

(2.68) Definizione (metodo di Jacobi).

Sia $A \in \mathbb{R}^{n \times n}$ invertibile con elementi diagonali $A(k,k)$ tutti diversi da zero. Posto:¹

$$D = \text{diag}(A) \quad , \quad M = A - D$$

la matrice D risulta invertibile e: $Ax = b$ è equivalente a $x = -D^{-1}Mx + D^{-1}b$.

Il *metodo di Jacobi* (applicato al sistema $Ax = b$) è il metodo iterativo definito da:
 $H_j = -D^{-1}M$ e $c_j = D^{-1}b$.

(2.69) Definizione (matrice a predominanza diagonale forte).

Sia $A \in \mathbb{R}^{n \times n}$. La matrice A è *a predominanza diagonale forte per righe* se

$$\text{per ogni } k: |A(k,k)| > \sum_{i \neq k} |A(k,i)|$$

(2.70) Teorema (predominanza diagonale forte \Rightarrow invertibilità).

Sia $A \in \mathbb{R}^{n \times n}$. Se A è a predominanza diagonale forte per righe allora A è invertibile.

(Dimostrazione: Per assurdo, se A fosse a predominanza diagonale forte per righe e non invertibile allora esisterebbe una colonna $y \neq 0$ tale che $Ay = 0$. Detta y_j la componente di y di massimo modulo (certamente diversa da zero), si avrebbe allora:

$$A(j,1)y_1 + \dots + A(j,j)y_j + \dots + A(j,n)y_n = 0 \quad \text{ovvero} \quad A(j,j)y_j = - \sum_{i \neq j} A(j,i)y_i$$

da cui:

$$|A(j,j)y_j| = \left| \sum_{i \neq j} A(j,i)y_i \right| \Rightarrow |A(j,j)| |y_j| \leq \sum_{i \neq j} |A(j,i)| |y_i|$$

Poiché per definizione $y_j \neq 0$ e per ogni $i \neq j$ è $|y_i| / |y_j| \leq 1$ si avrebbe infine:

$$|A(j,j)| \leq \sum_{i \neq j} |A(j,i)| \left| \frac{y_i}{y_j} \right| \leq \sum_{i \neq j} |A(j,i)|$$

assurdo.)

(2.71) Esempio.

Siano:²

$$A = \begin{bmatrix} 3 & & 1 \\ 1 & 3 & 1 \\ & 3 & 1 \\ 1 & & 3 \end{bmatrix} \quad , \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

1 Se $A \in \mathbb{R}^{n \times n}$, si indica con $\text{diag}(A)$ la matrice $[A(1,1); \dots; A(n,n)]$. La notazione è mutuata da *Scilab*.

2 Se nella scrittura di una matrice un elemento *non è indicato*, il suo valore è *zero*.

- La matrice A risulta a predominanza diagonale forte per righe, e quindi invertibile, e con elementi diagonali tutti diversi da zero. Il metodo di Jacobi è definito e si ha:

$$H_J = -\frac{1}{3} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad c_J = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

- Gli autovalori di H_J ($\lambda_1 = \lambda_2 = 0$, $\lambda_3 = 1/3$, $\lambda_4 = -1/3$) hanno tutti modulo minore di uno. Per il Teorema di caratterizzazione (2.66) della Lezione 22 il metodo risulta convergente. Per ogni g in \mathbb{R}^4 la successione generata dal metodo a partire da g è convergente alla soluzione x^* del sistema $Ax = b$.

(2.72) Teorema (condizione sufficiente di convergenza per il metodo di Jacobi).

Siano $A \in \mathbb{R}^{n \times n}$ a predominanza diagonale forte per righe e $b \in \mathbb{R}^n$. Allora il metodo di Jacobi applicato al sistema $Ax = b$ è convergente.

Il risultato è una semplice conseguenza del teorema e dell'osservazione seguenti.

(2.73) Teorema (norma e raggio spettrale).

Siano $A \in \mathbb{R}^{n \times n}$ e N una norma in \mathbb{R}^n . Allora: $\rho(A) \leq \|A\|_N$.

