Capitolo 3

Matrici polinomiali

Maggio 2004
Per affrontare lo studio dei sistemi con più ingressi ed uscite è necessario studiare la struttura delle matrici polinomiali e razionali che saranno impiegate nella descrizione dei legami funzionali tra le variabili in gioco e nella progettazione di leggi di controllo. Lo strumento fondamentale cui faremo ricorso è quello delle operazioni elementari sulle righe e sulle colonne. Esse permettono di costruire importanti forme canoniche e, più in generale, di operare su una matrice modifiche che ne evidenzino le caratteristiche intrinseche (quali la primalità) o le attribuiscano ulteriori proprietà, quali la "riduzione per colonne".
In questo capitolo ci occupiamo di matrici polinomiali. Le nozioni introdotte trovano una diretta applicazione nell'ambito di importanti problemi, quali la modellizzazione delle traiettorie di un sistema lineare nell'approccio "behavior", lo studio dei codici convoluzionali e, in un contesto più vicino alla teoria del controllo, la sintesi di controllori dead-beat per un sistema multivariabile. Inoltre esse costituiscono la necessaria premessa per studiare, nel capitolo seguente, le matrici di trasferimento razionali e trattare poi numerosi argomenti di tipo sistemistico.

Salvo contrario avviso, nel corso di questo capitolo $\mathbb{F}$ sarà un campo numerico arbitrario (tipicamente $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, ma anche un campo finito). Come è consuetudine, indicheremo con $\mathbb{F}[z]^{p\times m}$ l'insieme delle matrici $p \times m$ ad elementi in $\mathbb{F}[z]$.

3.1 Matrici elementari e forme canoniche

Definizione 3.1.1 [Matrice polinomiale singolare] *Una matrice $M(z) \in \mathbb{F}[z]^{m\times m}$ è* **singolare** *se il suo determinante, $\det M(z)$, è il polinomio nullo, cioè il polinomio a coefficienti tutti nulli. In caso contrario $M(z)$ è detta* **non singolare**.

Esempio 3.1.1 La matrice

$$M_1(z) = \begin{bmatrix} z+1 & 1 \\ z+1 & z+1 \end{bmatrix} \in \mathbb{F}[z]^{2\times 2}$$

è non singolare, dal momento che il suo determinante è il polinomio non nullo $\det M_1(z) = (z+1)^2 - (z+1) = z^2 + z$. Si noti che la funzione polinomiale $\det M_1(z)$ ha uno zero in $z = -1$ e uno in

$z = 0$ e quindi $M_1(-1)$ e $M_1(0)$ sono matrici singolari in $\mathbb{F}^{2\times 2}$, qualunque sia il campo $\mathbb{F}$. Pertanto una matrice polinomiale non singolare può dar luogo ad una matrice numerica singolare quando si interpreti l'indeterminata z come variabile e ad essa si attribuiscano opportuni valori. Nel caso in questione, se $\mathbb{F}$ è il campo di Galois a due elementi, la funzione polinomiale $z^2 + z$ assume anzi il valore zero per ogni valore della variabile: ciononostante, in base alla definizione data, la matrice polinomiale $M_1(z)$ è non singolare.

Invece

$$M_2(z) = \begin{bmatrix} z-1 & z+1 \\ z-1 & z+1 \end{bmatrix} \in \mathbb{F}[z]^{2\times 2}$$

ha $\det M_2(z) = (z-1)(z+1) - (z-1)(z+1) = 0$, il polinomio nullo, e quindi è una matrice polinomiale singolare.

Definizione 3.1.2 [Rango di una matrice polinomiale] *Una matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ ha* **rango** *r se i suoi minori di ordine $r+1$ sono polinomi nulli, mentre esiste (almeno) un minore di ordine r non nullo. Chiaramente il rango di una matrice $p \times m$ è minore o uguale al più piccolo tra p ed m; qualora esso coincida con tale limite superiore, la matrice è detta* **a rango pieno**.

La nozione di rango di una matrice si presta ad una immediata interpretazione in termini di spazi vettoriali: se infatti pensiamo alle colonne della matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ come a vettori in $\mathbb{F}(z)^p$, il rango di $M(z)$ rappresenta la dimensione del sottospazio di $\mathbb{F}(z)^p$ generato dalle colonne di $M(z)$ e del sottospazio vettoriale di $\mathbb{F}(z)^m$ generato dalle righe di $M(z)$.

Esempio 3.1.2 La matrice

$$M(z) = \begin{bmatrix} z+1 & z & z+2 \\ z^2-1 & z^2-z & z^2+z-2 \end{bmatrix}$$

ha rango 1. Infatti essa ha due righe non nulle, ma linearmente dipendenti su $\mathbb{R}(z)$, o, equivalentemente, contiene elementi non nulli ma tutti i suoi minori di ordine due sono polinomi nulli. Le righe della matrice sono dipendenti anche sul campo $\mathbb{R}$?

La prima classe di matrici polinomiali che considereremo è quella delle *matrici elementari*. Esse sono matrici quadrate polinomiali che assumono una delle seguenti tre forme

$$\tilde{E}_1(z) := \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & \alpha & \\ & & & & 1 \\ & & & & & \ddots \\ 0 & & \cdots & & 0 & 1 \end{bmatrix}, \quad \alpha \in \mathbb{F} \setminus \{0\},$$

$$\tilde{E}_2(z) := \begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 0 & \cdots & 1 & & & \\ & & & \vdots & \ddots & \vdots & & & \\ & & & 1 & \cdots & 0 & & & \\ & & & & & & \ddots & & \\ & & & & & & & 1 & \\ & & & & & & & & 1 \end{bmatrix},$$

$$\tilde{E}_3(z) := \begin{bmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & \dots & p(z) & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}, \qquad p(z) \in \mathbb{F}[z].$$

- Poiché il determinante delle matrici elementari è una costante non nulla, esse sono invertibili con inversa polinomiale. Tale inversa è una matrice elementare del medesimo tipo, ovvero l'inversa di $\tilde{E}_i$ è ancora una matrice di tipo i.

- Se $M(z)$ è un'arbitraria matrice polinomiale di dimensioni $p \times m$, la (post)moltiplicazione di $M(z)$ per le matrici elementari $\tilde{E}_i(z) \in \mathbb{F}[z]^{m\times m}$, $i = 1, 2, 3$, equivale ad una delle seguenti operazioni elementari, realizzate sulle colonne di $M(z)$:

  1) la moltiplicazione di $M(z)$ per $\tilde{E}_1$ corrisponde a moltiplicare la colonna k-esima di $M(z)$ per la costante non nulla α, se α è in posizione (k, k);

  2) se i termini unitari fuori diagonale in $\tilde{E}_2$ sono in posizione (h, k) e (k, h), la moltiplicazione di $M(z)$ per $\tilde{E}_2$ corrisponde a permutare la colonna k-esima con la colonna h-esima di $M(z)$;

  3) la moltiplicazione di $M(z)$ per $\tilde{E}_3$ corrisponde a sommare alla colonna k-esima di $M(z)$ la h-esima moltiplicata per $p(z)$, se $p(z)$ in $\tilde{E}_3$ appare in posizione (h, k).

  Di conseguenza, $\tilde{E}_i(z) \in \mathbb{F}[z]^{m\times m}$ può essere vista come un'operatore lineare che "agendo a destra" sullo spazio delle matrici polinomiali $\mathbb{F}[z]^{p\times m}$:

  $$\tilde{E}_i(z) \quad : \quad \mathbb{F}[z]^{p\times m} \to \mathbb{F}[z]^{p\times m} : M(z) \mapsto M(z)\tilde{E}_i(z).$$

  induce operazioni elementari sulle colonne delle matrici.

- Similmente, premoltiplicando una matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ per un'opportuna matrice elementare di dimensioni $p \times p$, è possibile realizzare le analoghe operazioni sulle righe di $M(z)$. In tal caso $\tilde{E}_i(z)$ viene associato ad un operatore lineare che "agisce a sinistra" sullo spazio delle matrici $\mathbb{F}[z]^{p\times m}$.

Definizione 3.1.3 [Equivalenza fra matrici polinomiali] *Siano $M(z)$ e $N(z)$ due matrici polinomiali in $\mathbb{F}[z]^{p\times m}$; diciamo che $M(z)$ è* **equivalente** *a $N(z)$ (e scriviamo*

$M(z) \sim N(z)$) se $M(z)$ può essere ottenuta da $N(z)$ attraverso una successione di operazioni elementari realizzate sia sulle righe che sulle colonne. In altre parole, devono esistere matrici elementari $E_1(z), E_2(z), \ldots, E_k(z) \in \mathbb{F}[z]^{p\times p}$ e $E'_1(z), E'_2(z), \ldots, E'_h(z) \in \mathbb{F}[z]^{m\times m}$ tali che

$$E_k(z) \ldots, E_2(z)E_1(z)N(z)E'_1(z)E'_2(z) \ldots E'_h(z) = M(z). \tag{3.1}$$

La relazione ora introdotta in $\mathbb{F}[z]^{p\times m}$ gode delle proprietà riflessiva, simmetrica e transitiva, ed è quindi effettivamente una relazione di equivalenza. L'insieme $\mathbb{F}[z]^{p\times m}$ ne risulta partizionato in classi disgiunte, a ciascuna delle quali appartengono tutte e sole le matrici polinomiali di dimensioni $p \times m$ che possono essere ottenute l'una dall'altra attraverso operazioni elementari su righe e colonne.

- ESERCIZIO 3.1.1 (*i*) Si verifichi che $\sim$ è una relazione di equivalenza.

  (*ii*) Si dimostri che la matrice elementare

  $$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

  è esprimibile come prodotto di matrici elementari 2×2 degli altri due tipi.
  (Suggerimento: si consideri il prodotto $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$*)*
  Si estenda il risultato al caso di arbitrarie matrici $n \times n$ del tipo $\tilde{E}_2(z)$, verificando così che la relazione di equivalenza $\sim$ è definibile con riferimento alle sole matrici di tipo $\tilde{E}_1(z)$ e $\tilde{E}_3(z)$.

  (*iii*) Se $f(z)$ e $g(z)$ sono polinomi in $\mathbb{F}[z]$, il secondo dei quali non nullo, allora esiste una matrice elementare che applicata al vettore $[f(z)\ g(z)]^T$ produce il vettore $[r(z)\ g(z)]^T$, dove $r(z)$ è il resto della divisione euclidea di $f(z)$ per $g(z)$.
  (Suggerimento: $[\, r(z) \quad g(z)\,] = [\, f(z) \quad g(z)\,]\begin{bmatrix} 1 & 0 \\ -q(z) & 1 \end{bmatrix}$*)*

  (*iv*) Se $f(z)$ e $g(z)$ sono due polinomi non nulli in $\mathbb{F}[z]$, l'algoritmo di Euclide per l'estrazione del MCD di $f(z)$ e $g(z)$ è esprimibile attraverso una successione di operazioni elementari che, applicate al vettore $[f(z)\ g(z)]^T$, portano al vettore $[d(z)\ 0]^T$, con $d(z) = \mathrm{MCD}(f(z), g(z))$.

Come è consuetudine per le relazioni di equivalenza, conviene ricercare all'interno di ciascuna delle classi $\mathbb{F}[z]^{p\times m}/\sim$ un elemento di struttura particolarmente semplice, che possa essere utilizzato come rappresentativo dell'intera classe. Il seguente teorema mostra come ogni matrice polinomiale $M(z)$ sia equivalente ad una matrice in *forma canonica di Smith* di $M(z)$, i cui unici elementi non nulli si trovano sulla "pseudodiagonale" passante per la posizione $(1, 1)$.

Teorema 3.1.4 [FORMA CANONICA DI SMITH] *Sia $M(z)$ una matrice in $\mathbb{F}[z]^{p\times m}$. Esiste una successione finita di trasformazioni elementari che, applicate alle righe e alle colonne di $M(z)$, la riducono alla forma:*

$$\Gamma(z) \quad = \quad E_k(z) \ldots E_1(z)M(z)E'_1(z) \ldots E'_h(z)$$

$$= \left[\begin{array}{ccccc|c} \gamma_1(z) & & & & & \\ & \gamma_2(z) & & & & \\ & & \ddots & & & 0 \\ & & & \gamma_r(z) & & \\ \hline & & 0 & & & 0 \end{array}\right]$$

$$= \operatorname{diag}\{\gamma_1(z), \ldots, \gamma_r(z)\}_{p\times m}, \tag{3.2}$$

dove r è il rango di $M(z)$ e $\gamma_1(z), \gamma_2(z), \ldots, \gamma_r(z) \in \mathbb{F}[z]$ sono polinomi monici che soddisfano $\gamma_i(z) \mid \gamma_{i+1}(z)$, per $i = 1, 2, ..., r-1$.
I polinomi $\gamma_i(z)$ sono univocamente determinati dalle condizioni precedenti.

DIMOSTRAZIONE [Esistenza] Proveremo l'esistenza della forma di Smith in modo costruttivo, fornendo un algoritmo che permette di ottenerla in un numero finito di passi.

(A) Se $M(z)$ è la matrice nulla, è già in forma di Smith.

(B) Altrimenti, con uno scambio di righe e/o di colonne si porta in posizione $(1,1)$ un suo elemento non nullo di grado minimo. Detta

$$M_1(z) = E_1(z)M(z)E_1'(z)$$

la matrice così ottenuta, se gli altri elementi della prima riga e della prima colonna sono nulli, si va al successivo passo (C). Altrimenti, esiste un elemento non nullo, p.es. nella prima colonna, $m_{i1}(z)$, di grado non inferiore a $\deg m_{11}(z)$. Applicando l'algoritmo di divisione dei polinomi esso può essere espresso nella forma

$$m_{i1}(z) = q_{i1}(z)m_{11}(z) + r_{i1}(z) \qquad \text{con } r_{i1} = 0 \text{ oppure } 0 \le \deg r_{i1} < \deg m_{11}.$$

Si somma allora alla riga i-esima di $M_1(z)$ la prima moltiplicata per $-q_{i1}(z)$ ottenendo una matrice $M_2(z)$ che ha $r_{i1}(z)$ in posizione $(i,1)$.
Si considera a questo punto l'intera matrice $M_2(z)$.
Se contiene elementi non nulli con grado minore di $\deg m_{11}$, se ne sceglie uno di grado minimo, lo si porta in posizione $(1,1)$ con scambi di righe e colonne e si ripete la procedura. Questa può essere applicata solo un numero finito di volte, dato che ogni volta diminuisce il grado dell'elemento in posizione $(1,1)$. Si arriva cosìa una matrice del tipo

$$M_3(z) = \begin{bmatrix} m_{11}(z) & 0 \\ 0 & M_{22}(z) \end{bmatrix}. \tag{3.3}$$

in cui gli elementi non nulli di $M_{22}(z)$ hanno tutti grado non inferiore a quello di $m_{11}(z)$.

(C) Se un elemento $m_{ij}(z)$ di $M_{22}(z)$ non è multiplo di $m_{11}(z)$, applicando l'algoritmo di divisione dei polinomi può essere espresso nella forma

$$m_{ij}(z) = q_{ij}(z)m_{11}(z) + r_{ij}(z) \qquad 0 \le \deg r_{i1} < \deg m_{11}$$

Sommando la i-esima riga alla prima e poi la prima colonna, moltiplicata per $-q_{ij}(z)$, alla j-esima, si ottiene in posizione $(1,j)$ l'elemento $r_{ij}(z)$.

A questo punto si ricomincia la procedura a partire dal punto (B), fino ad ottenere una matrice con la struttura (3.3), in cui tutti gli elementi di $M_{22}(z)$ sono multipli di $m_{11}(z)$ e quest'ultimo (eventualmente moltiplicando per una matrice elementare) può supporsi monico:

$$\begin{bmatrix} \gamma_1(z) & 0 \\ 0 & \gamma_1(z)H_2(z) \end{bmatrix}. \tag{3.4}$$

(D) Si ripetono le operazioni elementari dei passi (B) e (C) sulle righe $2, 3, ...p$ e sulle colonne $2, 3, ..., m$, che interessano solo la matrice $\gamma_1(z)H_2(z)$. Poiché tutti i suoi elementi sono multipli di $\gamma_1(z)$, si perviene ad una matrice

$$\begin{bmatrix} \gamma_1(z) & 0 & 0 \\ 0 & \gamma_2(z) & 0 \\ 0 & 0 & \gamma_2(z)H_3(z) \end{bmatrix}. \tag{3.5}$$

in cui $\gamma_1(z)|\gamma_2(z)$, $\gamma_1(z)$ e $\gamma_2(z)$ sono monici e $H_3(z) \in \mathbb{F}[z]^{(p-2)\times(m-2)}$, e così via fino ad ottenere la matrice (3.2).

[Unicità] Si tratta ora di dimostrare che il rango r di $M(z)$ coincide con il numero di polinomi non nulli $\gamma_i(z)$ e che i polinomi $\gamma_i(z)$ in $\Gamma(z)$, $i = 1, 2, ..., r$, forniti dal procedimento ora accennato, sono univocamente determinati. A tal fine è sufficiente verificare che valgono le seguenti relazioni

$$\begin{aligned} \gamma_1(z) &= \text{MCD monico dei minori di ordine 1 di } M(z) \\ \gamma_1(z)\gamma_2(z) &= \text{MCD monico dei minori di ordine 2 di } M(z) \\ \cdots &\quad \cdots \\ \gamma_1(z)\gamma_2(z)\cdots\gamma_r(z) &= \text{MCD monico dei minori di ordine r di } M(z). \end{aligned} \tag{3.6}$$

Poichè

$$\Gamma(z) \;=\; E_k(z)\dots E_1(z)M(z)E_1'(z)\dots E_h'(z) \;=\; U(z)M(z)V(z),$$

per il teorema di Binet-Cauchy [1] si ha che per $i = 1, 2, ..., r$ il MCD dei minori di ordine i di $M(z)$ è un divisore del MCD dei minori di ordine i di $\Gamma(z)$. D'altra parte, essendo

[1] Il teorema afferma che se F e G sono matrici ad elementi in un anello commutativo R, di dimensioni rispettivamente $n \times q$ e $q \times n$ con $n \le q$, allora $\det(FG) = \sum_{\mathbf{x}} \det(F_{\mathbf{x}}) \det(G_{\mathbf{x}})$, dove $\mathbf{x}$ descrive tutte le n-uple $(\nu_1, \nu_2, \dots, \nu_n)$ con $1 \le \nu_1 < \nu_2 \dots < \nu_n \le q$, $F_{\mathbf{x}}$ è la sottomatrice di F costituita dalle colonne indiciate da $\mathbf{x}$ e $G_{\mathbf{x}}$ è la sottomatrice di G costituita dalle righe indiciate da $\mathbf{x}$.
Si consideri ora il prodotto di tre matrici $M = FGH$, $F \in R^{p\times q}$, $G \in R^{q\times r}$, $H \in R^{r\times s}$. Se $n \le \min\{p, q, r, s\}$, se $\mathbf{a}$ e $\mathbf{b}$ sono due n-uple

$$\mathbf{a} = (\nu_1, \nu_2, \dots, \nu_n),\ 1 \le \nu_1 < \nu_2 \dots < \nu_n \le p, \quad \mathbf{b} = (\mu_1, \mu_2, \dots, \mu_n),\ 1 \le \mu_1 < \mu_2 \dots < \mu_n \le s$$

e se $m(M_{\mathbf{a},\mathbf{b}})$ denota il minore relativo alle righe indiciate da $\mathbf{a}$ e alle colonne indiciate da $\mathbf{b}$ nella matrice M, dal teorema di Binet si ha

$$m(M_{\mathbf{a},\mathbf{b}}) = \sum_{\mathbf{x},\mathbf{y}} m(F_{\mathbf{a},\mathbf{x}})m(G_{\mathbf{x},\mathbf{y}})m(H_{\mathbf{y},\mathbf{b}})$$

dove $\mathbf{x}$ e $\mathbf{y}$ descrivono tutte le n-uple crescenti estratte da $\{1, 2, \dots, q\}$ e da $\{1, 2, \dots, r\}$.

$E_1(z), \ldots, E_k(z), E'_1(z), \ldots, E'_h(z)$ matrici elementari, $U(z)$ e $V(z)$ sono dotate di inversa polinomiale e vale

$$M(z) = U^{-1}(z)\Gamma(z)V^{-1}(z),$$

da cui segue subito che il MCD dei minori di ordine i di $\Gamma(z)$ divide il MCD dei minori di ordine i di $M(z)$. Quindi i due MCD devono coincidere.
Per concludere la dimostrazione basta osservare che, in virtù della condizione di divisibilità $\gamma_i(z)|\gamma_{i+1}(z)$, per $i = 1, 2, ..., r-1$, il MCD dei minori di ordine i di $\Gamma(z)$, e quindi di $M(z)$, coincide con $\gamma_1(z)\gamma_2(z)\cdots\gamma_i(z)$. ■

Dal momento che per la precedente proposizione ogni matrice è equivalente a un'unica forma di Smith, ogni classe in $\mathbb{F}[z]^{p\times m}/\sim$ contiene una ed una sola forma di Smith. Pertanto le forme di Smith costituiscono una famiglia di *forme canoniche* in $\mathbb{F}[z]^{p\times m}/\sim$. I polinomi monici $\gamma_i(z)$ che compaiono nella forma di Smith di una matrice $M(z)$ sono detti *polinomi invarianti* di $M(z)$, e rappresentano una famiglia di invarianti completi per la relazione $\sim$, nel senso che ogni classe di equivalenza è univocamente individuata da essi.

