# 1 Lezione del 17-03-25

Avevamo visto i primi 2 teoremi di Gershgorin, cioè:

- **Primo teorema di Gershgorin:** ogni autovalore $\lambda$ è contenuto in:

$$\lambda \in \bigcup_{i=1}^{n} \mathcal{F}_i(A)$$

definiti i *cerchi di Gershgorin*:

$$\mathcal{F}_i(A) = \left\{ z \in \mathbb{C}, \quad |z - a_{ii}| < \sum_{j \neq i} a_{ij} \right\}$$

- **Secondo teorema di Gershgorin:** un $M_1$ è unione di $s$ cerchi e $M_2$ unione dei restanti $n - s$, allora: $M_1 \cap M_2 = \emptyset \implies M_1$ contiene $s$ autovalori e $M_2$ i restanti $n - s$.

Vediamo quindi il terzo teorema di Gershgorin. Per fare ciò, introduciamo il concetto di **matrice irriducibile**, e ancor prima di **grafo** associato alla matrice.

## 1.1 Matrici irriducibili

Diamo la definizione:

**Definizione 1.1: Grafo associato ad una matrice**

Data una matrice $A \in \mathbb{C}^{n \times n}$ si dice grafo associato ad $A$, indicato con $G(A) = (V, E)$, il grafo che ha come vertici l'insieme $V = \{1, ..., n\}$ e come archi l'insieme $E$ così definito:

$$(i, j) \in E \ \Leftrightarrow \ a_{ij} \neq 0$$

Solitamente si ignora la diagonale (notiamo che in caso di entrate $\neq 0$ risultano semplicemente in anelli aperti e chiusi sullo stesso nodo).

A noi interesseranno grafi **fortemente connessi**:

**Definizione 1.2: Grafo fortemente connesso**

Un grafo si dice fortemente connesso se per ogni scelta di vertici $i, j \in 1, ..., n$ esiste un cammino *orientato* sul grafo che parte da $i$ e arriva a $j$.

A questo punto possiamo definire le matrici **irriducibili**:

**Definizione 1.3: Matrice irriducibile**

Una matrice $A \in \mathbb{C}^{n \times n}$ si dice irriducibile se il grafo $G(A)$ è fortemente connesso.

Possiamo quindi enunciare il terzo teorema di Gershgorin:

**Teorema 1.1: Terzo teorema di Gershgorin**

Data una matrice $A \in \mathbb{C}^{n \times n}$ irriducibile si ha che, se $\lambda$ è un autovalore tale che:

$$\lambda \in \partial \left( \bigcup_{i=1}^{n} \mathcal{F}_i(A) \right)$$

dove con $\partial$ si indica il bordo, cioè esiste un indice, $\exists k \in \{1, ..., n\}$ tale che: $|\lambda - a_{kk}| = \sum_{j \neq k} a_{kj}$. Allora dovrà essere che:

$$\lambda \in \bigcap_{i=1}^{n} \partial \left( \mathcal{F}_i(A) \right)$$

cioè per *tutti* gli indici, $\forall k \in \{1, ..., n\}$ vale che: $|\lambda - a_{kk}| = \sum_{j \neq k} a_{kj}$.

Questo significa che, nel caso di matrici irriducibili, un autovalore che sta sul bordo di un disco di Gershgorin dovrà necessariamente stare sul bordo di *tutti* i dischi di Gershgorin.

Avevamo visto la definizione di matrici a predominanza diagonale forte. Esistono anche matrici a predominanza diagonale *debole* (o semplicemente *diagonali deboli*):

**Definizione 1.4: Matrice a predominanza diagonale debole**

Una matrice $A$ si dice a predominanza diagonale debole se vale:

$$|a_{ii}| \geq \sum_{j \neq i} |aij|$$

ed $\exists k \in \{1, ..., n\}$ tale per cui:

$$|a_{kk}| > \sum_{j \neq k} |akj|$$

Un **corollario** è che se $A$ è a predominanza diagonale debole ed è irriducibile, allora $A$ non è singolare.

