Richiami e nuove nozioni di Algebra Lineare III^a parte

Calcolo Numerico - Ing. Inf. - Lezione 5

Outline

Outline

Cerchi di Gershgorin

Data una matrice $A \in \mathbb{C}^{n \times n}$, defininiamo gli insiemi

$$\mathcal{F}_i = \{z \in \mathbb{C} \;\mid\; | z - a_{ii} | \leq \rho_i\}, \quad \text{dove} \quad \rho_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} | a_{ij} |$$

$i = 1, 2, \ldots, n$, chiamati **cerchi di Gershgorin**

Sul piano complesso il cerchio di Gershgorin $\mathcal{F}_i$ ha centro nel punto a_{ii} raggio ρ_i

I^o Teorema di Gershgorin

Se λ è autovalore di $A \in \mathbb{C}^{n \times n}$ risulta

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

Dimostrazione

Indichiamo con x un autovettore destro associato all'autovalore λ e sia x_k **una** sua componente di massimo modulo, cioè

$$|x_k| \geq |x_i|, \quad i = 1, 2, \ldots, n$$

Dalla k-esima equazione del sistema $Ax = \lambda x$ si ha

$$\sum_{j=1}^{n} a_{kj} \, x_j = \lambda \, x_k$$

da cui

$$(\lambda - a_{kk}) \, x_k = \sum_{\substack{j=1 \\ j \neq k}}^{n} a_{kj} \, x_j$$

Passando ai moduli dei due membri si ottiene

$$| \lambda - a_{kk} | \, | x_k | = \left| \sum_{\substack{j=1 \\ j \neq k}}^{n} a_{kj} \, x_j \right| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} | a_{kj} | \, | x_j |$$

Maggiorando nel secondo membro ogni $|\, x_j \,|$ con $|\, x_k \,|$, si ottiene

$$|\, \lambda - a_{kk} \,|\,|\, x_k \,| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |\, a_{kj} \,|\,|\, x_k \,|$$

Per la definizione di autovettore, $x_k \neq 0$ per cui

$$|\, \lambda - a_{kk} \,| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |\, a_{kj} \,| = \rho_k$$

L'ultimo risultato ci dice che l'autovalore λ appartiene al cerchio $\mathcal{F}_k$

A priori, non è noto quale sia l'indice k di una componente di massimo modulo dell'autovettore per cui si può solo concludere che sicuramente

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

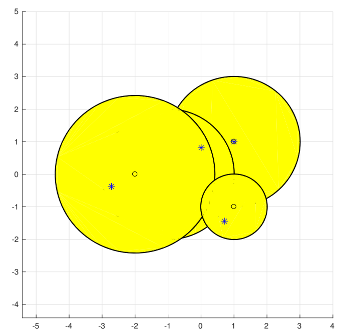
perchè appartiene ad almeno uno dei cerchi di Gershgorin

Esempio 1

Applichiamo il I^o Teorema di Gershgorin alla matrice

$$A = \begin{pmatrix} 1+i & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \\ 0 & 1+i & -2 & 1 \\ 0 & -1 & 0 & 1-i \end{pmatrix}$$

Esempio 1



Gli * di colore blu individuano gli autovalori della matrice

Conseguenza del I^o Teorema di Gershgorin

Corollario

Una matrice a predominanza diagonale forte **non è singolare**

Dimostrazione

Dal primo teorema di Gershgorin, tenendo presente che ogni cerchio di Gershgorin ha il centro a distanza $\mid a_{ii} \mid$ dall'origine degli assi ed il raggio $\rho_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \mid a_{ij} \mid$
Dall'ipotesi segue che nessuno dei detti cerchi contiene l'origine.
Lo zero non è quindi autovalore della matrice ed il determinante non può essere nullo

II^o e III^o Teorema di Gershgorin

II^o Teorema di Gershgorin

Se M_1 è l'unione di k cerchi di Gershgorin e M_2 è l'unione dei rimanenti $n-k$ ed è $M_1 \bigcap M_2 = \emptyset$
allora
k autovalori appartengono a M_1 e $n-k$ a M_2

III^o Teorema di Gershgorin

Se A è una matrice irriducibile
allora
se un autovalore appartiene alla frontiera dell'unione dei cerchi di Gershgorin esso appartiene alla frontiera di tutti i cerchi costituenti l'insieme $\mathcal{F}$

