Esercizi (& teoria): si svolgono dall'ultimo al primo (SMC-STM 2009)

$\square$ Siano

$$A = \begin{bmatrix} 0 & 1 & 2 \\ -2 & 3 & 5 \\ 1 & -1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 6 \\ 1 \end{bmatrix}.$$

i Risolvere il sistema lineare $A\mathbf{x} = \mathbf{b}$ con il metodo di Gauss

ii Trovare matrici Π di permutazione, D diagonale, L e U^T triangolari inferiori con $L_{ii} = U_{ii} = 1$, $i = 1, 2, 3$, tali che $A = \Pi LU$

iii Annullare la seconda e la terza componente del vettore $\mathbf{b}$ usando il metodo di Givens

iv Esistono le fattorizzazioni di Cholesky delle matrici $M = A + A^T$ e $N = AA^T$?

v Dire se -4 può essere autovalore di A (non calcolare il polinomio caratteristico di A)

vi Studiare la convergenza del metodo di Gauss-Seidel nella risoluzione di $A\mathbf{x} = \mathbf{b}$

... ...

? Maggio 2009

Condizionamento di un sistema lineare vs stabilità dell'algoritmo per la sua risoluzione (seconda parte)

Esempio. Gauss al sistema

$$A(\mathbf{x} + d\mathbf{x}) = \begin{bmatrix} .3 \cdot 10^{-3} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 + dx_1 \\ x_2 + dx_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{b} + d\mathbf{b}.$$

Anziché

$$E = \begin{bmatrix} 1 & 0 \\ -1/.3 \cdot 10^{-3} & 1 \end{bmatrix}$$

il calcolatore definisce

$$\text{“}E\text{”} = \begin{bmatrix} 1 & 0 \\ -.333 \cdot 10^4 & 1 \end{bmatrix}.$$

Poi effettua le moltiplicazioni macchina

$$\text{“}E\text{”}A = \begin{bmatrix} .3 \cdot 10^{-3} & 1 \\ 0 & -.333 \cdot 10^4 \end{bmatrix}, \quad \text{“}E\text{”}(\mathbf{b} + d\mathbf{b}) = \begin{bmatrix} 1 \\ -.333 \cdot 10^4 \end{bmatrix}$$

($-.333 \cdot 10^4 \text{“}+\text{”} 1 = -.333 \cdot 10^4$). Quindi “$\mathbf{x} + d\mathbf{x}$” = (cio' che la macchina ha effettivamente ottenuto calcolando $\mathbf{x} + d\mathbf{x}$) $= \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ ed

errore algoritmico $= \|\text{“}\mathbf{x} + d\mathbf{x}\text{”} - (\mathbf{x} + d\mathbf{x})\| / \|\mathbf{x} + d\mathbf{x}\| = 1/1 = 1$. Nota: è grande rispetto a $u = 0.5 \cdot 10^{-2}$ ($\mathbf{x} + d\mathbf{x} = \begin{bmatrix} 1 \\ 1 - .3 \cdot 10^{-3} \end{bmatrix}$ e si è usata la norma infinito).

In effetti gli elementi di

$$``L" = \begin{bmatrix} 1 & 0 \\ .333 \cdot 10^4 & 1 \end{bmatrix}, \ ``U" = \begin{bmatrix} .3 \cdot 10^{-3} & 1 \\ 0 & -.333 \cdot 10^4 \end{bmatrix}$$

sono in modulo molto più grandi di quelli di $|A|$.

Conclusione: il problema $A\mathbf{x} = \mathbf{b}$ è ben condizionato, ma l'algoritmo di Gauss per la sua risoluzione è instabile.

Modificando Gauss è possibile definire un algoritmo stabile per la risoluzione di $A\mathbf{x} = \mathbf{b}$:

Gauss modificato al sistema

$$A(\mathbf{x} + d\mathbf{x}) = \begin{bmatrix} .3 \cdot 10^{-3} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 + dx_1 \\ x_2 + dx_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{b} + d\mathbf{b}.$$

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ PA = \begin{bmatrix} 1 & 0 \\ .3 \cdot 10^{-3} & 1 \end{bmatrix}, \ P(\mathbf{b} + d\mathbf{b}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$E = \begin{bmatrix} 1 & 0 \\ -.3 \cdot 10^{-3} & 1 \end{bmatrix}, \ EPA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ EP(\mathbf{b} + d\mathbf{b}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$(-.3 \cdot 10^{-3} ``+" 1 = 1)$. Quindi $``\mathbf{x} + d\mathbf{x}" = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ed

errore algoritmo $= \| ``\mathbf{x} + d\mathbf{x}" - (\mathbf{x} + d\mathbf{x}) \| / \| \mathbf{x} + d\mathbf{x} \| = .3 \cdot 10^{-3} / 1 = .3 \cdot 10^{-3}$. Nota: è più piccolo di $u = 0.5 \cdot 10^{-2}$.

In effetti gli elementi di

$$``L" = \begin{bmatrix} 1 & 0 \\ .3 \cdot 10^{-3} & 1 \end{bmatrix}, \quad ``U" = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

sono dello stesso ordine di grandezza di quelli di $|A|$.

Conclusione: l'algoritmo di Gauss modificato per la risoluzione di $A\mathbf{x} = \mathbf{b}$ è stabile.

18 Maggio 2009, 9 Maggio 2009, 28 Aprile 2009, 21 Aprile 2009
18 Maggio 2009

Condizionamento di un sistema lineare vs stabilità dell'algoritmo per la sua risoluzione (prima parte)

"Risultato 1" (senza dim) sul condizionamento di un sistema lineare $A\mathbf{x} = \mathbf{b}$:

Siano $\mathbf{x}$ e $\mathbf{x} + d\mathbf{x}$ tali che $A\mathbf{x} = \mathbf{b}$ e $(A + dA)(\mathbf{x} + d\mathbf{x}) = \mathbf{b} + d\mathbf{b}$ dove dA e $d\mathbf{b}$ sono perturbazioni, rispettivamente, della matrice A e del vettore b. Allora

$$\frac{\|d\mathbf{x}\|}{\|\mathbf{x}\|} \leq \mu(A) \frac{\|dA\|/\|A\| + \|d\mathbf{b}\|/\|\mathbf{b}\|}{1 - \mu(A)\|dA\|/\|A\|}$$

dove $\mu(A) = \|A\| \|A^{-1}\|$ e la norma matriciale è quella indotta dalla norma vettoriale utilizzata.

[Nota: per $dA = 0$ ritroviamo il risultato visto in precedenza, relativo a perturbazioni sul solo vettore dei termini noti: $A\mathbf{x} = \mathbf{b}$ & $A(\mathbf{x} + d\mathbf{x}) = \mathbf{b} + d\mathbf{b}$ $\Rightarrow \|d\mathbf{x}\|/\|\mathbf{x}\| \leq \mu(A)\|d\mathbf{b}\|/\|\mathbf{b}\|$.]

Questo risultato è ovviamente utile per valutare l'errore inerente sulla soluzione di un sistema lineare: $\mu(A)$ grande $\Rightarrow$ errore inerente può essere molto più grande dell'errore inerente sui dati; $\mu(A)$ piccolo $\Rightarrow$ errore inerente deve essere dello stesso ordine dell'errore inerente sui dati.

Esempio. Il sistema

$$A\mathbf{x} = \begin{bmatrix} .3 \cdot 10^{-3} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 + .3 \cdot 10^{-3} \\ 1 \end{bmatrix} = \mathbf{b}$$

in $F(10, 3, \ldots)$ diventa

$$A(\mathbf{x} + d\mathbf{x}) = \begin{bmatrix} .3 \cdot 10^{-3} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 + dx_1 \\ x_2 + dx_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{b} + d\mathbf{b}$$

($1 + .3 \cdot 10^{-3} = .10003 \cdot 10^1$, $1 \text{“}+\text{”} .3 \cdot 10^{-3} = .1 \cdot 10^1$) quindi, per il “Risultato 1” con $dA = 0$, utilizzando la norma infinito, si ottiene la limitazione

$$\frac{\|d\mathbf{x}\|}{\|\mathbf{x}\|} \leq (1+e)^2 \frac{\|d\mathbf{b}\|}{\|\mathbf{b}\|}, \quad \frac{\|d\mathbf{b}\|}{\|\mathbf{b}\|} = \frac{.3 \cdot 10^{-3}}{1 + .3 \cdot 10^{-3}}.$$

Essendo $e = .3 \cdot 10^{-3}$, il numero $\mu_\infty(A) = (1+e)^2$ è poco più grande di 1 $\Rightarrow$ l'errore inerente sulla soluzione è dello stesso ordine di grandezza dell'errore inerente sul vettore dei termini noti $\Rightarrow$ il problema $A\mathbf{x} = \mathbf{b}$ è ben condizionato.

Vediamo ora come il “Risultato 1” può essere utile anche per valutare l'errore commesso nelle successive operazioni richieste dal metodo di Gauss (errore dovuto ai necessari arrotondamenti operati dalla macchina durante i calcoli) per la risoluzione di $A\mathbf{x} = \mathbf{b}$.

Premessa: “Risultato 2” (senza dim) sul metodo di Gauss :

Dati A $n \times n$ non singolare, $\mathbf{b}$ $n \times 1$, con a_{ij} e b_i numeri macchina, posto $\mathbf{x}$ = soluzione esatta di $A\mathbf{x} = \mathbf{b}$, “$\mathbf{x}$” = soluzione effettivamente calcolata con Gauss tramite i passi:

a) scrittura di “L”, “U” (i fattori della decomposizione LU di A effettivamente calcolati) e di “$\mathbf{c}$” (la soluzione di “L”$\mathbf{c} = \mathbf{b}$ effettivamente calcolata);

b) scrittura di “$\mathbf{x}$” (la soluzione di “U”$\mathbf{x}$ = “$\mathbf{c}$” effettivamente calcolata),

SI HA CHE “$\mathbf{x}$” è la soluzione esatta di un sistema ottenuto da $A\mathbf{x} = \mathbf{b}$ perturbando A di DA, cioè vale l'identità

$$(A + DA)\text{“}\mathbf{x}\text{”} = \mathbf{b}, \quad |DA| \leq 4nu(|A| + |\text{“}L\text{”}||\text{“}U\text{”}|) + O(u^2)$$

($|M|$ = matrice i cui elementi sono i moduli degli elementi di M).

Il “Risultato 2” permette di valutare l'errore dell'algoritmo di Gauss sfruttando la maggiorazione del “Risultato 1”:

$$\frac{\|\text{“}\mathbf{x}\text{”} - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \mu(A) \frac{\|DA\|/\|A\|}{1 - \mu(A)\|DA\|/\|A\|}$$

(usa Ris.1 per $dA = DA$, $d\mathbf{b} = \mathbf{0}$, $\mathbf{x} + d\mathbf{x}$ = “$\mathbf{x}$”). Quest'ultima disuguaglianza e la maggiorazione data per $|DA|$ ci dicono che $\|\text{“}\mathbf{x}\text{”} - \mathbf{x}\|/\|\mathbf{x}\|$ può essere grande quando $|\text{“}L\text{”}|$ e $|\text{“}U\text{”}|$ sono grandi rispetto a $|A|$; $\|\text{“}\mathbf{x}\text{”} - \mathbf{x}\|/\|\mathbf{x}\|$ è invece piccolo se $|\text{“}L\text{”}|$ e $|\text{“}U\text{”}|$ sono dello stesso ordine di grandezza di $|A|$.

Per risolvere $A\mathbf{x} = \mathbf{b}$ (sempre nel caso $\det(S_k) \neq 0$, $k = 1, \ldots, n$) si può usare anche l'algoritmo:

(1) calcolare L ed U tali che $A = LU$,

(2) risolvere il sistema $L\mathbf{c} = \mathbf{b}$ (con il m. di sostituzione in avanti),

(3) risolvere il sistema $U\mathbf{x} = \mathbf{c}$ (con il m. di sostituzione all'indietro).

NOTA: applicare le E_i oltre che ad A anche a $\mathbf{b}$ (cioè calcolare $E_{n-1} \cdots E_2 E_1 [A \,|\, \mathbf{b}]$), come si fa nel metodo di Gauss, è equivalente ad eseguire i passi (1) e (2) del presente algoritmo contemporaneamente, infatti $E_{n-1} \cdots E_2 E_1 \mathbf{b} = L^{-1}\mathbf{b}$.

9 Maggio 2009

□ Risolvere con il metodo di Gauss il sistema lineare

$$\begin{bmatrix} -2 & 4 & -1 \\ 4 & -9 & 0 \\ -4 & 5 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 12 \\ -32 \\ 3 \end{bmatrix}.$$

□ Calcolare le decomposizioni LU delle matrici

$$A = \begin{bmatrix} -2 & 4 \\ 4 & -8 \end{bmatrix}, \ A = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{bmatrix}, \ A = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}, \ A = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}$$

Nota: Le prime due matrici sono singolari, ad esempio nella prima matrice la seconda colonna è -2 volte la prima colonna, quindi è "linearmente dipendente" dalla prima. Ne segue che per tali matrici $u_{nn} = 0$.

Metodo di Gauss: decomposizione LU, risoluzione sistemi lineari

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & & . \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

IPOTESI: $a_{11} \neq 0$

$$E_1 = \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ -a_{21}/a_{11} & 1 & . & & . \\ . & 0 & . & . & . \\ . & . & . & . & 0 \\ -a_{n1}/a_{11} & 0 & \cdots & 0 & 1 \end{bmatrix}, \ E_1 A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^2 & \cdots & a_{2n}^2 \\ . & . & & . \\ 0 & a_{n2}^2 & \cdots & a_{nn}^2 \end{bmatrix} =: A^2,$$

$a_{ij}^2 = a_{ij} - (a_{i1}/a_{11})a_{1j}$, $2 \leq i, j \leq n$,
$a_{i1}^2 = 0$, $2 \leq i \leq n$,
$a_{ij}^2 = a_{ij}$ altrimenti (per i rimanenti i e j).
NOTA:

$$\begin{bmatrix} 1 & 0 \\ -a_{21}/a_{11} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22}^2 \end{bmatrix},$$

quindi, passando ai determinanti,

$$a_{22}^2 = \frac{1}{a_{11}} \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

IPOTESI: $a_{22}^2 \neq 0$ (ovvero $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \neq 0$)

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & -a_{32}^2/a_{22}^2 & 1 & . & . \\ . & . & . & . & 0 \\ 0 & -a_{n2}^2/a_{22}^2 & 0 & \cdots 0 & 1 \end{bmatrix}$$

$$E_2 E_1 A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22}^2 & a_{23}^2 & \cdots & a_{2n}^2 \\ 0 & 0 & a_{33}^3 & \cdots & a_{3n}^3 \\ . & . & . & & . \\ 0 & 0 & a_{n3}^3 & \cdots & a_{nn}^3 \end{bmatrix} =: A^3,$$

$a_{ij}^3 = a_{ij}^2 - (a_{i2}^2/a_{22}^2)a_{2j}^2$, $3 \leq i, j \leq n$,
$a_{i2}^2 = 0$, $3 \leq i \leq n$,
$a_{ij}^3 = a_{ij}^2$ altrimenti (per i rimanenti i e j).
NOTA:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -a_{32}^2/a_{22}^2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -a_{21}/a_{11} & 1 & 0 \\ -a_{31}/a_{11} & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^2 & a_{23}^2 \\ 0 & 0 & a_{33}^3 \end{bmatrix},$$

quindi, passando ai determinanti,

$$a_{33}^3 = \frac{1}{a_{11}a_{22}^2} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

...