Dimostrazione. Per definizione: $\|A\|_N = \max\{N(Av), N(v) = 1\}$. Siano poi $\lambda \in \mathbb{C}$ un autovalore di A e $w \in \mathbb{R}^n$ un autovettore associato. Allora, posto $w' = w / N(w)$ si ha:

$$N(w') = 1 \quad \text{e} \quad N(Aw') = N\left(A \frac{w}{N(w)}\right) = \frac{N(Aw)}{N(w)} = \frac{N(\lambda w)}{N(w)} = |\lambda| \frac{N(w)}{N(w)} = |\lambda|$$

quindi $|\lambda| \in \{N(Av), N(v) = 1\}$. Allora:

$$\rho(A) = \max\{|\lambda| \text{ t.c. } \lambda \in \sigma(A)\} \leq \max\{N(Av), N(v) = 1\} = \|A\|_N$$

(2.74) Osservazione.

Siano $A \in \mathbb{R}^{n \times n}$ e $b \in \mathbb{R}^n$. Se A è a predominanza diagonale forte per righe allora per la matrice H_J del metodo di Jacobi applicato al sistema $Ax = b$ si ha $\|H_J\|_\infty < 1$.

(Esercizio: dimostrare che l'asserto è conseguenza immediata della definizione di matrice a predominanza diagonale forte per righe.)

(2.75) Scilab (esempio precedente).

Si consideri l'Esempio (2.71). Per costruire la matrice A in *Scilab*, si utilizzano i seguenti assegnamenti:³

3 In *Scilab*: per ogni numero intero n , `eye(n,n)` è la matrice identica di ordine n ; se A è una matrice e m, k, l sono numeri interi allora: $A(m:k, l) = [A(m, l); \dots; A(k, l)]$.

```
--> A = 3 * eye(4,4)
```

```
A = [4x4 double]
```

```
3.    0.    0.    0.
0.    3.    0.    0.
0.    0.    3.    0.
0.    0.    0.    3.
```

```
--> A(2:4,1) = 1
```

```
A = [4x4 double]
```

```
3.    0.    0.    0.
1.    3.    0.    0.
1.    0.    3.    0.
1.    0.    0.    3.
```

```
--> A(1:3,4) = 1
```

```
A = [4x4 double]
```

```
3.    0.    0.    1.
1.    3.    0.    1.
1.    0.    3.    1.
1.    0.    0.    3.
```

```
--> b = [1;1;1;1]
```

```
b = [4x1 double]
```

```
1.
1.
1.
1.
```

Per costruire la matrice H_j e la colonna c_j :⁴

```
--> D = diag(diag(A))
```

```
D = [4x4 double]
```

```
3.    0.    0.    0.
0.    3.    0.    0.
0.    0.    3.    0.
0.    0.    0.    3.
```

4 In *Scilab*, se A è una matrice $n \times n$ allora $\text{diag}(A) = [A(1,1); \dots; A(n,n)]$; se $v = [v_1; \dots; v_n] \in \mathbb{R}^n$ allora $\text{diag}(v)$ è la matrice $M \in \mathbb{R}^{n \times n}$ diagonale tale che $M(1,1) = v_1, \dots, M(n,n) = v_n$.

```
--> M = A - D
```

```
M = [4x4 double]
```

```
0.    0.    0.    1.
1.    0.    0.    1.
1.    0.    0.    1.
1.    0.    0.    0.
```

```
--> HJ = - diag(1./diag(A)) * M
```

```
HJ = [4x4 double]
```

```
0.          0.    0.  -0.3333333
-0.3333333  0.    0.  -0.3333333
-0.3333333  0.    0.  -0.3333333
-0.3333333  0.    0.    0.
```

```
--> cJ = diag(1./diag(A)) * b
```

```
cJ = [4x1 double]
```

```
0.3333333
0.3333333
0.3333333
0.3333333
```

Un'approssimazione della soluzione del sistema $Ax = b$, calcolata utilizzando la funzione predefinita `backslash (\)`⁵ è:

```
--> y = A\b
```

```
y = [4x1 double]
```

```
0.25
0.1666667
0.1666667
0.2500000
```

Per ottenere un'approssimazione della soluzione con il metodo di Jacobi, si calcolano dieci elementi della successione generata dal metodo a partire dal vettore 0.⁶ Ad ogni iterazione l'istruzione `disp(norm(x - y,%inf))` mostra $\|x - y\|_\infty$ ovvero la distanza tra l'ultimo elemento calcolato, x , della successione e y .

```
--> x = zeros(4,1); for k = 1:10, x = HJ * x + cJ; disp(norm(x - y,%inf)); end
```

```
0.1666667
```

5 L'assegnamento $y = A \backslash b$ è equivalente alla sequenza: $(S,D,P) = \text{EGPP}_M(A)$; $w = SA_M(S,P b)$; $y = SI_M(D,w)$.

