

Analisi Numerica

(A.A. 2014-2015)

Appunti delle lezioni: Sistemi lineari

Docente Vittoria Bruni

Email: vittoria.bruni@sbai.uniroma1.it

Ufficio: Via A. Scarpa,

Pal. B, I piano, Stanza n. 16

Tel. 06 49766648

Ricevimento: Martedì 9.30-10.30

Testi consigliati:

Calcolo Numerico, L. Gori, Ed. Kappa, 2006

Esercizi di Calcolo Numerico, L. Gori-M.L. Lo Cascio, F. Pitolli, Ed. Kappa, 2007

Il materiale didattico è disponibile sul sito

<http://www.sbai.uniroma1.it/users/bruni-vittoria>

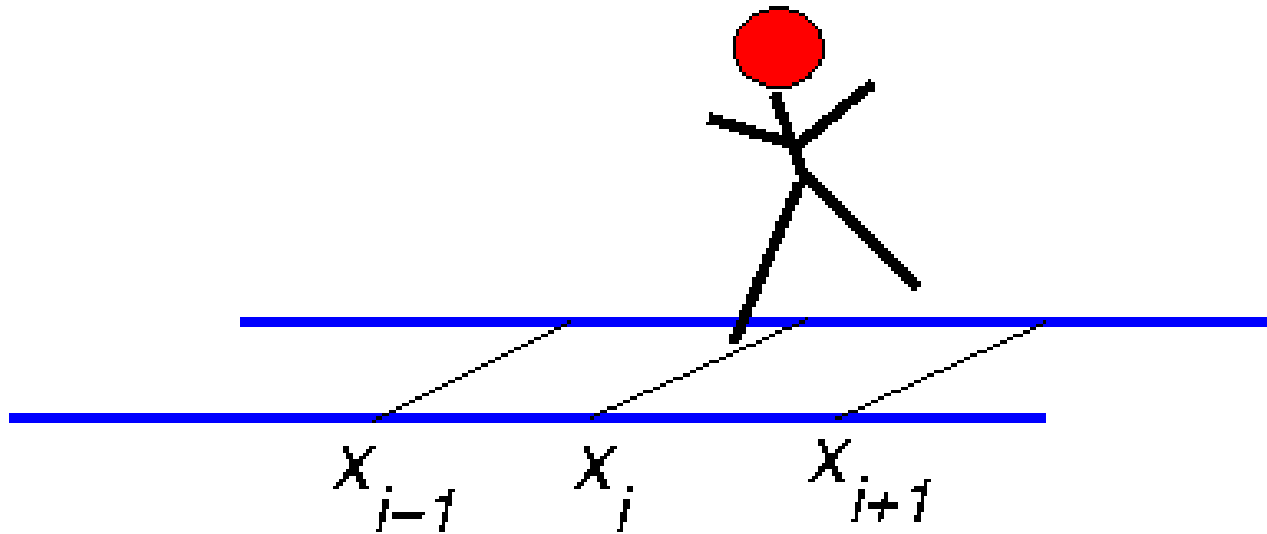
nella pagina dedicata al corso [Analisi Numerica](#)

Sistemi di equazioni lineari

Esempio 1

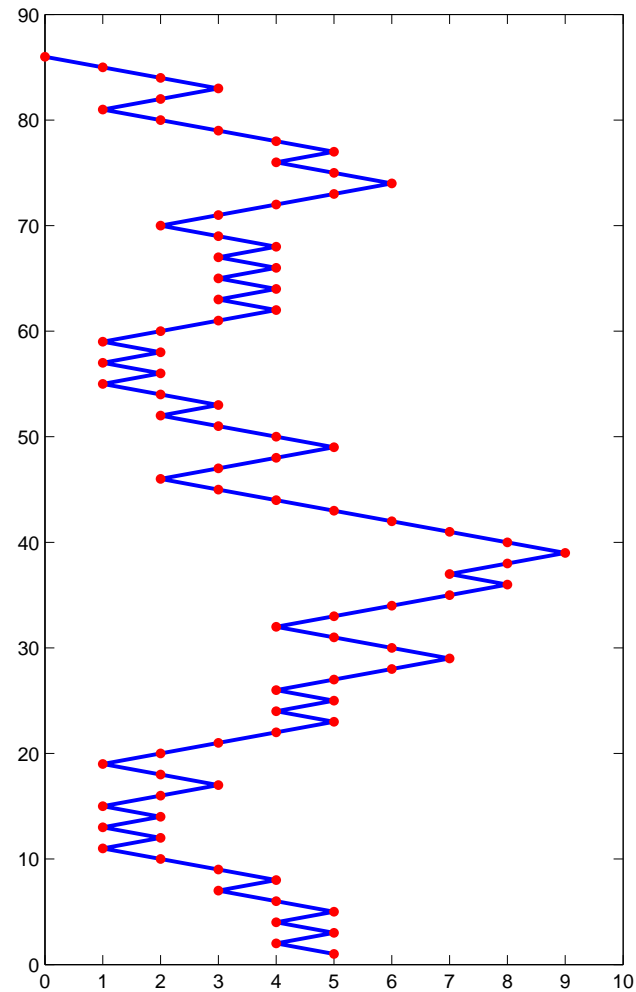
Un ubriaco compie una **passeggiata casuale**, facendo un passo a sinistra o a destra a caso lungo una strada rettilinea. Quando raggiunge una estremità della strada, si ferma.

Calcolare la **probabilità** che l'ubriaco raggiunga l'estremità sinistra della strada partendo dalla posizione i .



Esempio di passeggiata per $N = 10$

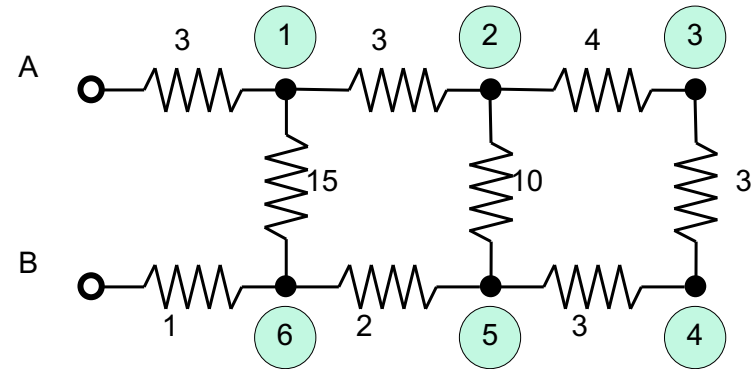
Si può simulare una **passeggiata casuale** tirando una moneta.



Punto di partenza: x_5

Esempio 2

Determinare i potenziali nei nodi 1 – 6 del circuito sapendo che tra A e B è applicata una differenza di potenziale pari a 100V (le resistenze sono misurate in *Ohm*)



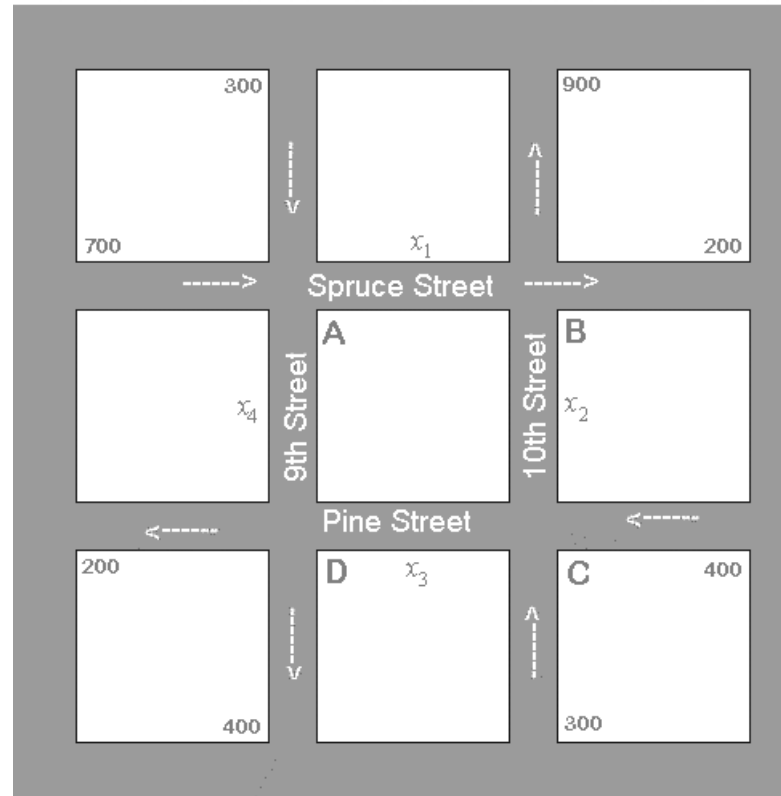
Soluzione

Applicando la **legge di Ohm** $\Delta V = RI$ e la **legge di Kirchhoff** $\sum_i I_i = 0$ in ogni nodo si ottiene il sistema lineare

$$\left\{ \begin{array}{l} 11v_1 - 5v_2 - v_6 = 500 \\ -20v_1 + 41v_2 - 15v_3 - 6v_5 = 0 \\ -3v_2 + 7v_3 - 4v_4 = 0 \\ -v_3 + 2v_4 - v_5 = 0 \\ -3v_2 - 10v_4 + 28v_5 - 15v_6 = 0 \\ -2v_1 - 15v_5 + 47v_6 = 0 \end{array} \right.$$

Esempio 3

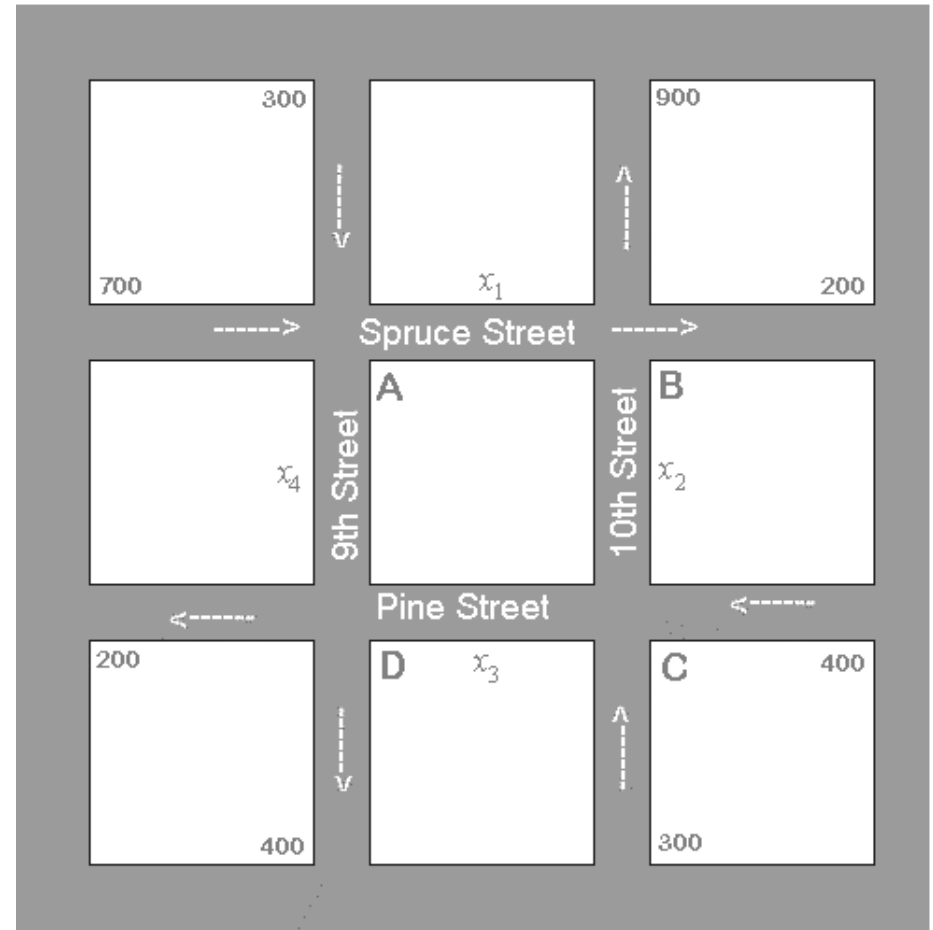
Nelle ore di punta il **traffico è congestionato** in corrispondenza degli incroci rappresentati nella figura di seguito



Tutte le strade sono a senso unico e la direzione di circolazione è indicata dalle frecce.

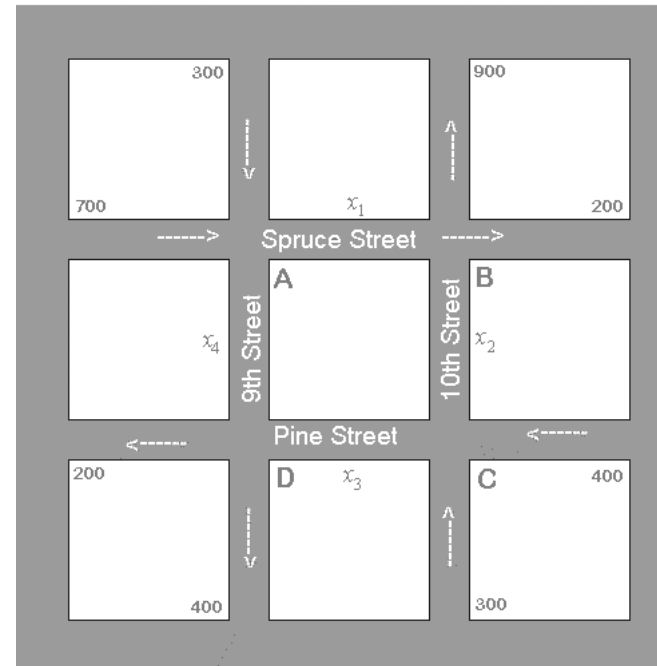
Inoltre:

1. **Incrocio A**: 700 macchine ogni ora provengono da Spruce Street mentre 300 provengono da 9th Street
2. **Incrocio B**: 200 macchine ogni ora attraversano l'incrocio B provengono da Spruce Street mentre 900 provengono da 10th Street.
3. **Incrocio C**: 400 macchine ogni ora entrano in Pine Street attraverso l'incrocio C mentre 300 provengono da 10th Street.
4. **Incrocio D**: 200 macchine ogni ora lasciano l'incrocio D in direzione Pine Street mentre 400 in direzione 9th Street fino all'incrocio A.



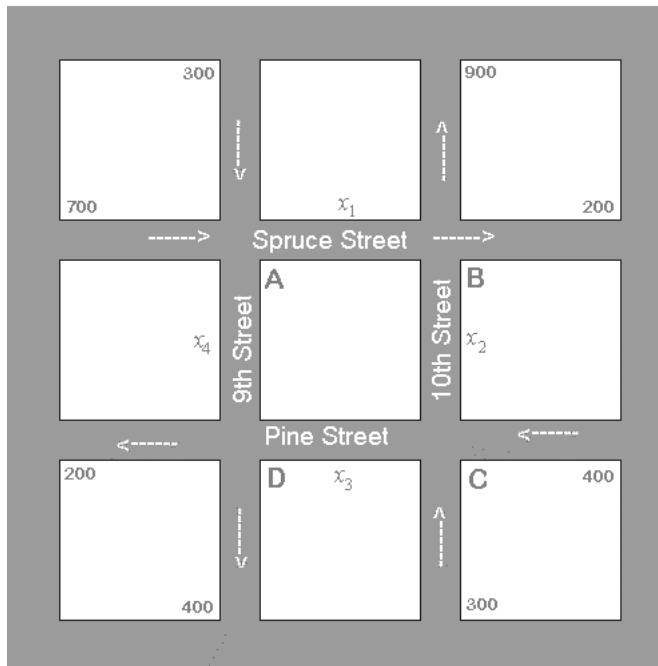
Indicando con

- x_1 il numero di macchine che lascia l'incrocio A percorrendo Spruce Street verso l'incrocio B ;
- x_2 il numero di macchine che arriva all'incrocio B percorrendo 10th Street dall'incrocio C;
- x_3 il numero di macchine che lascia l'incrocio C percorrendo Pine Street verso l'incrocio D;
- x_4 il numero di macchine che arriva all'incrocio D percorrendo 9th Street dall'incrocio A.



Assumendo che

1. per velocizzare il flusso del traffico ogni macchina che arriva ad un certo incrocio deve anche lasciarlo;
2. tutte le strade sono a senso unico
3. x_1, x_2, x_3 e x_4 sono numeri positivi



Si ha

Incrocio A

$$x_1 + x_4 = 700 + 300$$

Incrocio B

$$x_1 + x_2 = 900 + 200$$

Incrocio C

$$x_2 + x_3 = 400 + 300$$

Incrocio D

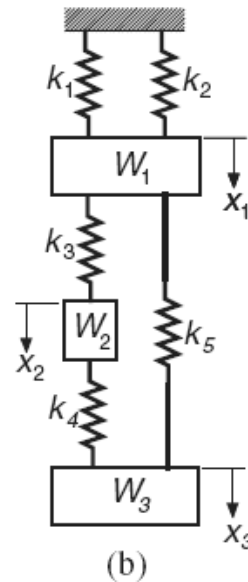
$$x_3 + x_4 = 400 + 200$$

Cioè è necessario risolvere il seguente sistema lineare

$$\left\{ \begin{array}{rclcl} x_1 & + & & & x_4 & = & 1000 \\ x_1 & + & x_2 & & & = & 1100 \\ & & x_2 & + & x_3 & = & 700 \\ & & & & x_3 & + & x_4 & = & 600 \end{array} \right.$$

Esempio 4

Il sistema in figura



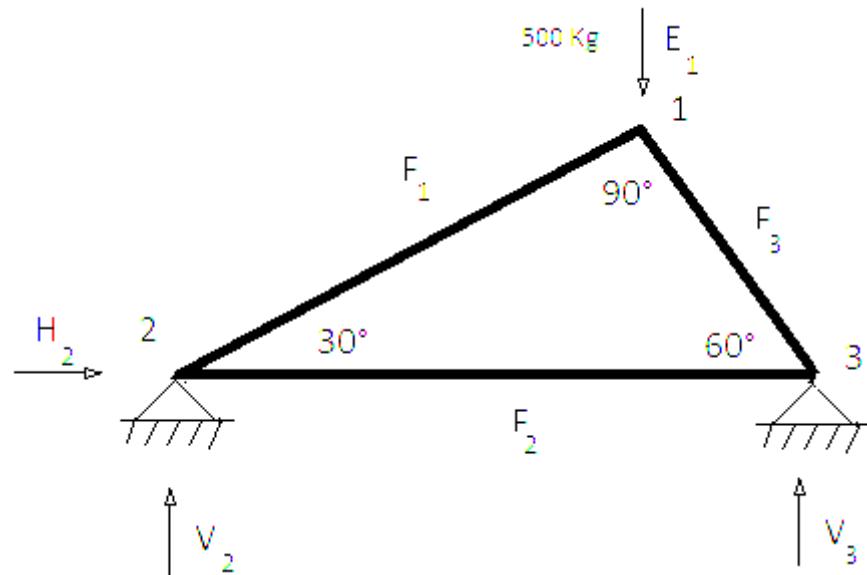
è costituito da 4 molle che sostengono 3 pesi $W_i, i = 1, \dots, 3$ e all'equilibrio soddisfa

$$\begin{cases} (k_1 + k_2 + k_3 + k_5)x_1 & -k_3 x_2 & -k_5 x_3 & = W_1 \\ -k_3 x_1 & +(k_3 + k_4) x_2 & -k_4 x_3 & = W_2 \\ -k_5 x_1 & -k_4 x_2 + & (k_4 + k_5) x_3 & = W_3 \end{cases}$$

con k_i costante elastica della molla i -esima e x_j spostamento della massa j -esima rispetto alla posizione nel sistema non deformato.

