

Metodi numerici per il calcolo degli autovalori I^a parte

Calcolo Numerico - Ing. Inf. - Lezione 12

Outline

- 1 Metodo delle potenze
 - Deflazione
- 2 Metodo di Jacobi per matrici simmetriche

La conoscenza degli autovalori e degli autovettori di una matrice quadrata è richiesta non solo nell'ambito di importanti teorie della matematica, ma anche in molte applicazioni, nelle quali si deve disporre di una loro buona approssimazione numerica

Per stimare gli autovalori e gli autovettori di una matrice A sembrerebbe naturale ricorrere alla approssimazione delle radici dell'equazione caratteristica

$$\det(A - \lambda I) = 0,$$

usando i metodi studiati e successivamente, per ogni autovalore λ trovato, risolvere il sistema lineare omogeneo

$$(A - \lambda I)x = 0$$

Tuttavia, tranne qualche caso speciale, è sconsigliabile seguire tale via, a causa degli inevitabili errori che si introducono nel calcolo dei coefficienti della [equazione caratteristica](#)

Infatti piccole variazioni nei coefficienti possono comportare forti variazioni delle radici, giungendo talvolta a mutare radici reali in complesse e viceversa

Inoltre, anche si disponesse di un autovalore esatto λ , i metodi di ricerca degli autovettori associati a λ mediante la risoluzione del sistema lineare non sempre risultano di semplice applicazione

Introdurremo alcuni metodi iterativi più comunemente usati

Il primo di essi serve ad approssimare un autovalore di modulo dominante ed un autovettore ad esso associato

Gli altri approssimano simultaneamente tutti gli autovalori, sfruttando la loro invarianza rispetto alle trasformazioni per similitudine e si deducono gli autovettori in base alla nota relazione tra autovettori di matrici simili

Outline

- 1 Metodo delle potenze
 - Deflazione
- 2 Metodo di Jacobi per matrici simmetriche

Teorema - Metodo delle potenze

Sia $A \in \mathbb{C}^{n \times n}$ una matrice **diagonalizzabile** con autovalori soddisfacenti le condizioni

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

sia $z^{(0)} \in \mathbb{C}^n$ un vettore arbitrario non nullo **allora** il processo iterativo

$$y^{(0)} = z^{(0)}$$

$$y^{(k)} = Ay^{(k-1)}, \quad k = 1, 2, \dots,$$

è tale che

$$\lim_{k \rightarrow \infty} \frac{y^{(k)}}{y_j^{(k)}} = v, \quad \lim_{k \rightarrow \infty} \frac{y^{(k)H} Ay^{(k)}}{y^{(k)H} y^{(k)}} = \lambda_1$$

dove j è un indice per cui $y_j^{(k)} \neq 0$ e v è l'**autovettore** associato a λ_1

Dimostrazione

La diagonalizzabilità di A implica l'esistenza di n autovettori $x^{(i)}$, $i = 1, 2, \dots, n$, linearmente indipendenti e quindi che il vettore $z^{(0)}$ possa rappresentarsi nella forma $z^{(0)} = \sum_{i=1}^n c_i x^{(i)}$, dove, per quanto seguirà, supponiamo che sia $c_1 \neq 0$

Dall'algoritmo del [metodo delle potenze](#) si ha, per ricorrenza,

$$y^{(k)} = A y^{(k-1)} = A^2 y^{(k-2)} = \dots = A^k y^{(0)}$$

Dimostrazione

Segue

$$\begin{aligned}y^{(k)} &= A^k y^{(0)} = A^k (c_1 x^{(1)} + \dots + c_n x^{(n)}) \\&= c_1 A^k x^{(1)} + \dots + c_n A^k x^{(n)} \\&= c_1 \lambda_1^k x^{(1)} + \dots + c_n \lambda_n^k x^{(n)} \\&= \lambda_1^k \left(c_1 x^{(1)} + \sum_{i=2}^n c_i \left(\frac{\lambda_i}{\lambda_1} \right)^k x^{(i)} \right) \\&= \lambda_1^k (c_1 x^{(1)} + \omega^{(k)})\end{aligned}$$

con $\lim_{k \rightarrow +\infty} \omega^{(k)} = 0$

Dimostrazione

Di conseguenza risulta

$$\begin{aligned}y_j^{(k)} &= \lambda_1^k \left(c_1 x_j^{(1)} + \sum_{i=2}^n c_i \left(\frac{\lambda_i}{\lambda_1} \right)^k x_j^{(i)} \right) \\ &= \lambda_1^k \left(c_1 x_j^{(1)} + \omega_j^{(k)} \right)\end{aligned}$$

