RICHIAMI DI ALGEBRA LINEARE E NORME DI MATRICI E VETTORI

LUCIA GASTALDI

1. Matrici. Operazioni fondamentali.

Una **matrice** A è un insieme di $m \times n$ numeri reali (o complessi) ordinati, rappresentato nella tabella rettangolare seguente:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Osserviamo che a_{ij} indica l'elemento di A che si trova sulla i-esima riga e sulla j-esima colonna. La notazione $A \in \mathbb{R}^{m\times n}$ significa che la matrice A ha m righe ed n colonne di elementi reali. Nel caso in cui $m = n$ la matrice si dice **quadrata** di ordine n. Se $n = 1$ abbiamo un vettore colonna che indicheremo con $x \in \mathbb{R}^m$, le cui componenti sono indicate con x_i.

Date due matrici della stessa dimensione $A \in \mathbb{R}^{m\times n}$ e $B \in \mathbb{R}^{m\times n}$, la matrice $C \in \mathbb{R}^{m\times n}$ di elementi $c_{ij} = a_{ij} + b_{ij}$ per $i = 1, \cdots, m$ e $j = 1, \cdots, n$ è la matrice **somma delle matrici** A e B. La matrice che ha tutti gli elementi nulli, indicata con $\mathbf{0}$, è l'elemento neutro rispetto alla somma di matrici. La matrice B di elementi $b_{ij} = -a_{ij}$ sommata ad A dà la matrice nulla.

Dato un numero reale α, il prodotto di A per lo scalare α è dato dalla matrice $C \in \mathbb{R}^{m\times n}$ di elementi $c_{ij} = \alpha a_{ij}$.

Pertanto l'insieme delle matrici in $\mathbb{R}^{m\times n}$ gode della struttura di spazio lineare.

Date due matrici $A \in \mathbb{R}^{m\times p}$ e $B \in \mathbb{R}^{p\times n}$, la **matrice prodotto** C ha dimensione $m \times n$ ed i suoi elementi si ottengono mediante la seguente relazione

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

Quindi per moltiplicare tra loro due matrici devono essere "compatibili", nel senso che la prima deve avere un numero di colonne pari al numero delle righe della seconda. Si può dimostrare che il prodotto di matrici gode della proprietà associativa e distributiva rispetto alla somma, ma non della proprietà commutativa.

Il prodotto di due matrici quadrate di ordine n è ancora una matrice quadrata di ordine n, quindi in $\mathbb{R}^{n\times n}$ l'operazione prodotto è ben definita. La **matrice identità** data da

$$I = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}$$

è l'elemento neutro rispetto al prodotto in $\mathbb{R}^{n\times n}$. Ossia se $A \in \mathbb{R}^{n\times n}$, allora $AI = IA = A$.

Definizione 1. Data una matrice $A \in \mathbb{R}^{m\times n}$, chiamiamo **trasposta** di A la matrice $A^T \in \mathbb{R}^{n\times m}$ che si ottiene scambiando tra loro le righe e le colonne di A.

Valgono le seguenti proprietà:

a. $(A^T)^T = A$;
b. $(A + B)^T = A^T + B^T$;
c. $(AB)^T = B^T A^T$;
d. $(\alpha A)^T = \alpha A^T$ per ogni $\alpha \in \mathbb{R}$.

Definizione 2. Una matrice $A \in \mathbb{R}^{n\times n}$ si dice **simmetrica** se $A = A^T$.
Se vale che $A^T A = AA^T = I$ allora la matrice si dice **ortogonale**.

Definizione 3. Sia $A \in \mathbb{C}^{m\times n}$, la matrice $B = A^H \in \mathbb{C}^{n\times m}$ è detta **matrice coniugata trasposta** o **aggiunta** di A se vale $b_{ij} = \bar{a}_{ji}$ essendo $\bar{a}_{ji}$ il numero complesso couniugato di A_{ji}.

Valgono le seguenti proprietà:

a. $(A^H)^H = A$;
b. $(A + B)^H = A^H + B^H$;
c. $(AB)^H = B^H A^H$;
d. $(\alpha A)^H = \bar{\alpha} A^H$ per ogni $\alpha \in \mathbb{C}$.

Definizione 4. Una matrice $A \in \mathbb{C}^{n\times n}$ si dice **hermitiana** se $A = A^H$.
Se vale che $A^H A = AA^H = I$ allora la matrice si dice **unitaria**.

