

Calcolo di autovalori e autovettori

Alvise Sommariva

Università degli Studi di Padova
Dipartimento di Matematica

9 maggio 2017

Autovalori

Il problema del calcolo degli autovalori di una matrice quadrata A di ordine n consiste nel trovare gli n numeri (possibilmente complessi) λ tali che

$$Ax = \lambda x, \quad x \neq 0 \quad (1)$$

Si osservi che a seconda delle esigenze

- talvolta è richiesto **solamente il calcolo di alcuni autovalori** (ad esempio quelli di massimo modulo, per determinare lo spettro della matrice),
- talvolta si vogliono determinare **tutti gli n autovalori** in \mathbb{C} .

Autovalori

Per semplicità, dopo i [teoremi di localizzazione di Gershgorin](#), mostreremo solo due metodi classici, uno per ognuna di queste classi, quello delle potenze e il metodo QR, rimandando per completezza alla monografia di Saad o a manuali di algebra lineare [1].

Nota.

Una interessante applicazione è l'algoritmo di [PageRank](#), utilizzato da Google per fornire i risultati migliori tra i siti web relativamente a certe parole chiave ed in prima approssimazione basato sul calcolo di un autovettore relativo all'autovalore 1 (ad esempio via metodo delle potenze) di una matrice stocastica di dimensioni enormi.

Teoremi di Gershgorin

In questo paragrafo mostriamo tre teoremi di localizzazione di autovalori dovuti a Gershgorin (cf. [1, p.76]).

Teorema (Primo teorema di Gershgorin, (Brauer, 1946))

Gli autovalori di una matrice A di ordine n sono tutti contenuti nell'unione dei cerchi di Gershgorin

$$K_i = \{z \in \mathbb{C} : |z - a_{i,i}| \leq \sum_{j=1, j \neq i}^n |a_{i,j}|\}$$

Teoremi di Gershgorin

Esempio

Vediamo quale esempio la matrice

$$A = \begin{pmatrix} 15 & -2 & 2 \\ 1 & 10 & -3 \\ -2 & 1 & 0 \end{pmatrix} \quad (2)$$

Il primo teorema di Gershgorin stabilisce che gli autovalori stanno nell'unione dei cerchi di Gershgorin

$$K_1 = \{z \in \mathbb{C} : |z - 15| \leq |-2| + |2| = 4\}$$

$$K_2 = \{z \in \mathbb{C} : |z - 10| \leq |1| + |-3| = 4\}$$

$$K_3 = \{z \in \mathbb{C} : |z - 0| \leq |-2| + |1| = 3\}$$

Teoremi di Gershgorin

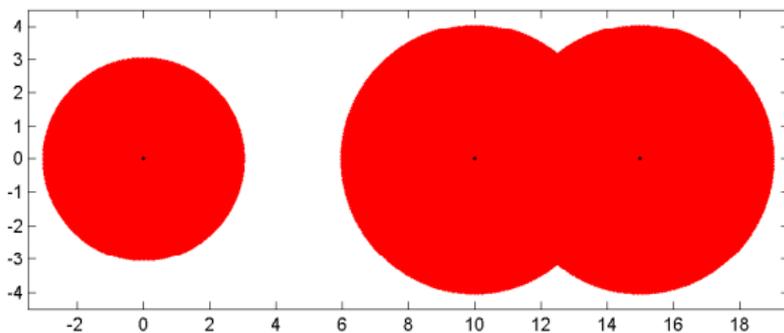


Figura : Cerchi di Gershgorin della matrice A definita in (5)

Teoremi di Gershgorin

Teorema (Secondo teorema di Gershgorin, (Brauer, 1948))

Se l'unione M_1 di k cerchi di Gershgorin è disgiunta dall'unione M_2 dei rimanenti $n - k$, allora k autovalori appartengono a M_1 e $n - k$ appartengono a M_2 .

Esempio

Relativamente a

$$A = \begin{pmatrix} 15 & -2 & 2 \\ 1 & 10 & -3 \\ -2 & 1 & 0 \end{pmatrix} \quad (3)$$

applicando il secondo teorema di Gershgorin, dal confronto con la figura abbiamo che un autovalore sta nel cerchio K_3 mentre due stanno nell'unione dei cerchi K_1, K_2 .

