

Corso di Laurea in Ingegneria Gestionale

Sede di Fermo

Corso di Calcolo Numerico

8 - METODI ITERATIVI PER I SISTEMI LINEARI

Lucio Demeio
DIISM

Norme e distanze

Autovalori ed autovettori

Matrici convergenti

Metodi iterativi

Norme

Una norma in \mathbb{R}^n è una funzione $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ tale che

- $\|\mathbf{x}\| \geq 0 \quad \forall \quad \mathbf{x} \in \mathbb{R}^n$;
- $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$;
- $\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|$;
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$;

Considereremo due norme notevoli in \mathbb{R}^n :

- La norma l_2 (o norma euclidea): se $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$
 $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$;
- La norma l_∞ (norma infinito): $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$

La norma l_2 rappresenta la (ben nota) distanza euclidea dall'origine. In \mathbb{R}^2 , il luogo dei punti (vettori) per i quali $\|\mathbf{x}\|_2 \leq R$ è un cerchio di raggio R , mentre il luogo dei punti per i quali $\|\mathbf{x}\|_\infty \leq R$ è un quadrato di lato $2R$.

Esempio con le norme

Sia $\mathbf{x} = (-1, -2, 0)$. Allora: $\|\mathbf{x}\|_2 = \sqrt{5}$ e $\|\mathbf{x}\|_\infty = 2$.

Verifiche

È facile verificare che le norme l_2 ed l_∞ soddisfano le proprietà elencate all'inizio. Per esempio,

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|_\infty &= \max_{1 \leq i \leq n} |x_i + y_i| \leq \max_{1 \leq i \leq n} (|x_i| + |y_i|) \\ &\leq \max_{1 \leq i \leq n} |x_i| + \max_{1 \leq i \leq n} |y_i| = \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty \\ \|\mathbf{x} + \mathbf{y}\|_2^2 &= \sum_{i=1, n} (x_i + y_i)^2 = \sum_{i=1, n} (x_i^2 + y_i^2 + 2x_i y_i) \\ &\leq \|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2 + 2\|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \\ \text{e quindi } \|\mathbf{x} + \mathbf{y}\|_2 &\leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2\end{aligned}$$

Nel penultimo passaggio è stata applicata la disuguaglianza di Cauchy-Schwarz: $\mathbf{x}^T \mathbf{y} \leq \|\mathbf{x}\| \|\mathbf{y}\|$.

Distanza

Dati due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, si definisce **distanza** tra \mathbf{x} e \mathbf{y} la norma del vettore differenza,

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

Ovviamente, ogni norma porta con sé una distanza.

Limiti e convergenza

Si dice che una successione di vettori $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty} \in \mathbb{R}^n$ converge ad un vettore $\mathbf{x} \in \mathbb{R}^n$ rispetto ad una data norma $\|\cdot\|$ se, per ogni $\varepsilon > 0$, esiste un intero $N(\varepsilon)$ tale che $\|\mathbf{x}^{(k)} - \mathbf{x}\| < \varepsilon$ per ogni $k \geq N(\varepsilon)$.

Theorem

La successione $\{\mathbf{x}^{(k)}\} \in \mathbb{R}^n$ converge ad $\mathbf{x} \in \mathbb{R}^n$ rispetto alla norma l_{∞} se e solo se $\lim_{k \rightarrow \infty} x_i^{(k)} = x_i$ per ogni $i = 1, 2, \dots, n$.

Dimostrazione

Supponiamo che $\{\mathbf{x}^{(k)}\} \rightarrow \mathbf{x}$ rispetto a $\|\cdot\|_\infty$. Allora, dato $\varepsilon > 0$, esiste $N(\varepsilon)$ t.c., per ogni $k \geq N(\varepsilon)$,

$\|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i^{(k)} - x_i| < \varepsilon$. Questo implica che $|x_i^{(k)} - x_i| < \varepsilon$ per ogni $i = 1, 2, \dots, n$, e quindi la tesi.

Viceversa, sia $\lim_{k \rightarrow \infty} x_i^{(k)} = x_i$ per ogni $i = 1, 2, \dots, n$. Allora, dato $\varepsilon > 0$, esistono $N_i(\varepsilon)$ t.c. $|x_i^{(k)} - x_i| < \varepsilon$ per $k > N_i(\varepsilon)$. Sia $N(\varepsilon) = \max_{1 \leq i \leq n} N_i(\varepsilon)$. Se $k > N(\varepsilon)$ allora

$\max_{1 \leq i \leq n} |x_i^{(k)} - x_i| = \|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty < \varepsilon$, che implica la convergenza della successione $\{\mathbf{x}^{(k)}\}$ ad \mathbf{x} .

