

1 Lezione del 14-03-25

1.1 Norme

Concludiamo il ripasso di algebra lineare parlando del concetto di **norma**.

1.1.1 Norme vettoriali

In molti casi è utile definire (quantificare) la "grandezza" di un vettore, prendendo il vettore 0 come nullo e ogni vettore come più grande di esso. Formalmente, quello che si fa è definire funzioni con proprietà specifiche dette **norme**.

Definizione 1.1: Norma

Una funzione $|\cdot| : \mathbb{C}^n \rightarrow \mathbb{R}^+$ è detta norma se verifica le 3 proprietà:

1. $|x| = 0 \Leftrightarrow x = 0$
2. $|\alpha x| = |\alpha| \cdot |x|, \quad \forall \alpha \in \mathbb{C}, \forall x \in \mathbb{C}^n$
3. $|x + y| \leq |x| + |y|, \quad \forall x, y \in \mathbb{C}^n$ (disuguaglianza del triangolo)

Possiamo dare alcuni esempi di norme:

- $|x|_1 = \sum_{i=1}^n |x_i|$ (norma di Manhattan)
- $|x|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$ (norma Euclidea)
- $|x|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ (p -norma)
- $|x|_\infty = \max_{i=1, \dots, n} |x_i|$ (norma a infinito)

Notiamo che l'insieme dei vettori di norma 1 cambia geometria al variare della norma che scegliamo. Ad esempio, in \mathbb{R}^2 , l'insieme dei vettori di norma 1 dà un quadrato ruotato di 45 gradi, dei vettori di norma 2 una circonferenza, e via via al crescere di p si approssima un quadrato unitario centrato sull'origine (che la norma a infinito effettivamente raggiunge).

Le norme sono equivalenti in quanto esiste il teorema:

Teorema 1.1: Norme equivalenti

Date due norme $|\cdot|_a$ e $|\cdot|_b$, si ha che $\exists \alpha, \beta \in \mathbb{R}^+$ tali che:

$$\alpha |x|_a \leq |x|_b \leq \beta |x|_a, \quad \forall x \in \mathbb{C}^n$$

Questa proprietà è sempre valida su spazi vettoriali finiti (non vale necessariamente lo stesso su spazi infiniti). Se vale il teorema 6.1, quindi, allora vale anche che:

$$\beta^{-1} |x|_b \leq |x|_a \leq \alpha^{-1} |x|_b, \quad \forall x \in \mathbb{C}^n$$

Per le norme più comuni ($|\cdot|_1, |\cdot|_2, |\cdot|_\infty$) valgono le disuguaglianze:

- $|x|_2 \leq |x|_1 \leq \sqrt{n} |x|_2$

Dimostriamo entrambe le disuguaglianze.

$$- |x|_2 \leq |x|_1$$

Questo si ha da:

$$\sqrt{\sum_{i=1}^n |x_i|^2} \leq \sum_{i=1}^n |x_i|$$

prendendo i quadrati:

$$\sum_{i=1}^n |x_i|^2 \leq \left(\sum_{i=1}^n |x_i| \right)^2 = \sum_{i=1}^n |x_i|^2 + 2 \sum_{i < j} |x_i| |x_j|$$

che è chiaramente vero con $2 \sum_{i < j} |x_i| |x_j| > 0$.

$$- |x|_1 \leq \sqrt{n} |x|_2$$

Questo si nota prendendo:

$$\sum_{i=1}^n |x_i| = (||x||, 1) = |x_1| \cdot 1 + \dots + |x_n| \cdot 1$$

dove $||x||$ rappresenta il modulo componente a componente di x . Varrà allora dalla disuguaglianza di Cauchy-Schwarz:

$$(|x|, 1) \leq ||x|| \cdot |1| = \sqrt{\sum_{i=1}^n |x_i|^2} \cdot \sqrt{\sum_{i=1}^n |1|^2} = \sqrt{n} \sqrt{\sum_{i=1}^n |x_i|^2}$$

che è la tesi. □

$$\bullet |x|_\infty \leq |x|_1 \leq n |x|_\infty$$

Questo risulta chiaro da:

$$\max_{i=1, \dots, n} |x_i| \leq \sum_{i=1}^n |x_i| \leq n \max_{i=1, \dots, n} |x_i|$$