2. Matrice inversa

Definizione 5. Data una matrice $A \in \mathbb{R}^{n\times n}$, se esiste una matrice $X \in \mathbb{R}^{n\times n}$ tale che $AX = XA = I$, allora A si dice **invertibile** e la matrice X viene detta **matrice inversa** di A e viene indicata con A^{-1}. Una matrice non invertibile viene detta **singolare**.

Se A è invertibile anche la sua inversa A^{-1} lo è, e vale $(A^{-1})^{-1} = A$. Inoltre anche la trasposta di A è invertibile e vale $(A^T)^{-1} = (A^{-1})^T = A^{-T}$. Infine se A e B sono due matrici quadrate di ordine n invertibili, anche il prodotto AB è invertibile e vale $(AB)^{-1} = B^{-1}A^{-1}$.

3. Determinante di una matrice

Definizione 6. Dicesi **determinante** di una matrice quadrata A di ordine n il numero

$$\det(A) = \sum_{\boldsymbol{\pi} \in P} \operatorname{sign}(\boldsymbol{\pi}) a_{1\pi_1} a_{2\pi_2} \cdots a_{n\pi_n},$$

dove $P = \{\boldsymbol{\pi} = (\pi_1, \cdots, \pi_n)^T\}$ è l'insieme degli $n!$ vettori ottenuti permutando il vettore $\mathbf{i} = (1, 2, \cdots, n)^T$ dei primi n interi e $\operatorname{sign}(\boldsymbol{\pi})$ è uguale a 1 (rispettivamente -1) se $\boldsymbol{\pi}$ si ottiene da $\mathbf{i}$ mediante un numero pari (rispettivamente dispari) di scambi.

Valgono le seguenti proprietà:

a. $\det(A^T) = \det(A)$;
b. $\det(AB) = \det(A)\det(B)$;
c. $\det(A^{-1}) = 1/\det(A)$;
d. $\det(\alpha A) = \alpha^n \det(A)$, per ogni $\alpha \in \mathbb{R}$.

Indichiamo con A_{ij} la matrice di ordine $n-1$ ottenuta sopprimendo la i-esima riga e la j-esima colonna, e diciamo **minore complementare** associato all'elemento a_{ij} il determinante della matrice A_{ij}. Chiamiamo **minore principale** k-esimo di A, il determinante della sottomatrice principale A_k di ordine k che si ottiene sopprimendo le ultime $n-k$ righe e colonne di A. Posto $\Delta_{ij} = (-1)^{i+j}\det(A_{ij})$, il **cofattore** dell'elemento a_{ij}, il determinante può essere ottenuto con la seguente relazione ricorsiva:

$$\det(A) = \begin{cases} a_{11} & \text{se } n = 1 \\ \displaystyle\sum_{j=1}^{n} \Delta_{ij} a_{ij} & \text{per } n > 1, \end{cases}$$

detta *formula di sviluppo del determinante secondo la i-esima riga e la j-esima colonna* o *regola di Laplace*.

Data una matrice $A \in \mathbb{R}^{m\times n}$, consideriamo le sottomatrici quadrate di ordine q che si ottengono sopprimendo in modo qualunque $m-q$ righe ed $n-q$ colonne. Il corrispondente determinante si dice **minore di ordine** q.

Definizione 7. Dicesi **rango** di una matrice l'ordine massimo dei minori non nulli della matrice. Esso viene indicato con $\operatorname{rank}(A)$. La matrice ha rango massimo o pieno se $\operatorname{rank}(A) = \min(m, n)$.

Il rango esprime il massimo numero di vettori colonna linearmente indipendenti di A.

Dato un vettore $x \in \mathbb{R}^n$, il vettore prodotto $y = Ax \in \mathbb{R}^m$ ha componenti date dalla formula

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad \text{per } i = 1, \cdots, m.$$

Osserviamo che, in questa formula, la componente j-esima del vettore x moltiplica gli elementi di A che si trovano sulla j-esima colonna. Indichiamo con $\mathbf{a}_j \in \mathbb{R}^m$ il vettore colonna dato dalla j-esima colonna di A. Allora il vettore $y = Ax$ può essere ottenuto come combinazione lineare dei vettori colonna di A come segue:

$$y = Ax = \sum_{j=1}^{n} x_j \mathbf{a}_j. \tag{1}$$

Consideriamo lo *spazio immagine* o *range* di A

$$\text{range}(A) = \{y \in \mathbb{R}^m : y = Ax \text{ per } x \in \mathbb{R}^n\}. \tag{2}$$

Da (1) si ricava che lo spazio immagine di A coincide con lo spazio vettoriale formato da tutte le combinazioni lineari delle colonne di A ossia

$$\text{range}(A) = \text{span}\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n\}.$$

Di conseguenza si ha $\text{rank}(A) = \dim(\text{range}(A))$.

