

Capitolo 5

Calcolo di autovalori e autovettori

La conoscenza degli autovalori e degli autovettori di una matrice quadrata (cfr. Capitolo 2) 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

$$\det(A - \lambda I) = 0, \quad (5.1)$$

usando i metodi studiati nel Capitolo 4 e successivamente, per ogni autovalore λ trovato, risolvere il sistema lineare omogeneo

$$(A - \lambda I)x = 0. \quad (5.2)$$

Tuttavia, tranne qualche caso speciale, (cfr. Complementi ed esempi del Capitolo 2 e Capitolo 4) è sconsigliabile seguire tale via, a causa degli inevitabili errori che si introducono nel calcolo dei coefficienti della (5.1). Infatti piccole variazioni nei coefficienti della (5.1) possono comportare forti variazioni delle radici, giungendo talvolta a mutare radici reali in complesse e viceversa. Inoltre, quand'anche si disponesse di un autovalore esatto λ , i metodi di ricerca degli autovettori associati a λ mediante la risoluzione del sistema (5.2) non sempre risultano di semplice applicazione.

Nei paragrafi che seguono sono esposti alcuni metodi iterativi più comunemente usati. Il primo di essi serve ad approssimare solo 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; si deducono poi gli autovettori in base alla nota relazione tra autovettori di matrici simili (cfr. Teorema 2.7.2).

5.1 Metodo delle potenze

Il metodo delle potenze si fonda sul seguente teorema.

Teorema 5.1.1 *Sia $A \in \mathcal{C}^{n \times n}$ una matrice diagonalizzabile con autovalori soddisfacenti le condizioni*

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|; \quad (5.3)$$

sia $z^{(0)} \in \mathcal{C}^n$ un vettore arbitrario; allora il processo iterativo

$$\begin{aligned} y^{(0)} &= z^{(0)} \\ y^{(k)} &= Ay^{(k-1)}, \quad k = 1, 2, \dots, \end{aligned} \quad (5.4)$$

è tale che

$$\lim_{k \rightarrow \infty} \frac{y^{(k)}}{y_j^{(k)}} = v, \quad \lim_{k \rightarrow \infty} \frac{y^{(k)H} Ay^{(k)}}{y^{(k)H} y^{(k)}} = \lambda_1, \quad (5.5)$$

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 è lecito supporre che sia $c_1 \neq 0$. Dalla (5.4) segue quindi

$$\begin{aligned} y^{(k)} &= A^k y^{(0)} = A^k (c_1 x^{(1)} + \dots + c_n x^{(n)}) \\ &= c_1 A^k x^{(1)} + \dots + c_n A^k x^{(n)} = c_1 \lambda_1^k x^{(1)} + \dots + c_n \lambda_n^k x^{(n)} \\ &= \lambda_1^k \left(c_1 x^{(1)} + \sum_{i=2}^n c_i \left(\frac{\lambda_i}{\lambda_1} \right)^k x^{(i)} \right) \end{aligned} \quad (5.6)$$

e anche, per ogni indice j ,

$$y_j^{(k)} = \lambda_1^k \left(c_1 x_j^{(1)} + \sum_{i=2}^n c_i \left(\frac{\lambda_i}{\lambda_1} \right)^k x_j^{(i)} \right). \quad (5.7)$$

In particolare, scegliendo $y_j^{(k)} \neq 0$, tenendo conto dell'ipotesi (5.3) e dividendo membro a membro la (5.6) per la (5.7), si ottiene

$$\lim_{k \rightarrow \infty} \frac{y^{(k)}}{y_j^{(k)}} = b_1 x^{(1)} = v, \quad (5.8)$$

dove $b_1 = 1/x_j^{(1)}$; perciò il vettore v è l'autovettore associato a λ_1 , come afferma la tesi. Si ha quindi

$$Av = \lambda_1 v$$

e anche

$$v^H Av = \lambda_1 v^H v$$

da cui

$$\frac{v^H Av}{v^H v} = \lambda_1;$$

infine, tenendo conto della (5.8), si ha

$$\lim_{k \rightarrow \infty} \frac{y^{(k)H} Ay^{(k)}}{y^{(k)H} y^{(k)}} = \lim_{k \rightarrow \infty} \frac{\left(y^{(k)}/y_j^{(k)}\right)^H A \left(y^{(k)}/y_j^{(k)}\right)}{\left(y^{(k)}/y_j^{(k)}\right)^H \left(y^{(k)}/y_j^{(k)}\right)} = \frac{v^H Av}{v^H v} = \lambda_1$$

per cui risulta dimostrata anche la seconda delle (5.5). \square

Il rapporto

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

dicesi *quoziente di Rayleigh*.

