

Sistemi lineari – Metodi diretti

Calcolo Numerico - Ing. Inf. - Lezione 6

Outline

Outline

È dato il sistema di n equazioni lineari $Ax = b$ ovvero il sistema di equazioni

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \cdots & & \cdots & & \cdots & & \cdots \\ a_{n1}x_1 & + & \cdots & + & a_{nn}x_n & = & b_n \end{array}$$

dove i **coefficienti** a_{ij} e i **termini noti** b_i sono numeri complessi

Si cerca un vettore $x^T = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ che verifichi tutte le equazioni

Se si scrive il sistema nella forma $Ax = b$, A è detta la **matrice dei coefficienti**, b è detto il **vettore dei termini noti** mentre x è il **vettore delle incognite**

Si studiano due tipi di metodi risolutivi per sistemi di equazioni lineari: i **metodi diretti** e i **metodi iterativi**

Nei **metodi diretti** si giunge alla soluzione esatta (a meno degli errori di arrotondamento) con un insieme finito di operazioni sui dati

Applicando i **metodi iterativi**, la soluzione viene **approssimata** dai termini di una successione di vettori di cui la soluzione cercata è il limite.

Outline

Un primo **metodo diretto** per la soluzione di un sistema lineare è il **metodo di Cramer**

Metodo di Cramer

Dato il sistema lineare $Ax = b$ la soluzione si ottiene calcolando

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

dove A_i è ottenuta da A sostituendo il vettore b alla i -esima colonna

Il **metodo di Cramer** è sicuramente utile per le dimostrazioni ma, anche se formalmente molto semplice, non risulta un **buon metodo** dal punto di vista **numerico** a causa del suo elevato **costo computazionale**

Nella tabella che segue vengono messi a confronto i **costi computazionali** del **Metodo di Cramer** e del **metodo di Gauss** (che ripeteremo a breve) considerando 10^{-6} **secondi** il tempo necessario per eseguire una operazione

n	metodo di Cramer	metodo di Gauss
12	103 minuti	$7.1 \cdot 10^{-4}$ secondi
13	24 ore	$8.9 \cdot 10^{-4}$ secondi
15	15 giorni	$1.1 \cdot 10^{-3}$ secondi
50	$4.9 \cdot 10^{52}$ anni	$4.4 \cdot 10^{-2}$ secondi

Il **metodo di Gauss** o **metodo di eliminazione** consiste nel trasformare il sistema $Ax = b$ in un sistema equivalente

$$R x = c$$

dove $R \in \mathbb{C}^{n \times n}$ è una matrice **triangolare superiore** con $r_{ii} \neq 0$, $i = 1, 2, \dots, n$

Il sistema $Rx = c$ è della forma

$$\begin{array}{ccccccc} r_{11}x_1 & + & r_{12}x_2 & + & \cdots & + & r_{1n}x_n & = & c_1 \\ & & r_{22}x_2 & + & \cdots & + & r_{2n}x_n & = & c_2 \\ & & & & \cdots & & \cdots & & \cdots \\ & & & & & & r_{nn}x_n & = & c_n \end{array}$$

La risoluzione, come tutti i sistemi con matrice dei coefficienti triangolare, risulta immediata utilizzando le formule

$$x_n = c_n / r_{nn}$$

$$x_i = (c_i - \sum_{j=i+1}^n r_{ij}x_j) / r_{ii}, \quad i = n-1, \dots, 1$$

Per passare dal sistema $AX = b$ al sistema equivalente $Rx = c$ occorre **eliminare** dalla i -esima equazione le incognite con indice minore di i , per $i = 2, 3, \dots, n$

Questo obiettivo si realizza utilizzando la proprietà che la soluzione non cambia se si sostituisce all'equazione i -esima una sua combinazione lineare con un'altra equazione del sistema

