniziale $t_0 \geq -\infty$ e sia associato a delle opportune condizioni iniziali.

Ci sono vari metodi per costruire una *realizzazione* del modello ingresso-uscita (2.1). Uno particolarmente semplice è il seguente. Sia

$$\mathbf{y}(t) = \frac{C^*(z^{-1})}{A^*(z^{-1})} \mathbf{n}(t) \quad (2.2)$$

la rappresentazione simbolica della (2.1) in forma di funzione di trasferimento con

$$C^*(z^{-1}) = c_0 + c_1 z^{-1} + \dots + c_n z^{-n} \quad (2.3a)$$

$$A^*(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_n z^{-n} \quad (2.3b)$$

Definendo la quantità

$$\boldsymbol{\xi}(t) := \frac{1}{A^*(z^{-1})} \mathbf{n}(t) \quad (2.4)$$

si può scrivere

$$\boldsymbol{\xi}(t) = \mathbf{n}(t) - a_1 z^{-1} \boldsymbol{\xi}(t) - \dots - a_n z^{-n} \boldsymbol{\xi}(t) \quad (2.5)$$

$$\mathbf{y}(t) = c_0 \boldsymbol{\xi}(t) + c_1 z^{-1} \boldsymbol{\xi}(t) + \dots + c_n z^{-n} \boldsymbol{\xi}(t) \quad (2.6)$$

Dalla (2.5) è evidente che la relazione (2.4) può essere rappresentata mediante lo schema a blocchi di fig. 2.1.

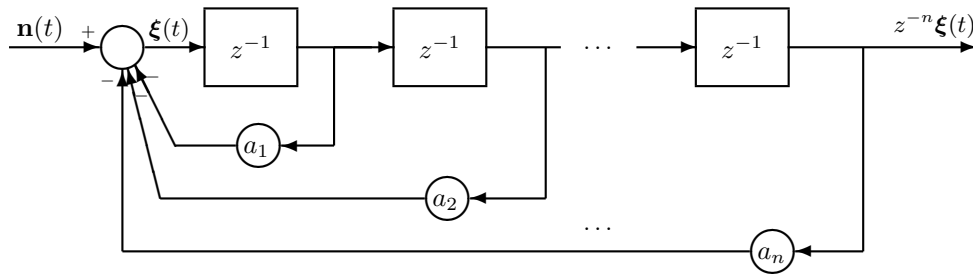


FIGURA 2.1. Realizzazione di un modello ingresso uscita.

In questo schema si possono definire delle variabili di stato scegliendo, ad esempio,

$$\mathbf{x}_1(t) = z^{-n} \boldsymbol{\xi}(t), \quad \mathbf{x}_2(t) = z^{-n+1} \boldsymbol{\xi}(t), \dots, \mathbf{x}_n(t) = z^{-1} \boldsymbol{\xi}(t) \quad (2.7)$$

che soddisfano l'equazione alle differenze

$$\begin{aligned} \mathbf{x}_1(t+1) &= \mathbf{x}_2(t) \\ \mathbf{x}_{n-1}(t+1) &= \mathbf{x}_n(t) \\ \mathbf{x}_n(t+1) &= -a_n \mathbf{x}_1(t) - \dots - a_1 \mathbf{x}_n(t) + \mathbf{n}(t) \quad (2.8) \end{aligned}$$

Usando ora la (2.6) si può rappresentare il processo $\{\mathbf{y}(t)\}$ nella forma

$$\begin{aligned} \mathbf{y}(t) &= c_0 \boldsymbol{\xi}(t) + [c_n \dots c_1] \mathbf{x}(t) \\ &= c_0 \mathbf{n}(t) - c_0 [a_n \dots a_1] \mathbf{x}(t) + [c_n \dots c_1] \mathbf{x}(t) \\ &= \left([c_n \dots c_1] - c_0 [a_n \dots a_1] \right) \mathbf{x}(t) + c_0 \mathbf{n}(t) \quad . \end{aligned}$$

Ne viene che $\{\mathbf{y}(t)\}$ può essere descritto mediante il sistema lineare di dimensione n (in forma canonica di controllo)

$$\mathbf{x}(t+1) = \begin{bmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ -a_n & & & -a_1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \mathbf{n}(t) \quad (2.9a)$$

$$\mathbf{y}(t) = [c_n - c_0 a_n, \dots, c_1 - c_0 a_1] \mathbf{x}(t) + c_0 \mathbf{n}(t) \quad (2.9b)$$

Questo modello, associato alle condizioni iniziali $\mathbf{x}(t_0)$ calcolate tramite la (2.7), è una *realizzazione* del processo $\{\mathbf{y}(t)\}$ in $t \geq t_0$. Se $\{\mathbf{y}(t)\}$ è stazionario, allora anche $\{\mathbf{x}(t)\}$ è un processo stazionario. La sua varianza può essere ricavata o risolvendo l'equazione di Lyapunov (1.24) o integrando la densità spettrale $S_{\mathbf{x}}(z)$ che è immediatamente ottenibile dalla definizione (2.7) e vale

$$\begin{aligned} S_{\mathbf{x}}(z) &= \frac{1}{A^*(z^{-1})} \begin{bmatrix} z^{-n} \\ z^{-n+1} \\ \vdots \\ z^{-1} \end{bmatrix} [z^n z^{n-1} \dots z] \frac{1}{A^*(z)} \\ &= \frac{1}{A^*(z^{-1}) A^*(z)} \begin{bmatrix} 1 & z^{-1} & \dots & z^{-n+1} \\ z & 1 & & \\ \vdots & & & \\ z^{n-1} & \dots & \dots & 1 \end{bmatrix} \quad (2.10) \end{aligned}$$

Naturalmente ci sono infinite altre realizzazioni possibili che si possono ricavare dalla (2.9) attraverso un cambio di base. Una, particolarmente conveniente, si ottiene attraverso la trasformazione di "dualità", ovvero ponendo

$$\bar{A} = A' \quad , \quad \bar{C} = B' \quad , \quad \bar{B} = C' \quad , \quad \bar{c}_0 = c_0$$

(con ovvio significato dei simboli). Si verifica immediatamente che la funzione di trasferimento (scalare) dei due modelli è la stessa, dato che

$$\begin{aligned} \bar{W}(z) &= \bar{c}_0 + \bar{c}(zI - \bar{A})^{-1} \bar{b} = \bar{c}_0 + [\bar{c}(zI - \bar{A})^{-1} \bar{b}]' \\ &= c_0 + c(zI - A)^{-1} b = W(z) \quad . \end{aligned}$$

Questa realizzazione di $W(z)$ (forma canonica di osservazione) assume allora la forma

$$\mathbf{x}(t+1) = \begin{bmatrix} 0 & -a_n \\ 1 & \vdots \\ & 1 & -a_1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} c_n - c_0 a_n \\ \vdots \\ c_1 - c_0 a_1 \end{bmatrix} \mathbf{n}(t) \quad (2.11a)$$

$$\mathbf{y}(t) = [0 \quad \dots \quad 0 \quad 1] \mathbf{x}(t) + c_0 \mathbf{n}(t) \quad (2.11b)$$

Il procedimento appena illustrato può essere naturalmente esteso al caso multivariabile, però lo schema di Figura 10.1 si generalizza in modo naturale solo se si parte da una rappresentazione matriciale fratta *destra*, del tipo $H(z) = N(z)D(z)^{-1}$, della matrice di trasferimento.

Nel caso nostro, in cui si ha a che fare con problemi di stima, sono più naturali le rappresentazioni matriciali fratte *sinistre*, del tipo

$$H(z) = A(z)^{-1} N(z) \quad , \quad \det A(z) \neq 0 \quad .$$

Come spiegato al capitolo 4, sezione 3, si può assumere senza perdita di generalità che $A(z)$ sia un polinomio monico ($m \times m$) del tipo

$$A(z) = I + A_1 z^{-1} + \dots + A_\nu z^{-\nu} \quad (2.12)$$

e, dato che ci occupiamo solo di funzioni di trasferimento *causali*, scriveremo $N(z)$ nella forma

$$N(z) = N_0 + N_1 z^{-1} + \dots + N_\nu z^{-\nu} \quad , \quad (2.13)$$

dove le N_k sono matrici reali $m \times p$.

Il termine costante $N_0 = H(\infty)$ si può estrarre da $H(z)$ ottenendo una matrice razionale strettamente propria che ha la r.p.f.

$$H_1(z) := H(z) - N_0 = A(z)^{-1} B(z) \quad , \quad (2.14)$$

dove

$$B(z) = N(z) - A(z) N_0 = (N_1 - A_1 N_0) z^{-1} + \dots + (N_\nu - A_\nu N_0) z^{-\nu} \quad . \quad (2.15)$$

Il procedimento di realizzazione che esporremo parte da r.p.f. sinistre di matrici *strettamente proprie* come la (2.14). Esso è basato sulla costruzione di uno schema a blocchi duale di quello di figura 2.1, che fa riferimento alla decomposizione dell'equazione ingresso-uscita $\mathbf{y}(t) = H_1(z) u(t)$ nella coppia di relazioni simboliche

$$\boldsymbol{\xi}(t) := B(z) u(t) \quad , \quad (2.16a)$$

$$A(z) \mathbf{y}(t) = \boldsymbol{\xi}(t) \quad . \quad (2.16b)$$

Notare che la struttura delle (2.16), e il significato della variabile intermedia $\boldsymbol{\xi}(t)$, sono in qualche senso “duali” delle (2.4)–(2.6).

Caso di uscita scalare ($m = 1$). Dato che, per $m = 1$, $B(z)$ è un vettore riga di polinomi scalari in z^{-1} (cfr. la (2.15)), esso si può scrivere nella forma

$$B(z) = [z^{-1} \quad \dots \quad z^{-\nu}] \begin{bmatrix} b'_1 \\ \vdots \\ b'_\nu \end{bmatrix} \quad , \quad (2.17)$$

dove i b'_k sono i coefficienti in \mathbb{R}^p dei relativi “monomi” (vettoriali) di grado k in z^{-1} . Da notare che se si tiene fede alla convenzione (usata sinora) di scrivere formalmente una funzione razionale propria come rapporto di polinomi di *uguale potenza massima* in z^{-1} , qualcuno degli ultimi coefficienti b'_k nella (2.17) potrebbe benissimo essere il vettore nullo.

Prendiamo allora $\nu = n$ *uguale al massimo tra i gradi dei polinomi* $A(z) \equiv a(z)$ e $B(z)$. Possiamo allora riscrivere le (2.16) nella forma

$$\begin{aligned} \mathbf{y}(t) &= (-a_1 z^{-1} \dots - a_n z^{-n}) \mathbf{y}(t) + \boldsymbol{\xi}(t) \\ &= z^{-1} [-a_1 \mathbf{y}(t) + b'_1 u(t)] + \dots + z^{-n} [-a_n \mathbf{y}(t) + b'_n u(t)] \end{aligned} \quad (2.18)$$

dalla quale si ottiene una rappresentazione di $\mathbf{y}(t)$ mediante uno schema a blocchi a retroazione multipla del tipo

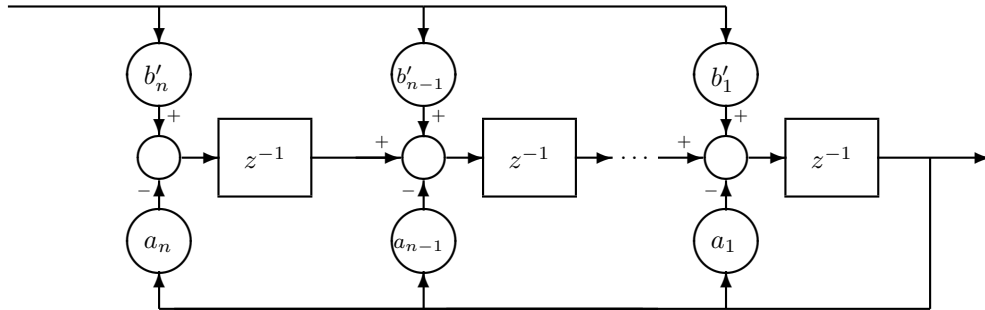


FIGURA 2.2. Realizzazione di una funzione di trasferimento con uscita scalare.

dal quale si possono scrivere le equazioni di stato per ispezione. Prendendo come variabili di stato le uscite degli elementi di ritardo, ad esempio definendo

$$\begin{aligned} \mathbf{x}_1(t) &= \mathbf{y}(t) = z^{-1} [-a_1 \mathbf{x}_1(t) + b'_1 u(t) + \mathbf{x}_2(t)] \\ \mathbf{x}_2(t) &= z^{-1} [-a_2 \mathbf{x}_1(t) + b'_2 u(t) + \mathbf{x}_3(t)] \\ &\vdots \\ \mathbf{x}_n(t) &= z^{-1} [-a_n \mathbf{x}_1(t) + b'_n u(t)] \quad , \end{aligned} \quad (2.19)$$

si ottiene la realizzazione in forma compagna

$$\mathbf{x}(t+1) = \begin{bmatrix} -a_1 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ \vdots & & & & 1 \\ -a_n & 0 & \dots & \dots & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} b'_1 \\ \vdots \\ b'_n \end{bmatrix} u(t) \quad , \quad (2.20a)$$

$$\mathbf{y}(t) = [1 \ 0 \ \dots \ 0] \mathbf{x}(t) \quad . \quad (2.20b)$$

Si vede subito che questa realizzazione è osservabile.

PROPOSIZIONE 9.6. *La realizzazione (2.20) è raggiungibile se e solo se $A(z)$ e $B(z)$ sono coprime (i.e., non ammettono fattori polinomiali comuni).*

DIMOSTRAZIONE. Supponiamo che la (2.20) non sia una rappresentazione completamente raggiungibile. Esiste allora un autovettore sinistro a , della matrice A , che è ortogonale alle colonne di B .

Supponiamo $a_n \neq 0$. Si verifica allora facilmente che ogni autovettore sinistro della matrice in forma compagna (A) che compare in (2.20) è della forma

$$a' \cong [\lambda^{-1}, \lambda^{-2}, \dots, \lambda^{-n}] \quad , \quad (2.21)$$

dove λ è un autovalore di A , ovvero una radice dell'equazione $\det(\lambda I - A) = \lambda^n + a_1 \lambda^{n-1} + \dots + a_n = 0$. Dato che $a_n \neq 0$ si ha $\lambda \neq 0$. Segue dalla (2.17) che a è ortogonale alle colonne di B se e solo se

$$a' B = B(\lambda) = [0 \ \dots \ 0] \quad ,$$

cioè tutti gli elementi di $B(\lambda)$ hanno in questo caso uno zero in $z = \lambda$. Se invece $a_n = a_{n-1} \dots = a_{k+1} = 0$, $a_k \neq 0$, si vede che ci sono al più k autovettori che hanno ancora la forma (2.21) con λ autovalore di A diverso da zero, mentre a $\lambda = 0$ corrisponde l'autovettore $v = [0 \dots 0 \ 1]'$. Ora, per le convenzioni fatte in precedenza, se $a_n = 0$ si deve necessariamente avere $b'_n \neq 0$ (cfr. la definizione di n), per cui gli unici autovettori che possono essere ortogonali alle colonne di B devono avere la forma (2.21) (ovviamente con $\lambda \neq 0$). In ogni caso quindi la non raggiungibilità implica la presenza di fattori comuni tra $A(z)$ e $B(z)$.

Viceversa, se $A(z)$ e $B(z)$ ammettono fattori comuni, è possibile, eliminandoli, ottenere una realizzazione del tipo (2.20) di ordine inferiore. In questo caso quindi la realizzazione non sarebbe minima. \square

Caso generale ($m > 1$). È abbastanza semplice generalizzare la costruzione appena vista al caso di più uscite. Per ottenere realizzazioni di *dimensione minima* occorre però tener conto esplicitamente dei *gradi riga* delle matrici $A(z)$ e $B(z)$.

Siano k_1, \dots, k_m e j_1, \dots, j_m i gradi riga di $A(z)$ e $B(z)$. Ricordiamo che nel nostro caso (in cui si ha a che fare con polinomi in z^{-1}) k_j è semplicemente la massima potenza di (z^{-1}) cui corrisponde un coefficiente non nullo nella riga i -sima.

Definiamo ancora

$$n_i = \max\{k_i, j_i\} \quad , \quad i = 1, \dots, m \quad . \quad (2.22)$$

Chiameremo n_i *ordine parziale* della componente i -sima $z_i(t)$ di $\mathbf{y}(t)$. Definendo $n := n_1 + \dots + n_m$, i vettori n -dimensionali,

$$\begin{aligned} \pi_1(z) &= [z^{-1} \dots z^{-n_1}, 0 \dots 0, 0 \dots 0] \\ \pi_m(z) &= [0 \dots 0, \dots, z^{-1} \dots z^{-n_m}] \end{aligned}$$

e la matrice $m \times n$,

$$\Pi(z) = \begin{bmatrix} \pi_1(z) \\ \vdots \\ \pi_m(z) \end{bmatrix} \quad , \quad (2.23)$$

si possono rappresentare $A(z)$ e $B(z)$ nella forma

$$A(z) = I + \Pi(z) \bar{A} \quad , \quad B(z) = \Pi(z) B \quad , \quad (2.24)$$

dove $\bar{A} \in \mathbb{R}^{n \times m}$ e $B \in \mathbb{R}^{n \times p}$ sono matrici di coefficienti che sarà utile scrivere in forme partizionate di tipo diverso, i.e.,

$$\bar{A} = \begin{bmatrix} \bar{A}_1 \\ \vdots \\ \bar{A}_m \end{bmatrix} = [a_1 \quad \dots \quad a_m] \quad , \quad \bar{A}_i \in \mathbb{R}^{n_i \times m} \quad , \quad a_j \in \mathbb{R}^n \quad , \quad (2.25a)$$

$$B = \begin{bmatrix} B_1 \\ \vdots \\ B_m \end{bmatrix} = [b_1 \quad \dots \quad b_p] \quad , \quad B_i \in \mathbb{R}^{n_i \times p} \quad , \quad b_j \in \mathbb{R}^n \quad . \quad (2.25b)$$

Da notare che $[z^{-1}, \dots, z^{-n_i}] \bar{A}_i$ è essenzialmente l' i -sima riga del polinomio $A(z)$, ecc. Le matrici \bar{A} e B si ottengono quindi per ispezione una volta noti i polinomi $A(z)$ e $B(z)$.

Come fatto per il caso scalare, usando la (2.24) si può ora riscrivere la relazione ingresso-uscita $\mathbf{y}(t) = A^{-1}(z) B(z) u(t)$ nella forma simbolica

$$\mathbf{y}(t) = \Pi(z) [-\bar{A} \mathbf{y}(t) + B u(t)] \quad , \quad (2.26)$$

ovvero

$$\begin{aligned} z_1(t) &= [z^{-1} \quad \dots \quad z^{-n_1}] [-\bar{A}_1 \mathbf{y}(t) + B_1 u(t)] \quad , \\ &\vdots \\ z_m(t) &= [z^{-1} \quad \dots \quad z^{-n_m}] [-\bar{A}_m \mathbf{y}(t) + B_m u(t)] \quad , \end{aligned} \quad (2.27)$$

che corrisponde ad una rappresentazione di $\mathbf{y}(t)$ mediante uno schema a retroazione multipla ottenibile mediante l'interconnessione di m schemi del tipo di Figura 10.2, uno per ciascuna componente scalare $z_1(t), \dots, z_m(t)$. La rappresentazione grafica di questa interconnessione è complicata e rinunciamo qui a tentare di darne una.

Il vettore di stato dell'interconnessione è composto di m sottovettori

$$\mathbf{x}(t)' = [\mathbf{x}_1(t)' \dots \mathbf{x}_m(t)'] \in \mathbb{R} \quad ,$$

di dimensioni n_1, \dots, n_m , pari al numero di elementi di ritardo necessari a rappresentare ciascuna componente scalare di $\mathbf{y}(t)$. Scriveremo

$$\mathbf{x}_i(t) := \begin{bmatrix} \mathbf{x}_{i1}(t) \\ \vdots \\ \mathbf{x}_{in_i}(t) \end{bmatrix} \in \mathbb{R}^{n_i} \quad , \quad i = 1, \dots, m \quad .$$

Con queste notazioni, scegliendo ad esempio $\mathbf{x}_{i1}(t) = z_i(t)$, si ottengono dalle (2.27) m equazioni di stato accoppiate della forma

$$\begin{cases} \mathbf{x}_i(t+1) = \begin{bmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & & & 0 \end{bmatrix} \mathbf{x}_i(t) - \bar{A}_i \begin{bmatrix} \mathbf{x}_{i1}(t) \\ \vdots \\ \mathbf{x}_{im_1}(t) \end{bmatrix} + B_i u(t) \\ \mathbf{y}(t) := \begin{bmatrix} 1 & 0 & \dots & 0 & & & \\ & 1 & 0 & \dots & 0 & & \\ & & 1 & 0 & \dots & 0 & \\ & & & 1 & 0 & \dots & 0 \\ & & & & 1 & 0 & \dots & 0 \end{bmatrix} \mathbf{x}(t) \quad . \end{cases} \quad i = 1, \dots, m \quad (2.28)$$

Indicando le variabili di stato che vengono riportate nel canale di reazione col simbolo $\mathbf{x}_r(t)$, col che, con la scelta appena fatta,

$$\mathbf{x}_r(t) := \begin{bmatrix} \mathbf{x}_{i1}(t) \\ \vdots \\ \mathbf{x}_{m1}(t) \end{bmatrix} = \mathbf{y}(t) \quad ,$$

e le matrici di traslazione temporale con

$$J_{n_i} := \begin{bmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & & & 0 \end{bmatrix} \quad (n_i \times n_i) \quad ,$$

si possono compattare un pochino le notazioni. La (2.28) si può riscrivere

$$\begin{cases} \mathbf{x}(t+1) = \begin{bmatrix} J_{n_1} & & \\ & \ddots & \\ & & J_{n_m} \end{bmatrix} \mathbf{x}(t) - \bar{A} \mathbf{x}_r(t) + B u(t) \\ \mathbf{y}(t) = C \mathbf{x}(t) \end{cases} \quad (2.29)$$

dove \bar{A} e B sono proprio le matrici dei coefficienti introdotte in (2.24). La matrice C è definita in (2.28).

Come si vede, gli elementi diversi da zero o uno nella matrice A della realizzazione (2.29) sono solo quelli delle colonne di indice

$$1, n_1 + 1, n_1 + n_2 + 1, \dots$$

Queste colonne coincidono in effetti con i vettori colonna a_1, a_2, \dots, a_m di \bar{A} .

ESEMPIO 9.1 (Realizzazione di un modello ARMA). Vogliamo costruire una realizzazione del processo $\{\mathbf{y}(t)\}$ descritto dal modello ARMA vettoriale

$$\begin{bmatrix} (1+z^{-1})^2(1+1/2z^{-1})^2 & 1+1/2z^{-1} \\ 0 & (1+1/2z^{-1})^2 \end{bmatrix} \mathbf{y}(t) = \begin{bmatrix} 1 & -z^{-2} \\ 1-z^{-1} & 1-z^{-1} \end{bmatrix} \mathbf{w}(t),$$

al quale corrispondono i polinomi matriciali

$$\begin{aligned} A(z) &= I + \begin{bmatrix} 3 & 1/2 \\ 0 & 1 \end{bmatrix} z^{-1} + \begin{bmatrix} 13/4 & 0 \\ 0 & 1/4 \end{bmatrix} z^{-2} \\ &\quad + \begin{bmatrix} 3/2 & 0 \\ 0 & 0 \end{bmatrix} z^{-3} + \begin{bmatrix} 1/4 & 0 \\ 0 & 0 \end{bmatrix} z^{-4} \end{aligned}$$

$$N(z) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -1 & -1 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} z^{-2}.$$

Seguendo il procedimento appena visto calcoliamo innanzitutto $B(z)$. I coefficienti sono

$$\begin{aligned} B_1 &= \begin{bmatrix} 0 & 0 \\ -1 & -1 \end{bmatrix} - \begin{bmatrix} 3 & 1/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -7/2 & -1/2 \\ -2 & -2 \end{bmatrix} \\ B_2 &= \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 13/4 & 0 \\ 0 & 1/4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -13/4 & -1 \\ -1/4 & -1/4 \end{bmatrix} \\ B_3 &= - \begin{bmatrix} 3/2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -3/2 & 0 \\ 0 & 0 \end{bmatrix} \\ B_4 &= \begin{bmatrix} -1/4 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

I gradi riga di $A(z)$ sono

$$k_1 = 4, \quad k_2 = 2,$$

uguali a quelli di $B(z)$. Pertanto $n_1 = 4$, $n_2 = 2$ e troveremo una realizzazione di ordine 6. Questo ordine coincide col grado del determinante di $A(z)$ e si potrebbe dimostrare (vedi ad esempio [Forney]) che questo è un fatto generale per matrici polinomiali moniche in z^{-1} .

Si individuano adesso i coefficienti \bar{A} e B nella (2.24),

$$\Pi(z) \bar{A} = \begin{bmatrix} z^{-1} & z^{-2} & z^{-3} & z^{-4} & 0 & 0 \\ 0 & 0 & 0 & 0 & z^{-1} & z^{-2} \end{bmatrix} \begin{bmatrix} 3 & 1/2 \\ 13/4 & 0 \\ 3/2 & 0 \\ 1/4 & 0 \\ 0 & 1 \\ 0 & 1/4 \end{bmatrix}$$

$$\Pi(z) B = \begin{bmatrix} z^{-1} & z^{-2} & z^{-3} & z^{-4} & 0 & 0 \\ 0 & 0 & 0 & 0 & z^{-1} & z^{-2} \end{bmatrix} \begin{bmatrix} -7/2 & -1/2 \\ -13/4 & -1 \\ -3/2 & 0 \\ -1/4 & 0 \\ -2 & -2 \\ -1/4 & -1/4 \end{bmatrix}$$

e si può scrivere così la realizzazione cercata

$$\mathbf{x}(t+1) = \begin{bmatrix} 3 & 1 & 0 & 0 & 1/2 & 0 \\ 13/4 & 0 & 1 & 0 & 0 & 0 \\ 3/2 & 0 & 0 & 1 & 0 & 0 \\ 1/4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1/4 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} -7/2 & -1/2 \\ -13/4 & -1 \\ -3/2 & 0 \\ -1/4 & 0 \\ -2 & -2 \\ -1/4 & -1/4 \end{bmatrix} \mathbf{w}(t)$$

$$\mathbf{y}(t) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \mathbf{w}(t) \quad .$$

◇

Esaminando la struttura di C e la forma “multicompagna” di A si vede facilmente che la realizzazione (2.29) è *sempre osservabile*. La raggiungibilità dipende ancora dalla presenza o meno di fattori destri non unimodulari comuni ad $A(z)$ e $B(z)$. L’enunciato seguente stabilisce in modo preciso questo fatto. La sua dimostrazione (che qui non verrà riportata) è una semplice (anche se un po’ noiosa) generalizzazione di quella della Proposizione 9.6.

PROPOSIZIONE 9.7. *Se gli ordini parziali n_i sono definiti come in (2.22), la realizzazione (2.29) è raggiungibile se e solo se i polinomi $A(z)$ e $B(z)$ sono coprimi a destra.*

Realizzazioni bilanciate. È bene mettere in guardia il lettore sul fatto che le realizzazioni in forma compagna o multicompagna che sono state introdotte in questo capitolo hanno *pessime proprietà numeriche*, nel senso che importanti quantità strutturali del sistema come ad esempio gli autovalori, gli indici di osservabilità e raggiungibilità, l’ordine, i valori singolari ecc. dipendono in modo estremamente sensibile dai coefficienti a_j e b_j delle matrici A e B .

Questo fatto si può in sostanza far risalire alla ben nota estrema sensibilità con cui le radici di un polinomio dipendono dai coefficienti, in particolare dagli *ultimi* coefficienti, quelli di grado zero, uno ecc. nell’indeterminata. Questa “ipersensibilità” a variazioni parametriche spiega il fatto (spesso sperimentalmente rilevabile e

fonte di grossi dispiaceri) che i modelli ARMA sono strumenti numericamente assai poco affidabili.

Per questo motivo si sono studiate realizzazioni di stato particolari che hanno la proprietà di avere bassa sensitività rispetto a variazioni parametriche. In particolare nel testo di Roberts e Mullis [124, Cap.9] è presentata una analisi molto esauriente del problema della costruzione di realizzazioni di una funzione di trasferimento data che hanno sensitività minima rispetto agli errori nella risposta dovuti a imprecisione nella rappresentazione dei coefficienti del modello (errori numerici di troncamento etc.). Questi errori si possono distinguere in errori nella risposta ingresso-stato e errori indotti sull'uscita nella risposta stato-uscita. Si mostra che tentare di diminuire gli errori di un tipo porta necessariamente ad esaltare quelli dell'altro. Per questo motivo la soluzione ottimale per ridurre al massimo entrambi le cause d'errore è di costruire realizzazioni che sono *bilanciate* cioè hanno risposte ingresso-stato e stato-uscita che hanno lo stesso "guadagno" (il termine essendo definito in modo opportuno). Queste particolari realizzazioni sono state successivamente pubblicizzate da B. C. Moore [104] nella letteratura controllistica e le loro notevoli proprietà sono state studiate poi da vari autori.

Supporremo dato un modello di tipo (1.1) in cui le matrici (A, B, C, D) sono costanti, gli autovalori di A sono strettamente dentro il cerchio unitario, (A, B) è una coppia raggiungibile e (C, A) è osservabile. In particolare i processi di stato e uscita sono asintoticamente stazionari e le matrici

$$\Pi := \sum_0^{+\infty} A^k B B' (A')^k, \quad \Omega := \sum_0^{+\infty} (A')^k C' C A^k \quad (2.30)$$

che chiameremo, rispettivamente, Gramiani (estesi) di raggiungibilità e di osservabilità, sono le soluzioni definite positive delle equazioni di Lyapunov duali,

$$\Pi = A \Pi A' + B B' \quad (2.31)$$

$$\Omega = A' \Omega A + C' C \quad (2.32)$$

Notiamo che per effetto di un cambio di base in (1.1) definito da una matrice non singolare T ,

$$\hat{A} = T^{-1} A T, \quad \hat{B} = T^{-1} B, \quad \hat{C} = C T$$

i due gramiani si trasformano secondo le relazioni

$$\hat{\Pi} = T^{-1} \Pi T^{-\top}, \quad \hat{\Omega} = T' \Omega T \quad (2.33)$$

dove l'apice $^{-\top}$ indica l'inversa trasposta.

DEFINIZIONE 9.1. *Un modello lineare del tipo (1.1) in cui valgono le ipotesi enunciate sopra, si dice in forma bilanciata se le due matrici gramiane Π e Ω sono diagonali e uguali.*

Ogni modello (1.1) per cui valgono le ipotesi enunciate, si può trasformare, mediante un cambiamento di base, in uno bilanciato.

Per far questo si può seguire il seguente procedimento.

ALGORITMO 9.1 (Trasformazione in Forma Bilanciata). *Una volta calcolate le soluzioni Π e Ω delle equazioni di Lyapunov (2.31), (2.32),*

- (1) Si calcola la decomposizione in autovalori e autovettori della matrice (simmetrica e definita positiva) Ω ,

$$\Omega = U\Lambda_o U'$$

dove Λ_o è la matrice diagonale degli autovalori di Ω .

- (2) Si introduce un primo cambio di base definito dalla matrice $T_1 := \Lambda_o^{-1/2} U'$ in modo da rendere $\hat{\Omega} = I$ e si calcola

$$\hat{\Pi} = U\Lambda_o^{1/2}\Pi\Lambda_o^{1/2}U'$$

- (3) Si calcola la decomposizione in autovalori e autovettori della matrice $\hat{\Pi}$,

$$\hat{\Pi} = V\Lambda^2 V'$$

dove adesso la matrice diagonale degli autovalori si denota col simbolo Λ^2 (il motivo di questa notazione diversa sarà chiaro tra un attimo),

- (4) Si introduce un secondo cambio di base definito dalla matrice $T_2 := V\Lambda^{1/2}$ in modo da rendere $\bar{\Pi} := T_2^{-1}\hat{\Pi}T_2^{-\top} = \Lambda$, diagonale.

Con questo cambio di base

$$\bar{\Omega} = T_2'\hat{\Omega}T_2 = \Lambda^{1/2}V'IV\Lambda^{1/2} = \Lambda$$

e entrambi i gramiani sono diagonali e uguali, i.e. $\bar{\Pi} = \bar{\Omega} = \Lambda$.

Notare che le matrici usate per il cambio di base sono prodotti di matrici diagonali per matrici ortogonali. Il calcolo dell'inversa di una matrice di questo tipo è molto semplice.

Si può dimostrare (ma qui ci esimeremo dal farlo) che la matrice diagonale Λ è un invariante del sistema lineare (1.1), cioè non dipende dalla scelta di base. In questo senso, la forma bilanciata è una specie di *forma canonica* del modello (1.1) (anche se questo non è letteralmente vero e servirebbe ancora fissare alcuni parametri strutturali del modello per renderla veramente tale [109]). Per questo motivo, si sarebbe arrivati allo stesso risultato seguendo l'algoritmo duale che parte dalla decomposizione in autovalori-autovettori di Π anziché da quella di Ω . Il significato di Λ è illustrato nel seguente teorema.

TEOREMA 9.1. *La matrice diagonale Λ è formata dai valori singolari non nulli della matrice di Hankel*

$$\mathbb{H} := \begin{bmatrix} CB & CAB & CA^2B & \dots \\ CAB & CA^2B & CA^3B & \dots \\ CA^2B & CA^3B & \dots & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} \quad (2.34)$$

associata al sistema lineare (1.1).

La decomposizione ai valori singolari (SVD) è richiamata in Appendice A. La realizzazione bilanciata è un punto di partenza naturale per costruire *modelli approssimati di ordine ridotto* [56, 113] e anche per questo motivo si usa molto nel filtraggio di segnali e in identificazione.

Campionamento di processi a tempo continuo. In genere i segnali con cui si ha a che fare nell'ingegneria sono segnali ottenuti per campionamento di segnali continui. Spesso il dispositivo di acquisizione e misura esegue questo campionamento automaticamente, ma la dinamica del segnale è pur sempre una dinamica a tempo continuo e per ottenere modelli di stato discreti è spesso necessario campionare i corrispondenti modelli continui. Descriveremo qui come ottenere la descrizione statistica dei corrispondenti modelli campionati. In questo paragrafo il simbolo t denoterà la variabile temporale continua ($t \in \mathbb{R}$).

Sia $\{\mathbf{y}(t)\}$ un processo stazionario a tempo continuo di densità spettrale $S(j\omega)$ nota. Supporremo per semplicità $\{\mathbf{y}(t)\}$ scalare.

Qualche volta gli spettri ottenuti da analizzatori di spettro vengono forniti in forma analitica come funzioni razionali di $j\omega$. In genere comunque possiamo supporre che $S(j\omega)$ sia approssimabile con una funzione densità spettrale razionale.

Le proprietà analitiche di uno spettro razionale sono analoghe a quelle degli spettri discreti che abbiamo discusso nei capitoli precedenti, una volta che si sostituisca il cerchio unità con l'asse immaginario del piano complesso. La teoria della fattorizzazione spettrale razionale, esposta nel Capitolo 5 si adatta facilmente a questo caso. In particolare si mostra che $S(j\omega)$ è una funzione reale e pari e dipende in realtà solo da $(j\omega)^2 = -\omega^2$. Pertanto uno spettro razionale scalare, pensato come funzione della variabile complessa s , ha poli e zeri simmetrici rispetto all'asse immaginario e si può sempre esprimere come

$$S(j\omega) = \sigma^2 |W(j\omega)|^2 = \sigma^2 W(j\omega) W(j\omega)^* \quad (2.35)$$

dove il fattore spettrale W (che abbiamo normalizzato in modo tale che $W(0) = 1$) è una funzione reale razionale che può sempre essere presa con tutti i poli (e eventualmente anche gli zeri) a sinistra dell'asse immaginario. Per motivi di integrabilità dello spettro, W deve avere grado del numeratore strettamente inferiore a quello del denominatore.

Una volta calcolata la fattorizzazione spettrale, si può esprimere il processo $\{\mathbf{y}(t)\}$ come uscita di un filtro lineare e stabile con ingresso un opportuno rumore bianco continuo

$$\begin{array}{ccc} \xrightarrow{\mathbf{v}(t)} & \boxed{W(s)} & \xrightarrow{\mathbf{y}(t)} \end{array}$$

e dalla espressione analitica di W si può scrivere per ispezione un modello di stato continuo del tipo

$$\begin{aligned} \dot{\mathbf{x}}(t) &= A \mathbf{x}(t) + b \mathbf{v}(t) \\ \mathbf{y}(t) &= c \mathbf{x}(t) \end{aligned}$$

dove $\{\mathbf{v}(t)\}$ è un rumore bianco continuo a intensità σ^2 .

Campionando questo modello con periodo T e scrivendo per semplicità $\mathbf{y}(k) := \mathbf{y}(kT)$, si ottiene

$$\begin{aligned} \mathbf{x}(k+1) &= e^{AT} \mathbf{x}(k) + \mathbf{w}(k) \\ \mathbf{y}(k) &= c \mathbf{x}(k) \end{aligned}$$

dove

$$\mathbf{w}(k) = \int_{kT}^{(k+1)T} e^{A((k+1)T-\tau)} b d\mathbf{v}(\tau) = \int_0^T e^{A\tau} b d\mathbf{v}((k+1)T - \tau) \quad .$$

Dato che $\{\mathbf{v}((k+1)T - \tau)\}$ è ancora rumore bianco a tempo continuo di intensità σ^2 , è chiaro che la successione $\{\mathbf{w}(k)\}$ è un rumore bianco stazionario. La sua varianza, Q_T , è data dalla formula

$$Q_T = \sigma^2 \int_0^T e^{A\tau} b b' e^{A'\tau} d\tau \quad .$$

Questa formula è complicata da usare in pratica e non è nemmeno immediato intravedere delle approssimazioni che siano semplici e sensate. Per questo motivo qui di seguito ci dilungheremo un pò a descrivere un possibile schema di calcolo per Q_T .

Notiamo innanzitutto che Q_T si ottiene risolvendo l'equazione di Lyapunov continua

$$\dot{Q} = A Q + Q A' + \sigma^2 b b'$$

con condizione iniziale nulla $Q(0) = 0$, valutando $Q_T = Q(T)$. L'equazione di Lyapunov è in realtà un'equazione lineare che si può riscrivere come un'ordinaria equazione differenziale lineare per il vettore n^2 -dimensionale $q := \text{vec}(Q)$ definito giustapponendo le colonne di Q nel loro ordinamento naturale. Dopo questa operazione di vettorizzazione l'equazione di Lyapunov ha la forma seguente

$$\dot{q} = A q + f, \quad q(0) = 0$$

dove $f := \sigma^2 \text{vec}(b b')$ e naturalmente una volta calcolata la soluzione $q(T)$, Q_T si ottiene in forma vettoriale come $\text{vec} Q_T = q(T)$. Per calcolare rapidamente $q(T)$ si può notare che

$$q(T) = \int_0^T e^{A\tau} f d\tau$$

e utilizzare per esempio la funzione MATLAB "c2d" di J. Little nel *Matlab Control System Toolbox* che discretizza modelli di stato lineari e continui.

PROBLEMA 9.1. *Dato che Q è simmetrica, la vettorizzazione si può fare in modo efficiente usando solo i sottovettori colonna di Q che stanno sopra la diagonale principale, riducendo così la dimensione del sistema lineare equivalente a $n(n+1)/2$ anzichè n^2 .*

Trovare la struttura di A in questo caso e scrivere una funzione MATLAB che trasforma A in A . ◇

3. Modelli di stato non stazionari

Quanto esposto finora mette in chiaro, ci sembra, con sufficiente evidenza, che un processo stazionario di densità spettrale razionale può sempre essere modellato nella forma (1.1). Dato che le equazioni di stato (2.9) e (2.11) si scrivono *per ispezione* a partire dal modello ARMA (non necessariamente d'innovazione) che descrive il processo, si vede che per i due tipi di modelli a cui fanno riferimento i due approcci, rispettivamente di Wiener-Kolmogorov e di Kalman, nel caso di segnali asintoticamente stazionari, possono essere riguardati come sostanzialmente equivalenti.

In realtà però il modello lineare a dimensione finita (1.1) è uno strumento assai più flessibile per descrivere i segnali che si incontrano nelle applicazioni. In molte applicazioni classiche del filtro di Kalman i segnali in gioco, ed in specie il segnale messaggio, sono raramente stazionari, però sono spesso facilmente descrivibili in modo diretto mediante equazioni dinamiche di stato del tipo (1.1) piuttosto che

attraverso parametri statistici come la funzione di covarianza o la densità spettrale di potenza.

Tracking e navigazione automatica. Una delle ragioni principali del successo del filtro di Kalman negli anni 1960-70 è stata la sua immediata applicabilità a problemi di inseguimento (*tracking* in inglese) e di navigazione automatica, specialmente in campo aeronautico e aereospaziale. Questi due problemi sono simili, in un certo senso duali uno dell'altro, e consistono nello stimare in linea le variabili di stato, tipicamente posizione e velocità, e eventualmente assetto nel caso di corpi rigidi, di un veicolo di dinamica nota, basandosi su misure in istanti discreti di certi parametri di posizione rispetto ad una o più stazioni di misura, fisse o mobili. Ad esempio, un tipico problema di inseguimento è stimare posizione, assetto e velocità di un aereomobile a partire da misure radar della sua distanza e velocità relativa rispetto ad una stazione fissa a terra. Questa ricostruzione potrebbe ad esempio servire per fornire automaticamente i comandi durante la fase di atterraggio oppure, eventualmente, per scopi meno "pacifici" facilmente immaginabili. Un altro problema di tracking è stimare la posizione e l'assetto di un satellite in orbita rispetto ad un altro veicolo che voglia eseguire una manovra di aggancio. La differenza tra tracking e navigazione è che nel primo problema la ricostruzione deve essere fatta dalla stazione di misura, senza conoscere i comandi e le forze agenti sul veicolo mentre nel problema di navigazione la ricostruzione dev'essere fatta a bordo, usando misure e sensori che si trovano a bordo del veicolo e che forniscono distanze e velocità rispetto a riferimenti fissi e noti, esterni al sistema.

Denotando ancora la variabile temporale col simbolo $t \in \mathbb{R}$, si può in generale supporre che il sistema veicolo-misura possa essere descritto per mezzo di un modello del tipo:

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t)) + \mathbf{v}(t) \quad (3.1)$$

$$\mathbf{y}(t_k) = h(\mathbf{x}(t_k)) + \mathbf{w}(t_k) \quad (3.2)$$

dove la prima equazione differenziale descrive la dinamica del veicolo soggetto a ingressi (comandi) noti $\mathbf{u}(t)$ e a ingressi aleatori non misurabili (noti solo in senso statistico) $\mathbf{v}(t)$. In problemi di tracking di solito $\mathbf{u}(t)$ non è noto e il suo effetto viene conglobato in un termine di rumore additivo come $\mathbf{v}(t)$. In casi semplici si suppone direttamente che $\mathbf{v}(t)$ sia bianco, ma anche nel caso più generale di disturbi "colorati" a spettro approssimativamente noto, si può ricorrere a tecniche di realizzazione del disturbo come quella vista nella sezione precedente e aumentare lo stato del sistema in modo da ricondursi ancora al caso di ingresso aleatorio bianco.

L'equazione di osservazione (3.2) descrive le misure di posizione e velocità effettuate sul veicolo dal sistema di rilevamento. Per motivi fisici le misure sono sempre acquisite in istanti discreti $\{t_k\}$. Anche in questo caso ci si può ricondurre al caso di rumore additivo di misura bianco, eventualmente introducendo una opportuna rappresentazione di stato del disturbo reale sovrapposto all'uscita ideale misurata $h(\mathbf{x}(t_k))$.

In queste applicazioni il modello (3.1), (3.2) si ottiene dalla fisica del sistema, generalmente usando le leggi della meccanica e le specifiche del dispositivo di misura. In problemi di tracking bisognerebbe in realtà supporre di conoscere troppi parametri del veicolo da inseguire (parametri di inerzia, attriti, forze agenti etc.) per ottenere un modello dinamico realistico e si ricorre a modelli semplificati in cui

si usano solo le leggi della cinematica e si concentra tutta la parte “dinamica” (gli effetti delle forze agenti) nel termine di rumore additivo. Questi modelli si chiamano *modelli cinematici* e hanno una struttura lineare (a differenza dei modelli dinamici veri e propri basati sulle leggi di Newton che si usano nei problemi di navigazione). Esaminiamo un semplice esempio in cui supponiamo per semplicità che il veicolo da inseguire si possa schematizzare come una massa puntiforme e il moto avvenga sul piano $\{x, y\}$.

Supponendo che il moto “nominale” del bersaglio sia un moto rettilineo uniforme a velocità v e discretizzando il sistema lineare

$$\frac{d}{dt} \begin{bmatrix} x \\ \dot{x} \\ y \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ y \\ \dot{y} \end{bmatrix} + \begin{bmatrix} 0 \\ f_x \\ 0 \\ f_y \end{bmatrix} \quad (3.3)$$

dove f_x e f_y indicano le componenti delle forze agenti, si trova un'equazione di stato del tipo

$$\mathbf{x}(t+1) = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}(t) + \mathbf{v}(t) \quad (3.4)$$

dove $\mathbf{v}(t)$ rappresenta il vettore discretizzato delle forze agenti. Questo processo vettoriale si suppone in genere bianco e a media zero (il che può essere ragionevole se il moto del bersaglio è abbastanza irregolare). Naturalmente, a questo punto l'essenza della modellizzazione sarà fornire una espressione ragionevole della covarianza del rumore di modello Q , che tenga conto dell'informazione a priori sul moto del bersaglio.

Supponiamo di indicare con (r, θ) le coordinate polari del veicolo rispetto alla stazione fissa di misura e di avere un'idea delle deviazioni standard σ_v, σ_ω , della velocità di variazione dei processi (discretizzati) delle velocità (scalare) longitudinale $v = \dot{r}$ e angolare $\omega = \dot{\theta}$ che supponiamo ancora a media nulla, bianchi e a incrementi scorrelati. Calcolando lo Jacobiano della trasformazione $\dot{x} = v \cos \theta, \dot{y} = v \sin \theta$,

$$\begin{bmatrix} d\dot{x} \\ d\dot{y} \end{bmatrix} = \begin{bmatrix} \cos \theta & -v \sin \theta \\ \sin \theta & v \cos \theta \end{bmatrix} \begin{bmatrix} dv \\ d\omega \end{bmatrix},$$

identificando $d\dot{x}(t)$ con $\mathbf{v}_2(t)$ e $d\dot{y}(t)$ con $\mathbf{v}_4(t)$ e prendendo la varianza delle variazioni, si trovano le espressioni

$$\begin{aligned} \sigma_{\mathbf{v}_2} &= \sigma_v^2 \sin^2 \theta + \sigma_\omega^2 v^2 \cos^2 \theta \\ \sigma_{\mathbf{v}_4} &= \sigma_v^2 \cos^2 \theta + \sigma_\omega^2 v^2 \sin^2 \theta \\ \sigma_{\mathbf{v}_2 \mathbf{v}_4} &= (\sigma_v^2 - \sigma_\omega^2 v^2) \sin \theta \cos \theta \end{aligned} \quad (3.5)$$

Queste espressioni evidentemente dipendono dalle coordinate correnti, ma $v(t)^2 = [\dot{x}(t)^2 + \dot{y}(t)^2]$ e $\theta(t) = \text{Arg}(x(t), y(t))$ possono essere approssimate, almeno per brevi periodi, con dei valori costanti ricavati da stime precedenti. I termini σ_v e σ_ω sono i valori efficaci delle variazioni di velocità lineare e angolare in un intervallo di campionamento T e si suppongono noti a priori. Le altre due componenti del rumore di modello $\mathbf{v}_1, \mathbf{v}_3$, si possono stimare impiegando un'approssimazione del tipo,

$$\mathbf{x}(t+1) \simeq \mathbf{x}(t) + \frac{T}{2}(\dot{\mathbf{x}}(t) + \dot{\mathbf{x}}(t+1)) = \mathbf{x}(t) + T(\dot{\mathbf{x}}(t) + \frac{1}{2}\mathbf{v}_2(t))$$

relazione che è corretta se la velocità media tra i due istanti di campionamento è $\frac{1}{2}(\dot{\mathbf{x}}(t) + \dot{\mathbf{x}}(t+1))$. Ne scende che il rumore di modello ha componenti

$$\mathbf{v}_1(t) = \frac{T}{2}\mathbf{v}_2(t), \quad \mathbf{v}_3(t) = \frac{T}{2}\mathbf{v}_4(t).$$

e quindi, usando le covarianze ricavate per $\mathbf{v}_2(t)$ e $\mathbf{v}_4(t)$ si può così facilmente ricavare un'espressione per Q .

Al modello di stato va accoppiata l'equazione che descrive il sistema di misura. Normalmente con un sistema di radar si riescono a misurare la distanza r e (mediante procedimenti di triangolazione) l'azimuth θ . Queste quantità vengono misurate con errori che supporremo scorrelati nel tempo, a media nulla e con deviazioni standard σ_r, σ_θ , note. Per evitare di introdurre non linearità nell'equazione di osservazione, si può "fingere" di misurare le coordinate cartesiane $x = r \cos \theta, y = r \sin \theta$ che sono una funzione nota delle quantità che abbiamo supposto misurabili. In questo modo l'equazione di misura si scrive

$$\mathbf{y}(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}(t) + \mathbf{w}(t)$$

Questo è lecito purchè si riescano a calcolare delle buone approssimazioni per le covarianze degli errori, \mathbf{w} , indotti dagli errori di misura reali sulle quantità trasformate. In effetti, usando lo stesso procedimento basato sul calcolo della matrice Jacobiana introdotto in precedenza, si trova

$$\begin{aligned} \sigma_{\mathbf{w}_1} &= \sigma_r^2 \sin^2 \theta + \sigma_\theta^2 r^2 \cos^2 \theta \\ \sigma_{\mathbf{w}_2} &= \sigma_r^2 \cos^2 \theta + \sigma_\theta^2 r^2 \sin^2 \theta \\ \sigma_{\mathbf{w}_1 \mathbf{w}_2} &= (\sigma_r^2 - \sigma_\theta^2 r^2) \sin \theta \cos \theta \end{aligned} \quad (3.6)$$

Come abbiamo già notato, anche questa matrice varianza è funzione dello stato e quindi il modello ricavato, a stretto rigore è non lineare. Vedremo più avanti come si possa estendere l'algoritmo di filtraggio di Kalman anche a modelli non lineari.

Modelli di disturbi periodici. Il problema del filtraggio e della soppressione di disturbi periodici (di frequenza generalmente nota) si incontra in diversissime circostanze. Un caso tipico è nel progetto di filtri per la soppressione di disturbi acustici generati da alberi rotanti, motori etc... [?]. In certe altre applicazioni si ha a che fare con segnali di bassissima potenza che sono acquisiti e trasmessi mediante supporti fisici esposti a fenomeni di induzione elettromagnetica esterna, tipicamente alla frequenza di rete (50 Hz). Il disturbo è in questi casi schematizzabile come un segnale sinusoidale a frequenza nota la cui ampiezza e fase possono variare in modo imprevedibile nel tempo

$$\begin{aligned} \mathbf{y}(t) &= A(t) \cos [\theta t + \varphi(t)] = A(t) \cos \theta t \cos \varphi(t) \\ &\quad - A(t) \sin \theta t \sin \varphi(t) := X(t) \cos \varphi(t) - Y(t) \sin \varphi(t) \quad . \end{aligned} \quad (3.7)$$

Nel caso di disturbi di rete, in questa espressione si pone $\theta = 2\pi 50/f_c$, dove f_c è la frequenza di campionamento del segnale.

Se ampiezza e fase fossero aleatorie ma *costanti* nel tempo si potrebbe modellare $\mathbf{y}(t)$ per mezzo del semplice sistema lineare invariante

$$\begin{aligned} \mathbf{x}(t+1) &= \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \mathbf{x}(t) \\ \mathbf{y}(t) &= [1 \ 0] \mathbf{x}(t) \quad , \end{aligned} \quad (3.8)$$

dove la condizione iniziale $\mathbf{x}(t_0)$ è aleatoria. (Le componenti di $\mathbf{x}(t_0)$ possono ovviamente essere interpretate come $A \cos \theta$ e $A \sin \theta$ rispettivamente).

Per tener conto della variabilità temporale di ampiezza e fase si può immaginare che ad ogni istante si sommi allo stato $\mathbf{x}(t)$ un disturbo $\mathbf{v}(t)$ che per semplicità si assume essere rumore bianco di varianza Q . Il modello (3.8) si trasforma perciò nel

$$\begin{aligned} \mathbf{x}(t+1) &= A \mathbf{x}(t) + \mathbf{v}(t) \\ \mathbf{y}(t) &= [1 \ 0] \mathbf{x}(t) \quad , \end{aligned} \quad (3.9)$$

dove A è la matrice (ortogonale) che compare nella prima equazione in (3.8). Notiamo che la varianza di $\mathbf{v}(t)$ dà un'idea di quanto possa variare l'energia media del segnale,

$$\|\mathbf{x}(t)\|^2 = E \mathbf{x}'(t) \mathbf{x}(t) = \text{tr } \Sigma_{\mathbf{x}}(t, t)$$

nell'intervallo di campionamento, dato che dalla (3.9) si ricava

$$\|\mathbf{x}(t+1)\|^2 = \|\mathbf{x}(t)\|^2 + \text{tr } Q \quad .$$

La quantità $\text{tr } Q/T_c$, dove T_c è il periodo di campionamento, è quindi interpretabile come la potenza del disturbo di induzione elettromagnetica a 50 Hz. Questo parametro può in qualche caso essere noto o misurabile.

Una conferma qualitativa che il nuovo modello riproduce segnali aventi l'andamento temporale indicato, si ha calcolando la risposta forzata del modello (3.9). Si trova

$$\begin{aligned} \mathbf{y}(t+1) &= [1 \ 0] \sum_0^t A^{t-k} v(k) = [1 \ 0] A^t \sum_0^t A^{-k} v(k) \\ &= [\cos \theta t - \sin \theta t] \sum_0^t A^{-k} v(k) \\ &= X(t) \cos \theta t - Y(t) \sin \theta t \quad , \end{aligned}$$

dove

$$\begin{aligned} X(t) &= \sum_0^t [\cos \theta k v_1(k) + \sin \theta k v_2(k)] \quad , \\ Y(t) &= - \sum_0^t [\sin \theta k v_1(k) - \cos \theta k v_2(k)] \quad . \end{aligned}$$

sono termini di modulazione aleatoria d'ampiezza la cui varianza cresce nel tempo e che rendono quindi l'evoluzione forzata preponderante rispetto all'evoluzione libera.

Modelli di questo genere, che possono facilmente essere complicati aggiungendo eventuali componenti armoniche del disturbo a frequenze multiple della fondamentale, vengono impiegati diffusamente. C'è da dire che nell'elaborazione di segnali la cui banda utile include la frequenza del disturbo (i 50 Hz ad esempio), l'impiego di filtri numerici tradizionali è assai critico perché, se non si vuole introdurre distorsione armonica a frequenze diverse da quella di rete la caratteristica del filtro deve essere il più vicina possibile all'"elimina-riga" ideale attorno alla frequenza di 50 Hz (filtri "ad ago").

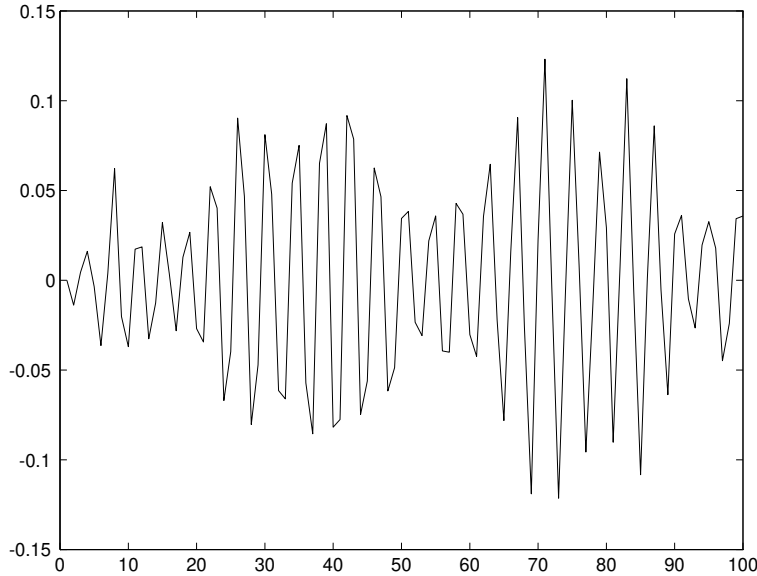


FIGURA 3.1. Oscillazione di ampiezza e fase aleatorie.

In questi casi si hanno normalmente problemi di stabilità e di durata inaccettabile del transitorio mentre invece un filtro di Kalman basato sul modello (3.9) permette di stimare in linea la componente a 50 Hz e di sottrarla successivamente al segnale utile realizzando questa operazione mediante uno schema di calcolo strutturalmente stabile (come vedremo). Inoltre il filtro di Kalman, non essendo basato su criteri di progetto euristici ed essendo ottimo dal punto di vista dell'elaborazione del segnale è l'operazione che fornisce la *minima varianza possibile* del segnale residuo a 50 Hz.

Rumore a bassissima frequenza (“drift”). In certi problemi di elaborazione del segnale i dati presentano sovrapposta una componente a bassissima frequenza (“drift”) dovuta ad artefatti e fenomeni indesiderati (ad esempio bassa reiezione degli amplificatori ecc.). Di norma questa componente non interessa (cioè non contiene informazione utile) e si cerca di eliminarla.

Le soluzioni più ovvie (filtro numerico di tipo passa-alto) danno risultati soddisfacenti solo se il segnale utile non ha componenti apprezzabili a bassa frequenza. In ogni caso i filtri usuali introducono distorsioni in banda estesa di frequenze e queste distorsioni possono essere indesiderate.

Sia $\mathbf{y}(t)$ il drift e $\mathbf{w}(t)$ il segnale utile. Scrivendo

$$\mathbf{y}(t) = \mathbf{z}(t) + \mathbf{w}(t) \quad (3.10)$$

si ha un semplice modello additivo per le osservazioni. Un andamento tipico del processo $\{\mathbf{y}(t)\}$ è quello rappresentato nella figura 3.2. Pensando il segnale (utile) sovrapposto al “drift” come *rumore* si può stimare il drift a partire dalle misure (3.10). Ottenuto lo stimatore ricorsivo per $\mathbf{z}(t)$ basterà togliere alle misure la

stima $\hat{z}(t | t)$ ad ogni istante di tempo. Applicando questo schema (o delle varianti un poco più elaborate) si riesce a non modificare apprezzabilmente il segnale utile $\mathbf{w}(t)$ alle basse frequenze.

La fase preliminare di questo procedimento è quella di procurarsi un modello matematico per $\mathbf{z}(t)$ che sia allo stesso tempo il più *realistico* e il più *semplice* possibile. Questa fase è decisamente la fase cruciale del problema e va ovviamente risolta in base alla *specifica* natura del drift $\mathbf{z}(t)$ che è presente nel segnale con cui si ha realmente a che fare.

Vi sono tuttavia alcune considerazioni generali da tener sempre presenti. Spesso esaminando l'andamento del segnale si possono individuare tratti, di durata molto lunga rispetto alla scala temporale in cui si verificano variazioni significative del segnale utile $\mathbf{w}(t)$, in cui $\mathbf{z}(t)$ è modellabile approssimativamente mediante andamenti temporali determinati, tipicamente del tipo seguente.

A) Segnali polinomiali del tipo

$$\mathbf{z}(t) = \theta_0 + \theta_1 t + \theta_2 t^2 + \dots + \theta_k t^k \quad (3.11a)$$

B) "trends" esponenziali (crescenti) come

$$\mathbf{z}(t) = A \lambda_0^t, \quad |\lambda_0| > 1 \quad (3.11b)$$

C) Componenti "stagionali" periodiche

$$\mathbf{z}(t) = A \cos(\omega_0 t + \theta) \quad (3.11c)$$

o pseudoperiodiche (crescenti)

$$\mathbf{z}(t) = A \rho_0^t \cos(\omega_0 t + \theta) \quad (3.11d)$$

È importante insistere sul carattere qualitativo di queste descrizioni. È sufficiente che il modello di $\mathbf{z}(t)$ renda approssimativamente ragione (meglio sarebbe dire "statisticamente") dell'andamento temporale del drift.

Ovviamente gli andamenti del tipo A), B) o C) vengono individuati in base a un campione finito di dati di misura e la scelta che si fa deve essere sufficientemente flessibile da permettere la descrizione del "trend" futuro, sull'andamento del quale si hanno solo delle convinzioni "a priori" che possono essere abbastanza vaghe.

I casi B) e C) sono tipici dei segnali che si incontrano in applicazioni di tipo econometrico. Spesso è possibile in questi casi ricavare dall'analisi dei dati delle buone stime di λ_0 oppure di ρ_0 e ω_0 , perché il trend e la componente stagionale del segnale sono spesso ben evidenti e talvolta anche note in base alla "fisica" del segnale.

In ogni caso gli andamenti temporali A)–B)–C) hanno la nota peculiarità di essere descrivibili come soluzioni di un'equazione alle differenze omogenea che si scrive nei diversi casi

$$A) \quad (1 - z^{-1})^{k+1} \mathbf{z}(t) = 0 \quad ,$$

$$B) \quad (1 - \lambda_0 z^{-1}) \mathbf{z}(t) = 0 \quad ,$$

$$C) \quad \left(1 - (2 \cos \omega_0) \rho_0 z^{-1} + \rho_0^2 z^{-2}\right) \mathbf{z}(t) = 0 \quad .$$

Queste equazioni sono di per sé già un modello del segnale $\{\mathbf{z}(t)\}$ i cui parametri possono considerarsi noti a priori (determinabili mediante un'analisi preliminare dei dati). Si noti che i tre casi sono stati distinti più per differenziare dei possibili

settori applicativi che per la presenza di differenze sostanziali di carattere matematico. In realtà si possono benissimo considerare delle “misure” dei tre andamenti, corrispondenti ad un modello generico del tipo

$$A^*(z^{-1}) \mathbf{z}(t) = 0 \quad , \quad (3.12)$$

dove $A^*(z^{-1})$ è fattorizzabile nel prodotto di termini polinomiali del tipo A)–B)–C) appena descritti.

