10 Metodi diretti per la risoluzione di sistemi lineari: fattorizzazione QR

10.1 Metodo di ortonormalizzazione di Gram-Schmidt

Per descrivere il metodo di ortonormalizzazione di Gram-Schmidt occorre la seguente definizione preliminare.

Definizione 10.1 *Vettori ortogonali*
Siano $\underline{x}, \underline{y} \in \mathcal{R}^n$. Si dice che che i vettori $\underline{x}$ e $\underline{y}$ sono ortogonali fra di loro se

$$\underline{x}^T \underline{y} = 0$$

Ora, siano $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_n \in \mathcal{R}^n$, n vettori linearmente indipendenti fra loro.
Si vogliono determinare $\underline{q}_1, \underline{q}_2, \ldots, \underline{q}_n \in \mathcal{R}^n$ vettori ortonormali, ossia tali che per ogni $i, j = 1, \ldots, n$

$$\underline{q}_i^T \underline{q}_j = 0 \quad \text{se } i \neq j$$

e per ogni $i = 1, \ldots, n$

$$\|\underline{q}_i\|_2 = 1$$

Il metodo di ortonormalizzazione di Gram-Schmidt procede nel modo seguente.

I passo. Si pone

$$\underline{q}_1 = \underline{a}_1 / \|\underline{a}_1\|_2$$

cosicché

$$\|\underline{q}_1\| = \|\underline{a}_1 / \|\underline{a}_1\|_2\|_2 = \|\underline{a}_1\|_2 / \|\underline{a}_1\|_2 = 1.$$

L'operazione è lecita in quanto il vettore $\underline{a}_1$ è diverso dal vettore nullo essendo gli $\underline{a}_i$ linearmente indipendenti fra loro, e quindi $\|\underline{a}_1\|_2 \neq 0$ (per le proprietà delle norme) e si può effettuare la normalizzazione.
Inoltre, i vettori $\underline{q}_1, \underline{a}_2, \ldots, \underline{a}_n$ sono a loro volta linearmente indipendenti fra loro.

$\Rightarrow A = [a_1, a_2 \ldots a_n]$
a_i l.i $\Leftrightarrow$ rank$(A) = n$
a_i sono basi di $R(A)$
$\Leftrightarrow$ null$(A) = \{0\}$
$a_i \notin$ null(A)

II passo. Si calcola la quantità

$$\alpha_1 = \underline{q}_1^T \underline{a}_2,$$

detta proiezione di $\underline{a}_2$ nella direzione di $\underline{q}_1$, e si pone

$$\tilde{\underline{q}}_2 = \underline{a}_2 - \alpha_1 \underline{q}_1 = a_2 - (q_1^T a_2) a_2$$

sottraiamo da a_2 le sue componente (proiez. nelle direzioni di q_1

cosicché i vettori $\tilde{\underline{q}}_2$ e $\underline{q}_1$ risultano essere ortogonali fra loro. Infatti, vale che

$$\begin{aligned}
\underline{q}_1^T \tilde{\underline{q}}_2 &= \underline{q}_1^T (\underline{a}_2 - \alpha_1 \underline{q}_1) \\
&= \underline{q}_1^T \underline{a}_2 - \alpha_1 \underline{q}_1^T \underline{q}_1 \\
&= \alpha_1 - \alpha_1 \|\underline{q}_1\|_2^2 \\
&= \alpha_1 - \alpha_1 \quad \text{essendo } \|\underline{q}_1\|_2 = 1 \\
&= 0.
\end{aligned}$$

Quindi si effettua la normalizzazione, ponendo

$$\underline{q}_2 = \tilde{\underline{q}}_2 / \|\tilde{\underline{q}}_2\|.$$

L'operazione è lecita in quanto il vettore $\tilde{\underline{q}}_2$ è diverso dal vettore nullo essendo i vettori $\underline{q}_1, \tilde{\underline{q}}_2, \underline{a}_3, \dots, \underline{a}_n$ linearmente indipendenti fra loro.

