Calcolo Numerico con elementi di programmazione

(A.A. 2014-2015)

Appunti delle lezioni sui metodi numerici
per la soluzione di sistemi lineari

Sistemi Lineari

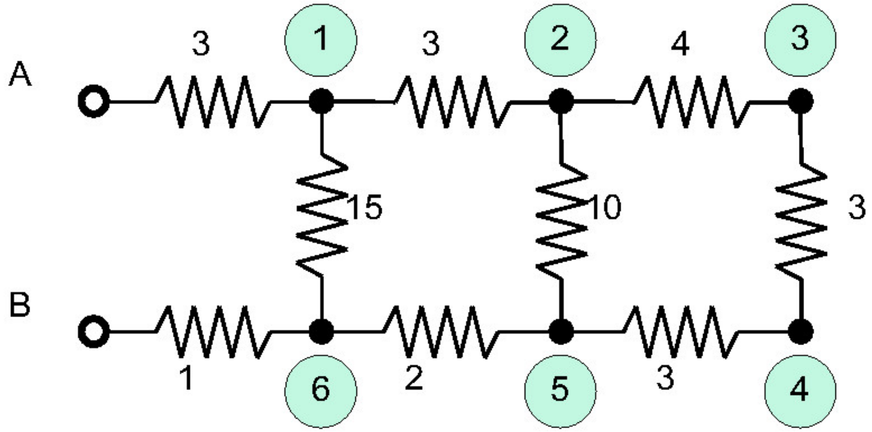
I sistemi lineari forniscono il **modello matematico** per la descrizione di numerosi fenomeni fisici. Si incontrano sistemi lineari nella:

- discretizzazione di problemi ai limiti per equazioni differenziali;
- approssimazione di funzioni o ricostruzione di dati sperimentali;
- trattazione numerica di reti;
- svariate situazioni in cui l'andamento del fenomeno è lineare (o linearizzato).

La grande varietà dei contesti, e quindi delle strutture dei sistemi lineari, ha determinato anche una notevole varietà di algoritmi. Le principali caratteristiche che si guarderanno nei vari metodi numerici sono: **stabilità** , **velocità di convergenza**, **costo computazionale**.

Esempio 1

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

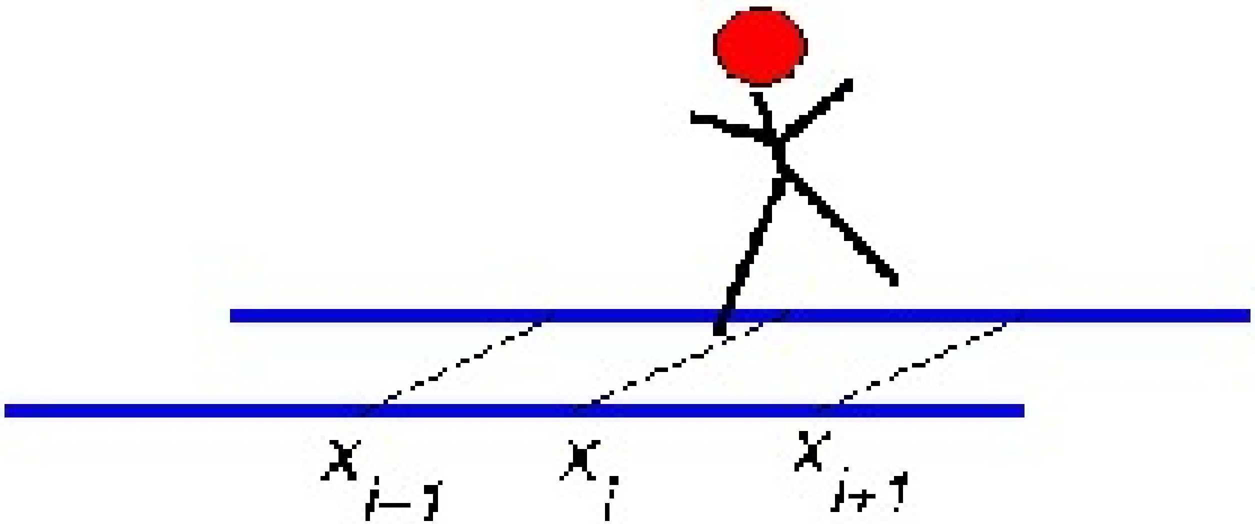
Applicando la **legge di Ohm** $\Delta V = RI$ e la **legge di Kirchoff** $\sum_i I_i = 0$ in ogni nodo si ottiene il **sistema lineare**