La (3.12) è un’equazione alle differenze che va accoppiata ad un insieme di condizioni iniziali le quali determinano le ampiezze (e fasi), $\theta_0, \theta_1, \dots, \theta_k, A, \rho_0$, dei modi (divergenti) in (3.11a).

Come tale essa è un modello “rigido” del segnale di drift nel senso che l’andamento temporale del segnale $\{\mathbf{z}(t)\}$ descritto dalla (3.12) associata ad un qualunque insieme di condizioni iniziali in t_0 ne risulta univocamente determinato per tutti i $t > t_0$.

Una descrizione più flessibile si ottiene ammettendo che le ampiezze $\theta_0, \theta_1, \dots, \theta_k, A, \rho_0$ possano *variare in modo aleatorio*. Un modo semplice per ottenere questa modulazione aleatoria delle ampiezze è di introdurre nel modello dinamico un rumore bianco $\{\mathbf{v}(t)\}$ di varianza opportuna. Si vede che usando ad esempio la tecnica di realizzazione illustrata in precedenza si può arrivare a descrizioni di $\{\mathbf{z}(t)\}$ del tipo

$$\begin{aligned} \mathbf{x}(t+1) &= \begin{bmatrix} 0 & & -a_n \\ 1 & & \\ & \ddots & \\ & & 1 & -a_1 \end{bmatrix} \mathbf{x}(t) + \mathbf{v}(t) \\ \mathbf{z}(t) &= [0 \dots 0 \quad 1] \mathbf{x}(t) \quad , \end{aligned} \quad (3.13)$$

dove a_1, \dots, a_n sono i coefficienti (noti!) del polinomio $A^*(z^{-1})$,

$$A^*(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_n z^{-n} \quad .$$

La scelta della matrice varianza, Q , di $\mathbf{v}(t)$ non è purtroppo un problema semplice. Si può qualche volta imporre a $\{\mathbf{v}(t)\}$ di avere la forma

$$\mathbf{v}(t) = \begin{bmatrix} b_n \\ \vdots \\ b_1 \end{bmatrix} \mathbf{n}(t) := b \mathbf{n}(t) \quad ,$$

con $\{\mathbf{n}(t)\}$ processo bianco scalare di varianza σ^2 (rango uguale a uno). Il termine b_1 può essere normalizzato e posto uguale ad 1. In questo caso il modello (3.13) corrisponde ad una relazione ingresso-uscita del tipo

$$\mathbf{z}(t) = \frac{z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} \mathbf{n}(t) \quad , \quad t \geq t_0 \quad ,$$

e la conoscenza di $Q = \sigma^2 b b'$ equivale a quella dei coefficienti del numeratore $\{b_k\}$ e della varianza σ^2 di $\mathbf{n}(t)$. Intuitivamente, si può dire che i coefficienti di $\{\mathbf{z}(t)\}$ sono funzioni “lentamente variabili” nel tempo e questo suggerisce di prendere valori molto piccoli per la varianza (scalare) di $\{\mathbf{v}(t)\}$.

Nelle applicazioni ingegneristiche molto spesso è sufficiente considerare segnali polinomiali (del tipo A) di grado k uguale a zero o uno. Il modello del drift per

$k = 0$ è semplicemente

$$\mathbf{z}(t+1) = \mathbf{z}(t) + \mathbf{n}(t) \quad .$$

Il processo che soddisfa a questa equazione in $t \geq t_0$ si chiama “passeggiata casuale”. Supponendo il processo $\{\mathbf{n}(t)\}$ i.i.d. si ricava, applicando il teorema ergodico, che

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t_0+1}^{t_0+T} \left(\mathbf{z}(t+1) - \mathbf{z}(t) \right)^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t_0+1}^{t_0+T} \mathbf{n}(t)^2 = \sigma^2 \quad ,$$

con probabilità 1. In questo caso quindi σ^2 ($= Q$) può essere stimata dai dati.

L'ordine del modello deve essere comunque fissato in base ad un accurato esame dei dati. Riportiamo qui due tipici andamenti.

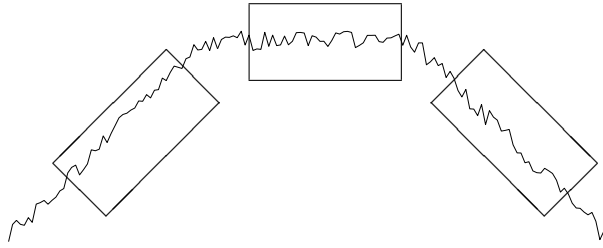


FIGURA 3.2. Individuazione dei trends.

Nel caso a) basterà verosimilmente prendere $k = 0$ (si ha approssimativamente $\mathbf{z}(t)$ costante per lunghi tratti), mentre nel caso b) sarà bene usare almeno $k = 1$ (ci sono dei tratti “a rampa”).

4. Le equazioni del Filtro di Kalman

Se si adotta il punto di vista di modellare tutti i segnali in gioco in un problema di stima come processi che ammettono realizzazioni lineari di dimensione finita, allora anche la modellizzazione di un qualunque problema di stima lineare si riporta naturalmente ad un modello standard del tipo (1.1). Come vedremo in questa sezione, usando l'idea di “aumentare lo stato” tutta la formulazione del problema di stima si riduce in essenza alla ricerca di un modello di stato che descriva il segnale di osservazione $\{\mathbf{y}(t)\}$. Dato che i modelli di stato possono essere combinati linearmente fra loro mantenendo la stessa struttura, scende che tutti i dati “a priori” sul problema di stima in esame possano essere condensati in un unico modello lineare di dimensione finita del tipo

$$\begin{cases} \mathbf{x}(t+1) &= A \mathbf{x}(t) + \mathbf{v}(t) \\ \mathbf{y}(t) &= C \mathbf{x}(t) + \mathbf{w}(t) \end{cases} \quad , \quad t \geq t_0 \quad , \quad (4.1)$$

dove $\{\mathbf{y}(t)\}$ è il processo m -dimensionale delle osservazioni, $\{\mathbf{v}(t)\}$ e $\{\mathbf{w}(t)\}$ sono *rumori bianchi* (di media zero) e di matrice di covarianza

$$E \left\{ \begin{bmatrix} \mathbf{v}(t) \\ \mathbf{w}(t) \end{bmatrix} [\mathbf{v}(s)', \mathbf{w}(s)'] \right\} = \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \delta(t-s) \quad . \quad (4.2)$$

Lo stato, $\{\mathbf{x}(t)\}$, del modello (4.1) è un processo n -dimensionale non osservabile direttamente, il cui valore iniziale $\mathbf{x}(t_0) = \mathbf{x}_0$ è un vettore casuale *scorrelato dall'andamento presente e futuro dei rumori di modello* $\{\mathbf{v}(t)\}$ e di misura $\{\mathbf{w}(t)\}$

$$E \left\{ \mathbf{x}_0 [\mathbf{v}(t)', \mathbf{w}(t)'] \right\} = 0 \quad , \quad \forall t \geq t_0 \quad . \quad (4.3)$$

Assumeremo anche che la media e la varianza di \mathbf{x}_0 siano note

$$E \mathbf{x}_0 = \mu_0 \quad , \quad \text{Var}\{\mathbf{x}_0\} = P_0 \quad . \quad (4.4)$$

La matrice P_0 è una generica matrice $n \times n$ simmetrica e semidefinita positiva. Il caso $P_0 = 0$ corrisponde all'esatta conoscenza dello stato iniziale $\mathbf{x}(t_0) = \mu_0$.

In quanto segue le matrici A, C, Q, S, R che caratterizzano il modello possono anche essere varianti nel tempo. L'unica restrizione che porremo sarà quella dell'*invertibilità di R* , ipotesi che può anche essere scritta come

$$R > 0 \quad . \quad (4.5)$$

Questa ipotesi, che corrisponde ad assumere tutte le componenti $\{\mathbf{y}_k(t)\}$, $k = 1, \dots, m$, del segnale che si misura affette da rumore, ha un carattere puramente tecnico e può anche essere rimossa al prezzo di qualche complicazione. Essa gioca qui un ruolo meno cruciale della positività stretta della densità spettrale di $\{\mathbf{y}(t)\}$ nel caso stazionario e serve soprattutto a garantire l'*invertibilità della varianza del processo di innovazione*

$$\mathbf{e}(t) := \mathbf{y}(t) - \hat{\mathbf{y}}(t | t-1) \quad , \quad t \geq t_0 \quad . \quad (4.6)$$

Prima di iniziare ad illustrare i (facili) calcoli che portano alla derivazione delle equazioni dinamiche del filtro cerchiamo di dare una qualche giustificazione "operativa" al modello (4.1) e di individuare un problema di stima "standard" al quale ricondursi.

Se si ha a che fare con un *problema di previsione* per $\{\mathbf{y}(t)\}$, è abbastanza ovvio che il modello (4.1) corrisponde alla descrizione di $\{\mathbf{y}(t)\}$ stesso come processo lineare a dimensione finita. Questa descrizione è stata illustrata e motivata nel capitolo precedente e non è qui necessario alcun commento ulteriore a questo riguardo. Notiamo poi che il predittore (di un passo) $\hat{\mathbf{y}}(t+1 | t)$ si ottiene dalla seconda delle (4.1) come

$$\hat{\mathbf{y}}(t+1 | t) = C \hat{\mathbf{x}}(t+1 | t) \quad , \quad (4.7)$$

dato che la proiezione di $\mathbf{w}(t+1)$ su $\mathbf{H}_t(\mathbf{y})$ è chiaramente zero essendo $\mathbf{w}(t+1) \perp \{\mathbf{v}(s), \mathbf{w}(s), \mathbf{x}(t_0) \mid s \leq t\}$ ed essendo

$$\mathbf{H}_t(\mathbf{y}) \subset \mathbf{H}(\mathbf{v}^{t-1}, \mathbf{w}^t, \mathbf{x}(t_0)) \quad , \quad (4.8)$$

come si riconosce facilmente dalle equazioni (4.1).

Quindi il predittore $\hat{\mathbf{y}}(t+1 | t)$ è noto non appena è noto il predittore dello stato $\hat{\mathbf{x}}(t+1 | t)$.

Supponendo invece di voler *filtrare* da $\{\mathbf{y}(t)\}$ un segnale (diverso), $\{\mathbf{z}(t)\}$, facciamo riferimento allo schema additivo

$$\mathbf{y}(t) = \mathbf{z}(t) + \boldsymbol{\xi}(t) \quad , \quad (4.9)$$

in cui $\{\boldsymbol{\xi}(t)\}$ è una componente di “rumore”, in generale correlato. (Lo schema $\mathbf{y}(t) = H\mathbf{z}(t) + \boldsymbol{\xi}(t)$, apparentemente più generale, conduce allo stesso tipo di conclusioni).

Modellando $\{\mathbf{z}(t)\}$ e $\{\boldsymbol{\xi}(t)\}$ come segnali lineari di dimensione finita, scorrelati, si può scrivere

$$\begin{cases} \mathbf{x}_1(t+1) = A_1 \mathbf{x}_1(t) + \mathbf{v}_1(t) \\ \mathbf{z}(t) = C_1 \mathbf{x}_1(t) + \mathbf{w}_1(t) \end{cases} \quad (4.10)$$

$$\begin{cases} \mathbf{x}_2(t+1) = A_2 \mathbf{x}_2(t) + \mathbf{v}_2(t) \\ \boldsymbol{\xi}(t) = C_2 \mathbf{x}_2(t) + \mathbf{w}_2(t). \end{cases} \quad (4.11)$$

Si noti che nel caso stazionario queste due rappresentazioni potrebbero ad esempio essere ricavate partendo dagli spettri di $\{\mathbf{z}(t)\}$ e $\{\boldsymbol{\xi}(t)\}$ col procedimento illustrato nel capitolo precedente.

Dalla (4.10) si ricava allora immediatamente il modello in forma standard

$$\mathbf{x}(t+1) = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \end{bmatrix} := A \mathbf{x}(t) + \mathbf{v}(t) \quad (4.12a)$$

$$\mathbf{y}(t) = [C_1 \ C_2] \mathbf{x}(t) + [I \ I] \begin{bmatrix} \mathbf{w}_1(t) \\ \mathbf{w}_2(t) \end{bmatrix} := C \mathbf{x}(t) + \mathbf{w}(t). \quad (4.12b)$$

Se $\begin{bmatrix} \mathbf{v}_1(t) \\ \mathbf{w}_2(t) \end{bmatrix}$ e $\begin{bmatrix} \mathbf{v}_2(t) \\ \mathbf{w}_1(t) \end{bmatrix}$ sono scorrelati (cosa che accade senz'altro se i due modelli per $\{\mathbf{z}(t)\}$ e $\{\boldsymbol{\xi}(t)\}$ non sono costruiti in modo artificiosamente “ridondante”) si può in effetti verificare che i rumori di modello, $\{\mathbf{v}(t)\}$, e di osservazione, $\{\mathbf{w}(t)\}$, nella (4.12) sono bianchi e hanno covarianza facilmente calcolabile a partire da quella di $\mathbf{v}_i(t)$ e $\mathbf{w}_i(t)$, $i = 1, 2$.

A questo punto $\mathbf{z}(t)$ si può esprimere come una funzione lineare dello stato (aumentato) $\mathbf{x}(t)$ del modello (4.12) più rumore bianco scorrelato da $\mathbf{x}(t)$,

$$\mathbf{z}(t) = [C_1 \ 0] \mathbf{x}(t) + \mathbf{w}_1(t) := \bar{C}_1 \mathbf{x}(t) + \mathbf{w}_1(t) \quad (4.13)$$

ed è immediato convincersi che questa rappresentazione di $\mathbf{z}(t)$ conduce all'espressione della stima $\hat{\mathbf{z}}(t | t)$ come

$$\hat{\mathbf{z}}(t | t) = \bar{C}_1 \hat{\mathbf{x}}(t | t) + \hat{\mathbf{w}}_1(t | t) \quad (4.14)$$

In questa formula $\hat{\mathbf{w}}_1(t | t)$ non è in generale zero, dato che $\mathbf{w}_1(t)$ ed $\mathbf{y}(t)$ sono correlati, essendo diversa da zero la covarianza tra $\mathbf{w}_1(t)$ e $\mathbf{w}(t)$,

$$E[\mathbf{w}_1(t) \mathbf{w}(t)'] := R_1 \quad .$$

Notiamo però che $\hat{\mathbf{w}}_1(t | t)$ può essere calcolato in modo molto semplice per mezzo di due proiezioni successive, prima sullo spazio $\text{span}\{\mathbf{x}_0, \mathbf{v}(s-1), \mathbf{w}(s); s \leq t\}$ e successivamente su $\mathbf{H}_t(\mathbf{y})$ che è un sottospazio di quest'ultimo (cfr. la (4.8)). Si ha³:

$$\hat{E}[\mathbf{w}_1(t) | \mathbf{x}_0 \mathbf{v}(s-1), \mathbf{w}(s); s \leq t] = \hat{E}[\mathbf{w}_1(t) | \mathbf{w}(t)] = R_1 R^{-1} \mathbf{w}(t) \quad (4.15)$$

dato che $\mathbf{w}_1(t) \perp \{\mathbf{v}(s), \mathbf{w}(s); s < t\}$. Proiettando ora la (4.15) su $\mathbf{H}_t(\mathbf{y})$ si trova

$$\hat{\mathbf{w}}_1(t | t) = R_1 R^{-1} \hat{\mathbf{w}}(t | t) \quad , \quad (4.16)$$

³Notare che $\text{span}\{\mathbf{x}_0, \mathbf{v}(s-1), \mathbf{w}(s); s \leq t\} = \mathbf{H}(\mathbf{x}_0) \oplus \mathbf{H}_{t-1}(\mathbf{v}, \mathbf{w}) \oplus \mathbf{H}(\mathbf{w}(t))$.

dove $\hat{\mathbf{w}}(t | t)$ si ricava immediatamente dalla (4.12)

$$\hat{\mathbf{w}}(t | t) = \hat{\mathbf{y}}(t | t) - C \hat{\mathbf{x}}(t | t) = \mathbf{y}(t) - C \hat{\mathbf{x}}(t | t) \quad .$$

Inserendo nella (4.14) si trova infine

$$\hat{\mathbf{z}}(t | t) = \bar{C}_1 \hat{\mathbf{x}}(t | t) + R_1 R^{-1} [\mathbf{y}(t) - C \hat{\mathbf{x}}(t | t)] \quad (4.17)$$

e questa formula esprime $\hat{\mathbf{z}}(t | t)$ in modo esplicito *come funzione della stima* (“filtrata”) *dello stato* $\hat{\mathbf{x}}(t | t)$. \diamond

Questi due esempi servono a mettere in evidenza un fatto generale: se le osservazioni sono descritte da un modello stocastico lineare del tipo (4.1) e lo stato $\mathbf{x}(t)$ ingloba tutta l'informazione necessaria per descrivere anche il processo messaggio $\{\mathbf{z}(t)\}$, allora la stima di $\mathbf{z}(t)$ si ottiene attraverso una semplice trasformazione lineare della stima del vettore di stato.

Questo “principio” consente quindi di ridurre ogni problema di stima lineare basato sulle osservazioni $\{\mathbf{y}(t)\}$ ad un problema di *stima dello stato* del modello (4.1). Questa è la ragione per cui nella teoria del filtro di Kalman ci si preoccupa unicamente di ricavare stimatori (lineari a minima varianza) del processo di stato.

Nei problemi concreti i rumori di modello e di osservazione sono spesso scorrelati. Per questa ragione molti autori danno prima una derivazione delle equazioni dinamiche del filtro nel caso di rumori scorrelati e presentano in una fase successiva una serie di trucchi per generalizzare l'algoritmo al caso di rumori correlati. Per cercare di ridurre al minimo le manipolazioni di formule presenteremo qui l'algoritmo di Kalman direttamente nel caso generale di rumori correlati e cercheremo di scrivere le formule in modo tale che il caso particolare in cui $S = 0$ sia immediatamente deducibile dalle formule generali.

Notiamo allo scopo che $\mathbf{v}(t)$ è correlata solo con $\mathbf{w}(t)$ essendo per ipotesi $\mathbf{v}(t) \perp \mathbf{w}(s)$ per $s \neq t$. Se ne deduce che “l'errore di stima”

$$\tilde{\mathbf{v}}(t) := \mathbf{v}(t) - \hat{E}[\mathbf{v}(t) | \mathbf{H}(\mathbf{w})] = \mathbf{v}(t) - \hat{E}[\mathbf{v}(t) | \mathbf{w}(t)] \quad (4.18)$$

è ortogonale all'intero processo rumore di misura $\{\mathbf{w}(t)\}$.

Inoltre $\tilde{\mathbf{v}}(t)$ si può scrivere come

$$\tilde{\mathbf{v}}(t) = \mathbf{v}(t) - SR^{-1} \mathbf{w}(t) \quad (4.19)$$

e da questa relazione si ricava immediatamente che $\{\tilde{\mathbf{v}}(t)\}$ è *rumore bianco di varianza*,

$$\tilde{Q} = Q - SR^{-1} S' \quad . \quad (4.20)$$

Sostituendo ora la (4.19) nel modello (4.1) e tenendo conto che $\mathbf{w}(t) = \mathbf{y}(t) - C\mathbf{x}(t)$, si ricava immediatamente che *il modello lineare* (4.1) *può essere riscritto nella forma*

$$\mathbf{x}(t+1) = F\mathbf{x}(t) + SR^{-1} \mathbf{y}(t) + \tilde{\mathbf{v}}(t) \quad (4.21a)$$

$$\mathbf{y}(t) = C\mathbf{x}(t) + \mathbf{w}(t) \quad , \quad (4.21b)$$

dove i rumori bianchi $\{\tilde{\mathbf{v}}(t)\}$ e $\{\mathbf{w}(t)\}$, di varianze rispettive \tilde{Q} ed R , sono tra loro scorrelati. La matrice F è data dall'espressione

$$F = A - SR^{-1} C \quad . \quad (4.22)$$

Siamo in questo modo riusciti a trasformare il modello a rumori correlati in uno in cui i rumori non lo sono, al prezzo di introdurre nell'equazione di stato un termine

forzante, $SR^{-1}\mathbf{y}(t)$, che ha il significato di una retroazione dall'uscita sullo stato (una retroazione di questo genere si chiama *output injection* in inglese). La presenza di un termine forzante di questo genere nelle equazioni di stato potrebbe in pratica aversi indipendentemente dalla correlazione dei rumori di modello e di misura e derivare dalla presenza di un ingresso che agisce sul sistema, che sia generato da una retroazione dall'uscita. Nei sistemi di controllo e in molti problemi di controllo con disturbi aleatori, capita in effetti di dover stimare lo stato in presenza di ingressi $\mathbf{u}(t) = F(\mathbf{y}^t)$, che possono essere ottenuti da una *retroazione lineare e causale sull'uscita*. Questi problemi di stima dello stato in *sistemi di controllo*, eventualmente operanti in retroazione, possono essere affrontati con una versione generalizzata del filtro di Kalman che richiede solo alcune semplici aggiunte alle formule. Si veda a questo proposito il problema 9.2.

ALGORITMO 9.2 (Filtro di Kalman). *Gli stimatori lineari a minima varianza $\hat{\mathbf{x}}(t+1 | t)$ e $\hat{\mathbf{x}}(t | t)$ dello stato del modello lineare (4.1) all'istante $t+1$ e t , in base alle osservazioni $\{\mathbf{y}(s) ; t_0 \leq s \leq t\}$, sono calcolabili mediante il seguente algoritmo ricorsivo.*

1. Stime a priori (Aggiornamento temporale)

$$\hat{\mathbf{x}}(t+1 | t) = F\hat{\mathbf{x}}(t | t) + SR^{-1}\mathbf{y}(t) \quad , \quad (4.23)$$

$$P(t+1 | t) = FP(t | t)F' + \tilde{Q} \quad , \quad t \geq t_0 \quad . \quad (4.24)$$

2. Stime a posteriori (Aggiornamento rispetto alle misure)

$$\hat{\mathbf{x}}(t+1 | t+1) = \hat{\mathbf{x}}(t+1 | t) + L(t+1)[\mathbf{y}(t+1) - C\hat{\mathbf{x}}(t+1 | t)] \quad (4.25)$$

$$P(t+1 | t+1) = P(t+1 | t) - P(t+1 | t)C'\Lambda(t+1)^{-1}CP(t+1 | t) \quad (4.26)$$

3. Condizioni iniziali,

$$\hat{\mathbf{x}}(t_0 | t_0 - 1) = \mu_0 \quad , \quad P(t_0 | t_0 - 1) = P_0 \quad . \quad (4.27)$$

Dove le matrici F e \tilde{Q} sono definite dalla (4.22) e dalla (4.20), $P(t+1 | t)$ e $P(t | t)$ sono le varianze degli errori di predizione e filtraggio,

$$P(t+1 | t) = E\tilde{\mathbf{x}}(t+1 | t)\tilde{\mathbf{x}}(t+1 | t)' \quad , \quad P(t | t) = E\tilde{\mathbf{x}}(t | t)\tilde{\mathbf{x}}(t | t)' \quad ,$$

$\Lambda(t)$ è la varianza del processo di innovazione $\mathbf{e}(t) = \mathbf{y}(t) - C\hat{\mathbf{x}}(t | t-1)$,

$$\Lambda(t) = CP(t | t-1)C' + R \quad (4.28)$$

e il "guadagno del filtro" $L(t)$ è definito dalla

$$L(t) = P(t | t-1)C'\Lambda^{-1}(t) \quad . \quad (4.29)$$

DIMOSTRAZIONE. Cominciamo col ricavare le equazioni di aggiornamento che descrivono l'evoluzione della stima a "informazione fissa".

La (4.23) si ricava proiettando l'equazione (4.21) su $\mathbf{H}_t(\mathbf{y})$ ⁴. Naturalmente, data l'ortogonalità

$$\tilde{\mathbf{v}}(t) \perp \text{span}\{\mathbf{x}_0, \tilde{\mathbf{v}}(s-1), \mathbf{w}(s); s \leq t\} \supset \mathbf{H}_t(\mathbf{y})$$

la proiezione di $\tilde{\mathbf{v}}(t)$ è zero.

La (4.24) si ricava sottraendo la (4.23) dalla (4.21), col che l'errore di predizione risulta descritto dalla

$$\tilde{\mathbf{x}}(t+1|t) = F\tilde{\mathbf{x}}(t|t) + \tilde{\mathbf{v}}(t) \quad . \quad (4.30)$$

Si noti che in questa equazione $\tilde{\mathbf{v}}(t)$ e $\tilde{\mathbf{x}}(t|t)$ sono *scorrelati* perché, come si vede dalla (4.19),

$$\mathbf{x}(t) \in \text{span}\{\mathbf{x}(t_0), \tilde{\mathbf{v}}(s), \mathbf{w}(s), ; s < t\} \quad (4.31)$$

e

$$\mathbf{H}_t(\mathbf{y}) \subset \text{span}\{\mathbf{x}(t_0), \tilde{\mathbf{v}}(s-1), \mathbf{w}(s), ; s \leq t\} \quad (4.32)$$

e dalle ipotesi poste segue che $\tilde{\mathbf{v}}(t)$ è ortogonale a entrambi i due sottospazi a secondo membro delle (4.30), (4.31). Calcolando allora la varianza dell'errore di predizione $\tilde{\mathbf{x}}(t+1|t)$ si ottiene la (4.24).

L'equazione di aggiornamento della stima di $\mathbf{x}(t+1)$ quando si aggiunge la nuova misura $\mathbf{y}(t+1)$, si ricava facilmente notando che:

$$\mathbf{H}_{t+1}(\mathbf{y}) = \mathbf{H}_t(\mathbf{y}) \oplus \mathbf{H}(\mathbf{e}(t+1)) \quad , \quad (4.33)$$

dove $\mathbf{e}(t+1)$ è l'innovazione all'istante $t+1$,

$$\begin{aligned} \mathbf{e}(t+1) &= \mathbf{y}(t+1) - C\hat{\mathbf{x}}(t+1|t) = C[\mathbf{x}(t+1) - \hat{\mathbf{x}}(t+1|t)] \\ &+ \mathbf{w}(t+1) = C\tilde{\mathbf{x}}(t+1|t) + \mathbf{w}(t+1) \quad . \end{aligned} \quad (4.34)$$

Proiettando $\mathbf{x}(t+1)$ su $\mathbf{H}_{t+1}(\mathbf{y})$ e usando la decomposizione ortogonale (4.33) si trova

$$\hat{\mathbf{x}}(t+1|t+1) = \hat{\mathbf{x}}(t+1|t) + L(t+1)\mathbf{e}(t+1) \quad ,$$

dove

$$L(t+1) = \text{Cov}[\mathbf{x}(t+1), \mathbf{e}(t+1)] \text{Var}\{\mathbf{e}(t+1)\}^{-1} \quad .$$

Per calcolare esplicitamente il guadagno $L(t+1)$ notiamo innanzitutto che

$$\begin{aligned} \text{Cov}[\mathbf{x}(t+1), \mathbf{e}(t+1)] &= \text{Cov}[\mathbf{x}(t+1), \tilde{\mathbf{x}}(t+1|t)] C' \\ &+ \text{Cov}[\mathbf{x}(t+1), \mathbf{w}(t+1)] \quad . \end{aligned}$$

L'ultimo addendo è zero per la (4.31) per cui ci rimane da calcolare solo $\text{Cov}[\mathbf{x}(t+1), \tilde{\mathbf{x}}(t+1|t)]$. Applicando il principio di ortogonalità si vede subito che questa quantità vale

$$\text{Cov}[\mathbf{x}(t+1), \tilde{\mathbf{x}}(t+1|t)] = \text{Cov}[\tilde{\mathbf{x}}(t+1|t), \tilde{\mathbf{x}}(t+1|t)] = P(t+1|t) \quad .$$

⁴Qui, come sempre quando si parla di proiezioni, si sottintende un'operazione preliminare di sottrazione delle medie dalle variabili in gioco e l'operazione di proiezione si intende effettuata sulle variabili a media nulla $\tilde{\mathbf{x}}(t) := \mathbf{x}(t) - \mu_{\mathbf{x}}(t)$, basandosi sui corrispondenti dati a media nulla $\tilde{\mathbf{y}}(t) := \mathbf{y}(t) - \mu_{\mathbf{y}}(t)$. Notare che per il modello (4.1), l'operazione di depurazione dalla media di entrambe le variabili, corrisponde semplicemente a considerare i dati generati da un modello *identico* al modello di partenza (4.1) in cui la condizione iniziale è $\tilde{\mathbf{x}}_0 := \mathbf{x}_0 - \mu_0$. Una volta calcolate le proiezioni e trovate le relative formule per le quantità barrate, l'operazione di reintroduzione della media si fa quindi semplicemente ri-assegnando alla condizione iniziale la sua media non nulla μ_0 .

Infine, la varianza dell'innovazione si può calcolare dalla (4.34), notando che $\tilde{\mathbf{x}}(t+1 | t) \perp \mathbf{w}(t+1)$,

$$\Lambda(t+1) = \text{Var}[\mathbf{e}(t+1)] = CP(t+1 | t)C' + R \quad .$$

Se ne conclude che $L(t+1)$ ha proprio la struttura (4.29).

Per trovare l'equazione di aggiornamento della varianza d'errore basta notare che dalla (4.25) segue la

$$\mathbf{x}(t+1) - \hat{\mathbf{x}}(t+1 | t+1) = \mathbf{x}(t+1) - \hat{\mathbf{x}}(t+1 | t) - L(t+1) \mathbf{e}(t+1) \quad ,$$

ovvero

$$\tilde{\mathbf{x}}(t+1 | t+1) = \tilde{\mathbf{x}}(t+1 | t) - L(t+1) \mathbf{e}(t+1) \quad .$$

Riscriviamo questa relazione come

$$\tilde{\mathbf{x}}(t+1 | t+1) + L(t+1) \mathbf{e}(t+1) = \tilde{\mathbf{x}}(t+1 | t)$$

e notiamo che i due termini a primo membro sono scorrelati essendo, per il principio di ortogonalità, $\tilde{\mathbf{x}}(t+1 | t+1) \perp \mathbf{H}_{t+1}(\mathbf{y}) = \mathbf{H}_{t+1}(\mathbf{e}) \ni \mathbf{e}(t+1)$. Ne viene che

$$P(t+1 | t+1) + L(t+1) \Lambda(t+1) L'(t+1) = P(t+1 | t)$$

e questa relazione coincide con la (4.26). Le condizioni iniziali corrispondono alle formule usuali della stima lineare a minima varianza (Bayesiana) particolarizzate al caso degenerare in cui non ci sono misure disponibili (il che si può far corrispondere a $\mathbf{H}(\mathbf{y}(t_0 - 1)) = \{0\}$). \square

L'algoritmo 9.2 può essere diviso in una parte "in linea" descritta dalle equazioni dello stimatore di $\mathbf{x}(t+1)$, (4.23) e (4.25), e in una parte "fuori linea" che non elabora le misure ma solo dati a priori del modello, descritta dalle equazioni (4.24), (4.26), (4.28), (4.29) e dalla condizione iniziale P_0 .

La parte fuori linea serve innanzitutto per un'analisi a priori del comportamento statistico del filtro, basata sull'andamento delle varianze degli errori di stima $P(t+1 | t)$ e $P(t | t)$, ma soprattutto a calcolare la successione dei guadagni $\{L(t) ; t \geq t_0\}$ che dovranno essere memorizzati e usati successivamente man mano che si elaborano le misure.

Notiamo che alla parte "fuori linea" dell'algoritmo compete sostanzialmente tutto il carico computazionale del filtro. È bene mettere subito in evidenza che l'equazione di aggiornamento (4.26) può dare luogo a seri problemi di carattere numerico, specialmente quando i due termini a secondo membro sono entrambi piccoli. In questo caso, a causa di fenomeni di accumulazione degli errori di arrotondamento, si finisce per distruggere la proprietà (essenziale) delle matrici $P(t | t)$ e $P(t+1 | t)$ di essere *semidefinite positive* e simmetriche. Questo fatto, come è facilmente intuibile, può avere conseguenze assai pericolose, tipicamente instabilità e "divergenza" delle stime.

È importante dunque trovare delle forme alternative della (4.26) meno sensibili agli errori di round-off. La variante più semplice è la cosiddetta *forma simmetrizzata*

$$\begin{aligned} P(t+1 | t+1) &= [I - L(t+1)C] P(t+1 | t) [I - L(t+1)C]' \\ &\quad + L(t+1)RL(t+1)', \end{aligned} \quad (4.35)$$

che si ricava agevolmente tenendo conto della definizione di $L(t)$ data nella (4.29). In questa espressione si sommano due quantità che sono certamente simmetriche e semidefinite positive. Si ha inoltre anche un miglioramento della sensitività dell'algoritmo rispetto agli errori di calcolo nel guadagno. Infatti perturbando il guadagno

$L(t+1)$ di ∂L si ha, usando l'espressione (4.35), una variazione (del prim'ordine) di $P(t+1 | t+1)$ che è *nulla*, mentre alla (4.26), che può essere sinteticamente riscritta come

$$P(t+1 | t+1) = \left(I - L(t+1)C \right) P(t+1 | t) \quad ,$$

corrisponde una variazione *diversa da zero*,

$$\partial P(t+1 | t+1) = -\partial LCP(t+1 | t) \quad .$$

Gli aspetti numerici dell'algoritmo di Kalman sono stati ampiamente discussi in letteratura; si veda ad esempio [20, 140].

Dinamica del Predittore e del Filtro. L'algoritmo di Kalman 9.2 può essere benissimo impiegato nella sua forma originale anche quando in realtà interessa calcolare solo il predittore di un passo $\hat{\mathbf{x}}(t+1 | t)$ oppure il "filtro" $\mathbf{x}(t | t)$. Per certi versi è però utile cercare di eliminare dalle equazioni le stime che non interessano, se non altro perché in questo modo si ottiene una percezione molto più chiara del modo in cui vengono elaborate le misure e in generale della struttura del filtro.

Per ricavare un algoritmo che *aggiorna solo la predizione di un passo dello stato* è sufficiente eliminare $\hat{\mathbf{x}}(t | t)$ nelle equazioni (4.23) e (4.25). Dopo semplici passaggi si ottiene il seguente risultato.

ALGORITMO 9.3 (Predittore di Kalman). *Il predittore di un passo dello stato del sistema (4.1), si può calcolare mediante un algoritmo ricorsivo descritto dall'equazione*

$$\hat{\mathbf{x}}(t+1 | t) = F\hat{\mathbf{x}}(t | t-1) + SR^{-1} \mathbf{y}(t) + K(t) [\mathbf{y}(t) - C\hat{\mathbf{x}}(t | t-1)] \quad , \quad (4.36)$$

in cui

$$K(t) := FL(t) = FP(t | t-1) C' \Lambda^{-1}(t) \quad (4.37)$$

dove $\Lambda(t) = CP(t | t-1) C' + R$ è la varianza dell'innovazione $\mathbf{e}(t) := \mathbf{y}(t) - C\hat{\mathbf{x}}(t | t-1)$. L'algoritmo si può rappresentare graficamente mediante lo schema a blocchi di figura 4.1. Questo è uno schema a controreazione in cui compare un ramo di

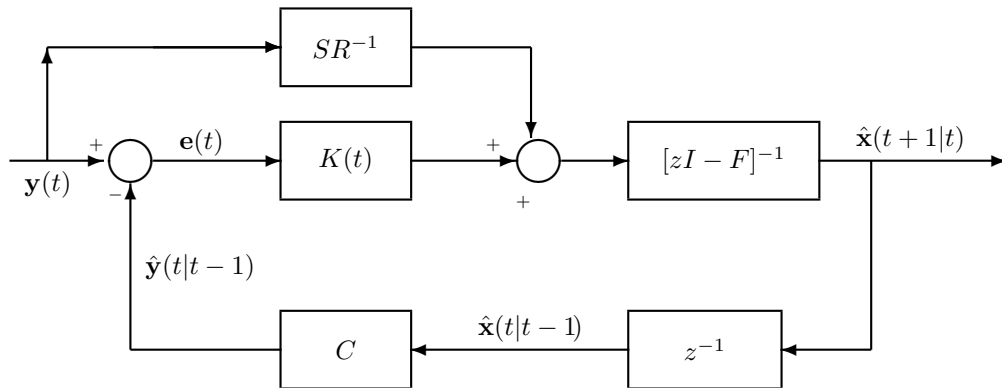


FIGURA 4.1. Schema a blocchi del predittore di Kalman.

“azione diretta” dell’ingresso $\mathbf{y}(t)$ dovuto all’effetto della correlazione tra i rumori di misura e di modello.

L’algoritmo (4.36) si può equivalentemente riscrivere nella forma

$$\hat{\mathbf{x}}(t+1 | t) = A\hat{\mathbf{x}}(t | t-1) + G(t) [\mathbf{y}(t) - C\hat{\mathbf{x}}(t | t-1)] \quad , \quad (4.38)$$

dove compare la matrice originale A del modello (4.1) e si è introdotto il guadagno del predittore,

$$G(t) := K(t) + SR^{-1} \quad . \quad (4.39)$$

questo schema corrisponde ancora ad una struttura a controreazione come quello di figura 4.3 in cui l’azione diretta non è più evidente ma è comunque implicita nell’espressione (4.39) del guadagno.

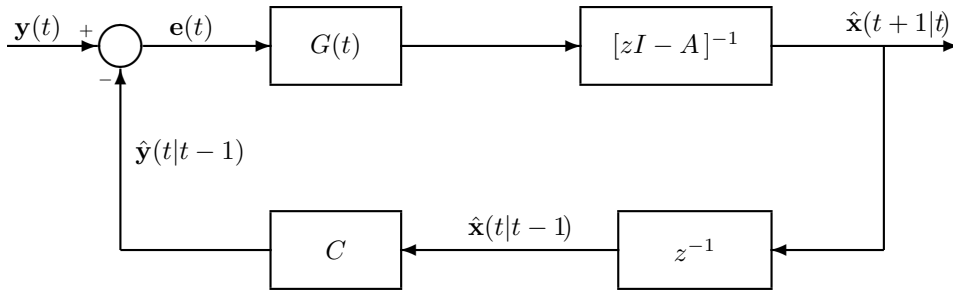


FIGURA 4.2. Schema alternativo del predittore di Kalman.

La matrice $K(t)$ oppure il guadagno del predittore $G(t)$ si possono calcolare fuori linea, noti i parametri del sistema e la matrice varianza dell’errore di predizione di un passo, $P(t | t-1)$. Quest’ultima si può aggiornare mediante l’equazione alle differenze

$$\begin{aligned} P(t+1 | t) &= F [P(t | t-1) - P(t | t-1) C' \Lambda^{-1}(t) C P(t | t-1)] F' + \tilde{Q} \\ P(t_0 | t_0 - 1) &= P_0 \quad . \end{aligned} \quad (4.40)$$

Anche la matrice $K(t)$ verrà talvolta chiamata impropriamente “guadagno” del predittore⁵).

L’equazione di aggiornamento della covarianza d’errore (4.40), che viene denominata *equazione di Riccati discreta*, per analogia con quella che si incontra nella teoria del filtraggio a tempo continuo, si ottiene facilmente combinando la (4.24) e la (4.26). Questo modo di aggiornare $P(t | t-1)$ soffre però degli stessi problemi numerici a cui si è accennato poco più sopra riguardo all’equazione (4.26). Per questo motivo la (4.40) viene usata solo a scopo di analisi. Nelle applicazioni pratiche conviene sempre ricorrere a schemi simmetrizzati o a varianti numericamente più stabili.

La versione simmetrizzata della (4.40) ha la seguente forma:

$$P(t+1 | t) = F (I - L(t) C) P(t | t-1) (I - L(t) C)' F' + FL(t) RL'(t) F' + \tilde{Q} \quad . \quad (4.41)$$

⁵Si ha $G(t) = K(t)$ nel caso di rumori scorrelati.

in cui si vede bene che si sommano solo termini simmetrici e semidefiniti positivi. Una semplice regola mnemonica per ricordare questa espressione è di far riferimento all'equazione dinamica per l'errore di predizione $\tilde{\mathbf{x}}(t+1 | t)$ che si ricava sottraendo la (4.36) dalla (4.21)

$$\begin{aligned}\tilde{\mathbf{x}}(t+1 | t) &= F\tilde{\mathbf{x}}(t | t-1) - FL(t) [\mathbf{y}(t) - C\hat{\mathbf{x}}(t | t-1)] + \tilde{\mathbf{v}}(t) \\ &= F\tilde{\mathbf{x}}(t | t-1) - FL(t) C\tilde{\mathbf{x}}(t | t-1) - FL(t) \mathbf{w}(t) + \tilde{\mathbf{v}}(t) \\ &= F(I - L(t)C) \tilde{\mathbf{x}}(t | t-1) - FL(t) \mathbf{w}(t) + \tilde{\mathbf{v}}(t) \quad .\end{aligned}$$

La (4.41) si può ottenere semplicemente calcolando la varianza dell'errore come somma delle varianze dei termini nell'ultimo membro di questa equazione (il lettore dovrebbe controllare che tutti i termini nell'ultima riga sono tra loro scorrelati).

Allo scopo di compattare un pochino le formule, conviene introdurre le quantità

$$\Gamma(t) := F - K(t)C \quad , \quad (4.42)$$

col che l'equazione di Riccati simmetrizzata assume l'aspetto

$$\begin{aligned}P(t+1 | t) &= \Gamma(t) P(t | t-1) \Gamma'(t) + K(t) RK'(t) + \tilde{Q} \\ P(t_0 | t_0 - 1) &= P_0 \quad .\end{aligned} \quad (4.43)$$

La matrice $\Gamma(t)$ ha un'importanza notevole nell'analisi del comportamento del predittore perché essa descrive la dinamica dell'errore di predizione, nel senso che si ha

$$\tilde{\mathbf{x}}(t+1 | t) = \Gamma(t) \tilde{\mathbf{x}}(t | t-1) - K(t) \mathbf{w}(t) + \tilde{\mathbf{v}}(t) \quad . \quad (4.44)$$

Ma il significato di $\Gamma(t)$ è soprattutto legato alla interpretazione dell'algoritmo di predizione (4.36) o (4.38) come il *sistema a controreazione* di figura ??

In questo schema *la matrice che descrive il comportamento dinamico del sistema in anello chiuso è proprio $\Gamma(t)$* , dato che,

$$A - G(t)C = A - K(t)C - SR^{-1}C = F - K(t)C = \Gamma(t) \quad , \quad (4.45)$$

come si verifica facilmente dalla definizione (4.39) di $G(t)$.

Se $\Gamma(t)$ descrive un sistema asintoticamente stabile, la struttura a controreazione del predittore gli conferisce i ben noti vantaggi di insensibilità alle variazioni parametriche e ad eventuali errori di modellizzazione e troncamento dei dati che si studiano nella teoria dei sistemi di controllo. Per questa ragione buona parte dello studio qualitativo del predittore di Kalman (che faremo più avanti) è, in ultima analisi, dedicata proprio alla stabilità (asintotica) dello schema a retroazione descritto dalla matrice $\Gamma(t)$.

La stima filtrata. Le formule per l'aggiornamento della stima filtrata $\hat{\mathbf{x}}(t | t)$ sono simili a quelle del predittore. Premettiamo un risultato ausiliario.

LEMMA 9.4. *Si consideri la stima filtrata $\hat{\mathbf{x}}(t | t)$ dello stato del modello lineare (4.1), assumendo che valgano le ipotesi standard, in particolare con rumori di modello e di misura possibilmente correlati. Il processo di errore del filtro*

$$\hat{\mathbf{e}}(t) := \mathbf{y}(t) - C\hat{\mathbf{x}}(t | t)$$

che per analogia chiameremo anche innovazione del filtro, soddisfa alla relazione

$$\hat{\mathbf{e}}(t) = (I - CL(t))\mathbf{e}(t) = R\Lambda(t)^{-1}\mathbf{e}(t) \quad (4.46)$$

dove $\mathbf{e}(t)$ è l'innovazione (del predittore) ed è quindi anch'esso rumore bianco. Gli altri simboli hanno il significato usuale. La varianza di $\hat{\mathbf{e}}(t)$ è uguale a $R\Lambda(t)^{-1}R$ e si ha

$$\Lambda(t) - R\Lambda(t)^{-1}R > 0 \quad (\text{definita positiva}) \quad (4.47)$$

per cui $\text{Var } \hat{\mathbf{e}}(t) \leq \text{Var } \mathbf{e}(t)$.

DIMOSTRAZIONE. Riscrivendo la (4.25) con t al posto di $t+1$ si trova

$$\mathbf{y}(t) - C\hat{\mathbf{x}}(t|t) = \mathbf{y}(t) - C\hat{\mathbf{x}}(t|t-1) - CL(t)[\mathbf{y}(t) - C\hat{\mathbf{x}}(t|t-1)]$$

e stante l'identità $C[I - L(t)C] = [I - CL(t)]C$ si trova

$$\hat{\mathbf{e}}(t) = \mathbf{y}(t) - C\hat{\mathbf{x}}(t|t) = [I - CL(t)][\mathbf{y}(t) - C\hat{\mathbf{x}}(t|t-1)] = [I - CL(t)]\mathbf{e}(t)$$

che è la formula (4.46). Da questa poi scende la disuguaglianza matriciale (4.47) data la nonnegatività della varianza d'errore dello stimatore $\mathbb{E}[\mathbf{e}(t) | \hat{\mathbf{e}}(t)]$ ovvero

$$\text{Var } \mathbf{e}(t) - \text{Cov}[\mathbf{e}(t), \hat{\mathbf{e}}(t)] \text{Var} \{\hat{\mathbf{e}}(t)\}^{-1} \text{Cov}[\hat{\mathbf{e}}(t), \mathbf{e}(t)] > 0.$$

Il significato intuitivo di questa disuguaglianza è ovvio. \square

Sostituendo nella (4.25) si trova un'espressione del filtro pilotato dal rumore bianco di innovazione (del filtro)

$$\hat{\mathbf{x}}(t+1|t+1) = F\hat{\mathbf{x}}(t|t) + SR^{-1}\mathbf{y}(t) + \hat{L}(t+1)\hat{\mathbf{e}}(t+1) \quad (4.48)$$

dove

$$\hat{L}(t+1) := P(t+1|t)C'R^{-1}$$

si può chiamare *guadagno del filtro*. Queste equazioni corrispondono ad uno schema a retroazione simile a quello del predittore di figura 4.1 dove occorrerà ovviamente sostituire $\hat{\mathbf{x}}(t|t-1)$ con $\hat{\mathbf{x}}(t|t)$ e $L(t)$ con $\hat{L}(t)$. Questo è uno schema a controreazio-

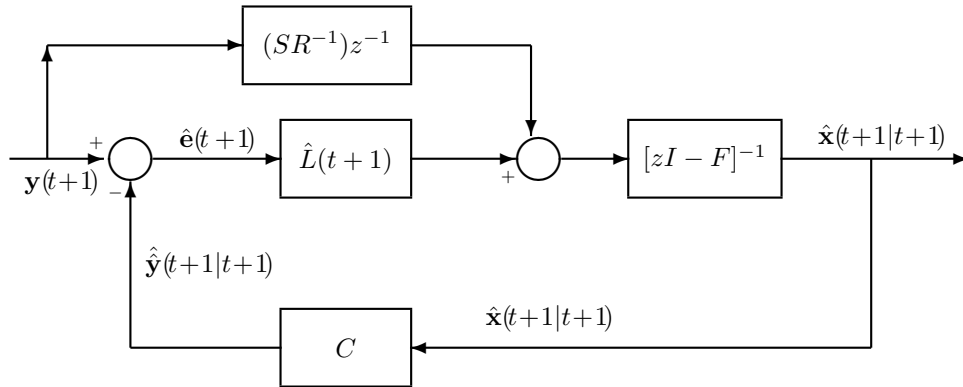


FIGURA 4.3. Schema a blocchi del Filtro di Kalman.

ne in cui compare un ramo di "azione diretta" dell'ingresso $\mathbf{y}(t)$ dovuto all'effetto della correlazione tra i rumori di misura e di modello.

Riporteremo qui sotto le formule relative al caso di rumori scorrelati ($S = 0$) in cui compaiono i parametri originali del modello.

ALGORITMO 9.4 (Stima Filtrata). *La stima “filtrata” $\hat{\mathbf{x}}(t | t)$ si può aggiornare mediante l’algoritmo ricorsivo,*

$$\hat{\mathbf{x}}(t+1 | t+1) = F\hat{\mathbf{x}}(t | t) + L(t+1) [\mathbf{y}(t+1) - CF\hat{\mathbf{x}}(t | t)] \quad . \quad (4.49)$$

dove il guadagno del filtro, $L(t+1)$, è in questo caso dato dalla formula

$$\begin{aligned} L(t+1) &= P(t+1 | t) C' \Lambda(t+1)^{-1} \\ \Lambda(t+1) &= CP(t+1 | t) C' + R \end{aligned}$$

e $P(t+1 | t)$ si può calcolare risolvendo l’equazione di Riccati (4.40) oppure la forma simmetrizzata (4.43).

Se i rumori di modello e di osservazione sono scorrelati ($S = 0$), la stima “filtrata” $\hat{\mathbf{x}}(t | t)$ si può aggiornare mediante l’algoritmo ricorsivo,

$$\hat{\mathbf{x}}(t+1 | t+1) = A\hat{\mathbf{x}}(t | t) + L(t+1) [\mathbf{y}(t+1) - CA\hat{\mathbf{x}}(t | t)] \quad . \quad (4.50)$$

dove il guadagno del filtro, $L(t+1)$, è in questo caso dato dalla formula

$$\begin{aligned} L(t+1) &= P(t+1 | t) C' \Lambda(t+1)^{-1} \\ \Lambda(t+1) &= CP(t+1 | t) C' + R \end{aligned}$$

e $P(t+1 | t)$ si può calcolare risolvendo l’equazione di Riccati (4.40) oppure la forma simmetrizzata (4.43).

La varianza dell’errore di filtraggio si può ricavare da $P(t+1 | t)$ mediante la formula

$$P(t+1 | t+1) = [I - L(t+1)C] P(t+1 | t) \quad . \quad (4.51)$$

Come si vede, anche il filtro (vero e proprio) ha struttura a controreazione. La matrice che ne determina il comportamento dinamico in catena chiusa è (per $S = 0$) data dalle formule

$$\Gamma_1(t) := A - L(t+1)CA = [I - L(t+1)C] A \quad . \quad (4.52)$$

Si noti che nelle stesse condizioni ($S = 0$) la matrice $\Gamma(t)$ che entra nel predittore vale

$$\Gamma(t) = A[I - L(t)C] \quad .$$

Si intuisce come il comportamento asintotico dei due algoritmi (stabilità ecc.) sia sostanzialmente identico.

Conviene ribadire ancora una volta che tutto quanto è stato detto fino a questo punto non richiede affatto l’invarianza temporale delle matrici (A, C, Q, R, S).

Questo non per accademico amore di “generalità”, ma perché esistono situazioni concrete in cui è assai conveniente modellare i segnali mediante schemi varianti nel tempo. Casi tipici sono quelli in cui si cambia il “sensore”, ovvero il dispositivo che misura lo stato. Si può ad esempio decidere ad un certo punto di misurare qualche variabile di stato in più allo scopo di migliorare l’accuratezza della stima. In questi casi basta modellare C come una matrice variante nel tempo nella quale gli elementi corrispondenti alle nuove variabili di stato sono posti uguali a zero per tutti i tempi precedenti l’istante in cui si cominciano ad acquisire le nuove componenti.

Conviene anche richiamare il fatto ben noto che se le distribuzioni di probabilità in gioco sono *Gaussiane*, lo stimatore lineare a minima varianza diventa automaticamente lo stimatore a minima varianza in senso assoluto, ovvero la funzione (non lineare) dei dati che ha varianza d’errore più piccola possibile. In particolare se i rumori di modello, di osservazione e la condizione iniziale sono congiuntamente

Gaussiani, il filtro di Kalman fornisce le stime ottime in senso assoluto, cioè le *medie condizionate dello stato date le osservazioni passate*.

L' equazione di Riccati nei parametri originali. In letteratura si trova spesso l'equazione di Riccati espressa mediante i parametri originali del modello (4.1). Per ricavare questa espressione senza eseguire il disaccoppiamento del rumore di modello da quello di misura, si ragiona come segue.

Siano $\Sigma(t)$ e $\hat{\Sigma}(t)$ le matrici varianza dello stato del modello, $\mathbf{x}(t)$, e del predittore di un passo $\hat{\mathbf{x}}(t | t-1)$. Dal principio di ortogonalità si ottiene la decomposizione

$$P(t | t-1) = E[\mathbf{x}(t) - \hat{\mathbf{x}}(t | t-1)] [\mathbf{x}(t) - \hat{\mathbf{x}}(t | t-1)]' = \Sigma(t) - \hat{\Sigma}(t) \quad , \quad (4.53)$$

dove $\Sigma(t)$ e $\hat{\Sigma}(t)$ soddisfano le equazioni alle differenze

$$\Sigma(t+1) = A \Sigma(t) A' + Q \quad , \quad \Sigma(t_0) = \Sigma_0 \quad , \quad (4.54)$$

$$\hat{\Sigma}(t+1) = A \hat{\Sigma}(t) A' + G(t) \Lambda(t) G(t)' \quad , \quad \hat{\Sigma}(t_0) = 0 \quad . \quad (4.55)$$

la seconda delle quali scende agevolmente calcolando la varianza dalla (4.38). Si trova così

$$P(t+1 | t) = AP(t | t-1)A' - G(t)\Lambda(t)G(t)' + Q \quad (4.56)$$

che va associata all'espressione di $G(t)$

$$G(t) = [AP(t | t-1)C' + S] \Lambda(t)^{-1} \quad , \quad (4.57)$$

che si può ricavare dalla (4.39) attraverso i passaggi algebrici che riportiamo qui di seguito.

$$\begin{aligned} G(t) \Lambda(t) &= (A - SR^{-1}C) P(t | t-1) C' + SR^{-1} \Lambda(t) \\ &= AP(t | t-1) C' - SR^{-1} CP(t | t-1) C' + SR^{-1} \Lambda(t) \\ &= AP(t | t-1) C' + S \quad . \end{aligned}$$

Questa forma dell' equazione di Riccati è forse quella che si incontra più spesso in letteratura.

ESEMPIO 9.2 (Filtro di Kalman di un processo stazionario descritto da un modello ARMA). Lo scopo di questo esempio è di discutere il comportamento asintotico del filtro di Kalman in un semplice esempio per il quale abbiamo a suo tempo applicato la teoria di Wiener-Kolmogorov. Molti dei risultati che ricaveremo qui dall'analisi diretta dell'equazione di Riccati, sono conseguenze ovvie del Teorema Fondamentale sul comportamento asintotico del filtro, che illustreremo nel prossimo capitolo.

Supponiamo di voler costruire il predittore di un processo stazionario $\{\mathbf{y}(t)\}$ descritto dal modello ARMA

$$\mathbf{y}(t) + \sum_1^n a_k \mathbf{y}(t-k) = \mathbf{w}(t) + \sum_1^n c_k \mathbf{w}(t-k) \quad (4.58)$$

in cui $\{\mathbf{w}(t)\}$ è bianco di varianza σ^2 . I dati sono disponibili a partire dall'istante $t = 0$.

Come già visto, $\{\mathbf{y}(t)\}$ è un processo di dimensione finita (n) e una sua possibile realizzazione è la

$$\mathbf{x}(t+1) = \begin{bmatrix} 0 & & -a_n \\ 1 & & \\ & \ddots & \\ & & 1 & -a_1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} c_n - a_n \\ \vdots \\ c_1 - a_1 \end{bmatrix} \mathbf{w}(t) \quad (4.59a)$$

$$\mathbf{y}(t) = [0 \dots 0 \ 1] \mathbf{x}(t) + \mathbf{w}(t) \quad (4.59b)$$

associata ad uno stato iniziale $\mathbf{x}(0)$ di media zero e matrice varianza P_0 (calcolabile a partire dal modello).

Scriveremo questa rappresentazione simbolicamente come

$$\begin{aligned} \mathbf{x}(t+1) &= A\mathbf{x}(t) + b\mathbf{w}(t) \\ \mathbf{y}(t) &= c\mathbf{x}(t) + \mathbf{w}(t) \end{aligned} ,$$

dove le quantità (A, b, c) corrispondono alla parametrizzazione usata nelle (4.59).

Notiamo che i rumori di modello e di misura sono correlati. Le varianze si calcolano immediatamente e risultano essere

$$Q = \sigma^2 bb' \quad , \quad S = \sigma^2 b \quad , \quad R = \sigma^2 \quad , \quad (4.60)$$

per cui:

$$\tilde{Q} = 0 \quad .$$

Notiamo anche che:

$$F = A - SR^{-1}C = \begin{bmatrix} 0 & & -c_n \\ 1 & & \\ & \ddots & \\ & & 1 & -c_1 \end{bmatrix} \quad , \quad (4.61)$$

che è la matrice compagna associata al polinomio a numeratore del modello (4.58).

L'equazione di Riccati per la varianza dell'errore di predizione di stato, $P(t) := P(t | t-1)$, si scrive così

$$\begin{aligned} P(t+1) &= F \left[P(t) - P(t) c' \frac{1}{cP(t) c' + \sigma^2} cP(t) \right] F' \\ P(0) &= P_0 \quad . \end{aligned} \quad (4.62)$$

ed è un'equazione omogenea.

Supponiamo ora che $C(z)$ abbia *tutti gli zeri dentro il cerchio unità* $\{|z| < 1\}$, nel qual caso come ben sappiamo, il modello ARMA è di innovazione.

Dall'equazione di Riccati si ricava direttamente la disuguaglianza

$$P(t) \leq FP(t-1)F' \leq \dots \leq F^t P(0) (F')^t$$

dalla quale, passando al limite per $t \rightarrow \infty$, si vede che per l'asintotica stabilità di F , $P(t) \rightarrow 0$, qualunque sia la varianza iniziale P_0 .

Se invece $C(z)$ non ha tutti gli zeri dentro il cerchio unità, si può ragionare come segue. Usando il lemma d'inversione di matrice si può scrivere

$$P(t) - P(t) c' [cP(t) c' + \sigma^2]^{-1} cP(t) = \left[P(t)^{-1} + c' \frac{1}{\sigma^2} c \right]^{-1} .$$

Questa relazione vale se $P(t)$ è invertibile e questo accade certamente se P_0 è invertibile e F è non singolare.

La prima condizione è verificata se (A, b) è raggiungibile ($A(z)$ e $C(z)$ coprimi). In questo caso infatti, $P_0 > 0$ giacchè la varianza di stato è soluzione dell'equazione di Lyapunov $P_0 = AP_0A' + \sigma^2bb'$. La seconda richiede che il polinomio $C(z) = z^n + c_1 z^{n-1} + \dots + c_n$ non abbia zeri nell'origine, ovvero che $c_n \neq 0$. Assumendo che queste condizioni siano verificate e ragionando per induzione sull'equazione (4.62), riscritta come

$$P(t+1) = F [P(t)^{-1} + c'c/\sigma^2]^{-1} F' \quad , \quad (4.63)$$

si vede poi che $P(t) > 0, \forall t$ se F è invertibile.

Prendendo l'inversa in entrambi i membri di (4.63) si trova

$$\begin{aligned} P(t+1)^{-1} &= (F')^{-1} P(t)^{-1} F^{-1} + (F')^{-1} \frac{c'c}{\sigma^2} F^{-1} \\ P^{-1}(0) &= P_0^{-1} \quad , \end{aligned}$$

che è un'equazione lineare e può essere risolta ottenendo

$$P(t)^{-1} = (F')^{-t} P_0^{-1} F^{-t} + \sum_1^t (F')^{-k} \frac{c'c}{\sigma^2} F^{-k} \quad .$$

Questa formula può essere utile per esaminare qualitativamente il comportamento asintotico di $P(t)$ per $t \rightarrow \infty$. Per esempio, se tutti gli zeri di $C(z)$ stanno fuori dal cerchio unità, ovvero $|\lambda_i(F)| > 1, i = 1, \dots, n$, si ha:

$$\lim_{t \rightarrow \infty} P(t)^{-1} = \sum_1^{\infty} (F^{-\top})^k \frac{c'c}{\sigma^2} (F^{-1})^k < \infty \quad ,$$

che è una quantità *definita positiva* (per l'osservabilità di (F, c)). Quindi, se $C(z)$ è "totalmente instabile",

$$\lim_{t \rightarrow \infty} P(t) = \bar{P} > 0 \quad .$$

Notiamo che se il modello di partenza è di innovazione, $P(t) \rightarrow 0$ e il guadagno del predittore

$$g(t) = [AP(t)c' + \sigma^2 b] \left(cP(t)c' + \sigma^2 \right)^{-1}$$

tende al vettore b ,

$$\lim_{t \rightarrow \infty} g(t) = b$$

Equivalentemente, $k(t) \rightarrow 0$. Ne risulta che l'equazione ricorsiva per il predittore dello stato può scriversi, per $t \rightarrow \infty$, nella forma tempo-invariante

$$\hat{\mathbf{x}}(t+1 | t) = F\hat{\mathbf{x}}(t | t-1) + b\mathbf{y}(t) = A\hat{\mathbf{x}}(t | t-1) + b\mathbf{e}(t) \quad , \quad (4.64)$$

la seconda delle quali è formalmente identica alla (4.59).

Ricaviamo ora la funzione di trasferimento $\mathbf{y}(t) \rightarrow \hat{\mathbf{y}}(t+1 | t)$. Notiamo allo scopo che il predittore di un passo di $\{\mathbf{y}(t)\}$ si scrive

$$\hat{\mathbf{y}}(t+1 | t) = c\hat{\mathbf{x}}(t+1 | t) \quad .$$

Tenendo conto che F è espressa dalla relazione (4.61), si trova

$$\begin{aligned}\hat{\mathbf{y}}(t+1|t) &= [cz(zI - F)^{-1}b] \mathbf{y}(t) \\ &= \frac{z[(c_1 - a_1)z^{-1} + \dots + (c_n - a_n)z^{-n}]}{1 + c_1 z^{-1} + \dots + c_n z^{-n}} \mathbf{y}(t) \\ &= \frac{z[C^*(z^{-1}) - A^*(z^{-1})]}{C^*(z^{-1})} \mathbf{y}(t) \\ &= \frac{C_1^*(z^{-1})}{C^*(z^{-1})} \mathbf{y}(t)\end{aligned}$$

che coincide con quella ricavata usando la teoria di Wiener-Kolmogorov.

