Analisi Numerica

Risoluzione numerica di sistemi lineari
Fattorizzazione di Cholesky
Risoluzione di sistemi sovradeterminati

Ángeles Martínez Calomardo
amartinez@units.it

Laurea Triennale in Intelligenza Artificiale e Data Analytics
A.A. 2024–2025

Fattorizzazione LDM^T

Teorema (LDM^T)

Sia $A \in \mathbb{R}^{n\times n}$. *Se* **tutti i minori principali di testa di** A **sono non nulli** *allora esiste ed è unica la fattorizzazione*

$$A = LDM^T$$

con L *ed* M *triangolari inferiori, con* $\ell_{ii} = m_{ii} = 1$ *e* D *matrice diagonale.*

Dimostrazione

Da $A = LU$,

posto

$$D = \text{diag}(U) = \text{diag}(u_{11}, u_{22}, \dots, u_{nn})$$

si ha, essendo D invertibile (poichè lo è A e quindi U)

$$LU = LD(D^{-1}U) = LDM^T \qquad (m_{ii} = 1)$$

$\square$

Fattorizzazione LDL^T

Teorema ($\boldsymbol{LDL^T}$)

Sia $A \in \mathbb{R}^{n\times n}$ tale che $A = A^T$. Se **tutti i minori principali di testa di A sono non nulli** *allora esiste ed è unica la fattorizzazione $A = LDL^T$ dove L è triangolare inferiore con $\ell_{ii} = 1$ e D diagonale.*

Dimostrazione

Per il risultato precedente si ha $A = LDM^T$.

Ma per la simmetria si ha anche $A = LDM^T = (LDM^T)^T = MDL^T$.

Per l'unicità della fattorizzazione segue $L = M,\ M^T = L^T$. $\square$

Fattorizzazione di Cholesky

Teorema (Fattorizzazione di Cholesky)

Sia A una matrice reale simmetrica e definita positiva. Allora A è fattorizzabile nella forma

$$A = RR^T$$

dove R è una matrice triangolare inferiore reale non singolare.

Dimostrazione

Essendo A simmetrica e definita positiva (verifica il criterio di Sylvester), ammette la fattorizzazione

$$A = LDL^T.$$

Dalla definizione di matrice definita positiva si ha, per ogni vettore $x \neq 0$,

$$0 < x^T A x = \underbrace{x^T L}_{y^T} D \underbrace{L^T x}_{y} = y^T D y,$$

con $y \neq 0$ perché L è invertibile, quindi D è definita positiva.

Dimostrazione (seguito)

Gli elementi diagonali della matrice D sono positivi in quanto, scegliendo $y = e_i$ l'i-esimo vettore della base canonica, si ha $0 < e_i^T D e_i = d_{ii}$.

La matrice D ammette quindi una radice quadrata (reale) ovvero

$$\sqrt{D} = \text{diag}(\sqrt{d_{11}}, \ldots, \sqrt{d_{nn}})$$

Pertanto

$$A = LDL^T = \underbrace{L\sqrt{D}}_{R}\underbrace{\sqrt{D}L^T}_{R^T} = RR^T.$$

$\square$

Il costo computazionale della fattorizzazione di Cholesky è pari alla metà del costo della fattorizzazione LU.

La fattorizzazione di Cholesky è numericamente stabile.

Calcolo pratico di R. Esempio

$$\text{Sia } A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Imponiamo l'uguaglianza $RR^T = A$

$$\begin{pmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} r_{11} & r_{21} & r_{31} \\ 0 & r_{22} & r_{32} \\ 0 & 0 & r_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{1l3} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Lo schema più efficiente è quello di calcolare R per colonne (e quindi R^T per righe):

$$\text{Prima colonna di } R: \begin{cases} r_{11}^2 & = a_{11} = 2 \\ r_{21}r_{11} & = a_{21} = 1 \\ r_{31}r_{11} & = a_{31} = 1 \end{cases} \Rightarrow \begin{cases} r_{11} = \sqrt{2} \\ r_{21} = 1/\sqrt{2} \\ r_{31} = 1/\sqrt{2}. \end{cases}$$

$$\text{Seconda colonna di } R: \begin{cases} r_{21}^2 + r_{22}^2 & = a_{22} = 2 \\ r_{21}r_{31} + r_{22}r_{32} & = a_{32} = 1 \end{cases} \Rightarrow \begin{cases} r_{22} = \sqrt{3/2} \\ r_{32} = 1/\sqrt{6}. \end{cases}$$

$$\text{Terza colonna di } R: \{ \ r_{31}^2 + r_{32}^2 + r_{33}^2 = \quad a_{33} = 2 \quad \Rightarrow \{ \ r_{33} = \sqrt{4/3}.$$

Esempio. Segue

Si è trovato

$$R = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & \sqrt{3/2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{6} & \sqrt{4/3} \end{pmatrix}.$$

Una volta in possesso della fattorizzazione, risolviamo il sistema lineare $Ax = b$ con $b^T = (1, 2, 3)$.

Si ha

$$Ax = R \underbrace{R^T x}_{y} = b \iff \begin{cases} Ry = b \\ R^T x = y \end{cases}$$

$$\Longrightarrow y = \begin{pmatrix} 1/\sqrt{2} \\ \sqrt{3/2} \\ \sqrt{3} \end{pmatrix}, \qquad x = \begin{pmatrix} -1/2 \\ 1/2 \\ 3/2 \end{pmatrix}.$$

Risoluzione di sistemi sovradeterminati

Sia $A \in \mathbb{R}^{m \times n}$, con $m \geq n$ e b $\in \mathbb{R}^m$. Il sistema lineare $Ax = b$ ammette soluzione se e solo se $b \in \mathrm{Im}(A)$. Inoltre la soluzione è unica se $\mathrm{Rg}(A) = n$.