Norme equivalenti (senza dimostrazione)

Tutte le norme in \mathbb{R}^n sono equivalenti rispetto alle proprietà di convergenza: se una successione $\{\mathbf{x}^{(k)}\} \rightarrow \mathbf{x}$ rispetto a $\|\cdot\|$, allora converge rispetto a qualunque altra norma $\|\cdot\|'$. Per le norme viste qui, vale la relazione $\|\cdot\|_\infty \leq \|\cdot\|_2 \leq \sqrt{n} \|\cdot\|_\infty$.

Convergenza: esempio

Sia

$$\{\mathbf{x}^{(k)}\} = \left(1, 2 + \frac{1}{k}, \frac{3}{k^2}, e^{-k} \sin k\right)^T$$

Abbiamo

$$\lim_{k \rightarrow \infty} 2 + \frac{1}{k} = 2$$

$$\lim_{k \rightarrow \infty} \frac{3}{k^2} = 0$$

$$\lim_{k \rightarrow \infty} e^{-k} \sin k = 0$$

e quindi $\{\mathbf{x}^{(k)}\} \rightarrow (1, 2, 0, 0)^T$.

Norme delle matrici

Sia M_n lo spazio vettoriale delle matrici reali $n \times n$. Una norma su M_n è una funzione $\|\cdot\| : M_n \rightarrow \mathbb{R}$ tale che

- $\|\mathbf{A}\| \geq 0 \quad \forall \mathbf{A} \in M_n$;
- $\|\mathbf{A}\| = 0 \iff \mathbf{A} = \mathbf{0}$;
- $\|\alpha\mathbf{A}\| = |\alpha| \|\mathbf{A}\|$;
- $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$;
- $\|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$;

Distanza

Si definisce **distanza** tra due matrici \mathbf{A} e \mathbf{B} la norma della differenza $\mathbf{A} - \mathbf{B}$:

$$d(\mathbf{A}, \mathbf{B}) = \|\mathbf{A} - \mathbf{B}\|$$

Norme delle matrici (teorema)

Se $\|\cdot\|$ è una norma sui vettori in \mathbb{R}^n , allora

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\| = \max_{\|\mathbf{z}\|\neq 0} \frac{\|\mathbf{Az}\|}{\|\mathbf{z}\|}$$

è una norma su M_n , che viene detta **norma naturale** o **indotta**.

Da qui in poi, salvo eccezioni, useremo sempre questa norma matriciale. L'ultima uguaglianza segue dal fatto che, se \mathbf{z} ha norma non nulla, allora $\mathbf{z}/\|\mathbf{z}\|$ ha norma unitaria.

Corollario

Dal teorema precedente, segue che, per ogni $\mathbf{z} \neq 0$, per ogni $\mathbf{A} \in M_n$ e per ogni norma naturale su M_n abbiamo che

$$\|\mathbf{Az}\| \leq \|\mathbf{A}\| \|\mathbf{z}\|.$$

Norme delle matrici

Le norme naturali che considereremo sono quelle indotte dalle norme già viste per i vettori:

$$\|\mathbf{A}\|_{\infty} = \max_{\|\mathbf{x}\|_{\infty}=1} \|\mathbf{A}\mathbf{x}\|_{\infty}$$

$$\|\mathbf{A}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2,$$

dette rispettivamente norma l_{∞} e norma l_2 .

Theorem (Norma infinito)

Sia $\mathbf{A} = \{a_{ij}\}$ una matrice $n \times n$. Si ha allora

$$\|\mathbf{A}\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

Dimostrazione

Dimostriamo che valgono contemporaneamente le due disuguaglianze

$$\|\mathbf{A}\|_{\infty} \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad \text{e} \quad \|\mathbf{A}\|_{\infty} \geq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

Per dimostrare la prima disuguaglianza, se \mathbf{x} è un vettore con $\|\mathbf{x}\|_{\infty} = 1$, allora

$$\begin{aligned} \|\mathbf{A}\mathbf{x}\|_{\infty} &= \max_{1 \leq i \leq n} |(\mathbf{A}\mathbf{x})_i| = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij}x_j \right| \\ &\leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| |x_j| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \end{aligned}$$

da cui segue la prima disuguaglianza.

Per dimostrare la seconda disuguaglianza, consideriamo dapprima le somme $\sum_{j=1}^n |a_{ij}|$ e sia $i = p$ l'intero per cui questa somma assume il massimo, cioè

$$\sum_{j=1}^n |a_{pj}| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Sia \mathbf{x} il vettore di componenti $x_j = 1$ se $a_{pj} \geq 0$ e $x_j = -1$ se $a_{pj} < 0$. Ovviamente è $\|\mathbf{x}\|_\infty = 1$, ed inoltre $a_{pj}x_j = |a_{pj}|$ per ogni $j = 1, 2, \dots, n$. Pertanto

$$\begin{aligned} \|\mathbf{A}\mathbf{x}\|_\infty &= \max_{1 \leq i \leq n} |(\mathbf{A}\mathbf{x})_i| = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij} x_j \right| \\ &\geq \left| \sum_{j=1}^n a_{pj} x_j \right| = \left| \sum_{j=1}^n |a_{pj}| \right| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \end{aligned}$$

da cui segue la seconda disuguaglianza.