Se il modello ARMA non è di innovazione, ad esempio se $C(z)$ è antistabile, la soluzione a regime trovata (\bar{P}) deve necessariamente corrispondere allo stesso predittore di Wiener. In un certo senso, risolvere l'equazione di Riccati deve perciò corrispondere ad eseguire la fattorizzazione spettrale che occorre fare per ottenere il modello ARMA d'innovazione. Per verificare che le cose stanno proprio così dovremo pazientare fino alla fine del prossimo capitolo. \diamond

Il Filtro di Kalman con Ingressi. In molti problemi di stima che si incontrano in pratica è necessario considerare anche la presenza di variabili di ingresso, o, più in generale di variabili esogene misurabili, ad esempio variabili di comando o regolazione nei sistemi di controllo. Supponendo che queste agiscano linearmente sullo stato, si può fare riferimento in questi casi ad un modello lineare del tipo

$$\begin{cases} \mathbf{x}(t+1) &= A\mathbf{x}(t) + B\mathbf{u}(t) + \mathbf{v}(t) \\ \mathbf{y}(t) &= C\mathbf{x}(t) + \mathbf{w}(t) \end{cases}, \quad t \geq t_0, \quad (4.65)$$

in cui $\mathbf{u}(t)$ è la variabile esogena (o di ingresso), che descriviamo come un processo del secondo ordine. A priori \mathbf{u} potrebbe benissimo essere correlato con altre variabili del modello. Per il resto, assumiamo che valgono le ipotesi standard e che tutti gli altri simboli abbiano il significato usuale.

PROBLEMA 9.2. Si assuma che $\mathbf{u}(t)$ sia funzione lineare della storia passata dell'uscita \mathbf{y} nell'intervallo $[t_0, t]$. Derivare le equazioni di aggiornamento per le stime dello stato $\hat{\mathbf{x}}(t+1|t)$, $\hat{\mathbf{x}}(t|t)$, in base all'osservazione della storia passata dell'uscita $\{\mathbf{y}(s) | t_0 \leq s \leq t\}$. Verificare che valgono gli stessi ragionamenti fatti per la derivazione dell'algoritmo di Kalman senza ingressi e le equazioni dell'algoritmo sono le stesse appena ricavate, eccettuata la (4.23), che assume ora la forma,

$$\hat{\mathbf{x}}(t+1|t) = F\hat{\mathbf{x}}(t|t) + SR^{-1}\mathbf{y}(t) + B\mathbf{u}(t). \quad (4.66)$$

\diamond

L'ipotesi che \mathbf{u} sia esattamente esprimibile come funzione lineare e causale dall'uscita è restrittiva. In genere ci sono altri segnali che concorrono a determinare l'andamento temporale dell'ingresso oltre a quello della variabile di uscita. Una situazione abbastanza realistica si ha quando nel modello lineare (4.65) l'ingresso è generato mediante una legge di controllo causale del tipo

$$\mathbf{u}(t) = H(\mathbf{y}^t) + \mathbf{r}(t) \quad (4.67)$$

dove H è una funzione lineare della storia passata e presente del processo \mathbf{y} e \mathbf{r} è un processo del secondo ordine *scorrelato da tutta la storia dei rumori* \mathbf{v} , \mathbf{w} del

modello (4.65). Se in particolare, $H \equiv 0$, l'ingresso \mathbf{u} è esso stesso scorrelato dai rumori di modello e di osservazione \mathbf{v} , \mathbf{w} . Notiamo però che non si può in generale assumere che gli ingressi futuri $\{\mathbf{u}; t \geq t_0\}$, siano scorrelati dallo stato iniziale $\mathbf{x}(t_0)$. Questo accade perchè lo stato nel modello (4.65) è una funzione causale degli ingressi passati, in particolare degli ingressi \mathbf{u} che sono stati applicati prima dell'istante t_0 e la storia passata e futura di \mathbf{u} sono in generale correlate (a meno che \mathbf{u} non sia un processo bianco).

D'ora in avanti supporremo che la storia passata della variabile di ingresso $\{\mathbf{u}(s) | t_0 \leq s \leq t\}$ non sia contenuta in $\mathbf{H}_t(\mathbf{y})$. Supporremo però che \mathbf{u} sia *osservabile* e quindi il passato $\{\mathbf{u}(s) | t_0 \leq s \leq t\}$ sia noto ad ogni istante t . Si vogliono trovare le equazioni del Filtro di Kalman che descrivono l'aggiornamento delle stime dello stato basate sull'osservazione *congiunta* dell'uscita \mathbf{y} e dell'ingresso \mathbf{u} fino all'istante t . Per ricavare queste equazioni è necessario in generale "aumentare" il modello (4.65) aggiungendo una descrizione della dinamica del processo di ingresso. Richiamandoci anche a quanto esposto nel Capitolo 7, considereremo allora il modello di stato congiunto ottenuto accoppiando (4.65) al modello del canale di reazione che genera l'ingresso (che supporremo di dimensione finita).

$$\begin{cases} \mathbf{x}_1(t+1) = \mathbf{A}\mathbf{x}_1(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{v}_1(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}_1(t) + \mathbf{w}_1(t) \end{cases} \quad (4.68)$$

$$\begin{cases} \mathbf{x}_2(t+1) = \mathbf{F}\mathbf{x}_2(t) + \mathbf{G}\mathbf{y}(t) + \mathbf{v}_2(t) \\ \mathbf{u}(t) = \mathbf{H}\mathbf{x}_2(t) + \mathbf{J}\mathbf{y}(t) + \mathbf{w}_2(t). \end{cases} \quad (4.69)$$

Per uniformità di notazioni abbiamo sostituito il simbolo \mathbf{x} (che denoterà lo stato aumentato) con \mathbf{x}_1 , \mathbf{v} con \mathbf{v}_1 (di varianza \mathbf{Q}_1), \mathbf{w} con \mathbf{w}_1 etc.. Per il modello che descrive la dinamica di \mathbf{u} abbiamo usato notazioni congruenti. I rumori bianchi con pedici diversi sono tra loro completamente scorrelati. Aumentando lo stato si ricava così il modello che descrive il processo congiunto delle osservazioni in forma standard,

$$\mathbf{x}(t+1) = \begin{bmatrix} \mathbf{A} + \mathbf{B}\mathbf{J}\mathbf{C} & \mathbf{B}\mathbf{H} \\ \mathbf{G}\mathbf{C} & \mathbf{F} \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} \bar{\mathbf{v}}_1(t) \\ \bar{\mathbf{v}}_2(t) \end{bmatrix} := \mathcal{A}\mathbf{x}(t) + \bar{\mathbf{v}}(t) \quad (4.70a)$$

$$\begin{bmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{J}\mathbf{C} & \mathbf{H} \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{J} & \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \mathbf{w}_1(t) \\ \mathbf{w}_2(t) \end{bmatrix} := \mathcal{C}\mathbf{x}(t) + \mathcal{D}\mathbf{w}(t) \quad (4.70b)$$

dove

$$\begin{bmatrix} \bar{\mathbf{v}}_1 \\ \bar{\mathbf{v}}_2 \end{bmatrix} := \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{B}\mathbf{J} & \mathbf{B} \\ \mathbf{G} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} := \mathbf{v} + \mathcal{B}\mathbf{w} \quad (4.71)$$

È allora ovvio che la stima dello stato del modello (4.68), ad esempio il predittore di un passo $\hat{\mathbf{x}}_1(t+1 | t)$, ottenuto proiettando lo stato $\mathbf{x}_1(t+1)$ sullo spazio $\mathbf{H}_t(\mathbf{y}, \mathbf{u})$, si può calcolare prendendo le prime n_1 componenti del predittore dello stato del modello congiunto (che ha dimensione n), basato sulla storia passata delle variabili di ingresso e di uscita $\{\mathbf{u}(s), \mathbf{y}(s) | t_0 \leq s \leq t\}$.

Naturalmente la dinamica di \mathbf{u} potrebbe non essere nota e comunque l'uso del modello aumentato porta ad un notevole aumento del carico computazionale. Ci si chiede se sia possibile, in analogia a quanto accade nel caso stazionario, calcolare il predittore di un passo $\hat{\mathbf{x}}_1(t+1 | t)$, usando solo il modello (4.65), assumendo noto che \mathbf{u} è generato da un canale di reazione causale come in (4.67), senza però conoscere la struttura di \mathbf{H} nè di \mathbf{r} . In particolare ci poniamo la seguente questione: è in

generale corretto affermare che le equazioni del filtro si modificano semplicemente sostituendo la (4.23), con la (4.66), senza fare alcuna ipotesi ulteriore su \mathbf{u} ?

Purtroppo per t_0 finito la risposta a questa domanda è in generale negativa, a meno che nella (4.67) non sia $H \equiv 0$, il che corrisponde intuitivamente all' *assenza di reazione* (sull'intervallo semi-infinito $[t_0, +\infty)$). In questo caso, definiamo

$$\hat{\mathbf{x}}(t_0) := \hat{E} [\mathbf{x}(t_0) | \mathbf{H}(\mathbf{u})], \quad \tilde{\mathbf{x}}(t_0) := \mathbf{x}(t_0) - \hat{\mathbf{x}}(t_0) \quad (4.72)$$

dove il sottospazio $\mathbf{H}(\mathbf{u})$ è generato da tutta la storia di \mathbf{u} ,

$$\mathbf{H}(\mathbf{u}) := \overline{\text{span}} \{ \mathbf{u}(s); s \geq t_0 \} = \overline{\text{span}} \{ \mathbf{r}(s); s \geq t_0 \}$$

È facile allora convincersi che vale la decomposizione ortogonale

$$\mathbf{H}(\mathbf{x}(t_0), \mathbf{v}, \mathbf{w}, \mathbf{u}) = \mathbf{H}(\mathbf{u}) \oplus \mathbf{H}(\tilde{\mathbf{x}}(t_0), \mathbf{v}, \mathbf{w}) \quad (4.73)$$

e proiettando l'uscita all'istante t del modello (4.65) su $\mathbf{H}(\mathbf{u})$ si ottiene così la decomposizione ortogonale corrispondente

$$\begin{aligned} \mathbf{y}(t) &= CA^{t-t_0} \hat{\mathbf{x}}(t_0) + \sum_{t_0}^{t-1} CA^{t-1-s} B \mathbf{u}(s) + \\ &+ CA^{t-t_0} \tilde{\mathbf{x}}(t_0) \sum_{t_0}^{t-1} CA^{t-1-s} \mathbf{v}(s) + \mathbf{w}(t) = \\ &:= \hat{\mathbf{y}}(t) + \tilde{\mathbf{y}}(t) \end{aligned} \quad (4.74)$$

dove $\hat{\mathbf{y}}$ è funzione solo dell' ingresso ed è descritta dall'equazione di stato

$$\begin{cases} \hat{\mathbf{x}}(t+1) &= A \hat{\mathbf{x}}(t) + B \mathbf{u}(t) \\ \hat{\mathbf{y}}(t) &= C \hat{\mathbf{x}}(t) \quad t \geq t_0, \end{cases}$$

mentre $\tilde{\mathbf{y}}$ è completamente scorrelata dall' ingresso ed è descritta da,

$$\begin{cases} \tilde{\mathbf{x}}(t+1) &= A \tilde{\mathbf{x}}(t) + \mathbf{v}(t) \\ \tilde{\mathbf{y}}(t) &= C \tilde{\mathbf{x}}(t) + \mathbf{w}(t), \quad t \geq t_0, \end{cases}$$

Da notare che anche lo stato del modello originale (4.65) si può decomporre nella somma ortogonale $\mathbf{x}(t) = \hat{\mathbf{x}}(t) + \tilde{\mathbf{x}}(t)$ degli stati dei due modelli (4.75) e (4.75).

Dato che le variabili di ingresso e le condizioni iniziali di questi due modelli sono completamente scorrelate, la proiezione ortogonale sulla storia passata dello stato "congiunto" \mathbf{x} all'istante $t+1$, si ottiene calcolando separatamente la proiezione delle componenti $\hat{\mathbf{x}}(t)$ e $\tilde{\mathbf{x}}(t)$ sui rispettivi spazi "passati" $\mathbf{H}_t(\hat{\mathbf{y}}) \subset \mathbf{H}(\mathbf{u})$ e su $\mathbf{H}_t(\tilde{\mathbf{y}}) \subset \mathbf{H}(\tilde{\mathbf{x}}(t_0), \mathbf{v}, \mathbf{w})$. In conclusione si ottiene il seguente risultato

PROPOSIZIONE 9.8. *Se $H \equiv 0$, le stime dello stato, $\hat{\mathbf{x}}(t+1 | t)$ e $\hat{\mathbf{x}}(t | t)$, del modello (4.65), basate sull'osservazione congiunta della variabile di ingresso $\{\mathbf{u}(s) | t_0 \leq s \leq t\}$ e di uscita $\{\mathbf{y}(s) | t_0 \leq s \leq t\}$, si aggiornano nel tempo secondo il seguente algoritmo ricorsivo*

$$\hat{\mathbf{x}}(t+1 | t) = F \hat{\mathbf{x}}(t | t) + SR^{-1} \mathbf{y}(t) + B \mathbf{u}(t) \quad (4.75)$$

$$\hat{\mathbf{x}}(t+1 | t+1) = \hat{\mathbf{x}}(t+1 | t) + L(t+1) \hat{\mathbf{e}}(t+1) \quad (4.76)$$

equivalente a

$$\hat{\mathbf{x}}(t+1 | t) = A \hat{\mathbf{x}}(t | t-1) + B \mathbf{u}(t) + G(t) \hat{\mathbf{e}}(t) \quad (4.77)$$

dove $\hat{\mathbf{e}}(t) := \mathbf{y}(t) - C \hat{\mathbf{x}}(t | t-1)$ è l'innovazione del processo \mathbf{y} rispetto al passato congiunto di \mathbf{u} , \mathbf{y} e le matrici $L(t)$ e $G(t)$ sono calcolabili mediante le stesse formule dell'algoritmo di Kalman senza ingressi.

La condizione iniziale dell'algoritmo è la stima non causale

$$\hat{\mathbf{x}}(t_0 | t_0 - 1) = \hat{\mathbf{x}}(t_0) = \hat{E} [\mathbf{x}(t_0) | \mathbf{H}(\mathbf{u})] .$$

Si può osservare che anche se questo algoritmo sembra l'esatto analogo del filtro di Kalman usuale, la sua inizializzazione richiederebbe la soluzione di un problema di interpolazione basato su dati futuri che di fatto non sono disponibili in nessun istante finito. Come vedremo in seguito però, in ipotesi abbastanza blande, l'effetto delle condizioni iniziali sulla stima tende a svanire al crescere di $t - t_0$ e in pratica si può inizializzare l'algoritmo prendendo ad esempio $\hat{\mathbf{x}}(t_0) = 0$. Notiamo anche che, se questo fosse veramente il caso, si avrebbe

$$\hat{E} [\mathbf{y}(t) | \mathbf{H}(\mathbf{u})] = \hat{E} [\mathbf{y}(t) | \mathbf{H}_t(\mathbf{u})]$$

che è l'analogo della condizione (3.3) del Cap. 7. Possiamo ben dire quindi che in questo caso il futuro di \mathbf{u} da t in poi è condizionatamente incorrelato dal passato di \mathbf{y} dato il passato (stretto) di \mathbf{u} relativo all'intervallo $[t_0, t]$. In altri termini, solo se $\hat{\mathbf{x}}(t_0) = 0$ si può dire che se $H \equiv 0$ non c'è reazione da \mathbf{y} a \mathbf{u} .

PROBLEMA 9.3. Si consideri il caso in cui nell'equazione di misura in (4.65), c'è accoppiamento diretto con la variabile di ingresso, i.e.

$$\mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t) + \mathbf{w}(t) \quad (4.78)$$

Ricavare le equazioni del filtro di Kalman supponendo che $H \equiv 0$.

Confrontare il risultato con le formule per il predittore di Wiener con ingressi osservabili discusso al Capitolo 7. \diamond

PROBLEMA 9.4. Derivare le equazioni del Filtro di Kalman per la stima dello stato del modello lineare

$$\begin{cases} \mathbf{x}(t) &= A\mathbf{x}(t-1) + B\boldsymbol{\nu}(t) \\ \mathbf{y}(t) &= C\mathbf{x}(t) + D\boldsymbol{\nu}(t) \end{cases} \quad t \geq t_0 \quad , \quad (4.79)$$

dove $\boldsymbol{\nu}(t)$ è rumore bianco normalizzato p -dimensionale il cui futuro in $t \geq t_0$ è scorrelato da $\mathbf{x}(t_0 - 1)$. Modelli scritti in questa forma si incontrano nella letteratura statistica perchè l'equazione di stato è scritta nella forma, più familiare agli statistici, dei modelli AR (oppure ARMA) di ordine uno.

Riscrivere il modello in forma standard e descrivere i rumori di modello e di misura \mathbf{v} e \mathbf{w} e la loro correlazione.

Soluzione: Alle zeta trasformate il modello AR si scrive

$$\mathbf{x}(t) = (I - z^{-1}A)^{-1}B\boldsymbol{\nu}(t) \quad (4.80)$$

e usando l'identità

$$(I - z^{-1}A)^{-1} = I + (zI - A)^{-1}A$$

la (4.81) si può riscrivere

$$\mathbf{x}(t) = B\boldsymbol{\nu}(t) + (zI - A)^{-1}AB\boldsymbol{\nu}(t) \quad (4.81)$$

per cui, definendo una nuova variabile di stato $\boldsymbol{\xi}(t) := (zI - A)^{-1}AB\boldsymbol{\nu}(t)$, il modello (4.79) si può riscrivere in forma standard

$$\begin{cases} \boldsymbol{\xi}(t+1) &= A\boldsymbol{\xi}(t) + AB\boldsymbol{\nu}(t) \\ \mathbf{y}(t) &= C\boldsymbol{\xi}(t) + (CB + D)\boldsymbol{\nu}(t) \end{cases} \quad t \geq t_0 \quad , \quad (4.82)$$

Alternativamente, mostrare che ci si può ricondurre al modello in forma standard scrivendo l'equazione di stato per l'istante $t+1$ anzichè t , assumendo come variabile

di uscita $\mathbf{z}(t) := \mathbf{y}(t+1)$ invece di $\mathbf{y}(t)$, di modo che il rumore normalizzato in ingresso al modello diventa $\mathbf{n}(t) := \boldsymbol{\nu}(t+1)$. Mostrare che la matrice D nel modello così ottenuto è ancora uguale a $\hat{D} = CB + D$. \diamond

Modelli con osservazioni perfette. Talvolta capita di dover considerare modelli in cui la covarianza del rumore di misura R , è singolare. Dato che questa singolarità potrebbe a priori implicare la singolarità della varianza dell'innovazione $\Lambda(t)$, le equazioni del filtro di Kalman andrebbero in generale riscritte tenendo conto di questa possibilità. La difficoltà può essere affrontata in varie modi, uno dei quali è quello di usare in tutte le formule dell'algoritmo, la pseudoinversa al posto dell'inversa di $\Lambda(t)$. Questa soluzione non è molto soddisfacente per vari motivi, uno dei quali è che il calcolo della pseudoinversa richiede la decomposizione ai valori singolari (SVD) e la determinazione del rango, che sono operazioni numericamente molto onerose se debbono essere fatte ad ogni passo. Inoltre il fatto che R sia singolare, ad esempio uguale a zero, non implica poi necessariamente che anche il processo di innovazione abbia varianza singolare, come si può vedere dal seguente esempio.

OSSERVAZIONE 9.1. Ogni modello standard del tipo (1.1) si può sempre riscrivere come un modello lineare in cui non c'è rumore di misura. Infatti, ponendo

$$F := \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}, \quad G := \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad H := [C \quad D]$$

e definendo lo stato "aumentato"

$$\mathbf{z}(t) := \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{n}(t) \end{bmatrix}$$

il modello può essere riscritto nella forma

$$\begin{cases} \mathbf{z}(t+1) &= F\mathbf{z}(t) + G\mathbf{n}(t+1) \\ \mathbf{y}(t) &= H\mathbf{z}(t) \end{cases}$$

dove si potrebbe anche riscrivere l'equazione di stato nel modo usuale, ponendo $\mathbf{n}(t+1) := \boldsymbol{\nu}(t)$, che è rumore bianco normalizzato la cui storia presente e futura è scorrelata dallo stato presente. Chiaramente questo modello ha $R = 0$ ma la varianza dell'innovazione, che deve coincidere con quella del modello originale, è non singolare se il modello originale aveva invece $R > 0$. \diamond

Similmente, se siamo in condizioni di stazionarietà asintotica e se lo spettro $S_{\mathbf{y}}$ (e quindi il processo \mathbf{y}) ha rango pieno, Λ è certamente non singolare anche se $R = 0$. La possibile singolarità di Λ ha quindi a che fare con il *rango del processo di osservazione*⁶ più che con la possibile singolarità della varianza del rumore \mathbf{w} .

Una classica tecnica euristica che qualche volta permette di risolvere il problema del rumore a covarianza singolare è quella della *differenziazione*. Per spiegare di cosa si tratta supponiamo che la componente k -sima del processo di osservazione abbia rumore trascurabile ($\mathbf{w}_k = 0$). Questa tecnica consiste nel sostituire alla k -sima equazione di misura, una sua versione "differenziata", i.e

$$\mathbf{y}_k(t+1) = C_k \mathbf{x}(t+1) = C_k A \mathbf{x}(t) + C_k \mathbf{v}(t)$$

⁶E quindi con modelli "non parsimoniosi" che descrivono le osservazioni come se fossero *ridondanti*.

che ha $R_{kk} \neq 0$ e può essere trattata come un'equazione di misura addizionale. Naturalmente si assume che $\mathbf{y}_k(t+1)$ sia la k -sima componente del vettore delle osservazioni all'istante t e bisogna ammettere un ritardo di un passo nelle stime. Se si tenta di descrivere in dettaglio un procedimento generale basato su questa semplice idea, bisogna introdurre notazioni complicate con una gran quantità di indici e si corre il rischio di rendere il procedimento alquanto oscuro. Ci accontenteremo pertanto di una spiegazione abbastanza "qualitativa".

L'idea è di lavorare per passi: ad ogni passo si fanno delle differenziazioni della componente dell'uscita che è "senza rumore" aggiungendo così del nuovo rumore all'equazione di osservazione.

Al passo k -simo si fattorizza la covarianza del rumore di misura

$$T_k R_k T_k' = \begin{bmatrix} J_k \\ 0 \end{bmatrix} \begin{bmatrix} J_k' & 0 \end{bmatrix}$$

dove J_k ha $m_k \leq m$ righe linearmente indipendenti. Si partiziona l'uscita ottenuta al passo k -simo

$$T_k \mathbf{y}_k(t) = \begin{bmatrix} \hat{\mathbf{y}}_k(t) \\ \bar{\mathbf{y}}_k(t) \end{bmatrix}$$

in modo tale che la componente $\bar{\mathbf{y}}_k(t)$ non ha rumore di osservazione. Differenziando $\bar{\mathbf{y}}_k(t)$ e sostituendo l'equazione relativa si aggiunge un termine di rumore all'uscita. Il procedimento viene iterato finché non si perviene ad una varianza R non singolare.

Naturalmente l'aspetto insoddisfacente del procedimento è che per poter disporre di certe componenti differenziate dell'uscita si deve lavorare con dati "futuri" rispetto all'istante corrente t , il che significa in pratica che le stime faranno riferimento ad un certo istante corrente "aumentato" $t+h$, ma questo solo per utilizzare alcune componenti di \mathbf{y} , di modo che non si usa mai tutta l'informazione disponibile nell'istante corrente.

Per una trattazione approfondita del problema rimandiamo il lettore alla letteratura ed in particolare all'articolo [45].

Esercizi

9-1 Supponiamo che gli autovalori della matrice A siano tutti interni al cerchio unitario e che $t_0 \rightarrow -\infty$. È possibile che l'equazione di stato nel modello (1.1), scritta nella forma

$$\mathbf{x}(t) - A\mathbf{x}(t-1) = B\mathbf{n}(t-1)$$

sia un modello ARMA d'innovazione? Se no, scrivere il modello d'innovazione corrispondente. Qual'è il rango del processo stazionario $\{\mathbf{x}(t)\}$? Quali sono le condizioni per cui \mathbf{x} è a rango pieno?

9-2 (Campionamento del rumore bianco) Considerare la versione integrata, \mathbf{w} , di un rumore bianco a tempo continuo $\{\mathbf{n}(t); t \in \mathbb{R}\}$ di intensità σ^2 ,

$$\dot{\mathbf{w}}(t) = \mathbf{n}(t), \quad \mathbf{w}(0) = \mathbf{w}_0$$

Questo processo si chiama *processo di Wiener* in letteratura; se ne consideri la versione campionata con periodo T , di modo che $\mathbf{w}(k) := \mathbf{w}(kT)$ soddisfa l'equazione alle differenze

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{v}(k), \quad \mathbf{w}(0) = \mathbf{w}_0$$

dove $\mathbf{v}(k)$ è rumore bianco a tempo discreto. Il processo $\mathbf{v}(k)$ si può considerare come la versione campionata di $\mathbf{n}(t)$. Calcolare la varianza di $\mathbf{v}(k)$.

9-3 (Generalizzazione) Si consideri un segnale aleatorio a tempo continuo $\{\mathbf{z}(t); t \in \mathbb{R}\}$, che è l'uscita del filtro lineare con ingresso un rumore bianco continuo

$$\begin{aligned}\dot{\mathbf{x}}(t) &= -a\mathbf{x}(t) + \mathbf{n}(t) \\ \mathbf{z}(t) &= \mathbf{x}(t)\end{aligned}$$

dove $\{\mathbf{n}(t)\}$ è un rumore bianco continuo di intensità σ^2 .

Campionando $\mathbf{z}(t)$ con periodo T e scrivendo per semplicità $\mathbf{z}(k) := \mathbf{z}(kT)$, etc. si ottiene un modello discreto con ingresso un processo bianco a tempo discreto $\{\mathbf{w}(k)\}$. Calcolare la varianza di $\mathbf{w}(k)$ e discutere cosa accade per $a \rightarrow 0$ (in questo caso $\mathbf{z}(k)$ diventa un noto processo non stazionario, quale?) e interpretare il risultato.

9-4 Siano

$$A := \begin{bmatrix} 0 & 1 \\ -0.72 & 1.7 \end{bmatrix} \quad b := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Calcolare la matrice di covarianza $\Sigma(t, s)$ del processo $\{\mathbf{x}(t)\}$ descritto dall'equazione alle differenze

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + b\mathbf{w}(t)$$

dove \mathbf{w} è bianco di varianza σ^2 . Dire se esiste una varianza di stato a regime e eventualmente calcolarla.

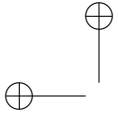
9-5 Mostrare che il calcolo della varianza di un segnale scalare stazionario (a media nulla) $\{\mathbf{z}(t)\}$, descritto da un modello di stato lineare minimo

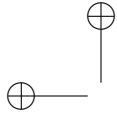
$$\begin{aligned}\mathbf{x}(t+1) &= A\mathbf{x}(t) + B\mathbf{w}(t) \\ \mathbf{z}(t) &= C\mathbf{x}(t).\end{aligned}$$

in cui $\{\mathbf{w}(t)\}$ è bianco di varianza Q , può essere ridotto alla soluzione di un'equazione di Lyapunov.

Si assumano A e B come nell'esercizio precedente e si confronti questo metodo di calcolo con quello basato sull'integrazione dello spettro di potenza.

9-6 Formulare il problema generale di predizione di $\mathbf{y}(t+k)$ ($k > 0$) in base alla storia passata fino all'istante t , per un processo descritto dal modello di stato lineare (4.1), in termini di stima dello stato e ricavare un algoritmo ricorsivo di predizione alla Kalman per $\hat{\mathbf{x}}(t+k | t)$.





CAPITOLO 10

IL FILTRO DI KALMAN A REGIME

La grande importanza applicativa del filtro di Kalman risiede, oltre che nella sua struttura particolarmente vantaggiosa dal punto di vista computazionale, nelle notevoli proprietà strutturali che esso possiede.

Quella di gran lunga più importante è la *stabilità*.

Per mostrare la stabilità del filtro di Kalman in generale, è necessaria una trattazione molto tecnica e poco intuitiva. Ci accontenteremo di studiare le proprietà asintotiche del filtro (al limite, quando $t \rightarrow \infty$) nel caso di *modelli a parametri invariabili nel tempo*, ovvero nel caso in cui (A, C, Q, S, R) sono costanti, rimandando il lettore al lavoro originale di Kalman [75], oppure al testo di Jazswinski [68], per l'analisi di stabilità nel caso generale.

1. Teoria Asintotica dell'Equazione di Riccati

Come già accennato in precedenza, la stabilità del predittore di un passo e quella del filtro (vero e proprio) sono intimamente legate. Sarà perciò sufficiente considerare uno solo dei due algoritmi. Come vedremo, i risultati dell'analisi asintotica di uno si trasportano molto facilmente all'altro. Fisseremo così le nostre idee sul *predittore di un passo*. Sarà utile semplificare un poco le notazioni. Scriveremo allo scopo semplicemente $\hat{\mathbf{x}}(t)$ al posto di $\hat{\mathbf{x}}(t | t-1)$ e $P(t)$ al posto di $P(t | t-1)$ *eliminando dalle notazioni l'indicazione del secondo argomento temporale*. Dato che ci occuperemo per il momento solo del predittore, non vi sarà pericolo di confusioni.

L'analisi che segue vale per il caso generale di rumori correlati. Ricordiamo che per particolarizzare i risultati che seguono al caso di rumori di modello e di osservazioni *scorrelati* basta evidentemente porre ovunque $S = 0$ nelle formule. Ovviamente in questo caso si ha da (4.22) e (4.20), $F = A$, $\tilde{Q} = Q$ e il guadagno del predittore $G(t)$ coincide con la matrice $K(t)$ definita nella (4.37). Ricordiamo che la matrice in catena chiusa $\Gamma(t)$ è *comunque* data dalla formula

$$\Gamma(t) = F - K(t)C \quad .$$

Con le nuove notazioni, il predittore di un passo evolve nel tempo secondo la relazione

$$\begin{aligned} \hat{\mathbf{x}}(t+1) &= F\hat{\mathbf{x}}(t) + K(t)[\mathbf{y}(t) - C\hat{\mathbf{x}}(t)] + SR^{-1}\mathbf{y}(t) \\ \hat{\mathbf{x}}(t_0) &= \bar{\mathbf{x}}_0 \quad . \end{aligned} \tag{1.1}$$

Qualche volta chiameremo (impropriamente se $S \neq 0$) anche la matrice $K(t)$ "guadagno" del predittore. Ricordiamo che essa è calcolabile mediante la formula

$$K(t) = FP(t)C' [CP(t)C' + R]^{-1} \quad , \tag{1.2}$$

dove $P(t)$ si ottiene risolvendo l'equazione di Riccati discreta (R.E.)

$$\begin{aligned} P(t+1) &= F[P(t) - P(t)C'(CP(t)C' + R)^{-1}CP(t)]F' + \tilde{Q} \\ P(t_0) &= P_0 \quad . \end{aligned} \quad (1.3)$$

Questa equazione alle differenze determina il comportamento asintotico del predittore.

Come si vede, il predittore (1.1) è sempre un sistema *variante* nel tempo dato che P e K sono in generale funzioni di t . Questa caratteristica del filtro è intimamente connessa al fatto che nel presente contesto i dati di misura sono disponibili *solo a partire dall'istante finito* t_0 in poi. Si noti che all'istante t_0 la stima a minima varianza di $\mathbf{x}(t_0)$ è la quantità deterministica $\mu_0 = E\mathbf{x}(t_0)$ (la media a priori). Questo perché non ci sono dati disponibili all'istante $t_0 - 1$, e questo fatto si può esprimere anche dicendo che $\mathbf{y}^{t_0-1} := \{\mathbf{y}(s) ; s < t_0\}$ generano un sottospazio $\mathbf{H}(\mathbf{y}^{t_0-1})$ contenente solo la v.a. nulla. Si può pensare così di calcolare $\hat{\mathbf{x}}(t_0)$ come la miglior stima lineare di $\mathbf{x}(t_0)$ in base a questi dati e quindi

$$\hat{\mathbf{x}}(t_0) = \mu_0 + \hat{E}(\mathbf{x}(t_0) - \mu_0 | \mathbf{y}^{t_0-1}) = \mu_0 \quad ,$$

dato che la proiezione di $\mathbf{x}(t_0) - \mu_0$ su $\mathbf{H}(\mathbf{y}^{t_0-1}) = \{0\}$ è necessariamente zero. In sostanza *la varianza di $\hat{\mathbf{x}}(t_0)$ è zero*, i.e.

$$\text{Var } \hat{\mathbf{x}}(t_0) = \text{Var } \mu_0 = 0 \quad .$$

Siano allora $\Sigma(t)$ e $\hat{\Sigma}(t)$ le matrici varianza dello stato del modello, $\mathbf{x}(t)$, e del predittore di un passo $\hat{\mathbf{x}}(t)$. Ricordiamo che dal principio di ortogonalità si ottiene la decomposizione

$$P(t) = E[\mathbf{x}(t) - \hat{\mathbf{x}}(t)][\mathbf{x}(t) - \hat{\mathbf{x}}(t)]' = \Sigma(t) - \hat{\Sigma}(t) \quad ,$$

(cf. (4.53)). All'istante iniziale la varianza dell'errore di stima è normalmente grande, dato che la stima a priori $\hat{\mathbf{x}}(t_0) = \bar{\mathbf{x}}_0$ è grossolana. Se le cose funzionano bene ci si aspetta che, man mano che vengono elaborati i dati di misura $\{\mathbf{y}(t)\}$, la stima $\hat{\mathbf{x}}(t)$ diventi sempre più accurata, ovvero $\hat{\Sigma}(t)$ tenda a diventare "prossima" a $\Sigma(t)$ con conseguente diminuzione della covarianza d'errore $P(t)$. Ciò spiega intuitivamente perché di norma $P(t)$ non può essere costante nel tempo.

Una situazione particolarmente semplice è quella in cui il processo di stato è *stazionario* in $t \geq t_0$. Questa situazione presuppone *A asintoticamente stabile* e P_0 soluzione dell'equazione di Lyapunov

$$P_0 = AP_0A' + Q \quad . \quad (1.4)$$

In queste ipotesi si ha $\Sigma(t) = P_0, \forall t \geq t_0$. Notiamo che $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ sono allora congiuntamente stazionari in $\{t \geq t_0\}$. Se fosse $t_0 = -\infty$ si potrebbe applicare la teoria di Wiener-Kolmogorov partendo dalla descrizione congiunta dei due processi attraverso i relativi spettri di potenza

$$\begin{aligned} S_{\mathbf{x}}(z) &= (zI - A)^{-1} Q(z^{-1}I - A')^{-1} \\ S_{\mathbf{x}\mathbf{y}}(z) &= S_{\mathbf{x}}(z)C' + (zI - A)^{-1}S \\ S_{\mathbf{y}}(z) &= CS_{\mathbf{x}}(z)C' + C(zI - A)^{-1}S + S'(z^{-1}I - A')^{-1}C' + R \quad . \end{aligned}$$

Se supponiamo che lo spettro del processo di osservazione non abbia zeri sulla circonferenza unità, i.e. $S_{\mathbf{y}}(e^{j\omega}) > 0 \forall \omega \in [-\pi, \pi]$, si ha stabilità stretta del predittore

di Wiener-Kolmogorov e stazionarietà del processo $\hat{\mathbf{x}}_\infty(t) := \hat{E}[\mathbf{x}(t) | \mathbf{y}^{t-1}]$. In queste condizioni si ha anche costanza nel tempo della varianza d'errore, $P_\infty := P_0 - \hat{\Sigma}_\infty$ (qui si è usato il simbolo $\hat{\Sigma}_\infty$ per indicare la matrice varianza di $\hat{\mathbf{x}}_\infty(t)$).

PROPOSIZIONE 10.1. *Se la matrice A nel modello (4.1) è asintoticamente stabile, allora la stima*

$$\hat{\mathbf{x}}(t) = \hat{E}[\mathbf{x}(t) | \mathbf{y}(t_0) \dots \mathbf{y}(t-1)]$$

converge, quando $t_0 \rightarrow -\infty$, in media quadratica, verso un limite $\hat{\mathbf{x}}_\infty(t)$, che è il predittore di Wiener-Kolmogorov di $\mathbf{x}(t)$ in base alla storia passata infinita, $\{\mathbf{y}(s); s < t\}$, del processo \mathbf{y} . Esiste allora il limite

$$\lim_{t_0 \rightarrow -\infty} P(t) = \bar{\Sigma} - \hat{\Sigma}_\infty := P_\infty \quad ,$$

dove $\hat{\Sigma}_\infty = \lim_{t_0 \rightarrow -\infty} \hat{\Sigma}(t)$ è la matrice varianza del processo stazionario $\{\hat{\mathbf{x}}_\infty(t)\}$ e $\bar{\Sigma}$ è la varianza di stato a regime, soluzione dell'equazione di Lyapunov (1.24). Pertanto P_∞ è soluzione dell'equazione algebrica di Riccati (A.R.E.)

$$P = F [P - PC'(CPC' + R)^{-1} CP] F' + \tilde{Q}. \quad (1.5)$$

La stima $\hat{\mathbf{x}}_\infty(t)$ soddisfa l'equazione alle differenze (il predittore di Kalman "a regime")

$$\hat{\mathbf{x}}_\infty(t+1) = F\hat{\mathbf{x}}_\infty(t) + K_\infty[\mathbf{y}(t) - C\hat{\mathbf{x}}_\infty(t)] + SR^{-1}\mathbf{y}(t) = \quad (1.6)$$

$$= \Gamma_\infty\hat{\mathbf{x}}_\infty(t) + G_\infty\mathbf{y}(t) \quad (1.7)$$

in cui K_∞ , si calcola mediante la formula

$$K_\infty = FP_\infty C' [CP_\infty C' + R]^{-1} \quad , \quad (1.8)$$

G_∞ è dato dalla formula $G_\infty = K_\infty + SR^{-1}$ e la matrice in catena chiusa a regime, è

$$\Gamma_\infty := F - K_\infty C \quad (1.9)$$

Se l'equazione (1.7) rappresenta un sistema dinamico completamente raggiungibile e $S_{\mathbf{y}}(e^{j\omega}) > 0$, allora Γ_∞ ha tutti gli autovalori dentro il cerchio unità e il predittore a regime è asintoticamente stabile.

DIMOSTRAZIONE. Nella dimostrazione supporremo che il processo di stato $\{\mathbf{x}(t)\}$ del modello (4.1) sia *stazionario* di varianza $\bar{\Sigma}$, soluzione dell'equazione di Lyapunov (1.24). In questo caso $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ sono congiuntamente stazionari e senza perdita di generalità li penseremo definiti su tutto l'asse dei tempi \mathbb{Z} .

Si noti che partendo da arbitrarie condizioni iniziali in $t_0 > -\infty$ e in ispecie da un'arbitraria matrice simmetrica e semidefinita positiva $\Sigma(t_0) = P_0$, l'asintotica stabilità di A garantisce, grazie al teorema 9.4, che al limite per $t_0 \rightarrow -\infty$ la coppia $\{\mathbf{x}(t)\}$, $\{\mathbf{y}(t)\}$ diventi congiuntamente stazionaria con varianza di stato $\Sigma(t) = \bar{\Sigma}$, $\forall t$ e asintoticamente si ricade quindi comunque nella condizione in cui ci siamo posti.

La stima $\hat{\mathbf{x}}(t)$, pensata funzione dell'indice discreto t_0 , converge per $t_0 \rightarrow -\infty$ in virtù del teorema generale di convergenza delle martingale, si veda il Teorema 5.4, Cap. 5) verso il limite, $\hat{\mathbf{x}}_\infty(t)$, che è il miglior stimatore lineare di $\mathbf{x}(t)$ in base a tutta la storia passata, $\mathbf{H}_{t-1}(\mathbf{y})$ del processo $\{\mathbf{y}(t)\}$. Ovviamente $\{\hat{\mathbf{x}}_\infty(t)\}$ è stazionario e ha quindi varianza costante, uguale a $\hat{\Sigma}_\infty$, per ogni t .

Notiamo ora che $\hat{\mathbf{x}}(t)$ si può pensare come la proiezione ortogonale di $\hat{\mathbf{x}}_\infty(t)$ su $\text{span}\{\mathbf{y}(t_0), \dots, \mathbf{y}(t-1)\} \subset \mathbf{H}_{t-1}(\mathbf{y})$. Ne segue che la varianza della differenza $\hat{\mathbf{x}}_\infty(t) - \hat{\mathbf{x}}(t)$ può, in base al principio di ortogonalità, scriversi come

$$\text{Var}\left(\hat{\mathbf{x}}_\infty(t) - \hat{\mathbf{x}}(t)\right) = \hat{\Sigma}_\infty - \hat{\Sigma}(t) \quad .$$

Prendendo allora un qualunque $a \in R^n$, si ha

$$\begin{aligned} a' \left(\hat{\Sigma}_\infty - \hat{\Sigma}(t) \right) a &= a' E \left[\hat{\mathbf{x}}_\infty(t) - \hat{\mathbf{x}}(t) \right] \left[\hat{\mathbf{x}}_\infty(t) - \hat{\mathbf{x}}(t) \right]' a \\ &= E \left\{ a' \left[\hat{\mathbf{x}}_\infty(t) - \hat{\mathbf{x}}(t) \right] \right\}^2 \leq E \left\{ |a|^2 \cdot \left| \hat{\mathbf{x}}_\infty(t) - \hat{\mathbf{x}}(t) \right|^2 \right\} \\ &= |a|^2 \cdot \left\| \hat{\mathbf{x}}_\infty(t) - \hat{\mathbf{x}}(t) \right\|^2 \quad , \end{aligned}$$

dove col simbolo “ $|a|$ ” si è denotata la norma euclidea di a . Da questa disuguaglianza si ricava che

$$\lim_{t_0 \rightarrow -\infty} \hat{\Sigma}(t) = \hat{\Sigma}_\infty \quad , \quad (1.10)$$

il che ovviamente implica

$$\lim_{t_0 \rightarrow -\infty} P(t) = \lim_{t_0 \rightarrow -\infty} \left(\bar{\Sigma} - \hat{\Sigma}(t) \right) = P_\infty \quad . \quad (1.11)$$

Passando ora al limite per $t_0 \rightarrow -\infty$ nell'espressione del predittore di Kalman (1.1) si trova la (1.6). Che P_∞ soddisfi l'A.R.E. (1.5) segue poi direttamente dalla (1.11), passando al limite su entrambi i membri dell'equazione di Riccati (1.3).

Ovviamente la (1.6) può essere pensata come una realizzazione della funzione di trasferimento del predittore di Wiener-Kolmogorov di $\mathbf{x}(t+1)$ in base al passato (infinito) \mathbf{y}^t . Per la positività stretta dello spettro di $\{\mathbf{y}(t)\}$ i poli della funzione di trasferimento devono essere tutti a modulo strettamente minore di uno.

Evidentemente la condizione di raggiungibilità completa dell'enunciato fa sì che i poli di $H(z)$ coincidano con gli autovalori della matrice Γ_∞ che regola la dinamica del sistema lineare (1.6). \square

Notiamo che si giunge a conclusioni sostanzialmente identiche se invece di far tendere l'istante iniziale t_0 a $-\infty$ si fa invece tendere $t - t_0$ a $+\infty$.

COROLLARIO 10.1. *Posto $\tau = t - t_0$, si ha anche*

$$\lim_{\tau \rightarrow \infty} \left\| \hat{\mathbf{x}}(\tau) - \hat{\mathbf{x}}_\infty(\tau) \right\| = 0 \quad ,$$

ovvero, il predittore di Kalman di $\mathbf{x}(\tau)$ in base a $\mathbf{y}(0), \dots, \mathbf{y}(\tau-1)$ converge in media quadratica quando $\tau \rightarrow \infty$ a quello di Wiener-Kolmogorov.

DIMOSTRAZIONE. In effetti, per la stazionarietà congiunta di $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ si può scrivere

$$\hat{\mathbf{x}}(t) = U_{t_0} \hat{E} \left[\mathbf{x}(t - t_0) \mid \mathbf{y}(0), \dots, \mathbf{y}(t - t_0 - 1) \right] \quad ,$$

dove U_t è l'operatore di traslazione temporale che opera sulle variabili casuali dello spazio di Hilbert $\mathbf{H} := \text{span}\{\mathbf{x}_k(t), \mathbf{y}_j(t) ; t \in \mathbb{Z} ; k = 1, \dots, n ; j = 1, \dots, m\}$ (supponiamo ancora che $\{\mathbf{x}(t)\}$ e $\{\mathbf{y}(t)\}$ siano definiti su tutto l'asse dei tempi).

Moltiplicando a sinistra per U_{-t_0} si trova

$$U_{-t_0} \hat{\mathbf{x}}(t) := \hat{\mathbf{x}}(t - t_0) = \hat{E} \left[\mathbf{x}(t - t_0) \mid \mathbf{y}(0), \dots, \mathbf{y}(t - t_0 - 1) \right] \quad ,$$

per cui $\hat{\mathbf{x}}(t - t_0) := U_{-t_0} \hat{\mathbf{x}}(t)$ rappresenta il predittore di $\mathbf{x}(t - t_0)$ in base ai dati $\mathbf{y}(0), \dots, \mathbf{y}(t - t_0 - 1)$. Ricordando che U_t è un operatore *unitario* ($\|U_t z\| = \|z\|$, $\forall z \in H$), si ha

$$\left\| \hat{\mathbf{x}}(t) - \hat{\mathbf{x}}_\infty(t) \right\| = \left\| U_{-t_0} \hat{\mathbf{x}}(t) - U_{-t_0} \hat{\mathbf{x}}_\infty(t) \right\| = \left\| \hat{\mathbf{x}}(t - t_0) - \hat{\mathbf{x}}_\infty(t - t_0) \right\| \quad ,$$

(giacché $\hat{\mathbf{x}}_\infty(t)$ è un processo stazionario). Ora il primo membro di questa uguaglianza tende a zero quando $t_0 \rightarrow -\infty$.

Per completare la dimostrazione basta ora notare che $P(\tau)$ soddisfa un'equazione alle differenze a coefficienti costanti. Ciò implica che nell'espressione della soluzione della (1.3) come funzione dei dati iniziali

$$P(t) = \Pi(t, t_0, P_0) \quad , \quad (1.12)$$

la funzione Π soddisfa la relazione

$$\Pi(t, t_0, P_0) = \Pi(t + \Delta, t_0 + \Delta, P_0) \quad , \quad \forall \Delta \in \mathbb{Z} \quad . \quad (1.13)$$

In particolare, posto $\Delta = -t_0$ si ha

$$P_\infty = \lim_{t_0 \rightarrow -\infty} P(t) = \lim_{t_0 \rightarrow -\infty} \Pi(t - t_0, 0, P_0) = \lim_{\tau \rightarrow \infty} P(\tau)$$

e pertanto la stima di Kalman $\hat{\mathbf{x}}(\tau)$ tende asintoticamente a soddisfare le (1.6), (1.8) e (1.5). \square

In pratica, *per segnali stazionari*, la stima calcolata col predittore di Kalman a “transitorio esaurito” coincide con quella che si otterrebbe mediante la teoria di Wiener-Kolmogorov.

OSSERVAZIONE 10.1. La condizione che lo spettro $S_{\mathbf{y}}$ sia privo di zeri sulla circonferenza unità è fondamentale per la stabilità asintotica del filtro a regime. Notare a questo proposito che nel caso di rumori di modello e di osservazione correlati, la condizione $R > 0$ *non implica necessariamente la positività dello spettro*.

Questo si può verificare per mezzo di un semplice esempio. Prendiamo uno spettro scalare definito dalla $S_{\mathbf{y}}(z) = W(z)W(1/z)$ dove $W(z)$ è la funzione di trasferimento

$$W(z) = \frac{(z-1)^2}{(z-1/2)^2} = \frac{1-2z^{-1}+z^{-2}}{1-z+1/4z^{-2}} = 1 + \frac{-z^{-1}+3/4z^{-2}}{1-z+1/4z^{-2}}$$

che, una volta realizzata in forma di stato usando una delle tecniche viste alla sezione X.2, fornisce un modello con $R = 1$. In questo esempio, nonostante R sia positivo, lo spettro ha uno zero di molteplicità quattro sulla circonferenza unitaria. Il filtro di Kalman a regime in questo caso sarebbe chiaramente instabile. \diamond

Risultati asintotici di questo genere, vista anche la semplicità computazionale della realizzazione (1.6) del predittore di Wiener-Kolmogorov, che è a parametri costanti, sono particolarmente importanti. Per mostrare la convergenza al filtro a regime ci siamo basati in modo essenziale sulla *stabilità asintotica* di A .

La stabilità asintotica di A è però una condizione troppo restrittiva in molte applicazioni. Si vorrebbe che anche per modelli instabili, per i quali cioè la varianza $\Sigma(t)$ diverge nel tempo, la stima $\hat{\mathbf{x}}(t)$ fosse ugualmente una buona ricostruzione di $\mathbf{x}(t)$, nel senso che l'errore di stima $\mathbf{x}(t) - \hat{\mathbf{x}}(t)$ si mantenesse a *varianza limitata* e possibilmente piccola (vedi figura 1.1).

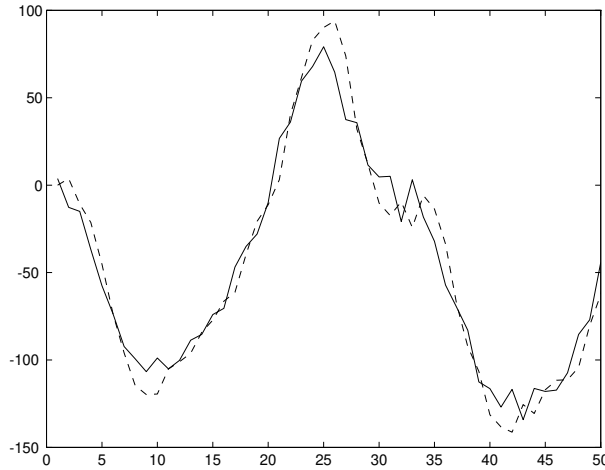


FIGURA 1.1. Predizione di un segnale non stazionario (linea tratteggiata)

Nel seguito di questo capitolo ci occuperemo della *teoria asintotica* del filtro (in realtà del predittore) di Kalman in ipotesi generali. La questione fondamentale sarà ancora quella dell'esistenza del limite

$$\lim_{t \rightarrow \infty} P(t)$$

e la stabilità asintotica del corrispondente filtro “a regime”.

Si noti che, venendo meno l'ipotesi di stabilità asintotica di A , il parallelo con la teoria di Wiener-Kolmogorov (che è basata in modo essenziale sull'ipotesi di stazionarietà) viene meno. In effetti, la teoria del filtro a regime è in un certo senso da riguardarsi come una generalizzazione nient'affatto ovvia della teoria di Wiener-Kolmogorov per processi *non stazionari*.

Notiamo innanzitutto che se si vuole che la stima $\hat{\mathbf{x}}(t)$ “corra dietro” a (tutte le componenti di) $\mathbf{x}(t)$, con errore finito, occorre che almeno i modi divergenti del modello siano accoppiati con l'uscita. Si dà a questo proposito la seguente definizione.

DEFINIZIONE 10.1. *Dato il sistema dinamico*

$$\begin{aligned} \mathbf{x}(t+1) &= F\mathbf{x}(t) + G\mathbf{u}(t) \\ \mathbf{y}(t) &= H\mathbf{x}(t) \end{aligned} ,$$

la coppia (F, H) si dice rivelabile se tutti i modi non asintoticamente stabili di F sono osservabili.

Dualmente, la coppia (F, G) si dice stabilizzabile se tutti i modi non asintoticamente stabili di F sono raggiungibili.

La rivelabilità e la stabilizzabilità sono ovviamente generalizzazioni dei concetti classici di osservabilità e raggiungibilità per i sistemi lineari.

LEMMA 10.1. *La coppia (F, H) è rivelabile se e solo se vale una qualunque delle seguenti condizioni.*

1) Se esiste una matrice non singolare T , tale per cui

$$TFT^{-1} = \begin{bmatrix} F_{11} & 0 \\ F_{21} & F_{22} \end{bmatrix}, \quad HT^{-1} = [H_1 \quad 0]$$

con (F_{11}, H_1) completamente osservabile, allora F_{22} è asintoticamente stabile.

- 2) Esiste una matrice K tale per cui $F + KH$ è asintoticamente stabile.
 3) Gli autovettori di F che appartengono allo spazio nullo di H corrispondono ad autovalori di modulo strettamente minore di 1. In formole

$$Ha = 0, \quad Fa = \lambda a, \quad a \neq 0 \Rightarrow |\lambda| < 1.$$

Dualmente, (F, G) è stabilizzabile se e solo se vale una delle seguenti condizioni.

1') Se esiste T non singolare, tale per cui

$$TFT^{-1} = \begin{bmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{bmatrix}, \quad TG = [G_1 \quad 0]$$

con (F_{11}, G_1) completamente raggiungibile, allora $|\lambda(F_{22})| < 1$.

- 2') Esiste una matrice L tale per cui $|\lambda(F + GL)| < 1$.
 3') Gli autovettori sinistri di F che sono ortogonali allo spazio immagine di G corrispondono ad autovalori di modulo strettamente minore di 1,

$$a'G = 0, \quad a'F = \lambda a', \quad a \neq 0 \Rightarrow |\lambda| < 1.$$

Per una dimostrazione delle proprietà 3) e 3') si può vedere ad esempio il testo di Kailath [72], pag. 136.

È da tener presente che se F è asintoticamente stabile allora (FH) è automaticamente rivelabile ed (F, G) è stabilizzabile, qualunque siano H e G .

Lo studio del comportamento asintotico del filtro di Kalman è sostanzialmente lo studio del comportamento asintotico delle soluzioni dell'equazione alle differenze di Riccati (1.3), sotto opportune ipotesi di rivelabilità e stabilizzabilità sulla matrice $F = A - SR^{-1}C$ e sui parametri C e \tilde{Q} . Il primo risultato di questo genere è il seguente.

LEMMA 10.2. Se la coppia (F, C) è rivelabile, allora, per ogni condizione iniziale $P(t_0) = P_0$ simmetrica e semidefinita positiva, la soluzione dell'equazione di Riccati (1.3) è limitata, ovvero, esiste una matrice $U \geq 0$ (eventualmente dipendente da P_0) tale che

$$P(t) \leq U, \quad \forall t \geq t_0. \quad (1.14)$$

DIMOSTRAZIONE. Supponiamo che (F, C) sia rivelabile. Allora per la proprietà 2), lemma 10.1, esiste una matrice costante L tale per cui

$$|\lambda(F - LC)| < 1.$$

Costruiamo allora un filtro lineare e invariante per il sistema (4.21) di guadagno (costante) L ,

$$\begin{aligned} \mathbf{s}(t+1) &= F\mathbf{s}(t) + L[\mathbf{y}(t) - C\mathbf{s}(t)] + SR^{-1}\mathbf{y}(t) \\ \mathbf{s}(t_0) &= \bar{\mathbf{x}}_0. \end{aligned} \quad (1.15)$$

il cui errore di predizione $\tilde{\mathbf{s}}(t) := \mathbf{x}(t) - \mathbf{s}(t)$ soddisfa l'equazione alle differenze

$$\tilde{\mathbf{s}}(t+1) = (F - LC)\tilde{\mathbf{s}}(t) + \tilde{\mathbf{v}}(t) - L\mathbf{w}(t)$$

Chiaramente la varianza, diciamola $U(t)$, dell'errore di predizione, non può essere mai inferiore alla varianza dell'errore del predittore di Kalman, che per definizione è lo stimatore a varianza *minima* tra gli stimatori lineari. Si ha quindi

$$P(t) \leq E \tilde{\mathbf{s}}(t) \tilde{\mathbf{s}}'(t) = U(t) \quad , \quad \forall t \geq t_0 \quad . \quad (1.16)$$

La varianza $U(t)$ è però limitata, essendo soluzione dell'equazione *lineare*

$$\begin{aligned} U(t+1) &= (F - LC) U(t) (F - LC)' + LRL' + \tilde{Q} \\ U(t_0) &= P_0 \quad , \end{aligned}$$

in cui la matrice $F - LC$ è asintoticamente stabile. Infatti, in virtù della stabilità asintotica di $F - LC$, esiste il limite $\bar{U} := \lim_{t \rightarrow \infty} U(t)$ e la (1.16) implica che per tutti i t maggiori o uguali di un opportuno $t_1 \geq t_0$ vale la

$$P(t) \leq \bar{U} \quad .$$

La (1.14) segue allora immediatamente. \square

La condizione di rivelabilità non è necessaria per la validità del lemma. In effetti basta considerare modelli in cui $Q = 0$ e la parte non osservabile di F ha solo autovalori semplici di modulo unitario. Ad esempio nel caso limite $F = I$, $C = 0$, si ha $P(t) = \Sigma(t) = P_0$, $\forall t \geq t_0$. Se però si suppone che i modi a modulo unitario di F siano accoppiati al rumore di ingresso (cioè, se si suppone $(F, \tilde{Q}^{1/2})$ *stabilizzabile*), allora il risultato si può invertire.

Vale a questo proposito la seguente proposizione.

PROPOSIZIONE 10.2. *La coppia $(F, \tilde{Q}^{1/2})$ sia stabilizzabile. Condizione necessaria e sufficiente per la limitatezza di $P(t)$, a partire da una qualunque condizione iniziale $P_0 = P'_0 \geq 0$, è che (F, C) sia rivelabile.*

La dimostrazione di questa equivalenza richiede l'uso di alcune conseguenze dell'ipotesi di stabilizzabilità che verranno ricavate nel seguito. Ce ne occuperemo un poco più avanti.

Ricordiamo qui la forma equivalente (simmetrizzata) dell'equazione di Riccati,

$$P(t+1) = \left(F - K(t)C \right) P(t) \left(F - K(t)C \right)' + K(t)RK'(t) + \tilde{Q}. \quad (1.17)$$

I risultati che seguono sono connessi al significato statistico della soluzione dell'equazione (1.17). Il primo è una caratterizzazione dall'ottimalità del filtro di Kalman e segue dal ragionamento fatto nella dimostrazione del lemma 10.2.

LEMMA 10.3. *Ad ogni istante t , il guadagno ottimo $K(t)$ si può ottenere minimizzando rispetto a L la varianza $P_L(t+1)$ definita dalla*

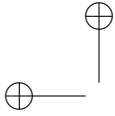
$$P_L(t+1) := (F - LC)P(t)(F - LC)' + LRL' + \tilde{Q} \quad , \quad (1.18)$$

dove $P(t)$ è l'usuale soluzione della R.E. (1.3) con condizione iniziale $P(t_0) = P_0$.

DIMOSTRAZIONE. Si può fare la verifica notando che la (1.18) si può riscrivere

$$\begin{aligned} P_L(t+1) &:= L(CP(t)C' + R)L' - FP(t)C'L - LCP(t)F' + FP(t)F' + \tilde{Q} \\ &= (L - K(t))\Lambda(t)(L - K(t))' - K(t)\Lambda(t)K(t)' + FP(t)F' + \tilde{Q}, \end{aligned}$$

e l'ultimo membro è ovviamente minimo per $L = K(t)$. \square



Naturalmente le dimostrazioni di questo tipo non aiutano molto a capire il significato dell'enunciato. Il risultato è in realtà interpretabile usando la teoria della Programmazione Dinamica [15], [68, p. 225]. La formula per $K(t)$ segue derivando formalmente rispetto ad L e imponendo che

$$\frac{\partial P_L(t+1)}{\partial L} = -CP(t) (F - LC)' + RL' = 0 \quad ,$$

da cui:

$$L = FP(t) C' \left(CP(t) C' + R \right)^{-1} .$$

Il lettore dimostri che questo è effettivamente un minimo. \diamond

LEMMA 10.4. $P(t)$ è una funzione monotona della varianza iniziale. Se $P_i(t)$, $i = 1, 2$, sono soluzioni della (1.17) corrispondenti alle condizioni iniziali

$$P(t_0) = P_i \quad , \quad i = 1, 2 \quad ,$$

si ha

$$P_2 \geq P_1 \Rightarrow P_2(t) \geq P_1(t) \quad , \quad \forall t \geq t_0 \quad . \quad (1.19)$$

DIMOSTRAZIONE. La (1.19) è ovvia per $t = t_0$. Assumiamo che sia verificata per $t = s > t_0$ e dimostriamo quindi che vale anche per $t = s + 1$.

Usando il lemma 10.3 e la $P_2(s) \geq P_1(s)$, si ha,

$$\begin{aligned} P_2(s+1) &= \min_L \left\{ (F - LC) P_2(s) (F - LC)' + LRL' + \tilde{Q} \right\} \\ &\geq \min_L \left\{ (F - LC) P_1(s) (F - LC)' + LRL' + \tilde{Q} \right\} \\ &= P_1(s+1) \quad . \end{aligned}$$

che è quanto si doveva dimostrare. \square

COROLLARIO 10.2. Se $P(t_0) = 0$, allora la corrispondente soluzione, $P_0(t)$, dell'equazione di Riccati è una funzione monotona non decrescente di t .

