

Costruzione di metodi iterativi mediante decomposizione della matrice

Vediamo una procedura generale che consente di costruire un metodo della forma (4.2) per risolvere il sistema $Ax=b$. Si considera una decomposizione di A del tipo

$$A = \underline{M} - (M - A), \quad (4.9)$$

con $M \in \mathbb{C}^{n \times n}$ invertibile detta *precondizionatore*. Si osserva che

$$\begin{aligned} Ax = b &\iff (M - (M - A))x = b \iff Mx = (M - A)x + b \\ &\iff x = M^{-1}(M - A)x + M^{-1}b \iff x = x + M^{-1}r(x), \\ &\iff x = \underbrace{(M \cdot M^{-1} - M^{-1}A)}_{\substack{\text{I} \\ \text{I}}} x + M^{-1}b = x - M^{-1}Ax + M^{-1}b = x + M^{-1}(-Ax + b) \end{aligned}$$

dove $r(y) = b - Ay$ è il *residuo in y* del sistema (4.1), e si definisce il metodo

$$\begin{aligned} x^{(0)} &\in \mathbb{C}^n \text{ dato,} \\ x^{(k+1)} &= \underbrace{M^{-1}(M - A)x^{(k)}}_{f_1} + \underbrace{M^{-1}b}_{f_2} = x^{(k)} + M^{-1}r^{(k)}, \quad k = 0, 1, 2, \dots \end{aligned} \quad (4.11)$$

dove $r^{(k)} = r(x^{(k)}) = b - Ax^{(k)}$. Il metodo (4.11) è della forma (4.2) e la sua matrice d'iterazione è $M^{-1}(M - A) = I - M^{-1}A$. Esso è sicuramente consistente con (4.1) grazie alla (4.10). Quindi, per il Teorema $Ax=b$, vale il seguente risultato.

TEOREMA \rightarrow Il metodo (4.11) per risolvere $Ax=b$ è convergente se e solo se $\rho(I - M^{-1}A) < 1$.

OSSERVAZIONE \rightarrow Il polinomio caratteristico di $I - M^{-1}A$ è dato da

$$C_{I-M^{-1}A}(\lambda) = \det(\lambda I - I + M^{-1}A) = \det(\underbrace{M^{-1}(\lambda M - M + A)}_{\substack{\det(A) \times \det(M^{-1}) \\ \text{Binet}}}) = \det(M^{-1}) \det(\lambda M + A - M),$$

dove nell'ultima uguaglianza abbiamo usato il teorema di Binet. Pertanto,

$$C_{I-M^{-1}A}(\lambda) = 0 \iff \det(\lambda M + A - M) = 0.$$

In conclusione, gli autovalori e il raggio spettrale della matrice d'iterazione $I - M^{-1}A$ del metodo (4.11) possono essere calcolati risolvendo l'equazione $(\det(\lambda M + A - M) = 0)$, senza quindi calcolare esplicitamente né l'inversa M^{-1} né la matrice $I - M^{-1}A$.

OSSERVAZIONE \rightarrow L'iterazione k -esima del metodo (4.11) viene normalmente calcolata mediante la formula (f2)

$x^{(k+1)} = x^{(k)} + M^{-1}r^{(k)}$ e richiede il calcolo del vettore $z^{(k)} = M^{-1}r^{(k)}$ detto *residuo preconditionato*. Tale calcolo viene fatto risolvendo il sistema lineare $Mz^{(k)} = r^{(k)}$ * che deve essere più facile/rapido da risolvere rispetto al sistema originario $Ax = b$, altrimenti non c'è nessun guadagno nel risolvere il sistema originario con il metodo (4.11)!

* e non calcolando l'inversa M^{-1} (cosa tipicamente sconsigliata dal punto di vista computazionale)

OSSERVAZIONE → Intuitivamente, quanto più il preconditionatore M “è vicino/assomiglia” alla matrice

A , tanto più il metodo (4.11) convergerà velocemente. Infatti, se $M \approx A$ allora intuitivamente $M - A \approx O$, $M^{-1}A \approx I$ e $M^{-1}(M - A) = I - M^{-1}A \approx O$, per cui ci si può aspettare che $\rho(I - M^{-1}A)$ sia piccolo. Il caso limite $M = A$ è quello in cui $I - M^{-1}A = O$ e il metodo converge in un’iterazione alla soluzione esatta \mathbf{x} , ma al prezzo che tale iterazione costa come la risoluzione del sistema originario $A\mathbf{x} = \mathbf{b}$. Nella scelta del preconditionatore M occorre quindi mediare fra **2 cose**:

- “qualità dell’approssimazione $M \approx A$ ^{circa A} una buona approssimazione $M \approx A$ generalmente assicura un buona velocità di convergenza;
- la facilità/rapidità della risoluzione di un sistema lineare di matrice M assicura che ogni iterazione del metodo (4.11) è rapida.

