

1 ALGEBRA LINEARE NUMERICA

1.1 RICHIAMI DI ALGEBRA LINEARE

R1 (Vettori). Si ricordi che un numero complesso è del tipo $z = a + ib$, con a e b numeri reali ed $i^2 = -1$. Il suo **coniugato** è $\bar{z} = a - ib$, il suo **modulo** è $|z| = (z\bar{z})^{\frac{1}{2}} = (a^2 + b^2)^{\frac{1}{2}}$. Se z è reale ($b = 0$) $|z|$ è il suo **valore assoluto**. L'insieme dei numeri reali \mathbb{R} e quello dei numeri complessi \mathbb{C} sono dei campi di scalari. Fissato un numero intero positivo n , si indica con \mathbb{C}^n (\mathbb{R}^n) lo spazio dei vettori di n componenti complesse (reali).

Ricordiamo che se $x \in \mathbb{C}^n$ ed indichiamo con x_k , per $k = 1, 2, \dots, n$, le sue componenti, per ogni scalare $\alpha \in \mathbb{C}$ resta definito il vettore αx di componenti $(\alpha x)_k = \alpha x_k$, per $k = 1, 2, \dots, n$. Inoltre, dato un vettore $y \in \mathbb{C}^n$ di componenti y_k , per $k = 1, 2, \dots, n$, resta definito il vettore somma $z = x + y$ di componenti $z_k = x_k + y_k$, per $k = 1, 2, \dots, n$. In \mathbb{C}^n lo zero indica il vettore nullo ossia quello che ha tutte le componenti nulle. Inoltre, k vettori, $v^{(1)}, v^{(2)}, \dots, v^{(k)}$, si diranno **linearmente dipendenti** se esistono coefficienti c_1, c_2, \dots, c_k non tutti nulli tali che

$$\sum_{i=1}^k c_i v^{(i)} = 0.$$

Altrimenti si diranno **linearmente indipendenti**. Si ricordi che la dimensione di uno spazio lineare è il massimo numero di elementi linearmente indipendenti.

Siano $v^{(1)}, v^{(2)}, \dots, v^{(k)}$, $k (\leq n) \in \mathbb{C}^n$ vettori dati. Si indica con

$$\text{Span} \{v^{(1)}, v^{(2)}, \dots, v^{(k)}\}$$

l'insieme formato da tutte le loro combinazioni lineari a coefficienti complessi, ossia vettori del tipo $\sum_{i=1}^k c_i v^{(i)}$.

Se i vettori sono linearmente indipendenti allora $\dim \text{Span} \{v^{(1)}, v^{(2)}, \dots, v^{(k)}\} = k$.

Per ogni $j = 1, 2, \dots, n$, si indica con $e_{.j}$ il vettore avente la j -esima componente uguale a 1 e le altre tutte nulle. I vettori e_1, e_2, \dots, e_n sono linearmente indipendenti e costituiscono la **base canonica** di \mathbb{R}^n e di \mathbb{C}^n (ogni vettore è rappresentabile come una loro combinazione lineare).

R2 (Matrici). Lo spazio delle matrici di m righe ed n colonne, ossia matrici $m \times n$, con elementi tutti reali si indica con $\mathbb{R}^{m \times n}$. Quello delle analoghe matrici con elementi complessi si indica con $\mathbb{C}^{m \times n}$. L'elemento generico, di riga i -esima e colonna j -esima, di una matrice A si indica con a_{ij} . Sinteticamente, si usa la notazione $A = [a_{ij}]_{m \times n}$. Fra due matrici $A = [a_{ij}]_{m \times n}$ e $B = [b_{ij}]_{m \times n}$ è definita la **matrice somma**

$$A + B = [c_{ij}]_{m \times n} : c_{ij} = a_{ij} + b_{ij},$$

per ogni i e j . Per ogni scalare α si definisce

$$\alpha A = [\alpha a_{ij}]_{m \times n}.$$

Fra due matrici $A = [a_{ij}]_{m \times p}$ e $B = [b_{ij}]_{p \times n}$ si definisce la **matrice prodotto** (prodotto righe per colonne):

$$AB = [c_{ij}]_{m \times n} : c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}.$$

Si dice **matrice identità** (o matrice identica) di ordine n , e si indica con I (oppure con I_n), la matrice quadrata $n \times n$ le cui colonne sono i vettori della base canonica, ossia la colonna j -esima è il vettore e_j , per $j = 1, 2, \dots, n$. Si ha $IA = AI = A$, per ogni matrice quadrata A di ordine n , cioè $n \times n$.

Data una matrice A il massimo numero di righe (colonne) linearmente indipendenti è detto il **rango** di A e si indica con $rank(A)$. Una matrice $m \times n$ si dice di rango massimo se $rank(A) = \min\{m, n\}$. Lo spazio generato dalle sue colonne, cioè l'insieme di tutti i vettori del tipo Ax , è detto il **range** di A (si indica con $R(A)$). Dunque

$$\dim R(A) = rank(A).$$

Si dice **nucleo** e si indica con $\ker(A)$ di una matrice A l'insieme dei vettori x tali che $Ax = 0$, ossia se $A \in C^{m \times n}$,

$$\ker(A) = \{x \in C^n : Ax = 0\}.$$

Per ogni $A \in C^{m \times n}$, si ha

$$rank(A) + \dim \ker(A) = n.$$

Data una matrice $A = [a_{ij}]_{m \times n}$ si indica con A^T la sua **trasposta**, ottenuta scambiando le righe con le colonne, cioè $A^T = [a_{ji}]_{n \times m}$. Si indica con A^H la sua **trasposta coniugata**, ottenuta scambiando le righe con le colonne e coniugando gli elementi, cioè $A^H = [\bar{a}_{ji}]_{n \times m}$. Se A è reale allora $A^H = A^T$.

Vale l'identità

$$rank(A) = rank(A^T) = rank(A^H).$$

Per il prodotto fra matrici valgono le seguenti proprietà:

$$\begin{aligned} (AB)^T &= B^T A^T, \\ (AB)^H &= B^H A^H. \end{aligned}$$

Una matrice quadrata tale che $A = A^T$ si dice **simmetrica**, se $A = A^H$ si dice **hermitiana**. Dunque se A è reale e simmetrica essa è hermitiana.

Una matrice quadrata tale che $A = -A^T$ si dice **antisimmetrica**, se $A = -A^H$ si dice **antihermitiana**.

Convenzionalmente un vettore x di n componenti viene rappresentato come **vettore colonna**, cioè come una matrice di n righe ed 1 colonna. Resta dunque definito il **vettore trasposto** x^T ed il **trasposto coniugato** x^H di un vettore x .

Se $A = [a_{ij}]_{m \times n}$ e indichiamo con x_j , per $j = 1, 2, \dots, n$, le componenti di $x \in \mathbb{C}^n$, resta definito il vettore Ax di m componenti

$$(Ax)_i = \sum_{k=1}^n a_{ik}x_k, \text{ per } i = 1, 2, \dots, m.$$

Dunque dire che le colonne di A sono linearmente dipendenti equivale a dire che esiste un vettore non nullo y per cui $Ay = 0$.

R3 (Prodotto scalare) Nello spazio \mathbb{C}^n si definisce **prodotto scalare** una operazione, che indichiamo qui con $\langle \cdot, \cdot \rangle$, che associa ad ogni coppia di vettori x, y uno scalare (ossia un numero) tale che:

- 1) $\langle x, y \rangle = \overline{\langle y, x \rangle}$;
 - 2) $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$, per ogni $\alpha \in \mathbb{C}$;
 - 3) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$, per ogni $z \in \mathbb{C}^n$;
 - 4) $\langle x, x \rangle$ è un numero reale non negativo ed è nullo se e solo se $x = 0$.
- Due vettori tali che

$$\langle x, y \rangle = 0,$$

si dicono **ortogonali** (rispetto a questo prodotto scalare). Vettori ortogonali non nulli sono linearmente indipendenti.

Usualmente in \mathbb{C}^n si considera il **prodotto scalare euclideo** definito, secondo la regola del prodotto tra matrici, dalla formula

$$\langle x, y \rangle = x^H y = \sum_{i=1}^n \overline{x_i} y_i, \text{ per ogni } x, y \in \mathbb{C}^n,$$

dove x_i e y_i , per $i = 1, 2, \dots, n$, sono le componenti di x ed y . Osserviamo che, per ogni $A \in \mathbb{C}^{n \times n}$,

$$x^H A y = (A^H x)^H y.$$

Si dice anche che A^H è la matrice aggiunta di A rispetto al prodotto scalare euclideo. Se A è hermitiana essa è autoaggiunta rispetto al prodotto scalare euclideo ed avremo per ogni vettore x

$$x^H A x = (A x)^H x = \overline{x^H A x}.$$

Pertanto se A è hermitiana ogni prodotto scalare del tipo $x^H A x$ è un **numero reale**. Osserviamo che se A è antihermitiana ($A = -A^H$) si ha

$$x^H A x = -(A x)^H x = -\overline{x^H A x}.$$

Dunque se A è antihermitiana ogni prodotto scalare del tipo $x^H A x$ è un **numero immaginario**.

