Capitolo 5

Funzioni di matrici

Lo studio dei sistemi lineari di equazioni alle differenze e differenziali, richiede l'introduzione di un importante strumento metodologico: le funzioni di matrici. Nel seguito, pertanto, cercheremo di dare un senso alla scrittura $f(A)$, essendo $f : \mathbb{C} \to \mathbb{C}$ ed A matrice quadrata di dimensione $\bar{m}$. Cominciamo con il caso, assai più semplice, in cui $f(z) = p(z) \in \Pi_k$:

$$p(z) = \sum_{i=0}^{k} p_i z^i.$$

In tal caso, si ha:

$$p(A) = \sum_{i=0}^{k} p_i A^i \in \mathbb{C}^{\bar{m}\times\bar{m}}.$$

Inoltre, se $\lambda \in \sigma(A)$, lo *spettro* di A, con autovettore $\mathbf{v}$, ovvero

$$A\mathbf{v} = \lambda\mathbf{v},$$

allora è facile dimostrare il seguente risultato.

Teorema 5.1 $\lambda \in \sigma(A) \;\Rightarrow\; p(\lambda) \in \sigma(p(A))$ *e l'autovettore corrispondente è lo stesso.*

Tuttavia, vi sono alcune differenze fondamentali tra i polinomi di una variabile ed i polinomi di matrice. Infatti è se $z \in \mathbb{C}$, allora $z \neq 0 \;\Leftrightarrow\; z^n \neq 0$, $n \geq 0$. Per le matrici, questo può non essere vero. Infatti, per esempio, denotando con O la matrice nulla, [1]

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq O \qquad \text{ma} \qquad A^2 = O.$$

Più in generale, vale il seguente risultato.

Teorema 5.2 (di Cayley-Hamilton) *Sia* $p(\lambda) = \det(A - \lambda I)$ *il* polinomio caratteristico *di A. Allora,* $p(A) = O$.

Supponiamo, ora che il *polinomio caratteristico* di A sia

$$p(\lambda) = \prod_{i=1}^{\nu} (\lambda - \lambda_i)^{\bar{m}_i}, \qquad \lambda_i \neq \lambda_j, \text{ se } i \neq j, \tag{5.1}$$

[1] La sua dimensione sarà sempre deducibile dal contesto.

ovvero $\{\lambda_1, \ldots, \lambda_\nu\}$ sono gli autovalori distinti di A aventi *molteplicità algebrica*, rispettivamente $\bar{m}_1, \ldots, \bar{m}_\nu$. Pertanto, esiste un polinomio monico di grado

$$\sum_{i=1}^{\nu} \bar{m}_i = \bar{m} \equiv \deg(p)$$

che, calcolato in A, vale la matrice nulla. Definiamo ora il *polinomio minimale* (o *minimo*) di A, sia esso $\psi(z)$, come il polinomio monico di grado minimo che si annulla in A. Pertanto,

$$m \equiv \deg(\psi) \leq \bar{m}.$$

Vale il seguente risultato.

Teorema 5.3 *Ogni radice di $p(z)$ è radice di $\psi(z)$, e viceversa.*

Dimostrazione. Poiché $m \leq \bar{m}$, è possibile dividere $p(z)$ per $\psi(z)$, ottenendo:

$$p(z) = q(z)\psi(z) + r(z),$$

con $\deg(r) < m$. Poiché $p(A) = \psi(A) = O$, segue che $r(A) = O$ e, pertanto, $r(z) \equiv 0$ perché, diversamente, si avrebbe $\deg(r) \geq m$, essendo $\psi(z)$ il polinomio minimale. Quindi, ogni radice di $\psi(z)$ è radice di $p(z)$. Sia ora λ un autovalore di A con autovetture $\mathbf{v}$. Pertanto, $p(\lambda) = 0$ e $A^j\mathbf{v} = \lambda^j\mathbf{v}$, $j \geq 0$. Pertanto, $\psi(\lambda)\mathbf{v} = \psi(A)\mathbf{v} = O$ e, quindi, $\psi(\lambda)=0$. □

Corollario 5.1 *Sia $p(z)$, come definito in (5.1), il polinomio caratteristico di A e sia*

$$\psi(z) = \prod_{i=1}^{\nu} (z - \lambda_i)^{m_i} \tag{5.2}$$

il suo polinomio minimale. Allora

$$1 \leq m_i \leq \bar{m}_i, \qquad i = 1, \ldots, \nu.$$

Il significato della molteplicità m_i di λ_i nel polinomio minimale sarà chiarito successivamente, sebbene il precedente corollario stabilisca che essa è non maggiore della molteplicità algebrica. Ovviamente, si avrà:

$$\sum_{i=1}^{\nu} m_i = m \equiv \deg(\psi). \tag{5.3}$$

Utilizzando argomenti analoghi a quelli visti nella dimostrazione del Teorema 5.3, si dimostra il seguente risultato.

Teorema 5.4 *Sia $h(z)$ un polinomio di grado maggiore di m, il grado del polinomio minimale di una data matrice A. Pertanto, dividendo $h(z)$ per $\psi(z)$, si ottiene:*

$$h(z) = q(z)\psi(z) + r(z), \qquad \deg(r) < m.$$

In tal caso, si ha: $h(A) = r(A)$.

