

Algebra lineare numerica

Alvise Sommariva

Università degli Studi di Padova

15 maggio 2025

Definizione (Norma vettoriale (spazio vettoriale sopra \mathbb{R}))

Dato uno spazio vettoriale V (sopra il campo \mathbb{R}), la funzione $\| \cdot \| : V \rightarrow \mathbb{R}$ è una **norma vettoriale** se

- 1 $\|x\| \geq 0$ per ogni $x \in V$ e $\|x\| = 0$ se e solo se $x = 0_V$,
- 2 $\|\alpha x\| = |\alpha|_{\mathbb{R}} \|x\|$ per ogni $x \in V$ e scalare $\alpha \in \mathbb{R}$,
- 3 vale la cosiddetta disuguaglianza triangolare

$$\|x + y\| \leq \|x\| + \|y\|$$

per ogni $x, y \in V$.

Definizione (Norma vettoriale (spazio vettoriale sopra \mathbb{C}))

Dato uno spazio vettoriale V (sopra il campo \mathbb{C}), la funzione $\| \cdot \| : V \rightarrow \mathbb{R}$ è una **norma vettoriale** se

- 1 $\|x\| \geq 0$ per ogni $x \in V$ e $\|x\| = 0$ se e solo se $x = 0_V$,
- 2 $\|\alpha x\| = |\alpha|_{\mathbb{C}} \|x\|$ per ogni $x \in V$ e scalare $\alpha \in \mathbb{C}$,
- 3 vale la cosiddetta disuguaglianza triangolare

$$\|x + y\| \leq \|x\| + \|y\|$$

per ogni $x, y \in V$.

Nota. (Norma in \mathbb{C})

Si ricorda che se $z = a + i \cdot b$ allora $|z|_{\mathbb{C}} = \sqrt{a^2 + b^2}$.

Esempio (Alcune norme vettoriali)

Se $V = \mathbb{C}^n$ o $V = \mathbb{R}^n$ e $x = (x_1, x_2, \dots, x_n) \in V$ alcuni esempi di norme sono

- per $p \in [1, +\infty)$, la **norma** p è definita come

$$\|x\|_p = \left(\sum_{j=1}^n |x_j|^p \right)^{1/p}$$

- la **norma del massimo** è definita per $p = \infty$ come

$$\|x\|_\infty = \max_{j=1, \dots, n} |x_j|$$

Nota.

Di seguito, non distingueremo se lo spazio vettoriale V , sia uguale \mathbb{C}^n o $V = \mathbb{R}^n$. Qualora necessario, per semplicità, si ponga $V = \mathbb{R}^n$.

Nota. (Alcuni casi notevoli della norma p)

Si osservi che essendo $\|x\|_p = (\sum_{j=1}^n |x_j|^p)^{1/p}$

- per $p = 1$, si definisce la importante **norma 1**

$$\|x\|_1 = \sum_{j=1}^n |x_j|$$

- per $p = 2$, si definisce la importante **norma 2**, detta anche **euclidea**

$$\|x\|_2 = \left(\sum_{j=1}^n |x_j|^2 \right)^{1/2}$$

Esempio

Si consideri il vettore $v = (4, -3) \in \mathbb{R}^2$. Allora

- $\|v\|_1 = |4| + |-3| = 7,$
- $\|v\|_2 = \sqrt{|4|^2 + |-3|^2} = 5,$
- $\|v\|_\infty = \max(|4|, |-3|) = 4.$

Esempio

Si considerino i vettori $u = (4, -3), v = (4.001, -2.999) \in \mathbb{R}^2$.

I due vettori sono numericamente *vicini*. Per verificarlo calcoliamo la loro distanza in termine di norme ovvero

$$\text{dist}(u, v) := \|u - v\|_p, \quad p = [1, \infty].$$

Allora, visto che $u - v = [-0.001, -0.001]$

- $\|u - v\|_1 = 0.002,$
- $\|u - v\|_2 = 0.0014,$
- $\|u - v\|_\infty = \max(|-0.001|, |-0.001|) = 0.001,$

risultati che sottolineano che u e v sono vicini rispetto alle distanze definite dalle norme.

Definizione

Sia $\|\cdot\|$ una norma vettoriale. Si definisce **norma vettoriale indotta** di una matrice in $\mathbb{C}^{n \times n}$ la funzione che mappa ogni matrice $A \in \mathbb{C}^{n \times n}$ nel valore

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Nota.

Tali norme si definiscono anche nel caso particolare che $A \in \mathbb{R}^{n \times n}$.

Definizione (Raggio spettrale)

Si definisce **raggio spettrale** di A la quantità

$$\rho(A) = \max_{k=1, \dots, n} (|\lambda_k|)$$

dove λ_k è un **autovalore** di A , ovvero esiste $x_k \neq 0$ detto **autovettore** di A , tale che $Ax_k = \lambda_k x_k$.

Esempio (Alcune norme matriciali)

Posto $A = (a_{i,j})$, alcuni esempi di norme matriciali risultano:

- per $p = 1$, partendo dalla norma vettoriale $\|\cdot\|_1$, si dimostra che la **norma 1** matriciale risulta uguale al massimo della somma dei moduli delle componenti di ogni **colonna** ovvero

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{i,j}|;$$

- per $p = 2$, partendo dalla norma vettoriale $\|\cdot\|_2$, si dimostra che la **norma 2** matriciale risulta uguale a

$$\|A\|_2 = \sqrt{\rho(A^*A)};$$

dove

- se $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ allora A^* è la trasposta A^T di A , ovvero $a_{i,j}^* = a_{j,i}$,
- se $A \in \mathbb{C}^{n \times n}$ allora A^* è la trasposta coniugata A^H di A ovvero $a_{i,j}^* = \overline{a_{j,i}}$;
- per $p = \infty$, partendo dalla norma vettoriale $\|\cdot\|_\infty$, si dimostra che la **norma ∞** matriciale risulta uguale al massimo della somma dei moduli delle componenti di ogni **riga** ovvero

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

Esempio (1)

Si consideri la matrice $A \in \mathbb{R}^{2 \times 2}$

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

Allora

- $\|A\|_1 = \max(|1| + |-3|, |-2| + |4|) = 6$;
- osservato che $A^* = A^T$ abbiamo

$$A^T A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} 10 & -14 \\ -14 & 20 \end{bmatrix}$$

i cui autovalori valgono rispettivamente

$$\lambda_1 \approx 0.1339312526814949, \quad \lambda_2 \approx 29.86606874731851$$

e di conseguenza

$$\|A\|_2 = \sqrt{\max(|\lambda_1|, |\lambda_2|)} \approx 5.464985704219044;$$

- $\|A\|_\infty = \max(|1| + |-2|, |-3| + |4|) = 7$.

Esempio (2)

Si considerino le matrici

$$A = \begin{bmatrix} 1.001 & -2.999 \\ -2.002 & 4.005 \end{bmatrix} \quad B = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}$$

Le due matrici sono numericamente *vicine*. Per verificarlo calcoliamo la loro distanza in termine di norme ovvero $\text{dist}(A, B) := \|A - B\|_p$, $p = [1, \infty]$.

Allora, visto che

$$A - B = \begin{bmatrix} 0.0010 & 0.0010 \\ -0.0020 & 0.0050 \end{bmatrix}$$

ricaviamo

- $\|A - B\|_1 = \max(0.003, 0.006) = 0.006$,
- $\|A - B\|_2 = 0.005415654779058332$,
- $\|A - B\|_\infty = \max(0.002, 0.007) = 0.007$,

risultati che sottolineano che A e B sono matrici vicine rispetto alle distanze definite dalle norme.

Teorema (Disuguaglianze norme indotte)

Per le norme matriciali indotte, se $A, B \in \mathbb{C}^{n \times n}$, $x \in \mathbb{C}^{n \times 1}$,

$$\|Ax\| \leq \|A\| \|x\|, \quad (1)$$

$$\|AB\| \leq \|A\| \|B\|. \quad (2)$$

Definizione (Indice di condizionamento)

Sia $A \in \mathbb{C}^{n \times n}$ una matrice invertibile ossia tale che $\det(A) \neq 0$. Si dice **indice di condizionamento** della matrice A , relativamente alla norma indotta $\|\cdot\|$,

$$k(A) := \|A\| \|A^{-1}\|.$$

Nota.

Si osservi che dalla definizione, se I è la matrice identica di $\mathbb{C}^{n \times n}$

$$\|I\| = \sup_{x \neq 0} \frac{\|Ix\|}{\|x\|} = \sup_{x \neq 0} \frac{\|x\|}{\|x\|} = 1$$

e quindi, se la norma $\|\cdot\|$ è indotta, essendo $\|ST\| \leq \|S\| \|T\|$ per ogni $S, T \in \mathbb{C}^{n \times n}$, ricaviamo

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\| = k(A)$$

ovvero **$k(A) \geq 1$** .

Valgono i seguenti risultati

Teorema (Stime dal basso dell'indice di condizionamento)

- 1 *Qualsiasi sia la norma matriciale indotta, e A invertibile,*

$$1 \leq \frac{|\lambda_{\max}|}{|\lambda_{\min}|} \leq k(A)$$

dove λ_{\min} e λ_{\max} sono rispettivamente gli autovalori di A di minimo e massimo modulo.

- 2 Se A è *simmetrica*, ovvero $A = A^T$,

$$k_2(A) = \frac{|\lambda_{\max}|}{|\lambda_{\min}|}.$$

Esempio (Matrice di Hilbert)

La matrice $H = (h_{i,j})_{i,j=1,\dots,n}$ con

$$h_{i,j} = \frac{1}{i+j-1}$$

è nota come matrice di Hilbert di ordine n . Per queste matrici gli indici di condizionamento crescono molto rapidamente, come si può vedere dalla tabella.

n	c_n	n	c_n	n	c_n
1	1.00e + 00	5	4.77e + 05	9	4.93e + 11
2	1.93e + 01	6	1.50e + 07	10	1.60e + 13
3	5.24e + 02	7	4.75e + 08	11	5.22e + 14
4	1.55e + 04	8	1.53e + 10	12	1.62e + 16

Tabella: Indice di condizionamento $c_n = \text{cond}_2(H_n)$ della matrici di Hilbert H_n di ordine n

Vale la seguente risposta della soluzione di un sistema lineare agli errori del vettore termine noto.

