Lezione 10 (10 novembre 2023)

MEG $\quad A \xrightarrow[\text{passi}]{n-1} U$

MEG equivalente a fattorizzare

$$A = LU$$

(o più generalmente $PA = LU$)

Se A è simmetrica e def +

$$A = H \cdot H^T \qquad \text{Fatt. di CHOLESKY}$$

$$H = \begin{pmatrix} h_{11} & & & \\ h_{21} & \ddots & & 0 \\ \vdots & & \ddots & \\ h_{n1} & & & h_{nn} \end{pmatrix} \qquad h_{ii} > 0$$

Osservazione: complessità di Cholesky è metà di quella della LU o MEG. $\Rightarrow O\left(\frac{n^3}{6}\right)$

Come costruisco H?

Basta fare il prodotto

$$H \cdot H^T = A$$

$$\begin{pmatrix} h_{11} & 0 & \cdots & 0 \\ h_{21} & h_{22} & \cdots & 0 \\ \vdots & & \ddots & \\ h_{n1} & \cdots & & h_{nn} \end{pmatrix} \begin{pmatrix} h_{11} & \cdots & h_{n1} \\ & \ddots & \vdots \\ 0 & & h_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \cdots & & a_{2n} \\ & & & \\ a_{n1} & \cdots & & a_{nn} \end{pmatrix}$$

$h_{11}^2 = a_{11} \implies h_{11} = \sqrt{a_{11}} > 0$

$h_{11} h_{21} = a_{12} \implies h_{21} = a_{12}/h_{11}$

$h_{11} h_{31} = a_{13} \implies h_{31} = a_{13}/h_{11}$

$\cdots$

In formule

$$\begin{cases} h_{11} = \sqrt{a_{11}} \\ h_{ij} = \dfrac{1}{h_{jj}} \left(a_{ij} - \displaystyle\sum_{k=1}^{j-1} h_{ik} h_{jk} \right) \qquad i = 2, \ldots, n \quad j = 1, \ldots, n \\ h_{ii} = \sqrt{a_{ii} - \displaystyle\sum_{k=1}^{i-1} h_{ik}^2} \qquad i = 2, \ldots, n \end{cases}$$

MATLAB $H = \mathrm{chol}(A)$

ALGORITMO di THOMAS

(caso di A tridiagonale)

$$A = \begin{pmatrix} a_1 & c_1 & & & 0 \\ b_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ 0 & & & b_n & a_n \end{pmatrix}$$

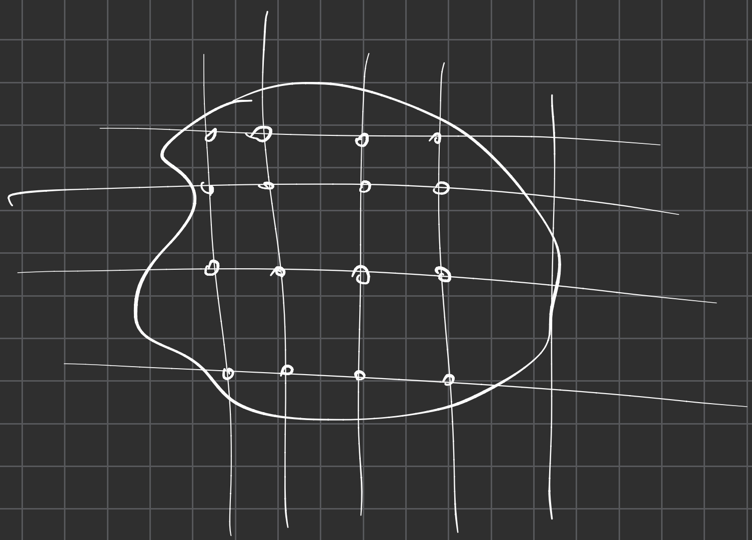
Se esiste la fatt. LU di A allora

$$L = \begin{pmatrix} 1 & & & & 0 \\ \beta_2 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & \beta_n & 1 \end{pmatrix} \qquad U = \begin{pmatrix} \alpha_1 & c_1 & & & 0 \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & c_{n-1} \\ 0 & & & & \alpha_n \end{pmatrix}$$

NOTA: $\bar{a}, \bar{b}, \bar{c}$ sono noti

$$\alpha_1 = a_1$$

$$\left.\begin{aligned} \beta_i &= \frac{b_i}{\alpha_{i-1}} \\ \alpha_i &= a_i - \beta_i c_{i-1} \end{aligned}\right\} \quad i = 2, \dots, n$$



DECOMPOSIONE QR

Caso A quadrata $n \times n$

$$A = QR$$

Q ortogonale / ortonormale $(QQ^T = I)$

R triangolare superiore

Se A è invertibile la fattorizzazione è unica se si richiede che gli elementi della diagonale di R siano $+$.

