

**COMPITO DI
IDENTIFICAZIONE DEI MODELLI ED ANALISI DEI DATI
19 febbraio 2010**

Domanda 1. Sia X una variabile aleatoria Gaussiana $\mathcal{N}(0, 1)$ e siano Y_1 e Y_2 due “misure rumorose” di X nel senso che $Y_1 = X + W_1$ e $Y_2 = X + W_2$ dove W_1 e W_2 sono Gaussiane, indipendenti, indipendenti da X , a media zero e varianza σ^2 .

1. Si trovi la densità di $Y := [Y_1 \ Y_2]^\top$ e la densità congiunta di X e Y .
2. Si calcoli il valore atteso condizionato $\hat{X}(y_1) = E[X|Y_1 = y_1]$ e $\hat{X}(y_1, y_2) = E[X|Y_1 = y_1, Y_2 = y_2]$
3. Si descriva come calcolare $P[X \in [\hat{X}(Y_1, Y_2) - \alpha, \hat{X}(Y_1, Y_2) + \alpha] | Y_1 = y_1, Y_2 = y_2]$. Si può dire (senza fare calcoli) se $P[X \in [-\alpha, \alpha]] > P[X \in [\hat{X}(Y_1) - \alpha, \hat{X}(Y_1) + \alpha] | Y_1 = y_1]$ o $P[X \in [-\alpha, \alpha]] < P[X \in [\hat{X}(Y_1) - \alpha, \hat{X}(Y_1) + \alpha] | Y_1 = y_1]$? Cosa succede per $P[X \in [\hat{X}(Y_1, Y_2) - \alpha, \hat{X}(Y_1, Y_2) + \alpha] | Y_1, Y_2]$. Si commenti il risultato.

Domanda 2. Si assuma che dei dati ingresso-uscita $\{u(t), y(t)\}_{t=1, \dots, N}$ siano campioni di una coppia di processi aleatori congiuntamente stazionari $\{y(t)\}_{t \in \mathbb{Z}}$, $\{u(t)\}_{t \in \mathbb{Z}}$, le cui statistiche del secondo ordine sono compatibili con il modello

$$y(t) = F_0(z)u(t) + G_0(z)e_0(t)$$

dove $e_0(t)$ è un rumore bianco Gaussiano, a media nulla e varianza $\sigma^2 = 1$ e $cov(e_0(t), u(s)) = 0, \forall t \geq s$. Si assuma che $F_0(z)$ sia la funzione di trasferimento di un sistema BIBO stabile e strettamente causale (i.e. $F_0(\infty) = 0$) e che $G_0(z)$ sia BIBO con inversa BIBO (i.e. $G_0(z)$ BIBO e $G_0^{-1}(z)$ BIBO) con $G_0(\infty) = 1$.

1. Si ricavi l'espressione del predittore ad un passo $\hat{y}(t|t-1) := \hat{E}[y(t)|H_t^-(y, u)]$.
2. Assumendo di fare identificazione utilizzando il metodo PEM con la classe di modelli di tipo “Output Error” (OE)

$$y(t) = F_\theta(z)u(t) + e(t)$$

dove

$$F_\theta(z) := \frac{\sum_{i=0}^n b_i z^{n-i}}{\sum_{i=0}^n a_i z^{n-i}} \quad \theta := [a_1, \dots, a_n, b_0, b_1, \dots, b_n], a_0 = 1$$

con il vincolo che $F_\theta(z)$ sia BIBO stabile.

Si assuma che $e_0(t)$ sia scorrelato da $u(s) \forall t, s \in \mathbb{Z}$. Si trovino condizioni sotto le quali lo stimatore $\hat{F}_N(z) := F_{\hat{\theta}_N^{PEM}}(z)$ converge, per $N \rightarrow \infty$, a $F_0(z)$.

Cosa si può dire se $e_0(t)$ è scorrelato da $u(s)$ solo per $t \geq s$?

Domanda 3. Si consideri il modello lineare

$$y(t) = \phi(t)^\top \theta + d(t) \tag{1}$$

dove $\theta \in \mathbb{R}^n$ è un vettore di parametri, $\phi(t)$ è un vettore di grandezze misurabili e $d(t)$ è un rumore bianco.

