Calcolo Numerico A - a.a. 2008/09

16 aprile 2009

Fattorizzazioni

Supponiamo che la matrice A del sistema lineare

$$Ax = b,$$

si scomponga nel prodotto di due matrici

$$A = BC, \tag{0.1}$$

allora il sistema lineare risulta equivalente a

$$\begin{cases} By = b, \\ Cx = y. \end{cases} \tag{0.2}$$

Fattorizzazioni classiche associate a diversi metodi di risoluzione sono le sequenti:

1. Fattorizzazione LU: L matrice triangolare inferiore (lower) con elementi diagonali uguali ad uno ($l_{ii} = 1$), U matrice triangolare superiore (upper).
   Questa fattorizzazione è associata al metodo di Gauss.
2. Fattorizzazione LL^T: L matrice triangolare inferiore con elementi diagonali positivi ($l_{ii} > 0$).
   Questa fattorizzazione è associata al metodo di Cholesky.
3. Fattorizzazione QR: Q matrice unitaria (ortogonale nel caso reale), R matrice triangolare superiore.
   Questa fattorizzazione è associata al metodo di Householder.

Osservazione

- *Se la matrice A è reale, allora anche le matrici delle fattorizzazioni lo sono.*
- *Il costo computazionale delle fattorizzazioni si vedrà che sarà dato da un numero di operazioni dell'ordine di n^3 (a meno di costante), mentre il costo computazionale della risoluzione dei sistemi (0.2) sarà dato da un numero di operazioni dell'ordine di n^2.*
- *La fattorizzazione QR esiste sempre, mentre non sempre si possono avere le fattorizzazione LU e LL^T.*

Fattorizzazione LU

Teorema

Sia $A \in \mathbb{R}^{n \times n}$ una matrice invertibile. Siano $A_k = (a_{ij})_{1 \le i,j \le k}$ i minori principali dominanti (detti anche "di testa") ottenuti dalla matrice A.
La fattorizzazione LU di A esiste, ed è anche unica unica, se e solo se le matrici A_k sono non singolari per ogni $k = 1, \dots, n-1$.

Osservazione

La fattorizzazione LU è sempre possibile per i seguenti tipi di matrici:

- *Matrici a dominanza diagonale stretta (infatti anche i loro minori principali sono di questo tipo e quindi invertibili per il corollario del teorema di Gerschgorin).*
- *Matrici simmetriche definite positive (segue dal criterio di Sylvester).*

Teorema

Data A matrice quadrata reale o complessa, esiste una matrice di permutazione Π per cui la matrice ΠA ammette la fattorizzazione LU, cioè:

$$\Pi A = LU.$$

Fattorizzazione LL^T

In questo fattorizzazione si richiede che L sia triangolare inferiore con elementi diagonali positivi.

Teorema

Una matrice reale $A \in \mathbb{R}^{n \times n}$ ammette la fattorizzazione LL^T se e solo se è simmetrica definita positiva. La fattorizzazione è unica .

Abbiamo già visto che se A è prodotto di una matrice invertibile e della sua trasposta, allora A è simmetrica e definita positiva.
Viceversa, supponiamo che A sia una matrice simmetrica definita positiva. Possiamo allora prendere la scomposizione $A = LU$ di tale matrice (dove L ha ora elementi 1 sulla diagonale). Scriviamo U come prodotto di una matrice D diagonale (invertibile) e una matrice V triangolare superiore con elementi 1 sulla diagonale:

$$U = DV.$$

Usando che A è simmetrica si trova:

$$LDV = LU = A = A^T = U^T L^T = V^T DL^T.$$

Si noti ora che le matrici V^T e DL^T forniscono di nuovo una scomposizione del tipo LU di A e quindi, per l'unicità di tale scomposizione, si ha:

$$L = V^T, \quad U = DL^T.$$

Sostituita nella relazione precedente, si ottiene $A = LDL^T$.

