



RISOLUZIONE DI SISTEMI LINEARI: METODI DIRETTI (Il Metodo di Eliminazione di Gauss)

Nella lezione precedente abbiamo visto come affrontare la risoluzione di sistemi lineari triangolari (inferiori e superiori) tramite gli algoritmi delle sostituzioni (in avanti e all'indietro). In questa lezione consideriamo il caso più generale, ovvero il caso in cui la matrice del sistema sia una matrice di dimensioni arbitrarie. Aniché utilizzare da subito le formule di sostituzione all'indietro o in avanti, utilizzeremo il metodo di eliminazione di Gauss (MEG).

Abbiamo osservato nella lezione precedente che l'obiettivo del MEG è quello di ridurre il sistema a un sistema triangolare superiore equivalente. Ciò significa che non abbiamo cambiato la soluzione del sistema. Una domanda preliminare è: quali sono le operazioni ammissibili sulla matrice affinché il sistema non venga modificato? In altre parole, quali operazioni sulle equazioni di un sistema garantiscono che il sistema finale non abbia soluzioni differenti da quelle di partenza? Si ottengono sistemi equivalenti se

- 1- Si aggiunge ad una riga una combinazione lineare delle altre. Ovvero, se aggiungiamo combinazioni lineari di altre righe non cambiamo la soluzione.
- 2- Si moltiplica una riga per un coefficiente diverso da zero. Ovvero, se moltiplichiamo la riga per un coefficiente, non cambiamo la soluzione.

Il metodo di eliminazione di Gauss sfrutta in maniera decisiva queste equivalenze, ossia il fatto che si possono operare queste manipolazioni sul sistema senza alterarne la soluzione.

Vediamo adesso come il MEG opera su un caso di esempio molto semplice, un sistema 3×3 , che scriviamo nella forma $Ax = b$

$$Ax = b = A^{(1)}x = b^{(1)} = \begin{cases} x_1 + 2x_2 + x_3 = 0 & \text{(I)} \\ 2x_1 + 2x_2 + 3x_3 = 3 & \text{(II)} \\ -x_1 - 3x_2 = 2 & \text{(III)} \end{cases}$$

Abbiamo evidenziato che $A^{(1)}x = b^{(1)}$ (in cui $A^{(1)} = A$ e $b^{(1)} = b$) per indicare la prima matrice e termine noto della trasformazione MEG. Facilmente si trova che questo è un sistema che ha come soluzione $x = (1, -1, 1)^T$. Il primo passo del MEG è quello di eliminare dalla seconda equazione e dalla terza equazione la prima incognita x_1 . In altre parole, vogliamo rimuovere i termini che contengono x_1 dalla seconda e dalla terza equazione. In questo caso di esempio, per farlo per la seconda equazione, possiamo sottrarre, alla seconda equazione, la prima equazione moltiplicata per due

$$(II) - 2(I) \rightarrow 2x_1 + 2x_2 + 3x_3 - 2(x_1 + 2x_2 + x_3) = 3 - 2$$

Infatti, moltiplicando la prima equazione per due, si ottiene un termine $2x_1$ che, sottratto al termine corrispondente della seconda equazione, annulla la prima incognita x_1 ($2x_1 - 2x_1 = 0$). In questo modo, abbiamo eliminato il termine x_1 dalla seconda equazione. Inoltre, se prendiamo la prima equazione e la sommiamo alla terza equazione, otteniamo un'altra manipolazione che elimina anche il termine x_1 dalla terza equazione

$$(III) + (I) = x_1 + 2x_2 + x_3 + (-x_1 - 3x_2) = 2 + 0.$$

Abbiamo ottenuto una nuova terza equazione (diversa da quella di partenza), ma con il termine x_1 annullato. In questo modo, abbiamo raggiunto il nostro primo obiettivo: annullare il termine x_1 dalla seconda e dalla terza equazione.