Esempio

$$\mathbf{A} = \begin{pmatrix} 4 & -1 & 4 & -3 & -1 \\ -1 & 3 & -2 & 3 & -3 \\ 3 & -3 & 3 & 0 & 2 \\ -2 & 2 & -2 & 3 & 1 \\ 3 & -3 & -2 & -4 & 2 \end{pmatrix}$$

$$\sum_{j=1,n} |a_{1j}| = 13 \qquad \sum_{j=1,n} |a_{2j}| = 12$$

$$\sum_{j=1,n} |a_{3j}| = 11 \qquad \sum_{j=1,n} |a_{4j}| = 10$$

$$\sum_{j=1,n} |a_{5j}| = 14$$

quindi $\|\mathbf{A}\|_{\infty} = 14$.

Esercizio 7.1.2

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|.$$

- (a) Verifica delle proprietà:
 - $\|\mathbf{x}\|_1 \geq 0$ perchè somma di numeri non negativi;
 - $\mathbf{x} = 0 \Rightarrow \|\mathbf{x}\|_1 = 0$ perchè $x_i = 0 \forall i$, e $\|\mathbf{x}\|_1 = 0 \Rightarrow \mathbf{x} = 0$, perchè se $x_i \neq 0$ per qualche i allora $\sum_{i=1}^n |x_i| > 0$;
 - $\|\alpha \mathbf{x}\|_1 = \sum_{i=1}^n |\alpha x_i| = |\alpha| \|\mathbf{x}\|_1$;
 - $\|\mathbf{x} + \mathbf{y}\|_1 = \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n (|x_i| + |y_i|) = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$.
- (c) Relazione con la norma l_2 :

$$\|\mathbf{x}\|_1^2 = \left(\sum_{i=1}^n |x_i| \right)^2 = \sum_{i=1}^n |x_i|^2 + \sum_{i \neq j} |x_i| |x_j| \geq \sum_{i=1}^n |x_i|^2 = \|\mathbf{x}\|_2^2$$

Esercizio 7.1.7

$$\|\mathbf{A}\|_s = \max_{i,j=1,n} |a_{ij}|.$$

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \mathbf{AB} = \begin{pmatrix} 3 & 3 \\ 2 & 2 \end{pmatrix}$$

Quindi

$$\|\mathbf{A}\|_s = 2 \quad \|\mathbf{B}\|_s = 1 \quad \|\mathbf{AB}\|_s = 3$$

in contraddizione con la proprietà $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$.

Autovalori ed autovettori

- Data una matrice quadrata \mathbf{A} di dimensione n , un vettore non nullo $\mathbf{x} \in \mathbb{R}^n$ si dice **autovettore** di \mathbf{A} se esiste un numero λ , reale o complesso, tale che

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$$

ed in tal caso λ si chiama **autovalore** di \mathbf{A} e l'equazione scritta sopra **equazione agli autovalori**. Tale equazione può anche essere scritta come

$$(\mathbf{A} - \lambda \mathbf{I}) \mathbf{x} = \mathbf{0}$$

e quindi gli autovettori sono le soluzioni non nulle di tale sistema. Siccome si tratta di un sistema omogeneo, esso ammette soluzioni non nulle se e solo se la matrice $\mathbf{A} - \lambda \mathbf{I}$ è singolare, ovvero ha determinante nullo.

- Gli autovettori sono quei vettori che vengono moltiplicati per uno scalare sotto l'azione della matrice, cioè conservano la direzione.

Autovalori ed autovettori

Questa condizione fornisce gli autovalori:

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} \equiv p_n(\lambda) = 0$$

che è un'equazione polinomiale di grado n in λ , detta **equazione caratteristica**. Per il teorema fondamentale dell'algebra, $p_n(\lambda)$ possiede n radici, reali o complesse, ciascuna contata con la propria molteplicità. Se la matrice è reale, allora, se z è una radice, lo è pure z^* . Ricordiamo anche che, se \mathbf{x} è un autovettore, anche $\alpha \mathbf{x}$ lo è, per qualunque α (reale o complesso), cioè gli autovettori sono determinati a meno di una costante arbitraria.