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

Esempio 2

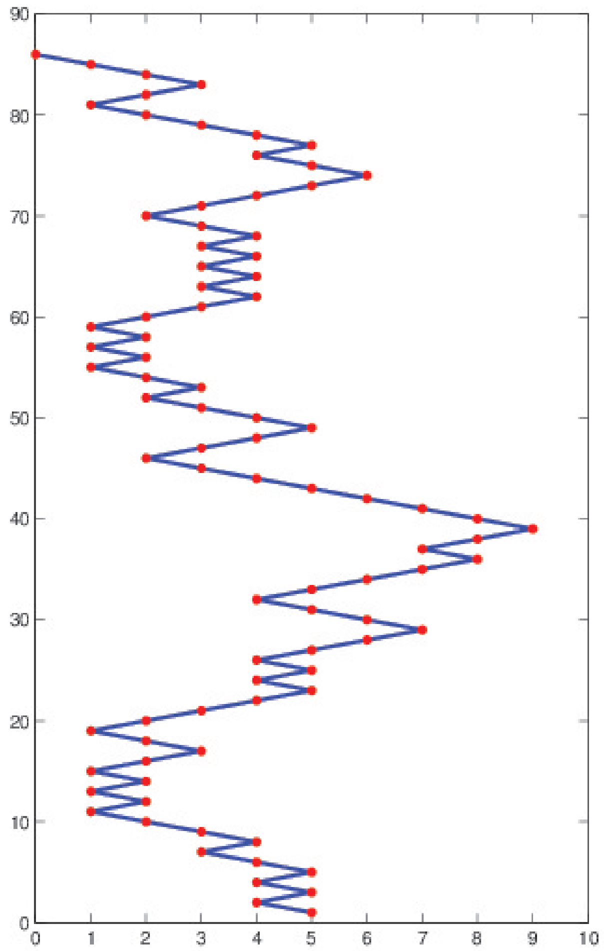
Un ubriaco compie una **passeggiata casuale**, facendo un passo a sinistra o a destra a caso lungo una strada rettilinea. Quando raggiunge una estremità della strada, si ferma.

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

Soluzione

La **probabilità** p_i, $i = 0, 1, ..., N$, di raggiungere l'estremo sinistro partendo dalla posizione i, soddisfa la relazione

$$\begin{cases} p_0 = 1 \quad p_N = 0 \\ \\ p_i = \frac{1}{2}p_{i-1} + \frac{1}{2}p_{i+1} \qquad i = 1, \ldots, N-1 \end{cases}$$

Si tratta di un **sistema lineare tridiagonale** nelle incognite

$$p_i,\ i = 1, \ldots, N-1$$

$$\begin{cases} \begin{array}{cccccccc} p_1 & -\frac{1}{2}p_2 & & & & & = & \frac{1}{2} \\ -\frac{1}{2}p_1 & +p_2 & -\frac{1}{2}p_3 & & & & = & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \\ & & & -\frac{1}{2}p_{N-3} & +p_{N-2} & -\frac{1}{2}p_{N-1} & = & 0 \\ & & & & -\frac{1}{2}p_{N-2} & +p_{N-1} & = & 0 \end{array} \end{cases}$$

Esempio 3

Punto di intersezione tra due rette

$$\begin{cases} 2x + y = 3 \\ \\ 4x + 2y = 6 \end{cases} \qquad \begin{cases} 2x + y = 3 \\ \\ 4x + 2y = 0 \end{cases} \qquad \begin{cases} 2x + y = 3 \\ \\ x + 2y = 3 \end{cases}$$

rette coincidenti **rette parallele** **rette incidenti**

infinite soluzioni **nessuna soluzione** **unica soluzione**

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} b = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \quad A = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} b = \begin{pmatrix} 3 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} b = \begin{pmatrix} 3 \\ 3 \end{pmatrix},$$