La procedura del MEG consiste quindi nel determinare opportuni moltiplicatori della prima riga in modo tale che, sottraendo la prima riga opportunamente moltiplicata dalle seconde e terze righe, si ottengano due nuove righe con elementi uguali a zero nella prima posizione

$$\begin{aligned}(II) - m_{21}(I) &= (II)', \\ (III) - m_{31}(I) &= (III)'. \end{aligned}$$

In questo caso specifico i moltiplicatori sono

$$m_{21} = 2, \quad m_{31} = -1,$$

dove m_{21} è il moltiplicatore che permette di trasformare la seconda riga, utilizzando l'informazione sulla prima riga, per fare annullare x_1 , mentre m_{31} gioca lo stesso ruolo di m_{21} , ma per la terza riga. Troviamo un nuovo insieme di equazioni, ovvero un nuovo sistema che ha questa forma

$$A^{(2)}\mathbf{x} = \mathbf{b}^{(2)} = \begin{cases} x_1 + 2x_2 + x_3 = 0 & (I) \\ -2x_2 + x_3 = 3 & (II)' \\ x_2 + x_3 = 2 & (III)' \end{cases}.$$

Osserviamo che, dalla nuova seconda e terza riga, è sparito il termine x_1 . Il sistema ottenuto in questa maniera è equivalente al precedente, poiché la soluzione \mathbf{x} non è cambiata, ma abbiamo un termine noto $\mathbf{b}^{(2)}$ e una matrice $A^{(2)}$ differenti. Abbiamo identificato questo cambiamento con un apice $\cdot^{(2)}$.

Se ora proseguiamo con questa strategia, al secondo passo del MEG dovremmo eliminare dalla terza riga la seconda incognita x_2 . A questo scopo, moltiplichiamo la seconda riga per il moltiplicatore $m_{31} = \frac{1}{2}$ ed eseguiamo la trasformazione

$$(III)'' = (III)' - m_{31}(II)', \quad m_{31} = \frac{1}{2}.$$

Questa nuova riga la chiamiamo $(III)''$, perché rappresenta la terza riga della seconda trasformazione. Dopo queste manipolazioni algebriche, il nuovo sistema lineare è

$$A^{(3)}\mathbf{x} = \mathbf{b}^{(3)} = \begin{cases} x_1 + 2x_2 + x_3 = 0 & (I) \\ -2x_2 + x_3 = 3 & (II)' \\ \frac{1}{2}x_3 = \frac{1}{2} & (III)'' \end{cases}.$$

Abbiamo, come si nota, raggiunto una struttura triangolare superiore

$$U\mathbf{x} = \mathbf{b}^{(3)} = \begin{cases} x_1 + 2x_2 + x_3 = 0 & (I) \\ -2x_2 + x_3 = 3 & (II)' \\ \frac{1}{2}x_3 = \frac{1}{2} & (III)'' \end{cases},$$

in cui il nuovo termine noto $\mathbf{b}^{(3)} = (0, 3, \frac{1}{2})^T$ è cambiato. La prima equazione è rimasta invariata perché è stata utilizzata per trasformare la seconda e la terza equazione. La seconda equazione è stata modificata al primo passo e poi è rimasta invariata, motivo per cui l'abbiamo etichettata come $(II)'$. La terza equazione è stata modificata anche al secondo passo, quindi la etichettiamo come $(III)''$. La matrice che abbiamo ottenuto è una matrice triangolare superiore, poiché gli elementi sotto la diagonale principale sono tutti nulli. Il sistema $U\mathbf{x} = \mathbf{b}^{(3)}$ si può risolvere quindi tramite il metodo delle sostituzioni all'indietro, che, come abbiamo visto, richiede un numero di operazioni dell'ordine di n^2 . Abbiamo visto come il MEG opera su una matrice 3×3 . Ora vediamo l'idea generale applicata a una matrice A di dimensione n . Iniziamo con una matrice A , che consideriamo a priori piena, e la definiamo $A = A^{(1)}$. Al primo passo, otteniamo una matrice $A^{(2)}$ che ha la stessa prima riga, ma con la prima sotto-colonna (ovvero, quella sotto la diagonale principale) annullata. Quindi tutti gli elementi sotto la



diagonale principale della prima colonna sono nulli. La seconda trasformazione $A^{(3)}$ metterà a zero la seconda sotto-colonna e così via. In generale, nella k -esima trasformazione $A^{(k+1)}$, avremo annullato tutti gli elementi sotto la diagonale principale fino alla k -esima colonna. Continuando questo processo, arriveremo alla fine con una matrice $A^{(n)}$ che avrà la parte superiore U diversa da zero e la parte inferiore L composta solo da zeri. Questo è lo schema generale di trasformazione che porta dalla matrice di partenza A alla matrice finale $A^{(n)}$.

