

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

---

---

SCUOLA DI SCIENZE  
Corso di Laurea in Matematica

Tesi di Laurea in Analisi Numerica

**Equazioni  
Matriciali  
Lineari**

**Relatrice:**  
**Chiar.ma Prof.ssa**  
**Valeria Simoncini**

**Presentata da:**  
**Francesco Mascari**

---

---

**Sessione Unica**  
**Anno Accademico 2017/2018**



# Indice

<b>Introduzione</b>	<b>3</b>
<b>Notazioni</b>	<b>5</b>
<b>1 Equazioni Matriciali Lineari</b>	<b>9</b>
1.1 Alcune operazioni matriciali e vettoriali . . . . .	9
1.2 Introduzione al problema . . . . .	14
1.3 Equazione matriciale lineare polinomiale . . . . .	15
1.4 Equazione di Sylvester . . . . .	16
<b>2 Analisi Spettrale</b>	<b>23</b>
2.1 Premesse teoriche . . . . .	23
2.1.1 Polinomio minimo . . . . .	23
2.1.2 Polinomio interpolante . . . . .	25
2.2 Funzione di matrice . . . . .	28
2.3 Matrici componente . . . . .	32
<b>3 Teorema di Krein</b>	<b>39</b>
3.1 Il Teorema di Cauchy per matrici . . . . .	39
3.2 Proiezioni . . . . .	42
3.3 Teorema di Krein . . . . .	46
3.4 Teorema di Krein per l'Equazione di Sylvester . . . . .	49
<b>4 Il Metodo dei Gradienti Coniugati (CG) per Equazioni Matriciali Lineari</b>	<b>53</b>
4.1 L'algoritmo . . . . .	53
4.2 Condizione sufficiente per applicare CG . . . . .	55
4.3 Un esempio . . . . .	56

4.3.1	Applicabilità di CG . . . . .	57
4.3.2	Risultati . . . . .	58
	<b>Conclusioni</b>	<b>69</b>

# Introduzione

Questa tesi si prefigge lo scopo di indagare le equazioni matriciali lineari e la loro risoluzione.

Nel Capitolo 1 si introducono alcune operazioni matriciali, in particolare l'operatore di Kronecker con alcune sue proprietà. Successivamente viene presentato il problema generale, la sua formulazione vettoriale e le condizioni di esistenza e unicità della soluzione. Analogamente si opera per la famiglia delle equazioni matriciali lineari polinomiali, di cui l'Equazione di Sylvester è un caso particolare.

La caratterizzazione della soluzione di un'equazione matriciale lineare polinomiale in termini di analisi complessa attraverso il Teorema di Krein viene affrontata nel Capitolo 3, dopo lo sviluppo della teoria spettrale necessaria nel Capitolo 2.

Nel Capitolo 4, si introduce il Metodo dei Gradienti Coniugati (CG), che si applica alla formulazione vettoriale di un'equazione matriciale lineare per ottenere una sua soluzione approssimata. Successivamente, si applica tale metodo a un set di dati di diverse dimensioni e se ne discutono i risultati.



# Notazioni

Introduciamo alcune notazioni che utilizzeremo nel corso della trattazione.

## Notazioni insiemistiche

Dati due insiemi  $A$  e  $B$  scriviamo:

- $A \subseteq B$  (rispettivamente  $A \supseteq B$ ) se  $x \in A \Rightarrow x \in B$  (rispettivamente  $x \in B \Rightarrow x \in A$ );
- $A = B$  se  $A \subseteq B$  e  $B \subseteq A$ ;
- $A \subset B$  (rispettivamente  $A \supset B$ ) se  $A \subseteq B$  e  $A \neq B$  (rispettivamente  $A \supseteq B$  e  $A \neq B$ ).

## Notazioni matriciali

Data una matrice  $A \in \mathbb{C}^{m \times n}$  denotiamo con:

- $a_{i,j}$  l'elemento sulla  $i$ -esima riga e sulla  $j$ -esima colonna;
- $A_{i,j}$  il blocco di posto  $i, j$  nel caso la matrice  $A$  sia a blocchi;
- $A_{i,*}$  l' $i$ -esima riga;
- $A_{*,j}$  la  $j$ -esima colonna;
- $R : \mathbb{C}^{m \times n} \rightarrow \mathbb{N}$  la funzione rango, che a ogni matrice associa la dimensione massima dei minori non nulli;
- $\text{Ker}(A) = \{x \in \mathbb{C}^n \text{ tale che } Ax = 0\}$  il nucleo (o *kernel*) di  $A$ ;

- $Im(A) = \{y \in \mathbb{C}^m \text{ tale che } \exists x \in \mathbb{C}^n, Ax = y\}$  l'immagine (o *range*) di  $A$ ;
- $\sigma(A) = \{\lambda \in \mathbb{C} \text{ tale che } \exists x \in \mathbb{C}^n \setminus \{0\} Ax = \lambda x\}$  lo spettro, l'insieme degli autovalori della matrice  $A$ , nel caso sia quadrata ( $m = n$ );
- $A^T$  (rispettivamente  $A^*$ ) la matrice trasposta (rispettivamente trasposta coniugata) di  $A$ .

Definiamo, inoltre:

- La norma  $p$ -esima di  $A$ :

$$\|A\|_p := \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{\|x\|_p=1} \|Ax\|_p$$

- La norma di Frobenius di  $A$ :

$$\|A\|_F := \sqrt{\sum_{i,j} |a_{i,j}|^2}$$

- Il numero di condizionamento per  $A \in \mathbb{C}^{n \times n}$  non singolare:

$$\kappa_\gamma(A) := \|A\|_\gamma \|A^{-1}\|_\gamma$$

per una norma matriciale  $\|\cdot\|_\gamma$ .

- La norma energia di un vettore  $x \in \mathbb{C}^n$  rispetto a una matrice  $A \in \mathbb{C}^{n \times n}$  simmetrica definita positiva:

$$\|x\|_A := \sqrt{x^T A x}$$

Scriviamo, infine,  $I_n$  per la matrice identità  $n \times n$  ed  $e_i \in \mathbb{C}^n$  per la sua  $i$ -esima colonna.

## Notazioni complesse

Dato un numero complesso  $\alpha \in \mathbb{C}$  denotiamo con:

- $\Re(\alpha)$  la parte reale di  $\alpha$



- $\Im(\alpha)$  la parte immaginaria di  $\alpha$



# Capitolo 1

## Equazioni Matriciali Lineari

In questo capitolo trattiamo le equazioni matriciali lineari. Dopo aver introdotto alcune operazioni matriciali e vettoriali utili per la trattazione, quindi definiamo un'equazione matriciale lineare e alcune condizioni per l'esistenza e l'unicità della soluzione. Faremo particolare riferimento alle equazioni matriciali lineari polinomiali e all'Equazione di Sylvester, trovando specifici criteri per la loro risolubilità

Il materiale qui contenuto è tratto da [L] e [LM].

### 1.1 Alcune operazioni matriciali e vettoriali

Introduciamo alcune operazioni non banali tra matrici e tra vettori con alcune conseguenti proprietà.

**Definizione 1.1** (Impilamento tra matrici). L'impilamento tra due matrici  $A \in \mathbb{C}^{m \times n}$  e  $B \in \mathbb{C}^{p \times q}$  di opportune dimensioni è la matrice:

$$[A; B] := \begin{pmatrix} A \\ B \end{pmatrix}$$

In particolare,  $A$  e  $B$  devono avere lo stesso numero di colonne, ossia  $n = q$ .

**Definizione 1.2** (Prodotto di Kronecker). Data una matrice  $A = (a_{i,j}) \in \mathbb{C}^{m \times n}$  e una matrice  $B \in \mathbb{C}^{p \times q}$  definiamo il prodotto di Kronecker tra  $A$  e

$B$  come:

$$A \otimes B := (a_{i,j}B)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} = \begin{pmatrix} a_{1,1}B & a_{1,2}B & \cdots & a_{1,n}B \\ a_{2,1}B & a_{2,2}B & \cdots & a_{2,n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}B & a_{m,2}B & \cdots & a_{m,n}B \end{pmatrix} \in \mathbb{C}^{mp \times nq}$$

dove la seconda e la terza scrittura rappresentano una matrice a blocchi.

**Proposizione 1.1.1.** *Date le matrici quadrate  $A_1, A_2, \dots, A_p \in \mathbb{C}^{m \times m}$  e  $B_1, B_2, \dots, B_p \in \mathbb{C}^{n \times n}$ , allora:*

$$(A_1 \otimes B_1)(A_2 \otimes B_2) \dots (A_p \otimes B_p) = A_1 A_2 \dots A_p \otimes B_1 B_2 \dots B_p$$

*Dimostrazione.* Basta considerare il caso  $p = 2$  e poi estendere per induzione.

Nel caso  $p = 2$ , ponendo  $A_1 = A, A_2 = C, B_1 = B, B_2 = D$ , dobbiamo provare:

$$(A \otimes B)(C \otimes D) = AC \otimes BD$$

Abbiamo  $A \otimes B = (a_{i,j}B)_{i,j=1}^m$  e  $C \otimes D = (c_{i,j}D)_{i,j=1}^n$ .

Sia  $(A \otimes B)(C \otimes D) = F = (F_{i,j})_{i,j=1}^m$ , allora il suo blocco di posto  $i, j$  per  $1 \leq i, j \leq m$  è:

$$F_{i,j} = \sum_{k=1}^m (a_{i,k}B)(c_{k,j}D) = \sum_{k=1}^m a_{i,k}c_{k,j}BD$$

Abbiamo inoltre  $AC \otimes BD = (\gamma_{i,j}BD)_{i,j=1}^m$  dove  $\gamma_{i,j}$  è l'elemento di  $AC$  di posto  $i, j$ , ossia  $\gamma_{i,j} = \sum_{k=1}^m a_{i,k}c_{k,j}$ .

Pertanto, otteniamo:

$$\begin{aligned} (A \otimes B)(C \otimes D) &= (F_{i,j})_{i,j=1}^m = \left( \sum_{k=1}^m a_{i,k}c_{k,j}BD \right)_{i,j=1}^m \\ &= (\gamma_{i,j}BD)_{i,j=1}^m = AC \otimes BD \end{aligned}$$

□

**Corollario 1.1.2.** *Sia  $A \in \mathbb{C}^{m \times m}$  e  $B \in \mathbb{C}^{n \times n}$ , allora*

$$1. A \otimes B = (A \otimes I_n)(I_m \otimes B) = (I_m \otimes B)(A \otimes I_n)$$

$$2. (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

**Proposizione 1.1.3.** *Siano  $A \in \mathbb{C}^{m \times m}$ ,  $B \in \mathbb{C}^{n \times n}$  matrici triangolari inferiori (rispettivamente superiori), allora  $A \otimes B$  è una matrice triangolare inferiore (rispettivamente superiore).*

*Dimostrazione.* Dimostriamo la proposizione per matrici triangolari inferiori; il caso di matrici triangolari superiori si prova analogamente.

Dalla Definizione 1.2 abbiamo:

$$A \otimes B = \begin{pmatrix} a_{1,1}B & a_{1,2}B & \cdots & a_{1,m}B \\ a_{2,1}B & a_{2,2}B & \cdots & a_{2,m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}B & a_{m,2}B & \cdots & a_{m,m}B \end{pmatrix} \in \mathbb{C}^{mn \times mn}$$

I blocchi sopra la diagonale sono nulli, ossia  $a_{i,j}B = 0$  per  $i < j$ , poiché  $A$  è triangolare inferiore; blocchi sulla diagonale, ossia  $a_{i,i}B$  per  $i = 1, 2, \dots, m$ , sono anch'essi triangolari inferiori poiché  $B$  è triangolare inferiore. Pertanto, la matrice  $A \otimes B$  è globalmente triangolare inferiore. □

**Proposizione 1.1.4.** *Siano  $A \in \mathbb{C}^{m \times m}$ ,  $B \in \mathbb{C}^{n \times n}$  due matrici simmetriche, allora  $A \otimes B$  è simmetrica.*

*Dimostrazione.* Dalla Definizione 1.2 abbiamo:

$$A \otimes B = \begin{pmatrix} a_{1,1}B & a_{1,2}B & \cdots & a_{1,m}B \\ a_{2,1}B & a_{2,2}B & \cdots & a_{2,m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}B & a_{m,2}B & \cdots & a_{m,m}B \end{pmatrix} \in \mathbb{C}^{mn \times mn}$$

$A$  è simmetrica, ossia  $a_{i,j} = a_{j,i}$  per  $i, j = 1, 2, \dots, m$ , allora  $A \otimes B$  è simmetrica a blocchi. Poiché, inoltre,  $B$  è simmetrica per ipotesi,  $A \otimes B$  è globalmente simmetrica. □

Dato il polinomio in due variabili di grado  $\ell$ :

$$\phi(x; y) = \sum_{j,k=0}^{\ell} \alpha_{j,k} y^k x^j$$

con  $\alpha_{j,k} \in \mathbb{C}$ , possiamo definire il polinomio matriciale come

$$\phi(A; B) := \sum_{j,k=0}^{\ell} \alpha_{j,k} B^k \otimes A^j$$

Per matrici quadrate, il seguente teorema lega lo spettro di  $A$  e  $B$  a quello di  $\phi(A; B)$ .

**Teorema 1.1.5.** *Date le matrici  $A \in \mathbb{C}^{m \times m}$  e  $B \in \mathbb{C}^{n \times n}$ , abbiamo che lo spettro della matrice  $\phi(A; B)$  è l'insieme di tutte le coppie della forma  $\phi(\lambda; \mu) = \sum_{j,k=0}^{\ell} \alpha_{j,k} \mu^k \lambda^j$  con  $\lambda \in \sigma(A)$  e  $\mu \in \sigma(B)$*

*Dimostrazione.* Considerando la decomposizione di Jordan di  $A$  e  $B$  abbiamo:

$$PAP^{-1} = J_1 \qquad QBQ^{-1} = J_2$$

con  $P \in \mathbb{C}^{m \times m}$  e  $Q \in \mathbb{C}^{n \times n}$  matrici invertibili,  $J_1$  e  $J_2$  matrici in forma di Jordan. Chiaramente  $J_1^i$  e  $J_2^j$  sono matrici triangolari superiori con elementi  $\lambda_1^i, \lambda_2^i, \dots, \lambda_m^i$  e  $\mu_1^j, \mu_2^j, \dots, \mu_n^j$  sulla diagonale rispettivamente.

Dalla Proposizione 1.1.3 si ricava che  $J_2^j \otimes J_1^i$  è triangolare superiore e dalla Definizione 1.2 si ha che gli elementi sulla diagonale sono  $\mu_s^j \lambda_r^i$  per  $r = 1, 2, \dots, m$ ,  $s = 1, 2, \dots, n$ ; pertanto, la matrice  $\phi(J_1; J_2)$  è triangolare superiore, poiché somma di matrici triangolari superiori. Gli elementi diagonali di  $\phi(J_1; J_2)$  sono dunque  $\phi(\lambda_r; \mu_s)$  per  $r = 1, 2, \dots, m$ ,  $s = 1, 2, \dots, n$  e corrispondono agli autovalori data la struttura triangolare della matrice. Ci basta ora mostrare che  $\phi(J_1; J_2)$  e  $\phi(A; B)$  hanno gli stessi autovalori.

Dalla Proposizione 1.1.1 e dal Corollario 1.1.2 abbiamo:

$$\begin{aligned} J_2^j \otimes J_1^i &= QB^j Q^{-1} \otimes PA^i P^{-1} \\ &= (Q \otimes P)(B^j \otimes A^i)(Q^{-1} \otimes P^{-1}) \\ &= (Q \otimes P)(B^j \otimes A^i)(Q \otimes P)^{-1} \end{aligned}$$

Da cui, essendo  $\phi$  un polinomio abbiamo che:

$$\phi(J_1; J_2) = (Q \otimes P)\phi(A; B)(Q \otimes P)^{-1}$$

Pertanto,  $\phi(J_1; J_2)$  e  $\phi(A; B)$  sono simili e hanno gli stessi autovalori.  $\square$

**Corollario 1.1.6.** *Siano  $A \in \mathbb{C}^{m \times m}$ ,  $B \in \mathbb{C}^{n \times n}$ , allora valgono le seguenti proprietà:*

1.  $\forall \lambda \in \sigma(A \otimes B)$ ,  $\lambda = \lambda_A \lambda_B$  con  $\lambda_A \in \sigma(A)$  e  $\lambda_B \in \sigma(B)$
2.  $\forall \lambda \in \sigma((I_n \otimes A) + (B \otimes I_m))$ ,  $\lambda = \lambda_A + \lambda_B$  con  $\lambda_A \in \sigma(A)$  e  $\lambda_B \in \sigma(B)$

**Proposizione 1.1.7.** *Se  $A \in \mathbb{C}^{m \times m}$ ,  $B \in \mathbb{C}^{n \times n}$  sono matrici simmetriche semidefinite positive, allora  $A \otimes B \in \mathbb{C}^{mn \times mn}$  è simmetrica semidefinita positiva. Nel caso in cui  $A$  e  $B$  siano entrambe definite positive,  $A \otimes B \in \mathbb{C}^{mn \times mn}$  è definita positiva.*

*Dimostrazione.* La simmetria è ottenuta dalla Proposizione 1.1.4.

Nel caso  $A$  e  $B$  siano semidefinite positive allora dal Corollario 1.1.6 gli autovalori di  $A \otimes B$  sono della forma  $\lambda = \lambda_A \lambda_B$  con  $\lambda_A \in \sigma(A)$  e  $\lambda_B \in \sigma(B)$ ; ma  $\lambda_A \geq 0 \forall \lambda_A \in \sigma(A)$  e  $\lambda_B \geq 0 \forall \lambda_B \in \sigma(B)$  poiché  $A$  e  $B$  sono definite positive, da cui la tesi. Analogo il caso in cui  $A$  e  $B$  siano entrambe definite positive.  $\square$

**Definizione 1.3** (Vettorizzazione). Sia  $A = [A_{*,1}, A_{*,2}, \dots, A_{*,n}] \in \mathbb{C}^{m \times n}$ , la sua vettorizzazione è il vettore dato dalle colonne di  $A$  impilate una sopra l'altra, ossia:

$$\text{Vec}(A) = [A_{*,1}; A_{*,2}; \dots; A_{*,n}] = \begin{pmatrix} A_{*,1} \\ A_{*,2} \\ \vdots \\ A_{*,n} \end{pmatrix} \in \mathbb{C}^{mn \times 1}$$

Il prodotto di Kronecker e la vettorizzazione sono legati da questa importante proprietà.