Esempio 5

Analisi di strutture



Nel telaio (*struttura statica bidimensionale*) rappresentato in figura, le forze F_i , $i = 1, 2, 3$ rappresentano le tensioni e compressioni che agiscono sulle aste della struttura, H_2 , V_2 , V_3 sono le forze relative all'interazione con il supporto; le forze E_i , $i = 1, 2, 3$ sono le forze esterne applicate ai nodi della struttura. Nell'esempio considerato, E_2, E_3 sono nulle, il nodo 2 è vincolato mentre il nodo 3 può scivolare.

In condizioni statiche, all' equilibrio si ha

$$\begin{cases} -F_1 \cos(30) + F_3 \cos(60) + E_{1,h} = 0 \\ -F_1 \sin(30) - F_3 \sin(60) + E_{1,v} = 0 \\ -F_1 \cos(30) + F_2 + H_2 + E_{2,h} = 0 \\ -F_1 \sin(30) + V_2 + E_{2,v} = 0 \\ -F_2 - F_3 \cos(60) + E_{3,h} = 0 \\ F_3 \sin(60) + V_3 + E_{3,v} = 0 \end{cases}$$

in cui $E_{1,h} = E_{2,h} = E_{2,v} = E_{3,h} = E_{3,v} = 0$ e $E_{1,v} = -500$.

Si tratta dunque di risolvere un sistema lineare di 6 equazioni nelle incognite $F_1, F_2, F_3, H_2, V_2, V_3$ ovvero

$$\mathbf{AX} = \mathbf{B},$$

con $\mathbf{X} = [F_1, F_2, F_3, H_2, V_2, V_3]^T$, $\mathbf{B} = [0, 500, 0, 0, 0, 0]^T$ e

$$\mathbf{A} = \begin{pmatrix} -\cos(30) & 0 & \cos(60) & 0 & 0 & 0 \\ -\sin(30) & 0 & -\sin(60) & 0 & 0 & 0 \\ -\cos(30) & 1 & 0 & 1 & 0 & 0 \\ -\sin(30) & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -\cos(60) & 0 & 0 & 0 \\ 0 & 0 & \sin(60) & 0 & 0 & 1 \end{pmatrix}$$

Esempio 6

Punto di intersezione tra due rette

$$\begin{cases} 2x + y = 3 \\ 4x + 2y = 6 \end{cases}$$

rette coincidenti

infinite soluzioni

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

$$\det(A) = 4 - 4 = 0$$

$$\begin{cases} 2x + y = 3 \\ 4x + 2y = 0 \end{cases}$$

rette parallele

non esistono soluzioni

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

$$\det(A) = 4 - 4 = 0$$

$$\begin{cases} 2x + y = 3 \\ x + 2y = 3 \end{cases}$$

rette incidenti

unica soluzione

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$\det(A) = 4 - 1 = 3 \neq 0$$

Metodi diretti per la soluzione di sistemi lineari

Sistema lineare $AX = B$

A = matrice dei coefficienti X = vettore delle incognite

B = vettore dei termini noti

- I **metodi diretti** sono basati sulla **trasformazione** del sistema di partenza in uno **equivalente** che abbia una struttura particolarmente **semplice** per cui è facile calcolarne la soluzione.
- La **soluzione numerica** viene calcolata in un **numero finito** di passi e, se non vi fossero errori di arrotondamento nei dati o durante i calcoli, la soluzione numerica sarebbe **esatta**.
- Data l'**occupazione di memoria** (RAM) richiesta nei passaggi dell'algoritmo, vengono utilizzati quando la **matrice dei coefficienti ha dimensione non "troppo" elevata**.

Costo computazionale di un algoritmo

Prima di implementare un algoritmo bisogna stimare il suo **costo computazionale**, cioè il numero di **operazioni pesanti** (moltiplicazioni o divisioni) necessarie per calcolare **numericamente** la soluzione.

Costo computazionale:

$C_c \approx$ numero di moltiplicazioni o divisioni

Richiami sul determinante di matrici

- Se A è una matrice di dimensione 1×1 con $a_{11} = a$, $\det(A) = a$

- Se A è una matrice di dimensione $n \times n$, allora

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} M_{ij}, \quad \forall i = 1, \dots, n$$

oppure $\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} M_{ij}, \quad \forall j = 1, \dots, n,$

con M_{ij} il determinante della matrice A di cui si trascurano la i -esima riga e la j -esima colonna

- Se A è una matrice quadrata di dimensione n :
 1. Se una riga o una colonna di A ha tutti gli elementi nulli, allora $\det(A) = 0$

2. Se A ha due righe o due colonne con gli stessi elementi, allora $\det(A) = 0$
3. Se \tilde{A} è ottenuta scambiando due delle righe di A allora $\det(\tilde{A}) = -\det(A)$
4. Se \tilde{A} è ottenuta moltiplicando una riga di A per lo scalare λ , allora $\det(\tilde{A}) = \lambda \det(A)$
5. Se \tilde{A} è ottenuta sommando ad una riga di A un'altra riga moltiplicata per lo scalare λ , allora $\det(\tilde{A}) = \det(A)$
6. Se B è una matrice quadrata di ordine n , allora $\det(AB) = \det(A)\det(B)$
7. $\det(A^T) = \det(A)$
8. Se esiste A^{-1} , allora $\det(A^{-1}) = 1/\det(A)$
9. Se A è una matrice triangolare superiore (o inferiore), allora $\det(A) = \prod_{i=1}^n a_{ii}$

Metodo di Cramer

Dall'**Algebra** sappiamo che la soluzione **esatta** di un sistema lineare si può ottenere con il **metodo di Cramer**.

Se A è **regolare**, allora

$$X = A^{-1}B \quad \Rightarrow \quad x_i = \frac{D_i}{\det A}, \quad i = 1, 2, \dots, n,$$

dove D_i è il determinante della matrice ottenuta da A sostituendo alla colonna i -esima il vettore B .

Esempio: Calcolare la soluzione del sistema $AX = B$, con

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} \quad \text{e} \quad b = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

$$\det(A) = 0 + 84 + 96 - 105 - 0 - 48 = 27 \neq 0$$

$$D_1 = \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 2 & 8 & 0 \end{pmatrix} = 24 - 30 - 48 = -54 \quad x_1 = \frac{D_1}{\det(A)} = \frac{-54}{27} = -2$$

$$D_2 = \det \begin{pmatrix} 1 & 1 & 3 \\ 4 & 0 & 6 \\ 7 & 2 & 0 \end{pmatrix} = 42 + 24 - 12 = 54 \quad x_2 = \frac{D_2}{\det(A)} = \frac{54}{27} = 2$$

$$D_3 = \det \begin{pmatrix} 1 & 2 & 1 \\ 4 & 5 & 0 \\ 7 & 8 & 2 \end{pmatrix} = 10 + 32 - 35 - 16 = -9 \quad x_3 = \frac{D_3}{\det(A)} = \frac{-9}{27} = -\frac{1}{3}$$

$$\text{e quindi } \mathbf{X} = \begin{pmatrix} -2 \\ 2 \\ -\frac{1}{3} \end{pmatrix}$$

Costo computazionale del metodo di Cramer

$$x_i = \frac{D_i}{\det A} \quad i = 1, 2, \dots, n$$

- $n + 1$ **determinanti** ($D_i, i = 1, \dots, n$ e $\det A$)
- $n!$ **prodotti** per ciascun determinante
- $n - 1$ **moltiplicazioni** per ciascun prodotto
- $+ n$ **divisioni** (trascurabili)

Costo computazionale: $C_c = (n + 1)n!(n - 1) + n \simeq (n + 1)n!(n - 1)$

$n = 15 \rightarrow C_c \simeq 3 \cdot 10^{14}$ moltiplicazioni $\rightarrow 3$ giorni
 $n = 20 \rightarrow C_c \simeq 3 \cdot 10^{21}$ moltiplicazioni $\rightarrow 3 \cdot 10^5$ anni } **Inutilizzabile!!**

(supponendo 10^{-9} secondi per operazione)

Calcolo dell'inversa di A

$$X = A^{-1}B$$

Si tratta di trovare la matrice M tale che

$$AM = I$$

con I la matrice identità .

Corrisponde alla soluzione di **tre sistemi lineari** ognuno dei quali ha A come matrice dei coefficienti, una colonna di I come vettore dei termini noti e una colonna di M come incognita.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Metodo di eliminazione di Gauss

Il **metodo di eliminazione di Gauss** trasforma, in $n-1$ passi, il sistema lineare

$$AX = B \quad X, B \in \mathbf{R}^n \quad A \in \mathbf{R}^{n \times n}$$

con matrice dei coefficienti A "**piena**", nel sistema **equivalente**

$$UX = \tilde{B} \quad X, \tilde{B} \in \mathbf{R}^n \quad U \in \mathbf{R}^{n \times n}$$

con matrice dei coefficienti U **triangolare superiore**.

Il metodo utilizza le seguenti **operazioni** "*lecite*":

- **scambio** di 2 equazioni
- **somma** di due equazioni di cui una è **moltiplicata** per una costante

Soluzione di sistemi triangolari

Sistemi triangolari superiori

$$\left\{ \begin{array}{l} u_{11}x_1 + u_{12}x_2 + \dots + u_{1k}x_k + \dots + u_{1n}x_n = b_1 \\ \quad u_{22}x_2 + \dots + u_{2k}x_k + \dots + u_{2n}x_n = b_2 \\ \quad \quad \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ \quad \quad \quad \quad u_{kk}x_k + \dots + u_{kn}x_n = b_k \\ \quad \quad \quad \quad \quad \quad \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad u_{nn}x_n = b_n \end{array} \right.$$



$$UX = B \rightarrow \left\{ \begin{array}{l} x_n = \frac{b_n}{u_{nn}} \\ x_k = \left(b_k - \sum_{i=k+1}^n u_{ki}x_i \right) \frac{1}{u_{kk}}, \quad k = n-1, n-2, \dots, 2, 1 \end{array} \right.$$

(**Algoritmo di sostituzione all'indietro**)

Sistemi triangolari inferiori

$$\begin{cases} l_{11}x_1 = b_1 \\ \dots \dots \dots \dots \dots \\ l_{k1}x_1 + l_{k2}x_2 + \dots + l_{kk}x_k = b_k \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ l_{n1}x_1 + l_{n2}x_2 + \dots + l_{nk}x_k + \dots + l_{nn}x_n = b_n \end{cases}$$



$$LX = B \rightarrow \begin{cases} x_1 = \frac{b_1}{l_{11}} \\ x_k = \left(b_k - \sum_{i=1}^{k-1} l_{ki}x_i \right) \frac{1}{l_{kk}} \quad k = 1, 2, \dots, n \end{cases}$$

(**Algoritmo di sostituzione in avanti**)

Nota: Se le matrici U e L sono **regolari**, sicuramente u_{kk} e l_{kk} sono $\neq 0$.

Costo computazionale

Algoritmo di sostituzione

Ad ogni passo ci sono

- $n - k$ **moltiplicazioni**
- **1** **divisione**

$$C_c = \sum_{k=1}^n (n - k + 1) \simeq \frac{n^2}{2}$$

Metodo di eliminazione di Gauss

Per risolvere il seguente sistema lineare

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ 4x_1 + 5x_2 + 6x_3 = 0 \\ 7x_1 + 8x_2 + = 2 \end{cases}$$

si considerano la matrice **A** e il termine noto **b** associati ad esso

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 0 & 2 \end{array} \right)$$

Metodo di eliminazione di Gauss

Si moltiplica la prima riga per 4 e si sottrae alla seconda, annullando così il secondo elemento della prima colonna di A .

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -4 \\ 7 & 8 & 0 & 2 \end{array} \right)$$

Si moltiplica la prima riga per 7 e si sottrae alla terza, annullando il terzo elemento della prima colonna di A .

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -4 \\ 0 & -6 & -21 & -5 \end{array} \right)$$

Metodo di eliminazione di Gauss

Si moltiplica la seconda riga per $6/3$ e si sottrae alla terza, annullando il terzo elemento della seconda colonna di A

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -4 \\ 0 & 0 & -9 & 3 \end{array} \right)$$

Si ottiene un sistema equivalente a quello dato in cui la matrice dei coefficienti è triangolare superiore

$$\left\{ \begin{array}{l} x_1 + 2x_2 + 3x_3 = 1 \\ \quad - 3x_2 - 6x_3 = -4 \\ \qquad \qquad - 9x_3 = 3 \end{array} \right.$$

Metodo di eliminazione di Gauss

Il sistema in questa forma diventa di facile soluzione.

Infatti, partendo dall'ultima equazione, si ha

$$\begin{aligned}x_3 &= -3/9 = -1/3 \\x_2 &= (-4 + 6x_3)/(-3) = 2 \\x_1 &= 1 - 3x_3 - 2x_2 = 1 + 1 - 4 = -2\end{aligned}$$

Metodo di eliminazione di Gauss

Generalizzando, se la matrice A è tale che $a_{ii} \neq 0, \quad \forall i = 1, 2, \dots, n$

I passo : annullare gli elementi della prima colonna di A al di sotto della diagonale principale

I.1 sottrarre alla seconda equazione la prima moltiplicata per $m_{21} = \frac{a_{21}}{a_{11}}$

I.2 sottrarre alla terza equazione la prima moltiplicata per $m_{31} = \frac{a_{31}}{a_{11}}$

⋮

⋮

I.n-1 sottrarre alla $n - \text{esima}$ equazione la prima moltiplicata per $m_{n1} = \frac{a_{n1}}{a_{11}}$

Al termine di questo passo si ottiene un sistema equivalente con matrice dei coefficienti

$$A^{(1)} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & \dots & a_{3n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}$$

dove

$$a_{ij}^{(1)} = a_{ij} - m_{i1}a_{1j}, \quad i = 2, \dots, n \quad j = 1, \dots, n$$

e vettore dei termini noti

$$b^{(1)} = \begin{pmatrix} b_1 \\ b_2^{(1)} \\ b_3^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix}$$

dove

$$b_i^{(1)} = b_i - m_{i1}b_1, \quad i = 2, \dots, n$$

II passo : annullare gli elementi della seconda colonna di $A^{(1)}$ al di sotto della diagonale principale

II.1 sottrarre alla terza equazione la seconda moltiplicata per $m_{32} = \frac{a_{32}^{(1)}}{a_{22}^{(1)}}$

II.2 sottrarre alla quarta equazione la seconda moltiplicata per $m_{42} = \frac{a_{42}^{(1)}}{a_{22}^{(1)}}$

⋮

⋮

II.n-2 sottrarre alla n -esima equazione la seconda moltiplicata per $m_{n2} = \frac{a_{n2}^{(1)}}{a_{22}^{(1)}}$

Al termine di questo passo si ottiene un sistema equivalente con

matrice dei coefficienti

$$A^{(2)} = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & \dots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix}$$

dove

$$a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i2}a_{2j}^{(1)}, \quad i = 3, \dots, n \quad j = 2, \dots, n$$

e vettore dei termini noti

$$b^{(2)} = \begin{pmatrix} b_1 \\ b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix}$$

dove

$$b_i^{(2)} = b_i^{(1)} - m_{i2}b_2^{(1)}, \quad i = 3, \dots, n$$

⋮

⋮

n-1 passo : annullare gli elementi della $(n - 1) -esima$ colonna di $A^{(n-2)}$ al di sotto della diagonale principale

II.1 sottrarre alla $n -esima$ equazione la $(n - 1) -esima$ moltiplicata per $m_{n\ n-1} = \frac{a_{n\ n-1}^{(n-2)}}{a_{n-1\ n-1}^{(n-2)}}$

Al termine di questo passo si ottiene un sistema equivalente con matrice dei coefficienti

$$A^{(n-1)} = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & \dots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn}^{(n-1)} \end{pmatrix}$$

dove

$$a_{ij}^{(n-1)} = a_{ij}^{(n-2)} - m_{i\ n-1} a_{n-1\ j}^{(n-2)}, \quad i = n \quad j = n - 1, \dots, n$$

e vettore dei termini noti

$$b^{(n-1)} = \begin{pmatrix} b_1 \\ b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_n^{(n-1)} \end{pmatrix}$$

dove

$$b_i^{(n-1)} = b_i^{(n-1)} - m_{i \ n-1} b_{n-1}^{(n-1)}, \quad i = n$$

Metodo di eliminazione di Gauss

Quindi, dopo $n - 1$ passi il sistema $Ax = b$ è stato trasformato nel sistema equivalente $A^{(n-1)}x = b^{(n-1)}$, con $A^{(n-1)}$ matrice triangolare superiore

al passo k si definiscono gli elementi

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - m_{ik}a_{kj}^{(k-1)}, \quad i = k + 1, \dots, n \quad j = k, k + 1, \dots, n$$

$$b_i^{(k)} = b_i^{(k-1)} - m_{ik}b_k^{(k-1)}, \quad i = k + 1, \dots, n$$

$$\text{con } m_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

Costo computazionale

Metodo di eliminazione di Gauss:

- triangolarizzazione $\simeq \frac{n^3}{3}$. Infatti:

$$\underbrace{\sum_{i=1}^{n-1}}_{\text{ognipasso}} \quad \underbrace{\sum_{l=i+1}^n}_{\text{righe}} \quad \left(\underbrace{1}_{\text{div.}} + \underbrace{n-i+2}_{\text{molt.}} \right) =$$
$$= \sum_{i=1}^{n-1} (n-i)(n+3-i) \simeq \frac{n^3}{3}$$

Sono state usate le seguenti identità :

$$\sum_{i=1}^n i = \frac{n(n+1)}{2}, \quad \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

- sostituzione all'indietro $\simeq \frac{n^2}{2}$

$$\Rightarrow C_c \simeq \frac{n^3}{3}$$

Esercizio

Risolvere il seguente sistema lineare

$$\begin{cases} 2x_1 + x_2 + 2x_3 = 10 \\ 4x_1 + x_2 + 2x_3 = 12 \\ x_1 + 2x_2 + 5x_3 = 20 \end{cases}$$

usando il metodo di eliminazione di Gauss

Fattorizzazione LU

Il **metodo di eliminazione di Gauss** può essere interpretato come la **fattorizzazione** della matrice di partenza A nel prodotto di due matrici triangolari.