□

$$\bullet |x|_\infty \leq |x|_2 \leq \sqrt{n} |x|_\infty$$

Anche questo risulta chiaro da:

$$\max_{i=1, \dots, n} |x_i| \leq \sqrt{\sum_{i=1}^n |x_i|^2} \leq \sqrt{n} \max_{i=1, \dots, n} |x_i|$$

prendendo i quadrati:

$$\left(\max_{i=1, \dots, n} |x_i| \right)^2 \leq \sum_{i=1}^n |x_i|^2 \leq n \left(\max_{i=1, \dots, n} |x_i| \right)^2$$

□

1.1.2 Norme matriciali

Si possono definire norme anche per matrici.

Definizione 1.2: Norma matriciale

Una funzione $|\cdot| : \mathbb{C}^{m \times n} \rightarrow \mathbb{R}^+$ è detta norma matriciale se verifica le 3 proprietà:

1. $|A| = 0 \Leftrightarrow A = 0$
2. $|\alpha A| = |\alpha| \cdot |A|, \quad \forall \alpha \in \mathbb{C}, \forall A \in \mathbb{C}^{m \times n}$
3. $|A + B| \leq |A| + |B|, \quad \forall A, B \in \mathbb{C}^{m \times n}$
4. Se $m = n$ chiediamo inoltre, visto che è ben definito il prodotto matriciale:
 $|A \cdot B| \leq |A| \cdot |B|, \quad \forall A, B \in \mathbb{C}^{m \times n}$
da qui in poi assumiamo solo questo caso.

Esiste un modo per trovare norme matriciali a partire da norme vettoriali.

Definizione 1.3: Norma indotta

Una norma matriciale $|\cdot|_m$ si dice indotta se esiste una norma vettoriale $|\cdot|_v$ tale che:

$$|A|_m = \sup_{|x|_v \neq 0} \frac{|Ax|_v}{|x|_v} = \sup_{|x|_v = 1} |Ax|_v$$

Intuitivamente, diciamo che una matrice è tanto più "grande" tanto più *allarga* i vettori a distanza unitaria.

Le norme matriciali indotte dalle norme vettoriali più comuni ($|\cdot|_1, |\cdot|_2, |\cdot|_\infty$) sono le seguenti:

- $|A|_1 = \max_{j=1, \dots, n} \sum_{i=1}^n |a_{ij}|$, cioè il massimo della somme delle colonne di A ;
- $|A|_\infty = \max_{i=1, \dots, n} \sum_{j=1}^n |a_{ij}|$, cioè il massimo delle somme delle righe di A ;
- $|A|_2 = \sqrt{\rho(A^H A)}$, cioè la radice dell'autovalore di modulo massimo di $A^H A$. Notiamo che questo metodo è sensibilmente più difficile di calcolare la somma delle entrate delle matrice in riga o in colonna.

Prendiamo $A^H A$ anzichè A in quanto:

$$|A|_2 = \sup_{|x|_2 = 1} |Ax|_2$$

e:

$$|Ax|_2 = \langle Ax, Ax \rangle = \langle x, A^H A x \rangle$$

il resto della dimostrazione non viene dato in quanto si basa su un risultato che esula dal corso (*quoziente di Rayleigh*), ma resta il fatto che ci serve una forma del tipo $A^H A$ anzichè la sola matrice A .

Un esempio di norma matriciale non indotta è la norma di Frobenius:

$$|A|_F = \sqrt{\sum_{ij} |\alpha_{ij}|^2}$$

che possiamo assumere come una grande norma $|\cdot|_2$ su tutta la matrice.