Se il sistema **non ammette soluzione**, il problema dei minimi quadrati consiste nel risolvere il problema

$$\min_{x \in \mathbb{R}^n} \|b - Ax\|_2$$

ovvero trovare $x \in \mathbb{R}^n$ tale che

$$\|b - Ax\|_2 \leq \|b - Ay\|_2, \quad \forall y \in \mathbb{R}^n \tag{1}$$

Teorema

Il vettore x è soluzione del sistema (delle equazioni normali)

$$A^T A x = A^T b \tag{2}$$

se e solo se x è soluzione di (1).

Dimostrazione del teorema

($\Longrightarrow$) Sia $x \in \mathbb{R}^n$ la soluzione di (2) cioè $A^T(b - Ax) = 0$. Per ogni $y \in \mathbb{R}^n$ si ha

$$b - Ay = b - Ax + Ax - Ay = b - Ax + A(x - y),$$

da cui prendendo norme

$$\begin{aligned}
\|b - Ay\|_2^2 &= \|b - Ax + A(x - y)\|_2^2 = \\
&= (b - Ax + A(x - y))^T (b - Ax + A(x - y)) = \\
&= \|b - Ax\|_2^2 + \|A(x - y)\|_2^2 + 2(x - y)^T \underbrace{A^T(b - Ax)}_{=0 \text{ per ipotesi}} = \\
&= \|b - Ax\|_2^2 + \|A(x - y)\|_2^2 \geq \|b - Ax\|_2^2.
\end{aligned}$$

($\Longleftarrow$) Dimostrare la tesi equivale a dimostrare che, se x soddisfa (1) allora $b - Ax$ è ortogonale a $\text{Im}(A)$.

$$\text{Scrivendo} \quad b = b_1 + b_2, \quad b_1 \in \text{Im}(A),\ b_2 \in \text{Im}(A)^\perp, \quad \text{si ottiene}$$

$$\|b - Ax\|_2^2 = \|b_1 - Ax\|_2^2 + \|b_2\|_2^2$$

Quindi la soluzione di (1) si ottiene per $b_1 - Ax = 0$, pertanto

$$b - Ax = b_1 + b_2 - Ax = b_2 \in \text{Im}(A)^\perp.$$

Soluzione del sistema dei minimi quadrati

La fattorizzazione QR

Il sistema delle equazioni normali ha soluzione **unica** se la matrice $A^T A$ è **non singolare**. Da $A^T Ax = A^T b$ si ha

$$x = (A^T A)^{-1} A^T b = A^+ b,$$

La matrice A^+ si definisce pseudoinversa di A.

In pratica, invece di formare la matrice $A^T A$ e risolvere il sistema delle equazioni normali si risolve direttamente il sistema sovradeterminato $Ax = b$ mediante la fattorizzazione QR.

Teorema

Data una matrice $A \in \mathbb{R}^{m \times n}$, $(m \geq n)$, *esiste una matrice ortogonale* $Q \in \mathbb{R}^{m \times m}$ *tale che*

$$A = QR = Q \begin{pmatrix} \tilde{R} \\ \bigcirc \end{pmatrix},$$

dove $R \in \mathbb{R}^{m \times n}$ *con* $(r_{i,j}) = 0, \; i > j$ $(\tilde{R} \in \mathbb{R}^{n \times n}$ *è triangolare superiore*$)$.

Sistemi sovradeterminati. Fattorizzazione QR

Osservando che le ultime $m-n$ righe di R sono nulle si può scrivere

$$A = QR = \begin{pmatrix} \tilde{Q} & Q_2 \end{pmatrix} \begin{pmatrix} \tilde{R} \\ \bigcirc \end{pmatrix} = \tilde{Q}\tilde{R},$$

dove $\tilde{Q} \in \mathbb{R}^{m \times n}$. Questa fattorizzazione ridotta è detta *skinny* o *light* QR.

Dunque si ha

$$Ax = b \iff \tilde{Q}\tilde{R}x = b \iff \tilde{Q}^T\tilde{Q}\tilde{R}x = \tilde{Q}^T b \iff \tilde{R}x = \tilde{Q}^T b.$$

Occorre quindi risolvere un sistema triangolare superiore.

Alternativa. Usare la SVD.

La decomposizione ai valori singolari (SVD) di una matrice

Teorema

$\forall A \in \mathbb{R}^{m\times n}$ esistono due matrici ortogonali $V \in \mathbb{R}^{m\times m}$, $U \in \mathbb{R}^{n\times n}$ e una matrice $\Sigma \in \mathbb{R}^{m\times n}$ con la forma seguente

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}, \qquad \Sigma_r = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix},$$

dove $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0,\ r \leq \min\{m,n\}$ tali che

$$A = V\Sigma U^T.$$

Gli scalari σ_i si chiamano valori singolari di A.

Le colonne di V e U si dicono vettori singolari sinistri e destri, rispettivamente, di A.

Sistemi sovradeterminati

La soluzione formale del sistema delle equazioni normali è

$$x = A^+ b,$$

dove $A^+ \in \mathbb{R}^{n\times m}$ si dice pseudo-inversa di Moore-Penrose di A.
A partire da una SVD di A:

$$A = V\Sigma U^T$$

si definisce

$$A^+ = U\Sigma^+ V^T, \text{ con la seguente definizione di } \Sigma^+ \in \mathbb{R}^{n\times m}$$

$$\sigma_{ii}^+ = \begin{cases} 1/\sigma_i, & i = 1,\dots,r \\ 0, & i = r+1,\dots,\min\{m,n\} \end{cases}$$