Esempi

[Mathematica file Iterativi1.nb](#)

Raggio spettrale

Sia \mathbf{A} una matrice $n \times n$ e siano λ_i , $i = 1, 2, \dots, n$ gli autovalori. Il massimo, in valore assoluto, degli autovalori è detto **raggio spettrale**:

$$\rho(\mathbf{A}) = \max_{1 \leq i \leq n} |\lambda_i|$$

Ad esempio, se

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

abbiamo

$$\lambda_1 = 1 + i\sqrt{2}, \quad \lambda_2 = 1 - i\sqrt{2}, \quad \lambda_3 = 4$$

con $|\lambda_1| = \sqrt{3}$, $|\lambda_2| = \sqrt{3}$, $|\lambda_3| = 4$ e quindi $\rho(\mathbf{A}) = 4$.

Theorem (Raggio spettrale)

Sia \mathbf{A} una matrice $n \times n$. Allora

- $\|\mathbf{A}\|_2 = [\rho(\tilde{\mathbf{A}}\mathbf{A})]^{1/2}$
- $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$ per ogni norma naturale $\|\cdot\|$.

Dimostrazione

Dimostriamo solo la seconda parte. Se \mathbf{x} è un autovettore appartenente all' autovalore λ , con $\|\mathbf{x}\| = 1$, abbiamo

$$\|\mathbf{A}\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda| \|\mathbf{x}\| = |\lambda|.$$

Ma $\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\|$ e quindi

$$\rho(\mathbf{A}) = \max |\lambda| \leq \|\mathbf{A}\|$$

Se \mathbf{A} è simmetrica, allora $\|\mathbf{A}\|_2 = \rho(\mathbf{A})$.

Autovalori ed autovettori

Calcolo
Numerico

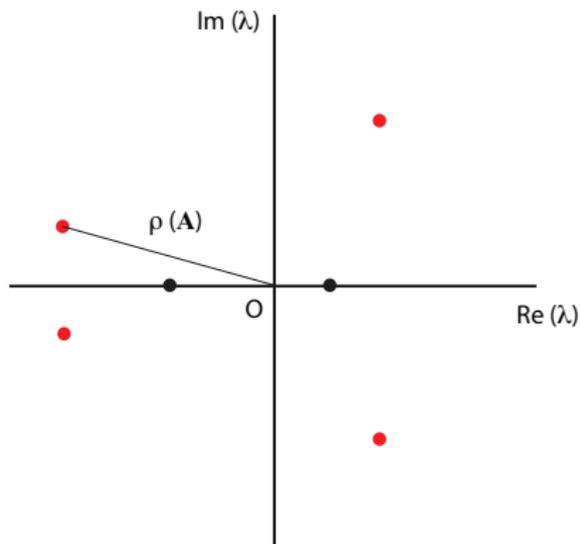
Lucio Demeio
DIISM

Norme e distanze

Autovalori ed
autovettori

Matrici
convergenti

Metodi iterativi



Matrici convergenti

In molte applicazioni, è importante studiare le potenze delle matrici, ed in particolare cosa succede delle potenze successive quando l'esponente va all'infinito, cioè $\lim_{k \rightarrow \infty} \mathbf{A}^k$. Si dà allora la seguente definizione: una matrice si dice che è **convergente** se

$$\lim_{k \rightarrow \infty} (\mathbf{A}^k)_{ij} = 0 \quad \forall 1 \leq i, j \leq n$$

Theorem (Matrici convergenti)

Le seguenti affermazioni sono equivalenti:

- \mathbf{A} è una matrice convergente;
- esiste una norma naturale t.c. $\lim_{n \rightarrow \infty} \|\mathbf{A}^n\| = 0$;
- $\lim_{n \rightarrow \infty} \|\mathbf{A}^n\| = 0$ per tutte le norme naturali;
- $\rho(\mathbf{A}) < 1$;
- $\lim_{n \rightarrow \infty} \mathbf{A}^n \mathbf{x} = \mathbf{0}$ per ogni \mathbf{x} .

Matrici convergenti

Ad esempio, la matrice

$$\begin{pmatrix} \frac{1}{2} & 0 \\ 16 & \frac{1}{2} \end{pmatrix}$$

è convergente (provare!!).

Metodi classici

- I metodi iterativi per la soluzione dei sistemi lineari vengono usati in alternativa ai metodi diretti (eliminazione Gaussiana) quando il numero di equazioni/incognite è elevato e la matrice dei coefficienti è sparsa. In tal caso, i metodi iterativi sono più veloci e richiedono meno memorizzazione dei metodi diretti.
- Un metodo iterativo per risolvere un sistema lineare $\mathbf{Ax} = \mathbf{b}$ parte, in generale, da una stima iniziale $\mathbf{x}^{(0)}$ e costruisce una successione di vettori $\{\mathbf{x}^{(k)}\}$ che converge alla soluzione.
- Vedremo tre metodi classici: quello di **Jacobi**, quello di **Gauss-Seidel** e quello di **Gauss-Seidel con rilassamento**.