Dalle relazioni (5.6) e (5.7) si deduce che gli errori $\frac{y^{(k)}}{y_j^{(k)}} - v$ ed $R(y^{(k)}) - \lambda_1$ tendono a zero come $\left(\frac{\lambda_2}{\lambda_1}\right)^k$.

Nel Teorema 5.1.1 l'ipotesi che A sia diagonalizzabile è, in generale, difficile da controllare; tuttavia tale ipotesi è certamente verificata per la vasta classe delle matrici normali, che sono riconoscibili in base alla proprietà $A^H A = A A^H$ (cfr. Capitolo 2). La convergenza del metodo delle potenze all'autovettore di modulo massimo e all'autovettore associato si può dimostrare anche se A non è diagonalizzabile, purché valga la condizione (5.3).

Un algoritmo basato sulla (5.4) 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. Pertanto si preferisce

ricorrere a qualche forma modificata delle iterazioni (5.4), introducendo una normalizzazione dei vettori. Un algoritmo che genera una successione di vettori $\{z^{(k)}\}$ normalizzati è il seguente,

$$\left. \begin{aligned} y^{(k)} &= Az^{(k-1)} \\ z^{(k)} &= \frac{y^{(k)}}{\alpha_k} \end{aligned} \right\} k = 1, 2, \dots, \quad (5.9)$$

dove è ancora $z^{(0)}$ arbitrario e α_k è una costante di normalizzazione opportuna. Se, ad esempio, α_k è una componente di massimo modulo di $y^{(k)}$, scelta, a partire da un certo k , sempre con lo stesso indice, risulta $\|z^{(k)}\|_\infty = 1$.¹

Nelle ipotesi del Teorema 5.1.1, si dimostra che

$$\lim_{k \rightarrow \infty} z^{(k)} = w \quad \text{e} \quad \lim_{k \rightarrow \infty} R(z^{(k)}) = \lambda_1 \quad (5.10)$$

dove w è l'autovettore associato a λ_1 .

In pratica si possono quindi assumere $R(z^{(k)})$ e $z^{(k)}$ come approssimazioni rispettivamente di λ_1 e dell'autovettore associato, adottando come criterio di arresto la condizione $|R(z^{(k)}) - R(z^{(k-1)})| < \epsilon$ con $\epsilon > 0$ prefissato.

Il Teorema 5.1.1 si può estendere al caso più generale in cui λ_1 abbia molteplicità $r \geq 1$, modificando l'ipotesi (5.3) 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 $z^{(k)}$; mentre, 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 $z^{(0)}$.

Nell'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 si limita qui 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 (5.9) con $\alpha_k = \|y^{(k)}\|_2$, e ottenuti i valori (approssimati) di $x^{(1)}$ e λ_1 , si considera la matrice

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

¹Per la (5.6) gli indici delle componenti di modulo massimo di $y^{(k)}$, per k sufficientemente grande, coincidono con gli indici delle componenti di massimo modulo di $x^{(1)}$.

Si constata subito che si ha

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

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

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

5.2 Metodo delle potenze inverse

La matrice A di ordine n sia diagonalizzabile ed abbia un solo autovalore di modulo minimo, cioè sia

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{n-1}| > |\lambda_n| > 0;$$

la matrice inversa A^{-1} avrà quindi autovalori verificanti la condizione

$$\left| \frac{1}{\lambda_n} \right| > \left| \frac{1}{\lambda_{n-1}} \right| \geq \dots \geq \left| \frac{1}{\lambda_1} \right|.$$

Ne segue che il metodo delle potenze, applicato alla matrice inversa A^{-1} , fornisce l'autovalore di massimo modulo $1/\lambda_n$ e, di conseguenza, anche l'autovalore di minimo modulo di A . Dalle (5.9) si ha, per $z^{(0)}$ arbitrario,

$$y^{(k)} = A^{-1} z^{(k-1)}, \quad z^{(k)} = \frac{y^{(k)}}{\alpha_k}, \quad k = 1, 2, \dots$$

