METODI ITERATIVI PER SISTEMI LINEARI

martedì 27 dicembre 2022 09:38

Introduzione

Descriveremo metodi che generano successioni di vettori $x^{(k)}$ che convergono alla soluzione del sistema cercata.
Il calcolo del prodotto matrice per vettore è l'operazione più costosa
Questa classe di metodi che approssimano la soluzione generando successioni di vettori è nota come classe dei metodi iterativi

Metodi stazionari

Dato $Ax=b$ consideriamo un partizionamento di A cioè $A=M-N$ con $\det M \neq 0$

Posso riscrivere il sistema nel seguente modo $Mx=Nx+b \Rightarrow x=M^{-1}Nx+M^{-1}b=:Px+q$ problema di punto fisso

matrice di iterazione

Il problema di punto fisso produce un metodo per generare una succ. di vettori una volta scelto $x^{(0)} \in \mathbb{C}^n$ cioè $x^{(k+1)}=Px^{(k)}+q$

metodo iterativo stazionario

P e q non dipendono da k

Se la succ ha un lim $x^* \Rightarrow x^*$ è sol. del sistema $Ax=b$ $\quad x^{(k+1)}=Px^{(k)}+q \;\downarrow\; x^*=Px^*+q$

⚠ Quindi ci basta dim la convergenza della succ. $x^{(k)}$ per concludere sull'efficacia del metodo

Convergenza

Teorema 1 $\rightarrow$ condizione sufficiente di convergenza

Se $\exists \|\cdot\|$ indotta t.c. $\|P\|<1 \Rightarrow A$ è invertibile $\Rightarrow \exists!\, x$ sol. di $Ax=b$.
Inoltre $\forall x^{(0)} \in \mathbb{C}^n$ iniziale la succ. generata converge a x

idea dim

① Per ass. $\rightarrow \det A=0$

$$Av=0 \leadsto Mv=Nv \leadsto v=\underbrace{M^{-1}N}_{P}v \Rightarrow \rho(P)\geq 1 \Rightarrow ⚡$$

v autovettore

$\|P\| \geq \rho(P) \geq 1$

② Sia x sol. sistema

(A) definiamo $e^{(k)}=x^{(k)}-x$ cioè errore di appros. al passo k

(B)
$$x^{(k+1)}=Px^{(k)}+q$$
$$e^{(k+1)}+\cancel{x}=P(e^{(k)}+x)+q=Pe^{(k)}+\cancel{Px}+\cancel{q}$$
$*$ x è sol $x=Px+q$
$$e^{(k+1)}=Pe^{(k)}$$

(C) itero $\Rightarrow e^{(k)}=Pe^{(k-1)}=\dots=P^k e^{(0)}$

(D) Passo alle norme $\Rightarrow \|e^{(k)}\| \leq \underbrace{\|P\|^k}_{\to 0 \text{ per } k\to+\infty \text{ perché } \|P\|<1}\|e^{(0)}\|$, per confronto $\|e^{(k)}\| \to 0$

Definiamo il metodo iterativo stazionario convergente se $\forall$ scelta di $x^{(0)}$ la succ $x^{(k)}$ conv. ad una sol. del sistema

Teorema 2 $\mapsto$ Condizione necessaria e sufficiente

Il metodo iterativo è convergente e $\det(A)\neq 0 \iff \rho(P)<1$

idea dim

$\boxed{\Leftarrow}$ Hp $\mapsto \rho(P)<1$

1. Usiamo Teo sulle norme indotte $\rho(P) \leq \|P\| \leq \rho(P)+\varepsilon < 1$ $\quad \exists\, \varepsilon$ t.c.
2. Applichiamo Teo 1

Tesi $\mapsto$ cioè metodo conv. e $\det A \neq 0$

$\boxed{\Rightarrow}$ Hp $\mapsto$ $\det A \neq 0$ e metodo converg.

1. la sol. x $\exists!$
2. arbitrarietà di $x^{(0)} \Rightarrow$ arbitrarietà di $e^{(0)} = x^{(0)} - x$
3. $e^{(k)} \longrightarrow 0 \quad \forall e^{(0)}$ perché $e^{(k)} = \underbrace{x^{(k)} - x}_{0 \text{ perché } x^{(k)} \to x}$
4. Scelgo $e^{(0)}$ autovettore di P cioè $Pe^{(0)} = \lambda e^{(0)}$
5. $\underset{\downarrow\, 0}{e^{(k)}} = \lambda^k \underset{\downarrow\, 0}{e^{(0)}} \Rightarrow \lim_k \lambda^k = 0 \Rightarrow |\lambda| < 1 \Rightarrow \overset{\text{Tesi}}{\rho(P) < 1}$

Osservazioni

- Se $\det A = 0 \Rightarrow x^{(k)} \downarrow \forall x^{(0)}$ ma il lim dipende $x^{(0)}$ e $\rho(P) = 1$
- Se un metodo è converg. vale $\Leftrightarrow$ nel Teo 1