Si noti che la determinazione della forma di Smith di una matrice non richiede di effettuare esplicitamente le operazioni elementari sopra descritte: qualora ci si proponga di ottenere la forma di Smith di $M(z)$, ma non la successione di operazioni elementari che realizza l'equivalenza fra $M(z)$ e $\Gamma(z)$, è possibile calcolare i polinomi invarianti direttamente dalla matrice $M(z)$, facendo uso delle (3.6).
Sottolineiamo, infine, che la forma canonica di Smith di una matrice è unica, ma il procedimento per ottenerla, e quindi le matrici elementari che compaiono nella (3.2), sono lungi dall'esserlo.

- ESERCIZIO 3.1.2 Sia $M(z)$ una matrice polinomiale quadrata non singolare e sia $\det M(z) = m_1(z)m_2(z)\cdots m_t(z)$ una arbitraria fattorizzazione del determinante di $M(z)$. Si dimostri, ricorrendo alla forma di Smith, che è sempre possibile trovare per $M(z)$ una fattorizzazione

  $$M(z) = M_1(z)M_2(z)\cdots M_t(z)$$

  con $M_i(z)$ matrici polinomiali quadrate con determinante $m_i(z)$, $i = 1, 2, \ldots, t$.

- ESERCIZIO 3.1.3* Sia F una matrice in $\mathbb{C}^{n\times n}$ e sia

  $$J = \mathrm{diag}\{J_{11}, \ldots, J_{1\nu_1}, J_{21}, \ldots, J_{2\nu_2}, \ldots, J_{t1}, \ldots, J_{t\nu_t}\}$$

  la sua forma di Jordan, dove J_{hk}, $k = 1, 2, \ldots, \nu_h$, è un miniblocco di Jordan relativo all'autovalore λ_h, ha dimensioni $m_{hk} \times m_{hk}$, e per $h = 1, 2, \ldots, t$ si ha

  $$m_{h1} \geq m_{h2} \geq \ldots \geq m_{h\nu_t}.$$

  Sia $c = \max\{\nu_1, \nu_2, \ldots, \nu_t\}$ e si ponga $m_{hk} = 0$ per $k > \nu_h$. Si provi che i polinomi

  $$\begin{aligned}
  \psi_1(z) &= (z-\lambda_1)^{m_{11}}(z-\lambda_2)^{m_{21}}\cdots(z-\lambda_t)^{m_{t1}} \\
  \psi_2(z) &= (z-\lambda_1)^{m_{12}}(z-\lambda_2)^{m_{22}}\cdots(z-\lambda_t)^{m_{t2}} \\
  \ldots &\quad \ldots \\
  \psi_c(z) &= (z-\lambda_1)^{m_{1c}}(z-\lambda_2)^{m_{2c}}\cdots(z-\lambda_t)^{m_{tc}}
  \end{aligned}$$

  coincidono con i polinomi invarianti non unitari della matrice $zI - F$. Questo giustifica il fatto di chiamare polinomi invarianti anche i $\psi_i(z)$, sebbene essi siano in genere un sottoinsieme proprio dei $\gamma_i(z)$, che non include i polinomi unitari.

Mediante operazioni elementari sulle righe è possibile ridurre ogni matrice polinomiale ad un'altra struttura della quale faremo frequente uso nel seguito, la *forma di Hermite per colonne*. Val la pena di notare che tale forma è ottenibile facendo uso soltanto di operazioni elementari sulle righe della matrice e pertanto non rappresenta una forma canonica per la relazione di equivalenza $\sim$ prima introdotta. Dimostreremo il risultato soltanto per matrici polinomiali il cui rango coincida con il numero delle colonne e accenneremo all'estensione di tale forma ad arbitrarie matrici polinomiali. Una discussione completa si può reperire in [4].

Teorema 3.1.5 [Forma di Hermite (per colonne)] *Sia $M(z)$ una matrice in $\mathbb{F}[z]^{p\times m}$ di rango m. Allora esiste una famiglia di matrici elementari $E_1(z)$, $E_2(z), \ldots, E_k(z) \in \mathbb{F}[z]^{p\times p}$ tali che*

$$H(z) := E_k(z)\cdots E_2(z)E_1(z)M(z) = \left[\begin{array}{cccc} h_{11}(z) & * & * & * \\ & h_{22}(z) & * & * \\ & & \ddots & * \\ & & & h_{mm}(z) \\ \hline & & 0 & \end{array}\right], \tag{3.7}$$

con $h_{jj}(z)$ polinomio monico soddisfacente[2] $\deg h_{jj} > \deg h_{ij}$ *per* $j = 1, 2, \ldots, m$ *e* $i < j$.

Dimostrazione Selezioniamo nella prima colonna di $M(z)$ un elemento non nullo di grado minimo e con una permutazione di righe lo portiamo in posizione $(1,1)$. Consideriamo ora la divisione euclidea di ogni elemento non nullo in prima colonna per l'elemento $m_{11}(z)$:

$$m_{i1}(z) = q_{i1}(z)m_{11}(z) + r_{i1}(z), \quad \deg r_{i1} < \deg m_{11}.$$

Con operazioni elementari sulle righe della matrice, e precisamente con premoltiplicazioni per matrici elementari del tipo $\tilde{E}_3(z)$, sostituiamo a ciascun elemento $m_{i1}(z)$ il corrispondente resto nella divisione per $m_{11}(z)$. Se almeno uno dei resti è diverso da zero operiamo una permutazione delle righe che porti in posizione $(1,1)$ un elemento non nullo di grado minimo della prima colonna e ripetiamo le operazioni di divisione. In un numero finito di passi giungiamo ad una matrice $M'(z)$ che in prima colonna ha tutti elementi nulli all'infuori del primo, e questo può esser reso monico mediante un'ulteriore operazione elementare sulle righe

$$M'(z) = \begin{bmatrix} m'_{11}(z) & *\,*\,* \\ 0 & M'_2(z) \end{bmatrix}.$$

[2] per convenzione, $\deg(0) = -\infty$

A questo punto applichiamo ricorsivamente alla sottomatrice $M_2'(z)$ le stesse operazioni applicate prima a $M(z)$, e così via fino ad ottenere una matrice triangolare superiore

$$M''(z) = \left[\begin{array}{cccc} m_{11}''(z) & * & * & * \\ & m_{22}''(z) & * & * \\ & & \ddots & * \\ & & & m_{mm}''(z) \\ \hline & & 0 & \\ & & & \end{array}\right].$$

Supponiamo ora che la j-esima colonna di $M''(z)$, $j \in \{2, 3, \ldots, m-1\}$, sia la prima in cui esiste un elemento $m_{ij}''(z)$ per il quale sia $\deg m_{ij}'' \geq \deg m_{jj}''$. Applicando l'algoritmo di divisione si ha $m_{ij}''(z) = q_{ij}(z)m_{jj}''(z) + r_{ij}(z)$. Sommando alla i-esima riga la j-esima moltiplicata per $-q_{ij}(z)$, si sostituisce all'elemento m_{ij}'' il resto r_{jj}, senza alterare nessuna delle colonne precedenti.Procedendo in questo modo si perviene in un numero finito di passi alla forma di Hermite. ■

Se la matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ non ha rango pieno di colonna, in particolare se il numero delle colonne eccede il numero delle righe, la sua forma di Hermite assume un aspetto leggermente diverso. Detto r il rango di $M(z)$, si permutano le colonne della matrice $M(z)$ assegnata, in modo da portare nelle prime r posizioni r colonne indipendenti : $M(z)\Pi = \bar{M}(z)$, con Π matrice di permutazione. Si applicano successivamente operazioni elementari di riga alla matrice$M(z)$, operando sulle sue prime r colonne analogamente a quanto già descritto nel Teorema 3.1.5. Si perviene così ad una matrice $\bar{H}(z) = E_1(z), E_2(z), \ldots, E_k(z)\bar{M}(z)$ della forma

$$\left[\begin{array}{cccccc} h_{11}(z) & * & * & * & * & * \\ & h_{22}(z) & * & * & * & * \\ & & \ddots & * & * & * \\ & & & h_{rr}(z) & \ldots & h_{rm}(z) \\ & & & & & \\ & & & & & \\ & & & 0 & & \end{array}\right] \begin{array}{l} \Big\}r \\ \\ \Big\}p-r \end{array}, \tag{3.8}$$

con $h_{jj}(z)$ polinomio monico soddisfacente $\deg h_{jj} > \deg h_{ij}$ per $j = 1, 2, \ldots, r$ e $i < j$. Sulle ultime $m - r$ colonne della matrice $\bar{H}(z)$ possiamo affermare (per motivi di rango) che le componenti relative alle ultime $p - r$ righe sono nulle, ma in generale non possiamo fare nessuna affermazione circa i gradi delle componenti non nulle.
Ovviamente, se non si effettuano permutazioni di colonna, la matrice cui si perviene con le sole operazioni di riga è $H(z) = \bar{H}(z)\Pi^{-1}$, in cui le colonne sono permutate rispetto a quelle di $\bar{H}(z)$.

- ESERCIZIO 3.1.4 i)* Si verifichi che la forma di Hermite è unica. (Suggerimento: se $H(z)$ e $H'(z)$ sono forme di Hermite della stessa matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ allora

$$H(z) = U(z)H'(z),$$

con

$$U(z) = \begin{bmatrix} I_m & U_{12}(z) \\ 0 & U_{22}(z) \end{bmatrix}.)$$

ii) Si discuta nei dettagli l'algoritmo per costruire la (3.8).

3.2 Matrici unimodulari

Le matrici unimodulari sono gli elementi invertibili nell'anello non commutativo delle matrici quadrate polinomiali e svolgono in tale ambito un ruolo analogo a quello che le costanti non nulle hanno nella fattorizzazione dei polinomi. È ovvio che le matrici elementari introdotte nel precedente paragrafo rientrano tra le matrici unimodulari; in realtà esse costituiscono gli ingredienti essenziali per costruire le matrici unimodulari dal momento che ogni matrice unimodulare in $\mathbb{F}[z]^{m\times m}$ fattorizza nel prodotto di matrici elementari.

Definizione 3.2.1 *Una matrice $U(z)$ in $\mathbb{F}[z]^{m\times m}$ è unimodulare se è invertibile in $\mathbb{F}[z]^{m\times m}$.*

La seguente proposizione fornisce varie caratterizzazioni delle matrici unimodulari, valide per un arbitrario campo $\mathbb{F}$.

Proposizione 3.2.2 [CONDIZIONI EQUIVALENTI ALLA UNIMODULARITÀ] *Sia $U(z)$ una matrice in $\mathbb{F}[z]^{m\times m}$. Sono equivalenti i seguenti fatti:*
i) $U(z)$ è una matrice unimodulare;
ii) $\det U(z)$ è una costante non nulla;
iii) la forma canonica di Smith di $U(z)$ è I_m;
iv) la forma di Hermite di $U(z)$ è I_m;
v) $U(z)$ è esprimibile come prodotto di matrici elementari.

DIMOSTRAZIONE i) $\Rightarrow$ ii) Se $U(z)$ è unimodulare, $U^{-1}(z)$ esiste ed è polinomiale. Pertanto, dalla relazione

$$1 = \det U(z) \det U^{-1}(z)$$

segue che $\det U$ e $\det U^{-1}$ sono entrambi costanti non nulle.

ii) $\Rightarrow$ iii) È sufficiente notare che $\det U$ è l'unico minore di ordine m e quindi coincide, a meno di una costante moltiplicativa non nulla, con il prodotto $\gamma_1(z)\gamma_2(z)\dots\gamma_m(z)$. Di conseguenza, i polinomi invarianti sono m e tutti unitari.

iii) $\Rightarrow$ iv) Poichè la forma di Smith $\Gamma(z) = I$ di $U(z)$ è legata ad $U(z)$ dalla relazione $I = E_k(z)\cdots E_1(z)U(z)E'_1(z)\dots E'_h(z)$, dove $E_1(z),\dots,E_k(z)$ e $E'_1(z),\dots,E'_h(z)$ sono opportune matrici elementari, si ha

$$E_k(z)\cdots E_1(z)\ U(z) = E'^{-1}_h(z)\cdots E'^{-1}_1(z) \tag{3.9}$$

e quindi

$$E'_1(z)\dots E'_h(z)E_k(z)\cdots E_1(z)\ U(z) = I_m \tag{3.10}$$

Pertanto la forma di Hermite di $U(z)$ è I_m.

iv) $\Rightarrow$ v) Per l'ipotesi iv) , esistono matrici elementari $E_1(z), \ldots, E_t(z)$ per cui vale l'eguaglianza $E_t(z) \cdots E_1(z)\, U(z) = I_m$. Quindi

$$U(z) = E_1^{-1}(z) \cdots E_t^{-1}(z)$$

è prodotto di matrici elementari.

v) $\Rightarrow$ i) Segue immediatamente dal fatto che le matrici elementari hanno inversa polinomiale. ■

Come diretta conseguenza dei punti i), iii) e iv) della precedente proposizione, due matrici polinomiali $M(z)$ e $N(z)$ sono equivalenti se e solo se $N(z) = U(z)M(z)V(z)$, dove $U(z)$ e $V(z)$ sono opportune matrici unimodulari; in particolare, i teoremi sulle forme di Smith e di Hermite possono essere riformulati nel seguente modo:

Data una matrice $M(z)$ in $\mathbb{F}[z]^{p\times m}$, esiste una coppia di matrici unimodulari $U(z) \in \mathbb{F}[z]^{p\times p}$ e $V(z) \in \mathbb{F}[z]^{m\times m}$ tali che

$$\Gamma(z) = U(z)M(z)V(z) = \left[\begin{array}{cccc} \gamma_1(z) & & & \\ & \gamma_2(z) & & \\ & & \ddots & 0 \\ & & & \gamma_r(z) \\ \hline & 0 & & 0 \end{array}\right], \tag{3.11}$$

dove r è il rango di $M(z)$ e $\gamma_1(z), \gamma_2(z), \ldots, \gamma_r(z) \in \mathbb{F}[z]$ sono polinomi monici univocamente determinati che soddisfano $\gamma_i(z)|\gamma_{i+1}(z)$, per $i = 1, 2, ..., r-1$.

Sia $M(z)$ una matrice in $\mathbb{F}[z]^{p\times m}$ di rango m. Allora esiste una matrice unimodulare $U(z) \in \mathbb{F}[z]^{p\times p}$ tale che

$$H(z) := U(z)M(z) = \left[\begin{array}{cccc} h_{11}(z) & * & * & * \\ & h_{22}(z) & * & * \\ & & \ddots & * \\ & & & h_{mm}(z) \\ \hline & & 0 & \end{array}\right], \tag{3.12}$$

con $h_{jj}(z)$ polinomio monico soddisfacente $\deg h_{jj} > \deg h_{ij}$ per $j = 1, 2, \ldots, m$ e $i < j$.

- Esercizio 3.2.1 Sia R un anello commutativo e $R^{m\times m}$ l'insieme delle matrici $m \times m$ a coefficienti in R. Se definiamo unimodulare una matrice U in $R^{m\times m}$ quando è dotata di inversa in $R^{m\times m}$, allora U è unimodulare se e solo se il suo determinante è invertibile in R.

- Esercizio 3.2.2 Sia $M(z) \in \mathbb{F}[z]^{p\times m}$ una matrice di rango m e sia $U(z) \in \mathbb{F}[z]^{m\times m}$ una matrice unimodulare. Dimostrare che i minori di ordine massimo corrispondenti in $M(z)$ e $M(z)U(z)$ differiscono per una costante moltiplicativa non nulla.

- ESERCIZIO 3.2.3 i) Se $A \in \mathbb{F}^{m\times m}$, la matrice $I - Az$ è unimodulare se e solo se A è nilpotente.
  ii) Se $U(z) = U_0 + U_1 z + \cdots + U_r z^r$, con $U_i \in \mathbb{F}^{m\times m}$ e $r > 1$, è unimodulare, allora U_0 è non singolare e U_r è singolare. È vero anche il viceversa?

Proposizione 3.2.3 [MINORI DELLE MATRICI UNIMODULARI] *Siano*

$$U(z) \text{ e } V(z)$$

una matrice unimodulare di dimensioni $n \times n$ e la sua inversa. Se $r < n$ è un intero positivo, siano $\mathbf{i} = (i_1, i_2, \ldots, i_r)$ e $\mathbf{j} = (j_1, j_2, \ldots, j_r)$ due r-uple ordinate di interi estratte dall'insieme $\{1, 2, \ldots, n\}$ e siano $\mathbf{i}_{(C)}$ e $\mathbf{j}_{(C)}$ le corrispondenti $(n-r)$-uple complementari.

1. *I determinanti delle sottomatrici*

$$U(z)_{\mathbf{ij}} \text{ e } (V^T(z))_{\mathbf{i}_{(C)}\mathbf{j}_{(C)}}$$

   ottenute selezionando rispettivamente in $U(z)$ righe di indici $i_1, i_2, \ldots, i_r$ e colonne di indici $j_1, j_2, \ldots, j_r$ e in $V^T(z)$ le righe e le colonne con gli indici complementari, sono due polinomi associati.

2. *I minori di ordine r della sottomatrice $U(z)_{\mathbf{j}} \in \mathbb{F}[z]^{n\times r}$, formata dalle colonne di $U(z)$ con indici $j_1, j_2, \ldots, j_r$, hanno MCD unitario.*

PROVA (1.) Partizioniamo $U(z)$ e $V(z)$ nella forma

$$U(z) = \begin{bmatrix} U_{11}(z) & U_{12}(z) \\ U_{21}(z) & U_{22}(z) \end{bmatrix}, \quad V(z) = \begin{bmatrix} V_{11}(z) & V_{12}(z) \\ V_{21}(z) & V_{22}(z) \end{bmatrix}$$

con $U_{11}(z)$ e $V_{11}(z)$ matrici $r \times r$.
Dall'identità

$$\begin{bmatrix} U_{11}(z) & U_{12}(z) \\ 0 & I_{n-r} \end{bmatrix} \begin{bmatrix} V_{11}(z) & V_{12}(z) \\ V_{21}(z) & V_{22}(z) \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ V_{21}(z) & V_{22}(z) \end{bmatrix}$$

segue

$$\det U_{11}(z) = \det V(z) \det V_{22}(z) = k \det V_{22}(z) = k \det V_{22}^T(z),\ 0 \neq k \in \mathbb{F} \tag{3.13}$$

Quindi l'enunciato è vero se $\mathbf{i} = \mathbf{j} = (1, 2, \ldots, r)$.
Per gli altri casi, ricordando che per ogni matrice di permutazione Π vale l'identità $\Pi^{-1} = \Pi^T$, si ha immediatamente

$$I_n = U(z)V(z) = [\Pi(\mathbf{i})U(z)\Pi(\mathbf{j})]\ [\Pi(\mathbf{i})V^T(z)\Pi(\mathbf{j})]^T$$

dove $\Pi(\mathbf{i})$ è la matrice di permutazione che porta in posizione $1, 2, \ldots, r$ le righe di indici $i_1, i_2, \ldots, i_r$ mantenendo l'ordine esistente fra le altre righe, mentre $\Pi(\mathbf{j})$ è la matrice di permutazione che porta in posizione $1, 2, \ldots, r$ le colonne di indici $j_1, j_2, \ldots, j_r$ mantenendo l'ordine esistente fra le altre colonne.
In tal modo,

- nella matrice $\bar{U}(z) := \Pi(\mathbf{i})U(z)\Pi(\mathbf{j})$ la sottomatrice $r \times r$ nell'angolo in alto a sinistra coincide con $U(z)_{\mathbf{ij}}$,
- nella matrice $\bar{V}^T(z) := \Pi(\mathbf{i})V^T(z)\Pi(\mathbf{j})$ la sottomatrice $(n-r) \times (n-r)$ nell'angolo in basso a destra coincide con $(V^T(z))_{\mathbf{i}_{(C)}\mathbf{j}_{(C)}}$.

Applicando alle matrici $\bar{U}(z)$ e $\bar{V}(z)$ il ragionamento svolto nella prima parte, si conclude che $U(z)_{\mathbf{ij}}$ e $(V^T(z))_{\mathbf{i}_{(C)}\mathbf{j}_{(C)}}$ hanno determinanti che sono polinomi associati.
(2.) Ricorrendo alla formula di Laplace per il calcolo del determinante, si ottiene

$$\det U(z) = \sum_{\mathbf{i}} \det U(z)_{\mathbf{ij}} \det U(z)_{\mathbf{i}_{(C)}\mathbf{j}_{(C)}} \tag{3.14}$$

dove $\mathbf{i}$ descrive tutte le r-uple di indici estratte da $\{1, 2, \ldots, n\}$. Poiché $\det U(z)$ è una costante non nulla, (3.14) implica che, per un fissato $\mathbf{j}$, i minori $\det U(z)_{\mathbf{ij}}$ hanno MCD unitario. ■

3.3 Fattorizzazione delle matrici polinomiali

Il concetto di divisore (o fattore) di un polinomio e di divisore comune di una coppia di polinomi si estende al caso matriciale. Ovvie ragioni dimensionali nel caso di matrici rettangolari e la non commutatività del prodotto nel caso di matrici quadrate impongono di distinguere tra fattori sinistri e destri.