Teorema. Se la matrice $A \in \mathbf{R}^{n \times n}$ ha **determinanti principali di testa** tali che

$$\det A_k \neq 0, \quad k = 1, 2, \dots, n,$$

allora

$$A = LU$$

dove $L \in \mathbf{R}^{n \times n}$ è una **matrice triangolare inferiore** con elementi diagonali pari a 1 e $U \in \mathbf{R}^{n \times n}$ è una **matrice triangolare superiore**.

Infatti, riprendendo l'esempio precedente in cui $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix}$

Moltiplicare per 4 la prima riga e sottrarla alla seconda e moltiplicare la prima riga per 7 e sottrarla alla terza equivale a moltiplicare a sinistra la matrice A per la matrice

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ -7 & 0 & 1 \end{pmatrix}$$

cioè

$$\begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ -7 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -21 \end{pmatrix}$$

Moltiplicare la seconda riga della matrice ottenuta al passo precedente per 2 e sottrarla alla terza equivale a moltiplicare a sinistra la matrice L_1A per la matrice

$$L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

cioè

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -21 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & -9 \end{pmatrix}$$

Ma allora la matrice triangolare superiore U ottenuta precedentemente con il metodo di eliminazione di Gauss è tale che

$$L_2 L_1 A = U$$

e quindi

$$A = (L_2 L_1)^{-1} U = L_1^{-1} L_2^{-1} U = LU$$

Si verifica facilmente che $L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 0 & 1 \end{pmatrix}$ e $L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$

Struttura delle matrici L e U

$$A = LU = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{n1} & m_{n2} & m_{n3} & \cdots & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & 0 & a_{33}^{(3)} & \vdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & a_{nn}^{(n)} \end{bmatrix}$$

Nota 1: Se A non soddisfa le ipotesi $\det A_k \neq 0$, ma è comunque **regolare**, tramite **scambi di righe** può essere riportata ad una matrice che soddisfa le ipotesi e che quindi può essere **fattorizzata**.

Nota 2: il **costo computazionale** della fattorizzazione è lo stesso di

quello dell'eliminazione di Gauss $C_c \simeq \frac{n^3}{3}$.

Applicazioni della fattorizzazione

Soluzione di un sistema lineare

Consideriamo il **sistema lineare** $AX = B$ e supponiamo che la matrice dei coefficienti A possa essere **fattorizzata**.

$$AX = B \xrightarrow{A=LU} L\underbrace{UX}_Y = B$$

$$\Rightarrow \begin{cases} LY = B & \text{sist. triang. inf. sost. in avanti} \\ UX = Y & \text{sistema triang. sup. sost. all'indietro} \end{cases}$$

Una volta **fattorizzata** A ,

la soluzione del sistema si ottiene risolvendo i due sistemi triangolari con **costo computazionale** $2\frac{n^2}{2}$.

Soluzione di più sistemi lineari

Se dobbiamo risolvere più sistemi lineari

$$AX_i = B_i \quad i = 1, 2, \dots, r$$

aventi la **stessa matrice** A e **diversi termini noti**, si fattorizza una volta per tutte la matrice $A = LU$ e si risolvono, per ogni vettore B_i , i due sistemi triangolari

$$\begin{cases} LY_i = B_i & i = 1, 2, \dots, r \\ UX_i = Y_i \end{cases}$$

per la soluzione dei quali il **costo computazionale** è "solo" n^2 .

Calcolo del determinante di A

$$\det A = \det(L U) = \det L \det U = \underbrace{\left(\prod_{k=1}^n l_{kk} \right)}_{\text{Teorema di Binet}} \left(\prod_{k=1}^n u_{kk} \right) = a_{11}^{(1)} a_{22}^{(2)} \cdots a_{nn}^{(n)}$$

$$\begin{array}{ccc} \downarrow & & \searrow \\ \text{Teorema di Binet} & & = 1 \end{array}$$

Se durante la fattorizzazione sono stati fatti s **scambi** di righe, allora

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

Calcolo dell'inversa di A

La matrice inversa A^{-1} di una matrice **regolare** A è la matrice tale che

$$A A^{-1} = I \quad I : \text{matrice identità}$$

$$I := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = [E_1 \ E_2 \ \cdots \ E_n] \quad E_i = [0, 0, \cdots, \underbrace{1}_i, 0, \cdots, 0]^T$$

↓
Vettori della **base canonica**

$$A A^{-1} = I \Rightarrow A X_i = E_i \quad i = 1, 2, \cdots, n \Rightarrow A^{-1} = [X_1 \ X_2 \ \cdots \ X_n]$$

Una volta nota la **fattorizzazione** di A , basta risolvere gli n sistemi

$$L Y_i = E_i, \quad U X_i = Y_i, \quad i = 1, \cdots, n$$

Costo computazionale: $C_c \simeq \frac{n^3}{3} + n \cdot n^2 = \frac{4}{3} n^3$

Calcolo del rango di A

- Se l'algoritmo di eliminazione applicato alla **matrice quadrata** $A \in \mathbf{R}^{n \times n}$ termina regolarmente dopo $n - 1$ passi \Rightarrow la matrice A ha **rango massimo** pari a n (**matrice regolare**).
- Se l'algoritmo di eliminazione applicato alla **matrice quadrata** $A \in \mathbf{R}^{n \times n}$ non può proseguire dopo q passi perchè $a_{rk}^{(k)} = 0, r = k, \dots, n \Rightarrow$ la matrice A ha **rango** pari a $q \leq n$ (**matrice singolare**).
- L'algoritmo di eliminazione applicato alla **matrice rettangolare** $A \in \mathbf{R}^{m \times n}$ termina necessariamente dopo $q \leq m$ passi \Rightarrow la matrice A ha **rango** pari a q .

Nota: le operazioni "lecite" conservano il rango.

Metodo di eliminazione di Gauss

Nel **metodo di Gauss**, come anche nella **fattorizzazione LU**, si richiedono divisioni per gli elementi della diagonale principale della matrice considerata. Se questi ultimi sono prossimi allo zero, la soluzione può non essere esatta a causa della **cancellazione numerica**.

Esempio: Supponiamo di dover risolvere il seguente sistema

$$\begin{cases} -0.0590x_1 + 0.2372x_2 = -0.3528 \\ 0.1080x_1 - 0.4348x_2 = 0.6452 \end{cases}$$

usando il metodo di eliminazione di Gauss e 4 decimali significativi.

Metodo di eliminazione di Gauss

Il rapporto

$$\frac{0.1080}{-0.0590} \approx -1.830508 \approx -1.831$$

moltiplicando questa quantità con la prima equazione del sistema e sottraendo alla seconda, il secondo elemento della seconda colonna diventa

$$-0.4348 - 0.2372(-1.831) = -0.4348 + 0.4343 = -0.0005$$

Mentre il secondo elemento del termine noto è

$$0.6452 - 0.3528(-1.831) = 0.6452 - 0.6460 = -0.0008$$

Metodo di eliminazione di Gauss

e quindi

$$\left(\begin{array}{cc|c} -0.0590 & 0.2372 & -0.3528 \\ 0 & -0.0005 & -0.0008 \end{array} \right)$$

da cui

$$x_2 = \frac{0.0008}{0.0005} = 1.6$$

$$x_1 = \frac{-0.3528 - 1.6(0.2372)}{-0.0590} = \frac{0.7323}{0.0590} = 12.41$$

Mentre la soluzione esatta è

$$x_1 = 1$$

$$x_2 = 10$$

Metodo di eliminazione di Gauss: pivoting parziale

Non si ha cancellazione numerica se si **scambiano le equazioni del sistema**, cioè

$$\begin{cases} 0.1080x_1 - 0.4348x_2 = 0.6452 \\ -0.0590x_1 + 0.2372x_2 = -0.3528 \end{cases}$$

Eseguendo un passo del metodo di eliminazione di Gauss, il sistema si riduce al sistema equivalente

$$\begin{cases} 0.1080x_1 - 0.4348x_2 = 0.6452 \\ 0x_1 + 0.3296x_2 = 0.3296 \end{cases}$$

da cui

$$x_1 = 1 \quad x_2 = 10$$

Pivoting parziale

Ad ogni passo k del metodo di eliminazione di Gauss si individua il valore $r \geq k$ per cui risulta

$$|a_{rk}^{(k)}| = \max_{k \leq s \leq n} |a_{sk}^{(k)}|,$$

dove $a_{ij}^{(k)}$ sono gli elementi della matrice del sistema al passo k ,

e si scambiano le righe r e k

Matrici tridiagonali

Nel caso in cui la matrice dei coefficienti A è **tridiagonale** la **fattorizzazione LU** è molto **semplice** in quanto le matrici L e U hanno una struttura semplice.

$$A = \begin{bmatrix} d_1 & s_1 & 0 & \cdots & 0 \\ a_2 & d_2 & s_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_{n-1} & s_{n-1} \\ 0 & 0 & \cdots & a_n & d_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha_2 & 1 & 0 & \cdots & 0 \\ 0 & \alpha_3 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & \alpha_n & 1 \end{bmatrix} \begin{bmatrix} u_1 & v_1 & 0 & \cdots & 0 \\ 0 & u_2 & v_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & u_{n-1} & v_{n-1} \\ 0 & 0 & \cdots & 0 & u_n \end{bmatrix} = LU$$

Algoritmo di Thomas

$$\left\{ \begin{array}{l} u_1 = d_1 \\ v_i = s_i \quad i = 1, 2, \dots, n-1 \\ \alpha_i = a_i/u_{i-1} \quad i = 2, 3, \dots, n \\ u_i = d_i - \alpha_i v_{i-1} \quad i = 2, 3, \dots, n \end{array} \right.$$

Costo computazionale: $C_c = 2n - 2$

Soluzione del sistema lineare $AX = B$

$$y_1 = b_1 \quad y_i = b_i - \alpha_i y_{i-1} \quad i = 2, 3, \dots, n$$

$$x_n = y_n/u_n \quad x_i = (y_i - v_i x_{i+1})/u_i \quad i = n-1, \dots, 1$$

Costo computazionale: $C_c = \underbrace{(n-1)}_{\text{moltiplicazioni}} + \underbrace{(n-1)}_{\text{moltiplicazioni}} + \underbrace{n}_{\text{divisioni}} = 3n - 2$

Esempio 1: soluzione

La **soluzione esatta** è $\bar{X} = [\bar{p}_0, \dots, \bar{p}_N]^T$, dove $\bar{p}_i = 1 - \frac{i}{N}$, $i = 0, \dots, N$.

Algoritmo di **eliminazione di Gauss**

N	$\ \bar{X} - X\ _\infty$	tempo di calcolo	occupazione di memoria
11	3.33e-016	0.000304 s	0.8 Kbyte
21	1.55e-015	0.000356 s	3.2 Kbyte
51	4.16e-015	0.000402 s	20 Kbyte
101	2.14e-014	0.001555 s	8 Kbyte
501	1.11e-013	0.070898 s	200 Kbyte
5001	1.84e-012	29.306639 s	2 Mbyte
10001	1.38e-011	225.971643 s	800 Mbyte

Algoritmo di **Thomas**

N	$\ \bar{X} - X\ _\infty$	tempo di calcolo	occupazione di memoria
11	1.11e-016	0.000061 s	624 bytes
21	1.11e-016	0.000089 s	1376 bytes
51	4.33e-015	0.000179 s	3536 bytes
101	9.10e-015	0.000454 s	7136 bytes
501	4.61e-014	0.002566 s	35936 bytes
5001	3.20e-012	0.093172 s	359936 bytes
10001	6.95e-012	0.364083 s	719936 bytes

Si osserva una notevole riduzione del tempo di calcolo, una discreta riduzione della memoria occupata e una maggiore precisione nella soluzione prodotta.

Condizionamento di un sistema lineare

Il **condizionamento** del problema della soluzione di un sistema lineare è indipendente dal **metodo numerico** scelto per risolverlo.

Il **condizionamento** "*misura*" quanto una **perturbazione** sui dati di input (matrice dei coefficienti e termine noto) influenzi i risultati (la soluzione).

Un **sistema lineare** si dice **ben condizionato** se a **piccole** variazioni sui dati corrispondono **piccole** variazioni sui risultati.

Viceversa, se a **piccole** variazioni sui dati corrispondono **grandi** variazioni sui risultati, si dice che il sistema è **mal condizionato**.

Quando si approssima la soluzione di un sistema lineare **mal condizionato** bisogna **ridurre** il più possibile gli **errori di arrotondamento**.

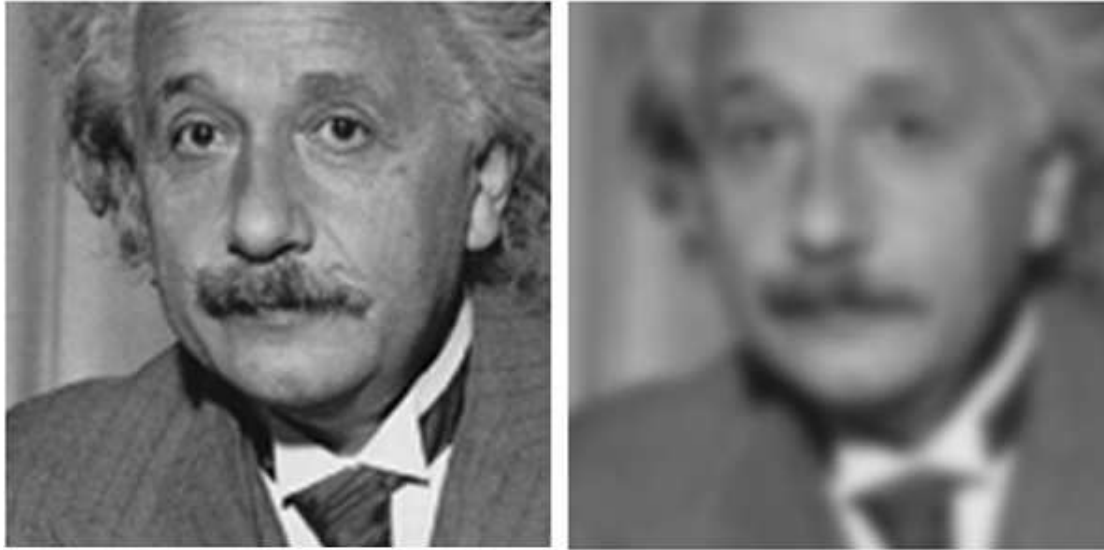
Condizionamento di un sistema lineare: Image deblurring

Un' **immagine digitale** è composta da elementi detti **pixels**. Ad ogni pixel è assegnato un valore di intensità che caratterizza il colore (livello di grigio) in una piccola porzione della scena rappresentata.

Un'immagine digitale a livelli di grigio (8 bits) è quindi una matrice i cui elementi sono numeri appartenenti all' intervallo **[0, 255]**

Piu' volte ci capita di scattare una foto e di non ottenere un' immagine nitida perchè la messa a fuoco non è stata corretta. Si ha così un' **immagine sfocata**. Una situazione molto simile si ha in diverse applicazioni come, per esempio, lo *imaging astronomico* in cui le immagini acquisite mediante telescopio possono essere disturbate dalla turbolenza dell' atmosfera.

L' **immagine sfocata** tipicamente si presenta mostrato di seguito



Ci chiediamo se si può ricostruire l' immagine originale a partire da quella sfocata (l' unico dato a disposizione)

Nel caso di **blurring lineare**, il problema consiste nel risolvere il sistema

$$AX = B$$

In cui A è l' operatore di blurring (dipendente dalla cosiddetta *point spread function*), X è l'immagine originale e B è l'immagine osservata (sfocata)

Tuttavia, se si risolve il sistema invertendo la matrice A , che si suppone nota, il risultato non è quello aspettato

Jonathan Swift
**Vision is the
art of seeing
what is
invisible to
others.** 

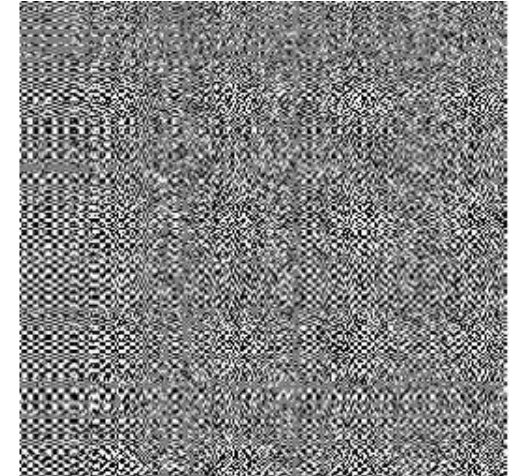


Immagine originale

Immagine sfocata

Immagine ricostruita

Cosa è accaduto???

L'inversione della matrice è molto sensibile agli errori sui dati, anche molto piccoli!

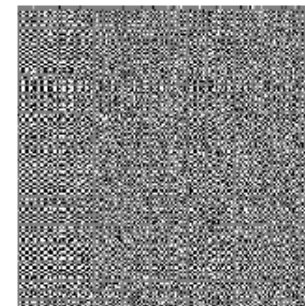
L'immagine B è sempre affetta da errori di arrotondamento che quindi possono propagarsi in maniera disastrosa sulla soluzione del sistema. In altre parole, invece di risolvere il sistema

$$AX = B$$

si risolve il sistema

$$AX = B + Err$$

e quindi nella soluzione $X = A^{-1}(B + Err) =$



il **rumore/errore** Err è molto amplificato!

Il problema si dice **mal condizionato**.