**Proposizione 1.1.8.** *Siano  $A \in \mathbb{C}^{m \times m}$ ,  $B \in \mathbb{C}^{n \times n}$ ,  $X \in \mathbb{C}^{m \times n}$ , allora:*

$$\text{Vec}(AXB) = (B^T \otimes A)\text{Vec}(X)$$

*Dimostrazione.* Per  $j = 1, 2, \dots, n$ , la  $j$ -esima colonna di  $AXB$  è:

$$\begin{aligned}
 (AXB)_{*,j} &= \sum_{k=1}^n b_{k,j} (AX)_{*,k} \\
 &= \sum_{k=1}^n (b_{k,j}A) X_{*,k} \\
 &= [b_{1,j}A \quad b_{2,j}A \quad \dots \quad b_{n,j}A] \begin{pmatrix} X_{*,1} \\ X_{*,2} \\ \vdots \\ X_{*,n} \end{pmatrix} \\
 &= [b_{1,j}A \quad b_{2,j}A \quad \dots \quad b_{n,j}A] \text{Vec}(X) \\
 &= (B^T \otimes A)_{j,*} \text{Vec}(X)
 \end{aligned}$$

da cui la tesi. □

## 1.2 Introduzione al problema

**Definizione 1.4** (Equazione matriciale lineare). Un'equazione matriciale lineare con matrici coefficiente  $A_1, A_2, \dots, A_p \in \mathbb{C}^{m \times m}$  e  $B_1, B_2, \dots, B_p \in \mathbb{C}^{n \times n}$  e con matrice di termine noto  $C \in \mathbb{C}^{m \times n}$  è una scrittura formale del tipo:

$$A_1XB_1 + A_2XB_2 + \dots + A_pXB_p = C \quad (1.1)$$

con  $X \in \mathbb{C}^{m \times n}$  incognita.

**Definizione 1.5** (Soluzione di un'equazione matriciale lineare). Diciamo che  $\tilde{X} \in \mathbb{C}^{m \times n}$  è soluzione di (1.1) se soddisfa l'equazione.

**Teorema 1.2.1.** (Formulazione vettoriale di un'equazione matriciale lineare) Una scrittura alternativa dell'equazione matriciale lineare (1.1) è:

$$Gx = c \quad (1.2)$$

dove  $G = B_1^T \otimes A_1 + B_2^T \otimes A_2 + \dots + B_p^T \otimes A_p$ ,  $x = \text{Vec}(X)$  e  $c = \text{Vec}(C)$



*Dimostrazione.* Dalla Proposizione 1.1.8 abbiamo che  $Vec(A_j X B_j) = (B_j^T \otimes A_j) Vec(X)$  per  $j = 1, 2, \dots, p$ . Poiché la vettorizzazione è lineare, abbiamo:

$$Vec(C) = Vec\left(\sum_{j=1}^p A_j X B_j\right) = \sum_{j=1}^p Vec(A_j X B_j) = \sum_{j=1}^p (B_j^T \otimes A_j) Vec(X)$$

da cui la tesi.  $\square$

Utilizzando la scrittura in (1.2) possiamo utilizzare risultati di algebra lineare (vedi [S]) per determinare se l'equazione (1.1) è risolubile e se la soluzione è unica.

**Proposizione 1.2.2.** *L'equazione (1.1) è risolubile se  $c \in Im(G)$ .*

**Proposizione 1.2.3.** *L'equazione (1.1) ha un'unica soluzione se e solo  $G$  è non singolare.*

### 1.3 Equazione matriciale lineare polinomiale

Nel caso in cui  $A_i$  sia un polinomio nella matrice  $A \in \mathbb{C}^{m \times m}$  e  $B_i$  un polinomio nella matrice  $B \in \mathbb{C}^{n \times n}$  per  $i = 1, 2, \dots, p$ , se il grado massimo di  $A_i$  e  $B_i$  è  $\ell$  abbiamo che (1.1) si può scrivere:

$$\sum_{j,k=0}^{\ell} \alpha_{j,k} A^j X B^k = C \quad (1.3)$$

con  $\alpha_{j,k} \in \mathbb{C}$ .

Possiamo, allora, considerare il polinomio in due variabili:

$$\phi(x; y) = \sum_{j,k=0}^{\ell} \alpha_{j,k} y^k x^j$$

così da riscrivere l'equazione matriciale (1.3) nella forma (1.2) con

$$G = \phi(A; B^T) = \sum_{j,k=0}^{\ell} \alpha_{j,k} (B^T)^k \otimes A^j$$

Poiché gli autovalori di  $B$  e  $B^T$  sono gli stessi, possiamo applicare il Teorema

1.1.5 al polinomio  $\phi(A; B^T)$ , così da poter legare lo spettro di  $G$  a quello di  $A$  e  $B$ .

Sempre per quanto riguarda l'esistenza della soluzione per un'equazione nella scrittura (1.3), abbiamo il seguente risultato:

**Teorema 1.3.1.** *Data un'equazione matriciale come in (1.3), una condizione necessaria per l'esistenza della soluzione è:*

$\forall \lambda \in \sigma(A), \mu \in \sigma(B)$  tali che  $\phi(\lambda; \mu) = 0$ ,

$$a^T C b = 0 \quad \forall a \in \text{Ker}(A^T - \lambda I_m), b \in \text{Ker}(B - \mu I_n)$$

*Dimostrazione.* Se  $X \in \mathbb{C}^{m \times n}$  è soluzione di

$$\sum_{j,k=0}^{\ell} \alpha_{j,k} A^j X B^k = C$$

allora se moltiplichiamo a sinistra per  $a^T$  e a destra per  $b$  con  $a \in \text{Ker}(A^T - \lambda I_m), b \in \text{Ker}(B - \mu I_n)$  otteniamo:

$$\begin{aligned} \sum_{j,k=0}^{\ell} \alpha_{j,k} a^T A^j X B^k b &= a^T C b \Leftrightarrow \sum_{j,k=0}^{\ell} \alpha_{j,k} \lambda^j \mu^k a^T X b = a^T C b \\ &\Leftrightarrow a^T X b \sum_{j,k=0}^{\ell} \alpha_{j,k} \mu^k \lambda^j = a^T C b \\ &\Leftrightarrow a^T X b \underbrace{\phi(\lambda, \mu)}_{=0} = a^T C b \\ &\Leftrightarrow a^T C b = 0 \end{aligned}$$

□

## 1.4 Equazione di Sylvester

Una caso particolare di equazione matriciale lineare polinomiale è l'equazione di Sylvester.

**Definizione 1.6** (Equazione di Sylvester). Data una matrice  $A \in \mathbb{C}^{m \times m}$ ,  $B \in \mathbb{C}^{n \times n}$ ,  $C \in \mathbb{C}^{m \times n}$ , l'equazione di Sylvester è un'equazione matriciale

lineare della forma:

$$AX + XB = C \quad (1.4)$$

con  $X \in \mathbb{C}^{m \times n}$  incognita.

Anche in questo caso possiamo scrivere l'equazione nella forma (1.2); l'equazione sarà della forma  $Gx = c$  con  $x = \text{Vec}(X)$ ,  $c = \text{Vec}(C)$  e

$$G = (I_n \otimes A) + (B^T \otimes I_m) \quad (1.5)$$

Sull'esistenza e unicità della soluzione dell'equazione di Sylvester abbiamo i seguenti teoremi.

**Teorema 1.4.1.** *L'equazione (1.4) ha un'unica soluzione se e solo se le matrici  $A$  e  $-B$  non hanno nessun autovalore in comune.*

*Dimostrazione.* Gli autovalori di  $G$  in (1.5) sono per il Corollario 1.1.6  $\lambda_A + \lambda_B \forall \lambda_A \in \sigma(A), \lambda_B \in \sigma(B)$ . Dalla Proposizione 1.2.3 l'equazione ha un'unica soluzione se  $G$  è non singolare, ossia se  $G$  non ha autovalori nulli. Pertanto, data la caratterizzazione degli autovalori di  $G$ , deve essere:

$$\begin{aligned} \lambda_A + \lambda_B \neq 0 \forall \lambda_A \in \sigma(A), \lambda_B \in \sigma(B) &\Leftrightarrow \lambda_A \neq -\lambda_B \forall \lambda_A \in \sigma(A), \lambda_B \in \sigma(B) \\ &\Leftrightarrow \lambda_A \neq \hat{\lambda}_B \forall \lambda_A \in \sigma(A), \hat{\lambda}_B \in \sigma(-B) \end{aligned}$$

□

**Teorema 1.4.2.** *Se l'equazione (1.4) è risolubile,*

- *o ha un'unica soluzione data dal Teorema 1.4.1, ossia nel caso  $\sigma(A) \cap \sigma(-B) = \emptyset$ ,*
- *o ha un'infinità di soluzioni date dalla formula:*

$$X = X_0 + \tilde{X}$$

*dove  $X_0$  è una soluzione particolare di (1.4), mentre  $\tilde{X}$  è una soluzione generale dell'equazione omogenea associata,  $AX + XB = 0$ .*

*Dimostrazione.* Se consideriamo la scrittura (1.2) abbiamo dalla teoria sui sistemi lineari (vedi [S]) che la soluzione di  $Gx = c$  sarà data da:

$$x = x_0 + \tilde{x}$$

dove  $x_0$  è una soluzione particolare, mentre  $\tilde{x}$  è una soluzione generale dell'equazione omogenea associata.

Pertanto, se prendiamo  $X_0$  e  $\tilde{X}$  tali che  $\text{Vec}(X_0) = x_0$  e  $\text{Vec}(\tilde{X}) = \tilde{x}$  abbiamo la tesi.  $\square$

**Teorema 1.4.3.** *L'equazione (1.4) è risolubile se e solo se le matrici*

$$\begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} \quad \begin{pmatrix} A & C \\ 0 & -B \end{pmatrix}$$

*sono simili.*

*Dimostrazione.* Se  $X \in \mathbb{C}^{m \times n}$  è soluzione di (1.4), allora la matrice di permutazione è

$$P = \begin{pmatrix} I_m & X \\ 0 & -I_n \end{pmatrix}$$

Infatti, abbiamo

$$\begin{aligned} P \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} &= \begin{pmatrix} A & C \\ 0 & -B \end{pmatrix} P; \\ \begin{pmatrix} I_m & X \\ 0 & -I_n \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} &= \begin{pmatrix} A & C \\ 0 & -B \end{pmatrix} \begin{pmatrix} I_m & X \\ 0 & -I_n \end{pmatrix}; \\ \begin{pmatrix} A & -XB \\ 0 & B \end{pmatrix} &= \begin{pmatrix} A & AX - C \\ 0 & B \end{pmatrix}; \end{aligned}$$

Le due matrici sono uguali se  $-XB = AX - C$ , ossia  $AX + XB = C$ , che è vero poiché  $X$  è soluzione di (1.4) per ipotesi.

Viceversa, se esiste una matrice invertibile  $P \in \mathbb{C}^{(m+n) \times (m+n)}$  tale che:

$$\begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} = P^{-1} \begin{pmatrix} A & C \\ 0 & -B \end{pmatrix} P \quad (1.6)$$

Possiamo definire le seguenti trasformazioni lineari su  $\mathbb{C}^{(m+n) \times (m+n)}$

$$T_1(Z) = \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} Z - Z \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix},$$

$$T_2(Z) = \begin{pmatrix} A & C \\ 0 & -B \end{pmatrix} Z - Z \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix},$$

Pertanto se troviamo una matrice in  $Ker(T_2)$  della forma

$$\begin{pmatrix} R & U \\ 0 & -I_n \end{pmatrix}$$

allora  $U$  è soluzione di (1.4). Infatti abbiamo:

$$\begin{aligned} \begin{pmatrix} A & C \\ 0 & -B \end{pmatrix} \begin{pmatrix} R & U \\ 0 & -I_n \end{pmatrix} - \begin{pmatrix} R & U \\ 0 & -I_n \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} &= 0; \\ \begin{pmatrix} AR & AU - C \\ 0 & B \end{pmatrix} - \begin{pmatrix} RA & -UB \\ 0 & B \end{pmatrix} &= 0; \\ \begin{pmatrix} AR - RA & AU + UB - C \\ 0 & 0 \end{pmatrix} &= 0; \end{aligned}$$

Da cui deve essere  $AU + UB - C = 0$ , ossia  $AU + UB = C$ .

La relazione (1.6) ci dà

$$T_1(Z) = P^{-1}T_2(PZ)$$

da cui

$$Ker(T_2) = \{PZ, Z \in \mathbb{C}^{(m+n) \times (m+n)}, Z \in Ker(T_1)\}$$

pertanto  $dim(Ker(T_1)) = dim(Ker(T_2))$ , poiché  $P$  è non singolare.

Sia

$$\mathcal{S} = \{(V, W), V \in \mathbb{C}^{n \times m}, W \in \mathbb{C}^{n \times n} \text{ tali che } BV + VA = 0, BW = WB\}$$

questo è banalmente un spazio lineare.

Possiamo considerare le trasformazioni lineari  $\varphi_i : Ker(T_i) \rightarrow \mathcal{S}$  per  $i = 1, 2$  così definite:

$$\varphi_i \left( \begin{pmatrix} R & U \\ V & W \end{pmatrix} \right) = (V, W) \quad \text{per } i = 1, 2$$

Abbiamo che:

- $Ker(\varphi_1) = Ker(\varphi_2) = \left\{ \begin{pmatrix} R & U \\ 0 & 0 \end{pmatrix}, AR = RA, AU + UB = 0 \right\}$  poiché, per  $i = 1, 2$

$$\begin{aligned} \begin{pmatrix} R & U \\ V & W \end{pmatrix} \in Ker(\varphi_i) &\Leftrightarrow \varphi_i\left(\begin{pmatrix} R & U \\ V & W \end{pmatrix}\right) = (V, W) = 0 \\ &\Leftrightarrow V = 0, W = 0 \end{aligned}$$

Inoltre,

$$\begin{aligned} \begin{pmatrix} R & U \\ 0 & 0 \end{pmatrix} \in Ker(T_2) &\Leftrightarrow \begin{pmatrix} A & C \\ 0 & -B \end{pmatrix} \begin{pmatrix} R & U \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} R & U \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} = 0 \\ &\Leftrightarrow \begin{pmatrix} AR & AU \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} RA & -UB \\ 0 & 0 \end{pmatrix} = 0 \\ &\Leftrightarrow \begin{pmatrix} AR - RA & AU + UB \\ 0 & 0 \end{pmatrix} = 0 \\ &\Leftrightarrow AR = RA, AU + UB = 0 \end{aligned}$$

Analogamente per  $Ker(T_1)$ .

- $Im(\varphi_1) = \mathcal{S}$ . Infatti:  
 $Im(\varphi_1) \subseteq \mathcal{S}$  è vero poiché abbiamo che:

$$\begin{aligned} \begin{pmatrix} R & U \\ V & W \end{pmatrix} \in Ker(T_1) &\Leftrightarrow \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} \begin{pmatrix} R & U \\ V & W \end{pmatrix} - \begin{pmatrix} R & U \\ V & W \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -B \end{pmatrix} = 0 \\ &\Leftrightarrow \begin{pmatrix} AR & AU \\ -BV & -BW \end{pmatrix} - \begin{pmatrix} RA & -UB \\ VA & -WB \end{pmatrix} = 0 \\ &\Leftrightarrow \begin{pmatrix} AR - RA & AU + UB \\ -BV - VA & -BW + WB \end{pmatrix} = 0 \\ &\Rightarrow BV + VA = 0, WB = BW \\ &\Rightarrow (V, W) \in \mathcal{S} \end{aligned}$$

da cui  $Im(\varphi_1) \subseteq \mathcal{S}$ .

$Im(\varphi_1) \supseteq \mathcal{S}$  poiché dati  $(V, W) \in \mathcal{S}$  abbiamo che  $BV + VA =$

$0, WB = BW$ , da cui

$$\begin{pmatrix} 0 & 0 \\ V & W \end{pmatrix} \in \text{Ker}(T_1)$$

da cui  $\text{Im}(\varphi_1) \supseteq \mathcal{S}$ .

Poiché vale che:

$$\dim(\text{Ker}(\varphi_1)) + \dim(\text{Im}(\varphi_1)) = \dim(\text{Ker}(T_1))$$

con  $\text{Ker}(\varphi_1) = \text{Ker}(\varphi_2)$  e  $\dim(\text{Ker}(T_1)) = \dim(\text{Ker}(T_2))$ ,

abbiamo  $\dim(\text{Im}(\varphi_1)) = \dim(\text{Im}(\varphi_2))$ .

Da questa uguaglianza e dal fatto che  $\text{Im}(\varphi_2) \subseteq \mathcal{S} = \text{Im}(\varphi_1)$ , otteniamo:

$$\text{Im}(\varphi_1) = \text{Im}(\varphi_2) = \mathcal{S}$$

Pertanto basta trovare una matrice in  $\text{Ker}(T_1)$  la cui immagine attraverso  $\varphi_1$  sia  $(0, -I_n)$ , ad esempio:

$$\begin{pmatrix} I_m & 0 \\ 0 & -I_n \end{pmatrix}$$

Allora poiché  $\text{Im}(\varphi_1) = \text{Im}(\varphi_2)$  esisterà una matrice in  $\text{Ker}(T_2)$  la cui immagine attraverso  $\varphi_2$  è  $(0, -I_n)$ , ossia una matrice del tipo:

$$\begin{pmatrix} R_0 & U_0 \\ 0 & -I_n \end{pmatrix}$$

che ha proprio la struttura desiderata. □





## Capitolo 2

# Analisi Spettrale

In questo capitolo analizziamo come estendere l'applicazione di funzioni complesse a valori complessi al caso matriciale. Introduciamo la nozione di polinomio minimo di una matrice e le sue proprietà. Successivamente, riprendiamo la definizione di polinomio interpolante; tale concetto ci servirà per definire la funzione di una matrice attraverso l'interpolazione sui valori di una funzione e delle sue derivate sullo spettro. Infine, vediamo una scomposizione della funzione di una matrice in matrici componente con alcune loro proprietà.