Come primo passo, se $a_{11} \neq 0$, si elimina x_1 da tutte le equazioni che seguono la prima, sottraendo membro a membro dalla i -esima equazione, $i = 2, 3, \dots, n$, la prima equazione i cui membri siano stati moltiplicati per il **moltiplicatore**

$$\ell_{i1} = \frac{a_{i1}}{a_{11}}$$

Ponendo per ragioni formali $a_{ij}^{(1)} = a_{ij}$, $i, j = 1, 2, \dots, n$, il sistema, dopo la prima eliminazione, assume la forma:

$$\begin{array}{ccccccccc} a_{11}^{(1)} x_1 & + & a_{12}^{(1)} x_2 & + & a_{13}^{(1)} x_3 & + & \cdots & + & a_{1n}^{(1)} x_n & = & b_1^{(1)} \\ & & a_{22}^{(2)} x_2 & + & a_{23}^{(2)} x_3 & + & \cdots & + & a_{2n}^{(2)} x_n & = & b_2^{(2)} \\ & & \cdots & & \cdots & & \cdots & & \cdots & & \cdots \\ & & a_{n2}^{(2)} x_2 & + & a_{n3}^{(2)} x_3 & + & \cdots & + & a_{nn}^{(2)} x_n & = & b_n^{(2)} \end{array}$$

dove

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \ell_{i1} a_{1j}^{(1)}, \quad b_i^{(2)} = b_i^{(1)} - \ell_{i1} b_1^{(1)}, \quad i, j = 2, 3, \dots, n$$

Dopo aver eseguito il primo passo dell'algoritmo, se nel sistema trovato risulta $a_{22}^{(2)} \neq 0$, si può eliminare x_2 da tutte le equazioni che seguono la seconda, utilizzando i moltiplicatori $\ell_{i2} = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}$, $i = 3, 4, \dots, n$

Iterando il procedimento, supposto che possa ripetersi $n - 1$ volte, si giunge al sistema

$$\begin{array}{ccccccc} a_{11}^{(1)} x_1 & + & a_{12}^{(1)} x_2 & + & \dots & + & a_{1n}^{(1)} x_n & = & b_1^{(1)} \\ & & a_{22}^{(2)} x_2 & + & \dots & + & a_{2n}^{(2)} x_n & = & b_2^{(2)} \\ & & & & \dots & & \dots & & \dots \\ & & & & & & a_{nn}^{(n)} x_n & = & b_n^{(n)} \end{array}$$

che è della forma cercata e risulta equivalente al sistema iniziale

Le condizioni perché l'algoritmo possa giungere al termine sono

$$a_{11}^{(1)} \neq 0, \quad a_{22}^{(2)} \neq 0, \dots, \quad a_{nn}^{(n)} \neq 0$$

Attenzione!!!

In mancanza di una di queste condizioni l'algoritmo si interrompe

Le condizioni citate equivalgono alla proprietà che la matrice A abbia i **minori principali di testa diversi da zero**, cioè

$$a_{11} \neq 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0, \dots, \det(A) \neq 0$$

In effetti, poche matrici godono di questa proprietà. Sicuramente verificano le condizioni richieste le **matrici simmetriche e definite** (ricorrono spesso nelle applicazioni)

Definizione

Una matrice $A \in \mathbb{C}^{n \times n}$ si dice **definita positiva (negativa)** se risulta

$$x^H A x > 0 (< 0), \quad \forall x \in \mathbb{C}^n - \{0\}$$

Vedremo tra non molto che, nella pratica del calcolo, l'algoritmo viene modificato sia per garantirne la completa esecuzione, sia per ridurre la propagazione degli errori di arrotondamento, anche quando le condizioni sui minori principali di testa fossero verificate

Costo computazionale

Costo computazionale

Per risolvere un sistema di n equazioni in n incognite applicando il **metodo di Gauss**, il numero delle operazioni (moltiplicazioni e

divisioni) necessarie è $\frac{n^3}{3} + n^2 - \frac{n}{3}$

Si ricorda che lo stesso sistema, risolto con il **metodo di Cramer**, richiede circa $(n+1)!(n-1)$ operazioni

Esempio 1

Calcolare la soluzione del sistema lineare

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

Soluzione

$$x = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

Esempio 2

Calcolare la soluzione del sistema lineare

$$\begin{pmatrix} 1 & -1 & 2 \\ -2 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix}$$

Soluzione

$$x = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$$

Supponiamo di dover risolvere k sistemi lineari con la stessa matrice dei coefficienti

$$A x^{(1)} = b^{(1)}, \quad A x^{(2)} = b^{(2)}, \quad \dots \quad A x^{(k)} = b^{(k)}$$