Osserviamo adesso come questo si riflette a livello algebrico sulle equazioni, poiché, alla fine, intendiamo fornire una formula che ci consenta di calcolare il generico passo k -esimo per il MEG. Al passo k , quindi, otteniamo un sistema della forma

$$A^{(k)}x = b^{(k)} \quad 1 \leq k \leq n,$$

dove $A^{(k)}$ è una matrice del tipo

$$A^{(k)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & a_{kk}^{(k)} & \dots & a_{kn}^{(k)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & a_{nk}^{(k)} & \dots & a_{nn}^{(k)} \end{pmatrix}.$$

La matrice risultante avrà tutti gli zeri sotto la diagonale principale fino alla riga $(k-1)$ -esima, con gli elementi sulla parte sopra diagonale che, a priori, non sono nulli. Per identificare gli elementi, utilizziamo un sopra-indice che caratterizza la riga: la prima riga avrà gli elementi indicati come riga $a_{11}^{(1)}, a_{12}^{(1)}, \dots, a_{1n}^{(1)}$, con il sopra-indice $\cdot^{(1)}$, la seconda riga come $0, a_{22}^{(2)}, \dots, a_{2n}^{(2)}$, con il sopra-indice $\cdot^{(2)}$, e così via fino alla k -esima riga, indicata come $0, \dots, a_{kk}^{(k)}, \dots, a_{kn}^{(k)}$, con il sopra-indice $\cdot^{(k)}$. L'indice superiore indica l'ultima volta in cui gli elementi sono stati modificati: l'indice $\cdot^{(1)}$, sulla prima riga indica che questi elementi non sono mai stati modificati, l'indice $\cdot^{(2)}$ sulla seconda riga indica che questi elementi sono stati modificati nel passaggio da $A^{(1)}$ a $A^{(2)}$ e l'indice $\cdot^{(k)}$ indica che questi elementi sono stati modificati, così come tutti quelli successivi, nel passaggio da $A^{(k-1)}$ ad $A^{(k)}$. Se, adesso, esprimiamo algebricamente il passaggio dalla matrice $A^{(k)}$ alla matrice $A^{(k+1)}$, questo può essere realizzato utilizzando le seguenti formule

$$\begin{aligned} a_{ij}^{(k+1)} &= a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)}, \\ b_i^{(k+1)} &= b_i^{(k)} - m_{ik}b_k^{(k)}, \\ m_{ik} &= \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}. \end{aligned} \quad i, j = k+1, \dots, n \quad k = 1, \dots, n-1,$$

Quindi la componente ij della nuova matrice $A^{(k+1)}$, ovvero $a_{ij}^{(k+1)}$ è la vecchia componente $a_{ij}^{(k)}$ alla quale abbiamo sottratto gli elementi della matrice vecchia k -esima in riga k , $a_{kj}^{(k)}$, opportunamente moltiplicatori per i fattori m_{ik} . La stessa operazione viene applicata al termine noto, poiché stiamo trasformando la riga del sistema, cioè l'equazione generica del sistema. Pertanto, non stiamo solo modificando gli elementi della matrice, ma anche gli elementi del termine noto. I moltiplicatori sono definiti come $m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$. Osserviamo che, in questo modo, stiamo dividendo per elementi $a_{kk}^{(k)}$ che



devono essere diversi da zero. Di conseguenza, otteniamo una formula compatta che riassume l'intero algoritmo MEG. Più precisamente, questa formula compatta indica come passare dalla matrice $A^{(k)}$ alla matrice $A^{(k+1)}$, lavorando sugli elementi individuali della matrice.