Il metodo di Jacobi

Consideriamo il sistema $\mathbf{A} \mathbf{x} = \mathbf{b}$ e scriviamo esplicitamente le equazioni:

$$a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n = b_1$$

$$a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n = b_2$$

...

$$a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nn} x_n = b_n$$

Risolviamo ora la i -esima equazione per l'incognita diagonale x_i , sotto l'ipotesi che $a_{ii} \neq 0$ per ogni $i = 1, 2, \dots, n$:

Il metodo di Jacobi

$$x_1 = [b_1 - (a_{12} x_2 + \dots + a_{1n} x_n)] / a_{11}$$

$$x_2 = [b_2 - (a_{21} x_1 + \dots + a_{2n} x_n)] / a_{22}$$

...

$$x_n = [b_n - (a_{n1} x_1 + a_{n2} x_2 + \dots)] / a_{nn}$$

ovvero

$$x_1 = \left(b_1 - \sum_{j \neq 1} a_{1j} x_j \right) / a_{11}$$

$$x_2 = \left(b_2 - \sum_{j \neq 2} a_{2j} x_j \right) / a_{22}$$

...

$$x_n = \left(b_n - \sum_{j \neq n} a_{nj} x_j \right) / a_{nn}$$

Norme e distanze

Autovalori ed
autovettoriMatrici
convergenti

Metodi iterativi

Il metodo di Jacobi

Supponiamo ora di aver ottenuto il vettore $\{\mathbf{x}^{(k)}\}$ della k -esima iterazione. L'iterazione successiva viene allora data da

$$\begin{aligned}x_1^{(k+1)} &= \left(b_1 - \sum_{j \neq 1} a_{1j} x_j^{(k)} \right) / a_{11} \\x_2^{(k+1)} &= \left(b_2 - \sum_{j \neq 2} a_{2j} x_j^{(k)} \right) / a_{22} \\&\dots \\x_n^{(k+1)} &= \left(b_n - \sum_{j \neq n} a_{nj} x_j^{(k)} \right) / a_{nn}\end{aligned}$$

L'ipotesi $a_{ii} \neq 0$ per ogni i è una condizione necessaria (ma non sufficiente !!) per la convergenza del metodo.

Il metodo di Jacobi - Esempio

$$7x_1 + 3x_2 + x_3 = 18$$

$$2x_1 - 9x_2 + 4x_3 = 12$$

$$x_1 - 4x_2 + 12x_3 = 6$$

Risolvendo per le ingognite diagonali otteniamo lo schema iterativo:

$$x_1^{(k+1)} = (18 - 3x_2^{(k)} - x_3^{(k)})/7$$

$$x_2^{(k+1)} = (2x_1^{(k)} + 4x_3^{(k)} - 12)/9$$

$$x_3^{(k+1)} = (6 - x_1^{(k)} + 4x_2^{(k)})/12$$

Mathematica file `Jacobi.nb`

Il metodo di Gauss-Seidel

Riprendiamo lo schema iterativo di Jacobi:

$$\begin{aligned}x_1^{(k+1)} &= \left(b_1 - \sum_{j \neq 1} a_{1j} x_j^{(k)} \right) / a_{11} \\x_2^{(k+1)} &= \left(b_2 - \sum_{j \neq 2} a_{2j} x_j^{(k)} \right) / a_{22} \\&\dots \\x_n^{(k+1)} &= \left(b_n - \sum_{j \neq n} a_{nj} x_j^{(k)} \right) / a_{nn}\end{aligned}$$

È chiaro che, una volta determinato $x_1^{(k+1)}$ dalla prima equazione, possiamo utilizzare il nuovo valore nelle equazioni successive, al posto di $x_1^{(k)}$. Lo stesso dicasi per $x_2^{(k+1)}$ nelle equazioni dalla terza in poi, e così via per le altre incognite.

Il metodo di Gauss-Seidel

Lo schema iterativo che si ottiene in tal guisa è detto metodo di Gauss-Seidel:

$$x_1^{(k+1)} = \left(b_1 - \sum_{j \neq 1} a_{1j} x_j^{(k)} \right) / a_{11}$$

$$x_2^{(k+1)} = \left(b_2 - a_{21} x_1^{(k+1)} - \sum_{j \geq 3} a_{2j} x_j^{(k)} \right) / a_{22}$$

$$x_3^{(k+1)} = \left(b_3 - a_{31} x_1^{(k+1)} - a_{32} x_2^{(k+1)} - \sum_{j \geq 4} a_{3j} x_j^{(k)} \right) / a_{33}$$

...

$$x_n^{(k+1)} = \left(b_n - \sum_{j < n} a_{nj} x_j^{(k+1)} \right) / a_{nn}$$

ovvero

$$x_i^{(k+1)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) / a_{ii}$$

Versioni matriciali: Jacobi

- Il metodo di Jacobi è equivalente a scrivere la matrice dei coefficienti, \mathbf{A} , come somma della sua parte diagonale e di quella non diagonale, $\mathbf{A} = \mathbf{D} + \mathbf{R}$, dove

$$\mathbf{D} = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix} \quad \mathbf{R} = \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & 0 & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 0 \end{pmatrix}$$