Un modo per verificare che questa norma non è effettivamente indotta è osservare che:

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

mentre per ogni norma indotta vale che:

$$|I|_i = \sup_{|x| \neq 0} \frac{|Ix|}{|x|} = \sup_{|x| \neq 0} \frac{|x|}{|x|} = 1 \neq \sqrt{n} \quad \forall n > 1$$

□

Notiamo che questa verifica vale in una sola direzione: se la norma dell'identità valesse effettivamente 1, potremmo dire che questa proprietà non è violata, ma non che di conseguenza la norma è indotta, cioè:

$$|I|_m \neq 1 \implies |\cdot|_m \text{ non è indotta}$$

$$|I|_m = 1 \not\implies |\cdot|_m \text{ è indotta (può non esserlo)}$$

Diamo quindi la definizione di compatibilità:

Definizione 1.4: Compatibilità

Una norma matriciale $|\cdot|_m$ si dice compatibile con una norma vettoriale $|\cdot|_v$ se vale:

$$|Ax|_v \leq |A|_m \cdot |x|_v, \quad \forall A \in \mathbb{C}^{n \times n}, \quad \forall x \in \mathbb{C}^n$$

Osserviamo che le norme matriciali indotte sono ovviamente compatibili con le norme vettoriali che le inducono. Infatti posto $x \neq 0$:

$$\frac{|Ax|_v}{|x|_v} \leq |A|_m \implies |Ax|_v \leq |A|_m \cdot |x|_v$$

□

1.1.3 Norme e autovalori

Le disuguaglianze precedenti e le proprietà delle norme ci dicono che gli autovalori non possono essere ovunque sul piano complesso, ma in modulo devono essere più piccoli di $|A|_m$ se $|\cdot|_m$ è compatibile (a maggior ragione indotta) con una norma vettoriale. Infatti, con (x, λ) autocoppia per A , cioè:

$$Ax = \lambda x, \quad \lambda \in \mathbb{C}, \quad x \neq 0$$

si ha:

$$|\lambda x| = |\lambda| \cdot |x| = |Ax|_v \leq |A|_m \cdot |x|_v \implies |\lambda| \leq |A|_m$$

□

Questo significa che se si prende il "disco" centrato sull'origine con raggio uguale alla norma della matrice A , allora tutti gli autovalori di A devono trovarsi all'interno di tale disco, cioè:

$$\lambda \text{ autovalore} \Leftrightarrow \lambda \in \{z \in \mathbb{C} : |z| \leq |A|_m\}$$

In realtà l'assunzione $|\cdot|_m$ compatibile può essere rimossa, ovvero vale il seguente teorema:

Teorema 1.2: di Hirsch

Data $A \in \mathbb{C}^{n \times n}$ allora $\forall |\cdot|$ norma matriciale vale:

$$\rho(A) = \max\{|\lambda| \text{ autovalore}\} \leq |A|$$

Questo si dimostra prendendo (λ, x) autocoppia per A e considerando la matrice:

$$B = \begin{pmatrix} x & | & 0 \end{pmatrix}$$

per cui vale:

$$A \cdot B = \begin{pmatrix} Ax & A0 & \dots & A0 \end{pmatrix} = \begin{pmatrix} \lambda x & 0 & \dots & 0 \end{pmatrix} = \lambda \begin{pmatrix} x & 0 & \dots & 0 \end{pmatrix} = \lambda B$$

allora vale:

$$|\lambda||B| = |\lambda B| = |A \cdot B| \leq |A| \cdot |B|$$

dividendo per B ($B \neq 0$ da $x \neq 0$ autovettore):

$$\frac{|\lambda||B|}{|B|} \leq \frac{|A| \cdot |B|}{|B|} \implies |\lambda| \leq |A|$$

e quindi $\rho(A) \leq |A|$. □

Questa misura potrebbe non essere molto utile in quanto la norma $|A|$ potrebbe essere arbitrariamente più grande del raggio spettrale $\rho(A)$

1.1.4 Teoremi di Gershgorin

Esistono alcuni teoremi più forti riguardo alle norme degli autovalori, che prendono il nome di **teoremi di Gershgorin**.

Iniziamo dalla definizione di *cerchio* di Gershgorin:

Definizione 1.5: Cerchio di Gershgorin

Data $A \in \mathbb{C}^{n \times n}$ definiamo gli insiemi:

$$\mathcal{F}_i(A) = \{z \in \mathbb{C} : |z - a_{ii}| \leq \rho_i\}, \quad \rho_i = \sum_{j \neq i}^n |a_{ij}| \geq 0$$

cerchi di Gershgorin di A associati alla riga i .

Questi sono effettivamente i cerchi di raggio uguale alla somma degli elementi fuori dalla diagonale, con centro sull'asse dei reali il valore presente sulla diagonale, per ogni riga.