3.3.1 Divisori e primalità

Definizione 3.3.1 [Divisori destri di una matrice] *Sia $M(z) \in \mathbb{F}[z]^{p\times m}$: la matrice quadrata $\Delta(z) \in \mathbb{F}[z]^{m\times m}$ è un divisore destro di $M(z)$ se esiste una matrice $\bar{M}(z) \in \mathbb{F}[z]^{p\times m}$ tale che*

$$M(z) = \bar{M}(z)\Delta(z). \tag{3.15}$$

In particolare, se per ogni fattorizzazione di $M(z)$

$$M(z) = \tilde{M}(z)\tilde{\Delta}(z)$$

esiste $Q(z) \in \mathbb{F}[z]^{m\times m}$ per cui vale

$$\Delta(z) = Q(z)\tilde{\Delta}(z) \tag{3.16}$$

$\Delta(z)$ è detto un divisore destro massimo di $M(z)$. Se in ogni fattorizzazione del tipo (3.15) la matrice $\Delta(z)$ è unimodulare, $M(z)$ è detta prima a destra.

Similmente $\nabla(z) \in \mathbb{F}[z]^{p\times p}$ è un *divisore sinistro* se esiste una matrice $\bar{M}(z) \in \mathbb{F}[z]^{p\times m}$ tale che

$$M(z) = \nabla(z)\bar{M}(z). \tag{3.17}$$

In particolare, se per ogni fattorizzazione di $M(z)$

$$M(z) = \tilde{\nabla}(z)\tilde{M}(z)$$

esiste $P(z) \in \mathbb{F}[z]^{m\times m}$ per cui vale

$$\nabla(z) = \tilde{\nabla}(z)P(z) \tag{3.18}$$

$\nabla(z)$ è un *divisore sinistro massimo* di $M(z)$. Se in ogni fattorizzazione del tipo (3.17) la matrice $\nabla(z)$ è unimodulare, $M(z)$ è detta *prima a sinistra*.

È immediato verificare che le matrici unimodulari sono fattori (sinistri o destri, a seconda dei casi) di ogni matrice, dal momento che

$$\begin{aligned} M(z) &= \Big(M(z)U^{-1}(z)\Big)U(z), \\ M(z) &= U(z)\Big(U^{-1}(z)M(z)\Big) \end{aligned}$$

per ogni $U(z)$ unimodulare.

Per motivi di brevità la maggior parte dei risultati che presenteremo nel seguito fa riferimento al caso "destro", risultando immediato l'adattamento al caso sinistro.

Proposizione 3.3.2 [Rango delle matrici prime] *Una matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ prima a destra ha rango m.*

Dimostrazione Se $M(z)$ non avesse rango m, esisterebbe un vettore razionale non nullo $\mathbf{q}(z) \in \mathbb{F}(z)^m$ tale che $M(z)\mathbf{q}(z) = 0$. Detto $d(z)$ il m.c.m. dei denominatori degli elementi di $\mathbf{q}(z)$, il vettore polinomiale non nullo $\mathbf{v}(z) := d(z)\mathbf{q}(z)$ soddisfa a sua volta la relazione $M(z)\mathbf{v}(z) = 0$. Possiamo supporre, senza perdita di generalità, che la m-esima componente di $\mathbf{v}(z)$, $v_m(z)$, sia non nulla. Distinguiamo due casi: *i)* $v_m(z)$ è un polinomio di grado positivo, oppure *ii)* $v_m(z)$ è una costante non nulla, ed in tal caso si può scalare $\mathbf{v}(z)$ in modo che $v_m(z)$ valga -1. Allora si ha

$$M(z) = M(z)\begin{bmatrix} & & v_1(z) \\ & & v_2(z) \\ & I_{m-1} & \vdots \\ & & v_{m-1}(z) \\ 0 & \dots\ 0 & 1+v_m(z) \end{bmatrix},$$

dove il fattore a destra è non unimodulare in entrambi i casi. ■

La nozione di primalità (a destra) può essere interpretata come una generalizzazione alle matrici rettangolari della nozione di unimodularità. Di fatto le caratterizzazioni della primalità, fornite dalla seguente proposizione, nel caso di matrici quadrate ripropongono o integrano le condizioni di unimodularità riportate nella Proposizione 3.2.2.

Proposizione 3.3.3 [Condizioni equivalenti alla primalità] *Sia $M(z)$ una matrice in $\mathbb{F}[z]^{p\times m}$. Sono fatti equivalenti:*

i) *$M(z)$ è prima a destra;*

ii) *la forma canonica di Smith di $M(z)$ è $\begin{bmatrix} I_m \\ 0 \end{bmatrix}$;*

iii) *la forma di Hermite di* $M(z)$ *è* $\begin{bmatrix} I_m \\ 0 \end{bmatrix}$;

iv) *la matrice* $M(z)$ *è completabile ad una matrice unimodulare* $p \times p$*, ovvero esiste una matrice* $K(z) \in \mathbb{F}[z]^{p\times(p-m)}$ *tale che* $[\, M(z) \quad K(z) \,]$ *è unimodulare;*

v) $M(z)$ *ammette un'inversa sinistra polinomiale, ovvero esiste* $X(z) \in \mathbb{F}[z]^{m\times p}$ *tale che*

$$X(z)M(z) = I_m; \tag{3.19}$$

vi) $M(z)$ *ha rango* $m \leq p$ *e il MCD dei minori di ordine* m *di* $M(z)$ *è l'unità;*

vii) *se* $\mathbf{u}(z)$ *è un vettore a componenti in* $\mathbb{F}((z))$ *e* $\mathbf{y}(z) = M(z)\mathbf{u}(z)$ *è un vettore polinomiale, allora* $\mathbf{u}(z)$ *appartiene a* $\mathbb{F}[z]^m$.

DIMOSTRAZIONE i) $\Rightarrow$ ii) Se $\Gamma(z)$ è la forma canonica di Smith di $M(z)$, esistono matrici unimodulari di opportune dimensioni, $U(z)$ e $V(z)$, tali che $M(z) = U(z)\Gamma(z)V(z)$. Detta $U'(z)$ la sottomatrice di $U(z)$ formata dalle sue prime m colonne, l'espressione

$$M(z) = U'(z)\Big[\mathrm{diag}\{\gamma_1(z), \gamma_2(z), \ldots, \gamma_m(z)\}_{m\times m} V(z)\Big] =: U'(z)\Delta(z),$$

fornisce una fattorizzazione della matrice $M(z)$.
Se non fosse l'identità, $\mathrm{diag}\{\gamma_1(z), \gamma_2(z), \ldots, \gamma_m(z)\}$ conterrebbe polinomi invarianti $\gamma_i(z)$ di grado positivo e $\Delta(z)$ sarebbe un divisore destro non unimodulare della matrice $M(z)$.

ii) $\Rightarrow$ iii) Siano $U(z)$ e $V(z)$ matrici unimodulari di opportune dimensioni che riducono $M(z)$ in forma canonica di Smith, ovvero

$$U(z)M(z)V(z) = \begin{bmatrix} I_m \\ 0 \end{bmatrix}.$$

Premoltiplicando entrambi i membri della precedente equazione per la matrice diagonale a blocchi $\mathrm{diag}\{V(z), I_{p-m}\}$ e postmoltiplicandoli per $V^{-1}(z)$ si ottiene

$$\begin{bmatrix} V(z) & 0 \\ 0 & I_{p-m} \end{bmatrix} U(z)M(z) = \begin{bmatrix} I_m \\ 0 \end{bmatrix},$$

che è la forma di Hermite di $M(z)$.

iii) $\Rightarrow$ iv) Se la forma di Hermite di $M(z)$ è $\begin{bmatrix} I_m \\ 0 \end{bmatrix}$, si ha

$$U(z)M(z) = \begin{bmatrix} I_m \\ 0 \end{bmatrix},$$

per qualche matrice unimodulare $U(z)$. Di conseguenza

$$M(z) = U^{-1}(z) \begin{bmatrix} I_m \\ 0 \end{bmatrix},$$

e quindi $M(z)$ può essere vista come la sottomatrice della matrice unimodulare $U^{-1}(z)$ formata dalle sue prime m colonne.

iv) $\Rightarrow$ v) Se $[M(z)\ K(z)]$ è una matrice unimodulare e $V(z)$ ne è l'inversa, le prime m righe di $V(z)$ forniscono un'inversa polinomiale sinistra di $M(z)$.

v) $\Rightarrow$ vi) Dall'ipotesi che esista un'inversa sinistra segue $m \leq p$. Applicando il teorema di Binet-Cauchy alla (3.19) si ottiene

$$1 = \det(X(z)M(z)) = \sum_i m_i(X)m_i(M), \tag{3.20}$$

dove $m_i(X)$ e $m_i(M)$ sono minori di ordine massimo m corrispondenti in $X(z)$ e $M(z)$, rispettivamente. Se i minori di ordine massimo di $M(z)$ avessero un fattore comune, esso dovrebbe essere fattore di ogni loro combinazione polinomiale e quindi comparire al primo membro della (3.20), un assurdo.

vi) $\Rightarrow$ vii) Sia S_i la matrice $m \times p$ ottenuta permutando le colonne di $[I_m\ 0]$ in modo che $S_iM(z)$ sia la sottomatrice relativa all'i-esimo minore $m_i(M)$ di $M(z)$, ovvero $m_i(M) = \det(S_iM(z))$. Dall'identità

$$\mathrm{adj}(S_iM(z))S_iM(z) = m_i(M)I_m,$$

tenendo conto che i minori $m_i(M)$ sono coprimi e quindi esistono polinomi $h_i(z)$ per cui si ha $\sum_i m_i(M)h_i(z) = 1$, otteniamo

$$\sum_i h_i(z)\mathrm{adj}\Big(S_iM(z)\Big)S_iM(z) = \sum_i h_i(z)m_i(M)I_m = I_m. \tag{3.21}$$

Postmoltiplicando per $\mathbf{u}(z)$ entrambi i membri di (3.21), si ricava

$$\sum_i h_i(z)\mathrm{adj}\Big(S_iM(z)\Big)S_i\mathbf{y}(z) = \mathbf{u}(z). \tag{3.22}$$

Poiché $\mathbf{y}(z)$ è polinomiale, tale è il primo membro della (3.22), quindi il secondo membro $\mathbf{u}(z)$ deve essere un vettore polinomiale.

vii) $\Rightarrow$ i) Supponiamo, per assurdo, che la matrice $M(z)$ non sia prima a destra.
Se $M(z)$ non ha rango m, esistono vettori razionali e non polinomiali $\mathbf{u}(z)$ per i quali $M(z)\mathbf{u}(z) = 0$, contro la vii).
Se $M(z)$ ha rango m, allora fattorizza nella forma

$$M(z) = \bar{M}(z)\Delta(z)$$

per qualche matrice non unimodulare e non singolare $\Delta(z) \in \mathbb{F}[z]^{m\times m}$. Dal momento che almeno una delle colonne di $\Delta^{-1}(z)$ è razionale (quindi di Laurent), ma non polinomiale, $\bar{M}(z) = M(z)\Delta^{-1}(z)$ implica l'esistenza di almeno un vettore costituto da serie di Laurent non polinomiali che viene trasformato dalla matrice $M(z)$ in un vettore polinomiale, contraddicendo ancora la vii). ■

- ESERCIZIO 3.3.1 Sia $M(z)$ una matrice in $\mathbb{F}[z]^{p\times m}$.

  (*i*) Se $M(z)$ è prima a destra, allora $M(0)$ ha rango m.

  (*ii*) Se $M(z)$ ha rango r, esiste una fattorizzazione

  $$M(z) = \bar{M}(z)T(z), \tag{3.23}$$

  con $\bar{M}(z)$ matrice prima a destra di dimensioni $p \times r$ e $T(z)$ matrice polinomiale di rango r e dimensioni $r \times m$ (Suggerimento: si riporti $M(z)$ in forma di Hermite).

  (*iii*) Nella fattorizzazione (3.23) $T(0)$ ha rango r?

Se $\mathbb{F}$ è un campo algebricamente chiuso, o comunque se si considerano i polinomi in $\mathbb{F}[z]$ come polinomi dell'anello $\bar{\mathbb{F}}[z]$, dove $\bar{\mathbb{F}}$ è la chiusura algebrica di $\mathbb{F}$, alle condizioni di primalità ora riportate se ne aggiunge un'altra: infatti la condizione vi), di coprimalità dei minori di ordine massimo di $M(z)$, è equivalente all'assenza in $\bar{\mathbb{F}}$ di zeri comuni ai minori stessi. Pertanto:

viii) una matrice polinomiale $M(z) \in \mathbb{F}[z]^{p\times m}$ è prima a destra se e solo $M(\alpha)$ ha rango pieno m per ogni valore della variabile α in $\bar{\mathbb{F}}$.

Vale infine la pena di osservare che, in base al punto ii) della precedente proposizione, due matrici polinomiali $p \times m$ sono prime a destra se e solo se sono entrambe equivalenti alla matrice $\begin{bmatrix} I_m \\ 0 \end{bmatrix}$. Di conseguenza, data una matrice $M(z)$ prima a destra, le matrici prime a destra di ugual dimensione sono tutte e sole quelle ottenibili premoltiplicando e/o postmoltiplicando $M(z)$ per matrici unimodulari e formano un'unica classe di equivalenza in $\mathbb{F}[z]^{p\times m}/\sim$.

3.3.2 Divisori massimi

Proposizione 3.3.4 [ESISTENZA DI UN DIVISORE DESTRO MASSIMO] *Se $M(z)$ è un'arbitraria matrice non nulla in $\mathbb{F}[z]^{p\times m}$, allora esiste una fattorizzazione di $M(z)$ nella forma*

$$M(z) = \bar{M}(z)\bar{\Delta}(z) \tag{3.24}$$

con $\bar{\Delta}(z) \in \mathbb{F}[z]^{m\times m}$ divisore destro massimo di $M(z)$.

DIMOSTRAZIONE Supponiamo dapprima $p \geq m$. Siano $U(z) \in \mathbb{F}[z]^{p\times p}$ una matrice unimodulare che riduce $M(z)$ in forma di Hermite, ovvero

$$U(z)M(z) = \begin{bmatrix} \bar{\Delta}(z) \\ 0 \end{bmatrix} \begin{matrix} m \\ p-m \end{matrix}, \tag{3.25}$$

e

$$V(z) = \begin{matrix} \\ m \\ p-m \end{matrix} \begin{matrix} \begin{matrix} m & p-m \end{matrix} \\ \begin{bmatrix} V_{11}(z) & V_{12}(z) \\ V_{21}(z) & V_{22}(z) \end{bmatrix} \end{matrix}$$

la matrice unimodulare inversa di $U(z)$. Si ha allora

$$M(z) = \begin{bmatrix} V_{11}(z) & V_{12}(z) \\ V_{21}(z) & V_{22}(z) \end{bmatrix} \begin{bmatrix} \bar{\Delta}(z) \\ 0 \end{bmatrix} = \begin{bmatrix} V_{11}(z) \\ V_{21}(z) \end{bmatrix} \bar{\Delta}(z) =: \bar{M}(z)\bar{\Delta}(z).$$

Poiché $\bar{M}(z)$ è completabile ad una matrice unimodulare, in base alla precedente proposizione essa è prima a destra. Inoltre $\bar{\Delta}(z)$ è un divisore destro massimo. Consideriamo infatti una generica fattorizzazione di $M(z)$ del tipo

$$M(z) = \tilde{M}(z)\tilde{\Delta}(z). \tag{3.26}$$

Da (3.25) e (3.26) segue

$$U(z)\tilde{M}(z)\tilde{\Delta}(z) = U(z)M(z) = \begin{bmatrix} \bar{\Delta}(z) \\ 0 \end{bmatrix}, \tag{3.27}$$

e partizionando per righe $U(z)$ nella forma

$$U(z) = \begin{bmatrix} U_1(z) \\ U_2(z) \end{bmatrix} \begin{matrix} \}m \\ \}p-m \end{matrix}$$

si ottiene

$$[U_1(z)\tilde{M}(z)]\tilde{\Delta}(z) = \bar{\Delta}(z), \tag{3.28}$$

che prova che $\bar{\Delta}(z)$ è multiplo di $\tilde{\Delta}(z)$ e quindi è un divisore destro massimo di $M(z)$.
Se $p < m$, da

$$M(z) = [\, I_p \quad 0\,] \begin{bmatrix} M(z) \\ 0 \end{bmatrix}$$

segue che

$$\begin{bmatrix} M(z) \\ 0 \end{bmatrix} \tag{3.29}$$

è un divisore destro di M. Per provarne la massimalità, se $M(z) = \tilde{M}(z)\Delta(z)$ è una fattorizzazione generica di $M(z)$, aggiungendo righe nulle si ha

$$\begin{bmatrix} M(z) \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{M}(z) \\ 0 \end{bmatrix} \Delta(z)$$

e $\Delta(z)$ è un divisore destro di (3.29) ■

Proposizione 3.3.5 [PROPRIETÁ DEL DIVISORE DESTRO MASSIMO] *Se $M(z)$ ha dimensione $p \times m$ e rango m e se*

$$M(z) = \bar{M}_1(z)\bar{\Delta}_1(z) = \bar{M}_2(z)\bar{\Delta}_2(z) \tag{3.30}$$

sono due fattorizzazioni con $\bar{\Delta}_1(z)$ e $\bar{\Delta}_2(z)$ divisori destri massimi, allora

(i) $\bar{\Delta}_1(z)$ e $\bar{\Delta}_2(z)$ differiscono per un fattore sinistro unimodulare (ovvero il divisore destro massimo di $M(z)$ è unico a meno di un fattore sinistro unimodulare);

(ii) $\bar{M}_1(z)$ e $\bar{M}_2(z)$ sono prime a destra e differiscono per un fattore destro unimodulare;

Se $M(z) \neq 0$ ha rango r minore di m, vale ancora (i), ma in (3.30) $\bar{M}_1(z)$ e $\bar{M}_1(z)$ non sono necessariamente prime a destra né differiscono per un fattore destro unimodulare.

Esiste peraltro una matrice $C(z) \in \mathbb{F}[z]^{p\times r}$, prima a destra e unica a meno di fattori destri $r \times r$ unimodulari, tale che

$$M(z) = [C(z) \;\; 0_{p\times(m-r)}]U_1(z)\bar{\Delta}_1(z) = [C(z) \;\; 0_{p\times(m-r)}]U_2(z)\bar{\Delta}_2(z) \tag{3.31}$$

con $U_i(z)$ unimodulari.

DIMOSTRAZIONE Supponiamo che $M(z)$ abbia rango m e quindi abbia rango m ogni suo fattore quadrato destro,
(i) Se $\bar{\Delta}_1(z)$ e $\bar{\Delta}_2(z)$ sono due divisori destri massimi di $M(z)$, si ha

$$\bar{\Delta}_1(z) = P(z)\bar{\Delta}_2(z), \qquad \bar{\Delta}_2(z) = Q(z)\bar{\Delta}_1(z), \tag{3.32}$$

quindi

$$\bar{\Delta}_1(z) = P(z)Q(z)\bar{\Delta}_1(z) \tag{3.33}$$

e da $\det \bar{\Delta}_1(z) \neq 0$ segue

$$\det[P(z)Q(z)] = 1 \tag{3.34}$$

che prova l'unimodularità di $P(z)$ e di $Q(z)$.
(ii) Nelle fattorizzazioni (3.30) sostituiamo $\bar{\Delta}_1(z)$ con $P(z)\bar{\Delta}_2(z)$, $P(z)$ unimodulare:

$$\bar{M}_1(z)P(z)\bar{\Delta}_2(z) = \bar{M}_2(z)\bar{\Delta}_2(z) \tag{3.35}$$

da cui $[\bar{M}_1(z)P(z) - \bar{M}_2(z)]\bar{\Delta}_2(z) = 0$ e infine

$$\bar{M}_1(z)P(z) = \bar{M}_2(z).$$

Se la matrice $\bar{M}_1(z)$ non fosse prima a destra, essa fattorizzerebbe nella forma $\bar{M}_1(z) = \tilde{M}(z)\tilde{D}(z)$ con $\deg\det\tilde{D} > 0$. Allora $\tilde{D}(z)\bar{\Delta}_1(z)$ sarebbe un divisore destro di $M(z)$, ma il divisore destro massimo $\bar{\Delta}_1(z)$ di $M(z)$ non potrebbe essere multiplo di $\tilde{D}(z)\bar{\Delta}_1(z)$.

