Trasformazioni elementari di Givens (rotazioni elementari)

Si dice trasformazione elementare di Givens o matrice di rotazione elementare G_{ij} di ordine n una matrice che coincide con l'identità di ordine n eccetto nelle posizioni ij, ji, ii, jj, ove stanno due valori c, s dipendenti da un solo parametro ϕ:

$$G_{ij} = \begin{pmatrix} 1 & & & & & & & & & & \\ & \ldots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & c & & & & s & & & \\ & & & & 1 & & & & & & \\ & & & & & \ldots & & & & & \\ & & & & & & 1 & & & & \\ & & & -s & & & & c & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ldots & \\ & & & & & & & & & & 1 \end{pmatrix}$$

ove $c^2 + s^2 = 1$, $c = \cos\phi$, $s = \sin\phi$. Essa esprime una rotazione di ampiezza ϕ nell'iperpiano individuato dai versori e_i, $e_j \in \mathbb{R}^n$.

Si osserva che:

$$G_{ij}x = y \quad \text{ove} \quad \begin{cases} y_k = x_k & k \neq i, j \\ y_i = c.x_i + s.x_j & \\ y_j = -s.x_i + c.x_j & \end{cases}$$

E' sempre possibile trovare un valore di ϕ per cui $y_j = 0$. Basta trovare c ed s tale che

$$\begin{array}{c} c^2 + s^2 = 1 \\ y_j = 0 = -s.x_i + c.x_j \end{array} \qquad \left\{ \begin{array}{l} c = \frac{x_i}{\sqrt{x_i^2 + x_j^2}} \\ s = \frac{x_j}{\sqrt{x_i^2 + x_j^2}} \end{array} \right.$$

Formule più stabili

Se $|x_i| < |x_j|$, si pone $t = \frac{x_i}{x_j} \Rightarrow s = \frac{1}{\sqrt{1+t^2}}, \quad c = t.s$

Se $|x_j| < |x_i|$, si pone $t = \frac{x_j}{x_i} \Rightarrow c = \frac{1}{\sqrt{1+t^2}}, \quad s = t.c$

La complessità della trasformazione è di 4 prodotti.

Si osserva che premoltiplicare una matrice per G_{ij} significa sostituire alle righe i e j una loro combinazione lineare:

$$G_{ij}A = B \qquad \begin{cases} b_{kl} = a_{kl} & k \neq i, j; l = 1, ..., n \\ b_{il} = c.a_{il} + s.a_{jl} & l = 1, ..., n \\ b_{jl} = -s.a_{il} + c.a_{jl} & l = 1, .., n \end{cases}$$

Pertanto, si può dimostrare che:

$$G_{ij}G_{ij}^T = \Delta \qquad \delta_{kl} = (I)_{kl} \quad k \neq i, j; l = 1, n$$

$\delta_{ii} = c^2 + s^2 = 1$

$\delta_{ij} = -cs + cs = 0$

$\delta_{ji} = -sc + cs = 0$

$\delta_{jj} = s^2 + c^2 = 1$

$\Rightarrow G_{ij}$ è ortogonale. Infatti l'inversa di G_{ij} è G_{ij}^T.

Postmoltiplicare A per una matrice G_{ij} vuol dire sostituire alle colonne i e j una loro combinazione lineare:

$$AG_{ij} = B \qquad \begin{cases} b_{kl} = a_{kl} & l \neq i, j; k = 1, ..., n \\ b_{ki} = c.a_{ki} - s.a_{kj} & k = 1, ..., n \\ b_{kj} = s.a_{ki} + c.a_{kj} & k = 1, .., n \end{cases}$$

La complessità del prodotto di una matrice con una rotazione elementare è pari a $\mathcal{O}(4n)$ prodotti.