Errore

$\frac{\|e^{(k)}\|}{\|e^{(k-1)}\|}$ = riduzione dell'errore al K-passo

media geometrica $\mapsto \theta_k(e^{(0)}) = \left(\frac{\|e^{(1)}\|}{\|e^{(0)}\|} \cdots \frac{\|e^{(k)}\|}{\|e^{(k-1)}\|}\right)^{\frac{1}{k}}$

$$= \left(\frac{\|e^{(k)}\|}{\|e^{(0)}\|}\right)^{\frac{1}{k}} = \left(\frac{\|P^k e^{(0)}\|}{\|e^{(0)}\|}\right)^{\frac{1}{k}} = \|P^k\|^{\frac{1}{k}}$$

$\theta(e^{(0)}) = \lim_K \theta_K(e^{0}) \leq \lim_K \|P^k\|^{\frac{1}{k}} = \rho(P)$

$\hookrightarrow$ riduzione asintotica media per passo

Teorema 3

Teorema 3 *La riduzione asintotica media per passo $\theta(e^{(0)})$ dell'errore di un metodo iterativo applicato con errore iniziale $e^{(0)}$ è minore o uguale al raggio spettrale della matrice di iterazione P. Inoltre, se $e^{(0)}$ è proporzionale a un autovettore corrispondente ad un autovalore di modulo massimo, allora $\theta(e^{(0)})$ coincide con il raggio spettrale di P.*

Un esempio significativo di metodo iterativo è dato dal *metodo di Richardson* definito dalla relazione

$$x^{(k+1)} = x^{(k)} - \alpha(Ax^{(k)} - b)$$

dove α è un opportuno parametro. Se la matrice è definita positiva, la scelta $\alpha = 1/\|A\|$, con $\|\cdot\|$ norma indotta, garantisce la convergenza del metodo.

Metodi di Jacobi e Gauss-Seidel

$A = D - B - C$

- D = diagonale $d_{i,i} = a_{i,i}$
- B = strett. triang. inf. con $b_{i,j} = -a_{i,j}$ per $i > j$
- C = strett. triang sup con $c_{i,j} = -a_{i,j}$ per $i < j$

$A = M - N$
- Se $M = D \wedge N = B + C \longmapsto$ metodo di Jacobi $\mapsto P = M^{-1}N = J = D^{-1}(B+C)$
- Se $M = D - B \wedge N = C \longmapsto$ metodo di Gauss-Seidel $\mapsto P = M^{-1}N = G = (D-B)^{-1}C$

$\det M \neq 0 \Rightarrow a_{ii} \neq 0 \; \forall i = 1..n$

Teorema 4

Se vale una delle seguenti condizioni $\Rightarrow \rho(J) < 1$ e $\rho(G) < 1$:

- A è fortemente dominante diagonale
- A^T è fortemente dominante diagonale
- A è irriducibilmente dominante diagonale
- A^T è irriducibilmente dominante diagonale

idea dim

Per ass.

1 $\rho(J) > 1 \Rightarrow \exists\ \overset{\text{autovalore}}{\lambda}$ tc. $|\lambda| \geq 1$

A $\det(\lambda I - J) = 0 \Longrightarrow \det H = 0$ con $H = \lambda D - B - C$ H si ottiene moltip la Δ di A per λ

B Quindi se A è F.D.D. $\Rightarrow$ H è F.D.D

C Per I/III Teo Gerschgorin $\Rightarrow \det H \neq 0$ ↯ (idem A^T)

2 $\rho(G) > 1 \quad \exists\ \overset{\text{autovalore}}{\lambda}$ tc. $|\lambda| \geq 1$

A $\det(\lambda I - G) = 0 \Longrightarrow \det H = 0$ con $H = D - B - \lambda^{-1} C$ H si ottiene moltip gli elt tran sup per $\underset{|\lambda^{-1}| \leq 1}{\lambda^{-1}}$

B Quindi se A è I.D.D. $\Rightarrow$ H è I.D.D

C Per I/III Teo Gerschgorin $\Rightarrow \det H \neq 0$ ↯ (idem A^T)

Aspetti computazionali

Jacobi

$x^{(k+1)} = D^{-1}((B+C)x^{(k)} + b) \longmapsto$ A $x_i^{(k+1)} = \frac{1}{a_{i,i}}(b_i - \sum_{j=1, j\neq i}^{n} a_{i,j} x_j^{(k)})$. $\mapsto$ Interviene sulle componenti di $x^{(k)}$ (metodo degli spostamenti simultanei)

costo $\underset{\cdot}{n^2} + \underset{+}{n^2}$

Gauss-Seidel

$x^{(k+1)} = (D-B)^{-1}(Cx^{(k)} + b) \longmapsto$ $x_i^{(k+1)} = \frac{1}{a_{i,i}}(b_i - \sum_{j=1}^{i-1} a_{i,j} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{i,j} x_j^{(k)})$. B Interviene sulle componenti aggiornate di $x^{(k+1)}$

$= D^{-1}(Cx^{(k)} + Bx^{(k+1)} + b)$ (metodo degli spostamenti successivi)

$(D-B)x^{(k+1)} = Cx^{(k)} + b$

costo $\underset{\cdot}{n^2} + \underset{+}{n^2}$

⚠ Gauss-Seidel ha migliori proprietà di convergenza

Matrici sparse

Se A è sparsa $\Rightarrow$ B e C sono sparse $\rightarrow$ costo è proporz. a m. invece di n^2

Fill in

Si perde la proprietà di sparsità (tipico delle matrici a banda) $\mapsto$ costo pari a m^3

Matrici a banda $\xrightarrow{\text{Gauss/Householder}}$ Complemento di Schur $\Rightarrow$ perdi la sparsità.