$\det(A) = 0,$ $\det(A) = 0,$ $\det(A) \neq 0$

Metodi diretti per la soluzione di sistemi lineari

Sistema lineare $AX = B, \quad A \in R^{M\times N}, \quad X \in R^N, \quad B \in R^M$

A = matrice dei coefficienti B = vettore dei termini noti
X = vettore delle incognite

- 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:

$$\mathbf{C_c} \approx \quad \text{numero di moltiplicazioni o divisioni}$$

Richiamo sui determinanti

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

- Se A è una matrice **quadrata** di dimensione n, allora

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

  oppure

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

  dove M_{ij} è il determinante della matrice che si ottiene da A trascurando 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 $\det(\tilde{A}) = \lambda \det(A)$

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

$$X = A^{-1}B \;\Rightarrow\; x_i = \frac{D_i}{\det(A)},\; i = 1, 2, ..., n$$

,

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

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

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

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

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

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

e quindi $X = \begin{pmatrix} -2 & 2 & -\frac{1}{3} \end{pmatrix}^T$

Costo computazionale del metodo di Cramer

$$\boxed{x_i = \frac{D_i}{\det(A)} \qquad i = 1, 2, ..., n}$$

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

Costo computazionale:

$$\mathbf{C_c} = (n+1)n!(n-1) + n \simeq (n+1)n!(n-1)$$

$$\left.\begin{array}{l} n = 15 \rightarrow C_c \simeq 3 \cdot 10^{14} \text{ moltiplicazioni} \rightarrow \text{circa 4 giorni} \\ n = 20 \rightarrow C_c \simeq 3 \cdot 10^{21} \text{ moltiplicazioni} \rightarrow \text{circa 4800 anni} \end{array}\right\} \textbf{Inutilizzabile!!}$$

(supponendo $0.5 \cdot 10^{-10}$ 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

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

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

$$\boxed{UX = \widetilde{B} \qquad X, \widetilde{B} \in \mathbf{R}^n \;\; U \in \mathbf{R}^{n \times n}}$$

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

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

- **scambio** di 2 equazioni

- **somma** di un'equazione con un'altra **moltiplicata** per una costante

Soluzione di sistemi triangolari

Sistemi triangolari superiori

$$
\begin{cases}
u_{11}\,x_1 + u_{12}\,x_2 + \cdots + u_{1k}\,x_k + \cdots + u_{1n}\,x_n = b_1 \\
\qquad\quad u_{22}\,x_2 + \cdots + u_{2k}\,x_k + \cdots + u_{2n}\,x_n = b_2 \\
\qquad\qquad\quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
\qquad\qquad\qquad\qquad u_{kk}\,x_k + \cdots + u_{kn}\,x_n = b_k \\
\qquad\qquad\qquad\qquad\qquad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad u_{nn}\,x_n = b_n
\end{cases}
$$

$$\Downarrow$$

$$
UX = B \rightarrow
\begin{cases}
x_n = \dfrac{b_n}{u_{nn}} \\[2ex]
x_k = \left(b_k - \displaystyle\sum_{i=k+1}^{n} u_{ki}\,x_i \right) \dfrac{1}{u_{kk}}, \quad k = n-1, n-2, ..., 2, 1
\end{cases}
$$

Algoritmo di sostituzione all'indietro

Sistemi triangolari inferiori

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

$$\Downarrow$$

$$
LX = B \rightarrow
\begin{cases}
x_1 = \dfrac{b_1}{l_{11}} \\[2ex]
x_k = \left(b_k - \displaystyle\sum_{i=1}^{k-1} l_{ki}\,x_i \right) \dfrac{1}{l_{kk}} \qquad k = 1, 2, \ldots, n
\end{cases}
$$

Algoritmo di sostituzione in avanti

OSS. Se le matrici U e L sono **regolari**, sicuramente u_{kk} e l_{kk} sono diversi da 0.