Teorema (Condizionamento ed errore nel termine noto)

Se

- 1 $A \in \mathbb{C}^{n \times n}$ è una matrice invertibile,
- 2 $b \in \mathbb{C}^n$, $b \neq 0$,
- 3 $Ax = b$,
- 4 $A(x + \delta x) = b + \delta b$,
- 5 $\|\cdot\|$ è una norma indotta,

allora vale la stima

$$\frac{\|\delta x\|}{\|x\|} \leq k(A) \frac{\|\delta b\|}{\|b\|}$$

dove $k(A) = \|A\| \|A^{-1}\|$.

Commento (Relazione tra $\|\delta x\|/\|x\|$ e $\|\delta b\|/\|b\|$)

- Questo teorema mostra che si fa un errore relativo sul termine noto pari a $\frac{\|\delta b\|}{\|b\|}$ allora si compie un errore relativo tra soluzione x e soluzione $x + \delta x$ del problema perturbato

$$\frac{\|x - (x + \delta x)\|}{\|x\|} = \frac{\|\delta x\|}{\|x\|}$$

tale che

$$\frac{\|\delta x\|}{\|x\|} \leq k(A) \frac{\|\delta b\|}{\|b\|}.$$

- Quindi, se $k(A) \gg 1$, può accadere che la soluzione $x + \delta x$ del problema perturbato sia molto distante da x e di conseguenza la soluzione del sistema lineare $Ax = b$ molto suscettibile a piccoli errori sul termine noto.
- In altri termini **più grande è $k(A)$** più i sistemi lineari $Ax = b$ sono **difficili da trattare** nel senso che piccoli errori sui dati b possono portare a grossi errori sulla soluzione x .

Esempio

Siano

$$A = \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0.7 \end{bmatrix}, \quad b + \delta b = \begin{bmatrix} 1.01 \\ 0.69 \end{bmatrix}.$$

Il sistema $Ax = b$ ha soluzione

$$x = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}$$

mentre $A(x + \delta x) = b + \delta b$ ha soluzione

$$x + \delta x = \begin{bmatrix} -0.17 \\ 0.22 \end{bmatrix}.$$

Di conseguenza seppure $\frac{\|\delta b\|_\infty}{\|b\|_\infty} = 0.01$ abbiamo $\frac{\|\delta x\|_\infty}{\|x\|_\infty} = 1.7$.

Questo significa che a un piccolo errore relativo sui dati abbiamo riscontrato un eccessivo errore relativo sulla soluzione.

Osserviamo che in questo caso $k(A) = \|A\|_\infty \|A^{-1}\|_\infty = 289$ e quindi non troppo vicino a 1.

Teorema (Condizionamento e errore sui dati della matrice)

Se

- 1 $A \in \mathbb{C}^{n \times n}$ è una matrice invertibile,
- 2 $b \in \mathbb{C}^n$, $b \neq 0$,
- 3 $Ax = b$,
- 4 $(A + \delta A)(x + \delta x) = b$, con $\det(A + \delta A) \neq 0$,

allora per una qualsiasi norma indotta $\|\cdot\|$ vale la stima

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq k(A) \frac{\|\delta A\|}{\|A\|}.$$

dove $k(A) = \|A\| \|A^{-1}\|$.

Commento

Il precedente teorema dice che in presenza di un forte indice di condizionamento, piccoli errori relativi sulla matrice possono indurre grandi errori relativi nella soluzione. Si noti che usualmente $\|x\| \approx \|x + \delta x\|$.

Esempio

Siano

$$A = \begin{bmatrix} 1.0000 & 0.5000 & 0.3333 \\ 0.5000 & 0.3333 & 0.2500 \\ 0.3333 & 0.2500 & 0.2000 \end{bmatrix}, A + \delta A = \begin{bmatrix} 1.0010 & 0.5030 & 0.3403 \\ 0.5040 & 0.3383 & 0.2520 \\ 0.3423 & 0.2520 & 0.2050 \end{bmatrix},$$

Il sistema $Ax = b$, per $b = [1, 2, 3]^T$ ha soluzione approssimativamente

$$x \approx \begin{bmatrix} 27.3232 \\ -193.6572 \\ 211.5374 \end{bmatrix}$$

mentre $(A + \delta A)(x + \delta x) = b$ ha soluzione

$$x + \delta x \approx \begin{bmatrix} 5.5489 \\ -75.3599 \\ 98.0063 \end{bmatrix}.$$

Di conseguenza seppure $\frac{\|\delta A\|_2}{\|A\|_2} = 9.35310e - 03$ abbiamo $\frac{\|\delta x\|_2}{\|x + \delta x\|_2} = 1.33653e + 00$. Questo significa che a un piccolo errore relativo sui dati abbiamo riscontrato un eccessivo errore relativo sulla soluzione.

Osserviamo che in questo caso $k(A) = \|A\|_\infty \|A^{-1}\|_\infty \approx 528.5409$ e quindi non troppo vicino a 1.

Proposito.

Il proposito di questa sezione è di introdurre il metodo dell'eliminazione gaussiana per risolvere sistemi lineari $Ax = b$ dove supponiamo A invertibile, cioè tale che $\det(A) \neq 0$.

Per semplificare l'esposizione cominciamo con un esempio.

Esempio

Si risolva il sistema lineare

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 6x_1 - 2x_2 + 2x_3 = 6 \\ -9x_1 + 7x_2 + x_3 = -1 \end{cases}$$

mediante eliminazione gaussiana.

Sappiamo che

- se moltiplichiamo ambo i membri di una stessa equazione per una costante non nulla, il nuovo problema ha le stesse soluzioni del precedente;
- se scambiamo due equazioni, il nuovo problema ha le stesse soluzioni del precedente;
- in generale **se moltiplichiamo la i -sima equazione per una costante non nulla e sottraiamo il risultato membro a membro alla j -sima**, ricaviamo una equazione che sostituita alla j -sima, propone un nuovo problema lineare che ha le stesse soluzioni del precedente.

Svolgimento.

Diremo che due sistemi sono equivalenti, se hanno le stesse soluzioni. Da:

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 6x_1 - 2x_2 + 2x_3 = 6 \\ -9x_1 + 7x_2 + x_3 = -1 \end{cases}$$

- moltiplichiamo la **prima equazione** per $m_{2,1} = 6/3 = 2$ e la sottraiamo alla seconda, sostituendo il risultato a quest'ultima, ottenendo il sistema lineare equivalente

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 0 \cdot x_1 + 2x_2 + 4x_3 = 6 \\ -9x_1 + 7x_2 + x_3 = -1 \end{cases}$$

- moltiplichiamo la **prima equazione** per $m_{3,1} = -9/3 = -3$ e la sottraiamo alla terza, sostituendo il risultato a quest'ultima, ottenendo il sistema lineare equivalente

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

in cui le componenti nulle della prima colonna, precedentemente ottenute, non vengono modificate.

- moltiplichiamo la **seconda equazione** per $m_{3,2} = 1/2$ e la sottraiamo alla terza, sostituendo il risultato a quest'ultima, ottenendo il sistema lineare equivalente

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 0x_1 + 2x_2 + 4x_3 = 6 \\ 0x_1 + 0x_2 - 4x_3 = -4 \end{cases}$$

in cui le componenti nulle della colonna precedente non vengono modificate.

Il sistema finale

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 0x_1 + 2x_2 + 4x_3 = 6 \\ 0x_1 + 0x_2 - 4x_3 = -4 \end{cases}$$

è facile da risolvere.

- L'ultima equazione $-4x_3 = -4$ ha una sola variabile e una sola incognita e comporta che

$$x_3 = 1.$$

- Inseriamo $x_3 = 1$ nella penultima equazione $2x_2 + 4x_3 = 6$ e otteniamo

$$2x_2 + 4 \cdot 1 = 6$$

da cui $2x_2 = 2$ e quindi $x_2 = 1$.

- Inseriamo $x_2 = x_3 = 1$ nella prima equazione $3x_1 - 2x_2 - x_3 = 0$ e otteniamo

$$3x_1 - 2 \cdot 1 - 1 = 0$$

da cui

$$x_1 = 1.$$

Quindi il sistema ha soluzioni $x_1 = 1$, $x_2 = 1$, $x_3 = 1$.

Nota. (Sistema matriciale)

Il sistema può essere scritto matricialmente come $Ax = b$ dove

$$A = \begin{bmatrix} 3 & -2 & -1 \\ 6 & -2 & 2 \\ -9 & 7 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix},$$

e abbiamo mostrato essere equivalente a $Ux = c$ dove

$$U = \begin{bmatrix} 3 & -2 & -1 \\ 0 & 2 & 4 \\ 0 & 0 & -4 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 6 \\ -4 \end{bmatrix}.$$

*Dal punto di vista matriciale, abbiamo generato una sequenza di sistemi lineari equivalenti, ma via via più **semplici**.*

*L'ultimo è un sistema $Ux = c$ con U triangolare superiore, ovvero tale che $U_{i,j} = 0$ qualora $i > j$, che abbiamo risolto facilmente mediante **sostituzione all'indietro**.*

Operativamente, al posto dello studio di

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 6x_1 - 2x_2 + 2x_3 = 6 \\ -9x_1 + 7x_2 + x_3 = -1 \end{cases}$$

si considera quello di $A^{(1)}x = b^{(1)}$ dove

$$A^{(1)} = A = \begin{bmatrix} 3 & -2 & -1 \\ 6 & -2 & 2 \\ -9 & 7 & 1 \end{bmatrix}, \quad b^{(1)} = b = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix},$$

I moltiplicatori sono $m_{2,1} = a_{2,1}^{(1)}/a_{1,1}^{(1)} = 2$ e $m_{3,1} = a_{3,1}^{(1)}/a_{1,1}^{(1)} = -3$. Definiamo ora

$$L^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ -m_{2,1} & 1 & 0 \\ -m_{3,1} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix},$$

e poniamo $A^{(2)} = L^{(1)} \cdot A^{(1)}$, $b^{(2)} = L^{(1)} \cdot b^{(1)}$. Si ricavano quindi

$$A^{(2)} = \begin{bmatrix} 3 & -2 & -1 \\ 0 & 2 & 4 \\ 0 & 1 & -2 \end{bmatrix}, \quad b^{(2)} = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix},$$

Dopo il primo passo si ottiene un sistema equivalente

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

ovvero matricialmente si considera il problema equivalente $A^{(2)}x = b^{(2)}$

$$A^{(2)} = \begin{bmatrix} 3 & -2 & -1 \\ 0 & 2 & 4 \\ 0 & 1 & -2 \end{bmatrix}, \quad b^{(2)} = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix},$$

in cui i termini della prima colonna di $A^{(1)}$, strettamente sotto la diagonale, sono nulli.