DIMOSTRAZIONE. Chiaramente $P_0(t_0 + 1) = \tilde{Q} \geq 0 = P_0(t_0)$, per cui l'asserto vale per $t = t_0 + 1$. Supponiamolo vero per $t = s$. In base all'invarianza temporale (1.13) si può allora scrivere

$$P_0(s+1) = \Pi(s+1, t_0, 0) = \Pi(s, t_0 - 1, 0) = \Pi(s, t_0, P(t_0 + 1)) \quad .$$

L'ultimo passaggio scende dalla cosiddetta "proprietà di semigrupp" della Π

$$\Pi(t, t_0, P_0) = \Pi(t, s, \Pi(s, t_0, P_0)) \quad , \quad t_0 \leq s \leq t \quad . \quad (1.20)$$

Dato ora che, per il lemma 10.4, si ha

$$\Pi(s, t_0, P(t_0 + 1)) \geq \Pi(s, t_0, 0) = P_0(s) \quad ,$$

ne scende che $P_0(s+1) \geq P_0(s)$ e quindi l'affermazione del corollario. \square

Ricordiamo ora che $P_0(t)$ è certamente limitata (nel senso dell'usuale ordinamento tra matrici) se la coppia (F, C) è rivelabile. D'altro canto, una successione monotona di matrici superiormente limitata ammette limite finito e si ha allora il seguente risultato.

PROPOSIZIONE 10.3. *Se (F, C) è rivelabile, esiste il*

$$\lim_{t \rightarrow \infty} P_0(t)$$

ed è una matrice simmetrica $\bar{P} \geq 0$ che soddisfa l'A.R.E. In altri termini, sotto l'ipotesi di rivelabilità, l'A.R.E. ammette almeno una soluzione simmetrica e semidefinita positiva.

In generale l'ARE potrebbe non ammettere soluzioni varianza se (F, C) non è rivelabile (il lettore dia un controesempio assumendo $C = 0$)¹.

Ad ogni soluzione $\bar{P} = \bar{P}' \geq 0$ dell'ARE corrisponde un "filtro a regime" di guadagno

$$\bar{K} := F \bar{P} C' (C \bar{P} C' + R)^{-1} \quad , \quad (1.21)$$

che per essere praticamente utilizzabile deve chiaramente essere *asintoticamente stabile*. Sorge così la questione della stabilità asintotica della matrice

$$\bar{\Gamma} := F - \bar{K} C \quad ,$$

che regola la dinamica del predittore a regime. Diamo a questo proposito la seguente definizione.

DEFINIZIONE 10.2. *La soluzione \bar{P} dell'A.R.E. (1.5) si dice stabilizzante se la corrispondente matrice in catena chiusa $F - \bar{K} C$ (con \bar{K} definito dalla (1.21)) è asintoticamente stabile.*

La struttura a reazione del filtro suggerisce che la questione della stabilità asintotica di $\bar{\Gamma}$ è legata alla stabilizzabilità. In effetti se $\bar{Q} = 0$ e si prende la soluzione banale dell'A.R.E., $\bar{P} = 0$, il corrispondente filtro "a regime" è semplicemente

$$\hat{\mathbf{x}}(t+1) = F \hat{\mathbf{x}}(t) + S R^{-1} \mathbf{y}(t) \quad ,$$

che è asintoticamente stabile solo se F stessa è asintoticamente stabile, ovvero $(F, 0)$ è stabilizzabile.

Nel seguito indicheremo col simbolo $Q^{1/2}$ una qualunque radice quadrata di Q , ovvero una qualunque matrice B tale che

$$B B' = Q \quad . \quad (1.22)$$

È chiaro che la stabilizzabilità della coppia (F, B) è equivalente a quella di (F, B_1) con B_1 un'altra qualsivoglia matrice soddisfacente la (1.22) (per convincersene basta calcolare il gramiano di raggiungibilità).

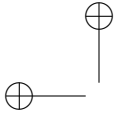
LEMMA 10.5. *Se $(F, \tilde{Q}^{1/2})$ è stabilizzabile, allora tutte le (eventuali) soluzioni simmetriche e semidefinite positive dell'A.R.E. sono stabilizzanti.*

DIMOSTRAZIONE. Supponiamo che $\bar{P} = \bar{P}' \geq 0$ sia una soluzione dell'A.R.E. e che $F - \bar{K} C$ abbia un autovalore, λ_0 , di modulo maggiore o uguale a 1. Mostriamo che questo porta a contraddire l'ipotesi di stabilizzabilità di $(F, \tilde{Q}^{1/2})$.

Esisterà allora un vettore (in generale complesso), a , tale per cui

$$a^* (F - \bar{K} C) = \lambda_0 a^* \quad , \quad |\lambda_0| \geq 1 \quad (1.23)$$

¹La condizione non è comunque necessaria. Nel caso in cui $\bar{Q} = 0$ l'A.R.E. è un'equazione omogenea in P che ammette sempre la soluzione varianza $\bar{P} = 0$ indipendentemente dalla rivelabilità di (F, C) .



(a^* è il trasposto coniugato di a). Moltiplicando a sinistra e a destra l'equazione algebrica di Riccati nella forma simmetrizzata

$$\bar{P} = (F - \bar{K}C) \bar{P}(F - \bar{K}C)' + \bar{K}R\bar{K}' + \bar{Q} \tag{1.24}$$

per (rispettivamente) a^* e a , si ricava allora

$$(1 - |\lambda_0|^2) a^* \bar{P} a = a^* \bar{K}R\bar{K}' a + a^* \bar{Q} a \quad .$$

In questa espressione il primo membro è non positivo e il secondo non negativo. Entrambi quindi devono essere nulli.

In particolare i due addendi a secondo membro (che sono non negativi) dovranno essere entrambi zero. Dato che $R > 0$ si avrà

$$a^* \bar{K} = 0 \tag{1.25}$$

e, presa un'arbitraria B soddisfacente la (1.22),

$$a^* B = 0 \quad . \tag{1.26}$$

Le (1.23) e (1.25) forniscono quindi

$$a^* F = \lambda_0 a^* \quad , \quad |\lambda_0| \geq 1 \quad ,$$

e questa relazione associata alla (1.26) contraddice la stabilizzabilità della coppia (F, B) . □

In realtà l'enunciato del lemma 10.5 può essere notevolmente rafforzato. Per fare ciò occorre però ricavare prima una formula notevole che descrive l'evoluzione temporale dello "scostamento" $P(t) - \bar{P}$ fra la soluzione della R.E. e una qualunque soluzione \bar{P} della corrispondente A.R.E..

Ricordiamo l'espressione (4.26) della varianza dell'errore di filtraggio $P(t | t) = E [\tilde{x}(t | t) \tilde{x}(t | t)']$ ricavata nel capitolo precedente. Scrivendo $\Pi(t)$ al posto di $P(t | t)$ si ha:

$$\Pi(t) = P(t) - P(t) C' [CP(t) C' + R]^{-1} CP(t) \quad , \tag{1.27}$$

da cui si ricava un'espressione notevole per il guadagno del filtro ².

LEMMA 10.6. *La matrice guadagno $K(t)$ può essere espresso mediante la formula*

$$K(t) = F \Pi(t) C' R^{-1} \quad . \tag{1.28}$$

DIMOSTRAZIONE. Si ha, ponendo $\Lambda(t) := [CP(t) C' + R]$,

$$\begin{aligned} \Pi(t) C' &= P(t) C' [I - \Lambda(t)^{-1} CP(t) C'] \\ &= P(t) C' \Lambda^{-1}(t) [\Lambda(t) - CP(t) C'] P(t) C' \Lambda^{-1}(t) R \quad , \end{aligned}$$

da cui scende immediatamente la (1.28). □

LEMMA 10.7. [3] *Sia $\bar{P} = \bar{P}'$ una soluzione dell'A.R.E. Vale allora la*

$$P(t+1) - \bar{P} = (F - \bar{K}C) (P(t) - \bar{P}) (F - K(t)C)' \quad , \tag{1.29}$$

dove \bar{K} e $K(t)$ sono legate a \bar{P} e $P(t)$ dalle (1.21) e (1.2) rispettivamente.

²Formalmente analoga a quella del filtro a tempo continuo.

DIMOSTRAZIONE. La (1.5) e la (1.3) possono essere riscritte come

$$\bar{P} = (F - \bar{K} C) \bar{P} F' + \tilde{Q}$$

e

$$P(t+1) = FP(t) [F - K(t) C]' + \tilde{Q} \quad .$$

Sottraendo membro a membro si ottiene

$$\begin{aligned} P(t+1) - \bar{P} &= FP(t) (F - K(t) C)' - (F - \bar{K} C) \bar{P} F' \\ &= (F - \bar{K} C) [P(t) - \bar{P}] (F - K(t) C)' \\ &\quad + \bar{K} C P(t) (F - K(t) C)' - (F - \bar{K} C) \bar{P} C' K'(t) \quad . \end{aligned}$$

Si tratta ora di far vedere che la differenza degli ultimi due termini è zero. Usando l'espressione (1.28) e denotando con i simboli $\bar{\Lambda}$ e $\bar{\Pi}$ le quantità

$$\bar{\Lambda} := C \bar{P} C' + R \quad , \quad \bar{\Pi} := \bar{P} - \bar{P} C' \bar{\Lambda}^{-1} C \bar{P} \quad , \quad (1.30)$$

si ottiene

$$\begin{aligned} &F \bar{\Pi} C' R^{-1} C P(t) (F' - C' \Lambda(t)^{-1} C P(t) F') \\ &\quad - (F - F \bar{P} C' \bar{\Lambda}^{-1} C) \bar{P} C' R^{-1} C \Pi(t) F' \\ &= F \bar{\Pi} C' R^{-1} C (P(t) - P(t) C' \Lambda(t)^{-1} C P(t)) F' \\ &\quad - F (\bar{P} - \bar{P} C' \bar{\Lambda}^{-1} C \bar{P}) C' R^{-1} C \Pi(t) F' \\ &= F \bar{\Pi} C' R^{-1} C \Pi(t) F' - F \bar{\Pi} C' R^{-1} C \Pi(t) F' = 0 \quad . \end{aligned}$$

che è quanto occorre dimostrare. \square

Questa formula ha una serie di importanti conseguenze.

PROPOSIZIONE 10.4. *Se $(F, \tilde{Q}^{1/2})$ è stabilizzabile, può esserci al più una soluzione $\bar{P} = \bar{P}' \geq 0$ dell'A.R.E. (necessariamente stabilizzante).*

DIMOSTRAZIONE. Siano \bar{P}_1 e \bar{P}_2 due soluzioni simmetriche e semidefinite positive dall'A.R.E. In base al lemma 10.5 le due matrici

$$\bar{\Gamma}_1 := F - \bar{K}_1 C \quad , \quad \bar{\Gamma}_2 := F - \bar{K}_2 C \quad ,$$

sono allora asintoticamente stabili. Consideriamo la formula (1.29) del lemma 10.7 in cui $P(t)$ è la soluzione dell'equazione di Riccati (1.3) con $P(t_0) = \bar{P}_1$ e $\bar{P} = \bar{P}_2$. Si ha ovviamente $P(t) = \bar{P}_1, \forall t \geq t_0$, per cui la (1.29) dà

$$\bar{P}_1 - \bar{P}_2 = \bar{\Gamma}_2 (\bar{P}_1 - \bar{P}_2) \bar{\Gamma}_1' \quad .$$

Moltiplicando a sinistra per $\bar{\Gamma}_2$ e a destra per $\bar{\Gamma}_1'$ si ottiene allora

$$\bar{P}_1 - \bar{P}_2 = \bar{\Gamma}_2^2 (\bar{P}_1 - \bar{P}_2) (\bar{\Gamma}_1')^2$$

e iterando il procedimento

$$\bar{P}_1 - \bar{P}_2 = \bar{\Gamma}_2^k (\bar{P}_1 - \bar{P}_2) (\bar{\Gamma}_1')^k \quad .$$

Dato che $\bar{\Gamma}_1$ e $\bar{\Gamma}_2$ sono asintoticamente stabili, passando al limite per $k \rightarrow \infty$ si ottiene $\bar{P}_1 - \bar{P}_2 = 0$. \square

Siamo ora in grado di enunciare il risultato centrale della teoria asintotica del filtro di Kalman.

TEOREMA 10.1 (Teorema fondamentale della teoria del filtro di Kalman). *Condizione necessaria e sufficiente affinché*

- (1) *esista un'unica soluzione $\bar{P} = \bar{P}' \geq 0$ dell'A.R.E.,*
- (2) *\bar{P} sia stabilizzante,*
- (3) *$\lim_{t \rightarrow \infty} P(t) = \bar{P}$ per qualunque condizione iniziale $P_0 = P'_0 \geq 0$,*

è che (F, C) sia rivelabile ed $(F, \tilde{Q}^{1/2})$ sia stabilizzabile.

DIMOSTRAZIONE. Che le prime due proprietà discendano dalla rivelabilità di (F, C) e dalla stabilizzabilità di $(F, \tilde{Q}^{1/2})$ è appena stato dimostrato. Per dimostrare anche che queste condizioni implicano la condizione (3) del Teorema, occorre innanzitutto stabilire una proprietà di stabilità dalla matrice in catena chiusa $\Gamma(t) := F - K(t)C$.

Facciamo vedere che se (F, C) è rivelabile, allora la matrice di transizione di stato del predittore

$$\Phi(t, s) := \prod_{i=s}^{t-1} \Gamma(i)$$

è limitata per ogni condizione iniziale $P_0 = P'_0$ strettamente definita positiva.

In effetti dalla (1.17) si ricava, per $t > s$,

$$P(t) = \Phi(t, t_0) P_0 \Phi(t, t_0)' + [\text{termine} \geq 0]$$

per cui, sotto l'ipotesi di rivelabilità,

$$U \geq P(t) \geq \Phi(t, t_0) P_0 \Phi(t, t_0)' \quad , \quad \forall t \geq t_0 \quad .$$

Sia ora ρ il più piccolo autovalore (necessariamente positivo) di P_0 . Dato che $\rho = \min \mathbf{x}' P_0 \mathbf{x} / \|\mathbf{x}\|^2$ e quindi $P_0 \geq \rho I$, si ha

$$U \geq \rho \Phi(t, t_0) \Phi(t, t_0)' \quad ,$$

da cui scende

$$\text{tr } U \geq \rho \text{tr } \Phi(t, t_0) \Phi(t, t_0)' = \rho \sum_{kj} \Phi_{kj}^2(t, t_0) \quad .$$

In conclusione, $\Phi_{kj}(t, t_0)$ è limitato $\forall t \geq t_0, \forall k, j$, purché, beninteso, $P_0 > 0$. Supponiamo ora anche che $(F, \tilde{Q}^{1/2})$ sia stabilizzabile.

Per la Proposizione 10.4 esiste allora un'unica soluzione $\bar{P} = \bar{P}' \geq 0$ dell'A.R.E. e questa soluzione è tale che $\bar{\Gamma} = F - \bar{K}C$ è asintoticamente stabile. Supponiamo che $P(t_0) = P_0 = P'_0 > 0$. Dalla formula (1.29) del lemma 10.7 scende allora che

$$P(t) - \bar{P} = \bar{\Gamma}^{t-t_0} (P_0 - \bar{P}) \Phi(t, t_0)' \quad (1.31)$$

e, dato che nell'ipotesi che sia $P_0 > 0$, $\Phi(t, t_0)$, è limitata, questa relazione implica immediatamente che

$$\lim_{t \rightarrow \infty} P(t) = \bar{P} \quad , \quad \forall P_0 = P'_0 > 0 \quad .$$

D'altro canto la soluzione $P_0(t)$ corrispondente a condizione iniziale nulla converge anch'essa (Proposizione 10.3) e necessariamente verso la stessa soluzione varianza \bar{P} dell'A.R.E.

Sia allora P_0 una qualunque matrice simmetrica e semidefinita positiva e λ il massimo autovalore di P_0 (che, per evitare di ricadere nel caso $P_0 = 0$ supponiamo sia un numero strettamente positivo) per cui

$$0 \leq P_0 \leq \lambda I \quad .$$

In base al lemma 10.4 la soluzione dell'equazione di Riccati con condizione iniziale $P(0) = \lambda I$, domina la soluzione con condizione iniziale $P(0) = P_0$ la quale a sua volta domina $P_0(t)$ per ogni t . Dato che la prima e l'ultima soluzione convergono allo stesso limite \bar{P} per $t \rightarrow \infty$, si può allora concludere che $P(t)$ converge a \bar{P} , e questo accade qualunque sia la varianza iniziale P_0 .

Supponiamo ora che valgano le tre condizioni del Teorema. È ovvio dalla condizione (2) che (F, C) deve essere rivelabile e rimane solo da far vedere che $(F, \tilde{Q}^{1/2})$ deve essere stabilizzabile.

Sia $\{\xi(t)\}$ il processo n -dimensionale soluzione dell'equazione di stato

$$\xi(t+1) = F \xi(t) + \tilde{v}(t) \quad , \quad \xi(t_0) = \mathbf{x}_0 \quad , \quad (1.32)$$

accoppiato al processo di osservazione $\{\eta(t)\}$ tramite la

$$\eta(t) = C \xi(t) + \mathbf{w}(t) \quad , \quad t \geq t_0 \quad .$$

In queste formule $\tilde{v}(t)$ e $\mathbf{w}(t)$ hanno il significato usuale. Il modello è quindi a rumori scorrelati e il predittore a minima varianza $\hat{\xi}(t)$, di $\xi(t)$, soddisfa allora alla

$$\begin{aligned} \hat{\xi}(t+1) &= F \hat{\xi}(t) + K(t) [\eta(t) - C \hat{\xi}(t)] \quad , \quad t \geq t_0 \\ \hat{\xi}(t_0) &= \bar{\mathbf{x}}_0 \quad , \end{aligned}$$

dove $K(t)$ è dato dalla solita formula (1.2). Notiamo ora che l'errore di predizione $\tilde{\xi}(t) = \xi(t) - \hat{\xi}(t)$ coincide con $\tilde{\mathbf{x}}(t)$ (perché soddisfa esattamente la stessa equazione alle differenze, dato che il termine aggiuntivo $SR^{-1}\mathbf{y}(t)$ nelle equazioni relative a $\mathbf{x}(t)$ e $\tilde{\mathbf{x}}(t)$ si elide calcolando la differenza $\tilde{\mathbf{x}}(t)$) e pertanto si ha

$$P(t) = E \tilde{\mathbf{x}}(t) \tilde{\mathbf{x}}(t)' = E \tilde{\xi}(t) \tilde{\xi}(t)' = V(t) - \hat{V}(t) \quad , \quad (1.33)$$

dove $V(t)$ e $\hat{V}(t)$ sono le matrici di varianza dei processi $\{\xi(t)\}$ e $\{\hat{\xi}(t)\}$, rispettivamente.

Sia allora a un vettore ortogonale al sottospazio raggiungibile di $(F, \tilde{Q}^{1/2})$, tale cioè che

$$a^* F^k \tilde{Q}^{1/2} = 0 \quad , \quad k = 0, 1, 2, \dots \quad (1.34)$$

Detta $V_0(t)$ la varianza di stato di $\xi(t)$ corrispondente alla condizione iniziale $V_0(t_0) = 0$, risolvendo l'equazione di Lyapunov $V_0(t+1) = F V_0(t) F' + \tilde{Q}$ e moltiplicando per a^* e a si trova

$$a^* V_0(t) a = a^* \sum_0^{t-t_0-1} F^k \tilde{Q} (F')^k a = 0 \quad , \quad \forall t \geq t_0 \quad ,$$

per la (1.34). La (1.33) implica allora che

$$a^* V_0(t) a = a^* P_0(t) a + a^* \hat{V}_0(t) a = 0 \quad ,$$

dove $P_0(t)$ è la soluzione della R.E. con condizione iniziale $P_0 = 0$. Dato che il secondo membro di questa espressione è la somma di due termini non negativi, si dovrà necessariamente avere

$$a^* P_0(t) a = 0 \quad , \quad \forall t \geq t_0 \quad ,$$

per cui, passando al limite per $t \rightarrow \infty$, si ottiene, per la condizione 3,

$$a^* \bar{P} a = 0 \quad . \quad (1.35)$$

Questa relazione ci dice che *ogni vettore del sottospazio non raggiungibile di $(F, \tilde{Q}^{1/2})$ sta nello spazio nullo di \bar{P} .*

Supponiamo ora che $(F, \tilde{Q}^{1/2})$ non sia stabilizzabile. Esisterà allora un vettore $a \neq 0$ per cui

$$\begin{aligned} a^* F &= \lambda_0 a^* \\ a^* \tilde{Q}^{1/2} &= 0 \quad , \end{aligned}$$

con $|\lambda_0| \geq 1$. Ovviamente a è ortogonale allo spazio raggiungibile per $(F, \tilde{Q}^{1/2})$. Calcolando $a^* \bar{K}$ si trova

$$a^* \bar{K} = a^* F \bar{P} C' \bar{\Lambda}^{-1} = \lambda_0 a^* \bar{P} C' \bar{\Lambda}^{-1} = 0 \quad (1.36)$$

dove l'ultima uguaglianza è una conseguenza della (1.35). Ne scende che

$$a^*(F - \bar{K} C) = a^* F = \lambda_0 a^* \quad ,$$

con $|\lambda_0| \geq 1$. Questo significa che $F - \bar{K} C$ non è asintoticamente stabile, contraddicendo la proprietà (2).

Abbiamo così dimostrato che $(F, \tilde{Q}^{1/2})$ dev'essere stabilizzabile e questo conclude la dimostrazione del Teorema. \square

Possiamo ora dare la dimostrazione della Proposizione 10.2.

Prova della Proposizione 10.2. Ciò che bisogna dimostrare è la rivelabilità di (F, C) nell'ipotesi di stabilizzabilità di $(F, \tilde{Q}^{1/2})$ e di limitatezza di $P(t)$.

Supponiamo allora che (F, C) non sia rivelabile. Esiste allora un vettore $a \neq 0$ per cui

$$Fa = \lambda_0 a \quad , \quad Ca = 0 \quad , \quad |\lambda_0| \geq 1 \quad .$$

Ma dalla seconda di queste relazioni scende immediatamente che

$$(F - \bar{K} C)a = Fa - \bar{K} Ca = \lambda_0 a \quad ,$$

il che implica che non può esistere alcuna soluzione $\bar{P} = \bar{P}' \geq 0$ dell'A.R.E.

In effetti se una tale soluzione esistesse (e dovrebbe allora essere unica!) dovrebbe necessariamente essere $|\lambda(F - \bar{K} C)| < 1$ (lemma 10.5).

Ma allora la soluzione dell'equazione di Riccati $P(t)$ non può essere limitata $\forall P_0$. Per esempio, la soluzione $P_0(t)$ corrispondente alla condizione iniziale $P_0 = 0$, è monotona crescente (Corollario (10.2)) e l'esistenza di una soluzione varianza \bar{P} dell'ARE è equivalente a quella di un limite superiore per $P_0(t)$. \square

Divergenza del Filtro di Kalman. Notiamo come la formula (1.35) ricavata nel corso della dimostrazione del teorema precedente metta chiaramente in luce la possibilità di predire esattamente (cioè con varianza d'errore uguale a zero) certe funzioni lineari di $\mathbf{x}(t)$ del tipo $a^* \mathbf{x}(t)$. Queste combinazioni lineari *non sono affette da rumore di modello* dato che, se il vettore a è ortogonale allo spazio raggiungibile per $(F, \tilde{Q}^{1/2})$, $a^* \mathbf{x}(t)$ evolve nel tempo in modo indipendente da $\{\tilde{v}(t)\}$. In questo caso il guadagno a regime \bar{K} si annulla nella direzione a , ovvero si ha $a^* \bar{K} = 0$.

Il fatto che \bar{K} si annulli è molto pericoloso, perché significa che la "correzione" apportata dalla reazione nel filtro scompare completamente. In altri termini, lungo la direzione a , la predizione di $\mathbf{x}(t)$ viene, a regime, fatta solo in base al modello dinamico di evoluzione dello stato, senza tener conto delle misure.

Il fenomeno è molto evidente nel caso in cui i rumori $\{\mathbf{v}(t)\}$ e $\{\mathbf{w}(t)\}$ sono scorrelati ($S = 0$). Se a è un vettore ortogonale allo spazio raggiungibile per $(A, Q^{1/2})$, allora la dinamica dello stato e del predittore a regime lungo la direzione a diventano

$$a^* \mathbf{x}(t+1) = a^* A \mathbf{x}(t) + a^* Q^{1/2} \mathbf{n}(t) = a^* A \mathbf{x}(t) \quad (1.37)$$

e

$$a^* \hat{\mathbf{x}}(t+1) = a^* A \hat{\mathbf{x}}(t) + a^* \bar{K} \mathbf{e}(t) = a^* A \hat{\mathbf{x}}(t) \quad (1.38)$$

Ora, se si ha stabilizzabilità di $(A, Q^{1/2})$, gli autovalori λ_i associati ad a ($a = \sum \lambda_i a_i$, $a_i^* A = \lambda_i a_i^*$) hanno necessariamente modulo minore di uno e quindi $a^* \tilde{\mathbf{x}}(t) \rightarrow 0$ per $t \rightarrow \infty$.

In pratica però i modelli sono sempre *descrizioni approssimate* della realtà e l'ipotetico stato "vero" del sistema "reale" potrebbe benissimo avere componenti lentamente variabili di drift o di deriva che non sono state modellata esplicitamente nell'equazione di stato perchè ritenute trascurabili. In questo caso l'errore di stima nella direzione a non tende affatto a zero per $t \rightarrow \infty$.

Per spiegare questo fenomeno, supponiamo che lo stato "vero" del sistema, detto $\{\mathbf{x}_0(t)\}$, sia descritto da un modello "vero", ancora lineare e invariante, del tipo

$$\mathbf{x}_0(t+1) = A_0 \mathbf{x}_0(t) + \mathbf{v}_0(t) \quad .$$

in cui (a puro titolo esemplificativo) supporremo che $\mathbf{x}_0(t)$ abbia la stessa dimensione di $\mathbf{x}(t)$ e sia legato alle osservazioni della solita relazione $\mathbf{y}(t) = C \mathbf{x}_0(t) + \mathbf{w}(t)$ con $\{\mathbf{v}_0(t)\}$ e $\{\mathbf{w}(t)\}$ scorrelati. Il modello "vero" (ammesso che questa espressione abbia un qualche senso) non è noto e si usa invece l'approssimazione (rivelabile e stabilizzabile)

$$\mathbf{x}(t+1) = A \mathbf{x}(t) + \mathbf{v}(t) \quad (1.39)$$

che, nel caso in esame, conduce alle relazioni (1.37) e (1.38).

Ora, "l'errore vero", $\tilde{\mathbf{x}}_0(t)$, anziché tendere a zero lungo la direzione a , come fa $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t)$, rimane sempre non nullo, dato che $\tilde{\mathbf{x}}_0(t) = \mathbf{x}_0(t) - \hat{\mathbf{x}}(t)$ soddisfa (se si usa il predittore a regime relativo al modello (1.39)) alla

$$a^* \tilde{\mathbf{x}}_0(t+1) = a^* A_0 \mathbf{x}_0(t) - a^* A \hat{\mathbf{x}}(t) + a^* \mathbf{v}_0(t)$$

inoltre, dato che $a^* \hat{\mathbf{x}}(t) \rightarrow 0$, si ha praticamente $a^* \tilde{\mathbf{x}}_0(t) = a^* \mathbf{x}_0(t)$ per t grandi. Se il modello "vero" ha qualche autovalore di modulo maggiore o uguale a uno, l'errore "vero" anziché tendere a zero ($a^* P a = 0$) può benissimo divergere.

Questo fenomeno si chiama *divergenza del filtro* ed è caratteristico di modelli troppo "ottimistici", nei quali si suppone che certe combinazioni lineari delle variabili di stato evolvano nel tempo in modo *deterministico* (la combinazione lineare non è affetta da incertezza, il rumore di modello relativo è zero).

In questi casi il filtro "crede troppo" al modello (che in realtà è sempre approssimato) proprio perchè gli si dice che il modello è "perfetto" (almeno per certe combinazioni lineari delle variabili di stato) e tende a eliminare la correzione nella stima dovuta alle misure. In effetti il guadagno a regime si annulla ($a^* \bar{K} = 0$) proprio nella direzione in cui il modello è "perfetto".

Per evitare il fenomeno della divergenza occorre far sì che *il guadagno a regime non si annulli mai*. In questo caso la reazione tende a compensare le discrepanze tra il modello che si usa nel progetto del filtro e il modello "vero". Si noti infatti che in ogni caso *le misure sono, per loro natura, sempre effettuate sullo stato "vero"*,

$\{\mathbf{x}_0(t)\}$, *del sistema*. La reazione tende così automaticamente a correggere la stima fornendo un segnale proporzionale (a meno del rumore $\mathbf{w}(t)$) a $\mathbf{x}_0(t) - \hat{\mathbf{x}}(t)$.

La condizione che esclude l'esistenza di direzioni lungo le quali il modello è "perfetto" è ovviamente la *completa raggiungibilità di* $(F, \tilde{Q}^{1/2})$ (nel qual caso tutte le componenti di $\mathbf{x}(t)$ sono affette da rumore di modello).

PROPOSIZIONE 10.5. *Se $(F, \tilde{Q}^{1/2})$ è completamente raggiungibile, allora ogni (eventuale) soluzione $\bar{P} = \bar{P}' \geq 0$ dell'A.R.E. è definita positiva.*

DIMOSTRAZIONE. Notiamo che \bar{P} è necessariamente unica e stabilizzante. Detto \bar{K} il relativo guadagno a regime, osserviamo che la raggiungibilità di $(F, \tilde{Q}^{1/2})$ implica banalmente quella di $(F, [\bar{K}R^{1/2} : \tilde{Q}^{1/2}])$ e questa a sua volta quella della coppia $(F - \bar{K}C, [\bar{K}R^{1/2} : \tilde{Q}^{1/2}])$ dato che $F - \bar{K}C$ corrisponde alla legge di retroazione sullo stato del sistema

$$\boldsymbol{\xi}(t+1) = F \boldsymbol{\xi}(t) + \begin{bmatrix} \bar{K}R^{1/2} & : & \tilde{Q}^{1/2} \end{bmatrix} \mathbf{u}(t) \quad ,$$

definita dalla relazione

$$\mathbf{u}(t) = - \begin{bmatrix} R^{-1/2} C \\ 0 \end{bmatrix} \boldsymbol{\xi}(t) + r(t) \quad .$$

Supponiamo che \bar{K} sia noto. La matrice \bar{P} è allora soluzione dell'equazione di Lyapunov

$$\bar{P} = (F - \bar{K}C) \bar{P} (F - \bar{K}C)' + \begin{bmatrix} \bar{K}R^{1/2} & : & \tilde{Q}^{1/2} \end{bmatrix} \begin{bmatrix} R^{1/2} \bar{K}' \\ (Q^{1/2})' \end{bmatrix} \quad , \quad (1.40)$$

in cui $(F - \bar{K}C)$ è asintoticamente stabile e si ha raggiungibilità completa. Segue che $\bar{P} > 0$. \square

La formula (1.29) fornisce un modo per valutare la durata del transitorio durante il quale $P(t)$ è da riguardarsi "effettivamente" variante nel tempo.

Nell'espressione esplicita (1.31) si vede che $P(t) - \bar{P}$ tende a zero con una velocità che è grosso modo determinata dall'*autovalore di* $F - \bar{K}C$ *a massimo modulo*, λ_M ,

$$P(t) - \bar{P} = O(|\lambda_M|^{t-t_0}) \quad . \quad (1.41)$$

Questo fatto risulta assai utile nel giudicare sull'opportunità di *usare il filtro a regime fin dall'inizio* risparmiando la memoria necessaria a immagazzinare i valori del guadagno in transitorio $K(t)$. \diamond

Quanto sopra esposto si può considerare una esposizione dei fatti essenziali della teoria asintotica del predittore di Kalman. Nel caso in cui si voglia considerare la *stima filtrata* $\hat{\mathbf{x}}(t | t)$ ci si può riportare facilmente ai risultati relativi al predittore. Per completezza di esposizione riportiamo qui le formule relative.

Le equazioni del filtro sono

$$\begin{aligned} \hat{\mathbf{x}}(t+1 | t+1) &= F \hat{\mathbf{x}}(t | t) + L(t+1) [\mathbf{y}(t+1) - CF \hat{\mathbf{x}}(t | t)] \\ &+ \left(I - L(t+1) C \right) SR^{-1} \mathbf{y}(t) \quad , \end{aligned} \quad (1.42)$$

dove il guadagno del filtro, $L(t)$, può essere espresso mediante la formula ricavata nel lemma (10.6)

$$L(t) = \Pi(t) C' R^{-1} \quad , \quad (1.43)$$

nella quale $\Pi(t) := E \tilde{\mathbf{x}}(t | t) \tilde{\mathbf{x}}(t | t)'$ soddisfa l'equazione

$$\Pi(t) = P(t) - P(t)C'[CP(t)C' + R]^{-1}CP(t) \quad . \quad (1.44)$$

Notiamo innanzitutto che il $\lim_{t \rightarrow \infty} \Pi(t)$ esiste se e solo se esiste il corrispondente limite della varianza dell'errore di predizione $P(t)$. In effetti dalla (1.44) segue che se $\bar{P} = \lim P(t)$ allora il corrispondente limite di $\Pi(t)$ si calcola immediatamente come

$$\bar{\Pi} = \bar{P} - \bar{P}C'[C\bar{P}C' + R]^{-1}C\bar{P} \quad .$$

Viceversa si ha

$$\bar{P} = F\bar{\Pi}F' + \tilde{Q}$$

in base alla (4.24).

Il guadagno a regime del filtro $\bar{L} = \bar{\Pi}C'R^{-1}$ entra nell'espressione della matrice $\bar{\Gamma}_1$ che descrive la dinamica dello stimatore $\hat{\mathbf{x}}(t | t)$ a regime,

$$\bar{\Gamma}_1 := F - \bar{L}CF = (I - \bar{L}C)F \quad .$$

Come già notato nel Capitolo (9), la matrice in catena chiusa del predittore a regime, $\bar{\Gamma}$, si può viceversa esprimere come

$$\bar{\Gamma} = F - \bar{K}C = F - F\bar{L}C = F(I - \bar{L}C) \quad .$$

Da queste due espressioni segue che $\bar{\Gamma}$ e $\bar{\Gamma}_1$ hanno gli stessi autovalori e pertanto la stabilità asintotica del filtro si ha esattamente sotto le stesse condizioni che garantiscono quella del predittore.

Nota. Per provare che AB e BA hanno lo stesso spettro basta moltiplicare a sinistra le relazioni

$$ABx = \lambda x \quad , \quad BAy = \mu y \quad , \quad x, y \neq 0 \quad ,$$

per B e per A rispettivamente, ottenendo

$$BA(Bx) = \lambda(Bx) \quad , \quad AB(Ay) = \mu(Ay) \quad ,$$

da cui

$$\sigma(BA) \subseteq \sigma(AB) \quad \text{e} \quad \sigma(AB) \subseteq \sigma(BA) \quad .$$

dove $\sigma(A)$ sta per l'insieme degli autovalori (lo spettro) di A . ◇

ESEMPIO 10.1 (Un esempio scalare). Cercheremo di discutere nel modo più completo possibile un semplice esempio scalare. Consideriamo il modello a rumori scorrelati

$$\begin{aligned} \mathbf{x}(t+1) &= a\mathbf{x}(t) + \mathbf{v}(t) \quad , \quad \text{var } \mathbf{v}(t) = q \geq 0 \quad , \\ \mathbf{y}(t) &= \mathbf{x}(t) + \mathbf{w}(t) \quad , \quad \text{var } \mathbf{w}(t) = r \geq 0 \quad . \end{aligned} \quad (1.45)$$

L'equazione di Riccati per la varianza dell'errore di predizione di un passo, $p(t)$, si scrive

$$p(t+1) = a^2 \frac{r}{p(t) + r} p(t) + q \quad (1.46)$$

e l'equazione dinamica per il predittore è così

$$\hat{\mathbf{x}}(t+1) = a\hat{\mathbf{x}}(t) + a \frac{p(t)}{p(t) + r} [\mathbf{y}(t) - \hat{\mathbf{x}}(t)] \quad . \quad (1.47)$$

Distingueremo un certo numero di casi possibili.

1. Caso di misure perfette ($r = 0$).

L'equazione di Riccati si riduce a

$$p(t+1) = q \quad p(0) = p_0 \quad ,$$

e così la varianza dell'errore di predizione coincide con quella del rumore di modello $\{\mathbf{v}(t)\}$ per tutti i $t > 0$.

L'equazione del predittore diventa

$$\hat{\mathbf{x}}(t+1) = a \mathbf{y}(t) \quad ,$$

la quale rappresenta ovviamente anche l'espressione del predittore di un passo di $\mathbf{x}(t+1)$ in base a $\mathbf{x}(t)$. Ciò rende ragione del fatto che $p(t) = q$. Notiamo che il filtro non ha dinamica.

Nel caso in cui $r > 0$, l'analisi che segue richiederà di discutere le soluzioni dell'ARE che si può scrivere nella forma

$$p = a^2 \frac{pr}{p+r} + q \quad , \quad (1.48)$$

e di studiare il comportamento asintotico del predittore.

2. Caso di modello perfetto ($q = 0$).

Considereremo dapprima la situazione in cui *si ha stabilizzabilità*, ovvero $|a| < 1$ (il modello è asintoticamente stabile).

L'A.R.E. diventa

$$p^2 + (1 - a^2)rp = 0 \quad ,$$

che ammette le due soluzioni $\bar{p} = 0$ e $\bar{p} = -(1 - a^2)r$.

Solo la prima è semidefinita positiva. Usando la (1.49) con $\bar{p} = 0$ si trova

$$p(t+1) = a^2 \frac{r}{p(t)+r} p(t) \quad , \quad p(0) = p_0 \geq 0 \quad .$$

Essendo sempre $a^2 r [p(t) + r]^{-1} < 1$ (dato che $p(t) \geq 0$), si vede allora che $p(t)$ decresce monotonicamente a zero per $t \rightarrow \infty$, qualunque sia $p_0 \geq 0$.

Se il modello *non è stabilizzabile* ($|a| \geq 1$), l'A.R.E. ha viceversa *due* soluzioni semidefinite

$$\bar{p} = 0 \quad , \quad \bar{p} = (a^2 - 1)r$$

(coincidenti nel caso in cui $|a| = 1$). Questo è un caso in cui non è applicabile il teorema fondamentale (10.1).

Bisogna studiare la stabilità delle due soluzioni di regime e le proprietà asintotiche del filtro relativo. Detta \bar{p} una soluzione dell'ARE, la formula (1.29) assume l'aspetto seguente

$$p(t+1) - \bar{p} = \frac{a^2 r}{\bar{p} + r} \frac{r}{p(t) + r} (p(t) - \bar{p}) \quad . \quad (1.49)$$

Prendendo $\bar{p} = (a^2 - 1)r$ nella (1.49), si trova

$$p(t+1) - \bar{p} = \frac{r}{p(t) + r} (p(t) - \bar{p}) \quad .$$

Se $p(0) > 0$, allora $p(t) > 0$ per tutti i t (cfr. la (1.46)) e il coefficiente $r/[p(t) + r]$ è strettamente minore di 1. Ne segue che per qualunque condizione iniziale $p_0 > 0$, $p(t)$ tende a $\bar{p} = (a^2 - 1)r$ quando $t \rightarrow \infty$.

Evidentemente allora il punto di equilibrio $\bar{p} = 0$ dovrà essere instabile (almeno per $|a| > 1$). Infatti, ponendo $\bar{p} = 0$ nella (1.49) si trova

$$p(t+1) = \frac{a^2 r}{p(t) + r} p(t)$$

e da questa si ricava

$$p(t+1) - p(t) = \frac{(a^2 - 1)r - p(t)}{p(t) + r} p(t) \quad .$$

Si vede che se $0 < p(t) < (a^2 - 1)r$, l'incremento $p(t+1) - p(t)$ deve essere positivo e quindi $p(t)$ cresce, allontanandosi dall'origine ($\bar{p} = 0$). In conclusione, se $|a| > 1$, qualunque sia la condizione iniziale $p_0 > 0$, si ha $\lim p(t) = (a^2 - 1)r$.

Il corrispondente filtro a regime ha guadagno $\bar{k} = a(1 - 1/a^2)$ ed è quindi descritto dalla matrice in catena chiusa

$$\bar{\gamma} = a - \bar{k} = 1/a \quad ,$$

che è *asintoticamente stabile*.

Se $|a| = 1$, l'unico punto di equilibrio dell'equazione di Riccati è $\bar{p} = 0$ e si può verificare che $p(t) \rightarrow 0$ per qualunque condizione iniziale. In questo caso però la matrice in catena chiusa coincide con a e il predittore a regime non è asintoticamente stabile.

3. Caso generale ($q > 0$, $r > 0$).

Si può applicare il teorema fondamentale; l'equazione algebrica (1.48) ha un'unica soluzione $\bar{p} > 0$ che è un punto di equilibrio asintoticamente stabile per l'equazione di Riccati (1.46). Questa affermazione si può verificare direttamente notando che se $\bar{p} > 0$ risolve la (1.48), necessariamente $(a^2 r / [\bar{p} + r]) \bar{p} < \bar{p}$. Segue che il coefficiente che moltiplica $p(t) - \bar{p}$ nella (1.49) è sempre strettamente minore di 1 e quindi $p(t) - \bar{p}$ deve tendere a zero per $t \rightarrow \infty$.

La matrice in catena chiusa è $\bar{\gamma} = ar / (\bar{p} + r)$, che è ovviamente asintoticamente stabile qualunque sia a (si ha $\bar{\gamma} < 1/|a|$ per $|a| \geq 1$ e $\bar{\gamma} < a$ per $|a| < 1$). \diamond

2. Fattorizzazione spettrale e ARE

Anche in questa sezione manterremo ferme le ipotesi di invarianza temporale dei parametri che definiscono il modello dei segnali in gioco. Converrà usare la notazione introdotta in (1.1) e scrivere

$$\begin{aligned} \mathbf{x}(t+1) &= A\mathbf{x}(t) + B\mathbf{n}(t) \\ \mathbf{y}(t) &= C\mathbf{x}(t) + D\mathbf{n}(t) \end{aligned} \quad (2.1)$$

con $\{\mathbf{n}(t)\}$ rumore bianco normalizzato ($\text{Var } \mathbf{n}(t) = I$) di dimensione p , ortogonale alla condizione iniziale \mathbf{x}_0 . Supporremo $E\mathbf{x}_0 = 0$ e assumeremo sempre il modello (4.21) associato al (2.1) *rivelabile* e *stabilizzabile* ed $R = DD'$ *definita positiva*.

Detta $W(z) = C(zI - A)^{-1}B + D$ la funzione di trasferimento del modello (2.1), si può definire la matrice

$$S(z) := W(z)W'(z^{-1}) \quad . \quad (2.2)$$

Se A è asintoticamente stabile e $t_0 = -\infty$, oppure $t_0 > -\infty$ e la varianza di $\mathbf{x}(t_0)$ soddisfa l'equazione di Lyapunov $P = APA' + Q$, la matrice $S(z)$ è effettivamente la densità spettrale del processo stazionario $\{\mathbf{y}(t)\}$. Più in generale questa interpretazione ha senso quando $\{\mathbf{y}(t)\}$ è asintoticamente stazionario. Ai fini di quanto ci interessa mettere in luce in questo capitolo, l'interpretazione di $S(z)$ come densità spettrale non è però essenziale e noi nel seguito non faremo l'ipotesi che sia $|\lambda(A)| < 1$. Chiameremo $S(z)$ "pseudospettro" in questo caso.

Consideriamo ora il predittore di Kalman a regime per il modello (2.1). Ricordando che l'innovazione (a regime) $\{\mathbf{e}(t)\}$ è definita dalla $\mathbf{e}(t) = \mathbf{y}(t) - C\hat{\mathbf{x}}(t)$, si possono scrivere le relazioni

$$\begin{aligned} \hat{\mathbf{x}}(t+1) &= A\hat{\mathbf{x}}(t) + \bar{G}\mathbf{e}(t) \\ \mathbf{y}(t) &= C\hat{\mathbf{x}}(t) + \mathbf{e}(t) \quad , \end{aligned} \quad (2.3)$$

dove $\bar{G} = (A\bar{P}C' + S)\bar{\Lambda}^{-1}$ è il guadagno *a regime* del predittore (si è qui usata la solita notazione $S = BD'$). Formalmente le condizioni iniziali associate al predittore sono nulle ($\hat{\mathbf{x}}(t_0) = 0$).

L'osservazione cruciale per quanto segue è che la (2.3) può essere riguardata, in perfetta analogia alla (2.1), come *una rappresentazione del processo* $\{\mathbf{y}(t)\}$ *come risposta del sistema lineare di funzione di trasferimento*

$$\hat{W}(z) = C(zI - A)^{-1}\bar{G} + I \quad (2.4)$$

all'ingresso bianco $\{\mathbf{e}(t)\}$. Notiamo per inciso che, a differenza di quella del modello di partenza (2.1), in cui normalmente $p > m$, la funzione di trasferimento del "modello" (2.4) è una matrice **quadrata** $m \times m$.

Ora, nel caso in cui $\{\mathbf{y}(t)\}$ è stazionario, segue dalla ben nota regola per il calcolo dello spettro dell'uscita di un sistema lineare (formla di Wiener-Kintchine), che $\hat{W}(z)$ deve soddisfare alla relazione di fattorizzazione spettrale

$$S(z) = \hat{W}(z)\bar{\Lambda}\hat{W}'(z^{-1})' \quad . \quad (2.5)$$

in cui $\bar{\Lambda}$ è la varianza dell'innovazione a regime.

Estrapolando questo risultato e dando per buono che esso valga anche per processi descritti da modelli non stazionari in cui A non ha tutti gli autovalori dentro al cerchio unità, si può immaginare per analogia, che il calcolo del predittore a regime possa essere associato alla ricerca di fattori spettrali quadrati $m \times m$ del

pseudospettro (2.2). Questa analogia fa pensare che la soluzione del problema di fattorizzazione spettrale (2.5) per $S(z)$, sia strettamente “imparentata” alla soluzione dell’equazione algebrica di Riccati che conduce al calcolo del predittore a regime. In questo capitolo vedremo che si tratta in realtà di molto di più che di una semplice analogia e stabiliremo una importante equivalenza tra i due problemi di fattorizzazione spettrale e soluzione dell’ARE. Questa equivalenza è uno dei fatti generali e profondi della teoria del filtraggio a minima varianza e ci porterà ad una comprensione migliore di tutto il problema. In particolare, i metodi per la risoluzione numerica dell’ARE che verranno esposti nel prossimo capitolo, sono basati su questa corrispondenza.

La relazione di fattorizzazione spettrale (2.5), che è ovvia nel caso di processi stazionari, richiede però una dimostrazione nel caso in cui A non è asintoticamente stabile. Il lemma che segue servirà in particolare a stabilire questo fatto, ma si tratta in realtà di un enunciato molto più generale. L’idea della dimostrazione risale a Kalman, Yakubovich e Popov [76, 158, 117].

LEMMA 10.8. *Una qualunque soluzione, \bar{P} , simmetrica ma non necessariamente definita dell’equazione algebrica di Riccati associata al modello lineare (2.1),*

$$P = APA' - G \Lambda G' + Q \quad G := (APC' + S)\Lambda^{-1} \quad (2.6)$$

definisce, attraverso la formula (2.4), dove

$$\bar{G} \bar{\Lambda} = A\bar{P}C' + S \quad \bar{\Lambda} = C\bar{P}C' + R, \quad (2.7)$$

un fattore spettrale quadrato ($m \times m$), normalizzato all’infinito ($W(\infty) = I$), del pseudospettro $S(z)$.

DIMOSTRAZIONE. Per dimostrare il lemma supponiamo che $\bar{P} = \bar{P}'$ sia una soluzione della (2.6). Notiamo innanzitutto che dall’identità

$$(zI - A)\bar{P}(z^{-1}I - A') = \bar{P} - A\bar{P}(z^{-1}I - A') - (zI - A)\bar{P}A' - A\bar{P}A'$$

segue che (nelle formule che seguono useremo l’abbreviazione $A(z) := (zI - A)$)

$$\bar{P} - A\bar{P}A' = A\bar{P}A'(z^{-1}) + A(z)\bar{P}A' + A(z)\bar{P}A'(z^{-1}) \quad (2.8)$$

Introducendo questa espressione nella (2.6) e moltiplicando a sinistra per $CA(z)^{-1}$ e a destra per $A(z^{-1})^{-\top}C'$ si trova

$$CA(z)^{-1}A\bar{P}C' + C\bar{P}A'A(z^{-1})^{-\top}C' + C\bar{P}C' + \\ + CA(z)^{-1}\bar{G}\bar{\Lambda}\bar{G}'A(z^{-1})^{-\top}C' = CA(z)^{-1}QA(z^{-1})^{-\top}C' \quad ,$$

dove $\bar{\Lambda} = C\bar{P}C' + R$, $\bar{G}\bar{\Lambda} = A\bar{P}C' + S$. Sostituendo ora al posto di $A\bar{P}C'$ la quantità $\bar{G}\bar{\Lambda} - S$ e riordinando i vari termini si perviene all’identità

$$\bar{\Lambda} + CA(z)^{-1}\bar{G}\bar{\Lambda} + \bar{\Lambda}\bar{G}'A(z^{-1})^{-\top}C' + CA(z)^{-1}\bar{G}\bar{\Lambda}\bar{G}'A(z^{-1})^{-\top}C' \\ = R + CA(z)^{-1}S + SA(z^{-1})^{-\top}C' + CA(z)^{-1}QA(z^{-1})^{-\top}C' .$$

Questa identità, si può riscrivere come una relazione di fattorizzazione spettrale del tipo (2.5). \square

Più esplicitamente, possiamo concludere affermando che *ad una qualunque soluzione simmetrica P dell’ARE corrisponde una matrice razionale quadrata*

$$W_P(z) = C(zI - A)^{-1}G(P) + I$$

in cui

$$G(P) := (APC' + S)\Lambda(P)^{-1}; \quad \Lambda(P) := CPC' + R,$$

che soddisfa la relazione di fattorizzazione spettrale

$$W_P(z)\Lambda(P)W_P(z^{-1})' = S(z) \quad . \quad (2.9)$$

PROBLEMA 10.1. *Dimostrare che per una qualunque soluzione simmetrica (ma non necessariamente definita) $P = P'$ dell'ARE, la matrice*

$$\Lambda(P) := CPC' + R$$

è definita positiva ed è quindi interpretabile come una varianza.

Suggerimento: *Dalla relazione di fattorizzazione spettrale (2.9) si ricava*

$$\Lambda(P) = W_P(z)^{-1}S(z)W_P(z^{-1})^{-\top} \quad .$$

Il secondo membro di questa eguaglianza dev'essere costante rispetto a z e quindi definito positivo perchè calcolandolo in un qualunque punto $z = e^{j\omega}$ che non sia uno zero di $S(z)$... \diamond

Nel caso in cui si prende \bar{P} uguale all'unica (nelle ipotesi poste) soluzione semi-definita positiva dell'A.R.E. si ottiene il particolare fattore ($m \times m$) uguale alla matrice di trasferimento della rappresentazione di stato di \mathbf{y} ricavato dal predittore di Kalman a regime.

Vediamo, in questo caso, quali proprietà caratterizzano il fattore spettrale $\hat{W}(z)$ definito dalla (2.4).

Innanzitutto notiamo che se fosse $|\lambda(A)| < 1$ allora $\hat{W}(z)$ sarebbe un fattore spettrale *analitico* giacchè i suoi poli sarebbero certamente tutti a modulo minore di uno. Per esaminare gli zeri di $\hat{W}(z)$ conviene procurarsi una realizzazione dell'inversa $\hat{W}(z)^{-1}$. Dalle equazioni (2.3) scende agevolmente che

$$\begin{cases} \hat{\mathbf{x}}(t+1) = (A - \bar{G}C)\hat{\mathbf{x}}(t) + \bar{G}\mathbf{y}(t) \\ \mathbf{e}(t) = -C\hat{\mathbf{x}}(t) + \mathbf{y}(t) \end{cases} \quad (2.10)$$

per cui

$$\hat{W}(z)^{-1} = -C(zI - \bar{\Gamma})^{-1}\bar{G} + I \quad , \quad (2.11)$$

dove $\bar{\Gamma} = A - \bar{G}C$ è la *matrice in catena chiusa del predittore*. Nelle ipotesi fatte gli autovalori di $\bar{\Gamma}$ sono tutti a modulo minore di uno e quindi gli zeri di $\hat{W}(z)$ (uguali in questo caso ai poli dell'inversa) stanno tutti dentro il cerchio unità del piano complesso. Si può dire che $\hat{W}(z)$ è a *fase minima*. Ne scende il seguente utile risultato.

TEOREMA 10.2. *Sotto le ipotesi di rivelabilità e stabilizzabilità menzionate in precedenza, la matrice di trasferimento, $\hat{W}(z)$, della rappresentazione di stato (2.3), di \mathbf{y} ricavata dal predittore di Kalman a regime dello stato del modello (2.1), è il fattore spettrale normalizzato a fase minima, del pseudospettro $S(z)$. In altri termini, $\hat{W}(z)$ è l'unico fattore di $S(z)$ che ha tutti gli zeri di modulo (strettamente) minore di uno ed è normalizzato all'infinito ($\hat{W}(\infty) = I$).*

Se in particolare il modello (2.3) descrive un processo stazionario, allora le equazioni (2.3) ne forniscono la rappresentazione d'innovazione (non normalizzata) in forma di realizzazione di stato.

Equivalentemente, posto

$$L(z) := \hat{W}(z)\bar{\Lambda}^{1/2} \quad (2.12)$$

e definito il processo d'innovazione normalizzata $\{\boldsymbol{\varepsilon}(t)\}$, tramite la

$$\boldsymbol{\varepsilon}(t) := \bar{\Lambda}^{-1/2} \mathbf{e}(t) \quad , \quad (2.13)$$

si ha $L(z)L'(z^{-1}) = S(z)$ ed $L(z)$ è il fattore spettrale canonico. La rappresentazione d'innovazione normalizzata di $\{\mathbf{y}(t)\}$ ha perciò la realizzazione

$$\begin{aligned} \hat{\mathbf{x}}(t+1) &= A \hat{\mathbf{x}}(t) + \bar{G} \bar{\Lambda}^{1/2} \boldsymbol{\varepsilon}(t) \\ \mathbf{y}(t) &= C \hat{\mathbf{x}}(t) + \bar{\Lambda}^{1/2} \boldsymbol{\varepsilon}(t) \quad . \end{aligned} \quad (2.14)$$

Questo enunciato stabilisce la sostanziale equivalenza tra il calcolo del fattore spettrale canonico e la soluzione dell'equazione algebrica di Riccati. Se si lascia cadere l'ipotesi che A abbia tutti gli autovalori dentro il cerchio, o, equivalentemente, la stazionarietà di $\{\mathbf{y}(t)\}$, il concetto di rappresentazione d'innovazione (non stazionaria) ha senso solo su un intervallo finito e alcune delle formule hanno scarso significato. Rimane comunque dimostrato il fatto che gli zeri del fattore spettrale $\hat{W}(z)$ o di $L(z)$, stanno dentro il cerchio unità e si potrebbe ancora dire che $\hat{W}(z)$ è a fase minima e $L(z)$ è il fattore spettrale canonico del pseudospettro $S(z)$ anche se questi fattori spettrali non sono più analitici in generale.

Notiamo che se $L(z)$ è noto nella forma (2.14) (cioè se $L(z)$ è noto tramite una sua realizzazione) il problema della predizione di $\{\mathbf{y}(t)\}$ in base al passato infinito è risolto già in partenza. Basta notare che la prima delle (2.14) definisce lo stato, $\hat{\mathbf{x}}(t)$, come una funzione del passato, $\boldsymbol{\varepsilon}^{t-1}$, dell'innovazione di $\{\mathbf{y}(t)\}$ per cui la seconda equazione è semplicemente la decomposizione ortogonale di $\mathbf{y}(t)$ come somma del predittore di un passo (in base al passato infinito) e dell'errore di predizione. In sostanza si ha

$$\hat{\mathbf{y}}(t | t-1) = C \hat{\mathbf{x}}(t) \quad . \quad (2.15)$$

PROBLEMA 10.2. Si consideri il modello (in $t \geq t_0$)

$$\begin{aligned} \mathbf{x}(t+1) &= A \mathbf{x}(t) + LR^{1/2} \mathbf{n}(t) \\ \mathbf{y}(t) &= C \mathbf{x}(t) + R^{1/2} \mathbf{n}(t) \quad , \end{aligned} \quad (2.16)$$

dove $\{\mathbf{n}(t)\}$ è rumore bianco normalizzato di dimensione m (uguale alla dimensione di $\mathbf{y}(t)$), R è non singolare, $R^{1/2}$ è una sua radice quadrata ed L è tale che $|\lambda(A - LC)| < 1$. Usando le notazioni introdotte nei paragrafi precedenti e senza assumere alcunché sulla stabilità di A , si mostri che

- (1) valgono le ipotesi del teorema (10.1) e l'unica soluzione simmetrica e semidefinita positiva dell'A.R.E. associata al modello (2.16) è $\bar{P} = 0$;
- (2) qualunque sia la condizione iniziale $\mathbf{x}(t_0) = \mathbf{x}_0$, si ha

$$\lim_{t \rightarrow \infty} \tilde{\mathbf{x}}(t) = 0$$

(in media quadratica) e quindi, per $t \rightarrow \infty$, $\hat{\mathbf{x}}(t) = \mathbf{x}(t)$. Inoltre il guadagno del predittore di un passo $G(t)$ tende alla costante L quando $t \rightarrow \infty$;

- (3) il modello (2.16) coincide con il predittore di Kalman a regime di $\mathbf{x}(t)$.

◇

In seguito, per evitare inutili giri di parole, sarà conveniente assumere che la realizzazione (2.3) del fattore spettrale a fase minima $\hat{W}(z)$ sia una *realizzazione minima* nel senso usuale della teoria dei sistemi (e cioè che la coppia (A, C) sia osservabile e (A, \bar{G}) sia raggiungibile). In questo caso gli autovalori di $\bar{\Gamma} = A - \bar{G}C$ sono tutti e soli gli zeri di $\hat{W}(z)$ e, per la formula di fattorizzazione (2.5) del Teorema 10.2, *questi autovalori vanno ricercati tra quegli zeri del “pseudo-spettro” $S(z)$ che hanno modulo strettamente minore di uno*. In altre parole la dinamica del predittore (e del filtro) a regime è determinata dagli zeri di $S(z)$ che stanno dentro al cerchio unità. Si noti che questo è sostanzialmente l’analogo di quanto accade per il predittore di Wiener.

Per inciso diciamo qui che l’osservabilità di (A, C) e la raggiungibilità della coppia (A, \bar{G}) sono interpretabili come le condizioni di *minimalità stocastica* della realizzazione di un processo $\{\mathbf{y}(t)\}$ mediante un modello di stato del tipo (1.45). La *minimalità stocastica* è definita imponendo che $\{\mathbf{y}(t)\}$ sia rappresentato mediante il più piccolo numero di variabili di stato possibile, in altri termini imponendo che lo *spazio di stato* del modello

$$\mathbf{X}_t := \text{span} \{\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)\}$$

abbia la minima dimensione possibile [94]. Essa può venire a mancare anche quando si parta da modelli del tipo (2.1) che sono osservabili e raggiungibili, cioè minimi in senso deterministico.

PROBLEMA 10.3. Consideriamo il seguente modello, che è ovviamente osservabile e raggiungibile,

$$\begin{aligned} \mathbf{x}(t+1) &= -a\mathbf{x}(t) + (1-a^2)\mathbf{n}(t) \quad , \quad |a| < 1 \\ \mathbf{y}(t) &= \mathbf{x}(t) + a\mathbf{n}(t) \quad , \end{aligned} \quad (2.17)$$

per il quale si trova $F = -1/a$, $\tilde{Q} = 0$, e la corrispondente equazione di Riccati si scrive:

$$p(t+1) = p(t)/(p(t) + a^2) \quad .$$

Si dimostri che se $p(0) > 0$, $p(t)$ converge al valore di regime $\bar{p} = (1-a^2)$, che è l’unica soluzione globalmente stabile e semidefinita positiva della corrispondente equazione algebrica di Riccati.

Alla soluzione $\bar{p} = (1-a^2)$ corrisponde il guadagno a regime $\bar{G} = 0$. Il predittore a regime quindi degenera nell’equazione $\hat{\mathbf{x}}(t+1) = -a\mathbf{x}(t)$ la quale (visto che $|a| < 1$) è equivalente per $t \rightarrow \infty$ alla $\hat{\mathbf{x}}(t) = 0$. Ne segue che, in questo esempio, la realizzazione di innovazione (2.3) di $\{\mathbf{y}(t)\}$ si può ridurre alla relazione $\mathbf{y}(t) = \mathbf{e}(t)$.

Si spieghi questo risultato mettendolo in relazione col fatto che la funzione di trasferimento $W(z)$ del modello (2.17) è la funzione “passa-tutto”,

$$W(z) = \frac{1+az}{z+a} \quad ,$$

alla quale corrisponde quindi lo spettro di potenza $S(z) = 1$, tipico del rumore bianco. \diamond

Si può mostrare che l’esempio appena discusso è, in un certo senso, rappresentativo del caso generale. La minimalità stocastica di una realizzazione del processo $\{\mathbf{y}(t)\}$ del tipo (2.1) richiede che la relativa funzione di trasferimento $W(z)$ non abbia fattori (destri) di tipo *inner*, cioè passa-tutto analitici [94]. Se questa condizione è soddisfatta e la realizzazione è minima in senso deterministico, allora il predittore

di Kalman a regime associato alla realizzazione (2.1) è ancora raggiungibile e la (2.3) fornisce ancora una realizzazione *stocasticamente minima* di $\{\mathbf{y}(t)\}$.

3. Analisi del filtro a regime nel dominio della frequenza

L'enunciato del teorema 10.2 contiene in essenza una tecnica di analisi del filtro a regime che può essere di qualche utilità pratica, nel caso (abbastanza frequente nelle applicazioni) in cui i parametri statistici del rumore di modello $\{\mathbf{v}(t)\}$ sono scarsamente noti.

Per illustrare questa tecnica faremo l'ipotesi che il processo di misura $\{\mathbf{y}(t)\}$ sia *scalare* ($m = 1$) e che i *rumori di modello e di osservazione siano scorrelati*.

In queste condizioni il pseudo-spettro $S(z)$ di $\{\mathbf{y}(t)\}$ si scrive

$$S(z) = r + c(zI - A)^{-1} Q(z^{-1}I - A')^{-1} c' \quad , \quad (3.1)$$

dove $r > 0$ è la varianza del rumore di misura e c è il vettore riga che accoppia lo stato del modello alle misure.

