

Calcolo Numerico A - a.a. 2008/09

Lunedì 18 maggio 2009

Metodi iterativi (lineari)

Data una matrice $B \in \mathbb{R}^{n \times n}$ e un vettore $c \in \mathbb{R}^n$, si consideri l'applicazione

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad F(x) = B \cdot x + c. \quad (0.1)$$

Se $c = 0$ l'applicazione è lineare (endomorfismo), altrimenti si dice affine.

Per ogni $s \in \mathbb{R}^n$, consideriamo la successione

$$\begin{cases} u_k = Bu_{k-1} + c, & k > 0, \\ u_0 = s. \end{cases} \quad (0.2)$$

Il vettore s si dice valore iniziale.

La matrice B si dice matrice del metodo iterativo.

Teorema (Criterio di convergenza dei metodi iterativi)

Sono fatti equivalenti:

- ❶ *L'applicazione F definita in (0.1) è una contrazione per una qualche norma su $\mathbb{R}^n$.*
- ❷ *La successione u_k è convergente ad un punto (fisso) $\bar{u}$ per qualsiasi valore iniziale s .*
- ❸ *$\rho(B) < 1$.*
- ❹ *$\|B\| < 1$ per almeno una norma matriciale $\|\cdot\|$ (indotta o no).*
- ❺ *$\lim_{k \rightarrow \infty} B^k = 0$.*
- ❻ *$\lim_{k \rightarrow \infty} B^k v = 0$ per ogni $v \in \mathbb{R}^n$.*

Se una delle condizioni sopra vale, allora l'applicazione F ammette un unico punto fisso.

Teorema

Sia B una matrice quadrata e $\|\cdot\|$ una norma matriciale qualsiasi. Allora:

$$\lim_{k \rightarrow \infty} \|B^k\|^{1/k} = \rho(B).$$

(cfr. P.G.Ciarlet, pag.25)

Esempio

Data la matrice

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

si consideri l'applicazione

$$F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad F(x) = Ax.$$

Si noti che:

$$\rho(A) = \frac{1}{\sqrt{2}} < 1 < \|A\|_1 = \|A\|_\infty = 1 + \frac{\sqrt{2}}{2} < \|A\|_2 = 1 + \frac{\sqrt{3}}{2}.$$

Si noti che ci sono applicazioni che hanno un unico punto fisso e non sono contrazioni rispetto a nessuna norma.

Esempio

Data una matrice $A \in O(n)$, $A \neq I$, l'applicazione

$$F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad F(x) = Ax$$

ha un unico punto fisso.

Essendo A ortogonale, i suoi autovalori hanno modulo uno, quindi $\rho(A) = 1$.

Già sappiamo che (3) e (4) sono equivalenti. Inoltre nella condizione (4) si può assumere che la norma $\|\cdot\|$ sia indotta. Abbiamo poi: $(4) \implies (5) \implies (6) \implies (3)$. Si noti che dalla (4) segue che F è una contrazione rispetto alla norma $\|\cdot\|$, cioè (1). Infatti, essendo

$$\|F(x) - F(x')\| = \|Bx - Bx'\| = \|B(x - x')\| \leq \|B\|\|x - x'\|,$$

dalla dimostrazione del teorema delle contrazioni segue che le successioni u_k sono convergenti (indipendentemente dal punto iniziale).

Infine, se le successioni u_k sono convergenti ad un certo $\bar{u}$, per ogni valore iniziale s , allora le successioni

$$e_k := u_k - \bar{u} = B^k e_0, \quad k \geq 0$$

sono infinitesime per ogni valore iniziale $e_0 = u_0 - \bar{u} = s - \bar{u}$ (cioè la (6)).

Esercizio

Dati $\alpha, \beta \in \mathbb{R}$, prendiamo l'applicazione (affine) $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definita da

$$F \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x_1 - \frac{1}{2}x_2 + \alpha \\ \frac{1}{2}x_1 + \frac{1}{2}x_2 + \beta \end{pmatrix}.$$

Studiare tale applicazione. In particolare dire se F è una contrazione per una qualche norma, se ammette un punto fisso e, in tal caso, se esso è unico.