In pratica si usa il seguente algoritmo, detto *metodo delle potenze inverse*,

$$Ay^{(k)} = z^{(k-1)}, \quad z^{(k)} = \frac{y^{(k)}}{\alpha_k}, \quad k = 1, 2, \dots, \quad (5.11)$$

nel quale ad ogni passo occorre risolvere un sistema lineare di matrice A per ottenere $y^{(k)}$. Questo metodo può essere vantaggiosamente usato per approssimare un qualunque autovalore λ_j di A quando se ne conosca già una

approssimazione iniziale $\tilde{\lambda}_j$. Partendo dall'osservazione che gli autovalori della matrice $A - \tilde{\lambda}_j I$ sono

$$\lambda_1 - \tilde{\lambda}_j, \dots, \lambda_j - \tilde{\lambda}_j, \dots, \lambda_n - \tilde{\lambda}_j,$$

mentre quelli della matrice $(A - \tilde{\lambda}_j I)^{-1}$ sono

$$\frac{1}{\lambda_1 - \tilde{\lambda}_j}, \dots, \frac{1}{\lambda_j - \tilde{\lambda}_j}, \dots, \frac{1}{\lambda_n - \tilde{\lambda}_j},$$

se $\tilde{\lambda}_j$ è abbastanza vicino a λ_j si può supporre che $\frac{1}{\lambda_j - \tilde{\lambda}_j}$ sia l'autovalore di massimo modulo per $(A - \tilde{\lambda}_j I)^{-1}$ e quindi si può applicare a questa matrice l'algoritmo (5.11) che ora assume la forma:

$$(A - \tilde{\lambda}_j I)y^{(k)} = z^{(k-1)}, \quad z^{(k)} = \frac{y^{(k)}}{\alpha_k}, \quad k = 1, 2, \dots \quad (5.12)$$

Arrestando, per esempio, le iterazioni per $k = m$, si ha

$$R(z^{(m)}) \simeq \frac{1}{\lambda_j - \tilde{\lambda}_j}$$

da cui si ottiene

$$\lambda_j \simeq \frac{1 + \tilde{\lambda}_j R(z^{(m)})}{R(z^{(m)})}.$$

Si noti che il vettore $z^{(m)}$ ottenuto da (5.12) approssima l'autovettore associato all'autovalore $\frac{1}{\lambda_j - \tilde{\lambda}_j}$ per la matrice $(A - \tilde{\lambda}_j I)^{-1}$, ma $z^{(m)}$ è anche un autovettore approssimato associato all'autovalore λ_j per la matrice A .

5.3 Metodo di Jacobi per matrici simmetriche

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

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

Precisamente si ha:

$$a_{rs}^{(k+1)} = a_{sr}^{(k+1)} = -\left(a_{rr}^{(k)} - a_{ss}^{(k)}\right) \sin \varphi \cos \varphi + a_{rs}^{(k)} \left(\cos^2 \varphi - \sin^2 \varphi\right), \quad (5.15)$$

$$\begin{aligned} a_{rj}^{(k+1)} &= a_{jr}^{(k+1)} = a_{rj}^{(k)} \cos \varphi + a_{sj}^{(k)} \sin \varphi, & j \neq r, s, \\ a_{sj}^{(k+1)} &= a_{js}^{(k+1)} = -a_{rj}^{(k)} \sin \varphi + a_{sj}^{(k)} \cos \varphi, & j \neq r, s, \\ a_{rr}^{(k+1)} &= a_{rr}^{(k)} \cos^2 \varphi + 2a_{rs}^{(k)} \sin \varphi \cos \varphi + a_{ss}^{(k)} \sin^2 \varphi, \\ a_{ss}^{(k+1)} &= a_{rr}^{(k)} \sin^2 \varphi - 2a_{rs}^{(k)} \sin \varphi \cos \varphi + a_{ss}^{(k)} \cos^2 \varphi. \end{aligned} \quad (5.16)$$

Dalla (5.15) segue che la (5.14) è verificata per ogni φ soluzione dell'equazione

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

In pratica si ricavano direttamente due valori, $\sin \varphi$ e $\cos \varphi$, per costruire la matrice $G_{rs}^{(k)}$, scrivendo la precedente equazione nella forma

$$t^2 + 2mt - 1 = 0, \quad (5.17)$$

dove si è posto $t := \tan \varphi$, $m := (a_{rr}^{(k)} - a_{ss}^{(k)})/2a_{rs}^{(k)}$.