Nota. (Importante)

Notiamo che matricialmente se $A^{(1)}x = b^{(1)}$, essendo $\det(L^{(1)}) = 1 \neq 0$ allora $L^{(1)}A^{(1)}x = L^{(1)}b$ ovvero $A^{(2)}x = b^{(2)}$.

In altri termini il vettore $x = (x_1, x_2, x_3)$ è soluzione tanto di $A^{(1)}x = b^{(1)}$ quanto del più semplice sistema lineare $A^{(2)}x = b^{(2)}$.

Consideriamo ora il sistema $A^{(2)}x = b^{(2)}$ con

$$A^{(2)} = \begin{bmatrix} 3 & -2 & -1 \\ 0 & 2 & 4 \\ 0 & 1 & -2 \end{bmatrix}, \quad b^{(2)} = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix},$$

L'unico moltiplicatore utile a porre uguali a 0 le componenti della seconda colonna sotto $a_{2,2}^{(2)}$ ottenuto è $m_{3,1} = a_{3,2}^{(2)}/a_{2,2}^{(2)} = 1/2$. Definiamo ora

$$L^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{3,1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix},$$

e poniamo $A^{(3)} = L^{(2)} \cdot A^{(2)}$, $b^{(3)} = L^{(2)} \cdot b^{(2)}$. Si ricavano quindi

$$A^{(3)} = \begin{bmatrix} 3 & -2 & -1 \\ 0 & 2 & 4 \\ 0 & 0 & -4 \end{bmatrix}, \quad b^{(3)} = c = \begin{bmatrix} 0 \\ 6 \\ -4 \end{bmatrix},$$

Nota. (Importante)

Notiamo che matricialmente se $A^{(2)}x = b^{(2)}$, essendo $\det(L^{(2)}) = 1 \neq 0$ allora $L^{(2)}A^{(2)}x = L^{(2)}b$ ovvero $A^{(3)}x = b^{(3)}$.

In altri termini il vettore $x = (x_1, x_2, x_3)$ è soluzione tanto di $A^{(1)}x = b^{(1)}$ quanto del più semplice sistema lineare $A^{(2)}x = b^{(2)}$, quanto del sistema triangolare $A^{(3)}x = b^{(3)}$.

Infine si studia il problema

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 0x_1 + 2x_2 + 4x_3 = 6 \\ 0x_1 + 0x_2 - 4x_3 = -4 \end{cases}$$

ovvero matricialmente si considera il problema equivalente $A^{(3)}x = b^{(3)}$ dove

$$A^{(3)} = U = \begin{bmatrix} 3 & -2 & -1 \\ 0 & 2 & 4 \\ 0 & 0 & -4 \end{bmatrix}, \quad b^{(3)} = c = \begin{bmatrix} 0 \\ 6 \\ -4 \end{bmatrix},$$

Notiamo in particolare che visto che $U = A^{(3)}$, $A = A^{(1)}$ e $b = b^{(1)}$

- $U = A^{(3)} = L^{(2)} \cdot A^{(2)} = L^{(2)} \cdot (L^{(1)} \cdot A^{(1)}) = L^{(2)} \cdot L^{(1)} \cdot A,$
- $b^{(3)} = L^{(2)} \cdot b^{(2)} = L^{(2)} \cdot (L^{(1)} \cdot b^{(1)}) = L^{(2)} \cdot L^{(1)} \cdot b.$

Di seguito mostreremo

- 1 come risolvere il sistema finale $Ux = c$;
- 2 come generare passo passo le matrici $A^{(1)}, A^{(2)}, \dots, A^{(n)}$ e i corrispondenti termini noti $b^{(1)}, b^{(2)}, \dots, b^{(n)}$.

In questa sezione mostriamo come si può effettuare matricialmente il processo di eliminazione gaussiana, per la risoluzione di sistemi lineari $Ax = b$ dove in generale si suppone

- $A \in \mathbb{C}^{n \times n}$, invertibile e non triangolare,
- $b \in \mathbb{C}^n$.

Poniamo

- $A = A^{(1)} = (a_{i,j}^{(1)})$,
- $b = b^{(1)} = (b_j^{(1)})$,

Intendiamo quindi proporre quanto evidenziato in precedenza al caso più generale di matrici di dimensione maggiore di 3.

Primo passo.

- Supponiamo sia l'elemento pivotale $a_{1,1}^{(1)} \neq 0$.
- Per $i = 2, \dots, n$, calcoliamo i *moltiplicatori*

$$m_{i,1} = \frac{a_{i,1}^{(1)}}{a_{1,1}^{(1)}},$$

- Definiamo la matrice

$$L^{(1)} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -m_{2,1} & 1 & 0 & \dots & 0 \\ -m_{3,1} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -m_{n,1} & 0 & 0 & \dots & 1 \end{bmatrix}$$

e poniamo

$$A^{(2)} = L^{(1)} \cdot A^{(1)}, \quad b^{(2)} = L^{(1)} \cdot b^{(1)}$$

- Dopo questa prima fase, essendo $\det(L^{(1)}) = 1 \neq 0$, come detto in precedenza il sistema $A^{(2)}x = b^{(2)}$ é equivalente a $A^{(1)}x = b^{(1)}$ ma piú semplice da risolvere;
- la matrice $A^{(2)} = L^{(1)} \cdot A^{(1)}$ risulta del tipo

$$A^{(2)} = \begin{bmatrix} a_{1,1}^{(2)} & a_{1,2}^{(2)} & \cdots & a_{1,n}^{(2)} \\ 0 & a_{2,2}^{(2)} & \cdots & a_{2,n}^{(2)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n,2}^{(2)} & \cdots & a_{n,n}^{(2)} \end{bmatrix}.$$

ovvero le componenti $a_{j,1}^{(2)}$ della prima colonna, strettamente sotto $a_{1,1}^{(2)}$, risultano nulle.

Passo k-simo.

A questo livello si suppone di aver applicato la strategia piú volte ottenendo

$$\blacksquare A^{(k)} = L^{(k-1)} \cdot A^{(k-1)} = \dots = L^{(k-1)} \cdot L^{(k-2)} A^{(k-2)} = \dots = L^{(k-1)} \cdot \dots \cdot L^{(1)} A^{(1)},$$

con

$$[A^{(k)}] = \begin{bmatrix} a_{1,1}^{(k)} & a_{1,2}^{(k)} & \dots & \dots & \dots & \dots & a_{1,n}^{(k)} \\ 0 & a_{2,2}^{(k)} & \dots & \dots & \dots & \dots & a_{2,n}^{(k)} \\ \dots & \ddots & \ddots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{k,k}^{(k)} & \dots & \dots & a_{k,n}^{(k)} \\ 0 & \dots & 0 & a_{k+1,k}^{(k)} & \dots & \dots & a_{k+1,n}^{(k)} \\ 0 & \dots & 0 & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{n,k}^{(k)} & \dots & \dots & a_{n,n}^{(k)} \end{bmatrix}.$$

ovvero la matrice $A^{(k)}$ ha componenti nulle sotto la diagonale delle prime $k - 1$ colonne.

$$\blacksquare b^{(k)} = L^{(k-1)} \cdot b^{(k-1)} = \dots = L^{(k-1)} \cdot L^{(k-2)} b^{(k-2)} = \dots = L^{(k-1)} \cdot \dots \cdot L^{(1)} b^{(1)}.$$

Supponiamo ora che l'elemento pivotale $a_{k,k}^{(k)} \neq 0$.

1. Definiamo i moltiplicatori

$$m_{i,k} = \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}, \quad i = k+1, \dots, n,$$

k. Definiamo la matrice

$$L^{(k)} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & \dots & \dots & 0 \\ \dots & \ddots & \ddots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & m_{k+1,k} & 1 & \dots & 0 \\ 0 & \dots & 0 & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & m_{n,k} & 0 & \dots & 1 \end{bmatrix}.$$

e poniamo

$$A^{(k+1)} = L^{(k)} \cdot A^{(k)}, \quad b^{(k+1)} = L^{(k)} \cdot b^{(k)}$$

Dopo questo ulteriore passo otteniamo un sistema lineare $A^{(k+1)}x = b^{(k+1)}$ avente quale soluzione la stessa di $A^{(1)}x = b^{(1)}$ ma in cui

$$A^{(k+1)} = \begin{bmatrix} a_{1,1}^{(k+1)} & a_{1,2}^{(k+1)} & \dots & \dots & \dots & \dots & a_{1,n}^{(k+1)} \\ 0 & a_{2,2}^{(k+1)} & \dots & \dots & \dots & \dots & a_{2,n}^{(k+1)} \\ \dots & \ddots & \ddots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{k,k}^{(k+1)} & \dots & \dots & a_{k,n}^{(k+1)} \\ 0 & \dots & 0 & 0 & a_{k+1,k+1}^{(k+1)} & \dots & a_{k+1,n}^{(k)} \\ 0 & \dots & 0 & 0 & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & a_{n,k+1}^{(k+1)} & \dots & a_{n,n}^{(k)} \end{bmatrix}.$$

Con questo procedimento, dopo $n - 1$ passi otteniamo un sistema $A^{(n)}x = b^{(n)}$ in cui la matrice $A^{(n)}$ è triangolare superiore con elementi diagonali $A_{k,k}^{(n)} \neq 0$, $k = 1, \dots, n$, problema che può essere facilmente risolto con l'algoritmo di sostituzione all'indietro.