L'insieme

$$\ker(A) = \{x \in \mathbb{R}^n : Ax = 0\} \tag{3}$$

si dice **nucleo** di A.

Valgono le seguenti relazioni:

$$\text{rank}(A) = \text{rank}(A^T) \tag{4}$$

$$\text{rank}(A) + \dim(\ker(A)) = n. \tag{5}$$

In particolare, se $m = n$ ed A è non singolare, si ha che $\det(A) \neq 0$, $\text{rank}(A) = n$ e $\ker(A)) = \{\mathbf{0}\}$.

Un caso che si presenta abbastanza frequentemente nelle applicazioni è quello delle matrici simmetriche e definite positive.

Definizione 8. Data una matrice $A \in \mathbb{R}^{n\times n}$ simmetrica essa è **definita positiva** se

$$x^T Ax > 0 \quad \text{per ogni } x \in \mathbb{R}^n \text{ con } x \neq \mathbf{0}. \tag{6}$$

Teorema 1. *(criterio di Sylvester) Una matrice simmetrica A di ordine n è definita positiva se e solo se*

$$\det(A_k) > 0, \quad \textit{per } k = 1, 2, \cdots, n,$$

essendo A_k il minore principale di ordine k.

4. Sistemi lineari

Dati una matrice $A \in \mathbb{R}^{m\times n}$ e un vettore $b \in \mathbb{R}^m$, il vettore $x \in \mathbb{R}^n$ si dice **soluzione** del **sistema lineare** di *matrice dei coefficienti* A e *termine noto* b se vale

$$Ax = b. \tag{7}$$

Dalla relazione (1) si vede che la condizione affinché esista $x \in \mathbb{R}^n$ soluzione di (7) è che

$$b \in \text{range}(A), \tag{8}$$

ossia che b sia combinazione lineare delle colonne di A. Questo può essere espresso in altri termini dicendo che la matrice A e la matrice *aumentata* $[A, b]$ hanno lo stesso rango, cioè $\text{rank}(A) = \text{rank}([A, b])$ (teorema di Rouchè-Capelli).

Quindi se la condizione (8) è soddisfatta *esiste* almeno un $x \in \mathbb{R}$ soluzione di (7). Per quanto riguarda l'unicità della soluzione si osserva che se $x_1 \in \mathbb{R}^n$ e $x_2 \in \mathbb{R}^n$ sono due soluzioni di (7) allora vale che

$$Ax_1 = b, \qquad \text{e} \qquad Ax_2 = b.$$

Sottraendo membro a membro si ottiene $Ax_1 - Ax_2 = A(x_1 - x_2) = 0$, quindi la differenza $x_1 - x_2$ è soluzione del *sistema omogeneo* $Ax = \mathbf{0}$ ed appartiene al nucleo di A. Quindi si ha unicità della soluzione se $\ker(A) = \{\mathbf{0}\}$. Dalla relazione fondamentale (5) si ha che $\ker(A) = \{\mathbf{0}\}$ equivale a $\text{rank}(A) = n$; se invece $\text{rank}(A) = r < n$ allora il sistema omogeneo ha infinite soluzioni. Più precisamente, ci sono $n-r$ soluzioni linearmente indipendenti, ogni altra soluzione si ottiene come combinazione lineare di esse.

In conclusione, se vale (8) e se si ha che $\text{rank}(A) = r < n$, allora la soluzione del sistema lineare $Ax = b$ è data da $x = y + x_0$ essendo $y \in \mathbb{R}^n$ una qualunque soluzione del sistema non omogeneo ($Ay = b$) e $x_0 \in \ker(A)$.

Nel caso in cui A sia un matrice quadrata di ordine n, la condizione $\text{rank}(A) = n$ garantisce che la condizione (8) sia sempre verificata e la relazione (5) implica che la soluzione del sistema lineare è unica. In questo caso, quindi, la condizione $\text{rank}(A) = n$ è equivalente a $\det(A) \neq 0$. Di conseguenza, la matrice A risulta invertibile e la soluzione del sistema lineare è espressa da $x = A^{-1}b$.