Scegliendo fra le due radici della (5.17) quella di modulo minore, corrispondente a $|\varphi| \leq \pi/4$, si ha

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

per $m = 0$ si sceglie $t = 1$. In ogni caso si ottiene

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

con cui si costruisce la $G_{rs}^{(k)}$ 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; tuttavia vale il seguente teorema.

Teorema 5.3.1 *La successione $\{A_k\}$ generata dalle (5.13) 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 .

In base a questo teorema, un criterio di arresto delle iterazioni (5.13) potrebbe essere il verificarsi della condizione

$$\max_{i>j} |a_{ij}^{(k+1)}| \leq \epsilon, \quad (5.18)$$

dove $\epsilon > 0$ è un numero prefissato.

Se il criterio (5.18) è verificato, allora la matrice A_{k+1} approssima la matrice diagonale D e, indicando con $G^{(i)}$ la matrice di rotazione della i -esima iterazione, si ha

$$G^{(k)T} G^{(k-1)T} \dots G^{(1)T} A G^{(1)} \dots G^{(k-1)} G^{(k)} = A_{k+1},$$

o anche, posto $H_k := G^{(1)} \dots G^{(k)}$,

$$A H_k = H_k A_{k+1}.$$

Quindi gli elementi $a_{ii}^{(k+1)}$ di A_{k+1} sono approssimazioni degli autovalori di A mentre i corrispondenti autovettori sono approssimati dalle colonne della matrice H_k che si può ottenere mediante l'algoritmo

$$H_0 = I, \quad H_j = H_{j-1} G^{(j)}, \quad j = 1, 2, \dots, k.$$

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 5.3.1.

5.4 Riduzione in forma tridiagonale e di Hessenberg

Se una matrice è hermitiana e tridiagonale, l'approssimazione dei suoi autovalori ed autovettori, sia col metodo di Jacobi che con altri metodi, risulta più agevole che per una matrice hermitiana qualunque non sparsa. Per questo motivo una generica matrice hermitiana A viene di solito trasformata in una matrice tridiagonale simile.

Fra i vari modi per effettuare la riduzione di A alla forma tridiagonale riportiamo qui il *metodo di Givens*, considerato, per semplicità, nel caso reale. In questo metodo si usano ancora le matrici di rotazione per ottenere termini nulli ma, a differenza di quanto si è visto nel metodo di Jacobi, un elemento che è stato annullato a un certo passo non viene più modificato nelle successive trasformazioni. Ciò si ottiene annullando ordinatamente i termini non nulli fra gli elementi a_{ij} con $i - j \geq 2$, considerati per colonne, nel seguente ordine

$$a_{31}, a_{41}, \dots, a_{n1}; a_{42}, a_{52}, \dots, a_{n2}; \dots; a_{n,n-2},$$

usando, rispettivamente, le matrici di rotazione

$$G_{23}, G_{24}, \dots, G_{2n}; G_{34}, G_{35}, \dots, G_{3n}; \dots; G_{n-1,n}. \quad (5.19)$$

Poiché ad ogni passo viene annullato un elemento e il suo simmetrico, bastano al più $\frac{(n-2)(n-1)}{2}$ rotazioni per trasformare A in una matrice simile A_1 di forma tridiagonale, conservando la simmetria.

In questo processo gli indici della matrice di rotazione e quelli dell'elemento da annullare non sono gli stessi come nel metodo di Jacobi, ma la matrice $G_{rs}^{(k)}$ viene costruita in modo che risulti

$$a_{s,r-1}^{(k+1)} = 0, \quad k = 1, 2, \dots, \frac{(n-2)(n-1)}{2}; \quad (5.20)$$

se fosse già $a_{s,r-1}^{(k)} = 0$ si pone $G_{rs}^{(k)} = I$ (cioè $\varphi = 0$).

Utilizzando la (5.16) ove si ponga $j = r - 1$, la (5.20) fornisce

$$a_{s,r-1}^{(k+1)} = a_{s,r-1}^{(k)} \cos \varphi - a_{r,r-1}^{(k)} \sin \varphi = 0,$$

che è soddisfatta per

$$\cos \varphi = \frac{a_{r,r-1}^{(k)}}{\sqrt{(a_{s,r-1}^{(k)})^2 + (a_{r,r-1}^{(k)})^2}}, \quad \sin \varphi = \frac{a_{s,r-1}^{(k)}}{\sqrt{(a_{s,r-1}^{(k)})^2 + (a_{r,r-1}^{(k)})^2}}.$$