Se $x, y \in \mathbb{R}^n$ il prodotto scalare euclideo è

$$\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i$$

ed sarà $x^T y = y^T x$. Inoltre se $A \in \mathbb{R}^{n \times n}$ si ha

$$x^T A y = (A^T x)^T y$$

e se A è simmetrica allora $x^T A y = (A x)^T y$.

R4 (Tipologie di matrici). Una matrice quadrata $A = [a_{ij}]_{n \times n}$, si dice:

Diagonale se $a_{ij} = 0$ per $i \neq j$;

Triangolare inferiore se $a_{ij} = 0$ per $i < j$;

Triangolare inferiore in senso stretto se $a_{ij} = 0$ per $i \leq j$;

Triangolare superiore se $a_{ij} = 0$ per $i > j$;

Triangolare superiore in senso stretto se $a_{ij} = 0$ per $i \geq j$;

Tridiagonale se $a_{ij} = 0$, per $|i - j| > 1$.

Il prodotto di due matrici triangolari inferiori (superiori) è triangolare inferiore (superiore).

Una matrice quadrata tale che $A^H A = A A^H = I$, si dice **unitaria**.

In particolare, una matrice reale tale che $A^T A = A A^T = I$, si dice **ortogonale**.

OSSERVAZIONE. Le matrici unitarie conservano il prodotto scalare euclideo, cioè, se A è unitaria, si ha

$$(Ax)^H Ay = x^H A^H Ay = x^H y.$$

□

Una matrice quadrata A tale che $A A^H = A^H A$ si dice **normale**. Dunque le matrici hermitiane, le antihermitiane e le unitarie (ortogonali) sono matrici normali.

R5 (Determinante ed inversa). Ad ogni matrice A quadrata di ordine n si associa il numero $\det(A)$ detto il **determinante** di A . Esso può essere definito tramite la formula ricorsiva (regola di Laplace):

$$\det(A) = a_{11}, \text{ se } n = 1$$

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}), \text{ se } n > 1,$$

dove i è un qualsiasi indice di riga e, per ogni j , A_{ij} è la sottomatrice quadrata di ordine $n - 1$ ottenuta da A eliminando la i -esima riga e la j -esima colonna

Proprietà:

Se A è triangolare $\det(A) =$ prodotto elementi diagonali $(= \prod_{i=1}^n a_{ii})$,

$\det(AB) = \det(A) \det(B)$,

$\det(\alpha A) = \alpha^n \det(A)$.

$\det(A^T) = \det(A)$,

$\det(A^H) = \overline{\det(A)}$.

Se si scambiano tra loro due righe (o due colonne) il determinante cambia di segno.

Per ogni $k = 1, 2, \dots, n - 1$, si dirà **sottomatrice principale** (di testa) di ordine k la matrice quadrata ottenuta da A eliminando le ultime $n - k$ righe

e le ultime $n - k$ colonne. I determinanti di tali matrici si dicono i **minori principali** di A .

Una matrice quadrata A si dice **non singolare** se $\det(A) \neq 0$. Una matrice A è non singolare se e solo se le sue colonne sono linearmente indipendenti ($Ax = 0 \iff x = 0$).

Si definisce **inversa** di una matrice quadrata A la matrice, se esiste, indicata con A^{-1} , tale che

$$AA^{-1} = A^{-1}A = I.$$

Se esiste A^{-1} allora A si dirà **invertibile**. Una matrice quadrata A è invertibile se e solo se è **non singolare**. Se A è quadrata non singolare, per ogni vettore b esiste uno ed un solo vettore $x = A^{-1}b$ tale che

$$Ax = b.$$

L' inversa di una matrice triangolare inferiore (superiore) è triangolare inferiore (superiore).

Per il prodotto di due matrici quadrate si ha

$$(AB)^{-1} = B^{-1}A^{-1},$$

inoltre $(A^T)^{-1} = (A^{-1})^T$, $(A^H)^{-1} = (A^{-1})^H$.

R6 (Autovalori). Sia A una matrice quadrata di ordine n . Siano λ un numero complesso e x un vettore complesso non nullo tali che

$$Ax = \lambda x,$$

allora λ si dirà un **autovalore** di A ed x un **autovettore** (a destra) ad esso corrispondente. L'insieme degli autovalori costituisce lo **spettro** della matrice A . Per quanto detto al punto **R5**, λ è un autovalore di A se e solo se $A - \lambda I$ è singolare ossia se e solo se

$$\det(A - \lambda I) = 0, \text{ (equazione caratteristica).}$$

Il polinomio

$$p_A(\lambda) = \det(A - \lambda I),$$

è detto il **polinomio caratteristico** di A . Le sue n radici sono tutti e soli gli autovalori di A , che indicheremo con λ_i , per $i = 1, \dots, n$. La molteplicità di un autovalore λ_i come radice di $p_A(\lambda)$ è detta la **molteplicità algebrica** di λ_i .

Ad uno stesso autovalore possono corrispondere più autovettori. Se x_1, x_2, \dots, x_k sono tutti autovettori corrispondenti allo stesso autovalore allora anche ogni loro combinazione lineare lo è. Il massimo numero di autovettori linearmente indipendenti che corrispondono allo stesso autovalore è detta la **molteplicità geometrica** di tale autovalore, essa è dunque la dimensione del sottospazio generato da tali autovettori. Si ha sempre

$$\text{molteplicità geometrica} \leq \text{molteplicità algebrica.}$$

Un autovalore con molteplicità geometrica inferiore a quella algebrica è detto difettivo. Una matrice che ha autovalori difettivi è detta matrice difettiva.

TEOREMA. *Autovettori associati ad autovalori distinti sono linearmente indipendenti.*

□

Si ha

$$p_A(0) = \det(A) = \prod_{i=1}^n \lambda_i.$$

Inoltre

$$\sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i.$$

Questa quantità è detta **traccia** di A . Con il simbolo $\rho(A)$ si indica il **raggio spettrale** di A , definito come il *massimo dei moduli dei suoi autovalori*, cioè

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|.$$

Gli autovalori di A^{-1} sono i reciproci di quelli di A . Dunque

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

A è non singolare se e solo se zero non è un suo autovalore.

Gli autovalori delle potenze di A , ossia delle matrici che si ottengono moltiplicando A più volte per se stessa, sono le relative potenze degli autovalori di A . Infatti se $Ax = \lambda x$, allora $A^2x = AAx = \lambda Ax = \lambda^2x$ e così via..

Gli autovalori di A^T sono gli stessi di A , quelli di A^H sono i coniugati. Infatti per i polinomi caratteristici si ha

$$\frac{\det(A^T - \lambda I)}{\det(A - \lambda I)} = \frac{\det(A - \lambda I)}{\det(A - \lambda I)},$$

$$\frac{\det(A^T - \lambda I)}{\det(A - \lambda I)} = \frac{\det((A - \bar{\lambda}I)^H)}{\det(A - \lambda I)} = \frac{\det(A^H - \bar{\lambda}I)}{\det(A - \lambda I)}.$$

Data $A \in \mathbb{C}^{n \times n}$ ed un vettore non nullo $x \in \mathbb{C}^n$ si definisce il **quoziente di Rayleigh** :

$$\mu = \frac{x^H Ax}{x^H x}.$$

Se x è un autovettore di A allora esso corrisponde all' autovalore μ . Quindi in base a quanto detto al punto **R3**, gli **autovalori di una matrice hermitiana sono tutti reali**, e quelli di una **matrice antihermitiana sono tutti immaginari**.

TEOREMA. *Una matrice $n \times n$ hermitiana possiede n autovettori tra loro ortogonali.*

□

TEOREMA (**Courant-Fisher**). *Sia A hermitiana e siano $\lambda_n \leq \lambda_{n-1} \leq \dots \leq \lambda_1$ i suoi autovalori disposti in ordine decrescente. Vale il seguente risultato:*

$$\lambda_n = \min_{0 \neq x \in \mathbb{C}^n} \frac{x^H Ax}{x^H x}, \quad \lambda_1 = \max_{0 \neq x \in \mathbb{C}^n} \frac{x^H Ax}{x^H x}.$$

□

R7 (Matrici definite). Una matrice A reale e simmetrica è **definita positiva** se $x^T Ax > 0$ per ogni vettore reale x non nullo (**semi definita positiva** se vale \geq). Più in generale una matrice A hermitiana è **definita positiva** se $x^H Ax > 0$ per ogni vettore complesso x non nullo.

Dunque una matrice A reale e simmetrica (o hermitiana) è definita positiva se e solo se i suoi autovalori sono tutti positivi.

Inoltre necessariamente anche gli elementi diagonali sono positivi, infatti, per ogni $j = 1, 2, \dots, n$, si ha $a_{jj} = e_j^T A e_j > 0$ con $e_j = j$ -esimo vettore della base canonica.