IPOTESI: $a_{n-1,n-1}^{n-1} \neq 0$ (ovvero $\det \begin{bmatrix} a_{11} & \cdots & a_{1n-1} \\ . & & . \\ a_{n-11} & \cdots & a_{n-1n-1} \end{bmatrix} \neq 0$)

$$E_{n-1} = \begin{bmatrix} 1 & 0 & \cdots & . & 0 \\ 0 & . & . & . & . \\ . & . & . & 1 & 0 \\ 0 & \cdots & 0 & -a_{nn-1}^{n-1}/a_{n-1n-1}^{n-1} & 1 \end{bmatrix}$$

$$E_{n-1} \cdots E_2 E_1 A = \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ & a_{22}^2 & a_{2n}^2 \\ & . & . \\ & & a_{nn}^n \end{bmatrix} =: A^n,$$

$a_{nn}^n = a_{nn}^{n-1} - (a_{nn-1}^{n-1}/a_{n-1n-1}^{n-1})a_{n-1n}^{n-1}$,
$a_{nn-1}^n = 0$,
$a_{ij}^n = a_{ij}^{n-1}$ altrimenti (per i rimanenti i e j).
NOTA: si osserva che

$$a_{nn}^n = \frac{1}{a_{11}a_{22}^2 ... a_{n-1n-1}^{n-1}} \det(A).$$

Sia A $n \times n$ e

$$S_k = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ . & & . \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}, \quad k = 1, \ldots, n.$$

Teorema. Se $\det(S_k) \neq 0$ per $k = 1 \ldots n-1$ allora

(1) esistono matrici $E_1, E_2, \ldots, E_{n-1}$ "elementari di Gauss" tali che $E_{n-1} \cdots E_2 E_1 A = A^{(n)} = (a^n_{ij})^n_{i,j=1}$ è triangolare superiore

(2) $\det(A) = 0$ se e solo se $a^n_{nn} = 0$

(3) esiste L $n \times n$ triangolare inferiore, $L_{ii} = 1$, ed U triangolare superiore tali che $A = LU$: $U = A^{(n)}$ e

$$L = E_1^{-1} E_2^{-1} \cdots E_{n-1}^{-1} = \begin{bmatrix} 1 & 0 & \cdots & . & 0 \\ a_{21}/a_{11} & 1 & & . & . \\ . & a^2_{32}/a^2_{22} & . & & . \\ . & . & & 1 & 0 \\ a_{n1}/a_{11} & a^2_{n2}/a^2_{22} & . & a^{n-1}_{nn-1}/a^{n-1}_{n-1n-1} & 1 \end{bmatrix}$$

(4) (metodo di Gauss (a)) considerato il sistema lineare $A\mathbf{x} = \mathbf{b}$, $\mathbf{b}$ $n \times 1$ generico, si può introdurre un sistema lineare triangolare superiore $U\mathbf{x} = \mathbf{c}$ equivalente ad $A\mathbf{x} = \mathbf{b}$ col seguente algoritmo:

$$\begin{aligned} &[A \mid \mathbf{b}] \\ &E_1[A \mid \mathbf{b}] = [E_1 A \mid E_1 \mathbf{b}] \\ &E_2 E_1 [A \mid \mathbf{b}] = [E_2 E_1 A \mid E_2 E_1 \mathbf{b}] \\ &\ldots \\ &E_{n-1} \cdots E_2 E_1 [A \mid \mathbf{b}] = [E_{n-1} \cdots E_2 E_1 A \mid E_{n-1} \cdots E_2 E_1 \mathbf{b}], \\ &U = E_{n-1} \cdots E_2 E_1 A, \quad \mathbf{c} = E_{n-1} \cdots E_2 E_1 \mathbf{b} \end{aligned}$$

ovvero:

$$\begin{aligned} &\text{per} \quad k = 1 \ldots n-1 \quad \{ \\ &a^{k+1}_{ij} = a^k_{ij} - (a^k_{ik}/a^k_{kk}) a^k_{kj}, \quad k+1 \leq i, j \leq n, \\ &a^{k+1}_{ik} = 0, \quad k+1 \leq i \leq n, \\ &a^{k+1}_{ij} = a^k_{ij}, \quad \text{altrimenti (per i rimanenti i,j)}, \\ &b^{k+1}_i = b^k_i - (a^k_{ik}/a^k_{kk}) b^k_k, \quad k+1 \leq i \leq n, \\ &b^{k+1}_i = b^k_i, \quad \text{altrimenti} \quad (i <= k); \; \} \\ &U = A^n = (a^n_{ij})_{i,j=1\ldots n}, \\ &\mathbf{c} = \mathbf{b}^n \; . \end{aligned}$$

Il costo di questo algoritmo è $n^3/3 + O(n^2)$ operazioni moltiplicative. Più precisamente, per k generico occorre eseguire: $n-k$ divisioni per calcolare i rapporti $l_{ik} = a^k_{ik}/a^k_{kk}$, cioè gli elementi della matrice L, le $(n-k)^2$ moltiplicazioni $l_{ik} a^k_{kj}$ e le $n-k$ moltiplicazioni $l_{ik} b^k_k$. Quindi L ed U possono essere calcolate con $\sum_{k=1\ldots n-1}(n-k) = n(n-1)/2$ divisioni e $\sum_{k=1\ldots n-1}(n-k)^2 = n(n-1)(2n-1)/6$ moltiplicazioni. Per calcolare anche $\mathbf{c}$ occorre eseguire altre $\sum_{k=1\ldots n-1}(n-k) = n(n-1)/2$ moltiplicazioni

(5) (metodo di Gauss (b)) se A è inoltre non singolare i sistemi $A\mathbf{x} = \mathbf{b}$ e $U\mathbf{x} = \mathbf{c}$ oltre ad essere equivalenti ammettono una unica soluzione

$\mathbf{x} = A^{-1}\mathbf{b} = U^{-1}\mathbf{c}$, calcolabile, con altre $n(n-1)/2$ moltiplicazioni e n divisioni, applicando il metodo di sostituzione all'indietro al sistema $U\mathbf{x} = \mathbf{c}$

□ Se A $n \times n$ è definita positiva allora $\det(S_k) > 0$, $k = 1 \dots n$, quindi esiste la decomposizione LU di A ed ogni sistema $A\mathbf{x} = \mathbf{b}$, $\mathbf{b}$ $n \times 1$, può essere risolto usando il metodo di Gauss

□ Un esercizio dell'esonero

A, B $n \times n$. Se λ è autovalore di AB allora è autovalore di BA.

Infatti λ autovalore di $AB \Rightarrow$ esiste $\mathbf{v} \neq \mathbf{0}$ tale che $AB\mathbf{v} = \lambda\mathbf{v} \Rightarrow B(AB\mathbf{v}) = B(\lambda\mathbf{v}) \Rightarrow (BA)B\mathbf{v} = \lambda B\mathbf{v}$, ovvero λ è autovalore di BA con corrispondente autovettore $B\mathbf{v}$. Quindi AB e BA hanno gli stessi autovalori e gli autovettori di BA si ottengono moltiplicando quelli di AB per B

A $n \times n$ invertibile $\Rightarrow \mu_2(A^*A) = \mu_2(A)^2$.

Infatti,

$$\begin{aligned}
\mu_2(A^*A) &= \|A^*A\|_2 \|(A^*A)^{-1}\|_2 = \rho(A^*A)\rho((A^*A)^{-1}) \\
&= \rho(A^*A)\rho(A^{-1}(A^*)^{-1}) = \rho(A^*A)\rho((A^*)^{-1}A^{-1}) \\
&= \rho(A^*A)\rho((A^{-1})^*A^{-1}) = \|A\|^2\|A^{-1}\|^2 = \mu_2(A)^2
\end{aligned}$$

(risultati utilizzati: A^*A è hermitiana; M hermitiana $\Rightarrow M^{-1}$ hermitiana; MN ed NM hanno gli stessi autovalori, quindi $\rho(MN) = \rho(NM)$; $(M^*)^{-1} = (M^{-1})^*$)

28 Aprile 2009

□ Usando il Teorema di Schur scrivere un algoritmo alternativo per la risoluzione di $A\mathbf{x} = \mathbf{b}$, A $n \times n$, $\mathbf{b}$ $n \times 1$. Si suppone $\det(A) =$ determinante di A diverso da zero. Qual'è l'inconveniente di tale algoritmo?

Risoluzione. Algoritmo:

(1) calcolare una matrice unitaria S tale che $S^{-1}AS =: U$ è triangolare superiore (essa esiste per il Teorema di Schur),

(2) risolvere i due sistemi lineari: $U\mathbf{y} = S^*\mathbf{b}$, $S^*\mathbf{x} = \mathbf{y}$.

Inconveniente: l'operazione di cui al punto (1) non è banale e comunque non può essere effettuata (in generale) con un numero finito di operazioni aritmetiche. Sappiamo invece che, usando Gauss o una delle sue varianti, o Givens, o Householder, è possibile introdurre un sistema lineare triangolare $U\mathbf{x} = \mathbf{c}$ equivalente ad $A\mathbf{x} = \mathbf{b}$ facendo *tot* $< +\infty$ o.a.. La U ottenuta con tali algoritmi non ha gli autovalori di A, ma che importa !

Risolvere sistemi lineari con matrice dei coefficienti triangolare superiore

U $n \times n$ triangolare superiore non singolare, ovvero $[U]_{ii} \neq 0$, per ogni i, e $\mathbf{c}$ vettore $n \times 1$: $U\mathbf{x} = \mathbf{c}$ iff

$$\begin{aligned}
u_{11}x_1 + u_{12}x_2 + \ldots + u_{1n}x_n &= c_1 \\
u_{22}x_2 + \ldots + u_{2n}x_n &= c_2 \\
&\vdots \\
u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n &= c_{n-1} \\
u_{nn}x_n &= c_n
\end{aligned}$$

iff

$$
\begin{aligned}
&x_n = c_n/u_{nn} \\
&x_{n-1} = (c_{n-1} - u_{n-1,n}x_n)/u_{n-1,n-1} \\
&\vdots \\
&x_2 = (c_2 - u_{23}x_3 ... - u_{2n}x_n)/u_{22} \\
&x_1 = (c_1 - u_{12}x_2 .... - u_{1n}x_n)/u_{11}
\end{aligned}
$$

iff per $i = n, \ldots, 1$: $x_i = (c_i - \sum_{j=i+1,\ldots,n} u_{ij}x_j)/u_{ii}$. Costo $= n(n+1)/2$ moltiplicazioni-divisioni.

Risolvere sistemi lineari con matrice dei coefficienti generica, triangolarizzazione di una matrice

A matrice $n \times n$ non singolare, $\mathbf{b}$ vettore $n \times 1$. E_k, $k = 1, \ldots, n-1$, matrici $n \times n$ non singolari tali che

$$
\begin{aligned}
&A\mathbf{x} = \mathbf{b} \text{ iff} \\
&E_1A\mathbf{x} = E_1\mathbf{b}, \quad E_1A = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ . & . & & . \\ 0 & * & \cdots & * \end{bmatrix} \text{ iff} \\
&E_2E_1A\mathbf{x} = E_2E_1\mathbf{b}, \quad E_2E_1A = \begin{bmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ . & 0 & * & \cdots & * \\ . & . & . & & . \\ 0 & 0 & * & \cdots & * \end{bmatrix} \text{ iff}
\end{aligned}
$$

... iff $E_{n-1} \cdots E_2E_1A\mathbf{x} = E_{n-1} \cdots E_2E_1\mathbf{b}$, ovvero $U\mathbf{x} = \mathbf{c}$, con

$$
U = E_{n-1} \cdots E_2E_1A = \begin{bmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ . & 0 & * & \cdots & * \\ . & . & . & . & . \\ 0 & . & . & 0 & * \end{bmatrix}, \quad \mathbf{c} = E_{n-1} \cdots E_2E_1\mathbf{b}.
$$

Definizione di E_k (passo k della triangolarizzazione)

Per $k = 1, \ldots, n-1$: Al termine del passo $k-1$ è stata prodotta la matrice

$$
E_{k-1} \cdots E_2E_1A = \begin{bmatrix} T & V \\ 0 & W \end{bmatrix},
$$

T $k-1 \times k-1$ triangolare superiore, W $n-k+1 \times n-k+1$ (per $k = 1$ si ha $A = W$).

Si vogliono annullare gli elementi $(k+1,k), \ldots, (n,k)$ di tale matrice, ovvero il secondo, ..., l'ultimo elemento della prima colonna di W:

$$
\begin{aligned}
&E_k = \begin{bmatrix} I & 0 \\ 0 & F \end{bmatrix}, \ I \ k-1 \times k-1, \quad F \ n-k+1 \times n-k+1, \\
&E_k(E_{k-1} \cdots E_2E_1A) = \begin{bmatrix} I & 0 \\ 0 & F \end{bmatrix} \begin{bmatrix} T & V \\ 0 & W \end{bmatrix} = \begin{bmatrix} T & V \\ 0 & FW \end{bmatrix} = \begin{bmatrix} T & V \\ 0 & [F\mathbf{w} \ FY] \end{bmatrix}
\end{aligned}
$$

dove $W = [\mathbf{w} \ Y]$. Se $\mathbf{w}$, la prima colonna di W, è gia' della forma $[* \ 0 \ \cdots \ 0]^T$ allora porre $F = I$, altrimenti scegliere F come in Gauss semplice, Gauss con pivot parziale, Givens, Householder per ottenere $F\mathbf{w} = [* \ 0 \ \cdots \ 0]^T$.

Costo $= \sum_{k=1\dots n-1}$[costo del prodotto matrice-matrice FY, F $n-k+1 \times n-k+1$, Y $n-k+1 \times n-k$]. Ad esempio, in Gauss costo$(FY) = (n-k)^2$ moltiplicazioni $\Rightarrow$ costo di Gauss $= n^3/3 + O(n^2)$.

Un esempio di triangolarizzazione

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix},$$

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \; E_1A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & -1 & 1 & 2 \\ 0 & -1 & -1 & 2 \end{bmatrix} \; (F = E_1, \; W = A),$$

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \; E_2E_1A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & -1 & 4 \end{bmatrix} \; (F = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \; W = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 2 \\ -1 & -1 & 2 \end{bmatrix}),$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \; E_3E_2E_1A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{bmatrix} =: U \; (F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \; W = \begin{bmatrix} 1 & 4 \\ -1 & 4 \end{bmatrix}).$$

In questo esempio di triangolarizzazione di una matrice si osserva che Gauss con pivot parziale coincide con Gauss semplice e che $\max |(FW)_{ij}| = 2 \max |(W)_{ij}|$, cioè la disuguaglianza ottenuta prima è ottimale. Nota: $\|A\|_F = \sqrt{13} < \|U\|_F = \sqrt{88}$.

Applicando Givens, per triangolarizzare A, si ottiene

$$E_3E_2E_1A = U = \begin{bmatrix} 2 & 1/2 & 0 & -1 \\ 0 & \sqrt{11}/2 & 0 & -1/\sqrt{11} \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2}/\sqrt{11} \end{bmatrix}.$$

(vedi più avanti). Nota: $\|A\|_2 = \|U\|_2 \le \|A\|_F = \|U\|_F = \sqrt{13}$.