1. Si scrivano le equazioni del filtro di Kalman per la stima ricorsiva dei parametri θ nel modello lineare (1).
2. Supponendo che i parametri θ varino nel tempo, e avendo a disposizione dei dati di “test” $y(t), \phi(t), t = 1, \dots, N$, come è opportuno “tarare” il filtro di Kalman?
3. Si descriva, in maniera qualitativa, come si potrebbe utilizzare il filtro di Kalman per rivelare un brusco cambiamento (dovuto, ad esempio, ad un malfunzionamento di un sistema) dei valori θ (incogniti).
4. Si illustri come si dovrebbe modificare l’algoritmo descritto sopra basato sul filtro di Kalman se $d(t)$, invece di essere bianco, fosse un processo colorato a spettro noto?

Domanda 4. I dati rilevati da un sensore si possono modellare, in condizioni di funzionamento normale come una variabile Gaussiana a media m_0 e varianza σ_0^2 . Tuttavia il sensore può essere soggetto ad un malfunzionamento. In tal caso, le misure si possono modellare come una variabile aleatoria Gaussiana a media m_1 e varianza σ_0^2 . Il sensore funziona in maniera corretta con probabilità p_c .

1. Sia X la variabile aleatoria che modella l’esito di una misura di questo sensore. Si ricavi la densità di probabilità di X . Supponendo che X_1, \dots, X_N siano N misure indipendenti, si ricavi la densità di probabilità del vettore $[X_1, \dots, X_N]^T$.
2. Si assuma che m_0, m_1, σ_0^2 e p_c siano noti. Sapendo che l’esito di una misura del sensore appartiene ad un insieme Δx , si calcoli la probabilità che il sensore abbia funzionato correttamente.
3. Si assuma che l’insieme Δx sia l’intervallo $\Delta x := [x, x + \Delta]$. Utilizzando l’approssimazione

$$\int_x^{x+\Delta} f(\tau) d\tau \simeq f(x)\Delta$$

ed indicando con C l’evento {“il sensore ha funzionato correttamente”}, si calcoli la probabilità condizionata

$$P[C|X = x]$$

SOLUZIONI

Domanda 1.[9 Punti]

1. Poichè X , W_1 e W_2 sono Gaussiane ed indipendenti, il vettore $Z := [X \ W_1 \ W_2]^\top$ è Gaussiano con media e matrice varianza

$$\mathbb{E}[Z] = \mathbb{E} \begin{bmatrix} X \\ W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Sigma_{zz} := Var \left\{ \begin{bmatrix} X \\ W_1 \\ W_2 \end{bmatrix} \right\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$$

Poichè $[X \ Y_1 \ Y_2]^\top$ si ottiene come trasformazione lineare da Z ,

$$\begin{bmatrix} X \\ Y_1 \\ Y_2 \end{bmatrix} = A \begin{bmatrix} X \\ W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ W_1 \\ W_2 \end{bmatrix}$$

allora anche $[X \ Y_1 \ Y_2]^\top$ è Gaussiano con media a varianza:

$$\mathbb{E} \begin{bmatrix} X \\ Y_1 \\ Y_2 \end{bmatrix} = A \mathbb{E} \begin{bmatrix} X \\ W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

e

$$Var \left\{ \begin{bmatrix} X \\ Y_1 \\ Y_2 \end{bmatrix} \right\} = A Var \left\{ \begin{bmatrix} X \\ W_1 \\ W_2 \end{bmatrix} \right\} A^\top$$

Il vettore $[Y_1 \ Y_2]^\top$ è Gaussiano perchè è un sottovettore del vettore Gaussiano $[X \ Y_1 \ Y_2]^\top$.
La media e la varianza di $[Y_1 \ Y_2]^\top$ risultano:

$$\mathbb{E} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad Var \left\{ \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \right\} = \begin{bmatrix} 1 + \sigma^2 & 1 \\ 1 & 1 + \sigma^2 \end{bmatrix}$$

2. Gli stimatori richiesti si ottengono usando le ben note formule:

$$\hat{X}(y_1) := \mathbb{E}[X|Y_1 = y_1] = E[X] + cov(X, Y_1) Var\{Y_1\}^{-1} (Y_1 - E[Y_1]) = \frac{y_1}{1 + \sigma^2}$$

$$\begin{aligned} \hat{X}(y_1, y_2) &:= \mathbb{E}[X|Y_1 = y_1, Y_2 = y_2] \\ &= E[X] + [cov(X, Y_1) \ cov(X, Y_2)] Var\{Y_1, Y_2\}^{-1} \begin{bmatrix} y_1 - E[Y_1] \\ y_2 - E[Y_2] \end{bmatrix} \\ &= [1 \ 1] \begin{bmatrix} 1 + \sigma^2 & 1 \\ 1 & 1 + \sigma^2 \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= \frac{1}{2 + \sigma^2} (y_1 + y_2) \end{aligned}$$