Teoremi di Gershgorin

Definizione

Una matrice di ordine $n \geq 2$ è **riducibile** se esiste una matrice di permutazione Π e un intero k , $0 < k < n$, tale che

$$B = \Pi A \Pi^T = \begin{pmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix}$$

in cui $A_{1,1} \in \mathbb{C}^{k \times k}$, $A_{2,2} \in \mathbb{C}^{(n-k) \times (n-k)}$.

Definizione

Una matrice si dice **irriducibile** se non è riducibile.

Teoremi di Gershgorin

Nota.

Per verificare se una matrice $A = (a_{i,j})$ sia irriducibile, ricordiamo che data una qualsiasi matrice, è possibile costruire un grafo avente come nodi gli indici della matrice. In particolare, il nodo i -esimo è connesso al nodo j -esimo se l'elemento $a_{i,j}$ è diverso da 0.

*Il grafo associato si dice **fortemente connesso** se per ogni coppia (i, j) posso raggiungere j a partire da i .*

*Una **matrice è irriducibile** se e solo se il grafo ad essa associata (detto di adiacenza) è **fortemente connesso**.*

In altre parole, una matrice è riducibile se e solo se il grafo di adiacenza ad esso associato non è fortemente connesso.

Teoremi di Gershgorin

Teorema (Terzo teorema di Gershgorin, (Brauer, 1948))

Se la matrice di ordine n è irriducibile e un autovalore λ sta sulla frontiera dell'unione dei cerchi di Gershgorin, allora sta sulla frontiera di ogni cerchio di Gershgorin.

Esempio

La matrice

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

è irriducibile, in quanto il grafico di adiacenza è fortemente connesso. Gli autovalori stanno nella disco centrato in $(2,0)$ e raggio 2 (primo teorema di Gershgorin), ma non sulla frontiera (terzo teorema di Gershgorin). Quindi è non singolare.

Teoremi di Gershgorin

Vediamo ora in Matlab quali sono effettivamente gli autovalori. si ha

```
>> A=[15 -2 2; 1 10 -3; -2 1 0]
A =
    15    -2     2
     1    10    -3
    -2     1     0

>> eig(A)
ans =
    0.5121
   14.1026
   10.3854

>>
```

a conferma di quanto stabilito dai primi due teoremi di Gershgorin.

Teoremi di Gershgorin

Nota.

- Ricordiamo che A è una matrice a coefficienti reali, allora A e A^T hanno gli stessi autovalori (cf. [1, p.47]) e quindi applicando i teoremi di Gershgorin alla matrice trasposta possiamo ottenere nuove localizzazioni degli autovalori.
- Nel caso A sia a coefficienti complessi, se λ è un autovalore di A allora il suo coniugato $\bar{\lambda}$ è autovalore della sua trasposta coniugata \bar{A} . Da qui si possono fare nuove stime degli autovalori di A .

Esercizio

Cosa possiamo dire relativamente agli autovalori di A se applichiamo i teoremi di Gershgorin ad A^T invece che ad A ?

Metodo delle potenze

Il **metodo delle potenze**, come vedremo, è particolarmente indicato per il calcolo dell'autovalore di massimo modulo di una matrice.

Sia A una matrice quadrata di ordine n con

- n autovettori x_1, \dots, x_n **linearmente indipendenti**,
- autovalori $\lambda_1, \dots, \lambda_n$ tali che

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|. \quad (4)$$

Metodo delle potenze

A tal proposito ricordiamo (cf. [2], p. 951) i seguenti risultati.

- Una matrice A è **diagonalizzabile** se e solo se possiede n **autovettori linearmente indipendenti**.
- Se tutti gli **autovalori di A sono distinti** la matrice è **diagonalizzabile**; l'opposto è ovviamente falso (si pensi alla matrice identica).
- Una matrice **simmetrica (hermitiana) è diagonalizzabile**. L'opposto è ovviamente falso, visto che la matrice

$$A = \begin{pmatrix} 15 & 0 \\ 1 & 10 \end{pmatrix} \quad (5)$$

è diagonalizzabile visto che ha tutti gli autovalori distinti ma non è simmetrica.