Osservazione 5.1 *Il precedente teorema dimostra come polinomi di grado diverso possano dare la stessa matrice, se calcolati in A. In particolare, qualunque polinomio calcolato in A si può ottenere calcolando in A un opportuno polinomio di grado* <u>minore</u> *del grado del polinomio minimale di A.*

Siano ora $h(z)$ e $g(z)$ due polinomi diversi tali che

$$h(A) = g(A). \tag{5.4}$$

Pertanto, anche la loro differenza,

$$d(z) = h(z) - g(z),$$

è un polinomio. Inoltre, si avrà $d(A) = h(A) - g(A) = O$. Se $\psi(z)$ è il polinomio minimale di A, come definito in (5.2), si avrà, dunque,

$$d(z) = q(z)\psi(z).$$

Poiché (vedi (5.2)) per ogni $i = 1\ldots,\nu$,

$$\psi^{(j)}(\lambda_i) = 0, \qquad j = 0,\ldots,m_i - 1,$$

segue che,

$$d^{(j)}(\lambda_i) = \sum_{k=0}^{j} \binom{j}{k} q^{(j-k)}(\lambda_i)\psi^{(k)}(\lambda_i) = 0, \qquad i = 1,\ldots,\nu, \quad j = 0,\ldots,m_i - 1.$$

ovvero,

$$h^{(j)}(\lambda_i) = g^{(j)}(\lambda_i), \qquad i = 1,\ldots,\nu, \quad j = 0,\ldots,m_i - 1. \tag{5.5}$$

Definizione 5.1 *Due polinomi $h(z)$ e $g(z)$ che soddisfino la (5.5) si dicono* assumere gli stessi valori sullo spettro di A.

Supponiamo, ora, che due polinomi $h(z)$ e $g(z)$ soddisfino la (5.5). Segue che la loro differenza è divisibile per $\psi(z)$, il polinomio minimale di A e, pertanto, risulta verificata la (5.4). I precedenti argomenti, dimostrano il seguente risultato.

Teorema 5.5 *Due polinomi $h(z)$ e $g(z)$ soddisfano $h(A) = g(A)$ se e solo se essi assumono gli stessi valori sullo spettro di A.*

Questo risultato, permette di dare le seguenti definizioni.

Definizione 5.2 *Una funzione $f(z)$ è* definita sullo spettro di A *se sono definiti:*

$$f^{(j)}(\lambda_i), \qquad i = 1,\ldots,\nu, \quad j = 0,\ldots,m_i - 1.$$

Osservazione 5.2 *Pertanto, se una funzione è olomorfa in un dominio contenente $\sigma(A)$, essa sarà definita sul suo spettro.*

Definizione 5.3 *Sia $f(z)$ una funzione definita sullo spettro di A, e sia $g(z)$ il polinomio che assume gli stessi valori di $f(z)$ sullo spettro di A, ovvero*

$$f^{(j)}(\lambda_i) = g^{(j)}(\lambda_i), \qquad i = 1,\ldots,\nu, \quad j = 0,\ldots,m_i - 1. \tag{5.6}$$

Allora

$$f(A) \equiv g(A). \tag{5.7}$$

Corollario 5.2 $\lambda \in \sigma(A) \Rightarrow f(\lambda) \in \sigma(f(A))$ *e l'autovettore corrispondente è lo stesso.*

Dimostrazione. Da (5.7) e dal Teorema 5.1 segue che se λ_i è autovalore di A con autovettore $\mathbf{x}_i$, allora $g(\lambda_i)$ è autovalore di $g(A) = f(A)$ con lo stesso autovettore. La tesi si completa osservando che dalla (5.6) segue che $f(\lambda_i) = g(\lambda_i)$. □

Il polinomio (5.6) è un polinomio interpolante di Hermite generalizzato, che si dimostra esistere ed essere unico. Esso si ottiene come

$$g(\lambda) = \sum_{i=1}^{\nu}\sum_{j=1}^{m_i} f^{(j-1)}(\lambda_i)\Phi_{ij}(\lambda), \tag{5.8}$$

dove $\Phi_{ij}(\lambda) \in \Pi_{m-1}$ è univocamente individuato dalle m condizioni:

$$\Phi_{ij}^{(r-1)}(\lambda_\ell) = \delta_{i\ell}\delta_{jr}, \qquad \ell = 1, \ldots, \nu, \quad r = 1, \ldots, m_\ell.$$

Osservazione 5.3 *Gli m polinomi (vedi (5.3)) di grado $m-1$*

$$\Phi_{ij}(\lambda) \qquad i = 1, \ldots, \nu, \quad j = 1, \ldots, m_i,$$

si dimostrano essere linearmente indipendenti tra loro e, pertanto, costituiscono una base per Π_{m-1}.

Pertanto, dalle (5.7)–(5.8) segue che:

$$f(A) = \sum_{i=1}^{\nu}\sum_{j=1}^{m_i} f^{(j-1)}(\lambda_i)\Phi_{ij}(A) \equiv \sum_{i=1}^{\nu}\sum_{j=1}^{m_i} f^{(j-1)}(\lambda_i)Z_{ij}, \tag{5.9}$$

dove le m matrici

$$Z_{ij}, \qquad i = 1, \ldots, \nu, \quad j = 1, \ldots, m_i,$$

sono dette *matrici componenti* di A. Quest'ultime non dipendono da $f(\lambda)$, ma solo dalla matrice A.

Vale il seguente risultato.