3. La probabilità richiesta si ottiene da

$$P \left[X \in [\hat{X}(Y_1, Y_2) - \alpha, \hat{X}(Y_1, Y_2) + \alpha] \mid Y_1 = y_1, Y_2 = y_2 \right] = \int_{\hat{X}(y_1, y_2) - \alpha}^{\hat{X}(y_1, y_2) + \alpha} f_{X|Y_1, Y_2}(x|y_1, y_2) dx$$

dove $f_{X|Y_1, Y_2}(x|y_1, y_2)$ è la densità condizionata di X dati Y_1 e Y_2 . Nel caso Gaussiano è noto dalla Teoria che $f_{X|Y_1, Y_2}(x|y_1, y_2)$ è una densità Gaussiana con media $\hat{X}(y_1, y_2)$ e varianza

$$\begin{aligned} Var\{X|Y_1, Y_2\} &= Var\{X\} - [cov(X, Y_1) \ cov(X, Y_2)] Var\{Y_1, Y_2\}^{-1} \begin{bmatrix} cov(X, Y_1) \\ cov(X, Y_2) \end{bmatrix} \\ &= 1 - \frac{2}{2 + \sigma^2} = \frac{\sigma^2}{2 + \sigma^2} \end{aligned}$$

da cui si calcola immediatamente la probabilità richiesta.

Le probabilità che si chiede di confrontare nel testo dell'esercizio sono per probabilità (a priori e a posteriori) di trovare X in un intervallo di ampiezza 2α intorno alla sua media (eventualmente condizionata). Poichè la varianza condizionata è minore della varianza a priori, la probabilità che X appartenga ad un intervallo di ampiezza 2α centrato sulla media è minore della probabilità, condizionata dalla misura di Y_1 , che X appartenga ad un intervallo di ampiezza 2α centrato sulla media condizionata di X dato Y_1 . In particolare la condizione

$$\text{Var}\{X|Y_1, Y_2\} = \frac{\sigma^2}{2 + \sigma^2} < \text{Var}\{X|Y_1\} = \frac{\sigma^2}{1 + \sigma^2} < \text{Var}\{X\} = 1$$

implica che

$$\begin{aligned} P \left[X \in \left[\hat{X}(Y_1, Y_2) - \alpha, \hat{X}(Y_1, Y_2) + \alpha \right] | Y_1, Y_2 \right] &> P \left[X \in \left[\hat{X}(Y_1) - \alpha, \hat{X}(Y_1) + \alpha \right] | Y_1 \right] > \\ &> P \left[X \in [-\alpha, \alpha] \right] \end{aligned}$$

Domanda 2.[8 Punti]

1. Le equazioni del predittore si ottengono come segue:

Poichè F_0 e G_0 sono BIBO allora $y(t)$ appartiene allo spazio $H_t^-(u) + H_{t+1}^-(e_0)$. Analogamente, dividendo per $G_0(z)$ entrambi i membri si ottiene

$$e_0(t) = G_0^{-1}(z)[y(t) - F_0(z)u(t)] \quad (2)$$

Poichè anche G_0^{-1} è BIBO anche $e_0(t)$ appartiene allo spazio $H_t^-(u) + H_{t+1}^-(y)$. Questo dimostra che

$$H_t^-(u, y) = H_t^-(u, e_0) \quad (3)$$

Si osservi che vale

$$y(t) = F_0(z)u(t) + (G_0(z) - 1)e_0(t) + e_0(t)$$

dove $F_0(z)$ e $G_0(z) - 1$ sono strettamente causali. Dalla condizione di ortogonalità $\text{cov}(e_0(t), u(s)) = 0$ per $t \geq s$ e $\text{cov}(e_0(t), e(s)) = 0$ per $t > s$ si ottiene:

$$\begin{aligned} \hat{y}(t|t-1) &= \hat{E}[y(t)|H_t^-(u, y)] = \hat{E}[y(t)|H_t^-(u, e_0)] \\ &= \hat{E}[F_0(z)u(t)|H_t^-(u, e_0)] + \hat{E}[(G_0(z) - 1)e_0(t)|H_t^-(u, e_0)] + \hat{E}[e_0(t)|H_t^-(u, e_0)] \\ &= F_0(z)u(t) + (G_0(z) - 1)e_0(t) \end{aligned}$$