- Il sistema $\mathbf{A} \mathbf{x} = \mathbf{b}$ si può scrivere allora: $(\mathbf{D} + \mathbf{R}) \mathbf{x} = \mathbf{b}$, ovvero $\mathbf{D} \mathbf{x} = \mathbf{b} - \mathbf{R} \mathbf{x}$. Quindi $\mathbf{x} = \mathbf{D}^{-1} (\mathbf{b} - \mathbf{R} \mathbf{x})$ e l'algoritmo iterativo di Jacobi si può scrivere come

$$\mathbf{x}^{(k+1)} = \mathbf{D}^{-1} (\mathbf{b} - \mathbf{R} \mathbf{x}^{(k)}) = \mathbf{T} \mathbf{x}^{(k)} + \mathbf{c}.$$

dove $\mathbf{T} \equiv -\mathbf{D}^{-1} \mathbf{R}$ e $\mathbf{c} = \mathbf{D}^{-1} \mathbf{b}$.

Versioni matriciali: Gauss-Seidel

Per il metodo di Gauss-Seidel scriviamo ancora $\mathbf{A} = \mathbf{D} + \mathbf{R}$,
ma poniamo inoltre $\mathbf{R} = -\mathbf{L} - \mathbf{U}$, dove

$$\mathbf{L} = - \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ a_{21} & 0 & 0 & \dots & 0 \\ a_{31} & a_{32} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 0 \end{pmatrix}$$
$$\mathbf{U} = - \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Il sistema $\mathbf{A} \mathbf{x} = \mathbf{b}$ si può scrivere allora: $(\mathbf{D} - \mathbf{L} - \mathbf{U}) \mathbf{x} = \mathbf{b}$,
ovvero $(\mathbf{D} - \mathbf{L}) \mathbf{x} = \mathbf{U} \mathbf{x} + \mathbf{b}$. Quindi $\mathbf{x} = (\mathbf{D} - \mathbf{L})^{-1} (\mathbf{U} \mathbf{x} + \mathbf{b})$
e l'algoritmo iterativo di Gauss-Seidel si può scrivere come

$$\mathbf{x}^{(k+1)} = (\mathbf{D} - \mathbf{L})^{-1} (\mathbf{U} \mathbf{x}^{(k)} + \mathbf{b}) = \mathbf{T} \mathbf{x}^{(k)} + \mathbf{c}.$$

dove $\mathbf{T} \equiv (\mathbf{D} - \mathbf{L})^{-1} \mathbf{U}$ e $\mathbf{c} = (\mathbf{D} - \mathbf{L})^{-1} \mathbf{b}$.

Forme matriciali

Le forme matriciali per il metodo di Jacobi e di Gauss-Seidel torneranno utili per l'analisi di convergenza dei metodi.

Convergenza dei metodi iterativi

Abbiamo visto che i metodi di Jacobi e di Gauss-Seidel in forma matriciale si scrivono $\mathbf{x}^{(k+1)} = \mathbf{T} \mathbf{x}^{(k)} + \mathbf{c}$, dove $\mathbf{T} \equiv -\mathbf{D}^{-1} \mathbf{R}$ per Jacobi e $\mathbf{T} \equiv (\mathbf{D} - \mathbf{L})^{-1} \mathbf{U}$ per Gauss-Seidel. Il teorema principale sulla convergenza dice che

Theorem

Per ogni vettore $\mathbf{x}^{(0)} \in \mathbb{R}^n$, la successione $\{\mathbf{x}^{(k)}\}$ definita da

$$\mathbf{x}^{(k+1)} = \mathbf{T} \mathbf{x}^{(k)} + \mathbf{c},$$

converge all'unica soluzione del problema $\mathbf{x} = \mathbf{T} \mathbf{x} + \mathbf{c}$ se e solo se $\rho(\mathbf{T}) < 1$.

Per la dimostrazione del teorema, abbiamo bisogno di un lemma preparatorio.