Se si utilizza il [metodo di Gauss](#) si possono risolvere i k sistemi contemporaneamente risolvendo il sistema lineare

$$A X = B$$

con

$$X = (x^{(1)} | x^{(2)} | \dots | x^{(k)}), \quad B = (b^{(1)} | b^{(2)} | \dots | b^{(k)})$$

Calcolo della matrice inversa

Caso particolare della matrice B è se risulta $B = I$ e cioè

$$B = (e^{(1)} | e^{(2)} | \dots | e^{(n)})$$

In questo caso si risolve il sistema lineare

$$AX = I$$

la cui soluzione è

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

I coefficienti $a_{11}^{(1)}$, $a_{22}^{(2)}$, $a_{33}^{(3)}$, ... si dicono **elementi pivotali**

Si introducono delle modifiche dell'algoritmo di base del metodo di Gauss per limitare la propagazione degli errori

Questi cambiamenti consistono nello stabilire a priori un criterio di scelta dell'elemento pivotale per ciascuna eliminazione

Un primo criterio, detto del **pivoting parziale**, prevede di calcolare

$$\max_{1 \leq i \leq n} |a_{i1}^{(1)}| = |a_{r1}^{(1)}|$$

Se $r \neq 1$, si scambiano di posto la prima e l' r -esima equazione e si considera un sistema in cui i coefficienti $a_{1j}^{(1)}$, $j = 1, 2, \dots, n$, sono i coefficienti $a_{rj}^{(1)}$, $j = 1, 2, \dots, n$, del sistema di partenza e viceversa

Effettuata la prima eliminazione, si supponga

$$\max_{2 \leq i \leq n} |a_{i2}^{(2)}| = |a_{s2}^{(2)}|$$

Se $s \neq 2$, si scambiano di posto l'equazione di indice s con quella di indice 2, quindi si procede alla seconda eliminazione e così via

Anche in questo caso, dopo $n - 1$ passi, si arriva ad un sistema della forma $Rx = c$ (R triangolare superiore) che è di risoluzione immediata

Osservazione

Ad ogni passo del metodo la tecnica di **pivoting parziale** prevede un possibile scambio tra due equazioni del sistema

Da quello che abbiamo visto nelle precedenti lezioni, uno scambio di righe di un sistema lineare si ottiene **premultiplicando** il sistema per una opportuna **matrice di permutazione**

Questo significa che se si ricorre a k ($k \leq n - 1$) scambi di equazioni utilizzando k matrici di permutazione P_1, P_2, \dots, P_k è come andare a risolvere il sistema

$$P_k P_{k-1} \cdots P_2 P_1 A x = P_k P_{k-1} \cdots P_2 P_1 b$$

Quindi, se fossimo partiti dal sistema lineare $\hat{A}x = \hat{b}$ con

$$\hat{A} = P_k P_{k-1} \cdots P_2 P_1 A \quad \text{e} \quad \hat{b} = P_k P_{k-1} \cdots P_2 P_1 b,$$

non ci sarebbe stata la necessità di ricorrere a scambi di equazioni nella risoluzione del sistema

Ricordando che il prodotto di k matrici di permutazione è una matrice di permutazione, ponendo $P = P_k P_{k-1} \cdots P_2 P_1$, il sistema da risolvere sarebbe della forma

$$P A x = P b$$

Un'altra strategia è quella del **pivoting totale** in cui il pivot è ancora l'elemento di massimo modulo, ma scelto ogni volta sull'intera matrice del sistema parziale da trasformare anziché su una sola colonna come nel **pivoting parziale**

È chiaro che in questo caso per portare il pivot selezionato nella posizione di testa può essere necessario un riordinamento delle equazioni e delle incognite

In entrambe le **tecniche di pivoting** l'obiettivo è quello di **contenere** i moltiplicatori del metodo di Gauss in modo tale che abbiano il modulo più piccolo possibile (sicuramente risultano di **modulo minore di 1**)

L'algoritmo di eliminazione può essere considerato come un procedimento che trasforma una data matrice A in una matrice triangolare superiore R