Conseguenza del II^o Teorema di Gershgorin

Corollario

Se una matrice presenta cerchi di Gershgorin due a due disgiunti **allora** gli autovalori sono due a due distinti

Questo risultato è una semplice generalizzazione del II^o Teorema di Gershgorin che si enuncia per due insiemi M_1 e M_2 disgiunti ma che può essere esteso ad insiemi due a due disgiunti

Conseguenza del III^o Teorema di Gershgorin

Corollario

Una matrice A a predominanza diagonale debole ed irriducibile è **non singolare**

Dimostrazione

Si osserva che la predominanza diagonale debole consente ai cerchi di Gershgorin di avere la circonferenza passante per l'origine, eccettuato uno almeno di tali cerchi

D'altra parte se lo zero fosse autovalore della matrice irriducibile A, per il terzo teorema di Gershgorin esso dovrebbe appartenere alla frontiera di tutti i cerchi, contrariamente a quanto osservato

Dunque lo zero non è autovalore e perciò $\det(A) \neq 0$

Matrice di Frobenius

Sia data l'equazione algebrica

$$x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0 = 0$$

con $a_i \in \mathbb{C}$, $i = 0, 1, \ldots, k-1$

Si consideri la matrice quadrata di ordine k

$$F = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{k-2} & -a_{k-1} \end{pmatrix}$$

detta **matrice di Frobenius** (o associata o compagna)

Matrice di Frobenius

Si verifica che l'equazione data è l'equazione caratteristica della matrice F per cui i suoi autovalori sono le radici dell'equazione

Quindi è possibile localizzare le radici dell'equazione algebrica facendo uso del I^o Teorema di Gershgorin

Esempio 2

Si consideri l'equazione algebrica

$$x^3 - 5\,x^2 + 3\,x - 1 = 0$$

Le radici dell'equazione sono gli autovalori della matrice

$$F = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 5 \end{pmatrix}$$

Esempio 2

Le radici dell'equazione appartengono all'insieme $\mathcal{F}$ unione dei cerchi

$$\mathcal{F}_1 = \mathcal{F}_2 = \{z \in \mathbb{C} \mid \mid z \mid \leq 1\}\,, \quad \mathcal{F}_3 = \{z \in \mathbb{C} \mid \mid z - 5 \mid \leq 4\}$$

Se si considera la matrice F^T e si applica ad essa il teorema di Gershgorin (ciò equivale ad operare sulle colonne di F) si ottiene l'insieme $\mathcal{G}$ unione dei cerchi

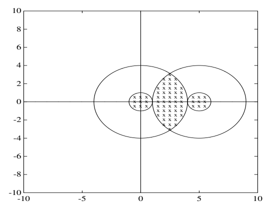
$$\begin{aligned} \mathcal{G}_1 &= \{z \in \mathbb{C} \mid \mid z \mid \leq 1\} \\ \mathcal{G}_2 &= \{z \in \mathbb{C} \mid \mid z \mid \leq 4\} \\ \mathcal{G}_3 &= \{z \in \mathbb{C} \mid \mid z - 5 \mid \leq 1\} \end{aligned}$$

Esempio 2

Ricordando che gli autovalori delle matrici F e F^T sono gli stessi, si deduce che gli autovalori, quindi le radici dell'equazione proposta, appartengono all'insieme

$$\mathcal{F} \bigcap \mathcal{G}$$

che è più ristretto sia di $\mathcal{F}$ che di $\mathcal{G}$



Outline

In molte applicazioni risulta necessario confrontare fra loro due vettori o due matrici
Per fare questo è utile il concetto di norma

Definizione

Si dice **norma vettoriale**, e si indica con $\|x\|$, una funzione

$$\|\cdot\| : \mathbb{C}^n \longmapsto \mathbb{R}_0^+$$

che verifica le seguenti condizioni:

$$\begin{aligned} &\|x\| = 0 \iff x = 0 \\ &\|\alpha x\| = |\,\alpha\,|\;\|x\|\,, \qquad \forall x \in \mathbb{C}^n,\ \forall \alpha \in \mathbb{C}\,; \\ &\|x+y\| \le \|x\| + \|y\|\,, \qquad \forall x, y \in \mathbb{C}^n \end{aligned}$$

