

Algoritmo QR per il calcolo degli autovalori

Studieremo dapprima l'algoritmo generale seguendo i seguenti passi:

1. algoritmo di base
2. Teorema di convergenza
3. riduzione del costo computazionale
4. convergenza in ipotesi più deboli
5. condizioni di arresto e riduzione dell'ordine
6. tecniche di traslazione dello spettro.

Vedremo poi come applicare questo algoritmo per il calcolo dei valori singolari.

1. Algoritmo di base (Francis, 1961)

È di una semplicità disarmante:

$$A_{k+1} = A_{k+1}$$

passo k

si fattorizza $A_k = Q_k R_k$

si calcola $A_{k+1} = R_k Q_k$

Le matrici sono tutte simili, infatti:

$$A_{k+1} = R_k Q_k = Q_k^T Q_k R_k Q_k = Q_k^T A_k Q_k$$

Il fatto singolare è che sotto determinate ipotesi la successione delle A_k converge ad una matrice triangolare superiore che, essendo simile alla matrice di partenza, contiene i suoi autovalori sulle diagonali.

~~Sp~~ $A_k \xrightarrow{k \rightarrow \infty} T$

Essendo $A_{n+1} = Q_n^T A_n Q_n$, se la matrice di partenza è simmetrica, tali saranno tutte le matrici della successione, e di conseguenza il limite sarà una matrice diagonale.

L'algoritmo fornisce la forma di Schur della matrice A , infatti

$$A_{n+1} = Q_n^T A_n Q_n = \dots Q_k^T \dots Q_1^T A Q_1 \dots Q_n$$

$$= U_n^T A U_n, \quad U_n = \prod_{j=1}^n Q_j$$

$$\lim_{n \rightarrow \infty} A_n = T = U^T A U, \quad U = \prod_{j=1}^{\infty} Q_j$$

~~diagonalizzabile~~
SCHUR

$$\Rightarrow \underline{A = U T U^T} \quad \underline{\text{FORMA DI SCHUR}} \quad (\text{esiste per ogni matrice})$$

Se A è simmetrica si ~~ottiene~~ ^{avrà} $T = D$, diagonale e quindi si ottiene la

$$\underline{A = U D U^T} \quad \underline{\text{Fattorizzazione Spettrale}}$$

con U matrice degli autovettori.

Per una A non simmetrica il calcolo degli autovettori si complica, anche perché la fattorizzazione potrebbe non esistere.

Per ottenere la fattorizzazione spettrale è sufficiente che la matrice A sia normale: $AA^T = A^T A$

2. Teorema di convergenza

Def matrice di fase $S = \text{diag}(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_n})$.

È una particolare matrice ortogonale. Osserviamo che S^*AS non modifica gli elementi diagonali di A , mentre apporta una modifica alle fase degli elementi non diagonali:

$$(S^*AS)_{kl} = a_{kl} e^{i(\theta_l - \theta_k)}$$

Teorema

Sia $A \in \mathbb{C}^{n \times n}$ dotata di autovalori distinti in modulo:

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$$

Sia X la matrice degli autovettori, cioè tale che

$$A = XDX^{-1}$$

con $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, e sia X^{-1} fattorizzabile LU senza pivoting:

$$X^{-1} = LU.$$

Esistono allora matrici di fase S_k tali che

$$\lim_{k \rightarrow \infty} S_k^* R_k S_k = I \quad \text{e} \quad \lim_{k \rightarrow \infty} S_k^* A_k S_{k-1} = T$$

dove T è una matrice triangolare superiore. Questo implica

$$\lim_{k \rightarrow \infty} S_{k-1}^* A_k S_{k-1} = T$$

e ~~ogni~~ gli elementi diagonali di T coincidono con gli autovalori λ_i di A . Quindi gli elementi diagonali di A_k tendono agli autovalori di A ($\text{diag}(A) \rightarrow (\lambda_1, \dots, \lambda_n)$ in ordine decrescente)
Se A è hermitiana, T è diagonale.

3. riduzione del costo computazionale

Il passo k dell'algoritmo richiederebbe n^3 operazioni per la fattorizzazione QR ed altrettanto per il prodotto matriciale

QR , conducendo ad un algoritmo di complessità troppo elevata.

Si può ottenere una forte riduzione osservando che la forma di Hessenberg è invariante per un passo dell'algoritmo QR:

$$\begin{pmatrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x & x \end{pmatrix} = \begin{pmatrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x & x \end{pmatrix} \begin{pmatrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x \end{pmatrix} \quad A_k = Q_k R_k$$

$$\begin{pmatrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x & x \end{pmatrix} \begin{pmatrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x \end{pmatrix} = \begin{pmatrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x & x \end{pmatrix} \quad Q_{k+1} = A_{k+1}$$

È possibile dunque effettuare il passo applicando $n-1$ trasformazioni di Householder o di Givens per annullare gli elementi della sottodiagonale, con una complessità $O(n^2)$.