5. Norma di vettore

L'insieme dei vettori $x \in \mathbb{R}^n$ e l'insieme delle matrici quadrate di ordine n formano due spazi vettoriali, in quanto in essi è definita l'operazione di somma e il prodotto per uno scalare reale. Introduciamo ora il nuovo concetto di **norma** che servirà a determinare la distanza fra elementi dello spazio lineare, la convergenza di successioni a valori nello spazio lineare e la grandezza delle quantità di interesse.

Ad esempio, se si considerano le soluzioni x e $\tilde{x}$ dei due sistemi lineari seguenti:

$$Ax = b, \qquad A\tilde{x} = b + \delta b$$

ci si può chiedere quale influenza ha la perturbazione del termine noto δb sull'*errore* $x - \tilde{x}$. In questo caso vorremmo stabilire se l'errore è *piccolo* ogni qual volta δb è piccolo.

Per misurare la *grandezza* di un elemento in uno spazio vettoriale usiamo le norme.

Definizione 9. Sia V uno spazio vettoriale di elementi x. Dicesi **norma** un'applicazione $\|\cdot\| : V \to \mathbb{R}$ che gode delle seguenti proprietà:

1. $\|x\| \geq 0$ per ogni $x \in V$;
2. $\|x\| = 0$ se e solo se $x = 0$;
3. $\|\alpha x\| = |\alpha|\|x\|$, per ogni $x \in V$ e $\alpha \in \mathbb{R}$;
4. $\|x + y\| \leq \|x\| + \|y\|$ (disuguaglianza triangolare).

Sia V uno spazio vettoriale e $\|\cdot\|$ una norma definita su esso. Dati $x_0 \in V$ ed $r > 0$, si dice **sfera centrata** in x_0 di raggio r l'insieme

$$B(x_0, r) = \{x \in V : \|x - x_0\| < r\}. \tag{9}$$

Iniziamo ad introdurre oltre alla ben nota **norma euclidea** di vettore, altre due norme che potranno essere utili in seguito: dato $x \in \mathbb{R}^n$ si pone

$$\begin{aligned} \|x\|_1 &= \sum_{i=1}^{n} |x_i|; \\ \|x\|_2 &= \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}; \\ \|x\|_\infty &= \max_{1 \leq i \leq n} |x_i|. \end{aligned} \tag{10}$$

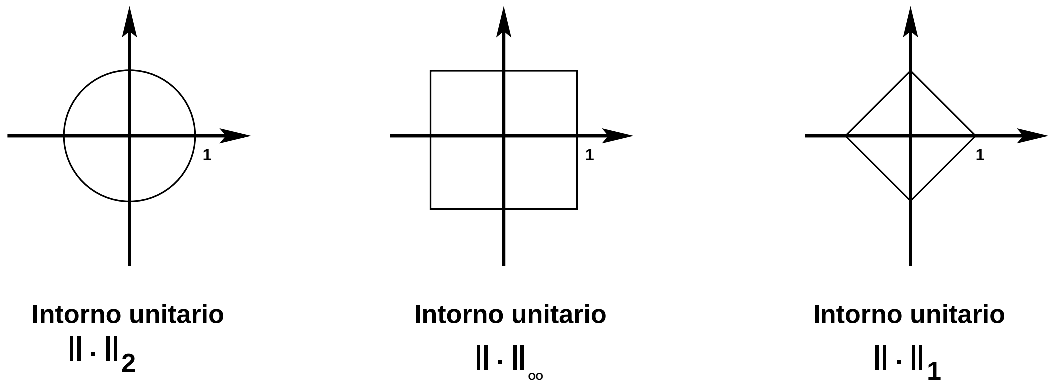

FIGURE 1. Intorni unitari nelle norme $\|\cdot\|_2$, $\|\cdot\|_\infty$ e $\|\cdot\|_1$.