Vedremo che per triangolarizzare una matrice A $n \times n$ occorre calcolare dei prodotti del tipo FW con F,W matrici $j \times j$, $j = n, \dots, 2$. La prima riga di FW coincide con $[u_{n-j+1,n-j+1} \dots u_{n-j+1,n}]$ essendo U $n \times n$ la matrice triangolare finale. Gli elementi della prima colonna di FW se si eccettua il primo sono tutti nulli. Inoltre l'elemento $(2, 2)$ dell'ultima FW coincide con u_{nn}.

In Givens e Householder F è unitaria, quindi $\|FW\|_2 = \|W\|_2$, ovvero gli elementi di FW e W hanno lo stesso ordine di grandezza.

In Gauss $F = GP$,

$$G = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ m_{21} & 1 & & \\ \vdots & & & 0 \\ m_{j1} & 0 & \dots 0 & 1 \end{bmatrix}, \quad P \; j \times j \text{ di permutazione}$$

$\Rightarrow \|FW\|_2 \leq \|F\|_2 \|W\|_2$, dove

$$\begin{aligned} \|F\|_2 &= \|G\|_2 = \max_{\mathbf{x}\, j\times 1\, \|\mathbf{x}\|_2=1} \|G\mathbf{x}\|_2 \\ &\geq \|G\mathbf{e}_1\|_2 = \sqrt{1+m_{21}^2+\ldots+m_{j1}^2} > 1 \ \text{ se } \ G \neq I, \end{aligned}$$

ovvero gli elementi di FW possono essere in media più grandi di quelli di W.

Nota: poiché $m_{k1} = -(P\mathbf{w})_k/(P\mathbf{w})_1$, $k = 2,\ldots,j$, dove $\mathbf{w}$ è il vettore che vogliamo trasformare in

$$F\mathbf{w} = \begin{bmatrix} * \\ 0 \\ \vdots \\ .0 \end{bmatrix}$$

($\mathbf{w}$=prima colonna di W), il numero $\sqrt{1+m_{21}^2+\ldots+m_{j1}^2}$ è minimo quando P è scelta in modo che $|(P\mathbf{w})_1| = \max_k |(P\mathbf{w})_k|$, ovvero in Gauss con pivot parziale. In Gauss con pivot parziale si ha inoltre: $\max|(FW)_{ij}| \leq 2\max|(W)_{ij}|$ (provare a dimostrare questa disuguaglianza) da cui segue il risultato $\max|u_{ij}| \leq 2^{n-1}\max|a_{ij}|$, cioè gli elementi di U non possono essere più di 2^{n-1} volte più grandi di quelli di A.

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & & \\ 1/\sqrt{2} & 1/\sqrt{2} & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 & \sqrt{2} \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix},$$

$$\begin{aligned} &\begin{bmatrix} \sqrt{2}/\sqrt{3} & 0 & -1/\sqrt{3} & \\ 0 & 1 & 0 & \\ 1/\sqrt{3} & 0 & \sqrt{2}/\sqrt{3} & \\ & & & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 & \sqrt{2} \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{3} & 0 & -1/\sqrt{3} & -1/\sqrt{3} \\ 0 & 1/\sqrt{2} & 0 & \sqrt{2} \\ 0 & -\sqrt{3}/\sqrt{2} & \sqrt{2}/\sqrt{3} & \sqrt{2}/\sqrt{3} \\ -1 & -1 & -1 & 1 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} &\begin{bmatrix} \sqrt{3}/2 & 0 & 0 & -1/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/2 & 0 & 0 & \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & -1/\sqrt{3} & -1/\sqrt{3} \\ 0 & 1/\sqrt{2} & 0 & \sqrt{2} \\ 0 & -\sqrt{3}/\sqrt{2} & \sqrt{2}/\sqrt{3} & \sqrt{2}/\sqrt{3} \\ -1 & -1 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1/2 & 0 & -1 \\ 0 & 1/\sqrt{2} & 0 & \sqrt{2} \\ 0 & -\sqrt{3}/\sqrt{2} & \sqrt{2}/\sqrt{3} & \sqrt{2}/\sqrt{3} \\ 0 & -\sqrt{3}/2 & -2/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 & 0 \\ 0 & \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1/2 & 0 & -1 \\ 0 & 1/\sqrt{2} & 0 & \sqrt{2} \\ 0 & -\sqrt{3}/\sqrt{2} & \sqrt{2}/\sqrt{3} & \sqrt{2}/\sqrt{3} \\ 0 & -\sqrt{3}/2 & -2/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1/2 & 0 & -1 \\ 0 & \sqrt{2} & -\sqrt{2}/2 & 0 \\ 0 & 0 & (\sqrt{2}/\sqrt{3})/2 & 2\sqrt{2}/\sqrt{3} \\ 0 & -\sqrt{3}/2 & -2/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}, \end{aligned}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2\sqrt{2}/\sqrt{11} & 0 & -\sqrt{3}/\sqrt{11} \\ 0 & 0 & 1 & 0 \\ 0 & \sqrt{3}/\sqrt{11} & 0 & 2\sqrt{2}/\sqrt{11} \end{bmatrix} \begin{bmatrix} 2 & 1/2 & 0 & -1 \\ 0 & \sqrt{2} & -\sqrt{2}/2 & 0 \\ 0 & 0 & (\sqrt{2}/\sqrt{3})/2 & 2\sqrt{2}/\sqrt{3} \\ 0 & -\sqrt{3}/2 & -2/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 1/2 & 0 & -1 \\ 0 & \sqrt{11}/2 & 0 & -1/\sqrt{11} \\ 0 & 0 & (\sqrt{2}/\sqrt{3})/2 & 2\sqrt{2}/\sqrt{3} \\ 0 & 0 & -(\sqrt{22}/\sqrt{3})/2 & 2\sqrt{2}/\sqrt{33} \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (1/\sqrt{3})/2 & -(\sqrt{11}/\sqrt{3})/2 \\ 0 & 0 & (\sqrt{11}/\sqrt{3})/2 & (1/\sqrt{3})/2 \end{bmatrix} \begin{bmatrix} 2 & 1/2 & 0 & -1 \\ 0 & \sqrt{11}/2 & 0 & -1/\sqrt{11} \\ 0 & 0 & (\sqrt{2}/\sqrt{3})/2 & 2\sqrt{2}/\sqrt{3} \\ 0 & 0 & -(\sqrt{22}/\sqrt{3})/2 & 2\sqrt{2}/\sqrt{33} \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 1/2 & 0 & -1 \\ 0 & \sqrt{11}/2 & 0 & -1/\sqrt{11} \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 4\sqrt{2}/\sqrt{11} \end{bmatrix}.$$

□ Mostrare che $\|\mathbf{y}\mathbf{w}^*\|_2 = \|\mathbf{y}\mathbf{w}^*\|_F = \|\mathbf{y}\|_2\|\mathbf{w}\|_2$, $\mathbf{y},\mathbf{w}$ vettori $n \times 1$, ricordando che $\|M\|_2 = \sqrt{\rho(M^*M)}$ e $\|M\|_F = \sqrt{\mathrm{tr}\,(M^*M)}$ ($\mathrm{tr}\,(A)$ = traccia di A $= \sum_i a_{ii}$)

Risoluzione.

$$\|\mathbf{x}\mathbf{y}^*\|_F^2 = \mathrm{tr}\,(\mathbf{y}\mathbf{x}^*\mathbf{x}\mathbf{y}^*) = \mathbf{x}^*\mathbf{x}[\,\mathrm{tr}\,(\mathbf{y}\mathbf{y}^*)] = \mathbf{x}^*\mathbf{x}[\mathbf{y}^*\mathbf{y}] = \|\mathbf{x}\|_2^2\|\mathbf{y}\|_2^2,$$

$$\|\mathbf{x}\mathbf{y}^*\|_2^2 = \rho(\mathbf{y}\mathbf{x}^*\mathbf{x}\mathbf{y}^*) = \mathbf{x}^*\mathbf{x}[\rho(\mathbf{y}\mathbf{y}^*)] = \mathbf{x}^*\mathbf{x}[\mathbf{y}^*\mathbf{y}] = \|\mathbf{x}\|_2^2\|\mathbf{y}\|_2^2$$

Nota: a entrambi i risultati si puo arrivare usando solo le definizioni di $\|\cdot\|_F$ e $\|\cdot\|_2$ e la disuguaglianza di Schwartz:

$$\|\mathbf{y}\mathbf{w}^*\|_2 = \max_{\mathbf{x}\ n\times 1\ \|\mathbf{x}\|_2=1} \|\mathbf{y}\mathbf{w}^*\mathbf{x}\|_2 = \max_{\mathbf{x}\ n\times 1\ \|\mathbf{x}\|_2=1} |\mathbf{w}^*\mathbf{x}|\|\mathbf{y}\|_2 = \|\mathbf{y}\|_2 \max_{\mathbf{x}\ n\times 1\ \|\mathbf{x}\|_2=1} |\mathbf{w}^*\mathbf{x}|.$$

Studiamo l'argomento del max: se $\mathbf{x} = \mathbf{w}/\|\mathbf{w}\|_2$, allora $|\mathbf{w}^*\mathbf{x}| = \|\mathbf{w}\|_2$; per ogni $\mathbf{x}$, invece, $\|\mathbf{w}^*\mathbf{x}\| \leq \|\mathbf{w}\|_2\|\mathbf{x}\|_2 = \|\mathbf{w}\|_2$. Ne segue che $\max_{\mathbf{x}\ n\times 1\ \|\mathbf{x}\|_2=1} |\mathbf{w}^*\mathbf{x}| = \|\mathbf{w}\|_2$.

Esercizi, II Teorema di Gershgorin, matrici normali e Teorema di Schur

□ A,B matrici $n \times n$ non singolari.

a $\mu(AB) \leq \mu(A)\mu(B)$, $\mu(M) = \|M\|\|M^{-1}\|$

b AB e BA hanno gli stessi autovalori (sugg.: dimostrare che AB e BA sono simili)

c $\mu_2(B^*B) \leq \mu_2(B)^2$ (sugg.: usare a) e b))

□

$$A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix},\ a \in \mathbb{C}.$$

a Trovare P 4×4 di permutazione tale che

$$PAP^T = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}$$

(Nota: righe di PA = permutazione delle righe di A; colonne di PAP^T = permutazione delle colonne di PA con la stessa legge)

b per quali valori di a la matrice A ha un autovalore uguale a $\rho(A)$?

□ Dati $\mathbf{y}, \mathbf{w}, \mathbf{z}$ vettori $n \times 1$ dimostrare che il prodotto matrice vettore $(\mathbf{y}\mathbf{w}^*)\mathbf{z}$ si può calcolare con $2n$ moltiplicazioni

□ Mostrare che in generale non è vera l'identità $AB = BA$, A, B $n \times n$

□ Per A $n \times n$ la misura $\|A\|_F = \sqrt{\sum_{i,j=1,\dots,n} |a_{ij}|^2}$ è una norma matriciale che non è indotta da nessuna norma vettoriale. Mostrare che $\|A\|_2 \leq \|A\|_F$.

□ Localizzare nel modo migliore possibile gli autovalori di

$$A = \begin{bmatrix} 10 & 2 & 3 \\ -1 & 2 & -1 \\ 0 & 1 & 3 \end{bmatrix}.$$

Nota1: A ha elementi reali $\Rightarrow$ il polinomio caratteristico di A ha coefficienti reali $\Rightarrow$ se λ è autovalore di A anche $\overline{\lambda}$ è autovalore di A. Nota2: A e A^T hanno gli stessi autovalori

□ II Teorema di Gershgorin: Se m cerchi di Gershgorin di A $n \times n$ sono disgiunti dai rimanenti $n - m$, allora m autovalori sono nell'unione di quegli m cerchi e $n - m$ nell'unione di quegli altri

□ La matrice

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$$

è diagonalizzabile con una trasformazione unitaria?

□ A $n \times n$ si dice "normale" se $AA^* = A^*A$, ovvero se commuta con la sua trasposta coniugata.

(1) Se A è normale e triangolare allora A è diagonale. (sugg.: fare il caso $n = 2$. Per n generico far vedere che A $n \times n$ triangolare $\Rightarrow$

$$A = \begin{bmatrix} a_{11} & \mathbf{w}^T \\ \mathbf{0} & B \end{bmatrix}$$

con B $n - 1 \times n - 1$ triangolare, e che, se A è anche normale allora il vettore $\mathbf{w}$ deve essere nullo e la matrice B deve essere normale.

(2) Sia $A = QTQ^*$ con Q unitaria, T triangolare. Allora A è normale se e solo se T è normale. *Teorema di Schur*: Data A $n \times n$ esiste Q unitaria tale che $A = QTQ^*$ con T triangolare, ovvero ogni matrice si può trasformare in triangolare con una trasformazione per similitudine unitaria. Dimostrazione: omessa.

(3) A $n \times n$ si può diagonalizzare con una trasformazione per similitudine unitaria se e solo se è normale (sugg. usare il teorema di Schur e i risultati (1) e (2))

□

$$A = \begin{bmatrix} 9 & 1 & -2 & 1 \\ 0 & 8 & 1 & 1 \\ -1 & 0 & 7 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -9 & 1 & -2 & 0 \\ 1 & 8 & 0 & 1 \\ -2 & 0 & 7 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

i osservare che A ha almeno due autovalori reali

ii trovare un limite superiore di $\mu_2(B)$

□ Sia p un numero naturale. Se $A^p = 0$, allora gli autovalori λ di A sono tutti nulli. Se $A^p = I$, allora gli autovalori λ di A devono risolvere l'equazione algebrica $z^p - 1 = 0$

21 Aprile 2009

Come annullare tutti gli elementi di un vettore eccetto il primo

□ Dato $\mathbf{w}$ vettore $j \times 1$ trovare F matrice $j \times j$ tale che $(F\mathbf{w})_k = 0$, $k = 2, \ldots, j$, cioè:

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_j \end{bmatrix} \quad \rightarrow \quad F\mathbf{w} = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Vediamo quattro matrici F con questa proprietà:

(1) Metodo di Gauss: $F = GP$ dove P di permutazione tale che $(P\mathbf{w})_1 \neq 0$ e

$$G = \begin{bmatrix} 1 & 0 & & \cdots & 0 \\ -(P\mathbf{w})_2/(P\mathbf{w})_1 & 1 & 0 & \cdots & 0 \\ . & 0 & \ddots & & \vdots \\ \vdots & \vdots & & & 0 \\ -(P\mathbf{w})_j/(P\mathbf{w})_1 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

(2) Metodo di Gauss con pivot parziale: come (1), ma si richiede $|(P\mathbf{w})_1| = \max_{k=1,\ldots,j} |(P\mathbf{w})_k|$. Con questa scelta di P gli elementi della matrice G hanno tutti modulo minore o uguale a 1 $\Rightarrow$ se $\mathbf{z}$ è un vettore affetto da errore questo errore non è amplificato nel calcolo di $G\mathbf{z}$ $\Rightarrow$ più stabile (meno accumulazione degli errori di arrotondamento) di Gauss

(3) Metodo di Givens (F unitaria): $F = Q_j \cdots Q_3 Q_2$, Q_k matrice $j \times j$ di Givens.