Osservazione

Se $|\det(B)| \geq 1$ e $|tr(A)| \geq n$, allora B ha almeno un autovalore λ con $|\lambda| \geq 1$, quindi $\rho(B) \geq 1$. Dunque

$|\det(B)| < 1$ e $|tr(A)| < n$ sono *condizioni necessarie* affinché $\rho(B) < 1$.

Definizione

Il numero $R_m(B) := -\frac{1}{m} \log_{10} \|B^m\|$ si dice *velocità media di convergenza* dopo m iterazioni.

Il numero $R := \lim_{k \rightarrow \infty} R_k(B) = -\log_{10} \rho(B)$ si dice *velocità (o tasso) asintotico di convergenza*.

Metodi iterativi

Sia $A \in \mathbb{R}^{n \times n}$ una matrice non singolare. Si consideri la scomposizione di A nella forma

$$A = M - N$$

dove M è una matrice non singolare. Dal sistema lineare

$$Ax = b \tag{0.3}$$

risulta allora

$$Mx - Nx = b,$$

cioè:

$$x = M^{-1}Nx + M^{-1}b.$$

Posto

$$B := M^{-1}N, \quad q := M^{-1}b,$$

si ottiene il seguente sistema

$$x = Bx + q, \tag{0.4}$$

equivalente a quello iniziale (0.3).

Definizione

La M si dice *matrice di preconditionamento*, mentre la matrice B si dice *matrice di iterazione* del sistema lineare (0.3).

Dato un vettore iniziale x_0 , si considera la successione la successione così definita:

$$x_k = Bx_{k-1} + q, \quad q \in \mathbb{N}. \quad (0.5)$$

Se la successione x_k è convergente ad un punto x_* , allora passando al limite si ottiene

$$x_* = Bx_* + q,$$

cioè x_* è soluzione del sistema lineare (0.4) e quindi del sistema (0.3).

Osservazione

Per applicare il metodo iterativo basato su uno splitting $A = M - N$ occorre avere:

- 1 M non singolare,
- 2 $\rho(B) < 1$, dove $B := M^{-1}N$.

Definizione

Una matrice $n \times n$ si dice **sparsa** se il numero dei suoi elementi non nulli è di ordine n .

Osservazione

Il costo computazionale di un metodo iterativo è dell'ordine di n^2 operazioni per ogni iterazione (contro le approssimativamente $2n^3/3$ richieste da un metodo diretto).

Per matrici **sparse** la moltiplicazione di B per un vettore richiede un numero di moltiplicazioni dell'ordine di n . Quindi per tali matrici (a meno che non siano strutturate) è in generale conveniente usare un metodo iterativo (i metodi diretti sono onerosi perché danno luogo al fenomeno del fill in, cioè fattori di una scomposizione di una matrice

Metodo di Jacobi

Dato il sistema $Ax = b$, sia

$$A = D - L - U$$

con:

- D matrice diagonale,
- L triangolare strettamente inferiore,
- U triangolare strettamente superiore.

Si considera lo splitting della matrice dei coefficienti $A = M - N$, dove

- $M = D$,
- $N = L + U$.

Osservazione

Per applicare il metodo di Jacobi occorre avere:

- 1 D non singolare, cioè $a_{ii} \neq 0, \forall i = 1, \dots, n$.
- 2 $\rho(B_J) < 1$, dove $B_J := M^{-1}N = D^{-1}(L + U)$.

Scelto un dato iniziale x_0 , la successione x_k delle iterate (0.5) nel metodo di Jacobi si calcola attraverso le formule:

$$(x_k)_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^n a_{ij}(x_{k-1})_j \right), \quad i = 1, \dots, n. \quad (1.6)$$

Metodo di Jacobi

Teorema

Se la matrice A dei coefficienti è **dominanza diagonale stretta** per righe o per colonne, allora (A e D sono invertibili e) $\rho(B) < 1$, dove $B := D^{-1}(L + U)$. Dunque, in tal caso, il metodo di Jacobi è convergente.

Infatti, se A è a dominanza diagonale stretta, $a_{ii} > 0$ e dunque D è invertibile e, per il teorema di Gerschgorin, (gli autovalori di A sono non nulli e quindi) anche A è invertibile. Essendo poi

$$\sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|} < 1, \quad i = 1, \dots, n,$$

poiché

$$\|B_J\|_{\infty} := \max_{1 \leq i \leq n} \sum_{j=1, j \neq i}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1,$$

che ci dice che il metodo è convergente.