Pseudocodice

Le seguenti righe di codice implementano la prima parte dell'algoritmo di eliminazione gaussiana senza pivoting, cui deve seguire la risoluzione di un sistema triangolare superiore mediante sostituzione all'indietro.

```
for k=1:n-1
    for i=k+1:n
        if a(k,k) == 0
            fprintf('Non posso completare il processo'); return;
        end
        m(i,k)=a(i,k)/a(k,k);
        for j=k+1:n
            a(i,j)=a(i,j)-m(i,k)*a(k,j);
        end
        b(i)=b(i)-m(i,k)*b(k);
    end
end
```

Notiamo che al termine del processo la matrice A coincide con L e che il codice produce errore se durante il processo qualche $a(k,k)$ è nullo.

Nota. (Complessità)

L'eliminazione gaussiana senza pivoting, relativamente ad $A \in \mathbb{R}^{n \times n}$, necessita al più $O(n^3/3)$ operazioni moltiplicative.

Definizione

Una matrice $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$ si dice

- **triangolare inferiore** se $a_{i,j} = 0$ qualora $i < j$;
- **triangolare superiore** se $a_{i,j} = 0$ qualora $i > j$.

Esempio

La matrice

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

è triangolare inferiore ($a_{i,j} = 0$ se l'indice di riga strettamente minore di quello di colonna).

Esempio

La matrice

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

è triangolare superiore ($a_{i,j} = 0$ se l'indice di riga strettamente maggiore di quello di colonna).

Supponiamo di dover risolvere un sistema lineare in forma triangolare superiore non singolare

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n = b_1 \\ a_{2,2}x_2 + \dots + a_{2,n}x_n = b_2 \\ \dots + \dots + \dots = \dots \\ a_{1,n}x_n = b_n \end{cases}$$

Evidentemente dall'ultima equazione

$$x_n^* = \frac{b_n}{a_n}$$

Supposti noti x_{k+1}^*, \dots, x_n^* , inserendoli nella k -sima equazione, per $k = 1, \dots, n-1$,

$$a_{k,k}x_k + a_{k,k+1}x_{k+1}^* + a_{k,k+2}x_{k+2}^* + \dots + a_{k,n}x_n^* = b_k$$

ricaviamo, essendo $a_{k,k} \neq 0$ in quanto A é non singolare,

$$x_k^* = \frac{b_k - (a_{k,k+1}x_{k+1}^* + a_{k,k+2}x_{k+2}^* + \dots + a_{k,n}x_n^*)}{a_{k,k}}.$$

Algoritmo. (Sostituzione all'indietro e in avanti)

Sia $b = (b_i) \in \mathbb{C}^n$ e supponiamo di dover risolvere il sistema $Ax = b$, con $A_{k,k} \neq 0$ per $k = 1, \dots, n$.

In tal caso,

- se A è triangolare superiore e invertibile, si usa l'algoritmo della **sostituzione all'indietro** ovvero

$$\begin{cases} x_n = \frac{b_n}{a_{n,n}} \\ x_k = \frac{b_k - (a_{k,k+1}x_{k+1} + \dots + a_{k,n}x_n)}{a_{k,k}}, \quad k = n-1, \dots, 1. \end{cases} \quad (3)$$

- se A è triangolare inferiore e invertibile, si usa l'algoritmo della **sostituzione in avanti** ovvero

$$\begin{cases} x_1 = b_1/a_{1,1}, \\ x_k = \frac{b_k - (a_{k,1}x_1 + \dots + a_{k,k-1}x_{k-1})}{a_{k,k}}, \quad k = 2, \dots, n. \end{cases} \quad (4)$$

Nota. (Complessità risoluzione sistemi triangolari con eliminazione gaussiana)

Si osservi che la **risoluzione di sistemi lineari triangolari** necessita approssimativamente $n^2/2$ **operazioni moltiplicative**, dove n è l'ordine della matrice.

Infatti il numero di operazioni moltiplicative, nel caso della sostituzione all'indietro risulta

$$1 + 2 + \dots + n = n(n+1)/2 \approx n^2/2.$$

Fattorizzazione $A = LU$

Definizione (Fattorizzazione LU)

Una matrice $A \in \mathbb{C}^{n \times n}$ è fattorizzabile LU se esistono

■ $L = (l_{i,j}) \in \mathbb{C}^{n \times n}$ triangolare inferiore con elementi diagonali $l_{i,i} = 1$,

■ $U = (u_{i,j}) \in \mathbb{C}^{n \times n}$ triangolare superiore,

tali che $A = LU$ (cf. [9]).

Importante. (Pivot non nulli)

Dall'applicazione della procedura precedente per la risoluzione di un sistema lineare, qualora tutti i pivot $a_{k,k}^{(k)}$ siano non nulli, otteniamo un sistema $A^{(n)}x = b^{(n)}$ equivalente all'originario e durante il calcolo introduciamo i moltiplicatori $m_{i,k}$, $k = 1, \dots, n-1$, $i = k+1, \dots, n$. Siano

■ $U = A^{(n)}$,

■ L la matrice

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ m_{2,1} & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & 0 \\ m_{n,1} & m_{n,2} & \dots & \dots & m_{n,n-1} & 1 \end{bmatrix}.$$

Allora $A = LU$ (cf. [1, p.511], [3, p.143, p.158]).

Questo risultato è rilevante, perchè tale fattorizzazione viene molto utilizzata, come vedremo in seguito con qualche esempio, indipendentemente dalla risoluzione di sistemi lineari.

Esempio

Sia, come nell'esempio iniziale,

$$A = \begin{bmatrix} 3 & -2 & -1 \\ 6 & -2 & 2 \\ -9 & 7 & 1 \end{bmatrix},$$

Alla fine dell'eliminazione gaussiana senza pivoting abbiamo determinato $U = A^{(n)}$ uguale a

$$U = \begin{bmatrix} 3 & -2 & -1 \\ 0 & 2 & 4 \\ 0 & 0 & -4 \end{bmatrix}.$$

Tenendo conto dei moltiplicatori $m_{2,1} = 2$, $m_{3,1} = -3$, $m_{2,2} = 0.5$,

$$L = \begin{bmatrix} 1 & 0 & 0 \\ m_{2,1} & 1 & 0 \\ m_{3,1} & m_{3,2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 0.5 & 1 \end{bmatrix},$$

e si verifica facilmente che $A = LU$, con L , U triangolari e della forma richiesta. In Matlab:

```
>> A=[3 -2 -1; 6 -2 2; -9 7 1];
>> U=[3 -2 -1; 0 2 4; 0 0 -4]; L=[1 0 0; 2 1 0; -3 0.5 1];
>> L*U % verifichiamo che A=L*U
ans =
     3     -2     -1
     6     -2      2
    -9      7      1
>> norm(A-L*U) % la norma di una matrice C nulla se e solo se C e' nulla
ans =
     0
>>
```

Definizione (Predominanza diagonale)

La matrice $A \in \mathbb{C}^{n \times n}$, $A = (a_{i,j})$, è a **predominanza diagonale per righe** se per ogni $i = 1, \dots, n$

$$|a_{i,i}| \geq \sum_{j=1, j \neq i}^n |a_{i,j}|$$

e per almeno un indice s si abbia

$$|a_{s,s}| > \sum_{j=1, j \neq s}^n |a_{s,j}|.$$

La matrice $A \in \mathbb{C}^{n \times n}$ è a **predominanza diagonale stretta per righe** se per ogni $i = 1, \dots, n$ risulta

$$|a_{i,i}| > \sum_{j=1, j \neq i}^n |a_{i,j}|.$$

La matrice $A \in \mathbb{C}^{n \times n}$ è a **predominanza diagonale (stretta) per colonne** se A^T è a predominanza diagonale (stretta) per righe.

Esempio

Ad esempio la matrice

$$A = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 4 & -1 \\ 0 & -0.5 & 5 \end{pmatrix}$$

è a predominanza diagonale per righe e a predominanza diagonale stretta per colonne.

Nota. (Matrici che hanno fattorizzazione $A = LU$)

Esistono classi di matrici come quelle

- 1** *a predominanza diagonale stretta per righe (cf. [10]), o*
- 2** *a predominanza diagonale stretta per colonne (cf. [10]), o*
- 3** *simmetriche e definite positive,*

per cui è sempre possibile la fattorizzazione $A = LU$.

Nota. (Fattorizzazione $PA = LU$ anche per matrici non invertibili)

Di seguito vedremo che una qualsiasi matrice invertibile possiede una fattorizzazione del tipo $PA = LU$, con P matrice di permutazione.

In realtà una matrice arbitraria possiede una fattorizzazione di questo tipo (cf. [3, p.145]).

Esempio

La matrice

$$\begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}.$$

è

- a predominanza diagonale stretta sia per righe che per colonne.
- simmetrica in quanto $A = A^T$ e definita positiva avendo per autovalori

$$\lambda_1 \approx 5.4142, \lambda_2 \approx 4.0000, \lambda_3 \approx 2.5858$$

A tal proposito ricordiamo che una matrice é **definita positiva** se e solo se ha tutti gli autovalori positivi.

n	cputime
1	$2.28081e - 04$
2	$1.14614e - 03$
4	$1.14304e - 04$
8	$4.85808e - 05$
16	$4.79512e - 04$
32	$1.15395e - 04$
64	$2.23282e - 03$
128	$2.85307e - 03$
256	$2.95903e - 03$
512	$1.71150e - 02$
1024	$4.20015e - 02$
2048	$2.45661e - 01$
4096	$1.55579e + 00$
8192	$9.08754e + 00$
16384	$6.18414e + 01$

Tabella: Secondi necessari per risolvere un sistema $Ax = b$, con una generica $A \in \mathbb{R}^{n \times n}$, mediante eliminazione gaussiana su un Mac Book Pro, con processore 2,7 GHz Intel Core i5 e 16 GB di memoria. Si ricordi che necessitano, al piú', $O(n^3/3)$ operazioni moltiplicative.

Commento

Nella precedente sezione, uno dei punti cruciali è risultato che i pivot $a_{k,k}^{(k)}$ fossero non nulli. Questo ci ha permesso di calcolare i moltiplicatori

$$m_{i,k} = \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}, \quad k = 1, \dots, n-1, \quad i = k+1, \dots, n$$

e di seguito ridurre il problema iniziale alla risoluzione di un sistema $A^{(n)}x = b^{(n)}$, con $A^{(n)}$ triangolare superiore.