Se inoltre A_k è tridiagonale (forma di Hessenberg simmetrica) anche A_{k+1} lo è, e la complessità del passo sono $O(n)$.

Se sono necessari anche gli autovettori è necessario accumulare anche le Q_k , e il costo computazionale resta n^2 anche per matrici simmetriche.

Per portare la matrice A in forma di Hessenberg sono necessarie $n-2$ trasformazioni di Householder, che si applicano una sola volta, prima di applicare il metodo iterativo.

$$A_H = H_{n-2} \dots H_2 H_1 A H_1 H_2 \dots H_{n-1} = Q^T A Q, \quad Q = H_1 \dots H_{n-2}$$

Se utilizziamo rotazioni di Givens, per l'algoritmo QR, si avrà:

$$R_n = \underbrace{G_{34} G_{23} G_{12}}_{Q_n^T} A_n$$

~~A_{n+1}~~

$$A_{n+1} = R_n \underbrace{G_{12}^T G_{23}^T G_{34}^T}_{Q_n}$$

Per l'associatività del prodotto possiamo anche scrivere:

$$A_{n+1} = G_{34} \left[G_{23} \left(G_{12} A_n G_{12}^T \right) G_{23}^T \right] G_{34}^T$$

In sostanza, una volta ottenuta una matrice di rotazione di Givens, possiamo applicarla ^{subito} a sinistra e a destra della matrice A_n e poi passare alle successive: non c'è bisogno di memorizzare tutte.

Osservando il modo in cui operano le matrici $G_{i,j+1}$ a sinistra e a destra è anche possibile dimostrare l'invarianza delle forme di Hessenberg.

4. Convergenza in ipotesi più deboli

Le ipotesi date per il teorema di convergenza sono molto stringenti, ma è possibile vedere che ribassandole il metodo resta applicabile, con qualche accorgimento aggiuntivo.

È stato richiesto, innanzitutto, che ^{l'inversa della} la matrice degli autovalori fosse fattorizzabile LU senza pivoting

$$X^{-1} = LU.$$

Se questo non è possibile, se cioè è necessario il pivoting, si può dimostrare che il metodo converge lo stesso, ma che gli autovalori di A non verranno ottenuti in ordine di modulo decrescente. Poco male.

L'ipotesi di autovalori distinti in modulo, invece, è essenziale. Infatti nel corso della dimostrazione si ottiene che gli elementi sottodiagonali (gli unici che ci interessano se partiamo in forma di Hessenberg) vanno a zero come $\left(\frac{|\lambda_{i+1}|}{|\lambda_i|}\right)^k$, con $|\lambda_{i+1}| < |\lambda_i|$.

Se per un certo r si ha $|\lambda_r| = |\lambda_{r+1}|$, il corrispondente elemento ~~sottodiagonale~~ non andrà a zero. È stato però osservato che la sottomatrice

$$A_r^{(k)} = \begin{bmatrix} a_{rr}^{(k)} & a_{r,r+1}^{(k)} \\ a_{r+1,r}^{(k)} & a_{r+1,r+1}^{(k)} \end{bmatrix}$$

pur non convergendo, ha autovalori che convergono a λ_r, λ_{r+1} .

Il resto della matrice, invece, converge regolarmente.

In caso di non convergenza, di conseguenza, è sufficiente analizzare i blocchi 2×2 che non convergono.

Questo capita, ad esempio, in presenza di autovalori complessi, o di autovalori reali uguali in modulo ed opposti in segno.

5. Condizioni di arresto e riduzione dell'ordine

È conveniente di partire in forma di Hessenberg, un buon criterio di arresto è

$$|a_{p+1,p}^{(k)}| < \tau (|a_{pp}^{(k)}| + |a_{p+1,p+1}^{(k)}|)$$

per tutti $i p = 1, 2, \dots, n-1$, dove τ è una tolleranza fissata

In realtà, se la condizione è verificata per un solo p , si ha

$$A_n = \begin{bmatrix} B_n & D_n \\ E & C_n \end{bmatrix} \begin{matrix} \} p \text{ righe} \\ \} n-p \end{matrix}$$

e dal punto di vista numerico possiamo supporre $E_n = O_{n-p,p}$.

Il problema si spezza dunque in due sottoproblemi: la ricerca degli autovalori delle matrici B_n e C_n , che possono essere trattati separatamente (e anche contemporaneamente \rightarrow parallelizzazione).