Formule numericamente più stabili si ottengono calcolando

$$t = \tan \varphi = \frac{a_{s,r-1}^{(k)}}{a_{r,r-1}^{(k)}}, \quad c = \cot \varphi = \frac{a_{r,r-1}^{(k)}}{a_{s,r-1}^{(k)}},$$

e ponendo

$$\begin{aligned} \cos \varphi &= \frac{1}{\sqrt{1+t^2}}, & \sin \varphi &= t \cos \varphi \quad \text{se } |t| < 1, \\ \sin \varphi &= \frac{1}{\sqrt{1+c^2}}, & \cos \varphi &= c \sin \varphi \quad \text{se } |t| > 1. \end{aligned}$$

La conoscenza delle matrici (5.19) consente di risalire dagli autovettori della matrice tridiagonale A_1 a quelli della matrice data. Infatti, sia y un autovettore di A_1 associato a λ e G la matrice prodotto di tutte le matrici (5.19) nell'ordine ivi considerato; G è ancora ortogonale e si ha $A_1 = G^T A G$ da cui, essendo $A_1 y = \lambda y$, segue $A_1 y = G^T A G y = \lambda y$ perciò $A G y = \lambda G y$ e $x = G y$ è l'autovettore di A corrispondente a y .

Se il processo di Givens si applica ad una matrice A non simmetrica, la matrice H che si ottiene al termine delle $(n-2)(n-1)/2$ trasformazioni è della forma seguente

$$H = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ & \ddots & \ddots & \vdots \\ & & h_{n-1,n} & h_{nn} \end{pmatrix}.$$

Si dice *matrice di Hessenberg superiore*.

La riduzione di una matrice A non simmetrica alla forma di Hessenberg, al pari di quella delle matrici simmetriche alla forma tridiagonale, è utile ai fini dell'ulteriore approssimazione di autovalori ed autovettori.

Un metodo per ottenere queste approssimazioni e che risulta particolarmente efficiente quando opera su matrici delle due forme suddette è il *metodo QR* esposto nel paragrafo seguente.

5.5 Schema del metodo QR

L'approssimazione simultanea di tutti gli autovalori di una matrice qualunque, viene generalmente effettuata mediante il metodo QR, di cui ci si

limita ad accennare le linee fondamentali nel caso reale partendo dal seguente teorema.

Teorema 5.5.1 *Per ogni matrice $A \in \mathbb{R}^{n \times n}$ esiste una decomposizione nel prodotto di una matrice Q ortogonale per una matrice R triangolare superiore.*

La dimostrazione si ottiene costruendo effettivamente le matrici Q ed R . Uno dei vari modi per raggiungere lo scopo consiste nel premoltiplicare successivamente A con matrici di rotazione G_{rs} scelte in modo che ad ogni passo si ottenga una matrice prodotto in cui risulti nullo un elemento di indici (s, r) situato sotto la diagonale principale, ammesso che non fosse già nullo (in tal caso si pone $G_{rs} = I$). La strategia che si segue è quella di premoltiplicare A ordinatamente per le $n(n-1)/2$ matrici

$$G_{12}, G_{13}, \dots, G_{1n}; G_{23}, \dots, G_{2n}; \dots; G_{n-1,n},$$

in modo che, nelle successive matrici prodotto, risultino nulli rispettivamente gli elementi di indici

$$(21), (31), \dots, (n1); (32), \dots, (n2); \dots; (n, n-1).$$

In questo modo non vengono modificati gli elementi nulli ottenuti nei passi precedenti.

Poiché gli elementi annullati sono tutti situati al disotto della diagonale principale, la matrice R ottenuta dopo $n(n-1)/2$ prodotti è triangolare superiore e si ha

$$G_{n-1,n} \cdots G_{13} G_{12} A = R,$$

infine, essendo le G_{rs} ortogonali, si ha

$$A = G_{12}^T G_{13}^T \cdots G_{n-1,n}^T R = QR$$

che fornisce la cosiddetta *fattorizzazione QR* della matrice A , con la matrice $Q = G_{12}^T G_{13}^T \cdots G_{n-1,n}^T$ ortogonale.

In particolare, se la matrice A è della forma di Hessenberg superiore oppure tridiagonale, la fattorizzazione QR richiede al più $n-1$ premoltiplicazioni per matrici di rotazione, essendo al più $n-1$ gli elementi non nulli di A al disotto della diagonale principale.