Proviamo che gli elementi diagonali $d_1, \ldots, d_n$ di D sono positivi (certamente sono non nulli). Questo segue dal fatto che:

$$d_j = e_j^T D e_j = e_j^T L^{-1} L D L^T (L^T)^{-1} e_j = e_j^T L^{-1} A (L^{-1})^T e_j > 0,$$

essendo A definita positiva.
Dunque possiamo scrivere $D = \Delta^2$, con Δ diagonale con elementi positivi sulla diagonale. Ne segue che:

$$A = L\Delta^2 L^T = L\Delta(\Delta L^T) = L\Delta(L\Delta)^T.$$

Unicità. Se $LDL^T = ND'N^T$ sono due scomposizioni con L e N triangolari inferiori aventi elementi diagonali uguali a 1 e matrici D e D' diagonali, allora

$$N^{-1} L D = D' N^T (L^T)^{-1},$$

da cui segue che $D = D'$ e $N = L$.

Forme compatte di fattorizzazione: caso del metodo di Cholesky

Sia $L = (\ell_{ij})$ con $\ell_{ij} = 0$ se $i < j$ e con $\ell_{jj} > 0$ per ogni j. Supponiamo $A = (a_{ij}) = LL^T$. Allora:

$$a_{ij} = \sum_{k=1}^{min(i,j)} \ell_{ik}\ell_{jk}.$$

Dunque:

$$\begin{cases} a_{jj} = \sum_{k=1}^{j} \ell_{jk}^2, & j = 1, \dots, n, \\ a_{ij} = \sum_{k=1}^{j} \ell_{ik}\ell_{jk}, & i > j, \end{cases}$$

da cui si ricavano le relazioni che definiscono il metodo di Cholesky:

$$\begin{cases} \ell_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} \ell_{jk}^2}, & j = 1, \dots, n, \\ \ell_{ij} = \frac{1}{\ell_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} \ell_{ik}\ell_{jk} \right), & i > j. \end{cases}$$

dove a sommatoria è intesa nulla quando il secondo estremo è nullo, cioè per $j = 1$.
Si noti che in questo medoto prima si calcola prima l'elemento ℓ_{11} e poi i sottostanti elementi della prima colonna; si passa quindi all'elemento ℓ_{22} e poi ai sottostanti elementi della seconda colonna, e così via.
Si tratta di un metodo stabile con costo computazionale dell'ordine di $n^3/6$.

FATTORIZZAZIONE QR

Teorema

Data $A \in \mathbb{R}^{n \times n}$, esistono una matrice Q ortogonale e una matrice R triangolare superiore tali che:

$$A = QR. \tag{3.3}$$

Inoltre si può richiedere che tutti gli elementi di R siano non negativi. Con tale richiesta, se la matrice A è invertibile, allora le matrici Q e R sono univocamente determinate, cioè in questi casi la fattorizzazione è unica.

Unicità: Supponiamo A non singolare. Siano $A = QR = Q'R'$ scomposizioni di A con la richiesta che R e R' abbiano elementi non negativi sulla diagonale. (Si noti che, essendo A non singolare, anche R e R' lo devono essere, quindi gli elementi sulla diagonale non possono essere nulli: stiamo quindi richiedendo che siano positivi). Allora abbiamo

$$QR = Q'R' \Rightarrow Q'^{-1}Q = R'R^{-1} \Rightarrow R'R^{-1} \text{ è triangolare e ortogonale,}$$

quindi $R'R^{-1}$ 'e diagonale con ± 1 sulla diagonale. Ma ora, per la richiestà di positività si ha questi elementi devono essere proprio 1, quindi $R = R'$ e di conseguenza $Q = Q'$.