Algoritmo di sostituzione: costo computazionale

Ad ogni passo ci sono

- $n - k$ **moltiplicazioni** (sost. indietro), $k - 1$ **moltiplicazioni** (sost. avanti)
- 1 **divisione**

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

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

OSS. $\sum_{k=1}^{n} k = \frac{1}{2} n(n+1)$

Metodo di elimazione di Gauss

Dato 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

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

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.

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

Si **moltiplica** la **seconda riga** per $\frac{6}{3}$ e si **sottrae** alla **terza**, annullando così 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

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ \quad\; - 3x_2 - 6x_3 = -4 \\ \qquad\qquad - 9x_3 = 3 \end{cases}$$

Il sistema in questa forma diventa di facile soluzione. Infatti, partendo dall'ultima equazione, si ha

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

Metodo di eliminazione di Gauss

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

passo 1: si annullano tutti gli elementi della **prima colonna** di A al di sotto della diagonale principale.

1.1 sottrarre alla **seconda** equazione la **prima** moltiplicata per $m_{21} = \frac{a_{21}}{a_{11}}$

1.2 sottrarre alla **terza** equazione la **prima** moltiplicata per $m_{31} = \frac{a_{31}}{a_{11}}$

$\vdots$

$\vdots$

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

Si ottiene così un **sistema equivalente** la cui matrice dei coefficienti e il vettore dei termini noti sono dati da

$$A^{(1)} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & \cdots & a_{3n}^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix} \qquad b^{(1)} = \begin{pmatrix} b_1 \\ b_2^{(1)} \\ b_3^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix}$$

dove

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

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

$$m_{i1} = \frac{a_{i1}}{a_{11}}, \quad i = 2, \ldots, n$$

passo 2 si annullano tutti gli elementi della **seconda colonna** di $A^{(1)}$ al di sotto della diagonale principale.

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

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

$\vdots$

$\vdots$

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

Si ottiene così un sistema equivalente la cui matrice dei coefficienti e il vettore dei termini noti sono dati da

$$A^{(2)} = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & \cdots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix} \qquad b^{(2)} = \begin{pmatrix} b_1 \\ b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix}$$

dove

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

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

$$m_{i2} = \frac{a_{i2}{}^{(1)}}{a_{22}^{(1)}}, \quad i = 3, \ldots, n$$

passo n-1: si annullano tutti gli elementi della $\mathbf{(n-1)-esima}$ colonna di $A^{(n-2)}$ al di sotto della diagonale principale.

(n-1).1 sottrarre all'equazione $\mathbf{n-esima}$ la $\mathbf{(n-1)-esima}$ moltiplicata per $m_{nn-1} = \frac{a^{(n-2)}_{n\ n-1}}{a^{(n-2)}_{n-1\ n-1}}$

Si ottiene così un sistema equivalente la cui matrice dei coefficienti e il vettore dei termini noti sono dati da

$$A^{(n-1)} = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & \cdots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{pmatrix} \qquad b^{(2)} = \begin{pmatrix} b_1 \\ b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_n^{(n-1)} \end{pmatrix}$$

dove

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

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

Dopo $n-1$ passi il sistema $Ax = b$ è stato trasformato nel **sistema triangolare equivalente**

$$A^{(n-1)}x = b^{(n-1)}$$

con $A^{(n-1)}$ **matrice triangolare superiore**.

Al passo k si definiscono gli elementi

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

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

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

MEG: costo computazionale

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

$$\underbrace{\sum_{i=1}^{n-1}}_{ogni\ \ passo} \quad \underbrace{\sum_{l=i+1}^{n}}_{righe} (\underbrace{1}_{div.} + \underbrace{n-i+2}_{molt.}) = \sum_{i=1}^{n-1} (n-i)(n-i+3) \approx \frac{n^3}{3}$$

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** $\approx \frac{n^2}{2}$

$$\Rightarrow \mathbf{C_c} \approx \frac{n^3}{3}$$

Esercizio 1

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

$$\det A_k \neq 0, \; k = 1, 2, \cdots, n,$$

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 destra la matrice A per la matrice

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

per cui

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

Moltiplicare per 2 la seconda riga della matrice ottenuta al passo precedente e sottrarla alla terza equivale a moltiplicare a destra la matrice $L_1 A$ per la matrice

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

per cui

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

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

$$L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 0 & 1 \end{pmatrix} \qquad L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

Struttura delle matrici L e U

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

matrice dei **moltiplicatori** matrice del **sistema triangolare**

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 $\boxed{\mathrm{C_c} \simeq \frac{n^3}{3}}$.