Metodo delle potenze

Teorema (Convergenza del metodo delle potenze, (Muntz, 1913))

Siano

- $A \in \mathbb{R}^{n \times n}$ una matrice diagonalizzabile con autovalori λ_k t.c.

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|.$$

- $u_k \neq 0$ autovettori relativi all'autovalore λ_k , cioè $Au_k = \lambda_k u_k$.
- $y_0 = \sum_k \alpha_k u_k$, $\alpha_1 \neq 0$.

La successione $\{y_s\}$ del metodo delle potenze definita da

$$y_{s+1} = Ay_s \tag{6}$$

converge ad un vettore parallelo a u_1 e il coefficiente di Rayleigh

$$\rho(y_s, A) := \frac{(y_s, Ay_s)}{(y_s, y_s)} \tag{7}$$

converge all'autovalore λ_1 .

Metodo delle potenze

Dimostrazione.

Per la dimostrazione si confronti [5, p.171]. Essendo la matrice A diagonalizzabile, esistono n autovettori u_k (relativi rispettivamente agli autovalori λ_k) che sono linearmente indipendenti e quindi formano una *base di \mathbb{R}^n* .

Sia $y_0 = \sum_k \alpha_k u_k$, $\alpha_1 \neq 0$.

Essendo $Au_k = \lambda_k u_k$ abbiamo

$$y_1 = Ay_0 = A\left(\sum_k \alpha_k u_k\right) = \sum_k \alpha_k Au_k = \sum_k \alpha_k \lambda_k u_k$$

$$y_2 = Ay_1 = A\left(\sum_k \alpha_k \lambda_k u_k\right) = \sum_k \alpha_k \lambda_k Au_k = \sum_k \alpha_k \lambda_k^2 u_k$$

e più in generale

$$y_{s+1} = Ay_s = A\left(\sum_k \alpha_k \lambda_k^s u_k\right) = \sum_k \alpha_k \lambda_k^s Au_k = \sum_k \alpha_k \lambda_k^{s+1} u_k.$$

Metodo delle potenze

Osserviamo ora che da $y_{s+1} = \sum_k \alpha_k \lambda_k^{s+1} u_k$ deduciamo

$$\frac{y_{s+1}}{\lambda_1^{s+1}} = \sum_k \alpha_k \frac{\lambda_k^{s+1}}{\lambda_1^{s+1}} u_k \quad (8)$$

per cui essendo $|\lambda_k| < |\lambda_1|$ qualora $k \neq 1$, necessariamente

$$\left| \frac{\lambda_k^{s+1}}{\lambda_1^{s+1}} \right| < 1, \quad k \neq 1$$

e di conseguenza

$$\lim_{s \rightarrow +\infty} \left(\frac{\lambda_k}{\lambda_1} \right)^s = 0$$

e quindi *la direzione di $\frac{y_s}{\lambda_1^s}$, che è la stessa di y_s , tende a quella dell'autovettore u_1 .*

Metodo delle potenze

Si osservi che il *coefficiente di Rayleigh* $\rho(\cdot, A) = (x, Ax)/(x, x)$ è *continuo in ogni* $x \neq 0$, $x \in \mathbb{R}^n$ in quanto tanto il numeratore quanto il denominatore sono funzioni (multivariate) polinomiali (quadratiche) delle componenti x_k di $x = (x_k)_k \in \mathbb{R}^n$, che sono appunto continue.

Per continuità, se $y_s/\lambda^s \rightarrow \alpha_1 u_1$ allora, essendo $\lambda_1 \neq 0$, da

$$\begin{aligned} \lim_s \rho(y_s, A) &:= \lim_s \frac{(y_s, Ay_s)}{(y_s, y_s)} = \lim_s \frac{(y_s/\lambda_1^s, A(y_s/\lambda_1^s))}{(y_s/\lambda_1^s, y_s/\lambda_1^s)} \\ &= \frac{(\alpha_1 u_1, A(\alpha_1 u_1))}{(\alpha_1 u_1, \alpha_1 u_1)} = \frac{(u_1, Au_1)}{(u_1, u_1)} = \lambda_1, \end{aligned} \quad (9)$$

ricaviamo che *il coefficiente di Rayleigh* $\rho(y_s, A)$ converge a λ_1 .