Errore sul termine noto

Supponiamo che il termine noto B sia affetto da un **errore** δB .

$$A X = B \xrightarrow{\delta B} A(X + \delta X) = B + \delta B$$

Per sottrazione si ricava

$$A \delta X = \delta B \rightarrow \boxed{\delta X = A^{-1} \delta B}$$

Per "*misurare*" la perturbazione δX indotta su X si ricorre alla norma.

$$\|\delta X\| = \|A^{-1} \delta B\| \leq \|A^{-1}\| \cdot \|\delta B\|$$

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

Dividendo termine a termine si trova una maggiorazione per l'**errore relativo** $\|\delta X\|/\|X\|$.

$$\frac{\|\delta X\|}{\|X\|} \leq \underbrace{\|A\| \cdot \|A^{-1}\|}_{K(A)} \frac{\|\delta B\|}{\|B\|} = K(A) \frac{\|\delta B\|}{\|B\|}$$

Numero di condizionamento

$K(A) := \|A\| \cdot \|A^{-1}\|$: **numero di condizionamento** della matrice A

$\frac{\|\delta X\|}{\|X\|} \leq K(A) \frac{\|\delta B\|}{\|B\|}$ Rappresenta il **coefficiente di amplificazione** dell'errore relativo sui dati

Si può dimostrare che

$$1 \leq K(A) \leq +\infty$$

Condizionamento **ottimo**
(matrici ortogonali)

Condizionamento **peggiore**
(matrici singolari)

Esempi di matrici malcondizionate: matrici di Hilbert

$$H = \begin{pmatrix} 1 & 1/2 & 1/3 & \cdots & 1/n \\ 1/2 & 1/3 & 1/4 & \cdots & 1/(n+1) \\ 1/3 & 1/4 & 1/5 & \cdots & 1/(n+2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1/n & 1/(n+1) & 1/(n+2) & \cdots & 1/(2n-1) \end{pmatrix}$$

Errore sulla matrice dei coefficienti

Se anche la matrice A è affetta da un **errore** δA si ha

$$\frac{\|\delta X\|}{\|X\|} \leq \underbrace{\frac{K(A)}{1 - K(A)} \frac{\|\delta A\|}{\|A\|}}_{\downarrow} \left(\frac{\|\delta B\|}{\|B\|} + \frac{\|\delta A\|}{\|A\|} \right)$$

Coefficiente di amplificazione

Oss: Se $\|\delta A\| \leq \frac{1}{2\|A^{-1}\|}$, allora $\frac{K(A)}{1 - K(A)} \frac{\|\delta A\|}{\|A\|} \leq 2K(A)$

Condizionamento in norma 2

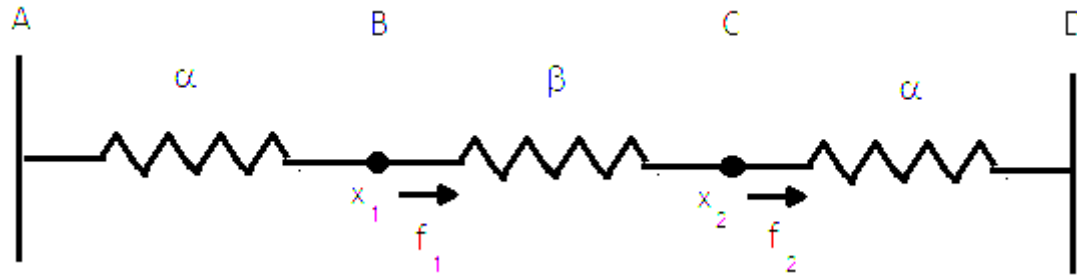
Se A è (simmetrica) **definita positiva** si ha $K_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\lambda_{max}}{\lambda_{min}}$

Infatti $\|A\|_2 = \sqrt{\rho(A^T A)} = \sqrt{\rho(A^2)} = \rho(A) = \lambda_{max}$

$$\|A^{-1}\|_2 = \rho(A^{-1}) = \max_i \frac{1}{\lambda_i} = \frac{1}{\lambda_{min}}$$

Esempio

Equilibrio delle forze elastiche



Nel sistema in figura f_1 , f_2 rappresentano le forze di stress applicate in B e C, mentre x_1 , x_2 sono gli spostamenti prodotti dalle due forze. Supponendo che le forze elastiche di reazione delle molle siano lineari (legge di Hooke) e indicando con α e β i coefficienti di stiffness (costanti elastiche), il sistema è in equilibrio se

$$\begin{cases} (\alpha + \beta)x_1 - \beta x_2 = f_1 \\ -\beta x_1 + (\alpha + \beta)x_2 = f_2 \end{cases}$$

Si tratta, dunque, di risolvere il sistema lineare rispetto al vettore delle incognite $\mathbf{X} = [x_1, x_2]^T$ avente

$$\mathbf{A} = \begin{pmatrix} \alpha + \beta & -\beta \\ -\beta & \alpha + \beta \end{pmatrix}$$

come matrice dei coefficienti e $\mathbf{B} = [f_1, f_2]^T$ come vettore dei termini noti.

Studiamo il condizionamento del sistema lineare rispetto alla norma 1 in dipendenza delle costanti elastiche α e β (equivalente a studiare il condizionamento rispetto alla norma infinito poichè \mathbf{A} è simmetrica).

$$\mathbf{A}^{-1} = \frac{1}{\alpha(\alpha + 2\beta)} \begin{pmatrix} \alpha + \beta & \beta \\ \beta & \alpha + \beta \end{pmatrix}$$

quindi, poichè $\alpha > 0$, $\beta > 0$,

$$K_1(\mathbf{A}) = \|\mathbf{A}\|_1 \|\mathbf{A}^{-1}\|_1 = (\alpha + 2\beta) \frac{1}{\alpha} = 1 + 2\frac{\beta}{\alpha}.$$

Ne segue che il problema non risulta ben condizionato per tutti i valori delle costanti elastiche: se la molla centrale è molto più rigida delle molle adiacenti, ovvero $\beta \gg \alpha$, il sistema risulta mal condizionato.

Per esempio, se $\alpha = 1$, $\beta = 1000$ e $f_1 = f_2 = 1$, allora $K_1(\mathbf{A}) = 2001$ mentre $\mathbf{X} = [1, 1]$. Se si perturba la matrice \mathbf{A} come segue

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1002 & -1000 \\ -1000 & 1002 \end{pmatrix}$$

mentre si lascia inalterato il vettore dei termini noti, si ha

$$\frac{\|\delta\mathbf{A}\|_1}{\|\mathbf{A}\|_1} = \frac{1}{2001} \approx 0.0005$$

e $\tilde{\mathbf{X}} = [.5, .5]$; quindi

$$\frac{\|\delta\mathbf{X}\|_1}{\|\mathbf{X}\|_1} \approx \frac{1}{2} = 0.5.$$

Esercizio: Studiare il condizionamento del sistema associato all'equilibrio delle forze elastiche nella struttura in figura quando le tre molle hanno tre diversi coefficienti di stiffness.

Esercizio 1

Determinare il numero di condizionamento, rispetto alla norma infinito, della seguente matrice

$$A = \begin{pmatrix} 1 & 1 + \delta \\ 1 - \delta & 1 \end{pmatrix}$$

con $\delta > 0$.

Posto $\delta = 0.01$, sia A la matrice dei coefficienti del sistema $AX = B$, con $B = (2.01, 1.99)^T$ e si consideri il sistema perturbato $A\bar{X} = \bar{B}$, con $\bar{B} = (2, 2)^T$. Dare una stima dell'errore relativo commesso sulla soluzione X .

Soluzione Si verifica facilmente che $A^{-1} = \frac{1}{\delta^2} \begin{pmatrix} 1 & -1 - \delta \\ -1 + \delta & 1 \end{pmatrix}$

e quindi $K_{\infty}(A) = \|A\|_{\infty} \|A^{-1}\|_{\infty} = (2 + \delta) \frac{2 + \delta}{\delta^2} = \frac{(2 + \delta)^2}{\delta^2}$

Si osserva che

$$\lim_{k \rightarrow +\infty} K_{\infty}(A) = 1$$

mentre

$$\lim_{k \rightarrow 0^+} K_{\infty}(A) = +\infty;$$

cioè per valori di δ molto piccoli, la matrice risulta malcondizionata.

Per esempio, se $\delta = 0.01$, $K_{\infty}(A) = (201^2) = 40401$, mentre se $\delta = 100$, $K_{\infty}(A) = \frac{10404}{10000} = 1.0404$.

$$\frac{\|\delta X\|_\infty}{\|X\|_\infty} \leq K_\infty(A) \frac{\|\delta B\|_\infty}{\|B\|_\infty}$$

e, quindi, posto $\delta = 0.01$ si ha

$$\frac{\|\delta X\|_\infty}{\|X\|_\infty} \leq 40401 \frac{\|B - \bar{B}\|_\infty}{2.01} = 40401 \frac{0.01}{2.01} = 201$$

Si verifica facilmente che la soluzione del sistema $AX = B$ è $X = (1, 1)^T$, da cui $\frac{\|\delta X\|_\infty}{\|X\|_\infty} \leq 201$

mentre $\bar{X} = (-200, 200)^T$, da cui $\|\delta X\|_\infty = 200$.

Esercizio 2

Data la matrice $A = \begin{pmatrix} 1 & -2 & -\lambda \\ -1 & 0 & -\lambda \\ 0 & 0 & -1 \end{pmatrix}$

a) studiare come varia il numero di condizionamento $K(A(\lambda))$ in norma 1 per $|\lambda| \leq \frac{1}{2}$ e trovarne il massimo.

b) Dato il sistema $A \left(\frac{1}{2}\right) X = B$, fornire una stima dell'errore relativo

$\frac{\|\delta X\|_1}{\|X\|_1}$ corrispondente a un errore relativo $\frac{\|\delta B\|_1}{\|B\|_1} = 10^{-2}$.

Traccia della soluzione

$$\mathbf{a)} \quad \|A(\lambda)\|_1 = \max(2, 2, 2|\lambda|+1) = \begin{cases} 2 & \text{per } 2|\lambda| + 1 \leq 2 \Rightarrow |\lambda| \leq \frac{1}{2} \\ 2|\lambda| + 1 & \text{per } 2|\lambda| + 1 > 2 \Rightarrow |\lambda| > \frac{1}{2} \end{cases}$$

$$\Rightarrow \|A(\lambda)\|_1 \leq 3 \text{ per } |\lambda| \leq 1$$

$$K(A(\lambda)) = \|A(\lambda)\|_1 \|A^{-1}(\lambda)\|_1$$

$$A^{-1}(\lambda) = \frac{1}{2} \begin{bmatrix} 0 & -2 & 2\lambda \\ -1 & 1 & 2\lambda \\ 0 & 0 & -2 \end{bmatrix} \quad \|A^{-1}(\lambda)\|_1 = \begin{cases} 3/2 & \text{per } |\lambda| \leq \frac{1}{4} \\ 2|\lambda| + 1 & \text{per } |\lambda| > \frac{1}{4} \end{cases}$$

$$K(A(\lambda)) = \begin{cases} 3 & \text{per } |\lambda| \leq \frac{1}{4} \\ 2(2|\lambda| + 1) & \text{per } \frac{1}{4} < |\lambda| \leq \frac{1}{2} \\ (2|\lambda| + 1)^2 & \text{per } \frac{1}{2} < |\lambda| \end{cases}$$

$$\Rightarrow \max_{|\lambda| \leq 1/2} K(A(\lambda)) = K(A(1/2)) = 4$$

$$\mathbf{b)} \quad \frac{\|\delta X\|_1}{\|X\|_1} \leq K(A(1/2)) \frac{\|\delta B\|_1}{\|B\|_1} = 4 \cdot 10^{-2}$$

Esercizio 3

Studiare il condizionamento del seguente sistema lineare

$$\begin{cases} 2x + y = 3 \\ 2x + 1.001y = 0 \end{cases}$$

Soluzione La matrice dei coefficienti del sistema è

$$A = \begin{pmatrix} 2 & 1 \\ 2 & 1.001 \end{pmatrix}$$

Il numero di condizionamento di A è : $K(A) = \|A\| \|A^{-1}\|$.

$K(A)$ dipende dalla norma di matrici scelta.

Valutiamo allora $K_1(A)$, $K_\infty(A)$ e $K_2(A)$, rispettivamente il numero di condizionamento di A rispetto alle norme 1 , ∞ e *spettrale*.

Si verifica facilmente che

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} 1.001 & -1 \\ -2 & 2 \end{pmatrix}$$

con $\det(A) = 0.002$.

Quindi

$$\|A\|_{\infty} = 3.001 \quad \|A^{-1}\|_{\infty} = \frac{4}{0.002} \quad \Rightarrow \quad K_{\infty}(A) = 3.001 \cdot 2 \cdot 1000 = 6002$$

$$\|A\|_1 = 4 \quad \|A^{-1}\|_1 = \frac{3.001}{0.002} \quad \Rightarrow \quad K_1(A) = 4 \cdot \frac{3.001}{0.002} = 6002$$

Per valutare la **norma spettrale** di A è necessario calcolare il **raggio spettrale** della matrice $A^T A$, infatti

$$\|A\|_2 = \sqrt{\rho(A^T A)}:$$

$$A^T A = \begin{pmatrix} 2 & 2 \\ 1 & 1.001 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 2 & 1.001 \end{pmatrix} = \begin{pmatrix} 8 & 4.002 \\ 4.002 & 2.002001 \end{pmatrix}$$

i cui autovalori λ_1 e λ_2 sono soluzione della seguente equazione di secondo grado

$$\lambda^2 - 10.002001\lambda + 0.000004$$

Poichè $\rho(A^T A) = \max\{|\lambda_1|, |\lambda_2|\} = 10.00200060008001$, risulta

$$\|A\|_2 = \sqrt{\rho(A^T A)} \approx 3.1626.$$

Analogamente

$$A^{-1T}A^{-1} = \frac{1}{(\det(A))^2} \begin{pmatrix} 1.001 & -2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1.001 & -1 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 5.002001 & -5.001 \\ -5.001 & 5 \end{pmatrix}$$

da cui risulta

$$\|A^{-1}\|_2 = \sqrt{\rho(A^{-1T}A^{-1})} \approx 1.5813 \cdot 10^3$$

e quindi

$$K_2(A) = \|A\|_2 \|A^{-1}\|_2 = 3.1626 \cdot 1.5813 \cdot 10^3 = 5.0010 \cdot 10^3.$$

In tutti e tre i casi il **numero di condizionamento è molto alto**, ne segue che una piccola perturbazione sui dati, produce un errore non trascurabile sulla soluzione.

Per esempio, la soluzione del sistema precedente è $\mathbf{x} = (x, y)$, con $x = 1501.5$ e $y = -3000$.

Supponiamo ora di aver un errore pari a 0.001 sul coefficiente a_{22} della matrice A , cioè di dover risolvere il sistema

$$\begin{cases} 2x + y = 3 \\ 2x + 1.002y = 0 \end{cases}$$

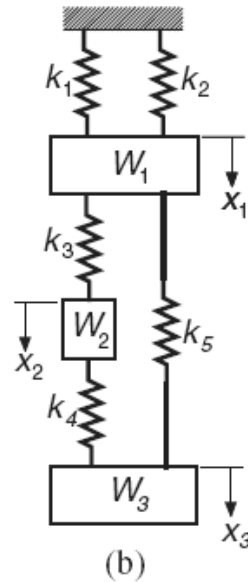
In questo caso la soluzione diventa $\tilde{\mathbf{x}} = (\tilde{x}, \tilde{y})$, con $\tilde{x} = 751.5$ e $\tilde{y} = -1500$, cioè, scegliendo la norma infinito,

$$\frac{\|\delta\mathbf{x}\|_{\infty}}{\|\mathbf{x}\|_{\infty}} = \frac{1500}{3000} = \frac{1}{2}$$

mentre

$$\frac{\|\delta\mathbf{A}\|_{\infty}}{\|\mathbf{A}\|_{\infty}} = \frac{0.001}{3.001} = 3.33 \cdot 10^{-4}.$$

Esercizio 4



Il sistema in figura, , costituito da 4 molle che sostengono 3 pesi W_i , $i = 1, \dots, 3$, all'equilibrio soddisfa

$$\begin{cases} (k_1 + k_2 + k_3 + k_5)x_1 & -k_3 x_2 & -k_5 x_3 & = W_1 \\ -k_3 x_1 & +(k_3 + k_4) x_2 & -k_4 x_3 & = W_2 \\ -k_5 x_1 & -k_4 x_2 + & (k_4 + k_5) x_3 & = W_3 \end{cases}$$

con k_i costante elastica della molla i -esima e x_j spostamento della massa j -esima rispetto alla posizione nel sistema non deformato.

Ponendo $k_1 = k_3 = k_4 = k$, $k_2 = k_5 = 2k$, $W_1 = W_3 = 2W$ e $W_2 = W$, stabilire se il numero di condizionamento del sistema, calcolato rispetto alle norme $1, 2$ e *infinito*, dipende dal valore del parametro k . Produrre una maggiorazione dell'errore relativo sulla soluzione nel caso in cui il valore del parametro W sia dato con errore pari a 0.03 .

Soluzione:

La matrice dei coefficienti del sistema è $A = \begin{pmatrix} 6k & -k & -2k \\ -k & 2k & -k \\ -2k & -k & 3k \end{pmatrix}$ da cui

$$A^{-1} = \frac{1}{k} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{14}{15} & \frac{8}{15} \\ \frac{1}{3} & \frac{8}{15} & \frac{11}{15} \end{pmatrix}$$

Poichè A , e quindi A^{-1} , è simmetrica, il numero di condizionamento di A rispetto alla norma 1 coincide con quello valutato rispetto alla norma infinito, cioè

$$K_1(A) = K_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty.$$

Poichè

$$\|A\|_{\infty} = \max\{9k, 4k, 6k\} = 9k$$

$$\|A^{-1}\|_{\infty} = \max\left\{\frac{1}{k}, \frac{27}{15k}, \frac{24}{15k}\right\} = \frac{27}{15k}$$

risulta

$$K_1(A) = K_{\infty}(A) = 9k \frac{27}{15k} = \frac{81}{5}$$

Quindi, il condizionamento del sistema rispetto alle norme 1 e infinito non dipende dal valore del parametro k .