Una situazione che si riesce ad analizzare agevolmente è quella in cui il processo di stato si descrive mediante un'equazione del tipo

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + b\mathbf{u}(t) \quad ,$$

dove $\{\mathbf{u}(t)\}$ è rumore bianco scalare di varianza q (non nota) e b è viceversa un vettore di accoppiamento fra il rumore e lo stato che si può considerare noto con esattezza. In questo caso quindi è nota la funzione di trasferimento che lega $\{\mathbf{u}(t)\}$ al "segnale messaggio" $c\mathbf{x}(t)$ (del quale $\mathbf{y}(t)$ rappresenta una versione rumorosa).

Dettala,

$$H(z) := c(zI - A)^{-1} b = N(z)/D(z) \quad ,$$

dove $D(z)$ è il polinomio caratteristico di A ed $N(z)$ ha la forma³ $N(z) = c \text{Adj}(zI - A)b$, la formula (3.1) si può riscrivere come

$$S(z) = r + q \frac{N(z)N(z^{-1})}{D(z)D(z^{-1})} \quad .$$

Introducendo il parametro $\rho := q/r$, l'equazione $S(z) = 0$ che determina in sostanza gli autovalori della matrice in catena chiusa del filtro, si può esplicitare nella forma

$$D(z)D(z^{-1}) + \rho N(z)N(z^{-1}) = 0 \quad , \quad \rho \geq 0 \quad . \quad (3.2)$$

Partendo da questa equazione si può analizzare, mediante la tecnica classica del *luogo delle radici*, la posizione degli autovalori di $\bar{\Gamma}$ nel piano complesso al variare del parametro non negativo ρ .

Notiamo che ρ può essere interpretato come un'indice di affidabilità delle misure rispetto al modello matematico del segnale. Se ρ è molto piccolo, $q \ll r$, e questo significa che il modello matematico del segnale che si vuole ricostruire è molto affidabile (c'è poco rumore di modello). In questo caso il filtro tenderà a tenere poco conto delle misure, assegnando valori piccoli al guadagno a regime \bar{K} . Così facendo si corregge di poco la stima "a priori" dello stato basata sul modello e contemporaneamente si introduce in modo molto attenuato il rumore di misura nelle stime (ricordare che la correzione dovuta alla retroazione è $\bar{K}\mathbf{e}(t) = \bar{K}(C\bar{\mathbf{x}}(t) + \mathbf{w}(t))$).

Dualmente, se ρ è molto grande, $q \gg r$, e il filtro tenderà a tenere poco conto del modello (che è scarsamente affidabile, dato che ha molto rumore sovrapposto).

³Qui $\text{Adj } A$ è l'aggiunta algebrica di A .

Il guadagno a regime sarà grande e le stime verranno aggiornate molto più in base ai dati di misura che alla dinamica del processo di stato. In questo caso si avrà un filtro a larga banda (dato che c'è poco rumore di misura da tagliare via) poco "accordato" sul segnale di stato.

Queste considerazioni sono ovviamente di natura del tutto qualitativa, ma trovano un certo riscontro nell'analisi dell'equazione (3.2).

Per $\rho \rightarrow 0$, ad esempio, si vede che la dinamica del filtro è all'incirca determinata dalle radici dell'equazione $D(z) D(z^{-1}) = 0$ che stanno in $|z| < 1$. Se il segnale messaggio è stazionario, i poli del filtro sono prossimi a quelli della funzione di trasferimento del segnale messaggio. Ciò è ovviamente in accordo col tendere a zero del guadagno a regime.

Per studiare le radici della (3.2) conviene supporre di conoscere $N(z)$ e $D(z)$ in forma fattorizzata

$$\begin{aligned} N(z) &= C z^\nu \prod_1^m (z - z_i) \quad , \quad z_i \neq 0 \\ D(z) &= z^\mu \prod_1^n (z - p_i) \quad , \quad p_i \neq 0 \quad , \end{aligned}$$

nel qual caso si può scrivere

$$\begin{aligned} N(z) N(z^{-1}) &= C \prod_1^m (z - z_i) (z^{-1} - z_i) = \bar{C}_1 z^{-m} \prod_1^m (z - z_i) (z - 1/z_i) \\ D(z) D(z^{-1}) &= \prod_1^n (z - p_i) (z^{-1} - p_i) = \bar{C}_2 z^{-n} \prod_1^n (z - p_i) (z - 1/p_i) \quad , \end{aligned}$$

dove \bar{C}_1 e \bar{C}_2 sono costanti opportune. Supponendo, tanto per fissare le idee, che $n \geq m$, si può riscrivere la (3.2) nella forma

$$\prod_1^n (z - p_i) (z - 1/p_i) + \rho \alpha z^{n-m} \prod_1^m (z - z_i) (z - 1/z_i) = 0 \quad , \quad (3.3)$$

dove $\alpha = \bar{C}_1/\bar{C}_2$. In questo caso il luogo delle radici associato alla (3.3) ha $2n$ rami che partono (per $\rho = 0$) dai poli $\{p_i, 1/p_i\}$ dello spettro del segnale messaggio. Per $\rho \rightarrow \infty$, $n - m$ di questi rami tendono all'origine, $n - m$ tendono all'infinito e i rimanenti $2m$ agli zeri $\{z_i, 1/z_i\}$. Si noti che il luogo deve essere composto di due parti "reciproche" rispetto alla circonferenza unita del piano complesso. La parte che descrive l'andamento degli autovalori di $\bar{\Gamma}$ è quindi quella interna al cerchio unita.

4. Risoluzione dell'Equazione Algebrica di Riccati

L'equazione algebrica di Riccati (A.R.E.) è un'equazione matriciale non lineare dall'aspetto abbastanza intimidatorio e sembra di primo acchito fuori discussione riuscire a dare delle formule esplicite per la sua soluzione. È possibile costruire esempi in cui l'A.R.E. ammette addirittura infinite soluzioni ed è perciò abbastanza stupefacente il fatto che si riesca, come vedremo in questo paragrafo, a dare in modo esplicito una formula molto semplice che permette di calcolare in maniera diretta una *qualsivoglia* soluzione dell'A.R.E. e in particolare quell'unica semidefinita positiva che interessa nelle applicazioni. La formula di cui parliamo (scoperta da A.J. McFarlane e J. Potter [99, 118] nel 1964) riduce la soluzione dell'A.R.E. a un problema di calcolo degli autovalori e autovettori per una matrice di struttura speciale di dimensione $(2n \times 2n)$ e all'inversione di una matrice $n \times n$ costruita

con una opportuna scelta di autovettori. Dato che sono disponibili oggi delle routines numeriche molto affidabili per risolvere problemi di autovalori-autovettori del tipo richiesto in queste applicazioni, il metodo di McFarlane-Potter (opportuna-mente riadattato) è diventato praticamente il metodo standard per la risoluzione numerica dell'A.R.E..

Prima di lanciarsi nella derivazione della formula è opportuno notare che per problemi di dimensioni modeste ($n \leq 6 \div 7$) esiste un algoritmo "naturale" da usare per la risoluzione dell'A.R.E. e questo algoritmo è proprio fornito dall'equazione alle differenze di Riccati, ad esempio la versione simmetrizzata vista nel capitolo 9. In effetti il teorema fondamentale 10.1, asserisce che (sotto le opportune ipotesi di stabilizzabilità e rivelabilità) questo algoritmo converge, a partire da una qualunque condizione iniziale, proprio verso la soluzione simmetrica e semidefinita dall'A.R.E.. Per problemi di dimensione non modesta questo algoritmo si rivela però poco affidabile. Il suo comportamento numerico peggiora tanto più quanto n è grande e quanto più gli autovalori della matrice in catena chiusa si avvicinano alla circonferenza unità. Questo è dovuto al fatto generale che l'equazione di Riccati può sostanzialmente essere vista come un metodo ricorsivo per risolvere le equazioni normali di un problema di minimi quadrati e, come è ben noto [135, 57], le equazioni normali possono in molte circostanze risultare mal condizionate.

È opportuno quindi tener sempre presente che l'equazione di Riccati, se pur implementata nella forma simmetrizzata illustrata nel capitolo precedente, non è un algoritmo molto affidabile per problemi di grandi dimensioni e può in qualche caso dare luogo a spiacevoli sorprese.

Per derivare la formula di McFarlane-Potter elaboreremo le considerazioni fatte nella sezione 2 relative al legame tra gli zeri dello pseudospettro $S(z)$ e la dinamica in catena chiusa del filtro.

Inizieremo col procurarci una realizzazione dell'inversa $S(z)^{-1}$ della matrice razionale $S(z)$. Partendo dallo schema a blocchi di figura 4.1,

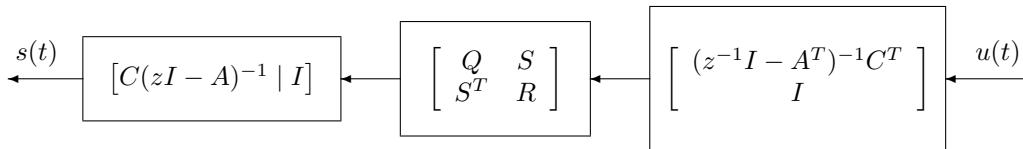


FIGURA 4.1. Schema a blocchi dell'inversa del pseudospettro $S(z)$.

in cui $u(t)$ ed $s(t)$ sono ingressi e uscite fittizie m -dimensionali, si trova facilmente una realizzazione della matrice di trasferimento $S(z)$ del tipo

$$\begin{aligned} \eta(t+1) &= A\eta(t) + Q\xi(t) + Su(t) \\ \xi(t-1) &= A'\xi(t) + C'u(t) \\ s(t) &= C\eta(t) + S'\xi(t) + Ru(t) \end{aligned} \quad (4.1)$$

Ricordando che R è invertibile e ricavando $u(t)$ in funzione dello stato $2n$ -dimensionale $[\eta(t)' \xi(t)']'$ e dell'uscita $s(t)$ dall'ultima equazione, si trova

$$\begin{aligned}\eta(t+1) &= F\eta(t) + \tilde{Q}\xi(t) + SR^{-1}s(t) \\ \xi(t-1) &= -C'R^{-1}C\eta(t) + F'\xi(t) + C'R^{-1}s(t) \\ u(t) &= -R^{-1}C\eta(t) - R^{-1}S'\xi(t) + R^{-1}s(t) \quad , \quad (4.2)\end{aligned}$$

dove F e \tilde{Q} sono definite dalle note formule (4.22) e (4.20). Queste equazioni forniscono una realizzazione di $S(z)^{-1}$. Notiamo che la seconda equazione di stato è scritta usando differenze all'indietro. Per studiare la dinamica di $S(z)^{-1}$ conviene avere un modello in cui compaiono solo differenze in avanti e occorre perciò eseguire una trasformazione di variabile. Allo scopo si trasla l'argomento temporale in avanti di un passo nelle due equazioni di stato e si pone $\tilde{\eta}(t) := \eta(t+1)$ ottenendo così

$$\begin{aligned}\tilde{\eta}(t+1) &= F\tilde{\eta}(t) + \tilde{Q}\xi(t+1) + SR^{-1}s(t+1) \\ \xi(t) &= -C'R^{-1}C\tilde{\eta}(t) + F'\xi(t+1) + C'R^{-1}s(t+1) \quad . \quad (4.3)\end{aligned}$$

La dinamica propria di $S(z)^{-1}$ si può dunque trascrivere mediante il sistema di equazioni alle differenze (nel quale abbiamo eliminato la "tilde" su $\eta(t)$ per semplicità di notazione)

$$\begin{bmatrix} F' & 0 \\ -\tilde{Q} & I \end{bmatrix} \begin{bmatrix} \xi(t+1) \\ \eta(t+1) \end{bmatrix} = \begin{bmatrix} I & C'R^{-1}C \\ 0 & F \end{bmatrix} \begin{bmatrix} \xi(t) \\ \eta(t) \end{bmatrix} \quad . \quad (4.4)$$

Poniamo

$$L = \begin{bmatrix} F' & 0 \\ -\tilde{Q} & I \end{bmatrix} \quad , \quad M = \begin{bmatrix} I & C'R^{-1}C \\ 0 & F \end{bmatrix} \quad (4.5)$$

e notiamo subito che queste due matrici sono invertibili se e solo se F è non singolare. Se F è non singolare il sistema (4.4) può essere allora riscritto nella forma usuale

$$\begin{bmatrix} \xi(t+1) \\ \eta(t+1) \end{bmatrix} = H \begin{bmatrix} \xi(t) \\ \eta(t) \end{bmatrix} \quad , \quad (4.6)$$

con

$$H = L^{-1}M = \begin{bmatrix} (F')^{-1} & (F')^{-1}C'R^{-1}C \\ Q(F')^{-1} & F + \tilde{Q}(F')^{-1}C'R^{-1}C \end{bmatrix} \quad . \quad (4.7)$$

Questa matrice ha una struttura molto particolare. Introducendo la matrice antisimmetrica $2n \times 2n$

$$J := \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} \quad , \quad (4.8)$$

che soddisfa alla

$$J' = J^{-1} = -J$$

e quindi $J^2 = -I_{2n}$ (notare che J si comporta come l'analogo matriciale dell'unità immaginaria), è facile verificare che H soddisfa la relazione $H'JH = J$ o, equivalentemente,

$$H' = JH^{-1}J' \quad . \quad (4.9)$$

Matrici (necessariamente di dimensione pari) per cui vale la (4.9) si chiamano *simplettiche*. Dalla (4.9) segue che gli autovalori di una matrice simplettica sono sempre a coppie *reciproche* (cioè λ e $1/\lambda$ devono contemporaneamente appartenere allo spettro) e questa proprietà è ovviamente in armonia col fatto che gli autovalori di H sono interpretabili come zeri dello pseudospettro $S(z)$. Il fatto più notevole è però il legame che esiste tra H e l'equazione algebrica di Riccati. Per illustrare questo

legame conviene prima introdurre alcune notazioni. Sia P una matrice quadrata $n \times n$ (che interpreteremo eventualmente come soluzione dell'A.R.E.). Poniamo

$$\Delta := C'R^{-1}C \quad , \quad \Omega := C'\Lambda^{-1}C \quad (4.10)$$

dove, come d'uso, $\Lambda := CPC' + R$. Una facile verifica permette di constatare che $(I + \Delta P)(I - \Omega P) = I$ e quindi

$$(I + \Delta P)^{-1} = (I - \Omega P) \quad . \quad (4.11)$$

Naturalmente la formula è valida solo *per matrici P per cui Λ è invertibile*. Consideriamo ora l'equazione di Riccati

$$P(t+1) = F P(t) [I - \Omega(t) P(t)] F' + \tilde{Q} \quad , \quad (4.12)$$

dove $\Omega(t)$ è definita dalla (4.10) con $P = P(t)$. Una successione di ovvi passaggi algebrici porge

$$\begin{aligned} FP[I - \Omega P]F' + \tilde{Q} &= FP(I + \Delta P)^{-1}F' + \tilde{Q} \\ &= FP(F^{-\top} + F^{-\top}\Delta P)^{-1} + Q(F^{-\top} + F^{-\top}\Delta P)(F^{-\top} + F^{-\top}\Delta P)^{-1} \\ &= [QF^{-\top} + (F + QF^{-\top}\Delta)P] [F^{-\top} + F^{-\top}\Delta P]^{-1} \end{aligned} \quad (4.13)$$

per cui l'equazione di Riccati (4.12) si riscrive nella forma

$$P(t+1) = [H_{21} + H_{22}P(t)] [H_{11} + H_{12}P(t)]^{-1} \quad , \quad (4.14)$$

dove H_{ij} , $i, j = 1, 2$, sono i blocchi $n \times n$ della matrice H che compaiono nella formula (4.7). Questo mostra che si può esprimere $P(t)$ nella forma

$$P(t) = Y(t) X(t)^{-1} \quad , \quad (4.15)$$

dove $X(t)$ e $Y(t)$ sono matrici quadrate $n \times n$ (mostreteremo tra un attimo che $X(t)$ deve essere invertibile!) che soddisfano il sistema di equazioni alle differenze *lineare*

$$\begin{aligned} X(t+1) &= H_{11} X(t) + H_{12} Y(t) \\ Y(t+1) &= H_{21} X(t) + H_{22} Y(t) \end{aligned} \quad , \quad (4.16)$$

ovvero

$$\begin{bmatrix} X(t+1) \\ Y(t+1) \end{bmatrix} = H \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} \quad , \quad (4.17)$$

al quale sono associate le condizioni iniziali

$$X(0) = X_0 \quad , \quad Y(0) = P_0 X_0 \quad , \quad (4.18)$$

con X_0 non singolare e $P_0 = P'_0 \geq 0$.

TEOREMA 10.3. *La matrice $X(t)$, soluzione del sistema lineare (4.16) con le condizioni iniziali (4.18), è non singolare per ogni $t \geq 0$ e, qualunque sia la condizione iniziale non singolare X_0 , il rapporto (4.15) fornisce la soluzione dell'equazione di Riccati (4.12) con condizione iniziale $P(0) = P_0 = P'_0 \geq 0$, per tutti $t \geq 0$.*

DIMOSTRAZIONE. Notiamo che la prima delle (4.16) si può equivalentemente riscrivere come

$$X(t+1) = F^{-\top} [X(t) + \Delta Y(t)] \quad . \quad (4.19)$$

Per $t = 0$, $Y(0)X_0^{-1} = P_0$ è simmetrica e semidefinita positiva per cui $\Lambda(0) = C'P_0C + R$ è non singolare e $\Omega(0)$ è ben definita. Ne segue che

$$X(1) = F^{-\top} (I + \Delta P_0) X_0$$

è invertibile e dalla rappresentazione (4.14) ottenuta in precedenza segue che effettivamente $P(1) = Y(1)X(1)^{-1}$. Procediamo allora per induzione. Sia $t > 1$ e assumiamo $X(t)$ invertibile e $P(t)$ data dalla formula (4.15). Dato che $P(t)$ risolve l'equazione di Riccati con condizione iniziale simmetrica e semidefinita positiva, $P(t) = P(t)' \geq 0$, per cui $\Lambda(t)$ è ancora invertibile e si può effettivamente formare $\Omega(t)$. Dalle (4.15) e (4.19) scende così che anche

$$X(t+1) = F^{-\top} [I + \Delta P(t)] X(t)$$

è non singolare. Segue allora dalla (4.14) che $P(t+1) = Y(t+1)X(t+1)^{-1}$. \square

Interpretazione Geometrica. Quanto abbiamo esposto si può interpretare usando un semplice modello geometrico bidimensionale. Consideriamo il sistema di equazioni alle differenze lineari in \mathbb{R}^2

$$\begin{aligned} x(t+1) &= h_{11}x(t) + h_{12}y(t) \\ y(t+1) &= h_{21}x(t) + h_{22}y(t) \end{aligned} \quad ,$$

e interpretiamo $p(t) := y(t)/x(t)$ come la pendenza della retta uscente dall'origine per il punto di coordinate $(x(t), y(t))$ del piano Cartesiano. Per quanto abbiamo appena visto $p(t)$ soddisfa un'equazione alle differenze non lineare del tipo di Riccati⁴. Naturalmente la pendenza è definita solo se $x(t) \neq 0$ ovvero la retta non è parallela all'asse delle y .

Per eliminare queste singolarità, conviene descrivere, invece che la pendenza p , la retta stessa. In geometria proiettiva le rette per l'origine si descrivono mediante *coordinante omogenee* che sono sostanzialmente coppie di numeri $[x, y]$, definiti a meno di un comune fattore moltiplicativo (che qui potremo chiamare x_0) non nullo. Con questa convenzione l'asse y si rappresenta ad esempio con le coordinate $[0, 1]$. Nel caso matriciale che ci interessa qui la "retta" $[X(t), Y(t)]$ può in realtà essere associata a soluzioni della R.E., $P(t)$, che non sono nemmeno semidefinite. \diamond

È naturale aspettarsi che le soluzioni di equilibrio dell'equazione di Riccati, cioè le soluzioni dell'A.R.E., si trovino mediante la stessa formula (4.15) ma questa volta applicata alle soluzioni di equilibrio del sistema lineare (4.16). Queste ultime si possono ovviamente identificare nei due blocchi \bar{X} , \bar{Y} di dimensione $n \times n$ in cui si partizionano le matrici \bar{U} , di dimensione $2n \times n$ che soddisfano la condizione di invarianza $\bar{U} = H\bar{U}$.

Notiamo anche che la soluzione $P = \bar{Y}\bar{X}^{-1}$ rimane invariata se si moltiplica a destra \bar{U} per un'arbitraria matrice nonsingolare di dimensione $n \times n$. In sostanza la soluzione dell'equazione di Riccati dipende solo dal *sottospazio di \mathbb{R}^{2n} generato dalle colonne di \bar{U}* . Si perviene così alla condizione di invarianza

$$H \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix} Z. \quad (4.20)$$

⁴Questa equazione è l'analogo discreto dell'equazione differenziale che il conte Jacopo Riccati trovò nel 1720 studiando la pendenza di una retta per un punto le cui coordinate $(x(t), y(t))$ soddisfacevano una coppia di equazioni differenziali lineari analoga al sistema discreto di cui stiamo occupando. Per una esauriente discussione di questo e innumerevoli altri fatti riguardanti l'equazione di Riccati si rimanda a [24].

dove Z è una matrice nonsingolare $n \times n$.

Naturalmente ci interessano qui le soluzioni in cui

- (1) X è non singolare,
- (2) YX^{-1} è simmetrica, il che si può imporre anche mediante la condizione di "ortogonalità",

$$\begin{bmatrix} X \\ Y \end{bmatrix}' J \begin{bmatrix} X \\ Y \end{bmatrix} = 0 \quad .$$

Un sottospazio di \mathbb{R}^{2n} generato dalle colonne di una matrice \bar{U} che soddisfa alle due condizioni precedenti, si chiama *Lagrangiano*. In altre parole, un sottospazio Lagrangiano può sempre pensarsi generato dai vettori colonna di una matrice \bar{U} , soluzione del sistema (4.20), che ha la forma $\begin{bmatrix} I \\ P \end{bmatrix}$, dove P è simmetrica di dimensione $n \times n$. Ovviamente uno spazio Lagrangiano ha sempre dimensione n .

Il problema della soluzione dell'ARE è così equivalente a quello della ricerca dei sottospazi Lagrangiani per la matrice H . A questo proposito notiamo che il problema (4.20) è un *problema di autovalori e autovettori per la matrice H* .

In effetti, se denotiamo con $A = [a_1, \dots, a_r]$ la matrice $2n \times r$ formata dagli r autovettori generalizzati (linearmente indipendenti) corrispondenti al generico autovalore λ_0 di H , si vede subito che vale una relazione del tipo

$$HA = A\hat{J} \quad , \quad (4.21)$$

dove $\hat{J} = \text{diag} \{ \hat{J}_1, \dots, \hat{J}_\nu \}$ è la matrice di Jordan corrispondente a λ_0 , formata dai ν miniblocchi ciascuno di dimensione $\ell(i) \times \ell(i)$

$$\hat{J}_i = \begin{bmatrix} \lambda_0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{bmatrix} \quad ,$$

con $\sum_{i=1}^{\nu} \ell(i) = r$. Come si vede, la ricerca dei sottospazi Lagrangiani H -invarianti si riduce alla ricerca di n opportuni autovalori (contati con la rispettiva molteplicità) di H e dei corrispondenti autovettori generalizzati. Come debbano essere scelti gli n autovalori perchè l'autospazio corrispondente sia (reale e) Lagrangiano è una questione che affronteremo tra poco in un contesto un poco più generale.

PROBLEMA 10.4. *Dimostrare che se $U := \begin{bmatrix} X \\ Y \end{bmatrix}$ genera un sottospazio invariante di dimensione n , la matrice Z in (4.20) è invertibile.*

Suggerimento: Dimostrare che Z è simile ad una matrice di Jordan, \hat{J} , associata ad n autovalori non necessariamente distinti di H . Questi autovalori debbono essere tutti diversi da zero perchè.... \diamond

Riassumiamo per il momento il risultato della discussione nel seguente importante enunciato.

TEOREMA 10.4. *Se F è non-singolare, tutte le soluzioni simmetriche (non necessariamente semidefinite positive) P dell'A.R.E. sono date dalla formula $P = YX^{-1}$, dove le colonne della matrice $U := \begin{bmatrix} X \\ Y \end{bmatrix}$ generano un sottospazio Lagrangiano invariante per H .*

Viceversa, ogni soluzione della (4.20) che genera un sottospazio Lagrangiano, rende $CPC' + R$ non singolare e definisce attraverso la $P := YX^{-1}$, una soluzione simmetrica dell'A.R.E.. C'è quindi una corrispondenza biunivoca tra le soluzioni simmetriche dell'A.R.E. e i sottospazi Lagrangiani invarianti per la matrice H .

DIMOSTRAZIONE. Che ogni soluzione dell'equazione algebrica di Riccati si possa esprimere come quoziente $P = YX^{-1}$ dove X e Y soddisfano la (4.20) con le condizioni 1 e 2, è già stato dimostrato nella discussione precedente.

Per mostrare l'implicazione inversa prendiamo una soluzione Lagrangiana (XY) della (4.20) e definiamo la matrice simmetrica $P = YX^{-1}$. Moltiplicando a destra la (4.20) per X^{-1} , si trova

$$\begin{aligned} [H_{11} + H_{12}P] &= XZX^{-1} \\ [H_{21} + H_{22}P] &= PZX^{-1} \end{aligned}$$

e sostituendo l'espressione di XZX^{-1} nella seconda equazione si ha,

$$P[H_{11} + H_{12}P] = [H_{21} + H_{22}P] \quad .$$

La verifica che P soddisfa l'A.R.E. si riduce quindi a mostrare che $XZX^{-1} = [H_{11} + H_{12}P] = F^{-\top} [I + \Delta P]$ è non singolare. Questo era l'oggetto del Problema (10.4) e possiamo qui dare questo fatto per scontato.

Che $\Lambda := CPC' + R$ sia non singolare scende poi immediatamente dalla (4.11). \square

PROBLEMA 10.5. *Dimostrare che P è una soluzione simmetrica (non necessariamente semidefinita positiva) dell'A.R.E. se e solo se soddisfa l'equazione*

$$[I \quad P] JH \begin{bmatrix} I \\ P \end{bmatrix} = 0 \quad , \quad (4.22)$$

dove J è la matrice simplettica elementare (4.8). \diamond

Come abbiamo visto, nell'ipotesi di *invertibilità della matrice F* , la ricerca delle soluzioni dell'A.R.E. è riconducibile al calcolo di n opportuni autovettori della matrice H .

L'ipotesi che F sia non singolare è però seriamente restrittiva. Per esempio, è facile verificare che per un generico modello ARMA scalare il polinomio caratteristico di F coincide con il numeratore $C(z)$ della relativa funzione di trasferimento (cfr. la (4.22)) e si vede così che se i coefficienti c_k, c_{k+1}, \dots, c_n sono nulli, F ha $n - k$ autovalori nulli. In particolare, per un modello AR *tutti* gli autovalori di F si trovano nell'origine.

Per ottenere un metodo di calcolo sufficientemente generale che non richieda l'invertibilità e il *calcolo* dell'inversa di F (che risulta normalmente molto impreciso se F è "quasi singolare") occorre fare una breve digressione preliminare sul *problema agli autovalori generalizzato*

$$\lambda La = Ma \quad (4.23)$$

in cui si cercano i numeri complessi λ a cui corrispondono soluzioni $a \neq 0$ dell'equazione lineare (4.23). Supporremo qui che il problema (4.23) sia *non singolare*, ovvero che il polinomio $\det(\lambda L - M)$ non si annulli identicamente. Gli *autovalori* (che più avanti verranno chiamati "finiti") del problema sono le radici dell'equazione caratteristica

$$\det(\lambda L - M) = 0 \quad . \quad (4.24)$$

A ciascuna soluzione λ_0 di molteplicità r dell'equazione caratteristica si possono associare, esattamente come nel caso usuale, delle opportune catene di autovettori generalizzati definiti dalla condizione

$$\begin{aligned} M a_1 &= \lambda_0 L a_1 \\ (M - \lambda_0 L) a_k &= L a_{k-1} \quad , \quad k = 2, 3, \dots, 1 \leq r \quad , \end{aligned} \quad (4.25)$$

ed è possibile mostrare che si possono sempre costruire ν catene di autovettori generalizzati indipendenti corrispondenti all'autovalore λ_0 in modo tale che la somma delle lunghezze $\ell(i)$, $i = 1, \dots, \nu$, di ciascuna catena è proprio uguale alla molteplicità r dell'autovalore.

Detta allora $A = [a_1, \dots, a_r]$ la matrice $2n \times r$ formata dagli r autovettori generalizzati (linearmente indipendenti) corrispondenti all'autovalore λ_0 , si ricava agevolmente dalla definizione (4.25) che

$$MA = LA \hat{J} \quad , \quad (4.26)$$

dove $\hat{J} = \text{diag} \{ \hat{J}_1, \dots, \hat{J}_\nu \}$ è la matrice di Jordan formata dai miniblocchi di dimensione $\ell(i) \times \ell(i)$ associati a λ_0 che hanno l'usuale struttura.

Come si vede, se L è non singolare, A è formata da autovettori (generalizzati) che generano l'autospazio corrispondente all'autovalore λ_0 della matrice $H = L^{-1}M$ e \hat{J} è allora il blocco di Jordan corrispondente. Le complicazioni che sorgono quando L è *singolare* sono connesse al fatto che il polinomio $\det(\lambda L - M)$ ha grado *strettamente minore* di $2n$ e sembrano quindi esserci meno di $2n$ autovalori (contati con la debita molteplicità) per il nostro problema. Per studiare questo fenomeno, conviene introdurre il problema "duale"

$$\mu M a = L a \quad , \quad a \neq 0 \quad , \quad (4.27)$$

i cui autovalori sono le radici dell'equazione

$$\det(\mu M - L) = 0 \quad .$$

Ragionando sulla definizione (4.27) si vede immediatamente che *se $\mu_0 \neq 0$ è un autovalore del problema duale (4.27) allora $\lambda_0 := 1/\mu_0$ è un autovalore del problema originario (4.23)* e, viceversa, se $\lambda_0 \neq 0$ è un autovalore del problema originario allora $\mu_0 := 1/\lambda_0$ è un autovalore del problema duale. Come si vede, il fatto che L sia singolare (e quindi $\det(\lambda L - M)$ abbia grado inferiore a $2n$) corrisponde proprio all'esistenza di *autovalori nulli* per il problema duale. È quindi naturale introdurre gli *autovalori infiniti* del problema (4.23) come quelli corrispondenti ai reciproci degli autovalori nulli del problema duale. Con questa convenzione ciascuno dei due problemi ha esattamente $2n$ autovalori. Si noti che gli autovettori corrispondenti ad autovalori infiniti per il problema (4.23) sono semplicemente una base per lo spazio nullo di L . In effetti se $La = 0$, con $a \neq 0$, allora $Ma = 0$ implica $\mu = 0$ dato che La ed Ma non possono annullarsi contemporaneamente essendo $\det(L - \mu M)$ non identicamente nullo.

La struttura simplettica del nostro problema permette di precisare ulteriormente queste considerazioni.

LEMMA 10.9. *Si consideri il problema agli autovalori generalizzato (4.23) in cui L ed M sono definite in (4.5). Se $\lambda_0 = 0$ è un autovalore di molteplicità r allora vi sono solo $2n - r$ autovalori finiti per il problema. Gli r autovalori mancanti*

sono autovalori infiniti (“reciproci di zero”). Gli autovalori finiti e non nulli sono a coppie reciproche e lo spettro si può separare nel modo seguente

$$\underbrace{0, \dots, 0}_r, \lambda_{r+1}, \dots, \lambda_n, \frac{1}{\lambda_n}, \dots, \frac{1}{\lambda_{r+1}}, \underbrace{\infty, \dots, \infty}_r, \quad (4.28)$$

dove i primi n autovalori $\{\lambda_k; k = 1, \dots, n\}$ possono essere presi a modulo minore o uguale a uno.

Per la dimostrazione rimandiamo a [112, 90], pp. 633–634. Possiamo in realtà dire esattamente chi sono gli autovalori (4.28).

LEMMA 10.10. *Si assumano (F, C) rivelabile ed $(F, \tilde{Q}^{1/2})$ stabilizzabile. Allora gli autovalori del problema (4.23) contati con le rispettive molteplicità sono quelli della matrice in catena chiusa $\bar{\Gamma}$ del predittore a regime e i rispettivi reciproci.*

DIMOSTRAZIONE. Denotiamo con simboli soprasssegnati (\bar{P} , $\bar{\Lambda}$ ecc.) le quantità relative al predittore a regime. Una dimostrazione intuitiva si può ottenere realizzando $S(z)$ espresso mediante la *fattorizzazione canonica* (2.5), la quale corrisponde evidentemente alla scelta

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} \bar{G} \bar{\Lambda} \bar{G}' & \bar{G} \bar{\Lambda} \\ \bar{\Lambda} \bar{G}' & \bar{\Lambda} \end{bmatrix} \quad (4.29)$$

nello schema a blocchi di Figura 12.3. Con Q, S, R date dalla (4.29) si ottiene

$$\begin{aligned} F &= A - \bar{G} C = \bar{\Gamma} \\ \tilde{Q} &= 0 \\ C' R^{-1} C &= C' \bar{\Lambda}^{-1} C := \bar{\Omega} \end{aligned}$$

e pertanto la dinamica propria di $S(z)^{-1}$ è anche descrivibile tramite le equazioni

$$\begin{bmatrix} \bar{\Gamma}' & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \xi(t+1) \\ \eta(t+1) \end{bmatrix} = \begin{bmatrix} I & C' \bar{\Lambda}^{-1} C \\ 0 & \bar{\Gamma} \end{bmatrix} \begin{bmatrix} \xi(t) \\ \eta(t) \end{bmatrix}, \quad (4.30)$$

i cui modi propri corrispondono agli autovalori del problema

$$\lambda \bar{L} a = \bar{M} a, \quad (4.31)$$

con

$$\bar{L} = \begin{bmatrix} \bar{\Gamma}' & 0 \\ 0 & I \end{bmatrix}, \quad \bar{M} = \begin{bmatrix} I & \bar{\Omega} \\ 0 & \bar{\Gamma} \end{bmatrix}. \quad (4.32)$$

Come si vede immediatamente

$$\det(\lambda \bar{L} - \bar{M}) = \det(\lambda \bar{\Gamma}' - I) \det(\lambda I - \bar{\Gamma}),$$

per cui gli autovalori del problema (4.31) sono proprio quelli di $\bar{\Gamma}$ e i reciproci. Naturalmente questo ragionamento non è a stretto rigore sufficiente per concludere che $\lambda L - M$ e $\lambda \bar{L} - \bar{M}$ hanno lo stesso spettro. (Occorrerebbe allo scopo un “teorema di isomorfismo” per realizzazioni della forma (4.2)). Facciamo allora vedere che esistono due matrici non singolari, U e T , di dimensione $2n \times 2n$ tali per cui

$$U(\lambda \bar{L} - \bar{M})T = \lambda L - M.$$

Com'è immediatamente verificabile (calcolando i determinanti) questa relazione di similitudine assicura l'uguaglianza degli spettri. (In particolare, se L è invertibile, essa equivale alla $T^{-1}HT = H$). Prendiamo allora

$$T = \begin{bmatrix} I & 0 \\ \bar{P} & I \end{bmatrix}, \quad U = \begin{bmatrix} (I + \Delta \bar{P})^{-1} & 0 \\ -F \bar{P} (I + \Delta \bar{P})^{-1} & I \end{bmatrix}$$

e notiamo che

$$LT = \begin{bmatrix} F' & 0 \\ \bar{P} - \tilde{Q} & I \end{bmatrix}, \quad MT = \begin{bmatrix} I + \Delta \bar{P} & \Delta \\ F \bar{P} & F \end{bmatrix}$$

dove, come al solito, $\Delta = C' R^{-1} C$.

Per verificare esplicitamente che vale l'equivalenza basta notare che

$$\begin{aligned} \bar{\Gamma} &= F(I - \bar{P}\bar{\Omega}) = F(I + P\Delta)^{-1} \\ \bar{\Gamma}' &= [(I + \bar{P}\Delta)']^{-1} F' = (I + \Delta\bar{P})^{-1} F' \\ \bar{P} - \tilde{Q} &= F\bar{P}(I + \Delta\bar{P})^{-1} F' \\ \bar{\Omega} &= (I + \Delta\bar{P})^{-1} \Delta, \end{aligned} \quad (4.33)$$

dove la terza equazione è semplicemente l'A.R.E. e l'ultima uguaglianza segue dall'identità

$$\begin{aligned} \Delta &= C' R^{-1} \bar{\Lambda} \bar{\Lambda}^{-1} C \\ &= C'(I + R^{-1} C \bar{P} C') (R + C' \bar{P} C)^{-1} C = (I + \Delta \bar{P}) C' \bar{\Lambda}^{-1} C. \end{aligned}$$

□

OSSERVAZIONE 10.2. Se si lasciano cadere le ipotesi di rivelabilità e stabilizzabilità, la dimostrazione continua a valere pur di prendere al posto di P una qualunque soluzione simmetrica, P , dell'A.R.E. e per Γ la corrispondente matrice $\Gamma = F(I + P\Delta)^{-1}$ (che ora non ha più ovviamente il significato di matrice in catena chiusa di un filtro a regime). Notiamo che se $P = P'$, la seconda relazione in (4.33) viene sostituita dalla $\Gamma' = (I + \Delta P)^{-1} F'$. L'enunciato del lemma si può così generalizzare dicendo che lo spettro di $\lambda L - M$ è dato dalla formula

$$\sigma[\lambda L - M] = \left\{ \lambda_k(\Gamma) \right\} \cup \left\{ 1/\lambda_k(\Gamma) \right\}, \quad (4.34)$$

dove $\{\lambda_k(\Gamma)\}$ sono gli autovalori della matrice in catena chiusa Γ corrispondente a una qualunque soluzione simmetrica dell'A.R.E.. \diamond

Invece di ricorrere alla matrice H , in questo contesto ci si riferisce direttamente al sistema di equazioni alle differenze matriciali, derivato dal sistema (4.4)

$$\begin{aligned} Y(t+1) &= F Y(t) + \tilde{Q} X(t+1) & Y(0) &= Y_0 \\ X(t) &= -\Delta Y(t) + F' X(t+1) & X(0) &= X_0 \end{aligned} \quad (4.35)$$

Prendendo condizioni iniziali Y_0 e X_0 (con X_0 non singolare) tali che $P_0 = Y_0 X_0^{-1}$ e supponendo che $X(t)$ si mantenga non singolare per ogni $t \geq 0$, è facile verificare che

$$P(t) := Y(t) X(t)^{-1} \quad (4.36)$$

risolve l'equazione di Riccati con condizione iniziale $P(0) = P_0$. Basta allo scopo usare la formula (4.11) e, dalla seconda equazione (4.35), ricavare che $X(t)$ evolve nel tempo secondo la

$$X(t) = (I + \Delta P(t))^{-1} F' X(t+1) = \Gamma'(t) X(t+1),$$

con un'ovvia interpretazione di $\Gamma(t)$. In questo caso però è molto più difficile dare condizioni per l'invertibilità di $X(t)$.

LEMMA 10.11. *Tutte le soluzioni dell'A.R.E. sono della forma $P = YX^{-1}$ dove $\begin{bmatrix} X \\ Y \end{bmatrix}$ è un insieme di n autovettori generalizzati corrispondenti ad autovalori finiti del problema (duale) $\mu Ma = La$.*

Viceversa, siano $\begin{bmatrix} X \\ Y \end{bmatrix}$ n vettori generalizzati del problema duale, corrispondenti ad autovalori finiti. Se X è invertibile e tale da rendere $\Lambda = R + CYX^{-1}C'$ non singolare allora $P = YX^{-1}$ risolve l'A.R.E..

DIMOSTRAZIONE. Sia P una qualunque soluzione dell'A.R.E. e Γ la corrispondente matrice "in catena chiusa" definita dalla relazione (cfr. la prima delle (4.33))

$$\Gamma = F(I + P\Delta)^{-1} \quad .$$

Definiamo $\tilde{\Gamma} := (I + \Delta P)^{-1} F'$ (ovviamente $\tilde{\Gamma} = \Gamma'$ se P è simmetrica). Esiste allora una matrice non singolare X che trasforma $\tilde{\Gamma}$ in forma canonica di Jordan, \hat{J} ,

$$\tilde{\Gamma} = X\hat{J}X^{-1} \quad .$$

Inserendo questa espressione nella $(I + \Delta P)\tilde{\Gamma} = F'$ si trova

$$X\hat{J}X^{-1} + \Delta PX\hat{J}X^{-1} = F' \quad .$$

Si ponga ora $Y := PX$, per cui $P = YX^{-1}$ e trasformiamo così questa equazione nella

$$X\hat{J} + \Delta Y\hat{J} = F'X \quad . \quad (4.37)$$

D'altro canto l'A.R.E. si può riscrivere come $FP\tilde{\Gamma} = -\tilde{Q} + P$ e pertanto

$$FY\hat{J} = -\tilde{Q}X + Y \quad . \quad (4.38)$$

Le (1.18) e (4.38) possono essere riscritte nella forma

$$\begin{bmatrix} I & \Delta \\ 0 & F \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \hat{J} = \begin{bmatrix} F' & 0 \\ -\tilde{Q} & I \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \quad (4.39)$$

con \hat{J} in forma di Jordan. Questa implica proprio che le colonne di $\begin{bmatrix} X \\ Y \end{bmatrix}$ sono autovettori (generalizzati) del problema duale con autovalori descritti da \hat{J} .

Per dimostrare la seconda parte del lemma basta partire dalla (4.39) (che vale per l'ipotesi di autovalori finiti) e ripercorrere all'indietro i calcoli. \square

Dalla dimostrazione si vede che se P è simmetrica la matrice di Jordan \hat{J} associata agli n autovettori $\begin{bmatrix} X \\ Y \end{bmatrix}$ è proprio la forma di Jordan della trasposta della matrice in catena chiusa Γ , corrispondente a P . Questa osservazione permette di stabilire rapidamente il seguente risultato fondamentale.

TEOREMA 10.5. *Si assumano (F, C) rivelabile e $(F, \tilde{Q}^{1/2})$ stabilizzabile. Allora esistono esattamente n autovalori (finiti) del problema $\mu Ma = La$, contati con la debita molteplicità, che hanno modulo strettamente minore di uno. Per ogni base n -dimensionale di vettori $\begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix}$ che generano il corrispondente autospazio, la matrice \bar{X} è non singolare e l'unica soluzione $\bar{P} = \bar{P}' \geq 0$ dell'A.R.E. è data dalla formula*

$$\bar{P} = \bar{Y}\bar{X}^{-1} \quad . \quad (4.40)$$

DIMOSTRAZIONE. La prima affermazione del teorema è una conseguenza immediata del lemma 10.10 una volta che si ricordi che gli autovalori di $\mu M - L$ sono i reciproci di quelli di $\lambda L - M$. In effetti gli autovalori a modulo minore di uno di $\mu M - L$ sono esattamente quelli di $\bar{\Gamma}$. Sia allora \bar{J} la forma canonica di Jordan di $\bar{\Gamma}'$. Detta \bar{X} la matrice $n \times n$ (non singolare!) che trasforma $\bar{\Gamma}'$ in forma di Jordan,

$$\bar{\Gamma}' = \bar{X} \bar{J} \bar{X}^{-1} \quad ,$$

ripercorrendo la dimostrazione del lemma 10.11 si arriva a esprimere la soluzione simmetrica e semidefinita positiva dell'A.R.E. come $\bar{P} = \bar{Y} \bar{X}^{-1}$, dove \bar{X} e \bar{Y} soddisfano l'equazione

$$M \begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} \bar{J} = L \begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} \quad .$$

Si vede quindi che gli n vettori colonna in $\begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix}$ generano l'autospazio di $\mu M - L$ corrispondente proprio agli autovalori asintoticamente stabili. D'altro canto, ogni base per questo sottospazio ha la forma

$$\begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} Z \quad ,$$

con Z $n \times n$ non singolare, per cui $U = \bar{X}Z$ è ancora invertibile e $\bar{P} = VU^{-1}$. \square

Su questo risultato si basano i metodi moderni di calcolo della soluzione simmetrica e semidefinita positiva dell'A.R.E.. Prima di discutere gli aspetti computazionali è forse utile cercare di capire un po' più a fondo la ragione dell'apparire del problema duale. La discussione che segue è organizzata come una successione di affermazioni la cui dimostrazione è generalmente molto semplice ed è lasciata per esercizio al lettore.

PROBLEMA 10.6. *Si assuma F non singolare. Mostrare, servendosi della proprietà simplettica di H , che la condizione (4.22) è equivalente alla*

$$[I \quad P] JH^{-1} \begin{bmatrix} I \\ P \end{bmatrix} = 0 \quad (4.41)$$

e quindi che le colonne di $\begin{bmatrix} I \\ P \end{bmatrix}$ generano un sottospazio di R^{2n} che è invariante anche per l'inversa, H^{-1} , di H . \diamond

PROBLEMA 10.7. *Sempre nel caso in cui F è non singolare, si mostri direttamente (oppure usando il lemma 10.11) che ogni soluzione dell'A.R.E. si esprime come $P = YX^{-1}$ dove $\begin{bmatrix} X \\ Y \end{bmatrix}$ sono n autovettori generalizzati di H^{-1} . Quest'ultimo fatto si può esprimere scrivendo*

$$M \begin{bmatrix} X \\ Y \end{bmatrix} \hat{J} = L \begin{bmatrix} X \\ Y \end{bmatrix} \quad , \quad (4.42)$$

ovvero

$$L \begin{bmatrix} \tilde{X} \\ \tilde{Y} \end{bmatrix} \hat{J}_1 = M \begin{bmatrix} \tilde{X} \\ \tilde{Y} \end{bmatrix} \quad , \quad (4.43)$$

dove $\hat{J}_1 = U^{-1} \hat{J}^{-1} U$ è la forma di Jordan di \hat{J}^{-1} e $\tilde{X} := XU$, $\tilde{Y} := YU$.

La soluzione simmetrica e semidefinita positiva dall'A.R.E. si può quindi ottenere calcolando gli n autovettori generalizzati di $\lambda L - M$ corrispondenti agli

autovalori di modulo maggiore di uno. Notare che se F è singolare, $\lambda L - M$ ha autovalori infiniti e la forma di Jordan (\bar{J}_1) corrispondente all'inversa di J non è definita. \diamond

Come si è visto, la risoluzione dell'A.R.E. è stata ridotta al problema di calcolare una qualunque base di vettori per l'autospazio asintoticamente stabile (cioè corrispondente agli autovalori a modulo minore di uno) del polinomio matriciale $\lambda M - L$. In pratica non conviene mai prendere una base di autovettori, sia perché si ha in genere a che fare con quantità complesse, sia perché il calcolo degli autovettori può essere un problema numericamente mal condizionato specie in presenza di autovalori multipli. Conviene usare, al posto della forma di Jordan, la cosiddetta *forma (reale) di Schur* [134, 58], [90].

TEOREMA 10.6 (Forma reale di Schur di $\lambda M - L$). *Esistono due matrici ortogonali reali, Q e Z , di dimensione $2n \times 2n$ che rendono simultaneamente*

- QMZ triangolare superiormente,
- QLZ quasi triangolare superiormente (cioè triangolare superiormente eccettuata al più la presenza di blocchi diagonali 2×2 corrispondenti a coppie di autovalori complessi coniugati).

Se è nota la forma di Schur di $\lambda M - L$,

$$\lambda \hat{M} - \hat{L} := Q(\lambda M - L)Z \quad ,$$

allora gli autovalori (eventualmente infiniti) di $\lambda M - L$ si ottengono calcolando i rapporti dei corrispondenti blocchi diagonali, cioè

$$\lambda_k = \hat{\ell}_{kk} / \hat{m}_{kk}$$

per i blocchi diagonali in \hat{L} di dimensione 1 e risolvendo un problema 2×2 del tipo

$$\det \left(\lambda \begin{bmatrix} \hat{m}_{11} & \hat{m}_{12} \\ 0 & \hat{m}_{22} \end{bmatrix} - \begin{bmatrix} \hat{\ell}_{11} & \hat{\ell}_{12} \\ \hat{\ell}_{21} & \hat{\ell}_{22} \end{bmatrix} \right) = 0$$

in corrispondenza di blocchi diagonali in \hat{L} di dimensione 2.

Si dimostra che è sempre possibile scegliere Q e Z in modo tale che *il blocco superiore sinistro di dimensione $n \times n$ in M ed L contenga tutti e soli gli n autovalori a modulo strettamente minore di uno.*

In questo modo la forma di Schur si può scrivere

$$\lambda \begin{bmatrix} \hat{M}_{11} & \hat{M}_{12} \\ 0 & \hat{M}_{22} \end{bmatrix} - \begin{bmatrix} \hat{L}_{11} & \hat{L}_{12} \\ 0 & \hat{L}_{22} \end{bmatrix} \quad ,$$

dove \hat{M}_{11} , \hat{M}_{22} sono triangolari superiormente, \hat{L}_{11} , \hat{L}_{22} quasi-triangolari superiormente e gli autovalori asintoticamente stabili di $\lambda L - M$ sono esattamente quelli di $\lambda \hat{M}_{11} - \hat{L}_{11}$. Dato che gli autovalori stabili sono tutti *finiti*, si vede che \hat{M}_{11} deve essere *invertibile* (ovvero $\hat{m}_{kk} \neq 0$ per $k = 1, \dots, n$) e quindi la matrice $n \times n$

$$\hat{S} = \hat{M}_{11}^{-1} \hat{L}_{11}$$

è reale, quasi-triangolare superiormente e ha tutti i suoi autovalori a modulo strettamente minore di uno. Notiamo ora che l'ovvia identità

$$\begin{bmatrix} \hat{M}_{11} \\ 0 \end{bmatrix} \hat{S} = \begin{bmatrix} \hat{L}_{11} \\ 0 \end{bmatrix}$$

si può riscrivere nella forma

$$QMZ \begin{bmatrix} I \\ 0 \end{bmatrix} \hat{S} = QLZ \begin{bmatrix} I \\ 0 \end{bmatrix} ,$$

che equivale alla

$$M \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \hat{S} = L \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} , \quad (4.44)$$

dove abbiamo partizionato il blocco delle prime n colonne di Z ponendo

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = Z \begin{bmatrix} I \\ 0 \end{bmatrix} .$$

La (4.44) dice che le prime n colonne di Z sono una base per l'autospazio asintoticamente stabile di $\lambda M - L$ (la base è ortonormale giacché $Z'Z = I_{2n}$) e pertanto

$$\bar{P} = Z_2 Z_1^{-1} .$$

Maggiori dettagli su questo algoritmo (che è disponibile oggi in molte librerie standard) si possono trovare negli articoli [9, 90] e [140].

L'algoritmo a iterazione esponenziale ("Doubling Algorithm"). La teoria che abbiamo sviluppato all'inizio di questo paragrafo e in particolare il Teorema 10.3 permettono di giustificare in poche righe un algoritmo iterativo per il calcolo della soluzione (simmetrica e semidefinita positiva) dell'A.R.E. che ha proprietà molto interessanti (alcune delle quali in realtà non sembrano essere state studiate e analizzate abbastanza a fondo). L'idea è quella di calcolare \bar{P} iterando, anziché l'equazione alle differenze di Riccati, lo schema lineare (4.17). Questa iterazione si può fare, come vedremo, in modo "accelerato" in maniera da produrre un algoritmo con eccezionale rapidità di convergenza.

Notiamo intanto che se H è una matrice simplettica la potenza H^t è ancora simplettica, qualunque sia il naturale t . Inoltre la (4.9) implica una relazione tra i blocchi $n \times n$ di H . Se

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} ,$$

con $H_{ij} \in R^{n \times n}$ e H_{11} invertibile, è simplettica, allora

$$H_{22} = (H'_{11})^{-1} + H_{21} H_{11}^{-1} H_{12} . \quad (4.45)$$

Consideriamo ora la successione

$$H_0 = H , H_1 = H_0^2 , \dots , H_{t+1} = (H_t)^2 , \dots \quad (4.46)$$

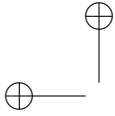
il cui termine t -simo è manifestamente uguale a H^{2^t} .

PROPOSIZIONE 10.6. *Si assuma F invertibile e $H_0 = H$ uguale alla matrice simplettica (4.7). Si può allora esprimere H_t nella forma*

$$H_t = \begin{bmatrix} A_t^{-1} & A_t^{-1} \Delta_t \\ Q_t A_t^{-1} & A_t' + Q_t A_t^{-1} \Delta_t \end{bmatrix} , \quad (4.47)$$

dove le matrici $n \times n$ A_t , Δ_t , Q_t si aggiornano mediante lo schema ricorsivo

$$\begin{aligned} A_{t+1} &= A_t (I + \Delta_t Q_t)^{-1} A_t \\ \Delta_{t+1} &= \Delta_t + A_t (I + \Delta_t Q_t)^{-1} \Delta_t A_t' \\ Q_{t+1} &= Q_t + A_t' Q_t (I + \Delta_t Q_t)^{-1} A_t , \end{aligned} \quad (4.48)$$



con le condizioni iniziali

$$A_0 = F' \quad , \quad \Delta_0 = C'R^{-1}C \quad , \quad Q_0 = \tilde{Q} \quad . \quad (4.49)$$

DIMOSTRAZIONE. Basta scrivere per esteso i blocchi $n \times n$ di

$$H_{t+1} = H_t \cdot H_t \quad ,$$

tenendo conto che $(H_{t+1})_{22}$ è dato dalla formula (4.45). Un facile ragionamento induttivo mostra che $Q'_t = Q_t \geq 0$ e quindi $(I + \Delta_t Q_t)$ è invertibile per ogni $t \geq 0$, grazie alla formula (4.11). \square

Notiamo allora che la soluzione del sistema (4.17) negli istanti $0, 2, 2^2 \dots 2^t, \dots$, si esprime tramite la

$$\begin{bmatrix} X(2^t) \\ Y(2^t) \end{bmatrix} = H_t \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} \quad , \quad (4.50)$$

dove H_t è data dalle (4.47), (4.48), (4.49). Se si ha rivelabilità e stabilizzabilità, detta come al solito $P_0(t)$ la soluzione dell'equazione di Riccati corrispondente alla condizione iniziale $P_0 = 0$, si ha

$$\lim_{t \rightarrow \infty} P_0(2^t) = \lim_{t \rightarrow \infty} Y(2^t) X(2^t)^{-1} = \bar{P} \quad ,$$

dove $X(2^t)Y(2^t)$ sono le soluzioni (4.50) corrispondenti alle condizioni iniziali $X(0) = I, Y(0) = 0$. Calcolando il rapporto si scopre così che

COROLLARIO 10.3 (“Doubling Algorithm”). *Se (F, C) è rivelabile ed $(F, \tilde{Q}^{1/2})$ è stabilizzabile la matrice simmetrica Q_t , soluzione dello schema ricorsivo (4.48) con le condizioni iniziali (4.49), converge per $t \rightarrow \infty$ alla soluzione simmetrica e semidefinita positiva dell'A.R.E., ovvero*

$$\lim_{t \rightarrow \infty} Q_t = \bar{P} \quad .$$

Notiamo che l'inversione della matrice $n \times n$ $(I + \Delta_t Q_t)$ può essere ridotta a quella di una matrice $m \times m$ simmetrica e definita positiva giacché, per la (4.11),

$$(I + \Delta_t Q_t)^{-1} = (I - \Omega_t Q_t) \quad ,$$

dove

$$\Omega_t = C'(R + CQ_t C')^{-1} C \quad .$$

L'ammontare di calcoli per iterazione richiesto dall'algoritmo (4.48) è pertanto dello stesso ordine di grandezze di quello richiesto dall'equazione di Riccati. La velocità di convergenza è però enormemente superiore.

Un aspetto curioso dello schema (4.48), (4.49) è che esso *non sembra affatto richiedere l'invertibilità della matrice F* (e in pratica sembra funzionare correttamente anche con F singolare!).

Esercizi

10-1 Realizzare in variabili di stato il processo stazionario descritto dal modello AR

$$\mathbf{y}(t) + \sum_{k=1}^n a_k \mathbf{y}(t-k) = \mathbf{w}(t)$$

e verificare che la matrice F è nilpotente (esiste un intero k tale per cui $F^k = 0$). Vale il teorema fondamentale sul comportamento asintotico del filtro di Kalman?

Scrivere l'equazione alle differenze di Riccati per la varianza dell'errore di predizione e dimostrare che dopo un numero finito di passi (quanti?) si ha $P(t) = 0$!

Dare una spiegazione intuitiva di questo fatto.

10-2 Sia φ una variabile aleatoria a distribuzione uniforme in $[-\pi, \pi]$. Mostrare che il processo

$$\mathbf{s}(t) := A \sin(\omega_0 t + \varphi), \quad t \in \mathbb{Z}$$

è debolmente stazionario, puramente deterministico e calcolare la sua funzione di covarianza $\sigma_s(\tau)$ e il relativo spettro di potenza.

Si vuole calcolare la stima filtrata $\hat{\mathbf{s}}(t | t)$ a partire da misure rumorose

$$\mathbf{y}(t) = \mathbf{s}(t) + \mathbf{w}(t)$$

disponibili in $[t_0, t]$ dove \mathbf{w} è rumore bianco di varianza σ^2 scorrelato da \mathbf{s} .

Si realizzi il processo $\{\mathbf{s}(t)\}$ mediante un opportuno modello di stato minimo e si formuli il problema usando la teoria del filtro di Kalman. Si verifichi se valgono le condizioni del teorema asintotico fondamentale del filtro di Kalman.

Si scriva l'equazione algebrica di Riccati per la varianza d'errore di predizione dello stato e si individui una eventuale soluzione che ha significato di varianza d'errore a regime per il problema in esame.

Studiare il comportamento a regime del filtro per $t_0 \rightarrow -\infty$ e interpretare il risultato.

10-3 Si vuole calcolare il predittore di Kalman $\hat{\mathbf{y}}(t+1 | t)$ di un processo stazionario $\{\mathbf{y}(t)\}$ descritto dal modello MA

$$\mathbf{y}(t) = (1 - z^{-1})\mathbf{w}(t)$$

dove \mathbf{w} è bianco di varianza σ^2 .

Si realizzi il processo mediante un opportuno modello di stato minimo e si verifichi se valgono le condizioni del teorema asintotico fondamentale del filtro di Kalman.

Si scriva l'equazione alle differenze di Riccati per la varianza d'errore di stato e la si risolva esplicitamente (si usi allo scopo la trasformazione $\pi(t) := 1/p(t)$ supponendo $p(0) > 0$). Si mostri direttamente che questa equazione ha un punto di equilibrio $\bar{p} \geq 0$ globalmente stabile.

Trovare la funzione di trasferimento del predittore a regime corrispondente a questa soluzione. Secondo voi ha senso usare questa espressione in pratica?

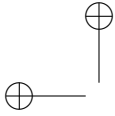
10-4 Realizzare il processo stazionario $\{\mathbf{y}(t)\}$ descritto dal modello MA

$$\mathbf{y}(t) = (1 - 7/6z^{-1} + 2/3z^{-2})\mathbf{w}(t)$$

dove \mathbf{w} è bianco di varianza σ^2 , in forma canonica di osservazione e discutere il comportamento della matrice varianza dell'errore di predizione dello stato P , per $t \rightarrow \infty$. Dire se esiste una varianza d'errore a regime e come si può calcolare.

10-5 Dare un metodo (e riportare le formule relative) per il calcolo del rapporto "disturbo-segnale" per la stima filtrata $\hat{\mathbf{z}}(t | t)$ a regime,

$$\rho := \frac{\text{var}[\mathbf{z}(t) - \hat{\mathbf{z}}(t | t)]}{\text{var} \mathbf{z}(t)}$$



per un segnale scalare stazionario a media nulla $\{\mathbf{z}(t)\}$, descritto da un modello lineare

$$\begin{aligned}\mathbf{x}(t+1) &= A\mathbf{x}(t) + \mathbf{v}(t) \\ \mathbf{z}(t) &= C\mathbf{x}(t).\end{aligned}$$

In cui $\{\mathbf{v}(t)\}$ è bianco di varianza Q . Il segnale è osservato in presenza di rumore additivo bianco $\{\mathbf{w}(t)\}$ di varianza R , scorrelato da $\{\mathbf{v}(t)\}$.

Svolgere i calcoli per

$$A^2 = 0.8, \quad Q = 0.2R, \quad C = 1.$$

10-6 Un segnale (non stazionario) $\{\mathbf{s}(t)\}$ descritto dal modello

$$(1 - z^{-1})\mathbf{s}(t) = \mathbf{v}(t)$$

dove $\{\mathbf{v}(t)\}$ è rumore bianco di varianza σ_1^2 , viene osservato per $t \geq 0$, con sovrapposto un altro rumore bianco scorrelato $\{\mathbf{w}(t)\}$ di varianza σ_2^2

$$\mathbf{y}(t) = \mathbf{s}(t) + \mathbf{w}(t).$$

Studiare il comportamento a regime del predittore di Kalman $\{\hat{\mathbf{s}}(t|t-1)\}$. Cosa accade a regime se $\sigma_1 \rightarrow 0$?

10-7 Si consideri il modello ingresso-uscita stazionario,

$$\begin{aligned}\mathbf{y}(t) &= \frac{(1 + 3/2z^{-1})}{(1 + 1/2z^{-1})} \mathbf{u}(t) + \mathbf{w}_1(t) \\ \mathbf{u}(t) &= \frac{(1 + 1/3z^{-1})}{(1 + 2/3z^{-1})} \mathbf{w}_2(t) \quad (*)\end{aligned}$$

dove \mathbf{w}_1 e \mathbf{w}_2 sono rumori bianchi scorrelati di varianza σ^2 .

A: Calcolare il predittore di Wiener

$$\hat{\mathbf{y}}(t+1|t) := \hat{E}[\mathbf{y}(t+1) | \mathbf{y}^t, \mathbf{u}^t]$$

basato sulla storia passata congiunta di \mathbf{y} e \mathbf{u} .