Supponiamo ora che $M(z)$ abbia rango $r < m$, e ricorrendo alla forma di Hermite

$$U(z)M(z) = \begin{bmatrix} H(z) \\ 0 \end{bmatrix} \begin{matrix} r \\ p-r \end{matrix}$$

con $H(z)$ di rango r, fattorizziamo $M(z)$ nella forma

$$M(z) = U^{-1}(z)\begin{bmatrix} H(z) \\ 0 \end{bmatrix} = C(z)H(z) \tag{3.36}$$

con $C(z)$ costituita dalle prime r colonne di $U^{-1}(z)$. "Orlando" $C(z)$ a destra con $m-r$ colonne nulle e $H(z)$ in basso con $m-r$ righe nulle, si ottiene

$$M(z) = [\, C(z) \quad | \quad 0\,]\begin{bmatrix} H(z) \\ 0 \end{bmatrix} \tag{3.37}$$

in cui $\begin{bmatrix} H(z) \\ 0 \end{bmatrix}$ è un MCD destro di $M(z)$. Se $M(z) = \bar{M}(z)\bar{\Delta}(z)$ è un'altra fattorizzazione, con $\bar{\Delta}(z)$ MCD destro, devono valere contemporaneamente

$$\bar{\Delta}(z) \;\; = \;\; P(z)\begin{bmatrix} H(z) \\ 0 \end{bmatrix} = [\, P_1(z) \quad | \quad \star\,]\begin{bmatrix} H(z) \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} H(z) \\ 0 \end{bmatrix} = Q(z)\bar{\Delta}(z) = \begin{bmatrix} Q_1(z) \\ 0 \end{bmatrix} \bar{\Delta}(z) \tag{3.38}$$

dove $P_1(z)$ e $Q_1(z)$ sono costituite rispettivamente dalle prime r colonne di $P(z)$ e dalle prime r righe di $Q(z)$. Da (3.38) si ricava facilmente $H(z) = Q_1(z)P_1(z)H(z)$ e, tenendo conto del rango di $H(z)$,

$$Q_1(z)P_1(z) = I_r,$$

ovvero $P_1(z)$ ha un'inversa sinistra polinomiale. Quindi $P_1(z)$ è completabile a una matrice unimodulare per aggiunzione di colonne e nella prima delle (3.38) le colonne di $P(z)$ indicate con $\star$ possono essere sostituite da colonne che rendono $P(z)$ unimodulare.
Poiché il generico fattore destro massimo $\bar{\Delta}(z)$ differisce per un fattore sinistro unimodulare da quello ottenuto mediante la forma di Hermite, anche nel caso in cui $M(z)$ non abbia rango di colonna pieno tutti i suoi fattori destri massimi differiscono fra loro per fattori unimodulari sinistri. Ciò dimostra che anche nel caso generale i divisori destri massimi fifferiscono per fattori sinistri unimodulari

Per dimostrare la (3.31), notiamo che il generico fattore destro massimo $\bar{\Delta}(z)$ di $M(z)$ è esprimibile nella forma

$$U(z)\bar{\Delta}(z) = \begin{bmatrix} H(z) \\ 0 \end{bmatrix}$$

con $U(z)$ unimodulare. Si ottiene allora

$$M(z) = [\, C(z) \quad | \quad 0\,] \begin{bmatrix} H(z) \\ 0 \end{bmatrix} = \Big([\, C(z) \quad | \quad 0\,] U(z) \Big) \bar{\Delta}(z)$$

che implica appunto (3.31).
La verifica dell' unicità di $C(z)$ a meno di fattori unimodulari destri $r \times r$ è lasciata per esercizio.
È altresí immediato che, quando $M(z)$ ha rango minore di m e si estrae un fattore destro massimo, la matrice $\bar{M}(z)$ non è unica a meno di fattori unimodulari. Basta notare che la fattorizzazione (3.37) rimane valida sostituendo le colonne nulle di $[\, C(z) \quad | \quad 0\,]$ con colonne arbitrarie. ■

- Esercizio 3.3.2* Dimostrare che la matrice

$$M(z) = \begin{bmatrix} z+1 & 1 \\ (z+1)z & z \end{bmatrix},$$

ammette due fattorizzazioni del tipo $M(z) = \bar{M}(z)\Delta_1(z) = \bar{M}(z)\Delta_2(z)$ con $\Delta_1(z)$ e $\Delta_2(z)$ fattori destri invertibili e distinti.

- Esercizio 3.3.3* Se $M(z)$ una matrice polinomiale $p \times m$, non nulla, ad elementi in $\mathbb{F}[z]$, sono fatti equivalenti:

  (*i*) $M(z)$ ha rango m;

  (*ii*) in ogni fattorizzazione $M(z) = \bar{M}(z)\Delta(z)$, con $\Delta(z) \in \mathbb{F}[z]^{m\times m}$, il divisore destro $\Delta(z)$ è non singolare;

  (*iii*) esiste una fattorizzazione $M(z) = \bar{M}(z)\Delta(z)$ con $\Delta(z) \in \mathbb{F}[z]^{m\times m}$ divisore destro massimo non singolare;

  (*iv*) in ogni fattorizzazione $M(z) = \bar{M}(z)\Delta(z)$, con $\Delta(z) \in \mathbb{F}[z]^{m\times m}$ divisore destro massimo, la matrice $\bar{M}(z)$ è prima a destra.

- ESERCIZIO 3.3.4* È vero che se esiste una fattorizzazione $M(z) = \bar{M}(z)\bar{\Delta}(z)$, con $\bar{\Delta}(z)$ divisore destro massimo di $M(z)$ e $\bar{M}(z)$ prima a destra allora $M(z)$ ha rango m ? (Suggerimento: si consideri l'ultima parte della dimostrazione della Prop.3.3.5)

Proposizione 3.3.6 [FATTORIZZAZIONE] *Sia $M(z)$ una matrice polinomiale $p \times m$ di rango r ad elementi in $\mathbb{F}[z]$. Allora esiste una fattorizzazione di $M(z)$ del tipo*

$$M(z) = L(z)C(z)R(z), \tag{3.39}$$

dove $L(z) \in \mathbb{F}^{p\times r}[z]$ è prima a destra, $R(z) \in \mathbb{F}^{r\times m}[z]$ è prima a sinistra e $C(z) \in \mathbb{F}^{r\times r}[z]$ è non singolare.

DIMOSTRAZIONE Riduciamo la matrice $M(z)$ in forma di Hermite per colonne, ottenendo

$$U(z)M(z) = \begin{bmatrix} H(z) \\ 0 \end{bmatrix},$$

con $H(z) \in \mathbb{F}^{r\times m}[z]$ di rango r e $U(z)$ unimodulare. Indicando con $L(z)$ la matrice prima a destra costituita dalle prime r colonne di $U^{-1}(z)$, si ha allora

$$M(z) = L(z)H(z).$$

Estraiamo poi da $H(z)$ un divisore sinistro massimo $C(z) \in \mathbb{F}[z]^{r\times r}$, necessariamente di rango r, ottenendo $H(z) = C(z)R(z)$. La matrice $R(z)$ è prima a sinistra, per (la versione duale del)la Proposizione 3.3.5. ■

- ESERCIZIO 3.3.5 Si considerino due fattorizzazioni $L_1(z)C_1(z)R_1(z)$ e $L_2(z)C_2(z)R_2(z)$ della matrice $M(z)$, soddisfacenti entrambe le condizioni della Proposizione 3.3.7. Si verifichi che esistono matrici unimodulari $r \times r$ $U(z)$ e $V(z)$ tali che $C_2(z) = U(z)C_1(z)V(z)$, $L_1(z) = L_2(z)U(z)$ e $V(z)R_2(z) = R_1(z)$.

Le definizioni ed i risultati sui divisori di una matrice polinomiale trovano un'estensione diretta al caso di coppie di matrici polinomiali.

Definizione 3.3.7 [DIVISORI DESTRI COMUNI] *Siano $M_1(z)$ e $M_2(z)$ matrici polinomiali con ugual numero di colonne m. Diciamo che $\Delta(z) \in \mathbb{F}[z]^{m\times m}$ è un* **divisore destro comune di** *$M_1(z)$ e $M_2(z)$ se*

$$M_1(z) = \bar{M}_1(z)\Delta(z) \qquad M_2(z) = \bar{M}_2(z)\Delta(z), \tag{3.40}$$

per opportune matrici polinomiali $\bar{M}_1(z)$ e $\bar{M}_2(z)$. Inoltre, $M_1(z)$ e $M_2(z)$ sono **coprime a destra** *se i loro unici divisori destri comuni sono matrici unimodulari.*

La definizione di divisore comune destro massimo è analoga a quella di divisore destro massimo per una singola matrice. Si osservi che $M_1(z)$ e $M_2(z)$ hanno $\Delta(z)$ come divisore destro comune se e solo se

$$\begin{bmatrix} M_1(z) \\ M_2(z) \end{bmatrix} = \begin{bmatrix} \bar{M}_1(z) \\ \bar{M}_2(z) \end{bmatrix} \Delta(z). \tag{3.41}$$

Pertanto cercare i divisori destri comuni di una coppia di matrici è equivalente a cercare i divisori destri della matrice ottenuta "impilando" $M_1(z)$ e $M_2(z)$, e i risultati sui fattori

destri e sulla primalità a destra di una matrice si applicano immediatamente ai fattori comuni destri e alla coprimalità a destra di una coppia di matrici. In particolare, come conseguenza della Proposizione 3.3.3, è immediato infatti provare il seguente

Corollario 3.3.8 [Condizioni di coprimalità a destra] *Se $M_1(z)$ e $M_2(z)$ sono matrici polinomiali di dimensioni $p_1 \times m$ e $p_2 \times m$, si equivalgono i seguenti fatti:*

i) *$M_1(z)$ e $M_2(z)$ sono coprime a destra;*

ii) *è prima a destra la matrice*

$$\begin{bmatrix} M_1(z) \\ M_2(z) \end{bmatrix};$$

iii) [Equazione di Bézout] *esistono due matrici polinomiali $X_1(z) \in \mathbb{F}[z]^{m \times p_1}$ e $X_2(z) \in \mathbb{F}[z]^{m \times p_2}$ che soddisfano l'equazione*

$$X_1(z)M_1(z) + X_2(z)M_2(z) = I_m. \qquad \blacksquare \tag{3.42}$$

Nel caso in cui $M_1(z)$ e $M_2(z)$ abbiano lo stesso numero di righe p si possono dare definizioni, analoghe a quelle di 3.3.7, di divisore comune sinistro e di matrici coprime a sinistra, e ridurre la coprimalità di tali matrici alla primalità a sinistra della matrice ottenuta giustapponendo le matrici $M_1(z)$ e $M_2(z)$.

- Esercizio 3.3.6 Sia $M(z) \in \mathbb{F}[z]^{p \times m}$, di rango r, e sia $\mathcal{M}$ l' $\mathbb{F}[z]$-modulo generato dalle colonne di $M(z)$. Si dimostri che

  (i) $\mathcal{M}$ è libero;

  (ii) se $V(z)$ è unimodulare, le colonne di $M(z)V(z)$ generano $\mathcal{M}$;

  (iii) se le colonne di $P(z)$ sono una base per $\mathcal{M}$, $P(z)$ ha rango di colonna pieno

  (iv) se $r = m$, le colonne di $M(z)$ sono una base per $\mathcal{M}$;

  (v) se $\gamma_1(z), \ldots \gamma_r(z)$ sono i polinomi invarianti di $M(z)$, una base per $\mathcal{M}$ è rappresentabile come insieme delle colonne di $\bar{P}(z)\text{diag}\{\gamma_1, \ldots, \gamma_r\}_{r \times r}$, con $\bar{P}(z) \in \mathbb{F}[z]^{p \times r}$ prima a destra;

  (vi) se le colonne di $P_1(z)$ e quelle di $P_2(z)$ costituiscono due basi per $\mathcal{M}$, allora le due matrici hanno lo stesso numero di colonne e $P_1(z) = P_2(z)V(z)$ dove $V(z)$ è una matrice unimodulare;

  (vii) $M(z)$ è prima a destra se e solo se ogni $\mathbb{F}[z]$-modulo che contenga propriamente $\mathcal{M}$ ha una base contenente più di m elementi.

3.4 Applicazioni

È interessante a questo punto soffermarci brevemente sul significato e su alcune conseguenze delle caratterizzazioni della primalità considerate nel paragrafo precedente.

3.4.1 Sistemi a risposta impulsiva finita

Consideriamo un sistema lineare e tempo invariante, a tempo discreto, causale con m ingressi e p uscite, di dimensione finita, e quindi descrivibile attraverso un modello di stato $\Sigma = (F, G, H, J)$ di tipo consueto. Diciamo che Σ ha *risposta impulsiva finita* (sistema FIR) se le uscite forzate corrispondenti agli ingressi impulsivi

$$\mathbf{u}^{(i)}(t) = \begin{cases} \mathbf{e}_i & \text{per } t = 0 \\ 0 & \text{per } t > 0, \end{cases} \qquad i = 1, 2, \ldots, m,$$

$\mathbf{e}_i$ l'i-esimo vettore della base canonica, hanno durata finita. È chiaro che condizione necessaria e sufficiente perchè ciò avvenga è che tutti gli elementi della matrice di trasferimento di Σ, $W(z) \in \mathbb{F}(z)^{p\times m}$, siano funzioni razionali proprie con valutazione negativa solo nell'origine, ovvero deve esistere un intero non negativo μ tale che

$$W(z) = W_0 + W_1 z^{-1} + \ldots + W_\mu z^{-\mu}, \qquad W_i \in \mathbb{F}^{p\times m}.$$

Spesso è conveniente studiare la dinamica di questi sistemi ponendo $d := z^{-1}$, così da avere una matrice di trasferimento

$$M(d) = \sum_{i=0}^{\mu} W_i d^i \in \mathbb{F}[d]^{p\times m}.$$

Conformemente, ogni segnale con supporto sul semiasse $t \geq 0$ sarà rappresentato da un vettore a componenti in $\mathbb{F}[[d]]$ e, se a durata finita, a componenti in $\mathbb{F}[d]$.

L'eventuale primalità a destra della matrice di trasferimento polinomiale $M(d)$ ha importanti conseguenze sulla struttura delle traiettorie di uscita del sistema Σ e sulla relazione che esse hanno con le corrispondenti traiettorie di ingresso.

• Osserviamo preliminarmente che per la Proposizione 3.3.2 la matrice prima a destra $M(d)$ ha rango m. Avendo colonne linearmente indipendenti, essa è un operatore iniettivo, ovvero *applica ingressi distinti in uscite forzate distinte*. Ciò vale sia per ingressi polinomiali che per arbitrari ingressi a supporto compatto nel passato, rappresentabili nella forma $\hat{\mathbf{u}}(d) = \sum_t \mathbf{u}(t)d^t \in \mathbb{F}((d))^m$.

• Dalla condizione ii) della Proposizione 3.3.3 segue che $M(0)$ ha rango m. Pertanto ogni segnale di ingresso $\sum_{t\geq 0} \mathbf{u}(t)d^t$ il cui primo campione non nullo sia $\mathbf{u}(0)$ produce un segnale di uscita $\sum_{t\geq 0} \mathbf{y}(t)d^t$ con $\mathbf{y}(0) \neq 0$, dal momento che

$$\mathbf{y}(0) = M(0)\mathbf{u}(0).$$

Quindi *l'uscita forzata del sistema non presenta ritardi rispetto all'ingresso forzante.*

• Esaminiamo ora la struttura della forma di Hermite data nella condizione iii):

$$\begin{bmatrix} I_m \\ 0 \end{bmatrix} = U(d)M(d) = \begin{bmatrix} U_1(d) \\ U_2(d) \end{bmatrix} M(d),$$

dove $U_1(d)$ è la sottomatrice costituita dalle prime m righe di $U(d)$. È chiaro che ogni uscita forzata prodotta da un ingresso $\hat{\mathbf{u}}(d)$ in $\mathbb{F}((d))^m$, ovvero ogni vettore $\hat{\mathbf{y}}(d) \in \mathbb{F}((d))^p$ appartenente all'immagine di $M(d)$ soddisfa alla condizione $U_2(d)\hat{\mathbf{y}}(d) = 0$.

Viceversa, supponiamo che $\hat{\mathbf{y}}(d)$ sia un vettore in $\mathbb{F}((d))^p$ che soddisfa $U_2(d)\hat{\mathbf{y}}(d) = 0$. Ponendo $\hat{\mathbf{u}}(d) := U_1(d)\hat{\mathbf{y}}(d)$ si ottiene

$$U(d)\hat{\mathbf{y}}(d) = \begin{bmatrix} \hat{\mathbf{u}}(d) \\ 0 \end{bmatrix} = \begin{bmatrix} I_m \\ 0 \end{bmatrix} \hat{\mathbf{u}}(d) = U(d)M(d)\hat{\mathbf{u}}(d).$$

Ciò implica $U(d)[\hat{\mathbf{y}}(d) - M(d)\hat{\mathbf{u}}(d)] = 0$ ed essendo $U(d)$ unimodulare si conclude che $\hat{\mathbf{y}}(d)$ sta nell'immagine di $M(d)$.

Possiamo quindi *caratterizzare le uscite forzate del sistema,* rappresentabili nel caso più generale da serie formali a supporto compatto a sinistra, come tutti e soli i vettori $\hat{\mathbf{y}}(d) \in \mathbb{F}((d))^p$ che soddisfano la condizione $U_2(d)\hat{\mathbf{y}}(d) = 0$, ovvero *come il nucleo della matrice* $U_2(d)$

• La condizione v) corrisponde all'*esistenza di un sistema inverso dotato anch'esso della proprietà FIR.* In evoluzione forzata tale sistema, alimentato dal segnale $\hat{\mathbf{y}}(d) = M(d)\hat{\mathbf{u}}(d)$, riproduce in uscita l'ingresso $\hat{\mathbf{u}}(d)$ al sistema di matrice di trasferimento $M(d)$

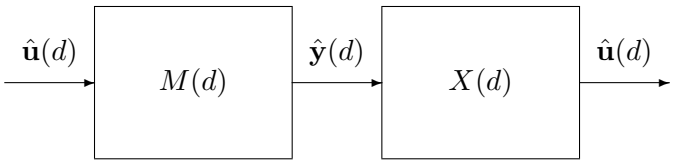

Fig.3.1

Mentre l'esistenza di inverse sinistre razionali (non necessariamente causali) per la matrice polinomiale $M(d)$ equivale al fatto che $M(d)$ abbia rango m, la sua primalità a destra equivale all'esistenza di inverse sinistre polinomiali. Tali inverse possono essere realizzate mediante sistemi a memoria finita, in cui l'evoluzione libera di stato si annulla in un numero finito di passi.

• Infine, la condizione vii) garantisce che *non esistano segnali di ingresso* $\hat{\mathbf{u}}(d)$ *a durata infinita che producono uscite forzate* $\hat{\mathbf{y}}(d) = M(d)\hat{\mathbf{u}}(d)$ *polinomiali e quindi di durata finita.* Per comprendere le conseguenze di questa proprietà, che discuteremo più diffusamente nel capitolo sulla codifica convoluzionale, basterà accennare qui al caso in cui $M(d)$ sia matrice di trasferimento di un dispositivo che elabora un segnale di errore $\hat{\mathbf{e}}(d)$ trasformandolo in un segnale di allarme $M(d)\hat{\mathbf{e}}(d)$. Se $M(d)$ non è prima, si può verificare la situazione - indesiderata - di un segnale di allarme che si estingue in un numero finito di passi nonostante il segnale di errore $\hat{\mathbf{e}}(d)$ si protragga indefinitamente nel tempo.

3.4.2 Raggiungibilità e controllabilità di sistemi a tempo discreto

Come è noto dal corso di Teoria dei Sistemi, un modello di stato a matrici reali $\Sigma = (F, G, H, J)$ è raggiungibile se e solo se la matrice del criterio PBH di raggiungibilità

$$[zI - F \mid G]$$

ha rango pieno per ogni $z \in \mathbb{C}$ ovvero, essendo $\mathbb{C}$ un campo algebricamente chiuso, se e solo se, vista come matrice polinomiale nell'indeterminata z, essa è prima a sinistra.
L'equivalenza fra la primalità a sinistra della matrice e la raggiungibilità del sistema a tempo discreto Σ può essere provata direttamente, basandosi sui risultati del precedente paragrafo, per sistemi definiti su un arbitrario campo numerico $\mathbb{F}$.
Poiché il sistema è tempo invariante, non è restrittivo supporre che l'istante finale in cui si raggiunge lo stato desiderato sia $t = 1$ e lasciare libero l'istante in corrispondenza al quale, a partire dallo stato nullo, il sistema inizia ad essere sollecitato.

Proposizione 3.4.1 [Criterio PBH di raggiungibilità] *Si consideri l'equazione di aggiornamento di stato di un sistema dinamico lineare*

$$\mathbf{x}(t+1) = F\mathbf{x}(t) + G\mathbf{u}(t), \tag{3.43}$$

con $F \in \mathbb{F}^{n\times n}$ e $G \in \mathbb{F}^{n\times m}$. Per ogni $\mathbf{x}_f \in \mathbb{F}^n$ esiste un ingresso di durata finita che porta il sistema dallo stato iniziale nullo nello stato finale $\mathbf{x}_f$ se e solo se la matrice polinomiale

$$[zI - F \mid G] \tag{3.44}$$

è prima a sinistra.