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**.

$$AX = B \overset{A=LU}{\longrightarrow} LUX = B \Rightarrow \begin{cases} LY = B & \text{sistema triang. inf.} \\ & \text{(algoritmo di sost. in avanti)} \\ UX = Y & \text{sistema triang. sup.} \\ & \text{(algoritmo di sost. all'indietro)} \end{cases}$$

Una volta **fattorizzata** A, la soluzione del sistema si ottiene risolvendo i due sistemi trangolari con **costo computazionale**

$$\mathbf{C_c} \approx \mathbf{2}\frac{n^2}{2} = \mathbf{n^2}$$

.

Soluzione di più sistemi lineari

Se dobbiamo risolvere più sistemi lineari

$$AX_i = B_i \qquad i = 1, 2, \ldots, r$$

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

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

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

Calcolo del determinante di A

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

$\downarrow$ **Teorema di Binet** $\qquad\qquad \searrow$ $= 1$

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

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

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

$\downarrow$

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

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 perché $a_{rk}^{(k)} = 0$, $r = k, \ldots, n \Rightarrow$ la matrice A ha **rango** pari a $q < n$ (**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**.

Supponiamo di voler risolvere il sitema

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

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

Moltiplichiamo la **prima equazione** per

$$m_{21} = \frac{a_{21}}{a_{11}} = \frac{0.1080}{-0.0590} \approx -1.830508 \approx -1.831$$

e sottraendola alla **seconda**, il secondo elemento della seconda colonna assume il valore

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

Mentre il secondo elemento del termine noto è

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

e quindi

$$\left(\begin{array}{ccc|c} -0.0590 & 0.2372 & & -0.3528 \\ 0 & -0.0005 & & -0.0004 \end{array} \right)$$

da cui otteniamo

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

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

mentre la soluzione esatta è

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

Pivoting parziale

Non si ha la 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

$$a_{rk}^{(k)} = \underbrace{\max}_{k \leq s \leq n} |a_{sk}^{(k)}|$$

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** (come accade in alcuni sistemi derivanti dalla soluzione numerica di problemi ai limiti per equazioni differenziali) la **fattorizzazione LU** è molto **semplice** in quanto le matrici L e U hanno una struttura semplice.

$$A = \begin{bmatrix} d_1 & s_1 & 0 & \cdots & 0 \\ a_2 & d_2 & s_2 & \cdots & 0 \\ \cdots & \ddots & \ddots & \ddots & \cdots \\ 0 & 0 & \cdots & d_{n-1} & s_{n-1} \\ 0 & 0 & \cdots & a_n & d_n \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha_2 & 1 & 0 & \cdots & 0 \\ 0 & \alpha_3 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n & 1 \end{bmatrix} \begin{bmatrix} u_1 & v_1 & 0 & \cdots & 0 \\ 0 & u_2 & v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{n-1} & v_{n-1} \\ 0 & 0 & \cdots & 0 & u_n \end{bmatrix} = LU$$

Algoritmo di Thomas

$$\begin{cases} u_1 = d_1 & \\ v_i = s_i & i = 1, 2, \ldots, n-1 \\ \alpha_i = a_i / u_{i-1} & i = 2, 3, \ldots, n \\ u_i = d_i - \alpha_i \, v_{i-1} & i = 2, 3, \ldots, n \end{cases}$$

Costo computazionale: $C_c = \underbrace{(n-1)}_{divisioni} + \underbrace{(n-1)}_{\substack{moltipli- \\ cazioni}} = 2n - 2$

Soluzione del sistema lineare $AX = B$

$LY = B \quad \Rightarrow y_1 = b_1 \qquad y_i = b_i - \alpha_i \, y_{i-1} \qquad i = 2, 3, \ldots, n$

$UX = Y \quad \Rightarrow x_n = y_n / u_n \quad x_i = (y_i - v_i \, x_{i+1}) / u_i \qquad i = n-1, \ldots, 1$

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

Esempio 2: soluzione

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

solutore di Matlab:

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

Algo 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.