Q ha n colonne che sono ortonormali se A ha rango n

altrimenti se A ha rango k, le prime k colonne di Q saranno ortogonali

A è rettangolare $m \times n$ $\quad m \geq n$

(caso di sistemi sovradeterminati es. soluzioni ai minimi quadrati)

100×3

$$A = QR = \underset{m\times m}{Q}\begin{bmatrix} R_1 \\ O \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}\begin{bmatrix} R_1 \\ O \end{bmatrix}$$

R_1 è $n \times n$

O è $(m-n) \times n$

Q_1 $m \times n$

Q_2 $m \times (m-n)$

che hanno colonne ortogonali

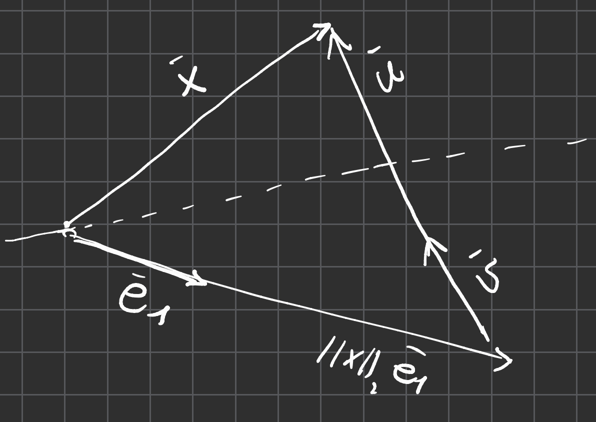
$$\begin{array}{r} n \left\{ \vphantom{R_1} \right. \\ m-n \left\{ \vphantom{O} \right. \end{array} \overbrace{\begin{pmatrix} R_1 \\ O \end{pmatrix}}^{n}$$

Domanda

Come costruire questa fattorizzazione?

L'idea è simile a quella della fattorizzazione LU. Importante è costruire la matrice Q

Per costruire Q ci sono varie tecniche (Gram-Schmidt) noi vedremo la tecnica delle matrici di riflessione di HOUSEHOLDER



Scopo: trovare una trasform. lineare che cambia il vettore $\bar{x}$ in un vettore della stessa lunghezza collineare a $\bar{e}_1$.

Q viene usata per riflettere un vettore x cosicché tutte le sue componenti eccetto una siano nulle

Dato x, $\|x\|_2 = \alpha$

$$u = x - \alpha e_1$$

$$v = \frac{u}{\|u\|} \qquad v^T v = 1$$

$$Q = I - 2 v v^T$$

<u>Osservazione</u> Q è simmetrica e ortogonale

- $$Q^T = (I - 2vv^T)^T = I - 2(vv^T)^T = I - 2\underbrace{(v^T)^T}_{v} v^T = I - 2vv^T = Q$$

- $$QQ^T = (I - 2vv^T)(I - 2vv^T) = I - 2vv^T - 2vv^T + 4\underbrace{(vv^T)(vv^T)}$$
$$= I - 4vv^T + 4v\underbrace{(v^Tv)}_{\|v\|_2^2 = 1}v^T$$
$$= I - 4vv^T + 4vv^T = I$$

$$Qx = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Quindi posso trasformare gradualmente A $m \times n$ in una forma triangolare superiore R mediante pre-moltiplicazioni per matrici di Householder