Jacobi v.s. Gauss-Seidel

Teorema di Stein-Rosenberg

Se A ha elt diagonali non nulli e J ha elt non negativi $\Rightarrow$ vale una sola delle seguenti proprietà:

- $\rho(J) = \rho(G) = 0$
- $0 < \rho(G) < \rho(J) < 1$
- $\rho(J) = \rho(G) = 1$
- $1 < \rho(J) < \rho(G)$

Teorema 6

Se A è una mat. tridiag. con elt diag $\neq 0$. Allora:

1. $\forall\, \lambda$ autov. di J $\exists\, \mu$ aut. di G t.c. $\mu = \lambda^2$
2. $\forall\, \mu \neq 0$ autova. di G $\exists\, \lambda$ autov. di J t.c. $\mu = \lambda^2$
3. In particolare vale $\rho(G) = \rho(J)^2$

Idea dim

dim 1 e 2

(1) λ e μ aut di J e G rispett $\Rightarrow$ $\det(\lambda I - J) = \det(\mu I - G) = 0$

$$\det(\lambda D - B - C) = \det(\mu D - \mu B - C) = 0$$

(a) (b)

(2) α parametro $\neq 0$ $D_\alpha = \operatorname{diag}(1, \alpha, \alpha^2, \dots, \alpha^{n-1})$ $\Rightarrow$
- $D_\alpha D D_\alpha^{-1} = D$
- $D_\alpha B D_\alpha^{-1} = \alpha B$ $\quad B$ strett. bidiag. inf
- $D_\alpha C D_\alpha^{-1} = \alpha^{-1} C$ $\quad C$ strett bidiag sup

(3) Calcolo in a $D_\alpha(\lambda D - B - C) D_\alpha^{-1} = \alpha^{-1}(\lambda\alpha D - \alpha^2 B - C)$ $\Rightarrow$ a diventa $\det(\lambda\alpha D - \alpha^2 B - C) = 0$

(4) Confronto l'espressione precedente con b.

Se scelgo $\alpha = \lambda$ e $\mu = \lambda^2$ $\Rightarrow$

- se $\lambda \neq 0$ aut di J $\Rightarrow$ $\mu = \lambda^2$ è aut di G $\square$ dim 1
- Se $\mu \neq 0$ aut di G $\Rightarrow$ $\exists\, \lambda$ t.c. $\lambda^2 = \mu$ sono aut. di J $\square$ dim 2

dim 3

Da 1+2 + Teo 3 segue che il metodo di Gauss-Seidel richiede la metà dei passi richiesti da Jacobi per ridurre l'errore di una quantità prefissata

Metodi a blocchi

Se la matrice A di dimensione $mn \times mn$ è partizionata in n^2 blocchi $m \times m$

$$A = (A_{i,j}) = \begin{bmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ A_{2,1} & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & \dots & A_{n,n} \end{bmatrix}$$

possiamo considerare una decomposizione additiva $A = D - B - C$ dove D è la matrice diagonale a blocchi con blocchi diagonali uguali a $A_{i,i}$, $i = 1, \dots, n$, $B = (B_{i,j})$ la matrice triangolare inferiore a blocchi tale che $B_{i,j} = -A_{i,j}$ se $i > j$ mentre $B_{i,j} = 0$ altrimenti, e $C = (C_{i,j})$ la matrice triangolare superiore a blocchi tale che $C_{i,j} = -A_{i,j}$ per $i < j$ e $C_{i,j} = 0$ altrimenti. In questo modo, se i blocchi diagonali $A_{i,i}$ di A sono non singolari, possiamo considerare i metodi iterativi definiti da $M = D$, $N = B + C$, e da $M = D - B$, $N = C$. Il primo metodo è detto metodo di *Jacobi a blocchi* mentre il secondo è detto metodo di *Gauss-Seidel a blocchi.*

Teorema 7

Analogo al Teo 6

Sia A è una mat. tridiag. a blocchi cioè t.c. $A_{ij} = 0$ se $|i-j| \geq 2$ con blocchi diag. A_{ii} non singolari

Siano J_B e G_B le mat. di iterazioni a blocchi.

Allora:

1. $\forall\, \lambda$ autov. di J_B $\exists\, \mu$ aut. di G_B t.c. $\mu = \lambda^2$
2. $\forall\, \mu \neq 0$ autova. di G_B $\exists\, \lambda$ autov. di J_B t.c. $\mu = \lambda^2$
3. In particolare vale $\rho(G) = \rho(J)^2$