La **memorizzazione** di A richiede l'uso di 3 soli vettori contenenti gli elementi della diagonale e delle codiagonali. Da un punto di vista computazionale, se non si è interessati a conservare gli elementi di A si possono memorizzare i vettori $[u_1, \ldots, u_n], [\alpha_2, \ldots, alpha_n]$ nell'area di memoria di A, sovrascrivendoli alla diagonale e alla sottodiagonale rispettivamente.

Nota. In realtà gli elementi diversi da zero della matrice A sono

$$N + (N-1) + (N-1) = 3N - 2 << N^2$$

$\rightarrow$ A è una matrice **sparsa**

Esercizio 2

Data la matrice tridiagonale $A = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 4 \end{bmatrix}$

calcolarne l'inversa con l'algoritmo di Thomas.

Soluzione

Poiché $\det A = 6 \neq 0$, la matrice A è **regolare** e quindi ammette l'inversa.

L'inversa è definita dalla relazione $A\,A^{-1} = I$, pertanto le colonne di $A^{-1} = [X_1 \; X_2 \; X_3]$ sono le soluzioni dei sistemi lineri $A\,X_i = E_i$, dove E_i, $i = 1, 2, 3$, sono i tre vettori della **base canonica** in $\mathbb{R}^3$.

Dall'uguaglianza $A = LU$, dove

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \alpha_2 & 1 & 0 \\ 0 & \alpha_3 & 1 \end{bmatrix} \qquad U = \begin{bmatrix} u_1 & v_1 & 0 \\ 0 & u_2 & v_2 \\ 0 & 0 & u_3 \end{bmatrix}$$

si ottiene

$$u_1 = 2 \quad v_1 = -2 \qquad\qquad v_2 = -1$$

$$\alpha_2 = -1 \quad u_2 = 3 - (-1)(-2) = 1$$

$$\alpha_3 = -1 \quad u_3 = 4 - (-1)(-1) = 3$$

$$\Rightarrow \quad L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 2 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{bmatrix}$$

La soluzione dei sistemi lineari

$$L\,Y_i = E_i \qquad U\,X_i = Y_i$$

con gli algoritmi di sostituzione in avanti e indietro dà

$$X_1 = \begin{bmatrix} \frac{11}{6} \\ \\ \frac{4}{3} \\ \\ \frac{1}{3} \end{bmatrix} \qquad X_2 = \begin{bmatrix} \frac{4}{3} \\ \\ \frac{4}{3} \\ \\ \frac{1}{3} \end{bmatrix} \qquad X_3 = \begin{bmatrix} \frac{1}{3} \\ \\ \frac{1}{3} \\ \\ \frac{1}{3} \end{bmatrix}$$

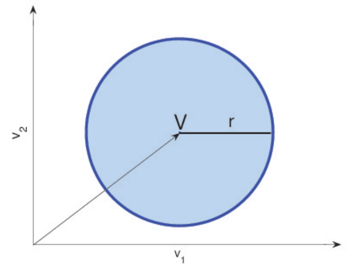
$$\Rightarrow \qquad A^{-1} = \begin{bmatrix} \frac{11}{6} & \frac{4}{3} & \frac{1}{3} \\ \\ \frac{4}{3} & \frac{4}{3} & \frac{1}{3} \\ \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Norma di vettore

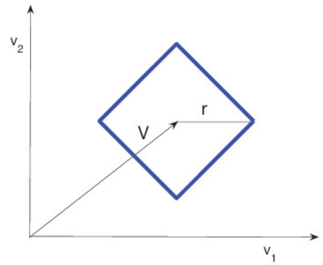
La **norma** di un vettore $V = [v_1, \ldots, v_n]^T$ viene utilizzata per "*misurare*" la sua **lunghezza**.

