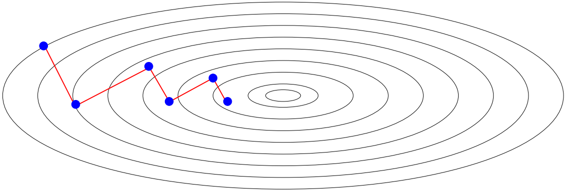
Il metodo del gradiente coniugato

Il metodo del gradiente puó avere un comportamento di tipo 'zig-zag'.



Nel metodo del gradiente $\mathbf{r}^{(k)^T}\mathbf{r}^{(k+1)} = 0$
Il metodo del gradiente coniugato sceglie le direzioni di discesa A-ortogonali

$$\mathbf{d}^{(k)^T} A\, \mathbf{d}^{(l)} = 0 \text{ se } k \neq l\,.$$

Questa scelta permette di ottenere, nel caso di matrici simmetriche definite positive di ordine n, la soluzione in un numero di passi $\leq n$.

Il metodo del gradiente coniugato

La nuova direzione di discesa $\mathbf{d}^{(k+1)}$ non sarà il residuo $\mathbf{r}^{(k+1)}$ ma sarà della forma

$$\mathbf{d}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{d}^{(k)}$$

Il numero reale β_k sarà tale che

$$\mathbf{d}^{(k)^T} A\, \mathbf{d}^{(k+1)} = \mathbf{d}^{(k)^T} A \left(\mathbf{r}^{(k+1)} + \beta_k \mathbf{d}^{(k)}\right) = 0$$

cioè

$$\beta_k = -\frac{\mathbf{d}^{(k)^T} A\, \mathbf{r}^{(k+1)}}{\mathbf{d}^{(k)^T} A\, \mathbf{d}^{(k)}}$$

Il metodo del gradiente coniugato

Scelta la direzione di discesa $\mathbf{d}^{(k)}$ si cerca il minimo di

$$\begin{aligned} J(\alpha) &= \Phi(\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)}) \\ &= \tfrac{1}{2}(\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)})^T A (\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)}) - (\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)})^T \mathbf{b} \end{aligned}$$

Siccome

$$J'(\alpha) = {\mathbf{d}^{(k)}}^T (A\mathbf{x}^{(k)} - \mathbf{b}) + \alpha {\mathbf{d}^{(k)}}^T A \mathbf{d}^{(k)}$$

il minimo si ha in α_k tale che $J'(\alpha_k) = 0$, quindi

$$\alpha_k = \frac{{\mathbf{d}^{(k)}}^T \mathbf{r}^{(k)}}{{\mathbf{d}^{(k)}}^T A \, \mathbf{d}^{(k)}}$$

Il metodo del gradiente coniugato

Dato $\mathbf{x}^{(0)} \in \mathbb{R}^n$, posto $\mathbf{d}^{(0)} = \mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$ per $k = 0, 1, \dots$
fino a convergenza

$$\alpha_k = \frac{{\mathbf{d}^{(k)}}^T \mathbf{r}^{(k)}}{{\mathbf{d}^{(k)}}^T A\, \mathbf{d}^{(k)}} = \frac{{\mathbf{r}^{(k)}}^T \mathbf{r}^{(k)}}{{\mathbf{d}^{(k)}}^T A\, \mathbf{d}^{(k)}}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A\mathbf{d}^{(k)}$$

$$\beta_k = -\frac{{\mathbf{d}^{(k)}}^T A\, \mathbf{r}^{(k+1)}}{{\mathbf{d}^{(k)}}^T A\, \mathbf{d}^{(k)}} = \frac{{\mathbf{r}^{(k+1)}}^T \mathbf{r}^{(k+1)}}{{\mathbf{r}^{(k)}}^T \mathbf{r}^{(k)}}$$

$$\mathbf{d}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{d}^{(k)}$$

Il metodo del gradiente coniugato

Teorema Se A è simmetrica definita positiva il metodo del gradiente coniugatoconverge al più in n iterazioni e

$$\|\mathbf{e}^{(k+1)}\|_A \leq 2c^k \|\mathbf{e}^{(0)}\|_A$$

dove

$$c = \frac{\sqrt{K_2(A)} - 1}{\sqrt{K_2(A)} + 1} \, .$$

La rapidità di convergenza dei metodi di discesa dipende dal numero di condizionamento della matrice del sistema.

