

LEZIONE 22

Meg come prodotto per matrici elementari di trasformazione, fattorizzazione LU di una matrice non singolare

Nella scorsa lezione abbiamo discusso del metodo di eliminazione gaussiana come metodo di riduzione di una matrice quadrata non singolare a forma triangolare, che permette di usarlo per calcolare il determinante (nel caso singolare ci si accorge durante l'esecuzione che il determinante è nullo).

In particolare abbiamo studiato l'implementazione in aritmetica floating-point con la tecnica del pivoting, che permette di superare il problema della comparsa di un elemento nullo sulla diagonale e dal punto di vista numerico ha un effetto stabilizzante sull'algoritmo (evitando la generazione di numeri arrotondati troppo grandi e quindi limitando la perdita di precisione nelle sottrazioni).

Sia l'operazione fondamentale di annullamento che l'operazione di scambio legata al pivoting si possono in realtà scrivere in forma matriciale, come prodotto a sinistra per opportune matrici elementari di trasformazione.

Cominciamo con lo scambio di 2 righe: lo scambio della riga k con la riga i di una generica matrice $B \in \mathbb{R}^{n \times n}$ si ottiene moltiplicando a sinistra B per la matrice identità in cui sono state scambiate le righe k e i (chiameremo $S_{k,i}$ tale matrice).

Facciamo un esempio 3×3

$$S_{2,1}B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} b_{21} & b_{22} & b_{23} \\ b_{11} & b_{12} & b_{13} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

ha esattamente l'effetto di scambiare la riga 2 con la riga 1 di B .

Si osservi che $\det(S_{2,1}) = -1$ e che $S_{2,1} \cdot S_{2,1} = I$ cioè $S_{2,1}^{-1} = S_{2,1}$.

In generale

$$S_{k,i} = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

e $S_{k,i}^{-1} = S_{k,i}$, $\det(S_{k,i}) = -1$.

Quest'ultima uguaglianza spiega in forma matriciale perchè scambiare 2 righe cambia segno al determinante, ricordando che il determinante del prodotto di due matrici è il prodotto dei determinanti

$$\det(S_{k,i}B) = \det(S_{k,i})\det(B) = -\det(B)$$

L'altra operazione vettoriale chiave del *meg* al passo i -esimo è

$$\mathcal{R}_k^{(i+1)} := \mathcal{R}_k^{(i)} + (\overbrace{-m_{ki}}^{\text{moltiplicatore}}) \mathcal{R}_i^{(i)}$$

dove $m_{ki} = \frac{a_{ki}^{(i)}}{a_{ii}^{(i)}}$, $i+1 \leq k \leq n$, detto "moltiplicatore" ha lo scopo di propagare la struttura di zeri

Annullare di zeri propagate

$$A^{(i)} = \begin{pmatrix} \boxed{0} & & & & \\ & \ddots & & & \\ & & a_{ii} & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} \rightarrow A^{(i+1)} = \begin{pmatrix} \boxed{0} & & & & \\ & \ddots & & & \\ & & a_{ii} & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}$$

In generale l'operazione tra righe

$$\mathcal{R}_k := \mathcal{R}_k + \alpha \mathcal{R}_i$$

di una matrice $B \in \mathbb{R}^{n \times n}$ corrisponde a moltiplicare a sinistra per la matrice elementare

$$T_{k,i}(\alpha) = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \alpha & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

che si ottiene dalla matrice identità mettendo α al punto di 0 come elemento k, i .

Facciamo un esempio 3×3

$$T_{3,1}(\alpha)B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ \alpha b_{11} + b_{31} & \alpha b_{12} + b_{32} & \alpha b_{13} + b_{33} \end{pmatrix}$$

Si osservi che si tratta di una matrice triangolare inferiore con 1 sulla diagonale principale e quindi $\det(T_{k,i}(\alpha)) = 1$ e che $T_{k,i}(-\alpha)T_{k,i}(\alpha) = I$ cioè $(T_{k,i}(\alpha))^{-1} = T_{k,i}(-\alpha)$.

Infatti, tornando per fissare le idee all'esempio 3×3

$$T_{3,1}(-\alpha)T_{3,1}(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\alpha & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha(-\alpha) & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Diventa chiaro che il passo i -esimo del *meg* si può scrivere come composizione di trasposizioni elementari tramite il prodotto a sinistra per la sequenza di matrici elementari che realizzano l'eventuale scambio del pivoting e gli annullamenti successivi in colonna i sotto la diagonale.