Intorno: $\|V - W\| \leq r$

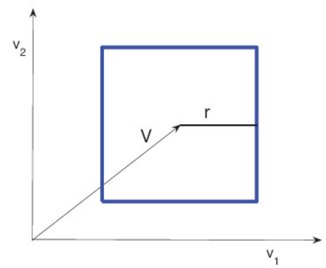
- **Norma due o euclidea:** $\|V\|_2 := \sqrt{\sum_{i=1}^{n} |v_i|^2}$



- **Norma uno:** $\|V\|_1 := \sum_{i=1}^{n} |v_i|$



- **Norma infinito:** $\|V\|_\infty := \max_{1 \leq i \leq n} |v_i|$



Nota. Tutte le norme sono **equivalenti**: $m\|V\|_p \leq \|V\|_q \leq M\|V\|_p$

Proprietà della norma di vettore

- $\|V\| \geq 0, \quad \|V\| = 0 \Longleftrightarrow V = 0$
- $\|\alpha V\| = |\alpha| \cdot \|V\| \quad \forall\, \alpha \in \mathbf{R},\ \forall\, V \in \mathbf{R}^n$
- $\|V + W\| \leq \|V\| + \|W\| \quad \forall\, V, W \in \mathbf{R}^n$ *(disuguaglianza triangolare)*

Distanza: in uno **spazio vettoriale normato** S è possibile introdurre la **distanza** tra due punti V e W in S

$$d(V, W) := \|V - W\|$$

Proprietà della distanza:

- $d(V, W) = 0 \iff V = W$
- $d(V, W) = d(W, V) \qquad \forall\, V, W \in S$
- $d(V, W) \leq d(V, Z) + d(Z, W) \qquad \forall\, V, W, Z \in S$

Norme di matrici

La **norma** di una matrice $A \in \mathbf{R}^{n \times n}$ soddisfa le seguenti

Proprietà

- $\|A\| \geq 0, \quad \|A\| = 0 \Longleftrightarrow A = 0$
- $\|\alpha A\| = |\alpha| \cdot \|A\|, \quad \forall \alpha \in \mathbf{R},\ \forall A \in \mathbf{R}^{n \times n}$
- $\|A + B\| \leq \|A\| + \|B\|, \quad \forall A, B \in \mathbf{R}^{n \times n}$ *(disuguaglianza triangolare)*
- $\|A \cdot B\| \leq \|A\| \cdot \|B\|, \quad \forall A, B \in \mathbf{R}^{n \times n}$

Definizione. Una matrice si dice **convergente** se $\lim_{k \to \infty} \|A^k\| = 0$

Norme indotte dalla norma di vettore

Ogni **norma di vettore** può essere utilizzata per definire una **norma di matrice** che permette di "*misurare*" come la matrice agisce sui vettori:

$$\|A\| = \max_{\|X\|=1} \|A\,X\| \qquad A \in \mathbf{R}^{n\times n} \quad X \in \mathbf{R}^{n}$$

$\|A\|$ misura la massima lunghezza del vettore AX

Le norme indotte soddisfano tutte le **proprietà delle norme** e, inoltre, soddisfano la **relazione di compatibilità** , che pone in relazione norme di vettori con norme di matrici:

$$\boxed{\|A\,X\| \leq \|A\| \cdot \|X\|}$$

dim. Infatti, se $X \neq 0$, si ha

$$\|A\| = \max_{\|X\|=1} \|A\,X\| = \max_{\|X\|\neq 0} \left\| \frac{A\,X}{\|X\|} \right\| = \max_{\|X\|\neq 0} \frac{\|A\,X\|}{\|X\|} \Rightarrow \|A\| \geq \frac{\|A\,X\|}{\|X\|}$$

Nota. Per tutte le norme indotte si ha $\|I\| = 1$ (I: matrice identità)

Norme indotte: esempi

- **Norma uno:** $\quad ||A||_1 := \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|$ *(per colonne)*

- **Norma infinito:** $\quad ||A||_\infty := \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$ *(per righe)*

- **Norma due o spettrale:** $\quad ||A||_2 := \sqrt{\rho(A^T A)}$

dove $\rho(M) := \max_i |\lambda_i|$ (λ_i: autovalori di M) è il **raggio spettrale** della matrice $M \in \mathbf{R}^{n \times n}$.
Se A è **simmetrica** $\Rightarrow \rho(A^T A) = \rho^2(A) \Rightarrow \|A\|_2 = \rho(A)$

Teorema. Per una norma verificante la **relazione di compatibilità** si ha $\rho(A) \le \|A\|$.

dim Infatti da $\lambda X = A\,X \Rightarrow \|\lambda X\| = \|A\,X\| \le \|A\| \cdot \|X\| \Rightarrow |\lambda| \le \|A\|$.