Questo si dimostra direttamente dal fatto che il valore $0$ appartiene ai cerchi per cui vale $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}$, ma esiste almeno un cerchio per cui vale $|a_{kk}| > \sum_{j \neq k} |akj|$, cioè non contiene lo $0$, e quindi per il terzo teorema di Gershgorin $0$ non può essere autovalore (matrice non singolare). $\square$

## 1.2 Matrici a blocchi

Talvolta è conveniente definire una matrice in termini di *sottomatrici*, o **blocchi**, al posto delle entrate scalari che la compongono, cioè si può scrivere:

$$A \in \mathbb{C}^{m \times n}, \quad A = \begin{pmatrix} A_{11} & \dots & A_{1r} \\ \dots & & \dots \\ A_{s1} & \dots & A_{sr} \end{pmatrix}$$

Si può sempre scrivere una matrice a blocchi, purché le sottomatrici abbiano dimensioni compatibili (sia le somme delle dimensioni devono corrispondere alle dimensioni effettive della matrice, sia le dimensioni delle matrici adiacenti devono corrispondere), cioè si deve avere:

$$A_{ij} \in \mathbb{C}^{m_j \times n_j}, \quad \sum_{i=1}^{s} m_i = m, \sum_{j=1}^{r} n_j = n$$

Notiamo che per noi sarà il importante il caso di matrici quadrate poste come sottomatrici quadrate, cioè $m = n$, $s = r$ e $m_i = n_i \ \forall i = 1, ..., r \cdot s$.

Vediamo quindi la definizione di matrici **triangolari a blocchi**:

**Definizione 1.5: Matrice triangolare a blocchi**

Nelle condizioni di cui sopra (in particolare, matrici quadrate) una matrice si dice triangolare a blocchi se sono $\neq 0$ tutti i blocchi lungo la diagonale, e sopra o sotto la diagonale (dove rispettivamente si parla di matrice triangolare superiore o inferiore), cioè:

$$A \in \mathbb{C}^{m \times n}, \quad A = \begin{pmatrix} A_{11} & ... & A_{1r} \\ 0 & ... & ... \\ 0 & 0 & A_{sr} \end{pmatrix} \text{ (superiore) o } A = \begin{pmatrix} A_{11} & 0 & 0 \\ ... & ... & 0 \\ A_{s1} & ... & A_{sr} \end{pmatrix} \text{ (inferiore)}$$

Vediamo anche le matrici **diagonali a blocchi**:

**Definizione 1.6: Matrice diagonale a blocchi**

Nelle condizoni della definizione 7.5, una matrice si dice diagonale a blocchi se sono $\neq 0$ i soli blocchi lungo la diagonale, cioè:

$$A \in \mathbb{C}^{m \times n}, \quad A = \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & ... & 0 \\ 0 & 0 & A_{sr} \end{pmatrix} \text{ (diagonale)}$$

### 1.2.1 Proprietà delle matrici a blocchi

Vediamo alcune proprietà delle matrici a blocchi:

1. Se $A$ è triangolare a blocchi, allora l'insieme degli autovalori di $A$ corrisponde all'unione degli insiemi di autovalori associati ai blocchi sulla diagonale $A_{jj}$;

2. Dalla scorsa proprietà valgono, a cascata, tutte le proprietà che avevamo visto su autovalori e determinanti. Ad esempio il determinante sarà:

$$\det(A) = \prod_{j=1}^{s} \det(A_{jj})$$

ricordando che:

$$\det(A_{jj}) = \prod_{i=1}^{\frac{n}{s}} \lambda_i$$

con i $\lambda_i$ autovalori di $A_{jj}$;

3. Se $A$ è diagonale, a blocchi, vale:

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 & 0 \\ 0 & ... & 0 \\ 0 & 0 & A_{sr}^{-1} \end{pmatrix}$$

4. Se si hanno 2 matrici a blcchi con lo stesso partizionamento, allora calcolare $A+B$ e $A \cdot B$ si può fare trattando i sottoblocchi come entrate scalari, cioè date:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

con $A_{11}$ corrispondente in dimensioni a $B_{11}$ e via dicendo, si ha che:

$$A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix}$$

riguardo alla somma, e considerazioni simili (che non riportiamo) riguardo al prodotto.