Nella Fig. 1 sono disegnati gli intorni di raggio unitario nelle norme sopra introdotte. Le norme definite in (10) sono tra loro **equivalenti** nel senso che

$$\begin{aligned} &\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2; \\ &\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty. \end{aligned} \tag{11}$$

6. Numero di condizionamento di una matrice

Dati $A \in \mathbb{R}^{n\times n}$ e $b \in \mathbb{R}^n$ con A non singolare, sia $x \in \mathbb{R}^n$ la soluzione del sistema lineare $Ax = b$. Risolvendo il sistema lineare mediante un metodo numerico si otterrà un valore non esatto a causa della propagazione degli errori di macchina. Indichiamo con $\hat{x}$ la soluzione del sistema ottenuta e confrontiamola con quella esatta. Innanzitutto, sostituendo la soluzione $\hat{x}$ al posto di x nel sistema di partenza non otterremo una uguaglianza ma avremo una discrepanza tra il valore di $A\hat{x}$ e di b. Definiamo il residuo come il vettore $r \in \mathbb{R}^n$ dato dalla seguente espressione:

$$r = b - A\hat{x}. \tag{12}$$

Da questa relazione si vede che possiamo interpretare la soluzione approssimata $\hat{x}$ come la soluzione esatta del sistema

$$A\hat{x} = b - r, \tag{13}$$

in cui il termine noto è stato perturbato dalla quantità $-r$. Ci proponiamo quindi di studiare l'effetto di tale perturbazione sulla soluzione e quindi di valutare le quantità

$$\|x - \hat{x}\| \quad \textbf{errore assoluto}, \qquad \frac{\|x - \hat{x}\|}{\|x\|} \quad \textbf{errore relativo}.$$

Data una matrice A le seguenti quantità tengono conto dell'effetto che può avere la moltiplicazione di un vettore $x \in \mathbb{R}^n$ per la matrice A

$$M = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} \qquad m = \min_{x\neq 0} \frac{\|Ax\|}{\|x\|}. \tag{14}$$

In particolare M esprime di quanto può amplificarsi la norma del vettore x per effetto della moltiplicazione per A. Si osserva che le quantità in (14) tengono conto del fatto che una matrice è un'applicazione da $\mathbb{R}^n$ in sè.

Per calcolare l'errore assoluto $\|x - \hat{x}\|$, sottraiamo membro a membro le due relazioni $Ax = b$ e $A\hat{x} = b - r$ e otteniamo

$$Ax - A\hat{x} = r \quad \text{da cui } A(x - \hat{x}) = r.$$

Quindi passando alle norme si ha

$$\|r\| = \|A(x - \hat{x})\| = \frac{\|A(x - \hat{x})\|}{\|x - \hat{x}\|} \|x - \hat{x}\| \geq m\|x - \hat{x}\|$$

ossia

$$\|x - \hat{x}\| \leq \frac{1}{m}\|r\|.$$

Per valutare l'errore relativo osserviamo che

$$\|b\| = \|Ax\| \leq \frac{\|A\|}{\|x\|} \|x\| \leq M\|x\| \quad \text{da cui} \frac{1}{\|x\|} \leq \frac{M}{\|b\|},$$

quindi

$$\frac{\|x - \hat{x}\|}{\|x\|} \leq \frac{M}{m} \| \frac{\|r\|}{\|b\|}. \tag{15}$$

Definizione 10. Il numero

$$K(A) = \frac{M}{m} \tag{16}$$

viene detto **numero di condizionamento della matrice** A.

Ovviamente si ha che $K(A) \geq 1$. Se il numero di condizionamento è molto grande allora la matrice si dice malcondizionata e dalla relazione (15) si vede che l'errore relativo sulla soluzione può essere *molto grande* anche se il residuo è *piccolo*.

Si noti che il numero di condizionamento è una quantità intrinseca della matrice e non dipende dal procedimento di calcolo che ha portato al valore approssimato $\hat{x}$. Pertanto nella risoluzione numerica di un sistema lineare si deve tenere conto di questo fatto ed operare di conseguenza. Per migliorare il condizionamento di una matrice si può utilizzare una tecnica di *precondizionamento*. Nella relazione (15), il residuo deve essere interpretato come una perturbazione δb del termine noto, di conseguenza l'errore relativo rappresenta la risposta della risoluzione all'effetto della perturbazione introdotta. Se si introduce una perturbazione anche agli elementi della matrice, il sistema perturbato risulta della forma:

$$(A + \delta A)\hat{x} = b + \delta b,$$

essendo δA una *matrice di perturbazione*, vale il seguente risultato più generale:

$$\frac{\|x - \hat{x}\|}{\|x\|} \leq \frac{K(A)}{1 - K(A)\|\delta A\|/\|A\|} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right),$$

sotto la condizione che $1 - K(A)\|\delta A\|/\|A\| > 0$.