Il materiale qui contenuto è tratto da [LM].

### 2.1 Premesse teoriche

#### 2.1.1 Polinomio minimo

**Definizione 2.1** (Polinomio minimo). Data una matrice  $A \in \mathbb{C}^{n \times n}$ , il suo polinomio minimo è il polinomio monico in  $\mathbb{C}$  non nullo di grado minimo tale che  $p_A(A) = 0$

**Definizione 2.2** (Indice di un autovalore). Data una matrice  $A \in \mathbb{C}^{n \times n}$  con spettro  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$ , per un autovalore  $\lambda_k$  definiamo il suo indice come:

$$m_k = \min\{m \in \mathbb{N} \text{ tale che } \text{Ker}(I\lambda_k - A)^m = \text{Ker}(I\lambda_k - A)^{m+1}\}$$

che corrisponde alla dimensione massima dei blocchi di Jordan associati a  $\lambda_k$ .

**Teorema 2.1.1** (Caratterizzazione del polinomio minimo). *Data una matrice  $A \in \mathbb{C}^{n \times n}$  il cui spettro è  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$ , allora*

$$p_A(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{m_k}$$

dove  $m_k$  è l'indice di  $\lambda_k$  per  $k = 1, 2, \dots, s$ .

Da cui otteniamo:

$$\deg(p_A(\lambda)) = m \quad \text{con } m = \sum_{k=1}^s m_k$$

*Dimostrazione.* Vedi [LM] pp.224-236. □

**Corollario 2.1.2.** *Data una matrice  $A \in \mathbb{C}^{n \times n}$  il cui spettro è  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  e ogni autovalore  $\lambda_k$  ha indice  $m_k$  per  $k = 1, 2, \dots, s$ ; sia  $p(\lambda)$  un polinomio tale che  $p(A) = 0$ , allora  $p_A(\lambda)$  divide  $p(\lambda)$ .*

*Dimostrazione.* Dividendo  $p(\lambda)$  per  $p_A(\lambda)$  abbiamo:

$$p(\lambda) = p_A(\lambda)q(\lambda) + r(\lambda)$$

dove  $\deg(r(\lambda)) < \deg(p_A(\lambda))$ . Calcolando questa espressione in  $A$  otteniamo:

$$\begin{aligned} p(A) &= p_A(A)q(A) + r(A); \\ 0 &= 0 + r(A) \end{aligned}$$

Ossia  $r(A) = 0$ , pertanto deve essere  $r(\lambda) = 0$  poiché se così non fosse  $r(\lambda)$  sarebbe un polinomio non nullo che annulla  $A$  di grado minore di  $p_A(\lambda)$ , andando contro la definizione di polinomio minimo. □

**Corollario 2.1.3.** *Data una matrice  $A \in \mathbb{C}^{n \times n}$  il cui spettro è  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  e ogni autovalore  $\lambda_k$  ha indice  $m_k$  per  $k = 1, 2, \dots, s$ , allora:  $\forall p(\lambda) \in \mathbb{P}^{m-1}(\mathbb{C})$ , spazio dei polinomi di grado strettamente minore di  $m$ ,*

$$p(A) = 0 \Leftrightarrow p(\lambda) = 0 \quad (\text{è il polinomio nullo})$$

**Corollario 2.1.4.** *Siano  $A, B \in \mathbb{C}^{n \times n}$  due matrici simili, ossia  $B = PAP^{-1}$  per una matrice invertibile  $P \in \mathbb{C}^{n \times n}$ , allora  $p_A(\lambda) = p_B(\lambda)$ .*

*Dimostrazione.* Segue dal fatto che  $A$  e  $B$  hanno la stessa decomposizione di Jordan. Infatti, sia  $A = DJD^{-1}$  con  $J$  matrice a blocchi di Jordan, allora:

$$B = PDJD^{-1}P^{-1} = PDJ(PD)^{-1}$$

□

### 2.1.2 Polinomio interpolante

Diamo ora una definizione di polinomio interpolante con alcune proprietà che useremo nelle prossime sezioni per interpolare una funzione sullo spettro di una matrice.

**Teorema 2.1.5.** *Dati  $\lambda_1, \lambda_2, \dots, \lambda_s \in \mathbb{C}$  distinti,  $m_1, m_2, \dots, m_s \in \mathbb{N}$  e un insieme di valori complessi  $\{f_{k,j} \in \mathbb{C}, k = 1, 2, \dots, s, j = 0, 1, \dots, m_k - 1\}$ , esiste un unico polinomio  $p(\lambda)$  di grado minore o uguale a  $m = \sum_{k=1}^s m_k$  tale che*

$$p^{(j)}(\lambda_k) = f_{k,j} \quad k = 1, 2, \dots, s \quad j = 0, 1, \dots, m_k - 1 \quad (2.1)$$

*Dimostrazione.* Definiamo  $p_k(\lambda) = \alpha_k(\lambda)\psi_k(\lambda)$  per  $k = 1, 2, \dots, s$  con:

$$\alpha_k(\lambda) = \alpha_{k,0} + \alpha_{k,1}(\lambda - \lambda_k) + \dots + \alpha_{k,m_k-1}(\lambda - \lambda_k)^{m_k-1}$$

con  $\alpha_{k,0}, \alpha_{k,1}, \dots, \alpha_{k,m_k-1}$  arbitrari

$$\psi_k(\lambda) = \prod_{\substack{j=1 \\ j \neq k}}^s (\lambda - \lambda_j)^{m_j}$$

ha grado minore di  $m$  e soddisfa:

$$p_k(\lambda_i) = p'_k(\lambda_i) = \dots = p_k^{(m_i-1)}(\lambda_i) = 0 \quad \text{per } k \neq i \quad (2.2)$$

Derivando  $\alpha_k$   $j$  volte e valutandolo in  $\lambda_k$  abbiamo:

$$\alpha_k^{(j)}(\lambda_k) = j! \alpha_{k,j}$$

Se definiamo  $p(\lambda) := \sum_{k=1}^s p_k(\lambda)$ , derivando  $p(\lambda)$   $j$  volte e valutandolo in  $\lambda_k$  per  $k = 1, 2, \dots, s$ ,  $j = 0, 1, \dots, m_k - 1$  abbiamo:

$$\begin{aligned} p^{(j)}(\lambda_k) &= p_k^{(j)}(\lambda_k) && \text{per (2.2)} \\ &= \sum_{i=0}^j \binom{j}{i} \alpha_k^{(i)}(\lambda_k) \psi_k^{(j-i)}(\lambda_k) \\ &= \sum_{i=0}^j \binom{j}{i} i! \alpha_{k,i} \psi_k^{(j-i)}(\lambda_k) \end{aligned}$$

pertanto, affinché valga (2.1) bisogna avere:

$$f_{k,j} = \sum_{i=0}^j \binom{j}{i} i! \alpha_{k,i} \psi_k^{(j-i)}(\lambda_k) \quad \text{per } k = 1, 2, \dots, s, j = 0, 1, \dots, m_k - 1$$

Da cui, per ogni  $k = 1, 2, \dots, s$ ,

$$\begin{aligned} j = 0 : f_{k,0} &= \alpha_{k,0} \psi_k(\lambda_k) \Leftrightarrow \alpha_{k,0} = \frac{f_{k,0}}{\psi_k(\lambda_k)} \\ j = 1 : f_{k,1} &= \alpha_{k,0} \psi_k'(\lambda_k) + \alpha_{k,1} \psi_k(\lambda_k) \Leftrightarrow \alpha_{k,1} = \frac{f_{k,1} - \alpha_{k,0} \psi_k'(\lambda_k)}{\psi_k(\lambda_k)} \\ &\vdots \end{aligned}$$

Procedendo per sostituzione, si possono trovare tutti i valori di  $a_{k,j}$  per  $k = 1, 2, \dots, s$   $j = 0, 1, \dots, m_k$  tali che  $p(\lambda)$  soddisfi (2.1). Tali coefficienti sono ben definiti poiché  $\psi_k(\lambda_k) \neq 0$  per  $k = 1, 2, \dots, s$  e sono unici; pertanto lo è anche il polinomio interpolante  $p(\lambda)$ .  $\square$

Un caso particolare è rappresentato dai polinomi in cui abbiamo un solo valore di interpolazione non nullo pari a uno, ossia  $\exists \bar{j}, \bar{k}$  tali che:

$$f_{k,j} = \begin{cases} 0 & \text{per } j \neq \bar{j} \\ 0 & \text{per } k \neq \bar{k} \\ 1 & \text{per } j = \bar{j}, k = \bar{k} \end{cases}$$

Per il Teorema 2.1.5 abbiamo che esiste un unico polinomio  $\varphi_{k,j}(\lambda)$  tale che:

$$\varphi_{k,j}^{(r)}(\lambda_i) = \begin{cases} \delta_{j,r} & \text{per } i = k, r = 0, 1, \dots, m_k - 1 \\ 0 & \text{per } i \neq k, r = 0, 1, \dots, m_k - 1 \end{cases}$$

Questi polinomi sono detti *polinomi fondamentali (o cardinali)*.

Un qualsiasi polinomio interpolante può essere ottenuto a partire dai polinomi fondamentali in tal modo:

$$p(\lambda) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f_{k,j} \varphi_{k,j}(\lambda) \quad (2.3)$$

**Proposizione 2.1.6.** *La famiglia  $\{\varphi_{k,j}(\lambda)\}_{\substack{1 \leq k \leq s \\ 0 \leq j \leq m_k - 1}}$  dei polinomi fondamentali è una famiglia linearmente indipendente; pertanto, forma una base di  $\mathbb{P}^{m-1}(\mathbb{C})$ , lo spazio dei polinomi di grado strettamente minore di  $m$ .*

*Dimostrazione.* Consideriamo una combinazione lineare di questi polinomi con coefficienti  $c_{k,j} \in \mathbb{C}$  che sia uguale a zero,

$$\sum_{k=1}^s \sum_{j=0}^{m_k-1} c_{k,j} \varphi_{k,j}(\lambda) = 0$$

se deriviamo  $r$  volte questa espressione, per un generico  $r = 0, 1, \dots, m_k - 1$ , otteniamo:

$$\sum_{k=1}^s \sum_{j=0}^{m_k-1} c_{k,j} \varphi_{k,j}^{(r)}(\lambda) = 0$$

In particolare, per  $\lambda = \lambda_i$ , abbiamo:

$$\begin{aligned} \sum_{k=1}^s \sum_{j=0}^{m_k-1} c_{k,j} \varphi_{k,j}^{(r)}(\lambda_i) = 0 &\Leftrightarrow \sum_{j=0}^{m_k-1} c_{i,j} \delta_{j,r} = 0 \\ &\Leftrightarrow c_{i,r} = 0 \end{aligned}$$

Ma questo vale per ogni  $i = 1, 2, \dots, s$  e per ogni  $r = 0, 1, \dots, m_k - 1$ ; pertanto tutti i coefficienti sono nulli e i polinomi  $\{\varphi_{k,j}(\lambda)\}_{\substack{1 \leq k \leq s \\ 0 \leq j \leq m_k - 1}}$  sono linearmente indipendenti.  $\square$

## 2.2 Funzione di matrice

**Definizione 2.3** (Funzione definita sullo spettro di una matrice). Data una funzione  $f : \mathbb{C} \rightarrow \mathbb{C}$  e una matrice  $A \in \mathbb{C}^{n \times n}$  diciamo che  $f$  è definita sullo spettro di  $A$ ,  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$ , se i valori

$$f(\lambda_k), f'(\lambda_k), \dots, f^{(m_k-1)}(\lambda_k)$$

, con  $m_k$  indice di  $\lambda_k$ , esistono per  $k = 1, 2, \dots, s$ ; tali valori sono detti valori di  $f$  sullo spettro di  $A$ .

**Definizione 2.4.** Data una matrice  $A \in \mathbb{C}^{n \times n}$  e una funzione  $f : \mathbb{C} \rightarrow \mathbb{C}$  definita su  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  diciamo che il polinomio  $p(\lambda)$  interpola  $f$  su  $\sigma(A)$  se è ottenuto attraverso l'interpolazione su

$$\{f_{k,j} = f^{(j)}(\lambda_k), k = 1, 2, \dots, s, j = 0, 1, \dots, m_k - 1\}$$

**Lemma 2.2.1.** Siano  $p_1(\lambda)$  e  $p_2(\lambda)$  due polinomi su  $\mathbb{C}$  e sia  $A \in \mathbb{C}^{n \times n}$ , allora  $p_1(A) = p_2(A)$  se e solo se  $p_1(\lambda)$  e  $p_2(\lambda)$  hanno gli stessi valori su  $\sigma(A)$ , ossia:

$$p_1^{(j)}(\lambda_k) = p_2^{(j)}(\lambda_k)$$

per  $k = 1, 2, \dots, s, j = 0, 1, \dots, m_k - 1$ .

*Dimostrazione.* Se  $p_1(A) = p_2(A)$ , definiamo  $p_0(\lambda) = p_1(\lambda) - p_2(\lambda)$  allora abbiamo  $p_0(A) = 0$ . Per il Lemma 2.1.2,  $p_A(\lambda)$  divide  $p_0(\lambda)$ , ossia esiste un polinomio  $q(\lambda)$  tale che  $p_0(\lambda) = q(\lambda)p_A(\lambda)$ . Calcolando le derivate di questo polinomio su  $\sigma(A)$  abbiamo

$$p_1^{(j)}(\lambda_k) - p_2^{(j)}(\lambda_k) = p_0^{(j)}(\lambda_k) = \sum_{i=0}^j \binom{j}{i} q^{(j-i)}(\lambda_k) p_A^{(i)}(\lambda_k) = 0$$

per  $k = 1, 2, \dots, s, j = 0, 1, \dots, m_k - 1$  poiché, per la caratterizzazione del polinomio minimo del Teorema 2.1.1,  $p_A(\lambda)$  e le sue derivate si annullano su  $\sigma(A)$ . Pertanto  $p_1(\lambda)$  e  $p_2(\lambda)$  hanno gli stessi valori su  $\sigma(A)$ .

Viceversa, se vale

$$p_1^{(j)}(\lambda_k) - p_2^{(j)}(\lambda_k) = p_0^{(j)}(\lambda_k) = 0$$

per  $k = 1, 2, \dots, s$ ,  $j = 0, 1, \dots, m_k - 1$ , allora  $p_0(\lambda)$  ha radici  $\lambda_k$  con molteplicità  $m_k$  per  $k = 1, 2, \dots, s$ . Pertanto  $p_0(\lambda)$  è divisibile per  $p_A(\lambda)$ , da cui  $p_0(A) = 0$ , ossia  $p_1(A) - p_2(A) = 0$   $\square$

**Definizione 2.5.** Data una matrice  $A \in \mathbb{C}^{n \times n}$  e una funzione  $f : \mathbb{C} \rightarrow \mathbb{C}$  definita su  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  definiamo:

$$f(A) := p(A)$$

dove  $p(\lambda)$  è un qualsiasi polinomio interpolante per  $f$  su  $\sigma(A)$ .

Il Lemma 2.2.1 ci assicura che la Definizione 2.5 sia ben definita.

Vogliamo caratterizzare l'immagine di una matrice mediante la sua forma di Jordan. I prossimi lemmi ci permettono di provare un teorema a riguardo.

**Lemma 2.2.2.** Sia  $A = \text{diag}([A_1, A_2, \dots, A_t]) \in \mathbb{C}^{n \times n}$  una matrice diagonale a blocchi e  $f : \mathbb{C} \rightarrow \mathbb{C}$  una funzione complessa definita sullo spettro di  $A$ , allora

$$f(A) = \text{diag}([f(A_1), f(A_2), \dots, f(A_t)])$$

*Dimostrazione.* La tesi vale per ogni polinomio. Infatti, essendo  $A$  a blocchi abbiamo che per  $k, j \in \mathbb{N}$  e  $\alpha, \beta \in \mathbb{C}$ :

$$\begin{aligned} \alpha A^k + \beta A^j &= \alpha \text{diag}([A_1^k, A_2^k, \dots, A_t^k]) + \beta \text{diag}([A_1^j, A_2^j, \dots, A_t^j]) \\ &= \text{diag}([\alpha A_1^k, \alpha A_2^k, \dots, \alpha A_t^k]) + \text{diag}([\beta A_1^j, \beta A_2^j, \dots, \beta A_t^j]) \\ &= \text{diag}([\alpha A_1^k + \beta A_1^j, \alpha A_2^k + \beta A_2^j, \dots, \alpha A_t^k + \beta A_t^j]) \end{aligned}$$

ossia, abbiamo  $p(A) = \text{diag}([p(A_1), p(A_2), \dots, p(A_t)])$ .

Se, quindi, consideriamo  $p(\lambda)$  il polinomio interpolante di  $f(\lambda)$  su  $\sigma(A)$  ottenuto dal Teorema 2.1.5, abbiamo  $f(A) = p(A)$ .