III passo. Si calcolano la quantità

$$\beta_1 = \underline{q}_1^T \underline{a}_3,$$

detta proiezione di $\underline{a}_3$ nella direzione di $\underline{q}_1$ e la quantità

$$\beta_2 = \underline{q}_2^T \underline{a}_3,$$

detta proiezione di $\underline{a}_3$ nella direzione di $\underline{q}_2$ e si pone

$$\tilde{\underline{q}}_3 = \underline{a}_3 - \beta_1 \underline{q}_1 - \beta_2 \underline{q}_2$$

cosicché il vettore $\tilde{\underline{q}}_3$ risulta essere ortogonale ai vettori $\tilde{\underline{q}}_2$ e $\underline{q}_1$.
Infatti, vale

$$\begin{aligned} \underline{q}_1^T \tilde{\underline{q}}_3 &= \underline{q}_1^T (\underline{a}_3 - \beta_1 \underline{q}_1 - \beta_2 \underline{q}_2) \\ &= \underline{q}_1^T \underline{a}_3 - \beta_1 \underline{q}_1^T \underline{q}_1 - \beta_2 \underline{q}_1^T \underline{q}_2 \\ &= \beta_1 - \beta 1 = 0, \end{aligned}$$

essendo $\underline{q}_1^T \underline{q}_1 = \|\underline{q}_1\|_2^2 = 1$ e $\underline{q}_1^T \underline{q}_2 = 0$ e vale pure

$$\begin{aligned} \underline{q}_2^T \tilde{\underline{q}}_3 &= \underline{q}_2^T (\underline{a}_3 - \beta_1 \underline{q}_1 - \beta_2 \underline{q}_2) \\ &= \underline{q}_2^T \underline{a}_3 - \beta_1 \underline{q}_2^T \underline{q}_1 - \beta_2 \underline{q}_2^T \underline{q}_2 \\ &= \beta_2 - \beta 2 = 0, \end{aligned}$$

$$q_1 = \frac{a_1}{\|a_1\|_2}$$

$$\tilde{q}_k = a_k - \sum_{i=1}^{k-1} (q_i^T a$$

$$q_k = \frac{\tilde{q}_k}{\|\tilde{q}_k\|_2}$$

essendo $\underline{q}_2^T \underline{q}_2 = \|\underline{q}_2\|_2^2 = 1$ e $\underline{q}_2^T \underline{q}_1 = 0$.
Quindi si effettua la normalizzazione, ponendo

$$\underline{q}_3 = \tilde{\underline{q}}_3 / \|\tilde{\underline{q}}_3\|$$

L'operazione è lecita in quanto il vettore $\tilde{\underline{q}}_3$ è diverso dal vettore nullo essendo i vettori $\underline{q}_1, \underline{q}_2, \tilde{\underline{q}}_3, \underline{a}_4, \dots, \underline{a}_n$ linearmente indipendenti fra loro.
Procedendo in questo modo, dopo n passi si ottengono i vettori $\underline{q}_1, \underline{q}_2, \dots, \underline{q}_n$ richiesti.

10.2 Metodo di fattorizzazione QR

La procedura di ortonormalizzazione di Gram-Schmidt appena illustrata, può essere reinterpretata come una procedura di fattorizzazione QR, ove Q è una matrice ortogonale (ossia tale che $QQ^T = Q^TQ = I$) e R è una matrice triangolare superiore.
In quest'ottica, occorre innanzitutto scrivere le relazioni che legano gli elementi

7.

j-1

di A a quelli della matrice prodotto QR, cui la prima viene uguagliata, ossia

$$
\begin{aligned}
a_{iK} &= \sum_{j=1}^{n} q_{ij} r_{jk} \\
&= \sum_{j=1}^{k} q_{ij} r_{jk}
\end{aligned}
$$

Prodotto righe x colonne

$$
\begin{aligned}
a_{ij} &= \sum_{k=1}^{m} q_{ik} r_{kj} \\
&= \sum_{k=1}^{j} q_{ik} r_{kj}
\end{aligned}
$$

$r_{kj} = 0 \quad k = j+1, \dots, m$

in quanto $r_{jk} = 0$ per $j = k+1, \dots, n$ essendo R triangolare superiore. Indicando con $\underline{a}_k$ la k-sima colonna della matrice A e riscrivendo tale formula in blocco per le colonne, si ha

$$
\begin{aligned}
\underline{a}_1 &= \underline{q}_1 r_{11} \\
\underline{a}_2 &= \underline{q}_1 r_{12} + \underline{q}_2 r_{22} \\
\underline{a}_3 &= \underline{q}_1 r_{13} + \underline{q}_2 r_{23} + \underline{q}_3 r_{33} \\
&\vdots \\
\underline{a}_k &= \underline{q}_1 r_{1k} + \underline{q}_2 r_{2k} + \ldots + \underline{q}_{k-1} r_{k-1k} + \underline{q}_k r_{kk} \\
&\vdots
\end{aligned}
$$