Nella fattispecie il vettore x su cui applico la trasformazione saranno le colonne di A

$$Q_1 A = \begin{bmatrix} \alpha & * \; * \; * \; * & \cdots & * \\ 0 & & & \\ | & & A' & \\ 0 & & & \end{bmatrix}$$

associata alla colonna 1

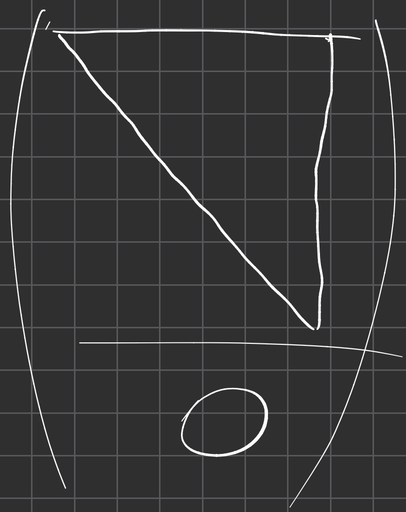
A' è la sotto matrice A dove tolgo la 1ª riga e la 1ª colonna

Posso applicare ad A' una trasformazione di Householder facendo attenzione al fatto che A' ha dimensione minore di A (infatti è $(m-1) \times (n-1)$)

Per evitare questo fatto, costruiremo

$$Q_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & Q'_k \end{bmatrix}$$

Ripeteremo l' operazione $t = \min(m-1, n)$

$R =$ 

Alla fine

$$R = Q_t \cdots Q_2 Q_1 A$$ (con R "triangolare superiore")

Detta

$$Q^T = Q_t \cdots Q_2 Q_1$$

$$Q = Q_1^T \cdots Q_t^T = Q_1 Q_2 \cdots Q_t = Q$$

$$\Longrightarrow A = QR$$

Si dimostra che la fatt. di Householder è molto più stabile che G-S

Infine, non lo dimostriamo, la complessità è $O\left(\frac{2}{3} n^3\right)$

<u>vantaggio</u>: si può applicare a matrici rettangolari

Esempio

$$A = \begin{bmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix}$$

1° passo: trovare una matrice di riflessione di H.

t.c. $a_1 = \begin{bmatrix} 12 \\ 6 \\ -4 \end{bmatrix} \longrightarrow \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix} = \|a_1\|_2 \, e_1$

Ora

$$u = x - \alpha e_1 \qquad v = \frac{u}{\|u\|_2}$$

$$\alpha = \|a_1\|_2 = 14$$

$$u = a_1 - 14 e_1 = [-2, 6, -4]^T$$

$$= 2 \, [-1, 3, -2]$$

$$v = \frac{1}{\sqrt{14}} [-1, 3, -2]^T$$

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{2}{\sqrt{14}\sqrt{14}} \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix} \begin{bmatrix} -1 & 3 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 6/7 & 3/7 & -2/7 \\ 3/7 & -2/7 & 6/7 \\ -2/7 & 6/7 & 3/7 \end{bmatrix}$$

$$Q_1 A = \begin{bmatrix} 14 & 21 & -14 \\ 0 & -49 & -14 \\ 0 & 168 & -77 \end{bmatrix}$$

Alla fine (dopo altri 2 passi di trasformazione)

$$R = \begin{bmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & -35 \end{bmatrix}$$

Per completezza si può ottenere la QR mediante TRASFORMAZIONI DI ROTAZIONE di GIVENS, soprattutto se A è SPARSA

Nota in MATLAB esiste la funzione

$$[Q,R] = qr(A)$$

Dalla QR alla SVD

SVD = Singular Value Decomposition

$A = QR$ Q $m \times m$ ortogonale

$m \geq n$

$$R = \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix} \begin{matrix} \}\, n \\ \}\, m-n \end{matrix}$$

$Ax = b$, $QRx = b$ ovvero $\tilde{R}x = Q^T b$

Idea: se invece di avere $\tilde{R}$ avessimo una matrice diagonale

$$R^* = \begin{pmatrix} * & & \\ & * & \\ & & * \\ \hline & 0 & \end{pmatrix}$$

$$\Longrightarrow \quad R^* x \simeq \begin{pmatrix} * \\ * \\ * \\ \vdots \\ * \end{pmatrix}$$

TEOREMA SVD

Data A $m \times n$ si può determinare una U $m \times m$ ortogonale una V $n \times n$ ortogonale e W $m \times n$ rettangolare "diagonale" t.c.