7. Norma di matrice

La definizione di *norma di matrice* è un po' più complessa in quanto si vuole che sia *compatibile* con una norma di vettore. Più precisamente data una norma di vettore ad essa si associa una norma di matrice in modo tale che valga la seguente relazione

$$\|Ax\| \leq \|A\|\|x\|.$$

Per ottenere la disuguaglianza desiderata si può definire una norma di matrice, detta **norma di matrice naturale associata alla norma di vettore** $\|\cdot\|$, come segue: per ogni matrice $A \in \mathbb{R}^{n\times n}$

$$\|A\| = \max_{x\in\mathbb{R}^n,\ x\neq\mathbf{0}} \frac{\|Ax\|}{\|x\|}. \tag{17}$$

Quindi la norma di matrice è proprio dato dal valore M definito in (14).

La norma così definita gode anche della *proprietà submoltiplicativa*:

$$\|AB\| \leq \|A\|\|B\|. \tag{18}$$

Si noti che $\|I\| = 1$ per ogni norma naturale di matrice associata ad una norma di vettore. Ad esempio la seguente *norma di Frobenius*

$$\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2\right)^{1/2}$$

non è una norma naturale di matrice come si verifica facilmente poiché $\|I\|_F = \sqrt{n}$.

Le norme naturali di matrice associate alle norme di vettore definite in (10) sono date da:

$$\begin{aligned}
\|A\|_1 &= \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|,\\
\|A\|_\infty &= \max_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|,\\
\|A\|_2 &= \sqrt{\rho(A^TA)}.
\end{aligned} \tag{19}$$

Valgono le seguenti disuguaglianze:

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1,\\
&\frac{1}{\sqrt{n}}\|A\|_2 \leq \|A\|_\infty \leq \sqrt{n}\|A\|_2,\\
&\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_2,\\
&\|Ax\|_2 \leq \|A\|_F\|x\|_2.
\end{aligned}$$

Le definizioni delle norme e le conseguenti disuguaglianze valgono anche nel caso di matrici rettangolari $A \in \mathbb{R}^{m\times n}$.

Infine osserviamo che anche la quantità m definita in (14) è legata ad una norma di matrice e precisamente si ha

$$\|A^{-1}\| = \frac{1}{m}.$$

Infatti siano $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^n$ legati dalla relazione

$$y = Ax \qquad\qquad x = A^{-1}y.$$

Per la definizione di norma si ha

$$\|A^{-1}\| = \max_{y\neq 0} \frac{\|A^{-1}y\|}{\|y\|} = \max_{x\neq 0} \frac{\|x\|}{\|Ax\|} = \max_{x\neq 0} \frac{1}{\dfrac{\|Ax\|}{\|x\|}} = \frac{1}{\min\limits_{x\neq 0} \dfrac{\|Ax\|}{\|x\|}} = \frac{1}{m}.$$

8. Autovalori ed autovettori

Definizione 11. Sia A una matrice quadrata di ordine n a coefficienti reali o complessi, il numero $\lambda \in \mathbb{C}$ si dice **autovalore** di A se esiste un vettore $x \in \mathbb{C}^n$ non nullo ($x \neq \mathbf{0}$) tale che

$$Ax = \lambda x. \tag{20}$$

Il vettore x è detto **autovettore** associato all'autovalore λ.

L'insieme degli autovalori della matrice A si chiama **spettro** di A e si indica con il simbolo $\sigma(A)$.

Dicesi **raggio spettrale** della matrice A il numero $\rho(A)$ dato dal massimo modulo degli autovalori e si pone

$$\rho(A) = \max_{\lambda\in\sigma(A)} |\lambda|.$$

Cerchiamo ora di caratterizzare gli autovalori di A. Dalla relazione (20) si ricava $(A - \lambda I)x = 0$. Quindi x è autovettore di A associato all'autovalore λ se appartiene al $\ker(A-\lambda I)$. Affinché $\ker(A - \lambda I)$ non sia ridotto al solo elemento nullo, la matrice $A - \lambda I$ deve essere singolare. Concludendo $\lambda \in \mathbb{C}$ è autovalore di A se vale

$$\det(A - \lambda I) = 0.$$

Quindi λ è autovalore della matrice A se e solo se è radice del *polinomio caratteristico* $p(\lambda) = \det(A - \lambda I)$. Dal teorema fondamentale dell'algebra segue che esistono esattamente n autovalori della matrice A se vengono contati tenendo conto della loro molteplicità. In particolare λ si dice autovalore semplice di A se λ è una radice semplice del polinomio caratteristico, inoltre si dice che λ ha *molteplicità* ν se λ è radice di molteplicità ν del polinomio caratteristico.