Il Teorema 5.5.1 può essere generalizzato al caso di una matrice ad elementi complessi; in tal caso la matrice Q della fattorizzazione è una matrice unitaria.

L'algoritmo QR per la ricerca degli autovalori di $A \in \mathbb{R}^{n \times n}$ utilizza la fattorizzazione QR secondo lo schema

$$\begin{aligned} A_1 &:= A, \\ A_k &= Q_k R_k, \\ A_{k+1} &= R_k Q_k, \quad k = 1, 2, \dots \end{aligned} \quad (5.21)$$

Tutte le matrici della successione $\{A_k\}$ generata dall'algoritmo (5.21) sono simili ad A ; infatti, per qualunque k , si ha $Q_k^T Q_k = I$ e anche

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

Ogni matrice della successione $\{A_k\}$ ha quindi gli stessi autovalori di A ; inoltre, se A è nella forma di Hessenberg, si può dimostrare che ogni matrice della successione $\{A_k\}$ si mantiene della stessa forma di A . In tal caso ad ogni passo la fattorizzazione richiede al più $(n-1)$ prodotti di matrici e il costo computazionale di una iterazione² è di circa $2n^2$ moltiplicazioni. Tale costo sale a n^3 moltiplicazioni se A è di forma qualsiasi.

La giustificazione teorica del metodo QR si completa con i due teoremi seguenti che ci si limita ad enunciare nel caso reale.

Teorema 5.5.2 (di Schur) *Per ogni matrice $A \in \mathbb{R}^{n \times n}$ esiste una matrice reale ortogonale B tale che la trasformata per similitudine $S = B^{-1}AB = B^T AB$ è una matrice triangolare a blocchi con i blocchi diagonali di ordine uno o due, della forma*

$$S = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1r} \\ & S_{22} & \cdots & S_{2r} \\ & & \ddots & \vdots \\ & & & S_{rr} \end{pmatrix}. \quad (5.22)$$

I blocchi diagonali di ordine 1 sono autovalori reali di A , mentre ogni blocco diagonale di ordine due ha come autovalori una coppia di autovalori complessi coniugati di A .

²Per "iterazione" si intende la fattorizzazione di A_k e il calcolo di A_{k+1} in base alle (5.21).

Teorema 5.5.3 *Se gli autovalori di $A \in \mathbb{R}^{n \times n}$ sono reali e distinti in modulo e quindi verificano la condizione*

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| \quad (5.23)$$

e se gli autovettori di A formano una matrice X tale che X^{-1} sia fattorizzabile LR, allora le matrici A_k per $k \rightarrow \infty$ tendono ad una matrice triangolare superiore e gli elementi diagonali $a_{ii}^{(k)}$ di A_k tendono agli autovalori λ_i di A ordinati per modulo decrescente.

Se A possiede qualche coppia di autovalori complessi coniugati (e quindi la (5.23) non è verificata) ma i moduli di ciascuna coppia e quelli degli autovalori reali sono distinti e se vale l'ipotesi fatta su X^{-1} , allora le matrici A_k convergono ad una matrice triangolare a blocchi del tipo (5.22), in cui gli autovalori dei vari blocchi sono ancora ordinati per modulo decrescente.

Osservazione 5.5.1 Mancando l'ipotesi della fattorizzabilità LR di X^{-1} , si ha ancora la convergenza del metodo ma può venire meno l'ordinamento per modulo decrescente degli autovalori.

In pratica, partendo da una matrice A in forma di Hessenberg superiore le iterazioni si arrestano al raggiungimento di una matrice A_m che possiede qualche elemento della codiagonale principale talmente piccolo in valore assoluto da potersi considerare nullo. Si supponga che sia

$$A_m = \begin{pmatrix} A_{11} & A_{12} \\ \mathbf{O} & A_{22} \end{pmatrix}$$

con A_{11} , A_{22} , entrambe in forma di Hessenberg superiore.

Gli autovalori di A_m sono quindi approssimati dagli autovalori di A_{11} e A_{22} . Se per una o entrambe le matrici A_{11} e A_{22} non si è ancora raggiunta la forma triangolare o (5.22) si dovrà applicare ancora il metodo QR a una o a ciascuna delle due matrici e così via.