Dimostrazione Supponiamo dapprima che la matrice (3.44) sia prima a sinistra. Esistono allora, in base all'analogo "sinistro" del Corollario 3.3.9, matrici polinomiali $Y_1(z)$ e $Y_2(z)$ soddisfacenti l'equazione di Bézout

$$(zI_n - F)Y_1(z) + G\ Y_2(z) = I_n, \tag{3.45}$$

e per ogni stato $\mathbf{x}_f$ si ha

$$(zI_n - F)Y_1(z)\mathbf{x}_f + G\ Y_2(z)\mathbf{x}_f = \mathbf{x}_f.$$

I vettori polinomiali $\tilde{X}(z) = -Y_1(z)\mathbf{x}_f$ e $\tilde{U}(z) = Y_2(z)\mathbf{x}_f$ soddisfano l'identità

$$z\tilde{X}(z) + \mathbf{x}_f = F\ \tilde{X}(z) + G\ \tilde{U}(z), \tag{3.46}$$

e possono essere espressi come polinomi a coefficienti vettoriali nella forma

$$\tilde{X}(z) := \sum_{t=0}^{\nu} \tilde{\mathbf{x}}(-t)z^t \qquad \tilde{U}(z) := \sum_{t=0}^{\nu} \tilde{\mathbf{u}}(-t)z^t,$$

per qualche intero positivo ν. Eguagliando i coefficienti relativi a potenze di ugual grado nella (3.46), si ottiene la relazione fra i coefficienti di $\tilde{X}(z)$ e $\tilde{U}(z)$

$$\tilde{\mathbf{x}}(-t+1) = F\tilde{\mathbf{x}}(-t) + G\tilde{\mathbf{u}}(-t),$$

per $t \geq 1$ e con condizione iniziale $\tilde{\mathbf{x}}(-\nu - 1) = 0$ Per $t = 0$ si ricava

$$\mathbf{x}_f = F\tilde{\mathbf{x}}(0) + G\tilde{\mathbf{u}}(0) = \mathbf{x}(1)$$

e di conseguenza la successione di ingresso $\tilde{\mathbf{u}}(-\nu), \ldots, \tilde{\mathbf{u}}(0)$, applicata a Σ in corrispondenza allo stato iniziale $\tilde{\mathbf{x}}(-\nu-1) = 0$, porta il sistema nello stato $\mathbf{x}(1) = \tilde{\mathbf{x}}_f$ in un numero finito di passi.

Viceversa, si supponga raggiungibile il sistema Σ e si consideri per ogni condizione finale $\mathbf{x}(1) = \mathbf{e}_i$, $\mathbf{e}_i$ vettore della base canonica di $\mathbb{F}^n$, una successione di ingresso a supporto finito $\mathbf{u}^{(i)}(-\nu_i), \ldots, \mathbf{u}^{(i)}(0)$ che piloti il sistema dallo stato nullo all'istante $-\nu_i$, passando attraverso una successione di stati intermedi $\mathbf{x}^{(i)}(-\nu_i+1), \mathbf{x}^{(i)}(-\nu_i+2), \ldots, \mathbf{x}^{(i)}(0)$, allo stato $\mathbf{e}_i$ all'istante 1. Ponendo

$$X^{(i)}(z) = \sum_{t\geq 0} \mathbf{x}^{(i)}(-t)z^t \in \mathbb{F}[z]^n \qquad U^{(i)}(z) = \sum_{t\geq 0} \mathbf{u}^{(i)}(-t)z^t \in \mathbb{F}[z]^m,$$

moltiplicando per z^t entrambi i membri di

$$\mathbf{x}^{(i)}(-t+1) = F\mathbf{x}^{(i)}(-t) + G\mathbf{u}^{(i)}(-t),$$

e sommando per tutti i valori di t maggiori o uguali a zero, si ottiene

$$zX^{(i)}(z) + \mathbf{e}_i = FX^{(i)}(z) + GU^{(i)}(z),$$

Ma allora da

$$\mathbf{e}_i = (zI_n - F)(-X^{(i)}(z)) + GU^{(i)}(z),$$

variando i tra 1 e n, si ottiene

$$I_n = (zI_n - F)\Big[-X^{(1)}(z) \mid \ldots \mid -X^{(n)}(z)\Big] + G\,\Big[U^{(1)}(z) \mid \ldots \mid U^{(n)}(z)\Big],$$

che dimostra che $[zI_n - F \mid G]$ è prima a sinistra. ■

La primalità a sinistra della matrice (3.44) risulta condizione necessaria e sufficiente per la raggiungibilità della coppia (F, G) e corrisponde, perciò, all'esistenza di soluzioni all'equazione di Bézout (3.45).
Abbiamo sfruttato tale caratterizzazione per ottenere esplicitamente, in corrispondenza ad una soluzione $(Y_1(z), Y_2(z))$ dell'equazione (3.45), una sequenza di ingresso che porta il sistema nello stato desiderato e la corrispondente evoluzione forzata dello stato.
È possibile dimostrare che, al variare della soluzione $(Y_1(z), Y_2(z))$ dell'equazione (3.45), le coppie $(-Y_1(z)\mathbf{x}_f, Y_2(z)\mathbf{x}_f)$ corrispondono a tutte le traiettorie ingresso-stato a supporto finito a partire da condizione iniziale nulla e con stato finale $\mathbf{x}(1) = \mathbf{x}_f$.

La controllabilità (a zero) di un sistema $\Sigma = (F, G, H, J)$ può essere caratterizzata in modo analogo. Nel caso di sistemi a matrici reali, infatti, il criterio PBH afferma che Σ è controllabile (a zero) se e solo se $[zI - F \mid G]$ ha rango pieno per ogni $z \in \mathbb{C} \setminus \{0\}$. Non è difficile verificare, attraverso un semplice cambio di variabili, che ciò equivale ad assumere che la matrice

$$[I - dF \mid dG] \tag{3.47}$$

abbia rango pieno per ogni $d \in \mathbb{C}$, ovvero che (3.47), vista come matrice polinomiale nell'indeterminata d, sia prima a sinistra. L'equivalenza fra la primalità a sinistra della matrice polinomiale (3.47) e la controllabilità del sistema Σ può essere anche provata

direttamente. In questo caso conviene imporre la condizione iniziale nell'istante $t = 0$ e interpretare i vettori polinomiali in d come segnali che evolvono per tempi positivi. I dettagli della dimostrazione vengono lasciati per esercizio.

- ESERCIZIO 3.4.1 Si dimostri che il sistema lineare (3.43) con equazione di uscita

$$\mathbf{y}(t) = H\mathbf{x}(t)$$

è osservabile se e solo se la matrice $\begin{bmatrix} zI - F \\ H \end{bmatrix}$ è prima a destra.

3.5 Grado di vettori e matrici polinomiali

Definizione 3.5.1 [GRADI DI UN VETTORE E DI UNA MATRICE POLINOMIALE] *(i) Il grado di un vettore polinomiale non nullo* $\mathbf{v}(z) \in \mathbb{F}[z]^p$ *è il massimo dei gradi delle sue componenti non nulle.*
(ii) Data una matrice polinomiale $M(z) \in \mathbb{F}[z]^{p\times m}$*, di rango* m*, indichiamo con* $k_1, k_2, \ldots, k_m$ *i gradi delle colonne di* $M(z)$*. Il grado esterno (di colonna) di* $M(z)$

$$\text{extdeg}M = \sum_{i=1}^{m} k_i$$

è la somma dei gradi delle sue colonne mentre il grado interno di $M(z)$, intdeg$M$*, è il grado massimo dei suoi minori di ordine* m*.*

- ESERCIZIO 3.5.1 [CARATTERIZZAZIONE "ESTREMALE" DELLE MATRICI PRIME A DESTRA] Una matrice polinomiale $M(z) \in \mathbb{F}[z]^{p\times m}$ di rango m è prima a destra se e solo se, fra tutte le matrici polinomiali della forma $M(z)R(z)$ con $R(z) \in \mathbb{F}(z)^{m\times m}$ matrice razionale non singolare, essa ha grado interno minimo. (Suggerimento: se $M(z)$ è prima a destra, $R(z)$ deve essere polinomiale e i gradi dei minori di $M(z)R(z)$ non possono essere inferiori a quelli dei corrispondenti minori in $M(z)$. Se non è prima a destra e $\Delta(z)$ è un fattore destro non unimodulare, si consideri $M(z)\Delta^{-1}(z)$)

Come è noto, il determinante di una matrice quadrata $P = [p_{ij}]$ di ordine m, può essere espresso nella forma

$$\det P = \sum_{\sigma} \text{sgn}(\sigma) p_{\sigma(1),1} p_{\sigma(2),2} \cdots p_{\sigma(m),m}, \tag{3.48}$$

dove la somma è estesa a tutte le permutazioni σ degli interi $1, 2, \ldots, m$, e sgn(σ) è la segnatura [3] della permutazione σ.
Dalla definizione dei gradi di colonna k_i segue che, quando P è una sottomatrice $m \times m$ di $M(z)$ ciascuno degli addendi della sommatoria in (3.81), e quindi il determinante, ha grado al più $\sum_{i=1}^{m} k_i$. Poiché tale numero costituisce un confine superiore al grado di tutti i minori di ordine m di $M(z)$, si ha sempre

$$\text{intdeg}M \leq \text{extdeg}M \tag{3.49}$$

[3] Con *segnatura di* σ si intende il numero $+1$ oppure -1 a seconda che sia pari o dispari il numero di scambi necessari per ottenere da $(1, 2, \ldots, m)$ la m-pla $(\sigma(1), \sigma(2), \ldots, \sigma(m))$.

3.5.1 Matrici ridotte per colonne

Definizione 3.5.2 [MATRICI RIDOTTE PER COLONNE] *Una matrice polinomiale $M(z) \in \mathbb{F}[z]^{p\times m}$ di rango m è ridotta per colonne se almeno un suo minore di ordine m ha grado $\sum_{i=1}^{m} k_i$, ovvero se*

$$\text{intdeg}M = \text{extdeg}M$$

In particolare, una matrice quadrata $M(z)$ di dimensione m è ridotta per colonne se è non singolare e

$$\deg \det M(z) = \sum_{i=1}^{m} k_i. \tag{3.50}$$

Esempio 3.5.1 (*i*) La matrice

$$M_1(z) = \begin{bmatrix} z+1 & z-1 \\ z+4 & z-1 \end{bmatrix}$$

ha gradi di colonna $k_1 = k_2 = 1$ e determinante $\det M_1(z) = -3z + 3$, e quindi non è ridotta per colonne. Invece la matrice

$$M_2(z) = \begin{bmatrix} z-1 & 3z \\ z-4 & z+1 \end{bmatrix}$$

è ridotta per colonne; i gradi di colonna, infatti, sono $k_1 = k_2 = 1$ e il determinante ha grado 2.

(*ii*) Ogni vettore polinomiale non nullo è una matrice ridotta per colonne.

La proprietà di una matrice di essere ridotta per colonne può essere ricondotta ad alcune condizioni equivalenti che ci proponiamo di esaminare. Per $i = 1, 2, \ldots, m$ indichiamo con $\mathbf{c}_i \in \mathbb{F}^p$ il vettore dei coefficienti dei monomi di grado k_i nella colonna i-esima della matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ e costruiamo per giustapposizione dei vettori $\mathbf{c}_i$ la *matrice conduttrice*

$$M_{hc} = [\, \mathbf{c}_1 \; \mathbf{c}_2 \; \ldots \; \mathbf{c}_m \,].$$

$M(z)$ può essere espressa allora nella forma

$$M(z) = M_{hc} \begin{bmatrix} z^{k_1} & & & \\ & z^{k_2} & & \\ & & \ddots & \\ & & & z^{k_m} \end{bmatrix} + M_{rem}(z), \tag{3.51}$$

dove M_{hc} appartiene a $\mathbb{F}^{p\times m}$ e i gradi di colonna della "matrice resto" $M_{rem}(z)$ sono strettamente minori dei corrispondenti gradi di colonna in $M(z)$. Nella proposizione che segue si mostra come la decomposizione (3.51) fornisca un semplice criterio per verificare se $M(z)$ è ridotta per colonne.

Proposizione 3.5.3 [RANGO PIENO DELLA MATRICE CONDUTTRICE] *La matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ è ridotta per colonne se e solo se la matrice conduttrice M_{hc} ha rango m.*

DIMOSTRAZIONE Sia $\mathbf{i} = (i_1, i_2, \ldots, i_m)$ una m-upla ordinata di interi, con $1 \le i_1 < i_2 < \ldots < i_m \le p$, e indichiamo con $M^{(\mathbf{i})}(z)$, $M_{hc}^{(\mathbf{i})}$ e $M_{rem}^{(\mathbf{i})}(z)$ le sottomatrici ottenute

selezionando rispettivamente in $M(z)$, M_{hc} e $M_{rem}(z)$ le righe di indici $i_1, i_2, \ldots, i_m$. Dalla (3.51) si ricava[4]

$$M^{(\mathbf{i})}(z) = M_{hc}^{(\mathbf{i})} \begin{bmatrix} z^{k_1} & & & \\ & z^{k_2} & & \\ & & \ddots & \\ & & & z^{k_m} \end{bmatrix} + M_{rem}^{(\mathbf{i})}(z). \tag{3.52}$$

Se identifichiamo nella (3.48) la matrice P con $M^{(\mathbf{i})}(z)$, vediamo che il monomio di grado $\sum_i k_i$ in $\det M^{(\mathbf{i})}(z)$ proviene dai prodotti dei monomi di grado k_1 in $[M^{(\mathbf{i})}(z)]_{\sigma(1),1}$, di grado k_2 in $[M^{(\mathbf{i})}(z)]_{\sigma(2),2}$, ... di grado k_m in $[M^{(\mathbf{i})}(z)]_{\sigma(m),m}$ ed è dato da

$$\sum_{\sigma} \operatorname{sgn}(\sigma) \left([M_{hc}^{(\mathbf{i})}]_{\sigma(1),1} z^{k_1}\right) \left([M_{hc}^{(\mathbf{i})}]_{\sigma(2),2} z^{k_2}\right) \cdots \left([M_{hc}^{(\mathbf{i})}]_{\sigma(m),m} z^{k_m}\right), \tag{3.53}$$

Quindi esso coincide con il determinante di $M_{hc}^{(\mathbf{i})} \operatorname{diag}\{z^{k_1}\ z^{k_2}\ \ldots\ z^{k_m}\}$ ed ha per coefficiente $\det M_{hc}^{(\mathbf{i})}$.
Se $M(z)$ è ridotta per colonne, esiste almeno una m-pla $\mathbf{i} = (i_1, i_2, \ldots, i_m)$ tale che $M^{(\mathbf{i})}(z)$ abbia determinante di grado $\sum_{i=1}^m k_i$. In corrispondenza, $M_{hc}^{(\mathbf{i})}$ è non singolare e la matrice conduttrice M_{hc} ha rango m.
Viceversa, se M_{hc} ha rango m, esiste una sua sottomatrice $M_{hc}^{(\mathbf{i})}$ di ordine m non singolare, la corrispondente sottomatrice $M^{(\mathbf{i})}(z)$ ha determinante di grado $\sum_{i=1}^m k_i$ e $M(z)$ è ridotta per colonne. ■

Esempio 3.5.2 Le matrici

$$M(z) = \begin{bmatrix} z^2 - z & z \\ z & z^2 \\ 2z^2 & 2z^2 - 3 \end{bmatrix}, \quad N(z) = \begin{bmatrix} z^2 - 1 & 3z^2 \\ z & 1 \\ z^2 - z & 3z^2 - 2 \end{bmatrix}$$

hanno, rispettivamente,

$$M_{hc} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{bmatrix}, \quad N_{hc} = \begin{bmatrix} 1 & 3 \\ 0 & 0 \\ 1 & 3 \end{bmatrix}$$

come matrici dei coefficienti conduttori di colonna. In base alla proposizione precedente, $M(z)$ è ridotta per colonne, mentre $N(z)$ non lo è.

Un'ulteriore caratterizzazione delle matrici ridotte per colonne è fornita dalla seguente proprietà, nota come *predicibilità del grado* (predictable degree property), che mette in relazione il grado di un vettore polinomiale $\mathbf{u}(z)$ con il grado del corrispondente vettore trasformato $\mathbf{y}(z) := M(z)\mathbf{u}(z)$.

Proposizione 3.5.4 [PREDICIBILITÀ DEL GRADO] *Una matrice $M(z) \in \mathbb{F}[z]^{p\times m}$, di rango m e gradi di colonna $k_1, k_2, \ldots, k_m$, è ridotta per colonne se e solo se per ogni vettore non nullo $\mathbf{u}(z) \in \mathbb{F}[z]^m$ si ha*

$$\deg(M\mathbf{u}) = \max_{i:u_i(z)\neq 0} \{k_i + \deg u_i\}. \tag{3.54}$$

[4] se M_{hc} è la matrice conduttrice di $M(z)$, non è necessariamente vero che $M_{hc}^{(\mathbf{i})}$ sia la matrice conduttrice di $M^{(\mathbf{i})}(z)$.

DIMOSTRAZIONE* Per ogni matrice polinomiale $M(z) = [m_{ji}(z)]$, posto $\mathbf{y}(z) = M(z)\mathbf{u}(z)$. l'elemento in posizione j-esima di $\mathbf{y}(z)$ soddisfa la diseguaglianza

$$\begin{aligned} \deg y_j(z) &= \deg(\sum_{i=1}^{m} m_{ji}u_i) \leq \max_{i:\ u_i(z)\neq 0}\{\deg m_{ji} + \deg u_i\} \\ &\leq \max_{i:\ u_i(z)\neq 0}\{k_i + \deg u_i\}, \end{aligned}$$

e quindi

$$\deg \mathbf{y}(z) \leq \max_{i:u_i(z)\neq 0}\{k_i + \deg u_i\} =: K. \tag{3.55}$$

Se $M(z)$ è ridotta per colonne, vale anche la diseguaglianza opposta, ovvero il grado di $\mathbf{y}(z)$ è esattamente K. Per provarlo, consideriamo un arbitrario ingresso $\mathbf{u}(z)$ per cui si abbia $\max_{i:u_i(z)\neq 0}\{k_i + \deg u_i\} = K$, rappresentiamo $\mathbf{y}(z) = M(z)\mathbf{u}(z)$ nella forma

$$\mathbf{y}(z) = \mathbf{y}_{hc}\ z^K + \mathbf{y}_{rem}(z), \qquad \mathbf{y}_{hc} \in \mathbb{F}^m,\ \ \deg \mathbf{y}_{rem} < K,$$

e mostriamo che $\mathbf{y}_{hc}$ è un vettore non nullo. Tenuto conto del fatto che $K - k_i$ è un confine superiore per il grado della i-esima componenente di $\mathbf{u}(z)$, per $i = 1, 2, \ldots, m$, e che per qualche valore di i tale confine viene raggiunto, possiamo porre

$$u_i(z) = \alpha_i\ z^{K-k_i} + r_i(z), \qquad \deg r_i < K - k_i,$$

e affermare che uno almeno dei coefficienti α_i deve essere non nullo. Dalla scrittura

$$\mathbf{y}(z) = \left\{ M_{hc} \begin{bmatrix} z^{k_1} & & & \\ & z^{k_2} & & \\ & & \ddots & \\ & & & z^{k_m} \end{bmatrix} + M_{rem}(z) \right\} \begin{bmatrix} \alpha_1 z^{K-k_i} + r_1(z) \\ \alpha_2 z^{K-k_2} + r_2(z) \\ \vdots \\ \alpha_m z^{K-k_m} + r_m(z) \end{bmatrix}$$

si deduce allora che il vettore $\mathbf{y}_{hc}$ è dato dal prodotto

$$M_{hc} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix},$$

e non può essere nullo dal momento che $M(z)$ è ridotta per colonne e quindi M_{hc} ha rango m.

Viceversa, ipotizziamo che $M(z)$ non sia ridotta per colonne, ovvero M_{hc} abbia rango minore di m. Esiste allora un vettore non nullo $[\beta_1\ \beta_2\ \ldots \beta_m]^T$ appartenente al nucleo di M_{hc}. Scelto un arbitrario numero $k \geq \max_i k_i$, ogni vettore $\mathbf{u}(z)$ del tipo

$$\mathbf{u}(z) = \begin{bmatrix} \beta_1 z^{k-k_1} \\ \beta_2 z^{k-k_2} \\ \vdots \\ \beta_m z^{k-k_m} \end{bmatrix}$$

non soddisfa la (3.54).