Osserviamo quindi come si presenta la matrice $A^{(n)}$ al passo n

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

Questa è una matrice triangolare superiore ed il sistema da risolvere sarà $A^{(n)}\mathbf{x} = \mathbf{b}^{(n)}$

$$A^{(n)}\mathbf{x} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \cdots & 0 & a_{nn}^{(n)} \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(n)} \end{bmatrix} = \mathbf{b}^{(n)}$$

Dove il vettore soluzione \mathbf{x} è rimasto invariato e $\mathbf{b}^{(n)}$ è il nuovo termine noto. Dunque, se poniamo $\mathbf{f} = \mathbf{b}^{(n)}$ e $U = A^{(n)}$, abbiamo ritrovato il sistema triangolare superiore

$$U\mathbf{x} = \mathbf{f}.$$

Abbiamo visto che il MEG è un metodo diretto, in quanto conduce a formule che permettono di calcolare la soluzione in un numero finito di passi. Ora, siamo interessati a calcolare quante operazioni sono necessarie per applicare il MEG al fine di trovare la soluzione del sistema. In altre parole, dobbiamo determinare il costo computazionale del MEG. Si può verificare che il MEG richiede circa $\frac{2}{3}n^3$ operazioni (o flops, che sta per "floating point operations" o "operazioni in virgola mobile", in quanto si tratta di operazioni effettuate da un calcolatore). Una verifica rigorosa è presentata nell'Approfondimento di questa lezione.

Per riassumere, la regola di Cramer per la risoluzione di un sistema lineare richiederebbe un numero di operazioni pari a $n!$, che comporta un numero straordinariamente elevato di operazioni. Il MEG, invece, trasforma il sistema $A\mathbf{x} = \mathbf{b}$ in un sistema equivalente $U\mathbf{x} = \mathbf{f}$, con U una matrice triangolare superiore, in $n - 1$ passaggi. Il costo computazionale del MEG è di circa $\frac{2}{3}n^3$ operazioni. A queste vanno aggiunte n^2 operazioni per risolvere il sistema finale. Pertanto, per n "grande", il numero totale di operazioni per risolvere il sistema lineare utilizzando il MEG è dell'ordine di $\frac{2}{3}n^3$. Come abbiamo visto questo rappresenta un guadagno considerevole rispetto al passaggio dalla regola di Cramer all'uso del MEG.

Approfondimento:

Abbiamo visto che il Metodo di Eliminazione di Gauss (MEG) è un metodo diretto, in quanto conduce a formule che permettono di calcolare la soluzione in un numero finito di passi. Ora, siamo interessati a calcolare quante operazioni sono necessarie per applicare il MEG al fine di trovare la soluzione del sistema. In altre parole, dobbiamo determinare il costo computazionale del MEG. Si può verificare che il MEG richiede circa $\frac{2}{3}n^3$ operazioni (o flops, che sta per "floating point operations" o "operazioni in virgola mobile", in quanto si tratta di operazioni effettuate da un calcolatore).



mobile", in quanto si tratta di operazioni effettuate da un calcolatore). Vale la pena notare che stiamo calcolando queste operazioni trascurando i termini di ordine inferiore, ovvero i termini di ordine n^2 e n . Infatti, quando n è "grande", l'esponente 3 domina sugli altri termini, e i termini di ordine inferiore (come n^2 o n) diventano irrilevanti. In altre parole, stiamo trattando questi termini come ordini di grandezza infinitesimi, come si farebbe nell'analisi asintotica.