Osservazione

La fattorizzazione QR di una matrice A (in generale, senza nessuna richiesta) non è unica, infatti, per ogni matrice unitaria (ortogonale nel caso reale) e diagonale, detta matrice di fase,

$$S = \begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{pmatrix}, \quad \text{con } |\theta_i| = 1,\ i = 1, \ldots, n,$$

risulta

$$QR = QS\bar{S}R = (QS)(\bar{S}R) = Q'R'.$$

Se però A è non singolare, allora la matrice triangolare R deve essere invertibile e quindi con elementi diagonali non nulli. Allora si può richiedere che essi siano reali positivi. Tale scomposizione è unica.

Fattorizzazione QR: matrici di Householder

Una matrice (simmetrica) della forma

$$H = I - \beta v\, v^T, \text{ con } \beta \in \mathbb{R}, v \in \mathbb{R}^n,$$

si dice matrice di Householder se (oltre ad essere simmetrica) è anche ortogonale. Imponendo tale condizione, si trova $v = 0$ e $\beta = 0$, oppure $v \neq 0$ con

$$\beta = \frac{2}{v^T v} = \frac{2}{\|v\|^2}.$$

Infatti imponendo $I = HH^T = H^2$ si trova:

$$I = H^2 = I - 2\beta vv^T + \beta^2 v(v^T v)v^T = I - \beta vv^T(2 - \beta vv^T).$$

Osservazione

Le matrici di Householder sono simmetriche e ortogonali, quindi si ha sempre $H^2 = I$.

Si noti che (se $v \neq 0$):

$$Hv = v - \beta v v^T v = v - 2v = -v,$$
$$w \in v^{\perp} \Rightarrow Hw = w - \beta v v^T w = w.$$

Osservazione

Le matrici di Householder associate ad vettore $v \neq 0$ sono riflessioni rispetto al piano ortogonale al vettore v. (Il loro determinante vale -1 e la loro traccia è $n-2$).

Lemma

Per ogni $x \in \mathbb{R}^n, n \neq 0$, *è possibile trovare una matrice di Householder*

$$H = I - \frac{2}{v^T v} vv^T$$

tale che:

$$Hx = \alpha e_1. \tag{3.4}$$

Si noti che H ortogonale implica $\|x\|_2 = \|Hx\|_2 = \|\alpha e_1\| = |\alpha|$, *quindi deve essere*

$$\alpha = \pm\|x\|.$$

Imponendo le condizione (3.4) si trova che basta porre:

$$v = x - \alpha e_1.$$

Matrici di Householder: esempio

Prendiamo il vettore:

$$x = \begin{pmatrix} 4 \\ 7 \\ 4 \end{pmatrix}, \qquad \text{con norma} \|x\| = 9.$$

Consideriamo delle matrici che moltiplicate con x danno il vettore e_1. Ricordo che nel caso del metodo di Gauss si prende:

$$M = \begin{pmatrix} 1 & 0 & 0 \\ -7/4 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

per cui si ottiene:

$$Mx = 4e_1.$$

In particolare la prima componente del vettore colonna x è rimasta inalterata.

(continua esempio)

Nel caso del medoto di Householder si considerano invece le matrici:

$$H_{\pm} = I_3 - \beta v v^T, \quad \text{dove } v = \begin{pmatrix} 4 \pm \|x\| \\ 7 \\ 4 \end{pmatrix}, \; \beta = \frac{2}{\|v\|^2} = \frac{2}{(4 \pm \|x\|)^2 + 49 + 16}.$$

In particolare:

$$\begin{aligned} H_+ &= I_3 - \frac{2}{234} \begin{pmatrix} 13 \\ 7 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 13 & 7 & 4 \end{pmatrix} = I_3 - \frac{1}{117} \begin{pmatrix} 169 & 91 & 52 \\ 91 & 49 & 28 \\ 52 & 28 & 16 \end{pmatrix} \\ &= \frac{1}{117} \begin{pmatrix} -52 & -91 & -52 \\ -91 & 49 & -28 \\ -52 & -28 & 101 \end{pmatrix}. \end{aligned}$$