Teorema 5.6 *Le m matrici componenti $\{Z_{ij}\}$ sono linearmente indipendenti.*[2]

Dimostrazione. Supponiamo che sia $\sum_{i=1}^{\nu}\sum_{j=1}^{m_i} c_{ij}Z_{ij} = O$. In tal caso, se non fossero nulli tutti i coefficienti c_{ij}, si otterrebbe che il polinomio

$$\phi(z) = \sum_{i=1}^{\nu}\sum_{j=1}^{m_i} c_{ij}\Phi_{ij}(z) \in \Pi_{m-1}$$

non sarebbe identicamente nullo, poichè i polinomi $\{\Phi_{ij}\}$ sono linearmente indipendenti. Tuttavia, considerando che $Z_{ij} = \Phi_{ij}(A)$, si otterrebbe

$$\phi(A) = O$$

con $\deg(\phi) < \deg(\psi) = m$, essendo ψ il polinomio minimale di A, il che è chiaramente assurdo. □

[2] E, pertanto, nonnulle.

Un esempio

Sia

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{5.10}$$

Questa matrice ha due autovalori semplici uguali a $\lambda_1 = 1$ e $\lambda_2 = 3$. In tal caso, quindi, il polinomio minimale e il polinomio caratteristico coincideranno. I corrispondenti polinomi $\Phi_{i1}(\lambda)$, $i = 1, 2$, sono dati dai polinomi di base di Lagrange:

$$\Phi_{11}(\lambda) = \frac{\lambda - 3}{1 - 3} = \frac{1}{2}(3 - \lambda), \qquad \Phi_{21}(\lambda) = \frac{\lambda - 1}{3 - 1} = \frac{1}{2}(\lambda - 1).$$

Le matrici componenti sono quindi date da:

$$Z_{11} = \frac{1}{2}(3I - A) = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \qquad Z_{21} = \frac{1}{2}(A - I) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{5.11}$$

Pertanto, data una generica funzione $f(z)$, si ottiene

$$f(A) = f(1)Z_{11} + f(3)Z_{21}.$$

Ad esempio:

$$e^A = e^1 Z_{11} + e^3 Z_{21} = \frac{1}{2}\begin{pmatrix} e^3 + e^1 & e^3 - e^1 \\ e^3 - e^1 & e^3 + e^1 \end{pmatrix},$$

$$\sqrt{A} = Z_{11} + \sqrt{3}Z_{21} = \frac{1}{2}\begin{pmatrix} \sqrt{3} + 1 & \sqrt{3} - 1 \\ \sqrt{3} - 1 & \sqrt{3} + 1 \end{pmatrix},$$

$$A^{-1} = Z_{11} + 3^{-1}Z_{21} = \frac{1}{2}\begin{pmatrix} \frac{1}{3} + 1 & \frac{1}{3} - 1 \\ \frac{1}{3} - 1 & \frac{1}{3} + 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

5.1 Proprietà delle matrici componenti

Preliminarmente, osserviamo che le matrici componenti *commutano* tra loro, essendo polinomi della stessa matrice. Si consideri $f(z) \equiv 1$. Pertanto, dalla (5.9) si ottiene:

$$I = \sum_{i=1}^{\nu} Z_{i1}. \tag{5.12}$$

Se si considera, invece, $f(z) = z$, si ottiene, allo stesso modo:[3]

$$A = \sum_{i=1}^{\nu} (\lambda_i Z_{i1} + Z_{i2}). \tag{5.13}$$

Da questa si ottiene:

$$A^2 = \left(\sum_{i=1}^{\nu} (\lambda_i Z_{i1} + Z_{i2}) \right)^2 = \sum_{i=1}^{\nu} \sum_{k=1}^{\nu} (\lambda_i \lambda_k Z_{i1} Z_{k1} + 2\lambda_i Z_{i1} Z_{k2} + Z_{i2} Z_{k2}). \tag{5.14}$$

[3]Nel seguito assumeremo, per semplicità di notazione, che $Z_{ij} = O$, se $j \geq m_i$. Questa condizione è infatti vera, come dimostreremo successivamente.

Tuttavia, considerando $f(z) = z^2$, dalla (5.9) si ottiene:

$$A^2 = \sum_{i=1}^{\nu}(\lambda_i^2 Z_{i1} + 2\lambda_i Z_{i2} + 2Z_{i3}).$$

Uguagliando il secondo membro di questa equazione all'ultimo membro della (5.14), si ottiene:

$$\begin{aligned} Z_{i1}Z_{k1} &= \delta_{ik}Z_{i1}, &(5.15)\\ Z_{i1}Z_{k2} &= \delta_{ik}Z_{i2}, &(5.16)\\ Z_{i2}Z_{k2} &= 2\delta_{ik}Z_{i3}. &(5.17)\end{aligned}$$

Procedendo in maniera analoga, è possibile dimostrare che:

$$\begin{aligned} Z_{ij}Z_{kr} &= O, \quad \text{per} \quad i \neq k, &(5.18)\\ Z_{i1}Z_{ij} &= Z_{ij}, \quad \text{per} \quad j \geq 1, &(5.19)\\ Z_{i2}Z_{ij} &= jZ_{i,j+1}, \quad \text{per} \quad j \geq 1. &(5.20)\end{aligned}$$

Osservazione 5.4 *In particolare, dalla (5.19) segue, per $j = 1$:*

$$Z_{i1}^2 = Z_{i1}.$$

Ovvero, le matrici $\{Z_{i1}\}$ sono idempotenti. *Questo fatto, unito alla (5.12) ed alla (5.15) ci dice che è possibile decomporre un generico vettore $\mathbf{v} \in \mathbb{C}^{\bar{m}}$ come somma di ν vettori:*

$$\mathbf{v} = \sum_{i=1}^{\nu}\mathbf{v}_i, \qquad \mathbf{v}_i = Z_{i1}\mathbf{v}, \quad i = 1, \ldots, \nu.$$

