Metodi numerici per il calcolo degli autovalori
II^a parte

Calcolo Numerico - Ing. Inf. - Lezione 13

Outline

Outline

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 illustriamo il **metodo di Givens**, considerato, per semplicità, nel caso reale

Nel metodo di Givens 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

$$G_{23}, G_{24}, \dots, G_{2n}; G_{34}, G_{35}, \dots, G_{3n}; \dots ; G_{n-1,n}$$

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 $G_{rt}^{(k)}$ viene costruita in modo che risulti

$$a_{t,r-1}^{(k+1)} = 0 \qquad k = 1, 2, \ldots, \frac{(n-2)(n-1)}{2}$$

Se fosse già $a_{t,r-1}^{(k)} = 0$ si pone $G_{rt}^{(k)} = I$

Utilizzando la terza formula della Lezione 12 – Slide 23 dove si pone $j = r - 1$, risulta

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

che è soddisfatta per

$$\cos\varphi \;=\; \frac{a_{r,r-1}^{(k)}}{\sqrt{\left(a_{t,r-1}^{(k)}\right)^2 + \left(a_{r,r-1}^{(k)}\right)^2}}$$

$$\sin\varphi \;=\; \frac{a_{t,r-1}^{(k)}}{\sqrt{\left(a_{t,r-1}^{(k)}\right)^2 + \left(a_{r,r-1}^{(k)}\right)^2}}.$$

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

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

H si dice **matrice di Hessenberg superiore**

La riduzione di una matrice A non simmetrica alla forma di Hessenberg è utile ai fini della approssimazione degli autovalori

Outline

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

Teorema (Fattorizzazione QR)

Per ogni matrice $A \in \mathbb{R}^{n \times n}$ esiste una fattorizzazione data dal prodotto di una matrice Q ortogonale per una matrice R triangolare superiore

La dimostrazione si ottiene costruendo effettivamente le matrici Q ed R

Un modo (non è l'unico) di calcolare una fattorizzazione QR consiste nel premoltiplicare successivamente A con matrici di rotazione G_{rt} scelte in modo che ad ogni passo si ottenga una matrice prodotto in cui risulti nullo un elemento di indici (t, r) situato sotto la diagonale principale, ammesso che non fosse già nullo (in tal caso si pone $G_{rt} = I$)

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

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

Osservazione

Seguendo la procedura esposta, non vengono modificati gli elementi nulli ottenuti nei passi precedenti

Poiché gli elementi annullati sono tutti sotto la 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, ponendo $Q = G_{12}^T G_{13}^T \cdots G_{n-1,n}^T$ (ortogonale), si ha

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

che fornisce la **fattorizzazione** QR della matrice A

Osservazione

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

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

$$\begin{aligned} A_1 &= A\,, \\ A_k &= Q_k R_k\,, \\ A_{k+1} &= R_k Q_k\,, \qquad k = 1, 2, \ldots \end{aligned}$$

Le matrici della successione $\{A_k\}$ generata dall'algoritmo 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

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 iterazione è di circa $2n^2$ moltiplicazioni

Tale costo sale a n^3 moltiplicazioni se A è di forma qualsiasi

La teoria del metodo QR si completa con due teoremi che ci si limita ad enunciare nel caso reale

Teorema 1 (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

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

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 di A complessi coniugati

Teorema 2

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

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

e se gli autovettori di A formano una matrice X tale che X^{-1} sia fattorizzabile LR, **allora** le matrici A_k per $k \to \infty$ 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 non è verificata la condizione sugli autovalori) 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 in cui gli autovalori dei vari blocchi sono ancora ordinati per modulo decrescente

Osservazione

Nel caso in cui manchi 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

Conclusione

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 riportate nel Teorema 2

Metodo QR con A matrice di Hessenberg

In generale, come si è ricordato, prima di applicare il metodo QR si preferisce riportare la matrice A nella forma di Hessenberg superiore attraverso trasformazioni per similitudine

Si potrebbe dimostrare che tutte le matrici della successione generata col metodo QR risultano tutte matrici di Hessenberg superiori

In questo contesto faremo solo una veloce verifica nel caso in cui sia $A \in \mathbb{R}^{4\times 4}$

Metodo QR con A matrice di Hessenberg

Sia

$$A = \begin{pmatrix} X & X & X & X \\ X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{pmatrix}$$

Per ottenere la fattorizzazione QR della matrice A, si devono azzerare gli elementi sulla codiagonale inferiore premoltiplicando per opportune matrici di rotazione G_{12}, G_{23} e G_{34}

Metodo QR con A matrice di Hessenberg

Siano

$$G_{12} = \begin{pmatrix} X & X & 0 & 0 \\ X & X & 0 & 0 \\ 0 & 0 & X & 0 \\ 0 & 0 & 0 & X \end{pmatrix}, \quad G_{23} = \begin{pmatrix} X & 0 & 0 & 0 \\ 0 & X & X & 0 \\ 0 & X & X & 0 \\ 0 & 0 & 0 & X \end{pmatrix},$$

$$G_{34} = \begin{pmatrix} X & 0 & 0 & 0 \\ 0 & X & 0 & 0 \\ 0 & 0 & X & X \\ 0 & 0 & X & X \end{pmatrix}$$

Metodo QR con A matrice di Hessenberg

Quindi risulta

$$G_{34}\ G_{23}\ G_{12}A = \begin{pmatrix} X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \\ 0 & 0 & 0 & X \end{pmatrix} = R$$

La matrice Q della fattorizzazione è

$$Q = G_{12}^T\ G_{23}^T\ G_{34}^T = \begin{pmatrix} X & X & X & X \\ X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{pmatrix}$$

Metodo QR con A matrice di Hessenberg

L'algoritmo del metodo QR prevede di costruire una nuova matrice della successione eseguendo il prodotto RQ

In questo caso risulta

$$R\,Q \;=\; \begin{pmatrix} X & X & X & X \\ X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{pmatrix}$$

che è ancora una matrice di Hessenberg superiore