Gli autovalori di A e A^{-1} non dipendono dal parametro k in quanto esso moltiplica tutti gli elementi delle due matrici, quindi anche $K_2(A)$ non dipende dal valore del parametro k e per calcolarlo basta valutare il numero di condizionamento delle matrici

$$A_1 = \begin{pmatrix} 6 & -1 & -2 \\ -1 & 2 & -1 \\ -2 & -1 & 3 \end{pmatrix} \quad \text{e} \quad A_1^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{14}{15} & \frac{8}{15} \\ \frac{1}{3} & \frac{8}{15} & \frac{11}{15} \end{pmatrix}$$

L'errore relativo della soluzione del sistema valutato rispetto alla norma infinito soddisfa

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} < K_\infty(A) \frac{\|\delta b\|_\infty}{\|b\|_\infty}$$

in quanto la matrice A si suppone non affetta da errori. Poichè W ha errore 0.03, $\|\delta b\|_\infty = \max\{2 \cdot 0.03, 0.03, 2 \cdot 0.03\} = 0.06$, e quindi

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} < \frac{81}{5} \frac{6}{100} \frac{1}{2W} = \frac{243}{500 W}$$

Esercizio 5

Data la matrice $A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 3 & 4 & 5 & 6 \\ 0 & 0 & 5 & 6 & 7 \\ 0 & 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 0 & 9 \end{pmatrix}$

e il termine noto $b = \begin{pmatrix} 15 \\ 18 \\ 18 \\ 15 \\ 9 \end{pmatrix}$

1. Risolvere il sistema $AX = b$ con un opportuno metodo diretto
2. Sia $\delta b = \text{rand}(5, 1) * 10^{-3}$ una perturbazione di b ; dare una stima dell'errore relativo commesso risolvendo il sistema $AX = b + \delta b$ (usare le norme 1, 2 e ∞)

Sistemi lineari: Metodi iterativi

- la soluzione si ottiene tramite approssimazioni successive (**metodi del punto unito**)
- soluzione esatta si ottiene in un numero infinito di passi (**errore di troncamento**)
- bassa occupazione di memoria
- problemi sparsi e/o di elevate dimensioni

Sistemi non lineari

$$F(X) = 0$$

$$X, F \in \mathbf{R}^n$$

$$\rightarrow F(X) = AX - B$$

Sistemi lineari

$$AX = B$$

$$X, B \in \mathbf{R}^n \quad A \in \mathbf{R}^{n \times n}$$



$$X = [x_1, x_2, \dots, x_n]^T$$

$$F = [f_1(X), f_2(X), \dots, f_n(X)]^T$$

$$0 = [0, 0, \dots, 0]^T$$

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \dots\dots\dots \\ \dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases}$$



$$X = [x_1, x_2, \dots, x_n]^T$$

$$A = [a_{ij}]_{i,j=0}^n$$

$$B = [b_1, b_2, \dots, b_n]^T$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots\dots \\ \dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}$$

Metodo del punto unito

Sistemi non lineari

$$F(X) = 0 \Leftrightarrow X = \Phi(X) \text{ con } \Phi = [\varphi_1(X), \varphi_2(X), \dots, \varphi_n(X)]^T$$

Se $\bar{X} \in \mathbb{R}^n$ è **radice** di F allora è **punto unito** di Φ :

$$F(\bar{X}) = 0 \Leftrightarrow \bar{X} = \Phi(\bar{X})$$

Sistemi lineari

$$AX = B \Leftrightarrow X = CX + Q \text{ con } Q = [q_1, q_2, \dots, q_n]^T$$

$$C = [c_{ij}]_{i,j=0}^n$$

Se $\bar{X} \in \mathbb{R}^n$ è **soluzione** di $AX = B$ allora è **punto unito** di $\Phi = CX + Q$:

$$A\bar{X} = B \Leftrightarrow \bar{X} = C\bar{X} + Q$$

Metodi iterativi a un punto

Il **punto unito** $\bar{X} = \Phi(\bar{X})$, $\bar{X} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n]^T$, può essere **approssimato** generando la successione

$$\left\{ \begin{array}{l} X^{(0)} \text{ dato} \\ X^{(k)} = \Phi(X^{(k-1)}) \\ k = 1, 2, \dots \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} X^{(0)} = [x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}]^T \text{ dato} \\ x_1^{(k)} = \varphi_1(x_1^{(k-1)}, x_2^{(k-1)}, \dots, x_n^{(k-1)}) \\ x_2^{(k)} = \varphi_2(x_1^{(k-1)}, x_2^{(k-1)}, \dots, x_n^{(k-1)}) \\ \dots\dots\dots \\ x_n^{(k)} = \varphi_n(x_1^{(k-1)}, x_2^{(k-1)}, \dots, x_n^{(k-1)}) \end{array} \right.$$

Le funzione φ_i sono chiamate **funzioni di iterazione**.

Metodi iterativi per sistemi lineari

Nel caso **lineare** il **punto unito** $\bar{X} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n]^T$, può essere **approssimato** generando la successione

$$\begin{cases} X^{(k)} = \mathbf{C}X^{(k-1)} + Q & k = 1, 2, \dots \\ X^{(0)} \in \mathbf{R}^n & \text{dato} \end{cases}$$

$$\Rightarrow x_i^{(k)} = \sum_{j=1}^n c_{ij} x_j^{(k-1)} + q_i \quad i = 1, \dots, n$$

La matrice $\mathbf{C} \in \mathbf{R}^{n \times n}$ è chiamata **matrice di iterazione**.

Convergenza

Per poter definire la **convergenza** di un metodo iterativo dobbiamo prima di tutto definire l'**errore di troncamento**

Errore di troncamento: $E^{(k)} = \bar{X} - X^{(k)} \in \mathbb{R}^n$

\swarrow \searrow

soluzione esatta **soluzione approssimata**

Per "*misurare*" la lunghezza di un vettore $V \in \mathbb{R}^n$ si ricorre alla **norma di vettore**:

$$\|V\| = \left(\sum_{i=1}^n |v_i|^p \right)^{1/p}$$

Convergenza: $\lim_{k \rightarrow \infty} \|E^{(k)}\| = 0 \iff \lim_{k \rightarrow \infty} X^{(k)} = \bar{X}$

Se il metodo iterativo è **convergente**, in assenza di errori di arrotondamento si ottiene la **soluzione esatta** dopo un **numero infinito** di passi.

Nota. In pratica ci si arresta quando $\|E^{(k)}\| \leq \epsilon$ (**criterio di arresto**)

Convergenza: condizione necessaria

Tramite la **norma di vettore** si può "misurare" la **lunghezza** del vettore errore di troncamento, cioè la **distanza** tra la soluzione esatta e quella approssimata.

$$\text{Convergenza: } \lim_{k \rightarrow \infty} \|E^{(k)}\| = 0 \iff \lim_{k \rightarrow \infty} X^{(k)} = \bar{X}$$

Teorema. Sia S uno **spazio vettoriale normato** e sia $\Phi : S \rightarrow S$. Se la successione $\{X^{(k)}\} = \{\Phi(X^{(k-1)})\}$ è **convergente** a un valore $\bar{X} \in S$ e l'applicazione Φ è **continua** in $\bar{X} \Rightarrow \bar{X}$ è **punto unito** di Φ , cioè $\bar{X} = \Phi(\bar{X})$.

Dim.

$$\bar{X} = \lim_{k \rightarrow \infty} X^{(k)} = \lim_{k \rightarrow \infty} \Phi(X^{(k-1)}) = \Phi\left(\lim_{k \rightarrow \infty} X^{(k-1)}\right) = \Phi(\bar{X})$$

Convergenza: condizione sufficiente

Definizione. Un'applicazione $\Phi : S \rightarrow S$, dove S è uno **spazio normato** è detta **contrazione**, se esiste $\lambda \in (0, 1)$ tale che

$$\|\Phi(X) - \Phi(Y)\| \leq \lambda \|X - Y\| < \|X - Y\| \quad \forall X, Y \in S$$

Teorema. Sia $D \subset \mathbb{R}^n$. Se $\Phi : D \rightarrow D$ è una **contrazione**

\Rightarrow • esiste un **unico punto unito** $\bar{X} \in D$ di Φ

- la successione $\{X^{(k)}\} = \{\Phi(X^{(k-1)})\}$ è **convergente** a \bar{X} per ogni **approssimazione iniziale** $X^{(0)} \in D$

Contrazione: condizione sufficiente

Matrice Jacobiana di Φ

$$J(X) = \begin{bmatrix} \frac{\partial \varphi_1(X)}{\partial x_1} & \frac{\partial \varphi_1(X)}{\partial x_2} & \dots & \frac{\partial \varphi_1(X)}{\partial x_n} \\ \frac{\partial \varphi_2(X)}{\partial x_1} & \frac{\partial \varphi_2(X)}{\partial x_2} & \dots & \frac{\partial \varphi_2(X)}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \varphi_n(X)}{\partial x_1} & \frac{\partial \varphi_n(X)}{\partial x_2} & \dots & \frac{\partial \varphi_n(X)}{\partial x_n} \end{bmatrix}$$

Teorema. Se *i)* le **funzioni di iterazione** $\varphi_1, \varphi_2, \dots, \varphi_n$ sono **continue** e **parzialmente derivabili** in D ;

ii) esiste $\lambda \in (0, 1)$ tale che $\|J(X)\| \leq \lambda$ per $X \in D$

$\Rightarrow \Phi$ è una **contrazione** in D

dim: $\forall X, Y \in D$, si può valutare $\|\Phi(Y) - \Phi(X)\|$ considerando lo sviluppo in serie di Taylor arrestato al primo ordine della funzione Φ attorno a X , da cui risulta

$$\|\Phi(Y) - \Phi(X)\| \leq \|J(X)\| \|Y - X\| \leq \lambda \|Y - X\|.$$

Metodi iterativi per sistemi lineari: condizione sufficiente di convergenza

Matrice Jacobiana di $\Phi = \mathbf{C}X + Q$

$$J(X) = \begin{bmatrix} \frac{\partial \varphi_1(X)}{\partial x_1} & \frac{\partial \varphi_1(X)}{\partial x_2} & \cdots & \frac{\partial \varphi_1(X)}{\partial x_n} \\ \frac{\partial \varphi_2(X)}{\partial x_1} & \frac{\partial \varphi_2(X)}{\partial x_2} & \cdots & \frac{\partial \varphi_2(X)}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial \varphi_n(X)}{\partial x_1} & \frac{\partial \varphi_n(X)}{\partial x_2} & \cdots & \frac{\partial \varphi_n(X)}{\partial x_n} \end{bmatrix} = \mathbf{C}$$

Corollario. Se esiste $\lambda \in (0, 1)$ tale che $\|\mathbf{C}\| \leq \lambda$
 $\Rightarrow \Phi$ è una **contrazione** per ogni X
 \Rightarrow il metodo iterativo è **convergente**

Dim:

$$\forall X \neq Y, \quad \|\Phi(X) - \Phi(Y)\| = \|\mathbf{C}(X - Y)\| \leq \|\mathbf{C}\| \|X - Y\| \leq \lambda \|X - Y\|$$

Condizione sufficiente di convergenza

Teorema. Condizione sufficiente affinché un metodo iterativo sia **convergente** a \bar{X} per **qualsiasi scelta** del vettore iniziale $X^{(0)}$, è che

$$\|C\| < 1$$

$$E^{(k)} = \bar{X} - X^{(k)} = (C\bar{X} + Q) - (CX^{(k-1)} + Q) =$$

$$= C(\bar{X} - X^{(k-1)}) = CE^{(k-1)} \Rightarrow \boxed{E^{(k)} = CE^{(k-1)}}$$

$$\|E^{(k)}\| = \|CE^{(k-1)}\| = \|C^2E^{(k-2)}\| = \dots = \|C^k E^{(0)}\| \leq \|C^k\| \cdot \|E^{(0)}\|$$

↓
Relazione di compatibilità

$$\|E^{(k)}\| \leq \|C^k\| \cdot \|E^{(0)}\| = \underbrace{\|C \cdot C \dots C\|}_{k \text{ volte}} \cdot \|E^{(0)}\| \leq \|C\|^k \cdot \|E^{(0)}\|$$

$$\boxed{\text{se } \|C\| < 1 \Rightarrow \lim_{k \rightarrow \infty} \|E^{(k)}\| = 0}$$

Condizione necessaria e sufficiente di convergenza

Definizione. Una matrice $A \in \mathbb{R}^{n \times n}$ si dice **convergente** se $\lim_{k \rightarrow \infty} A^k = 0$
(o, in modo equivalente, se $\lim_{k \rightarrow \infty} \|A^k\| = 0$)

Teorema. Una matrice $A \in \mathbb{R}^{n \times n}$ è **convergente** se e solo se $\rho(A) < 1$

dim: \Rightarrow Se $\lim_{k \rightarrow \infty} A^k = 0$, allora definitivamente $\|A^k\| < 1$. Poichè, dalla relazione di compatibilità risulta $\rho(A^k) < \|A^k\|$, segue che $\rho(A^k) < 1$; ma $\rho^k(A) = \rho(A^k)$, quindi $\rho^k(A) < 1$, ovvero $\rho(A) < 1$.

\Leftarrow Se $\rho(A) < 1$ allora $\exists \epsilon : \rho(A) < 1 - \epsilon$. Inoltre, poichè $\rho(A) \leq \|A\|$, esiste una norma compatibile per cui $\|A\| < \rho(A) + \epsilon$; quindi $\|A\| < 1 - \epsilon + \epsilon = 1$. Ne segue che $\lim_{k \rightarrow \infty} \|A\|^k = 0$. Il teorema risulta dimostrato osservando che $0 \leq \|A^k\| \leq \|A\|^k$.

Teorema. Un metodo iterativo **converge** per **qualunque scelta** del vettore iniziale $X^{(0)}$, **se e solo se**

$$\rho(C) < 1$$

dove $\rho(C) = \max_{1 \leq i \leq n} |\lambda_i|$ è il **raggio spettrale** della matrice di iterazione C

$$E^{(k)} = C E^{(k-1)} = \dots = C^k E^{(0)}$$

$$\lim_{k \rightarrow \infty} E^{(k)} = \lim_{k \rightarrow \infty} C^k E^{(0)} = 0 \iff \underbrace{\lim_{k \rightarrow \infty} C^k = 0}_{\text{matrice convergente}} \iff \rho(C) < 1$$

Criterio d'arresto

Se il metodo iterativo è **convergente**, si arresta il procedimento quando

$$\|X^{(k+1)} - X^{(k)}\| < \varepsilon \quad \varepsilon: \text{tolleranza prefissata}$$

$$\begin{aligned} \bullet \quad \|E^{(k)}\| &= \|X^{(k)} - \bar{X}\| = \|X^{(k)} - X^{(k+1)} + X^{(k+1)} - \bar{X}\| = \\ &= \|X^{(k)} - X^{(k+1)} + E^{(k+1)}\| \leq \underbrace{\|X^{(k+1)} - X^{(k)}\| + \|E^{(k+1)}\|}_{\downarrow} \\ \bullet \quad \|E^{(k+1)}\| &\leq \|C\| \cdot \|E^{(k)}\| \leq \|C\| (\|X^{(k+1)} - X^{(k)}\| + \|E^{(k+1)}\|) \end{aligned}$$

$$\|E^{(k+1)}\| \leq \frac{\|C\|}{1 - \|C\|} \cdot \|X^{(k+1)} - X^{(k)}\|$$

d'altra parte

$$\|E^{(k+1)}\| \leq \|C\|^k E^{(1)} \leq \frac{\|C\|^{k+1}}{1 - \|C\|} \cdot \|X^{(1)} - X^{(0)}\| \leq \varepsilon$$

Stima a priori: il numero di iterazioni K necessario affinché

$\|E^{(K)}\| < \varepsilon$, è dato da

$$K > \log \left(\frac{(1 - \|C\|) \varepsilon}{\|X^{(1)} - X^{(0)}\|} \right) \frac{1}{\log \|C\|}$$

Velocità asintotica di convergenza

Se $A \in \mathbf{R}^{n \times n}$ è una **matrice convergente**, vale la proprietà

$$\lim_{k \rightarrow \infty} \sqrt[k]{\|A^k\|} = \rho(A)$$

Dalla relazione

$$\|E^{(k)}\| \leq \|C^k\| \cdot \|E^{(0)}\|$$

per k "grande" si ottiene

$$\frac{\|E^{(k)}\|}{\|E^{(0)}\|} \leq \|C^k\| \approx \rho^k(C)$$

⇒ L'errore si riduce di un fattore 10^{-m} all'iterazione

$$K \simeq -\frac{m}{\text{Log } \rho(C)}$$

Velocità asintotica di convergenza: $V = -\text{Log } \rho(C)$

Esercizio 1

Data la matrice di iterazione $C(\beta) = \begin{pmatrix} 0 & \beta & -\frac{\beta}{2} \\ 0 & \beta & -\frac{\beta}{2} \\ 0 & -\frac{1}{2} & \frac{1}{4} \end{pmatrix}$, $\beta \in \mathbf{R}$, e il vettore $Q = (7/8, 7/8, -1/2)^T$,

1.1) determinare per quali valori di β il procedimento iterativo

$$\begin{cases} X^{(k+1)} = C(\beta)X^{(k)} + Q, & k = 0, 1, \dots & X^{(k)} \in \mathbf{R}^3 \\ X^{(0)} \text{ dato} \end{cases}$$

risulta **sicuramente convergente**;

1.2) posto $\beta = 1/2$, $X^{(0)} = (0, 0, 0)^T$, dare una stima del numero di iterazioni necessarie affinché l'approssimazione abbia 5 decimali esatti.