Per vedere in quale relazione sono le matrici A ed R si supponga che la matrice A abbia i minori principali di testa diversi da zero e quindi che si possa applicare l'algoritmo di eliminazione senza effettuare scambi tra le righe

Teorema

Nell'ipotesi che valgano le condizioni sui minori principali di testa, l'algoritmo di eliminazione produce la **fattorizzazione**

$$A = LR$$

dove R è la matrice triangolare superiore ottenuta alla fine della applicazione del metodo ed L ha la forma

$$L = \begin{pmatrix} 1 & & & & & \\ \ell_{21} & 1 & & & & \\ \ell_{31} & \ell_{32} & 1 & & & \\ \vdots & \vdots & \ddots & \ddots & & \\ \vdots & \vdots & & \ddots & 1 & \\ \ell_{n1} & \ell_{n2} & \cdots & \cdots & \ell_{n,n-1} & 1 \end{pmatrix}$$

in cui gli elementi al disotto della diagonale principale coincidono con i moltiplicatori dell'algoritmo di eliminazione

Dimostrazione

Siano $A_1, A_2, \dots, A_{n-1} = R$ le matrici dei successivi sistemi equivalenti a $Ax = b$ che si ottengono dopo ciascuna eliminazione

Risulta

$$A_1 = H_1 A, \quad A_2 = H_2 A_1, \dots, \quad A_{n-1} = H_{n-1} A_{n-2} = R$$

con

$$H_i = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\ell_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & -\ell_{n,i} & & & 1 \end{pmatrix}$$

Dimostrazione

Posto $L_i := H_i^{-1}$ e tenuto conto che

$$H_i^{-1} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \ell_{i+1,i} & \ddots & \\ & & \vdots & & \ddots \\ & & \ell_{n,i} & & & 1 \end{pmatrix}$$

e che $L_1 L_2 \dots L_{n-1} = L$, segue

$$H_{n-1} H_{n-2} \dots H_1 A = R$$

Dimostrazione

Dall'ultima uguaglianza si ricava

$$\begin{aligned} A &= H_1^{-1} H_2^{-1} \cdots H_{n-2}^{-1} H_{n-1}^{-1} R \\ &= L_1 L_2 \cdots L_{n-2} L_{n-1} R \\ &= L R \end{aligned}$$

che prova la tesi.....se gli elementi di L sono effettivamente quelli dichiarati nell'enunciato del Teorema

Elementi della matrice L

Non dimostreremo che in generale gli elementi di L sono quelli dichiarati ma lo verichiamo direttamente solo nel caso $n = 4$

Da quanto detto avremo

$$L = L_1 L_2 L_3$$

dove $L_i = H_i^{-1}$, $i = 1, 2, 3$

Elementi della matrice L

Svolgiamo il calcolo necessario

$$\begin{aligned} L &= L_1 L_2 L_3 \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ \ell_{2,1} & 1 & 0 & 0 \\ \ell_{3,1} & 0 & 1 & 0 \\ \ell_{4,1} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \ell_{3,2} & 1 & 0 \\ 0 & \ell_{4,2} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \ell_{4,3} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ \ell_{2,1} & 1 & 0 & 0 \\ \ell_{3,1} & 0 & 1 & 0 \\ \ell_{4,1} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \ell_{3,2} & 1 & 0 \\ 0 & \ell_{4,2} & \ell_{4,3} & 1 \end{pmatrix} \end{aligned}$$

Elementi della matrice L

Da cui, eseguendo l'ultimo prodotto, otteniamo

$$\begin{aligned} L &= L_1 L_2 L_3 \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{2,1} & 1 & 0 & 0 \\ l_{3,1} & l_{3,2} & 1 & 0 \\ l_{4,1} & l_{4,2} & l_{4,3} & 1 \end{pmatrix} \end{aligned}$$

Esempio 3

Calcolare la fattorizzazione LR della matrice

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

Soluzione

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{pmatrix} \quad R = \begin{pmatrix} 4 & 1 & 1 & 1 \\ 0 & 3 & 2 & 2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Esempio 4

Calcolare la fattorizzazione LR della matrice

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

Soluzione

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix} \quad R = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix}$$