$A_{*j} = \sum_{k=1}^{m} q_{*k} r_{kj} \qquad a_j = \sum_{k=1} q \; r_k$

$a_1 = q_1 r_{11}$

$a_2 = q_1 r_{12} + q_2 r_{22}$

$a_k = q_1 r_{1k} + q_2 r_{2k} + q_{k-1} r_{k-1\,k} + q_k r_{kk}$

Quindi si può scrivere

$$
\underline{q}_k r_{kk} = \underline{a}_k - (\underline{q}_1 r_{1k} + \underline{q}_2 r_{2k} + \ldots + \underline{q}_{k-1} r_{k-1k}) = \underline{a}_k - \sum_{j=1}^{k-1} \underline{q}_j r_{jk}.
$$

Pertanto, scegliendo i coefficienti $r_{jk}, j = 1, \ldots, k-1$ come le proiezioni del vettore $\underline{a}_k$ nella direzione dei vettori $\underline{q}_j, j = 1, \ldots, k-1$, il vettore

$$
\tilde{\underline{q}}_k = \underline{a}_k - \sum_{j=1}^{k-1} \underline{q}_j r_{jk} \qquad a_k^T q_j
$$

risulta essere ortogonale ai vettori $\underline{q}_j$, $j = 1, \ldots, k-1$. Inoltre, scegliendo il coefficiente $r_{kk} = \|\tilde{\underline{q}}_k\|_2$ il vettore $\underline{q}_k = \tilde{\underline{q}}_k / r_{kk}$ risulta avere norma 2 unitaria. L'intera procedura corrisponde a quella precedentemente descritta come ortonormalizzazione di Gram-Schmidt.

Una volta ottenuta la fattorizzazione QR la si può utilizzare per la risoluzione del sistema lineare

$$
A\underline{x} = \underline{b},
$$

procedendo in modo analogo a quanto visto nel caso della fattorizzazione LU. Più precisamente, si ottiene che si devono risolvere i due sistemi lineari

$$
\begin{aligned}
Q\underline{y} &= \underline{b} \\
R\underline{x} &= \underline{y}
\end{aligned}
$$

Il vantaggio dell'introduzione della matrice Q al posto della matrice A originaria, consiste nel fatto che, essendo Q ortogonale, ossia valendo $QQ^T = Q^TQ = I$ si ha che, per unicità dell'inversa,

$$Q^{-1} = Q^T,$$

cosicché il vettore $\underline{y} = Q^{-1}\underline{b}$ viene ottenuto semplicemente considerando il prodotto matrice-vettore

$$\underline{y} = Q^T\underline{b}$$

(e non la risoluzione del sistema lineare).
Per quanto riguarda il sistema $R\underline{x} = \underline{y}$, essendo R una matrice triangolare superiore, esso viene risolto con la procedura di risoluzione backward, già vista nel caso della fattorizzazione LU.
Prima di procedere alla scrittura dell'algoritmo per la determinazione della fattorizzazione QR con il procedimento indicato, è opportuno riflettere sulla seguente questione: si può modificare l'ordine delle operazioni di normalizzazione così da rendere più agevole l'introduzione di un'eventuale tecnica di pivot per colonne.
In effetti, il risultato è inalterato se si procedere per "aggiornamenti" successivi delle colonne di A, cosicché al $k-$simo passo le prime k colonne saranno ortonormali fra loro, mentre le rimanenti colonne lo saranno già rispetto alle prime k, ma non fra di loro.
Si può quindi scrivere il seguente algoritmo

```
per k=1,...,n
    r_kk = ||a_k||_2
    a_k = a_k/r_kk

    per j=k+1,...,n
        r_kj = a_k^T a_j
        a_j = a_j - r_kj a_k
```

che trasforma la matrice A nella matrice Q e contemporaneamente costruisce la matrice R.
NO L'algoritmo, scritto in questo modo, permette l'introduzione di un'eventuale tecnica di pivot, in cui al passo k si sceglie la miglior colonna fra le colonne rimanenti, ossia dalla $k-$sima all' $n-$sima.
Il costo della fattorizzazione in termini di operazioni moltiplicative è dell'ordine di n^3; mentre quello di risoluzione è dell'ordine di $n^2 + n^2/2$.
Occorre sottolineare che l'algoritmo della fattorizzazione QR in genere utilizzato non è quello sopra descritto, esso risulta in genere molto stabile e questo motiva il maggior costo computazionale rispetto alla fattorizzazione LU.

SEGUE METODO QR x il calcolo autovalori

11 ~~Metodi diretti per la risoluzione di sistemi lineari: fattorizzazione LL^H~~

~~Sia $A \in C^{n\times n}$ una matrice hermitiana definita positiva, ossia tale che $A^H = A$ e per ogni $\underline{x} \neq \underline{0}$ e $\underline{x}^HA\underline{x} > 0$, vale il seguente teorema di fattorizzazione.~~