Metodo delle potenze

Nota.

Il metodo converge anche nel caso in cui

$$\lambda_1 = \dots = \lambda_r$$

per $r > 1$, tuttavia non è da applicarsi quando l'autovalore di modulo massimo non sia unico.

Nota.

In virtù di possibili underflow e overflow si preferisce normalizzare il vettore y_k precedente definito. Così l'algoritmo diventa

$$\begin{aligned} u_k &= At_{k-1} \\ t_k &= \frac{u_k}{\beta_k}, \quad \beta_k = \|u_k\|_2 \\ l_k &= \rho(t_k, A) \end{aligned} \tag{10}$$

dove $\rho(t_k, A)$ è il coefficiente di Rayleigh definito in (7).

Metodo delle potenze inverse

Una variante particolarmente interessante del metodo delle potenze, detta **delle potenze inverse** è stata scoperta da Wielandt nel 1944 ed è particolarmente utile nel caso in cui A sia una matrice quadrata con n autovettori linearmente indipendenti,

$$|\lambda_1| \geq |\lambda_2| \geq \dots > |\lambda_n| > 0. \quad (11)$$

e si desidera calcolare il **più piccolo autovalore in modulo**, cioè λ_n , applicando il metodo delle potenze ad A^{-1} .

Metodo delle potenze inverse

Si ottiene così la successione $\{t_k\}$ definita da

$$\begin{aligned} Au_k &= t_{k-1} \\ t_k &= \frac{u_k}{\beta_k}, \\ \beta_k &= \|u_k\|_2 \end{aligned}$$

e convergente ad un vettore parallelo a x_n . La successione di coefficienti di Rayleigh è tale che

$$\rho(t_k, A^{-1}) := \frac{(t_k, A^{-1}t_k)}{(t_k, t_k)} = \frac{(t_k, u_{k+1})}{(t_k, t_k)} \rightarrow 1/\lambda_n. \quad (12)$$

da cui è immediato calcolare λ_n .

Metodo delle potenze inverse

Nota.

- Se $\{\xi_i\}$ sono gli autovalori di A^{-1} con

$$|\xi_1| > |\xi_2| \geq |\xi_3| \geq \dots \geq |\xi_n|$$

allora il metodo delle potenze inverse calcola un'approssimazione di ξ_1 e di un suo autovettore x .

- Si osserva subito che se $A^{-1}x = \xi_i x$ (con $\xi_i \neq 0$) allora $x = A^{-1}x/\xi_i$ e $Ax = A(A^{-1}x/\xi_i) = \xi_i^{-1}x$ cioè ξ_i^{-1} è un autovalore di A e che se x è autovettore di A^{-1} relativo all'autovalore ξ_i , allora x è autovettore di A relativo all'autovalore ξ_i^{-1} . Conseguentemente se ξ_1 è l'autovalore di massimo modulo di A^{-1} e λ_n è l'autovalore di minimo modulo di A si ha $\lambda_n = \xi_1^{-1}$ e che

$$A^{-1}x = \xi_1 x \Rightarrow Ax = \xi_1^{-1}x = \lambda_n x.$$

Notiamo che il metodo delle potenze inverse, calcola $\xi_1 = \lambda_n^{-1}$ e il relativo autovettore x .

Metodo delle potenze inverse

Nota.