Lemma

Sia \mathbf{T} una matrice $n \times n$ e sia $\rho(\mathbf{T})$ il suo raggio spettrale. Se $\rho(\mathbf{T}) < 1$ allora la matrice $\mathbf{I} - \mathbf{T}$ è invertibile e

$$(\mathbf{I} - \mathbf{T})^{-1} = \mathbf{I} + \mathbf{T} + \mathbf{T}^2 + \dots = \sum_{n=0}^{\infty} \mathbf{T}^n$$

Dimostrazione

Cominciamo con l'osservare che, se $\mathbf{T}\mathbf{x} = \lambda\mathbf{x}$ (cioè se λ è un autovalore di \mathbf{T}) allora $1 - \lambda$ è un autovalore di $\mathbf{I} - \mathbf{T}$, infatti $(\mathbf{I} - \mathbf{T})\mathbf{x} = (1 - \lambda)\mathbf{x}$. Ma, per ipotesi, $\rho(\mathbf{T}) < 1$ e quindi $|\lambda| < 1$ per ogni autovalore. Di conseguenza, la matrice $\mathbf{I} - \mathbf{T}$ non ha autovalori nulli ed è invertibile. Sia ora

$\mathbf{S}_m = \mathbf{I} + \mathbf{T} + \mathbf{T}^2 + \dots + \mathbf{T}^m = \sum_{n=0}^m \mathbf{T}^n$. Allora

$$(\mathbf{I} - \mathbf{T})\mathbf{S}_m = (\mathbf{I} + \mathbf{T} + \mathbf{T}^2 + \dots + \mathbf{T}^m) - (\mathbf{T} + \mathbf{T}^2 + \dots + \mathbf{T}^{m+1}) = \mathbf{I} - \mathbf{T}^{m+1}.$$

Ora, dall'ipotesi $\rho(\mathbf{T}) < 1$ segue che \mathbf{T} è convergente, e quindi

$$\lim_{m \rightarrow \infty} (\mathbf{I} - \mathbf{T}) \mathbf{S}_m = \lim_{m \rightarrow \infty} (\mathbf{I} - \mathbf{T}^{m+1}) = \mathbf{I}$$

$$\text{da cui } (\mathbf{I} - \mathbf{T})^{-1} = \lim_{m \rightarrow \infty} \mathbf{S}_m = \sum_{n=0}^{\infty} \mathbf{T}^n$$

Dimostrazione del teorema

Dimostriamo prima che se $\rho(\mathbf{T}) < 1$ la successione converge.

Consideriamo la successione

$$\begin{aligned} \mathbf{x}^{(k)} &= \mathbf{T} \mathbf{x}^{(k-1)} + \mathbf{c} \\ &= \mathbf{T} (\mathbf{T} \mathbf{x}^{(k-2)} + \mathbf{c}) + \mathbf{c} \\ &= \mathbf{T}^2 \mathbf{x}^{(k-2)} + (\mathbf{T} + \mathbf{I}) \mathbf{c} \\ &\dots \\ &= \mathbf{T}^k \mathbf{x}^{(0)} + (\mathbf{T}^{k-1} + \mathbf{T}^{k-2} + \dots + \mathbf{I}) \mathbf{c} \end{aligned}$$

Riassumendo,

$$\mathbf{x}^{(k)} = \mathbf{T}^k \mathbf{x}^{(0)} + (\mathbf{T}^{k-1} + \mathbf{T}^{k-2} + \dots + \mathbf{I}) \mathbf{c}$$

Ma $\rho(\mathbf{T}) < 1$, quindi \mathbf{T} è convergente e quindi

$$\lim_{k \rightarrow \infty} \mathbf{T}^k \mathbf{x}^{(0)} = \mathbf{0}$$

e quindi, per il lemma precedente,

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \left(\sum_{n=0}^{\infty} \mathbf{T}^n \right) \mathbf{c} = (\mathbf{I} - \mathbf{T})^{-1} \mathbf{c}$$

La successione $\{\mathbf{x}^{(k)}\}$ converge pertanto al vettore $\mathbf{x} = (\mathbf{I} - \mathbf{T})^{-1} \mathbf{c}$, ed abbiamo $\mathbf{x} = \mathbf{T} \mathbf{x} + \mathbf{c}$.

Supponiamo ora che la successione $\{\mathbf{x}^{(k)}\}$ converga alla soluzione (unica) di $\mathbf{x} = \mathbf{T}\mathbf{x} + \mathbf{c}$ e dimostriamo che deve essere $\rho(\mathbf{T}) < 1$. Dimostriamo in realtà l'affermazione equivalente che \mathbf{T} è convergente. Sia allora $\mathbf{z} \in \mathbb{R}^n$ arbitrario e sia \mathbf{x} la soluzione (unica) di $\mathbf{x} = \mathbf{T}\mathbf{x} + \mathbf{c}$. Poniamo $\mathbf{x}^{(0)} = \mathbf{x} - \mathbf{z}$ e costruiamo la successione $\mathbf{x}^{(k)} = \mathbf{T}\mathbf{x}^{(k-1)} + \mathbf{c}$; per ipotesi, $\{\mathbf{x}^{(k)}\}$ converge ad \mathbf{x} e

$$\begin{aligned}\mathbf{x} - \mathbf{x}^{(k)} &= (\mathbf{T}\mathbf{x} + \mathbf{c}) - (\mathbf{T}\mathbf{x}^{(k-1)} + \mathbf{c}) = \mathbf{T}(\mathbf{x} - \mathbf{x}^{(k-1)}) \\ &= \mathbf{T}^2(\mathbf{x} - \mathbf{x}^{(k-2)}) = \dots = \mathbf{T}^k(\mathbf{x} - \mathbf{x}^{(0)}) = \mathbf{T}^k \mathbf{z}\end{aligned}$$