Data una matrice A, è possibile trovare G_{ij} tale che $B = G_{ij}A$ abbia l'elemento $b_{ji} = 0$.

$$\begin{array}{c} c^2 + s^2 = 1 \\ b_{ji} = 0 = -s.a_{ii} + c.a_{ji} \end{array} \qquad \left\{ \begin{array}{l} c = \frac{a_{ii}}{\sqrt{a_{ii}^2 + a_{ji}^2}} \\ s = \frac{a_{ji}}{\sqrt{a_{ii}^2 + a_{ji}^2}} \end{array} \right.$$

(è preferibile usare le formule stabili). Inoltre, con $\mathcal{O}(4n)$ prodotti, si trova:

$$\begin{array}{rl} b_{kl} = a_{kl} & k \neq i, j; l = 1, ..., n \\ b_{il} = c.a_{il} + s.a_{jl} & l = 1, ..., n \\ b_{jl} = -s.a_{il} + c.a_{jl} & l = 1, .., n \end{array}$$

Fattorizzazione QR

Teorema. Sia A una matrice $m \times n$. Esiste una matrice Q ortogonale di ordine m tale che $A = QR$ ove R è una matrice trapezoidale superiore $m \times n$. Inoltre $rango(A) = rango(R)$.

La dimostrazione dell'esistenza di Q è costruttiva.

Data $A \in \mathbb{R}^{m\times n}$, è possibile costruire $G_{12}, G_{13}, \ldots, G_{1m}$ tali che

$$G_{1m} \ldots G_{13} G_{12} A = \begin{pmatrix} x & x & \ldots & x \\ 0 & x & \ldots & x \\ 0 & x & \ldots & x \\ 0 & x & \ldots & x \end{pmatrix}$$

ove G_{12} annulla l'elemento $(2, 1)$, G_{13} annulla $(3, 1)$,...In seguito, premoltiplicando quanto ottenuto per opportune matrici:

$$G_{2m} \ldots G_{24} G_{23} (G_{1m} \ldots G_{13} G_{12} A)$$

si annulla la seconda colonna al di sotto della diagonale; premoltiplicando per:

$G_{3m} \ldots G_{35} G_{34}$	si annulla la terza colonna
$G_{4m} \ldots G_{45}$	si annulla la quarta colonna
$\ldots$	
$G_{rm} \ldots G_{rr+1}$	

ove, $r = \min(m - 1, n)$ si ottiene una matrice trapezoidale superiore. In conclusione

$$\prod_{i=r,\ldots,1;j>i} G_{ij} A = R$$

Se $m = n$, la complessità computazionale è dell'$\mathcal{O}(4n^3/3)$ prodotti e somme e dell'$\mathcal{O}(n^2/2)$ radici quadrate.

Inoltre, posto $Q^T = \prod_{i=r,\ldots,1;j>i} G_{ij}$, si ottiene

$$A = QR$$

Le rotazioni di Givens sono efficienti per ottenere la fattorizzazione QR di matrici sparse. Per esempio, per matrici tridiagonali sono sufficienti $n-1$ rotazioni di Givens e si ottiene una R triangolare superiore con solo tre diagonali non nulle:

$$G_{n-1n} \ldots G_{23} G_{12} A = R = \begin{pmatrix} x & x & x & \\ & x & x & x \\ & & x & x \\ & & & x \end{pmatrix}$$

Si ottiene la fattorizzazione con una complessità pari a $\mathcal{O}(12n)$ prodotti e somme e $\mathcal{O}(n-1)$ radici quadrate.

In generale,

$$Q = \prod_{i=1,\ldots,r;j>i} G_{ij}^T$$

e si può calcolare nel seguente modo:

- posto $Q = I$;
- per $i = 1, 2, \ldots, r$
-     per $j = i+1, \ldots, m$
-         $Q = Q * G_{ij}^T$
-     end;
- end;

Risoluzione di un sistema

$$Ax = b$$

Poichè

$$A = QR \Rightarrow QRx = b$$

$$Rx = Q^T b = \prod_{i=n-1,\dots,1;j>i} G_{ij} b$$

Esempio.