Traccia della soluzione

1.1) Condizione sufficiente di convergenza: $\|C\|_1 < 1$ oppure $\|C\|_\infty < 1$

$$\|C\|_1 = \max\left(2|\beta| + \frac{1}{2}, |\beta| + \frac{1}{4}\right) = 2|\beta| + \frac{1}{2} \quad \|C\|_1 < 1 \Rightarrow |\beta| < \frac{1}{4}$$

$$\|C\|_\infty = \max\left(\frac{3}{2}|\beta|, \frac{3}{4}\right) = \begin{cases} \frac{3}{4} & |\beta| \leq \frac{1}{2} \\ \frac{3}{2}|\beta| & |\beta| > \frac{1}{2} \end{cases} \quad \|C\|_\infty < 1 \Rightarrow |\beta| < \frac{2}{3}$$

Attenzione: le condizioni su β sono diverse a seconda della norma che si sceglie.

$$1.2) \beta = \frac{1}{2} \Rightarrow \|C(\frac{1}{2})\|_1 = \frac{3}{2} > 1 \quad \|C(\frac{1}{2})\|_\infty = \frac{3}{4} < 1$$

Attenzione: in **norma uno** non si può stabilire se il metodo converge, si ha invece convergenza in **norma infinito**, infatti

$$\lim_{k \rightarrow \infty} \|E^{(k)}\|_\infty = 0$$

$$K > \log \left(\frac{(1 - \|C(\frac{1}{2})\|_\infty)\varepsilon}{\|X^{(1)} - X^{(0)}\|_\infty} \right) \frac{1}{\log \|C(\frac{1}{2})\|_\infty}$$

$$\|C(\frac{1}{2})\|_\infty = \frac{3}{4}, \quad \|X^{(1)} - X^{(0)}\|_\infty = \|Q\|_\infty = \max |q_i| = \frac{7}{8}, \quad \varepsilon = 0.5 \cdot 10^{-5}$$

$$\Rightarrow K \geq 47$$

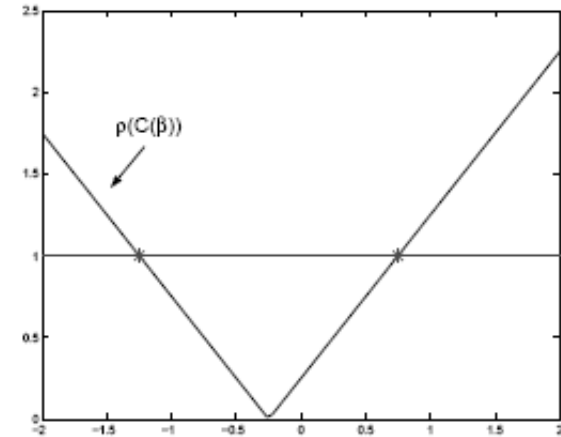
Condizione necessaria e sufficiente di convergenza:

$$\rho(C(\beta)) < 1$$

Autovalori di $C(\beta)$: 0 con molteplicità 2 , $\frac{1}{4} + \beta$

Raggio spettrale di $C(\beta)$: $\rho(C(\beta)) = \left| \frac{1}{4} + \beta \right|$

$$\rho(C(\beta)) < 1 \Rightarrow \boxed{-\frac{5}{4} < \beta < \frac{3}{4}}$$



Velocità di convergenza: $V(\beta) = -\text{Log}(\rho(C(\beta))) = -\text{Log} \left| \frac{1}{4} + \beta \right|$

Nota: questo intervallo di β **contiene** entrambi gli intervalli trovati con la condizione sufficiente.

Esercizio 2

Supponendo di applicare il metodo iterativo

$$X^{(k+1)} = X^{(k)} + \beta(AX^{(k)} - B)$$

al sistema lineare

$$\begin{cases} 2x_1 + x_2 = 5 \\ x_1 + 2x_2 = 1 \end{cases}$$

determinare un intervallo di valori di β per il quale il metodo sia convergente e calcolare in funzione di β la costante di contrazione in una norma a scelta.

Soluzione

Il metodo iterativo si può riscrivere come

$$X^{(k+1)} = (I + \beta A)X^{(k)} - \beta B$$

con matrice di iterazione è $C = I + \beta A$, cioè

$$C = \begin{pmatrix} 1 + 2\beta & \beta \\ \beta & 1 + 2\beta \end{pmatrix}$$

Poichè C è simmetrica, $\|\cdot\|_\infty = \|\cdot\|_1$.

Condizione sufficiente per la convergenza

$$\|C\|_\infty = |1 + 2\beta| + |\beta| < 1 \Leftrightarrow \beta \in \left(-\frac{2}{3}, 0\right)$$

Condizione necessaria e sufficiente per la convergenza $\rho(C) < 1$, cioè

$$\det \begin{pmatrix} 1 + 2\beta - \lambda & \beta \\ \beta & 1 + 2\beta - \lambda \end{pmatrix} = 0 \Leftrightarrow (1 + 2\beta - \lambda)^2 - \beta^2 = 0$$

$$\Leftrightarrow \lambda^2 - 2\lambda(1 + 2\beta) + (1 + 2\beta)^2 - \beta^2 = 0$$

$$\Leftrightarrow \lambda_{1,2} = (1 + 2\beta) \pm \sqrt{(1 + 2\beta)^2 - (1 + 2\beta)^2 + \beta^2}$$

$$\Leftrightarrow \lambda_{1,2} = 1 + 2\beta \pm \beta.$$

I due autovalori di C sono $\lambda_1 = 1 + \beta$ e $\lambda_2 = 1 + 3\beta$ da cui

$$\rho(C) = \max\{|1 + \beta|, |1 + 3\beta|\} < 1 \Leftrightarrow \beta \in \left(-\frac{2}{3}, 0\right)$$

Costruzione di metodi iterativi

$$AX = B \quad \longrightarrow \quad X = CX + Q$$

Splitting di A : $A = M + N$ dove M è una matrice **invertibile**.

$$AX = B \rightarrow (M + N)X = B \rightarrow MX = -NX + B \rightarrow$$

$$\rightarrow X = -M^{-1}NX + M^{-1}B \Rightarrow \boxed{C = -M^{-1}N \quad Q = M^{-1}B}$$

Una possibile **decomposizione** di A è

$$\boxed{A = L + D + U}$$

dove $D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$ (elementi **diagonali** di A)

$$L = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ a_{21} & 0 & \cdots & \cdots & 0 \\ a_{31} & a_{32} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{bmatrix}$$

(elementi di A al di **sotto**
della diagonale principale)

$$U = \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & a_{n-1,n} \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}$$

(elementi di A al di **sopra**
della diagonale principale)

Metodo di Jacobi

$$M = D, N = L + U$$

$$\begin{cases} C_J = -D^{-1}(L + U) \\ Q_J = D^{-1}B \end{cases}$$

$$C_J = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{bmatrix} \quad Q_J = \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \dots \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$

Algoritmo J

$$X^{(k)} = C_J X^{(k-1)} + Q_J \Rightarrow x_i^{(k+1)} = \frac{1}{a_{ii}} \left(- \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} x_j^{(k)} + b_i \right)$$

$$i = 1, 2, \dots, n; k \geq 0$$

Infatti, D^{-1} è una matrice diagonale i cui elementi diagonali sono $-\frac{1}{a_{ii}}$, $i = 1, \dots, n$, mentre la matrice $(L + U)$ ha gli elementi sulla diagonale principale tutti uguali a 0.

Moltiplicare a sinistra la matrice $(L + U)$ per D^{-1} , corrisponde a dividere ogni elemento della i -esima riga di $(L + U)$ per $-\frac{1}{a_{ii}}$

Metodo di Gauss-Seidel

$$M = D + L, N = U \Rightarrow \begin{cases} C_{GS} = -(D + L)^{-1}U \\ Q_{GS} = (D + L)^{-1}B \end{cases}$$

$$X^{(k)} = C_{GS}X^{(k-1)} + Q_{GS}$$



Algoritmo GS

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(-\sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_i \right)$$

$$i = 1, 2, \dots, n; k \geq 0$$

Infatti, riscrivendo $X^{(k+1)} = C_{GS}X^{(k)} + Q_{GS}$ come $(L + D)X^{(k+1)} = -UX^{(k)} + B$ e considerando che $D + L$ è una matrice triangolare inferiore mentre U è triangolare superiore con elementi nulli sulla diagonale principale, si ha

$$\sum_{j=1}^i a_{ij} x_j^{(k+1)} = -\sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_i.$$

L' **algorithm GS** si ottiene, quindi, isolando la componente $x_i^{(k+1)}$ al primo membro.

Convergenza dei metodi di Jacobi e Gauss-Seidel

- Per **verificare** la convergenza si possono applicare alle matrici di iterazione C_J e C_{GS} le condizioni già viste:

$$\left\{ \begin{array}{l} \text{C.S.:} \quad \|C_J\|_1, \|C_J\|_\infty < 1 \quad \text{e} \quad \|C_{GS}\|_1, \|C_{GS}\|_\infty < 1 \\ \text{C.N.S.:} \quad \rho(C_J) < 1 \quad \text{e} \quad \rho(C_{GS}) < 1 \end{array} \right.$$

Attenzione: Le **condizioni di convergenza** per le matrici di iterazione C_J e C_{GS} vanno **verificate** di volta in volta.

- Per alcune matrici A potrebbe convergere **solo uno** dei due metodi.
- Se convergono entrambi i metodi, quello di **Gauss-Seidel** converge **più velocemente**.

Esercizio

Dato il sistema lineare $AX = B$ dove

$$A = \begin{pmatrix} 1 & 1 & -2 \\ 2 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix}, \quad B = (1, 0, 0)^T,$$

1.1) verificare quale tra i metodi di Jacobi e Gauss-Seidel risulta convergente;

1.2) approssimare la soluzione del sistema lineare con 6 decimali esatti.

Traccia della soluzione

$$\mathbf{1.1)} \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad L = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 2 & 1 & 0 \end{pmatrix} \quad U = \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$
$$C_J = \begin{pmatrix} 0 & -1 & 2 \\ -2 & 0 & -2 \\ -2 & -1 & 0 \end{pmatrix} \quad C_{GS} = \begin{pmatrix} 0 & -1 & 2 \\ 0 & 2 & -6 \\ 0 & 0 & 2 \end{pmatrix}$$

Entrambe le matrici hanno norma **maggiore di 1**, quindi la **condizione sufficiente** di convergenza **non è soddisfatta**. Per verificare se è soddisfatta la **condizione necessaria e sufficiente** bisogna calcolare il **raggio spettrale** delle matrici di iterazione.

Autovalori di C_J : 0 con molteplicità 3

$\Rightarrow \rho(C_J) = 0 \Rightarrow$ il metodo di Jacobi **converge**

Autovalori di C_{GS} : 0 con molteplicità 1, 2 con molteplicità 2

$\Rightarrow \rho(C_{GS}) = 2 \Rightarrow$ il metodo di Gauss-Seidel **non converge**

1.2) Per approssimare la soluzione del sistema lineare con **6 decimali esatti**, si può utilizzare il **metodo di Jacobi** arrestando le iterazioni quando

$$\|E^{(k+1)}\| \simeq \|X^{(k+1)} - X^{(k)}\| \leq 0.5 \cdot 10^{-6} \text{ (criterio di arresto)}$$

Iterazioni

$$X^{(0)} = Q_J = D^{-1}B = (1, 0, 0)^T$$

$$X^{(1)} = C_J X^{(0)} + Q_J = (1, -2, -2)^T \quad \|X^{(1)} - X^{(0)}\|_{\infty} = 2$$

$$X^{(2)} = C_J X^{(1)} + Q_J = (1, 2, 0)^T \quad \|X^{(2)} - X^{(1)}\|_{\infty} = 4$$

$$X^{(3)} = C_J X^{(2)} + Q_J = (1, 2, 0)^T \quad \|X^{(3)} - X^{(2)}\|_{\infty} = 0$$

Nota 1. In questo caso la soluzione è **esatta**.

Nota 2. La **velocità di convergenza** è $V_J = -\text{Log}(\rho(C_J)) = \infty$.

Convergenza per matrici A con struttura speciale

Matrici diagonalmente dominanti:

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad i = 1, 2, \dots, n$$

(diagonale dominante per righe)

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ji}| \quad i = 1, 2, \dots, n$$

(diagonale dominante per colonne)

Condizione sufficiente di convergenza:

Teorema. Se A è **diagonalmente dominante** per righe o per colonne, i **metodi di Jacobi** e **Gauss-Seidel** sono entrambi **convergenti** per qualunque scelta dell'approssimazione iniziale $X^{(0)}$.

Per esempio, se A è **diagonalmente dominante per righe** si ha

$$\|C_J\|_\infty = \max_{1 \leq i \leq n} \sum_{\substack{j=1 \\ j \neq i}}^n \left| -\frac{a_{ij}}{a_{ii}} \right| = \max_{1 \leq i \leq n} \left(\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right) \frac{1}{|a_{ii}|} < 1$$

Matrici (simmetriche) definite positive:

Una matrice quadrata $A \in \mathbf{R}^{n \times n}$ **simmetrica** è **definita positiva** se

$$X^T A X = \sum_{i,j=1}^n a_{ij} x_i x_j > 0 \quad \forall X \in \mathbf{R}^n$$

Per riconoscere se una matrice è definita positiva si può usare il **criterio di Sylvester**:

Affinché una matrice $A \in \mathbf{R}^{n \times n}$ **simmetrica** sia **definita positiva**, è **necessario e sufficiente** che

$$\det A_k > 0 \quad k = 1, 2, \dots, n,$$

dove A_k sono le **sottomatrici principali di testa** di A .

Condizione sufficiente di convergenza:

Teorema. Se A è (simmetrica) **definita positiva**, il **metodo di Gauss-Seidel** è **convergente** per qualunque scelta dell'approssimazione iniziale $X^{(0)}$.

Esercizio

Dato il sistema

$$\begin{cases} 10x_1 + x_2 + x_3 = 2 \\ x_1 + \alpha x_2 + x_3 = 1 \\ 2x_1 + x_2 + \beta x_3 = 3 \end{cases}$$

dipendente dai parametri α e β . Stabilire per quali valori dei parametri α e β i metodi iterativi di **Jacobi** e **Gauss-Seidel** convergono sicuramente per ogni scelta del vettore iniziale $\mathbf{X}^{(0)}$.

Condizione sufficiente perchè un metodo iterativo converga rispetto ad una norma di matrici $\|\cdot\|$ è che la norma della matrice di iterazione C sia strettamente minore di 1, cioè $\|C\| < 1$.

Sia $A = \begin{pmatrix} 10 & 1 & 1 \\ 1 & \alpha & 1 \\ 2 & 1 & \beta \end{pmatrix}$ la matrice dei coefficienti del sistema,

$$L = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 10 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix} \quad \text{e} \quad U = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

allora la matrice di iterazione del metodo di Jacobi C_J è data da

$$C_J = -D^{-1}(L + U),$$

cioè

$$C_J = \begin{pmatrix} 0 & -\frac{1}{10} & -\frac{1}{10} \\ -\frac{1}{\alpha} & 0 & -\frac{1}{\alpha} \\ -\frac{2}{\beta} & -\frac{1}{\beta} & 0 \end{pmatrix}$$

con $\alpha \neq 0$ e $\beta \neq 0$.

Considerando la $\|\cdot\|_\infty$ risulta

$$\|C_J\|_\infty = \max\left\{\frac{2}{10}, \frac{2}{|\alpha|}, \frac{3}{|\beta|}\right\} < 1 \quad \Leftrightarrow \quad \frac{2}{|\alpha|} < 1 \quad \wedge \quad \frac{3}{|\beta|} < 1,$$

cioè

$$|\alpha| > 2 \quad \wedge \quad |\beta| > 3.$$

Si osserva che alla stessa conclusione si giunge imponendo che la matrice A del sistema sia a **diagonale dominante per righe**.

Considerando, invece, la $\|\cdot\|_1$ risulta

$$\|C_J\|_1 = \max\left\{\frac{1}{|\alpha|} + \frac{2}{|\beta|}, \frac{1}{10} + \frac{1}{|\beta|}, \frac{1}{10} + \frac{1}{|\alpha|}\right\} < 1$$

$$\Leftrightarrow \begin{cases} \frac{1}{|\alpha|} + \frac{2}{|\beta|} < 1 \\ \frac{1}{10} + \frac{1}{|\beta|} < 1 \\ \frac{1}{10} + \frac{1}{|\alpha|} < 1 \end{cases}$$

$$\Leftrightarrow \begin{cases} \frac{|\beta|+2|\alpha|}{|\alpha||\beta|} < 1 \\ |\beta| > \frac{10}{9} \\ |\alpha| > \frac{10}{9} \end{cases}$$

da cui $\frac{10}{3} < |\beta| + 2|\alpha| < |\alpha||\beta|$.

E' opportuno notare che in questo caso non si arriva alla stessa conclusione imponendo la dominanza diagonale per colonne.

Inoltre, scegliendo $\alpha = \frac{9}{4}$ e $\beta = \frac{13}{4}$ risulta $\|C_J\|_\infty < 1$ mentre $\|C_J\|_1 > 1$. Infatti,

$$\frac{13}{4} + 2\frac{9}{4} = \frac{31}{4} = 7.75 > \frac{9}{4} \frac{13}{4} = \frac{117}{16} \approx 7.3125.$$

In questo caso, per esempio, sono necessarie almeno

$$K = 184$$

iterazioni per ottenere un errore $\|E^{(K)}\|_\infty$ tra due approssimazioni successive inferiore a $0.5 \cdot 10^{-5}$ avendo scelto $\mathbf{x}^{(0)} = [0 \ 0 \ 0]^T$ come approssimazione iniziale.

Infatti risulta

$$\|E^{(K)}\|_{\infty} < \varepsilon$$

quando

$$K > \log \left(\frac{(1 - \|C\|_{\infty}) \varepsilon}{\|X^{(1)} - X^{(0)}\|_{\infty}} \right) \frac{1}{\log \|C\|_{\infty}}$$

dove $\|C\|_{\infty} = \frac{12}{13}$, $\varepsilon = 0.5 \cdot 10^{-5}$,

$$\|X^{(1)} - X^{(0)}\|_{\infty} = \|X^{(1)}\|_{\infty} = \|Q_J\|_{\infty} = \frac{12}{13}$$

e $Q = D^{-1}B = \begin{bmatrix} \frac{b_1}{a_{11}} & \frac{b_2}{a_{22}} & \frac{b_3}{a_{33}} \end{bmatrix}^T = \begin{bmatrix} \frac{2}{10} & \frac{4}{9} & \frac{12}{13} \end{bmatrix}^T$,

con $B = [b_1 \quad b_2 \quad b_3]^T = [2 \quad 1 \quad 3]^T$ vettore dei termini noti del sistema.

Viceversa, se si pone $\alpha = 4$ e $\beta = 3$, il metodo di Jacobi non soddisfa la condizione sufficiente rispetto alla norma $\|\cdot\|_\infty$ mentre converge sicuramente rispetto alla norma $\|\cdot\|_1$. Inoltre

$$\|E^{(K)}\|_1 < 0.5 \cdot 10^{-5}$$

se

$$K > 166.03$$

avendo scelto $\mathbf{X}^{(0)} = [0 \quad 0 \quad 0]^T$

ed essendo $\|X^{(1)}\|_1 = \left\| \begin{bmatrix} \frac{2}{10} & \frac{1}{4} & \frac{1}{3} \end{bmatrix}^T \right\|_1 = \frac{2}{10} + \frac{1}{4} + \frac{1}{3} = \frac{47}{60}$

e $\|C\|_1 = \frac{11}{12}$

La matrice di iterazione del metodo di Gauss-Seidel è data da

$$C_{GS} = -(L + D)^{-1}U,$$

con

$$(L + D)^{-1} = \begin{pmatrix} \frac{1}{10} & 0 & 0 \\ -\frac{1}{10\alpha} & \frac{1}{\alpha} & 0 \\ \frac{1}{10\beta} \left(\frac{1}{\alpha} - 2\right) & -\frac{1}{\alpha\beta} & \frac{1}{\beta} \end{pmatrix}.$$

Allora

$$C_{GS} = \begin{pmatrix} 0 & -\frac{1}{10} & -\frac{1}{10} \\ 0 & \frac{1}{10\alpha} & -\frac{9}{10\alpha} \\ 0 & -\frac{1}{10\alpha\beta} + \frac{1}{5\beta} & \frac{9}{10\alpha\beta} + \frac{1}{5\beta} \end{pmatrix}.$$

$$\|C_{GS}\|_{\infty} = \max\left\{\frac{2}{10}, \frac{1}{|\alpha|}, \frac{|2\alpha - 1| + |9 + 2\alpha|}{10|\alpha||\beta|}\right\} < 1$$

se

$$\begin{cases} |\alpha| > 1 \\ |2\alpha - 1| + |9 + 2\alpha| < 10|\alpha||\beta| \end{cases}$$

Si osserva che $\alpha = \frac{9}{4}$ e $\beta = \frac{13}{4}$ soddisfano la condizione precedente.