dove nell'ultima equazione si è utilizzato il fatto che $F_0(z)u(t)$ e $(G_0(z) - 1)e_0(t)$ sono funzionali lineari e stabili di (rispettivamente) $H_t^-(u)$ e $H_t^-(e_0)$ e quindi appartengono entrambi a $H_t^-(u, e_0)$. Utilizzando (2) si ottiene

$$\hat{y}(t|t-1) = G_0^{-1}(z) [F_0(z)u(t) + (G_0(z) - 1)y(t)]$$

2. Utilizzando il metodo PEM con la classe di modelli output error

$$y(t) = F_\theta(z)u(t) + e(t),$$

lo stimatore $\hat{\theta}_N^{PEM}$ converge all'insieme dei punti di minimo di

$$\mathbb{E}[(y(t) - \hat{y}_\theta(t|t-1))^2] = \mathbb{E}[(F_0(z) - F_\theta(z))u(t) + G_0(z)e_0(t)]^2$$

se $u(t)$ e $e_0(s)$ sono scorrelati per ogni scelta di t ed s , allora vale

$$\mathbb{E}[(F_0(z) - F_\theta(z))u(t) + G_0(z)e_0(t)]^2 = \mathbb{E}[(F_0(z) - F_\theta(z))u(t)]^2 + \mathbb{E}[G_0(z)e_0(t)]^2 \quad (4)$$

Poichè il secondo addendo della somma non dipende da θ , si può considerare solo il primo termine. L'insieme dei punti di minimo di

$$\mathbb{E}[(F_0(z) - F_\theta(z))u(t)]^2 = \frac{1}{2\pi} \int_0^{2\pi} (F_0(e^{j\theta}) - F_\theta(e^{j\theta}))S_u(e^{j\theta})(F_0(e^{-j\theta}) - F_\theta(e^{-j\theta}))d\theta$$

dipende dallo spettro $S_u(e^{j\theta})$ di $u(t)$ e dalla classe di funzioni F_θ scelta (ad esempio dal grado n nel testo dell'esercizio).

In particolare

- (a) se l'ordine n viene scelto maggiore o uguale al grado di $F_0(z)$ (in modo che esista un vettore di parametri θ_0 per i quali $F_0(z) = F_{\theta_0}(z)$) e,
- (b) se l'ingresso è "persistentemente eccitante" (come abbiamo visto a lezione una condizione sufficiente è che $S_u(e^{j\theta}) > 0$ per ogni θ)

allora

- i) per $\theta = \theta_0$, $\mathbb{E}[(F_0(z) - F_{\theta_0}(z))u(t)]^2 = 0$
- ii) $\mathbb{E}[(F_0(z) - F_\theta(z))u(t)]^2 = 0$ se e solo se $F_0(z) = F_\theta(z)$.

Di conseguenza, sotto le condizioni (a) e (b) sopra lo stimatore $\hat{F}_N(z)$ converge a $F_0(z)$.

Se, come menzionato nel testo, $e_0(t)$ fosse scorrelato solo dal $u(s)$, con $s \leq t$, (questo ad esempio succede se il sistema opera in retroazione, cioè con $u(t)$ che dipende dai valori passati di $y(t)$) allora non sarebbe possibile separare i due termini in (4) e, in generale, si perde la consistenza di $\hat{F}_N(z)$.

Domanda 3.[9 Punti]

1. È sufficiente considerare il modello di stato

$$\begin{cases} \theta(t+1) &= \theta(t) \\ y(t) &= \phi(t)^\top \theta(t) + d(t) \end{cases}$$

e applicare il filtro di Kalman come visto a lezione (vedi note).

2. se il parametro θ varia nel tempo bisogna aggiungere un rumore bianco $n(t)$ all'equazione di stato

$$\theta(t+1) = \theta(t) + n(t)$$

La varianza di $n(t)$ si può aggiustare, ad esempio, utilizzando il test del periodogramma cumulato

3. Il filtro di Kalman fornisce il predittore $\hat{\theta}(t|t-1)$ dal quale si può costruire il predittore dell'uscita

$$\hat{y}(t|t-1) = \phi(t)^\top \hat{\theta}(t|t-1)$$

Si osservi che $y(t) - \hat{y}(t|t-1)$ dovrebbe avere media nulla e varianza

$$\text{Var}\{y(t) - \hat{y}(t|t-1)\} = \phi(t)^\top \text{Var}\{\hat{\theta}(t|t-1)\}\phi(t) + \text{Var}\{d(t)\}$$

Confrontando $\sqrt{\text{Var}\{y(t) - \hat{y}(t|t-1)\}}$ con l'errore $y(t) - \hat{y}(t|t-1)$ si può capire se il parametro ha subito una brusca variazione.