Possono sorgere i seguenti problemi:

- qualche $a_{k,k}^{(k)}$ **risulta nullo** e quindi il processo precedente è inapplicabile;
- qualche $a_{k,k}^{(k)}$ **risulta molto piccolo in modulo**, questione che rende il metodo soggetto a una cattiva propagazione degli errori (cf. [1, p.516], [3, p.174]).

La soluzione risulta la seguente, detta del **pivoting (parziale)** o del **massimo pivot**.

Metodo di eliminazione gaussiana con pivoting parziale (esempio)

Mostriamo come utilizzare in un esempio l'eliminazione gaussiana con pivoting parziale (per colonne).

Di seguito **colonna** sarà sinonimo di variabile, mentre **riga** sarà sinonimo di equazione.

Esempio

Si risolva il sistema lineare

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 6x_1 - 2x_2 + 2x_3 = 6 \\ -9x_1 + 7x_2 + x_3 = -1 \end{cases}$$

mediante il metodo di eliminazione gaussiana con pivoting (parziale).

Il **termine maggiore in modulo della prima colonna** è il terzo e vale $|-9| = 9$. Scriviamo $\phi(1) = 3$, intendendo che nella prima colonna, la componente di massimo modulo è la terza. Scambiamo la prima equazione, con la terza e ricaviamo il sistema equivalente

$$\begin{cases} -9x_1 + 7x_2 + x_3 = -1 \\ 6x_1 - 2x_2 + 2x_3 = 6 \\ 3x_1 - 2x_2 - x_3 = 0 \end{cases}$$

Applichiamo ora i primi passi dell'eliminazione gaussiana relativamente a

$$\begin{cases} -9x_1 + 7x_2 + x_3 = -1 \\ 6x_1 - 2x_2 + 2x_3 = 6 \\ 3x_1 - 2x_2 - x_3 = 0 \end{cases}$$

- Moltiplichiamo la **prima riga**, ovvero equazione, per $m_{2,1} = (-6/9) = -2/3$ e la sottraiamo alla seconda, ottenendo

$$\begin{cases} -9x_1 + 7x_2 + x_3 = -1 \\ 0x_1 + (8/3)x_2 + (8/3)x_3 = 16/3 \\ 3x_1 - 2x_2 - x_3 = 0 \end{cases}$$

- Moltiplichiamo la **prima riga**, ovvero equazione, per $m_{3,1} = (-1/3) = -1/3$ e la sottraiamo alla terza, ottenendo

$$\begin{cases} -9x_1 + 7x_2 + x_3 = -1 \\ 0x_1 + (8/3)x_2 + (8/3)x_3 = 16/3 \\ 0x_1 + (1/3)x_2 - (2/3)x_3 = -1/3 \end{cases}$$

Partiamo da

$$\begin{cases} -9x_1 + 7x_2 + x_3 = -1 \\ 0x_1 + (8/3)x_2 + (8/3)x_3 = 16/3 \\ 0x_1 + (1/3)x_2 - (2/3)x_3 = -1/3 \end{cases}$$

- Osserviamo che il **termine maggiore in modulo della seconda colonna**, di indice di riga maggiore o uguale di quello di colonna, é quello della seconda riga visto che $\max(8/3, 1/3) = 8/3$.

Con le notazioni introdotte, $\phi(2) = 2$ e quindi non si richiedono scambi di equazioni (si dovrebbe scambiare la seconda equazione con la seconda equazione).

Moltiplichiamo la **seconda riga** per $m_{3,2} = (1/3)/(8/3) = 1/8$ e la sottraiamo alla terza, ottenendo

$$\begin{cases} -9x_1 + 7x_2 + x_3 = -1 \\ 0x_1 + (8/3)x_2 + (8/3)x_3 = 16/3 \\ 0x_1 + 0x_2 - x_3 = -1 \end{cases}$$

Il sistema finale, è facile da risolvere mediante sostituzione all'indietro.

- L'ultima equazione ha una sola variabile e una sola incognita e comporta che $x_3 = 1$.
- Inseriamo questo risultato nella penultima equazione e otteniamo

$$(8/3)x_2 + (8/3) \cdot 1 = 16/3$$

da cui $x_2 = 1$.

- Inseriamo questi risultati nella prima equazione e otteniamo

$$-9x_1 + 7 \cdot 1 + 1 = -1$$

da cui $x_1 = 1$.

Il processo di eliminazione gaussiana con pivoting, è lo stesso dell'eliminazione gaussiana, però in aggiunta dobbiamo scambiare ad ogni passo certe equazioni.

Matricialmente questo corrisponde a scambiare (matematicamente si dice **permutare**) certe righe della matrice, e per fare questo risulta fondamentale definire la matrice di permutazione.

Definizione (Matrice di permutazione)

Una matrice $P = (p_{i,j})$ si dice di **permutazione** se

- 1 le sue componenti sono 0 o 1,
- 2 ogni riga e colonna contengono una sola componente uguale a 1.

Supponiamo sia $P \in \mathbb{R}^{n \times n}$ una matrice di permutazione.

- Sia $x \in \mathbb{R}^n$ un vettore colonna e $y = Px$. Allora

$$P_{i,j} = 1 \Rightarrow y_i = x_j.$$

- Sia $A \in \mathbb{R}^n$ una matrice e $C = PA$. Allora

$$P_{i,j} = 1 \Rightarrow C_{i,:} = A_{j,:}$$

ovvero la i -sima riga di C è la j -sima riga di A .

Esempio

Consideriamo la matrice di permutazione

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

e il vettore $x = (x_1, x_2, x_3)^T$. Allora

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \cdot x_1 + 1 \cdot x_2 + 0 \cdot x_3 \\ 1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 \\ 0 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \\ x_3 \end{pmatrix}.$$

Piú in generale, per $A = (a_{i,j})$,

$$PA = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = \begin{pmatrix} a_{2,1} & a_{2,2} & a_{2,3} \\ a_{1,1} & a_{1,2} & a_{1,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}.$$

Esempio

La matrice $P_{i,j}$ che é come la matrice identica eccetto per le righe i, j che sono tutte nulle eccetto per le componenti $P_{i,j} = P_{j,i} = 1$, é tale che $C = PA$ coincide con la matrice A eccetto per il fatto che le righe i, j sono scambiate.

Cosí

$$\begin{aligned}
 P_{2,3}A &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{pmatrix} \\
 &= \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{pmatrix}
 \end{aligned}$$

ha le stesse righe di A , ma con la seconda permutata con la terza.

Illustriamo l'importanza delle matrici di permutazione, nel processo di eliminazione gaussiana con pivoting parziale (per colonne).

Abbiamo visto che il precedente sistema lineare

$$\begin{cases} 3x_1 - 2x_2 - x_3 = 0 \\ 6x_1 - 2x_2 + 2x_3 = 6 \\ -9x_1 + 7x_2 + x_3 = -1 \end{cases}$$

è equivalente al sistema lineare $A^{(1)}x = b^{(1)}$ con

$$A^{(1)} = \begin{bmatrix} 3 & -2 & -1 \\ 6 & -2 & 2 \\ -9 & 7 & 1 \end{bmatrix} \quad b^{(1)} = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix}$$

- Osserviamo che l'elemento di massimo modulo della prima colonna è $a_{3,1}^{(1)}$.
- Scambiamo quindi la prima riga con la terza.
- La matrice di permutazione usata risulta quindi

$$P^{(1)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

- Consideriamo il sistema equivalente $P^{(1)}A^{(1)}x = P^{(1)}b^{(1)}$ in cui

$$P^{(1)}A^{(1)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & -2 & -1 \\ 6 & -2 & 2 \\ -9 & 7 & 1 \end{bmatrix} = \begin{bmatrix} -9 & 7 & 1 \\ 6 & -2 & 2 \\ 3 & -2 & -1 \end{bmatrix}$$

$$P^{(1)}b^{(1)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 6 \\ 0 \end{bmatrix}$$

- Osserviamo che i moltiplicatori in $P^{(1)}A^{(1)}$, relativamente alla prima colonna sono $m_{2,1} = 6/(-9) = -2/3$ e $m_{3,1} = 3/(-9) = -1/3$.

- Introduciamo la matrice $L^{(1)}$ basata sui moltiplicatori $m_{2,1} = 6/(-9) = -2/3$ e $m_{3,1} = 3/(-9) = -1/3$ ovvero

$$L^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ -m_{2,1} & 1 & 0 \\ -m_{3,1} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & 0 & 1 \end{bmatrix}$$

- Introduciamo il sistema $A^{(2)}x = b^{(2)}$ equivalente a $A^{(1)}x = b^{(1)}$ dove

$$\begin{aligned} A^{(2)} &= L^{(1)}P^{(1)}A^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & 0 & 1 \end{bmatrix} \begin{bmatrix} -9 & 7 & 1 \\ 6 & -2 & 2 \\ 3 & -2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -9 & 7 & 1 \\ 0 & 8/3 & 8/3 \\ 0 & 1/3 & -2/3 \end{bmatrix} \end{aligned}$$

$$b^{(2)} = L^{(1)}P^{(1)}b^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 16/3 \\ -1/3 \end{bmatrix}$$

Il sistema lineare precedente, dopo il primo passo di l'eliminazione gaussiana (con pivoting), é risultato equivalente a

$$\begin{cases} -9x_1 + 7x_2 + x_3 = -1 \\ 0x_1 + (8/3)x_2 + (8/3)x_3 = 16/3 \\ 0x_1 + (1/3)x_2 - (2/3)x_3 = -1/3 \end{cases}$$

Alternativamente, per i calcoli effettuati in versione matriciale, é equivalente al sistema lineare $A^{(2)}x = b^{(2)}$ definito come

$$A^{(2)}x := \begin{bmatrix} -9 & 7 & 1 \\ 0 & 8/3 & 8/3 \\ 0 & 1/3 & -2/3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 16/3 \\ -1/3 \end{bmatrix} := b^{(2)}$$

Tanto la formulazione matriciale che quella per equazioni hanno portato quindi allo stesso sistema lineare.