TEOREMA (Criterio di Sylvester). Sia A quadrata reale e simmetrica di ordine n . Per $k = 1, 2, \dots, n$, indichiamo con A_k le sottomatrici principali (di testa) di ordine k di A . Allora A è definita positiva se e solo se $\det(A_k) > 0$ per ogni k , ossia se e solo se tutti i minori principali sono positivi.

□

Se A è reale simmetrica e definita positiva allora, per ogni coppia di indici i, j ($i \neq j$) vale la maggiorazione

$$|a_{ij}|^2 < a_{ii}a_{jj}.$$

Sia $A \in \mathbb{R}^{m \times n}$. Allora $A^T A$ è quadrata, simmetrica e semi-definita positiva. Infatti si ha $0 \leq (Ax)^T Ax = x^T A^T A x$, per ogni x .

TEOREMA. La matrice $A^T A$ è non singolare, quindi definita positiva, se e solo se le colonne di A sono linearmente indipendenti.

Dim. Supponiamo che $A^T A$ sia non singolare. Se le colonne di A fossero linearmente dipendenti allora esisterebbe un vettore $y (\neq 0)$ tale che $Ay = 0$ e dunque $A^T Ay = 0$, il che contraddice l'ipotesi.

Viceversa, supponiamo che le colonne di A siano linearmente indipendenti. Se $A^T A$ fosse singolare allora esisterebbe $y (\neq 0)$ tale che: $A^T Ay = 0$ e dunque $y^T A^T Ay = (Ay)^T Ay = 0$, cioè $Ay = 0$ il che contraddice l'ipotesi.

□

Le radici quadrate degli autovalori di $A^T A$ sono dette i **valori singolari** di A e si indicano con $\sigma_j = \sqrt{\lambda_j(A^T A)}$, $j = 1, 2, \dots, n$.

Nel caso che A sia a valori complessi, quanto detto sopra vale se si considera la trasposta coniugata e si sostituisce il termine "simmetrica" con "hermitiana".

R8 (Matrici simili). Due matrici quadrate A e B si dicono **simili** se esiste una matrice non singolare S tale che $S^{-1}AS = B$.

Due matrici simili hanno lo stesso polinomio caratteristico. Infatti

$$\begin{aligned} p_B(\lambda) &= \det(B - \lambda I) = \det(S^{-1}(A - \lambda I)S) \\ &= \det(S^{-1}) \det(A - \lambda I) \det(S) \\ &= \det(A - \lambda I) = p_A(\lambda). \end{aligned}$$

Una matrice quadrata A si dice **diagonalizzabile** se essa è simile ad una matrice diagonale, ossia se esiste una matrice non singolare S tale che $S^{-1}AS = D$ (diagonale).

TEOREMA. Una matrice A di ordine n è diagonalizzabile se e solo se possiede n autovettori linearmente indipendenti.

□

TEOREMA (Schur). Una qualunque matrice quadrata A può essere rappresentata nella forma

$$A = UTU^H$$

dove U è una matrice unitaria ($U^H U = I$) e T è una matrice triangolare superiore.

□

TEOREMA. Una matrice A è normale ($AA^H = A^H A$) se e solo se $A = UDU^H$ dove D è una matrice diagonale formata dagli autovalori. Gli autovettori corrispondenti ad autovettori distinti sono ortogonali.

□

R9 (Matrici elementari) Una matrice elementare reale è una matrice del tipo

$$H(\alpha, u, v) = I - \alpha uv^T,$$

dove $u, v \in \mathbb{R}^n$ ed $\alpha \in \mathbb{R}$. Si osservi che $\text{rank}(uv^T) = 1$. Dati $\alpha, \beta \in \mathbb{R}$,

$$H(\alpha, u, v)H(\beta, u, v) = H(\gamma, u, v),$$

con

$$\gamma = \alpha + \beta - \alpha\beta(v^T u).$$

Dunque se $\gamma = 0$, $H(\alpha, u, v)H(\beta, u, v) = I$. Pertanto se $\alpha(v^T u) - 1 \neq 0$, $H(\alpha, u, v)$ risulta invertibile e

$$H(\alpha, u, v)^{-1} = H(\beta, u, v),$$

con

$$\beta = \frac{\alpha}{\alpha(v^T u) - 1}.$$

Applichiamo questo risultato al calcolo dell' inversa di

$$B = A - zw^T,$$

con A invertibile e $z, w \in \mathbb{R}^n$. Poichè $B = A(I - A^{-1}zw^T)$, se la matrice elementare $(I - A^{-1}zw^T)$ è invertibile allora

$$B^{-1} = (I - A^{-1}zw^T)^{-1}A^{-1}.$$

Applicando quanto visto sopra (con $u = A^{-1}z$, $v = w$), nell' ipotesi che $w^T A^{-1}z - 1 \neq 0$, avremo che $(I - A^{-1}zw^T)$ è invertibile e si ottiene

$$B^{-1} = \left[I - \frac{A^{-1}zw^T}{w^T A^{-1}z - 1} \right] A^{-1} \quad (\text{formula di Scherman-Morrison}).$$

Si dice matrice di permutazione P una matrice ottenuta dalla matrice identità tramite permutazioni di righe (o colonne). Una matrice elementare di permutazione è del tipo

$$I_{i,j} = I - (e_i - e_j)(e_i - e_j)^T.$$

Premoltiplicando per $I_{i,j}$ si scambiano tra loro le righe i -esima e j -esima.

DEFINIZIONE. Una matrice quadrata A si dice **riducibile** se esiste una matrice di permutazione P tale che

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

con A_{11} e A_{22} quadrate.

□

ESEMPIO IMPORTANTE. Una matrice **tridiagonale** è irriducibile se e solo se tutti gli elementi della sotto-diagonale e della sopra-diagonale sono non nulli.

1.2 NORME DI VETTORI E DI MATRICI

DEFINIZIONE. Indichiamo con $\|\cdot\|$ una funzione da \mathbb{C}^n in \mathbb{R} tale che

- a) $\|x\| \geq 0$, per ogni $x \in \mathbb{C}^n$, e $\|x\| = 0$ se e solo se $x = 0$ (vettore nullo),
- b) $\|\alpha x\| = |\alpha| \|x\|$, $\alpha \in \mathbb{C}$, $x \in \mathbb{C}^n$,
- c) $\|x + y\| \leq \|x\| + \|y\|$, $x, y \in \mathbb{C}^n$.

Una tale funzione è detta **norma vettoriale**.

□

ESEMPLI. Le seguenti sono le norme vettoriali (in \mathbb{C}^n e in \mathbb{R}^n) più comunemente usate.

$$\|x\|_1 = \sum_{i=1}^n |x_i| \text{ (norma 1),}$$

$$\|x\|_2 = (x^H x)^{\frac{1}{2}} = \sqrt{\sum_{i=1}^n |x_i|^2} \text{ (norma 2 o norma euclidea),}$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \text{ (norma } \infty \text{ o norma del massimo).}$$

□

Queste norme sono casi particolari delle p -norme definite, per $p \geq 1$, dalla formula

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}.$$

Per la norma 2 vale la disuguaglianza di Cauchy-Schwarz:

$$|x^H y| \leq \|x\|_2 \|y\|_2.$$

TEOREMA. Ogni norma vettoriale è una funzione continua delle componenti del vettore stesso.

□

TEOREMA (equivalenza delle norme vettoriali). Siano $\|\cdot\|^*$ e $\|\cdot\|^{**}$ due norme vettoriali. Allora esistono due costanti positive α, β tali che per ogni x ,

$$\alpha \|x\|^{**} \leq \|x\|^* \leq \beta \|x\|^{**}.$$

□

DEFINIZIONE. Indichiamo con $\|\cdot\|$ una funzione tale che:

a) $\|A\| \geq 0$, per ogni matrice $A \in \mathbb{C}^{n \times n}$, e si ha $\|A\| = 0$ se e solo se $A = 0$ (matrice nulla),

b) $\|\alpha A\| = |\alpha| \|A\|$, $\alpha \in \mathbb{C}$, $A \in \mathbb{C}^{n \times n}$,

c) $\|A + B\| \leq \|A\| + \|B\|$, $A, B \in \mathbb{C}^{n \times n}$,

d) $\|AB\| \leq \|A\| \|B\|$, $A, B \in \mathbb{C}^{n \times n}$.

Una tale funzione si dirà una **norma matriciale**.

□

TEOREMA. Per ogni norma vettoriale $\|\cdot\|$, la funzione definita da

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|,$$

è una norma matriciale e si dirà una **norma indotta** (o naturale).

□

Si osservi che per ogni norma indotta $\|I\| = 1$ (I è la matrice identità).

Si osservi anche che, per la definizione di norma indotta, per ogni vettore x si ha

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

Questa relazione rappresenta la proprietà di **consistenza** fra una norma matriciale e una vettoriale. Dunque ogni norma matriciale indotta è **consistente** con la norma vettoriale che la induce.