$a = w_1/\sqrt{w_1^2 + w_2^2}$, $b = w_2/\sqrt{w_1^2 + w_2^2}$,

$$Q_2\mathbf{w} = \begin{bmatrix} a & b & & & \\ -b & a & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_j \end{bmatrix} = \begin{bmatrix} \sqrt{w_1^2 + w_2^2} \\ 0 \\ w_3 \\ \vdots \\ w_j \end{bmatrix}$$

$a = \sqrt{w_1^2 + w_2^2}/\sqrt{w_1^2 + w_2^2 + w_3^2}$, $b = w_3/\sqrt{w_1^2 + w_2^2 + w_3^2}$,

$$Q_3 Q_2 w = \begin{bmatrix} a & 0 & b & & & \\ 0 & 1 & 0 & & & \\ -b & 0 & a & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} \sqrt{w_1^2 + w_2^2} \\ 0 \\ w_3 \\ \vdots \\ w_j \end{bmatrix} = \begin{bmatrix} \sqrt{w_1^2 + w_2^2 + w_3^2} \\ 0 \\ 0 \\ w_4 \\ \vdots \\ w_j \end{bmatrix}$$

e cosi' via...

(4) Metodo di Householder (F unitaria): F = matrice di Householder $= I - (2/\mathbf{u}^*\mathbf{u})\mathbf{u}\mathbf{u}^*$, $\mathbf{u}$ vettore $j \times 1$ opportuno

□ Utilizzando i metodi (1),(2),(3) trasformare i vettori

$$\mathbf{w} = \begin{bmatrix} 0 \\ 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1/\sqrt{2} \\ 1/2 \\ 1/2 \end{bmatrix}$$

in un vettore del tipo

$$F\mathbf{w} = \begin{bmatrix} * \\ 0 \\ 0 \end{bmatrix}$$

Numero di condizionamento di una matrice

□ Mostrare che le seguenti due matrici

$$\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}$$

sono definite positive. Calcolare il loro numero di condizionamento

□ Usando l'identità $1/(i+j-1) = \int_0^1 x^{i+j-2} dx$, mostrare che la matrice "di Hilbert"

$$A = \begin{bmatrix} 1 & 1/2 & \cdots & 1/n \\ 1/2 & 1/3 & \cdots & 1/(n+1) \\ & & & \\ 1/n & 1/(n+1) & \cdots & 1/(2n-1) \end{bmatrix} = (1/(i+j-1))_{i,j=1\dots n}$$

è definita positiva. La matrice di Hilbert ha un numero di condizionamento che cresce molto rapidamente con n

□ Per A $n \times n$ non singolare sia $\mu(A) = \|A\|\|A^{-1}\|$ il "numero di condizionamento" di A. Dal risultato: $A\mathbf{x} = \mathbf{b}$ & $A(\mathbf{x} + d\mathbf{x}) = \mathbf{b} + d\mathbf{b} \Rightarrow \|d\mathbf{x}\|/\|\mathbf{x}\| \leq \mu(A)\|d\mathbf{b}\|/\|\mathbf{b}\|$ (valido se la norma matriciale è indotta da quella vettoriale), segue che $\mu(A)$ misura la sensibilità della soluzione di un sistema lineare a una eventuale perturbazione del vettore dei termini noti. Ad esempio la soluzione del sistema

$$A\mathbf{x} = \begin{bmatrix} 1 & 900 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

è molto sensibile a variazioni di b_1, b_2 perché $\mu(A) \approx 10^6$ e, quindi, una perturbazione su $\mathbf{b}$ può causare una grande perturbazione su $\mathbf{x}$

☐ Valgono le seguenti affermazioni:

(1) $\mu(A) \geq \max|\lambda_i|/\min|\lambda_i| \geq 1$, λ_i autovalori di A

(2) Se A è hermitiana e $\|\cdot\|$ è la norma spettrale (cioè la norma matriciale indotta dalla norma vettoriale euclidea), allora $\mu(A) = \max|\lambda_i|/\min|\lambda_i|$

(3) Se A è definita positiva e $\|\cdot\|$ è la norma spettrale, allora $\mu(A) = \max\lambda_i/\min\lambda_i$

☐ Scrivere una limitazione superiore per $\mu(A)$ essendo A la matrice tridiagonale

$$A = \begin{bmatrix} 2+h^2p_1 & -1 & & & \\ -1 & 2+h^2p_2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2+h^2p_n \end{bmatrix},$$

con p_i reali positivi per ogni i

☐ Il numero di condizionamento delle matrici

$$\begin{bmatrix} 1 & 0 \\ m_{21} & 1 \end{bmatrix}, \ |m_{21}| \leq 1, \ \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & 0 & 1 \end{bmatrix}, \ |m_{i1}| \leq 1,$$

è maggiore di 1.

☐ Il numero di condizionamento di ogni matrice unitaria, misurato con la norma spettrale, è uguale a 1

Norme matriciali

☐ Data $\|\cdot\|$ norma vettoriale, per A $n \times n$ si può usare come misura di A il numero $\|A\| = \sup_{\mathbf{z}\ n\times 1 \neq \mathbf{0}}(\|A\mathbf{z}\|/\|\mathbf{z}\|)$. Questo numero soddisfa le seguenti proprietà:

(1) $\|A\| \geq 0$

(2) $\|A\| = 0$ iff $A = 0$

(3) $\|aA\| = |a|\|A\|$

(4) $\|A + B\| \leq \|A\| + \|B\|$

(5) $\|AB\| \leq \|A\|\|B\|$

(A, B $n \times n$, a 1×1)

Ogni numero che soddisfa tali proprietà è detto “norma matriciale”

☐ Dimostrare che $\rho(A) = \max|\lambda_i|$, λ_i autovalori di A $n\times n$, non è una norma matriciale

☐ Per le norme matriciali “indotte” dalle norme vettoriali $\|\cdot\|_2$, $\|\cdot\|_1$, $\|\cdot\|_\infty$, si dimostra che $\|A\|_2 = \sqrt{\rho(A^*A)}$, $\|A\|_1 = \max_j \sum_i |a_{ij}|$, $\|A\|_\infty = \max_i \sum_j |a_{ij}|$.

Se $\|\cdot\|$ è una norma matriciale indotta da una norma vettoriale ed I è la matrice identica allora $\|I\| = 1$. Esempi di norme matriciali che non sono indotte da norme vettoriali sono: $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2} = \sqrt{\operatorname{tr}(A^*A)}$ "norma di Frobenius" e $\|A\| = n \max_{i,j} |a_{ij}|$, infatti se si misura la matrice identica con tali norme si ottengono numeri maggiori di 1. Le norme matriciali, come quelle vettoriali, sono equivalenti tra loro.

□ La maggiorazione per il raggio spettrale (di una matrice) ottenuta in precedenza si può riscrivere come segue: $\rho(A) \leq \min\{\|A\|_1, \|A\|_\infty\}$. Tale risultato si può generalizzare: per ogni norma matriciale $\|\cdot\|$ indotta da una norma vettoriale si ha $\rho(A) \leq \|A\|$. In realtà la limitazione $\rho(A) \leq \|A\|$ è vera per ogni norma matriciale $\|\cdot\|$. Quindi, ogni cerchio del piano complesso del tipo $\{z \in \mathbb{C} : \ |z| \leq \|A\|\}$ (di centro $(0,0)$ e raggio $\|A\|$) con $\|\cdot\|$ norma matriciale qualsiasi deve contenere tutti gli autovalori di A

□ Dimostrare che per A $n \times n$ il numero $\max_{i,j} |a_{i,j}|$ è una misura di A che non è una norma matriciale

□ Sia A $n \times n$ generica e Q $n \times n$ unitaria. Allora

(1) $\|Q\|_2 = 1$

(2) $\|QA\|_2 = \|AQ\|_2 = \|A\|_2$

(3) $\|Q\|_F = \sqrt{n}$

(4) $\|QA\|_F = \|AQ\|_F = \|A\|_F$

(5) $\|A\|_{2,F} = \|A^H\|_{2,F}$

(6) $(\|A\|_2)^2 \leq \|A\|_1 \|A\|_\infty$

(7) esiste A $n \times n$ tale che $\|A\|_2 > \|A\|_\infty$

Risoluzione. (5) $\|A\|_2^2 = \rho(A^*A) \leq \|A^*A\|_2 \leq \|A^*\|_2 \|A\|_2 \Rightarrow \|A\|_2 \leq \|A^*\|_2$. Si dimostra allo stesso modo che $\|A^*\|_2 \leq \|A\|_2$.

(6) $\|A\|_2^2 = \rho(A^*A) \leq \|A^*A\|_1 \leq \|A^*\|_1 \|A\|_1 = \|A\|_\infty \|A\|_1$.

(7) Sia

$$A = \begin{bmatrix} 1 & 10 \\ 1 & 10 \end{bmatrix}$$

Allora $\|A\|_\infty = 11$, $\|A\|_1 = 20$, $\|A\|_2 \geq \frac{\|A\mathbf{e}_2\|_2}{\|\mathbf{e}_2\|_2} = 10\sqrt{2} \Rightarrow \|A\|_2 > \|A\|_\infty$.

Un altro esempio. Sia $A = \alpha \mathbf{e}\mathbf{e}^T + \beta \mathbf{e}\mathbf{e}_j^T$, $\mathbf{e}^T = [1\ 1 \cdots 1]$, $\alpha \geq 0$, $\beta > 0 \Rightarrow \|A\|_\infty = n\alpha + \beta$, $\|A\|_1 = n(\alpha+\beta)$, $\|A\|_2 \geq \frac{\|A\mathbf{e}_j\|_2}{\|\mathbf{e}_j\|_2} = \sqrt{n}(\alpha+\beta)$. Quindi, se $\beta > \sqrt{n}\alpha$ allora $\|A\|_2 > \|A\|_\infty$.

Nota: per (6), ogni volta che $\|A\|_2 > \|A\|_\infty$ deve essere necessariamente verificata la disuguaglianza $\|A\|_2 < \|A\|_1$.

□ Calcolare il numero $\mu_1(A) = \|A\|_1 \|A^{-1}\|_1$ per

$$A = \begin{bmatrix} 1 & 900 \\ 0 & 1 \end{bmatrix}$$

7 Aprile 2009, 4 Aprile 2009, 22 Marzo 2009, 15 Marzo 2009, 7 Marzo 2009

7 aprile era perFra

4 Aprile 2009

☐ Sia $A = \mathbf{u}\mathbf{v}^*$, con $\mathbf{u}, \mathbf{v}$ vettori $n \times 1$. Dimostrare:

(1) $\sqrt{\rho(A^*A)} = \|\mathbf{u}\|_2 \|\mathbf{v}\|_2$

(2) $\max_i \sum_j |a_{ij}| = \|\mathbf{u}\|_\infty \|\mathbf{v}\|_1$

(3) $\max_j \sum_i |a_{ij}| = \|\mathbf{u}\|_1 \|\mathbf{v}\|_\infty$

(4) $\rho(A) = |\mathbf{v}^*\mathbf{u}| \leq \min\{\|\mathbf{u}\|_2\|\mathbf{v}\|_2, \|\mathbf{u}\|_1\|\mathbf{v}\|_\infty, \|\mathbf{u}\|_\infty\|\mathbf{v}\|_1\}$

(5) Esistono $\mathbf{u}, \mathbf{v}$ per cui $|\mathbf{v}^*\mathbf{u}| \leq \|\mathbf{u}\|_1\|\mathbf{v}\|_\infty < \|\mathbf{u}\|_2\|\mathbf{v}\|_2$

☐ Usando il I teorema di Gershgorin osservare che esiste un intervallo positivo della retta reale che contiene i due autovalori di

$$A = \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix}$$

(non calcolare gli autovalori). Essendo A anche hermitiana ne segue che A è definita positiva.

☐ E' possibile stabilire usando il I teorema di Gershgorin se

$$A = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}$$

è definita positiva? Dimostrare che A è definita positiva.

☐ Data

$$A = \begin{bmatrix} 1 & 1/4 & 2 & 1 \\ 1/2 & 2 & 0 & 1-\sqrt{2} \\ 0 & 1/2 & -3 & 0 \\ -\mathbf{i}/2 & 1/2 & 1 & 1+\mathbf{i} \end{bmatrix}$$

disegnare i seguenti insiemi, tutti contenenti gli autovalori di A:

(1) $\{z : \ |z| \leq \min\{\max_i \sum_j |a_{ij}|, \max_j \sum_i |a_{ij}|\}\}$

(2) $U_{i=1\ldots n} K_i$, K_i cerchi di Gershgorin di A

(3) $U_{i=1\ldots n} H_i$, H_i cerchi di Gershgorin di A^T

(4) $(U_{i=1\ldots n} K_i) \bigcap (U_{i=1\ldots n} H_i)$

Si può stabilire se A è non singolare senza calcolare il suo determinante?

☐ I T. di Gershgorin: Sia A $n \times n$ generica. Gli autovalori di A sono nell'insieme $U_{i=1\ldots n} K_i$, $K_i = \{z : \ |z - a_{ii}| \leq \sum_{j, j\neq i} |a_{ij}|\}$

☐ Rafforzare il I T.di Gershgorin usando l'osservazione che A e A^T hanno gli stessi autovalori.

☐ Localizzare gli autovalori della matrice

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1/2 & 1 \\ 4 & 2 & 1 \end{bmatrix}$$

☐ Dimostrare, senza calcolare gli autovalori, che la matrice

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$$

ha autovalori reali

☐ Sia Q unitaria $n \times n$. Dimostrare l'identità $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$, ovvero applicando una matrice unitaria a un vettore la misura in norma euclidea del vettore non cambia.

☐ Sia $\mathbf{w}$ $n \times 1$. Se F $n \times n$ unitaria è tale che

$$Fw = \begin{bmatrix} a \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

cioè $(F\mathbf{w})_i = 0$, $i = 2 \ldots n$, allora $|a| = \|w\|_2$.

☐ Dare una maggiorazione per $\rho(A)$ essendo

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1/2 & 1 \\ 4 & 2 & 1 \end{bmatrix}$$

☐ Scrivere l'inversa della matrice A $n \times n$

$$a_{i,i} = 1,\ i = 1 \ldots n,\ a_{i,i+1} = -1,\ i = 1 \ldots n-1,\ a_{n,1} = a,\ a_{i,j} = 0 \text{ altrimenti.}$$

☐ Sia $A = A^*$ $n \times n$. Allora le seguenti affermazioni sono equivalenti:

(1) $\mathbf{z}^* A \mathbf{z} > 0$ per ogni vettore $\mathbf{z}$ non nullo

(2) gli autovalori di A sono positivi

(3) le sottomatrici di A

$$\begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}$$

hanno determinante positivo per $k = 1, \ldots, n$.

(A tale che $A = A^*$ e vale (1) si dice definita positiva).

Vedremo che in certi casi è possibile dire che gli autovalori sono positivi senza calcolarli e questo, per la (2), è sufficiente per stabilire che A è definita positiva

☐ (PER ME) Sia A la matrice 2×2

$$A = \begin{bmatrix} c & a \\ \overline{a} & b \end{bmatrix},$$

a numero complesso, b, c reali. Si osserva che A è hermitiana.

1) Condizione necessaria affinché A sia definita positiva è che b e c siano positivi.