In linea di principio il metodo QR fornisce anche gli autovettori di una matrice A . Infatti si supponga che, dopo $m - 1$ trasformazioni ortogonali, si sia ottenuta una matrice A_m (della forma (5.22) o triangolare superiore) e sia λ uno degli autovalori di A fornito da A_m ; posto

$$Q = Q_1 Q_2 \cdots Q_{m-1},$$

Q è ancora una matrice ortogonale.

Ripetendo per Q , A_m ed A , il ragionamento fatto in 5.4 per le matrici G , A_1 ed A , segue che, se y è un autovettore di A_m associato a λ , $x = Qy$ è il corrispondente autovettore di A . Tuttavia il costo computazionale per il calcolo della matrice Q è elevato per cui, disponendo dell'approssimazione λ di un autovalore, conviene utilizzare il metodo delle potenze inverse descritto in 5.2 per ottenere un autovettore corrispondente di A .

Il metodo QR , qui esposto in una delle sue versioni più semplici, viene adoperato anche in forme modificate più efficienti; esso presenta comunque una notevole stabilità numerica e la proprietà di convergere sotto ipotesi anche più deboli di quelle del Teorema 5.5.3.

5.6 Complementi ed esempi

5.6.1 Precisazione sul metodo delle potenze

Nel metodo delle potenze, applicato nella forma (5.4), la scelta del vettore $z^{(0)}$ è arbitraria. La diagonalizzabilità di A , e quindi l'esistenza di n autovettori $x^{(i)}$ linearmente indipendenti, implica che, in ogni caso, risulti $z^{(0)} = \sum_{i=1}^n c_i x^{(i)}$.

Se per una data scelta di $z^{(0)}$ risulta $c_1 = 0$ e si suppone $|\lambda_2| > |\lambda_3|$, la successione dei vettori $y^{(k)}$ converge, come si riscontra dalla (5.6), ad un autovettore associato a λ_2 mentre $R(y^{(k)})$ tende a λ_2 . Ciò accade effettivamente quando si ricorre a precisioni molto elevate. In caso contrario, gli errori di arrotondamento fanno sì che la successione $\{y^{(k)}\}$ finisca per convergere ancora, sia pure lentamente, ad un autovettore associato a λ_1 così come $R(y^{(k)})$ a λ_1 . Infatti, in questo secondo caso, la (5.4) si scrive

$$y^{(k)} = Ay^{(k-1)} + \delta^{(k)}, \quad k = 1, 2, \dots, \quad (5.24)$$

dove $\delta^{(k)}$ è il vettore di errore, anch'esso rappresentabile come una combinazione degli autovettori di A . Posto $\delta^{(1)} = d_1 x^{(1)} + \dots + d_n x^{(n)}$, il vettore $y^{(1)}$, fornito dalla (5.24) per $k = 1$, si può esprimere nella forma

$$y^{(1)} = (c_1 \lambda_1 + d_1) x^{(1)} + \dots + (c_n \lambda_n + d_n) x^{(n)},$$

dove ora il coefficiente di $x^{(1)}$ è, in generale, diverso da zero e quindi, a partire da $y^{(1)}$, le ipotesi del Teorema 5.1.1 sono soddisfatte. Una situazione analoga si verifica per il metodo applicato nella forma normalizzata (5.9).

Esempio 5.6.1 Sia

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

e si assuma $z^{(0)} = (1, 1, 1, 1)^T$: con questa scelta, nell'uguaglianza $z^{(0)} = \sum_{i=1}^4 c_i x^{(i)}$ risulta $c_1 = 0$.

Si è applicato il metodo delle potenze in forma normalizzata con tre precisioni di macchina: $u_s \simeq 4.75 \times 10^{-7}$ (precisione semplice), $u_d \simeq 1.11 \times 10^{-16}$ (precisione doppia) e $u_q \simeq 2.58 \times 10^{-26}$ (precisione quadrupla). Si è adottato il criterio di arresto $|R(z^{(k)}) - R(z^{(k-1)})| < \frac{1}{2}10^{-5}$. I risultati ottenuti sono nella tavola che segue.

	k	$R(z^{(k)})$
u_s	6	2.6180295...
u_d	6	2.6180339...
u_q	6	2.6180339...

Le otto cifre significative del valore riportato per la precisione doppia e la precisione quadrupla corrispondono a quelle dell'autovalore λ_2 e sono corrette. Tuttavia, proseguendo le iterazioni in precisione semplice e in precisione doppia, l'effetto degli errori di arrotondamento fa sì che, per $k \geq k^*$, $R(z^{(k)})$ si stabilizzi definitivamente su valori che approssimano λ_1 . Mentre, la precisione quadrupla, essendo trascurabili gli errori di arrotondamento, continua a fornire l'approssimazione di λ_2 .