6 Se m,n sono numeri interi, `zeros(m,n)` è la matrice di ordine $m \times n$ di elementi tutti uguali a zero.

0.0555556

0.0185185

0.0061728

0.0020576

0.0006859

0.0002286

0.0000762

0.0000254

0.0000085

Si osservi che, come ci si doveva aspettare dalla convergenza della successione, la distanza $\|x - y\|_\infty$ è decrescente.

(2.76) Definizione (metodo di Gauss-Seidel).

Sia $A \in \mathbb{R}^{n \times n}$ invertibile con elementi diagonali $A(k,k)$ tutti diversi da zero. Posto:⁷

$$T = \text{tril}(A) , \quad M = A - T$$

la matrice T risulta invertibile e: $Ax = b$ è equivalente a $x = -T^{-1}Mx + T^{-1}b$.

Il *metodo di Gauss-Seidel* (applicato al sistema $Ax = b$) è il metodo iterativo definito da:
 $H_{GS} = -T^{-1}M$ e $c_{GS} = T^{-1}b$.

(2.77) Esempio.

Siano A e b come nell'Esempio (2.71).

- La matrice A risulta a predominanza diagonale forte per righe, e quindi invertibile, e con elementi diagonali tutti diversi da zero. Il metodo di Gauss-Seidel è definito e si ha:

$$H_{GS} = \begin{bmatrix} & -1/3 \\ & -2/9 \\ & -2/9 \\ & 1/9 \end{bmatrix} \in \mathbb{R}^{4 \times 4} , \quad c_{GS} = \begin{bmatrix} 1/3 \\ 2/9 \\ 2/9 \\ 2/9 \end{bmatrix}$$

- Gli autovalori di H_{GS} ($\lambda_1 = \lambda_2 = \lambda_3 = 0$, $\lambda_4 = 1/9$) hanno tutti modulo minore di uno. Per il Teorema di caratterizzazione (2.66) della Lezione 22 il metodo risulta convergente. Per ogni g in \mathbb{R}^4 la successione generata dal metodo a partire da g è convergente alla soluzione x^* del sistema $Ax = b$.

⁷ Se $A \in \mathbb{R}^{n \times n}$, si indica con $\text{tril}(A)$ la *parte strettamente triangolare inferiore* di A , ovvero la matrice B (triangolare inferiore) tale che: $i \leq j \Rightarrow B(i,j) = A(i,j)$ e $i > j \Rightarrow B(i,j) = 0$. La notazione è mutuata da *Scilab*.

(2.78) Teorema (condizione sufficiente di convergenza per il metodo di Gauss-Seidel).

Siano $A \in \mathbb{R}^{n \times n}$ e $b \in \mathbb{R}^n$. Se:

(1) A è a predominanza diagonale forte per righe

oppure:

(2) A è simmetrica definita positiva

allora il metodo di Gauss-Seidel applicato al sistema $Ax = b$ è convergente.

(2.4) COSTO DELLA SOLUZIONE DI UN SISTEMA DI EQUAZIONI LINEARI CON UN METODO ITERATIVO

(2.79) Osservazione.

Siano $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ e x' , x'' le approssimazioni della soluzione x^* del sistema $Ax = b$ ottenute, rispettivamente, con un *metodo diretto* e con un *metodo iterativo* (dotato, come vedremo, di un opportuno criterio d'arresto). Vogliamo confrontare x' ed x'' dal punto di vista del *costo aritmetico*.

Supponiamo x' calcolata con il procedimento che utilizza EGPP. Il costo asintotico del calcolo è allora: $(2/3) n^3$.

Il costo del calcolo di x'' è:

$$(\text{costo per iterazione}) * (\text{numero di iterazioni})$$

Dobbiamo quindi determinare il costo di una *singola* iterazione.