2) Condizione necessaria e sufficiente affinché A sia definita positiva è che $c > 0$, $b/c > 0$, $\sqrt{b/c} - |a/c| > 0$ ovvero che $c > 0$ & $bc - |a|^2 > 0$.

Suggerimento: applicare il risultato ottenuto nell'esercizio precedente

Risoluzione

$$A = \begin{bmatrix} c & a \\ \overline{a} & b \end{bmatrix}, \ b, c \in \mathbb{R}, \quad A = c \begin{bmatrix} 1 & a/c \\ \overline{a/c} & b/c \end{bmatrix}$$

definita positiva iff

$$c > 0 \ \ \& \ \ \begin{bmatrix} 1 & a/c \\ \overline{a/c} & b/c \end{bmatrix} \ \text{p.d.}$$

iff $c > 0$, $b/c > 0$, $\sqrt{b/c} - |a/c| > 0$ iff $c > 0$ & $bc - |a|^2 > 0$

□ (PER ME) Sia A la matrice 2×2

$$A = \begin{bmatrix} 1 & a \\ \overline{a} & b \end{bmatrix},$$

a numero complesso, b reale. Si osserva che A è hermitiana

1) Condizione necessaria affinché A sia definita positiva è che b sia positivo.

2) Condizione necessaria e sufficiente affinché A sia definita positiva è che b sia positivo & $\sqrt{b} - |a| > 0$ o, equivalentemente, che $b - |a|^2$ sia positivo.

Trovare una formula per $\mathbf{z}^* A\mathbf{z}$ che consenta di dedurre tale affermazione senza calcolare esplicitamente gli autovalori.

Risoluzione

a, b numeri complessi $\Rightarrow$

$$\mathbf{z}^* A\mathbf{z} = |z_1|^2 + b|z_2|^2 + z_1\overline{z_2 a} + \overline{z_1} z_2 a.$$

b reale positivo $\Rightarrow$

$$\mathbf{z}^* A\mathbf{z} = (1 - \frac{|a|}{\sqrt{b}})(|z_1|^2 + b|z_2|^2) + \frac{|a|}{\sqrt{b}}|z_1 + \sqrt{b}e^{\mathbf{i}\arg(a)} z_2|^2.$$

Ne segue che A è definita positiva se e solo se $b > 0$ & $\sqrt{b} - |a| > 0$ se e solo se $b - |a|^2 > 0$

□ La matrice

$$A = \begin{bmatrix} 1 & a \\ \overline{a} & 1 \end{bmatrix},$$

a numero complesso, è definita positiva se e solo se $1 - |a| > 0$ (o, equivalentemente, se e solo se $1 - |a|^2 > 0$). Trovare una formula per $\mathbf{z}^* A\mathbf{z}$ che consenta di dedurre (senza calcolare esplicitamente gli autovalori) tale affermazione. Suggerimento: studiare i casi a reale negativo (vedi $a = -1/2$ del precedente esercizio), poi a reale positivo e infine a generico.

Risoluzione

$$\mathbf{z}^* A\mathbf{z} = |z_1|^2 + |z_2|^2 + z_1\overline{z_2 a} + z_2\overline{z_1} a.$$

a reale positivo: $\mathbf{z}^* A\mathbf{z} = (1 - a)(|z_1|^2 + |z_2|^2) + a|z_1 + z_2|^2$,

a reale negativo: $\mathbf{z}^* A\mathbf{z} = (1 + a)(|z_1|^2 + |z_2|^2) - a|z_1 - z_2|^2$,

a generico: $\mathbf{z}^* A\mathbf{z} = (1-|a|)(|z_1|^2+|z_2|^2) + |a||z_1 + e^{\mathbf{i}\arg(a)} z_2|^2$.

Ne segue che A è definita positiva se e solo se $1-|a|>0$.

☐ Dimostrare, senza calcolare gli autovalori, che la matrice

$$A = \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix}$$

è definita positiva.

Risoluzione: 1) A è reale simmetrica, quindi è hermitiana. 2) Inoltre,

$$\begin{aligned}
\mathbf{z}^* A\mathbf{z} &= [\overline{z_1}\ \overline{z_2}] \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = [\overline{z_1} - 1/2\overline{z_2}\ \ -1/2\overline{z_1} + \overline{z_2}] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\
&= |z_1|^2 - 1/2 z_1\overline{z_2} - 1/2\overline{z_1}z_2 + |z_2|^2 = 1/2(|z_1|^2+|z_2|^2) + 1/2(z_1-z_2)(\overline{z_1}-\overline{z_2}) \\
&= 1/2(|z_1|^2+|z_2|^2) + 1/2|z_1-z_2|^2 >= 1/2(|z_1|^2+|z_2|^2).
\end{aligned}$$

Ne segue che $\mathbf{z}^* A\mathbf{z} > 0$ per ogni vettore $\mathbf{z} = [z_1\ z_2]^T$ non nullo.

Norme vettoriali:

☐ a scalare (1×1), $|a|$ = modulo di a = misura della distanza tra i punti $(0,0)$ e $(\Re(a), \Im(a)) = \sqrt{\Re(a)^2 + \Im(a)^2}$.

☐ $\mathbf{z}$ $n\times 1$, $\|\mathbf{z}\| = \|\mathbf{z}\|_2 = (\sum_i |z_i|^2)^{1/2}$ = misura "euclidea" di $\mathbf{z}$. Proprietà:

(1) $\|\mathbf{z}\| \geq 0$ per ogni $\mathbf{z}$

(2) $\|\mathbf{z}\| = 0$ iff $\mathbf{z} = \mathbf{0}$

(3) $\|a\mathbf{z}\| = |a|\|\mathbf{z}\|$, per ogni a scalare e $\mathbf{z}$

(4) $\|\mathbf{z}+\mathbf{w}\| \leq \|\mathbf{z}\| + \|\mathbf{w}\|$, per ogni $\mathbf{z}, \mathbf{w}$

Ogni misura che soddisfa tali proprietà è detta "norma vettoriale".

Ad esempio sono norme vettoriali anche $\|\mathbf{z}\| = \|\mathbf{z}\|_1 = \sum_i |z_i|$ e $\|\mathbf{z}\| = \|\mathbf{z}\|_\infty = \max_i |z_i|$ (verificarlo).

Per ogni $\mathbf{z}$ $n\times 1$ non nullo il vettore $\mathbf{y} = \mathbf{z}/\|\mathbf{z}\|$ ha norma uguale a 1, infatti, per la terza proprietà delle norme vettoriali, $\|\mathbf{z}/\|\mathbf{z}\|\| = (1/\|\mathbf{z}\|)\|\mathbf{z}\| = 1$.

Esempi:

$$\mathbf{z} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \Rightarrow \|\mathbf{z}\|_2 = \frac{1}{\sqrt{2}},\ \|\mathbf{z}\|_1 = 1,\ \|\mathbf{z}\|_\infty = 1/2$$

$$\Rightarrow \mathbf{y} = \mathbf{z}/\|\mathbf{z}\|_2 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix},\ \mathbf{y} = \mathbf{z}/\|\mathbf{z}\|_1 = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix},\ \mathbf{y} = \mathbf{z}/\|\mathbf{z}\|_\infty = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

E' semplice far vedere che, per ogni vettore $\mathbf{z}$ $n\times 1$, $(1/n)\|\mathbf{z}\|_1 \leq \|\mathbf{z}\|_\infty \leq \|\mathbf{z}\|_1$.

Più in generale comunque prese due norme vettoriali $\|\cdot\|_a$ e $\|\cdot\|_b$ queste sono equivalenti, cioè esistono c_1 e c_2 costanti positive per cui $c_1\|\mathbf{z}\|_b \leq \|\mathbf{z}\|_a \leq c_2\|\mathbf{z}\|_b$ per ogni $\mathbf{z}$. Ne segue che, se un vettore ha misura grande (piccola) in una norma allora ha misura grande (piccola) in ogni altra norma. Inoltre, se per una certa norma $\|\cdot\|_b$ si ha che $\|\mathbf{z}_k - \mathbf{z}\|_b \to 0$ (per k che tende a infinito) allora sarà anche vero che $\|\mathbf{z}_k - \mathbf{z}\| \to 0$ in ogni norma vettoriale $\|\cdot\|$.

Due limitazioni superiori per il raggio spettrale:

$\square$
$$A\ n\times n\ \Rightarrow \rho(A)\le \max_i \sum_j |a_{ij}|,\quad \rho(A)\le \max_j \sum_i |a_{ij}|$$

Dim: $A\mathbf{v}=\lambda\mathbf{v}\Rightarrow \sum_j a_{ij}v_j=\lambda v_i$ per ogni $i\Rightarrow |\lambda||v_i|\le\sum_j|a_{ij}||v_j|\le \|\mathbf{v}\|_\infty\sum_j|a_{ij}|$ per ogni $i\Rightarrow$ esiste k tale che $|\lambda|\le\sum_j|a_{kj}|\le\max_i\sum_j|a_{ij}|$. La seconda disuguaglianza si ottiene applicando la prima ad A^T ed osservando che $\rho(A)=\rho(A^T)$ (gli autovalori di A^T coincidono con quelli di A).

Teoria(metododellepotenze) [per il calcolo dell'autovalore in modulo dominante di una matrice e del corrispondente autovettore]:

$A\ n\times n$ diagonalizzabile, ovvero $A\mathbf{v}_i=\lambda_i\mathbf{v}_i$, $i=1,\dots,n$, con $\mathbf{v}_1,\dots,\mathbf{v}_n$ linearmente indipendenti. $|\lambda_1|>|\lambda_j|$ per ogni $j\ne1$, ovvero c'è un solo autovalore il cui modulo è uguale al raggio spettrale di A e tale autovalore è semplice (cioè è radice semplice del polinomio caratteristico di A). $\mathbf{x}\ n\times1$ tale che a_1 in $\mathbf{x}=a_1\mathbf{v}_1+a_2\mathbf{v}_2+\ldots+a_n\mathbf{v}_n$ è diverso da zero. In queste ipotesi si ha, per k che tende a infinito,

- $(1/\lambda_1^k)A^k\mathbf{x}\to a_1\mathbf{v}_1$, quindi $(e^{-\mathbf{i}a})^kA^k\mathbf{x}/\|A^k\mathbf{x}\|\to a_1\mathbf{v}_1/\|a_1\mathbf{v}_1\|$ dove a è l'argomento di λ_1 ($\lambda_1=|\lambda_1|e^{\mathbf{i}a}$).
- per ogni $\mathbf{v}$, $\mathbf{v}^T\mathbf{v}_1\ne0$, $\mathbf{v}^TA^{k+1}\mathbf{x}/\mathbf{v}^TA^k\mathbf{x}\to\lambda_1$ (nota: $\mathbf{v}^T\mathbf{v}_1\ne0\Rightarrow \mathbf{v}^TA^k\mathbf{x}\ne0$ per k grande).

Ne segue che le successioni $\{\mathbf{z}_k\},\{\mathbf{y}_k\}$ generate dall'algoritmo:

$\mathbf{y}_{-1}=\mathbf{x}$
Per $k=0,1\ldots$
$\mathbf{z}_k=\mathbf{y}_{k-1}/\|\mathbf{y}_{k-1}\|,\ \mathbf{y}_k=A\mathbf{z}_k$ (NOTA : $\mathbf{z}_k=A^k\mathbf{x}/\|A^k\mathbf{x}\|$)

soddisfano le proprietà:

(1) $(e^{-\mathbf{i}a})^k\mathbf{z}_k\to a_1\mathbf{v}_1/\|a_1\mathbf{v}_1\|$

(2) $\mathbf{v}^T\mathbf{y}_k/\mathbf{v}^T\mathbf{z}_k\to\lambda_1$ ($\mathbf{v}^T\mathbf{v}_1\ne0\Rightarrow\mathbf{v}^T\mathbf{z}_k\ne0$ per k grande)

Dim della NOTA. Per induzione:

$$\mathbf{z}_0=\frac{\mathbf{y}_{-1}}{\|\mathbf{y}_{-1}\|}=\frac{\mathbf{x}}{\|\mathbf{x}\|},\quad ok!$$

$$\mathbf{z}_{k+1}=\frac{\mathbf{y}_k}{\|\mathbf{y}_k\|}=\frac{A\mathbf{z}_k}{\|A\mathbf{z}_k\|}=\frac{AA^k\mathbf{x}}{\|AA^k\mathbf{x}\|}=\frac{A^{k+1}\mathbf{x}}{\|A^{k+1}\mathbf{x}\|},\quad ok!$$

Dim di (1) e (2). Usare la NOTA. Ad esempio, per (2):

$$\frac{\mathbf{v}^T\mathbf{y}_k}{\mathbf{v}^T\mathbf{z}_k}=\frac{\mathbf{v}^TA\mathbf{z}_k}{\mathbf{v}^T\mathbf{z}_k}=\frac{\mathbf{v}^TA^{k+1}\mathbf{x}/\|A^k\mathbf{x}\|}{\mathbf{v}^TA^k\mathbf{x}/\|A^k\mathbf{x}\|}=\frac{\mathbf{v}^TA^{k+1}\mathbf{x}}{\mathbf{v}^TA^k\mathbf{x}}.$$

Esercizi:

$\square$ Sia A la matrice
$$\begin{bmatrix}1 & -1/2\\ -1/2 & 1\end{bmatrix}.$$

Scrivere un algoritmo per il calcolo dell'autovalore di modulo massimo e del corrispondente autovettore (usare il metodo delle potenze con $\|\mathbf{z}\| = \sum_i |z_i|$)

Risoluzione: A è diagonalizzabile e $|\lambda_1| > |\lambda_2|$: $A = VDV^{-1}$,

$$D = \begin{bmatrix} 1/2 & 0 \\ 0 & 3/2 \end{bmatrix}, V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

(Nota: $V = V^* = V^{-1}$). In altri termini $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, $A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\lambda_1 = 3/2$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\lambda_2 = 1/2$ e $\mathbf{v}_1, \mathbf{v}_2$ sono linearmente indipendenti. Sia $\mathbf{x}$ il vettore $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. a_1 in $\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2$ è diverso da zero, infatti

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Quindi, se $\mathbf{y}_{-1} = \mathbf{x}$ e per $k = 0, 1, \dots$ si definisce la coppia di vettori $\mathbf{z}_k = \mathbf{y}_{k-1}/\|\mathbf{y}_{k-1}\|$, $\mathbf{y}_k = A\mathbf{z}_k$ (nota: $\mathbf{z}_k = A^k\mathbf{x}/\|A^k\mathbf{x}\|$) allora, per k che tende a infinito,

$$\mathbf{z}_k \to a_1\mathbf{v}_1/\|a_1\mathbf{v}_1\| = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix}, \mathbf{v}^T\mathbf{y}_k/\mathbf{v}^T\mathbf{z}_k \to \lambda_1 = 3/2 \;\; \forall\, \mathbf{v},\, \mathbf{v}^T\mathbf{v}_1 \neq 0$$

(nota: $\mathbf{v}^T\mathbf{v}_1 \neq 0 \Rightarrow \mathbf{v}^T\mathbf{z}_k \neq 0$ per k grande).