Prendendo l'esempio 3×3 della scorsa lezione

$$A^{(1)} = A = \begin{pmatrix} 1 & 2 & 1 \\ \underbrace{2}_{pivot} & 2 & 3 \\ -1 & -3 & 0 \end{pmatrix}$$

abbiamo (stavolta applichiamo il pivoting per coerenza con l'implementazione anche se operiamo con frazioni)

$$\begin{aligned}
A &\xrightarrow[\substack{\text{scambio} \\ R_2 \leftrightarrow R_1}]{S_{2,1}} \begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ -1 & -3 & 0 \end{pmatrix} \xrightarrow[\substack{R_2 := R_2 + (-\frac{1}{2})R_1}]{T_{2,1}(-\frac{1}{2})} \begin{pmatrix} 2 & 2 & 3 \\ 0 & 1 & -1/2 \\ -1 & -3 & 0 \end{pmatrix} \\
&\xrightarrow[\substack{R_3 := R_3 + (\frac{1}{2})R_1}]{T_{3,1}(1/2)} A^{(2)} = \begin{pmatrix} 2 & 2 & 3 \\ 0 & 1 & -1/2 \\ 0 & -2 & 3/2 \end{pmatrix} \xrightarrow[\substack{\text{scambio} \\ R_3 \leftrightarrow R_2}]{S_{3,2}} \begin{pmatrix} 2 & 2 & 3 \\ 0 & -2 & 3/2 \\ 0 & 1 & -1/2 \end{pmatrix} \\
&\xrightarrow[\substack{R_3 := R_3 + (\frac{1}{2})R_2}]{T_{3,2}(1/2)} A^{(3)} = U = \begin{pmatrix} 2 & 2 & 3 \\ 0 & -2 & 3/2 \\ 0 & 0 & 1/4 \end{pmatrix}
\end{aligned}$$

UPPER TRIANGULAR

cioè

$$U = A^{(3)} = T_{3,2}\left(\frac{1}{2}\right) S_{3,2} \underbrace{T_{3,1}\left(\frac{1}{2}\right) T_{2,1}\left(-\frac{1}{2}\right) S_{2,1} A}_{A^{(2)}}$$

dove $-m_{3,2} = -m_{3,1} = \frac{1}{2}$, $-m_{2,1} = -\frac{1}{2}$ e $\det(U) = 2(-2)\frac{1}{4} = -1 = \det(A)$ essendoci stato un numero pari di scambi.

In generale si avrà che il passo i -esimo del meg in forma matriciale diventa

$$A^{(i+1)} = T_{n,i}(-m_{n,i}) T_{n-1,i}(-m_{n-1,i}) \dots T_{i+2,i}(-m_{i+2,i}) T_{i+1,i}(-m_{i+1,i}) S_{p_i,i} A^{(i)}$$

con p_i l'indice del pivot sottodiagonale in colonna i , mentre la sequenza completa di trasformazioni (A non singolare) è

$$\begin{aligned}
U = A^{(n)} &= T_{n,n-1}(-m_{n,n-1}) S_{p_{n-1},n-1} \\
&T_{n,n-2}(-m_{n,n-2}) T_{n-1,n-2}(-m_{n-1,n-2}) S_{p_{n-2},n-2} \\
&\dots S_{p_2,2} T_{n,1}(-m_{n,1}) \dots T_{2,1}(-m_{2,1}) S_{p_1,1} A
\end{aligned}$$

A questo punto possiamo far vedere che il meg permette di ottenere una fattorizzazione della matrice di partenza (a meno di una permutazione delle righe) nel prodotto di due matrici triangolari, una triangolare inferiore per una triangolare superiore.

Per semplicità ci limitiamo al caso "senza scambi", cioè in cui il pivoting trova l'elemento di max modulo già sulla diagonale.

Ci sono classi di matrici per cui si può dimostrare che questo vale (lo accettiamo), ad esempio:

- matrici a diagonale strettamente dominante

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}| \quad \forall i$$

- matrici simmetriche definite positive.