Questo consente di ridurre la complessità.

Il procedimento è detto DEFLAZIONE.

~~Questo~~ Potrebbe essere un problema gestire ~~sta~~ questa scomposizione caotica in sottoproblemi, dal punto di vista dell'implementazione.

6. Tecnica di traslazione dello spettro

La velocità di convergenza dell'algoritmo dipende dal numero

$$\max_{i=1, \dots, n-1} \left| \frac{\lambda_{i+1}}{\lambda_i} \right|.$$

Più ~~lontano~~ questo rapporto è vicino a 1, più la convergenza è lenta.

Supponiamo che al passo k -esimo sia disponibile una approssimazione

τ_k di un autovalore; ad esempio dell' n -esimo.

La matrice $A - \tau_k I$ avrebbe allora n autovalori $\{\lambda_i - \tau_k\}_{i=1}^n$

e l' n -esimo sarebbe prossimo a zero. In queste condizioni

l'elemento $a_{n,n-1}^{(k)}$ tenderebbe a zero molto in fretta.

L'algoritmo viene così modificato:

passo k

fattorizzazione

$$A_n - \tau_k I = Q_k R_k$$

prodotto

$$A_{n+1} = R_k Q_k + \tau_k I.$$

Infatti si ottiene

$$A_{n+1} = Q_k^T Q_k R_k Q_k + \tau_k I = Q_k^T A_n Q_k + \tau_k I$$

e le matrici della successione restano tutte simili.

Questo stratagemma conserva ovviamente le forme di Hessenberg (e la simmetria) e non modifica la complessità.

Una buona scelta per il parametro di shift τ_k è data da

$$\tau_k = \alpha_{nn}^{(k)}$$

Alternativamente si può prendere l'autovalore più prossimo ad $\alpha_{nn}^{(k)}$ della sottomatrice ~~precedente~~

$$\begin{bmatrix} \alpha_{n-1, n-1}^{(k)} & \alpha_{n-1, n}^{(k)} \\ \alpha_{n, n-1}^{(k)} & \alpha_{n, n}^{(k)} \end{bmatrix} \cdot \begin{array}{l} \text{Shift di} \\ \text{Wilkinson} \end{array}$$

In questo caso τ_k potrebbe essere complesso, e la tecnica di traslazione andrebbe modificata (si fanno due passi ~~alla~~ alla volta).

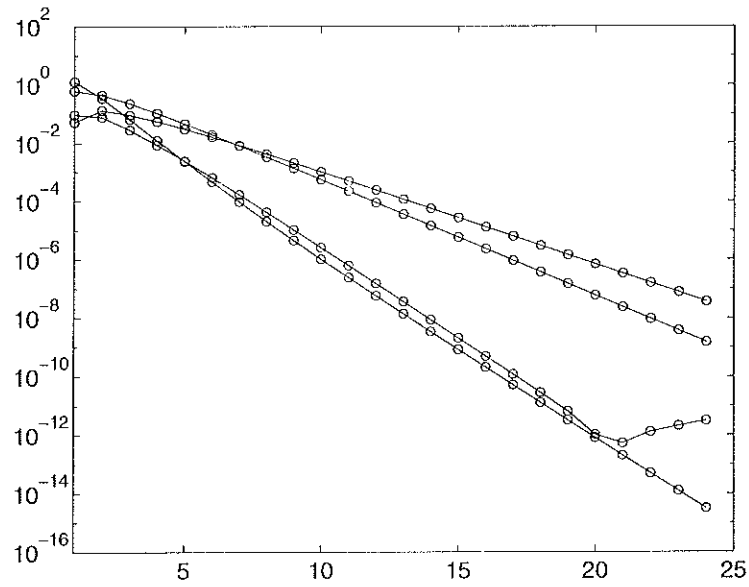
In entrambi i casi si può dimostrare che la convergenza a zero dell'elemento $\alpha_{n, n-1}^{(k)}$ è del terzo ordine: $O\left(\left|\frac{\lambda_n}{\lambda_{n-1}}\right|^3\right)$

Una volta che questo elemento verifica la condizione di stop, si può passare alla matrice di ordine $n-1$ ottenuta cancellando l'ultima riga e l'ultima colonna di A_k , ed applicare l'algoritmo QR con traslazione dello spettro in maniera analoga.