Osservando che $Z_{k1}\mathbf{v}_i = \delta_{ik}\mathbf{v}_i$, abbiamo che le matrici $\{Z_{i1}\}$ proiettano un generico vettore in ν sottospazi disgiunti, la cui unione vale $\mathbb{C}^{\bar{m}}$. Per questo, si denominano proiettori. *Più in dettaglio, avremo che i sottospazi*

$$\mathcal{M}_i = \{Z_{i1}\mathbf{v} \,:\, \mathbf{v} \in \mathbb{C}^{\bar{m}}\} \equiv Z_{i1}\mathbb{C}^{\bar{m}}, \qquad i = 1, \ldots, \nu, \tag{5.21}$$

soddisfano le condizioni:

$$\mathcal{M}_i \cap \mathcal{M}_j = \emptyset, \quad i \neq j, \qquad \cup_{i=1}^{\nu}\mathcal{M}_i = \mathbb{C}^{\bar{m}}. \tag{5.22}$$

Inoltre, si dimostra che

$$\dim(\mathcal{M}_i) = \bar{m}_i, \qquad i = 1, \ldots, \nu. \tag{5.23}$$

Infine, le matrici idempotenti Z_{i1}, $i = 1, \ldots, \nu$, possono essere riguardate come applicazioni lineari:

$$Z_{i1} \,:\, \mathbb{C}^{\bar{m}} \longrightarrow \mathcal{M}_i, \qquad i = 1, \ldots, \nu.$$

Ancora, dalla (5.20) segue, ragionando per induzione, che:

$$Z_{ij} = \frac{1}{(j-1)!}Z_{i2}^{j-1}, \qquad j \geq 2. \tag{5.24}$$

Consideriamo, ora, $f(z) = z - \lambda_k$, dove $\lambda_k \in \sigma(A)$. Si ottiene, dalle (5.15) e (5.18):

$$Z_{k1}(A - \lambda_k I) = \sum_{i=1}^{\nu}(\lambda_i - \lambda_k)Z_{k1}Z_{i1} + \sum_{i=1}^{\nu} Z_{k1}Z_{i2} \equiv Z_{k2}. \tag{5.25}$$

Da questa, tenendo conto della (5.26), si ottiene:

$$Z_{ij} = \frac{1}{(j-1)!}(A - \lambda_i I)^{j-1} Z_{i1}, \qquad j \geq 1. \tag{5.26}$$

Questa consente di riscrivere la (5.9) come:

$$f(A) = \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{f^{(j-1)}(\lambda_i)}{(j-1)!} Z_{i1}(A - \lambda_i I)^{j-1}. \tag{5.27}$$

Il prossimo risultato stabilisce che la matrice $\{Z_{i2}\}$ è *nilpotente* con *indice di nilpotenza* pari a m_i, $i = 1, \ldots, \nu$, essendo m_i la molteplicità di λ_i nel polinomio minimale.

Teorema 5.7 $Z_{i2}^{m_i} = O$, $i = 1, \ldots, \nu$.

Dimostrazione. Sia $\lambda \neq \sigma(A)$. Segue che entrambe le funzioni $(\lambda - z)$ e $(\lambda - z)^{-1}$ sono definite sullo spettro di A. Pertanto, si ottiene che:

$$(\lambda I - A) = \sum_{i=1}^{\nu} [(\lambda - \lambda_i) Z_{i1} - Z_{i2}]$$

e, inoltre, per la (5.26),

$$\begin{aligned}
(\lambda I - A)^{-1} &= \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{(j-1)!}{(\lambda - \lambda_i)^j} Z_{ij} = \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{1}{(\lambda - \lambda_i)^j} Z_{i1}(A - \lambda_i I)^{j-1} \\
&= \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{1}{(\lambda - \lambda_i)^j} Z_{i1} Z_{i2}^{j-1}.
\end{aligned}$$

Di conseguenza, dalla (5.12):

$$\begin{aligned}
I &= \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} [(\lambda - \lambda_i) Z_{i1} - Z_{i2}] \frac{1}{(\lambda - \lambda_i)^j} Z_{i1} Z_{i2}^{j-1} \\
&= \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \left[\frac{1}{(\lambda - \lambda_i)^{j-1}} Z_{i1} Z_{i2}^{j-1} - \frac{1}{(\lambda - \lambda_i)^j} Z_{i1} Z_{i2}^{j} \right] \\
&= \sum_{i=1}^{\nu} \left[Z_{i1} - \frac{1}{(\lambda - \lambda_i)^{m_i}} Z_{i1} Z_{i2}^{m_i} \right],
\end{aligned}$$

ovvero

$$Z_{i1} = Z_{i1} - \frac{1}{(\lambda - \lambda_i)^{m_i}} Z_{i2}^{m_i}, \qquad i = 1, \ldots, \nu,$$

da cui segue, evidentemente, che $Z_{i2}^{m_i} = O$, $i = 1, \ldots, \nu$. □

Come corollario di questo risultato, ricordando che $m_i \leq \bar{m}_i$ e tenendo in conto della (5.25), possiamo quindi riscrivere la (5.27) equivalentemente come:

$$f(A) = \sum_{i=1}^{\nu} \sum_{j=1}^{\bar{m}_i} \frac{f^{(j-1)}(\lambda_i)}{(j-1)!} Z_{i1}(A - \lambda_i I)^{j-1}. \tag{5.28}$$

In altri termini, nell'espressione di $f(A)$ possiamo utilizzare indifferentemente le molteplicità algebriche $\{\bar{m}_i\}$ o le molteplicità nel polinomio minimale $\{m_i\}$ degli autovalori $\{\lambda_i\}$.