- Osserviamo che l'elemento di massimo modulo della seconda colonna, sulla diagonale o al di sotto della stessa é $a_{2,2}^{(2)}$.
- In virtù di quanto visto, non serve fare scambi.
- La matrice di permutazione usata risulta quindi

$$p^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Consideriamo il sistema equivalente $P^{(2)}A^{(2)}x = P^{(2)}b^{(1)}$ in cui

$$P^{(2)}A^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -9 & 7 & 1 \\ 0 & 8/3 & 8/3 \\ 0 & 1/3 & -2/3 \end{bmatrix} = \begin{bmatrix} -9 & 7 & 1 \\ 0 & 8/3 & 8/3 \\ 0 & 1/3 & -2/3 \end{bmatrix}$$

$$P^{(2)}b^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 16/3 \\ -1/3 \end{bmatrix} = \begin{bmatrix} -1 \\ 16/3 \\ -1/3 \end{bmatrix}$$

- Osserviamo che il moltiplicatore in $P^{(2)}A^{(2)}$, relativamente alla seconda colonna é $m_{3,2} = (1/3)/(8/3) = 1/8$.

- Introduciamo la matrice $L^{(2)}$ basata sul moltiplicatore $m_{3,2} = 1/8$ ovvero

$$L^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{3,2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/8 & 1 \end{bmatrix}$$

- Definiamo il sistema $A^{(3)}x = b^{(3)}$ equivalente a $A^{(2)}x = b^{(2)}$ e quindi a $A^{(1)}x = b^{(1)}$ dove

$$\begin{aligned} A^{(3)} &= L^{(2)}P^{(2)}A^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/8 & 1 \end{bmatrix} \begin{bmatrix} -9 & 7 & 1 \\ 0 & 8/3 & 8/3 \\ 0 & 1/3 & -2/3 \end{bmatrix} \\ &= \begin{bmatrix} -9 & 7 & 1 \\ 0 & 8/3 & 8/3 \\ 0 & 0 & -1 \end{bmatrix} \end{aligned}$$

$$b^{(3)} = L^{(2)}P^{(2)}b^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/8 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 16/3 \\ -1/3 \end{bmatrix} = \begin{bmatrix} -1 \\ 16/3 \\ -1 \end{bmatrix}$$

- Utilizzando la metodologia per singole equazioni abbiamo ottenuto al termine del processo di eliminazione gaussiana lo studio del sistema triangolare

$$\begin{cases} -9x_1 + 7x_2 + x_3 = -1 \\ 0x_1 + (8/3)x_2 + (8/3)x_3 = 16/3 \\ 0x_1 + 0x_2 - x_3 = -1 \end{cases} \quad (5)$$

equivalente a quello iniziale.

- Matricialmente, abbiamo ricavato similmente

$$A^{(3)}_X := \begin{bmatrix} -9 & 7 & 1 \\ 0 & 8/3 & 8/3 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 16/3 \\ -1 \end{bmatrix} := b^{(3)}$$

che ovviamente é equivalente al sistema (5).

- Il sistema finale può essere risolto con la sostituzione all'indietro.

Nota.

Il sistema lineare precedente può essere riscritto direttamente in forma matriciale, utilizzando le operazioni di moltiplicazione per matrici di permutazione nonché l'utilizzo passo passo dell'eliminazione gaussiana dal punto di vista matriciale.

Passiamo alla fattorizzazione LU della matrice

$$A = \begin{bmatrix} 3 & -2 & -1 \\ 6 & -2 & 2 \\ -9 & 7 & 1 \end{bmatrix}$$

utilizzata nell'esempio precedente per risolvere $Ax = b$, con $b = (0, 6, -1)^T$. Per quanto visto, analizzando il sistema finale del processo, $U = A^{(3)}$ e dunque

$$U = \begin{bmatrix} -9 & 7 & 1 \\ 0 & 8/3 & 8/3 \\ 0 & 0 & -1 \end{bmatrix}.$$

- 1 Ricordando i moltiplicatori $m_{2,1} = -2/3$, $m_{3,1} = -1/3$, $m_{3,2} = 1/8$, abbiamo

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -2/3 & 1 & 0 \\ -1/3 & 1/8 & 1 \end{bmatrix}$$

- 2 visto che le matrici di permutazione sono risultate

$$P^{(1)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

mentre nel passo relativo alla seconda colonna, visto che non abbiamo effettuato scambi

$$P^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(ovvero la matrice identica),

si può mostrare, con qualche ragionamento non banale, che per

$P = P^{(2)} \cdot P^{(1)} = I \cdot P_1 = P_1$ si ha

$$PA = LU.$$

Avendo calcolato la fattorizzazione $PA = LU$, verifichiamo i risultati in Matlab.

```
>> P=[0 0 1; 0 1 0; 1 0 0];  
>> A=[3 -2 -1; 6 -2 2; -9 7 1];  
>> PA=P*A;  
>> L=[1 0 0; -2/3 1 0; -1/3 1/8 1];  
>> U=[-9 7 1; 0 8/3 8/3; 0 0 -1];  
>> LU=L*U;  
>> norm(PA-LU,2)  
ans =  
    4.9651e-16  
>>
```

Si afferma che $\|P * A - L * U\|_2 \approx 4.9651e - 16$ ovvero che anche numericamente $P * A = L * U$ (la distanza euclidea matriciale tra PA e LU é molto piccola).

Metodo di eliminazione gaussiana con pivoting parziale (fattorizzazione $PA = LU$)

Importante. (Fattorizzazione $PA=LU$)

Dall'applicazione della procedura dell'eliminazione gaussiana con pivoting per la risoluzione di un sistema lineare, abbiamo ottenuto un sistema $A^{(n)}x = b^{(n)}$ equivalente all'originario. Poniamo

- P matrice di **permutazione** ottenuta come $P = P^{(n-1)} \cdot \dots \cdot P^{(1)}$ dove $P^{(k)}$ è la matrice di permutazione che la scambia la riga i_k con la k -sima, ovvero la matrice identica eccetto per le righe i_k, k in cui

$$P_{i_k, i_k} = 0, P_{i_k, k} = 1, P_{k, i_k} = 1, P_{k, k} = 0;$$

- $U = A^{(n)}$ di forma **triangolare superiore**,
- $L = (l_{i,j})$ la matrice

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ m_{2,1} & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ m_{n,1} & m_{n,2} & \dots & \dots & m_{n,n-1} & 1 \end{bmatrix}.$$

di forma **triangolare inferiore**, con componenti diagonali $l_{k,k} = 1, k = 1, \dots, n$.

Allora $PA = LU$ (cf. [1, p.511], [3, p.143, p.158]).

Questo risultato è rilevante, perchè tale fattorizzazione viene molto utilizzata indipendentemente dalla risoluzione di sistemi lineari.

Nota. (storica)

L'algoritmo di eliminazione gaussiana è stato scoperto in Cina circa 2000 anni fa, con più recenti contributi in lavori di Lagrange e Jacobi. La descrizione moderna è stata introdotta solo nel 1930. [7, p.607].

Nota. (Importante)

1. Visto che risolvere ognuno dei sistemi triangolari costa circa $n^2/2$ mentre la parte precedente necessita di circa $n^3/3$ operazioni moltiplicative, deduciamo che risolvere il sistema $Ax = b$ con l'eliminazione gaussiana necessita di circa $n^3/3$ operazioni moltiplicative, visto che

$$O(n^3/3) + 2 \cdot O(n^2/2) = O(n^3/3).$$

2. La complessità della fattorizzazione $PA = LU$ è essenzialmente la stessa della eliminazione gaussiana, ovvero $\frac{n^3}{3}$.

3. In generale l'algoritmo di eliminazione gaussiana per la risoluzione di un sistema lineare $Ax = b$ con A invertibile, può essere implementato come segue

- 1 Calcola la fattorizzazione $PA = LU$. Osserviamo che essendo P invertibile (si dimostra che $|\det(P)| = 1$)

$$Ax = b \Leftrightarrow PAx = Pb \Leftrightarrow LUx = Pb \Leftrightarrow Ly = \tilde{b} \text{ con } Ux = y, \tilde{b} = Pb.$$

- 2 Calcola $\tilde{b} = Pb$.
- 3 Calcoli la soluzione y^* del sistema triangolare inferiore $Ly = \tilde{b}$ con la sostituzione in avanti.
- 4 Calcoli la soluzione x^* del sistema triangolare superiore $Ux = y^*$ con la sostituzione all'indietro. Tale x^* è quindi soluzione di $Ax = b$.

La complessità dell'algoritmo così implementato è di $\frac{n^3}{3}$ operazioni moltiplicative, ma offre il vantaggio di avere a disposizione la fattorizzazione $PA = LU$.

Nota. (Decomposizione di Cholesky)

Se la A matrice è *simmetrica e definita positiva*, ovvero

- $A = A^T$,
- tutti gli autovalori di A sono positivi,

è più conveniente utilizzare la *decomposizione di Cholesky* [1, p.524], in cui

$$A = LL^T$$

con $L = (l_{i,j})$ triangolare inferiore con elementi diagonali $l_{k,k} > 0$, $k = 1, \dots, n$.

La determinazione di questa fattorizzazione richiede approssimativamente $\frac{n^3}{6}$ operazioni moltiplicative.

Di conseguenza, per calcolare la soluzione di $Ax = b$, basta risolvere i due sistemi triangolari $Ly = b$, $L^T x = y$. Quindi in generale la soluzione del sistema lineare si può ottenere in

$$O\left(\frac{n^3}{6}\right) + O\left(\frac{n^2}{2}\right) + O\left(\frac{n^2}{2}\right) = O\left(\frac{n^3}{6}\right)$$

operazioni moltiplicative.

Esempio

La matrice simmetrica

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

è definita positiva, in quanto l'autovalore minimo è $\lambda_{\min} = 0.271 \dots > 0$.