Per ottenere λ_n viene naturale calcolare $\xi_1^{-1} = (1/\lambda_1)^{-1}$, ma usualmente essendo x autovettore di A relativo a λ_n si preferisce calcolare λ_n via il coefficiente di Rayleigh

$$\rho(x, A) := \frac{(x, Ax)}{(x, x)}.$$

Quindi se $x_n \approx x$ è autovalore di ξ_1^{-1} allora

$$\rho(x_n, A) := \frac{(x_n, Ax_n)}{(x_n, x_n)} \approx \lambda_n.$$

Metodo delle potenze inverse

Metodo (Potenze inverse con shift, [4], p. 181)

In generale, fissato $\mu \in \mathbb{C}$ è possibile calcolare, se esiste unico, l'autovalore λ più vicino a μ tramite il seguente algoritmo:

$$\begin{aligned} (A - \mu I) z_k &= q_{k-1} \\ q_k &= z_k / \|z_k\|_2 \\ \sigma_k &= q_k^H A q_k. \end{aligned} \tag{13}$$

Metodo delle potenze inverse

- Ricordiamo che se λ è autovalore di A allora

$$Ax = \lambda x \Rightarrow (A - \mu I)x = \lambda x - \mu x = (\lambda - \mu)x$$

e quindi $\lambda - \mu$ è autovalore di $A - \mu I$.

- Il metodo delle potenze inverse applicato a $A - \mu I$ calcola il minimo autovalore $\sigma = \lambda - \mu$ in modulo di $A - \mu I$ cioè il σ che rende minimo il valore di $|\sigma| = |\lambda_i - \mu|$, dove λ_i sono gli autovalori di A .

Quindi essendo $\lambda_i = \sigma_i - \mu$ si ottiene pure il λ_i più vicino a μ .

- Per versioni più sofisticate di questa tecnica detta di *shift* (o in norma infinito invece che in norma 2) si confronti [1, p.379].

Problema. Si può applicare il metodo delle potenze inverse con shift μ nel caso μ sia proprio un autovalore di A ?

Metodo QR

Il **metodo QR**, considerato tra i 10 algoritmi più rilevanti del ventesimo secolo, cerca di calcolare tutti gli autovalori di una matrice A .

Lemma (Fattorizzazione QR)

Sia A una matrice quadrata di ordine n . Esistono

- Q **unitaria** (cioè $Q^T * Q = Q * Q^T = I$),
- R **triangolare superiore**

tali che

$$A = QR.$$

Citiamo alcune cose:

- La matrice A ha quale sola particolarità di essere quadrata. Nel caso generale però la sua fattorizzazione **QR in generale non è unica** bensì determinata a meno di una matrice di *fase* (cf. [1, p.149]) .
- Nel caso sia non singolare, allora tale fattorizzazione è unica qualora si chieda che i **coefficienti diagonali di R siano positivi**.
- La routine Matlab **qr** effettua tale fattorizzazione. Si consiglia di consultare l'`help` di Matlab, per consultare le particolarità di tale routine.

Metodo QR

- Se la matrice H è **simile** a K (cioè esiste una matrice non singolare S tale che $H = S^{-1}KS$) allora H e K hanno gli **stessi autovalori**.
- Si può vedere facilmente che la relazione di **similitudine è transitiva**, cioè se H_1 è simile ad H_2 e H_2 è simile ad H_3 allora H_1 è simile ad H_3 .

Il metodo QR venne pubblicato indipendentemente nel 1961 da Francis e da Kublanovskaya e successivamente implementato in EISPACK. Ci limiteremo a considerare versioni di base del metodo.

Metodo QR

Lemma

Sia

$$A_0 = A = Q_0 R_0$$

e

$$A_1 := R_0 Q_0.$$

Le matrici A_0 e A_1 sono simili e quindi hanno gli stessi autovalori.

Dimostrazione.

Basta notare che

$$Q_0 A_1 Q_0^T = Q_0 R_0 Q_0 Q_0^T = A_0$$

e quindi la matrice A_1 è simile ad A_0 (si ponga $S = Q_0^{-1} = Q_0^T$)

Metodo QR

Definiamo il seguente metodo, detto QR, dovuto a Francis (1961).

Metodo (QR)

$$\begin{aligned}A_k &= Q_k R_k \\ A_{k+1} &= R_k Q_k\end{aligned}$$

Per un lemma precedente A_{k+1} è simile ad A_k , che è simile ad A_{k-1}, \dots, A_0 . Quindi A_{k+1} essendo per transitività simile ad A_0 ha gli stessi autovalori di A_0 .