In $\mathbb{C}^n$ si possono definire molte norme in modo arbitrario ma ci sono alcune norme cosiddette classiche

$\|x\|_1 = \sum_{i=1}^{n} \mid x_i \mid$ *norma 1*;

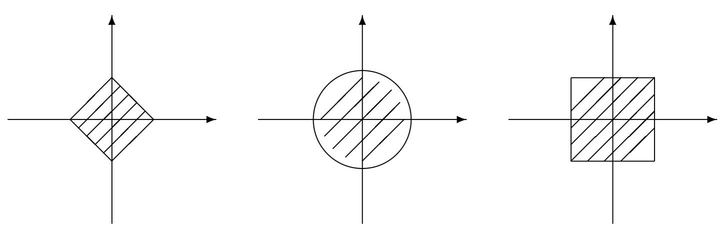
$\|x\|_2 = \sqrt{\sum_{i=1}^{n} \mid x_i \mid^2}$ *norma 2* o *norma euclidea*;

$\|x\|_\infty = \max_{1\le i\le n} \mid x_i \mid$ *norma* ∞.

Nella figura è riportato l'insieme

$$S = \{x \in \mathbb{R}^2 \mid \|x\| \le 1\},$$

detto sfera unitaria di $\mathbb{R}^2$, per le tre norme classiche



Teorema

Ogni norma vettoriale è uniformemente continua su $\mathbb{C}^n$

Teorema di equivalenza fra norme

Date due norme vettoriali $\|x\|_p$ e $\|x\|_q$, esistono $\alpha,\ \beta \in \mathbb{R}^+$ tali che

$$\alpha\|x\|_p \le \|x\|_q \le \beta\|x\|_p \qquad \forall x \in \mathbb{C}^n$$

Per le norme classiche valgono le relazioni

$$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2$$
$$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$$
$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$$

Definizione

Si dice **norma matriciale**, e si indica con $\|A\|$, una funzione

$$\|\cdot\| : \mathbb{C}^{n\times n} \longmapsto \mathbb{R}_0^+$$

che verifica le seguenti condizioni:

$$\|A\| = 0 \Longleftrightarrow A = \mathbf{O}\,;$$

$$\|\alpha A\| = |\,\alpha\,|\,\|A\|\,, \qquad \forall A \in \mathbb{C}^{n\times n},\ \forall \alpha \in \mathbb{C}\,;$$

$$\|A + B\| \le \|A\| + \|B\|\,, \qquad \forall A, B \in \mathbb{C}^{n\times n}\,;$$

$$\|AB\| \le \|A\|\,\|B\|\,, \qquad \forall A, B \in \mathbb{C}^{n\times n}\,.$$

Norme matriciali indotte

Si possono ottenere norme matriciali facendo ricorso alle norme vettoriali definendo

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

In questo caso la norma matriciale si dice **naturale** o **indotta** dalla norma vettoriale considerata

Norme coerenti

Una norma matriciale si dice **coerente** o **compatibile** con una norma vettoriale se si ha

$$\|Ax\| \le \|A\|\,\|x\|\,, \qquad \forall A \in \mathbb{C}^{n\times n},\ \forall x \in \mathbb{C}^n$$

Le norme indotte sono coerenti con le rispettive norme vettoriali poiché

$$\frac{\|Ax\|}{\|x\|} \le \|A\| \Longrightarrow \|Ax\| \le \|A\|\,\|x\|$$

Norma matriciali classiche

Le norme matriciali indotte dalle tre norme vettoriali classiche sono

$\|A\|_1 = \max_j \sum_{i=1}^{n} \mid a_{ij} \mid,$ *norma 1;*

$\|A\|_2 = \sqrt{\rho(A^H A)},$ *norma 2 o norma euclidea;*

$\|A\|_\infty = \max_i \sum_{j=1}^{n} \mid a_{ij} \mid,$ *norma* ∞.