Poiché  $\sigma(A_j) \subseteq \sigma(A)$ ,  $f$  è definita anche su  $\sigma(A_j)$  per  $j = 1, 2, \dots, t$ . Inoltre, poiché  $p$  e  $f$  assumono gli stessi valori nello spettro di  $A$ , hanno anche gli stessi valori nello spettro di  $A_j$  per  $j = 1, 2, \dots, t$ . Pertanto abbiamo:

$$\begin{aligned} f(A) = p(A) &= \text{diag}([p(A_1), p(A_2), \dots, p(A_t)]) \\ &= \text{diag}([f(A_1), f(A_2), \dots, f(A_t)]) \end{aligned}$$

$\square$

**Lemma 2.2.3.** *Siano  $A, B \in \mathbb{C}^{n \times n}$  matrici simili, ossia tali che  $\exists P \in \mathbb{C}^{n \times n}$  non singolare per cui vale  $B = PAP^{-1}$ , sia  $f : \mathbb{C} \rightarrow \mathbb{C}$  una funzione complessa definita sullo spettro di  $A$ , allora:*

$$f(B) = Pf(A)P^{-1}$$

*Dimostrazione.* Per il Corollario 2.1.4  $A$  e  $B$  hanno lo stesso polinomio minimo, ossia hanno stessi autovalori e indici a essi associati; pertanto ogni polinomio  $p(\lambda)$  interpolante per  $f$  su  $\sigma(A)$  è anche interpolante per  $f$  su  $\sigma(B)$ , da cui

$$f(A) = p(A) \quad f(B) = p(B) \quad (2.4)$$

Inoltre, abbiamo

$$\begin{aligned} \alpha B^k + \beta B^j &= \alpha(PAP^{-1})^k + \beta(PAP^{-1})^j \\ &= \alpha PA^k P^{-1} + \beta PA^j P^{-1} \\ &= P(\alpha A^k)P^{-1} + P(\beta A^j)P^{-1} \\ &= P(\alpha A^k + \beta A^j)P^{-1} \end{aligned}$$

da cui, essendo  $p(\lambda)$  un polinomio, abbiamo  $p(B) = Pp(A)P^{-1}$ , da cui  $f(B) = Pf(A)P^{-1}$  per la (2.4).  $\square$

**Teorema 2.2.4.** *Sia  $A \in \mathbb{C}^{n \times n}$  matrice con decomposizione di Jordan  $A = PJP$ ,  $J = \text{diag}([J_1, J_2, \dots, J_t])$ ; sia  $f : \mathbb{C} \rightarrow \mathbb{C}$  una funzione complessa definita su  $\sigma(A)$ , allora:*

$$f(A) = P \text{diag}([f(J_1), f(J_2), \dots, f(J_t)])P^{-1} \quad (2.5)$$

*Inoltre, se  $J_i$  è l' $i$ -esimo blocco di Jordan, di dimensione  $\ell_i$  associato all'autovalore  $\lambda_k$ :*

$$J_i = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & \cdots & 0 & \lambda_k \end{pmatrix}$$

*Poiché  $f$  è derivabile  $m_k - 1$  volte in un intorno di  $\lambda_k$ , dove  $m_k$  è la dimensio-*



ne massima dei blocchi di Jordan associati a  $\lambda_k$ , abbiamo che  $f$  è derivabile  $\ell_i - 1$  volte in un intorno di  $\lambda_k$  e vale:

$$f(J_i) = \begin{pmatrix} f(\lambda_k) & \frac{1}{1!}f'(\lambda_k) & 0 & \cdots & \frac{1}{\ell_i-1!}f^{\ell_i-1}(\lambda_k) \\ 0 & f(\lambda_k) & \frac{1}{1!}f'(\lambda_k) & \cdots & \frac{1}{\ell_i-2!}f^{\ell_i-2}(\lambda_k) \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \ddots & \ddots & \frac{1}{1!}f'(\lambda_k) \\ 0 & 0 & \cdots & 0 & f(\lambda_k) \end{pmatrix} \quad (2.6)$$

*Dimostrazione.* Il risultato in (2.5) è facile da ottenere a partire dai lemmi precedenti:

$$\begin{aligned} A = PJP^{-1} &\Rightarrow f(A) = Pf(J)P^{-1} \text{ dal Lemma 2.2.3} \\ &\Rightarrow f(A) = P \text{diag}([f(J_1), f(J_2), \dots, f(J_t)])P^{-1} \\ &\text{dal Lemma 2.2.2} \\ &\text{poiché } J \text{ è diagonale a blocchi.} \end{aligned}$$

Proviamo ora la struttura dei blocchi di Jordan.

Il polinomio minimo di  $J_i$  è  $(\lambda - \lambda_k)^{\ell_i}$  dal Teorema 2.1.1; pertanto i valori di  $f$  sullo spettro di  $J_i$  sono  $f(\lambda_k), f'(\lambda_k), \dots, f^{\ell_i-1}(\lambda_k)$ . Dal Teorema 2.1.5 otteniamo che il polinomio interpolante per  $f$  sullo spettro di  $J_i$  è:

$$p(\lambda) = \sum_{s=0}^{\ell_i-1} \frac{1}{s!} f^{(s)}(\lambda_k) (\lambda - \lambda_k)^s$$

di conseguenza abbiamo:

$$f(J_i) = p(J_i) = \sum_{s=0}^{\ell_i-1} \frac{1}{s!} f^{(s)}(\lambda_k) (J_i - \lambda_k I_{\ell_i})^s$$

E calcolando le potenze di  $J_i - \lambda_k I_{\ell_i}$  otteniamo:

$\underbrace{\hspace{2cm}}_{s \text{ volte}}$

$$(J_i - \lambda_k I_{\ell_i})^s = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^s = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Pertanto abbiamo:

$$f(J_i) = \begin{pmatrix} f(\lambda_k) & 0 & \cdots & 0 \\ 0 & f(\lambda_k) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & f(\lambda_k) \end{pmatrix} + \cdots \\ \cdots + \begin{pmatrix} 0 & \cdots & 0 & \frac{1}{(\ell_i-1)!} f^{(\ell_i-1)}(\lambda_k) \\ 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

da cui la tesi.  $\square$

## 2.3 Matrici componente

**Teorema 2.3.1.** *Data  $A \in \mathbb{C}^{n \times n}$  con spettro  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  e una funzione  $f : \mathbb{C} \rightarrow \mathbb{C}$  definita sullo spettro di  $A$ , allora esistono matrici  $Z_{k,j} \in \mathbb{C}^{n \times n}$  indipendenti da  $f(\lambda)$  tali che*

$$f(A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j} \quad (2.7)$$

dove  $m_k$  è l'indice di  $\lambda_k$  per  $k = 1, 2, \dots, s$ .

Inoltre, le matrici  $\{Z_{k,j}\}_{\substack{1 \leq k \leq s \\ 0 \leq j \leq m_k-1}}$  sono linearmente indipendenti e commutano con  $A$  e tra loro.

*Dimostrazione.* La funzione  $f(\lambda)$  e il polinomio  $p(\lambda)$  definito da (2.3) hanno

gli stessi valori su  $\sigma(A)$ . Pertanto:

$$\begin{aligned} f(A) = p(A) &\Leftrightarrow f(A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) \varphi_{k,j}(A) \\ &\Leftrightarrow f(A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j} \quad \text{con } Z_{k,j} = \varphi_{k,j}(A) \end{aligned}$$

Segue inoltre che le  $\{Z_{k,j}\}_{\substack{1 \leq k \leq s \\ 0 \leq j \leq m_k-1}}$  commutano sia con  $A$  che tra di loro, poiché sono prodotto di polinomi in  $A$ .

Per dimostrare che queste matrici sono linearmente indipendenti consideriamo dei coefficienti  $c_{k,j} \in \mathbb{C}$  supponiamo che:

$$\sum_{k=1}^s \sum_{j=0}^{m_k-1} c_{k,j} Z_{k,j} = 0$$

e denotiamo  $h(\lambda) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} c_{k,j} \varphi_{k,j}(\lambda)$  che è un polinomio di grado minore di  $m-1$ . Poiché  $Z_{k,j} = \varphi_{k,j}(A)$  abbiamo  $h(A) = 0$ , allora  $h$  è il polinomio nullo dal Corollario 2.1.3. Pertanto,  $\sum_{k=1}^s \sum_{j=0}^{m_k-1} c_{k,j} \varphi_{k,j}(\lambda) = 0$  e siccome  $\{\varphi_{k,j}\}_{\substack{1 \leq k \leq s \\ 0 \leq j \leq m_k-1}}$  sono linearmente indipendenti dal Teorema 2.1.6,  $c_{k,j} = 0 \forall k, j$ , che implica che  $Z_{k,j}$  sono linearmente indipendenti.  $\square$

**Definizione 2.6** (Matrici componente). Data una matrice  $A \in \mathbb{C}^{n \times n}$  con spettro  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  con  $m_k$  indice di  $\lambda_k$  per  $k = 1, 2, \dots, s$ , definiamo  $Z_{k,j}$  le matrici componente di  $A$  come le matrici tali che:

$$f(A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j}$$

per ogni funzione  $f : \mathbb{C} \rightarrow \mathbb{C}$  definita sullo spettro di  $A$ .

Le seguenti proposizioni e i seguenti teoremi ci permettono di caratterizzare le matrici componente e di trovare ulteriori proprietà

**Proposizione 2.3.2.** *Data una matrice  $A \in \mathbb{C}^{n \times n}$  la cui forma di Jordan è  $A = P \text{diag}([J_1, J_2, \dots, J_t]) P^{-1}$ , allora*

$$Z_{k,j} = \frac{1}{j!} P \text{diag}([0, 0, \dots, 0, N_{1_k}^j, N_{2_k}^j, \dots, N_{q_k}^j, 0, \dots, 0]) P^{-1}$$

per  $k = 1, 2, \dots, s$  e  $j = 0, 1, \dots, m_k$  e dove  $J_{1_k}, J_{2_k}, \dots, J_{q_k}$  sono i blocchi di Jordan associati all'autovalore  $\lambda_k$ , mentre  $N_i$  per  $i = 1_k, 2_k, \dots, q_k$  è l'equivalente blocco di Jordan associato a 0.

*Dimostrazione.* Per definizione,

$$Z_{k,j} = \varphi_{k,j}(A) = P \text{diag}([\varphi_{k,j}(J_1), \varphi_{k,j}(J_2), \dots, \varphi_{k,j}(J_t)]) P^{-1}$$

con  $k = 1, 2, \dots, s$  e  $j = 0, 1, \dots, m_k - 1$ .

Per ogni blocco di Jordan  $J_i$  (di dimensione  $\ell_i$ ) abbiamo due casi:

**Caso 1**  $J_i$  è associato all'autovalore  $\lambda_k$ .

Poiché  $\varphi_{k,j}^{(r)}(\lambda_k) = \delta_{j,r}$  ( $r, j = 0, 1, \dots, m_k - 1$ ), abbiamo da (2.6):

$$\varphi_{k,j}(J_i) = \frac{1}{j!} N_i^j$$

dove  $N_i^j$  è il blocco di Jordan di dimensione  $\ell_i$  associato all'autovalore 0.

**Caso 2**  $J_i$  è associato all'autovalore  $\lambda_h \neq \lambda_k$ .

Poiché  $\varphi_{k,j}^{(r)}(\lambda_h) = 0$  ( $h \neq k, r, j = 0, 1, \dots, m_k - 1$ ), abbiamo da (2.6):

$$\varphi_{k,j}(J_i) = 0$$

Pertanto, per  $k = 1$ , se supponiamo che i blocchi di Jordan associati a  $\lambda_1$  siano i primi  $q_1$ , abbiamo:

$$Z_{1,j} = \frac{1}{j!} P \text{diag}([N_{1_1}^j, N_{2_1}^j, \dots, N_{q_1}^j, 0, 0, \dots, 0]) P^{-1} \quad (2.8)$$

per  $j = 0, 1, \dots, m_1 - 1$ .

Analogo il risultato per un  $k$  generico.  $\square$

Se in (2.8) consideriamo il caso  $j = 0$  abbiamo:

$$Z_{1,0} = P \text{diag}([I_r, 0]) P^{-1} \quad (2.9)$$

con  $r = \ell_{1_1} + \ell_{2_1} + \dots + \ell_{q_1}$

**Corollario 2.3.3** (Proprietà delle matrici componente). *Data una matrice  $A \in \mathbb{C}^{n \times n}$ , le matrici componente  $Z_{k,j}$  godono delle seguenti proprietà:*

1.  $\sum_{k=1}^s Z_{k,0} = I_n$
2.  $Z_{k,p}Z_{l,q} = 0 \ \forall l \neq k, \forall p, q$
3.  $Z_{k,j}^2 = Z_{k,j}$  se e solo se  $j = 0$
4.  $Z_{k,0}Z_{k,r} = Z_{k,r}$  per  $r = 0, 1, \dots, m_k - 1$

**Teorema 2.3.4.** *Sia  $A \in \mathbb{C}^{n \times n}$  con spettro  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  con  $m_k$  indice di  $\lambda_k$  per  $k = 1, 2, \dots, s$ , allora*

$$\text{Im}(Z_{k,0}) = \text{Ker}(A - \lambda_k I_n)^{m_k}$$

*Dimostrazione.* Studiamo il caso  $k = 1$  senza perdere di generalità.

Se consideriamo la decomposizione di Jordan di  $A$ , abbiamo  $A = PJP^{-1}$  con  $P \in \mathbb{C}^{n \times n}$  matrice invertibile e  $J = \text{diag}([J_1, J_2, \dots, J_t]) \in \mathbb{C}^{n \times n}$  matrice a blocchi di Jordan di dimensione  $\ell_i$  per  $i = 1, 2, \dots, t$ . Pertanto, otteniamo:

$$(J - \lambda_1 I_n)^{m_1} = \text{diag}([(J_1 - \lambda_1 I_{\ell_1})^{m_1}, (J_2 - \lambda_1 I_{\ell_2})^{m_1}, \dots, (J_t - \lambda_1 I_{\ell_t})^{m_1}])$$

Supponiamo che i primi  $q$  blocchi siano associati all'autovalore  $\lambda_1$ ; poiché  $m_1$  è la dimensione massima dei blocchi di Jordan associati a  $\lambda_1$ , abbiamo  $\ell_i \leq m_1$  per  $i = 1, 2, \dots, q$  e quindi  $(J_i - \lambda_1 I_{\ell_i})^{m_1} = 0$  per  $i = 1, 2, \dots, q$ . Pertanto otteniamo:

$$(J - \lambda_1 I_n)^{m_1} = \text{diag}([0, \dots, 0, (J_{q+1} - \lambda_1 I_{\ell_{q+1}})^{m_1}, \dots, (J_t - \lambda_1 I_{\ell_t})^{m_1}])$$

I blocchi  $(J_i - \lambda_1 I_{\ell_i})^{m_1}$  per  $i = q + 1, q + 2, \dots, t$  sono non singolari poiché  $J_i$  non è associato a  $\lambda_1$  per  $i = q + 1, q + 2, \dots, t$ , pertanto il nucleo di  $(J - \lambda_1 I_n)^{m_1}$  è generato dai vettori unitari  $e_1, e_2, \dots, e_r \in \mathbb{C}^n$  con  $r = \ell_1 + \ell_2 + \dots + \ell_q$ ; poniamo  $\mathcal{S} = \text{Ker}((J - \lambda_1 I_n)^{m_1})$ .

Poiché  $(A - \lambda_1 I_n)^{m_1} = P(J - \lambda_1 I_n)^{m_1}P^{-1}$  e  $P$  è non singolare,

$$\begin{aligned} x \in \text{Ker}((A - \lambda_1 I_n)^{m_1}) &\Leftrightarrow P(J - \lambda_1 I_n)^{m_1}P^{-1}x = 0 \\ &\Leftrightarrow P^{-1}x \in \text{Ker}(P(J - \lambda_1 I_n)^{m_1}) \\ &\Leftrightarrow P^{-1}x \in \text{Ker}((J - \lambda_1 I_n)^{m_1}) \end{aligned}$$

$$\begin{aligned} & \text{poiché } P \text{ non singolare} \\ & \Leftrightarrow P^{-1}x \in \mathcal{S} \end{aligned}$$

Da (2.9) si deduce che:

$$\begin{aligned} x \in \text{Im}(Z_{1,0}) & \Leftrightarrow x \in \text{Im}(P \text{diag}([I_r, 0])P^{-1}) \\ & \Leftrightarrow P^{-1}x \in \text{Im}(\text{diag}([I_r, 0])P^{-1}) \\ & \Leftrightarrow P^{-1}x \in \text{Im}(\text{diag}([I_r, 0])) \text{ poiché } P \text{ non singolare} \\ & \Leftrightarrow P^{-1}x \in \mathcal{S} \end{aligned}$$

Pertanto,  $\text{Ker}((A - \lambda_1 I_n)^{m_1}) = \text{Im}(Z_{1,0})$ . □

**Teorema 2.3.5.** Per  $k = 1, 2, \dots, s$  e  $j = 0, 1, \dots, m_k - 1$ ,

$$Z_{k,j} = \frac{1}{j!} (A - \lambda_k I_n)^j Z_{k,0} \quad (2.10)$$

*Dimostrazione.* Supponiamo  $k = 1$  senza perdere di generalità.

Il risultato è ovvio per  $j = 0$ .

Se consideriamo la decomposizione di Jordan di  $A$ , abbiamo  $A - \lambda_1 I_n = P(J - \lambda_1 I_n)P^{-1}$  con  $P \in \mathbb{C}^{n \times n}$  matrice invertibile e  $J = \text{diag}([J_1, J_2, \dots, J_t]) \in \mathbb{C}^{n \times n}$  matrice a blocchi di Jordan di dimensione  $\ell_i$  per  $i = 1, 2, \dots, t$ . Se supponiamo che i blocchi di Jordan associati all'autovettore  $\lambda_1$  siano i primi  $q$ , abbiamo che:

$$\begin{aligned} P^{-1}(A - \lambda_1 I_n)P &= J - \lambda_1 I_n \\ &= \text{diag}([J_1 - \lambda_1 I_{\ell_1}, J_2 - \lambda_1 I_{\ell_2}, \dots, J_t - \lambda_1 I_{\ell_t}]) \\ &= \text{diag}([N_1, \dots, N_q, J_{q+1} - \lambda_1 I_{\ell_{q+1}}, \dots, J_t - \lambda_1 I_{\ell_t}]) \end{aligned}$$

dove  $N_1, N_2, \dots, N_q$  sono blocchi di Jordan associati a 0 poiché  $J_1, J_2, \dots, J_q$  sono gli unici blocchi di Jordan associati a  $\lambda_1$ .