Definizione 5.4 *Sia $\lambda_i \in \sigma(A)$. Diremo che l'autovalore è:*

- semplice *se $\bar{m}_i = 1$;*
- degenere *se $\bar{m}_i > 1$;*
- semisemplice *se $m_i = 1$.*

Chiaramente, un autovalore semplice è anche semisemplice. Viceversa, un autovalore semisemplice può essere degenere.

Osservazione 5.5 *È evidente che, nel caso in cui λ_i sia semplice o semisemplice, allora $Z_{i2} = O$. Pertanto, l'i-esimo termine della prima sommatoria in (5.27) e (5.28) si riduce a $f(\lambda_i)Z_{i1}$.*

Alcuni esempi

La matrice identità I $\bar{m} \times \bar{m}$ ha un unico autovalore uguale ad 1. Infatti il suo polinomio caratteristico vale

$$p(\lambda) = (\lambda - 1)^{\bar{m}}.$$

Il suo polinomio minimale, invece, è

$$\psi(\lambda) = \lambda - 1.$$

Pertanto 1 è un autovalore semisemplice, in quanto la sua molteplicità nel polinomio minimale vale 1, ma degenere, poiché la sua molteplicità algebrica è $\bar{m}$.

Una *matrice di Frobenius* (o di *compagnia*, già incontrata in precedenza) è una matrice del tipo

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \\ -a_s & -a_{s-1} & \dots & -a_2 & -a_1 \end{pmatrix}_{s\times s}. \tag{5.29}$$

Il suo polinomio caratteristico è dato (vedi Lemma 3.1) da

$$p(\lambda) = \lambda^s + \sum_{i=1}^{s} a_i \lambda^{s-i}. \tag{5.30}$$

Per queste matrici, molteplicità algebrica e molteplicità nel polinomio minimale di un autovalore coincidono. Vale, infatti, il seguente risultato.

Teorema 5.8 *Per la matrice (5.29), il polinomio minimale $\psi(\lambda)$ coincide con il polinomio carateristico $p(\lambda)$.*

Dimostrazione. Sia, per assurdo,

$$\psi(\lambda) = \sum_{i=0}^{\ell} \alpha_i \lambda^i, \qquad \ell < s,$$

con $\alpha_\ell = 1$, il polinomio minimale (monico) di A. Denotando con $\mathbf{e}_i$ l'i-esimo versore in $\mathbb{R}^s$, segue facilmente, dalla (5.29), che

$$\mathbf{e}_i^T A = \mathbf{e}_{i+1}, \qquad i = 1, \dots, \ell,$$

poiché $\ell < s$. Pertanto,

$$\mathbf{e}_1 A^i = \mathbf{e}_{i+1}, \qquad i = 0, \dots, \ell.$$

D'altronde, essendo $\psi(A) = O$, segue che

$$\mathbf{0}^T = \mathbf{e}_1^T \psi(A) = \sum_{i=0}^{\ell} \alpha_i \mathbf{e}_1^T A^i = \sum_{i=0}^{\ell} \alpha_i \mathbf{e}_{i+1} = (\underbrace{\alpha_0 \dots \alpha_\ell}_{\ell+1} \overbrace{0 \dots 0}^{s-\ell-1}).$$

Questo è chiaramente un assurdo. Poiché l'assurdo è derivato dall'ipotesi $\ell < s$, si conclude che $\deg(\psi) = s$ e, quindi, $p(\lambda) = \psi(\lambda)$. $\square$

Come semplice conseguenza, vale il seguente risultato.

Corollario 5.3 *Una matrice di Frobenius non ammette autovalori semisemplici e degeneri.*

Esercizio 5.1 *Determinare gli autovettori della matrice (5.29), nel caso in cui il polinomio caratteristico (5.30) abbia tutte radici semplici.*

5.2 Successioni di funzioni di matrici

Sia ora assegnata una successione di funzioni $\{f_k(z)\}$ definita sullo spettro di A. Diremo che la successione *converge sullo spettro di A* se:

$$\lim_{k\to\infty} f^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \qquad i = 1, \dots, \nu, \quad j = 0, \dots, m_i - 1. \tag{5.31}$$

Vale il seguente risultato.