$$A = U W V^T$$

con

$$w_{ij} = \begin{cases} 0 & i \neq j \\ w_i & i = j \end{cases}$$

w_i sono detti valori singolari di A e sono tali che

$$\underbrace{w_1 \geq w_2 \geq \dots \geq w_r}_{} > w_{r+1} = \dots = w_p = 0$$

$$p = \min\{m, n\}$$

Proprietà delle matrici U e V

- Le colonne di U, $u_1, \dots, u_m$ sono dette <u>vettori singolari sx</u> di A e sono autovettori di AA^T

- Le colonne di V, $v_1, \dots, v_n$ sono vettori singolari dx di A e sono autovettori di A^TA

<u>Infatti</u>

$$A^T A = V \underbrace{W^T}_{U^TU = I} W V^T$$

$$\Downarrow$$

$$V^T (A^T A) V = W^T W \quad \text{che è diagonale}$$

- I numeri $w_1, \dots, w_p$ sono le radice degli autovalori di A^TA

$$w_i = \sqrt{\lambda_i(A^TA)} = \sqrt{\lambda_i(AA^T)} \quad \text{sono positivi}$$

Sono univocamente determinati e $\operatorname{rank}(A) = r$

Utilità della SVD

(i) Calcolo dell'inversa

$$A^{-1} = (UWV^T)^{-1} = (V^T)^{-1} W^{-1} U^{-1}$$
$$= V W^{-1} U^T$$

perché V e U sono ortogonali

sapendo che

$$W = diag(w_1, \dots, w_r, 0 \dots 0)$$

$$\underset{\text{pseudoinversa}}{\underset{\parallel}{W^{-1}}} = diag\left(\frac{1}{w_1}, \dots, \frac{1}{w_r}, 0, \dots 0\right)$$

Pseudoinversa

Si definisce pseudoinversa di MOORE - PENROSE

A^+ di A se $rank(A) = n$

$$A^+ = (A^T A)^{-1} A^T \qquad (*)$$

che usando la SVD diventa

$$A^+ = V W^+ U^T$$

con $W^+ = diag\left(\frac{1}{w_1}, \dots, \frac{1}{w_r}, 0, \dots, 0\right)$

$(*)$ $Ax = b$ con A $m \times n$

se A ha $rank(A) = n$

$$A^T A x = A^T b \implies x = (A^T A)^{-1} A^T b$$

$$A^+$$

(ii)

$$\text{cond}_2(A) = \frac{w_{max}}{w_{min}}$$

detta norma spettrale (nel caso di matrici quadrate $w_{max} = \lambda_{max}$, $w_{min} = \lambda_{min}$)

$$\text{Cond}_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2$$

$$\|A\|_2 = \lambda_{max} \qquad \|A^{-1}\|_2 = 1/\lambda_{min}$$

(iii) Serve a risolvere sistemi sovradeterminati

$$Ax = b \qquad A \quad m \times n$$

$$\underbrace{A^T A}_{\substack{\text{simm.} \\ \text{def} +}} x = A^T b \longrightarrow \boxed{\begin{aligned} x^T A^T A x &= (Ax)^T Ax \\ &= \lambda x^T \lambda x \\ &= \lambda^2 x^T x > 0 \\ \text{essendo } x &\text{ autovettore} \end{aligned}}$$

$$x = \underbrace{(A^T A)^{-1} A^T}_{A^+} b$$

(iv) Calcolo della norma di Frobenius di A

$$\|A\|_F^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 = \text{tr}(A^T A)$$

Usando la SVD

$$\|A\|_F^2 = \sum_{i=1}^{r} w_i^2$$

(v) se A è quadrata e invertibile

$$w_j = \lambda_j$$

MATLAB

$$[U,S,V] = svd(A)$$

$S = svd(A)$ → S è il vettore dei valori singolari

$$[U,S,V] = svd(A,0)$$

decomposizione "ECONOMICA"

Cioè se A $m \times n$ $m > n$ allora solo le prime n colonne di U sono determinate, S è $n \times n$

se $m \leq n$ $svd(A,0)$ è equivalente $S = svd(A)$

Applicazione della SVD alla compressione di immagini

ma la vedremo la prossima lezione