Dalla Proposizione 2.3.2 abbiamo che

$$j!P^{-1}Z_{1,j}P = \text{diag}([N_1^j, N_2^j, \dots, N_q^j, 0, 0, \dots, 0])$$

otteniamo:

$$\begin{aligned}
j!P^{-1}(A - \lambda_1 I_n)Z_{1,j}P &= \\
&= \text{diag}([N_1, \dots, N_q, J_{q+1} - \lambda_1 I_{\ell_{q+1}}, \dots, J_t - \lambda_1 I_{\ell_t}]) \\
&\quad \text{diag}([N_1^j, \dots, N_q^j, 0, \dots, 0]) \\
&= \text{diag}([N_1^{j+1}, \dots, N_q^{j+1}, 0, \dots, 0]) \\
&= (j+1)!P^{-1}Z_{1,j+1}P
\end{aligned}$$

Da questa uguaglianza, otteniamo la formula per ricorrenza:

$$Z_{1,j+1} = \frac{1}{j+1}(A - \lambda_1 I_n)Z_{1,j}$$

che ci permette di provare (2.10) per induzione. □

**Corollario 2.3.6.** *Data una matrice  $A$  in  $\mathbb{C}^{n \times n}$*

$$A = \sum_{k=1}^s \lambda_k Z_{k,0} + Z_{k,1}$$

*Dimostrazione.* Usando il Teorema 2.3.5 abbiamo:

$$\begin{aligned}
\sum_{k=1}^s \lambda_k Z_{k,0} + Z_{k,1} &= \sum_{k=1}^s \lambda_k Z_{k,0} + (A - \lambda_k I_n)Z_{k,0} \\
&= \sum_{k=1}^s A Z_{k,0} \\
&= A \underbrace{\sum_{k=1}^s Z_{k,0}}_{=I_n} = A \\
&\quad \text{per il Corollario 2.3.3}
\end{aligned}$$

□





## Capitolo 3

# Teorema di Krein

In questo capitolo, dopo aver esteso il Teorema di Cauchy alle funzioni matriciali, introduciamo le proiezioni definite da una matrice in un sottoinsieme del suo spettro, che ci permettono di generalizzare il Teorema di Cauchy a un maggior numero di contorni. Studiamo le proiezioni sia da un punto di vista analitico, come integrale complesso, sia da un punto di vista geometrico, come sommatoria di matrici componente.

Queste nozioni saranno poi applicate al Teorema di Krein per la risoluzione di un'equazione matriciale lineare polinomiale e, in particolare, per l'Equazione di Sylvester.

Il materiale qui contenuto è tratto da [L].

### 3.1 Il Teorema di Cauchy per matrici

**Definizione 3.1** (Risolvente di una matrice). Data una matrice  $A \in \mathbb{C}^{n \times n}$ , definiamo il risolvente della matrice  $A$  come:

$$R(z, A) = (zI_n - A)^{-1}$$

.

**Lemma 3.1.1** (Equazione del risolvente). *Data una matrice  $A \in \mathbb{C}^{n \times n}$ , abbiamo*

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda)R(\lambda, A)R(\mu, A)$$

*Dimostrazione.*

$$\begin{aligned}
R(\lambda, A) - R(\mu, A) &= (\lambda I_n - A)^{-1} - (\mu I_n - A)^{-1} \\
&= (\lambda I_n - A)^{-1}(\mu I_n - A)(\mu I_n - A)^{-1} + \\
&\quad - (\mu I_n - A)^{-1}(\lambda I_n - A)(\lambda I_n - A)^{-1} \\
&= [(\lambda I_n - A)^{-1}(\mu I_n - A)^{-1}][(\mu I_n - A) - (\lambda I_n - A)] \\
&= [(\lambda I_n - A)^{-1}(\mu I_n - A)^{-1}](\mu I_n - \lambda I_n) \\
&= (\mu - \lambda)R(\lambda, A)R(\mu, A)
\end{aligned}$$

□

**Lemma 3.1.2.** *Data una matrice  $A \in \mathbb{C}^{n \times n}$ , siano  $f$  e  $g$  due funzioni definite su  $\sigma(A)$ , siano inoltre  $h(z) = \alpha f(z) + \beta g(z)$  con  $\alpha, \beta \in \mathbb{C}$  e  $k(z) = f(z)g(z)$ , allora anche  $h$  e  $g$  sono definite su  $\sigma(A)$  e vale:*

$$\begin{aligned}
h(A) &= \alpha f(A) + \beta g(A) \\
k(A) &= f(A)g(A)
\end{aligned}$$

*Dimostrazione.* Le derivate  $j$ -esime di  $h$  e  $k$  sono:

$$\begin{aligned}
h^{(j)}(z) &= \alpha f^{(j)}(z) + \beta g^{(j)}(z) \\
k^{(j)}(z) &= \sum_{i=0}^j \binom{j}{i} f^{(j-i)}(z)g^{(i)}(z)
\end{aligned}$$

Pertanto  $h$  e  $k$  sono definite su  $\sigma(A)$  in quanto  $f$  e  $g$  lo sono.

Siano  $p_1$  e  $p_2$  i polinomi che interpolano  $f$  e  $g$  su  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$  rispettivamente, allora abbiamo che  $p_1^{(j)}(\lambda_k) = f^{(j)}(\lambda_k)$ ,  $p_2^{(j)}(\lambda_k) = g^{(j)}(\lambda_k)$  per  $k = 1, 2, \dots, s$ ,  $j = 0, 1, \dots, m_k - 1$  da cui abbiamo:

$$\begin{aligned}
h(\lambda) &= \alpha f(\lambda) + \beta g(\lambda) = \alpha p_1(\lambda) + \beta p_2(\lambda) && \text{su } \sigma(A) \\
k(\lambda) &= f(\lambda)g(\lambda) = p_1(\lambda)p_2(\lambda) && \text{su } \sigma(A)
\end{aligned}$$

Da cui otteniamo la tesi. □

**Lemma 3.1.3.** *Data una matrice  $A \in \mathbb{C}^{n \times n}$ , sia  $f(z) = (a - z)^{-1}$  con  $a \notin \sigma(A)$  allora  $f(A) = (aI_n - A)^{-1}$*

*Dimostrazione.* Sia  $p$  il polinomio interpolante di  $f$  su  $\sigma(A)$ , allora  $p(\lambda) = f(\lambda) = (a - \lambda)^{-1} \quad \forall \lambda \in \sigma(A)$ , tale valore è ben definito poiché  $a \notin \sigma(A)$ ; da cui  $(a - \lambda)p(\lambda) = (a - \lambda)f(\lambda) = 1$ .

Poiché la funzione  $(a - z)$  è definita su  $\sigma(A)$  abbiamo che per il Lemma 3.1.2  $(aI_n - A)p(A) = (aI_n - A)f(A) = I_n$ , cioè  $f(A) = (aI_n - A)^{-1}$ , che esiste poiché  $a \notin \sigma(A)$ .  $\square$

**Proposizione 3.1.4.** *Sia  $A \in \mathbb{C}^{n \times n}$  matrice con  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  e  $m_k$  indice di  $\lambda_k$  per  $k = 1, 2, \dots, s$ , se  $\lambda \notin \sigma(A)$ , allora*

$$R(\lambda, A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} \frac{j!}{(\lambda - \lambda_k)^{j+1}} Z_{k,j} \quad (3.1)$$

*Dimostrazione.* Consideriamo  $f(z) = (\lambda - z)^{-1}$ , la cui derivata  $j$ -esima è:

$$f^{(j)}(z) = (-1)^j (-1)^j j! (\lambda - z)^{-(j+1)} = \frac{j!}{(\lambda - z)^{(j+1)}}$$

Allora  $f$  è definita sullo spettro di  $A$  poiché  $\lambda \notin \sigma(A)$  e abbiamo che  $f(A) = R(\lambda, A)$  per il Lemma 3.1.3:

$$\begin{aligned} R(\lambda, A) &= f(A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j} \\ &= \sum_{k=1}^s \sum_{j=0}^{m_k-1} \frac{j!}{(\lambda - \lambda_k)^{j+1}} Z_{k,j} \end{aligned}$$

$\square$

**Definizione 3.2** (Contorno). Un contorno  $\Gamma$  è un'unione finita di curve semplici e differenziabili.

**Definizione 3.3** (Contorno chiuso). Un contorno il cui punto di inizio e di fine coincidono si chiama contorno chiuso.

Dato un contorno chiuso, denotiamo con:

- $\check{\Gamma}$ , l'insieme dei punti interni a  $\Gamma$ ;
- $\hat{\Gamma} = \Gamma \cup \check{\Gamma}$ , la regione di  $\mathbb{C}$  racchiusa da  $\Gamma$ .

**Teorema 3.1.5** (Teorema di Cauchy matriciale). *Data  $A \in \mathbb{C}^{n \times n}$  con  $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_s\}$  e dato un contorno  $\Gamma$  tale che  $\sigma(A) \subseteq \check{\Gamma}$ ; sia  $f$  una funzione continua su  $\hat{\Gamma}$  e olomorfa su  $\check{\Gamma}$ , allora*

$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(z) R(z, A) dz$$

*Dimostrazione.* Se moltiplichiamo entrambi i lati di (3.1) per  $\frac{f(z)}{2\pi i}$ , abbiamo:

$$\frac{f(\lambda)}{2\pi i} R(\lambda, A) = \frac{f(\lambda)}{2\pi i} \sum_{k=1}^s \sum_{j=0}^{m_k-1} \frac{j!}{(\lambda - \lambda_k)^{j+1}} Z_{k,j}$$

Poiché  $\sigma(A) \subseteq \check{\Gamma}$  abbiamo che  $R(\lambda, A)$  non ha singolarità su  $\Gamma$  e possiamo integrare lungo  $\Gamma$ :

$$\frac{1}{2\pi i} \int_{\Gamma} f(\lambda) R(\lambda, A) d\lambda = \frac{j!}{2\pi i} \sum_{k=1}^s \sum_{j=0}^{m_k-1} \int_{\Gamma} \frac{f(\lambda)}{(\lambda - \lambda_k)^{j+1}} d\lambda Z_{k,j}$$

Dall'analisi complessa (vedi [R]) abbiamo  $f^{(j)}(\lambda_k) = \frac{j!}{2\pi i} \int_{\Gamma} \frac{f(\lambda)}{(\lambda - \lambda_k)^{j+1}} d\lambda$ ; sostituendo otteniamo:

$$\frac{1}{2\pi i} \int_{\Gamma} f(\lambda) R(\lambda, A) d\lambda = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j} = f(A)$$

□

## 3.2 Proiezioni

**Definizione 3.4** (Proiezione). *Data una matrice  $A \in \mathbb{C}^{n \times n}$  e dato un contorno  $\Gamma$  tale che  $\Gamma \cap \sigma(A) = \emptyset$ , l'operatore*

$$P := \frac{1}{2\pi i} \int_{\Gamma} R(z, A) dz$$

è la proiezione definita da  $A$  e  $\sigma_1$ , dove  $\sigma_1 = \hat{\Gamma} \cap \sigma(A) = \check{\Gamma} \cap \sigma(A)$  sono i poli di  $R(z, A)$  in  $\hat{\Gamma}$ . Tale operatore è ben definito poiché  $\Gamma \cap \sigma(A) = \emptyset$ .

Si vede facilmente come per  $\sigma_1 = \sigma(A) \subseteq \check{\Gamma}$  abbiamo  $P = f(A)$  con  $f(\lambda) = 1$  funzione costante.

**Proposizione 3.2.1.** *Data una matrice  $A \in \mathbb{C}^{n \times n}$  e dato un contorno  $\Gamma$  tale che  $\Gamma \cap \sigma(A) = \emptyset$ , allora*

$$P = \sum_{k|\lambda_k \in \sigma_1} Z_{k,0}$$

con  $\sigma_1 = \hat{\Gamma} \cap \sigma(A) = \check{\Gamma} \cap \sigma(A)$

*Dimostrazione.*

$$\begin{aligned} P &= \frac{1}{2\pi i} \int_{\Gamma} R(z, A) dz \\ &= \frac{1}{2\pi i} \sum_{k=1}^s \sum_{j=0}^{m_k-1} \int_{\Gamma} \frac{j!}{(z - \lambda_k)^{j+1}} dz Z_{k,j} \\ &= \frac{1}{2\pi i} \sum_{k=1}^s \sum_{j=0}^{m_k-1} \int_{\Gamma} f_j(z) dz Z_{k,j} \quad \text{con } f_j(z) = \frac{j!}{(z - \lambda_k)^{j+1}} \end{aligned}$$

Se  $\lambda_k \notin \check{\Gamma}$ , abbiamo che  $f_j$  è olomorfa in una regione contenente  $\Gamma$ , da cui  $\int_{\Gamma} f_j(z) dz = 0$ .

Se  $\lambda_k \in \check{\Gamma}$ , abbiamo due casi:

- $j = 0 \Rightarrow \text{Res}(f_j, \lambda_k) = 1$
- $j \neq 0 \Rightarrow \text{Res}(f_j, \lambda_k) = 0$

Da cui otteniamo

$$\begin{aligned} P &= \frac{1}{2\pi i} \sum_{k|\lambda_k \in \sigma_1} \sum_{j=0}^{m_k-1} 2\pi i \text{Res}(f_j, \lambda_k) Z_{k,j} \\ &= \frac{1}{2\pi i} \sum_{k|\lambda_k \in \sigma_1} 2\pi i Z_{k,0} \\ &= \sum_{k|\lambda_k \in \sigma_1} Z_{k,0} \end{aligned}$$

□

Possiamo estendere il Teorema di Cauchy al caso in cui  $\check{\Gamma}$  non contenga tutto lo spettro di  $A$ , ma semplicemente un suo sottoinsieme  $\sigma_1$ , utilizzando le proiezioni.

**Teorema 3.2.2.** *Siano  $A \in \mathbb{C}^{n \times n}$  e  $\Gamma$  un contorno tale che  $\sigma(A) \subseteq \check{\Gamma}$ , data una proiezione  $P$  definita da  $A$  e  $\sigma_1 = \{\lambda_1, \lambda_2, \dots, \lambda_r\} \subseteq \sigma(A)$ , allora abbiamo che*

$$Pf(A) = f(A)P = \frac{1}{2\pi i} \int_{\Gamma_1} f(z)R(z, A)dz = \sum_{k=1}^r \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k)Z_{k,j}$$

per ogni  $f$  continua su  $\hat{\Gamma}$  e regolare su  $\check{\Gamma}$ , dove  $\Gamma_1$  è un contorno tale che  $\Gamma_1 \cap \sigma(A) = \emptyset$ , con  $\sigma_1 \subseteq \check{\Gamma}_1$ .

*Dimostrazione.* Dal Teorema 3.1.5 abbiamo:

$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda)R(\lambda, A)d\lambda$$

e dalla Definizione 3.4

$$P = \frac{1}{2\pi i} \int_{\Gamma_1} R(z, A)dz$$

Se moltiplichiamo  $f(A)$  a destra per  $P$  otteniamo

$$\begin{aligned} f(A)P &= \left[ \frac{1}{2\pi i} \int_{\Gamma} f(\lambda)R(\lambda, A)d\lambda \right] \left[ \frac{1}{2\pi i} \int_{\Gamma_1} R(z, A)dz \right] \\ &= -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma} f(\lambda)R(\lambda, A)R(z, A)d\lambda dz \end{aligned}$$

Poiché  $\sigma_1 \subseteq \sigma(A)$ ,  $\sigma_1 \subseteq \check{\Gamma}_1$  e  $\sigma(A) \subseteq \check{\Gamma}$ , possiamo deformare  $\Gamma_1$  in maniera tale che  $\Gamma_1 \subset \hat{\Gamma}$ , da cui otteniamo:

$$\begin{aligned} f(A)P &= \frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma} f(\lambda) \frac{R(\lambda, A) - R(z, A)}{\lambda - z} d\lambda dz \\ &= \frac{1}{4\pi^2} \int_{\Gamma} f(\lambda)R(\lambda, A) \left( \int_{\Gamma_1} \frac{1}{\lambda - z} dz \right) d\lambda + \\ &\quad - \frac{1}{4\pi^2} \int_{\Gamma_1} \left( \int_{\Gamma} \frac{f(\lambda)}{\lambda - z} d\lambda \right) R(z, A) dz \end{aligned}$$

Tuttavia, poiché  $\Gamma_1 \subset \hat{\Gamma}$ , abbiamo che:

- la funzione  $\frac{1}{\lambda - z}$  è olomorfa in  $z$  su  $\check{\Gamma}_1$ ;
- la funzione  $\frac{f(\lambda)}{\lambda - z}$  ha un polo di ordine uno per  $\lambda = z$ .

Pertanto, dall'analisi complessa (vedi [R]) abbiamo:

$$\int_{\Gamma_1} \frac{1}{\lambda - z} dz = 0 \qquad \int_{\Gamma} \frac{f(\lambda)}{\lambda - z} d\lambda = 2\pi i f(z)$$

Sostituendo otteniamo che:

$$Pf(A) = \frac{1}{2\pi i} \int_{\Gamma_1} f(z)R(z, A)dz$$

Dal Teorema 2.3.1 abbiamo:

$$f(A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j}$$

e dalla Proposizione 3.2.1 abbiamo:

$$P = \sum_{k=1}^r Z_{k,0}$$

Se moltiplichiamo  $f(A)$  a destra per  $P$  otteniamo:

$$\begin{aligned} f(A)P &= \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j} \sum_{k=1}^r Z_{k,0} \\ &= \sum_{k=1}^r \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j} \end{aligned}$$

Dove la seconda uguaglianza deriva direttamente dal Corollario 2.3.3 sulle proprietà delle matrici componente.

Questa dimostrazione può essere sviluppata analogamente moltiplicando  $f(A)$  a sinistra per  $P$ .  $\square$

**Definizione 3.5.** Data  $A \in \mathbb{C}^{n \times n}$ , un contorno  $\Gamma$  tale che  $\sigma(A) \subseteq \check{\Gamma}$  e una proiezione  $P$  definita da  $A$  e  $\sigma_1 = \{\lambda_1, \lambda_2, \dots, \lambda_r\} \subseteq \sigma(A)$ , allora definiamo la proiezione rispetto a  $P$  della funzione  $f$  su  $\sigma_1$  come:

$$f_P(A) := Pf(A) = f(A)P = \frac{1}{2\pi i} \int_{\Gamma_1} f(z)R(z, A)dz = \sum_{k=1}^r \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) Z_{k,j}$$

per ogni  $f$  continua su  $\hat{\Gamma}$  e regolare su  $\check{\Gamma}$ , dove  $\Gamma_1$  è un contorno tale che  $\Gamma_1 \cap \sigma(A) = \emptyset$ , con  $\sigma_1 \subseteq \check{\Gamma}_1$ .