Limitiamoci per capire meglio a matrici 3×3 con

$$U = A^{(3)} = T_{3,2}(-m_{3,2}) T_{3,1}(-m_{3,1}) T_{2,1}(-m_{2,1}) A$$

Posto $\mathcal{L} = T_{3,2}(-m_{3,2}) T_{3,1}(-m_{3,1}) T_{2,1}(-m_{2,1})$ abbiamo

$$U = \mathcal{L}A \Rightarrow A = LU \text{ con } L = \mathcal{L}^{-1}$$

Ma come è fatta L ?

$$L = \mathcal{L}^{-1} = \overbrace{(T_{2,1}(-m_{2,1}))^{-1} (T_{3,1}(-m_{3,1}))^{-1} (T_{3,2}(-m_{3,2}))^{-1}}^{\text{ordine invertito}} \\ = T_{2,1}(m_{2,1}) T_{3,1}(m_{3,1}) T_{3,2}(m_{3,2})$$

ricordando che $(T_{k,i}(\alpha))^{-1} = T_{k,i}(-\alpha)$ cioè

$$L = \begin{pmatrix} 1 & 0 & 0 \\ m_{2,1} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ m_{3,1} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & m_{3,2} & 0 \end{pmatrix} \\ = \begin{pmatrix} 1 & 0 & 0 \\ m_{2,1} & 1 & 0 \\ m_{3,1} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & m_{3,2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ m_{2,1} & 1 & 0 \\ m_{3,1} & m_{3,2} & 1 \end{pmatrix}$$

ovvero L è una matrice triangolare inferiore (Lower Triangular in inglese)

$$L = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}$$

che ha $l_{ii} = 1$, $l_{ki} = m_{k,i}$ per $i + 1 \leq k \leq n$ e $l_{ki} = 0$ per $k < i$. Questo è vero anche in generale

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{2,1} & 1 & 0 & \cdots & 0 \\ m_{3,1} & m_{3,2} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & m_{n,3} & \cdots & 1 \end{pmatrix}$$

in cui il triangolo sotto la diagonale principale contiene i moltiplicatori del meg.

In presenza di scambi imposti dal pivoting, si può dimostrare (non richiesto) che il meg produce comunque una fattorizzazione del tipo

$$PA = LU$$

matrice di permutazione (punta a P)
triang. sup, $u_{ii} \neq 0$ (punta a U)
triang. inf, $l_{ii} = 1$ (punta a L)

dove

- U è la matrice triangolare superiore ottenuta alla fine del meg
- P è una matrice di permutazione cioè una matrice invertibile che contiene solo 0 o 1 ed è ottenuta come prodotto di matrici di scambio.

$$P = S_{p_{n-1}, n-1} \cdot S_{p_{n-2}, n-2} \cdot \dots \cdot S_{p_2, 2} \cdot S_{p_1, 1}$$

Ad esempio nel caso 3×3 , P può essere:

$$P = I, \quad P = S_{2,1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad P = S_{3,1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ P = S_{2,1}S_{3,2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad P = S_{3,1}S_{3,2} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ P = S_{3,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

- L è una matrice triangolare inferiore con $l_{ii} = 1 \forall i$ e con i moltiplicatori opportunamente rimescolati nel triangolo inferiore sotto la diagonale (accettiamo questo fatto senza dimostrarlo per semplicità).

Quindi l'aspetto è:

$$U = \begin{pmatrix} u_{11} & \dots & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & u_{nn} \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ l_{21} & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & \dots & l_{nn-1} & 1 \end{pmatrix} \quad u_{ii} \neq 0 \forall i$$

e con una interpretazione

$$PA = \begin{pmatrix} \boxed{1} & & & 0 \\ & \boxed{1} & & \\ & & \ddots & \\ & & & \boxed{1} \end{pmatrix} \begin{pmatrix} \boxed{0} & & & \\ & \boxed{0} & & \\ & & \boxed{0} & \\ & & & \boxed{0} \end{pmatrix}$$

$L \qquad U$

Perché le fattorizzazioni sono importanti? Perché permettono di decomporre una matrice senza struttura nel prodotto di matrici strutturate. Nella prossima lezione vedremo infatti come risolvere sistemi lineari tramite la fattorizzazione LU generata dal meg.

Concludiamo la lezione osservando che il costo computazionale della fattorizzazione $PA = LU$ ottenuta col meg coincide col costo del meg, infatti le matrici P e L sono costruite con gli scambi e con i moltiplicatori utilizzati durante il processo di eliminazione.

Quindi:

$$c_n^{LU} = c_n^{meg} \sim \frac{2}{3}n^3, \quad n \rightarrow \infty$$

come parte dominante in termini di # di flops.