N.B: \approx significa ASSOMIGLIA

METODO DI JACOBI

Supponiamo che A abbia elementi diagonali non nulli. In tal caso, la parte diagonale di A , cioè la matrice diagonale D ottenuta ricopiando la parte diagonale di A , è invertibile e possiamo definire il metodo di Jacobi, cioè il metodo (4.11) con $M = D$:

$$\begin{aligned} \mathbf{x}^{(0)} &\in \mathbb{C}^n \text{ dato,} \\ \mathbf{x}^{(k+1)} &= D^{-1}(D - A)\mathbf{x}^{(k)} + D^{-1}\mathbf{b} = \mathbf{x}^{(k)} + D^{-1}\mathbf{r}^{(k)}, \quad k = 0, 1, 2, \dots \end{aligned}$$

Il metodo di Jacobi è convergente se e solo se $\rho(J) < 1$, dove $J = D^{-1}(D - A) = I - D^{-1}A$. L’iterazione k -esima del metodo di Jacobi richiede di calcolare il vettore $\mathbf{z}^{(k)} = D^{-1}\mathbf{r}^{(k)}$ risolvendo il sistema diagonale $D\mathbf{z}^{(k)} = \mathbf{r}^{(k)}$, il che è facilissimo:

$$\begin{cases} a_{11}z_1^{(k)} &= r_1^{(k)} \\ &a_{22}z_2^{(k)} &= r_2^{(k)} \\ &&\ddots &\vdots \\ &&&a_{nn}z_n^{(k)} &= r_n^{(k)} \end{cases} \iff \begin{cases} z_1^{(k)} = r_1^{(k)} / a_{11} \\ z_2^{(k)} = r_2^{(k)} / a_{22} \\ \vdots \\ z_n^{(k)} = r_n^{(k)} / a_{nn} \end{cases} \quad (4.14)$$

Il costo del calcolo di $\mathbf{z}^{(k)}$ è nD .¹⁹

METODO DI GAUSS-SEIDEL

Supponiamo che A abbia elementi diagonali non nulli. In tal caso, la parte triangolare inferiore di A , cioè la matrice triangolare inferiore E ottenuta ricopiando la parte triangolare inferiore di A (inclusa la diagonale), è invertibile e possiamo definire il metodo di Gauss-Seidel, cioè il metodo (4.11) con $M = E$:

$$\begin{aligned} \mathbf{x}^{(0)} &\in \mathbb{C}^n \text{ dato,} \\ \mathbf{x}^{(k+1)} &= \underbrace{E^{-1}(E - A)\mathbf{x}^{(k)}}_{\mathbf{f}_1} + \underbrace{E^{-1}\mathbf{b}}_{\mathbf{f}_2} = \mathbf{x}^{(k)} + E^{-1}\mathbf{r}^{(k)}, \quad k = 0, 1, 2, \dots \end{aligned} \quad (4.15)$$

Il metodo di Gauss-Seidel è convergente se e solo se $\rho(G) < 1$, dove $G = E^{-1}(E - A) = I - E^{-1}A$. L'iterazione k -esima del metodo di Gauss-Seidel richiede di calcolare il vettore $\mathbf{z}^{(k)} = E^{-1}\mathbf{r}^{(k)}$ risolvendo il sistema triangolare inferiore $E\mathbf{z}^{(k)} = \mathbf{r}^{(k)}$, il che è facile (la soluzione si ottiene per sostituzione in avanti):

$$\begin{cases} a_{11}z_1^{(k)} & = r_1^{(k)} \\ a_{21}z_1^{(k)} + a_{22}z_2^{(k)} & = r_2^{(k)} \\ a_{31}z_1^{(k)} + a_{32}z_2^{(k)} + a_{33}z_3^{(k)} & = r_3^{(k)} \\ \vdots & \vdots \\ a_{n1}z_1^{(k)} + a_{n2}z_2^{(k)} + \dots + a_{nn}z_n^{(k)} & = r_n^{(k)} \end{cases} \iff \begin{cases} z_1^{(k)} = r_1^{(k)} / a_{11} \\ z_2^{(k)} = (r_2^{(k)} - a_{21}z_1^{(k)}) / a_{22} \\ z_3^{(k)} = (r_3^{(k)} - a_{31}z_1^{(k)} - a_{32}z_2^{(k)}) / a_{33} \\ \vdots \\ z_n^{(k)} = \underbrace{(r_n^{(k)} - a_{n1}z_1^{(k)} - a_{n2}z_2^{(k)} - \dots - a_{n,n-1}z_{n-1}^{(k)})}_{Q_{nn}} \end{cases} \quad (4.16)$$

Per ogni $i = 1, \dots, n$, il costo del calcolo di

$$z_i^{(k)} = \frac{r_i - a_{i1}z_1^{(k)} - a_{i2}z_2^{(k)} - \dots - a_{i,i-1}z_{i-1}^{(k)}}{a_{ii}}$$

è $1D + (i-1)M + (i-1)A$, per cui il costo complessivo del calcolo di $\mathbf{z}^{(k)}$ è

$$\sum_{i=1}^n (1D + (i-1)M + (i-1)A) = nD + \frac{n(n-1)}{2}M + \frac{n(n-1)}{2}A.$$

Questo costo può ridursi se la parte triangolare inferiore E di A ha molti zeri.

OSSERVAZIONE \longrightarrow Confrontando i preconditionatori D ed E dei metodi di Jacobi e Gauss-Seidel, osserviamo quanto segue:

- L'approssimazione $E \approx A$ è migliore dell'approssimazione $D \approx A$ perché $E - A$ ha più zeri di $D - A$. Questo spiega perché molto spesso il metodo di Gauss-Seidel converge più velocemente del metodo di Jacobi (cioè $\rho(G) < \rho(J)$, essendo J e G le matrici d'iterazione di Jacobi e Gauss-Seidel).
- La risoluzione di un sistema lineare di matrice E è più costosa della risoluzione di un sistema lineare di matrice D (cfr. (4.16) e (4.14)). Pertanto, un'iterazione di Gauss-Seidel costa di più di un'iterazione di Jacobi.