Quindi il metodo di Gauss-Seidel sicuramente converge per ogni scelta dell'approssimazione iniziale.

Scegliendo $\mathbf{X}^{(0)} = [0 \ 0 \ 0]^T$, sono necessarie

$$K = 16$$

affinchè $\|E^{(K)}\|_{\infty} < \varepsilon = 0.5 \cdot 10^{-5}$.

Infatti, $\|C_{GS}\|_{\infty} = \frac{4}{9}$

mentre

$$\begin{aligned} \|\mathbf{X}^{(1)} - \mathbf{X}^{(0)}\|_{\infty} &= \|\mathbf{X}^{(1)}\|_{\infty} = \|\mathbf{Q}_{GS}\|_{\infty} = \\ &= \|(\mathbf{D} + \mathbf{L})^{-1}\mathbf{B}\|_{\infty} = \left\| \begin{bmatrix} 1 & 16 & 404 \\ 5 & 45 & 585 \end{bmatrix}^T \right\|_{\infty} = \frac{404}{585} \end{aligned}$$

dove

$$(L + D)^{-1} = \begin{pmatrix} \frac{1}{10} & 0 & 0 \\ -\frac{2}{45} & \frac{4}{9} & 0 \\ -\frac{28}{585} & -\frac{16}{117} & \frac{4}{13} \end{pmatrix}.$$

Si osserva che il **metodo di Gauss-Seidel converge molto più velocemente del metodo di Jacobi**

Ripetere per $\|C_{GS}\|_1$.

Esercizio

Dato il sistema

$$\begin{cases} x_1 + 4x_2 & = & 5 \\ 3x_1 + x_2 + x_3 & = & 2 \\ 2x_2 + 4x_3 & = & 20 \end{cases}$$

stabilire se è possibile risolverlo usando il **metodo di Jacobi** per ogni scelta dell'approssimazione iniziale.

Soluzione

La matrice dei coefficienti del sistema è $A = \begin{pmatrix} 1 & 4 & 0 \\ 3 & 1 & 1 \\ 0 & 2 & 4 \end{pmatrix}$.

Si osserva subito che A non è a diagonale dominante nè per righe nè per colonne.

La matrice di iterazione del metodo di Jacobi è

$$C_J = \begin{pmatrix} 0 & -4 & 0 \\ -3 & 0 & -1 \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}$$

per la quale risulta

$$\|C_J\|_\infty = \max\{4, 4, 1/2\} = 4 > 1$$

$$\|C_J\|_1 = \max\{3, 9/2, 1\} = 3 > 1$$

$$\rho(C_J) = \max_i |\lambda_i| = 5/\sqrt{2} > 1$$

infatti $\det(C_J - \lambda I) = -\lambda^3 + \lambda/2 + 12\lambda = \lambda(-\lambda^2 + 25/2) = 0 \Leftrightarrow \lambda_1 = 0, \lambda_{2,3} = \pm 5/\sqrt{2}$ e quindi $\max_i |\lambda_i| = 5/\sqrt{2}$.

Quindi il metodo di Jacobi **non converge**.

Tuttavia, **scambiando la prima e la seconda equazione del sistema**, si ottiene un sistema equivalente la cui matrice dei coefficienti è data da

$$\hat{A} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 4 & 0 \\ 0 & 2 & 4 \end{pmatrix}$$

alla quale è associata la seguente matrice di iterazione del metodo di Jacobi

$$\hat{C}_J = \begin{pmatrix} 0 & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}$$

per la quale risulta

$$\|\hat{C}_J\|_\infty = \max\left\{\frac{2}{3}, \frac{1}{4}, \frac{1}{2}\right\} = \frac{2}{3} < 1$$

$$\|\hat{C}_J\|_1 = \max\left\{\frac{1}{4}, \frac{5}{6}, \frac{1}{3}\right\} = \frac{5}{6} < 1$$

$$\rho(\hat{C}_J) = \max_i |\lambda_i| = 0.4257 < 1.$$

In questo caso il metodo di Jacobi converge per ogni scelta della approssimazione iniziale.

Inoltre, la **velocità di convergenza** del metodo è

$$-\text{Log}(\rho(\hat{C}_J)) = -\text{Log}(0.4257) = 0.8540$$

e l'errore di approssimazione si riduce di un fattore **10^{-m}** all'iterazione

$$K \approx -\frac{m}{\text{Log}(\rho(\hat{C}_J))} = \frac{m}{0.8540}.$$

Metodo di rilassamento (S.O.R.)

Il metodo di Gauss-Seidel può essere generalizzato introducendo un *parametro di rilassamento* $\omega \neq 0$ nella decomposizione della matrice A :

$$A = L + \frac{D}{\omega} - \frac{D}{\omega} + D + U \quad \omega \in \mathbb{R}^+$$

e ponendo

$$M = \frac{D}{\omega} + L$$

e

$$N = U - \frac{D}{\omega}(1 - \omega)$$

da cui

$$\begin{cases} C_\omega = -\left(\frac{D}{\omega} + L\right)^{-1} \left(U - \frac{D}{\omega}(1 - \omega)\right) \\ Q_\omega = \left(\frac{D}{\omega} + L\right)^{-1} B \end{cases}$$

e quindi

$$\begin{aligned} X^{(k)} &= C_\omega X^{(k-1)} + Q_\omega = \\ &= (D + \omega L)^{-1} (-\omega U + (1 - \omega)D) X^{(k)} + \omega (D + \omega L)^{-1} B \end{aligned}$$



Algoritmo S.O.R. (successive over relaxation)

$$v_i^{(k)} = \frac{1}{a_{ii}} \left(- \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} + b_i \right)$$

$$x_i^{(k+1)} = \omega v_i^{(k)} + (1 - \omega) x_i^{(k)}$$

$$i = 1, 2, \dots, n; \quad k \geq 0$$

$v_i^{(k)}$ è la soluzione del metodo di Gauss-Seidel mentre

$x_i^{(k+1)}$ è la somma pesata di $v_i^{(k)}$ e $x_i^{(k)}$ usando i pesi ω e $1 - \omega$,

con $x_i^{(k)}$ *i-esima* componente della soluzione calcolata al passo precedente

Se $\omega = 1$, il metodo S.O.R. coincide con il metodo di Gauss-Seidel.

Convergenza del metodo di sovrarilassamento

Teorema. Se A è (simmetrica) **definita positiva**, la condizione $\omega \in (0, 2)$ è **necessaria** e **sufficiente** per la convergenza.

Teorema. Se A è (simmetrica) **definita positiva** e **tridiagonale** allora la **velocità di convergenza** è **massima** per

$$\omega = \omega_0 = \frac{2}{1 + \sqrt{1 - \rho(C_{GS})}}$$

La soluzione esatta è $p_i = 1 - \frac{i}{N}$, $i = 0, \dots, N$.

Per il **metodo di Jacobi** si ha:

N	$\ C_J\ _1$	$\ C_J\ _\infty$	$\rho(C_J)$	$\ X^{(50)} - X^{(49)}\ _\infty$	$\ E^{(50)}\ _\infty$	decimali esatti
11	1	1	0.9595	$0.64 \cdot 10^{-2}$	$0.77 \cdot 10^{-1}$	0
21	1	1	0.9888	$0.95 \cdot 10^{-2}$	0.36	0
51	1	1	0.9981	$0.95 \cdot 10^{-2}$	0.68	0
101	1	1	0.9995	$0.95 \cdot 10^{-2}$	0.81	0

Per il **metodo di Gauss-Seidel** si ha:

N	$\ C_{GS}\ _1$	$\ C_{GS}\ _\infty$	$\rho(C_{GS})$	$\ X^{(50)} - X^{(49)}\ _\infty$	$\ E^{(50)}\ _\infty$	decimali esatti
11	0.9990	0.9980	0.9206	$0.69 \cdot 10^{-3}$	$0.80 \cdot 10^{-2}$	1
21	1	1	0.9778	$0.42 \cdot 10^{-2}$	0.18	0
51	1	1	0.9962	$0.44 \cdot 10^{-2}$	0.55	0
101	1	1	0.9990	$0.44 \cdot 10^{-2}$	0.73	0

Per il **metodo S.O.R.** si ha:

N	ω	$\ C_\omega\ _1$	$\ C_\omega\ _\infty$	$\rho(C_\omega)$	$\ X^{(50)} - X^{(49)}\ _\infty$	$\ E^{(50)}\ _\infty$	decimali esatti
11	0.5	0.9999	0.9999	0.9733	$0.41 \cdot 10^{-2}$	0.15	0
11	1.5604	2.34	1.34	0.56	$1.02 \cdot 10^{-12}$	$1.24 \cdot 10^{-12}$	11
11	1.8	5.21	1.7	0.8	$8.2 \cdot 10^{-6}$	$6.29 \cdot 10^{-6}$	5
21	0.5	0.9999	0.9999	0.9926	$0.4 \cdot 10^{-2}$	0.43	0
21	1.7406	5.35	1.61	0.7400	$0.32 \cdot 10^{-6}$	$0.89 \cdot 10^{-6}$	6
21	1.8	7.02	1.7	0.8	$2.65 \cdot 10^{-6}$	$6.29 \cdot 10^{-6}$	5
51	0.5	1	1	0.9987	$0.46 \cdot 10^{-2}$	0.71	0
51	1.8849	14.47	1.82	0.8840	$0.22 \cdot 10^{-2}$	$0.41 \cdot 10^{-2}$	2
51	1.8	7.95	1.7	0.9634	$0.26 \cdot 10^{-2}$	$0.69 \cdot 10^{-1}$	1
101	0.5	1	1	0.9997	$0.45 \cdot 10^{-2}$	0.84	0
101	1.9397	29.70	1.90	0.9396	$0.69 \cdot 10^{-2}$	0.05	1
101	1.8	8	1.70	0.9912	$0.37 \cdot 10^{-2}$	0.34	0

Si può verificare che dopo **15** iterazioni, scegliendo $\omega = \omega_{ott}$, il numero di decimali esatti della soluzione è **3** nel caso $N = 11$.

Esercizi d'esame

ESERCIZIO 1

Dato il sistema lineare

$$GY = F \quad G = \begin{bmatrix} 3 & \alpha & 0 \\ 1 & 2 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \beta \end{bmatrix} \quad \alpha, \beta \in \mathbf{R} \quad F = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$$

- 1.1) individuare per quali valori dei parametri α e β la matrice G è definita positiva;
- 1.2) posto $\alpha = 1$ e $\beta = 1/4$ stabilire se i metodi di Jacobi e di Gauss-Seidel sono adatti ad approssimare la soluzione del sistema lineare;
- 1.3) in caso di convergenza, specificare per ciascun metodo la scelta dell'approssimazione iniziale.

Soluzione ESERCIZIO 1

1.1) La matrice G è definita positiva se è simmetrica e se $X^T G X \geq 0$ per ogni $X \in \mathbf{R}^3$. Per la simmetria deve essere $\alpha = 1$. Per verificare se G è definita positiva si può utilizzare il criterio di Sylvester:

$$\det G_1 = 3 \quad \det G_2 = 5 \quad \det G_3 = \det G = 5\beta - \frac{3}{4}$$

I primi due determinanti principali di testa sono positivi. Il determinante di G è positivo se e solo se

$$5\beta - \frac{3}{4} > 0 \quad \Rightarrow \quad \beta > \frac{3}{20} = 0.15$$

1.2) Per $\alpha = 1$ e $\beta = 1/4$ la matrice

$$G = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

è definita positiva, quindi il metodo di Gauss-Seidel converge sicuramente. Poiché la matrice G non è diagonalmente dominante né per righe né per colonne, non si può trarre nessuna conclusione sulla convergenza del metodo di Jacobi analizzando le proprietà della matrice dei coefficienti.

Per quanto riguarda la matrice di iterazione del metodo, data da

$$C_J = -D^{-1}(L + U) = \begin{bmatrix} 0 & -\frac{1}{3} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 2 & 0 \end{bmatrix},$$

sia la norma 1 che la norma uniforme sono maggiori di 1. Per verificare la convergenza bisogna quindi calcolare il raggio spettrale, cioè il massimo dei moduli degli autovalori di C_j :

$$\rho(C_J) = \max\left(0, \sqrt{\frac{2}{3}}\right) = \sqrt{\frac{2}{3}} \approx 0.8165 < 1,$$

quindi anche il metodo di Jacobi è convergente.

1.3) La convergenza dei metodi è indipendente dalla scelta dell' approssimazione iniziale $X^{(0)} \in \mathbf{R}^3$. Una buona scelta può essere $X^{(0)} = 0$ per il metodo di Gauss-Seidel, del quale non è stata calcolata la matrice di iterazione, e $X^{(0)} = D^{-1}F = Q_J$ per il metodo di Jacobi.

Si suggerisce di fare qualche iterazione di entrambi i metodi per confrontare le differenti velocità di convergenza.

ESERCIZIO 2

Sia dato un sistema lineare con matrice

$$\mathbf{A} = \begin{bmatrix} 2 & \alpha & 2 & 1 \\ -\alpha & 5 & 2 & \\ 1 & 1 & \alpha & \end{bmatrix}$$

dipendente dal parametro reale α , e vettore dei termini noti

$$\mathbf{b} = [5 \quad -2 \quad 1]^T:$$

- 2.1) stabilire per quali valori del parametro α il metodo di Jacobi converge per ogni scelta della approssimazione iniziale $\mathbf{X}^{(0)}$;
- 2.2) scelti $\alpha = \frac{5}{2}$ e $\mathbf{X}^{(0)} = [0 \quad 0 \quad 0]^T$, produrre una stima superiore del numero di iterazioni necessarie affinché il metodo fornisca una soluzione con almeno 5 decimali esatti.

Soluzione Esercizio 2

La matrice di iterazione del metodo di Jacobi è

$$C = \begin{bmatrix} 0 & -\frac{1}{\alpha} & -\frac{1}{2\alpha} \\ \frac{\alpha}{5} & 0 & -\frac{2}{5} \\ -\frac{1}{\alpha} & -\frac{1}{\alpha} & 0 \end{bmatrix}$$

$\forall \alpha \neq 0$, mentre il termine noto è $Q = [\frac{5}{2\alpha} \quad -\frac{2}{5} \quad \frac{1}{\alpha}]^T$.

Affinchè il metodo di Jacobi converga per ogni scelta del punto iniziale, è necessario che $\|C\|_{\infty} < 1$.

$$\|C\|_{\infty} = \max\left\{\frac{1}{|\alpha|} + \frac{1}{2|\alpha|}, \frac{|\alpha| + 2}{5}, \frac{2}{|\alpha|}\right\} = \max\left\{\frac{3}{2|\alpha|}, \frac{|\alpha| + 2}{5}, \frac{2}{|\alpha|}\right\}.$$

Poichè $\frac{3}{2|\alpha|} < \frac{2}{|\alpha|}$, $\forall \alpha \neq 0$,

$$\|C\|_{\infty} = \max\left\{\frac{|\alpha| + 2}{5}, \frac{2}{|\alpha|}\right\}.$$

Quindi,

$$\|C\|_{\infty} = \begin{cases} \frac{2}{|\alpha|} & \text{se } \frac{|\alpha|+2}{5} < \frac{2}{|\alpha|} \\ \frac{|\alpha|+2}{5} & \text{viceversa} \end{cases}$$

cioè

$$\|C\|_{\infty} = \begin{cases} \frac{2}{|\alpha|} & \text{se } |\alpha|^2 + 2|\alpha| - 10 < 0 \\ \frac{|\alpha|+2}{5} & \text{viceversa} \end{cases}$$

da cui

$$\|C\|_{\infty} = \begin{cases} \frac{2}{|\alpha|} & \text{se } |\alpha| < \sqrt{11} - 1 \\ \frac{|\alpha|+2}{5} & \text{se } |\alpha| \geq \sqrt{11} - 1 \end{cases}$$

Imponendo $\|C\|_\infty < 1$ risulta

$$\begin{cases} |\alpha| > 2 & \text{se } |\alpha| < \sqrt{11} - 1 \\ |\alpha| < 3 & \text{se } |\alpha| \geq \sqrt{11} - 1 \end{cases}$$

da cui si deduce che il metodo di Jacobi converge in norma infinito se $2 < |\alpha| < 3$.

Il numero di iterazioni k necessarie affinché la soluzione abbia la precisione ϵ è maggiorabile come segue

$$k > \log \left(\frac{(1 - \|C\|_\infty)\epsilon}{\|\mathbf{X}^{(1)} - \mathbf{X}^{(0)}\|_\infty} \right) \frac{1}{\log(\|C\|_\infty)}.$$

In questo caso, $\epsilon = 0.5 \cdot 10^{-5}$, $\|C\|_\infty = \frac{5+2}{5} = \frac{9}{10}$ mentre

$$\mathbf{X}^{(1)} = \mathbf{C}\mathbf{X}^{(0)} + \mathbf{Q} = \mathbf{Q} = \left[1 \quad -\frac{2}{5} \quad \frac{2}{5}\right]^T$$

e quindi $\|\mathbf{X}^{(1)} - \mathbf{X}^{(0)}\|_\infty = \|\mathbf{X}^{(1)}\|_\infty = 1$. Ne risulta che

$$k > \frac{\log((1 - 9/10) \cdot 0.5 \cdot 10^{-5})}{\log(9/10)} = 137.70.$$

Quindi, il numero minimo di iterazioni per raggiungere la precisione richiesta è $k = 138$.