Precondizionamento

Sostituire $A\mathbf{x} = \mathbf{b}$ col sistema equivalente $P^{-1}A\mathbf{x} = P^{-1}\mathbf{b}$.

- La matrice P si sceglie in modo che $K(P^{-1}A) << K(A)$.
- Assumiamo P^{-1} simmetrica definita positiva. In generale $P^{-1}A$ non sarà simmetrica definita positiva.
- Il sistema lineare

$$P^{-1/2}AP^{-1/2}\mathbf{y} = P^{-1/2}\mathbf{b} \qquad (1)$$

  con $\mathbf{y} = P^{1/2}\mathbf{x}$ è equivalente al sistema $A\mathbf{x} = \mathbf{b}$ e la matrice $P^{-1/2}AP^{-1/2}$ è simmetrica definita positiva.
    - $P^{-1/2}$ è una matrice simmetrica definita positiva tale che $(P^{-1/2})^2 = P^{-1}$.
    - $P^{1/2}$ è l'inversa di $P^{-1/2}$.
- Il metodo del gradiente coniugato applicato al sistema (1) non richiede di valutare $P^{-1/2}$ o $P^{1/2}$.

Il metodo del gradiente coniugato precondizionato

Dato $\mathbf{x}^{(0)} \in \mathbb{R}^n$, posto $\mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$, $\mathbf{z}^{(0)} = P^{-1}\mathbf{r}^{(0)}$ e $\mathbf{d}^{(0)} = \mathbf{z}^{(0)}$ per $k = 0, 1, \ldots$ fino a convergenza

$$\alpha_k = \frac{{\mathbf{z}^{(k)}}^T \mathbf{r}^{(k)}}{{\mathbf{d}^{(k)}}^T A\, \mathbf{d}^{(k)}}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A\mathbf{d}^{(k)}$$

$$P\mathbf{z}^{(k+1)} = \mathbf{r}^{(k+1)}$$

$$\beta_k = \frac{{\mathbf{z}^{(k+1)}}^T \mathbf{r}^{(k+1)}}{{\mathbf{z}^{(k)}}^T \mathbf{r}^{(k)}}$$

$$\mathbf{d}^{(k+1)} = \mathbf{z}^{(k+1)} + \beta_k \mathbf{d}^{(k)}$$

Il metodo del gradiente coniugato precondizionato

Per il metodo del gradiente coniugato precondizionato

$$\|\mathbf{e}^{(k+1)}\|_A \leq 2c^k \|\mathbf{e}^{(0)}\|_A$$

dove

$$c = \frac{\sqrt{K_2(P^{-1}A)} - 1}{\sqrt{K_2(P^{-1}A)} + 1}.$$

In generale non è possibile individuare a priori il precondizionatore ottimale.

La scelta del precondizionatore deve essere guidata da considerazione computazionali come il costo e la memoria richiesta.

Esercizi

Dato il sistema lineare

$$\begin{bmatrix} 4 & -1 & 2 \\ 1 & 3 & 1 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ -1 \end{bmatrix}$$

- Verificare che il metodo di Jacobi e il metodo di Gauss-Seidel sono convergenti.
- Partendo da $x^{(0)} = [1, 0, 0]^T$ fare due iterazioni del metodo di Jacobi e due iterazioni del metodo di Gauss-Seidel.

Esercizi

Date le matrici

$$A_1 = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & -1 \\ -2 & -2 & 1 \end{bmatrix} \quad \text{e} \quad A_2 = \begin{bmatrix} 2 & -1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & 2 \end{bmatrix}$$

- Calcolare il raggio spettrale della matrice d'iterazione del metodo di Jacobi, $\rho(B_J)$, e della matrice d'iterazione del metodo di Gauss-Sidel, $\rho(B_{GS})$.
- Confrontare la convergenza dei due metodi.

Esercizi

Data la matrice

$$A = \begin{bmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix}$$

- Calcolare la matrice d'iterazione del metodo di Jacobi, B_J.
- Calcolare il raggio spettrale di B_J, $\rho(B_J)$.
- È convergente il metodo di Jacobi per un sistema lineare con matrice A?
- Verificare che il metodo SOR per un sistema lineare con matrice A è convergente se $0 < \omega < 2$.
- Si consideri il sistema lineare $A\mathbf{x} = \begin{bmatrix} 24 \\ 30 \\ -24 \end{bmatrix}$. Partendo da $x^{(0)} = [1, 1, 1]^T$ fare un'iterazione del metodo SOR con $\omega = 1.25$.