## 1.3 Matrici riducibili

Vediamo ora una definizione che è effettivamente *duale* a quella di matrice irriducibile: quella di **matrice riducibile**:

**Definizione 1.7: Matrice riducibile**

Una matrice $A \in \mathbb{C}^{n \times n}$, $n \geq 2$, si dice riducibile se $\exists \Pi$ matrice di permutazione tale che:

$$\Pi A \Pi^T = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

a blocchi, con $A_{11} \in \mathbb{C}^{k \times k}$, $A_{22} \in \mathbb{C}^{(n-k) \times (n-k)}$, $k \in \{1, ..., n\}$.

Osserviamo che ci possono essere più di una matrice di permutazione $\Pi$ che portano $A$ in una forma triangolare a blocchi, con diversi valori di $k$.

Osserviamo poi che la forma triangolare superiore non è imperativa, cioè se $\exists \Pi_1$ tale che:

$$\Pi_1 A \Pi_1^T = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

triangolare superiore, allora vale anche che $\exists \Pi_2$ tale che:

$$\Pi_2 A \Pi_2^T = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$$

triangolare inferiore. Vediamo anzi che in verità queste due forme sono equivalenti.

### 1.3.1 Sistemi lineari con matrici riducibili

Supponiamo di voler risolvere $Ax = b$ con $A \in \mathbb{C}^n \times n$, con $A$ riducibile, e conoscendo $T$. In questo caso conviene procedere sfruttando:

$$Ax = b \Leftrightarrow \Pi A x = \Pi b \Leftrightarrow \Pi A \Pi^T \Pi x = \Pi b \Leftrightarrow By = c$$

dove si è detto $B = \Pi A \Pi^T$, $c = \Pi b$, e $y = \Pi x$. $By = c$ sarà in forma:

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

con $c_1, y_1 \in \mathbb{C}^k$ e $c_2, y_2 \in \mathbb{C}^{n-k}$.

Una volta risolto $By = C$, si ritrova $x$ mediante la relazione:

$$\Pi x = y \ \rightarrow \ x = \Pi^T y$$

Vediamo come risolvere effettivamente questo sistema. Scrivendo il sistema $2 \times 2$ a blocchi per esteso si ottiene:

$$\begin{cases} A_{11}y_1 + A_{12}y_2 = C_1 \\ A_{22}y_2 = c_2 \end{cases}$$

L'ultima equazione rappresenta un sistema $(n-k) \times (n-k)$, che risolviamo per trovare $y_2$ (notiamo non c'è dipendenza da $y_1$, dalla definizione come triangolare superiore a blocchi della matrice $B$). A questo punto si ottiene un altro sistema lineare:

$$A_{11}y_1 = c_1 - A_{12}y_2$$

con $y_2$ stavolta noto.

Il *guadagno* di questo metodo è che risolvere un sistema $n \times n$ costa generalmente $O(n^3)$. Effettuare il *preprocessing* reso disponibile dalla riducibilità della matrice ci permette di rendere la complessità uguale a:

$$O(k^3 + (n-k)^3 + k(n-k)) = O(k^3 + (n-k)^3)$$

dove $k(n-k)$ alla prima equazione rappresenta il costo della moltiplicazione matrice-vettore $A_{11}y_2$ nella seconda equazione di risoluzione del sistema a blocchi.

Se si hanno molti $0$ da sfruttare, e quindi si può scegliere $k = \frac{n}{2}$, si ottiene un costo:

$$O\left(\left(\frac{n}{2}\right)^3 + \left(n - \frac{n}{2}\right)^3\right) = O\left(2\left(\frac{n}{2}\right)^3\right) = O\left(\frac{n^3}{4}\right)$$

cioè si guadagna di un fattore di 4 in termini di tempo di esecuzione.

Osserviamo inoltre che se i blocchi $A_{11}$, $A_{22}$ sono a loro volta matrici riducibili, il processo si può applicare **ricorsivamente** sulle matrici della diagonale, in modo da ridurre ulteriormente la complessità del sistema (a patto di dover effettuare più passi di preprocessing).