Esempi. Le norme matriciali più comunemente utilizzate sono quelle indotte rispettivamente dalle norme vettoriali 1, 2 ed ∞ . Esse risultano definite come segue:

1) **norma matriciale 1:**

$$\|A\|_1 = \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |a_{ij}| \right).$$

2) **norma matriciale 2 o norma spettrale:**

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sqrt{\rho(A^H A)}$$

(ρ = raggio spettrale).

3) **norma matriciale ∞**

$$\begin{aligned}\|A\|_\infty &= \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} \\ &= \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |a_{ij}| \right).\end{aligned}$$

Si noti che le norme $\|A\|_1$ e $\|A\|_\infty$ sono direttamente calcolabili tramite gli elementi di A .

Dimostrazioni:

1) Dimostriamo che

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Sia x , con $\|x\|_\infty = 1$, tale che $\|A\|_\infty = \|Ax\|_\infty$. Allora avremo

$$\begin{aligned}\|A\|_\infty &= \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij} x_j \right| \\ &\leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.\end{aligned}$$

Consideriamo ora l'indice i^* tale che $\max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n |a_{i^*j}|$, quindi consideriamo il vettore x di componenti:

$$\begin{aligned}x_k &= \frac{\overline{a_{i^*k}}}{|a_{i^*k}|}, \text{ se } a_{i^*k} \neq 0, \\ &= 0 \quad \text{se } a_{i^*k} = 0.\end{aligned}$$

Ovviamente $\|x\|_\infty = 1$ (se $A \neq 0$) e avremo

$$|(Ax)_{i^*}| = \left| \sum_{j=1}^n a_{i^*j} \frac{\overline{a_{i^*j}}}{|a_{i^*j}|} \right| = \left| \sum_{j=1}^n |a_{i^*j}| \right|.$$

Pertanto

$$\begin{aligned}\max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| &= \sum_{j=1}^n |a_{i^*j}| \leq \|Ax\|_\infty \\ &\leq \|A\|_\infty.\end{aligned}$$

Vale dunque la doppia inclusione e quindi l'uguaglianza.
□

2) Dimostriamo che

$$\|A\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |a_{ij}| \right).$$

Sia x , con $\|x\|_1 = 1$, tale che $\|A\|_1 = \|Ax\|_1$. Allora avremo

$$\begin{aligned} \|A\|_1 &= \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{j=1}^n \sum_{i=1}^n |x_j| |a_{ij}| \\ &= \sum_{j=1}^n |x_j| \sum_{i=1}^n |a_{ij}| \\ &\leq \left(\sum_{j=1}^n |x_j| \right) \left(\max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \right) \\ &= \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |a_{ij}| \right). \end{aligned}$$

Vediamo il viceversa. Sia j^* tale

$$\sum_{i=1}^n |a_{ij^*}| = \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |a_{ij}| \right).$$

Prendiamo $x = e_{j^*}$. Avremo $\|A\|_1 \geq \|Ae_{j^*}\|_1 = \sum_{i=1}^n |a_{ij^*}|$.

3) Dimostriamo che

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

Ricordiamo che una matrice quadrata $n \times n$ hermitiana possiede n autovettori u_1, u_2, \dots, u_n ortogonali. Siano dunque u_1, u_2, \dots, u_n , tali che $u_i^H u_i = 1$ e $u_i^H u_j = 0$ se $j \neq i$, n autovettori ortogonali di $A^H A$, vale a dire che per ogni $k = 1, 2, \dots, n$ esiste un autovalore $\lambda_k \geq 0$ di $A^H A$ tale che

$$A^H A u_k = \lambda_k u_k.$$

Sia y , con $\|y\|_2 = 1$, tale che $\|A\|_2 = \|Ay\|_2$. Possiamo rappresentare y nella base u_1, u_2, \dots, u_n :

$$y = \sum_{j=1}^n \alpha_j u_j,$$

dove necessariamente $\sum_{j=1}^n |\alpha_j|^2 = 1$ (poichè $y^H y = 1$). Allora avremo

$$\begin{aligned} \|A\|_2^2 &= \|Ay\|_2^2 = (Ay)^H (Ay) = \\ &= \sum_{k=1}^n \overline{\alpha_k} u_k^H A^H A \sum_{j=1}^n \alpha_j u_j \\ &= \sum_{k=1}^n \lambda_k |\alpha_k|^2 \leq \max_{1 \leq k \leq n} \lambda_k \sum_{j=1}^n |\alpha_j|^2 \\ &= \max_{1 \leq k \leq n} \lambda_k. \end{aligned}$$

Dunque

$$\|A\|_2 \leq \sqrt{\rho(A^H A)}.$$

Sia ora $\lambda_s = \rho(A^H A)$ e sia u_s un autovettore normalizzato ad esso corrispondente, ossia $A^H A u_s = \lambda_s u_s$ con $\|u_s\|_2 = 1$.

Allora avremo

$$\begin{aligned} \|A u_s\|_2 &= [(A u_s)^H (A u_s)]^{\frac{1}{2}} \\ &= [u_s^H A^H A u_s]^{\frac{1}{2}} \\ &= \lambda_s^{\frac{1}{2}} = \sqrt{\rho(A^H A)} \\ &\leq \|A\|_2. \end{aligned}$$

Vale dunque la doppia inclusione e quindi l'uguaglianza.

□

OSSERVAZIONE. Se A è unitaria: $\|A\|_2 = \|A^H\|_2 = 1$.

□

TEOREMA (equivalenza delle norme matriciali). Siano $\|\cdot\|^*$ e $\|\cdot\|^{**}$ due

norme matriciali. Allora esistono due costanti positive α, β tali che per ogni matrice quadrata:

$$\alpha \|A\|^{**} \leq \|A\|^* \leq \beta \|A\|^{**}.$$

□

La funzione

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

è detta norma di **Frobenius**, ma **non** è una norma indotta ($\|I\|_F = \sqrt{n}$). Valgono le seguenti relazioni:

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2,$$

e

$$\|Ax\|_2 \leq \|A\|_F \|x\|_2,$$

cioè la norma di Frobenius è *consistente* con la norma euclidea.

TEOREMA. Il raggio spettrale di A è l'estremo inferiore dell'insieme delle norme indotte di A . Vale a dire che, per ogni matrice quadrata A :

- 1) $\rho(A) \leq \|A\|$, per ogni norma indotta;
- 2) per ogni $\varepsilon > 0$ esiste una norma indotta $\|\cdot\|^*$ tale che $\|A\|^* \leq \rho(A) + \varepsilon$.

□

Teoremi di Gershgorin.

TEOREMA. Sia A una matrice quadrata $n \times n$. Per $k = 1, 2, \dots, n$, definiamo i dischi di Gershgorin:

$$D_k = \left\{ z \in \mathbb{C} : |z - a_{kk}| \leq \sum_{j=1, j \neq k}^n |\alpha_{k,j}| \right\}.$$

Se λ è un autovalore di A allora λ appartiene alla riunione degli n dischi D_k .

Dim. Sia $Au = \lambda u$, con $\|u\|_\infty = 1$. Sia k tale che $|u_k| = \|u\|_\infty$. Poichè $(Au)_k = \sum_{j=1}^n \alpha_{k,j} u_j = \lambda u_k$ avremo

$$(\lambda - a_{kk})u_k = \sum_{j=1, j \neq k}^n \alpha_{k,j} u_j,$$

e quindi

$$|\lambda - a_{kk}| \leq \sum_{j=1, j \neq k}^n |\alpha_{k,j}|.$$

□

TEOREMA. Se l'unione M_1 di k dischi di Gershgorin è disgiunta dall'unione M_2 dei rimanenti $n - k$ dischi, allora k autovalori appartengono a M_1 e $n - k$ a M_2 .

□

TEOREMA. Sia A una matrice quadrata $n \times n$, irriducibile. Se un autovalore λ di A è situato sulla frontiera della riunione dei dischi D_k , allora tutti i dischi passano per λ .

□

1.3 CONDIZIONAMENTO DI SISTEMI LINEARI

Qui e nel seguito consideriamo una matrice A , reale quadrata $n \times n$. Assumendo che A sia invertibile, consideriamo il sistema lineare

$$Ax = b,$$

dove $b \in \mathbb{R}^n$ è il vettore dei termini noti e $x \in \mathbb{R}^n$ quello delle incognite. Dunque $x = A^{-1}b$.

Qui e nel seguito, indichiamo con $\|\cdot\|$ una norma vettoriale e la norma matriciale indotta.

Per prima cosa vogliamo analizzare il *condizionamento* di questo problema. Si ricordi che un problema si dirà "bene o male condizionato" se a piccole perturbazioni sui dati (in questo caso A e b) corrispondono piccole o grandi perturbazioni sui risultati x .

Per valutare il **condizionamento** del sistema cominciamo col perturbare il termine noto b , ossia consideriamo il sistema:

$$A(x + \Delta x) = b + \Delta b.$$

dove, rispetto al sistema di partenza, il vettore Δb rappresenta la perturbazione apportata al dato b e Δx rappresenta la conseguente perturbazione sulla soluzione.