B: Trovare il modello ARMA d'innovazione per il processo \mathbf{y} e calcolare il predittore di Wiener basato sulla storia passata del solo processo \mathbf{y} . Confrontare le varianze dell'errore di predizione nei due casi A e B.

C: Introdurre una opportuna descrizione in variabili di stato del modello (*) e fornire uno schema ricorsivo per il calcolo del predittore di Kalman

$$\hat{\mathbf{y}}(t+1|t) := \hat{E}[\mathbf{y}(t+1) | \mathbf{y}^t, \mathbf{u}^t] \quad t \geq 0$$

basato sulla storia passata *congiunta* di \mathbf{y} e \mathbf{u} in $t \geq 0$. Discutere il comportamento a regime del predittore e confrontare con quanto trovato ai punti precedenti.

10-8 L'algoritmo di Bauer (Teorema 4.3 Cap. V) nel caso scalare, risolve il problema di calcolare il polinomio $\hat{C}(z)$ a zeri in $\{|z| < 1\}$, spettralmente equivalente ad un assegnato polinomio (supposto monico) $C(z)$, i.e.

$$\lambda^2 \hat{C}^*(z^{-1}) \hat{C}^*(z) = C^*(z^{-1}) C^*(z)$$

Usando una rappresentazione di stato per il processo stazionario

$$\mathbf{y}(t) = C^*(z^{-1}) \mathbf{w}(t)$$

dove \mathbf{w} è bianco di varianza unitaria, si dimostri, facendo uso della teoria del predittore di Kalman a regime, che l'algoritmo di Bauer calcola la rappresentazione d'innovazione di \mathbf{y} , i.e.

$$\mathbf{y}(t) = \hat{C}^*(z^{-1}) \mathbf{e}(t)$$

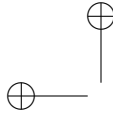
dove \mathbf{e} è il processo di innovazione (stazionaria) di \mathbf{y} .

10-9 Si consideri un segnale aleatorio $\{\mathbf{y}(t)\}$, che è l'uscita di un filtro lineare stabile con ingresso un rumore bianco

$$\mathbf{x}(t+1) = A \mathbf{x}(t) + B \mathbf{n}(t)$$

$$\mathbf{y}(t) = C \mathbf{x}(t) + D \mathbf{n}(t)$$

dove $\{\mathbf{n}(t)\}$ è un rumore bianco normalizzato. Riscrivendo il modello nell'usuale forma a rumori scorrelati ($R > 0$), si suppongano $(F, \tilde{Q}^{1/2})$ stabilizzabile e (C, F) rivelabile. Usando il teorema 10.2 che lega la fattorizzazione a fase minima del pseudospettro, diciamolo $S(z)$, di $\mathbf{y}(t)$, alla soluzione semidefinita positiva dell'A-RE, mostrare che $S(z)$ è *strettamente definito positivo* sulla circonferenza unitaria, i.e. $S(e^{j\omega}) > 0$.



CAPITOLO 11

IL FILTRO DI KALMAN MESSO IN PRATICA

In questo capitolo ci occuperemo di alcune questioni che capita frequentemente di dover affrontare quando si vogliono utilizzare gli algoritmi di stima visti nei due capitoli precedenti in problemi pratici. Tratteremo dell'impiego della teoria in problemi descritti da modelli non lineari (il cosiddetto filtro di Kalman “esteso”) e del problema dell’“accordo” del filtro, cioè dell’aggiustamento sperimentale della varianza del rumore di modello nel caso (frequente) in cui questa non sia nota. Discuteremo anche alcune classi di applicazioni particolarmente significative.

1. Il Filtro di Kalman Esteso

Nelle applicazioni a problemi pratici, la fase più importante per la risoluzione di un problema di stima è quella della *modellizzazione*. Questa è una fase per cui purtroppo non si possono dare ricette generali e conta in massimo grado l’esperienza e l’intuizione del progettista. Ogni campo specifico di applicazione (elaborazione di segnali per comunicazioni, radar, sonar, inseguimento di bersagli ecc.) ha la sua letteratura specifica e i relativi trucchi del mestiere. Si può comunque affermare che sono abbastanza rari i casi concreti che si lasciano descrivere in modo soddisfacente per mezzo di un modello dinamico lineare come quello trattato al Cap. X. Ci si chiede allora cosa fare in queste situazioni. A questo riguardo è importante ribattere ancora l’ovvia affermazione che un problema di stima è la formulazione matematica di un problema di *misura* in cui conta soprattutto la *precisione* dei risultati. Per questa ragione, in questo tipo di problemi la pratica della *linearizzazione*, che si presenta spontaneamente come la soluzione più ovvia per superare la difficoltà, va usata con molto giudizio. Linearizzare in modo ingenuo un modello può significare produrre degli errori di predizione irrealisticamente grandi perchè le misure sono descritte troppo approssimativamente. In questo modo, si trattano in sostanza i dati come se fossero acquisiti da uno strumento (lineare) di classe scadente anziché da quello (non lineare) originale che potrebbe essere invece di classe elevata.

Il *filtro di Kalman esteso* (E.K.F.) è un algoritmo di impiego molto generale che risolve in modo approssimato il problema di stima dello stato di un modello non lineare. Dato che è praticamente impossibile risolvere esattamente in termini di algoritmi a memoria finita un problema di filtraggio non lineare [68, 95] la tecnica che (inevitabilmente) si deve usare per derivare anche le equazioni dell’E.K.F. è quella della linearizzazione. Si tratta però di una linearizzazione “intelligente”. Come vedremo, il principio fondamentale su cui è basato l’E.K.F. è di linearizzare ad ogni istante di campionamento le equazioni dinamiche *attorno alla miglior stima di $\mathbf{x}(t)$ disponibile a quel momento*. Questo principio permette in linea generale di mantenere gli errori di linearizzazione a un livello accettabile (se le non linearità sono sufficientemente regolari il “rumore di linearizzazione” è dello stesso ordine di

grandezza dei rumori di modello e di misura) ed è la chiave per spiegare il buon comportamento di questo algoritmo in molti casi pratici.

Per trattare un caso di interesse generale ci metteremo nella situazione (che è forse la più comune in pratica) in cui le misure sono ottenute per campionamento di un segnale a tempo continuo. In questo paragrafo la variabile t designerà quindi la variabile temporale continua ($t \in \mathbb{R}$). Il modello cui faremo riferimento è il seguente

$$\begin{aligned}\dot{\mathbf{x}}(t) &= f(\mathbf{x}(t)) + \boldsymbol{\xi}(t) \\ \mathbf{y}(t_k) &= h(\mathbf{x}(t_k)) + \mathbf{w}(t_k)\end{aligned}\quad (1.1)$$

dove $\{\boldsymbol{\xi}(t)\}$ è un rumore bianco a tempo continuo n -dimensionale di media nulla e intensità $Q = Q' \geq 0$; in notazione simbolica,

$$E \boldsymbol{\xi}(t) \boldsymbol{\xi}(s)' = Q \delta(t - s) \quad . \quad (1.2)$$

Il processo rumore di misura $\{\mathbf{w}(t_k) ; k = 0, 1, \dots\}$ è una successione bianca m -dimensionale di varianza R strettamente definita positiva. Supporremo $\{\boldsymbol{\xi}(t)\}$ e $\{\mathbf{w}(t_k)\}$ scorrelati fra loro e dallo stato iniziale $\mathbf{x}(t_0)$.

Le funzioni f e h verranno supposte derivabili ovunque e con derivata continua. Per semplificare le notazioni, f , h , Q ed R sono state ipotizzate invarianti nel tempo, ma quanto diremo è immediatamente generalizzabile a una dinamica tempo-variante e a funzioni $Q = Q(t)$ ed $R = R(t_k)$ variabili in modo arbitrario. In particolare R e Q potrebbero essere funzioni dello stato corrente del sistema. Per coprire alcune applicazioni interessanti supporremo poi che gli istanti di campionamento t_k , $k = 0, 1, \dots$, nei quali viene “letta” la misura k -sima $\mathbf{y}(t_k)$, possano anche non essere necessariamente equispaziati (campionamento aperiodico).

Sia ora $\bar{\mathbf{x}}(t)$ una traiettoria di stato di riferimento (che verrà poi cambiata ad ogni istante di campionamento) soddisfacente l'equazione differenziale

$$\dot{\bar{\mathbf{x}}}(t) = f(\bar{\mathbf{x}}(t)) \quad . \quad (1.3)$$

Definiamo le matrici Jacobiane

$$F(\bar{\mathbf{x}}(t)) = \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\mathbf{x}=\bar{\mathbf{x}}(t)} \quad (1.4)$$

$$H(\bar{\mathbf{x}}(t)) = \left. \frac{\partial h}{\partial \mathbf{x}} \right|_{\mathbf{x}=\bar{\mathbf{x}}(t)} \quad (1.5)$$

e introduciamo gli scostamenti

$$\begin{aligned}\delta_{\mathbf{x}}(t) &:= \mathbf{x}(t) - \bar{\mathbf{x}}(t) \\ \delta_{\mathbf{y}}(t) &:= \mathbf{y}(t) - \bar{\mathbf{y}}(t)\end{aligned}$$

dove $\bar{\mathbf{y}}(t) = h(\bar{\mathbf{x}}(t))$. Per piccoli scostamenti potremo scrivere

$$\dot{\delta}_{\mathbf{x}}(t) = F(\bar{\mathbf{x}}(t)) \delta_{\mathbf{x}}(t) + \boldsymbol{\xi}(t) \quad (1.6)$$

e discretizzando (trattando $\bar{\mathbf{x}}(t)$ come una funzione *nota* del tempo) si ottiene da questa equazione lineare la

$$\delta_{\mathbf{x}}(t_{k+1}) = \Phi(t_{k+1}, t_k, \bar{\mathbf{x}}) \delta_{\mathbf{x}}(t_k) + \mathbf{v}(t_k) \quad (1.7)$$

dove $\Phi(t, s, \bar{\mathbf{x}})$ è la matrice fondamentale associata al sistema lineare $\dot{z} = F(\bar{\mathbf{x}}(t))z$ e $\{\mathbf{v}(t_k)\}$ è una successione di vettori aleatori scorrelati (rumore bianco in generale

non stazionario) definiti dalla

$$\mathbf{v}(t_k) = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, s) \boldsymbol{\xi}(s) ds \quad . \quad (1.8)$$

La matrice varianza del processo $\{\mathbf{v}(t_k)\}$ è calcolabile tramite la formula

$$Q(t_k) = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, s) Q \Phi(t_{k+1}, s)' ds \quad . \quad (1.9)$$

Passo di predizione

All'istante t_k , dopo aver elaborato la k -sima misura, la migliore stima dello stato disponibile è $\hat{\mathbf{x}}(t_k | t_k)$. Si sceglie come traiettoria di riferimento $\bar{\mathbf{x}}$ in (t_k, t_{k+1}) la soluzione di

$$\dot{\bar{\mathbf{x}}}(t) = f(\bar{\mathbf{x}}(t)) \quad , \quad \bar{\mathbf{x}}(t_k) = \hat{\mathbf{x}}(t_k | t_k) \quad . \quad (1.10)$$

Da notare che con questa scelta si ha $\hat{\delta}_{\mathbf{x}}(t_k | t_k) := \hat{E}[\delta_{\mathbf{x}}(t_k) | \mathbf{y}^{t_k}] = 0$, per cui, proiettando ambo i membri del modello di stato lineare (1.7) su $\mathbf{H}(\mathbf{y}^{t_k})$, si ottiene

$$\hat{\delta}_{\mathbf{x}}(t_{k+1} | t_k) = \Phi(t_{k+1}, t_k, \bar{\mathbf{x}}) \hat{\delta}_{\mathbf{x}}(t_k | t_k) = 0 \quad , \quad (1.11)$$

che è equivalente a $\hat{\mathbf{x}}(t_{k+1} | t_k) = \bar{\mathbf{x}}(t_{k+1})$, dove $\bar{\mathbf{x}}(t)$ è ottenuto risolvendo la (1.10) con la condizione iniziale $\hat{\mathbf{x}}(t_k | t_k)$.

L'errore di predizione $\tilde{\mathbf{x}}(t_{k+1} | t_k) = \mathbf{x}(t_{k+1}) - \hat{\mathbf{x}}(t_{k+1} | t_k)$ si può allora esprimere come

$$\tilde{\mathbf{x}}(t_{k+1} | t_k) = \bar{\mathbf{x}}(t_{k+1}) + \delta_{\mathbf{x}}(t_{k+1}) - \bar{\mathbf{x}}(t_{k+1}) = \delta_{\mathbf{x}}(t_{k+1}) \quad ,$$

e la varianza dell'errore di predizione si può quindi calcolare usando la relazione lineare (1.7), ottenendo

$$P(t_{k+1} | t_k) = \Phi(t_{k+1}, t_k, \bar{\mathbf{x}}) P(t_k | t_k) \Phi(t_{k+1}, t_k, \bar{\mathbf{x}})' + Q(t_k) \quad . \quad (1.12)$$

Passo di aggiornamento rispetto alle misure

All'istante t_{k+1} è disponibile la nuova misura $\mathbf{y}(t_{k+1})$ e la stima $\hat{\mathbf{x}}(t_{k+1} | t_k)$ con la relativa matrice varianza d'errore $P(t_{k+1} | t_k)$. Si linearizza l'equazione di osservazione attorno a $\bar{\mathbf{x}}(t_{k+1}) = \hat{\mathbf{x}}(t_{k+1} | t_k)$ ottenendo

$$\begin{aligned} \delta_{\mathbf{y}}(t_{k+1}) &= h(\mathbf{x}(t_{k+1})) - h(\hat{\mathbf{x}}(t_{k+1} | t_k)) + \mathbf{w}(t_{k+1}) \\ &= H(\hat{\mathbf{x}}(t_{k+1} | t_k)) \delta_{\mathbf{x}}(t_{k+1}) + \mathbf{w}(t_{k+1}) \quad . \end{aligned} \quad (1.13)$$

Questa espressione è lineare in $\delta_{\mathbf{x}}(t_{k+1})$. Usando le formule per l'aggiornamento delle misure del Cap. X per il sistema lineare (1.7)–(1.13) si ottiene

$$\hat{\delta}_{\mathbf{x}}(t_{k+1} | t_{k+1}) = \hat{\delta}_{\mathbf{x}}(t_{k+1} | t_k) + L(t_{k+1}) [\delta_{\mathbf{y}}(t_{k+1}) - H(\hat{\mathbf{x}}(t_{k+1} | t_k)) \hat{\delta}_{\mathbf{x}}(t_{k+1} | t_k)].$$

Notiamo subito che $\hat{\delta}_{\mathbf{x}}(t_{k+1} | t_k) = 0$ per la (1.11) e che per definizione $\hat{\delta}_{\mathbf{x}}(t_{k+1} | t_{k+1}) = \hat{\mathbf{x}}(t_{k+1} | t_{k+1}) - \bar{\mathbf{x}}(t_{k+1})$. Questa relazione può pertanto essere riscritta nella forma

$$\hat{\mathbf{x}}(t_{k+1} | t_{k+1}) = \hat{\mathbf{x}}(t_{k+1} | t_k) + L(t_{k+1}) [\mathbf{y}(t_{k+1}) - h(\hat{\mathbf{x}}(t_{k+1} | t_k))] \quad (1.14)$$

e questa è l'equazione di aggiornamento rispetto alle misure del filtro di Kalman esteso. La matrice guadagno si calcola con le stesse formule date al Cap. X e si ha così

$$L(t_{k+1}) = P(t_{k+1} | t_k) H(\hat{\mathbf{x}}(t_{k+1} | t_k)) \Lambda(t_{k+1})^{-1} \quad , \quad (1.15)$$

dove

$$\Lambda(t_{k+1}) = H(\hat{\mathbf{x}}(t_{k+1} | t_k)) P(t_{k+1} | t_k) H(\hat{\mathbf{x}}(t_{k+1} | t_k))' + R \quad . \quad (1.16)$$

La matrice $\Lambda(t_k)$ è interpretabile come una approssimazione per la varianza della "pseudo-innovazione"

$$\mathbf{e}(t_k) = \mathbf{y}(t_k) - h(\hat{\mathbf{x}}(t_k | t_{k-1})) \quad . \quad (1.17)$$

Naturalmente il processo $\{\mathbf{e}(t_k)\}$ non è più rigorosamente bianco ma solo approssimativamente tale.

La varianza dell'errore di filtraggio,

$$\begin{aligned} \tilde{\mathbf{x}}(t_{k+1} | t_{k+1}) &= \mathbf{x}(t_{k+1}) - \hat{\mathbf{x}}(t_{k+1} | t_{k+1}) \\ &= \mathbf{x}(t_{k+1}) - \bar{\mathbf{x}}(t_{k+1}) - [\hat{\mathbf{x}}(t_{k+1} | t_{k+1}) - \bar{\mathbf{x}}(t_{k+1})] \\ &= \delta_{\mathbf{x}}(t_{k+1}) - \hat{\delta}_{\mathbf{x}}(t_{k+1} | t_{k+1}) = \delta_{\mathbf{x}}(t_{k+1}) \quad , \end{aligned}$$

si può calcolare facendo ancora riferimento al modello linearizzato (1.7)–(1.13). La formula di aggiornamento (simmetrizzata) assume così la forma

$$\begin{aligned} P(t_{k+1} | t_{k+1}) &= (I - L(t_{k+1}) \hat{H}(t_{k+1} | t_k)) P(t_{k+1} | t_k) \\ &\quad \cdot (I - L(t_{k+1}) \hat{H}(t_{k+1} | t_k))' + L(t_{k+1}) R L(t_{k+1})' \quad . \quad (1.18) \end{aligned}$$

Prima di discutere la struttura di questo algoritmo e i problemi di calcolo connessi alla sua implementazione, riassumiamo le formule ottenute nel seguente enunciato. Allo scopo di semplificare le notazioni utilizzeremo i simboli

$$\hat{\Phi}(k | k) := \Phi(t_{k+1}, t_k, \hat{\mathbf{x}}(t_k | t_k)), \quad \hat{H}(k+1 | k) := H(\hat{\mathbf{x}}(t_{k+1} | t_k)) \quad . \quad (1.19)$$

e sostituiremo dappertutto l'indice discreto k all'argomento temporale t_k nelle variabili; ad esempio scriveremo $\hat{\mathbf{x}}(k+1 | k)$ al posto di $\hat{\mathbf{x}}(t_{k+1} | t_k)$ etc.

ALGORITMO 11.1 (Filtro di Kalman esteso). *Per il modello non lineare descritto dalle equazioni dinamiche (1.1), gli stimatori $\hat{\mathbf{x}}(k+1 | k)$, $\hat{\mathbf{x}}(k | k)$ e le rispettive varianze d'errore $P(k+1 | k)$, $P(k | k)$ si possono calcolare ricorsivamente mediante il seguente algoritmo.*

- Si risolve l'equazione differenziale

$$\dot{\hat{\mathbf{x}}}(t) = f(\bar{\mathbf{x}}(t)) \quad (1.20a)$$

$$\bar{\mathbf{x}}(t_k) = \hat{\mathbf{x}}(k | k) \quad . \quad (1.20b)$$

nell'intervallo $[t_k, t_{k+1}]$ e si pone $\hat{\mathbf{x}}(k+1 | k) := \bar{\mathbf{x}}(t_{k+1})$.

- Si calcola la varianza a priori,

$$P(k+1 | k) = \hat{\Phi}(k | k) P(k | k) \hat{\Phi}(k | k)' + Q(k) \quad . \quad (1.21)$$

- Calcolo della stima a posteriori:

$$\hat{\mathbf{x}}(k+1 | k+1) = \hat{\mathbf{x}}(k+1 | k) + L(k+1) [\mathbf{y}(k+1) - h(\hat{\mathbf{x}}(k+1 | k))] \quad (1.22)$$

dove il guadagno $L(k+1)$ si calcola mediante le

$$\begin{aligned}\Lambda(k+1) &= \hat{H}(k+1|k)P(k+1|k)\hat{H}(k+1|k)' + R \\ L(k+1) &= P(k+1|k)\hat{H}(k+1|k)'\Lambda(k+1)^{-1} .\end{aligned}\quad (1.23)$$

• *Calcolo della varianza a posteriori*

$$P(k+1|k+1) = [I - L(k+1)\hat{H}(k+1|k)]P(k+1|k)[I - L(k+1)\hat{H}(k+1|k)]' + L(k+1)RL(k+1)'. \quad (1.24)$$

• *Condizioni iniziali:*

$$\hat{\mathbf{x}}(0|-1) = E\mathbf{x}(t_0) \quad P(0|-1) = \text{Var}\mathbf{x}(t_0) . \quad (1.25)$$

OSSERVAZIONE 11.1. .

- (1) Una prima osservazione importante da fare è che l'algoritmo appena descritto usa il modello linearizzato *solo per il calcolo del guadagno $L(k)$ e delle matrici varianza d'errore.*

L'elaborazione dei dati di misura viene effettuata invece sulla base del modello non lineare originario. In particolare il passo di predizione viene effettuato usando l'equazione differenziale non lineare originaria del modello e la stima "a priori" $\hat{\mathbf{x}}(k+1|k)$ viene corretta in base alla descrizione "vera" (non linearizzata) del sensore di misura.

- (2) Gli stimatori $\hat{\mathbf{x}}(k+1|k)$ e $\hat{\mathbf{x}}(k|k)$ calcolati con l'E.K.F. sono quindi essi stessi funzioni *non lineari* dei dati di misura. Il problema di quale interpretazione statistica si possa dare a queste quantità è connesso al problema generale del *filtraggio non lineare* del quale non possiamo assolutamente occuparci in questo corso (il testo di Jazwinski [68] contiene fra l'altro anche un'eccellente introduzione al filtraggio non lineare). Ci limitiamo qui a menzionare il fatto che le equazioni dell'E.K.F. sono "molto simili" a quelle che descrivono l'evoluzione della moda condizionata di \mathbf{x} (cioè il vettore $\hat{\mathbf{x}}$ che massimizza la densità condizionata $p(\mathbf{x}|\mathbf{y}^t)$).
- (3) Dal punto di vista realizzativo è evidente che l'algoritmo è computazionalmente molto più "costoso" del corrispettivo lineare. In effetti per il calcolo del guadagno $L(k+1)$ bisogna usare le matrici $\Phi(k|k)$ e $\hat{H}(k+1|k)$ e queste devono essere ricalcolate *ad ogni passo di campionamento* in corrispondenza alle stime correnti $\hat{\mathbf{x}}(k|k)$ e $\hat{\mathbf{x}}(k+1|k)$. Ne viene che in particolare la successione dei guadagni $\{L(k)\}$ *non può essere precalcolata fuori linea.*
- (4) Come per il caso lineare, la presenza di un termine additivo di ingresso nell'equazione di stato (1.1) del tipo

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) + B u(t) + \boldsymbol{\xi}(t) , \quad (1.26)$$

dove u è un segnale deterministico oppure ottenuto per retroazione *causale* sull'uscita (cioè $u(t) = k(\mathbf{y}(s))$; $s < t$), non dà luogo a complicazioni apprezzabili nella struttura dell'algoritmo. Basta allo scopo solamente modificare l'equazione di predizione definendo ora $\mathbf{x}(t)$ come soluzione della

$$\dot{\bar{\mathbf{x}}}(t) = f(\bar{\mathbf{x}}(t)) + B u(t) , \quad \bar{\mathbf{x}}(t_k) = \hat{\mathbf{x}}(t|k) . \quad (1.27)$$

◇

Il fatto che il predittore $\hat{\mathbf{x}}(k+1 | k)$ debba essere calcolato risolvendo in linea l'equazione differenziale (1.10) può sembrare particolarmente gravoso e poco attraente, ma questo è realmente necessario solo in applicazioni in cui gli intervalli di campionamento sono particolarmente lunghi (ad esempio in problemi di ricostruzione di orbite di satelliti le misure radar possono avere periodi di ripetizione dell'ordine dell'ora). Se il campionamento è sufficientemente fitto (rispetto alla velocità di variazione dei segnali in gioco), si può sempre sostituire all'equazione differenziale (1.10) un opportuno schema discreto,

$$\bar{\mathbf{x}}(k+1) = \hat{f}_k(\bar{\mathbf{x}}(k)) \quad , \quad (1.28)$$

in cui la funzione \hat{f}_k risulta dall'applicazione di uno dei classici schemi di risoluzione numerica di equazioni differenziali ordinarie (predittore-correttore ecc.) noti dall'analisi numerica. Il più primitivo di questi schemi è la ben nota discretizzazione di Eulero,

$$\bar{\mathbf{x}}(k+1) = \bar{\mathbf{x}}(k) + (t_{k+1} - t_k) f(\bar{\mathbf{x}}(k)) \quad , \quad (1.29)$$

che può essere di pratico impiego in problemi in cui la dinamica del sistema è molto lenta rispetto al periodo di campionamento.

La discretizzazione dell'equazione differenziale (1.10) permette così di ridurre il passo di calcolo della stima a priori (predittore) alla (cfr. la (1.28))

$$\hat{\mathbf{x}}(k+1 | k) = \hat{f}_k(\hat{\mathbf{x}}(k | k)) \quad . \quad (1.30)$$

Si noti che così facendo abbiamo in realtà automaticamente fatto riferimento a una versione campionata *anche dell'equazione di stato* in (1.1) alla quale può equivalentemente sostituirsi la

$$\mathbf{x}(k+1) = \hat{f}_k(\mathbf{x}(k)) + \mathbf{v}(k) \quad , \quad (1.31)$$

dove $\{\mathbf{v}(k)\}$ è un opportuno processo bianco a tempo discreto (che congloba eventualmente anche gli errori di discretizzazione del modello a tempo continuo). Ripercorrendo la derivazione della formula dell'E.K.F. fatta in precedenza facendo riferimenti ora al *modello a tempo discreto*

$$\begin{aligned} \mathbf{x}(k+1) &= \hat{f}_k(\mathbf{x}(k)) + \mathbf{v}(k) \\ \mathbf{y}(k) &= h(\mathbf{x}(k)) + \mathbf{w}(k) \quad , \end{aligned} \quad (1.32)$$

è in effetti immediato constatare che tutte le formule date nel Teorema 11.1 rimangono invariate con la sola eccezione del calcolo del predittore di un passo che va effettuato utilizzando invece la formula (1.29).

ALGORITMO 11.2 (E.K.F. a tempo discreto). *Per il modello discreto descritto dalle equazioni (1.32) in cui $\{\mathbf{v}(k)\}$ e $\{\mathbf{w}(k)\}$ sono rumori bianchi scorrelati fra loro e dallo stato iniziale \mathbf{x}_0 , di varianze rispettive $Q(k)$ ed R , le stime (approssimate) $\hat{\mathbf{x}}(k+1 | k)$ e $\hat{\mathbf{x}}(k | k)$ si possono calcolare con lo stesso algoritmo 11.1, eccezion fatta per il calcolo del predittore di un passo che viene invece aggiornato mediante l'equazione (1.29).*

La definizione di $\hat{\Phi}(k | k)$ è in questo caso

$$\hat{\Phi}(k | k) = \left. \frac{\partial \hat{f}_k}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}(k|k)} \quad . \quad (1.33)$$

Notiamo ancora che nel caso di dinamica di stato a tempo continuo il calcolo di $\Phi(k | k)$ mediante la (1.19) introduce una complicazione ulteriore, giacché occorrerebbe ad ogni passo di campionamento calcolare la matrice fondamentale associata al sistema lineare *tempo-variante* $\dot{z} = F(\bar{\mathbf{x}}(t)) z$. Questo calcolo può essere spesso fatto in modo approssimato senza portare a deterioramenti apprezzabili nel comportamento del filtro giacché, come si è fatto notare in precedenza, il sistema linearizzato influisce solo sul calcolo del guadagno del filtro e non interviene direttamente nella elaborazione dei dati di misura. Un modo tipico di procedere è di considerare $F(\bar{\mathbf{x}}(t))$ costante nell'intervallo di campionamento e uguale a $F(\hat{\mathbf{x}}(k | k))$. L'esponenziale di questa matrice può in qualche caso essere calcolato analiticamente oppure a sua volta approssimato in modo tale da ridurre il calcolo di $\hat{\Phi}(k | k)$ (che deve essere fatto ad ogni passo di campionamento) alla valutazione di una funzione di struttura nota in corrispondenza alla stima corrente $\hat{\mathbf{x}}(k | k)$.

Considerazioni dello stesso genere possono essere fatte per quanto riguarda il calcolo della varianza del rumore di modello $Q(k)$ data dalla (1.9). Notiamo per inciso che questa espressione si può ritenere anche un modo ragionevole per approssimare la varianza del rumore di modello $\mathbf{v}(k)$ nello schema alle differenze (1.31) ottenuto per campionamento dall'equazione di stato in (1.1). In questo caso è poi ancora meno essenziale un calcolo accurato, dato che di norma l'intensità del rumore di modello Q è un parametro molto incerto (a differenza di R che solitamente si ricava in modo abbastanza accurato dalle specifiche di precisione dei sensori).

Chiudiamo quindi questa discussione mettendo in guardia il lettore sul fatto che il filtro di Kalman esteso tende a essere molto più vulnerabile a problemi di *divergenza* dell'algoritmo lineare. In genere questi problemi si risolvono aumentando artificialmente la varianza dell'errore di modello ma qualche volta la causa della divergenza può essere insita nella struttura stessa delle equazioni. In questi casi occorre riconsiderare tutto il processo di modellizzazione eventualmente impiegando sistemi di coordinate diversi ecc...

ESEMPIO 11.1 (La demodulazione di frequenza). Come esempio di applicazione del filtro di Kalman esteso a un tipico problema non lineare consideriamo il problema della demodulazione di frequenza (o più in generale di fase). Ammettiamo che il segnale messaggio, diciamolo $\{\mathbf{a}(t)\}$, che si vuole trasmettere, si possa in generale descrivere come un processo aleatorio a tempo continuo stazionario a media zero e di spettro approssimativamente noto. È sempre possibile approssimare lo spettro $S(i\omega)$, di \mathbf{a} , con una funzione razionale e quindi rappresentare $\{\mathbf{a}(t)\}$ come l'uscita di un sistema dinamico di dimensione finita n , del tipo

$$\begin{aligned} \dot{\mathbf{z}}(t) &= F \mathbf{z}(t) + b \boldsymbol{\xi}(t) \\ \mathbf{a}(t) &= h' \mathbf{z}(t) \quad , \end{aligned} \quad (1.34)$$

dove $\{\boldsymbol{\xi}(t)\}$ è un rumore bianco scalare a tempo continuo di intensità $q > 0$. (Il modello (1.34) si può ottenere attraverso un procedimento di fattorizzazione spettrale e successiva realizzazione del tutto analogo a quello descritto alla sezione X.1.

In uno schema di modulazione di frequenza il messaggio $\{\mathbf{a}(t)\}$ viene prima integrato ottenendo il segnale

$$\boldsymbol{\theta}(t) = \beta \int_{-\infty}^t \mathbf{a}(r) dr \quad , \quad \beta > 0 \quad , \quad (1.35)$$

e successivamente inviato a un modulatore di fase che modula una portante di frequenza ω_0 . Il segnale trasmesso ha così la forma

$$\mathbf{s}(t) = A \sin(\omega_0 t + \boldsymbol{\theta}(t)) \quad .$$

In ricezione si assume di disporre di una versione campionata (con periodo T opportuno) e rumorosa del segnale trasmesso, diciamola

$$\mathbf{y}(k) = A \sin(\Delta_0 k + \boldsymbol{\theta}(k)) + \mathbf{w}(k) \quad , \quad (1.36)$$

dove abbiamo posto $\Delta_0 := \omega_0 T$, $\boldsymbol{\theta}(k) := \boldsymbol{\theta}(kT)$ e $\mathbf{w}(k) := \mathbf{w}(kT)$. Il processo $\{\mathbf{w}(k)\}$ verrà supposto bianco e di varianza (scalare) nota $r > 0$. Notiamo che in realtà in ricezione la frequenza portante viene di solito convertita a una opportuna frequenza intermedia. L'ampiezza A dell'onda modulata viene poi normalizzata attraverso degli opportuni circuiti rigeneratori. Noi nel seguito trascureremo questo tipo di dettagli. Accoppiando le (1.34) e (1.35) si arriva a un modello di stato del segnale modulante $\{\boldsymbol{\theta}(t)\}$ del tipo

$$\begin{aligned} \begin{bmatrix} \dot{\mathbf{z}}(t) \\ \dot{\mathbf{z}}_{n+1}(t) \end{bmatrix} &= \begin{bmatrix} F & 0 \\ h' & 0 \end{bmatrix} \begin{bmatrix} \mathbf{z}(t) \\ \mathbf{z}_{n+1}(t) \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} \boldsymbol{\xi}(t) \\ \boldsymbol{\theta}(t) &= \begin{bmatrix} 0 & \beta \end{bmatrix} \begin{bmatrix} \mathbf{z}(t) \\ \mathbf{z}_{n+1}(t) \end{bmatrix} \quad . \end{aligned} \quad (1.37)$$

Denotando lo stato $(n+1)$ -dimensionale di questo modello con $\mathbf{x}(t) = [\mathbf{z}(t)' \mathbf{z}_{n+1}(t)]'$, lo riscriviamo come

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \bar{F} \mathbf{x}(t) + \bar{b} \boldsymbol{\xi}(t) \\ \boldsymbol{\theta}(t) &= c' \mathbf{x}(t) \quad , \end{aligned} \quad (1.38)$$

dove c' è il vettore riga $[0 \quad \beta]$ e gli altri simboli hanno un ovvio significato. Notiamo che il segnale messaggio è descritto dalla

$$\mathbf{a}(t) = [h' \quad 0] \mathbf{x}(t) \quad , \quad (1.39)$$

per cui la stima in linea del segnale messaggio campionato $\hat{\mathbf{a}}(k | k) := \hat{E}[\mathbf{a}(kT) | \mathbf{y}^k]$ si ottiene come trasformazione lineare di $\mathbf{x}(k | k)$. *Il demodulatore ottimo* (o approssimativamente tale) è quindi descritto dalle equazioni del filtro di Kalman esteso per il modello continuo (1.38) e l'equazione di osservazione non lineare (1.36).

Studiamo ora la struttura del demodulatore. Campionando la (1.38) con periodo T si trova

$$\begin{aligned} \mathbf{x}(k+1) &= \Phi(T) \mathbf{x}(k) + \mathbf{v}(k) \\ \boldsymbol{\theta}(k) &= c' \mathbf{x}(k) \quad , \end{aligned} \quad (1.40)$$

dove $\Phi(T) = \exp(\bar{F}T)$ e $\{\mathbf{v}(k)\}$ è un processo bianco di varianza $Q(T)$ calcolabile tramite la formula (1.9) ponendovi $Q = qb b'$. Usando la formula (1.19) si ha poi

$$\begin{aligned} \hat{\Phi}(k | k) &= \Phi(T) \\ \hat{H}(k | k-1) &= A \cos(\Delta_0 k + c' \hat{\mathbf{x}}(k | k-1)) c' \end{aligned} \quad (1.41)$$

e il guadagno del filtro può essere espresso nella forma

$$L(k) = A \frac{P(k | k-1) c}{\lambda(k)} \cos(\Delta_0 k + c' \hat{\mathbf{x}}(k | k-1)) \quad , \quad (1.42)$$

dove la varianza scalare $\lambda(k)$ è data dalla

$$\lambda(k) = A^2 c' P(k | k-1) c \cos^2 (\Delta_0 k + c' \hat{\mathbf{x}}(k | k-1)) + r \quad . \quad (1.43)$$

Le varianze $P(k | k-1)$ e $P(k | k)$ si aggiornano in linea mediante la formula del Teorema 11.1 opportunamente particolarizzata al caso in esame.

Lo schema di elaborazione del segnale in ricezione si può così rappresentare mediante le equazioni:

$$\begin{aligned} \hat{\mathbf{x}}(k | k) &= \hat{\mathbf{x}}(k | k-1) + L_0(k) \cos(\Delta_0 k + \hat{\boldsymbol{\theta}}(k | k-1)) \cdot \\ &\quad \cdot [\mathbf{y}(k) - A \sin(\Delta_0 k + \hat{\boldsymbol{\theta}}(k | k-1))] \\ \hat{\mathbf{x}}(k+1 | k) &= \Phi(T) \hat{\mathbf{x}}(k | k) \\ \hat{\boldsymbol{\theta}}(k+1 | k) &= c' \hat{\mathbf{x}}(k+1 | k) \quad , \end{aligned} \quad (1.44)$$

dove nella prima relazione abbiamo posto

$$L_0(k) = A \frac{P(k | k-1) c}{\lambda(k)} \quad . \quad (1.45)$$

La struttura di questo algoritmo è quella di un *anello di retroazione ad aggancio di fase* (phase locked loop) a guadagno variabile. Questi schemi sono usati diffusamente nelle comunicazioni elettriche specialmente per costruire dei sincronizzatori, si veda ad esempio il testo di Van Trees [142], Parte II. Lo schema a blocchi corrispondente ha la struttura riportata in Figura 12.1, dove il secondo blocco circolare rappresenta un moltiplicatore.

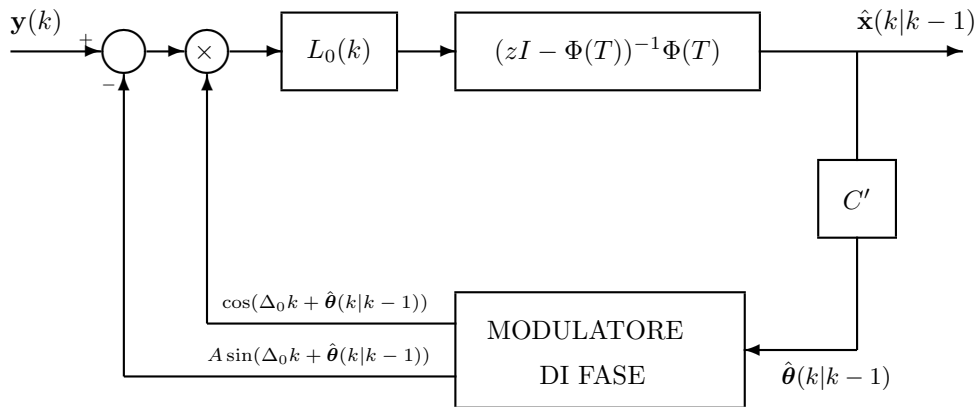


FIGURA 1.1. Schema a blocchi del filtro di Kalman esteso.

Per rendersi intuitivamente conto del funzionamento del demodulatore notiamo che in base all'identità trigonometrica $\cos \alpha \sin \beta = 1/2 \sin(\alpha + \beta) - 1/2 \sin(\alpha - \beta)$ e

alla (1.36) si può scrivere

$$\begin{aligned}
 & \cos \left(\Delta_0 k + \hat{\theta}(k | k-1) \right) [y(k) - A \sin(\Delta_0 k + \hat{\theta}(k | k-1))] \\
 &= \frac{A}{2} \sin \left(\theta(k) - \hat{\theta}(k | k-1) \right) + \frac{A}{2} \sin \left(2\Delta_0 k + \theta(k) + \hat{\theta}(k | k-1) \right) \\
 &+ \frac{A}{2} \sin \left(2\Delta_0 k + 2\hat{\theta}(k | k-1) \right) + w(k) \cos \left(\Delta_0 k + \hat{\theta}(k | k-1) \right).
 \end{aligned} \tag{1.46}$$

I termini a frequenza $2\Delta_0$ vengono attenuati e praticamente eliminati dal filtro lineare che agisce nell'anello di reazione, per cui tutto avviene come se a valle del moltiplicatore si avesse solo il segnale $A/2 \sin(\theta(k) - \hat{\theta}(k | k-1))$ corrotto da un rumore additivo $\tilde{w}(k)$ uguale all'ultimo addendo nella somma (1.46). Il sistema funziona quindi come l'anello non lineare di Figura 12.2, dove si è indicata simbolicamente con $H(z)$ la "funzione di trasferimento"

$$H(z) = c'(zI - \Phi(T))^{-1} \Phi(T) L_0(k) . \tag{1.47}$$

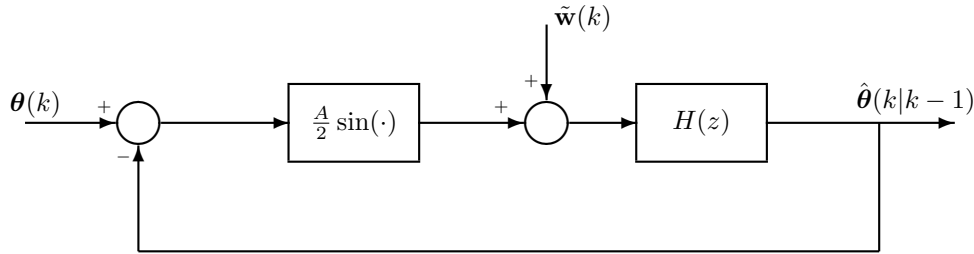


FIGURA 1.2. Schema a blocchi del predittore di $\theta(k)$.

Da notare che se l'errore di stima $\tilde{\theta}(k | k-1)$ diventa grande il sistema può "perdere il passo" (phase slipping). Nei circuiti classici di sincronizzazione in cui il segnale che interessa ricostruire è proprio la fase $\theta(t)$, anziché la sua derivata $\mathbf{a}(t)$, il filtro $H(z)$ è a parametri costanti e il problema principale è di progettare $H(z)$ in modo da minimizzare la probabilità di perdere il passo. Come si vede il filtro di Kalman, oltre a giustificare su basi rigorose l'uso dello schema ad aggancio di fase, fornisce anche la struttura ottima del filtro $H(z)$ (che dovrebbe in teoria essere a guadagno variabile). Naturalmente l'aggiornamento in linea di $L_0(k)$ mediante la formula dell'E.K.F. può essere estremamente dispendioso specie se la dinamica di $\theta(t)$ ha ordine elevato. In ogni caso queste formule forniscono una base di partenza rigorosa per procedere a delle approssimazioni. \diamond

2. Filtraggio Bayesiano non lineare a tempo discreto

In questa sezione accenneremo brevemente al problema del filtraggio non lineare. Questo argomento può diventare molto tecnico e purtroppo le formule che si trovano in letteratura sono scarsamente utili in pratica ad eccezione di alcuni semplicissimi esempi che sono di scarso interesse nelle applicazioni. Se però ci si

restringe a processi *a stati finiti* si possono fortunatamente applicare direttamente le formule più semplici e generali.

N.B. In questa sezione ritorneremo ad usare il simbolo t per denotare la variabile temporale discreta.

Ricordiamo che gli ingredienti fondamentali di un problema di filtraggio statistico sono

- un processo di Markov $\mathbf{x}(t)$, lo *stato del sistema*, che prende valori in un certo *spazio di stato* Σ . Lo stato può essere osservato solo indirettamente.
- un processo di osservazione $\mathbf{y}(t)$, che supporremo a valori reali o vettoriali.

Il fatto che $\mathbf{x}(t)$ sia lo stato del sistema e $\mathbf{y}(t)$ il processo delle osservazioni, si traduce in una condizione di *indipendenza condizionata* del futuro congiunto delle osservazioni \mathbf{y} e dello stato \mathbf{x} dal loro passato congiunto, *dato lo stato all'istante presente*. Questa condizione si può formalmente esprimere come

$$\{\mathbf{x}(s), \mathbf{y}(s); s \geq t\} \parallel \{\mathbf{x}(s), \mathbf{y}(s); s \leq t\} \mid \mathbf{x}(t) \quad (2.1)$$

dove il simbolo \parallel denota indipendenza.

Usando la definizione di indipendenza si può verificare che la relazione (2.1) implica separatamente la proprietà di Markov, i.e.

$$\{\mathbf{x}(s); s \geq t\} \parallel \{\mathbf{x}(s); s \leq t\} \mid \mathbf{x}(t)$$

e la proprietà di memoria dinamica per il processo di osservazione, i.e.

$$\{\mathbf{y}(s); s \geq t\} \parallel \{\mathbf{y}(s); s \leq t\} \mid \mathbf{x}(t)$$

In queste ipotesi si possono dare delle formule ricorsive generali che descrivono l'evoluzione temporale delle distribuzioni di probabilità condizionate a priori e a posteriori, di $\mathbf{x}(t)$ date le osservazioni fino all'istante di tempo corrente. Abuseremo un pò delle notazioni denotando col simbolo

$$p(x(t) \mid \mathbf{y}^t)$$

la distribuzione di probabilità condizionata della variabile $\mathbf{x}(t)$, date le osservazioni \mathbf{y}^t . Il simbolo $x(t)$ ha qui la funzione di una variabile corrente. Nel caso più interessante che considereremo in dettaglio più avanti, lo stato sarà un processo di Markov a stati finiti, cioè una *Catena di Markov* e p denoterà quindi una *distribuzione discreta*.

Si suppongono note la distribuzione $p(y(t) \mid \mathbf{x}(t))$ che descrive l'accoppiamento stato-osservazione e la distribuzione condizionata $p(x(t+1) \mid \mathbf{x}(t))$ che descrive la dinamica dello stato. Per semplicità supporremo che entrambe questa distribuzioni siano tempo invarianti. Lo stato iniziale ha distribuzione assegnata $p(x_0)$.

TEOREMA 11.1. *Nelle ipotesi fatte, le distribuzioni condizionate a priori e a posteriori, dello stato $\mathbf{x}(t)$, date le osservazioni correnti, si aggiornano mediante le formule ricorsive*

$$p(x(t+1) \mid \mathbf{y}^t) = \int_{\mathbf{x}(t) \in \Sigma} p(x(t+1) \mid x(t)) p(x(t) \mid \mathbf{y}^t) dx(t) \quad (2.2)$$

$$p(x(t+1) \mid \mathbf{y}^{t+1}) = \frac{p(\mathbf{y}(t+1) \mid x(t+1)) p(x(t+1) \mid \mathbf{y}^t)}{p(\mathbf{y}(t+1) \mid \mathbf{y}^t)} \quad (2.3)$$

con la condizione iniziale $p(x(t_0) \mid \mathbf{y}^{t_0-1}) = p(x_0)$.

Notare che il denominatore in (2.3) è semplicemente un fattore di normalizzazione che si ottiene integrando il numeratore rispetto a $x(t+1)$.

DIMOSTRAZIONE. La prima relazione si ottiene facilmente facendo uso della (2.1), che si particularizza in

$$p(x(t+1) | x(t), \mathbf{y}^t) = p(x(t+1) | x(t)).$$

La seconda si ottiene con la regola di Bayes applicata alla probabilità congiunta (condizionata) $p(x(t+1), y(t+1) | \mathbf{y}^t) = p(y(t+1) | x(t+1), \mathbf{y}^t)p(x(t+1) | \mathbf{y}^t)$, tenendo ancora conto di (2.1) che permette di scrivere

$$p(y(t+1) | x(t+1), \mathbf{y}^t) = p(y(t+1) | x(t+1))$$

dato che la conoscenza di $x(t+1)$ rende $\mathbf{y}(t+1)$ condizionatamente indipendente dal suo passato \mathbf{y}^t . \square

Le equazioni dinamiche del filtro di Kalman ricavate al capitolo X, hanno la stessa struttura. Se si assume che tutte le variabili in gioco siano Gaussiane, si possono in effetti ritrovare (con un certo numero di passaggi algebrici), partendo dalle (2.2) (2.3).

Le formule diventano particolarmente semplici nel caso in cui il processo di stato è una catena di Markov a N stati con matrice di transizione P , nota. Ricordiamo a questo scopo che la matrice di transizione ha il significato seguente,

$$P_{kj} = p(\mathbf{x}(t+1) = k | \mathbf{x}(t) = j), \quad k, j = 1, \dots, N$$

Sarà conveniente usare notazioni vettoriali e definire i vettori di probabilità condizionate $\mathbf{p}(t+1 | t)$, $\mathbf{p}(t | t)$ nel seguente modo

$$\mathbf{p}_k(t+1 | t) := p(x(t+1) = k | \mathbf{y}^t), \quad \mathbf{p}_k(t | t) := p(x(t) = k | \mathbf{y}^t) \quad k = 1, \dots, N \quad (2.4)$$

Definiamo inoltre la matrice diagonale $N \times N$, $D(y(t))$, ponendo

$$D(y(t)) := \text{diag} \{p(y(t) | x(t) = 1), \dots, p(y(t) | x(t) = N)\}. \quad (2.5)$$

Con queste notazioni le formule del Teorema 11.1 assumono l'aspetto seguente.

COROLLARIO 11.1. *Nel caso di un processo di stato a stati finiti, le probabilità condizionate del Teorema 11.1 si aggiornano ricorsivamente mediante le*

$$\mathbf{p}(t+1 | t) = P \mathbf{p}(t | t) \quad (2.6)$$

$$\mathbf{p}(t+1 | t+1) = \frac{D(\mathbf{y}(t+1))\mathbf{p}(t+1 | t)}{\|D(\mathbf{y}(t+1))\mathbf{p}(t+1 | t)\|_1} \quad (2.7)$$

dove $\|x\|_1 := \sum_1^N x_k$.

Queste formule permettono di calcolare in tempo reale lo stimatore MAP dello stato, semplicemente attraverso una massimizzazione sulle componenti dei vettori delle probabilità condizionate, ad esempio

$$\hat{\mathbf{x}}(t | t) = \text{Arg} \max_k \{\mathbf{p}_k(t | t); k = 1, \dots, N\} \quad (2.8)$$

Queste semplici formule si possono applicare direttamente a problemi di stima di segnali numerici, ma possono anche essere usate per approssimare la soluzione di problemi di filtraggio molto "non lineari" per i quali il filtro di Kalman esteso dà prestazioni insoddisfacenti. Naturalmente occorre discretizzare opportunamente lo

spazio di stato del sistema e calcolarsi preliminarmente la matrice di transizione P a partire dal modello di stato continuo. Si possono in questo modo trattare solo problemi di filtraggio di piccola dimensione perchè all'aumentare di n il procedimento di discretizzazione diventa sempre più complesso.

Applicazione all'eliminazione di interferenza intersimbolo. Un esempio di applicazione delle formule appena viste viene dal problema dell'eliminazione dell'*interferenza intersimbolo* nelle comunicazioni elettriche. In questo contesto si ha un segnale in ricezione che ha la forma

$$\mathbf{y}(t) = \sum_{k=0}^N g_k \mathbf{s}(t - k) + \mathbf{w}(t) \quad (2.9)$$

dove $\{\mathbf{s}(t)\}$ è il segnale numerico trasmesso dalla sorgente, $g_k := g(kT)$ sono i valori campionati della risposta del canale all'impulso base usato per la modulazione (ad esempio un impulso rettangolare di durata T o una semionda sinusoidale di periodo T , etc..). Si assume normalmente che la risposta impulsiva del canale sia finita e che per $k > N$ si abbia $g_k \sim 0$.

Il segnale numerico \mathbf{s} è un segnale a stati finiti, spesso rappresentabile come un processo stazionario a variabili indipendenti. Per fissare le idee supporremo che \mathbf{s} sia un segnale binario i.i.d.. Il rumore additivo \mathbf{w} , introdotto dal canale (analogico) e dalla demodulazione è bianco, Gaussiano, di varianza nota σ^2 .

L'interferenza intersimbolo non si avrebbe se il canale fosse istantaneo (a banda infinita), cioè fosse $g_k = 0$, $k > 0$, perchè allora il segnale in ricezione sarebbe $\mathbf{y}(t) = g_0 \mathbf{s}(t) + \mathbf{w}(t)$ e quindi (a meno del rumore) sarebbe proporzionale solo al simbolo corrente, $\mathbf{s}(t)$. Per eliminarla, bisogna quindi modellare la dinamica del canale. Questo si può fare mediante schemi del tipo rappresentato in figura

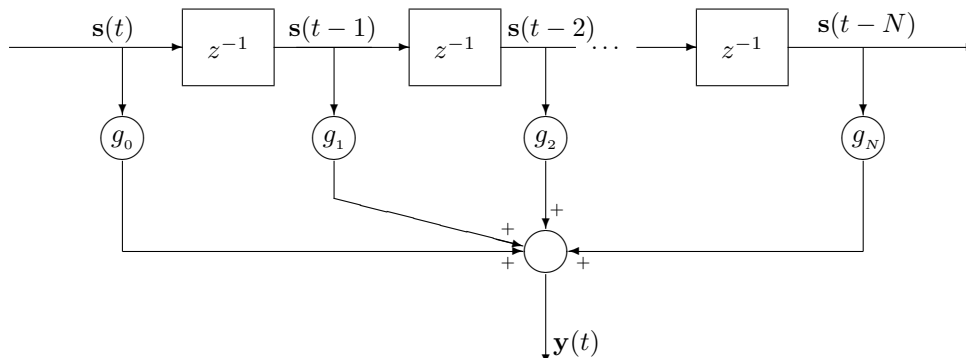


FIGURA 2.1. Schema a blocchi del canale numerico.

dove i coefficienti $\{g_k\}$ sono le risposte campionate del canale all'impulso base, possibilmente scalate per fattori noti, e si suppongono noti. Come stato del canale

si può prendere il vettore

$$\mathbf{x}(t) := \begin{bmatrix} \mathbf{s}(t-N) \\ \dots \\ \mathbf{s}(t-1) \\ \mathbf{s}(t) \end{bmatrix}$$

che soddisfa alla ricorsione

$$\mathbf{x}(t+1) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & \ddots & 1 \\ 0 & \dots & & & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} \mathbf{s}(t+1) \quad (2.10)$$

$$\mathbf{y}(t) = [g_N \dots g_0] \mathbf{x}(t) + \mathbf{w}(t) \quad (2.11)$$

Si può cercare di adattare la teoria della stima lineare a questo caso, ma evidentemente lo stimatore lineare ricostruisce stime dello stato che sono numeri (o vettori) *reali* e non ci si può aspettare che da questi si sia in grado poi di individuare in modo affidabile un particolare stato discreto in cui si trova il canale. Conviene quindi cercare di impostare direttamente il problema come stima dello stato discreto in cui si trova il canale.

Notiamo a questo proposito che se si riguarda lo stato all'istante t , $\mathbf{x}(t)$, semplicemente come una $N+1$ -pla ordinata di possibili simboli dell'alfabeto in cui prende valori $\mathbf{s}(t)$, esso è un processo (catena) di Markov stazionario a $2^{(N+1)}$ possibili stati. La matrice di transizione di questa catena di Markov si può ricavare facilmente dalla specificazione delle probabilità con cui vengono emessi i simboli $\mathbf{s}(t)$ e si possono così applicare direttamente le formule di stima MAP viste a suo tempo. Per esempio, supponiamo che la probabilità dell'evento $\{\mathbf{s}(t) = 0\}$ sia p e quella di $\{\mathbf{s}(t) = 1\}$ sia $q = 1 - p$, indipendenti da t . Se il canale ha memoria $N = 1$, ci sono 4 possibili stati $\{x_k, k = 1, 2, 3, 4\}$ in cui si può trovare. Nel caso specifico, lo spazio di stato della catena è quindi $\{00, 01, 10, 11\}$ e la matrice di transizione P è facilmente calcolabile individuando le transizioni di stato possibili del sistema,

$$P = \begin{matrix} & \begin{matrix} 00 & 01 & 10 & 11 \end{matrix} \\ \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix} & \begin{bmatrix} p & p & 0 & 0 \\ 0 & 0 & p & p \\ q & q & 0 & 0 \\ 0 & 0 & q & q \end{bmatrix} \end{matrix}$$

dove, sopra la matrice sono elencati gli stati di partenza e a sinistra quelli di arrivo.

Gli elementi della matrice D in (2.5) sono le probabilità condizionate

$$p(y(t) | \mathbf{x}(t) = x_k) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\frac{1}{2} \left\{ \frac{(y(t) - g(x_k))^2}{\sigma^2} \right\}$$

dove $g(x_k)$ è il valore dell'uscita in assenza di rumore additivo, quando lo stato del canale è x_k . Le probabilità vanno calcolate sostituendo alla variabile corrente $y(t)$ il valore osservato di $\mathbf{y}(t)$. \diamond

3. Accordo (“tuning”) del filtro di Kalman

Come è già stato messo in evidenza a più riprese, il “rumore di modello” $\{\mathbf{v}(t)\}$ che compare nelle equazioni di stato serve molto più a tener conto delle inevitabili approssimazioni che si introducono in fase di modellistica dei segnali in gioco che non a descrivere dei veri e propri termini forzanti aleatori di origine “fisica”. Spesso occorre tenere presenti dei limiti a priori sulla potenza di calcolo disponibile (specialmente in problemi non lineari in cui l’equazione di Riccati va risolta in linea) e quindi è necessario ricorrere a modelli di ordine ridotto che non riescono a descrivere completamente tutte le caratteristiche salienti dei segnali. L’errore che si commette usando un modello semplificato viene così schematizzato introducendo un rumore bianco additivo nell’equazione di stato. In altri casi la dinamica del segnale che si vuole stimare è solo approssimativamente nota e il meglio che si può fare è descriverne matematicamente in modo deterministico un andamento qualitativo medio. Il termine di fluttuazione aleatoria $\mathbf{v}(t)$ aggiunto al modello serve così a rendere a priori possibili le deviazioni dall’andamento medio qualitativo che si impone al segnale mediante il modello di stato deterministico. Alcuni esempi di questo modo di procedere sono stati già visti al Capitolo 9, sezione 3. Altri esempi verranno illustrati qui di seguito. In tutti questi esempi è in genere assai difficile avere a priori un’idea (per quanto poco precisa) della varianza Q che meglio si adatta a quantificare gli errori di modello. In questi casi è necessario effettuare delle “tarature” sperimentali di questo parametro, basandosi su simulazioni.

Stima di disturbi lentamente variabili (*bias*). Spesso non è nemmeno certo che l’errore di modello, $\{\mathbf{v}(t)\}$, che si introduce in fase di descrizione del sistema mediante un modello di stato semplificato, abbia *media zero*, come viceversa dovrebbe essere in teoria. In queste circostanze è assai opportuno *separarne la componente lentamente variabile* (che ha media significativamente non nulla su intervalli di tempo dell’ordine dei tempi di risposta o delle costanti di tempo tipiche del sistema). L’errore di modellizzazione, che adesso per convenienza chiameremo $\boldsymbol{\varepsilon}(t)$, viene così decomposto nella forma

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{\beta}(t) + \mathbf{v}(t) \quad , \quad (3.1)$$

dove $\boldsymbol{\beta}(t)$ rappresenta la componente lentamente variabile (“bias”) e $\{\mathbf{v}(t)\}$ è un rumore bianco a *media zero*. Per la modellizzazione di $\boldsymbol{\beta}$ tipicamente si ricorre a descrizioni del tipo “passeggiata aleatoria”,

$$\boldsymbol{\beta}(t+1) = \boldsymbol{\beta}(t) + \mathbf{n}(t) \quad , \quad (3.2)$$

dove la varianza di $\mathbf{n}(t)$ è scelta opportunamente piccola, oppure se è il caso, a modelli del tipo “rampa aleatoria” ecc. Questi modelli sono stati studiati nella sezione 3 del Capitolo 9. C’è da aggiungere che in pratica la separazione (3.1) e la conseguente “aggiunta” di $\{\boldsymbol{\beta}(t)\}$ al processo di stato si rivela spesso uno strumento utile per combattere fenomeni di *divergenza* dinamica del filtro che normalmente si verificano quando si assume che $\boldsymbol{\varepsilon}(t)$ sia bianco e a media nulla. Qui sotto descriveremo alcuni casi in cui l’introduzione di una componente lentamente variabile nel modello di stato, gioca un ruolo particolarmente significativo.

Impiego del filtro di Kalman nei Controlli Automatici. Nelle applicazioni nel campo dei controlli automatici l’inclusione di un termine lentamente variabile del tipo (3.2) fra le variabili di stato può risultare dalla modellizzazione

di disturbi effettivamente agenti sul processo. L'idea di descrivere statisticamente i disturbi fornisce, come abbiamo già visto nel caso stazionario, dei criteri per il progetto razionale dei regolatori. Vedremo che in questo caso la struttura del regolatore è basata essenzialmente sulla stima in linea della componente media del disturbo che agisce sul sistema da controllare.

Per chiarire meglio questo punto supponiamo che la dinamica del processo da controllare sia descritta dal modello lineare

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + B u(t) + D\mathbf{d}(t) \quad , \quad (3.3)$$

dove $\mathbf{d}(t)$ rappresenta il vettore dei disturbi non misurabili e $u(t)$ è la variabile di ingresso manipolabile. In genere, per cause di varia natura ma principalmente dovute alle inerzie degli attuatori e a limiti inerenti il dispendio di energia di controllo, non è consigliabile “opporsi” alle componenti a frequenza elevata dei disturbi che vengono d'altro canto normalmente tagliate dalla dinamica “passa basso” del processo. Si cerca così di progettare il sistema di controllo in modo tale che esso agisca efficacemente solo sulla componente lentamente variabile di $\mathbf{d}(t)$.

Si decompone allora $\mathbf{d}(t)$ usando lo schema

$$\mathbf{d}(t) = \boldsymbol{\beta}(t) + \tilde{\mathbf{v}}(t) \quad , \quad (3.4)$$

dove $\boldsymbol{\beta}(t)$ è descritto ad esempio dal modello di tipo “passeggiata aleatoria” (3.2) e $\tilde{\mathbf{v}}(t)$ è un termine bianco (rapidamente variabile) a media zero. Naturalmente il processo $\{\boldsymbol{\beta}(t)\}$ è variabile “lentamente” solo se la varianza del rumore bianco $\{\mathbf{n}(t)\}$ nella (3.2) è scelta *opportunamente piccola*.

Aggregando il modello del processo (3.3) e la (3.2) si perviene a un modello a “stato aumentato” del tipo

$$\begin{bmatrix} \mathbf{x}(t+1) \\ \boldsymbol{\beta}(t+1) \end{bmatrix} = \begin{bmatrix} A & D \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\beta}(t) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} \mathbf{v}(t) \\ \mathbf{n}(t) \end{bmatrix} \quad , \quad (3.5)$$

dove il rumore bianco $\mathbf{v}(t) := D\tilde{\mathbf{v}}(t)$ si assume scorrelato da $\{\mathbf{n}(t)\}$. Il filtro di Kalman, che ha normalmente la funzione di *ricostruttore dello stato* in queste applicazioni, opera su delle misure legate allo stato originario che possono quindi essere descritte da una relazione del tipo

$$\mathbf{y}(t) = [C \ 0] \begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\beta}(t) \end{bmatrix} + \mathbf{w}(t) \quad (3.6)$$

Usando questo modello si possono aggiornare in linea stime a minima varianza d'errore anche della componente lenta del disturbo.