Osserviamo che se x è un autovettore di A associato all'autovalore λ anche αx, con $\alpha \in \mathbb{C}$, è autovettore di A associato all'autovalore λ, infatti vale:

$$A(\alpha x) = \alpha Ax = \alpha\lambda x = \lambda(\alpha x).$$

Inoltre uno stesso autovettore non può essere associato a due autovalori distinti. Infatti se $Ax = \lambda_1 x$ e $Ax = \lambda_2 x$ con $\lambda_1 \neq \lambda_2$, allora $(\lambda_1 - \lambda_2)x = \mathbf{0}$ e quindi necessariamente $x = \mathbf{0}$.

Si definisce **autospazio** associato all'autovalore λ lo spazio lineare

$$V(\lambda) = \{x \in \mathbb{C}^n : Ax - \lambda x = 0\}.$$

Si ha che la dimensione di $V(\lambda)$ è minore o uguale alla molteplicità $\nu(\lambda)$.

Definizione 12. Date due matrici $A \in \mathbb{C}^{n\times n}$ e $S \in \mathbb{C}^{n\times n}$, con $\det(S) \neq 0$, si consideri la matrice $B \in \mathbb{C}^{n\times n}$ data da $B = S^{-1}AS$. Si dice allora che A e B sono matrici **simili** e la trasformazione da A a B si dice **trasformazione per similitudine**.

Due matrici simili hanno lo stesso spettro e lo stesso polinomio caratteristico, infatti se λ è un autovalore di A e x è un autovettore associato, vale $Ax = \lambda x$. Moltiplicando a sinistra per la matrice S^{-1} si ricava

$$S^{-1}Ax = \lambda S^{-1}x, \text{ da cui segue } S^{-1}ASS^{-1}x = BS^{-1}x = \lambda S^{-1}x,$$

quindi λ è autovalore di B e $S^{-1}x$ è l'autovettore associato.

Definizione 13. Data $A \in \mathbb{C}^{n\times n}$, se esiste una matrice $X \in \mathbb{C}^{n\times n}$ invertibile tale che la matrice $\Lambda = X^{-1}AX$ è diagonale, allora la matrice A si dice **diagonalizzabile** e la matrice Λ contiene sulla sua diagonale tutti gli autovalori di A.

Se A è diagonalizzabile vale che $X^{-1}AX = \Lambda$ e moltiplicando a sinistra per X si ottiene $AX = X\Lambda$. Denotando con $\mathbf{x}_i$ la i-esima colonna di X, si ricava che $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$. Quindi le colonne di X sono gli autovettori di A. Poiché X ha rango pieno tutte le sue colonne sono linearmente indipendenti, ne consegue che gli autovettori di una matrice diagonalizzabile formano una base per $\mathbb{C}^n$.

Il caso della matrice $A \in \mathbb{R}^{n\times n}$ simmetrica è particolarmente importante. Infatti se A è simmetrica allora ha autovalori reali ed autovettori a due a due ortogonali. Se inoltre la matrice è anche definita positiva allora tutti gli autovalori sono positivi.

Allo scopo di calcolare numericamente gli autovalori è utile avere un'indicazione sulle regioni del piano complesso che li contengono.

Teorema 2. *(Teorema di Gershgorin) Sia A una matrice quadrata di ordine n. Per ogni $i = 1, 2, \cdots, n$ consideriamo i dischi R_i e C_i del piano complesso costruiti come segue:*

$$R_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1, j\neq i}^{n} |a_{ij}|\}$$
$$C_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1, j\neq i}^{n} |a_{ji}|\}.$$

Allora gli autovalori di A appartengono sia all'unione dei dischi R_i che a quella dei C_i ossia

$$\lambda \in \Big(\bigcup_{i=1}^{n} R_i\Big) \bigcap \Big(\bigcup_{i=1}^{n} C_i\Big).$$

Dipartimento di Matematica, Università di Brescia, Italy
E-mail address: `gastaldi@ing.unibs.it`
URL: `http://dm.ing.unibs.it/gastaldi/`