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**.

Errore sul termine noto

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

$$A\,X = B \stackrel{\delta B}{\to} A(X + \delta X) = B + \delta B$$

Per sottrazione si ricava

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

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** $||\delta X||/||X||$.

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

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

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

$$\downarrow$$

Coefficiente di amplificazione

Condizionamento in norma 2

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

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

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

Esercizio 3

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

$$A^{-1} = \frac{1}{\delta^2} \begin{pmatrix} 1 & -1-\delta \\ -1+\delta & 1 \end{pmatrix}$$

per cui il **numero di condizionamento** (rispetto alla norma infinito) è dato da:

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

Si osserva

$$\lim_{\delta \to \infty} K_\infty(A) = 1$$

$$\lim_{\delta \to 0} K_\infty(A) = +\infty$$

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

Per esempio per $\delta = 0.01$, $K_\infty(A) = (201)^2 = 40401$, mentre per $\delta = 100$, $K_\infty(A) = 1.0404$

Volendo dare una stima dell'**errore relativo** commesso **sulla soluzione** X, si ha

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

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 $||\delta X||_\infty = 201$.

Esercizio 4

L. Gori, M.L. Lo Cascio, F. Pitolli, Esercizi di Calcolo Numerico, 7.52

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

a) individuare i valori di λ per cui $\|A(\lambda)\|_1 \leq 3$;

b) studiare come varia il numero di condizionamento $K(A(\lambda))$ in **norma 1** per $|\lambda| \leq \dfrac{1}{2}$ e trovarne il massimo;

c) 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

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

$\Rightarrow \; \|A(\lambda)\|_1 \le 3$ per $|\lambda| \le 1$

b) $K(A(\lambda)) = \|A(\lambda)\|_1 \; \|A^{-1}(\lambda)\|_1, \quad |\lambda| \le \frac{1}{2}$

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

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

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

c) $\frac{||\delta X||_1}{||X||_1} \le K(A(1/2)) \; \frac{||\delta B||_1}{||B||_1} = 4 \cdot 10^{-2}$

Esercizio 5

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 è

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

e dipende dalla norma di matrici scelta.

Calcoliamo $K_1(A), K_\infty(A), K_2(A)$, ossia il numero di condizionamento di A rispetto alle norme $1, \infty$ e spettrale.

La matrice inversa di A è data da

$$A^{-1} = \frac{1}{0.002} \begin{pmatrix} 1.001 & -1 \\ -2 & 2 \end{pmatrix}$$

Per cui avremo

$$||A||_\infty = 3.001, \quad ||A^{-1}||_\infty = \frac{4}{0.002}, \quad \Rightarrow K_\infty(A) = 3.001 \cdot 2000 = 6002$$

$$||A||_1 = 4, \quad ||A^{-1}||_\infty = \frac{3.001}{0.002}, \quad \Rightarrow K_1(A) = 4 \cdot \frac{3.001}{2000} = 6002$$

La **norma spettrale** di A è data da $||A||_2 = \sqrt{\rho(A^TA)}$, per cui si deve calcolare il raggio spettrale della matrice A^TA

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

i cui autovalori λ_1, λ_2 sono soluzione dell'equazione di secondo grado

$$\lambda^2 - 10.002001\lambda + 0.000004 = 0$$

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

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

Analogamente

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

da cui risulta

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

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

$$(x, y) = (1501.5, -3000)$$

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

$$(\tilde{x}, \tilde{y}) = (751.5, -1500)$$

L'errore relativo, rispetto alla norma infinito, è

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

mentre

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

Riferimenti bibliografici

L. Gori, *Calcolo Numerico*:
Cap. 2 §§ **2.1-2.5, 2.8-2.11**
Cap. 4 §§ **4.1-4.3, 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, 7.15, 7.52