PROBLEMA 11.1. *Il lettore verifichi che il modello aumentato è osservabile se lo è la coppia (C, A) del modello originario.* \diamond

Le stime $\hat{\boldsymbol{\beta}}(t+1 | t)$ o $\hat{\boldsymbol{\beta}}(t | t)$ possono essere impiegate per una compensazione in linea della componente lenta dei disturbi. Per esempio, le uscite $\hat{\mathbf{x}}(t | t)$ e $\hat{\boldsymbol{\beta}}(t | t)$ del filtro di Kalman costruito sul modello aumentato (3.5) possono venire utilizzate da una legge di controllo a retroazione lineare del tipo

$$u(t) = -K \hat{\boldsymbol{\beta}}(t | t) - G \hat{\mathbf{x}}(t | t) \quad , \quad (3.7)$$

dove K è un'opportuna matrice guadagno che si vuole determinare in modo da minimizzare l'effetto della componente lenta del disturbo sulle variabili regolate del sistema. Il termine di retroazione $G \hat{\mathbf{x}}(t | t)$ può essere progettato ad esempio

per assegnare i poli in catena chiusa della parte raggiungibile del sistema, oppure G può essere determinato minimizzando un indice di costo quadratico [89] ecc.¹. Chiaramente se B fosse invertibile, prendendo $K = B^{-1}D$ si riuscirebbe a eliminare completamente dal disturbo $D\mathbf{d}(t)$ almeno la componente media stimata $D\hat{\boldsymbol{\beta}}(t|t)$. Supponendo gli errori di stima $\tilde{\boldsymbol{\beta}}(t|t)$ piccoli, si raggiungerebbe così l'obiettivo di togliere ai disturbi agenti proprio la componente media $\boldsymbol{\beta}(t)$ descritta nelle (3.4) e (3.2).

È facile verificare che a regime la stima $\hat{\boldsymbol{\beta}}(t|t)$ si ottiene mediante lo schema ricorsivo

$$\hat{\boldsymbol{\beta}}(t+1|t+1) = \hat{\boldsymbol{\beta}}(t|t) + \bar{L}_{\boldsymbol{\beta}}[\mathbf{y}(t+1) - \hat{\mathbf{y}}(t+1|t)], \quad (3.8)$$

dove $\bar{L}_{\boldsymbol{\beta}}$ è la sottomatrice formata dalle ultime n righe del guadagno a regime \bar{L} . Il predittore $\hat{\mathbf{y}}(t|t-1)$ è descritto dalla

$$\begin{aligned} \hat{\mathbf{x}}(t+1|t) &= A\hat{\mathbf{x}}(t|t) + D\hat{\boldsymbol{\beta}}(t|t) + B\mathbf{u}(t) \\ \hat{\mathbf{y}}(t|t-1) &= C\hat{\mathbf{x}}(t|t-1) \end{aligned} \quad (3.9)$$

e può essere interpretato come l'uscita del sistema (3.5) in condizioni di *disturbo agente uguale a $\hat{\boldsymbol{\beta}}(t|t)$* , supponendo che i rumori $\mathbf{v}(t)$ e $\mathbf{w}(t)$ diano contributi trascurabili all'uscita rispetto a quello dato dal disturbo $\hat{\boldsymbol{\beta}}(t|t)$.

Confrontando queste relazioni con il modello nominale

$$\begin{aligned} \mathbf{x}(t+1) &= A\mathbf{x}(t) + D\mathbf{d}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) &= C\mathbf{x}(t) \end{aligned} \quad (3.10)$$

in cui abbiamo supposto ancora gli errori di modello e di misura in (3.5) (3.6) trascurabili rispetto a $\mathbf{d}(t)$, si trova che la differenza $\mathbf{e}(t) := \mathbf{y}(t) - \hat{\mathbf{y}}(t|t-1)$ è l'errore sull'uscita provocato dalla deviazione aleatoria del disturbo $\mathbf{d}(t)$ dal valore stimato $\hat{\boldsymbol{\beta}}(t|t)$. Quindi:

Nel caso di modello ragionevolmente accurato e misure poco rumorose, rispetto al disturbo agente, l'innovazione $\mathbf{e}(t) = \mathbf{y}(t) - \hat{\mathbf{y}}(t|t-1)$ è causata sostanzialmente dalle fluttuazioni del disturbo agente attorno al valore stimato $\hat{\boldsymbol{\beta}}(t|t)$.

Risolvendo ora esplicitamente l'equazione (3.8) si ottiene

$$\hat{\boldsymbol{\beta}}(t|t) = \bar{L}_{\boldsymbol{\beta}} \sum_{t_0}^t \mathbf{e}(s) \quad (3.11)$$

(qui $t_0 \rightarrow -\infty$) e la legge di controllo (3.7) può essere riscritta nella forma

$$\mathbf{u}(t) = -K\bar{L}_{\boldsymbol{\beta}} \sum_{t_0}^t \mathbf{e}(s) - G\hat{\mathbf{x}}(t|t) \quad (3.12)$$

Come si vede, questa legge di controllo contiene una *azione integrale* (dell'errore) la quale, come è ben noto dalla teoria dei controlli automatici, ha proprio lo scopo

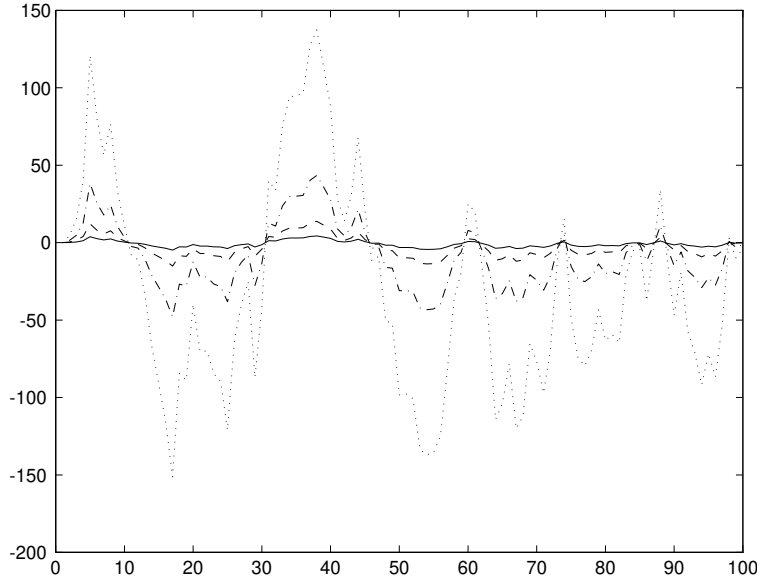


FIGURA 3.1. Passeggiata aleatoria corrispondente alla stesso sequenza di rumore bianco con varianza 1 (tratto continuo), 10 (tratteggiato), 100 (punto linea), e 1000 (punto punto).

di ridurre a zero l'errore a regime dovuto a disturbi costanti (si noti che per la (3.2) $\beta(t)$ è “costante” a meno della fluttuazione aleatoria $\mathbf{n}(t)$).

Nella teoria classica dei controlli il guadagno dell'azione integrale, K_I , viene aggiustato usando regole semiempiriche. In questo caso si vede che $K_I = K \bar{L}_\beta$ dipende in modo noto dai parametri del modello ed in particolare dalle varianze dei processi bianchi $\{\mathbf{n}(t)\}$ e $\{\mathbf{v}(t)\}$ con cui si descrive il disturbo. L'aggiustamento di K_I (ovvero del guadagno \bar{L}_β) viene effettuato in pratica *agendo sulle varianze dei rumori bianchi nel modello del disturbo*. Si ricordi che è proprio la scelta di questi parametri, e in particolare la varianza di $\{\mathbf{n}(t)\}$, che determina la velocità di variazione della componente lenta $\beta(t)$ e quindi la dinamica della stima $\hat{\beta}(t | t)$ prodotta dal filtro di Kalman.

In pratica, dato che il modello (3.4)–(3.2) è solo una descrizione imposta in modo aprioristico della dinamica di $\mathbf{d}(t)$, queste varianze devono essere scelte “sperimentalmente” (cioè mediante simulazioni) e per quanto detto in precedenza il criterio ovvio da seguire è quello di aggiustare $\text{Var } \mathbf{n}(t) := Q_n$ e $\text{Var } \mathbf{v}(t) := Q$ in modo tale che la differenza tra il disturbo simulato $\mathbf{d}(t)$ e la sua stima $\hat{\beta}(t | t)$ sia il più possibile assimilabile a un rumore bianco. In questo modo si raggiunge

¹La versione deterministica di questo problema, chiamata in inglese *Disturbance Decoupling Problem*, (DDP), assume completa ignoranza del disturbo e cerca di determinare G in modo tale che il corrispondente sistema in catena chiusa abbia funzione di trasferimento disturbo-uscita uguale a zero [13, 157]. Le condizioni per cui questo avviene si ritrovano verificate raramente in pratica.

automaticamente lo scopo di aggiustare il guadagno K_I in modo da eliminare il massimo di componente lenta dal disturbo agente.

Accordo del filtro di Kalman. Possiamo a questo punto cercare di estrarre dalla discussione precedente una formulazione precisa del problema che ci interessa. Si abbiano delle misure m -dimensionali $\{\mathbf{y}(t)\}$ che sono la sovrapposizione di un processo messaggio $\{\mathbf{z}_0(t)\}$ e di un rumore bianco di misura $\{\mathbf{w}(t)\}$ di varianza nota $R = R' > 0$

$$\mathbf{y}(t) = \mathbf{z}_0(t) + \mathbf{w}(t) \quad ; \quad (3.13)$$

supporremo che $\{\mathbf{z}_0(t)\}$ e $\{\mathbf{w}(t)\}$ siano scorrelati. Il processo messaggio viene descritto mediante un modello *approssimato* del tipo

$$\begin{aligned} \mathbf{x}(t+1) &= A\mathbf{x}(t) + \mathbf{v}(t) \\ \mathbf{z}(t) &= C\mathbf{x}(t) \end{aligned} \quad (3.14)$$

in cui le matrici A e C sono fissate e la varianza $Q = Q' \geq 0$ del rumore bianco $\{\mathbf{v}(t)\}$ è un *parametro variabile*. Il modello (3.14) viene usato per stimare il segnale messaggio $\{\mathbf{z}_0(t)\}$ a partire dalle misure (3.13).

Tanto per fissare le idee supponiamo di voler costruire il predittore a regime $\hat{\mathbf{z}}_0(t+1 | t)$ usando un filtro di Kalman a regime basato sul modello approssimato (3.14) (naturalmente supponiamo che il filtro a regime per il modello (3.14) esista). Dato che $\{\mathbf{z}(t)\}$ è solo una approssimazione del segnale vero $\{\mathbf{z}_0(t)\}$, ci aspettiamo che il *predittore approssimato* basato sul modello (3.14), diciamolo $\hat{\mathbf{z}}_Q(t+1 | t)$, sia diverso da $\hat{\mathbf{z}}_0(t+1 | t)$. Nasce quindi in modo ovvio il seguente problema.

PROBLEMA 11.2. *Trovare la matrice varianza Q tale per cui la varianza della differenza $\hat{\mathbf{z}}_0(t+1 | t) - \hat{\mathbf{z}}_Q(t+1 | t)$ tra il predittore vero e quello approssimato sia minima.*

Diciamo subito che non c'è da aspettarsi di riuscire a dare una soluzione esplicita soddisfacente a questo problema, perché una qualunque trattazione analitica richiede necessariamente di introdurre una descrizione matematica di $\{\mathbf{z}_0(t)\}$ e quindi delle ipotesi non verificabili su un ipotetico “modello vero” che lo generi.

Questo d'altra parte potrebbe essere plausibile solo nel caso assai particolare in cui si usano modelli di ordine ridotto, ma il caso che ci interessa di più in questo momento è proprio quello in cui si hanno informazioni a priori su $\{\mathbf{z}_0(t)\}$ che possono essere anche un pò vaghe (si vedano gli esempi precedentemente riportati). Tanto più che \mathbf{z}_0 non è mai accessibile direttamente ma si può misurare solo tramite un sensore rumoroso.

Si può comunque arrivare a una regola pratica di procedere ragionando nel modo seguente.

Innanzitutto notiamo che dall'ortogonalità dei due segnali in (3.13) segue immediatamente che

$$\hat{\mathbf{z}}_0(t+1 | t) = \hat{\mathbf{y}}(t+1 | t) \quad , \quad (3.15)$$

per cui il Problema 12.1 potrebbe equivalentemente essere riproposto come ricerca dell'approssimazione ottima di $\hat{\mathbf{y}}(t+1 | t)$. Definiamo poi il processo degli *errori di predizione* di $\mathbf{y}(t)$

$$\boldsymbol{\varepsilon}_Q(t) := \mathbf{y}(t) - \hat{\mathbf{z}}_Q(t | t-1) \quad ; \quad (3.16)$$

si ha allora il seguente fatto fondamentale.

TEOREMA 11.2. *Se per qualche valore di $Q = Q' \geq 0$ il processo $\{\varepsilon_Q(t)\}$ è bianco, allora il corrispondente predittore basato sul modello (3.14) coincide con $\hat{\mathbf{y}}(t | t - 1)$.*

DIMOSTRAZIONE. Per semplicità di notazioni non indicheremo esplicitamente la dipendenza da Q delle varie grandezze in gioco. Il predittore di Kalman basato sul modello (3.14) ha la forma

$$\begin{aligned}\hat{\mathbf{x}}(t + 1 | t) &= A \hat{\mathbf{x}}(t | t - 1) + K(t) \varepsilon(t) \\ \hat{\mathbf{z}}(t + 1 | t) &= C \hat{\mathbf{x}}(t + 1 | t) \\ \hat{\mathbf{x}}(t_0 | t_0 - 1) &= E \mathbf{x}_0 = \mu_0\end{aligned}$$

e da queste equazioni scende immediatamente che $\hat{\mathbf{z}}(t + 1 | t)$ è funzione di $\varepsilon(t_0), \dots, \varepsilon(t)$ ovvero

$$\hat{\mathbf{z}}(t + 1 | t) \in \mathbf{H}(\varepsilon^t) \quad . \quad (3.17)$$

Analogamente, impiegando la definizione (3.16) di $\{\varepsilon(t)\}$ e sostituendo nell'equazione del predittore si vede che

$$\hat{\mathbf{z}}(t + 1 | t) \in \mathbf{H}(\mathbf{y}^t) \quad . \quad (3.18)$$

Riscrivendo la (3.16) come $\mathbf{y}(t) = \varepsilon(t) + \hat{\mathbf{z}}(t | t - 1)$ scende allora dalla (3.17) che

$$\mathbf{y}(t) \in \mathbf{H}(\varepsilon^t) \quad . \quad (3.19)$$

Dualmente, la $\varepsilon(t) = \mathbf{y}(t) - \hat{\mathbf{z}}(t | t - 1)$ implica, in base alla (3.18), che

$$\varepsilon(t) \in \mathbf{H}(\mathbf{y}^t) \quad . \quad (3.20)$$

Le (3.19) e (3.20) insieme permettono così di affermare che

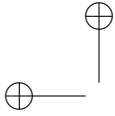
$$\mathbf{H}(\mathbf{y}^t) = \mathbf{H}(\varepsilon^t) \quad , \quad (3.21)$$

ovvero $\{\mathbf{y}(t)\}$ ed $\{\varepsilon(t)\}$ sono legati fra loro da una trasformazione lineare causale e causalmente invertibile. Supponiamo ora che l'errore di predizione sia un processo bianco. Evidentemente la (3.21) implica $\varepsilon(t) \perp \mathbf{H}(\mathbf{y}^{t-1})$ e la decomposizione $\mathbf{y}(t) = \hat{\mathbf{z}}(t | t - 1) + \varepsilon(t)$ è una somma di due termini ortogonali in cui $\hat{\mathbf{z}}(t | t - 1) \in \mathbf{H}(\mathbf{y}^{t-1})$. Per il teorema delle proiezioni $\hat{\mathbf{z}}(t | t - 1)$ deve allora necessariamente coincidere con $\hat{\mathbf{y}}(t | t - 1)$. \square

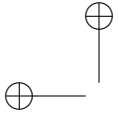
OSSERVAZIONE 11.2. È importante notare che il Teorema 11.2 esprime una condizione molto generale e la sua dimostrazione continua a valere pari pari anche nel caso in cui il parametro che individua il generico modello nella classe (3.14) non è necessariamente la varianza del rumore di modello. L'enunciato vale infatti anche nel caso di una classe arbitraria di modelli di dimensione finita, ad esempio una classe parametrica del tipo $A = A(\theta)$, $C = C(\theta)$, $Q = Q(\theta)$, in cui θ può addirittura avere dimensione variabile. \diamond

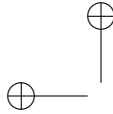
Notiamo che la scelta di un valore di Q (o più in generale del parametro θ) che rende bianco il processo degli errori di predizione ne rende anche automaticamente *minima la varianza*. In effetti se $\{\varepsilon(t)\}$ è bianco allora esso coincide necessariamente con l'errore commesso usando lo stimatore lineare a minima varianza (d'errore) di $\{\mathbf{y}(t)\}$ (cioè con l'innovazione di $\{\mathbf{y}(t)\}$).

In pratica il Teorema 11.2 asserisce che *la bianchezza dell'errore di predizione è una condizione sufficiente* (e ovviamente anche necessaria) *per l'ottimalità del filtro*.



Il modello approssimato va pertanto “accordato” variando Q in modo tale che l’errore di predizione $\{\varepsilon(t)\}$ risulti il più possibile assimilabile a un processo di rumore bianco. Per eseguire concretamente questa operazione di “accordo” occorre avere a disposizione dei criteri statistici che permettono di decidere se una sequenza finita di dati $\{\varepsilon(t) ; t = 1, \dots, T\}$ (ottenuti sperimentalmente mediante simulazione) sia ragionevolmente interpretabile come un tratto di realizzazione di un processo di rumore bianco. Questi procedimenti sono chiamati “tests di bianchezza” e sono ampiamente discussi nella letteratura statistica. Alcuni tra quelli di impiego più frequente sono illustrati ad esempio nel testo [132]. Riferiremo qui solamente il fatto che le informazioni più utili per giungere rapidamente all’accordo del filtro di Kalman si ottengono probabilmente mediante il cosiddetto *test del periodogramma cumulato* di Bartlett [12].





APPENDICE A

RICHIAMI DI ALGEBRA DELLE MATRICI

Il Lemma di inversione. Il risultato seguente, detto *Lemma di inversione di matrice* è di grande utilità in molte manipolazioni di formule matriciali.

LEMMA A.1. *Siano A e C matrici quadrate invertibili. Si ha*

$$[A + BCD]^{-1} = A^{-1} - A^{-1}B [C^{-1} + DA^{-1}B]^{-1} DA^{-1} \quad (0.1)$$

DIMOSTRAZIONE. Si tratta di una verifica:

$$\begin{aligned} [A + BCD] \{A^{-1} - A^{-1}B [C^{-1} + DA^{-1}B]^{-1} DA^{-1}\} &= \\ I + BCDA^{-1} - [B + BCDA^{-1}B] [C^{-1} + DA^{-1}B]^{-1} DA^{-1} &= \\ I + BCDA^{-1} - BC [C^{-1} + DA^{-1}B] [C^{-1} + DA^{-1}B]^{-1} DA^{-1} &= I. \end{aligned}$$

□

Una applicazione classica della formula che ne mette in evidenza l'utilità, è nel calcolo dell'inversa di una matrice soggetta ad una "perturbazione di rango uno". Supponiamo di conoscere già A^{-1} e di voler calcolare l'inversa di $A + b\gamma d'$, dove γ è uno scalare e b e d sono vettori colonna di dimensione n . La (0.1) porge

$$[A + b\gamma d']^{-1} = A^{-1} - A^{-1}b \frac{1}{\gamma^{-1} + d'A^{-1}b} d'A^{-1}$$

da cui si vede che l'inversa può essere aggiornata semplicemente eseguendo delle moltiplicazioni e calcolando il reciproco di uno scalare.

Un'altra formula notevole, simile a quella appena vista, è quella dell' *inversione a blocchi* di una matrice che descriviamo nel lemma seguente.

LEMMA A.2. *Siano A e D matrici quadrate; se A è invertibile, $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ è invertibile se e solo se lo è $D - CA^{-1}B$ e vale la relazione*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix} (D - CA^{-1}B)^{-1} [-CA^{-1} \ I]. \quad (0.2)$$

La matrice $D - CA^{-1}B$ si chiama complemento di Schur di A .

DIMOSTRAZIONE. La formula si ottiene calcolando l'inversa dalla diagonalizzazione a blocchi

$$\begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix} = \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}.$$

□

In realtà il lemma A.1 può anche essere visto come un corollario della formula di inversione a blocchi.

La Decomposizione ai Valori Singolari (SVD). Richiameremo un risultato di algebra delle matrici che, nonostante sia estremamente utile, spesso non viene insegnato nei corsi di base. Si tratta della cosiddetta *Decomposizione ai Valori Singolari (SVD)* di una matrice.

TEOREMA A.1. *Sia $A \in \mathbb{R}^{m \times p}$ una matrice di rango $n \leq \min(m, p)$. Esistono due matrici ortogonali $U \in \mathbb{R}^{m \times m}$ e $V \in \mathbb{R}^{p \times p}$ e una successione ordinata di numeri reali positivi $\{\sigma_1 \geq \dots \geq \sigma_n\}$, detti valori singolari di A , tali che*

$$A = U\Delta V' \quad (0.3)$$

dove Δ ha la struttura quasi diagonale:

$$\Delta = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma = \text{diag} \{ \sigma_1, \dots, \sigma_n \} \quad (0.4)$$

La matrice $U = [u_1, \dots, u_m]$ si può costruire prendendo come colonne gli autovettori normalizzati di AA' ; analogamente, $V := [v_1, \dots, v_p]$ si può costruire prendendo come colonne gli autovettori normalizzati di $A'A$. I quadrati dei valori singolari $\{\sigma_1^2 \geq \dots \geq \sigma_n^2\}$ sono gli autovalori non nulli di AA' (o di $A'A$).

DIMOSTRAZIONE. Siano $[v_1, \dots, v_p]$, p autovettori ortonormali di $A'A$ di modo che

$$A'Av_k = \sigma_k^2 v_k \quad k = 1, \dots, n$$

e $A'Av_k = 0$ per $k > n$. Notare che gli ultimi $p - n$ autovettori possono essere scelti in modo sostanzialmente arbitrario. Moltiplicando a sinistra per A si ottiene

$$AA'(Av_k) = \sigma_k^2 (Av_k) \quad k = 1, \dots, n$$

Si verifica che gli autovettori di AA'

$$u_k := \frac{1}{\sigma_k} Av_k \quad k = 1, \dots, n$$

sono ortonormali. Infatti

$$\langle u_k, u_j \rangle = \frac{v_k' A' Av_j}{\sigma_k \sigma_j} = \frac{\sigma_j^2}{\sigma_k \sigma_j} \langle v_k, v_j \rangle = \frac{\sigma_j^2}{\sigma_k \sigma_j} \delta_{kj}$$

Completiamo ora la famiglia $\{u_1, \dots, u_n\}$ con altri $m - n$ (auto)vettori nello spazio nullo di AA' in modo da ottenere una base ortonormale in \mathbb{R}^m . Un semplice calcolo fornisce

$$u_k' Av_j = \frac{v_k' A' Av_j}{\sigma_k} = \frac{\sigma_j^2}{\sigma_k} \langle v_k, v_j \rangle = \frac{\sigma_j^2}{\sigma_k} \delta_{kj}$$

per $k, j \leq n$ e $u_k' Av_j = 0$ altrimenti. Queste relazioni sono equivalenti alla $U'AV = \Delta$ e quindi alla relazione (0.3). \square

La SVD fornisce la descrizione più completa che si conosca della struttura di una trasformazione lineare. Dalla (0.3) si ricava, eliminando i prodotti con i blocchi nulli di Δ , la seguente *fattorizzazione a rango pieno* di A

$$A = [u_1, \dots, u_n] \Sigma [v_1, \dots, v_n]' := U_n \Sigma V_n' \quad (0.5)$$

dove U_n, V_n sono le sottomatrici di U, V ottenute eliminando le ultime $m - n$ e $p - n$ colonne. Notiamo che

$$U_n' U_n = I_n = V_n' V_n$$

Ricordiamo che la norma ℓ^2 di una matrice $A \in \mathbb{R}^{m \times p}$ è definita dalla relazione

$$\|A\|_2 := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

dove $\|x\|$ è l'ordinaria norma Euclidea. La cosiddetta *norma di Frobenius* $\|A\|_F$ è invece la radice quadrata della somma dei quadrati degli elementi, i.e. $\|A\|_F^2 = \sum_{i,j} a_{i,j}^2 = \text{Tr } AA' = \text{Tr } A'A$.

COROLLARIO A.1. *Lo spazio immagine e lo spazio nullo di A sono dati rispettivamente da:*

$$\text{Im}(A) = \text{Im}(U_n), \quad \ker(A) = \text{Im}([v_{n+1}, \dots, v_p])$$

Inoltre,

$$\|A\|_2 = \|\Sigma\|_2 = \sigma_1, \quad \|A\|_F^2 = \|\Sigma\|_F^2 = \sigma_1^2 + \dots + \sigma_n^2 \quad (0.6)$$

La matrice

$$A_k := \sum_{i=1}^k \sigma_i u_i v_i', \quad k \leq n \quad (0.7)$$

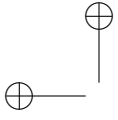
è la miglior approssimante di rango k di A ; infatti

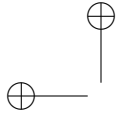
$$\min_{\text{rango}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1} \quad (0.8)$$

e inoltre

$$\min_{\text{rango}(B)=k} \|A - B\|_F^2 = \|A - A_k\|_F^2 = \sigma_{k+1}^2 + \dots + \sigma_n^2 \quad (0.9)$$

La dimostrazione di queste proprietà si può trovare ad esempio nel testo [58, p. 584].





APPENDICE B

RICHIAMI SUGLI SPAZI DI HILBERT

Spazi di Hilbert. Assumeremo che il lettore abbia una qualche conoscenza di spazi vettoriali con prodotto interno. In questa appendice richiameremo semplicemente i fatti essenziali. Se v è un vettore di \mathbb{R}^n o di \mathbb{C}^n , $|v|$ denoterà di solito la sua norma Euclidea (ovviamente uguale al valore assoluto se v è uno scalare).

Ricordiamo che un *prodotto scalare* (o *prodotto interno*) su uno spazio vettoriale \mathbf{V} è una funzione

$$\langle \cdot, \cdot \rangle : \mathbf{V} \times \mathbf{V} \rightarrow \mathbb{C}$$

che è

- (1) lineare nel primo argomento

$$\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle \quad x, y, z \in \mathbf{V}$$

- (2) antisimmetrica

$$\langle y, x \rangle = \overline{\langle x, y \rangle}$$

dove la barra denota complesso coniugato,

- (3) positiva, nel senso che

$$\|x\|^2 := \langle x, x \rangle > 0 \quad \text{per } x \neq 0$$

La quantità $\|x\|$ è chiamata la *norma* indotta dal prodotto scalare $\langle \cdot, \cdot \rangle$. Ogni prodotto scalare soddisfa la *disuguaglianza di Schwartz*

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

È facile verificare che $\|x\|$ soddisfa tutti gli assiomi di una norma e in particolare la disuguaglianza triangolare

$$\|x + y\| \leq \|x\| + \|y\| \quad x, y \in \mathbf{V}.$$

Uno *spazio di Hilbert* è uno spazio con prodotto interno $(\mathbf{H}, \langle \cdot, \cdot \rangle)$ che è completo rispetto alla metrica indotta dal prodotto interno. In altre parole ogni successione di Cauchy ha limite in \mathbf{H} . Esempi di spazi di Hilbert che sono usati di frequente in questo libro sono:

- (1) *Lo spazio Euclideo n -dimensionale*, ovvero \mathbb{R}^n dotato dell'usuale prodotto interno Euclideo, $\langle x, y \rangle := \sum_{t=1}^n x_t y_t$.
- (2) *Lo spazio dei segnali m -dimensionali a quadrato sommabile*, ℓ_m^2 . Gli elementi di questo spazio sono successioni reali o complesse m -dimensionali $x = \{x(t)\}_{t \in \mathbb{Z}}$ dove $x(t)$ è scritto per convenienza come un *vettore riga*, indicate dal parametro intero t , tali che

$$\|x\|^2 := \sum_{t=-\infty}^{+\infty} x(t) x(t)^* < \infty$$

(* denota complesso-coniugato trasposto). Questa norma è talvolta chiamata “energia” del segnale x . Essa è indotta dal prodotto scalare

$$\langle x, y \rangle := \sum_{t=-\infty}^{+\infty} x(t) y(t)^*$$

Una dimostrazione del fatto che ℓ_m^2 è completo si trova per esempio nel libro [160, p. 21].

- (3) *Lo spazio delle variabili del second'ordine di un esperimento aleatorio* $\{\Omega, \mathcal{A}, P\}$.

Una variabile aleatoria reale, f , è una funzione reale misurabile definita sullo spazio di probabilità $\{\Omega, \mathcal{A}, P\}$ (P è la misura di probabilità su Ω e \mathcal{A} la σ -algebra¹ degli eventi). Il simbolo $E\{f\}$ denota l'integrale $\int_{\Omega} f dP$, ed è chiamato *l'aspettazione* della variabile f . Consideriamo l'insieme, $L^2(\Omega, \mathcal{A}, P)$, delle variabili aleatorie (reali o complesse) f definite sullo stesso spazio, $\{\Omega, \mathcal{A}, P\}$, che hanno momento del secondo ordine finito, $E\{|f|^2\} < \infty$. Queste variabili sono chiamate *variabili del secondo ordine* dello spazio di probabilità.

Lo spazio $L^2(\Omega, \mathcal{A}, P)$ è ovviamente uno spazio vettoriale reale (o complesso) rispetto alle operazioni usuali di somma e moltiplicazione per un numero reale (o complesso). Esso è dotato del prodotto scalare

$$\langle \xi, \eta \rangle = E\xi\bar{\eta},$$

che è la *correlazione* delle due variabili ξ, η . Da notare che la norma $\|\xi\| = \langle \xi, \xi \rangle^{1/2}$ indotta da questo prodotto scalare (la radice quadrata del momento del secondo ordine di ξ) è positiva, i.e. $\|\xi\| = 0 \Leftrightarrow \xi = 0$, se si conviene di identificare variabili casuali che differiscono tra di loro al più in un insieme di probabilità zero. La convergenza rispetto a questa norma si chiama *convergenza in media quadratica*. È un fatto fondamentale dell'analisi che $L^2(\Omega, \mathcal{A}, P)$ è chiuso rispetto alla convergenza in media quadratica [125] ed è quindi uno spazio di Hilbert.

NOTA: Lo spazio di probabilità relativo all'esperimento aleatorio su cui sono definite le variabili, non verrà mai specificato e gli eventi elementari di Ω non verranno mai menzionati esplicitamente in questo libro. Noi riserveremo il simbolo ω per denotare la *frequenza (angolare)*, che è una variabile corrente sull'intervallo $[-\pi, \pi]$ (tempo discreto) oppure $(-\infty, +\infty)$ (tempo continuo).

- (4) *Lo spazio di Lebesgue* $L_m^2([a, b], d\mu)$.

Sia $[a, b]$ un intervallo (non necessariamente limitato) della retta reale. Denoteremo con $L_m^2([a, b], d\mu)$ lo spazio delle funzioni m -dimensionali (reali o complesse) a quadrato integrabile sull'intervallo $[a, b]$ rispetto ad una misura μ non necessariamente finita (i.e. $\mu([a, b]) \leq \infty$). In particolare se μ è la misura di Lebesgue, scriveremo semplicemente $L_m^2([a, b])$. Anche in questo caso L_m^2 è uno spazio di Hilbert rispetto al prodotto scalare

$$\langle f, g \rangle := \int_a^b f(t) g(t)^* d\mu \quad .$$

¹Il concetto di σ -algebra e in generale la teoria della probabilità saranno usate solo in modo superficiale in questo testo.

- (5) Lo spazio di Hilbert $\mathbf{H}(\mathbf{y})$ è generato linearmente da un processo del secondo ordine \mathbf{y} . Questo spazio è un sottospazio di $L^2(\Omega, \mathcal{A}, P)$ (lo spazio di probabilità è quello dove è definito il processo) che è stato definito e studiato nel capitolo 3.

Nel caso di segnali scalari, gli spazi ℓ^2 e L^2 verranno indicati senza specificare il pedice. Altri esempi di spazi di Hilbert verranno introdotti più avanti.

In generale il termine *sottospazio* di uno spazio di Hilbert significherà *sottospazio chiuso*. Il *sottospazio generato linearmente da una famiglia di elementi* $\{x_\alpha\}_{\alpha \in \mathbb{A}} \subset \mathbf{H}$, denotato $\overline{\text{span}}\{x_\alpha | \alpha \in \mathbb{A}\}$, è la chiusura dello spazio vettoriale generato dalla famiglia. Quest'ultimo a sua volta consiste di tutte le combinazioni lineari finite degli elementi x_α ed è denotato $\text{span}\{x_\alpha | \alpha \in \mathbb{A}\}$.

La *somma vettoriale* di due sottospazi \mathbf{X} , \mathbf{Y} , è la chiusura dell'insieme $\{x + y | x \in \mathbf{X}, y \in \mathbf{Y}\}$. In questo testo i simboli \vee , $+$ e \oplus denoteranno somma vettoriale, somma vettoriale *diretta* e somma vettoriale *ortogonale*. Il simbolo \mathbf{X}^\perp denota il complemento ortogonale del sottospazio \mathbf{X} in qualche spazio ambiente predefinito.

Esempi importanti di sottospazi di ℓ_p^2 sono gli spazi delle successioni *causali*, ℓ_p^{2+} che sono nulle per valori negativi dell'indice ($f(t) = 0, t < 0$) e quello dei segnali *anticausali* ℓ_p^{2-} che sono viceversa nulli per valori positivi dell'indice ($f(t) = 0, t > 0$). Questi due sottospazi hanno ovviamente intersezione non nulla. Il complemento ortogonale, $\ell_p^{2+\perp}$, di ℓ_p^{2+} in ℓ_p^2 è lo spazio dei segnali *strettamente anticausali* che sono nulli anche per $t = 0$. Ovviamente si ha la decomposizione in somma diretta ortogonale

$$\ell_p^2 = \ell_p^{2+} \oplus \ell_p^{2+\perp}. \quad (0.1)$$

Un *operatore lineare* T da uno spazio di Hilbert \mathbf{H}_1 in \mathbf{H}_2 , è semplicemente una mappa lineare tra i due spazi. È da notare però che in generale T potrebbe non essere definita su tutto \mathbf{H}_1 (si pensi ad esempio all'operatore di derivazione in L_m^2). Se viceversa questo è il caso, diremo che T è definito *su* \mathbf{H}_1 . I più semplici operatori da maneggiare sono quelli *continui*, chiamati spesso anche *limitati*, che sono definiti su tutto lo spazio e soddisfano ad una disuguaglianza del tipo

$$\|Tx\|_2 \leq k\|x\|_1, \quad x \in \mathbf{H}_1$$

per una qualche costante k . I pedici si riferiscono alle norme nei relativi spazi. Come si vede un operatore continuo è in realtà uniformemente continuo. L'estremo superiore dei valori di k per cui vale questa disuguaglianza, si chiama *norma* dell'operatore T e si denota col simbolo $\|T\|$. Si ha così

$$\|T\| = \sup_{\|x\|_1 \neq 0} \frac{\|Tx\|_2}{\|x\|_1} \quad (0.2)$$

Consideriamo un operatore lineare T tra spazi di dimensioni finita. Assumiamo che gli elementi dei due spazi siano rappresentati rispetto a due basi ortonormali, per cui la norma di elementi di questi spazi è la norma Euclidea delle loro coordinate. L'operatore sia rappresentato dalla matrice A . Allora le norme di T e di A (come operatore tra i due spazi Euclidei delle coordinate) coincidono e si ha

$$\|T\| = \|A\|_2 := [\lambda_{MAX}(A' A)]^{1/2} \quad (0.3)$$

dove λ_{MAX} sta per autovalore massimo. Il pedice $_2$ è una notazione convenzionale per indicare la *norma indotta* dalle norme Euclidee. Si chiama qualche volta anche *norma* ℓ^2 .

Si dimostra che esiste un unico operatore lineare e continuo $T^* : \mathbf{H}_2 \rightarrow \mathbf{H}_1$, che soddisfa l'equazione

$$\langle Tx, z \rangle_2 = \langle x, T^*z \rangle_1 \quad \forall x \in \mathbf{H}_1, z \in \mathbf{H}_2$$

L'operatore T^* si chiama *l'aggiunto* di T . Un operatore (da \mathbf{H} in sè stesso) per cui $T^* = T$ si chiama *autoaggiunto*. È ovvio che il concetto di aggiunto corrisponde alla nozione di matrice trasposta in algebra lineare.

Un operatore lineare T che preserva la norma, i.e.

$$\|Tx\|_2 = \|x\|_1, \quad \forall x \in \mathbf{H}_1$$

si chiama *isometria*.

TEOREMA B.1. *Ogni operatore isometrico definito su una famiglia di elementi $\{x_\alpha \mid \alpha \in \mathbb{A}\} \subset \mathbf{H}$ può essere esteso per continuità a tutto lo spazio di Hilbert $\overline{\text{span}}\{x_\alpha \mid \alpha \in \mathbb{A}\}$ generato linearmente dagli $\{x_\alpha\}$, preservandone la proprietà di isometria. L'estensione isometrica è unica.*

La dimostrazione di questo importante risultato è basata (vedi [126, p.14-15]) sul seguente fatto: un qualunque operatore che preserva la norma, definito su una famiglia di elementi $\{x_\alpha\}_{\alpha \in \mathbb{A}} \subset \mathbf{H}$, è automaticamente lineare e può essere esteso per linearità allo spazio vettoriale $\text{span}\{x_\alpha \mid \alpha \in \mathbb{A}\}$ come operatore isometrico. \diamond

È facile mostrare che un'isometria è sempre un operatore iniettivo (1:1) e in base alla definizione di operatore aggiunto, si vede che $\langle x, T^*Tx \rangle_1 = \langle x, x \rangle_1$, per cui $T^*T = I$ (l'identità). Se T è anche suriettivo ($T\mathbf{H}_1 = \mathbf{H}_2$) si trova allora che

$$T^* = T^{-1}.$$

Un operatore che gode di questa proprietà si chiama *unitario*. Due operatori lineari $A : \mathbf{H}_1 \rightarrow \mathbf{H}_1$ e $B : \mathbf{H}_2 \rightarrow \mathbf{H}_2$ che siano legati dalla relazione,

$$A = T^{-1}BT$$

dove T è unitario, si chiamano *unitariamente equivalenti*. L'equivalenza unitaria è una relazione di equivalenza che preserva le caratteristiche fondamentali di un operatore, tra cui lo spettro. La *Trasformata di Fourier* che definiremo tra poco è un esempio di operatore unitario.

Un sottospazio \mathbf{X} è invariante per l'operatore T , se $T\mathbf{X} \subset \mathbf{X}$. In questo caso denoteremo con $T|_{\mathbf{X}}$ la *restrizione* di T a \mathbf{X} . Si dice che un sottospazio \mathbf{X} *riduce* un operatore lineare T se è invariante per T e ammette un sottospazio complementare, \mathbf{Y} , i.e.

$$\mathbf{H} = \mathbf{X} + \mathbf{Y}$$

anch'esso invariante. In questo caso T ha, rispetto a questa decomposizione in somma diretta, la rappresentazione matriciale

$$T = \begin{bmatrix} T|_{\mathbf{X}} & 0 \\ 0 & T|_{\mathbf{Y}} \end{bmatrix}$$

LEMMA B.1. *Sia T un operatore lineare su uno spazio di Hilbert \mathbf{H} . Vale l'equivalenza*

$$T\mathbf{X} \subset \mathbf{X} \Leftrightarrow T^*\mathbf{X}^\perp \subset \mathbf{X}^\perp$$

Se T è autoaggiunto, sia \mathbf{X} che \mathbf{X}^\perp riducono T .

DIMOSTRAZIONE. Infatti \mathbf{X} è T -invariante se e solo se $\langle Tx, y \rangle = 0$ per tutti gli $x \in \mathbf{X}$ e $y \in \mathbf{X}^\perp$. Basta allora applicare la definizione di operatore aggiunto. \square

Avremo spesso l'occasione di incontrare serie infinite di elementi ortogonali in uno spazio di Hilbert. Una caratterizzazione semplice ma fondamentale della convergenza di queste serie è la seguente.

LEMMA B.2. *Una serie di elementi ortogonali in uno spazio di Hilbert converge se e solo se è convergente la serie numerica dei quadrati delle norme degli elementi. Simbolicamente, se $x_k \perp x_j, k \neq j$,*

$$\sum_{k=0}^{\infty} x_k < \infty \Leftrightarrow \sum_{k=1}^{\infty} \|x_k\|^2 < \infty \quad (0.4)$$

DIMOSTRAZIONE. Infatti la serie converge se e solo se

$$\left\| \sum_{k=0}^m x_k - \sum_{k=0}^{n-1} x_k \right\| \rightarrow 0$$

per $n, m \rightarrow \infty$. L'affermazione del lemma scende quindi dall'equivalenza

$$\left\| \sum_{k=n}^m x_k \right\|^2 \rightarrow 0 \Leftrightarrow \sum_{k=n}^m \|x_k\|^2 \rightarrow 0$$

per $n, m \rightarrow \infty$. □

COROLLARIO B.1. *Sia $\{e_k; k \in \mathbb{Z}\}$ una successione ortonormale in uno spazio di Hilbert \mathbf{H} e $c := \{c_k; k \in \mathbb{Z}\}$ una successione di numeri complessi. La serie $\sum_{-\infty}^{+\infty} c_k e_k$ converge se e solo se*

$$\sum_{-\infty}^{+\infty} |c_k|^2 < \infty$$

ovvero, se e solo se $c \in \ell^2$.

Ricordiamo che un sistema di elementi ortonormali $\{e_\alpha; \alpha \in \mathbb{A}\}$ in uno spazio di Hilbert \mathbf{H} si dice *completo* se

$$\langle x, e_\alpha \rangle = 0, \quad \forall \alpha \in \mathbb{A} \Rightarrow x = 0$$

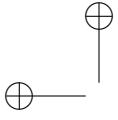
i.e. l'unico elemento di \mathbf{H} che può essere ortogonale a tutti gli $\{e_\alpha; \alpha \in \mathbb{A}\}$ è l'elemento nullo. Un sistema completo numerabile, $\{e_k; k \in \mathbb{Z}\}$, si chiama una *base ortonormale* di \mathbf{H} . Gli spazi di Hilbert che ammettono una base ortonormale si dicono *separabili*.

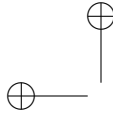
TEOREMA B.2. *In uno spazio di Hilbert separabile \mathbf{H} , ogni elemento x ammette la rappresentazione*

$$x = \sum_{-\infty}^{+\infty} c_k e_k, \quad c_k = \langle x, e_k \rangle \quad (0.5)$$

rispetto alla base ortonormale $\{e_k\}$. La successione $c := \{c_k\}$ dei coefficienti è un segnale a energia finita. Viceversa, per ogni segnale a energia finita c , e base ortonormale $\{e_k\}$, la serie in (0.5) converge ad un elemento di \mathbf{H} .

Questo risultato è talvolta chiamato *teorema di Riesz-Fisher*.





APPENDICE C

TRASFORMATA Z E SPAZI DI HARDY

La Trasformata di Fourier. È un noto risultato fondamentale dell'analisi che le funzioni trigonometriche

$$e_t(\omega) := e^{j\omega t} \quad t \in \mathbb{Z}$$

formano un sistema completo ortonormale (cioè una base ortonormale) nello spazio di Hilbert $L^2([-\pi, \pi], \frac{d\omega}{2\pi})$. Per una dimostrazione si può consultare [160, p. 45].

Da questo fatto scende immediatamente il seguente risultato fondamentale, talvolta noto come Teorema di Fourier-Plancherel.

TEOREMA C.1. *L'operatore lineare (Trasformata di Fourier)*

$$\mathfrak{F} : \ell_m^2 \rightarrow L_m^2([-\pi, \pi], \frac{d\omega}{2\pi}),$$

definito, per ogni $f \in \ell_m^2$, dalla somma

$$\mathfrak{F}(f) := \sum_{t=-\infty}^{+\infty} e^{-j\omega t} f(t) \quad (0.1)$$

che è convergente nella metrica dello spazio $L_m^2([-\pi, \pi], \frac{d\omega}{2\pi})$, è isometrico e suriettivo, ovvero è un operatore unitario da ℓ_m^2 su $L_m^2([-\pi, \pi], \frac{d\omega}{2\pi})$. L'operatore inverso è dato dalla formula

$$\mathfrak{F}^{-1}(\hat{f})(t) := \int_{-\pi}^{\pi} e^{j\omega t} \hat{f}(\omega) \frac{d\omega}{2\pi}, \quad \hat{f} \in L_m^2([-\pi, \pi], \frac{d\omega}{2\pi}).$$

La proprietà di preservare la norma, si traduce nell'eguaglianza tra l'energia del segnale f e l'integrale rispetto alla frequenza $\frac{\omega}{2\pi}$ del modulo quadrato della sua trasformata di Fourier $\hat{f}(\omega) := \mathfrak{F}(f)(\omega)$,

$$\sum_{t=-\infty}^{+\infty} |f(t)|^2 = \int_{-\pi}^{\pi} |\hat{f}(\omega)|^2 \frac{d\omega}{2\pi}$$

Questa relazione è nota anche come *teorema di Parseval*.

Una ragione dell'importanza della trasformata di Fourier nello studio di fenomeni dinamici che riguardano trasformazioni di segnali a energia finita, è il fatto che l'operatore di *traslazione temporale* $\sigma : \ell_m^2 \rightarrow \ell_m^2$, definito dalla

$$\sigma(f)(t) := f(t+1)$$

corrisponde, nel dominio delle trasformate, all'operazione algebrica di moltiplicazione per la funzione esponenziale $e(\omega) : \omega \rightarrow e^{j\omega}$, pensato come operatore lineare in $L_m^2([-\pi, \pi], \frac{d\omega}{2\pi})$. In altre parole

$$\mathfrak{F}(\sigma f) = e(\omega)\mathfrak{F}(f) = e(\omega)\hat{f}(\omega).$$

Su questa proprietà è basato tutto il calcolo operatorio ben noto in teoria dei segnali.

Trasformata Z e Spazi di Hardy

La trasformata di Fourier-Plancherel, \hat{f} , di un segnale a energia finita, è un limite in media quadratica in $L^2([-\pi, \pi], d\omega/2\pi)$. Anche se spesso si dice che \hat{f} è una “funzione” questa affermazione in realtà non è corretta perchè un limite in media può essere modificato in modo arbitrario in un sottoinsieme di misura nulla dell'intervallo $[-\pi, \pi]$ e pertanto, in particolare il valore in un punto, $\hat{f}(\omega)$, può essere arbitrario. Formalmente si può cercare di associare ad \hat{f} una funzione $F(z)$ sul piano complesso $\{z\}$, sostituendo la variabile complessa z a $e^{j\omega}$ nella serie di Fourier di f ,

$$\hat{f}(\omega) = \sum_{t=-\infty}^{+\infty} f(t)e^{-j\omega t} \Rightarrow F(z) = \sum_{t=-\infty}^{+\infty} f(t)z^{-t}$$

In questo modo il secondo membro assume formalmente l'aspetto di una serie di Laurent. $F(z)$ è chiamata la *Z-transformata* (bilatera) del segnale f .

Da notare comunque che questa operazione di cosmesi non muta la sostanza delle cose giacchè la serie di Laurent a secondo membro in generale è garantita convergente (nella metrica di L^2 !) solo sulla circonferenza unitaria del piano complesso e potrebbe benissimo non convergere puntualmente in alcun punto del piano complesso. In questo caso non ha significato pensare a $F(z)$ come una funzione definita punto per punto. In certe circostanze però si può mostrare che \hat{f} deve avere delle proprietà di regolarità e questo permette di ricondurre la classe di equivalenza di \hat{f} ad una funzione. Come vedremo più sotto, la *causalità* del segnale f è una di questa circostanze. La causalità porta ad una Z-trasformata che ha una struttura analitica particolarmente ricca.

DEFINIZIONE C.1. *Lo spazio vettoriale delle funzioni p -dimensionali della variabile complessa z , $F(z)$, che sono analitiche nella regione $\{|z| > 1\}$ del piano complesso e i cui coefficienti $f(k)$ dello sviluppo di Laurent attorno all'infinito sono a quadrato sommabile, i.e. delle funzioni*

$$F(z) = \sum_{k=0}^{+\infty} f(k)z^{-k}, \quad |z| > 1, \quad \sum_{k=0}^{+\infty} |f(k)|^2 < \infty \quad (0.2)$$

è chiamato lo spazio di Hardy p -dimensionale ed è denotato col simbolo H_p^2 . Le funzioni di questo spazio verranno per brevità chiamate analitiche nel seguito.

Dualmente, lo spazio vettoriale delle funzioni p -dimensionali della variabile complessa z , $F(z)$, che sono analitiche nella regione $\{|z| < 1\}$ del piano complesso e i cui coefficienti $f(k)$ dello sviluppo di Taylor attorno all'origine sono a quadrato sommabile, i.e. delle funzioni

$$F(z) = \sum_{k=0}^{+\infty} f(k)z^k, \quad |z| < 1, \quad \sum_{k=0}^{+\infty} |f(k)|^2 < \infty \quad (0.3)$$

è chiamato lo spazio di Hardy coniugato p -dimensionale ed è denotato col simbolo $\overline{H_p^2}$. Le funzioni di questo spazio verranno per brevità chiamate co-analitiche nel seguito.

Le funzioni degli spazi H_p^2 e $\overline{H_p^2}$ verranno chiamate *reali* se i coefficienti nelle serie di Laurent hanno valori in \mathbb{R}^p .

Gli spazi di Hardy hanno proprietà notevoli e sono stati molto studiati. Menzioniamo qui solo alcuni testi facilmente accessibili [65, 64, 41, 51, 160, 108]. La loro importanza nelle scienze applicate scende da un fondamentale teorema di Paley e Wiener che li lega alle trasformate di Fourier di segnali causali a energia finita. Il teorema verrà riportato tra un attimo qui di seguito. Le definizioni che abbiamo dato sono standard e si possono trovare comunemente nei testi a indirizzo ingegneristico. Normalmente i matematici trovano più piacevole esteticamente lavorare all'interno del disco unitario anzichè in $\{|z| > 1\}$ e per questo motivo definiscono la trasformata di Fourier con il fattore $e^{j\omega t}$ in (0.1). La nostra scelta è dettata invece da considerazioni di coerenza con la struttura spettrale dei sistemi dinamici che operano sui segnali e sui processi stocastici di cui ci occupiamo.

Si noti che H_p^2 e $\overline{H_p^2}$ sono spazi vettoriali lineari. Si può introdurre una norma in H_p^2 nel modo seguente. Poniamo $z = \rho e^{j\omega}$ e consideriamo, per $\rho > 1$, la norma L^2 della famiglia di funzioni $F_\rho : \omega \mapsto F(\rho e^{j\omega})$. Si ha

$$\begin{aligned} \|F_\rho\|^2 &:= \int_{-\pi}^{\pi} |F(\rho e^{j\omega})|^2 d\omega / 2\pi = \int_{-\pi}^{\pi} F(\rho e^{j\omega}) F(\rho e^{j\omega})^* d\omega / 2\pi = \\ &= \sum_{k=0}^{+\infty} |f(k)|^2 \rho^{-2k} \quad (0.4) \end{aligned}$$

ed è evidente che il membro a sinistra di questa uguaglianza è una funzione monotona non decrescente di ρ , limitata superiormente dal quadrato della norma l^2 della successione dei coefficienti $f = \{f(k)\}$. Ne segue che il limite

$$\|F\|^2 := \lim_{\rho \downarrow 1} \|F_\rho\|^2$$

esiste ed è proprio uguale a $\|f\|_p^2$. Con questa definizione, la mappa che associa ad F la successione dei suoi coefficienti di Laurent, $F \mapsto f$, è una mappa lineare da H_p^2 su l_p^{2+} che preserva la norma. Con un procedimento simile si può in realtà definire in H_p^2 anche un prodotto interno e in questo modo si introduce in H_p^2 una struttura di spazio di Hilbert, isometrico allo spazio, l_p^{2+} , delle successioni dei coefficienti della serie di Laurent, $f = \{f(k)\}$ delle relative funzioni.

Un argomento completamente analogo porta a introdurre una struttura di spazio di Hilbert sullo spazio coniugato $\overline{H_p^2}$ che lo rende isometrico allo spazio delle successioni dei coefficienti di Taylor delle funzioni co-analitiche. Questo spazio di coefficienti, nelle convenzioni che abbiamo stabilito, si identifica con l_p^{2-} .

Riportiamo di seguito la versione discreta del teorema di Paley-Wiener [111]¹.

TEOREMA C.2. *Supponiamo che $F \in H_p^2$ abbia coefficienti di Laurent $f = \{f(k)\}$ e sia \hat{f} la trasformata di Fourier di f . Allora \hat{f} è il valore al contorno di F sulla circonferenza unità, nel senso che*

$$\lim_{\rho \downarrow 1} F_\rho = \hat{f} \quad (0.5)$$

¹In realtà questo risultato è solo un corollario del teorema originale, che è un risultato molto più generale e non ci interessa direttamente in questo ambito.

sia nella metrica di $L_p^2([-\pi, \pi], d\omega/2\pi)$ che puntualmente quasi ovunque in ω . Più precisamente, $F(z) \rightarrow \hat{f}(\omega)$ quasi ovunque, quando $z \rightarrow e^{j\omega}$ lungo ogni curva regolare in $\{|z| > 1\}$ che intersechi il cerchio unità non-tangenzialmente. Viceversa, la trasformata di Fourier \hat{f} di ogni segnale causale in l^2 può essere estesa per mezzo della formula integrale di Cauchy ad una funzione analitica di H_p^2 , preservandone la norma, i.e. $\|F\| = \|\hat{f}\|$. La corrispondenza $F \leftrightarrow \hat{f}$ è quindi unitaria. In questo senso i due spazi di funzioni possono essere identificati e si può scrivere,

$$\mathfrak{F}(l_p^{2+}) = H_p^2. \quad (0.6)$$

Dualmente, ogni funzione G in $\overline{H_p^2}$ ammette come valore al contorno sulla circonferenza unitaria (sia in L^2 che quasi ovunque in $e^{j\omega}$, lungo curve interne al cerchio unitario e intersecanti la sua frontiera non-tangenzialmente) la trasformata di Fourier \hat{g} dei suoi coefficienti di Taylor $g = \{g(-k)\} \in l_p^{2-}$. La funzione G è univocamente determinata in $\{|z| < 1\}$ dai suoi valori al contorno \hat{g} e la corrispondenza preserva la norma, cosicchè si può scrivere

$$\mathfrak{F}(l_p^{2-}) = \overline{H_p^2}. \quad (0.7)$$

La dimostrazione di questo teorema si può trovare in [65] o in [51, p.172].

Dato che la trasformata di Fourier di un segnale causale in l^2 può essere identificata con la sua estensione analitica di classe H^2 (e dualmente per i segnali anticausali), è pratica comune di trattare i due concetti (la trasformata di Fourier e la sua estensione analitica) come se fossero la stessa cosa. In questo spirito lo spazio H_p^2 viene identificato con il sottospazio di $L_p^2([-\pi, \pi], d\omega/2\pi)$ delle trasformate di Fourier dei segnali causali, i.e.

$$H_p^2 = \{\hat{f} \in L_p^2([-\pi, \pi], d\omega/2\pi) \mid \int_{-\pi}^{\pi} e^{j\omega t} \hat{f}(e^{j\omega}) d\omega/2\pi = 0, t < 0\}.$$

Naturalmente lo spazio coniugato $\overline{H_p^2}$ può essere descritto in modo duale.

NOTAZIONI: In seguito sarà più conveniente pensare le trasformate di Fourier di segnali l^2 come funzioni definite sulla circonferenza unità del piano complesso e denotare la dipendenza di \hat{f} da ω come $\hat{f}(e^{j\omega})$. L'estensione analitica (quando esiste) è quindi esattamente $\hat{f}(z)$, $z \in \mathbb{C}$. Per uniformarci all'uso della teoria in problemi di fattorizzazione spettrale, penseremo le funzioni di H_p^2 e di $\overline{H_p^2}$ come *vettori riga*. Ricordiamo che se si considerano segnali reali, l'estensione del coniugato Hermitiano $\hat{f}^*(e^{j\omega})$ di $\hat{f}(e^{j\omega})$ è $\hat{f}'(1/z)$. $L_p^2(\mathbb{C})$

Parte causale e proiezioni ortogonali su H^2 . Nella formula per il calcolo del filtro di Wiener (1.6) del Cap. VII, compare l'operazione di calcolo della "parte causale" di una funzione di L^2 . In generale, se $f(\cdot)$ è un segnale a energia finita con trasformata di Fourier \hat{f} , la parte causale di \hat{f} è $[\hat{f}]_+(e^{j\omega}) := \sum_0^{+\infty} f(k) e^{-jk\omega}$. Come si vede, $[\hat{f}]_+$ è semplicemente la *proiezione ortogonale* di $\hat{f} \in L_p^2(\mathbb{C})$ su H^2 .

Il complemento ortogonale di H_p^2 , $H_p^{2\perp}$ è composto di Z-trasformate di segnali a energia finita che sono *strettamente anticausali*, i.e. sono nulli per $t \geq 0$. La proiezione ortogonale sullo spazio $H_p^{2\perp}$ di $F(z)$ si denota

$$[[F(z)]]_- := F(z) - [F(z)]_+$$

Chiaramente questa funzione è analitica dentro il cerchio unitario e nulla nell'origine, $z = 0$. \diamond

DEFINIZIONE C.2. Denoteremo col simbolo $H_{m \times p}^\infty$ lo spazio vettoriale delle funzioni matriciali che sono analitiche e uniformemente limitate nella regione $\{|z| > 1\}$ e con $\overline{H_{m \times p}^\infty}$ quello delle funzioni che sono analitiche e limitate nel cerchio unitario $\{|z| < 1\}$.

La norma L^2 di una funzione limitata in $[-\pi, \pi]$ è ovviamente finita, per cui le righe di una funzione matriciale in $H_{m \times p}^\infty$ appartengono in particolare ad H_p^2 . Segue che ogni $F \in H_{m \times p}^\infty$ (o, rispettivamente allo spazio coniugato) ammette limite quasi ovunque per $z \rightarrow e^{j\omega}$ lungo ogni curva che incrocia la circonferenza unità non-tangenzialmente. I valori al contorno sul cerchio unità di queste funzioni appartengono quindi allo spazio (di Banach) delle funzioni (essenzialmente) limitate sulla circonferenza unità $L_{m \times p}^\infty([-\pi, \pi], d\omega/2\pi)$, i cui coefficienti di Fourier ad indice negativo (o, rispettivamente, positivo) sono nulli.

Se $\hat{f} \in H_m^2$ e $\hat{A} \in H_{m \times p}^\infty$, è ovvio che $\hat{f}\hat{A} \in H_p^2$. Lo spazio $H_{m \times p}^\infty$ gioca il ruolo di spazio degli operatori lineari causali $A : H_m^2 \rightarrow H_p^2$.

Il seguente teorema stabilisce formalmente questo fatto. La dimostrazione è semplice ma verrà saltata per ragioni di spazio, si veda [51, p.]. Ricordiamo che l'operatore di moltiplicazione per la trasformata di Fourier dell'operatore di traslazione all'indietro $\bar{e} : \omega \mapsto e^{-j\omega}$ mappa lo spazio H_p^2 in sè stesso.

TEOREMA C.3 (Bochner-Chandrasekharan). Ogni operatore lineare limitato $A : H_m^2 \rightarrow H_p^2$ che commuta con l'operatore di moltiplicazione per $\bar{e} : \omega \mapsto e^{-j\omega}$ è un operatore di moltiplicazione per una funzione matriciale $m \times p$ -dimensionale \hat{A} in $H_{m \times p}^\infty$.

La proprietà di commutare con $e^{-j\omega}$ è l' *invarianza temporale*. Nel dominio del tempo, questo risultato dice che il più generale operatore lineare tempo-invariante sullo spazio dei seganli causali a energia finita è la convoluzione con una matrice di funzioni causali la cui trasformata di Fourier sta in $H_{m \times p}^\infty$.

DEFINIZIONE C.3. Una funzione matriciale $Q \in H_{p \times p}^\infty$ i cui valori al contorno sul cerchio unitario sono matrici unitarie, i.e.

$$Q(e^{i\omega})Q(e^{i\omega})^* = I \tag{0.8}$$

si chiama *inner*. Dualmente, una funzione matriciale $\bar{Q} \in \overline{H_{p \times p}^\infty}$ i cui valori al contorno sul cerchio unitario soddisfano (0.8) si chiama *inner-coniugata* (o *co-inner*).

Le funzioni inner scalari sono state descritte e classificate completamente, si consulti ad esempio il libro [65]. Si dimostra che una funzione scalare Q è inner se e solo se è della forma $Q(z) = \theta B(z)S(z)$ dove

- θ è una costante di modulo unitario,
- $B(z)$ è un *Prodotto di Blaschke*, i.e. una funzione del tipo

$$B(z) = \prod_{k=1}^{+\infty} \frac{1 - \alpha_k z}{z - \alpha_k}, \quad |\alpha_k| < 1.$$

dove, per rendere il prodotto infinito convergente in $\{|z| > 1\}$, si richiede che $\sum_{k=1}^{+\infty} (1 - |\alpha_k|) < \infty$. Questa condizione allo stesso tempo prescrive la velocità di accumulazione degli zeri di $B(z)$ nel caso in cui $|\alpha_k| \rightarrow 1$

- $S(z)$ è una funzione *inner singolare* che ha la rappresentazione

$$S(z) = \exp\left\{-\int_{-\pi}^{\pi} \frac{z + e^{i\omega}}{z - e^{i\omega}} d\mu_s(\omega)\right\}$$

dove μ_s è una misura finita e positiva sul cerchio unità che è singolare rispetto alla misura di Lebesgue (in altri termini, μ_s è concentrata su un sottoinsieme di misura di Lebesgue zero).