□ Eseguiamo i primi passi con $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ($\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ non ok perché $\mathbf{v}^T\mathbf{v}_1 = 0$)

$$\begin{aligned}
\mathbf{z}_0 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; \mathbf{y}_0 = \begin{bmatrix} 1 \\ -1/2 \end{bmatrix}, \; \mathbf{v}^T\mathbf{y}_0/\mathbf{v}^T\mathbf{z}_0 = 1 \\
\mathbf{z}_1 &= \begin{bmatrix} 2/3 \\ -1/3 \end{bmatrix}, \; \mathbf{y}_1 = \begin{bmatrix} 5/6 \\ -4/6 \end{bmatrix}, \; \mathbf{v}^T\mathbf{y}_1/\mathbf{v}^T\mathbf{z}_1 = 5/4 \\
\mathbf{z}_2 &= \begin{bmatrix} 5/9 \\ -4/9 \end{bmatrix}, \; \mathbf{y}_2 = \begin{bmatrix} 14/18 \\ -13/18 \end{bmatrix}, \; \mathbf{v}^T\mathbf{y}_2/\mathbf{v}^T\mathbf{z}_2 = 14/10
\end{aligned}$$

...

□ Ripetere l'esercizio precedente per le matrici

$$\begin{bmatrix} 1 & -1/2 \\ -4 & 1 \end{bmatrix}, \;\; \begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix}.$$

□ Sia A $n \times n$ definita positiva e λ_j, $j = 1, \dots, n$, i suoi autovalori. Nell'ipotesi $\lambda_1 > \lambda_2 \geq \dots \geq \lambda_{n-1} > \lambda_n$ proporre un algoritmo per il calcolo del numero $\max \lambda_j / \min \lambda_j$

□ Sia A $n \times n$ hermitiana invertibile e λ_j, $j = 1, \dots, n$, i suoi autovalori. Proporre un algoritmo per il calcolo del numero $\max |\lambda_j| / \min |\lambda_j|$

□ (PER ME) Sia A $n \times n$ hermitiana e $\lambda_1, \lambda_2, \dots, \lambda_n$ i suoi autovalori dove λ_1 si suppone diverso da 0. Sia $\mathbf{y}_1$ $n \times 1$ non nullo tale che $A\mathbf{y}_1 = \lambda_1\mathbf{y}_1$ e $W = A - (\lambda_1/\mathbf{y}_1^*\mathbf{y}_1)\mathbf{y}_1\mathbf{y}_1^*$. Allora W è hermitiana e i suoi autovalori sono $0, \lambda_2, \dots, \lambda_n$. Dedurre da questo risultato un algoritmo per il calcolo di tutti gli autovalori di A

□ (PER ME) Sia A $n \times n$ e $\lambda_1, \lambda_2, \ldots, \lambda_n$ i suoi autovalori dove λ_1 si suppone diverso da 0. Sia $\mathbf{y}_1$ $n \times 1$ non nullo tale che $A\mathbf{y}_1 = \lambda_1 \mathbf{y}_1$ e $W = A - (\lambda_1/\mathbf{w}^*\mathbf{y}_1)\mathbf{y}_1\mathbf{w}^*$ con $\mathbf{w}$ $n \times 1$ tale che $\mathbf{w}^*\mathbf{y}_1 \neq 0$. Allora gli autovalori di W sono, per ogni scelta di $\mathbf{w}$, $0, \lambda_2, \ldots, \lambda_n$.

ESEMPIO. Se $\mathbf{w}^* = \mathbf{e}_i^T A$ = riga i di A con i tale che $(\mathbf{y}_1)_i \neq 0$, allora $\mathbf{w}^*\mathbf{y}_1 = \lambda_1(\mathbf{y}_1)_i \neq 0$, quindi gli autovalori di $W = A - (1/(\mathbf{y}_1)_i)\mathbf{y}_1\mathbf{e}_i^T A$ sono $0, \lambda_2, \ldots, \lambda_n$. Notare che per tale scelta di $\mathbf{w}$, la riga i di W è nulla, cioè

$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ 0..0 & 0 & 0..0 \\ W_{31} & W_{32} & W_{33} \end{bmatrix}$$

con W_{11} $i-1 \times i-1$, W_{12} $i-1 \times 1$, W_{32} $n-i \times 1$, W_{33} $n-i \times n-i$. Ne segue che $\lambda_2, \ldots, \lambda_n$ sono gli autovalori della matrice $n-1 \times n-1$

$$\begin{bmatrix} W_{11} & W_{13} \\ W_{31} & W_{33} \end{bmatrix}.$$

FINE ESEMPIO

Si verifica inoltre che $W\mathbf{y}_1 = 0$, cioè $\mathbf{y}_1$ (autovettore di A relativo a λ_1) è autovettore di W relativo a 0. Se poi $\mathbf{y}_i$ son tali che $A\mathbf{y}_i = \lambda_i \mathbf{y}_i$, $i = 2, \ldots, n$, allora per gli i tali che $\lambda_i \neq 0, \lambda_1$, posto $\mathbf{w}_i = \mathbf{y}_i - (\lambda_1/\lambda_i)(\mathbf{w}^*\mathbf{y}_i/\mathbf{w}^*\mathbf{y}_1)\mathbf{y}_1$ si ha $W\mathbf{w}_i = \lambda_i \mathbf{w}_i$. Però, se A è singolare ci sono valori di i in $2, \ldots, n$ per cui $\lambda_i = 0$. In questo caso $A\mathbf{y}_i = \mathbf{0} \Rightarrow W\mathbf{w}_i = \mathbf{0}$ con $\mathbf{w}_i = ?$ ... Inoltre, λ_1 può essere radice multipla del polinomio caratteristico di A, cioè ci possono essere valori di i in $2, \ldots, n$ tali che $\lambda_i = \lambda_1$. In questo caso $A\mathbf{y}_i = \lambda_1 \mathbf{y}_i$ $i \neq 1 \Rightarrow W\mathbf{w}_i = \lambda_1 \mathbf{w}_i$ con $\mathbf{w}_i = ?$ ...

Teoria(matrici stocastiche):

Se A $n \times n$ è stocastica per colonne, cioè se $a_{ij} \geq 0$ per ogni i, j e $\sum_i a_{ij} = 1$ per ogni j, allora:

(1) 1 è autovalore di A, cioè esiste $\mathbf{z}$ non nullo tale che $A\mathbf{z} = \mathbf{z}$ (ma è banale calcolarlo solo se A è anche stocastica per righe)

(2) $\rho(A) = 1$, cioè $1 \geq |\lambda|$ per ogni autovalore λ di A

(3) per ogni vettore $\mathbf{x}$ $n \times 1$ con tutte le componenti non negative e di somma uguale a 1 si ha che $A\mathbf{x} \geq \mathbf{0}$ & $\sum_i (A\mathbf{x})_i = 1$

(4) Teoria(metododellepotenzematricistocastiche):

se A, inoltre, è diagonalizzabile,
se $\lambda_1 = 1 > |\lambda_j|$ per ogni $j \neq 1$,
se $\mathbf{x}$ è tale che a_1 in $\mathbf{x} = a_1\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n$ è diverso da 0,
se $\|\cdot\|$ è la norma $\|\mathbf{z}\| = \sum_i |z_i|$, e
se il vettore iniziale $\mathbf{x}$ è scelto tale che $x_i \geq 0$ & $\sum_i x_i = \|\mathbf{x}\| = 1$,

ALLORA la successione $\{\mathbf{z}_k\}$ generata dall'algoritmo:

$$\mathbf{z}_0 = \mathbf{x}, \quad \mathbf{z}_{k+1} = A\mathbf{z}_k, \; k = 0, 1 \ldots$$

soddisfa le proprietà:

- $\mathbf{z}_k = A^k\mathbf{x} \geq 0$ e $\|\mathbf{z}_k\| = 1$ per ogni $k = 0, 1, \dots$ e, per k che tende a infinito, $\mathbf{z}_k \to a_1\mathbf{v}_1/\|a_1\mathbf{v}_1\| = \mathbf{z}$, $A\mathbf{z} = \mathbf{z}$, $\mathbf{z} \geq \mathbf{0}$, $\|\mathbf{z}\| = 1$,
- $\mathbf{e}^T\mathbf{z}_{k+1}/\mathbf{e}^T\mathbf{z}_k = 1/1 \to 1$ ($\mathbf{e}^T = [1\ 1\ \dots\ 1]$).

Esercizi:

☐ Sia

$$A = \begin{bmatrix} 1/2 & 0 & 1/3 \\ 0 & 1/2 & 2/3 \\ 1/2 & 1/2 & 0 \end{bmatrix}.$$

Osservare che per $\mathbf{z}_0 = \mathbf{x} = [1/3\ 1/3\ 1/3]^T$ si ottiene

$$\mathbf{z}_1 = A\mathbf{z}_0 = \begin{bmatrix} 5/18 \\ 7/18 \\ 1/3 \end{bmatrix}, \quad \mathbf{z}_2 = A\mathbf{z}_1 = \begin{bmatrix} 3/12 \\ 5/12 \\ 4/12 \end{bmatrix}, \dots$$

quindi, effettivamente $\mathbf{z}_k$ sta convergendo al vettore $\mathbf{z}$ tale che $A\mathbf{z} = \mathbf{z}$, cioè $\mathbf{z} = [2/9\ 4/9\ 3/9]^T$.

Verificare che tutte le ipotesi della teoria sulla convergenza sono soddisfatte (gli autovalori di A sono: $1, 1/2, -1/2 \dots$)

☐ Ripetere l'esercizio precedente per

$$A = \begin{bmatrix} 1/6 & 1/2 & 1/3 \\ 1/6 & 1/4 & 1/3 \\ 2/3 & 1/4 & 1/3 \end{bmatrix}$$

Teoria(Google)

[In un modello per la ricerca di documenti nel web è determinante una corretta valutazione delle importanze delle sue pagine. Per eseguirla è sufficiente calcolare l'autovettore relativo all'autovalore dominante di una matrice, solo che questa matrice è $n \times n$ con n = numero delle pagine del web(...)]:

- $G = \{1, 2, \dots, n\}$ le pagine del web
- p_i = importanza della pagina i
- $\deg(j)$ = numero delle pagine puntate da j

$p_i = \sum_{j:j\to i} p_j/\deg(j)$ (p_i è proporzionale alle importanze delle pagine che la puntano e inversamente proporzionale al numero delle pagine puntate da queste).

$a_{ij} = 1/\deg(j)$ se $j \to i$, $a_{ij} = 0$ altrimenti $\Rightarrow p_i = \sum_{j=1\dots n} a_{ij}p_j$, cioè, in forma vettoriale, $\mathbf{p} = A\mathbf{p}$. Quindi, per determinare il vettore $\mathbf{p} = [p_1\ p_2 \dots p_n]^T$ delle importanze delle pagine web è sufficiente trovare l'autovettore di A relativo all'autovalore 1.

Se A fosse stocastica per colonne e soddisfacesse le ipotesi in Teoria(metododellepotenzematricistocastiche), $\mathbf{p}$ si potrebbe calcolare con il metodo delle potenze. Questo in generale non è vero (ad esempio, è semplice vedere che $\sum_i a_{ij} = 1$ se $\deg(j) > 0$ e $\sum_i a_{ij} = 0$ altrimenti, cioè se vi sono pagine j che non puntano a nulla A non è nemmeno stocastica per colonne). Quindi A va sostituita con una matrice A' che soddisfa le proprietà dette, con il metodo delle potenze si calcola $\mathbf{p}'$ tale che $\mathbf{p}' = A'\mathbf{p}'$ e si usa $\mathbf{p}'$ come vettore delle importanze (...)

22 Marzo 2009

□ $\mathbf{e}_1$ vettore $n \times 1$ con elementi tutti nulli eccetto il primo che è uguale a 1, $\mathbf{y}$ vettore $n \times 1$ con prima componente nulla. Allora $(I + \mathbf{y}\mathbf{e}_1^T)^{-1} = I - \mathbf{y}\mathbf{e}_1^T$

□ Scrivere l'inversa delle matrici

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/3 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 1 & 0 \\ 1/2 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 1/2 & 0 & 3 \end{bmatrix}$$

□ Il prodotto di due matrici T, U triangolari superiori (inferiori) è triangolare superiore (inferiore). Una matrice T triangolare è invertibile se e solo se i suoi elementi diagonali sono tutti non nulli. L'inversa di una matrice T triangolare superiore (inferiore) invertibile è triangolare superiore (inferiore). Inoltre gli elementi diagonali di T^{-1} sono gli inversi degli elementi diagonali di T. (T, U matrici $n \times n$)

□ Trovare i valori di a numero complesso per cui la matrice A 2×2

$$\begin{bmatrix} a & a \\ a & -a \end{bmatrix}$$

è unitaria

□ A hermitiana iff esiste Q tale che $Q^{-1}AQ = D$, $Q^* = Q^{-1}$, D diagonale reale (dimostrare almeno una delle due implicazioni)

□ Dimostrare che le matrici A

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}, \begin{bmatrix} 2 & \mathbf{i}+4 \\ \mathbf{i}+4 & 7 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1/2 & 0 & 1/3 \\ 0 & 1/2 & 2/3 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

sono diagonalizzabili, cioè esiste S non singolare tale che $S^{-1}AS = D$, D diagonale. Dire quali tra di esse è diagonalizzabile tramite trasformazione unitaria, cioè esiste S, $S^H = S^{-1}$, per cui $S^{-1}AS = D$, con D diagonale. Dire quali di esse hanno autovalori tutti reali.

□ Mostrare che le matrici

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

non sono diagonalizzabili. Calcolare A^k, B^k, $k = 1, 2, \ldots$ Dimostrare che ogni matrice A $n \times n$ hermitiana che abbia tutti gli autovalori nulli deve essere la matrice nulla.

□ Descrivere i seguenti insiemi di numeri complessi z: $\mathcal{I}_1 = \{z : \ |z-1| = 1\}$, $\mathcal{I}_2 = \{z : \ |z+1| = 3\}$, $\mathcal{I}_3 = \{z : \ |z-1| > 1\}$, $\mathcal{I}_4 = \{z : \ |z+1| < 3\}$, $\mathcal{I}_3 \bigcup \mathcal{I}_4$, $\mathcal{I}_3 \bigcap \mathcal{I}_4$, $\mathcal{I}_1 \bigcap \mathcal{I}_2$, $\mathcal{I}_2 \bigcup \mathcal{I}_4$, $\partial \mathcal{I}_4$, $\overline{\mathcal{I}_4}$.

□ Se $A = I + \mathbf{y}\mathbf{w}^*$ è unitaria allora $\mathbf{w} = a\mathbf{y}$ con a complesso

□ I valori di b numero complesso per cui la matrice $A = I + b\mathbf{y}\mathbf{y}^*$ è unitaria sono $b = -2\cos(a)e^{\mathbf{i}a}/\mathbf{y}^*\mathbf{y}$, a reale. Per $a = 0$ (oppure $a = \pi$) si ottengono, in particolare, le matrici di Householder

□ Usando la formula $(I + \mathbf{y}\mathbf{w}^*)^{-1} = I - (1/(\mathbf{w}^*\mathbf{y} + 1))\mathbf{y}\mathbf{w}^*$, trovare una formula per $(A + \mathbf{y}\mathbf{w}^*)^{-1}$, A $n \times n$ non singolare (Sherman-Morrison). Suggerimento: $(A + \mathbf{y}\mathbf{w}^*)^{-1} = (A(I + A^{-1}\mathbf{y}\mathbf{w}^*))^{-1}$

□ Se $S^{-1}MS = \operatorname{diag}(a, b, 0 \dots 0)$, $S^* = S^{-1}$, allora esistono $\mathbf{y}$, $\mathbf{w}$ tali che $M = \mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^*$.