Scegliendo $y_j^{(k)} \neq 0$, dividendo membro a membro si ottiene

$$\lim_{k \rightarrow \infty} \frac{y^{(k)}}{y_j^{(k)}} = \frac{x^{(1)}}{x_j^{(1)}} = v$$

dove il vettore v è l'autovettore associato a λ_1

Dimostrazione

Infine si ha

$$\begin{aligned}\lim_{k \rightarrow \infty} \frac{y^{(k)H} A y^{(k)}}{y^{(k)H} y^{(k)}} &= \lim_{k \rightarrow \infty} \frac{\left(y^{(k)} / y_j^{(k)}\right)^H A \left(y^{(k)} / y_j^{(k)}\right)}{\left(y^{(k)} / y_j^{(k)}\right)^H \left(y^{(k)} / y_j^{(k)}\right)} \\ &= \frac{v^H A v}{v^H v} \\ &= \lambda_1\end{aligned}$$

per cui risulta dimostrata anche la seconda tesi

Criterio di arresto

Si ricordi che il rapporto

$$R(y^{(k)}) = \frac{y^{(k)H} A y^{(k)}}{y^{(k)H} y^{(k)}}$$

dicesi **quoziente di Rayleigh**

In pratica si adotta come criterio di arresto la condizione

$$|R(y^{(k)}) - R(y^{(k-1)})| < E$$

dove $E \in \mathbb{R}^+$ è un valore prefissato

Nel **Teorema delle potenze** l'ipotesi che A sia diagonalizzabile è, in generale, difficile da controllare (tale ipotesi è certamente verificata le **matrici normali**)

La convergenza del metodo delle potenze all'autovalore di modulo massimo e all'autovettore associato si può dimostrare anche se A non risulta diagonalizzabile, purché valga la condizione sugli autovalori

L'algoritmo del metodo delle potenze può dar luogo a fenomeni di overflow o di underflow in quanto può produrre vettori con componenti di valore assoluto eccessivamente grande o eccessivamente piccolo

Nella pratica si preferisce ricorrere a qualche forma modificata delle iterazioni del metodo delle potenze, introducendo una normalizzazione dei vettori

Metodo delle potenze normalizzato

Un algoritmo che genera una successione di vettori $z^{(k)}$ normalizzati è il seguente

$$\left. \begin{aligned} y^{(k)} &= Az^{(k-1)} \\ z^{(k)} &= \frac{y^{(k)}}{\alpha_k} \end{aligned} \right\} k = 1, 2, \dots,$$

dove è ancora $z^{(0)}$ arbitrario (non nullo) e α_k è una costante di normalizzazione opportuna

Se, ad esempio, α_k è una componente di massimo modulo di $y^{(k)}$ risulta

$$\|z^{(k)}\|_{\infty} = 1$$

Nelle ipotesi del [Teorema delle potenze](#), si può dimostrare che

$$\lim_{k \rightarrow \infty} z^{(k)} = w \quad \text{e} \quad \lim_{k \rightarrow \infty} R(z^{(k)}) = \lambda_1$$

dove w è l'autovettore associato a λ_1

Il Teorema del metodo delle potenze si può estendere al caso più generale in cui λ_1 abbia molteplicità $r \geq 1$, modificando l'ipotesi del teorema nella forma

$$\lambda_1 = \lambda_2 = \dots = \lambda_r, \quad |\lambda_1| > |\lambda_{r+1}| \geq \dots \geq |\lambda_n|$$

Si osservi che per $r = 1$ l'unico autovettore associato a λ_1 è approssimato da $z^{(k)}$