Di conseguenza possiede una decomposizione di Cholesky $A = LL^T$. In questo caso particolare risulta:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

In Matlab la routine `chol` calcola la trasposta di tale matrice L . Vediamo in effetti che i risultati sono corretti.

```
>> A=gallery('minij',5);
```

```
>> LT=chol(A); L=LT';
```

```
>> L
```

```
L =
```

```
    1    0    0    0    0
    1    1    0    0    0
    1    1    1    0    0
    1    1    1    1    0
    1    1    1    1    1
```

```
>> norm(A-L*L',inf)
```

```
ans =
```

```
    0
```


Motivazione. (Metodi diretti e iterativi)

- I metodi come l'eliminazione gaussiana calcolano in aritmetica esatta la soluzione x^* di $Ax = b$, $A \in \mathbb{C}^{n \times n}$ in un numero finito di passi, sono detti metodi **diretti**.
- Osserviamo che
 - (a) la matrice n è di **grandi dimensioni** e quindi ci sono sia problemi di immagazzinamento dati che di complessità computazionale (ovvero $n^3/3$ può essere proibitivo in termini di costo);
 - (b) la matrice A può possedere **molti zeri**, che si verifica spesso non aiutare a migliorare la performance del metodo diretto.
- Una alternativa consiste nei **metodi iterativi** che tipicamente non è detto determinino la soluzione esatta, ma possono approssimarla **arbitrariamente bene**. Usualmente **ogni iterazione** calcola un prodotto matrice-vettore, e quindi ha complessità di circa **n^2 operazioni moltiplicative**.
- Se i metodi iterativi forniscono l'approssimazione della soluzione richiesta in un **$k \ll n/3$** iterazione, essendo $kn^2 \ll n^3/3$ risultano **convenienti** rispetto alla generale fattorizzazione LU.

Sia $A \in \mathbb{C}^{n \times n}$ e definiamo lo *splitting*

$$A = D - E - F$$

dove

- D è una matrice diagonale, non singolare,
- E è triangolare inferiore con elementi diagonali nulli,
- F è triangolare superiore con elementi diagonali nulli.

Esempio

Sia A la matrice

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 2 \\ 8 & 9 & 1 & 2 \\ 3 & 4 & 5 & 1 \end{bmatrix},$$

allora può essere descritta in termini di D , E , F con

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, E = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -5 & 0 & 0 & 0 \\ -8 & -9 & 0 & 0 \\ -3 & -4 & -5 & 0 \end{bmatrix}, F = \begin{bmatrix} 0 & -2 & -3 & -4 \\ 0 & 0 & -7 & -2 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Se $A \in \mathbb{C}^{n \times n}$ è una matrice quadrata, risulta importante anche lo *splitting*

$$A = P - N, \quad \det(P) \neq 0. \quad (6)$$

Allora

$$Ax = b \Leftrightarrow (P - N)x = b \Leftrightarrow Px - Nx = b \Leftrightarrow Px = Nx + b$$

da cui, per l'invertibilità di P , moltiplicando ambo i membri per P^{-1} , ricaviamo

$$x = P^{-1}Px = P^{-1}(Nx + b) = \underbrace{P^{-1}Nx}_B + \underbrace{P^{-1}b}_c = Bx + c := \phi(x). \quad (7)$$

Quindi, posti $B := P^{-1}N$, $c := P^{-1}b$, la soluzione x^* di $Ax = b$ risolve pure $x = Bx + c$ (e viceversa).

Conseguentemente, vista la struttura di punto fisso $x = \phi(x)$ con $\phi(x) = Bx + c$, si considera il **metodo iterativo (stazionario)**

$$x^{(k+1)} = \phi(x^{(k)}) = Bx^{(k)} + c. \quad (8)$$

(Metodi di Jacobi e Gauss-Seidel)

Sia $A = D - E - F$ con

- 1 D matrice diagonale,
- 2 E triangolare inferiore,
- 3 F triangolare superiore.

Allora

- il **metodo di Jacobi** corrisponde a scegliere, se D è invertibile,

$$P = D, \quad N = E + F$$

- il **metodo di Gauss-Seidel** corrisponde a scegliere, se D è invertibile,

$$P = D - E, \quad N = F$$

Nota.

Il metodo di Jacobi fu pubblicato dall'autore nel 1845 in un giornale di astronomia, mentre quello che conosciamo come metodo di Gauss-Seidel fu il risultato di una collaborazione di Seidel con Jacobi nel 1874. Gauss aveva già descritto un metodo analogo nel 1845 [4, p.466].

Commento (Osservazioni sul metodo di Jacobi e Gauss-Seidel)

Se $A = (a_{i,j})_{i,j=1,\dots,n} \in \mathbb{C}^{n \times n}$, $\det(A) \neq 0$ ed $Ax = b$, visto che $(Ax)_i = \sum_{j=1}^n a_{i,j}x_j$, allora

$$\sum_{j=1}^n a_{i,j}x_j = b_i, \quad i = 1, \dots, n$$

e quindi, per $i = 1, \dots, n$, evidenziando il termine i -simo,

$$\sum_{j=1}^{i-1} a_{i,j}x_j + a_{i,i}x_i + \sum_{j=i+1}^n a_{i,j}x_j = \sum_{j=1}^n a_{i,j}x_j = b_i,$$

da cui portando a secondo membro $\sum_{j=1}^{i-1} a_{i,j}x_j + \sum_{j=i+1}^n a_{i,j}x_j$

$$a_{i,i}x_i = b_i - \sum_{j=1}^{i-1} a_{i,j}x_j - \sum_{j=i+1}^n a_{i,j}x_j,$$

e quindi nel caso $a_{i,i} \neq 0$ per $i = 1, \dots, n$

$$x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{i,j}x_j - \sum_{j=i+1}^n a_{i,j}x_j}{a_{i,i}}, \quad i = 1, \dots, n.$$

Da

$$x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{i,j}x_j - \sum_{j=i+1}^n a_{i,j}x_j}{a_{i,i}}, \quad i = 1, \dots, n.$$

il metodo di **Jacobi** si ricava ponendo nelle occorrenze del membro di

- sinistra $x_i^{(k+1)}$ al posto di x_i ,
- destra $x_j^{(k)}$ al posto di ogni occorrenza di x_j (con $j = 1, \dots, n$),

ricavando

$$x_i^{(k+1)} = (b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)})/a_{ii}, \quad i = 1, \dots, n, \quad (9)$$

Nota.

Il metodo ottenuto e' lo stesso descritto matricialmente come

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

con $P = D$ e $N = E + F$.

Esempio (Metodo di Jacobi)

Applicando

$$x_i^{(k+1)} = \frac{1}{a_{i,i}} \left(b_i - \left(\sum_{j<i} a_{i,j} x_j^{(k)} + \sum_{j>i} a_{i,j} x_j^{(k)} \right) \right), \quad i = 1, \dots, n$$

al caso di $A \in \mathbb{C}^{3 \times 3}$, otteniamo, supponendo disponibile $x^{(k)}$:

$$1 \quad x_1^{(k+1)} = \frac{1}{a_{1,1}} (b_1 - a_{1,2} x_2^{(k)} - a_{1,3} x_3^{(k)});$$

$$2 \quad x_2^{(k+1)} = \frac{1}{a_{2,2}} (b_2 - a_{2,1} x_1^{(k)} - a_{2,3} x_3^{(k)});$$

$$3 \quad x_3^{(k+1)} = \frac{1}{a_{3,3}} (b_3 - a_{3,1} x_1^{(k)} - a_{3,2} x_2^{(k)}).$$

Si nota che

- dopo aver calcolato $x_1^{(k+1)}$ tale nuova approssimazione della prima componente x_1^* della soluzione x^* non é sfruttata nel determinare $x_2^{(k+1)}$;
- dopo aver calcolato $x_2^{(k+1)}$ tale nuova approssimazione della prima componente x_2^* della soluzione x^* non é sfruttata nel determinare $x_3^{(k+1)}$.

Ogni $x_j^{(k+1)}$, $j = 1, 2, 3$ é calcolato direttamente dall'iterazione precedente $x^{(k)} = (x_i^{(k)})_i$, senza richiedere la conoscenza di qualche $x_i^{(k+1)}$ con $i = 1, \dots, j-1$.

Da

$$x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{i,j}x_j - \sum_{j=i+1}^n a_{i,j}x_j}{a_{i,i}}, \quad i = 1, \dots, n.$$

il metodo di *Gauss-Seidel* si ricava ponendo nelle occorrenze del membro di

- sinistra $x_i^{(k+1)}$ al posto di x_i ,
- destra nella prima sommatoria $x_i^{(k+1)}$ al posto di x_i e nella seconda sommatoria $x_i^{(k)}$

ricavando

$$x_i^{(k+1)} = (b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)})/a_{ii}, \quad i = 1, \dots, n, \quad (10)$$

Nota.

Il metodo ottenuto e' lo stesso descritto matricialmente come

$$x^{(k+1)} = P^{-1}Nx^{(k)} + P^{-1}b, \quad P = D - E, \quad N = F.$$

Una sua generalizzazione porta ai metodi SOR (cf. [12]).

Esempio (Metodo di Gauss-Seidel)

Applicando

$$x_i^{(k+1)} = \frac{1}{a_{i,i}} \left(b_i - \left(\sum_{j<i} a_{i,j} x_j^{(k+1)} + \sum_{j>i} a_{i,j} x_j^{(k)} \right) \right), \quad i = 1, \dots, n$$

al caso di $A \in \mathbb{C}^{3 \times 3}$, otteniamo:

- 1 $x_1^{(k+1)} = \frac{1}{a_{1,1}} (b_1 - a_{1,2} x_2^{(k)} - a_{1,3} x_3^{(k)});$
- 2 $x_2^{(k+1)} = \frac{1}{a_{2,2}} (b_2 - a_{2,1} x_1^{(k+1)} - a_{2,3} x_3^{(k)});$
- 3 $x_3^{(k+1)} = \frac{1}{a_{3,3}} (b_3 - a_{3,1} x_1^{(k+1)} - a_{3,2} x_2^{(k+1)}).$

Si nota che

- dopo aver calcolato $x_1^{(k+1)}$ tale nuova approssimazione della prima componente x_1^* della soluzione x^* è sfruttata nel determinare $x_2^{(k+1)}$;
- dopo aver calcolato $x_2^{(k+1)}$ tale nuova approssimazione della prima componente x_2^* della soluzione x^* è sfruttata nel determinare $x_3^{(k+1)}$.

Teorema (Convergenza globale dei metodi iterativi)

Siano

- $B \in \mathbb{C}^{n \times n}$,
- si supponga che la soluzione x^* di $Ax = b$ sia l'unica soluzione di $x = Bx + c$.
- sia $\rho(B) = \max_{k=1, \dots, n} |\lambda_k|$ il raggio spettrale di B (dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di B).