Teorema 5.9 *Una successione di funzioni di matrici $\{f_k(A)\}$ converge a $f(A)$ se e solo se la successione $\{f_k(z)\}$ converge sullo spettro di A.*

Dimostrazione. Il fatto che se la successione converge sullo spettro di A allora

$$f_k(A) \longrightarrow_{k\to\infty} f(A),$$

discende facilmente dal fatto che è sufficiente la convergenza puntuale fornita dalle (5.31). Viceversa, se

$$f_k(A) \equiv \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{f_k^{(j-1)}(\lambda_i)}{(j-1)!} Z_{i1} Z_{i2}^{j-1} \longrightarrow_{k\to\infty} \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{f^{(j-1)}(\lambda_i)}{(j-1)!} Z_{i1} Z_{i2}^{j-1} \equiv f(A), \tag{5.32}$$

allora, moltiplicando membro a membro per $Z_{i1} Z_{i2}^{m_i-1}$ si ottiene, in virtù del Teorema 5.7,

$$f_k(\lambda_i) Z_{i1} Z_{i2}^{m_i-1} \longrightarrow_{k\to\infty} f(\lambda_i) Z_{i1} Z_{i2}^{m_i-1},$$

ovvero

$$f_k(\lambda_i) \longrightarrow_{k\to\infty} f(\lambda_i). \tag{5.33}$$

Procedendo in modo analogo, moltiplicando membro a membro la (5.32) per $Z_{i1}Z_{i2}^{m_i-2}$, si arriva a concludere che, oltre alla (5.33), vale anche

$$f_k'(\lambda_i) \longrightarrow_{k\to\infty} f'(\lambda_i).$$

La tesi si completa ripetendo il procedimento utilizzando, nell'ordine, $Z_{i1}Z_{i2}^{m_i-3}, \ldots, Z_{i1}Z_{i2}^{0}$. □

Corollario 5.4 *Sia A una matrice avente autovalori di modulo minore di 1. Allora:*

$$(I-A)^{-1} = \sum_{n=0}^{\infty} A^n. \tag{5.34}$$

Dimostrazione. Consideriamo la successione di funzioni

$$f_k(z) = \sum_{n=0}^{k} z^n, \qquad k = 0, 1, \ldots.$$

Questa successione converge a $f(z) = (1-z)^{-1}$, per $k \to \infty$, se e solo se $|z| < 1$. Quindi converge sullo spettro di A e, pertanto, la (5.34) segue. □

Corollario 5.5 *Sia A una qualunque matrice quadrata. Allora*

$$\mathrm{e}^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Dimostrazione. La dimostrazione è analoga alla precedente, considerando la successione di funzioni

$$f_k(z) = \sum_{n=0}^{k} \frac{z^n}{n!}, \qquad k = 0, 1, \ldots. \square$$

Valgono i seguenti risultati.

Teorema 5.10 $\lim_{t\to\infty} \|\mathrm{e}^{At}\|$:

- *tende a 0 se e solo se tutti gli autovalori di A hanno parte reale negativa;*
- *è limitato se e solo se tutti gli autovalori si A hanno parte reale non positiva, essendo semisemplici quelli con parte reale nulla;*
- *$\uparrow \infty$, altrimenti.*

Dimostrazione. La tesi discende facilmente, osservando che

$$\mathrm{e}^{At} = \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} t^{j-1} \mathrm{e}^{\lambda_i t} Z_{ij}. \square$$

Teorema 5.11 $\lim_{t\to\infty} \|A^n\|$:

- *tende a 0 se e solo se tutti gli autovalori di A hanno modulo minore di 1;*
- *è limitato se e solo se tutti gli autovalori si A hanno modulo non maggiore di 1, essendo semisemplici quelli di modulo 1;*

- $\uparrow \infty$, *altrimenti.*

Dimostrazione. La tesi discende facilmente, osservando che

$$A^n = \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} n^{(j-1)} \lambda_i^{n-j+1} Z_{ij}. \square$$

Teorema 5.12 *Sia dato il sistema di equazioni differenziali lineari*

$$y' = Ay, \qquad y(0) = y_0.$$

Allora, la sua soluzione è data da

$$y(t) = \mathrm{e}^{At} y_0, \qquad t \geq 0.$$

Dimostrazione. Infatti, si ottiene:

$$\begin{aligned} A\mathrm{e}^{At} &= \sum_{i=1}^{\nu} (\lambda_i Z_{i1} + Z_{i2}) \cdot \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{t^{j-1}}{(j-1)!} \mathrm{e}^{\lambda_i t} Z_{i1} Z_{i2}^{j-1} \\ &= \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{t^{j-1}}{(j-1)!} \lambda_i \mathrm{e}^{\lambda_i t} Z_{i1} Z_{i2}^{j-1} + \frac{t^{j-1}}{(j-1)} \mathrm{e}^{\lambda_i t} Z_{i1} Z_{i2}^{j} \\ &= \sum_{i=1}^{\nu} \left[\mathrm{e}^{\lambda_i t} Z_{i1} + \sum_{j=2}^{m_i} \frac{t^{j-1}\lambda_i + (j-1)t^{i-2}}{(j-1)!} \mathrm{e}^{\lambda_i t} Z_{i1} Z_{i2}^{j-1} \right] \\ &= \sum_{i=1}^{\nu} \sum_{j=1}^{m_i} \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{t^{j-1}}{(j-1)!} \mathrm{e}^{\lambda_i t} \right) Z_{i1} Z_{i2}^{j-1} \equiv \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{e}^{At}. \square \end{aligned}$$

Da questo risultato, e dal Teorema 5.10, segue immediatamente il seguente corollario.

Corollario 5.6 *La soluzione nulla dell'equazione*

$$y' = Ay$$

è:

- asintoticamente stabile *se e solo se tutti gli autovalori di A hanno parte reale negativa;*
- stabile *se e solo se tutti gli autovalori si A hanno parte reale non positiva, essendo semisemplici quelli con parte reale nulla;*
- instabile, *altrimenti.*

Analogamente, nel caso discreto si dimostra il seguente risultato.