Possiamo quindi enunciare i teoremi di Gershgorin:

Teorema 1.3: di Gershgorin 1

Se $\lambda \in \mathbb{C}$ è autovalore vale che:

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

cioè gli autovalori stanno nell'unione dei cerchi di Gershgorin. Per dimostrare questo risultato si prende l'autocoppia di A (λ, x) e si indica con k l'indice della componente di massimo modulo in x . Quindi $x \neq 0$ perchè x autovettore, e k è scelto tale che $|x_k| > |x_j|$, $\forall j = 1, \dots, n$. Sappiamo quindi che vale $Ax = \lambda x$ dalla definizione di autovalore, e consideriamo la riga k nell'uguaglianza:

$$a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kn}x_n = \lambda x_k \Leftrightarrow \sum_{j=1}^n a_{kj}x_j = \lambda x_k \Leftrightarrow \sum_{j \neq k}^n a_{kj}x_j = \lambda x_k - a_{kk}x_k = (\lambda - a_{kk})x_k$$

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

Dato $|x_k| > 0$, si divide da entrambi i lati per $|x_k|$:

$$\Rightarrow |\lambda - a_{kk}| \leq \sum_{j \neq k}^n |a_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{j \neq k}^n |a_{kj}| = \rho_k$$

Quindi:

$$|\lambda - a_{kk}| \leq \rho_k \Rightarrow \lambda \in \mathcal{F}_k(A)$$

e, dato che l'argomento utilizzato vale per un qualsiasi autovalore, abbiamo la tesi del teorema:

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

□

Osserviamo che in realtà avremmo dimostrato che $\lambda \in \mathcal{F}_k(A)$ con k indice dell'elemento di massimo modulo nell'autovettore. Tuttavia, nella situazione usuale in cui non conosciamo né λ né x , sapere k non è plausibile.

Definizione 1.6: Matrice a predominanza diagonale forte

Una matrice A si dice a predominanza diagonale forte se vale:

$$|a_{ii}| > \sum_{j \neq i}^n |a_{ij}| = \rho_i, \quad \forall i = 1, \dots, n$$

Un **corollario** del teorema 6.3 è che se A è a predominanza diagonale forte, allora A è non singolare ($\det(A) \neq 0$).

Questo si dimostra dal fatto che tutti i cerchi di Gershgorin di A non comprendono l'origine, e quindi:

$$0 \notin \bigcup_{i=1}^n \mathcal{F}_i(A)$$

e 0 non può essere autovalore di A .

Teorema 1.4: di Gershgorin 2

Se M_1 è unione di s cerchi di Gershgorin ed M_2 è unione di $n - s$ cerchi di Gershgorin e vale $M_1 \cap M_2 = \emptyset$, allora M_1 contiene esattamente s autovalori e M_2 ne contiene esattamente altri $n - s$.

Osserviamo che il primo teorema di Gershgorin (6.3) non ci dice che necessariamente c'è un autovalore per ogni cerchio $\mathcal{F}_i(A)$, cioè può succedere che $\mathcal{F}_i(A)$ non ne contiene nemmeno uno. Il secondo teorema di Gershgorin (6.4) invece ci dice che se abbiamo un cerchio isolato, allora questo contiene esattamente un autovalore.

Un **corollario** del secondo teorema di Gershgorin è che se A ha n cerchi disgiunti, allora A ha n autovalori distinti ed è diagonalizzabile.

1.1.5 Matrici reali e Gershgorin

Ci sono delle note da fare sulle **matrici reali**. Osserviamo infatti che se $A \in \mathbb{R}^{n \times n}$, se λ è autovalore di A allora lo è anche $\bar{\lambda}$. Quindi, se un cerchio di Gershgorin contiene un autovalore complesso, dovrà necessariamente contenere anche il suo coniugio. Ora, visto che un cerchio isolato contiene necessariamente un solo autovalore, sarà che un cerchio isolato in una matrice a entrate reali contiene necessariamente un autovalore reale, e quindi la matrice ha necessariamente un autovalore reale.