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

$$G_{12} = \begin{pmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

$$-2s + 1c = 0 \Rightarrow \begin{cases} \quad = 2/\sqrt{5} \\ s = 1/\sqrt{5} \end{cases}$$

$$G_{12}A = \begin{pmatrix} 5/\sqrt{5} & 4/\sqrt{5} & 1/\sqrt{5} \\ 0 & 3/\sqrt{5} & 2/\sqrt{5} \\ 0 & 0 & 1 \end{pmatrix} = R$$

$$Q = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} & 0 \\ 1/\sqrt{5} & 2/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Inversa di una matrice

Se A è fattorizzabile nella forma $A = QR$, occorre risolvere gli n sistemi

$$AX = I \Rightarrow QRX = I \Rightarrow RX = Q^T$$

Ciò comporta $4\mathcal{O}(n^3/3)$ prodotti per la fattorizzazione (matrici di Givens), e $\mathcal{O}(n^3/2)$ prodotti per la soluzione. Si può anche calcolare A^{-1} mediante l'inversione della matrice R ($\mathcal{O}(n^3/6)$ prodotti), eseguendo poi il prodotto dell'inversa di R con Q^T ($\mathcal{O}(n^3/2)$ prodotti):

$$A^{-1} = R^{-1}Q^T$$

Osservazioni

Sia A una matrice $m \times n$, di rango n $(m \geq n)$. Allora esiste una e una sola matrice Q_1 $m \times n$ a colonne ortonormali $(Q_1^T Q_1 = I)$ e una e una sola matrice R_1 triangolare superiore di ordine n a elementi diagonali positivi tale che $A = Q_1 R_1$.

Dim. Poichè $A^T A$ è simmetrica definita positiva (A è di rango n), per il teorema di Cholesky esiste una e una sola matrice R_1 triangolare superiore R_1 di ordine n a elementi diagonali positivi tale che

$$A^T A = R_1^T R_1$$

Da $R_1^{-T} A^T A = R_1$, posto $Q_1^T = R_1^{-T} A^T \Rightarrow Q_1 = A R_1^{-1}$, si prova che

- Q_1 è una matrice $m \times n$;
- $Q_1^T Q_1 = R_1^{-T} A^T A R^{-1} = R_1^{-T} R_1^T R_1 R_1^{-1} = I$, ossia Q_1 è a colonne ortonormali;
- Q_1 è unica; se esistesse Q_2 a colonne ortonormali tale che $A = Q_2 R_2$, allora $R_2 = R_1$ per l'unicità del fattore di Cholesky e $Q_2 = A R_1^{-1} = Q_1$ segue l'unicità di Q_1.

Poichè $A = QR$ per il teorema di fattorizzazione generale, segue che

$$A = (Q_1 \; Q_2) \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q_1 R_1 = (Q_1 D_1)(D_1 R_1)$$

ove D_1 è una matrice diagonale con $\delta_i = 1$ se $r_{ii} > 0$ e $\delta_i = -1$ se $r_{ii} < 0$. Allora la fattorizzazione $Q_1 R_1$, che è unica, è un caso particolare del teorema di fattorizzazione generale.

$$A = Q_1 R_1$$

$$(a_1 ... a_n) = (q_1 ... q_n) \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ & r_{22} & \dots & r_{2n} \\ & & \ddots & \\ & & & r_{nn} \end{pmatrix}$$

$$\begin{aligned} \Rightarrow a_1 &= q_1 r_{11} \\ a_2 &= q_1 r_{12} + q_2 r_{22} \\ &\dots \\ a_i &= q_1 r_{1i} + q_2 r_{2i} + \dots + q_i r_{ii} \\ &\dots \\ a_n &= q_1 r_{1n} + q_2 r_{2n} + \dots + q_n r_{nn} \end{aligned}$$

Per l'ortogonalità dei q_i segue:

$$\begin{aligned} r_{ij} &= q_i^T a_j \quad i \neq j \\ q_i &= \frac{a_i - \sum_{j=1}^{i-1} r_{ji} q_j}{\| a_i - \sum_{j=1}^{i-1} r_{ji} q_j \|}, \quad r_{ii} = \| a_i - \sum_{j=1}^{i-1} r_{ji} q_j \| \end{aligned}$$