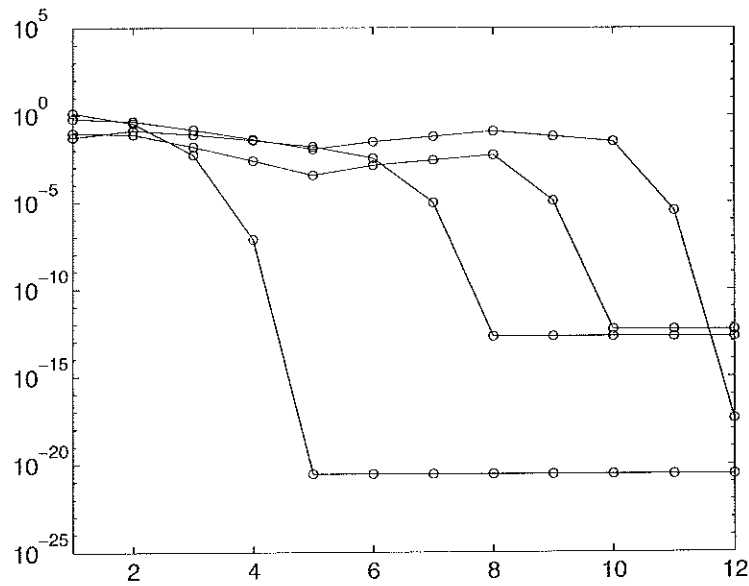
La traslazione dello spettro consente quindi anche di gestire le ridondanze dell'ordine in maniera ordinata.

Metodo QR per la determinazione degli autovalori

andamento della sottodiagonale per una matrice hermitiana
dimensione: 5, tolleranza richiesta: $1e-8$
autovalori reali e distinti



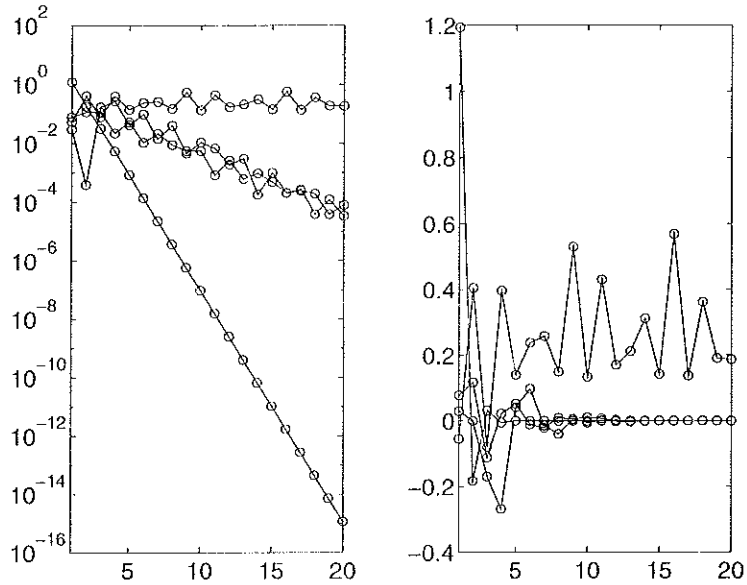
senza traslazione dello spettro



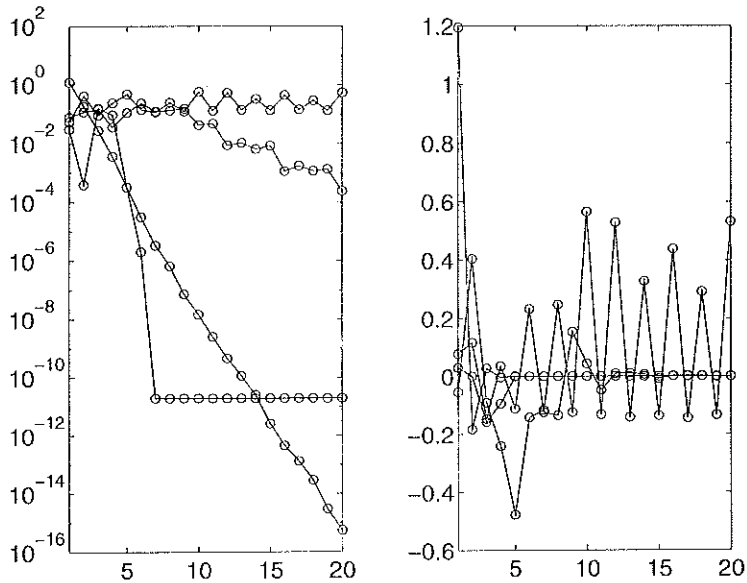
con traslazione dello spettro

Metodo QR per la determinazione degli autovalori

andamento della sottodiagonale per una matrice hermitiana
dimensione: 5, tolleranza richiesta: $1e-8$
3 autovalori reali e 2 complessi coniugati



senza traslazione dello spettro



con traslazione dello spettro