Teorema 5.13 *La soluzione dell'equazione*

$$y_{n+1} = Ay_n, \qquad n \geq n_0,$$

è:

$$y_n = A^{n-n_0} y_{n_0},$$

dove y_{n_0} è la condizione iniziale. La soluzione nulla di tale equazione sarà:

- asintoticamente stabile *se e solo se tutti gli autovalori di A hanno autovalori di modulo minore di 1;*
- stabile *se e solo se tutti gli autovalori si A hanno modulo non maggiore di 1, essendo semisemplici quelli di modulo 1;*
- instabile, *altrimenti.*

Vale, inoltre, il seguente risultato.

Teorema 5.14 *Sia $\mathbf{v} \neq \mathbf{0}$, e sia A con autovalori tali $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_\nu|$, con λ_1 semplice. Allora $A^n\mathbf{v}$ tende, al crescere di n, ad allinearsi con l'autovettore relativo all'autovalore di modulo massimo.*

Dimostrazione. Preliminarmente osserviamo che, poiché

$$A = \lambda_1 Z_{11} + \sum_{i=2}^{\nu}(\lambda_i Z_{i1} + Z_{i2}),$$

essendo $Z_{12} = O$. Pertanto,

$$A(Z_{11}\mathbf{v}) = Z_{11}A\mathbf{v} = Z_{11}\left[\lambda_1 Z_{11} + \sum_{i=2}^{\nu}(\lambda_i Z_{i1} + Z_{i2})\right]\mathbf{v} = \lambda_1(Z_{11}\mathbf{v}).$$

Ovvero, $Z_{11}\mathbf{v}$ è autovettore relativo a λ_1. La tesi si completa considerando che, per $n \gg 1$:

$$A^n\mathbf{v} = \lambda_1^n\left[Z_{11}\mathbf{v} + \sum_{i=2}^{\nu}\sum_{j=1}^{m_i} n^{(j-1)}\frac{\lambda_i^{n-j+1}}{\lambda_1^n}Z_{ij}\right] \approx \lambda_1^n Z_{11}\mathbf{v}. \square$$

Osservazione 5.6 *Il risultato del Teorema 5.14 è alla base del* metodo delle potenze *per il calcolo dell'autovalore dominante di una matrice.*

5.3 Forma canonica di Jordan

Consideriamo i sottospazi $\mathcal{M}_i$, definiti in (5.21), aventi dimensione $\bar{m}_i$, ovvero pari alla molteplicità algebrica dell'autovalore λ_i, $i = 1\ldots,\nu$, come esposto in (5.23). Ricordiamo, inoltre, che essi soddisfano le proprietà (5.22). Inoltre, se $\mathbf{v}_i \in \mathcal{M}_i$, allora esiste $\mathbf{v} \in \mathbb{C}^{\bar{m}}$ tale che $\mathbf{v}_i = Z_{i1}\mathbf{v}$. Vale il seguente risultato.

Teorema 5.15 *Sia $Z_{i2}^{j-1}Z_{i1}\mathbf{v} \neq \mathbf{0}$, $i = 1,\ldots,p$ e $Z_{i2}^p Z_{i1}\mathbf{v} = \mathbf{0}$. Allora i vettori $Z_{i2}^{j-1}Z_{i1}\mathbf{v}$, $j = 1,\ldots,p$, sono linearmente indipendenti.*

Dimostrazione. Infatti, se fosse

$$\sum_{j=1}^{p} c_j Z_{i2}^{j-1}Z_{i1}\mathbf{v} = \mathbf{0},$$

seguirebbe che

$$Z_{i2}^{p-1}\sum_{j=1}^{p} c_j Z_{i2}^{j-1}Z_{i1}\mathbf{v} = c_1 Z_{i1}Z_{i2}^{p-1}\mathbf{v} = \mathbf{0},$$

ovvero $c_1 = 0$. Moltiplicando, quindi, per Z_{i2}^{p-2}, si perviene a concludere che $c_2 = 0$, ecc. $\square$

Definizione 5.5 *Ricordando che* $\mathbf{v}_i = Z_{i1}\mathbf{v} \neq \mathbf{0}$, *i vettori*

$$\mathbf{x}_{ij} \equiv Z_{i2}^{j-1}\mathbf{v}_i, \qquad i = 1, \ldots, p, \tag{5.35}$$

costituiscono una catena di Jordan *che parte da* $\mathbf{v}_i \equiv \mathbf{x}_{i1}$.

Osservazione 5.7 *Osserviamo che, dalle (5.25) e (5.35), segue che*

$$\mathbf{x}_{i,j+1} = Z_{i2}\mathbf{x}_{ij} = A\mathbf{x}_{ij} - \lambda_i\mathbf{x}_{ij}, \quad i = 1, \ldots, p-1, \qquad e \qquad Z_{i2}\mathbf{x}_{ip} = \mathbf{0},$$

ovvero:

$$A\mathbf{x}_{ij} = \lambda_i\mathbf{x}_{ij} + \mathbf{x}_{i,j+1}, \quad j = 1, \ldots, p-1, \qquad e \qquad A\mathbf{x}_{ip} = \lambda_i\mathbf{x}_{ip}.$$

Per questo motivo, la catena di Jordan (5.35) sarà costituita da $p-1$ autovettori generalizzati, $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{i,p-1}$, *e da un autovettore*, $\mathbf{x}_{ip}$ *(tra loro linearmente indipendenti). Evidentemente, dal Teorema 5.7, segue che* $p \le m_i \le \bar{m}_i$.