PS: questo procedimento si può rendere rigoroso utilizzando la Teoria dei Test delle Ipotesi; può essere utile guardare gli errori $y(\tau) - \hat{y}(\tau|\tau-1)$ per una "finestra" di valori $\tau \in [t, t+T]$ in modo da rendere più robusta la procedura.

4. Se $d(t)$, invece di essere bianco, ha uno spettro colorato noto (diciamo $S_d(z)$), si può pensare che $d(t)$ sia l'uscita di un sistema lineare tempo invariante, con funzione di trasferimento $W(z)$, i.e.

$$d(t) = W(z)\epsilon(t)$$

dove $\epsilon(t)$ é un rumore bianco a media nulla e varianza unitaria e

$$S_d(z) = W(z)W^\top(1/z).$$

Basta ora procurarsi una realizzazione di stato di $W(z)$, cioè un sistema lineare

$$\begin{cases} x(t+1) &= Ax(t) + B\epsilon(t) \\ d(t) &= Cx(t) + D\epsilon(t) \end{cases}$$

tale che $W(z) = C(zI - A)^{-1}B + D$.

Di conseguenza $y(t)$ si può rappresentare nella forma:

$$\begin{cases} \xi(t+1) &= F\xi(t) + G\epsilon(t) \\ y(t) &= H(t)\xi(t) + D\epsilon(t) \end{cases} \quad (5)$$

dove $\xi(t) := [x^\top(t) \theta^\top(t)]^\top$,

$$F := \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \quad G := \begin{bmatrix} B \\ 0 \end{bmatrix} \quad H(t) := [C \quad \phi^\top(t)].$$

Lo stimatore di $\theta(t)$ si può ottenere applicando l'algoritmo del filtro di Kalman per la stima dello stato $\xi(t)$ nel modello (5).

Domanda 4.[8 Punti]

1. Indicando con C l'evento "il sensore ha funzionato correttamente", ed utilizzando il Teorema della probabilità totale, si ottiene:

$$P[X \leq x] = P[X \leq x|C]P[C] + P[X \leq x|\bar{C}]P[\bar{C}].$$

Quindi la funzione di distribuzione di X è

$$F_X(x) = F_{X|C}(x|C)p_C + F_{X|\bar{C}}(x|\bar{C})(1 - p_C)$$

e quindi la densità di X

$$f_X(x) = f_{X|C}(x|C)p_C + f_{X|\bar{C}}(x|\bar{C})(1 - p_C).$$

Le densità $f_{X|C}(x|C)$ ed $f_{X|\bar{C}}(x|\bar{C})$ sono, dal testo dell'esercizio:

$$f_{X|C}(x|C) = \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2} \frac{(x-m_0)^2}{\sigma_0^2}}$$

$$f_{X|\bar{C}}(x|\bar{C}) = \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2} \frac{(x-m_1)^2}{\sigma_0^2}}$$

e quindi:

$$f_X(x) = p_C \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2} \frac{(x-m_0)^2}{\sigma_0^2}} + (1 - p_C) \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2} \frac{(x-m_1)^2}{\sigma_0^2}}$$

Se X_1, \dots, X_N sono N misure indipendenti, la densità congiunta si ottiene come prodotto delle densità delle variabili X_i , i.e.

$$f_{X_1, \dots, X_N}(x_1, \dots, x_N) = \prod_{i=1}^N \left[p_C \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2} \frac{(x_i-m_0)^2}{\sigma_0^2}} + (1 - p_C) \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2} \frac{(x_i-m_1)^2}{\sigma_0^2}} \right]$$

2. La probabilità richiesta si ottiene utilizzando la regola di Bayes:

$$\begin{aligned} P[C|X \in \Delta x] &= \frac{P[x \in \Delta x|C]P[C]}{P[x \in \Delta x]} \\ &= \frac{p_C \int_{\Delta x} f_{X|C}(x|C) dx}{\int_{\Delta x} f_X(x) dx} \end{aligned}$$

3. Utilizzando l'approssimazione suggerita si ha:

$$P[C|X \in [x, x + \Delta)] \simeq \frac{f_{X|C}(x|C)p_C \Delta}{(f_{X|C}(x|C)p_C + f_{X|\bar{C}}(x|\bar{C})(1-p_C))\Delta}$$

e facendo tendere Δ a zero si ottiene:

$$P[C|X = x] = \frac{f_{X|C}(x|C)p_C}{f_{X|C}(x|C)p_C + f_{X|\bar{C}}(x|\bar{C})(1-p_C)}$$