E pertanto

$$\lim_{k \rightarrow \infty} \mathbf{T}^k \mathbf{z} = \lim_{k \rightarrow \infty} (\mathbf{x} - \mathbf{x}^{(k)}) = \mathbf{0}$$

che, per l'arbitrarietà di \mathbf{z} , implica la convergenza di \mathbf{T} .

Alcuni fatti importanti

Elenchiamo di seguito alcuni teoremi sulla convergenza dei metodi iterativi senza dimostrarli; comunque essi seguono dai teoremi visti in precedenza.

- Sia $\|\mathbf{T}\| < 1$ in qualunque norma naturale e sia $\mathbf{c} \in \mathbb{R}^n$ un vettore dato. Allora la successione $\{\mathbf{x}^{(k)}\}$ data da $\mathbf{x}^{(k+1)} = \mathbf{T}\mathbf{x}^{(k)} + \mathbf{c}$ converge, per ogni $\mathbf{x}^{(0)} \in \mathbb{R}^n$ ad un vettore $\mathbf{x} \in \mathbb{R}^n$ e valgono le seguenti limitazioni sull'errore:
 - 1 $\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \|\mathbf{T}\|^k \|\mathbf{x}^{(0)} - \mathbf{x}\|;$
 - 2 $\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \frac{\|\mathbf{T}\|^k}{1 - \|\mathbf{T}\|} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|;$
- Se \mathbf{A} è a dominanza diagonale stretta, allora i metodi di Jacobi e di Gauss-Seidel convergono per qualunque scelta della stima iniziale $\mathbf{x}^{(0)}$;
- il raggio spettrale dà una misura della rapidità della convergenza del metodo iterativo, $\|\mathbf{x}^{(k)} - \mathbf{x}\| \approx \rho(\mathbf{T})^k \|\mathbf{x}^{(0)} - \mathbf{x}\|.$

Alcuni fatti importanti: teorema di Stein-Rosenberg

Indichiamo con \mathbf{T}_J e \mathbf{T}_G le matrici \mathbf{T} rispettivamente del metodo di Jacobi e di Gauss-Seidel e sia \mathbf{A} la matrice dei coefficienti del sistema. Allora, se $a_{ii} > 0$ per ogni i e $a_{ij} \leq 0$ per ogni $i \neq j$, una e solo una delle seguenti affermazioni è vera:

- $0 \leq \rho(\mathbf{T}_G) < \rho(\mathbf{T}_J) < 1$;
- $1 < \rho(\mathbf{T}_J) < \rho(\mathbf{T}_G)$;
- $\rho(\mathbf{T}_J) = \rho(\mathbf{T}_G) = 0$
- $\rho(\mathbf{T}_J) = \rho(\mathbf{T}_G) = 1$.

Vediamo che, sotto le ipotesi particolari enunciate nel teorema, quando uno dei due metodi converge allora convergono entrambi, Gauss-Seidel facendolo con maggior rapidità (vedi prima affermazione) e che quando un metodo diverge allora divergono entrambi (seconda affermazione).

Esercizi

- **Esercizio 7.3.12:** Se \mathbf{A} è a dominanza diagonale stretta, allora $\|\mathbf{T}_J\|_\infty < 1$.
- **Esercizio 7.3.14:** Un oggetto si trova in uno qualunque dei punti equispaziati $\{x_0, x_1, \dots, x_n\}$. Quando si trova in x_i , può saltare in x_{i+1} o in x_{i-1} con la stessa probabilità. Ora, indichiamo con P_i la probabilità che l'oggetto, partendo dalla posizione x_i , raggiunga l'estremo x_0 prima di x_n . Ovviamente, $P_0 = 1$ e $P_n = 0$. Siccome l'oggetto può giungere in x_i solo da x_{i+1} o da x_{i-1} e con la stessa probabilità, ne segue che

$$P_i = \frac{P_{i-1} + P_{i+1}}{2}$$

Dimostrare che le probabilità $\mathbf{p} = (P_1, P_2, \dots, P_{n-1})$ obbediscono ad un sistema lineare $\mathbf{A}\mathbf{p} = \mathbf{b}$, e determinare \mathbf{A} e \mathbf{b} . Risolvere il sistema per $n = 10$, $n = 50$ ed $n = 100$. Cambiare quindi le probabilità di salto in α ed $1 - \alpha$.