Abbiamo che:

$$H_+ x = \begin{pmatrix} -9 \\ 0 \\ 0 \end{pmatrix} = -\|x\| e_1.$$

Metodo di Householder per la fattorizzazione QR

Componendo trasformazioni di Householder si può ottenere una matrice triangolare R,

$$H_{n-1} \cdots H_1 A = R.$$

quindi una scomposizione QR della matrice A (ricordando che le matrici di Householder sono involutive e ortogonali, e che la composizione di matrici ortogonali è ancora ortogonale):

$$A = H_1 \cdots H_{n-1} R,$$

con $Q = H_1 \cdots H_{n-1}$ ortogonale.
Un altro medoto per ottenere la scomposizione QR (almeno per A invertibili) è dato dal procedimento di Gram-Schmidt, ma tale processo non è stabile.

Considerando A ortogonale, si trova che R deve essere anche ortogonale, quindi diagonale. Dunque $R = diag(1, \ldots 1, \pm 1)$ è una riflessione. Abbiamo cosìottenuto il seguente

Corollario (Teorema di Cartan-Dieudonné)

Ogni matrice ortogonale di ordine n è prodotto di (al più) n riflessioni.

Metodo di Householder per la fattorizzazione QR: esempio

Sia:

$$A = \begin{pmatrix} 72 & -144 & -144 \\ -144 & -36 & -360 \\ -144 & -360 & 450 \end{pmatrix}$$

Passo 1: Poniamo $A = A_1$.
Abbiamo $|\alpha_1| = 216$, cioè la norma 2 della prima colonna di A_1. Dunque possiamo scegliere:

$$v_1 = \begin{pmatrix} 288 \\ -144 \\ -144 \end{pmatrix}, \quad \beta_1 = \frac{2}{\|v_1\|^2} = \frac{1}{62208},$$

$$H_1 = I - \beta_1 v_1 v_1^T = \frac{1}{6} \begin{pmatrix} -2 & 4 & 4 \\ 4 & 4 & -2 \\ 4 & -2 & 4 \end{pmatrix}$$

Nota che $H_1 X_1 = -216 e_1$, quindi:

$$A_2 = H_1 A_1 = \begin{pmatrix} -216 & -216 & 108 \\ 0 & 0 & -486 \\ 0 & -324 & 324 \end{pmatrix}$$

Abbiamo $|\alpha_2| = 324$, che è la norma 2 della prima colonna del minore ottenuto da A_2 eliminando la prima riga e la prima colonna. Dunque possiamo scegliere:

$$v_2 = \begin{pmatrix} 0 \\ 324 \\ -324 \end{pmatrix}, \quad \beta_1 = \frac{2}{\|v_2\|^2} = \frac{1}{104976},$$

$$H_2 = I - \beta_1 v_2 v_2^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Si trova:

$$A_3 = H_2 A_2 = \begin{pmatrix} -216 & -216 & 108 \\ 0 & -324 & 324 \\ 0 & 0 & -486 \end{pmatrix} = R.$$

Dunque $P_2 P_1 A = R$, da cui segue:

$$A = H_1 H_2 R = QR, \qquad \text{con } Q = H_1 H_2 = \frac{1}{6} \begin{pmatrix} -2 & 4 & 4 \\ 4 & -2 & 4 \\ 4 & 4 & -2 \end{pmatrix}.$$

Costi computazionali

Table: Ordine di grandezza del numero di operazioni richieste

Medoto	molt.	add.	rad.q.	div.
Sostituzione	$\frac{1}{2}n^2$	$\frac{1}{2}n^2$	-	n
Gauss	$\frac{1}{3}n^3$	$\frac{1}{3}n^3$	-	$\frac{1}{2}n^2$
Gauss-Jordan	$\frac{1}{2}n^3$	$\frac{1}{2}n^3$	-	
Cholesky	$\frac{1}{6}n^3$	$\frac{1}{6}n^3$	n	$\frac{1}{2}n^2$
Householder	$\frac{2}{3}n^3$	$\frac{2}{3}n^3$	n	$\frac{1}{2}n^2$