### 3.3 Teorema di Krein

**Lemma 3.3.1.** *Sia  $A \in \mathbb{C}^{n \times n}$  e sia  $N$  un intorno di  $\sigma(A)$ ; date  $F_1 : N \rightarrow \mathbb{C}^{\ell \times n}$  e  $F_2 : N \rightarrow \mathbb{C}^{n \times m}$  funzioni globalmente olomorfe e  $\Gamma \subseteq N$  un contorno con  $\sigma(A) \cap \Gamma = \emptyset$  allora:*

$$\begin{aligned} \left[ \frac{1}{2\pi i} \int_{\Gamma} F_1(\lambda) R(\lambda, A) d\lambda \right] \left[ \frac{1}{2\pi i} \int_{\Gamma} R(\lambda, A) F_2(\lambda) d\lambda \right] &= \\ &= \frac{1}{2\pi i} \int_{\Gamma} F_1(\lambda) R(\lambda, A) F_2(\lambda) d\lambda \end{aligned}$$

*Dimostrazione.* Possiamo allargare il contorno  $\Gamma$  attraverso il contorno  $\Gamma_1$  in maniera tale che  $\Gamma \subseteq \check{\Gamma}_1$ ,  $\hat{\Gamma}_1 \subseteq N$  e  $\check{\Gamma} \cap \sigma(A) = \check{\Gamma}_1 \cap \sigma(A)$ , possiamo allora scrivere:

$$\int_{\Gamma} R(\lambda, A) F_2(\lambda) d\lambda = \int_{\Gamma_1} R(\mu, A) F_2(\mu) d\mu$$

Allora abbiamo:

$$\begin{aligned} &\left[ \frac{1}{2\pi i} \int_{\Gamma} F_1(\lambda) R(\lambda, A) d\lambda \right] \left[ \frac{1}{2\pi i} \int_{\Gamma} R(\mu, A) F_2(\mu) d\mu \right] \\ &= -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma} F_1(\lambda) R(\lambda, A) R(\mu, A) F_2(\mu) d\lambda d\mu = \\ &= \frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma} F_1(\lambda) \frac{R(\lambda, A) - R(\mu, A)}{\lambda - \mu} F_2(\mu) d\lambda d\mu = \\ &= \frac{1}{4\pi^2} \int_{\Gamma} F_1(\lambda) R(\lambda, A) \left( \int_{\Gamma_1} \frac{F_2(\mu)}{\lambda - \mu} d\mu \right) d\lambda + \\ &\quad - \frac{1}{4\pi^2} \int_{\Gamma_1} \left( \int_{\Gamma} \frac{F_1(\lambda)}{\lambda - \mu} d\lambda \right) R(\mu, A) F_2(\mu) d\mu \end{aligned}$$

Dalla definizione di  $\Gamma_1$  abbiamo che:

- $\frac{F_2(\lambda)}{\lambda - \mu}$  è olomorfa in  $\mu$  su  $\check{\Gamma}_1$  tranne per  $\mu = \lambda$ , dove ha un polo di ordine uno;
- $\frac{F_1(\lambda)}{\lambda - \mu}$  è olomorfa in  $\lambda$  su  $\check{\Gamma}$ .



Da cui:

$$\int_{\Gamma_1} \frac{F_2(\mu)}{\lambda - \mu} d\mu = -2\pi i F_2(\lambda) \quad \int_{\Gamma} \frac{F_1(\lambda)}{\lambda - \mu} d\lambda = 0$$

Pertanto, otteniamo:

$$\frac{1}{2\pi i} \int_{\Gamma} F_1(\lambda) R(\lambda, A) F_2(\lambda) d\lambda$$

□

Il Teorema di Krein applica quanto visto finora per risolvere una funzione matriciale lineare polinomiale, vedi Sezione 1.3, ossia della forma:

$$\sum_{j,k=0}^{\ell} \alpha_{j,k} A^j X B^k = C \quad (3.2)$$

cui può essere associato il polinomio

$$\phi(x; y) = \sum_{j,k=0}^{\ell} \alpha_{j,k} y^k x^j$$

che ci permette di scrivere (3.2) nella forma  $Gx = c$  con

$$G = \phi(B^T; A) = \sum_{j,k=0}^{\ell} \alpha_{j,k} (B^T)^k \otimes A^j$$

con  $c = \text{Vec}(C)$  e  $x = \text{Vec}(X)$  incognita.

**Teorema 3.3.2** (Teorema di Krein). *Date*  $A \in \mathbb{C}^{m \times m}$ ,  $B \in \mathbb{C}^{n \times n}$ ,  $C \in \mathbb{C}^{m \times n}$ , *supponiamo* esistano  $\sigma_1 \subseteq \sigma(A)$ ,  $\sigma_2 \subseteq \sigma(B)$  *tali che*  $\phi(\lambda; \mu) \neq 0$   $\forall \lambda \in \sigma_1, \mu \in \sigma_2$ . *Siano*

$$P_1 = \frac{1}{2\pi i} \int_{\Gamma_1} R(\lambda, A) d\lambda \quad P_2 = \frac{1}{2\pi i} \int_{\Gamma_2} R(\mu, B) d\mu$$

*ossia*,  $P_1$  *è la proiezione definita da*  $A$  *e*  $\sigma_1$  *e*  $P_2$  *è la proiezione definita da*  $B$  *e*  $\sigma_2$ .

Se  $P_1CP_2 = C$ , allora la soluzione di (3.2) è:

$$X = -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \frac{R(\lambda, A)CR(\mu, B)}{\phi(\lambda; \mu)} d\mu d\lambda$$

*Dimostrazione.* Dal Teorema 3.2.2 abbiamo:

$$A^j P_1 = \frac{1}{2\pi i} \int_{\Gamma_1} \lambda^j R(\lambda, A) d\lambda \quad P_2 B^k = \frac{1}{2\pi i} \int_{\Gamma_2} \mu^k R(\mu, B) d\mu$$

Assumendo  $C = P_1CP_2$  abbiamo:

$$\begin{aligned} A^j X B^k &= A^j P_1 \left[ -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \frac{R(\lambda, A)CR(\mu, B)}{\phi(\lambda; \mu)} d\mu d\lambda \right] P_2 B^k = \\ &= \left[ \frac{1}{2\pi i} \int_{\Gamma_1} \lambda^j R(\lambda, A) d\lambda \right] \\ &\quad \left[ \frac{1}{2\pi i} \int_{\Gamma_1} R(\lambda, A) C \left( \frac{1}{2\pi i} \int_{\Gamma_2} \frac{R(\mu, B)}{\phi(\lambda; \mu)} d\mu \right) d\lambda \right] \\ &\quad \left[ \frac{1}{2\pi i} \int_{\Gamma_2} \mu^k R(\mu, B) d\mu \right] \end{aligned}$$

Poiché per ipotesi  $\phi(\lambda; \mu) \neq 0 \forall \lambda \in \sigma_1, \mu \in \sigma_2$ , possiamo scegliere  $\Gamma_1, \Gamma_2$  vicini a  $\sigma_1, \sigma_2$  rispettivamente tali che  $\frac{1}{2\pi i} \int_{\Gamma_2} \frac{R(\mu, B)}{\phi(\lambda; \mu)} d\mu$  sia una funzione regolare in  $\lambda$  su  $\check{\Gamma}_1$ .

Applicando, il Lemma 3.3.1 abbiamo

$$\begin{aligned} A^j X B^k &= \left[ \frac{1}{2\pi i} \int_{\Gamma_1} \lambda^j R(\lambda, A) C \left( \frac{1}{2\pi i} \int_{\Gamma_2} \frac{R(\mu, B)}{\phi(\lambda; \mu)} d\mu \right) d\lambda \right] \\ &\quad \left[ \frac{1}{2\pi i} \int_{\Gamma_2} \mu^k R(\mu, B) d\mu \right] \\ &= \left[ \frac{1}{2\pi i} \int_{\Gamma_2} \left( \frac{1}{2\pi i} \int_{\Gamma_1} \frac{1}{2\pi i} \frac{\lambda^j R(\lambda, A)}{\phi(\lambda; \mu)} d\lambda \right) CR(\mu, B) d\mu \right] \\ &\quad \left[ \frac{1}{2\pi i} \int_{\Gamma_2} \mu^k R(\mu, B) d\mu \right] \\ &= \frac{1}{2\pi i} \int_{\Gamma_2} \left( \frac{1}{2\pi i} \int_{\Gamma_1} \frac{\lambda^j R(\lambda, A)}{\phi(\lambda; \mu)} d\lambda \right) CR(\mu, B) \mu^k d\mu \\ &= -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \frac{\lambda^j \mu^k}{\phi(\lambda; \mu)} R(\lambda, A) CR(\mu, B) d\mu d\lambda \end{aligned}$$

Da cui:

$$\begin{aligned}
 \sum_{j,k=0}^{\ell} \alpha_{j,k} A^j X B^k &= \sum_{j,k=0}^{\ell} \alpha_{j,k} \left( -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \frac{\lambda^j \mu^k}{\phi(\lambda; \mu)} R(\lambda, A) C R(\mu, B) d\mu d\lambda \right) \\
 &= -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \underbrace{\left( \frac{\sum_{j,k=0}^{\ell} \alpha_{j,k} \lambda^j \mu^k}{\phi(\lambda; \mu)} \right)}_{=1} R(\lambda, A) C R(\mu, B) d\mu d\lambda \\
 &= -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} R(\lambda, A) C R(\mu, B) d\mu d\lambda \\
 &= \left( \frac{1}{2\pi i} \int_{\Gamma_1} R(\lambda, A) d\lambda \right) C \left( \frac{1}{2\pi i} \int_{\Gamma_2} R(\mu, B) d\mu \right) \\
 &= P_1 C P_2 = C
 \end{aligned}$$

□

### 3.4 Teorema di Krein per l'Equazione di Sylvester

Se applichiamo il Teorema di Krein all'Equazione di Sylvester

$$AX + XB = C \quad (3.3)$$

abbiamo che  $\phi(A; B) = I_n \otimes A + B^T \otimes I_m$ .

Inoltre, dal Teorema 3.3.2 abbiamo che per  $\sigma_1 \subseteq \sigma(A)$ ,  $\sigma_2 \subseteq \sigma(B)$  tali che  $\phi(\lambda; \mu) \neq 0 \forall \lambda \in \sigma_1, \mu \in \sigma_2$ , se  $P_1$  è la proiezione definita da  $A$  e  $\sigma_1$  e  $P_2$  è la proiezione definita da  $B$  e  $\sigma_2$  e vale  $P_1 C P_2 = C$ , allora la soluzione dell'equazione (3.3) è:

$$\begin{aligned}
 X &= -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \frac{R(\lambda, A) C R(\mu, B)}{\phi(\lambda; \mu)} d\mu d\lambda \\
 &= -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \frac{R(\lambda, A) C R(\mu, B)}{\lambda + \mu} d\mu d\lambda
 \end{aligned} \quad (3.4)$$

Abbiamo un caso in cui  $X$  si può esprimere alternativamente attraverso un integrale reale.

**Teorema 3.4.1.** *Nelle ipotesi precedenti, se  $\Re(\lambda + \mu) < 0 \forall \lambda \in \sigma_1, \mu \in \sigma_2$  allora:*

$$X = -\int_0^{+\infty} e^{At} C e^{Bt} dt$$

è soluzione di  $AX + XB = C$

*Dimostrazione.* Possiamo scegliere  $\Gamma_1, \Gamma_2$  in maniera tale che  $Re(\lambda + \mu) < 0$   $\forall \lambda \in \Gamma_1, \mu \in \Gamma_2$  così da poter scrivere:

$$\frac{1}{\lambda + \mu} = - \int_0^{+\infty} e^{(\lambda + \mu)t} dt$$

Da cui, sostituendo in (3.4)

$$\begin{aligned} X &= \frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} R(\lambda, A) C R(\mu, B) \int_0^{+\infty} e^{(\lambda + \mu)t} dt d\mu d\lambda \\ &= \frac{1}{4\pi^2} \int_0^{+\infty} \left( \int_{\Gamma_1} e^{\lambda t} R(\lambda, A) d\lambda \right) C \left( \int_{\Gamma_2} e^{\mu t} R(\mu, B) d\mu \right) dt \\ &= - \int_0^{+\infty} \left( \frac{1}{2\pi i} \int_{\Gamma_1} e^{\lambda t} R(\lambda, A) d\lambda \right) C \left( \frac{1}{2\pi i} \int_{\Gamma_2} e^{\mu t} R(\mu, B) d\mu \right) dt \end{aligned}$$

Usando il Teorema 3.2.2 otteniamo:

$$X = - \int_0^{+\infty} e^{At} P_1 C P_2 e^{Bt} dt = - \int_0^{+\infty} e^{At} C e^{Bt} dt$$

□

**Teorema 3.4.2.** Sia  $\Gamma$  qualsiasi contorno con nessun punto di  $\sigma(A) \cap \sigma(-B)$  su  $\Gamma$ . Sia  $\tau_1 = \sigma(A) \cap \check{\Gamma}$  e  $\tau_2 = \sigma(-B) \cap \check{\Gamma}$  con  $K_1$  la proiezione definita da  $A$  e  $\tau_1$  e  $K_2$  la proiezione definita da  $-B$  e  $\tau_2$ . Sia  $\delta$  un qualsiasi sottoinsieme di  $\sigma(-B)$  e  $J$  la proiezione definita da  $-B$  e  $\delta$ . Allora

$$X = \frac{1}{2\pi i} \int_{\Gamma} (I\lambda - A)^{-1} C (I\lambda + B)^{-1}_J d\lambda$$

è una soluzione di  $AX + XB = C$  se e solo se  $K_1 C J = C$  e  $CK_2 J = 0$  ( $K_2 J$  è la proiezione definita da  $-B$  e  $\tau_2 \cap \delta$ ).

*Dimostrazione.* Notiamo che:

- $A(I_m \lambda - A)^{-1} = \lambda(I_m \lambda - A)^{-1} - I_m$ , infatti:

$$A(I_m \lambda - A)^{-1} = \lambda(I_m \lambda - A)^{-1} - I_m;$$

$$A(I_m \lambda - A)^{-1} (I_m \lambda - A) = (\lambda(I_m \lambda - A)^{-1} - I_m) (I_m \lambda - A);$$

$$A = \lambda I_m - (I_m \lambda - A);$$

$$A = A$$

Questo vale poiché tutte le matrici che otteniamo commutano con  $A$  poiché somma e prodotto di potenze di  $A$ .

- $B(I_n\lambda + B)_J^{-1} = J - \lambda(I_n\lambda + B)_J^{-1}$ , infatti:

$$\begin{aligned} B(I_n\lambda + B)_J^{-1} &= J - \lambda(I_n\lambda + B)_J^{-1}; \\ BJ(I_n\lambda + B)^{-1} &= J - \lambda J(I_n\lambda + B)^{-1}; \\ BJ(I_n\lambda + B)^{-1}(I_n\lambda + B) &= (J - \lambda J(I_n\lambda + B)^{-1})(I_n\lambda + B); \\ BJ &= J(I_n\lambda + B) - J\lambda \\ BJ &= BJ \end{aligned}$$

Questo vale poiché tutte le matrici che otteniamo commutano con  $B$  poiché somma e prodotto di matrici che commutano con  $B$ .

Da queste osservazioni abbiamo:

$$\begin{aligned} AX + XB &= \frac{1}{2\pi i} \int_{\Gamma} A(I_m\lambda - A)^{-1} C(I_n\lambda + B)_J^{-1} d\lambda + \\ &\quad + \frac{1}{2\pi i} \int_{\Gamma} (I_m\lambda - A)^{-1} C(I_n\lambda + B)_J^{-1} B d\lambda \\ &= \frac{1}{2\pi i} \int_{\Gamma} [\lambda(I_m\lambda - A)^{-1} - I_m] C(I_n\lambda + B)_J^{-1} d\lambda + \\ &\quad + \frac{1}{2\pi i} \int_{\Gamma} (I_m\lambda - A)^{-1} C[J - \lambda(I_n\lambda + B)_J^{-1}] d\lambda \\ &= -\frac{1}{2\pi i} C \int_{\Gamma} (I_n\lambda + B)_J^{-1} d\lambda + \frac{1}{2\pi i} \int_{\Gamma} (I_m\lambda - A)^{-1} d\lambda \cdot CJ \\ &= K_1 CJ - CK_2 J = C - 0 = C \end{aligned}$$

□



## Capitolo 4

# Il Metodo dei Gradienti Coniugati (CG) per Equazioni Matriciali Lineari

Il Teorema di Krein ci offre uno strumento per poter calcolare direttamente la soluzione di un'equazione matriciale lineare polinomiale. Tuttavia, se consideriamo il caso di equazioni matriciali lineari a coefficienti reali, possiamo trovare la soluzione con il metodo iterativo dei Gradienti Coniugati (CG) in casi particolari.

In questo capitolo introduciamo l'algoritmo e alcuni criteri per valutarne l'applicabilità. Successivamente, eseguiremo l'algoritmo sfruttando un set di dati di diverse dimensioni, discutendone i risultati; in particolare, studieremo il numero di iterazioni, il tempo di esecuzione e la convergenza del residuo avvalendoci di risultati sia teorici sia sperimentali.

### 4.1 L'algoritmo

Ricapitolando quanto visto nel Capitolo 1, abbiamo che un'equazione matriciale lineare è una scrittura della forma:

$$A_1XB_1 + A_2XB_2 + \dots + A_pXB_p = C$$

con  $A_i \in \mathbb{C}^{m \times m}$ ,  $B_i \in \mathbb{C}^{n \times n}$  per  $i = 1, 2, \dots, p$  e  $C \in \mathbb{C}^{m \times n}$ ; con  $X \in \mathbb{C}^{m \times n}$  incognita.

Per il Teorema 1.2.1, tale equazione può essere riscritta nella forma:

$$Gx = c \quad (4.1)$$

con  $G = B_1^T \otimes A_1 + B_2^T \otimes A_2 + \dots + B_p^T \otimes A_p$  e  $c = \text{Vec}(C)$ , con  $x = \text{Vec}(X)$  incognita.