Se $r > 1$, $z^{(k)}$ approssima un autovettore appartenente allo spazio generato dagli r autovettori associati a λ_1 e l'autovettore approssimato cambia in dipendenza dalla scelta del vettore iniziale $z^{(0)}$

Nella ipotesi

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \geq \dots \geq |\lambda_n|$$

il metodo delle potenze consente il calcolo dell'autovalore λ_2 e del corrispondente autovettore utilizzando la conoscenza di λ_1 e $x^{(1)}$

Per semplicità, ci limiteremo al caso leggermente più restrittivo delle **matrici normali**, le quali hanno autovettori due a due ortogonali; supponendoli normalizzati a lunghezza unitaria, per essi risulta

$$x^{(i)H} x^{(j)} = \delta_{ij}, \quad 1 \leq i, j \leq n$$

Applicando il metodo nella forma normalizzata con $\alpha_k = \|y^{(k)}\|_2$, e ottenuti i valori (approssimati) di $x^{(1)}$ e λ_1 , si considera la matrice

$$A_1 = A - \lambda_1 x^{(1)} x^{(1)H}$$

Si constata facilmente che si ha

$$A_1 x^{(1)} = 0, \quad A_1 x^{(i)} = \lambda_i x^{(i)}, \quad i = 2, 3, \dots, n;$$

la matrice A_1 , quindi, ha autovalori

$$0, \lambda_2, \lambda_3, \dots, \lambda_n$$

e gli stessi autovettori di A

Il metodo, applicato ora a A_1 , converge all'autovettore $x^{(2)}$ e all'autovalore λ_2

Questo procedimento, detto **deflazione**, può essere ripetuto, nell'ipotesi $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$, per calcolare tutti gli autovalori e tutti gli autovettori, considerando successivamente le matrici

$$A_i = A - \sum_{r=1}^i \lambda_r x^{(r)} x^{(r)H}, \quad i = 1, 2, \dots, n-1$$

Outline

- 1 Metodo delle potenze
 - Deflazione
- 2 Metodo di Jacobi per matrici simmetriche

Il **metodo di Jacobi** permette di approssimare tutti gli autovalori di una matrice hermitiana. Per semplicità considereremo solo matrici **A reali e simmetriche**

Il metodo consiste nell'operare successive trasformazioni per similitudine mediante matrici di rotazione G_{rt} della forma

$$G_{rt} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \cos \varphi & & & \\ & & & & 1 & & \\ & & & & & -\sin \varphi & \\ & & & & & & \ddots & \\ & & & & & & & 1 & \\ & & & & & \sin \varphi & & & \cos \varphi & \\ & & & & & & & & & 1 & \\ & & & & & & & & & & \ddots & \\ & & & & & & & & & & & 1 \end{pmatrix} \begin{matrix} \leftarrow r \\ \\ \\ \\ \leftarrow t \\ \\ \end{matrix}$$

Le matrici G_{rt} sono determinate dagli interi r ed t , con $1 \leq r < t \leq n$, e dal parametro φ ; inoltre, come abbiamo già visto, sono ortogonali, cioè si ha $G_{rt}G_{rt}^T = G_{rt}^T G_{rt} = I$ e quindi $G_{rt}^{-1} = G_{rt}^T$

Il **metodo di Jacobi** è un metodo iterativo in cui al passo k -esimo si trasforma una matrice A_k mediante una matrice ortogonale $G_{rt}^{(k)}$ secondo l'algoritmo

$$A_1 := A$$

$$A_{k+1} = G_{rt}^{(k)T} A_k G_{rt}^{(k)}, \quad k = 1, 2, \dots,$$

dove gli indici r ed t variano al variare di k

Nella versione classica del metodo si scelgono gli indici di $G_{rt}^{(k)}$ coincidenti con quelli di un elemento non diagonale $a_{rt}^{(k)}$ di A_k avente modulo massimo e il valore di φ viene determinato in modo che nella matrice trasformata A_{k+1} risulti

$$a_{rt}^{(k+1)} = 0$$

La trasformazione $A_{k+1} = G_{rt}^{(k)\top} A_k G_{rt}^{(k)}$ lascia inalterati tutti gli elementi di A_k non appartenenti alle righe e alle colonne di indici r ed t , cioè si ha