ESERCIZIO 3

Si consideri il sistema lineare $\mathbf{A}_n \mathbf{x} = \mathbf{b}_n$ dipendente da $n \in \mathbb{N}$, con

$$\mathbf{A}_n = \begin{bmatrix} \frac{n}{2} & \frac{n}{2} + 1 \\ \frac{n}{2} + 1 & \frac{n}{2} + 2 \end{bmatrix} \quad \mathbf{b}_n = \begin{bmatrix} b_{1,n} \\ b_{2,n} \end{bmatrix}$$

- 3.1) Valutare il numero di condizionamento di \mathbf{A}_n al variare di n ;
- 3.2) stabilire per quali n il numero di condizionamento risulta inferiore a 50;
- 3.3) sia \bar{n} il valore più grande di n verificante la condizione al punto (2.2). Dare una maggiorazione in norma ∞ dell'errore relativo su

\mathbf{x} nel caso in cui la soluzione del sistema sia $\mathbf{x} = [1, 1]^T$ e il vettore $\mathbf{b}_{\bar{n}}$ subisca una perturbazione pari a

$$\delta \mathbf{b}_{\bar{n}} = \begin{bmatrix} 5 \cdot 10^{-2} \\ 10^{-3} \end{bmatrix}.$$

Soluzione ESERCIZIO 3

3.1) La matrice \mathbf{A}_n è simmetrica e tutti i suoi elementi sono non negativi. Ne segue che $\|\mathbf{A}_n\|_1 = \|\mathbf{A}_n\|_\infty = \max\{n+1, n+3\} = n+3$.

Inoltre, risulta

$$\mathbf{A}_n^{-1} = \frac{1}{\det(\mathbf{A}_n)} \begin{bmatrix} \frac{n}{2} + 2 & -\frac{n}{2} - 1 \\ -\frac{n}{2} - 1 & \frac{n}{2} \end{bmatrix}$$

e quindi

$$\|\mathbf{A}_n^{-1}\|_1 = \|\mathbf{A}_n^{-1}\|_\infty$$

in quanto si verifica facilmente che $\det(\mathbf{A}_n) = -1$, $\forall n$. Ne segue che il numero di condizionamento di \mathbf{A}_n risulta

$$K(\mathbf{A}_n) = (n+3)^2,$$

rispetto alle norme $\|\cdot\|_1$ e $\|\cdot\|_\infty$, quindi il sistema diventa mal condizionato anche per piccoli valori di n .

3.2)

$$K(\mathbf{A}_n) = (n + 3)^2 < 50 \Leftrightarrow n < \sqrt{50} - 3.$$

Poichè $n \in \mathbf{N}$, risulta $n \leq 4$.

3.3) Si osserva che $\bar{n} = 4$. Poichè $\mathbf{x} = [1 \quad 1]^T$, risulta

$$\mathbf{b}_{\bar{n}} = [\bar{n} + 1 \quad \bar{n} + 3]^T.$$

Quindi, $\frac{\|\delta \mathbf{b}_{\bar{n}}\|_{\infty}}{\|\mathbf{b}_{\bar{n}}\|_{\infty}} = \frac{5 \cdot 10^{-2}}{\bar{n} + 3}$ e

$$\frac{\|\delta \mathbf{x}\|_{\infty}}{\|\mathbf{x}\|_{\infty}} \leq K(\mathbf{A}_{\bar{n}}) \frac{\|\delta \mathbf{b}_{\bar{n}}\|_{\infty}}{\|\mathbf{b}_{\bar{n}}\|_{\infty}} = (\bar{n} + 3)^2 \frac{5 \cdot 10^{-2}}{\bar{n} + 3} = 5 \cdot 10^{-2} (\bar{n} + 3) = 3.5 \cdot 10^{-1}.$$

ESERCIZIO 4

Si consideri la matrice

$$\mathbf{A} = \begin{pmatrix} -1 & \frac{7}{2} & -1 \\ 0 & \frac{1}{\gamma} & \frac{7}{2} \\ \frac{7}{2} & \gamma & \delta \end{pmatrix} \quad \gamma, \delta \in \mathbf{R}.$$

- 4.1) Dopo una opportuna permutazione delle righe, individuare i valori dei parametri γ e δ per i quali la matrice \mathbf{A} risulti definita positiva;
- 4.2) Sia \mathbf{B} la matrice definita positiva costruita al punto precedente. Determinare per quali valori dei parametri il numero di condizionamento della matrice \mathbf{B} risulta minore di 7.

Soluzione ESERCIZIO 4

4.1) Con una permutazione delle righe della matrice \mathbf{A} si ottiene la matrice

$$\mathbf{B} = \begin{pmatrix} \frac{7}{2} & \gamma & \delta \\ -1 & \frac{7}{2} & -1 \\ 0 & \frac{1}{\gamma} & \frac{7}{2} \end{pmatrix}$$

che risulta simmetrica per $\gamma = -1$ e $\delta = 0$. Si verifica facilmente che per questi valori dei parametri la matrice \mathbf{B} è definita positiva.

4.2) Poichè la matrice \mathbf{B} è simmetrica, il numero di condizionamento in norma 1 o infinito sono coincidenti:

$$K(\mathbf{B}) = \|\mathbf{B}\| \cdot \|\mathbf{B}^{-1}\| \approx 2.95.$$

Si suggerisce di calcolare il numero di condizionamento in norma 2.

ESERCIZIO 5 Dato il sistema $\mathbf{Ax} = \mathbf{b}$ con

$$\mathbf{A} = \begin{pmatrix} \alpha & 4 & 1/2 \\ 0 & 2\alpha & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

e $\alpha \neq 0$,

5.1) stabilire per quali valori del parametro α i metodi iterativi di Jacobi e Gauss-Seidel convergono alla soluzione del sistema per qualsiasi scelta della approssimazione iniziale;

5.2) posto $\alpha = 8$, in caso di convergenza dei metodi, stimare il numero minimo di iterazioni necessarie affinché l'errore relativo soddisfi

$$\frac{\|\mathbf{x} - \mathbf{x}_k\|}{\|\mathbf{x}\|} < 10^{-8}$$

dove \mathbf{x}_k è l'approssimazione prodotta alla k -esima iterazione.

ESERCIZIO 6

Dato il sistema lineare

$$H A = Y \quad H = \begin{bmatrix} \theta & 1 & 0 \\ 1 & 2 & -\frac{1}{2} \\ 0 & \gamma & 3 \end{bmatrix} \quad \gamma, \theta \in \mathbf{R} \quad Y = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

- 6.1) individuare per i quali valori dei parametri γ e θ la matrice H è definita positiva;
- 6.2) posto $\gamma = -1/2$ e $\theta = 2/3$ stabilire se i metodi di Jacobi e di Gauss-Seidel sono adatti ad approssimare la soluzione del sistema lineare;
- 6.3) in caso di convergenza, specificare per ciascun metodo la scelta dell'approssimazione iniziale.

ESERCIZIO 7

- 7.1) Illustrare in dettaglio un metodo di stima dell'errore commesso sulla soluzione di un sistema lineare in presenza di errori sui dati. Dare un esempio di matrice mal condizionata e uno di matrice ben condizionata.
- 7.2) Si consideri il sistema lineare di n equazioni in n incognite $\mathbf{AX} = \mathbf{B}$, con matrice dei coefficienti

$$\mathbf{A} = \begin{pmatrix} 3 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & \dots & 0 & -1 & 3 \end{pmatrix}$$

e vettore dei termini noti $\mathbf{B} = (1, 0, \dots, 0)^T$.

- Stabilire se esistono valori di n per cui il metodo di Jacobi sicuramente converge alla soluzione del sistema perturbato $(\mathbf{A} + \delta\mathbf{A})(\mathbf{X} + \delta\mathbf{X}) = \mathbf{B}$, con $\delta\mathbf{A} = \text{diag}(-\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}, -\frac{1}{n})$;
- scelto n pari all'intero per cui $\|\delta\mathbf{A}\|_\infty = \frac{1}{2}$, dare una maggioranza dell'errore relativo $\frac{\|\delta\mathbf{X}\|_\infty}{\|\mathbf{X}\|_\infty}$.

ESERCIZIO 8

- 8.1 Illustrare i teoremi di convergenza dei metodi iterativi per la soluzione di sistemi lineari e descrivere un metodo di stima a priori del numero di iterazioni necessarie affinché l'approssimazione abbia un numero stabilito di decimali esatti.
- 8.2 Discutere la convergenza dei metodi di Jacobi e Gauss Seidel del seguente sistema lineare

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

- 8.3 Detto $\bar{A}\mathbf{X} = \bar{\mathbf{b}}$ un sistema lineare equivalente al precedente per il quale il metodo di Jacobi converge per ogni scelta dell'approssimazione iniziale, stabilire se esistono valori del parametro $\beta \in \mathbf{R}$ per cui il metodo iterativo

$$\mathbf{X}_{k+1} = \mathbf{X}_k - \beta(\bar{A}\mathbf{X}_k - \bar{\mathbf{b}})$$

converge più velocemente del metodo di Jacobi.

Soluzione

Siano \mathbf{A} la matrice del sistema, \mathbf{x} il vettore della soluzione e \mathbf{b} il vettore dei termini noti. La matrice \mathbf{A} non soddisfa le condizioni sufficienti di convergenza dei metodi di Jacobi e di Gauss Seidel. In particolare, \mathbf{A} non è a diagonale dominante per righe o per colonne (una condizione sufficiente per la convergenza dei due metodi iterativi), e neanche simmetrica e definita positiva (condizione sufficiente per la convergenza del metodo di Gauss Seidel). Tuttavia, scambiando la seconda e la terza riga della matrice \mathbf{A} , si ottiene la matrice

$$\bar{\mathbf{A}} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

che risulta a diagonale dominante per righe e per colonne. Inoltre è simmetrica e definita positiva; infatti il determinante dei minori di testa di $\bar{\mathbf{A}}$ è positivo:

$$\det(\bar{\mathbf{A}}_1) = 2 > 0; \quad \det(\bar{\mathbf{A}}_2) = 6 - 1 = 5 > 0; \quad \det(\bar{\mathbf{A}}) = 12 - 4 = 8 > 0.$$

Ne segue che i metodi di Jacobi e di Gauss-Seidel convergono alla soluzione del sistema $\bar{\mathbf{A}}\mathbf{X} = \bar{\mathbf{b}}$, con

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{e} \quad \bar{\mathbf{b}} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

In questo caso, la matrice di iterazione del metodo di Jacobi è

$$\mathbf{C}_J = \begin{pmatrix} 0 & -1/2 & 0 \\ -1/3 & 0 & -1/3 \\ 0 & -1/2 & 0 \end{pmatrix},$$

il cui raggio spettrale $\rho(\mathbf{C}_J) = \max \{|\lambda_i|\}_{1 \leq i \leq 3} = 1/\sqrt{3}$. Infatti, gli autovalori λ_i di \mathbf{C}_J sono le soluzioni dell'equazione caratteristica $-\lambda^3 + \lambda/3 = 0$.

La matrice di iterazione del metodo iterativo

$$\mathbf{X}_{k+1} = \mathbf{X}_k - \beta(\bar{\mathbf{A}}\mathbf{X}_k - \bar{\mathbf{b}})$$

è $\mathbf{C}(\beta) = \mathbf{I} - \beta\bar{\mathbf{A}}$, con \mathbf{I} matrice identità. Affinchè il metodo converga più velocemente del metodo di Jacobi, deve esistere almeno un β per cui

$$\rho(\mathbf{C}(\beta)) < \rho(\mathbf{C}_J), \quad (*)$$

Gli autovalori della matrice $\mathbf{C}(\beta)$ sono le soluzioni dell'equazione caratteristica

$$(1 - 2\beta - \lambda)^2(1 - 3\beta - \lambda) - 2\beta^2(1 - 2\beta - \lambda) = 0,$$

ovvero

$$(1 - 2\beta - \lambda)(\lambda^2 - \lambda(2 - 5\beta) + 4\beta^2 - 5\beta + 1) = 0$$

cioè $\lambda_1 = 1 - 2\beta$, $\lambda_{2,3} = \frac{2-5\beta \pm 3\beta}{2}$ e quindi $\lambda_2 = 1 - \beta$, $\lambda_3 = 1 - 4\beta$. Ne segue che $\rho(\mathbf{C}(\beta)) < 1 \Leftrightarrow 0 < \beta < 1/2$ e

$$\rho(\mathbf{C}(\beta)) = \begin{cases} 1 - \beta & \beta \leq 2/5 \\ 4\beta - 1 & \beta > 2/5 \end{cases}$$

Poichè

$$1 - \beta < 1/\sqrt{3} \Leftrightarrow \beta > 1 - 1/\sqrt{3} \approx 0.4226 > 2/5$$

e

$$4\beta - 1 < 1/\sqrt{3} \Leftrightarrow \beta < (1 + 1/\sqrt{3})/4 \approx 0.3943 < 2/5,$$

non esistono valori di β per cui il metodo iterativo considerato converge più velocemente del metodo di Jacobi.

Oss: Alla stessa conclusione si giunge ragionando come segue.

Poichè per una generica matrice \mathbf{M} risulta $\rho(\mathbf{M}) \leq \|\mathbf{M}\|$, con $\|\cdot\|$ norma di matrice compatibile, la (*) è sicuramente verificata per i valori di β tali che

$$\|\mathbf{C}(\beta)\|_{\infty} < \rho(\mathbf{C}_J),$$

cioè

$$\max |1 - 2\beta| + |\beta|, |1 - 3\beta| + 2|\beta| < 1/\sqrt{3}$$

ovvero

$$|1 - 3\beta| + 2|\beta| < 1/\sqrt{3}$$

che non è mai soddisfatta.

Tuttavia, in questo caso non si può concludere che non esiste alcun valore di β per cui il metodo iterativo considerato converge più velocemente del metodo di Jacobi.

ESERCIZIO 9

9.1 Illustrare il problema della soluzione di sistemi di equazioni lineari con particolare riferimento ai metodi diretti.

9.2 Si considerino i seguenti sistemi lineari:

$$\left\{ \begin{array}{l} \varepsilon x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{array} \right. , \quad \left\{ \begin{array}{l} x_1 + 3x_2 + 5x_3 + 3x_4 = -1 \\ 2x_2 + 4x_3 + 3x_4 = -1 \\ 2x_3 + x_4 = 1 \\ 2x_4 = 2 \end{array} \right. , \quad \left\{ \begin{array}{l} 3x_1 - x_2 + 2x_3 = 1 \\ 9x_1 - x_2 - 4x_3 = 0 \\ -6x_1 + 6x_3 = 1 \end{array} \right. .$$

Stabilire per quali di essi il metodo di eliminazione di Gauss senza pivoting fornisce la soluzione esatta. Motivare la risposta.

ESERCIZIO 10

10.1 Illustrare il problema della soluzione di sistemi di equazioni lineari con particolare riferimento ai metodi diretti.

10.2 Dati i seguenti sistemi lineari dipendenti dal parametro $k \in \mathbf{R}$:

$$\begin{cases} x_1 - kx_2 = 1 - k \\ -kx_1 + x_2 - kx_3 = 1 - 2k \\ -x_2 + x_3 = 1 - k \end{cases}, \quad \begin{cases} -x_1 + x_2 = -1 \\ 0.5x_1 + 2x_2 - 0.5x_3 = 2 \\ \sqrt{2}x_2 + kx_3 = 1 \end{cases}, \quad \begin{cases} kx_1 - x_2 + kx_3 = 1 \\ -x_1 + kx_2 - x_3 = 0 \\ kx_1 - x_2 = 1 \end{cases},$$

stabilire per quali valori di k è possibile calcolarne la soluzione in modo esatto usando il metodo di Thomas.

ESERCIZIO 11

11.1 Illustrare dettagliatamente i metodi iterativi per la soluzione di sistemi lineari.

11.2 Stabilire se i seguenti metodi iterativi

$$\mathbf{X}^{(k+1)} = (\mathbf{I} - \alpha\mathbf{A})\mathbf{X}^{(k)} + \beta\mathbf{B}, \quad \alpha, \beta \in \mathbf{R}$$

$$\mathbf{X}^{(k+1)} = \beta((1 - \alpha)\mathbf{I} - \mathbf{A})\mathbf{X}^{(k)} + \beta\mathbf{B}, \quad \alpha, \beta \in \mathbf{R}$$

$$\mathbf{X}^{(k+1)} = \mathbf{D}^{-1}(\mathbf{D} + \mathbf{A})\mathbf{X}^{(k)} + \mathbf{B}, \quad \text{con } \mathbf{D} \text{ matrice diagonale: } D_{ii} = A_{ii}, \forall i$$

possono convergere alla soluzione del sistema lineare $\mathbf{AX} = \mathbf{B}$ per ogni scelta dell'approssimazione iniziale \mathbf{X}^0 . Motivare la risposta.

ESERCIZIO 12 Si consideri il seguente procedimento iterativo dipendente dal parametro reale ω :

$$\mathbf{X}^{(k+1)} = \mathbf{D}(\omega)\mathbf{X}^{(k)} + \mathbf{F}(\omega),$$

$$\text{con } \mathbf{D}(\omega) = \frac{1}{4} \begin{pmatrix} 2\omega^2 + 2\omega + 1 & -2\omega^2 + 2\omega + 1 \\ -2\omega^2 + 2\omega + 1 & 2\omega^2 + 2\omega + 1 \end{pmatrix} \text{ e } \mathbf{F}(\omega) = \begin{pmatrix} \frac{1}{2} - \omega \\ \frac{1}{2} - \omega \end{pmatrix}.$$

12.1 Stabilire se esiste almeno un valore di ω per cui il procedimento iterativo risulta adatto ad approssimare la soluzione del sistema lineare $\mathbf{AX} = \mathbf{B}$ con

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \text{ e } \mathbf{B} = \begin{pmatrix} 3 \\ 5 \end{pmatrix};$$

12.2 in caso di risposta affermativa, determinare il valore di ω per cui la velocità di convergenza al vettore \mathbf{X} è massima;

12.3 stabilire se il procedimento iterativo è adatto ad approssimare la soluzione del sistema perturbato $\mathbf{AX} = \bar{\mathbf{B}}$ con $\bar{\mathbf{B}} = \begin{pmatrix} 3.1 \\ 5.1 \end{pmatrix}$. Motivare la risposta.

Riferimenti bibliografici

L. Gori, *Calcolo Numerico*:

Cap. 2 §§ 2.1-2.5, 2.8-2.11

Cap. 4 §§ 4.1-4.6, 4.8, 4.9 (escluso il pivoting totale) 4.10 (solo enunciati dei teoremi), 4.12

L. Gori, M.L. Lo Cascio, F. Pitolli, *Esercizi di Calcolo Numerico*:

2.1-2.5, 2.10-2.13, 2.17, 2.19 - 2.26, 2.30, 7.7, 7.15, 7.16, 7.19, 7.35, 7.41, 7.49, 7.52, 7.57, 7.58, 7.64