Poiché $Ax = b$, avremo che

$$A\Delta x = \Delta b.$$

e dunque

$$\Delta x = A^{-1}\Delta b,$$

da cui, per la consistenza delle norme,

$$\|\Delta x\| \leq \|A^{-1}\| \|\Delta b\|.$$

Questa disuguaglianza mette in relazione gli errori assoluti. Consideriamo ora quelli relativi. Poiché

$$\|b\| \leq \|A\| \|x\|,$$

si ottiene la seguente relazione tra gli errori relativi

$$\frac{\|\Delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|}.$$

Il numero

$$K(A) = \|A\| \|A^{-1}\|$$

è detto **indice o numero di condizionamento** di A (nella particolare norma considerata).

OSSERVAZIONE. Dato un vettore z consideriamo il **vettore residuo**, del sistema di partenza, calcolato in z , ossia

$$r(z) = b - Az.$$

Dunque

$$Az = b - r(z).$$

Applicando il risultato di cui sopra, con $\Delta b = -r(z)$, si ottiene

$$\frac{\|z - x\|}{\|x\|} \leq K(A) \frac{\|r(z)\|}{\|b\|}.$$

Il che ci dice che se $K(A)$ è grande, ad un "piccolo" residuo può corrispondere un errore "grande".

□

LEMMA(di perturbazione). Sia ΔA una matrice tale che

$$\|\Delta A\| \|A^{-1}\| < 1,$$

allora la matrice $(A + \Delta A)$ è invertibile e

$$\|(A + \Delta A)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|\Delta A\| \|A^{-1}\|}.$$

Dim. Sia E una matrice tale che $\|E\| < 1$ allora $(I - E)$ è invertibile, poichè $\rho(E) < 1$. Inoltre, essendo

$$(I - E) \sum_{k=0}^n E^k = (I - E^{n+1}),$$

si ha

$$(I - E)^{-1} = \sum_{k=0}^{\infty} E^k \text{ (serie di Neumann).}$$

Da questa passando alle norme si ottiene

$$\|(I - E)^{-1}\| \leq (1 - \|E\|)^{-1}.$$

Applicando questi risultati alla matrice $(A + \Delta A) = A(I + A^{-1}\Delta A)$ si ottiene la tesi.

□

TEOREMA. Sia

$$\|\Delta A\| \|A^{-1}\| < 1.$$

Considerato il sistema

$$(A + \Delta A)(x + \Delta x) = b + \Delta b.$$

si ha

$$\frac{\|\Delta x\|}{\|x\|} \leq K(A) \frac{(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|})}{(1 - K(A) \frac{\|\Delta A\|}{\|A\|})}. \quad (1)$$

Dim. Per il Lemma di perturbazione la matrice $(A + \Delta A)$ è non singolare. Avremo inoltre

$$(A + \Delta A)\Delta x = \Delta b - (\Delta A)x,$$

ossia

$$\Delta x = (A + \Delta A)^{-1}(\Delta b - (\Delta A)x).$$

Quindi

$$\begin{aligned}\|\Delta x\| &\leq \|(A + \Delta A)^{-1}\| [\|\Delta b\| + \|\Delta A\| \|x\|] \\ &\leq \frac{\|A^{-1}\| [\|\Delta b\| + \|\Delta A\| \|x\|]}{1 - \|\Delta A\| \|A^{-1}\|}.\end{aligned}$$

Poichè $\|b\| \leq \|A\| \|x\|$, si ottiene

$$\frac{\|\Delta x\|}{\|x\|} \leq \|A^{-1}\| \frac{(\|\Delta A\| + \frac{\|\Delta b\| \|A\|}{\|b\|})}{(1 - \|\Delta A\| \|A^{-1}\|)},$$

e quindi la tesi.

□

Le relazioni di cui sopra ci indicano dunque che se il numero $K(A)$ (che dipende da A e dalla norma scelta) è grande, possiamo avere delle "grandi" perturbazioni sui risultati anche se quelle sui dati sono "piccole".

Si ha, per ogni norma indotta,

$$K(A) = \|A\| \|A^{-1}\| \geq \|AA^{-1}\| = \|I\| = 1.$$

Più precisamente, qualsiasi sia la norma indotta scelta, avremo

$$K(A) \geq \rho(A)\rho(A^{-1}) = \frac{\max |\lambda_i|}{\min |\lambda_i|},$$

dove, per $i = 1, 2, \dots, n$, i λ_i sono gli autovalori di A .

Si osservi che, se A è reale e simmetrica e si sceglie la norma 2, vale l'uguaglianza, cioè

$$\begin{aligned}K_2(A) &= \|A\|_2 \|A^{-1}\|_2 = \rho(A^T A)^{1/2} \rho((A^{-1})^T A^{-1})^{1/2} \\ &= \rho(A)\rho(A^{-1}) = \frac{\max |\lambda_i|}{\min |\lambda_i|}.\end{aligned}$$

1.4 SISTEMI TRIANGOLARI

I sistemi lineari a matrice triangolare sono di semplice risoluzione. Infatti, sia L una matrice **triangolare inferiore** di elementi l_{ij} ($l_{ij} = 0$, se $j > i$) non singolare, ossia con tutti gli elementi diagonali non nulli, allora il sistema

$$Ly = b,$$

si risolve tramite il seguente:

algoritmo di sostituzione in avanti:

$$y_1 = b_1/l_{11},$$

Se

$$a_{kk}^{(k)} \neq 0,$$

allora, dalla k -esima equazione possiamo ricavare l' espressione dell' incognita x_k in funzione delle x_{k+1}, \dots, x_n , cioè

$$x_k = \frac{b_k^{(k)} - \sum_{j=k+1}^n a_{kj}^{(k)} x_j}{a_{kk}^{(k)}},$$

ed inserire questa nelle successive equazioni (dalla $k+1$ -esima alla n -esima) che verranno così trasformate e non conterranno più i termini in x_k (assumendo così una forma più "triangolare"). In questo modo otteniamo un nuovo sistema $A^{(k+1)}x = b^{(k+1)}$.

Per realizzare questo poniamo

$$l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \text{ per } i = k+1, \dots, n. \quad (2)$$

e definiamo gli elementi del nuovo sistema $A^{(k+1)}x = b^{(k+1)}$ nel modo seguente:

$$\begin{aligned} a_{ij}^{(k+1)} &= a_{ij}^{(k)}, \text{ se } i \leq k, \\ a_{ij}^{(k+1)} &= a_{ij}^{(k)} - l_{i,k} a_{kj}^{(k)}, \text{ se } i > k, \\ b_i^{(k+1)} &= b_i^{(k)}, \text{ se } i \leq k, \\ b_i^{(k+1)} &= b_i^{(k)} - l_{i,k} b_k^{(k)}, \text{ se } i > k. \end{aligned}$$

□

Il procedimento completo dell' algoritmo di eliminazione richiede in generale circa $n^3/3$ operazioni aritmetiche.

L' elemento $a_{kk}^{(k)}$ è detto **pivot** (perno) della trasformazione al passo k .

Se $a_{kk}^{(k)} = 0$ l' algoritmo non si può realizzare a meno che non si ridefinisca il sistema $A^{(k)}x = b^{(k)}$ (prima di procedere come sopra) portando ad elemento pivot un elemento non nullo della matrice $A^{(k)}$ opportunamente scelto.

L' algoritmo di eliminazione gaussiana può presentare notevoli problemi di **instabilità numerica**. Tale instabilità può condurre, anche in casi molto semplici, a risultati del tutto errati.

Per ovviare a questo inconveniente, le analisi teoriche (vedi anche il paragrafo successivo) suggeriscono di introdurre sistematicamente (anche se $a_{kk}^{(k)} \neq 0$) una scelta opportuna dell' elemento pivot ad ogni passo, cioè di adottare sistematicamente una cosiddetta **strategia di pivoting**.

Le strategie più comuni sono:

a) **Pivoting parziale**. Per ogni k , si sceglie come elemento pivot al passo k quello di modulo massimo nella sottocolonna k -esima cioè: si determina l' indice i^* , compreso fra k ed n , per cui

$$\left| a_{i^*k}^{(k)} \right| = \max_{k \leq i \leq n} \left| a_{ik}^{(k)} \right|.$$

Quindi si scambia l' equazione k -esima di $A^{(k)}x = b^{(k)}$ con la i^* -esima; cioè si scambia la riga k -esima di $A^{(k)}$ con la i^* -esima e le relative componenti del vettore $b^{(k)}$.

b) **Pivoting totale.** Per ogni k , si sceglie come elemento pivot al passo k , quello di modulo massimo nella sottomatrice quadrata inferiore destra di $A^{(k)}$ di ordine $n - k + 1$, cioè si determinano gli indici i^* e j^* , compresi fra k ed n , per cui

$$\left| a_{i^*j^*}^{(k)} \right| = \max_{k \leq i, j \leq n} \left| a_{ij}^{(k)} \right|.$$

Quindi si scambia la riga k -esima di $A^{(k)}$ con la i^* -esima e le relative componenti del vettore $b^{(k)}$ ed inoltre la colonna k -esima di $A^{(k)}$ con la j^* -esima e le relative componenti del vettore x .

c). **Pivoting parziale scalato.** Si calcolano i valori

$$d_i = \max_{1 \leq j \leq n} |a_{ij}|, \text{ per } i = 1, 2, \dots, n.$$

Quindi, per ogni k , al passo k -esimo si determina l' indice i^* , compreso fra k ed n , per cui

$$\frac{\left| a_{i^*k}^{(k)} \right|}{d_{i^*}} = \max_{k \leq i \leq n} \frac{\left| a_{ik}^{(k)} \right|}{d_i}.$$

Quindi si eseguono le permutazioni come per il pivoting parziale. Questa strategia va applicata in particolare quando la matrice A possiede elementi molto diversi tra loro in valore assoluto.