Consideriamo, ad esempio, il metodo di Gauss-Seidel. Per calcolare la colonna $x(k+1)$ si hanno (almeno) due alternative:

- (1) calcolare $-T^{-1}Mx(k) + T^{-1}b$;
- (2) calcolare la soluzione del sistema $Tx = -Mx(k) + b$.

Per il costo della prima alternativa si ha:

(1.a) $2n^2 - 3n$ operazioni per calcolare $-T^{-1}Mx(k)$

(1.b) n operazioni per calcolare la somma $-T^{-1}Mx(k) + T^{-1}b$

in totale: $2n^2 - 2n$ operazioni.

Per il costo della seconda alternativa si ha:

(2.a) $n^2 - 2n + 1$ operazioni per calcolare $-Mx(k)$

(2.b) n operazioni per calcolare la somma $-Mx(k) + b$

(2.c) n^2 operazioni per calcolare la soluzione del sistema

in totale: $2n^2 - n + 1$ operazioni.

In entrambi i casi il *costo asintotico* è $2n^2$. Dunque: se x'' è stata calcolata con k iterazioni dal metodo di Gauss-Seidel, il costo asintotico del calcolo è $2kn^2$. Il metodo di Gauss-Seidel risulta più economico del metodo diretto che usa EGPP se $k < n/3$.

(Esercizio: verificare i costi per entrambe le alternative.)

Occorre studiare la *rapidità di convergenza* di un metodo iterativo.

(2.80) Esempio.

Siano $H = \text{diag}(s_1, s_2)$ con $|s_2| < |s_1| < 1$ e $c = 0$. Per il Teorema di caratterizzazione dei metodi convergenti (vedi Teorema (2.66) della Lezione 22), il metodo iterativo definito da H e zero è convergente: per ogni g in \mathbb{R}^2 la successione $x(k)$ generata converge a zero. Quanto rapidamente?

Sia:

$$g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \neq 0.$$

Allora:

$$x(k) = H^k g = \text{diag}(s_1^k, s_2^k) g = \begin{bmatrix} s_1^k g_1 \\ s_2^k g_2 \end{bmatrix}$$

e, utilizzando la norma uno:

$$\|x(k)\|_1 = |s_1^k g_1| + |s_2^k g_2|$$

- Se $g_1 \neq 0$:

$$\|x(k)\|_1 = |s_1|^k |g_1| (1 + |s_2/s_1|^k |g_2/g_1|)$$

da cui:

$$\frac{\|x(k)\|_1}{|s_1|^k} \rightarrow |g_1| \neq 0$$

e:

$$\|x(k)\|_1 \text{ tende a zero con la stessa rapidità di } |s_1|^k$$

- Se $g_1 = 0$, invece:

$$\frac{\|x(k)\|_1}{|s_2|^k} \rightarrow |g_2| \neq 0$$

e:

$$\|x(k)\|_1 \text{ tende a zero con la stessa rapidità di } |s_2|^k$$

dunque, essendo $|s_2| < |s_1|$, *più rapidamente* di $|s_1|^k$.

(2.81) Teorema (sulla rapidità di convergenza).

Quanto accade nell'Esempio (2.80) si ritrova in generale.

Si consideri il metodo iterativo convergente definito da $H \in \mathbb{R}^{n \times n}$ e $c \in \mathbb{R}^n$. Detta x^* la soluzione del sistema $(I - H)x = c$ e detta $x(k)$ la successione generata dal metodo a partire da $g \in \mathbb{R}^n$, allora, indicato con $\rho(H)$ il raggio spettrale di H :⁸

$$\|x(k) - x^*\| \text{ converge a zero } \textit{almeno con la stessa rapidità di } \rho(H)^k$$

Inoltre, se il vettore iniziale g è scelto in modo aleatorio, la probabilità che la successione converga a zero più rapidamente di $\rho(H)^k$ è nulla.

(2.82) Esempio.

In base a quanto ottenuto negli esempi (2.71) e (2.77) in cui $\rho(H_J) = 1/3$ e $\rho(H_{GS}) = 1/9$: scelto $g \in \mathbb{R}^2$ in modo aleatorio, la successione generata dal metodo di Gauss-Seidel converge a x^* *più rapidamente* di quella generata dal metodo di Jacobi.

8 Vedere la Definizione (2.65) della Lezione 22.