I prodotti di Blaschke finiti sono proprio le funzioni inner razionali. Sfortunatamente, per le funzioni inner nel caso matriciale non ci sono espressioni così semplici ed eleganti. Comunque è immediato vedere che il determinante di una matrice inner dev'essere una funzione inner scalare (per il teorema di Binet). Si possono così classificare le matrici inner basandosi sulle proprietà del determinante. Per maggiori dettagli rimandiamo alla letteratura [64] p. 80-89.

Notiamo che le espressioni analitiche riportate sopra hanno in generale senso anche nella regione interna al cerchio unitario $\{|z| < 1\}$ del piano complesso. Si può in effetti mostrare che le funzioni inner hanno sempre un prolungamento analitico attraverso la circonferenza unità, eccettuati al più un insieme (di misura nulla) di punti singolari. In particolare si può pensare ad una funzione inner come ad una funzione che è definita e analitica quasi ovunque sul piano complesso, eccettuato al più un insieme di punti singolari nella regione $\{|z| \leq 1\}$, che sono proprio *i reciproci dei suoi zeri*, questi tutti localizzati nella regione $\{|z| \geq 1\}$ del piano complesso. È in effetti possibile specificare completamente una funzione inner assegnando i suoi zeri in $\{|z| \geq 1\}$ (obbedendo naturalmente ai vincoli posti dalla convergenza del prodotto di Blaschke e alla condizione sul supporto di μ).

Notiamo che la condizione (0.8), scritta per il prolungamento analitico $Q(z)$ di una funzione inner matriciale diventa

$$Q(z)Q(1/z)' = I.$$

Gli zeri di una funzione inner matriciale Q sono i punti in $\{|z| \geq 1\}$ dove o $\det Q(z)$ si annulla, oppure tende a zero per $z \rightarrow \alpha = e^{i\omega_0}$ non-tangenzialmente in $\{|z| \geq 1\}$.

Chiudiamo questa sezione riportando il celebre teorema di fattorizzazione *inner-outer* di Beurling [65]. La definizione di *funzione a fase minima*² è stata data al Capitolo IV, Def. 3.2.

TEOREMA C.4. *Ogni funzione matriciale $m \times p$ a rango pieno p e righe in H_p^2 , ammette la fattorizzazione*

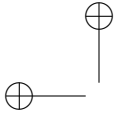
$$F = GQ \tag{0.9}$$

dove G è a fase minima e Q è una matrice inner $p \times p$. In questa fattorizzazione G e Q sono uniche a meno di moltiplicazione a destra per matrici costanti ortogonali $p \times p$.

Il fattore inner Q nella fattorizzazione (0.9) è univocamente determinato dagli zeri della funzione F nella regione $\{|z| > 1\}$ (gli zeri "instabili")³. In particolare, si vede che una funzione a fase minima non può avere zeri in $\{|z| > 1\}$.

²Che in inglese si chiama anche *outer*, cioè esterna.

³Come siano definiti questi zeri è una faccenda che lasceremo per il momento nel vago. Daremo delle definizioni precise nel caso di matrici razionali, nel Capitolo V.



APPENDICE D

ANALISI ARMONICA DEI PROCESSI STAZIONARI

Processi Debolmente Stazionari. Il processo m -dimensionale \mathbf{y} è *stazionario in senso debole*, se la sua matrice di covarianza

$$\Sigma(t, s) := E\{[\mathbf{y}(t) - m(t)][\mathbf{y}(s) - m(s)]^*\}, t, s \in \mathbb{Z} \quad (0.1)$$

è una funzione della differenza dei due argomenti temporali $t - s$, nel qual caso, con un piccolo abuso di notazione scriveremo

$$\Sigma(t, s) = \Sigma(t - s).$$

In seguito la stazionarietà in senso debole verrà senz'altro chiamata "stazionarietà *tout-court*". Dato che la media $m(t)$ si assume nota per tutti i t , essa può essere sottratta dalla variabile $\mathbf{y}(t)$ e, senza perdita di generalità assumeremo che tutti i processi del secondo ordine che considereremo in questa appendice abbiano *media zero*.

Se il process \mathbf{y} è stazionario si può introdurre in $\mathbf{H}(\mathbf{y})$ un operatore lineare isometrico $U : \mathbf{H}(\mathbf{y}) \rightarrow \mathbf{H}(\mathbf{y})$, chiamato *traslazione temporale* o *shift* del processo \mathbf{y} , mediante la posizione

$$U\mathbf{y}_k(t) := \mathbf{y}_k(t+1) \quad k = 1, \dots, m \quad t \in \mathbb{Z}.$$

Più in generale lo shift di t unità di tempo, U_t , $t \in \mathbb{Z}$, è definito dapprima sull'insieme $Y := \{\mathbf{y}_k(t); k = 1, 2, \dots, m, t \in \mathbb{Z}\}$ ponendo

$$U_t\mathbf{y}_k(s) := \mathbf{y}_k(s+t) \quad k = 1, \dots, m \quad (0.2)$$

e poi esteso su tutto $\mathbf{H}(\mathbf{y})$ per linearità e continuità.

PROPOSIZIONE D.1. *La famiglia $\{U_t | t \in \mathbb{Z}\}$ forma un gruppo di operatori unitari sullo spazio di Hilbert $\mathbf{H}(\mathbf{y})$. Si ha $U_t = U^t$ per tutti i $t \in \mathbb{Z}$, U essendo lo shift di un passo in avanti nel tempo.*

DIMOSTRAZIONE. Notiamo che per la stazionarietà si ha

$$\langle U_t\xi, U_t\eta \rangle = \langle \xi, \eta \rangle \quad \xi, \eta \in Y$$

di modo che U_t è isometrico (e quindi *lineare* su Y , [126, p.14-15]). Per il teorema B.1, la mappa lineare U_t può allora essere estesa per continuità come operatore isometrico allo spazio $\mathbf{H}(\mathbf{y})$: Infatti, se $\xi \in \mathbf{H}(\mathbf{y})$ allora esiste una successione $\{\xi_k\}$ in $\text{span}\{\mathbf{y}(t); t \in \mathbb{Z}\}$ che converge a ξ e quindi $\|\xi_j - \xi_k\| = \|U_t\xi_j - U_t\xi_k\| \rightarrow 0$ per $j, k \rightarrow \infty$. Quindi $\{U_t\xi_k\}$ è convergente e si può definire $U_t\xi$ per un arbitrario elemento $\xi \in \mathbf{H}(\mathbf{y})$ come

$$U_t\xi = \lim_{k \rightarrow \infty} U_t\xi_k.$$

Questo mostra che l'estensione è un'isometria definita su tutto $\mathbf{H}(\mathbf{y})$. Inoltre è evidente che ogni elemento di $\mathbf{H}(\mathbf{y})$ della forma $\sum \alpha_{kj}\mathbf{y}_k(t_j)$ è l'immagine attraverso

U_t dell'elemento $\sum \alpha_{kj} \mathbf{y}_k(t_j - t)$. Mediante l'argomento limite usato in precedenza si mostra così che U_t è suriettivo e quindi unitario. Il suo aggiunto, U_t^* , è quindi dato dalla $U_t^* = U_t^{-1}$. In particolare, $U_t^* \mathbf{y}_k(s) = \mathbf{y}_k(s - t)$ per $k = 1, \dots, m$ a ogni $t, s \in \mathbb{Z}$. \square

È ben noto che la covarianza di un processo del secondo ordine è una funzione di *tipo positivo* nel senso che

$$\sum_{k,j=1}^N a_k^* \Sigma(t_k, t_j) a_j \geq 0 \quad (0.3)$$

per tutte le scelte dei tempi t_1, \dots, t_N e per coefficienti $a_k \in \mathbb{C}^m$ arbitrari. La proprietà di positività della matrice di covarianza $\Sigma(\tau)$ di un processo stazionario è caratterizzabile mediante una rappresentazione integrale (il *teorema di Herglotz*) che assomiglia in un certo senso alla trasformata inversa di Fourier. L'enunciato che si trova di solito in letteratura è relativo al caso di processi scalari.

TEOREMA D.1 (Herglotz). *Sia $\sigma(\tau)$, $\tau \in \mathbb{Z}$, la funzione covarianza di un processo stazionario reale a tempo discreto \mathbf{y} . Esiste allora una funzione $\omega \rightarrow F(\omega)$ definita sull'intervallo $[-\pi, \pi]$ che è monotona non-decrescente in ω i.e.*

$$\omega_2 \geq \omega_1 \Rightarrow F(\omega_2) - F(\omega_1) \geq 0$$

tale che

$$\sigma(\tau) = \int_{-\pi}^{\pi} e^{j\omega\tau} dF(\omega), \quad \tau \in \mathbb{Z}. \quad (0.4)$$

La funzione F può essere presa continua da destra; in questo caso essa è univocamente determinata da σ a meno di una costante arbitraria. Si può rendere F unica imponendo che sia $F(0) = 0$. In questo caso, il fatto che il processo sia reale (i.e. che $\sigma(-\tau) = \sigma(\tau)$) implica che F dev'essere una funzione dispari di ω , i.e. $F(-\omega) = -F(\omega)$.

Per la dimostrazione rimandiamo al trattato di Doob [39, p.474].

La funzione monotona F è chiamata la *distribuzione spettrale di potenza* del processo \mathbf{y} . Da notare che

$$\infty > E\{|\mathbf{y}(t)|^2\} = \sigma(0) = \int_{-\pi}^{\pi} dF(\omega) = F(\pi),$$

e quindi la funzione F è limitata ¹. Ricordiamo che per un processo scalare, la quantità $\sigma(0) = E\{|\mathbf{y}(t)|^2\}$ è convenzionalmente chiamata *potenza statistica* del segnale. Il significato di F scende dal fatto che $F(\omega_2) - F(\omega_1)$ è la potenza statistica del segnale contenuta nella banda di frequenza $(\omega_1, \omega_2]$. In particolare se F ha una discontinuità (che può solo essere un salto di ampiezza finita) alla frequenza $\pm\omega_k$, la funzione σ ha una componente oscillatoria della stessa frequenza,

$$\sigma_k(\tau) = \int_{\{\pm\omega_k\}} e^{j\omega\tau} dF(\omega) = \Delta F(+\omega_k) e^{j\omega_k\tau} - \Delta F(-\omega_k) e^{-j\omega_k\tau} = 2\Delta F(+\omega_k) \cos \omega_k\tau$$

dove abbiamo usato la notazione $\Delta F(+\omega_k) := F(\omega_k) - F(\omega_k-)$.

L'estensione al caso vettoriale (reale) si basa sull'applicazione del risultato al processo scalare $a'\mathbf{y}$ dove a è un vettore arbitrario in \mathbb{R}^m .

¹Il che è equivalente a dire che $F(\omega) = F(\omega+) = \mu([-\pi, \omega])$ dove μ è una misura di Borel positiva e finita sull'intervallo $[-\pi, \pi]$.

COROLLARIO D.1. Sia $\Sigma(\tau)$, $\tau \in \mathbb{Z}$, la matrice di covarianza di un processo stazionario reale m -dimensionale a tempo discreto \mathbf{y} . Esiste allora una funzione a valori matriciali $m \times m$, $\omega \rightarrow F(\omega)$ definita sull'intervallo $[-\pi, \pi]$ che è monotona non-decrescente² in ω i.e.

$$\omega_2 \geq \omega_1 \Rightarrow F(\omega_2) - F(\omega_1) \geq 0$$

e tale che

$$\Sigma_{kj}(\tau) = \int_{-\pi}^{\pi} e^{j\omega\tau} dF_{kj}(\omega), \quad \tau \in \mathbb{Z} \quad k, j = 1, \dots, m. \quad (0.5)$$

Se si richiede che F sia continua da destra, allora essa è univocamente determinata da Σ a meno di una matrice costante arbitraria. Il fatto che il processo sia reale (i.e. che $\Sigma(-\tau)' = \Sigma(\tau)$) implica che le differenze $\{F(\omega_2) - F(\omega_1); \omega_2 \geq \omega_1\}$ sono matrici Hermitiane (i.e. $F(\Delta)^* = F(\Delta)$). Se si fissa $F(0) = 0$, allora F dev'essere una funzione dispari di ω , nel senso che $F(\omega)^* = -F(\omega)$.

Come è noto dall'analisi reale, [125], ogni funzione monotona F ammette una decomposizione canonica in parte assolutamente continua e in parte singolare (questo risultato continua a valere anche per funzioni a valori matriciali)

$$F = F_a + F_d + F_s \quad (0.6)$$

in cui F_a è la componente *assolutamente continua*, che è strettamente crescente e ammette derivata in un sottoinsieme di $[-\pi, \pi]$ di misura $d\omega/(2\pi) = 1$ e $F_d + F_s$ costituiscono la *parte singolare di F* , in cui F_d è una funzione costante a tratti, le cui discontinuità sono quindi salti di ampiezza finita (al più un insieme numerabile), e F_s è una *funzione monotona singolare*, continua con derivata nulla quasi ovunque, che è strettamente crescente solo nell'insieme di misura nulla dove la sua derivata F_s' o non esiste oppure vale $+\infty$. La componente assolutamente continua è esprimibile come l'integrale

$$F_a(\omega) = \int_{-\pi}^{\omega} S(\lambda) \frac{d\lambda}{2\pi}$$

di una funzione matriciale $m \times m$ (ci riferiamo qui direttamente al caso di processi vettoriali), $S(\omega)$ si chiama *densità spettrale di potenza*. Questa funzione matriciale ha le seguenti proprietà,

- (1) $S(\omega)$ è Hermitiana, ovvero $S^*(\omega) = S(\omega)$
- (2) $S(\omega)$ è semidefinita positiva, ovvero $S(\omega) \geq 0$, $\forall \omega \in [-\pi, \pi]$
- (3) S è integrabile, ovvero $\int_{-\pi}^{\pi} S(\omega) \frac{d\omega}{2\pi} < \infty$,

che si dimostrano facilmente riportandosi al caso del processo scalare $a'y$ (dove a è un vettore arbitrario in \mathbb{R}^m) di densità spettrale $a'S(\omega)a$. È immediato che, viceversa, l'integrale di una qualunque funzione (eventualmente matriciale) di ω che gode di queste proprietà è una distribuzione spettrale di tipo assolutamente continuo e definisce quindi, attraverso la formula (0.4)

$$\sigma(\tau) = \int_{-\pi}^{\pi} e^{j\omega\tau} S(\omega) \frac{d\omega}{2\pi}, \quad \tau \in \mathbb{Z}. \quad (0.7)$$

una (matrice) covarianza di un processo stazionario. In questo caso la rappresentazione di Herglotz coincide la formula che esprime Σ come l'antitrasformata di Fourier di S .

²Nel senso del solito ordinamento tra matrici. Se A, B sono matrici $m \times m$, in generale a valori complessi, allora $A \geq B \equiv A - B \geq 0$ significa che $A - B$ è semidefinita positiva.

OSSERVAZIONE D.1. Dalla rappresentazione integrale (0.4) si vede che danno contributo alla covarianza solo le frequenze ω , che chiameremo *punti di crescita di F* in cui $F(\omega + \epsilon) - F(\omega - \epsilon) > 0$, $\forall \epsilon > 0$. Questo insieme di frequenze viene talvolta chiamato *spettro del processo \mathbf{y}* . Per definizione, i punti di crescita della parte singolare $F_d + F_s$ formano un insieme $N \subset [\pi, \pi]$ di misura nulla. In particolare, se F_s è assente, N è un insieme di frequenze al più numerabile. Sia $\bar{N} := [-\pi, \pi] - N$, χ_N la funzione caratteristica dell'insieme N e $\chi_{\bar{N}} := 1 - \chi_N$ la sua complementare. Dato che in \bar{N} non vi sono punti di crescita di $F_d + F_s$, la funzione

$$\bar{F}_a(\omega) := \int_{-\pi}^{\omega} \chi_{\bar{N}}(\lambda) dF(\lambda) = \int_{\{\bar{N} \cap [-\pi, \omega]\}} dF(\lambda) = \int_{\{\bar{N} \cap [-\pi, \omega]\}} dF_a(\lambda)$$

è assolutamente continua. È inoltre ovvio, dato che $\chi_{\bar{N}} = 1$ quasi ovunque, che $\bar{F}_a = F_a$. Abbiamo così decomposto lo spettro del processo in due componenti disgiunte, quella *singolare*, N , e quella *continua*, costituita dai punti di crescita di F_a che formano un sottoinsieme di \bar{N} . \diamond

Siano F_a e F_s le parti assolutamente continua e singolare della distribuzione spettrale del processo \mathbf{y} definite nell'Osservazione D.1. Usando la teoria della rappresentazione spettrale (trasformata di Fourier) dei processi stazionari [126] si mostra che il processo \mathbf{y} si può decomporre nella somma di due componenti, diciamo \mathbf{y}_a e \mathbf{y}_s dette *componente assolutamente continua e singolare*, che hanno rispettivamente distribuzioni spettrali F_a e F_s e sono tra loro completamente scorrelate, cioè

$$E\mathbf{y}_a(t)\mathbf{y}_s(\tau)^\top = 0, \quad t, \tau \in \mathbb{Z}.$$

Relazione con la Trasformata di Fourier. Se Σ è la covarianza di un processo a spettro assolutamente continuo, applicando alla (0.4) un classico risultato in analisi reale, il cosiddetto *lemma di Riemann-Lebesgue* (vedere ad es. [127, p. 109], [42, p. 39]), si mostra che

$$\lim_{\tau \rightarrow \pm\infty} \Sigma(\tau) = 0$$

In questo caso si vorrebbe esprimere S come *trasformata di Fourier della covarianza*, in altri termini si desidererebbe che valesse una relazione del tipo

$$S(\omega) = \sum_{\tau=-\infty}^{+\infty} e^{-j\omega\tau} \Sigma(\tau) \quad (?)$$

Portroppo però, senza ipotesi aggiuntive l'eguaglianza *non vale nemmeno come limite in $L^1([-\pi, \pi], d\omega/(2\pi))$* . Il risultato classico in questo contesto è un teorema "alla Fejèr" che assicura la convergenza (in $L^1([-\pi, \pi], d\omega/(2\pi))$) solo delle medie aritmetiche delle ridotte, i.e.

$$S_N(\omega) = \frac{1}{N} \sum_{n=1}^N \left(\sum_{\tau=-n}^{+n} e^{-j\omega\tau} \Sigma(\tau) \right) \xrightarrow{L^1} S(\omega).$$

per $N \rightarrow \infty$. Si veda [42, p. 38]. Da tener presente che senza ipotesi aggiuntive non ha nemmeno senso chiedersi se la serie converga puntualmente. È però noto (e abbastanza facile verificare) che se $\sum_{\tau=-\infty}^{+\infty} |\Sigma(\tau)| < \infty$, la serie di Fourier converge puntualmente alla densità spettrale che risulta allora essere una funzione *continua* di ω ([39, p. 476]). Questo risultato è in realtà la caratterizzazione di una classe

di funzioni continue in $[-\pi, \pi]$ che si chiama *Algebra di Wiener*, si veda ad es. [42, p. 43]. La positività di S qui non gioca alcun ruolo.

Rappresentazione spettrale

Il teorema di Herglotz permette di stabilire un notevole isomorfismo tra gli spazi di Hilbert $\mathbf{H}(\mathbf{y})$ e lo spazio $L^2([-\pi, \pi], dF)$. Supponiamo per il momento che \mathbf{y} sia un processo scalare. Facciamo corrispondere alla variabile del processo al tempo t , $\mathbf{y}(t)$, la funzione $e_t := \omega \rightarrow e^{j\omega t}$. Questa corrispondenza si estende per linearità a tutte le combinazioni lineari finite per cui la mappa

$$I_{\mathbf{y}} : \sum_{\text{finita}} a_k \mathbf{y}(t_k) \rightarrow \sum_{\text{finita}} a_k e^{j\omega t_k}, \quad a_1, \dots, a_k \in \mathbb{R}$$

è una trasformazione lineare tra gli spazi vettoriali $\text{span}\{\mathbf{y}(t); t \in \mathbb{Z}\}$ e lo spazio $\text{span}\{e_t; t \in \mathbb{Z}\}$ dei cosiddetti *polinomi trigonometrici* che vogliamo qui considerare come un sottoinsieme di $L^2([-\pi, \pi], dF)$. Ora, per la formula (0.4), si ha

$$\langle \mathbf{y}(t), \mathbf{y}(s) \rangle_{\mathbf{H}(\mathbf{y})} = \langle e_t, e_s \rangle_{L^2([-\pi, \pi], dF)}$$

per cui $I_{\mathbf{y}}$ è un'isometria definita su uno spazio vettoriale che è denso in $\mathbf{H}(\mathbf{y})$.

Vogliamo ora stabilire che le funzioni trigonometriche $e_t(\omega) := e^{j\omega t}$ $t \in \mathbb{Z}$ oltre a formare un sistema completo ortonormale nello spazio di Hilbert $L^2([-\pi, \pi], \frac{d\omega}{2\pi})$ formano un sistema completo anche in $L^2([-\pi, \pi], dF)$, dove dF è una qualunque funzione distribuzione spettrale sull'intervallo $[-\pi, \pi]$. Più precisamente, dimostriamo che le combinazioni lineari di queste funzioni esponenziali (i.e. i polinomi trigonometrici) sono *dense* in $L^2([-\pi, \pi], dF)$.

LEMMA D.1. *Sia dF una funzione monotona limitata definita sull'intervallo $[-\pi, \pi]$. Data una qualunque funzione $\hat{f} \in L^2([-\pi, \pi], dF)$ esiste una successione $\{\varphi_k\}$ di combinazioni lineari finite della famiglia $\{e_t \mid t \in \mathbb{Z}\}$*

$$\varphi_k(\omega) = \sum_{h=-k}^{+k} a_{kh} e^{j\omega t_h} \tag{0.8}$$

tali che

$$\lim_{k \rightarrow \infty} \|\hat{f} - \varphi_k\|_{L^2([-\pi, \pi], dF)} \rightarrow 0 \tag{0.9}$$

per $k \rightarrow \infty$.

DIMOSTRAZIONE. La dimostrazione segue [39, p. 487-488]. Dato che ogni \hat{f} è approssimabile in media quadratica mediante funzioni semplici, combinazioni lineari di un numero finito di funzioni indicatrici di intervalli, è sufficiente dimostrare il lemma per una funzione indicatrice $\hat{f} = \chi_{\Delta}$, dove $\Delta = [\lambda_1 \lambda_2]$ è un intervallo. La funzione χ_{Δ} ha uno sviluppo di Fourier

$$\chi_{\Delta}(\omega) = \sum_{-\infty}^{+\infty} c_n e^{j\omega n}, \quad c_n = \int_{\lambda_1}^{\lambda_2} e^{-j\omega n} \frac{d\omega}{2\pi} = \frac{e^{-j\lambda_2 n} - e^{-j\lambda_1 n}}{-jn}$$

Per un noto teorema di Fejér [42, p. 34-38], la successione delle medie aritmetiche delle somme parziali della serie di Fourier converge, per $N \rightarrow \infty$, puntualmente a χ_{Δ} , i.e.

$$s_N(\omega) := \frac{1}{N} \sum_{n=1}^N \left(\sum_{k=-n}^n c_k e^{j\omega k} \right) \rightarrow \chi_{\Delta}(\omega)$$

per ogni ω , pur di ridefinire i valori di χ_Δ agli estremi dell'intervallo $[\lambda_1, \lambda_2]$ prendendo $\chi_\Delta(\lambda_1) = \chi_\Delta(\lambda_2) = 1/2$. Dato che la successione di polinomi trigonometrici $\{s_N\}$ è uniformemente limitata, anche la successione di funzioni $|\chi_\Delta(\omega) - s_N(\omega)|^2$ è limitata (da una costante) e converge puntualmente a zero per $N \rightarrow \infty$. Si ha allora, in base al teorema della convergenza dominata,

$$\int_{-\pi}^{\pi} |\chi_\Delta(\omega) - s_N(\omega)|^2 dF(\omega) \rightarrow 0$$

che è quanto si doveva dimostrare. \square

Ne segue, per un ragionamento ormai noto, che

TEOREMA D.2 (Rappresentazione Spettrale). *La corrispondenza $I_{\mathbf{y}}$ definisce una mappa unitaria*

$$I_{\mathbf{y}} : \mathbf{H}(\mathbf{y}) \rightarrow L^2([-\pi, \pi], dF) \quad (0.10)$$

che trasforma l'operatore di traslazione temporale U_t sulle variabili del processo \mathbf{y} , nell'operatore di moltiplicazione per la funzione e_t, M_{e_t} , in $L^2([-\pi, \pi], dF)$.

La seconda affermazione scende dal fatto che l'operatore $I_{\mathbf{y}}$ applicato alla relazione $\mathbf{y}(t) = U_t \mathbf{y}(0)$ porge $e_t = M_{e_t} e_0$ dove e_0 è evidentemente la funzione identicamente uguale a uno che rappresenta $\mathbf{y}(0)$. In sostanza, ogni variabile aleatoria ξ di $\mathbf{H}(\mathbf{y})$ è rappresentata da una funzione della frequenza \hat{f}_ξ , che ha la stessa norma.

Il lemma D.1 e il teorema D.2 si estendono al caso di processi vettoriali, definendo $I_{\mathbf{y}}$ componente per componente: $I_{\mathbf{y}} : \mathbf{y}_i(t_k) \rightarrow [0 \dots 1 \dots 0] e^{j\omega t_k}$ (l'1 è all' i -esimo posto), e considerando lo spazio immagine contenuto in $L_m^2([-\pi, \pi], dF)$, lo spazio popolato da funzioni *riga* m -dimensionali della variabile ω che sono a quadrato integrabile rispetto alla distribuzione spettrale a valori matriciali $m \times m$, del processo stazionario m -dimensionale \mathbf{y} , i.e. tali che

$$\int_{-\pi}^{\pi} f(\omega) dF(\omega) f^*(\omega) < \infty \quad (0.11)$$

dove l'asterisco sta per coniugata trasposta. L'enunciato del teorema di rappresentazione spettrale nel caso vettoriale è formalmente analogo a quello del teorema D.2 pur di sostituire lo spazio vettoriale $L_m^2([-\pi, \pi], dF)$ a quello scalare.

Generalizzazione della trasformata di Fourier

Sia Σ la funzione di covarianza di un processo stazionario m -dimensionale \mathbf{y} e dF la relativa matrice distribuzione spettrale sulla circonferenza unita, corrispondenti attraverso la formula di Herglotz. Supponiamo che f e g siano funzioni a supporto finito (ovvero, $f(t), g(t) \neq 0$ solo per un numero finito di indici t). Dato che Σ è una funzione di tipo positivo, la forma bilineare definita dalla somma (che è estesa ad un numero finito di termini)

$$\langle f, g \rangle_\Sigma := \sum_{t,s} f(t) \Sigma(t-s) g(s)' = E \left(\sum_t f(t) \mathbf{y}(t) \sum_s g(s) \mathbf{y}(s) \right) \quad (0.12)$$

è non negativa e pertanto può essere usata per definire una norma, $\|\cdot\|_\Sigma$, sullo spazio vettoriale delle successioni m -dimensionali f a supporto finito.

Definiamo allora lo spazio di Hilbert $\ell_m^2(\Sigma)$ come la chiusura di questo spazio vettoriale rispetto alla norma $\|\cdot\|_\Sigma$. Ovviamente $\ell_m^2(\Sigma)$ contiene tutte le successioni

infinite f per cui

$$\sum_{t,s=-\infty}^{+\infty} f(t)\Sigma(t-s)f(s)' < \infty$$

però non ci è ancora noto se queste successioni esauriscano lo spazio o se questo potrebbe a priori contenere anche altri oggetti.

Nota Bene: È necessario fare una precisazione sul significato che si attribuisce alla parola “successione” quando ci si riferisce ad elementi di $\ell_m^2(\Sigma)$. Ogni elemento f di questo spazio è in realtà una *classe di equivalenza* di successioni g che sono indistinguibili da f nella norma dello spazio, tali cioè che $\|f - g\|_\Sigma = 0$. Gli elementi di norma nulla sono le successioni appartenenti al nucleo della matrice di Toeplitz infinita $\Sigma_\infty := \{\Sigma_{t,s} = \Sigma(t-s) \mid t, s \in \mathbb{Z}\}$.

ESEMPIO D.1. Consideriamo un processo p.d. elementare consistente di n oscillazioni aleatorie elementari scorrelate di frequenze ω_k , $k = 1, \dots, n$. Si verifica facilmente che in questo caso la matrice di Toeplitz Σ_∞ ha rango finito, uguale ad n . Ne segue che il nucleo di Σ_∞ è praticamente tutto lo spazio ... In questo caso $\ell_m^2(\Sigma)$ (che ha dimensione n) consiste di classi di equivalenza che hanno dimensione infinita! \diamond

ESEMPIO D.2. Sia \mathbf{w} un rumore bianco di varianza σ^2 . Come visto all'esempio 2.5, la matrice di covarianza $\Sigma_t := E \mathbf{y}^t (\mathbf{y}^t)'$, $t = 1, 2$, del processo stazionario

$$\mathbf{y}(t) = \mathbf{w}(t) + \mathbf{w}(t-1)$$

ha la struttura tridiagonale

$$\Sigma_t = \sigma^2 \begin{bmatrix} 2 & 1 & 0 & \dots & 0 \\ 1 & 2 & \ddots & 1 & \\ 0 & & \ddots & \ddots & \\ \vdots & & & \ddots & 1 \\ 0 & & & 1 & 2 \end{bmatrix}$$

Notare che nonostante Σ_t sia non singolare per ogni t , tutte le successioni f che soddisfano l'equazione alle differenze

$$f(t) + 2f(t-1) + f(t-2) = 0$$

appartengono allo spazio nullo della matrice (doppiamente) infinita

$$\Sigma_\infty = \sigma^2 \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \dots & \dots \\ \dots & 0 & 1 & 2 & 1 & 0 & \dots \\ & \dots & 0 & 1 & 2 & 1 & 0 & \dots \\ & & \dots & \dots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

L'elemento nullo dello spazio $\ell^2(\sigma)$ è quindi tutto il sottospazio bidimensionale

$$\mathcal{O} = \{f \mid f(t) = A(-1)^t + B(t-1)(-1)^{t-1}; A, B \in \mathbb{R}, t \in \mathbb{Z}\}$$

e due successioni f, g sono lo stesso elemento di $\ell^2(\sigma)$ se e solo se $f - g \in \mathcal{O}$. \diamond

Introdurremo ora una generalizzazione della trasformazione di Fourier che torna utile per discutere le proprietà del filtro sbiancante di un processo p.n.d.

PROPOSIZIONE D.2. *L'operatore lineare*

$$\mathfrak{F}_\Sigma : \ell_m^2(\Sigma) \rightarrow L_m^2([-\pi, \pi], dF),$$

definito, per ogni f a supporto compatto in $\ell_m^2(\Sigma)$, i.e. $f = \sum_k f_k \delta(t - t_k)$ dove la somma è estesa ad un numero finito di termini, dalla somma finita

$$\mathfrak{F}_\Sigma(f) := \sum_{t=-\infty}^{+\infty} e^{-j\omega t} f(t) = \sum_k f_k e^{-j\omega t_k} \quad (0.13)$$

è isometrico e può essere esteso in modo unico ad un operatore unitario che applica $\ell_m^2(\Sigma)$ su $L_m^2([-\pi, \pi], dF)$. L'estensione si chiama Trasformata di Fourier Generalizzata.

DIMOSTRAZIONE. La proprietà di preservare la norma (isometria), è una conseguenza dalla rappresentazione di Herglotz, dalla quale scende che per tutte le f a supporto compatto,

$$\begin{aligned} \sum_{t,s=-\infty}^{+\infty} f(t)\Sigma(t-s)f(s)' &= \sum_{hk} f_h \Sigma(t_h - t_k) f_k' = \\ \sum_{hk} f_h \int_{-\pi}^{\pi} e^{j\omega(t_h - t_k)} dF(\omega) f_k' &= \int_{-\pi}^{\pi} \hat{f}(\omega) dF(\omega) \hat{f}(\omega)^* \end{aligned}$$

Dato che $\|\mathfrak{F}_\Sigma(f)\|_{L_m^2([-\pi, \pi], dF)} = \|f\|_{\ell_m^2(\Sigma)}$, l'operatore \mathfrak{F}_Σ è isometrico e può essere esteso in modo unico a tutti i limiti di successioni a supporto compatto (cioè a tutto $\ell_m^2(\Sigma)$) preservando il carattere di isometria (Teorema B.1). D'altro canto, l'immagine di una f a supporto compatto è un polinomio trigonometrico, per cui i limiti delle trasformate esauriscono tutto $L_m^2([-\pi, \pi], dF)$ e l'estensione di \mathfrak{F}_Σ è un operatore suriettivo. \square

L'estensione di \mathfrak{F}_Σ verrà denotata con lo stesso simbolo.

COROLLARIO D.2. Se $\hat{f} \in L_m^2([-\pi, \pi], dF)$ e $f := \mathfrak{F}_\Sigma^{-1}(\hat{f})$ è l'antitrasformata di Fourier generalizzata di \hat{f} , allora la somma

$$\sum_{-\infty}^{+\infty} f(t) \mathbf{y}(t) \quad (0.14)$$

è convergente in media quadratica.

DIMOSTRAZIONE. Sia $\hat{f}_N := \sum_{k=-N}^N f_{Nk} e^{-j\omega t_k}$ una successione di polinomi trigonometrici convergenti a \hat{f} e $\{f_N\}$ la corrispondente successione di antitrasformate generalizzate. Ciascun termine f_N ha supporto in $[-N, N]$ e, in base all'isometria della trasformata di Fourier generalizzata dimostrata più sopra, si ha $f_N \rightarrow f$ in $\ell_m^2(\Sigma)$. Dobbiamo dimostrare che la successione di variabili aleatorie $\{I_N\}$ definita ponendo

$$I_N := \sum_{-\infty}^{+\infty} f_N(t) \mathbf{y}(t)$$

è convergente in media quadratica.

La cosa scende facilmente notando che, se ad esempio $M \leq N$, si ha

$$I_N - I_M = \sum_{-\infty}^{+\infty} \varphi_N(t) \mathbf{y}(t)$$

con $\varphi_N := f_N - f_M$ ancora a supporto compatto in $[-N, N]$ e pertanto

$$E [I_N - I_M]^2 = \sum_{-\infty}^{+\infty} \sum_{-\infty}^{+\infty} \varphi_N(t) \Sigma(t-s) \varphi_N(s)' = \|\varphi_N\|_{\ell_m^2(\Sigma)}^2 = \|f_N - f_M\|_{\ell_m^2(\Sigma)}^2$$

che tende a zero per $N, M \rightarrow \infty$. Questo implica che I_N converge in media quadratica e quindi la somma (0.14) è convergente. \square

Una conseguenza immediata di questo risultato è che se $f := \mathfrak{F}_\Sigma^{-1}(\hat{f})$ è l'anti-trasformata di Fourier generalizzata di \hat{f} , allora anche la somma di convoluzione

$$\mathbf{z}(t) := \sum_{-\infty}^{+\infty} f(t-s) \mathbf{y}(s) \quad (0.15)$$

è convergente in media quadratica. Chiameremo in questo caso \hat{f} la *funzione di trasferimento del "filtro" di risposta impulsiva* f . Ovviamente si possono introdurre filtri con funzione di trasferimento matriciale e riposta un processo a valori vettoriali, purchè beninteso le righe della matrice di trasferimento del filtro siano tutte elementi di $L_m^2([-\pi, \pi], dF)$.

PROPOSIZIONE D.3 (Formula di Wiener-Kintchine). *Nelle ipotesi poste, l'uscita $\{\mathbf{z}(t)\}$, del filtro di funzione di trasferimento $\hat{f} \in L_m^2([-\pi, \pi], dF)$, definita dalla (0.15), è un processo stazionario di distribuzione spettrale*

$$F_{\mathbf{z}}(\omega) = \int_{-\pi}^{\omega} \hat{f}(\lambda) dF(\lambda) \hat{f}(\lambda)^* \quad (0.16)$$

che si può riscrivere in forma simbolica come $dF_{\mathbf{z}}(\omega) = \hat{f}(\omega) dF(\omega) \hat{f}(\omega)^*$.

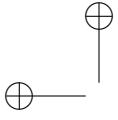
DIMOSTRAZIONE. Per l'unicità della rappresentazione di Herglotz, la formula (0.16) è equivalente alla scrittura

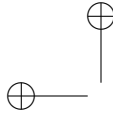
$$\sigma_{\mathbf{z}}(t-s) = \langle e_t \hat{f}, e_s \hat{f} \rangle_{L_m^2([-\pi, \pi], dF)}$$

dove $\sigma_{\mathbf{z}}$ è la funzione di covarianza del processo \mathbf{z} . Quest'ultima è a sua volta una immediata conseguenza della

$$I_{\mathbf{y}} \mathbf{z}(t) = e_t \hat{f} \quad (*)$$

che implica tra l'altro la stazionarietà di $\{\mathbf{z}(t)\}$. Per provare la (*), basta allora considerare una successione di polinomi trigonometrici convergenti ad \hat{f} nello spazio $L_m^2([-\pi, \pi], dF)$ e notare che per ogni polinomio approssimante della successione la relazione (*) è vera per definizione. \square





Bibliografia

- [1] H.Akaike, "Markovian representation of stochastic processes by canonical variables," *SIAM J. Control* **13**, 162–173, (1975).
- [2] Akhiezer N.I., Glazman I.M. *Theory of Linear Operators in Hilbert Spaces*, voll. I e II, Ungar, New, York, 1966.
- [3] Anderson B.D.O., "The inverse problem of stationary covariance generation," *J. Statistical Physics*, **1** (1969), 133–147.
- [4] Anderson, B.D.O. Gevers M. "Identifiability of Linear Stochastic Systems Operating Under Linear Feedback", *Automatica*, **18**: 195-213, 1982.
- [5] Anderson B.D.O. Hitz K. L. Diem N.D. "Recursive Algorithm for Spectral Factorization", *IEEE Trans. Circuits and Systems*, **CAS-21**: 742-750, 1974.
- [6] Anderson, B.D.O., Moore J.B., *Optimal Filtering*, Prentice Hall. Englewood Cliffs, 1979.
- [7] Anderson, B.D.O., Moore J.B., *Optimal Control: Linear Quadratic Methods*, Prentice Hall. Englewood Cliffs, 1989.
- [8] Anderson, B.D.O., Choi A.C., "Connecting forward and backward autoregressive models" *IEEE Trans. Automatic Control*, **AC-29**: 917-925, 1984.
- [9] Arnold W.F. and Laub A.J., "Generalized Eigenproblem Algorithms and software for the Algebraic Riccati equation", *Proc. of the IEEE*, **79**, 1746–1754, 1984.
- [10] Åstrom K.J., *Introduction to Stochastic Control Theory*, Ac. Press, 1970.
- [11] Åstrom, K.J., Wittenmark B., *Computer Controlled Systems*, Prentice-Hall, 1984.
- [12] Bartlett M.S., *Introduction to stochastic processes with special reference to methods and applications*, Cambridge U.P., 1955.
- [13] Basile G. and Marro G., *Controlled and conditioned invariants in linear system theory*, Prentice-Hall, 1991.
- [14] Bauer F.L. "Ein direktes Iterationsverfahren zur Hurwitz Zerlegung eines Polynoms" *Arch. Elektr. Übertr.*, **9**, pp. 285-290, 1955.
- [15] Bellman R. *Dynamic Programming*, Princeton U.P., 1957.
- [16] Ben-Israel A., Greville T. N. E., *Generalized Inverses: Theory and Applications*, Wiley, New York, 1974.
- [17] Benveniste A., Chaure C., A.R. and ARMA identification of Levinson Type: an Innovations Approach, *IEEE Tr. Aut. Control.*, **A.C.-26**, pp. 1243-1261, 1981.
- [18] Bertero M. Linear inverse and ill-posed problems, *Advances in Electronics and Electr. Physics*, **75**, pp 1-121, 1989.
- [19] Bertero M. and Boccacci P. *Introduction to inverse problems in imaging*, Inst. of Physics Pub. London, 1998.
- [20] Bierman J., *Factorization Methods for Discrete Sequential Estimation*, Ac. Press 1977.
- [21] Birkhoff G.D., "Proof of the ergodic hypothesis, *Proc. Nat. Acad Sci (U.S.A.)* **17**, pp. 565-600. 1931.
- [22] Bittanti S., Scattolini R., *Controllo Stocastico a minima Varianza*, Clup Milano 1982.
- [23] S. Bittanti, *Count Riccati and the early years of the Riccati Equation*, ..., 1991.
- [24] S. Bittanti, A Laub, J.C. Willems (editors), *The Riccati Equation*, Springer Verlag, 1991.
- [25] Bittanti S., Picci G. (editors) *Identification, Adaptation, Learning (the science of learning models from data)*, Springer Verlag NATO-ASI series, vol 153, 1996.
Bittanti S., Scattolini R., *Controllo Stocastico a minima Varianza*, Clup Milano 1982.
- [26] H.W. Bode, C.E. Shannon, "A Simplified Derivation of Linear Least-Squares Smoothing and Prediction Theory", *Proceedings of the IRE*, **38**: 417-425, 1950.
- [27] R. W. Brockett, Lie Algebras and Lie Groups in Control Theory, in *Geometric Methods in Control*, R. W. Brockett and D. Mayne eds. Reidel, Dordrecht, 1973.

- [28] R. W. Brockett, Notes on Stochastic Processes on Manifolds, in *Control Theory in the 21st Century*, C.I Byrnes, C. Martin, B. Datta eds. Birkhauser, 1997.
- [29] Byrnes C. I., Lindquist A., Gusev S.V., Matveev A.S.: A complete parametrization of all positive real extensions of a covariance sequence, *IEEE Trans. Automatic Control*, **AC-40**, pp 1841-1857, 1995.
- [30] Callier F., "On polynomial matrix spectral factorization by symmetric extraction", *IEEE Transactions on Automatic Control*, **AC-30**: 453-454, 1985.
- [31] Caines P.E., Chan C.W., "Estimation, Identification and Feedback", in *System Identification: Advances and case studies*, R.K. Mehra and D. Lainiotis eds, Ac. Press, 1976.
- [32] Chan S.W., Goodwin G.C., Sin K.S., "Convergence properties of the Riccati Difference Equation in Optimal Filtering of Nonstabilizable Systems", *IEEE Transactions on Automatic Control*, **AC-29**, 110-118, 1984.
- [33] Childers D.G. (Ed.), *Modern Spectrum Analysis*, IEEE Press 1978.
- [34] A. Chiuso and G. Picci, Tracking of point features in computer vision as estimation on the unit sphere, in *The Confluence of Vision and Control*, D. Kriegman, W. Haeger and S. Morse eds, Springer LNCIS series, 1998.
- [35] Cover T.M. and J.A. Thomas *Elements of Information Theory*, Wiley, 1991.
- [36] Davis M.H.A., Vinter R.B., *Stochastic Modelling and Control*, Chapman and Hall, 1985.
- [37] Delsarte P., Genin Y. and Kamp Y.G., "Orthogonal Polynomial Matrices on the Unit Circle", *IEEE Transactions on Circuits and Systems*, **CAS-25**: 149-160, 1978.
- [38] De Souza C., Gevers M. and Goodwin G.C., "Riccati Equations in Optimal Filtering of Nonstabilizable Systems having Singular State Transition Matrix", *IEEE Transactions on Automatic Control*, **AC-31**, 831-838, 1986.
- [39] Doob J.L., *Stochastic Processes*, Wiley 1953.
- [40] Durbin J. "The fitting of time series models", *Rev. of the International Statistical Institute*, **28**, 233-244, 1960.
- [41] Duren P. L. *Theory of H^p spaces*, Academic Press, New York, 1970 (ristampato da Dover).
- [42] H.Dym and H.P.McKean, *Fourier Series and Integrals*, Academic Press (1972).
- [43] Edward J.A., Fitelson M.M., "Note on Maximun Entropy Processing", *IEEE Tr. Inf. Theory*, **IT-19**, pp. 232-234, 1973.
- [44] P.Faurre, M.Clerget and F.Germain, *Opérateurs Rationnels Positifs*, Dunod, Paris (1979).
- [45] A. Ferrante, G. Picci, and S. Pinzoni "Silverman algorithm and the structure of discrete-time stochastic systems, *Linear Algebra and its Applications*, special issue on systems and control, **351-352**, pp. 219-242 (2002).
- [46] T. L. Fine *Theories of Probability*, Ac Press, New York, 1973.
- [47] R. A. Fisher, Dispersion on a sphere, *Proc. Royal Soc. London*, **A 217**, p. 295-305, 1953.
- [48] Fraser D., Potter J.E., "The Optimum Linear Smoother as a Combination of two Optimum Linear Filters", *IEEE Transactions on Automatic Control*, **AC-14**, 387-390, 1969.
- [49] Freiburger W., Grenander U., *A short Course in Computational Probability and Statistics*, Springer Verlag 1971.
- [50] Friedlander B., "lattice Filters for Adaptive Processing", *Proc. IEEE*, 70, pp. 829-867, 1982.
- [51] Fuhrmann P.A. *Linear Systems and Operators in Hilbert Space*, Mac Graw Hill, 1984.
- [52] Gantmacher F. R. *Theory of Matrices*, Voll. I e II, Chelsea, New York 1959.
- [53] Georgiou T. T.: Realization of power spectra from partial covariance sequences, *IEEE Trans. on Acoustic Speech and Signal Processing*, **ASSP-35**, pp 438-449, 1987.
- [54] Gevers M., Anderson B.D.O., "Representations of jointly Stationary Stochastic Feedback Processes", *Int. Journal of Control*, **33**, pp. 777-809, 1981.
- [55] Gevers M., Anderson B.D.O., "On Jointly Stationary, Feedback-Free Stochastic Processes", *IEEE Trans. Aut. Control*, **AC-27**, pp. 431-436, 1982.
- [56] Glover K.: All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ error bounds, *International Journal of Control*, **39**, pp. 1115-1193, 1984.
- [57] Golub G.H. "Numerical methods for solving linear least-squares problems", *Numerische Mathematik*, **7**, pp.206-216, 1965.
- [58] Golub G.H, Van Loan C. F., *Matrix Computations*, Johns Hopkins U. P., Baltimore, 1989.
- [59] Granger C.W.J., "Economic Processes Involving Feedback", *Information and Control*, **6**, pp. 28-48, 1963.
- [60] Grenander U. and G. Szegö. *Toeplitz Forms and their Applications*, Chelsea, New York, 1958.
- [61] Halmos P.R., *Introduction to Hilbert Space*, Chelsea 1957.

- [62] Hannan E.J., *Multiple Time Series*, Wiley 1970.
- [63] Hannan, E.J. and D.S. Poskitt (1988). Unit canonical correlations between future and past. *The Annals of Statistics* **16**, 784–790.
- [64] Helson H., *Lectures on Invariant Subspaces*, Academic Press, (1976).
- [65] Hoffmann K. *Banach Spaces of Analytic Functions*, Prentice-Hall, 1962.
- [66] Horn R.A and Johnson C.R. *Matrix Analysis*, Cambridge Univ. press , 1985.
- [67] Jayant N.S. and Noll P., *Digital Coding of Waveforms*, Prentice-Hall, 1984.
- [68] Jazwinski A. H. *Stochastic processes and Filtering Theory* Academic Press, New York, 1970.
- [69] Jezek J., Kucera, V. “Efficient Algorithm for Matrix Spectral Factorization”, *Automatica*, **21**: 663-669, 1985.
- [70] Kailath T., An innovations approach to least squares estimation, Part. I', *IEEE Trans. Aut. Control*, **AC-13**, pp. 646-655, 1968.
- [71] Kailath T., "A view of three decades of linear filtering theory", *IEEE Tr. Inf. Theory*, **IT-20**, pp. 146-181, 1974.
- [72] Kailath T., *Linear Systems Theory*, Prentice Hall, 1980.
- [73] Kailath T., Vieira A., Morf M., "Inverses of Toeplitz Operators, Innovations and Orthogonal Polynomials", *SIAM Review*, **20**, pp. 106-119, 1978,
- [74] Kalman R.E., "A new Approach to linear Filtering and Prediction problems", *Trans. A.S.M.E., J. of Basic Engineering*, **82**, pp. 35-45, 1960.
- [75] Kalman R.E., "New methods in Wiener filtering theory", *Proc. Ist Symp. Eng. Appl. of Random Functions Theory Probability*, J.L. Boganoff, F. Kozin eds, Wiley 1963.
- [76] Kalman R.E. "Lyapunov functions for the problem of Lur'e in Automatic Control, *Proc. natl. Acad. Sci. USA* **49**:201-205, 1963.
- [77] R.E.Kalman, P.L.Falb, and M.A.Arbib, *Topics in Mathematical Systems Theory*, McGraw-Hill (1969).
- [78] Kalman, R.E. Realization of Covariance Sequences, *Proc. Toeplitz Memorial Conference*, Tel Aviv, Birkhauser, 1981.
- [79] Kamen E.W., Su J.K., *Introduction to Optimal Estimation*, Springer Verlag, 1999.
- [80] Kolmogorov A.N., *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlino, 1933 (traduz. inglese in *Foundations of the Theory of Probability*, Chelsea, 1950, 1956).
- [81] Kolmogorov A.N., "Sur l' interpolation et extrapolation des suites stationnaires", *C. R. Acad. Sci. Paris*, **208**, pp. 2043-2045 1939.
- [82] Kolmogorov A.N., "Stationary Sequences in Hilbert Spaces", *Bull. Moskov Gos. Univ. Mat.* (in russo), **2**, pp. 1-40, 1941 (traduz. Inglese in *Linear Least Squares Estimation*, T. Kailath ed., Benchmark papers in Electrical Engineering and Computere Sciences, Dowden Hutchinson & Ross, inc 1977).
- [83] Kolmogorov A.N., "Interpolation und Extrapolation von stationaren zufalligen Folgen", *Izv. Akad. Nauk U.S.S.R. Ser. Mat.*, **5**, pp. 3-14, 1941 (traduz. Inglese in *Linear Least Squares Estimation*, T. Kailath ed., Benchmark papers in Electrical Engineering and Computere Sciences, Dowden Hutchinson & Ross, inc 1977).
- [84] Kucera V., "The discrete Riccati equation of optimal control", *Kybernetika* (Prague), **8**, 430-447, 1972.
- [85] Kucera V., "On nonnegative definite solutions to matrix quadratic equations", *Automatica*, **8**, 413–423, 1972.
- [86] Kucera V. *Discrete Linear Control, The Polynomial Equation Approach*, Wiley-Chichester, 1979.
- [87] Kucera V. "Algebraic Riccati Equation: Hermitian and Definite Solutions" in *The Riccati equation*, S. Bittanti, A Laub, J.C. Willems eds, Springer Verlag, 1991.
- [88] Kullback S.: "A lower bound for discrimination information in terms of variation", *IEEE Trans. on Information Theory*, **4**, pp. 126-127, 1967.
- [89] Kwakernaak H. and Sivan R. *Linear optimal control systems*, Wiley, New York, 1972.
- [90] Laub A. "A Schur method for solving algebraic Riccati equations", *IEEE Transactions Aut. Control*, **AC-24**, pp 913-921, 1979.
- [91] P. Langevin, Magnetisme et theorie des electrons, *Ann. de Chim et de Phys.*, **5**, p. 70-127, 1905.
- [92] Levinson N., "The Wiener r.m.s. (root means square) Error Criterion in Filter Design and Prediction", *J. Math. Phys.* **25**, pp. 261-278, 1947 (pubblicato anche in Appendice B al libro di Wiener [150]).

- [93] Lindquist A., "A new Algorithm for optimal Filtering of discrete-time stationary processes", *SIAM J., Control*, **12**, pp. 736-746, 1974.
- [94] Anders Lindquist and Giorgio Picci. Realization theory for multivariate stationary Gaussian processes. *SIAM J. Control Optim.*, **23**: 809–857, 1985.
- [95] Liptser R., Shirayev A.N., *Statistics of Random Processes*, Voll I e II, Springer Verlag, 1977.
- [96] Ljung L., *System Identification (theory for the user)*, Prentice Hall, 1987.
- [97] J. Lo and A. Willsky, Estimation for rotational processes with one degree of freedom, parts I, II, III, *IEEE Transactions on Automatic Control*, **AC-20**, pp. 10-33, 1975.
- [98] H. P. McKean, Brownian Motion on the Three-Dimensional Rotation Group, *Mem. Coll. Sci. University of Kyoto, Series A*, **XXXIII**, N. 1, pp. 25-38, 1960.
- [99] MacFarlane A.G.J. "An eigenvector solution to the optimal linear regulator", *Journal Electr. and Control*, **14**, pp. 643-654, 1963.
- [100] Makhoul J., Linear Prediction: A tutorial review, *Proc. IEEE*, 63. pp. 501-580, 1995.
- [101] Masani P., "The prediction theory of of multivariate stationary processes, Part III: Unbounded spectral densities", *Acta Mathematica*, **104**, 141–162, 1960.
- [102] Mayne D.Q. "A Solution to the Smoothing Problem for Linear Dynamical Systems", *Automatica*, **4**, 73–92, 1966.
- [103] Molinari B. P. "The stabilizing solution of the discrete algebraic Riccati equation", *IEEE Trans. Automatic Control*, **AC-20**, pp. 396–399, 1975.
- [104] Moore B.C.: Principal component analysis in linear systems: controllability, observability and model reduction, *IEEE Trans. Automatic Control*, **AC-26**, pp. 17-32, 1981.
- [105] Morf, M., Vieira A. and Kailath T., "Covariance characterization by partial autocorrelation matrices", *The Annals of Statistics*, **6** pp. 643–648, 1978.
- [106] Mullis C.T. and Roberts R.A., "The use of second order information in the approximation of discrete-time linear systems", *IEEE Transactions on Acustics Speech and Signal Processing*, **ASSP-24**: 226–238, 1976.
- [107] Neveu J., *Processus Aleatoires Gaussiens*, Les Presses de l'Université de Montreal, 1968.
- [108] Nikolskii N.K., *Operators, Functions and Systems: An Easy Reading Volume I: Hardy, Hankel and Toeplitz*, A.M.S. Mathematical Surveys and Monographs, Providence R.I., 2002.
- [109] Ober R.J.: Balanced realizations.: canonical forms, parametrizations, model reduction, *International Journal of Control*, **46**, pp. 643-670, 1987.
- [110] Øksendal, *Stochastic Differential Equations*, Springer, 1990.
- [111] Paley R. E. A. C. and Norbert Wiener, *Fourier transforms in the complex domain*, volume 19 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1987. (Reprint of the 1934 original).
- [112] Pappas T., Laub A., Sandell N., "On the numerical solution of the discrete-time Algebraic Riccati Equation", *IEEE Tr. Aut. Control*, **AC-25**, pp. 631-641, 1980.
- [113] Pernebo L. Silverman L.M.: Model reduction via balanced state space representations, *IEEE Transactions on Automatic Control*, **AC-37**, pp382-387, 1982.
- [114] F. Perrin, étude Mathématique du Mouvement Brownien de Rotation, *Ann. Ecole Normale Supérieure*, (3), **XLV**: 1-51, 1928.
- [115] G. Picci, Dynamic Vision and Estimation on Spheres, in *Proceedings of the 36th Conf. on Decision and Control*, p. 1140-1145, IEEE Press, 1997.
- [116] G. Picci, *Il concetto di stima e la sua applicazione alle misure*, in *Metodologie Statistiche per il Trattamento delle Misure*, M. Gasparetto e G. Iuculano eds. CISM, Udine, 1993.
- [117] V.M.Popov, "Hyperstability and optimality of automatic systems with several control functions," *Revue Roumaine des Sciences Techniques, Sèrie Electrotechnique et Energétique* **9** (1964), 629–690.
- [118] Potter J.E., "Matrix quadratic solutions", *SIAM. J. Appl Math*, **14**, pp. 496-501, 1966.
- [119] Rabiner L., Schafer B.W., *Digital Processing of Speech Signals*, Prentice Hall, 1978.
- [120] Rauch H.E., Tung F., Striebel C.T., "Maximum Likelihood Estimation of Linear Dynamic Systems", *AIAA Journal*, **3**, 1445–1450, 1965.
- [121] Rissanen J., "Algorithms for Triangular Decomposition of Block Hankel and Toeplitz matrices with applications to factoring positive matrix Polynomials", *Math. Comp.*, **27**, pp. 147-154, 1973.
- [122] Rissanen J., "Solution of Linear Equations with Hankel or Toeplitz Matrices", *Numer. Mathematik*, **22**: 361-366, 1974.

- [123] Rissanen J., Barbosa L., "Properties of Infinite Covariance matrices and stability of Optimum Predictors", *Inf. Sciences*, **1**, pp. 221-236, 1969.
- [124] Roberts, R.A. and Mullis C.T. *Digital Signal Processing*, Addison Wesley, 1987.
- [125] Royden H.L. Real Analysis, 3rd ed., Prentice Hall, 1988.
- [126] Rozanov Y.A., *Stationary Random Processes*, Holden-Day 1967.
- [127] Rudin W. Real and Complex Analysis, 3rd ed. McGraw Hill, 1988.
- [128] Shannon C. E., "A Mathematical Theory of Communication" *Bell System Tech. Journal*, **27**, 379-423, 623-656, 1948.
- [129] Shannon C.E., Weaver W., *The Mathematical Theory of Communication*, Univ. of Illinois, Press 1949.
- [130] Sherman S., "Non-Mean Square Error Criteria", *IRE Trans. Inform. Theory*, **4**, pp. 125-126, 1958.
- [131] Smith *History of Mathematics*, Dover, 1958.
- [132] Söderström T. and Stoica P. *System Identification*, Prentice-Hall, 1989.
- [133] Sorenson H. W. (ed.) *IEEE Tr. Aut. Control Special Issue on Applications of Kalman Filtering*, **AC-28**, March 1983.
- [134] Stewart G.W., *Introduction to Matrix Computations*, Ac. Press 1973.
- [135] Strang G. *Linear Algebra and its Applications*, Academic Press, 1976.
- [136] Strobach P. *Linear Prediction Theory*, Springer Verlag, 1990.
- [137] Szegő G., *Orthogonal Polynomials* volume 23 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1967. (Reprint of the 1939 original).
- [138] Trench W.F., "An Algorithm for the inversion of finite Toeplitz matrices", *SIAM J. Appl. Math*, **12**, pp. 515-522, 1964.
- [139] Van Der Bos A., "Alternative Interpretation of M.E. Spectral Analysis" *IEEE Transactions on Inf. Theory*, **IT-17**, pp. 493-494, 1971.
- [140] Van Dooren P., "A Generalized Eigenvalue Approach for solving Riccati Equations", *SIAM J. Sci. Stat. Comput.*, **2**, pp. 121-135, 1981.
- [141] Overschee, P. Van and B. De Moor (1996). *Subspace Identification for Linear Systems*. Kluwer Academic Publications.
- [142] Van Trees H., *Detection, Estimation and Modulation Theory*, (3 voll.), Wiley, 1968.
- [143] Vidyasagar M. *Control System Synthesis, A Factorization Approach*, MIT Press, 1985.
- [144] Van Dooren P. "A generalized eigenvalue approach for solving Riccati equations", —em *SIAM Journal Sci. Stat. Comput.*, **2**, pp. 121-135, 1981.
- [145] Von Mises R., *Mathematical Theory Probability and Statistics*, Ac. Press 1964.
- [146] G. S. Watson, *Statistics on Spheres*, Wiley, N.Y 1983.
- [147] Whittle P., *Prediction and Regulation*, Blackwell, Oxford, 1963.
- [148] Whittle P., "On the fitting of multivariate autoregressions and the approximate canonical factorization of a spectral density matrix". *Biometrika*, **50**, pp. 129-134, 1963.
- [149] Wiener N., *Final Report on section D2, Project N. 6*, Nat. Def. Res. Comm., Dic. 1942, (versione espansa in *Statistical Methods of Prediction in Fire Control* Report N. 59, 1945)
- [150] Wiener N., *The Extrapolation Interpolation and Smoothing of Stationary Time Series with engineering Applications*, J. Wiley & Sons, 1949.
- [151] Wiener N., "The Theory of Prediction", in *Modern Mathematics for the Engineer*, E.F. Beckenbach ed., McGraw-Hill, 1956.
- [152] Wiener N., Masani P., "The Prediction Theory of multivariate stochastic processes", I, *Acta Math.*, **98**, pp. 111-150, 1957; II *ibidem*, **99**, pp. 93-137, 1958.
- [153] Wiggins R.A. Robinson E.A. " Recursive Solution to the Multichannel Filing Problem" *J. Geophys. Research*, **70**: 1885-1891, 1965.
- [154] Wilson G.T. "The Factorization of Matricial Spectral Densities", *SIAM J. Applied Math.*, **23**: 420, 1972.
- [155] Wold H., *A Study in the Analysis of stationary Time Series*, Almquist and Wiksell, Stockholm, 1938.
- [156] Wong E. *Stochastic Processes in Information and Dynamical Systems*, Mc Graw-Hill, 1971.
- [157] Wonham M. *Linear multivariable control: a geometric approach*, Springer, Berlin, Heidelberg New York, 1985.
- [158] V. A. Yakubovich: The solution of some matrix inequalities encountered in automatic control theory. *Dokl. Akad. Nauk SSSR*, **143**: 1304-1307, 1962.

- [159] Yang B., "Projection Approximation Subspace Tracking", *IEEE Transactions on Signal Processing*, **43**, 95-107, 1995.
- [160] Young N., *An introduction to Hilbert Space*, Cambridge U.P., 1988.
- [161] Youla D.C., "On the Factorization of Rational Matrices", *IRE Trans. Inf. Theory*, **IT-7**, 3, 172-189, 1961.
- [162] Youla D.C., Kazanjian, N.N. "Bauer-type Factorization of positive Matrices and the Theory of Matrix Polynomials Orthogonal on the Unit Circle", *IEEE Trans. Circuits and Systems*, **CAS-25**: 57-65, 1978.