Si verifica che se la matrice A è non singolare, l' algoritmo di eliminazione con una strategia di pivoting è sempre realizzabile.

1.6 FATTORIZZAZIONE TRIANGOLARE

TEOREMA (di fattorizzazione LU). Per $k = 1, 2, \dots, n - 1$, indichiamo

con A_k le sottomatrici principali (di testa) di ordine k di A . Se $\det(A_k) \neq 0$ per ogni $k = 1, 2, \dots, n - 1$, allora esiste la fattorizzazione

$$A = LU,$$

con L triangolare inferiore ed U triangolare superiore.

Inoltre, se A è non singolare, tale fattorizzazione è unica se imponiamo che L (oppure U) sia a diagonale unitaria (tutti gli elementi sulla diagonale = 1).

□

Se $A = LU$ avremo

$$A^{-1} = U^{-1}L^{-1}$$

ed anche

$$\det(A) = \det(L)\det(U).$$

Inoltre $Ax = b$ diventa $LUx = b$ e quindi la soluzione x si ottiene risolvendo in sequenza i sistemi triangolari

$$Ly = b, Ux = y. \quad (3)$$

TEOREMA. *Qualora sia realizzabile, ossia se $a_{kk}^{(k)} \neq 0$ per ogni $k = 1, 2, \dots, n-1$, l' algoritmo di eliminazione di Gauss, senza permutazioni, consente di ottenere la fattorizzazione $A = LU$, con L a diagonale unitaria. Precisamente L è la matrice di elementi $l_{i,k}$ (vedi formula (2)) mentre $U = A^{(n)}$.*

Dim. Introduciamo le matrici elementari di Gauss, definite per $k = 1, 2, \dots, n$, dalla formula

$$L_k = H(1, l_k, e_k) = I - l_k e_k^T,$$

dove e_k è il k -esimo vettore della base canonica di \mathbb{R}^n ed l_k è un vettore del tipo $l_k = (0, 0, \dots, 0, l_{k+1,k}, l_{k+2,k}, \dots, l_{n,k})^T$.

Osserviamo che, dato un vettore $a = (a_1, a_2, \dots, a_n)^T$, con $a_k \neq 0$, se poniamo $l_{i,k} = \frac{a_i}{a_k}$, per $i = k+1, \dots, n$, allora

$$L_k a = (a_1, a_2, \dots, a_k, 0, 0, \dots, 0)^T.$$

Notiamo inoltre che, per le proprietà delle matrici elementari, per ogni k si ha

$$L_k^{-1} = H(-1, l_k, e_k),$$

ed inoltre, per ogni coppia di indici i, j ($i < j$),

$$L_i L_j = I - l_i e_i^T - l_j e_j^T,$$

per cui

$$L_1^{-1} L_2^{-1} \dots L_k^{-1} \dots L_{n-1}^{-1} = I + l_1 e_1^T + l_2 e_2^T + \dots + l_{n-1} e_{n-1}^T.$$

Questa, come si può facilmente verificare, è una matrice triangolare inferiore a diagonale unitaria formata con gli $l_{i,k}$, per $i = k+1, \dots, n$.

Consideriamo ora l' algoritmo di eliminazione Gaussiana al generico passo k , cioè quando si passa da $A^{(k)}$ ad $A^{(k+1)}$, nell' ipotesi che $a_{kk}^{(k)} \neq 0$. Definendo come sopra

$$L_k = I - l_k e_k^T,$$

dove ora poniamo $l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$, per $i \geq k+1$, si ottiene

$$A^{(k+1)} = L_k A^{(k)},$$

e

$$b^{(k+1)} = L_k b^{(k)}.$$

Pertanto se l'algoritmo è realizzabile completamente, cioè fino a $k = n - 1$, avremo

$$U := A^{(n)} = L_{n-1} L_{n-2} \dots L_k \dots L_1 A,$$

ossia

$$A = (L_1^{-1} L_2^{-1} \dots L_k^{-1} \dots L_{n-1}^{-1}) U = LU,$$

con U triangolare superiore ed L triangolare inferiore a diagonale unitaria.

□

Vediamo ora che: **se, per $k = 1, 2, \dots, n - 1$, $\det(A_k) \neq 0$ (minori principali $\neq 0$) allora l'algoritmo di eliminazione è realizzabile.** In questo modo avremo provato il Teorema di fattorizzazione.

Infatti, pensiamo di essere giunti al passo k con

$$A^{(k)} = L_{k-1} L_{k-2} \dots L_1 A.$$

Diciamo $A_k^{(k)}$ la sottomatrice principale di ordine k di $A^{(k)}$. Si vede che (essendo tutte le matrici L_j a diagonale unitaria) :

$$\det(A_k^{(k)}) = \prod_{j=1}^k a_{jj}^{(j)} = \det(A_k) \neq 0.$$

Quindi poichè $a_{jj}^{(j)} \neq 0$, per $j = 1, \dots, k - 1$, dovrà essere anche $a_{kk}^{(k)} \neq 0$ e quindi l'algoritmo di eliminazione può procedere.

□

Anche quando si adotta una strategia di pivoting si realizza lo schema risolutivo di cui sopra, ma in questo caso è come se si operasse su un sistema ottenuto da quello di partenza attraverso le permutazioni effettuate. Ad esempio, nel caso del pivoting parziale, si realizza la fattorizzazione LU , con L a diagonale unitaria, di una matrice ottenuta da A permutandone opportunamente le righe. Ossia

$$LU = PA,$$

dove P indica la matrice di permutazione che opera il riordinamento delle righe di A in corrispondenza degli elementi pivot via via scelti. In definitiva, quando si opera col pivoting parziale è come se si operasse sul sistema equivalente

$$PAx = Pb.$$

Quanto ora detto ci consente di dare una giustificazione più precisa dell'introduzione delle strategie di pivoting nell'algoritmo di eliminazione.

Dovendo calcolare una fattorizzazione di matrice LU , supponiamo che nel farlo si commettano degli errori e che vengano in realtà calcolate due matrici

$$\bar{L} = L + \delta L, \bar{U} = U + \delta U,$$

per cui, anzichè LU , abbiamo ottenuto la fattorizzazione della matrice

$$\bar{L}\bar{U} = (L + \delta L)(U + \delta U) = LU + (\delta L)U + L(\delta U) + (\delta L)(\delta U).$$

Quindi, trascurando il termine $(\delta L)(\delta U)$, l'errore (assoluto) è

$$(\delta L)U + L(\delta U).$$

Anche se δL e δU sono di piccola entità, tale errore può essere molto grande se tali sono gli elementi delle matrici L ed U , e pertanto la matrice in realtà fattorizzata, ossia $\bar{L}\bar{U}$, può essere molto diversa da quella che si vuole fattorizzare.

Dunque bisogna evitare di calcolare fattorizzazioni con matrici L ed U arbitrariamente grandi. Pertanto, si dirà che una fattorizzazione LU è **numericamente stabile** se è possibile limitare convenientemente a priori gli elementi di L e di U . Si vede che ciò non è sempre realizzabile quando si usa l'algoritmo di eliminazione di Gauss senza pivoting e questa è proprio la causa della possibile instabilità numerica del metodo.

Invece le fattorizzazioni ottenute adottando le strategie di pivoting sono stabili. La strategia del pivoting totale rende in teoria l'algoritmo di Gauss più stabile rispetto a quella del pivoting parziale, ma quest'ultima viene solitamente preferita in quanto meno dispendiosa.

OSSERVAZIONE. Consideriamo il problema del calcolo della matrice inversa A^{-1} . Indicando con α_j , per $j = 1, \dots, n$, i vettori che compongono le colonne di A^{-1} , avremo che

$$A\alpha_j = e_j, \text{ per } j = 1, \dots, n, \text{ con } e_j = j\text{-esimo vettore della base canonica di } \mathbb{R}^n.$$

Dunque calcolare l'inversa è equivalente a risolvere questi n sistemi lineari.

□

Dal Teorema di fattorizzazione LU si ottiene il seguente

TEOREMA (di fattorizzazione LDM^T). Data una matrice reale A , se, per $k = 1, \dots, n$, $\det(A_k) \neq 0$ allora esistono due matrici triangolari inferiori a diagonale unitaria L ed M e una matrice diagonale D tali che

$$A = LDM^T.$$

Inoltre se A è simmetrica sarà $M = L$.