Nel caso di una matrice A qualunque si può dimostrare che l'algoritmo di Gauss con l'eventuale uso del pivoting parziale conduce ancora ad una fattorizzazione della forma

$$P A = L_p R_p$$

dove P è una matrice di permutazione definita dagli scambi di righe richiesti dall'algoritmo, R_p è triangolare superiore ed L_p è triangolare inferiore con elementi diagonali unitari

Una conseguenza delle fattorizzazioni $A = LR$ e $PA = L_p R_p$ è data, rispettivamente, dalle uguaglianze

$$\det(A) = \det(R), \quad \det(A) = (-1)^s \det(R_p)$$

dove s è il numero degli scambi di righe dovuti all'uso del pivoting, mentre i determinanti di R ed R_p sono dati dal prodotto dei termini diagonali

Si osservi che il costo computazionale di $\det(A)$ mediante la definizione è di circa $n!$ operazioni mentre il numero di operazioni per costruire R ed R_p con l'eliminazione gaussiana è di circa $n^3/3$ operazioni.

Metodo diretto di fattorizzazione

La conoscenza di una fattorizzazione della matrice A può essere utile ai fini della risoluzione del sistema $Ax = b$

Infatti se ad esempio si conosce a priori la fattorizzazione $A = LR$, il sistema si può scrivere

$$L R x = b$$

la sua risoluzione si riconduce a quella immediata dei due sistemi triangolari

$$L c = b, \quad R x = c$$

Metodo di Gauss-Jordan

Una variante del metodo di Gauss è il **metodo di Gauss-Jordan** che consiste nell'operare sulla matrice dei coefficienti del sistema $Ax = b$ delle combinazioni tra le righe in modo da ottenere un sistema lineare equivalente la cui matrice dei coefficienti sia diagonale

Per fare ciò, basta effettuare, dal secondo passo in poi, le combinazioni lineari opportune anche con le righe che precedono la riga a cui appartiene l'elemento pivotale
In altre parole, al passo i -esimo del metodo di Gauss-Jordan si elimina l'incognita x_i da tutte le equazioni esclusa l' i -esima

Metodo di Gauss-Jordan

Il risultato finale è un sistema del tipo

$$Dx = b'$$

dove D è una matrice diagonale

Come per il metodo base di Gauss, è possibile che uno o più elementi pivotali risultino nulli

Non si presenta questo caso se e solo se valgono le condizioni riportate in precedenza sui minori principali di testa che assicurano l'applicabilità del metodo senza dover ricorrere a scambi di righe

Metodo di Gauss-Jordan

Per ridurre la propagazione degli errori di arrotondamento si ricorre alle tecniche di pivoting

L'applicazione del metodo di Gauss-Jordan a un sistema di ordine n comporta un costo computazionale di

$$\frac{n^3}{2} + n^2 - \frac{n}{2}$$

operazioni, cioè superiore a quello del metodo di Gauss ($O(n^3/3)$)

Calcolo della matrice A^{-1}

Come accennato in precedenza, data una matrice A di ordine n non singolare, la sua matrice inversa A^{-1} è la soluzione del sistema matriciale

$$AX = I$$

Si tratta quindi di risolvere n sistemi lineari $Ax^{(i)} = e^{(i)}$, $i = 1, 2, \dots, n$, dove $x^{(i)}$ e $e^{(i)}$ sono la i -esima colonna, rispettivamente, della matrice X e della matrice I

Calcolo della matrice A^{-1}

Tali sistemi vengono risolti simultaneamente considerando la matrice completa $(A \mid I)$ ed effettuando su di essa le operazioni di eliminazione gaussiana

Applicando il [metodo di Gauss-Jordan](#) alla matrice completa del sistema e riportando la matrice diagonale finale ad essere la matrice identica si ha il sistema

$$I X = C$$

dove $C = A^{-1}$

Esempio 5

Calcolare la matrice inversa di

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

Risultato

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

Esempio 6

Calcolare la matrice inversa di

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

Risultato

$$A^{-1} = \frac{1}{8} \begin{pmatrix} -3 & 4 & 1 \\ 4 & -8 & 4 \\ 1 & 4 & -3 \end{pmatrix}$$