Nella tavola che segue si danno i valori dell'indice k^* a partire dal quale le otto cifre significative riportate per $R(z^{(k)})$ rimangono fisse.

	k^*	$R(z^{(k)}), k \geq k^*$
u_s	79	3.6180315...
u_d	144	3.6180339...
u_q	6	2.6180339...

L'approssimazione di λ_1 ottenuta in precisione doppia è corretta nelle otto cifre significative riportate nella tavola. \square

5.6.2 Accelerazione della convergenza

Si può dimostrare che per gli elementi della matrice A_k del metodo QR (5.21), nelle ipotesi del Teorema 5.5.3, risulta

$$a_{ij}^{(k)} = O\left(\left(\frac{\lambda_i}{\lambda_j}\right)^k\right), \quad i > j.$$

Ne segue che per questo metodo la convergenza può essere molto lenta se $\left|\frac{\lambda_i}{\lambda_j}\right| \simeq 1$; ad analoga conclusione si giunge, a causa della (5.6), per il metodo delle potenze se $\lambda_1 \simeq \lambda_2$.

Una tecnica che consente di ovviare a questo inconveniente consiste nell'effettuare una traslazione dello spettro di A (cfr. Teorema 2.7.6). Si considera, cioè, in luogo di A , la matrice $B = A + qI$, che ha gli stessi autovettori e i cui autovalori sono $\mu_i = \lambda_i + q$, $i = 1, 2, \dots, n$, scegliendo il parametro q tale che sia ancora $|\mu_1| > |\mu_2| > \dots > |\mu_n|$, ma risulti

$$\left|\frac{\mu_2}{\mu_1}\right| < \left|\frac{\lambda_2}{\lambda_1}\right|.$$

In tal caso il metodo delle potenze ed il metodo QR , applicati a B , convergono più rapidamente.

Esempio 5.6.2 Si consideri la matrice

$$A = \begin{pmatrix} 21 & -1 & 1 \\ -1 & 21 & 1 \\ -1 & 1 & 21 \end{pmatrix}.$$

Dal teorema di Gershgorin si ricava $19 \leq |\lambda_3| \leq |\lambda_2| \leq |\lambda_1| \leq 23$, per cui la convergenza sarà, nella migliore delle ipotesi, come quella del termine $(19/23)^k \simeq (0.826)^k$.

Effettuando una traslazione di spettro con $q = -19$, si ha la matrice B con $0 \leq |\mu_3| \leq |\mu_2| \leq |\mu_1| \leq 4$. Essendo $\mu_i = \lambda_i - q$, risulta $|\mu_2/\mu_1| \leq |\lambda_2/\lambda_1|$.

Si è applicato il metodo delle potenze in precisione doppia alle due matrici A e B , con il criterio di arresto $|R(z^{(k)}) - R(z^{(k-1)})| < 10^{-9}$ ottenendo i risultati riportati nella tavola che segue.

	k	$R(z^{(k)})$
matrice A	401	22.0000000118...
matrice B	48	3.0000000052...

Il metodo QR , applicato con i criteri di arresto $\max_{i>j} |a_{ij}^{(k)}| < 10^{-6}$ e $\max_{i>j} |b_{ij}^{(k)}| < 10^{-6}$, ha fornito i seguenti risultati:

$$\begin{aligned} a_{11}^{(286)} &= 22.00000166\dots, & b_{11}^{(33)} &= 3.00000154\dots, \\ a_{22}^{(286)} &= 19.99999944\dots, & b_{22}^{(33)} &= 1.00000058\dots, \\ a_{33}^{(286)} &= 20.99999888\dots, & b_{33}^{(33)} &= 1.99999787\dots \end{aligned}$$

Gli autovalori di A sono $\lambda_1 = 22$, $\lambda_2 = 21$, $\lambda_3 = 20$.

Si noti inoltre che nelle matrici $A^{(k)}$ e $B^{(k)}$ gli autovalori compaiono sulla diagonale principale non ordinati per modulo decrescente: ciò è dovuto al fatto che la matrice X^{-1} di cui nel Teorema 5.5.3 non è fattorizzabile LR (cfr. Osservazione 5.5.1); si ha infatti

$$X^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

□

Bibliografia: [2], [12], [14], [28], [30].