■

- ESERCIZIO 3.5.2* Supponiamo che $M(z) \in \mathbb{F}[z]^{p\times m}$ sia ridotta per colonne, con gradi di colonna $k_1, k_2, \ldots, k_m$.

  i) Se $\mathbf{u}(z) \in \mathbb{F}[z]^m$, allora il vettore $\mathbf{y}_{hc}$ in $\mathbf{y}(z) = M(z)\mathbf{u}(z)$ appartiene a $\mathrm{Im}M_{hc}$.

  ii) Se $T(z) \in \mathbb{F}[z]^{m\times m}$ ha rango pieno e $h_1, h_2, \ldots, h_m$ sono i gradi di colonna di $M(z)T(z)$, esiste una permutazione $(p_1, p_2, \ldots, p_m)$ di $(1, 2, \ldots, m)$ tale che $k_{p_1} \leq h_1,\ k_{p_2} \leq h_2, \ldots, k_{p_m} \leq h_m$ (Suggerimento: si consideri una permutazione tale che $T_{\nu,p_\nu}(z) \neq 0,\ \nu = 1, 2, \ldots, m$ e si applichi la relazione $h_i = \max_{j:T_{ij}\neq 0}\{k_j + \mathrm{deg}T_{ij}\} \geq k_{p_i} + \mathrm{deg}T_{i,p_i}$)

Se $M(d) \in \mathbb{F}[d]^{p\times m}$ è, come nel paragrafo precedente, la matrice di trasferimento di un sistema FIR, la proprietà di essere ridotta per colonne, e quindi la predicibilità del grado, hanno un'interpretazione piuttosto suggestiva. Se $\mathbf{u}(\cdot)$ è un segnale d'ingresso di durata limitata, rappresentato da un vettore polinomiale $\hat{\mathbf{u}}(d)$, l'ultimo campione non nullo ai terminali d'ingresso si presenta all'istante $\max_i\{\deg u_i\}$. Quando $M(d)$ è ridotta per colonne, il corrispondente segnale di uscita forzata dura esattamente fino all'istante $\max_i\{\deg u_i + k_i\}$, ovvero i gradi delle componenti del segnale di ingresso, e non i loro valori, consentono di "predire" il grado, e quindi la durata, del segnale di uscita $\hat{\mathbf{y}}(d) = M(d)\hat{\mathbf{u}}(d)$.

È sempre possibile applicare alle colonne di un'arbitraria matrice $M(z) \in \mathbb{F}[z]^{p\times m}$ di rango m una successione di operazioni elementari in modo da ridurla per colonne, ovvero esiste una matrice unimodulare $U(z) \in \mathbb{F}[z]^{m\times m}$ tale che

$$\bar{M}(z) = M(z)U(z)$$

abbia grado esterno ed interno coincidenti.
*Supponiamo infatti che tutti i minori di ordine massimo di $M(z)$ abbiano grado strettamente minore di $\sum_i k_i$, con k_i il grado della colonna i-esima di $M(z)$, $i = 1, 2, \ldots, m$. Allora M_{hc} ha rango minore di m ed esiste un vettore non nullo $\tilde{\mathbf{a}} := [\tilde{a}_1\ \tilde{a}_2\ \ldots\ \tilde{a}_m]^T \in \mathbb{F}^m$ tale che

$$M_{hc}\ \tilde{\mathbf{a}} = \mathbf{0}.$$

Tra le colonne di $M(z)$ corrispondenti a indici i per i quali $\tilde{a}_i$ è diverso da zero, ne scegliamo una, la j-esima, di grado massimo k_j.
Allora il vettore $\mathbf{a}(z) := [\,\tilde{a}_1 z^{k_j-k_1}\ \ \ldots\ \ \tilde{a}_j\ \ \ldots\ \ \tilde{a}_m z^{k_j-k_m}\,]^T$ è polinomiale e soddisfa

$$M_{hc}\begin{bmatrix} z^{k_1} & & & & \\ & \ddots & & & \\ & & z^{k_j} & & \\ & & & \ddots & \\ & & & & z^{k_m} \end{bmatrix}\begin{bmatrix} \tilde{a}_1 z^{k_j-k_1} \\ \vdots \\ \tilde{a}_j \\ \vdots \\ \tilde{a}_m z^{k_j-k_m} \end{bmatrix} = \mathbf{0}.$$

È immediato verificare che nella matrice

$$M'(z) = M(z)[\mathbf{e}_1 \mid \ldots \mid \mathbf{e}_{j-1} \mid \mathbf{a}(z) \mid \mathbf{e}_{j+1} \mid \ldots \mid \mathbf{e}_m] =: M(z)U_1(z)$$

la colonna j-esima $M(z)\mathbf{a}(z)$ ha grado minore di k_j mentre le altre colonne coincidono con le corrispondenti in $M(z)$. Quindi la somma $\sum_i k'_i$ dei gradi di colonna k'_i di $M'(z)$ è

minore di $\sum_i k_i$. D'altra parte, poichè la matrice $U_1(z)$ è unimodulare i minori di ordine m di $M'(z)$ coincidono a meno di una costante moltiplicativa non nulla con i minori di $M(z)$.
Di conseguenza in un numero finito di passi si perviene ad una matrice $\bar{M}(z) = M(z)U(z)$, con $U(z)$ matrice unimodulare, nella quale esiste un minore di ordine massimo il cui grado coincide con la somma dei gradi di colonna.

- ESERCIZIO 3.5.3 [CARATTERIZZAZIONE "ESTREMALE" DELLE MATRICI RIDOTTE PER COLONNE] Si dimostri che una matrice $M(z)$ a rango di colonna pieno è ridotta per colonne se e solo se, fra tutte le matrici della forma $M(z)U(z)$, con $U(z)$ unimodulare, essa ha grado esterno minimo.

Il procedimento di riduzione sopra descritto lascia invariante ad ogni passo il grado interno, abbassando progressivamente il grado esterno fino a renderlo eguale a quello interno. La successione di operazioni elementari di colonna non è unica, ed è possibile giungere, a seconda della successione utilizzata, a diverse matrici ridotte per colonna. In proposito, è naturale chiedersi allora se, a partire dalla medesima matrice, si possano ottenere due matrici ridotte per colonna, dotate del medesimo grado esterno, ma con insiemi di gradi di colonna diversi. La proposizione 3.5.6 mostra che una tale eventualità non si verifica mai, dato che due matrici $M(z)$ e $M'(z)$ ridotte per colonna e che differiscono per un fattore destro unimodulare hanno, a meno dell'ordine, gli stessi gradi di colonna.

Lemma 3.5.5 *Siano $M(z)$ e $M'(z)$ in $\mathbb{F}[z]^{p\times m}$, con gradi di colonna rispettivamente $k_1 \geq k_2 \geq \ldots \geq k_m$ e $k'_1 \geq k'_2 \geq \ldots \geq k'_m$. Se $M'(z)$ è ridotta per colonne e se esiste una matrice unimodulare $U(z) \in \mathbb{F}[z]^{m\times m}$ tale che*

$$M(z) = M'(z)U(z)$$

allora si ha $k_i \geq k'_i$, per $i = 1, 2, \ldots, m$.

<< $*$ DIMOSTRAZIONE Supponiamo che risulti $k_\nu < k'_\nu$. Poiché $M'(z)$ è ridotta per colonne, per la Proposizione 3.5.4 si ha per ogni i

$$k_i = \deg \mathrm{col}_i M = \deg(M' \mathrm{col}_i U) = \max_{h:u_{hi}(z)\neq 0}\{k'_h + \deg u_{hi}\}. \tag{3.56}$$

Se supponiamo che per qualche indice di colonna $i \geq \nu$ uno almeno tra i polinomi $u_{1i}(z), u_{2i}(z), \ldots, u_{\nu i}(z)$ sia non nullo, la determinazione di k_i in (3.56) come $\max_{h:u_{hi}(z)\neq 0}\{k'_h + \deg u_{hi}\}$ coinvolge almeno un valore $\bar{h} \in \{1, 2, \ldots, \nu\}$ dell'indice h, e si ha quindi

$$k_i \leq k_\nu < k'_\nu \leq k'_{\bar{h}} \leq \bar{k}'_{\bar{h}} + \deg u_{\bar{h}i} \leq \max_{h:u_{hi}(z)\neq 0}\{k'_h + \deg u_{hi}\}. \tag{3.57}$$

Poiché (3.56) e (3.57) sono contraddittorie, dobbiamo supporre $u_{hi}(z) = 0$ per ogni $h \leq \nu$ e per ogni $i \geq \nu$. In tal caso, però, la matrice $U(z)$ ha la struttura partizionata

$$U(z) = \begin{bmatrix} U_{11}(z) & 0 \\ U_{21}(z) & U_{22}(z) \end{bmatrix},$$

con $U_{11}(z)$ di dimensioni $\nu \times (\nu - 1)$, e quindi è singolare. È quindi assurdo supporre $k_\nu < k'_\nu$. $*$ >> ■

Di conseguenza, quando una matrice $M(z)$ viene ridotta per colonne, nella matrice risultante $M'(z)$, nessuno dei gradi di colonna (opportunamente riordinati) può superare quello corrispondente nella matrice $M(z)$.

Se nel lemma precedente si suppone che anche $M(z)$ sia ridotta per colonne, si perviene immediatamente alla

Proposizione 3.5.6 [Invarianza dei gradi di colonna] *Siano $M(z)$ e $M'(z)$ matrici ridotte per colonne e appartenenti a $\mathbb{F}[z]^{p\times m}$. Se esiste una matrice unimodulare $U(z) \in \mathbb{F}[z]^{m\times m}$ tale che*

$$M(z) = M'(z)U(z)$$

allora i gradi di colonna delle matrici coincidono a meno di una permutazione. ■

È possibile ridurre per colonne una matrice $M(z)$ $p \times m$ di rango m anche effettuando operazioni elementari di riga, ad esempio riducendo $M(z)$ alla forma di Hermite per colonne $H(z)$. In tale forma, infatti, ciascun termine sulla diagonale, $h_{ii}(z)$, ha grado strettamente superiore ai gradi di tutti gli altri elementi della stessa colonna, e quindi $k_i = \deg h_{ii}$, e l'unico minore di ordine m non nullo è $h_{11}(z)h_{22}(z)\dots h_{mm}(z)$.

Tuttavia matrici unimodulari differenti applicate alla sinistra della matrice $M(z)$ possono dar luogo a matrici ridotte per colonne con famiglie di gradi di colonna diverse, non riconducibili l'una all'altra mediante permutazioni. Si consideri, ad esempio, la matrice colonna

$$M(z) = \begin{bmatrix} z \\ z-1 \end{bmatrix}.$$

Essa è evidentemente ridotta per colonne ed ha grado $k_1 = 1$, tuttavia applicando alla sua sinistra la matrice unimodulare

$$\begin{bmatrix} 1-z & z \\ 1 & -1 \end{bmatrix},$$

si può portare $M(z)$ alla forma ridotta per colonne

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

il cui grado di colonna è zero.

La definizione di matrice ridotta per righe è del tutto analoga alla Definizione 3.5.3; se $h_1, h_2, \dots, h_p$ sono i gradi di riga della matrice $M(z) \in \mathbb{F}[z]^{p\times m}$, diciamo che $M(z)$ è *ridotta per righe* se ha rango p ed esiste almeno un minore di ordine p il cui grado coincide con $\sum_{i=1}^{p} h_i$.
I risultati visti in precedenza sulla riduzione per colonne trovano un esatto corrispondente in termini di righe e vengono riportati per completezza. La prova è immediata una volta che si passi alle matrici trasposte.

Proposizione 3.5.7 [Matrici ridotte per righe] *Sia $M(z)$ una matrice in $\mathbb{F}[z]^{p\times m}$ di rango p, con gradi di riga $h_1, h_2, \dots, h_p$. Sono equivalenti i seguenti fatti:*

i) $M(z)$ è ridotta per righe;

ii) nella scomposizione

$$M(z) = \begin{bmatrix} z^{h_1} & & & \\ & z^{h_2} & & \\ & & \ddots & \\ & & & z^{h_p} \end{bmatrix} M_{hr} + M_{rem}(z), \tag{3.58}$$

con $M_{rem}(z)$ matrice i cui gradi di riga sono strettamente inferiori ai corrispondenti gradi di riga in $M(z)$, M_{hr} ha rango p;

iii) per ogni vettore $\mathbf{u}(z) \in \mathbb{F}[z]^p$ si ha

$$\deg(\mathbf{u}(z)^T M(z)) = \max_{i:u_i(z)\neq 0}\{\deg u_i + h_i\}. \qquad \blacksquare \tag{3.59}$$

- ESERCIZIO 3.5.4 i) Una matrice $M(z) \in \mathbb{F}[z]^{m\times m}$ con

  $$M(z) = M_0 + M_1 z + \ldots + M_h z^h, \qquad M_h \neq 0,$$

  è detta *regolare* se M_h è una matrice invertibile. Dimostrare che ogni matrice regolare è ridotta sia per righe che per colonne. Qual è il grado di $\det M$?

  ii) Sia $M(z) \in \mathbb{F}[z]^{m\times m}$ una matrice ridotta per colonne e sia $K \in \mathbb{F}^{m\times m}$ un'arbitraria matrice costante invertibile. Dimostrare che $KM(z)$ è sempre ridotta per colonne, mentre $M(z)K$ può non esserlo.

- ESERCIZIO 3.5.5 Sia $M(z) \in \mathbb{F}[z]^{p\times m}$ una matrice di rango m, con gradi di colonna $\nu_1, \nu_2, \ldots, \nu_m$ e sia $M(z) = \bar{M}(z)\Delta(z)$ una fattorizzazione con $\bar{M}(z)$ ridotta per colonne. Allora il grado della colonna i-esima di $\Delta(z)$ non può eccedere ν_i (Suggerimento: $\nu_i = \max_{j:\Delta_{ji}(z)\neq 0}\{\deg \mathrm{col}_j(\bar{M}) + \deg\Delta_{ji}\}$);

  In particolare, se $M(z)$ è prima a destra e $U(z)$ è una matrice unimodulare che, operando sulle colonne di $M(z)$, la riduce per colonne, il grado della i-esima colonna di $U(z)^{-1}$ non eccede ν_i.

- ESERCIZIO 3.5.6 Se $\tilde{M}(z) \in \mathbb{F}[z]^{p\times m}$ ha rango m e $U(z)$ è una matrice unimodulare tale che $M(z) := \tilde{M}(z)U(z)$ sia ridotta per colonne con gradi di colonna $k_1, k_2, \ldots k_m$, per ogni vettore non nullo $\mathbf{u}(z) \in \mathbb{F}[z^m$ si ha $\deg(\tilde{M}(z)\mathbf{u}(z)) \geq \min k_i$ ed esistono vettori per i quali la relazione precedente vale con il segno di eguaglianza.

3.5.2 $<< *$*Divisionedivettoripermatrici*

Dedicheremo questo paragrafo a descrivere un algoritmo per ridurre il grado di un vettore polinomiale mediante aggiunzione di multipli polinomiali delle colonne di una assegnata matrice "divisore" $p\times m$, ridotta per colonne $D(z)$. L'algoritmo consente, fra l'altro, di decidere se un vettore polinomiale $\mathbf{y}(z) \in \mathbb{F}[z]^p$ è combinazione polinomiale delle colonne di $D(z)$, ovvero può essere espresso nella forma

$$\mathbf{y}(z) = D(z)\mathbf{u}(z) \tag{3.60}$$

per qualche $\mathbf{u}(z) \in \mathbb{F}[z]^m$. Il procedimento fornisce infatti un vettore polinomiale $\mathbf{y}_{\min}(z)$ di grado minimo per cui vale la relazione

$$\mathbf{y}(z) = D(z)\mathbf{q}(z) + \mathbf{y}_{\min}(z), \qquad \mathbf{q}(z) \in \mathbb{F}[z]^m, \tag{3.61}$$

e la condizione $\mathbf{y}_{\min}(z) = \mathbf{0}$ corrisponde al verificarsi di (3.60).

Per taluni aspetti l'algoritmo ricorda la divisione euclidea dei polinomi e sarà detto, seppure impropriamente, "divisione" del vettore $\mathbf{y}(z)$ per la matrice $D(z)$.

3.5.7 Algoritmo per la costruzione di $\mathbf{y}_{\min}(z)$

Riscriviamo la matrice $D(z)$ nella forma (3.51) e il vettore $\mathbf{y}(z) \neq \mathbf{0}$ nella forma

$$\mathbf{y}(z) = \mathbf{y}_{hc} z^\nu + \mathbf{y}_{rem}(z), \qquad \mathbf{y}_{hc} \in \mathbb{F}^p, \quad \deg \mathbf{y}_{rem}(z) < \nu. \tag{3.62}$$

Se $\mathbf{y}(z)$ ammette una scomposizione

$$\mathbf{y}(z) = D(z)\mathbf{q}(z) + \mathbf{y}'(z), \tag{3.63}$$

con $\deg \mathbf{y}' < \deg \mathbf{y}$, allora $D(z)\mathbf{q}(z)$ ha grado ν e, dal momento che $D(z)$ è ridotta per colonne, deve essere (cfr. la Proposizione 3.5.6)

$$\nu = \max_{i:q_i(z)\neq 0} \{k_i + \deg q_i\}. \tag{3.64}$$

Perciò in (3.63) il monomio $\mathbf{y}_{hc}z^{\nu}$ di (3.62) può provenire soltanto da combinazioni delle colonne di $D(z)$ con indice i per cui $k_i + \deg q_i = \nu$, e la colonna conduttrice soddisfa

$$\mathbf{y}_{hc} \in \mathrm{Im}_{i:k_i\leq\nu}\{\mathrm{col}_i D_{hc}\}. \tag{3.65}$$

Viceversa, se la colonna conduttrice $\mathbf{y}_{hc}$ di un vettore polinomiale $\mathbf{y}(z)$ di grado ν soddisfa la (3.65), si può scomporre $\mathbf{y}(z)$ nella forma

$$\begin{aligned}\mathbf{y}(z) &= \mathbf{y}_{hc}z^{\nu} + \mathbf{y}_{rem}(z)\\ &= D_{hc}\begin{bmatrix} z^{k_1} & & \\ & \ddots & \\ & & z^{k_m}\end{bmatrix}\begin{bmatrix}\beta_1 z^{\nu-k_1}\\ \vdots \\ \beta_m z^{\nu-k_m}\end{bmatrix} + \mathbf{y}_{rem}(z)\end{aligned}$$

con $\beta_i = 0$ se $k_i > \nu$, e quindi anche

$$\begin{aligned}\mathbf{y}(z) &= \left\{D_{hc}\begin{bmatrix} z^{k_1} & & \\ & \ddots & \\ & & z^{k_m}\end{bmatrix} + D_{rem}(z)\right\}\begin{bmatrix}\beta_1 z^{\nu-k_1}\\ \vdots \\ \beta_m z^{\nu-k_m}\end{bmatrix}\\ &+ \left\{\mathbf{y}_{rem}(z) - D_{rem}(z)\begin{bmatrix}\beta_1 z^{\nu-k_1}\\ \vdots \\ \beta_m z^{\nu-k_m}\end{bmatrix}\right\} = D(z)\mathbf{q}^{(1)}(z) + \mathbf{y}^{(1)}(z),\end{aligned}$$

in cui $\deg \mathbf{y}^{(1)}(z) < \nu = \deg \mathbf{y}(z)$.
È ora chiaro come a partire da $\mathbf{y}(z)$ si possa costruire una successione di vettori $\mathbf{y}(z), \mathbf{y}^{(1)}(z), \mathbf{y}^{(2)}(z), \ldots$ di grado decrescente e soddisfacenti le condizioni

$$\begin{aligned}&\mathbf{y}_{hc},\ \mathbf{y}^{(1)}_{hc},\ \mathbf{y}^{(2)}_{hc}\ \ \ldots \in \mathrm{Im}D_{hc}\\ &\mathbf{y}^{(i-1)}(z) = D(z)\mathbf{q}^{(i)}(z) + \mathbf{y}^{(i)}(z),\end{aligned}$$

fino ad ottenere un vettore $\mathbf{y}^{(\ell)}(z)$ nullo, oppure di grado $\nu^{(\ell)}$ con colonna conduttrice $\mathbf{y}^{(\ell)}_{hc}$ non appartenente a $\mathrm{Im}_{i:k_i\leq\nu^{(\ell)}}\{\mathrm{col}_i D_{hc}\}$. Si ha così

$$\mathbf{y}(z) = D(z)[\mathbf{q}^{(1)}(z) + \ldots + \mathbf{q}^{(\ell)}(z)] + \mathbf{y}^{(\ell)}(z) = D(z)\mathbf{q}(z) + \mathbf{y}^{(\ell)}(z). \tag{3.66}$$

Se $\mathbf{y}^{(\ell)}(z) \neq \mathbf{0}$ ha grado $\nu^{(\ell)}$, vogliamo provare che non esiste alcun vettore $\bar{\mathbf{y}}(z)$ con $\deg \bar{\mathbf{y}} < \nu^{(\ell)}$ che soddisfi un'equazione del tipo

$$\mathbf{y}(z) = D(z)\bar{\mathbf{q}}(z) + \bar{\mathbf{y}}(z). \tag{3.67}$$

Da (3.66) e (3.67) si ottiene

$$D(z)[\bar{\mathbf{q}}(z) - \mathbf{q}(z)] = \mathbf{y}^{(\ell)}(z) - \bar{\mathbf{y}}(z). \tag{3.68}$$

A secondo membro della (3.68), la colonna conduttrice $\left(\mathbf{y}^{(\ell)} - \bar{\mathbf{y}}\right)_{hc}$ è relativa ai termini di grado $\nu^{(\ell)}$ e coincide con $\mathbf{y}^{(\ell)}_{hc}$. Quindi

$$\left(\mathbf{y}^{(\ell)} - \bar{\mathbf{y}}\right)_{hc} \notin \mathrm{Im}_{i:k_i\leq\nu^{(\ell)}}\{\mathrm{col}_i D_{hc}\}.$$

A primo membro, per la predicibilità del grado, la medesima colonna è data da una combinazione lineare delle colonne della matrice conduttrice D_{hc} corrispondenti, nella matrice $D(z)$, a colonne di grado $k_i \leq \nu^{(\ell)}$: un assurdo.∗ >>

3.6 Equazioni diofantee

Al pari di quanto avviene per i polinomi e le funzioni razionali proprie e stabili, anche per le matrici polinomiali si possono considerare equazioni diofantee in cui coefficienti ed incognite sono matrici di dimensioni opportune. In questo paragrafo studieremo le equazioni del tipo

$$X(z)A(z) + Y(z)B(z) = C(z), \tag{3.69}$$

dove $A(z) \in \mathbb{F}[z]^{p\times m}$, $B(z) \in \mathbb{F}[z]^{q\times m}$ e $C(z) \in \mathbb{F}[z]^{\ell\times m}$ sono matrici note e le matrici soluzione $X(z)$ e $Y(z)$, di dimensioni rispettivamente $\ell \times p$ e $\ell \times q$ devono avere elementi polinomiali.