Dim. Per le ipotesi fatte esistono L triangolare inferiore a diagonale unitaria e U triangolare superiore tali che $A = LU$, inoltre, poichè per ipotesi $A (= A_n)$ è non singolare, anche U è non singolare. Sia D la matrice diagonale estratta da U , cioè $D = \text{diag}(u_{11}, u_{22}, \dots, u_{nn})$. Essa è invertibile. La matrice $M^T = D^{-1}U$ è triangolare superiore a diagonale unitaria e dunque $A = LDM^T$.

Supponiamo ora $A = A^T$. Allora da $A = LDM^T$ si ottiene che $M^{-1}AM^{-T} = M^{-1}LD$. Questa matrice è evidentemente simmetrica ed è triangolare inferiore (si ricordi che l'inversa di una triangolare inferiore è triangolare inferiore così

come il prodotto di due triangolari inferiori), pertanto essa è necessariamente diagonale. Quindi $M^{-1}L = M^{-1}LDD^{-1}$ è diagonale ma è anche a diagonale unitaria (perchè lo sono M ed L) e quindi necessariamente $M^{-1}L = I$.

□

1.7 METODO DI CHOLESKY

Il metodo di **Cholesky** è un metodo diretto (di fattorizzazione triangolare) per matrici simmetriche e definite positive. Si basa sul seguente risultato.

TEOREMA. *Sia A una matrice reale simmetrica e definita positiva. Allora A è fattorizzabile nella forma*

$$A = RR^T, \quad (4)$$

dove R è una matrice triangolare inferiore reale non singolare.

Dim. Valgono (criterio di Sylvester) le ipotesi del teorema di fattorizzazione LDM^T e pertanto $A = LDL^T$. Sia $D = \text{diag}(d_1, \dots, d_n)$. Dal fatto che A è definita positiva consegue facilmente che $d_i > 0$ per ogni i . Pertanto possiamo considerare la matrice diagonale $D^{1/2} = \text{diag}(d_1^{1/2}, \dots, d_n^{1/2})$. Avremo dunque $A = RR^T$ con $R = LD^{1/2}$.

□

Per il calcolo di R si può utilizzare il seguente:

Algoritmo di Cholesky.

Per $j = 1, \dots, n$:

$$r_{jj} = (a_{jj} - \sum_{k=1}^{j-1} r_{jk}^2)^{1/2} \quad (r_{11} = a_{11}^{1/2})$$

$$r_{i,j} = (a_{ij} - \sum_{k=1}^{j-1} r_{ik}r_{jk})/r_{jj}, \text{ per } i = j + 1, \dots, n \text{ e } j \neq n.$$

□

Il costo dell' algoritmo è di circa $\frac{n^3}{6}$ operazioni aritmetiche. Inoltre la fattorizzazione è numericamente stabile. Infatti si ha, per ogni i e j ,

$$r_{ij}^2 \leq a_{ii}.$$

Vale a dire che gli elementi di R sono limitabili a priori tramite quelli di A .

La risoluzione del sistema $Ax = b$ procederà dunque risolvendo

$$Ry = b, R^T x = y. \quad (5)$$

1.8 ANALISI DELL' ERRORE

Riportiamo qui di seguito alcuni risultati circa gli errori prodotti dagli algoritmi considerati. Tali risultati provengono dalla così detta analisi dell' errore all' indietro. Vale a dire che si interpreta la soluzione calcolata di un certo problema come la soluzione esatta di un problema con dati perturbati. Per risalire all'

errore ci si può quindi avvalere di una relazione di condizionamento come la (1).
 elle formule che seguono u indica la precisione di macchina.

1. Siano \bar{L} e \bar{U} le matrici triangolari calcolate con l' algoritmo di eliminazione
 Gaussiana con pivoting (parziale o totale). Allora, se L ed U sono le matrici
 teoriche da calcolare, si ha

$$\bar{LU} = LU + E,$$

dove la matrice E è tale che

$$\|E\|_{\infty} \leq k_1(n)u \|LU\|_{\infty},$$

dove

$$k_1(n) = n^2 g_n.$$

con:

- $g_n \leq 2^{n-1}$ per il pivoting parziale;
- $g_n \leq 1.8n^{0.25 \ln n}$ per il pivoting totale;
- $g_n \leq 1$ se A è simmetrica e definita positiva.

□

2. Siano \bar{L} e \bar{U} come sopra e sia \bar{x} la soluzione ottenuta numericamente
 applicando i relativi algoritmi di sostituzione (in avanti ed all' indietro) alla
 risoluzione dei sistemi triangolari con matrici \bar{L} e \bar{U} (vedi 3). Allora esiste una
 matrice ΔA tale che

$$(A + \Delta A)\bar{x} = b,$$

con

$$\|\Delta A\|_{\infty} \leq k_2(n)u \|A\|_{\infty}$$

dove

$$k_2(n) = (n^3 + 3n^2)g_n,$$

con g_n limitato come in precedenza.

□

3. Per il metodo di Cholesky (A simmetrica e definita positiva). Sia \bar{x} la
 soluzione calcolata eseguendo lo schema (4)-(5). Allora esiste una matrice ΔA
 tale che

$$(A + \Delta A)\bar{x} = b,$$

con

$$\|\Delta A\|_2 \leq c_n u \|A\|_2$$

dove c_n è un numero "piccolo" dipendente da n .

□

1.9 METODI ITERATIVI(cenni)

Lo scopo di questi metodi è quello di costruire successioni di vettori convergenti alla soluzione del sistema $Ax = b$. Un modo per costruire tali successioni è il seguente. Consideriamo una decomposizione (splitting) di A nella forma

$$A = M - N,$$

con M matrice invertibile. Allora il sistema si scrive

$$Mx = Nx + b,$$

cioè

$$x = M^{-1}Nx + M^{-1}b.$$

Per ogni vettore iniziale $x^{(0)}$ (vettore d' innesco) resta definito il processo iterativo

$$Mx^{(k)} = Nx^{(k-1)} + b, \text{ per } k = 1, 2, \dots,$$

ossia

$$x^{(k)} = M^{-1}Nx^{(k-1)} + M^{-1}b.$$

La matrice $P = M^{-1}N$ è detta **matrice di iterazione** del metodo.

Si tratta di valutare se $\lim_{k \rightarrow \infty} (x - x^{(k)}) = 0$, qualsiasi sia il vettore d' innesco $x^{(0)}$. Se ciò accade il metodo si dirà **convergente**.

TEOREMA. *Il metodo iterativo converge ad x , per ogni vettore d' innesco $x^{(0)}$, se e solo se*

$$\rho(P) < 1.$$

□

1.10 PRECONDIZIONAMENTO

In molti casi pratici, prima di applicare un qualsiasi metodo per la risoluzione di un sistema lineare risulta opportuno cercare di migliorarne il condizionamento tramite una tecnica di **precondizionamento**. Si tratta di ottenere un nuovo sistema che sia meglio condizionato di quello di partenza per poi applicare ad esso il metodo numerico scelto.

Per essere conveniente, la fase di preconditionamento non deve richiedere un elevato costo computazionale.

In generale si procede come segue. Consideriamo una matrice di preconditionamento C (che in qualche senso "assomigli" ad A) opportunamente scelta, espressa nella forma

$$C = C_1 C_2,$$

con C_1 e C_2 matrici invertibili e di forma semplice. Per prima cosa premoltiplichiamo il sistema di partenza $Ax = b$ per C_1^{-1} ottenendo

$$C_1^{-1}Ax = C_1^{-1}b.$$

Si noti che, se si potesse scegliere $C_1 = A$, questo sistema sarebbe *perfettamente* condizionato.

Posto quindi $y = C_2x$, si ottiene

$$C_1^{-1}AC_2^{-1}y = C_1^{-1}b.$$

Questo nuovo sistema sarà (sperabilmente) meglio condizionato di quello di partenza. Risolvendolo si otterrà x da $C_2x = y$. Se $C_2 = I$ si parla di preconditionamento solo a sinistra, analogamente se $C_1 = I$, di preconditionamento solo a destra.

1.11 MINIMI QUADRATI

PROBLEMA (dei minimi quadrati). Sia $A \in R^{m \times n}$, con $m \geq n$ e $b \in R^m$. Si cerca $x \in R^n$ che rende minima la norma euclidea in R^m del vettore residuo $r(x) = b - Ax$, vale a dire $x \in R^n$ tale che

$$\|b - Ax\|_2 \leq \|b - Ay\|_2, \text{ per ogni } y \in R^n.$$

Si dice che x è una *soluzione ai Minimi Quadrati* (del sistema sovradimensionato $Ax = b$).