Vogliamo calcolare, quindi, esattamente il numero di operazioni richieste dal MEG. Ricordiamo che la prima formula del MEG è

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

che ci dice come calcolare i nuovi coefficienti della matrice. Per ogni scelta di indici i, j, k , abbiamo

$$1 \text{ molt.} + 1 \text{ sottr.} = 2 \text{ oper.,}$$

una moltiplicazione e una sottrazione, quindi due operazioni. Inoltre, tenendo conto della variazione degli indici $i, j = k+1, \dots, n$ e $k = 1, \dots, n-1$, abbiamo in totale

$$\sum_{k=1}^{n-1} \sum_{i=k+1}^n \sum_{j=k+1}^n 2 = 2 \sum_{k=1}^{n-1} (n-k)^2.$$

Facendo il calcolo delle operazioni necessarie, troviamo due volte la somma di $(n-k)^2$, con $k = 1, \dots, n-1$. Questa somma rappresenta le operazioni necessarie per calcolare i coefficienti delle matrici di trasformazione, dalla prima matrice prima $A^{(2)} (k=1)$ all'ultima $A^{(n)} (k=n-1)$. Analogamente, per calcolare il termine noto, dovremo utilizzare una formula simile che tiene conto delle trasformazioni applicate alla matrice e ai termini noti ad ogni passo del MEG

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

Questa consiste di una moltiplicazione e una sottrazione. Quindi, considerando la variazione degli indici si ha

$$\sum_{k=1}^{n-1} \sum_{j=k+1}^n 2 = 2 \sum_{k=1}^{n-1} (n-k).$$

Infine, dovremo calcolare i moltiplicatori

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

Questa operazione consiste di una sola divisione da fare per ogni indice i e per ogni k , quindi

$$\sum_{k=1}^{n-1} \sum_{j=k+1}^n 1 = \sum_{k=1}^{n-1} (n-k).$$

Possiamo ora riassumere tutte le operazioni necessarie per eseguire i tre passaggi fondamentali del MEG e calcolare il costo computazionale complessivo del metodo. Per fare ciò, teniamo in considerazione i seguenti risultati derivanti dall'analisi

- $3 \sum_{k=1}^{n-1} (n-k) = \frac{3}{2} n(n-1),$
- $2 \sum_{k=1}^{n-1} (n-k)^2 = \frac{1}{3} n(n-1).$

Se, in definitiva, sommiamo tutte le operazioni del MEG, troviamo



$$\frac{3}{2}n(n-1) + \frac{1}{3}(n-1)n(2n-1) = \frac{2}{3}n^3 + \frac{n^2}{2} - \frac{7}{6}n \text{ operazioni}$$

Se trascuriamo i termini in n^2 e n , abbiamo

$$\frac{2}{3}n^3 \text{ operazioni.}$$

Questo sono le operazioni complessive richieste dal MEG.

Concludiamo, dando una stima sui tempi di calcolo che servirebbero per risolvere un sistema $n \times n$ con il MEG, supponendo di utilizzare un calcolatore moderno in grado di effettuare un miliardo di operazioni al secondo (10^9 opz/s, 1 Giga flop). I tempi di calcolo per risolvere un sistema lineare $n \times n$ con $n = 10, 15, 20$, utilizzando il MEG, sono di seguito riportati

$$n = 10 \rightarrow t \sim 0.67 \mu s \text{ (microsecondi),}$$

$$n = 15 \rightarrow t \sim 2.25 \mu s,$$

$$n = 20 \rightarrow t \sim 5.33 \mu s.$$

Siamo quindi passati da tempi dell'ordine di migliaia di anni per $n = 20$ (come osservato nella lezione precedente) con la regola di Cramer a tempi dell'ordine dei microsecondi utilizzando il MEG, ottenendo un miglioramento significativo. Tuttavia, nelle applicazioni ingegneristiche è comune lavorare con matrici di dimensioni molto grandi, con centinaia, migliaia o addirittura milioni di righe. In questi casi, anche il MEG diventa estremamente oneroso e impraticabile. Ad esempio, supponendo di utilizzare un calcolatore capace di eseguire 1 Giga flop, la risoluzione di un sistema di dimensioni $n = 10^6$ richiederebbe circa 7700 giorni (21 anni). Per affrontare sistemi di tali dimensioni, è quindi necessario adottare metodi più efficienti, come algoritmi iterativi o fattorizzazioni più avanzate, tra cui la decomposizione LU.