Più in generale, ricordando che gli elementi della catena di Jordan (5.35) sono vettori linearmente indipendenti di $\mathcal{M}_i$, la cui dimensione è $\bar{m}_i \ge p$, è possibile dimostrare il seguente risultato.[4]

Teorema 5.16 *Se* λ_i *è un autovalore di* A *con molteplicità algebrica* $\bar{m}_i$ *e molteplicità nel polinomio minimo* m_i, *allora* $\exists k_i \ge 1$ *tale che:*

- $\mathbf{x}_{i1}^{(k)} \in \mathcal{M}_i$ *origina una catena di Jordan di lunghezza* p_k, $k = 1, \ldots, k_i$.
- $m_i = p_1 \ge p_2 \ge \ldots \ge p_{k_i} \ge 1$;
- $\sum_{k=1}^{k_i} p_k = \bar{m}_i$;
- *gli* $\bar{m}_i$ *vettori*

$$\left\{\mathbf{x}_{i1}^{(1)}, \ldots, \mathbf{x}_{ip_1}^{(1)}, \ldots, \mathbf{x}_{i1}^{(k_i)}, \ldots, \mathbf{x}_{ip_{k_i}}^{(k_i)}\right\} \tag{5.36}$$

  sono linearmente indipendenti. In particolare, $\left\{\mathbf{x}_{ip_1}^{(1)}, \ldots, \mathbf{x}_{ip_{k_i}}^{(k_i)}\right\}$ *sono* k_i *autovettori di* A *linearmente indipendenti relativi all'autovalore* λ_i*; gli altri vettori sono autovettori generalizzati.*

Definizione 5.6 *Il numero di autovettori linearmente indipendenti relativi ad un autovalore (e.g.,* k_i *nel precedente teorema) è detto* molteplicità geometrica *dell'autovalore.*

Definendo la matrice $X_i \in \mathbb{C}^{\bar{m} \times \bar{m}_i}$, le cui colonne sono i vettori in (5.36), nell'ordine specificato, si verifica facilmente che

$$AX_i = X_i J_i,$$

con

$$J_i = \begin{pmatrix} J_i^{(1)} & & \\ & \ddots & \\ & & J_i^{(k_i)} \end{pmatrix} \in \mathbb{C}^{\bar{m}_i \times \bar{m}_i},$$

[4]Si veda, ad esempio: *P. Lancaster, M. Tismenetsky. The Theory of Matrices, 2nd Ed. Academic Press, San Diego, 1985.*

e

$$J_i^{(k)} = \begin{pmatrix} \lambda_i & & & \\ 1 & \lambda_i & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_i \end{pmatrix} \in \mathbb{C}^{p_k \times p_k}, \qquad k = 1, \dots, k_i.$$

Questi ultimi sono detti *miniblocchi di Jordan* relativi all'autovalore λ_i. Evidentemente, la molteplicità geometrica dell'autovalore coincide con il loro numero.

Similmente, definendo la matrice

$$X = \begin{pmatrix} X_1 & \dots & X_\nu \end{pmatrix} \in \mathbb{C}^{\bar{m} \times \bar{m}},$$

si ha che questa matrice è nonsingolare e, inoltre,

$$AX = XJ \equiv \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\nu \end{pmatrix}, \tag{5.37}$$

che definisce la *forma canonica di Jordan* di A. Evidentemente, la (5.37) può essere riscritta come

$$X^{-1}AX = J. \tag{5.38}$$

Alcune osservazioni:

- se $\bar{m}_i = 1$, cioè l'autovalore è semplice, allora anche $m_i = k_i = 1$;
- se $m_i = 1 < \bar{m}_i$, cioè l'autovalore è semisemplice ma degenere, allora $J_i = \lambda_i I_{\bar{m}_i}$, ovvero $k_i = \bar{m}_i$. In altri termini, le molteplicità geometrica ed algebrica dell'autovalore coincidono;
- per contro, se $m_i = \bar{m}_i$, allora $k_i = 1$.

La matrice al primo membro della (5.38) si dice *trasformata per similitudine* di A. Le matrici ai due membri della (5.38) si dicono *simili.*

Osservazione 5.8 *È facile verificare che due matrici simili hanno lo stesso polinomio caratteristico e polinomio minimale. Inoltre, poiché $(X^{-1}AX)^i = X^{-1}A^iX$, segue facilmente che le matrici componenti di due matrici simili sono legate dalla stessa trasformazione di similitudine. Pertanto, se $f(z)$ è definita sullo spettro di A e T è una matrice nonsingolare avente la stessa dimensione, si ottiene:*

$$T^{-1}f(A)T = f\left(T^{-1}AT\right).$$

Definizione 5.7 *Una matrice si dirà* diagonalizzabile *se la matrice J in (5.37) è diagonale.*

Vale, evidentemente, il seguente risultato.

Teorema 5.17 *Una matrice è diagonalizzabile se e solo se tutti i suoi autovalori sono semplici o semisemplici.*

Esercizio 5.2 *Verificare che, data la matrice di Frobenius (5.29), se il polinomio caratteristico (5.30) ha s radici distinte $\{\lambda_1, \dots, \lambda_s\}$, allora essa è diagonalizzata dalla matrice di Vandermonde*

$$X = \begin{pmatrix} \lambda_1^0 & \dots & \lambda_s^0 \\ \vdots & & \vdots \\ \lambda_1^{s-1} & \dots & \lambda_s^{s-1} \end{pmatrix}.$$