□

TEOREMA. *Esiste sempre almeno una soluzione ai Minimi Quadrati.*

□

TEOREMA. *Se x è una soluzione del sistema*

$$A^T Ax = A^T b,$$

*detto delle **Equazioni Normali**, allora x è una soluzione ai Minimi Quadrati.*

Dim. Sia $x \in R^n$ tale che

$$A^T(b - Ax) = 0.$$

Per ogni $y \in R^n$ scriviamo

$$b - Ay = b - Ax + A(x - y),$$

per cui

$$\begin{aligned} \|b - Ay\|_2^2 &= [b - Ax + A(x - y)]^T [b - Ax + A(x - y)] \\ &= \|b - Ax\|_2^2 + 2[A(x - y)]^T [b - Ax] + \|A(x - y)\|_2^2 \\ &= \|b - Ax\|_2^2 + 2(x - y)^T A^T [b - Ax] + \|A(x - y)\|_2^2 \\ &\geq \|b - Ax\|_2^2. \end{aligned}$$

□

Si può provare che vale anche il viceversa, ossia che se $x \in R^n$ è una soluzione del problema allora necessariamente

$$A^T(b - Ax) = 0.$$

Questo significa che il corrispondente vettore residuo $r(x) = b - Ax$ è tale che $A^T r(x) = 0$, cioè $r(x)$ è ortogonale al Range di A . Si osservi che l'uguaglianza $Ax = b$ vale se e solo se b appartiene al Range di A . Inoltre la soluzione x è unica se e solo se $A^T A$ è non singolare, ossia se le colonne di A sono linearmente indipendenti.

Ricordiamo che il numero di condizionamento spettrale (in norma 2) di $A^T A$ è

$$K_2(A^T A) = \frac{\max \lambda_i(A^T A)}{\min \lambda_i(A^T A)}.$$

Dunque se $K_2(A^T A)$ è grande il sistema delle equazioni normali è mal condizionato. Bisognerà in tal caso cercare di evitare di risolvere tale sistema.

TEOREMA Una matrice $A \in R^{m \times n}$, $m \geq n$, di rango n , si può fattorizzare nella forma

$$A = QR,$$

dove $Q \in R^{m \times n}$ è tale che $Q^T Q = I$ ed $R \in R^{n \times n}$ è una triangolare superiore non singolare.

□

La fattorizzazione QR si può realizzare con circa mn^2 operazioni aritmetiche (tramite gli algoritmi di Householder e di Givens).

Ricordando che i valori singolari di una matrice $A \in R^{m \times n}$, sono

$$\sigma_i(A) = \sqrt{\lambda_i(A^T A)}, \text{ per } i = 1, 2, \dots, n,$$

possiamo estendere la definizione di numero di condizionamento spettrale alle matrici rettangolari definendo (se $A^T A$ è invertibile)

$$K_2(A) = \frac{\max \sigma_i(A)}{\min \sigma_i(A)}.$$

Dunque

$$K_2(A) = \sqrt{K_2(A^T A)}.$$

Osserviamo che se $A = QR$ avremo che $A^T A = R^T R$ e quindi

$$K_2(R) = K_2(A).$$

Tramite la fattorizzazione QR possiamo ottenere x risolvendo un sistema a matrice R . Infatti $A^T Ax = A^T b$ si può scrivere come $R^T Q^T QRx = R^T Q^T b$. Quindi il vettore x si ottiene risolvendo il sistema

$$Rx = Q^T b.$$

Il residuo è $b - AR^{-1}Q^T b = b - QQ^T b$.

1.12 CALCOLO DI AUTOVALORI

PROBLEMA: trovare $\lambda \in \mathbb{C}$ e un vettore $x \neq 0$ per cui

$$Ax = \lambda x.$$

Condizionamento degli autovalori.

TEOREMA (Bauer-Fike).

Introduciamo una norma matriciale indotta $\|\cdot\|$ tale che per ogni matrice diagonale $D = \text{diag}(d_1, d_2, \dots, d_n)$, si abbia

$$\|D\| = \max_{1 \leq i \leq n} |d_i|.$$

Sia A una matrice quadrata diagonalizzabile cioè

$$S^{-1}AS = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Considerata una matrice $A + E$, se λ è un autovalore di $A + E$ allora

$$\min_{1 \leq i \leq n} |\lambda - \lambda_i| \leq \|S\| \|S^{-1}\| \|E\|.$$

Dim. Supponiamo $\lambda \neq \lambda_i$, per ogni i .

Sia $(A + E)x = \lambda x$ per cui $(\lambda I - A)x = Ex$. Pertanto $(\lambda I - SDS^{-1})x = Ex$, ossia $(\lambda I - D)S^{-1}x = S^{-1}ES(S^{-1}x)$.

Poichè $(\lambda I - D)^{-1} = \text{diag}((\lambda - \lambda_1)^{-1}, (\lambda - \lambda_2)^{-1}, \dots, (\lambda - \lambda_n)^{-1})$ avremo

$$S^{-1}x = (\lambda I - D)^{-1}S^{-1}ES(S^{-1}x),$$

da cui

$$\|S^{-1}x\| \leq \|(\lambda I - D)^{-1}\| \|S^{-1}ES\| \|S^{-1}x\|,$$

e quindi

$$1 \leq \max_{1 \leq i \leq n} |(\lambda - \lambda_i)^{-1}| \|S^{-1}\| \|S\| \|E\|,$$

da cui la tesi.

□

COROLLARIO. Se A è normale ($A^H A = AA^H$) allora

$$\min_{1 \leq i \leq n} |\lambda - \lambda_i| \leq \|E\|_2.$$

Dim. Si ricordi che se A è normale S è unitaria (decomposizione di Schur).

Metodo delle potenze.

Sia A una matrice quadrata $n \times n$ con autovalori $\lambda_1, \lambda_2, \dots, \lambda_n$. Supponiamo che a questi corrispondano n autovettori tra loro linearmente indipendenti

x_1, x_2, \dots, x_n , con $\|x_k\|_2 = 1$, per ogni k . Inoltre supponiamo che l'autovalore di modulo massimo λ_1 sia unico cioè

$$|\lambda_n| \leq |\lambda_{n-1}| \leq \dots < |\lambda_1|$$

e che sia semplice (molteplicità algebrica=1). Per approssimare λ_1 consideriamo il seguente metodo detto **Metodo delle potenze**.

Dato un vettore z_0 consideriamo la successione di vettori ottenuta tramite il seguente processo iterativo:

$$\begin{aligned} y_0 &= z_0, \\ y_{k+1} &= Ay_k, \text{ per } k = 0, 1, 2, \dots \end{aligned}$$

Studiamo la convergenza della successione così costruita.

Poichè x_1, x_2, \dots, x_n sono linearmente indipendenti, possiamo rappresentare z_0 come

$$z_0 = \sum_{j=1}^n \alpha_j x_j.$$

Facciamo l'ipotesi che z_0 sia scelto in modo che $\alpha_1 \neq 0$. Allora, per ogni k , ricordando che gli autovalori di A^k sono λ_j^k , per $j = 1, 2, \dots, n$,

$$\begin{aligned} y_k &= A^k z_0 = \sum_{j=1}^n \alpha_j A^k x_j \\ &= \lambda_1^k [\alpha_1 x_1 + \sum_{j=2}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^k x_j]. \end{aligned}$$

Poichè $\left|\frac{\lambda_j}{\lambda_1}\right| < 1$, per $j = 2, \dots, n$, avremo $\lim_{k \rightarrow \infty} \left|\frac{\lambda_j}{\lambda_1}\right|^k = 0$, per ogni $j = 2, \dots, n$. Pertanto la direzione di y_k tende a quella dell'autovettore x_1 . Possiamo quindi ottenere l'approssimazione μ_k di λ_1 tramite il **quoziente di Rayleigh**:

$$\mu_k = \frac{y_k^H A y_k}{y_k^H y_k}.$$

Nella applicazione pratica del metodo è opportuno adottare uno scalaggio ottenendo il seguente algoritmo:

$$\begin{aligned} &\text{per } k = 0, 1, 2, \dots \\ t_k &= \frac{y_k}{\|y_k\|_2}, \\ \mu_k &= \frac{t_k^H A t_k}{t_k^H t_k}, \\ y_{k+1} &= A t_k. \end{aligned}$$

In generale gli errori $\|t_k - x_1\|_2$ e $|\mu_k - \lambda_1|$ sono proporzionali a $\left|\frac{\lambda_2}{\lambda_1}\right|^k$ (a $\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}$ se A è simmetrica).

OSSERVAZIONE. Il metodo converge anche se λ_1 non è semplice, purchè esso sia l'unico autovalore di modulo massimo.

Metodo delle potenze inverse.

Supponiamo che nella situazione precedente sia

$$0 < |\lambda_n| < |\lambda_{n-1}| \leq \dots$$

e si voglia determinare l' autovalore di modulo minimo λ_n . Il metodo delle potenze inverse non è altro che il metodo delle potenze applicato alla matrice A^{-1} .

Più in generale volendo, sotto le opportune ipotesi, approssimare l' autovalore di A più vicino a un certo λ^* , il metodo verrà applicato alla matrice $(A - \lambda^*I)^{-1}$.