□ (PER ME) $A\mathbf{x} = \mathbf{b}$, A hermitiana con autovalori qualsiasi, Sistema equivalente: $R\mathbf{x} = \mathbf{c}$, R hermitiana con autovalori tutti uguali a 1 eccetto due, $1 + a$ e $1 + b \Rightarrow R - I$ hermitiana con autovalori tutti nulli eccetto due, a e $b \Rightarrow R - I = \mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^* \dots$

□ Siano $\mathbf{y}$, $\mathbf{w}$ vettori indipendenti.

$$(\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^*)(a\mathbf{y} + b\mathbf{w}) = c(a\mathbf{y} + b\mathbf{w}),$$

con a, b non entrambi nulli, se e solo se

$$\begin{aligned} a(\mathbf{y}^*\mathbf{y} - c) + b(\mathbf{y}^*\mathbf{w}) = 0, \\ a(\mathbf{w}^*\mathbf{y}) + b(\mathbf{w}^*\mathbf{w} - c) = 0, \end{aligned}$$

con a, b non entrambi nulli, se e solo se il determinante della matrice 2×2

$$\begin{bmatrix} \mathbf{y}^*\mathbf{y} - c & \mathbf{y}^*\mathbf{w} \\ \mathbf{w}^*\mathbf{y} & \mathbf{w}^*\mathbf{w} - c \end{bmatrix} \tag{$eq(c)$}$$

è zero. In altre parole, i due autovalori significativi di $\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^*$ (cioè quelli non nulli) sono i due autovalori della matrice

$$\begin{bmatrix} \mathbf{y}^*\mathbf{y} & \mathbf{y}^*\mathbf{w} \\ \mathbf{w}^*\mathbf{y} & \mathbf{w}^*\mathbf{w} \end{bmatrix}$$

e cioè $c_\pm = \frac{1}{2}(\mathbf{y}^*\mathbf{y} + \mathbf{w}^*\mathbf{w} \pm \sqrt{(\mathbf{y}^*\mathbf{y} - \mathbf{w}^*\mathbf{w})^2 + 4|\mathbf{y}^*\mathbf{w}|^2})$. Poiché $\mathbf{y}$ e $\mathbf{w}$ sono indipendenti, si ha $|\mathbf{y}^*\mathbf{w}| < \|\mathbf{y}\|\|\mathbf{w}\|$ (Nota: $(\mathbf{y}^*\mathbf{y})(\mathbf{w}^*\mathbf{w}) = |\mathbf{y}^*\mathbf{w}|^2$ iff $\mathbf{w} = a\mathbf{y}$,infatti $\mathbf{y}^*[\mathbf{w} - (\mathbf{y}^*\mathbf{w}/\mathbf{y}^*\mathbf{y})\mathbf{y}] = 0$, $\mathbf{w}^*[\dots] = 0 \Rightarrow [\dots] = \mathbf{0}$). Ne segue che $0 < c_- \le c_+$.

Agli autovalori c_+ e c_- corrispondono, rispettivamente, gli autovettori $\mathbf{v}_+ = a_+\mathbf{y} + b_+\mathbf{w}$ e $\mathbf{v}_- = a_-\mathbf{y} + b_-\mathbf{w}$ con $a_\pm, b_\pm$ soluzioni di $eq(c_\pm)$.

$(\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^*)\mathbf{z} = 0\mathbf{y} + 0\mathbf{w} = 0$ se $\mathbf{y}^*\mathbf{z} = \mathbf{w}^*\mathbf{z} = 0$. Quindi 0 è autovalore di $\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^*$ e i corrispondenti autovettori sono i $\mathbf{z} \neq \mathbf{0}$ per cui $\mathbf{y}^*\mathbf{z} = \mathbf{w}^*\mathbf{z} = 0$. Calcoliamoli. Poiché $\mathbf{y}$ e $\mathbf{w}$ sono indipendenti, esiste $i < j$ tale che $y_i w_j - y_j w_i \neq 0$. Quindi gli autovettori corrispondenti all'autovalore 0 sono i $\mathbf{z} \neq \mathbf{0}$ tali che

$$\begin{bmatrix} z_i \\ z_j \end{bmatrix} = \begin{bmatrix} \overline{y_i} & \overline{y_j} \\ \overline{w_i} & \overline{w_j} \end{bmatrix}^{-1} \begin{bmatrix} -\sum_{k \neq i,j} \overline{y_k} z_k \\ -\sum_{k \neq i,j} \overline{w_k} z_k \end{bmatrix}.$$

Tra di essi se ne possono trovare $n - 2$ linearmente indipendenti, $\mathbf{v}_1, \dots, \mathbf{v}_{n-2}$. Riassumendo

$$\begin{aligned} &(\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^*)[\mathbf{v}_+|\mathbf{v}_-|\mathbf{v}_1| \ \dots \ |\mathbf{v}_{n-2}] \\ &\qquad = [\mathbf{v}_+|\mathbf{v}_-|\mathbf{v}_1| \ \dots \ |\mathbf{v}_{n-2}] \operatorname{diag}(c_+, c_-, 0, \dots, 0) \end{aligned}$$

(Nota: $\mathbf{y}^*\mathbf{w} = 0 \Rightarrow (\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^*)\mathbf{w} = (\mathbf{w}^*\mathbf{w})\mathbf{w}$ e $(\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^*)\mathbf{y} = (\mathbf{y}^*\mathbf{y})\mathbf{y}$.)

☐ (PER ME) Siano $\mathbf{y}, \mathbf{w}, \mathbf{z}$ vettori indipendenti.

$$(\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^* + \mathbf{z}\mathbf{z}^*)(a\mathbf{y} + b\mathbf{w} + c\mathbf{z}) = d(a\mathbf{y} + b\mathbf{w} + c\mathbf{z}),$$

con a, b, c non tutti nulli, se e solo se

$$\begin{array}{l} a(\mathbf{y}^*\mathbf{y} - d) + b(\mathbf{y}^*\mathbf{w}) + c(\mathbf{y}^*\mathbf{z}) = 0 \\ a(\mathbf{w}^*\mathbf{y}) + b(\mathbf{w}^*\mathbf{w} - d) + c(\mathbf{w}^*\mathbf{z}) = 0 \\ a(\mathbf{z}^*\mathbf{y}) + b(\mathbf{z}^*\mathbf{w}) + c(\mathbf{z}^*\mathbf{z} - d) = 0, \end{array}$$

con a, b, c non tutti nulli, se e solo se il determinante della matrice 3×3

$$\begin{bmatrix} \mathbf{y}^*\mathbf{y} - d & \mathbf{y}^*\mathbf{w} & \mathbf{y}^*\mathbf{z} \\ \mathbf{w}^*\mathbf{y} & \mathbf{w}^*\mathbf{w} - d & \mathbf{w}^*\mathbf{z} \\ \mathbf{z}^*\mathbf{y} & \mathbf{z}^*\mathbf{w} & \mathbf{z}^*\mathbf{z} - d \end{bmatrix}$$

è zero. In altre parole, i tre autovalori significativi di $\mathbf{y}\mathbf{y}^* + \mathbf{w}\mathbf{w}^* + \mathbf{z}\mathbf{z}^*$ (cioè quelli non nulli) sono i tre autovalori della matrice

$$\begin{bmatrix} \mathbf{y}^*\mathbf{y} & \mathbf{y}^*\mathbf{w} & \mathbf{y}^*\mathbf{z} \\ \mathbf{w}^*\mathbf{y} & \mathbf{w}^*\mathbf{w} & \mathbf{w}^*\mathbf{z} \\ \mathbf{z}^*\mathbf{y} & \mathbf{z}^*\mathbf{w} & \mathbf{z}^*\mathbf{z} \end{bmatrix}$$

(Nota: quest'ultima matrice ha 0 come autovalore se e solo se ... $\mathbf{y}, \mathbf{w}, \mathbf{z}$ linearmente dipendenti ?)

☐ (PER ME) Siano $\mathbf{y}, \mathbf{w}$ vettori indipendenti e $\mathbf{u}, \mathbf{v}$ vettori indipendenti.
$(\mathbf{y}\mathbf{u}^* + \mathbf{w}\mathbf{v}^*)(a\mathbf{y} + b\mathbf{w}) = c(a\mathbf{y} + b\mathbf{w})$, con a, b non entrambi nulli, se e solo se

$$\begin{array}{l} a(\mathbf{u}^*\mathbf{y} - c) + b(\mathbf{u}^*\mathbf{w}) = 0, \\ a(\mathbf{v}^*\mathbf{y}) + b(\mathbf{v}^*\mathbf{w} - c) = 0, \end{array}$$

con a, b non entrambi nulli, se e solo se il determinante della matrice 2×2

$$\begin{bmatrix} \mathbf{u}^*\mathbf{y} - c & \mathbf{u}^*\mathbf{w} \\ \mathbf{v}^*\mathbf{y} & \mathbf{v}^*\mathbf{w} - c \end{bmatrix}$$

è zero. In altre parole, i due autovalori significativi di $\mathbf{y}\mathbf{u}^* + \mathbf{w}\mathbf{v}^*$ (cioè quelli non nulli) sono i due autovalori della matrice

$$\begin{bmatrix} \mathbf{u}^*\mathbf{y} & \mathbf{u}^*\mathbf{w} \\ \mathbf{v}^*\mathbf{y} & \mathbf{v}^*\mathbf{w} \end{bmatrix}$$

(Nota: quest'ultima matrice ha 0 come autovalore se e solo se $(\mathbf{u}^*\mathbf{y})(\mathbf{v}^*\mathbf{w}) = (\mathbf{u}^*\mathbf{w})(\mathbf{v}^*\mathbf{y}) \Rightarrow \mathbf{v}^*(\mathbf{w} - \frac{\mathbf{u}^*\mathbf{w}}{\mathbf{u}^*\mathbf{y}}\mathbf{y}) = \mathbf{0}$, $\mathbf{u}^*(\mathbf{w} - \frac{\mathbf{u}^*\mathbf{w}}{\mathbf{u}^*\mathbf{y}}\mathbf{y}) = \mathbf{0} \Rightarrow$ piano $\mathbf{u}, \mathbf{v}$ ortogonale a piano $\mathbf{y}, \mathbf{w}$ ... 0 autovalore $\Rightarrow \mathbf{y}\mathbf{u}^* + \mathbf{w}\mathbf{v}^*$ ha rango 1 ? Nota sul viceversa: $\mathbf{y}$ e $\mathbf{w}$ ($\mathbf{u}$ e $\mathbf{v}$) dipendenti $\Rightarrow$ 0 è autovalore.)

☐ $\mathbf{y}\mathbf{u}^* + \mathbf{w}\mathbf{v}^*$ è uguale a $\mathbf{x}\mathbf{z}^*$ se e solo se $\mathbf{w}$ e $\mathbf{y}$ (o $\mathbf{v}$ e $\mathbf{u}$) sono linearmente dipendenti. E' vera tale affermazione ?

15 Marzo 2009

☐ A $n \times n$ di dice "definita positiva" se soddisfa le due proprietà:

(1) $A = A^*$ (A è hermitiana)

(2) $\mathbf{z}^* A\mathbf{z} > 0$, per ogni vettore $n \times 1$ $\mathbf{z}$ non nullo

(Nota: (1) $\Rightarrow \overline{\mathbf{z}^* A\mathbf{z}} = \mathbf{z}^* A\mathbf{z}$, cioè $\mathbf{z}^* A\mathbf{z}$ reale $\Rightarrow$ ha senso richiedere una condizione sul segno di $\mathbf{z}^* A\mathbf{z}$)

□ Trovare una condizione su B $n \times n$ per cui $A = B^*B$ (oppure $A = BB^*$) sia definita positiva

□ Se A soddisfa (1) allora valgono le seguenti due implicazioni:
(2) $\Rightarrow$ autovalori di A positivi (Dim: $\lambda = \mathbf{v}^* A\mathbf{v}/\mathbf{v}^*\mathbf{v}$, con $\mathbf{v}$ autovettore relativo all'autovalore λ di A),
autovalori di A positivi $\Rightarrow$ (2) (Dim: omessa).
In altre parole, se A è hermitiana, per verificare se A soddisfa la condizione (2) è necessario e sufficiente verificare che il segno degli autovalori di A sia positivo

□ Dimostrare che $A = I + \mathbf{y}\mathbf{y}^*$ è definita positiva ed osservare che $A^{-1} = I - (1/(\mathbf{y}^*\mathbf{y}+1))\mathbf{y}\mathbf{y}^*$. Sotto quali condizioni su b reale la matrice $A = I + b\mathbf{y}\mathbf{y}^*$ è definita positiva? Scrivere l'inversa della matrice $3I + \mathbf{y}\mathbf{y}^*$. Risolvere il sistema

$$\begin{bmatrix} 4 & -1 & -\mathbf{i} & 2 \\ -1 & 4 & \mathbf{i} & -2 \\ \mathbf{i} & -\mathbf{i} & 4 & 2\mathbf{i} \\ 2 & -2 & -2\mathbf{i} & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

□ Dimostrare che $A = I + \mathbf{y}\mathbf{w}^*$ è non singolare se $\mathbf{w}^*\mathbf{y} \neq -1$ ed osservare che $A^{-1} = I - (1/(\mathbf{w}^*\mathbf{y}+1))\mathbf{y}\mathbf{w}^*$. Sotto quali condizioni su b reale la matrice $A = I + b\mathbf{y}\mathbf{w}^*$ è definita positiva? Scrivere, quando esiste, l'inversa della matrice $3I + \mathbf{y}\mathbf{w}^*$. Risolvere il sistema

$$\begin{bmatrix} 4 & -2 & 1 & -1 \\ -1 & 5 & -1 & 1 \\ \mathbf{i} & -2\mathbf{i} & \mathbf{i}+3 & -ui \\ 2 & -4 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

□ Siano M $n \times n$ hermitiana e b reale. Mostrare che esistono sempre valori di a reale per cui $A = aI + bM$ è definita positiva

□ Calcolare autovalori, determinante e inversa della matrice A $n \times n$, con $a_{ii} = 2$, per ogni i, e $a_{ij} = 1$ se i è diverso da j.

□ Q $n \times n$ si dice unitaria se $Q^* = Q^{-1}$ (ovvero, se per calcolarne l'inversa basta farne la trasposta coniugata) o, equivalentemente, se $Q^*Q = I$

□ Se Q è unitaria allora gli autovalori di Q hanno modulo 1

□ Gli autovalori della "matrice di Givens"

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \ a, b \in \mathbb{R}, \ a^2 + b^2 = 1$$

sono $a - \mathbf{i}b$ e $a + \mathbf{i}b$. Le matrici di Givens hanno determinante 1

☐ La matrice $A = I + b\mathbf{y}\mathbf{y}^*$, b reale, è unitaria se $b = -2/\mathbf{y}^*\mathbf{y}$. In tal caso i suoi autovalori sono 1 ($n-1$ volte) e -1. Le "matrici di Householder" $A = I - (2/\mathbf{y}^*\mathbf{y})\mathbf{y}\mathbf{y}^*$ hanno determinante -1.