Possiamo, pertanto, utilizzare il metodo dei gradienti coniugati (CG) (vedi [QSSG]) nel caso del sistema lineare (4.1) quando la matrice  $G$  è simmetrica definita positiva.

Mantenendo la scrittura in (4.1) possiamo utilizzare l'algoritmo canonico; tuttavia, poiché  $x = \text{Vec}(X)$  e  $c = \text{Vec}(C)$  sono ottenute per impilamento, possiamo lavorare con le matrici, ossia contemporaneamente su tutte le colonne. Così facendo l'algoritmo diventa:

Fissati:

un dato iniziale  $X_0 \in \mathbb{R}^{m \times n}$

una tolleranza  $tol > 0$

un limite di iterazioni  $maxit \in \mathbb{N}$

Posti:

**Costo computazionale(flops)**

$$R_0 = C - \sum_{i=1}^p A_i X_0 B_i \quad 2p(\text{costo MxM}) + p(mn)$$

$$P_0 = R_0$$

$$\beta_0 = 0$$

$$\rho_0 = \|R_0\|_F^2 \quad mn + (mn - 1)$$

$$norm_0 = \sqrt{\rho_0} \quad 1$$

Per  $k = 0, 1, \dots, maxit$

$$W = \sum_{i=1}^p A_i P_k B_i \quad 2p(\text{costo MxM}) + (p - 1)mn$$

$$\alpha_k = \frac{\rho_k}{\text{tr}(P_k^T W)} \quad \text{costo MxM} + (n - 1) + 1$$

$$X_{k+1} = X_k + \alpha_k P_k \quad 2mn$$

$$R_{k+1} = R_k - \alpha_k W \quad 2mn$$

$$\rho_{k+1} = \|R_{k+1}\|_F^2 \quad mn + (mn - 1)$$

$$\text{Se } \frac{\sqrt{\rho_{k+1}}}{norm_0} < tol \quad 1 + 1$$



$$\begin{array}{ll}
\text{stop} & \\
\beta_k = \frac{\rho_{k+1}}{\rho_k} & 1 \\
P_{k+1} = R_{k+1} + \beta_k P_k & 2mn
\end{array}$$

Il costo computazionale a ogni iterazione pertanto sarà:  
il costo di  $(2p+1)$  prodotti MxM di cui:  $p$  tra matrici di dimensione  $m \times m$  e  $m \times n$ ;  $p$  tra matrici di dimensione  $m \times n$  e  $n \times n$ ; 1 tra matrici di dimensione  $n \times m$  e  $m \times n$ ; al costo di tali prodotti vanno aggiunti altri  $(p+7)mn + n + 2$  flops, che  $\mathcal{O}(mn)$  ed è dominato dai prodotti matriciali.  
Inoltre, il dispendio di memoria è molto alto in quanto prima delle iterazioni sono allocati in memoria i valori  $tol$ ,  $maxit$  e  $norm_0$  e a ogni iterazione sono memorizzati e sovrascritti i valori  $\rho$  e  $\beta$  e le matrici  $W$ ,  $R$ ,  $X$  e  $P$ , tutte  $m \times n$ ; pertanto, la memoria utilizzata cresce fortemente all'aumentare di  $m$  e  $n$ .

## 4.2 Condizione sufficiente per applicare CG

Vogliamo trovare una condizione sufficiente su  $A_i$  e  $B_i$  affinché  $G$  sia simmetrica definita positiva, ossia sia possibile applicare CG.

**Lemma 4.2.1.** *Siano  $M_1, M_2, \dots, M_p$  matrici simmetriche semidefinite positive, allora  $M = \sum_{i=1}^p M_i$  è simmetrica semidefinita positiva. Inoltre, se esiste  $\bar{i}$  tale che  $M_{\bar{i}}$  sia simmetrica definita positiva, allora  $M$  è simmetrica definita positiva.*

*Dimostrazione.* La simmetria è banale.

$M$  è semidefinita positiva poiché se consideriamo il quoziente di Rayleigh di  $M$  per un vettore  $x \neq 0$  abbiamo:

$$\frac{x^T M x}{x^T x} = \sum_{i=1}^p \frac{x^T M_i x}{x^T x} \geq \sum_{i=1}^p \lambda_{min}(M_i) \geq 0$$

poiché  $\lambda_{min}(M_i) \geq 0$  poiché  $M_i$  è semidefinita positiva per  $i = 1, 2, \dots, p$ ; da cui la tesi. La disuguaglianza è stretta nel caso una delle matrici sia definita positiva.  $\square$

**Proposizione 4.2.2.** *Condizione sufficiente affinché  $G$  sia simmetrica definita positiva è che  $A_i$  e  $B_i$  siano simmetriche semidefinite positive per  $i = 1, 2, \dots, p$  ed esista  $\bar{i}$  tale che  $A_{\bar{i}}$  e  $B_{\bar{i}}$  siano entrambe definite positive.*

*Dimostrazione.* Se  $A_i$  e  $B_i$  sono simmetriche semidefinite positive per  $i = 1, 2, \dots, p$  ed esiste  $\bar{i}$  tale che  $A_{\bar{i}}$  e  $B_{\bar{i}}$  siano entrambe definite positive: dalla Proposizione 1.1.7 abbiamo che  $B_{\bar{i}}^T \otimes A_i$  è simmetrica semidefinita positiva per  $i = 1, 2, \dots, p$  con  $B_{\bar{i}}^T \otimes A_{\bar{i}}$  definita positiva. Pertanto, dal Lemma 4.2.1 abbiamo che  $G = \sum_{i=1}^p B_{\bar{i}}^T \otimes A_i$  è simmetrica definita positiva.  $\square$

### 4.3 Un esempio

Consideriamo le seguenti matrici coefficiente  $A_i \in \mathbb{C}^{m \times m}$  e  $B_i \in \mathbb{C}^{n \times n}$  per  $i = 1, 2, 3$ :

$$A_1 = \begin{pmatrix} 3 & -1 & 0 & \dots & 0 \\ -1 & 3 & -1 & \ddots & \vdots \\ 0 & -1 & 3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & -1 & 3 \end{pmatrix} \quad B_1 = I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 3 & 1 & 0 & \dots & 0 \\ 1 & 3 & 1 & \ddots & \vdots \\ 0 & 1 & 3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & 1 & 3 \end{pmatrix} \quad B_2 = \begin{pmatrix} 2 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \ddots \\ 0 & 0 & 2 & 0 & \ddots \\ 0 & 0 & 0 & 1 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 5 & 1 & -\frac{1}{2} & 0 & \dots & 0 \\ 1 & 5 & 1 & -\frac{1}{2} & \ddots & \vdots \\ -\frac{1}{2} & 1 & 5 & 1 & \ddots & 0 \\ 0 & -\frac{1}{2} & 1 & 5 & \ddots & -\frac{1}{2} \\ \vdots & \ddots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & -\frac{1}{2} & 1 & 5 \end{pmatrix} \quad B_3 = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}$$

e la seguente matrice di termine noto:

$$C = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

#### 4.3.1 Applicabilità di CG

Per poter applicare il metodo dei gradienti coniugati, vogliamo che  $G = \sum_{i=1}^3 B_i^T \otimes A_i$  sia simmetrica definita positiva; pertanto, controlliamo se le matrici coefficiente soddisfano la Proposizione 4.2.2, ossia se  $A_i$  e  $B_i$  sono simmetriche semidefinite positive per  $i = 1, 2, 3$  con almeno una coppia  $A_{\bar{i}}, B_{\bar{i}}$  di matrici definite positive.

Preliminarmente, dall'algebra lineare (vedi [S]) sappiamo che una matrice simmetrica a valori reali ha autovalori reali. Pertanto, possiamo provare a utilizzare i dischi di Gerschgorin (vedi [QSSG]) per righe (che coincidono con quelli per colonne data la simmetria) per determinare una regione del piano complesso contenente gli autovalori reali.

Per la matrice

- $A_1$ : i dischi di Gerschgorin per righe sono:

$$\begin{aligned} \mathcal{G}_1^{(r)} = \mathcal{G}_m^{(r)} &= \{z \in \mathbb{C} \text{ tale che } |z - 3| \leq 1\} \\ \mathcal{G}_i^{(r)} = \mathcal{G}_j^{(r)} &= \{z \in \mathbb{C} \text{ tale che } |z - 3| \leq 2\} \quad \text{per } i, j \neq 1, m \end{aligned}$$

Pertanto, gli autovalori sono strettamente positivi e la matrice è definita positiva.

- $A_2$ : i dischi di Gerschgorin per righe sono:

$$\begin{aligned} \mathcal{G}_1^{(r)} = \mathcal{G}_m^{(r)} &= \{z \in \mathbb{C} \text{ tale che } |z - 3| \leq 1\} \\ \mathcal{G}_i^{(r)} = \mathcal{G}_j^{(r)} &= \{z \in \mathbb{C} \text{ tale che } |z - 3| \leq 2\} \quad \text{per } i, j \neq 1, m \end{aligned}$$

Pertanto, gli autovalori sono strettamente positivi e la matrice è definita positiva.

- $A_3$ : i dischi di Gerschgorin per righe sono:

$$\mathcal{G}_1^{(r)} = \mathcal{G}_m^{(r)} = \{z \in \mathbb{C} \text{ tale che } |z - 5| \leq \frac{3}{2}\}$$

$$\mathcal{G}_2^{(r)} = \mathcal{G}_{m-1}^{(r)} = \{z \in \mathbb{C} \text{ tale che } |z - 5| \leq \frac{5}{2}\}$$

$$\mathcal{G}_i^{(r)} = \mathcal{G}_j^{(r)} = \{z \in \mathbb{C} \text{ tale che } |z - 5| \leq 3\} \quad \text{per } i, j \neq 1, 2, m-1, m$$

Pertanto, gli autovalori sono strettamente positivi e la matrice è definita positiva.

- $B_1$ : è la matrice identità  $I_n$ , che è banalmente definita positiva.
- $B_2$ : è una matrice diagonale, pertanto il suo spettro è dato dagli elementi diagonali, ossia  $\sigma(B_2) = \{1, 2\}$  e la matrice è definita positiva.
- $B_3$ : i dischi di Gerschgorin per righe sono:

$$\mathcal{G}_1^{(r)} = \mathcal{G}_m^{(r)} = \{z \in \mathbb{C} \text{ tale che } |z - 2| \leq 1\}$$

$$\mathcal{G}_i^{(r)} = \mathcal{G}_j^{(r)} = \{z \in \mathbb{C} \text{ tale che } |z - 2| \leq 2\} \quad \text{per } i, j \neq 1, m$$

Pertanto, gli autovalori sono non negativi e la matrice è semidefinita positiva

Abbiamo che  $A_i$  e  $B_i$  sono semidefinite positive per  $i = 1, 2, 3$  e definite positive per  $i = 1, 2$ ; pertanto, dalla Proposizione 4.2.2 abbiamo che  $G = \sum_{i=1}^3 (B_i^T \otimes A_i)$  è simmetrica definita positiva.

### 4.3.2 Risultati

Abbiamo applicato il Metodo dei Gradienti Coniugati all'esempio descritto precedentemente. La macchina utilizzata per l'esecuzione dell'algoritmo in linguaggio MATLAB<sup>®</sup> ha le seguenti specifiche tecniche:

- Costruttore: HP inc.
- Modello: Elite Slice
- Processore: Intel<sup>®</sup> Core<sup>™</sup> i5-6500T @ 2.50 GHz
- RAM: 8,00 GB
- Sistema operativo: Windows 10 Pro

Abbiamo fatto variare le dimensioni della matrici  $A_i$  ponendo  $m = 100, 1000, 10,000$ ; inoltre, per ogni  $m$ , abbiamo variato le dimensioni  $n$  delle

matrici  $B_i$  da 10 a 50 con passo 10, da 100 a 500 con passo 100, da 1000 a 5000 con passo 1000 e infine pari a 10,000. I parametri utilizzati sono stati  $maxit = 5000$  e  $tol = 10^{-6}$ ; inoltre, il dato iniziale  $X_0$  è stato ottenuto in maniera casuale con valori in  $[0, 1]$  attraverso il comando `rand`.

Nel nostro esempio, la struttura a banda delle matrici  $A_i$  e  $B_i$ , ossia la loro sparsità, può essere sfruttata sia per diminuire il consumo di memoria sia per abbassare notevolmente il costo computazionale.

Possiamo memorizzare solamente gli elementi non nulli delle matrici  $A_i$  e  $B_i$  utilizzando alcuni comandi di MATLAB<sup>®</sup> dedicati alla costruzione di matrici sparse. Pertanto, per definire le matrici utilizziamo il seguente codice:

```
em=ones(m,1);
A1=spdiags([-1*em 3*em -1*em],[-1:1,m,m]);
A2=spdiags([em 3*em em],[-1:1,m,m]);
A3=spdiags([-0.5*em em 5*em em -0.5*em],[-2:2,m,m]);
A={A1,A2,A3};
en=ones(n,1);
B1=speye(n);
B2=sparse(kron(eye(n/2),[2,0;0,1]));
B3=spdiags([-1*en 2*en -1*en],[-1:1,n,n]);
B={B1,B2,B3};
C=ones(m,1)*ones(1,n);
X0=rand([m,n]);
```

Questo codice permette, inoltre, che la struttura delle matrici  $A_i$  e  $B_i$  sia anche sfruttata per ridurre il costo della valutazione dell'espressione matriciale:

$$\sum_{i=1}^3 (A_i X B_i)$$

che è presente a ogni iterazione per il calcolo di  $W$ .

Avremo che i costi computazionali dei singoli addendi sono i seguenti:

- $A_1 X B_1$  è dato dal prodotto  $A_1 X$ , il cui costo è  $5nm - 4n$  poiché  $A_1$  è tridiagonale, e dal prodotto  $(A_1 X) B_1$ , il cui costo è 0 poiché  $B_1$  è la matrice identità; il costo totale è  $5nm - 4n$ .
- $A_2 X B_2$  è dato dal prodotto  $A_2 X$ , il cui costo è  $5nm - 4n$  poiché  $A_2$  è tridiagonale, e dal prodotto  $(A_2 X) B_2$ , il cui costo è  $mn$  poiché  $B_2$  è

diagonale; il costo totale è  $6mn - 4n$ .

- $A_3XB_3$  è dato dal prodotto  $A_3X$ , il cui costo è  $9nm - 12n$  poiché  $A_3$  è pentadiagonale, e dal prodotto  $(A_2X)B_2$ , il cui costo è  $5mn - 4m$  poiché  $B_3$  è tridiagonale; il costo totale è  $14mn - 4m - 12n$ .

La somma dei 3 termini costa  $2mn$ ; pertanto, il costo computazionale totale della valutazione è:  $27mn - 4m - 20n$ .

Quindi, è stato ridotto il costo della valutazione dell'algoritmo, poiché il prodotto tra 6 coppie di matrici è stato portato all'ordine di  $\mathcal{O}(nm)$ . Così facendo, il costo computazionale dell'algoritmo per iterazione diventa:

costo di 1 prodotto MxM (prodotto tra matrici  $n \times m$  e  $m \times n$ ) +  $34mn - 4m - 19n + 1$ .

### Iterazioni, tempo di esecuzione e residuo relativo

Utilizzando i comandi `tic` e `toc`, abbiamo misurato il tempo di esecuzione dell'algoritmo.

Possiamo, inoltre, studiare il residuo assoluto calcolando:

$$resabs = \|C - \sum_{i=1}^3 A_i \tilde{X} B_i\|_F$$

per la soluzione  $\tilde{X}$  fornita dall'algoritmo.

Analogamente possiamo studiare il residuo relativo calcolando:

$$resrel = \frac{\|C - \sum_{i=1}^3 A_i \tilde{X} B_i\|_F}{\|C - \sum_{i=1}^3 A_i X_0 B_i\|_F}$$

dove  $X_0$  è il dato iniziale. Essendo il residuo relativo il criterio di arresto dell'algoritmo, si può prendere in considerazione il valore assunto a convergenza.