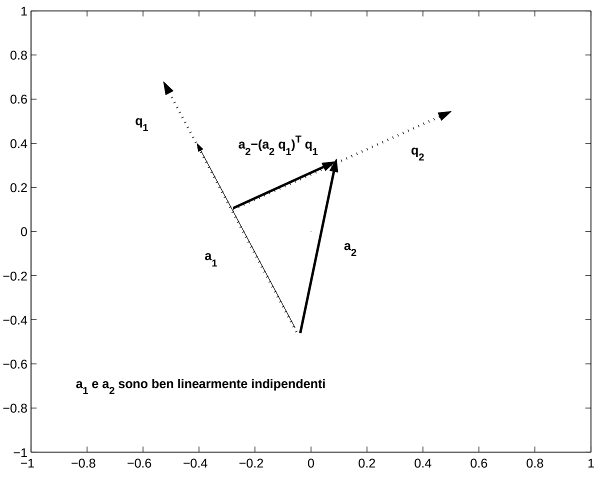
Pertanto si possono ricavare i q_i e gli r_{ij}:

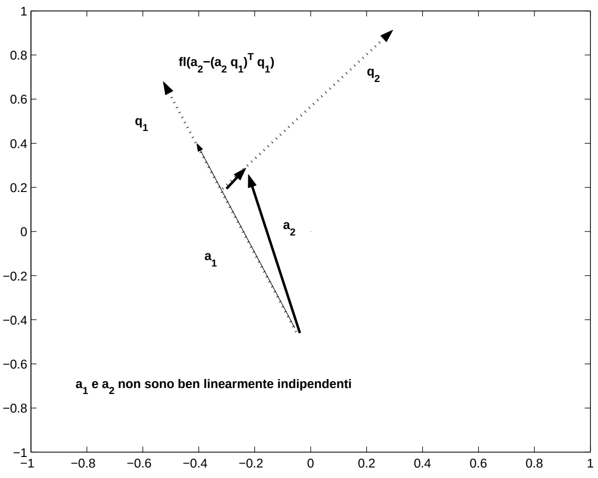
$$
\begin{array}{lll}
q_1' = a_1 & q_1 = \frac{q_1'}{\|q_1'\|} & r_{11} = \|q_1'\| \\
span\{a_1\} = span\{q_1\} & & \\
q_2' = a_2 - (a_2^T q_1) q_1 & q_2 = \frac{q_2'}{\|q_2'\|} & r_{12} = a_2^T q_1 \\
 & & r_{22} = \|q_2'\| \\
span\{a_1, a_2\} = span\{q_1, q_2\} & & \\
q_3' = a_3 - (a_3^T q_1) q_1 - (a_3^T q_2) q_2 & q_3 = \frac{q_3'}{\|q_3'\|} & r_{13} = a_3^T q_1 \\
 & & r_{23} = a_3^T q_2 \\
 & & r_{33} = \|q_3'\| \\
span\{a_1, a_2, a_3\} = span\{q_1, q_2, q_3\} & & \\
\ldots & &
\end{array}
$$

Processo di ortogonalizzazione di Gram-Schmidt: Dati n vettori linearmente indipendenti è sempre possibile generare n vettori ortonormali tali che

$$span\{a_1, \ldots, a_i\} = span\{q_1, \ldots, q_i\} \quad i = 1, \ldots, n$$

Tuttavia l'algoritmo di Gram Schmidt non è stabile se i vettori $q_1, \ldots, q_n$ non sono ben linearmente indipendenti. Per esempio se a_2 non è ben linearmente indipendente da a_1, $fl(a_2 - (a_2^T q_1) q_1)$ può non essere ortogonale a q_1 per un problema di cancellazione che amplifica l'errore commesso nel calcolo di $fl((a_2^T q_1) q_1)$ ($fl(a_2)$ e $fl((a_2^T q_1) q_1)$ sono circa uguali, ma il secondo termine può essere affetto da errore). In questo caso si ha che q_1 e $fl(q_2)$ non sono ortogonali. Di conseguenza $fl(Q_1)^T fl(Q_1) \neq I$.





Tuttavia, poichè Q_1 e R_1 sono unici, si possono determinare dal teorema di fattorizzazione generale che è stabile.