Proposizione 3.6.1 [Condizioni di risolubilità] *L'equazione (3.69) ammette una soluzione* $(X(z), Y(z)) \in \mathbb{F}[z]^{\ell\times p} \times \mathbb{F}[z]^{\ell\times q}$ *se e solo se ogni divisore destro comune* $\Delta(z)$ *di* $A(z)$ *e* $B(z)$ *è divisore destro di* $C(z)$*. In particolare, ha soluzione per ogni* $C(z) \in \mathbb{F}[z]^{\ell\times m}$ *se e solo se* $A(z)$ *e* $B(z)$ *sono coprime a destra.*

Dimostrazione Supponiamo che $(\bar{X}(z), \bar{Y}(z))$ sia una soluzione di (3.69) e sia $\Delta(z)$ un divisore destro comune di $A(z)$ e $B(z)$, ovvero

$$A(z) = \bar{A}(z)\Delta(z) \qquad B(z) = \bar{B}(z)\Delta(z). \tag{3.70}$$

Sostituendo le (3.70) in

$$\bar{X}(z)A(z) + \bar{Y}(z)B(z) = C(z),$$

si ottiene

$$[\bar{X}(z)\bar{A}(z) + \bar{Y}(z)\bar{B}(z)]\Delta(z) = C(z),$$

e quindi $\Delta(z)$ è divisore destro di $C(z)$.

Viceversa, supponiamo che ogni divisore destro di $A(z)$ e $B(z)$ divida a destra anche $C(z)$. Se $p + q \geq m$, per il teorema sulla forma di Hermite, esiste una matrice unimodulare $U(z) \in \mathbb{F}[z]^{(p+q)\times(p+q)}$ tale che

$$U(z)\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = \begin{bmatrix} U_{11}(z) & U_{12}(z) \\ U_{21}(z) & U_{22}(z) \end{bmatrix}\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = \begin{bmatrix} \bar{\Delta}(z) \\ 0 \end{bmatrix}, \tag{3.71}$$

e per la Proposizione 3.3.5 $\bar{\Delta}(z) \in \mathbb{F}[z]^{m\times m}$ è un divisore comune destro massimo di $A(z)$ e $B(z)$. Per ipotesi esiste allora $\bar{C}(z)$ tale che

$$C(z) = \bar{C}(z)\bar{\Delta}(z). \tag{3.72}$$

Dalle (3.71) e (3.72) si ricava

$$C(z) = \bar{C}(z)\bar{\Delta}(z) = [\bar{C}(z)U_{11}(z)]A(z) + [\bar{C}(z)U_{12}(z)]B(z),$$

e quindi $(\bar{C}(z)U_{11}(z), \bar{C}(z)U_{12}(z))$ è una soluzione di (3.69).
Se, invece, $p + q < m$, è sufficiente completare la matrice

$$\begin{bmatrix} A(z) \\ B(z) \end{bmatrix}$$

ad una matrice quadrata, aggiungendo un opportuno numero di righe nulle, ed applicare a tale matrice i ragionamenti fatti nel caso precedente.

Se $A(z)$ e $B(z)$ sono matrici coprime a destra, allora, come conseguenza del Corollario 3.3.9, l'equazione diofantea è risolubile per $C(z) = I_m$ e quindi per ogni altra matrice $C(z)$. Viceversa, se l'equazione è risolubile per ogni $C(z)$, lo è in particolare per $C(z) = I_m$ e quindi $A(z)$ e $B(z)$ devono essere coprime a destra. ■

Vogliamo ora analizzare la struttura della soluzione generale di (3.69), ovviamente nell'ipotesi che essa sia risolubile. È chiaro che se le coppie di matrici $(X(z), Y(z))$ e $(\bar{X}(z), \bar{Y}(z))$ risolvono entrambe la (3.69) si ha

$$[X(z) - \bar{X}(z)]A(z) + [Y(z) - \bar{Y}(z)]B(z) = 0,$$

ovvero

$$[X(z) - \bar{X}(z) \mid Y(z) - \bar{Y}(z)] \begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = 0.$$

L'esistenza di più di una soluzione è equivalente al fatto che il rango r della matrice

$$\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} \tag{3.73}$$

sia minore di $p + q$, e ciò si verifica se e solo se ha dimensione positiva il nucleo sinistro della medesima matrice, i.e. lo spazio dei vettori riga $\mathbf{v}(z) \in \mathbb{F}(z)^{p+q}$ tali che

$$\mathbf{v}(z) \begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = 0.$$

Il nucleo sinistro ha dimensione $p + q - r$ sopra il campo $\mathbb{F}(z)$ e ammette basi di vettori polinomiali (basta moltiplicare i vettori di una base razionale per il minimo comune multiplo dei denominatori). Se $V(z) \in \mathbb{F}[z]^{(p+q-r)\times(p+q)}$ è una matrice le cui righe formano una base polinomiale per il nucleo sinistro di (3.73), e fattorizziamo $V(z)$ nella forma

$$V(z) = \nabla(z)\bar{V}(z), \tag{3.74}$$

con $\bar{V}(z)$ prima a sinistra e $\nabla(z)$ non singolare, allora anche le righe di $\bar{V}(z)$ forniscono una base e risulta ovviamente

$$\bar{V}(z) \begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = 0.$$

Se $\mathbf{v}(z) \in \mathbb{F}[z]^{p+q}$ è un arbitrario vettore polinomiale nel nucleo sinistro di (3.73), esso è combinazione a coefficienti razionali delle righe di $\bar{V}(z)$. Ma allora, in virtù della primalità della matrice $\bar{V}(z)$ e della Proposizione 3.3.3, $\mathbf{v}(z)$ è esprimibile come combinazione polinomiale delle righe di $\bar{V}(z)$, ovvero

$$\mathbf{v}(z) = \mathbf{q}(z)\bar{V}(z),$$

con $\mathbf{q}(z)$ vettore riga in $\mathbb{F}[z]^{p+q-r}$.

Poichè tutte le righe della matrice $[X(z) - \bar{X}(z) \mid Y(z) - \bar{Y}(z)]$ sono nel nucleo di (3.73), esiste una matrice polinomiale $Q(z) \in \mathbb{F}[z]^{\ell\times(p+q-r)}$ tale che

$$[X(z) - \bar{X}(z) \mid Y(z) - \bar{Y}(z)] = Q(z)\bar{V}(z).$$

Quindi, date due arbitrarie soluzioni $(X(z), Y(z))$ e $(\bar{X}(z), \bar{Y}(z))$ di (3.69), esiste una matrice polinomiale $Q(z)$ per cui vale

$$[X(z) \mid Y(z)] = [\bar{X}(z) \mid \bar{Y}(z)] + Q(z)\bar{V}(z).$$

Viceversa, se $(\bar{X}(z), \bar{Y}(z))$ è una soluzione particolare di (3.69), per ogni scelta di $Q(z)$ la formula precedente fornisce una soluzione. Abbiamo così provato la seguente proposizione.

Proposizione 3.6.2 [Parametrizzazione delle soluzioni] *Supponiamo che l'equazione (3.69) sia risolubile e sia $(\bar{X}(z), \bar{Y}(z))$ una sua soluzione. Siano inoltre r il rango della matrice (3.73) e $\bar{V}(z) \in \mathbb{F}[z]^{(p+q-r)\times(p+q)}$ una matrice polinomiale prima a sinistra le cui righe costituiscono una base per il nucleo sinistro di (3.73). La soluzione generale di (3.69) è data da*

$$[X(z) \mid Y(z)] = [\bar{X}(z) \mid \bar{Y}(z)] + Q(z)\bar{V}(z), \tag{3.75}$$

al variare di $Q(z)$ sull'insieme delle matrici polinomiali $\ell \times (p+q-r)$. ■

La discussione che abbiamo svolto fin qui per le equazioni diofantee a coefficienti matriciali ha seguito le medesime linee della trattazione del primo capitolo per il caso scalare, in particolare la struttura della soluzione generale viene parametrizzata da una matrice arbitraria $Q(z)$, analoga al parametro polinomiale $q(z)$ che compare nella Proposizione 1.42. Per quanto riguarda il problema di ridurre il grado delle soluzioni, che nel caso scalare abbiamo affrontato minimizzando il grado di una delle componenti della soluzione, adotteremo qui un punto di vista diverso, cercando di minimizzare i gradi di riga della matrice $[X(z)\ Y(z)]$, rendendo "uniformemente bassi" i gradi delle due componenti della soluzione.

<< $*$ **Proposizione 3.6.3** [Soluzione di grado minimo] *Tra tutte le coppie risolutive $(X(z), Y(z))$ dell'equazione (3.69) ce n'è almeno una $(X_{\min}(z), Y_{\min}(z))$ per la quale sono simultaneamente soddisfatte le condizioni*

$$\deg \mathrm{riga}_i[X_{\min}(z) \mid Y_{\min}(z)] \le \deg \mathrm{riga}_i[X(z) \mid Y(z)] \tag{3.76}$$

per $i = 1, 2, \ldots, \ell$, e per ogni altra coppia risolutiva $(X(z), Y(z))$. Una tale soluzione è detta soluzione a gradi di riga minimi.

Dimostrazione La prova dell'esistenza è costruttiva: viene fornito qui di seguito l'algoritmo per ottenere una soluzione con le proprietà suddette.

Sia $(\bar{X}(z), \bar{Y}(z))$ una soluzione particolare di (3.69) e si consideri la soluzione generale data in (3.75), dove non è restrittivo supporre che $\bar{V}(z)$ sia non solo prima a sinistra ma anche ridotta per righe. Notiamo che la riga i-esima della generica soluzione $(X(z), Y(z))$ può essere espressa in funzione della riga i-esima della soluzione $(\bar{X}(z), \bar{Y}(z))$ nella forma

$$\mathrm{riga}_i[X(z) \mid Y(z)] = \mathrm{riga}_i[\bar{X}(z) \mid \bar{Y}(z)] + \mathrm{riga}_i Q(z)\bar{V}(z), \tag{3.77}$$

e quindi per minimizzare il grado della riga i-esima della soluzione è sufficiente far variare la riga i-esima della matrice parametro $Q(z)$. In base all'algoritmo 3.5.11, adattato al caso di vettori riga e di matrici ridotte per righe, si determina un vettore riga $\tilde{\mathbf{q}}_i(z)$ in modo che il grado di

$$\mathrm{riga}_i[\bar{X}(z) \mid \bar{Y}(z)] + \tilde{\mathbf{q}}_i(z)\bar{V}(z)$$

sia minimo. Posto allora

$$\tilde{Q}(z) := \begin{bmatrix} \tilde{\mathbf{q}}_1(z) \\ \tilde{\mathbf{q}}_2(z) \\ \vdots \\ \tilde{\mathbf{q}}_\ell(z) \end{bmatrix},$$

la soluzione a gradi di riga minimi è data allora da

$$[X_{\min}(z) \mid Y_{\min}(z)] := [\bar{X}(z) \mid \bar{Y}(z)] + \tilde{Q}(z)\bar{V}(z). \qquad * >> \blacksquare$$

3.7 Matrici polinomiali di Laurent

I risultati sulle matrici polinomiali contenuti nei precedenti paragrafi possono essere adattati alle matrici polinomiali di Laurent. La differenza più importante deriva dal fatto che elementi invertibili dell'anello $\mathbb{F}[z, z^{-1}]$ sono tutti i monomi αz^h, $\alpha \in \mathbb{F} \setminus \{0\}$ e $h \in \mathbb{Z}$.

Anche la classe delle trasformazioni elementari è più ampia, dal momento che le matrici $\tilde{E}_1$ possono avere sulla diagonale arbitrari monomi non nulli e nelle matrici $\tilde{E}_3$ il polinomio $p(z, z^{-1})$ è un elemento di $\mathbb{F}[z, z^{-1}]$.
Applicando trasformazioni elementari di questo tipo alle righe e alle colonne di una matrice polinomiale di Laurent $M(z, z^{-1}) \in \mathbb{F}[z, z^{-1}]^{p\times m}$ è possibile ridurla alla forma canonica di Smith

$$\Gamma(z, z^{-1}) = \mathrm{diag}\{\gamma_1(z, z^{-1}), \ldots, \gamma_r(z, z^{-1})\}_{p\times m},$$

dove r è il rango di $M(z, z^{-1})$ e i polinomi $\gamma_i(z, z^{-1})$ sono univocamente determinati dalla condizione che appartengano a $\mathbb{F}[z]$, siano monici, abbiano termine noto non nullo e soddisfino $\gamma_i|\gamma_{i+1}$.
Ricorrendo soltanto a trasformazioni elementari di riga, si riduce $M(z, z^{-1})$ in forma di Hermite. In particolare, se $M(z, z^{-1})$ ha rango m si ottiene una matrice

$$H(z, z^{-1}) = \begin{bmatrix} h_{11}(z, z^{-1}) & * & * & * \\ & h_{22}(z, z^{-1}) & * & * \\ & & \ddots & * \\ & & & h_{mm}(z, z^{-1}) \\ & & 0 & \\ & & & \end{bmatrix},$$

dove i polinomi $h_{ij}(z, z^{-1}) \in \mathbb{F}[z, z^{-1}]$, $i < j$, soddisfano la condizione $\delta(h_{ij}(z, z^{-1})) < \delta(h_{jj}(z, z^{-1}))$.
Sono matrici unimodulari le matrici quadrate il cui determinante è un monomio non nullo. Si verifica anche in questo caso che sono unimodulari tutti e soli i prodotti di matrici elementari e che la loro forma di Smith è la matrice identità.

Le proprietà di fattorizzazione delle matrici polinomiali di Laurent sono essenzialmente le stesse delle matrici ad elementi in $\mathbb{F}[z]$.
In particolare, la matrice $\Delta(z, z^{-1}) \in \mathbb{F}[z, z^{-1}]^{m\times m}$ è un *divisore (o fattore) destro di* $M(z, z^{-1}) \in \mathbb{F}[z, z^{-1}]^{p\times m}$ se esiste una matrice $\bar{M}(z, z^{-1}) \in \mathbb{F}[z, z^{-1}]^{p\times m}$ tale che

$$M(z, z^{-1}) = \bar{M}(z, z^{-1})\Delta(z, z^{-1}), \tag{3.78}$$

ed è un *divisore destro massimo* di $M(z,z^{-1})$ se per ogni altra fattorizzazione di $M(z,z^{-1})$

$$M(z,z^{-1}) = \tilde{M}(z,z^{-1})\tilde{\Delta}(z,z^{-1})$$

esiste $Q(z,z^{-1}) \in \mathbb{F}[z,z^{-1}]^{m\times m}$ per cui vale

$$\Delta(z,z^{-1}) = Q(z,z^{-1})\tilde{\Delta}(z,z^{-1}). \tag{3.79}$$

$M(z,z^{-1})$ è *prima a destra* se in ogni fattorizzazione del tipo (3.78) la matrice $\Delta(z,z^{-1})$ è unimodulare.

Le matrici prime a destra ammettono un insieme di descrizioni equivalenti analogo a quello della Proposizione 3.3.3. Anche l'algoritmo per la determinazione di un divisore destro massimo si ottiene estendendo quello descritto nel paragrafo 3.3.2, a condizione di considerare la forma di Hermite su $\mathbb{F}[z,z^{-1}]$.

- ESERCIZIO 3.7.1 (*i*) $M(z,z^{-1}) \in \mathbb{F}[z,z^{-1}]^{p\times m}$ è prima a destra se e solo se esiste una matrice unimodulare $V(z,z^{-1}) \in \mathbb{F}[z,z^{-1}]^{m\times m}$ tale che $M(z,z^{-1})V(z,z^{-1}) = \tilde{M}(z) \in \mathbb{F}[z]^{p\times m}$ è prima a destra come matrice polinomiale "standard".

  (*ii*) Se $M(z) \in \mathbb{F}[z]^{p\times m}$ è prima a destra come matrice a elementi nell'anello $\mathbb{F}[z]$ allora lo è come matrice a elementi in $\mathbb{F}[z,z^{-1}]$ (Suggerimento: considerare l'inversa sinistra);

  (*iii*) Se $M(z) \in \mathbb{F}[z]^{p\times m}$ è prima a destra come matrice a elementi nell'anello $\mathbb{F}[z,z^{-1}]$, allora il fattore destro massimo di $M[z]$ nell'anello $\mathbb{F}[z]$ ha per determinante una potenza di z (Suggerimento: esiste una matrice $L(z) \in \mathbb{F}[z]^{m\times p}$ soddisfacente $L(z)M(z) = \mathrm{diag}\{z^{\nu_1}, z^{\nu_2}, \ldots, z^{\nu_m}\}$ e nella fattorizzazione $M(z) = \bar{M}(z)\Delta(z)$ il fattore destro massimo $\Delta(z)$ è un divisore destro di $\mathrm{diag}\{z^{\nu_1}, z^{\nu_2}, \ldots, z^{\nu_m}\}$).

  (*iv*) Se, viceversa, il fattore destro massimo $\Delta(z)$ di $M[z]$ nell'anello $\mathbb{F}[z]$ ha per determinante una potenza di z, allora $M(z) \in \mathbb{F}[z]^{p\times m}$ è prima a destra come matrice a elementi nell'anello $\mathbb{F}[z,z^{-1}]$ (Suggerimento: se $\bar{L}(z)$ è l'inversa sinistra polinomiale di $\bar{M}(z)$, si verifichi $[\Delta(z)^{-1}\bar{L}(z)]M(z) = I_m$).

Infine, le condizioni di risolubilità e la struttura della soluzione generale di un'equazione diofantea su $\mathbb{F}[z]$ presentate nel paragrafo 3.6 valgono anche nel caso di equazioni diofantee su $\mathbb{F}[z,z^{-1}]$.

<< $*$ Per quanto riguarda il grado, ad ogni vettore non nullo in $\mathbb{F}[z,z^{-1}]^p$ (in particolare, ad ogni polinomio non nullo in $\mathbb{F}[z,z^{-1}]$)

$$\mathbf{v}(z,z^{-1}) = \mathbf{v}_\ell z^\ell + \mathbf{v}_{\ell+1} z^{\ell+1} + \ldots + \mathbf{v}_L z^L, \qquad \mathbf{v}_\ell, \mathbf{v}_L \neq \mathbf{0},$$

associamo il *grado minimo*

$$\min\deg \mathbf{v} := \ell$$

e il *grado massimo*

$$\max\deg \mathbf{v} := L$$

Inoltre, se $\mathbf{i} = (i_1, i_2, \ldots, i_m)$ è una m-upla ordinata, con $1 \le i_1 < i_2 < \ldots < i_m \le p$, indichiamo con $M^{(\mathbf{i})}(z,z^{-1})$, $M_{hc}^{(\mathbf{i})}$ e $M_{lc}^{(\mathbf{i})}(z)$ le sottomatrici ottenute selezionando rispettivamente in $M(z,z^{-1})$, M_{hc} e M_{lc} le righe di indici $i_1, i_2, \ldots, i_m$.

Definizione 3.7.1 *Sia $M(z,z^{-1}) \in \mathbb{F}[z,z^{-1}]^{p\times m}$ e siano $K_1, K_2, \ldots, K_m$ e $k_1, k_2, \ldots, k_m$ rispettivamente i gradi massimi ed i gradi minimi delle sue colonne. Diciamo che $M(z,z^{-1})$ è ridotta per colonne se*

$$\mathrm{rank} M(z,z^{-1}) = m$$

e

$$\max_{\mathbf{i}}\{\max\deg(\det M^{(\mathbf{i})})\} - \min_{\mathbf{i}}\{\min\deg(\det M^{(\mathbf{i})})\} = \sum_{j=1}^{m} K_j - \sum_{j=1}^{m} k_j \tag{3.80}$$

dove $\mathbf{i}$ *descrive tutte le* m*-uple ordinate, con* $1 \le i_1 < i_2 < \ldots < i_m \le p$ *e* $\det M^{(\mathbf{i})} \neq 0$.