Teorema (cfr. QSS, p.117)

Se A e $2D - A$ sono **simmetriche definite positive**, allora il metodo di Jacobi converge. (Inoltre si ha: $\rho(B_J) = \|B_J\|_A = \|B_J\|_D$).

Metodo di Gauss-Seidel¹

Dato il sistema $Ax = b$, sia

$$A = D - L - U$$

con:

- D matrice diagonale,
- L triangolare strettamente inferiore,
- U triangolare strettamente superiore.

Si considera lo splitting della matrice dei coefficienti $A = M - N$, dove

- $M = D - L$,
- $N = U$.

Osservazione

Per applicare il metodo di Jacobi occorre avere:

- 1 $D - L$ non singolare, cioè $a_{ii} \neq 0, \forall i = 1, \dots, n$.
- 2 $\rho(B_{GS}) < 1$, dove $B_{GS} := M^{-1}N = (D - L)^{-1}U$.

Scelto un dato iniziale x_0 , la successione x_k delle iterate (0.5) nel metodo di Gauss-Seidel si calcola attraverso le formule:

$$(x_k)_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}(x_k)_j - \sum_{j=i+1}^n a_{ij}(x_{k-1})_j \right), \quad i = 1, \dots, n. \quad (2.7)$$

¹Philipp Ludwig Seidel (Zweibrücken 1821 - München 1896)

Metodo di Gauss-Seidel

Teorema (cfr. BCM, p.249)

Se la matrice A dei coefficienti è **a dominanza diagonale stretta** per righe o per colonne, allora (A e $M = D - L$ sono invertibili e) $\rho(B) < 1$, dove $B := (D - L)^{-1}U$. Dunque, il metodo di Gauss-Seidel è convergente.

Teorema (cfr. QSS, p.117)

Se A è una matrice **simmetrica definita positiva**, allora il metodo di Gauss-Seidel converge (in modo monotono rispetto alla norma $\|\cdot\|_A$).

Criteri di arresto

1) Controllo dell'incremento.

Fissata una tolleranza ε , ci si ferma al primo k per cui

$$\|x_{k+1} - x_k\| < \varepsilon.$$

Corrispondentemente, se $\bar{x}$ è soluzione esatta del sistema $Ax = b$, ci chiediamo quanto sia grande il vero errore:

$$\|e_{k+1}\|, \text{ dove } e_{k+1} := x_{k+1} - \bar{x}.$$

Se B è la matrice di iterazione, allora

$$e_{k+1} = Be_k.$$

Abbiamo:

$$\|x_{k+1} - x_k\| = \|(x - x_k) - (x - x_{k+1})\| \geq \|e_k\| - \|e_{k+1}\| \geq (1 - \|B\|)\|e_k\|.$$

Dunque:

$$\|e_{k+1}\| \leq \|B\|\|e_k\| \leq \frac{\|B\|}{1 - \|B\|} \|x_{k+1} - x_k\| \leq \frac{\|B\|}{1 - \|B\|} \varepsilon.$$

Di conseguenza l'errore $\|e_{k+1}\|$ è piccolo purché $\|B\| < 1$ non sia troppo vicino a 1.

Criteri di arresto

2) Controllo del residuo.

Fissata una tolleranza ε , ci si ferma al primo k per cui la norma del **residuo** $r_k := b - Ax_k$ risulta

$$\|r_k\| = \|b - Ax_k\| < \varepsilon.$$

Dalle relazioni:

$$r_k = b - Ax_k, \quad 0 = b - A\bar{x},$$

si ottiene

$$r_k = Ae_k,$$

quindi:

$$\|e_k\| \leq \|A^{-1}\| \|r_k\| = \|K(A)\| \frac{\|r_k\|}{\|A\|}.$$

Passando agli errori relativi ed osservando che $\|b\| \leq \|A\| \|\bar{x}\|$, si ottiene:

$$\frac{\|e_k\|}{\|\bar{x}\|} = \frac{\|x_k - \bar{x}\|}{\|\bar{x}\|} \leq \|K(A)\| \frac{\|r_k\|}{\|b\|}.$$