Allora qualsiasi sia $x^{(0)} \in \mathbb{C}^n$, la successione $\{x^{(k)}\}$ definita da

$$x^{(k+1)} = Bx^{(k)} + c, \quad k = 0, 1, 2, \dots$$

converge a x^* **se e solo se** $\rho(B) < 1$.

Nota.

- 1 Si noti che in questo asserto compare **un se e solo se**.
- 2 Si noti compare il **raggio spettrale di B** (e non di A).
- 3 Si noti che richiede il calcolo dell'autovalore di modulo massimo di B , un problema potenzialmente complicato. Di solito, noto la struttura della matrice A ci sono dei teoremi che offrono risultati sul raggio spettrale di B .

A partire dai precedenti teoremi si è potuto dimostrare (non facile!)

Teorema (Convergenza Jacobi per matrici a predominanza diagonale stretta)

Se la matrice A è a **predominanza diagonale (stretta) per righe**, allora il metodo di **Jacobi** converge.

Teorema (Convergenza Gauss-Seidel per matrici a predominanza diagonale stretta)

Se la matrice A è a **predominanza diagonale (stretta) per righe**, allora il metodo di **Gauss-Seidel** converge.

Teorema (Convergenza G.-S. per matrici simmetriche e definite positive)

Se la matrice A è **simmetrica e definita positiva** allora il metodo di **Gauss-Seidel** converge.

Nota. (Grandezza raggio spettrale)

In generale, con stime quantitative, si vede che sono da preferire metodi con raggio spettrale $\rho(B)$ il più piccolo possibile.

Esempio (Matrice a predominanza diagonale stretta)

Consideriamo il sistema lineare $Ax = b$ dove

$$A = \begin{bmatrix} 11 & -5 & -5 \\ 5 & 12 & 6 \\ 6 & -4 & 11 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 23 \\ 13 \end{bmatrix},$$

che ha soluzione $x^* = [1, 1, 1]^T$ (cf. [3, p.252]). Si dica se convergono i metodi di Jacobi e Gauss-Seidel.

- La matrice A è a predominanza diagonale in senso stretto e quindi tanto il metodo di Jacobi, quanto quello di Gauss-Seidel, risultano convergenti. In effetti $\rho(B_J) \approx 0.79$, $\rho(B_{GS}) \approx 0.83$ e questo si dimostra implicare che asintoticamente il metodo di Jacobi converge più rapidamente di quello di Gauss-Seidel.
- Gli errori assoluti $\|x^{(k)} - x^*\|_\infty$ compiuti rispettivamente dal metodo di Jacobi e Gauss-Seidel sono esposti nella tabella che segue, da cui si evince la convergenza dei metodi.

n	E_J	E_{GS}
10	$1.6e - 01$	$2.2e - 01$
30	$1.5e - 03$	$6.2e - 03$
50	$1.4e - 05$	$1.7e - 04$
70	$1.3e - 07$	$4.8e - 06$
90	$1.2e - 09$	$1.3e - 07$
110	$1.1e - 11$	$3.8e - 09$
130	$1.0e - 13$	$1.1e - 10$
150	$1.1e - 15$	$2.9e - 12$
170	—	$8.2e - 14$
190	—	$2.2e - 15$

Tabella: Errori assoluti $\|x^{(k)} - x^*\|_\infty$ compiuti dalle k -sime iterazioni di Jacobi e Gauss-Seidel, nel risolvere un certo problema $Ax = b$ con A a predominanza diagonale stretta, con vettore iniziale $x_0 = (0, 0, 0)^T \in \mathbb{R}^3$.

Esempio (Matrice definita positiva)

Si studi la convergenza dei metodi di Jacobi e Gauss-Seidel, relativamente al sistema lineare $Ax = b$ dove

$$A = \left[\begin{array}{ccc|ccc|ccc} 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 4 & 0 & 0 & -1 & 0 & 0 & 0 \\ \hline -1 & 0 & 0 & 4 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 4 & 0 & 0 & -1 \\ \hline 0 & 0 & 0 & -1 & 0 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 \end{array} \right], \quad b = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \\ 0 \\ 1 \\ 2 \\ 1 \\ 2 \end{bmatrix},$$

che ha soluzione $x^* = [1, \dots, 1]^T \in \mathbb{R}^9$.

La matrice A è un esempio di matrice di Poisson ed è definita positiva, in quanto il suo autovalore più piccolo vale

$$\lambda_{\min} \approx 1.171572875253810e + 00 > 0.$$

Le linee mostrano come $A \in \mathbb{R}^{9 \times 9}$ segua *pattern* particolari con *blocchi* pari a

$$B = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix}, \quad -I = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad 0_{3 \times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

- Gli errori assoluti $\|x^{(k)} - x^*\|_\infty$ compiuti rispettivamente dal metodo di Jacobi e Gauss-Seidel partendo da $x_0 = [0, \dots, 0]^T \in \mathbb{R}^9$. Entrambi i metodi convergono, e per quanto concerne il metodo di Gauss-Seidel, ciò era garantito dal teorema di convergenza, in quanto A è simmetrica e definita positiva.
- Si vede che $\rho(B_J) \approx 0.7071$, mentre $\rho(B_{GS}) \approx 0.5$ e quindi ci si aspetta che asintoticamente il metodo di Gauss-Seidel converga in meno iterazioni di quello di Jacobi.

n	E_J	E_{GS}
10	$4.7e - 02$	$1.8e - 03$
30	$4.6e - 05$	$1.7e - 09$
50	$4.5e - 08$	$1.8e - 15$
70	$4.4e - 11$	—
90	$4.3e - 14$	—

Tabella: Errori assoluti $\|x^{(k)} - x^*\|_\infty$ compiuti dalle k -sime iterazioni di Jacobi e Gauss-Seidel, nel risolvere un certo problema $Ax = b$ con A matrice di Poisson, con vettore iniziale $x_0 = (0, 0, 0, 0, 0, 0, 0, 0, 0)^T \in \mathbb{R}^9$.

Esempio (Mancata convergenza dei metodi di Jacobi e Gauss-Seidel)

Consideriamo il sistema $Ax = b$ con

$$A = \begin{bmatrix} 3 & -2 & -1 \\ 6 & -2 & 2 \\ -9 & 7 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix},$$

avente soluzione $x^* = (1, 1, 1)^T$. Si verifichi la mancata convergenza globale dei metodi di Jacobi e Gauss-Seidel partendo da $x^{(0)} = (1 - \epsilon, 1 - \epsilon, 1 - \epsilon)^T$ con $\epsilon = 10^{-14}$.

- Utilizziamo i metodi di Jacobi e Gauss-Seidel con dato iniziale

$$x^{(0)} = (1 - 10^{-14}, 1 - 10^{-14}, 1 - 10^{-14})^T$$

e quindi molto prossimo al vettore soluzione x^* in quanto $\|x^* - x^{(0)}\|_\infty = 10^{-14}$.

- Entrambi i metodi citati, pur partendo da un tal $x^{(0)}$, non convergono, come si può vedere dalla relativa tabella. In effetti, si verifica che $1 < \rho(B_J) \approx 8.66$, $1 < \rho(B_{GS}) \approx 9.62$, e quindi entrambi i metodi iterativi non sono globalmente convergenti.

n	E_J	E_{GS}
10	$2.6e - 12$	$1.4e - 05$
30	$6.7e - 09$	$6.3e + 14$
50	$1.8e - 05$	$2.9e + 34$
70	$4.5e - 02$	$1.4e + 54$
90	$1.2e + 02$	$6.3e + 73$
110	$2.9e + 05$	$2.9e + 93$
130	$7.3e + 08$	$1.4e + 113$
150	$1.8e + 12$	$6.3e + 132$
170	$4.4e + 15$	$2.9e + 152$

Tabella: Errori assoluti $\|x^{(k)} - x^*\|_\infty$ compiuti dalle n -sime iterazioni di Jacobi e Gauss-Seidel, nel risolvere un certo problema $Ax = b$ con A matrice utilizzata come esempio per l'eliminazione gaussiana, con vettore iniziale $x^{(0)} = (1 - 10^{-14}, 1 - 10^{-14}, 1 - 10^{-14})^T \in \mathbb{R}^3$.

Sia $\epsilon > 0$ una tolleranza fissata dall'utente. Il metodo iterativo

$$x^{(k+1)} = Bx^{(k)} + c$$

viene arrestato utilizzando

- il **criterio dello step**, ovvero si interrompe il processo quando

$$\|x^{(k+1)} - x^{(k)}\| \leq \epsilon,$$

- **criterio del residuo**, ovvero si interrompe il processo quando

$$\|b - Ax^{(k)}\| \leq \epsilon,$$

- **criterio del residuo relativo**, ovvero si interrompe il processo quando

$$\frac{\|b - Ax^{(k)}\|}{\|b\|} \leq \epsilon, \quad b \neq 0_{\mathbb{C}^n}$$

- [1] K. Atkinson, *Introduction to Numerical Analysis*, Wiley, 1989.
- [2] K. Atkinson, W. Han, *Elementary Numerical Analysis*, Wiley, 2003.
- [3] D. Bini, M. Capovani, O. Menchi, *Metodi numerici per l'algebra lineare*, Zanichelli, 1988.
- [4] J.F. Epperson, *Introduzione all'analisi numerica. Teoria, metodi, algoritmi.*, McGraw-Hill, 2009.
- [5] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd edition, John Hopkins University Press, (1996).
- [6] C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM Frontiers in Applied Mathematics. SIAM, Philadelphia, (1995).
- [7] L.N. Trefethen, *Numerical Analysis*, Princeton Companion to Mathematics, 2008.
- [8] [Wikipedia, Decomposizione di Cholesky.](#)
- [9] [Wikipedia, Decomposizione LU.](#)
- [10] [Wikipedia, Matrice a diagonale dominante.](#)

- [11] Wikipedia, Matrice definita positiva.
- [12] Wikipedia, SuccessiveOverRelaxation.
- [13] Wikipedia, Teorema di Binet.
- [14] Wikipedia, Teorema di Laplace.