1.1.6 Matrici trasposte e Gershgorin

Dato che A e A^T hanno gli stessi autovalori (dallo stesso polinomio caratteristico), si può prendere come regione dove stanno gli autovalori l'intersezione:

$$\left[\bigcup_{i=1}^n \mathcal{F}_i(A) \right] \cap \left[\bigcup_{i=1}^n \mathcal{F}_i(A^T) \right]$$

dove notiamo bene che:

$$\left[\bigcup_{i=1}^n \mathcal{F}_i(A) \right] \cap \left[\bigcup_{i=1}^n \mathcal{F}_i(A^T) \right] \neq \bigcup_{i=1}^n (\mathcal{F}_i(A) \cap \mathcal{F}_i(A^T))$$

infatti solitamente il termine a sinistra è ben più grande. Osserviamo poi che:

$$\mathcal{F}_I(A^T) = \{z \in \mathbb{C} : |z - a_{ii}| \leq \rho_i\}, \quad \rho_i = \sum_{j \neq i}^n |a_{ij}| \geq 0$$

cioè semplicemente si prendono le colonne anziché le righe.

1.1.7 Note sul calcolo dei dischi di Gershgorin

Vediamo come si può usare MATLAB per il calcolo e la visualizzazione dei dischi di Gershgorin di una matrice.

Potremmo pensare di scrivere una funzione per il calcolo e la rappresentazione grafica dei cerchi di Gershgorin di una matrice arbitraria (sia reale che complessa), ad esempio come:

```

1 function draw_gersh(A, draw_eigen)
2     % draws a circle
3     function circle(center, radius, col)
4         t = linspace(0, 2*pi);
5         patch(radius * cos(t) + real(center), ...
6               radius * sin(t) + imag(center), col);
7     end
8
9     if nargin < 2
10         draw_eigen = true;

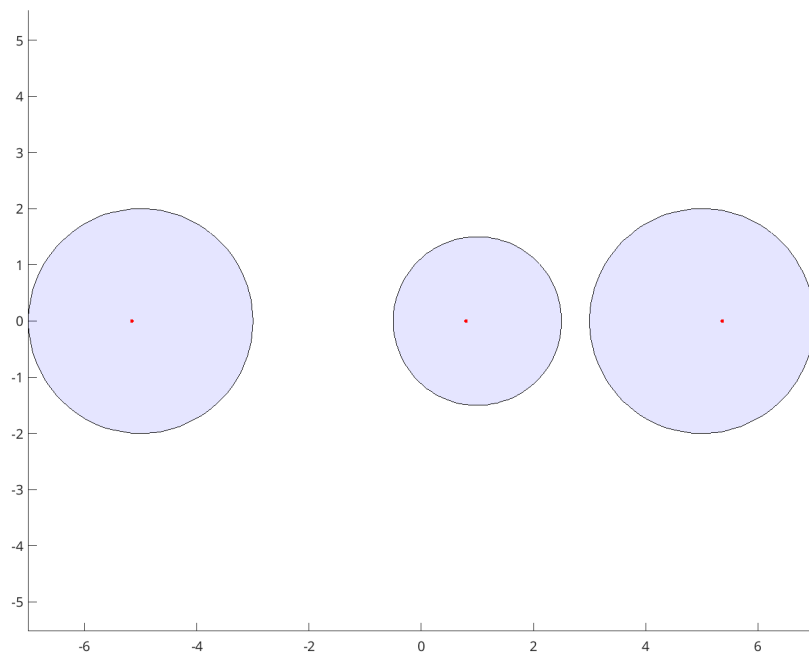
```

```
11 end
12
13 n = size(A, 1);
14
15 % graph setup
16 close all;
17 hold on;
18 axis('equal');
19
20 for k = 1:n
21     center = A(k, k);
22     radius = sum(abs(A(k, [1:k-1, k+1:end])));
23     circle(center, radius, 'blue');
24 end
25
26 alpha(.1)
27
28 if draw_eigen
29     eigenvals = eig(A);
30     plot(real(eigenvals), imag(eigenvals), 'red.');
```

A questo punto potremo tracciare i cerchi di Gershgorin come:

```
1 >> A = [5, 1, 1; 1, -5, 1; 1, 0.5, 1];
2 >> draw_gersh(A)
```

da cui si ottiene il grafico:



Notiamo come, nel caso specifico, a 3 cerchi di Gershgorin distinti corrispondono 3 autovalori distinti, e quindi la matrice A è diagonalizzabile.