Metodo QR

Teorema (Convergenza metodo QR, [4], p. 169)

Se $A \in \mathbb{R}^{n \times n}$ ha *autovalori tutti distinti in modulo*, con

$$|\lambda_1| > \dots > |\lambda_n| \quad (14)$$

allora l'alg. QR converge a $A_\infty = (a_{i,j}^\infty)$ triangolare sup., cioè

$$\lim_k A_k = \begin{pmatrix} a_{1,1}^\infty & a_{1,2}^\infty & \dots & \dots & a_{1,n}^\infty \\ 0 & a_{2,2}^\infty & a_{2,3}^\infty & \dots & a_{2,n}^\infty \\ 0 & 0 & a_{3,3}^\infty & \dots & a_{3,n}^\infty \\ 0 & 0 & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{n-1,n-1}^\infty & a_{n-1,n}^\infty \\ 0 & 0 & 0 & 0 & a_{n,n}^\infty \end{pmatrix} \quad (15)$$

con $\lambda_k = a_{k,k}^\infty$.

Inoltre

- Se la condizione (14) *non è verificata* si può dimostrare che la successione $\{A_k\}$ tende a una forma *triangolare a blocchi*.
- se $A_k = (a_{ij}^{(k)})$, e $\lambda_{i-1} \neq 0$

$$|a_{i,i-1}^{(k)}| = \mathcal{O} \left(\frac{|\lambda_i|}{|\lambda_{i-1}|} \right)^k, \quad i = 2, \dots, n, \quad k \rightarrow \infty. \quad (16)$$

- se A è una *matrice Hessenberg superiore* allora l'algoritmo QR converge ad A_∞ triangolare a blocchi, simile ad A e con gli autovalori di ogni blocco diagonale tutti uguali in modulo.

Implementazione del metodo QR

Nelle implementazioni si calcola con un metodo scoperto da [Householder](#) (ma esiste un metodo alternativo dovuto a Givens) una matrice di Hessenberg T

$$T = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ 0 & a_{3,2} & a_{3,3} & \dots & a_{3,n} \\ 0 & 0 & a_{4,3} & \dots & a_{4,n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_{n,n-1} & a_{n,n} \end{pmatrix}$$

simile ad A ed in seguito si applica il metodo QR relativamente alla matrice T .

Nota.

Se A è simmetrica la matrice di Hessenberg T simile ad A risulta tridiagonale simmetrica.

Implementazione del metodo QR

Si può mostrare che

- se A è una matrice di Hessenberg superiore, allora $A = QR$ con Q di Hessenberg superiore.
- se A è tridiagonale allora $A = QR$ con Q di Hessenberg e R triangolare superiore con $r_{i,j} = 0$ qualora $j - i \geq 2$.
- le iterazioni mantengono la struttura, cioè
 - se $A_0 = T$ è di Hessenberg, allora A_k è di Hessenberg,
 - se $A_0 = T$ è tridiagonale allora A_k è tridiagonale.

Implementazione del metodo QR

Nota. (Operazioni trasformazioni in matrici di Hessenberg)

Il numero di moltiplicazioni necessarie

- *all'algoritmo di Givens per calcolare tale matrice T a partire da A è approssimativamente $10n^3/3$;*
- *all'algoritmo di Householder per calcolare tale matrice T a partire da A è approssimativamente $5n^3/3$.*

Nota. (Costo QR)

*Il **metodo QR** applicato ad una matrice A in forma di Hessenberg superiore ha ad ogni passo un costo di $2n^2$ operazioni moltiplicative.*

Bibliografia



D. Bini, M. Capovani e O. Menchi, *Metodi numerici per l'algebra lineare*, Zanichelli, 1988.



V. Comincioli, *Analisi Numerica, metodi modelli applicazioni*, Mc Graw-Hill, 1990.



G.H. Golub e C.F. Van Loan, *Matrix Computation, 3rd Edition*, The John Hopkins University Press 1996.



A. Quarteroni, R. Sacco, F. Saleri, *Matematica numerica*, 2001.



A. Quarteroni e F. Saleri, *Introduzione al calcolo scientifico*, Springer Verlag, 2006.