☐ Q $n \times n$ si dice "matrice di permutazione" se $Q = [\mathbf{e}_{i_1}\, \mathbf{e}_{i_2}\, \dots\, \mathbf{e}_{i_n}]$ dove $\{i_1, i_2, \dots, i_n\}$ è una permutazione di $\{1, 2, \dots, n\}$ e $\mathbf{e}_j$ è il vettore con componenti tutte nulle eccetto la j-esima che è uguale a 1. Osservare che le matrici di permutazione sono matrici unitarie (cioè $Q^*Q = I$) e che il loro determinante deve essere uguale a 1 oppure a -1

Esercizi di base

- Se A $m \times n$, B $n \times k$ allora $(AB)^* = B^*A^*$
- Se A, B sono matrici $n \times n$ allora $(AB)^{-1} = B^{-1}A^{-1}$
- a numero complesso, A matrice $n \times n \Rightarrow (aA)^{-1} = a^{-1}A^{-1}$
- se λ è autovalore di M $n \times n$ allora $a_0 + a_1\lambda + \ldots + a_k\lambda^k$ è autovalore di $A = a_0 I + a_1 M + \ldots + a_k M^k$ (a_j numeri complessi)
- $\det(A) = \det(A^T)$, $\det(\overline{A}) = \overline{\det(A)}$, $\det(A^*) = \overline{\det(A)}$
- se λ è autovalore di A allora λ è autovalore di A^T
- se λ è autovalore di A allora $\overline{\lambda}$ è autovalore di $\overline{A}$ e, quindi, di A^*

7 Marzo 2009

☐ Dato $\mathbf{y}$, vettore colonna $n \times 1$, trasposto di $\mathbf{y}$: $\mathbf{y}^T = [y_1\ y_2\ \dots\ y_n]$, trasposto coniugato di $\mathbf{y}$: $\mathbf{y}^* = [\overline{y_1}\ \overline{y_2}\ \dots\ \overline{y_n}]$
($\overline{y_1} = \overline{\Re(y_1) + i\Im(y_1)} = \Re(y_1) - i\Im(y_1), \ldots$)

☐ Sia $\rho(*)$ il raggio spettrale di $*$. Siano A, B matrici $n \times n$, a scalare. Quali delle seguenti proprietà sono vere?

(1) $\rho(A) \geq 0$

(2) $\rho(A) = 0$ iff $A = 0$

(3) $\rho(aA) = |a|\rho(A)$

(4) $\rho(A+B) \leq \rho(A) + \rho(B)$

(5) $\rho(AB) \leq \rho(A)\rho(B)$

Risoluzione
Laproprietà (4) è falsa, infatti

$$A = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 0 \\ -b & a \end{bmatrix},\ a, b \in \mathbb{R},\ a^2+b^2 = 1, \quad \rho(A+B) = 1,\ \rho(A)+\rho(B) = 2|a|$$

($A + B$ is real unitary). Thus, if $|a| < 1/2$ then $\rho(A + B) > \rho(A) + \rho(B)$.
In the complex field:

$$A = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 0 \\ -\overline{b} & \overline{a} \end{bmatrix},\ a, b \in \mathbb{C},\ |a|^2+|b|^2 = 1, \quad \rho(A+B) = 1,\ \rho(A)+\rho(B) = 2|a|$$

($A+B$ is unitary). Thus, if $|a|<1/2$ then $\rho(A+B)>\rho(A)+\rho(B)$.

La proprietà (5) è falsa, infatti

$$A=\begin{bmatrix}1&0\\0&0\end{bmatrix},\ B=\begin{bmatrix}0&0\\0&1\end{bmatrix},\quad \rho(AB)=0<\rho(A)\rho(B)=1\cdot 1=1,$$

$$A=\begin{bmatrix}1&1\\1&1\end{bmatrix},\ B=\begin{bmatrix}1&1\\1&1\end{bmatrix},\quad \rho(AB)=\rho(2A)=2\rho(A)=2\cdot 2=4=\rho(A)\rho(B),$$

$$A=\begin{bmatrix}1&1\\0&0\end{bmatrix},\ B=\begin{bmatrix}1&0\\1&0\end{bmatrix},\quad \rho(AB)=2>\rho(A)\rho(B)=1\cdot 1=1.$$

□ Discutere se, per ogni A, B $n\times n$, (1) $AB=BA$ (2) autovalori di AB coincidono con quelli di BA, quindi $\rho(AB)=\rho(BA)$

□ Supponendo $\mathbf{y}$, $\mathbf{w}$ indipendenti, scrivere l'insieme dei $\mathbf{z}$ per cui $\mathbf{y}^*\mathbf{z}=0$ & $\mathbf{w}^*\mathbf{z}=0$

□ Discutere l'affermazione: $|\mathbf{y}^*\mathbf{w}|=\|\mathbf{y}\|\|\mathbf{w}\|$ iff $\mathbf{w}=a\mathbf{y}$

□ Siano $\mathbf{y}$, $\mathbf{w}$ vettori $n\times 1$ indipendenti. Calcolare esplicitamente

(1) determinante, traccia, autovalori e autovettori di $\mathbf{y}\mathbf{y}^*$

(2) *idem* di $\mathbf{y}\mathbf{w}^*$

(3) *idem* di $\mathbf{y}\mathbf{y}^*+\mathbf{w}\mathbf{w}^*$

(4) *idem* di $I+\mathbf{y}\mathbf{w}^*$

(5) inversa di $A=I+\mathbf{y}\mathbf{w}^*$ (cercarla del tipo $I+a\mathbf{y}\mathbf{w}^*$)

Risoluzione

(1): $(\mathbf{y}\mathbf{y}^*)\mathbf{y}=\mathbf{y}(\mathbf{y}^*\mathbf{y})=(\mathbf{y}^*\mathbf{y})\mathbf{y}\Rightarrow \mathbf{y}^*\mathbf{y}$ è autovalore di $\mathbf{y}\mathbf{y}^*$ con corrispondente autovettore $\mathbf{y}$.

$(\mathbf{y}\mathbf{y}^*)\mathbf{z}=0\mathbf{y}=\mathbf{0}$ se $\mathbf{y}^*\mathbf{z}=0$. Quindi 0 è autovalore di $\mathbf{y}\mathbf{y}^*$ e i corrispondenti autovettori sono i $\mathbf{z}\neq\mathbf{0}$ per cui $\mathbf{y}^*\mathbf{z}=0$. Calcoliamoli. Poiché $\mathbf{y}\neq\mathbf{0}$, esiste i t.c. $y_i\neq 0$. Quindi gli autovettori corrispondenti all'autovalore 0 sono i $\mathbf{z}\neq\mathbf{0}$ tali che $z_i=(-\sum_{k\neq i}\overline{y_k}z_k)/\overline{y_i}$. Tra di essi se ne possono trovare $n-1$ linearmente indipendenti, $\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}$. Riassumendo

$$\mathbf{y}\mathbf{y}^*[\mathbf{y}|\mathbf{v}_1|\ldots|\mathbf{v}_{n-1}]=[\mathbf{y}|\mathbf{v}_1|\ldots|\mathbf{v}_{n-1}]\,\mathrm{diag}\,(\mathbf{y}^*\mathbf{y},0,\ldots,0).$$

In particolare si ha: $\rho(\mathbf{y}\mathbf{y}^*)=\mathbf{y}^*\mathbf{y}$

(2): $(\mathbf{y}\mathbf{w}^*)\mathbf{y}=\mathbf{y}(\mathbf{w}^*\mathbf{y})=(\mathbf{w}^*\mathbf{y})\mathbf{y}\Rightarrow \mathbf{w}^*\mathbf{y}$ è autovalore di $\mathbf{y}\mathbf{w}^*$ con corrispondente autovettore $\mathbf{y}$. $\mathbf{w}^*\mathbf{y}$ in generale non è reale.

$(\mathbf{y}\mathbf{w}^*)\mathbf{z}=0\mathbf{y}=0$ se $\mathbf{w}^*\mathbf{z}=0$. Quindi 0 è autovalore di $\mathbf{y}\mathbf{w}^*$ e i corrispondenti autovettori sono i $\mathbf{z}\neq\mathbf{0}$ per cui $\mathbf{w}^*\mathbf{z}=0$. Calcoliamoli. Poiché $\mathbf{w}\neq\mathbf{0}$, esiste i t.c. $w_i\neq 0$. Quindi gli autovettori corrispondenti all'autovalore 0 sono i $\mathbf{z}\neq\mathbf{0}$ tali che $z_i=(-\sum_{k\neq i}\overline{w_k}z_k)/\overline{w_i}$. Tra di essi se ne possono trovare $n-1$ linearmente indipendenti, $\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}$. Riassumendo

$$\mathbf{y}\mathbf{w}^*[\mathbf{y}|\mathbf{v}_1|\ldots|\mathbf{v}_{n-1}]=[\mathbf{y}|\mathbf{v}_1|\ldots|\mathbf{v}_{n-1}]\,\mathrm{diag}\,(\mathbf{w}^*\mathbf{y},0,\ldots,0).$$

In particolare si ha: $\rho(\mathbf{y}\mathbf{w}^*)=|\mathbf{w}^*\mathbf{y}|$

Nota: $\rho(\mathbf{y}\mathbf{w}^*) = |\mathbf{w}^*\mathbf{y}| \leq \|\mathbf{y}\mathbf{w}^*\|_2 = \|\mathbf{y}\|_2\|\mathbf{w}\|_2$, mentre $\rho(\mathbf{y}\mathbf{y}^*) = \mathbf{y}^*\mathbf{y} = \|\mathbf{y}\mathbf{y}^*\|_2 = \|\mathbf{y}\|_2^2$.

(3) (vedi più avanti).

☐ Studiare l'andamento del numero $\rho(A)\rho(A^{-1})$, $A = I + \mathbf{y}\mathbf{w}^*$

☐ Scrivere le matrici

$$\begin{bmatrix} 2 & 1 \\ 1 & 1/2 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 2\mathbf{i} & -3 \\ 2 & -2 & 4\mathbf{i} & -6 \\ i & -i & -2 & -3\mathbf{i} \\ 4 & -4 & 8\mathbf{i} & -12 \end{bmatrix}$$

nella forma $\mathbf{y}\mathbf{y}^*$ o $\mathbf{y}\mathbf{w}^*$, e calcolare il loro raggio spettrale

☐ Calcolare $\rho(A)$ dove A è la matrice 4×4

$$\begin{bmatrix} 1 & -2 & 5 & -1 \\ -2 & 4 & 10 & 2 \\ 5 & 10 & 25 & -5 \\ -1 & 2 & -5 & 1 \end{bmatrix}$$

☐ Calcolare gli autovalori λ della matrice 2×2

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix},$$

con a, b reali. Determinare una CNES affinché $|\lambda| = 1$

Risoluzione

Gli autovalori sono $a - \mathbf{i}|b|$ e $a + \mathbf{i}|b|$, e il loro modulo è $\sqrt{a^2 + b^2}$. Quindi, $|\lambda| = 1$ se e solo se $a^2 + b^2 = 1$.

Gli autovalori di

$$A = \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}, \; a, b \in \mathbb{C},$$

sono $\Re(a) \pm \mathbf{i}\sqrt{\Im(a)^2 + |b|^2}$, e il loro modulo vale $\sqrt{|a|^2 + |b|^2}$. Quindi, $|\lambda| = 1$ se e solo se $|a|^2 + |b|^2 = 1$.

☐ Per A $n \times n$ sia A^* la matrice $[A^*]_{ij} = \overline{[A]_{ji}}$. A $n \times n$ si dice hermitiana, se $a_{ij} = \overline{a_{ji}}$ $\forall\, i, j$ (ovvero se $A = A^*$).

☐ $\mathbf{y}\mathbf{w}^*$ è hermitiana se e solo se $\mathbf{w} = a\mathbf{y}$, a reale

☐ A $n \times n$ con elementi non negativi tale che $\sum_j a_{ij} = 1$ per ogni i si dice stocastica per righe. Se A è stocastica per righe allora

(1) 1 è autovalore di A con autovettore $[1\ 1\ \ldots\ 1]^T$

(2) $\rho(A) = 1$

(3) A^k è stocastica per righe (per induzione su k)

(4) $\sum_{k=0}^{r} y_k A^k$ è stocastica per righe, se $y_k \geq 0$ per ogni k e $\sum_k y_k = 1$

Esercizi di base

- Ad autovalori distinti corrispondono autovettori indipendenti
- Gli autovalori di aA si ottengono moltiplicando gli autovalori di A per a
- Gli autovalori di A^k si ottengono facendo le potenze k degli autovalori di A
- Gli autovalori di A^{-1} si ottengono invertendo quelli di A
- Se $B = S^{-1}AS$ allora A e B hanno gli stessi autovalori
- Se T è triangolare gli autovalori di T sono i suoi elementi diagonali
- Se $S^{-1}AS = T$ con T triangolare allora gli autovalori di A sono gli elementi diagonali di T
- Se $S^{-1}AS = D$ con D diagonale allora gli autovalori di A sono gli elementi diagonali di D e i corrispondenti autovettori sono le colonne di S (e A si dice diagonalizzabile)
- Esiste sempre S, $S^* = S^{-1}$, per cui $S^{-1}AS = T$ con T triangolare. Esiste S, $S^* = S^{-1}$, per cui $S^{-1}AS = D$ se e solo se $AA^* = A^*A$ (ovvero A è normale)
- $A = A^* \Rightarrow$ autovalori di A reali
- Autovalori di A reali & A normale $\Rightarrow A = A^*$

Inizio:

$\mathbf{y}$ vettore colonna $n \times 1$,

$\mathbf{y}^T = [y_1 \ \dots \ y_n]$ trasposto di $\mathbf{y}$,

$\mathbf{y}^* = [\overline{y_1} \ \dots \ \overline{y_n}]$ trasposto coniugato di $\mathbf{y}$

Esempio

$$\mathbf{y} = \begin{bmatrix} 1 \\ 1 - \mathbf{i} \end{bmatrix}, \quad \mathbf{y}^T = [1 \ \ 1 - \mathbf{i}], \quad \mathbf{y}^H = [1 \ \ 1 + i]$$

$\mathbf{y}^*\mathbf{y} = \sum_i |y_i|^2$ è un numero reale non negativo (che si indica anche $\|y\|_2^2$)

$\mathbf{y}\mathbf{y}^*$ è una matrice $n \times n$, ha determinante nullo ($\Rightarrow$ ha almeno un autovalore uguale a zero), ha traccia uguale a $\|y\|_2^2$, è un esempio di matrice hermitiana (def di traccia e di matrice hermitiana)

matrici hermitiane hanno gli elementi diagonali reali

def di autovalori, autovettori, polinomio caratteristico, A $n \times n$ ha n autovalori, gli autovalori di una matrice triangolare sono i suoi elementi diagonali, autovettori di autovalori distinti sono indipendenti, se A è simile a T triangolare allora gli autovalori di A sono gli elementi diagonali di T

def di raggio spettrale di A $n \times n$: $\rho(A) = \max_i |\lambda_i|$ dove λ_i, $i = 1, \dots, n$ sono gli autovalori di A.

es. $\mathbf{y}$ vettore $n \times 1 \Rightarrow \rho(\mathbf{y}\mathbf{y}^*) = \mathbf{y}^*\mathbf{y}$

es. trovare due matrici 2×2 A, B per cui $\rho(AB) > \rho(A)\rho(B)$, $<$, $=$