- ESERCIZIO 3.7.2 Si dimostri l'invarianza di

$$\max\deg(\det M^{(\mathbf{i})}) - \min\deg(\det M^{(\mathbf{i})})$$

e di

$$\max_{\mathbf{i}}\{\max\deg(\det M^{(\mathbf{i})})\} - \min_{\mathbf{i}}\{\min\deg(\det M^{(\mathbf{i})})\}$$

rispetto alla postmoltiplicazione di $M(z, z^{-1})$ per matrici unimodulari.

La condizione che $M(z, z^{-1})$ sia ridotta per colonne si riconduce a condizioni di rango su matrici di scalari, associate ai monomi di grado massimo e di grado minimo.

Per $i = 1, 2, \ldots, m$, siano $\mathbf{c}_i$ e $\mathbf{d}_i$ in $\mathbb{F}^p$ i vettori dei coefficienti dei monomi rispettivamente di grado L_i e di grado ℓ_i nella colonna i-esima della matrice $M(z, z^{-1})$ e poniamo

$$M_{hc} = [\,\mathbf{c}_1\ \mathbf{c}_2\ \ldots\ \mathbf{c}_m\,], \qquad M_{lc} = [\,\mathbf{d}_1\ \mathbf{d}_2\ \ldots\ \mathbf{d}_m\,].$$

Se $M_{hc}^{(\mathbf{i})}$ denota le sottomatrice ottenuta selezionando in M_{hc} e M_{lc} le righe indiciate da $\mathbf{i} = (i_1, i_2, \ldots, i_m)$, applicando la definizione di determinante data in (3.81) si vede che il monomio di grado $\sum_j K_j$ in $\det M^{(\mathbf{i})}(z)$ proviene solo dai prodotti dei monomi di grado K_1 in posizione $(i_{\sigma(1)}, 1)$, di grado K_2 in posizione $(i_{\sigma(2)}, 2)$, ... di grado K_m in posizione $(i_{\sigma(m)}, m)$. Esso coincide quindi con

$$\sum_{\sigma} \operatorname{sgn}(\sigma) c_{i_{\sigma(1)},1} z^{K_1} c_{i_{\sigma(2)},2} z^{K_2} \ldots c_{i-\sigma(m),m} z^{K_m}, \tag{3.81}$$

ovvero con il determinante di $M_{hc}^{(\mathbf{i})}\operatorname{diag}\{z^{K_1}\ z^{K_2}\ \ldots\ z^{K_m}\}$, ed ha coefficiente $\det M_{hc}^{(\mathbf{i})}$.
Similmente il monomio di grado $\sum_j k_j$ in $\det M^{(\mathbf{i})}(z)$ coincide con il determinante di $M_{lc}^{(\mathbf{i})}\operatorname{diag}\{z^{k_1}\ z^{k_2}\ \ldots\ z^{k_m}\}$, ed ha coefficiente $\det M_{lc}^{(\mathbf{i})}$.
Poiché nella diseguaglianza

$$\max_{\mathbf{i}}\{\max\deg(\det M^{(\mathbf{i})})\} \le \sum_{j=1}^{m} K_j$$

vale il segno di eguale se e solo se M_{hc} ha rango m, e nella

$$\min_{\mathbf{i}}\{\min\deg(\det M^{(\mathbf{i})})\} \ge \sum_{j=1}^{m} k_j$$

vale il segno di eguale se e solo se M_{lc} ha rango m, si conclude con la seguente

Proposizione 3.7.2 *$M(z, z^{-1}) \in \mathbb{F}(z, z^{-1})^{p\times m}$ è ridotta per colonne se e solo se hanno rango m sia M_{hc} che M_{lc}.* ■

Anche la condizione di predicibilità del grado può essere estesa al caso di matrici di Laurent, sia pure con una formulazione leggermente più elaborata.

Proposizione 3.7.3 *Sia $M(z, z^{-1}) \in \mathbb{F}[z, z^{-1}]^{p\times m}$ e siano $K_1, K_2, \ldots, K_m$ e $k_1, k_2, \ldots, k_m$ rispettivamente i gradi massimi ed i gradi minimi delle sue colonne. $M(z, z^{-1})$ è ridotta per colonne se e solo se, per ogni vettore non nullo $\mathbf{u}(z, z^{-1}) \in \mathbb{F}[z, z^{-1}]^m$ risulta*

$$\max\deg(M\mathbf{u}) - \min\deg(M\mathbf{u}) \tag{3.82}$$

$$= \max_{j:u_j\neq 0}\{K_j + \max\deg u_j\} - \min_{j:u_j\neq 0}\{k_j + \min\deg u_j\} \tag{3.83}$$

DIMOSTRAZIONE Se $M(z,z^{-1})$ è ridotta per colonne, M_{hc} e M_{lc} hanno rango m. Di conseguenza, con considerazioni analoghe a quelle svolte nel par. 3.5, si verifica che i polinomi di Laurent nel vettore $M(z,z^{-1})\mathbf{u}(z,z^{-1})$ hanno grado massimo $\max_{j:u_j\neq 0}\{K_j + \max\deg u_j\}$ e grado minimo $\min_{j:u_j\neq 0}\{k_j + \min\deg u_j\}$, da cui la (3.83).
Viceversa, supponiamo che una almeno fra M_{hc} e M_{lc} abbia rango minore di m, p.es. M_{lc}. Se $[\beta_1\ \beta_2\ \dots\ \beta_m]^T$ è un vettore non nullo nel nucleo di M_{lc}, nel prodotto $M(z,z^{-1})\mathbf{u}(z,z^{-1})$ con

$$\mathbf{u}(z,z^{-1}) = \begin{bmatrix} \beta_1 z^{-k_1} \\ \beta_2 z^{-k_2} \\ \vdots \\ \beta_m z^{-k_m} \end{bmatrix}$$

si ha

$$\min\deg(M\mathbf{u}) > 0 = \min_{j:u_j\neq 0}\{k_j + \min\deg u_j\}$$

Poiché comunque risulta

$$\max\deg(M\mathbf{u}) \leq \max_{j:u_j\neq 0}\{K_j + \max\deg u_j\}$$

la (3.83) non può essere verificata. ■

- ESERCIZIO 3.7.3 (*i*) Una matrice $M(z)\in\mathbb{F}[z]^{p\times m}$ prima a destra e ridotta per colonne come matrice ad elementi in $\mathbb{F}[z]$ è anche prima a destra e ridotta per colonne come matrice a elementi in $\mathbb{F}[z,z^{-1}]$ (Suggerimento: si applichi la Proposizione 3.7.4, tenendo conto degli esercizi 3.3.4 e 3.7.1).

  (*ii*) Si interpreti la condizione espressa nella Proposizione 3.7.5 come una proprietà di predicibilità del grado per il sistema a risposta impulsiva finita (non necessariamente causale) descritto dalla matrice $M(z,z^{-1})$.

Infine, è importante notare che ogni matrice $M(z,z^{-1})\in\mathbb{F}[z,z^{-1}]^{p\times m}$ di rango m può essere ridotta per colonne postmoltiplicandola per un'opportuna matrice unimodulare di Laurent. Riportiamo qui di seguito la procedura da seguire.

1) Ricorrendo eventualmente a matrici unimodulari del tipo

$$\mathrm{diag}\{z^{\nu_1}, z^{\nu_2}, \dots, z^{\nu_m}\} \tag{3.84}$$

non è restrittivo supporre che $M(z,z^{-1})$ non contenga potenze negative. Inoltre, ricorrendo alla tecnica di riduzione esposta in par. 3.5, possiamo anche ritenere che M_{hc} abbia rango m. Infine, moltiplicando a destra per una ulteriore matrice (3.84), si può anche ritenere che i gradi minimi di colonna k_i nella matrice $M(z,z^{-1})$ siano tutti nulli.

2) Se M_{lc} non ha rango m, sia

$$\mathbf{b} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} \in \mathbb{F}^m$$

un vettore non nullo nel nucleo di M_{lc} e, tra le colonne di $M(z,z^{-1})$ corrispondenti ad indici i per i quali β_i è diverso da zero, ne scegliamo una, la j-esima, tale che il grado K_j sia massimo. Nella matrice

$$M'(z,z^{-1}) = M(z,z^{-1})[\mathbf{e}_1 \mid \dots \mid \mathbf{e}_{j-1} \mid \mathbf{b} \mid \mathbf{e}_{j+1} \mid \dots \mid \mathbf{e}_m] =: M(z,z^{-1})U$$

la colonna j-esima $M(z,z^{-1})\mathbf{b}$ ha grado massimo $K'_j = K_j$, grado minimo $k'_j > 0$, e le altre colonne coincidono con le corrispondenti in $M(z,z^{-1})$. Ma allora in

$$M''(z,z^{-1}) = M'(z,z^{-1})\mathrm{diag}\{1,\dots,1,z^{-k'_j},1,\dots,1\}$$

i gradi massimi di colonna sono $K_1,\dots,K_j-k'_j,\dots K_m$, la matrice $M''_{hc} = M_{hc}/U$ ha rango m, e i gradi minimi di colonna sono tutti nulli.

3) Se M''_{lc} ha rango m si è pervenuti ad una matrice ridotta per colonne, altrimenti si ripete la procedura. Poiché ad ogni iterazione diminuisce il valore della somma dei gradi massimi di colonna, si ottiene una matrice ridotta per colonne in un numero finito di passi. $*>>$

3.8 << ∗ Complementi

In questo paragrafo vogliamo accennare alla possibilità di estendere le nozioni di matrice elementare, di matrice unimodulare e di forma di Smith ad arbitrari domini euclidei e, più in generale, a qualsiasi dominio a ideali principali. Mentre dal punto di vista formale il procedimento che seguiremo ricalca molto da vicino (e nel caso dei domini euclidei, quasi pedissequamente) quello dei paragrafi 3.1 e 3.2, i risultati hanno un notevole interesse, in quanto nel seguito ci serviremo anche di PID diversi da $\mathbb{F}[z]$, ad esempio l'anello $\mathcal{S}$ delle funzioni razionali stabili, l'anello $\mathbb{F}_\infty(z)$ delle funzioni razionali proprie, l'anello $\mathbb{F}_\alpha(z)$ delle funzioni razionali prive di poli in α, etc.

Definizione 3.8.1 *Due matrici M ed N in $D^{p\times m}$, D un PID, si dicono equivalenti se esistono una matrice invertibile $U \in D^{p\times p}$ e una matrice invertibile $V \in D^{m\times m}$ tali che*

$$N = UMV.$$

Tra le matrici invertibili, possiamo anche qui includere le matrici elementari, con una delle seguenti strutture:

$$\tilde{E}_1 := \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & u & \\ & & & & 1 \\ & & & & & \ddots \\ 0 & & \cdots & & 0 & 1 \end{bmatrix}, \quad \tilde{E}_2 := \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 0 & \cdots & 1 & & \\ & & & & \ddots & & & \\ & & & 1 & \cdots & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \\ & & & & & & & & 1 \end{bmatrix},$$

$$\tilde{E}_3 := \begin{bmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & d & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix},$$

dove u e d sono rispettivamente un elemento invertibile e uno arbitrario di D. È chiaro che $\tilde{E}_1, \tilde{E}_2, \tilde{E}_3$ inducono operazioni elementari sulle righe o sulle colonne della matrice M, a seconda che vengano applicate a sinistra o a destra di M.

In generale, tuttavia, dovremo far ricorso anche ad altre matrici invertibili, del tipo

$$\tilde{F} = \begin{bmatrix} f_{11} & f_{12} & & 0 \\ f_{21} & f_{22} & & \\ & 0 & & I \\ & & & \end{bmatrix}, \tag{3.85}$$

dove $\begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \in D^{2\times 2}$ è invertibile. Esse sono dette matrici secondarie, e possono non essere riconducibili a prodotti di matrici elementari quando D non è euclideo.

Proposizione 3.8.2 *Se M è in $D^{p\times m}$, D un dominio a ideali principali, allora esistono due matrici invertibili $U \in D^{p\times p}$ e $V \in D^{m\times m}$ per cui si ha*

$$UMV = \operatorname{diag}\{\gamma_1, \gamma_2, \ldots, \gamma_r\}_{p\times m} := \Gamma \tag{3.86}$$

dove $\gamma_i|\gamma_j$ *se* $i|j$.

DIMOSTRAZIONE L'argomento si basa sulla lungezza degli elementi non nulli di un PID, anziché sulla nozione di grado di cui ci siamo serviti nel caso dei polinomi. La lunghezza $\ell(a)$ di un elemento $a \neq 0$ è il numero dei fattori che compaiono in qualsiasi fattorizzazione del tipo $a = p_1 p_2 \cdots p_r$, con p_i irriducibili. Inoltre, poniamo per convenzione $\ell(u) = 0$ per ogni elemento invertibile $u \in D$.
Se $M = 0$, non c'è nulla da provare. Se $M \neq 0$, con operazioni elementari sulle righe e sulle colonne di M portiamo in posizione $(1,1)$ un elemento non nullo di lunghezza minima. Avremo così

$$U^{(1)} M V^{(1)} = M^{(1)} = [m_{ij}], \tag{3.87}$$

con $\ell(m_{11}^{(1)}) \leq \ell(m_{ij}^{(1)})$ per ogni $m_{ij}^{(1)} \neq 0$.
Se $m_{12}^{(1)} = m_{11}^{(1)}\beta$, sommando alla seconda colonna di $M^{(1)}$ la prima moltiplicata per $-\beta$ si ottiene una matrice in cui è nullo l'elemento in posizione $(1,2)$. Se $m_{11}^{(1)} \not| m_{12}^{(1)}$, allora $d := \mathrm{MCD}(m_{11}^{(1)}, m_{12}^{(1)})$ ha lunghezza $\ell(d) < \ell(m_{11}^{(1)})$, ed esistono $x, y \in D$ per cui vale $d = x m_{11}^{(1)} + y m_{12}^{(1)}$. Inoltre, se $m_{11}^{(1)} = ad$, $m_{12}^{(1)} = bd$, si ha anche $1 = xa + yb$.
Ma allora $\begin{bmatrix} x & -b \\ y & a \end{bmatrix}$ è invertibile in $D^{2\times 2}$, e nel prodotto

$$M^{(2)} = M^{(1)} \begin{bmatrix} x & -b & \\ y & a & \\ & & I \end{bmatrix} \tag{3.88}$$

la prima riga è $\begin{bmatrix} d & 0 & m_{13}^{(1)} & \cdots & m_{1m}^{(1)} \end{bmatrix}$, con $\ell(d) < \ell(m_{11}^{(1)})$. Permutando le colonne, si porta in posizione $(1,2)$ un elemento non nullo fra $m_{1i}^{(1)}$, $i = 3, \ldots, m$ e si ripete il procedimento precedente per annullarlo. In un numero finito di passi si ottiene una matrice $M^{(r)} = MU^{(r)}$ in cui la prima riga è $\begin{bmatrix} m_{11}^{(r)} & 0 & \cdots & 0 \end{bmatrix}$ e $\ell(m_{11}^{(r)}) < \ell(m_{ij}^{(r)})$ per ogni $m_{ij}^{(r)} \neq 0$.
Con operazioni elementari e secondarie sulle colonne, si ottiene poi

$$M^{(s)} = V^{(s)} M U^{(s)} = \begin{bmatrix} m_{11}^{(s)} & 0 & 0 & \cdots & 0 \\ 0 & & & & \\ 0 & & & & \\ \vdots & & M_2^{(s)} & & \\ 0 & & & & \end{bmatrix} \tag{3.89}$$

Si prosegue infine come nella prova della forma di Smith per l'anello $\mathbb{F}[z]$. L'unica differenza è che, anziché il grado, si riduce progressivamente la lunghezza. ■

Gli elementi $\gamma_1, \gamma_2, \ldots, \gamma_r$ di Γ non sono individuati univocamente, ma a meno di un elemento invertibile di D. La prova, lasciata per esercizio, si basa sulle relazioni

$$\begin{aligned} &\gamma_1 = \text{MCD minori di } M \text{ di ordine } 1 \\ &\gamma_2 = \text{MCD minori di } M \text{ di ordine } 2/\text{MCD minori di } M \text{ di ordine } 1 \\ &\cdots\cdots \end{aligned} \tag{3.90}$$

- ESERCIZIO 3.8.1 Sono fatti equivalenti:
  (i) $M \in D^{m\times m}$ ha un'inversa in $D^{m\times m}$;
  (ii) $\det(M)$ è elemento invertibile di D;
  (iii) M è prodotto di matrici elementari e secondarie;
  (iv) la forma di Smith di M è $\Gamma = \mathrm{diag}\{\gamma_1, \gamma_2, \ldots, \gamma_m\}$, γ_i invertibili, e quindi anche $\Gamma = I_m$.

  Esempio 3.8.1 Sia $\mathbb{F}_\infty(z)$ l'insieme delle funzioni razionali proprie, i.e. con valutazione all'infinito non negativa. È facile constatare che $\mathbb{F}_\infty(z)$ è un dominio. Gli elementi invertibili sono tutte le funzioni razionali $f(z)$ con $v_\infty(f) = 0$, ovvero con grado del numeratore e del denominatore eguali.

Ogni elemento di $\mathbb{F}_\infty(z)$ differisce per un fattore invertibile da una potenza di z^{-1}, essendo

$$f(z) = \frac{n(z)}{d(z)} = \left(\frac{n(z)z^\nu}{d(z)}\right)\frac{1}{z^\nu}$$

con $\nu = \deg d - \deg n \geq 0$.
$\mathbb{F}_\infty(z)$ è un PID, con ideali principali

$$(1),\ \left(\frac{1}{z}\right),\ \left(\frac{1}{z^2}\right), \ldots$$

e con $1/z$ unico elemento irriducibile (a meno di fattori invertibili). Quindi $v_\infty(f)$ può essere assunta come lunghezza di ogni elemento non nullo $f(z) \in \mathbb{F}_\infty(z)$.
La forma di Smith di una matrice a elementi in $\mathbb{F}_\infty(z)$ è sempre riconducibile a una matrice del tipo

$$\Gamma = \mathrm{diag}\left\{\frac{1}{z^{\nu_1}}, \frac{1}{z^{\nu_2}}, \ldots \frac{1}{z^{\nu_r}}\right\}_{p\times m}, \quad \nu_1 \leq \nu_2 \leq \ldots \leq \nu_r$$

- ESERCIZIO 3.8.2 Indichiamo con $\mathbb{F}_\alpha(z)$ l'insieme delle funzioni razionali prive di poli in α, $\alpha \in \bar{\mathbb{F}}$, ovvero delle funzioni razionali $f(z) \in \mathbb{F}(z)$ per cui $v_\alpha \geq 0$.
  i) Si verifichi che $\mathbb{F}_\alpha(z)$ è un dominio, in cui gli elementi invertibili sono quelli per cui v_α è nullo;
  ii) Ogni elemento di $\mathbb{F}_\alpha(z)$ si esprime nella forma

$$f(z) = (z-\alpha)^\nu \ \frac{n(z)}{d(z)}$$

  con $\nu = v_\alpha(f)$ e $n(z)/d(z)$ invertibile in $\mathbb{F}_\alpha(z)$;
  iii) $\mathbb{F}_\alpha(z)$ è un PID. Quali sono i suoi elementi irriducibili?
  iv) La forma di Smith di una matrice con elementi in $\mathbb{F}_\alpha(z)$ è del tipo

$$\Gamma = \mathrm{diag}\{(z-\alpha)^{\nu_1}, (z-\alpha)^{\nu_2}, \ldots (z-\alpha)^{\nu_r}\}_{p\times m}, \quad 0 < \nu_1 \leq \nu_2 \leq \ldots \leq \nu_r \qquad * >>$$

3.9 Riferimenti bibliografici

Un riferimento classico per le trasformazioni elementari e le forme di Smith e di Hermite è

1. F.R.Gantmacher "The Theory of Matrices", voll.I e II, Chelsea, New York, 1959.

Interessanti estensioni a matrici con elementi in un generico PID si possono trovare in

2. S.MacLane e G.Birkhoff "Algebra", MacMillan, Londra, 1967

3. M.Vidyasagar "Control System Synthesis: a Factorization Approach" MIT Press, 1985.

Per un quadro generale sull'impiego delle matrici polinomiali conviene consultare

4. T.Kailath "Linear Systems", Prentice Hall, Englewood Cliffs, 1980

che fornisce un'informazione ampia, ma non sempre sistematica, e il meno recente

5. H.H.Rosenbrock "State-space and Multivariable Theory", Wiley, New York, 1970

scritto da uno dei pionieri del metodo polinomiale. Sui problemi di riduzione per colonne o per righe si vedano, oltre alle ultime due opere citate

6. F.M.Callier e C.A.Desoer "Multivariable Feedback Systems", Springer-Verlag, New York, 1982

7. W.A.Wolovich "Linear Multivariable Systems", Springer-Verlag, New York, 1974.

Infine, sulle equazioni diofantee matriciali, su vari aspetti algoritmici connessi e sul controllo dead-beat sono molto utili

8. V.Kučera "Discrete Linear Control: the Polynomial Equation Approach" Wiley, 1979.

9. V.Kučera "Analysis and Design of Discrete Linear Control Systems", Academia, Praga, 1991.