Norma di Frobenius

Esistono norme matriciali che non sono indotte da norme vettoriali
Un esempio è dato dalla **norma matriciale di Frobenius** definita da

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \mid a_{ij} \mid^2}$$

Norma di Frobenius

La norma di Frobenius non è indotta da alcuna norma vettoriale poiché risulta

$$\|I\|_F = \sqrt{n}$$

mentre si ha

$$\|I\| = \sup_{x \neq 0} \frac{\|Ix\|}{\|x\|} = 1$$

qualunque sia la norma vettoriale considerata

Ad una matrice $A \in \mathbb{C}^{n\times n}$ possiamo abbinare tre numeri particolari: determinante, raggio spettrale e norma

Concentriamo la nostra attenzione su raggio spettrale e norma che, per definizione, sono numeri reali non negativi

Esiste una relazione tra $\rho(A)$ e $\|A\|$ che sia indipendente dalla particolare matrice considerata?

La risposta è data dal Teorema che segue

Teorema di Hirsh

Teorema di Hirsh

Sia $A \in \mathbb{C}^{n\times n}$; per ogni norma matriciale (indotta o no) vale la relazione

$$\rho(A) \;\le\; \|A\|$$

Dimostrazione

Sia λ un autovalore di A; quindi si ha $Ax = \lambda\, x$ con x autovettore destro associato a λ

Consideriamo la matrice $B = (x|0|0|\cdots|0) \in \mathbb{C}^{n\times n}$ per cui vale l'uguaglianza

$$A\,B \;=\; \lambda\, B$$

Teorema di Hirsh

Dimostrazione

Passando alle norme si ha

$$\|\, A\,B \,\| \;=\; \|\, \lambda\, B \,\|$$

e, per le proprietà delle norme matriciali,

$$\|A\|\,\|B\| \;\ge\; |\lambda|\,\|B\|$$

La matrice B non è uguale alla matrice nulla per cui $\|B\| \ne 0$ e quindi risulta

$$\|\, A \,\| \;\ge\; |\lambda|$$

Poiché λ è un qualunque autovalore di A, la relazione precedente è valida anche per l'autovalore il cui modulo coincide con $\rho(A)$, da cui la tesi

Conseguenze del Teorema di Hirsh

Corollario

Affinché una matrice sia convergente è sufficiente che una sua norma risulti minore di 1

Corollario

Se $A \in \mathbb{C}^{n\times n}$, gli autovalori di A appartengono al cerchio

$$\{z \in \mathbb{C} \mid |z| \leq \|A\|\}$$

dove $\|\cdot\|$ è una qualunque norma matriciale

Un esempio in cui vale la relazione

$$\rho(A) = \|A\|$$

è dato dalle matrici hermitiane ($A = A^H$)

Infatti si ha

$$\|A\|_2 = \sqrt{\rho(A^H A)} = \sqrt{\rho(A^2)} = \sqrt{\rho^2(A)} = \rho(A)$$

Esempio

Le matrici della forma

$$G_{rt} = \begin{pmatrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& c &&&& -s &&& \\ &&&& 1 &&&&&& \\ &&&&& \ddots &&&&& \\ &&&&&& 1 &&&& \\ &&& s &&&& c &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow r \\ \\ \\ \\ \leftarrow t \\ \\ \\ \\ \end{array}$$

dove $G_{rt} \in \mathbb{R}^{n\times n}$ e $c^2 + s^2 = 1$, si dicono **matrici di rotazione**

Esempio

La matrice G_{rt} è ortogonale per cui, da $\det(G_{rt}G_{rt}^T) = 1$, risulta

$$\det(G_{rt}) = \lambda_1 \, \lambda_2 \cdots \lambda_n = \pm 1$$

Dal Teorema di Hirsh si ottiene

$$|\, \lambda_i \,| \leq \|G_{rt}\|_2 = \sqrt{\rho(G_{rt}G_{rt}^T)} = 1$$

Segue che può essere solo

$$|\, \lambda_i \,| = 1 \qquad i = 1, 2, \ldots, n$$

Esempio

L'equazione caratteristica della matrice G_{rt} è

$$(1-\lambda)^{n-2}(\lambda^2 - 2c\lambda + 1) = 0$$

pertanto gli autovalori di G_{rt} sono

$$\lambda_1 = \cdots = \lambda_{n-2} = 1, \quad \lambda_{n-1} = c - \sqrt{-s^2}, \quad \lambda_n = c + \sqrt{-s^2}$$