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} \quad \text{per } i, j \neq r, t,$$

mentre vengono modificati solo gli elementi delle righe e colonne di indici r e t e inoltre conserva la simmetria iniziale

Risulta

$$a_{rt}^{(k+1)} = a_{tr}^{(k+1)} = - \left(a_{rr}^{(k)} - a_{tt}^{(k)} \right) \sin \varphi \cos \varphi + a_{rt}^{(k)} (\cos^2 \varphi - \sin^2 \varphi)$$

$$a_{rj}^{(k+1)} = a_{jr}^{(k+1)} = a_{rj}^{(k)} \cos \varphi + a_{tj}^{(k)} \sin \varphi, \quad j \neq r, t,$$

$$a_{tj}^{(k+1)} = a_{jt}^{(k+1)} = -a_{rj}^{(k)} \sin \varphi + a_{tj}^{(k)} \cos \varphi, \quad j \neq r, t,$$

$$a_{rr}^{(k+1)} = a_{rr}^{(k)} \cos^2 \varphi + 2a_{rt}^{(k)} \sin \varphi \cos \varphi + a_{tt}^{(k)} \sin^2 \varphi,$$

$$a_{tt}^{(k+1)} = a_{rr}^{(k)} \sin^2 \varphi - 2a_{rt}^{(k)} \sin \varphi \cos \varphi + a_{tt}^{(k)} \cos^2 \varphi$$

Dalla relazione che definisce $a_{rt}^{(k+1)}$ segue che tale valore risulta nullo per ogni φ soluzione dell'equazione

$$-\left(a_{rr}^{(k)} - a_{tt}^{(k)}\right) \sin \varphi \cos \varphi + a_{rt}^{(k)} (\cos^2 \varphi - \sin^2 \varphi) = 0$$

Si ricavano due valori, $\sin \varphi$ e $\cos \varphi$, per costruire la matrice $G_{rt}^{(k)}$, scrivendo la precedente equazione nella forma

$$y^2 + 2my - 1 = 0,$$

dove si è posto $y = \tan \varphi$, $m := (a_{rr}^{(k)} - a_{tt}^{(k)})/2a_{rt}^{(k)}$

Scegliendo fra le due radici della equazione di secondo grado quella di modulo minore, corrispondente a $|\varphi| \leq \pi/4$, si ha

$$y = \begin{cases} -m + \sqrt{1 + m^2} & \text{se } m > 0, \\ -m - \sqrt{1 + m^2} & \text{se } m < 0; \end{cases}$$

per $m = 0$ si sceglie $y = 1$

In ogni caso si ottiene

$$\cos \varphi = \frac{1}{\sqrt{1 + y^2}}, \quad \sin \varphi = y \cos \varphi,$$

con cui si costruisce la matrice $G_{rt}^{(k)}$ voluta

In questo modo ad ogni passo si annulla un elemento non diagonale (e il suo simmetrico); ciò non esclude che nei passi successivi un elemento non diagonale nullo possa essere modificato e assumere un valore non nullo

Teorema (di Jacobi)

La successione $\{A_k\}$ generata col **metodo di Jacobi classico** converge alla matrice diagonale

$$D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di A

Criterio di arresto

In base al precedente teorema, un criterio di arresto delle iterazioni potrebbe essere il verificarsi della condizione

$$\max_{i>j} |a_{ij}^{(k+1)}| \leq E$$

dove $E > 0$ è un numero prefissato

Una variante del metodo di Jacobi classico consiste nel sopprimere la ricerca dell'elemento di modulo massimo da annullare ad ogni passo, evitando così tutti i necessari confronti fra elementi

Ciò si ottiene annullando sistematicamente tutti gli elementi non nulli che si incontrano percorrendo per righe gli elementi al disopra della diagonale principale, cioè quelli di indici

$$(12), (13), \dots, (1n); (23), \dots, (2n); \dots; (n-1, n)$$

L'operazione è ciclica nel senso che si ripete eseguendo gli annullamenti sempre nello stesso ordine

Per questo metodo, noto come **metodo di Jacobi ciclico**, si può dimostrare un teorema di convergenza analogo al Teorema di Jacobi