Dall'esecuzione dell'algoritmo abbiamo ottenuto la seguente tabella con residuo relativo (*resrel*), tempo di esecuzione (*time*) e numero di iterazioni (*it*) al variare di  $m$  e  $n$ :

$m =$	100			1000			10,000		
$n =$	<i>resrel</i>	<i>time(s)</i>	<i>it</i>	<i>resrel</i>	<i>time(s)</i>	<i>it</i>	<i>resrel</i>	<i>time(s)</i>	<i>it</i>
10	$4.29 \cdot 10^{-7}$	0.0038	15	$4.38 \cdot 10^{-7}$	0.0135	15	$4.40 \cdot 10^{-7}$	0.0559	15
20	$6.90 \cdot 10^{-7}$	0.0043	15	$7.02 \cdot 10^{-7}$	0.0137	15	$7.06 \cdot 10^{-7}$	0.1738	15
30	$6.85 \cdot 10^{-7}$	0.0045	15	$6.78 \cdot 10^{-7}$	0.0191	15	$6.83 \cdot 10^{-7}$	0.2475	15
40	$6.85 \cdot 10^{-7}$	0.0049	15	$6.86 \cdot 10^{-7}$	0.0253	15	$6.90 \cdot 10^{-7}$	0.3248	15
50	$6.86 \cdot 10^{-7}$	0.0065	15	$6.91 \cdot 10^{-7}$	0.0271	15	$6.87 \cdot 10^{-7}$	0.4004	15
100	$6.73 \cdot 10^{-7}$	0.0084	15	$6.83 \cdot 10^{-7}$	0.0533	15	$6.83 \cdot 10^{-7}$	0.7931	15
200	$6.74 \cdot 10^{-7}$	0.0165	15	$6.78 \cdot 10^{-7}$	0.1701	15	$6.78 \cdot 10^{-7}$	1.5878	15
300	$6.67 \cdot 10^{-7}$	0.0247	15	$6.75 \cdot 10^{-7}$	0.2572	15	$6.75 \cdot 10^{-7}$	2.4385	15
400	$6.77 \cdot 10^{-7}$	0.0403	15	$6.75 \cdot 10^{-7}$	0.3569	15	$6.73 \cdot 10^{-7}$	3.3174	15
500	$6.76 \cdot 10^{-7}$	0.0515	15	$6.71 \cdot 10^{-7}$	0.4514	15	$6.73 \cdot 10^{-7}$	4.2375	15
1000	$6.76 \cdot 10^{-7}$	0.1214	15	$6.71 \cdot 10^{-7}$	1.0480	15	$6.71 \cdot 10^{-7}$	9.5702	15
2000	$6.70 \cdot 10^{-7}$	0.3934	15	$6.70 \cdot 10^{-7}$	2.4769	15	$6.70 \cdot 10^{-7}$	22.6931	15
3000	$6.69 \cdot 10^{-7}$	0.7093	15	$6.71 \cdot 10^{-7}$	4.4284	15	$6.70 \cdot 10^{-7}$	40.2469	15
4000	$6.70 \cdot 10^{-7}$	1.0766	15	$6.69 \cdot 10^{-7}$	6.4934	15	$6.69 \cdot 10^{-7}$	60.2097	15
5000	$6.68 \cdot 10^{-7}$	1.5461	15	$6.69 \cdot 10^{-7}$	9.1776	15	$6.70 \cdot 10^{-7}$	84.2343	15
10,000	$6.69 \cdot 10^{-7}$	5.3601	15	$6.69 \cdot 10^{-7}$	29.0904	15	$6.69 \cdot 10^{-7}$	265.3309	15

Dalla tabella sopra riportata si evince come i residui siano molto simili e il numero di iterazioni non vari; questo può essere ricondotto a una convergenza simile al variare delle dimensioni e sarà oggetto di successivo studio. Si nota, inoltre, che:

- per  $m, n$  piccole (almeno una  $\leq 100$ ): il tempo di esecuzione è più influenzato dalla variabilità del dato iniziale  $X_0$  che dalla effettiva complessità del problema dovuta alle dimensioni delle matrici  $A_i$  e  $B_i$ , ad esempio il tempo di esecuzione per  $m = 1000$  e  $n = 10$  è circa uguale al tempo per  $m = 1000$  e  $n = 20$ .
- per  $m, n$  grandi (entrambe  $> 100$ ): il tempo di esecuzione è fortemente condizionato dall'aumento delle dimensioni a causa del conseguente aumento del costo computazionale per iterazione e della memoria utilizzata.

### Studio della convergenza

Il fatto che il numero di iterazioni sia costante e il residuo relativo non cambi notevolmente al variare delle dimensioni, ci fa sospettare che la convergenza possa essere analoga per ogni valore delle di  $m$  e  $n$ .

Per  $m = 100$  e  $n = 10, 100, 1000, 10,000$ , se creiamo un grafico che tiene conto sull'asse delle ascisse del numero di iterazioni e sull'asse delle ordinate

(in scala logaritmica) del residuo relativo a ogni iterazione, otteniamo:

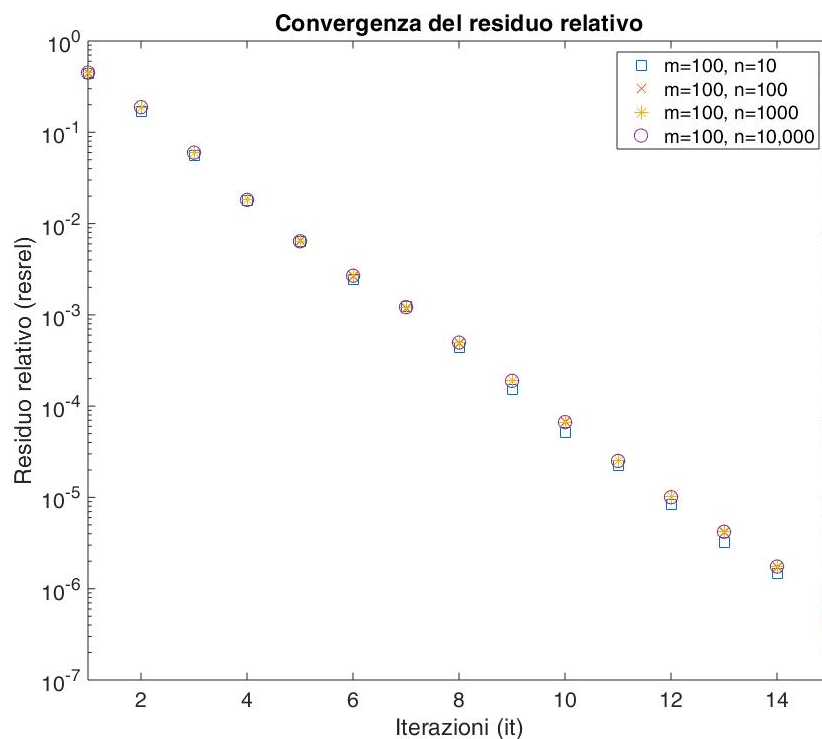


Figura 4.1: Grafico del residuo relativo per  $m = 100$  e  $n = 10, 100, 1000, 10,000$ .

Si può notare come l'aumento delle dimensioni non influenzi l'andamento della convergenza. Un risultato analogo si vede per  $n = 1000, 10,000$ ; pertanto, possiamo dire che la convergenza è indipendente dai valori di  $m$  e  $n$ .

Questo risultato ha un riscontro teorico; infatti, sulla convergenza dell'errore per il Metodo dei Gradienti Coniugati per un sistema lineare abbiamo la seguente proposizione:

**Proposizione 4.3.1.** *Dato un sistema del tipo  $Gx = c$  con  $G$  matrice quadrata reale simmetrica definita positiva, con soluzione  $\hat{x}$ ; sia  $x_k$  la soluzione al passo  $k$  ottenuta applicando il Metodo dei Gradienti Coniugati, allora*



l'errore al passo  $k$  gode della seguente proprietà

$$\|\hat{x} - x_k\|_G \leq 2 \left( \frac{\sqrt{\kappa_2(G)} - 1}{\sqrt{\kappa_2(G)} + 1} \right)^k \|\hat{x} - x_0\|_G \quad (4.2)$$

*Dimostrazione.* Si veda [QSSG].  $\square$

Dalle ipotesi della Proposizione 4.3.1, possiamo ricavare un analogo risultato per il residuo.

Se denotiamo  $r_k = c - Ax_k$ ,  $r_0 = c - Ax_0$  per i residui e  $e_k = \hat{x} - x_k$ ,  $e_0 = \hat{x} - x_0$  per gli errori, abbiamo  $r_k = Ge_k$  e  $r_0 = Ge_0$ . Poiché  $G$  è simmetrica definita positiva, abbiamo che  $G^{-1}$ ,  $G^{\frac{1}{2}}$  e  $G^{-\frac{1}{2}}$  sono ben definite, simmetriche e si possono facilmente ottenere dalla decomposizione spettrale della matrice; tale risultato può essere utilizzato per avere i seguenti risultati.

Per il residuo e l'errore al passo  $k$  abbiamo:

$$\begin{aligned} \|r_k\|_2^2 &= \|Ge_k\|_2^2 \\ &\leq \|G^{\frac{1}{2}}\|_2^2 \|G^{\frac{1}{2}}e_k\|_2^2 \\ &= \|G^{\frac{1}{2}}\|_2^2 (e_k^T (G^{\frac{1}{2}})^T G^{\frac{1}{2}} e_k) \\ &= \|G^{\frac{1}{2}}\|_2^2 (e_k^T Ge_k) \\ &= \|G^{\frac{1}{2}}\|_2^2 \|e_k\|_G^2 \end{aligned}$$

Per il residuo e l'errore al passo iniziale abbiamo:

$$\begin{aligned} \|e_0\|_G^2 &= \|G^{-1}r_0\|_G^2 \\ &= r_0^T (G^{-1})^T GG^{-1}r_0 \\ &= r_0^T G^{-1}r_0 \\ &= r_0^T (G^{-\frac{1}{2}})^T G^{-\frac{1}{2}}r_0 \\ &= \|G^{-\frac{1}{2}}r_0\|_2^2 \\ &\leq \|G^{-\frac{1}{2}}\|_2^2 \|r_0\|_2^2 \end{aligned}$$

Da cui, considerando le radici quadrate, otteniamo:

$$\|e_k\|_G \geq \frac{1}{\|G^{\frac{1}{2}}\|_2} \|r_k\|_2 \quad \|e_0\|_G \leq \|G^{-\frac{1}{2}}\|_2 \|r_0\|_2 \quad (4.3)$$

Se consideriamo il rapporto tra la norma dell'errore al passo  $k$  e la norma dell'errore iniziale e applichiamo le maggiorazioni in (4.3) otteniamo:

$$\frac{\|e_k\|_G}{\|e_0\|_G} \geq \frac{1}{\|G^{\frac{1}{2}}\|_2 \|G^{-\frac{1}{2}}\|_2} \frac{\|r_k\|_2}{\|r_0\|_2} = \frac{1}{\kappa_2(G^{\frac{1}{2}})} \frac{\|r_k\|_2}{\|r_0\|_2} \quad (4.4)$$

Se sostituiamo tale risultato in (4.2) otteniamo:

$$2 \left( \frac{\sqrt{\kappa_2(G)} - 1}{\sqrt{\kappa_2(G)} + 1} \right)^k \geq \frac{\|\hat{x} - x_k\|_G}{\|\hat{x} - x_0\|_G} \geq \frac{1}{\kappa_2(G^{\frac{1}{2}})} \frac{\|r_k\|_2}{\|r_0\|_2}$$

da cui ricaviamo il seguente risultato sul residuo:

$$\|r_k\|_2 \leq 2\kappa_2(G^{\frac{1}{2}}) \left( \frac{\sqrt{\kappa_2(G)} - 1}{\sqrt{\kappa_2(G)} + 1} \right)^k \|r_0\|_2$$

abbiamo, quindi, che il fattore di convergenza per il residuo è:

$$\frac{\sqrt{\kappa_2(G)} - 1}{\sqrt{\kappa_2(G)} + 1}$$

lo stesso dell'errore, come abbiamo visto nella Proposizione 4.3.1

Dobbiamo allora stimare  $\kappa_2(G)$  e calcolarne il valore per poter confrontare l'andamento del fattore di convergenza con quello del residuo.

Per la stima possiamo utilizzare quanto già sappiamo da [QSSG]:

$$\kappa_2(G) = \frac{\lambda_{max}(G)}{\lambda_{min}(G)}$$

pertanto, per stimare  $\kappa_2(G)$  dobbiamo stimare l'intervallo spettrale di  $G$ .

Applicando la funzione `eig` alle matrici  $A_i$  e  $B_i$  per  $i = 1, 2, 3$  possiamo stimare il raggio spettrale di ogni matrice al variare di  $m$  e  $n$ . Dall'esecuzione otteniamo che i raggi spettrali di  $A_i$  e  $B_i$  non sono influenzati dalle dimensioni  $m$  e  $n$  e abbiamo:

- $\sigma(A_1) \subseteq [1, 5]$
- $\sigma(A_2) \subseteq [1, 5]$
- $\sigma(A_3) \subseteq [2, 6.5]$
- $\sigma(B_1) = 1$

- $\sigma(B_2) = 1, 2$
- $\sigma(B_3) \subseteq [0, 4]$

Dal Corollario 1.1.6 abbiamo che gli autovalori di un prodotto di Kronecker tra due matrici sono tutti i possibili prodotti degli autovalori delle matrici; pertanto abbiamo:

- $\sigma(B_1^T \otimes A_1) \subseteq [1, 5]$
- $\sigma(B_2^T \otimes A_2) \subseteq [1, 10]$
- $\sigma(B_3^T \otimes A_3) \subseteq [0, 26]$

Se, infine, consideriamo il quoziente di Rayleigh di  $G$  per un suo generico vettore non nullo  $x \in \mathbb{C}^{mn \times mn}$  otteniamo:

$$\begin{aligned} \frac{x^T G x}{x^T x} &= \sum_{i=1}^3 \frac{x^T (B_i^T \otimes A_i) x}{x^T x} \\ &\geq \sum_{i=1}^3 \lambda_{\min}(B_i^T \otimes A_i) \geq 2 \end{aligned}$$

Procedendo analogamente e maggiorando con  $\lambda_{\max}$  ricaviamo:

$$\frac{x^T G x}{x^T x} \leq \sum_{i=1}^3 \lambda_{\max}(B_i^T \otimes A_i) \leq 41$$

Da cui otteniamo che

$$\sigma(G) \subseteq [2, 41]$$

per ogni dimensione  $m$  e  $n$ .

Possiamo allora stimare dall'alto il numero di condizionamento in norma-2:

$$\kappa_2(G) = \frac{\lambda_{\max}(G)}{\lambda_{\min}(G)} \leq \frac{41}{2} = 20.5$$

il quale a partire dalla Proposizione 4.3.1 e dalle successive osservazioni, ci dà un controllo sull'errore e sul residuo.

Per avere, invece, un calcolo esatto del valore di  $\kappa_2(G)$  possiamo utilizzare la funzione `cond` per ottenere il numero di condizionamento di  $G$ . Poiché abbiamo visto dalla Figura 4.3.2 che la convergenza di CG non varia con

le dimensioni, ci aspettiamo che anche il numero di condizionamento non varierà. Infatti, al variare di  $m$  e  $n$  abbiamo sempre che:

$$\kappa_2(G) \simeq 5.2$$

Possiamo verificare la correttezza di questo valore aggiungendo alla Figura 4.3.2 un grafico che decresce di un fattore  $\frac{\sqrt{\kappa_2(G)-1}}{\sqrt{\kappa_2(G)+1}}$  per  $\kappa_2(G) = 5.2$  a ogni iterazione. Così facendo otteniamo:

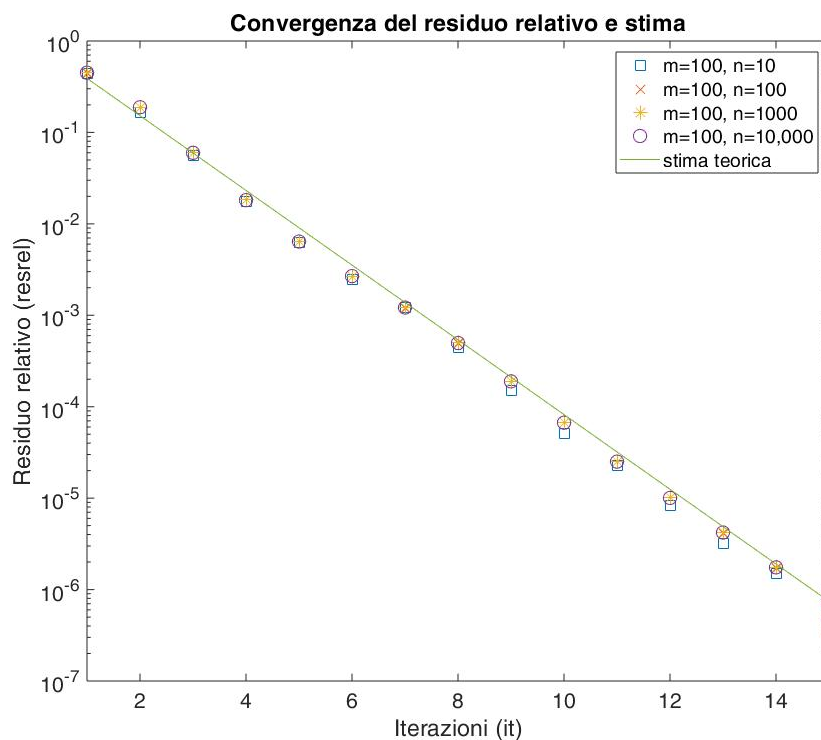


Figura 4.2: Grafico del residuo relativo per  $m = 100$  e  $n = 10, 100, 1000, 10,000$  e del fattore di convergenza.

Pertanto, abbiamo che il fattore di convergenza ricalca perfettamente la convergenza empirica data dall'esecuzione dell'algoritmo.



# Conclusioni

Abbiamo visto come le equazioni matriciali lineari possano essere studiate da diversi punti di vista.

Un approccio legato all'algebra lineare e all'analisi complessa ci permette di calcolare la soluzione di un'equazione matriciale lineare polinomiale attraverso il Teorema di Krein, che può essere esteso all'Equazione di Sylvester.

Un approccio di tipo numerico, sviluppato nel Capitolo 4, mostra come il Metodo dei Gradienti Coniugati possa essere applicato con successo alle equazioni matriciali sfruttando la loro formulazione vettoriale mediante l'operatore di Kronecker.

L'algoritmo CG converge in poche iterazioni anche per matrici di grandi dimensioni. Infatti, l'aver sfruttato la sparsità delle matrici nel nostro esempio ha portato a una riduzione del tempo di esecuzione, ma non ha influito sul numero di iterazioni, che sarebbe rimasto tale anche se fossero state memorizzate come matrici piene.

Inoltre, un grande vantaggio di questo metodo è la stima che abbiamo sull'errore relativo, che ci dà pieno controllo sulla convergenza del metodo. Pertanto, possiamo concludere che il Metodo dei Gradienti Coniugati è un ottimo metodo non solo per i sistemi ma anche per le equazioni matriciali lineari.

Ciononostante, CG ha il limite di poter essere applicato solo alle equazioni la cui scrittura sotto forma di sistema lineare ha matrice dei coefficienti simmetrica definita positiva. Questa verifica può diventare estremamente complessa nel caso in cui le matrici coefficiente siano piene, di grandi dimensioni e i termini dell'equazione siano numerosi.





# Bibliografia

- [LM] Peter Lancaster, Miron Tismenetsky, *The Theory of Matrices*, Academic Press, 1985, pp. 570.
- [L] Peter Lancaster, “*Explicit Solutions of Linear Matrix Equations*” in *SIAM Review Vol. 12 No. 4*, 1970, pp. 544-566.
- [QSSG] Alfio Quarteroni, Riccardo Sacco, Fausto Saleri, Paola Gervasio, *Matematica Numerica*, Springer, 2014, pp.520.
- [S] Edoardo Sernesi, *Geometria 1*, Bollati Boringhieri, 1989, pp.480.
- [R] Walter Rudin, *Analisi reale e complessa*, Bollati Boringhieri, 1974, pp.451.