• Data una matrice A simmetrica e definita positiva si considerino, per $A \in \mathcal{M}(n \times n)$, $b \in \mathbb{R}^n$:

1) il sistema lineare $Ax = b$

2) il funzionale $\phi(y) = \frac{1}{2} y^T A y - y^T b = \frac{1}{2}(Ay, y) - (b, y)$

dove $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$

• Problema di minimizzazione del funzionale:

Trovare $x^* \in \mathbb{R}^n$ tale che

$$\phi(x^*) = \min_{y \in \mathbb{R}^n} \phi(y)$$

Teorema.

x è soluzione di $Ax = b \iff x = x^*$, cioè x è la soluzione del problema di minimizzazione del funzionale.

$\boxed{\Rightarrow}$ x soluzione di $Ax = b \Rightarrow x = x^*$

$$\boxed{\phi(y) - \phi(x)} = \frac{1}{2}(Ay, y) - (b, y) - \frac{1}{2}(Ax, x) + (b, x) =$$

$$\frac{1}{2}(Ay, y) - (Ax, y) - \frac{1}{2}(Ax, x) + (Ax, x) =$$

$$\frac{1}{2}(Ay - Ax, y) + \frac{1}{2}(Ax, x) - \frac{1}{2}(Ax, y) =$$

N.B. A simmetrica $(Ax, y) = (x, Ay)$ $\forall x, y$

$$= \frac{1}{2}\left(A(y-x), y\right) - \frac{1}{2}(Ax, y-x) = \frac{1}{2}\left(A(y-x), y\right) - \frac{1}{2}\left(A(y-x), x\right) =$$

$$\frac{1}{2}\left(A(y-x), y-x\right)$$

Essendo A definita positiva

$\boxed{\left(A(y-x), y-x\right)} \boxed{\geq 0} \quad \forall\, y-x \Rightarrow \phi(y) \geq \phi(x)$

$\left(A(y-x), y-x\right) = 0 \iff y-x=0 \iff y=x$

Dunque $\phi(y) \geq \phi(x) \quad \forall y$

$\phi(y) = \phi(x) \iff y=x$

Dunque x è soluzione del problema di minimo di ϕ

$\boxed{\Leftarrow}$ Sia y soluzione del problema di minimo per ϕ
Allora $y=x$, cioè $Ay=b$ (la soluzione è unica, essendo $\det A \neq 0$)

Se x è soluzione del problema di minimo per ϕ, allora $\nabla\phi(x)=0$

$$\phi(y) = \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}\, y_i\, y_k - \sum_{i=1}^{n} b_i y_i$$

$$= \frac{1}{2}\sum_{i=1}^{n} y_i \sum_{k=1}^{n} a_{ik}\, y_k - \sum_{i=1}^{n} b_i y_i$$

$$\frac{\partial\phi(y)}{\partial y_j} = \frac{1}{2}\left(\sum_{k=1}^{n} a_{jk}\, y_k + \sum_{i=1}^{n} y_i\, a_{ij}\right) - b_j$$

Essendo A simmetrica:

$$= \sum_{k=1}^{n} a_{jk}\, y_k - b_j = 0 \iff (A\underline{y})_j - b_j = 0$$

$$\forall j = 1, \dots, n$$

In forma vettoriale:

$$\nabla\phi = 0 \iff Ay - b = 0 \iff Ay = b$$

Il sistema $Ax = b$ ammette una ed una sola soluzione. Dunque $y = x$ [N.B. $\underset{\downarrow \text{ Hessiana}}{H(y)} = A$ che è definita positiva $\Downarrow$ Minimo assoluto]

CONCLUSIONE:

- Se x è di minimo, $\nabla\phi(x) = 0 \Rightarrow Ax = b$
- Se $Ax = b$, allora $\phi(y) \geq \phi(x) \quad \forall y \in \mathbb{R}^n$

  e $\quad \phi(y) = \phi(x) \iff y = x$

Inoltre: $\quad Ay - b = -r$ (residuo associato a y relativo al sistema $Ax = b$)

Richiami: derivata direzionale

• Derivata direzionale:

Dato $d \in \mathbb{R}^n$, $\|d\|_2 = 1$:

$$\lim_{\lambda \to 0} \frac{\phi(x+\lambda d) - \phi(x)}{\lambda} = \nabla\phi(x) \cdot d$$

• Si dice che d è direzione di discesa in x per ϕ se $\exists\ \bar{\lambda} > 0$ t.c.

$$\phi(x+\lambda d) < \phi(x) \quad \forall\ 0 < \lambda < \bar{\lambda}$$

$$\Downarrow$$

Teorema

Se la derivata direzionale è negativa allora d è direzione di discesa (in x per ϕ)

Oss.ni

•) $\nabla\phi(x) \cdot d > 0$ direzione di salita (ϕ crescente)

•) $\nabla\phi(x) \cdot d = 0$ $\nabla\phi(x) \perp d \Rightarrow$ ortogonalità (ϕ costante sulle curve / superfici di livello)

• $\nabla\phi(x) \cdot d < 0$ direzione di discesa (ϕ decrescente)

Richiami: angolo tra due vettori

Dati $x, y \in \mathbb{R}^n$:

$$\cos\theta = \frac{x^T y}{\|x\|_2 \|y\|_2}$$

$\cos\theta = 0$: $x \perp y$ (ortogonali)

$\cos\theta = 1$: $y = \alpha x \quad \alpha > 0$

$\cos\theta = -1$: $y = \alpha x \quad \alpha < 0$

Applicazione:

$$d = -\nabla\phi(x) \Rightarrow \nabla\phi(x) \cdot d = -\nabla\phi(x) \cdot \phi(x)$$

$$= -\|\nabla\phi(x)\|_2^2 < 0$$

$d = -\nabla\phi(x)$ è una direzione di discesa in x per ϕ, ove $d = -\nabla\phi(x) = r = b - Ax$

RICHIAMI:

Disuguaglianza di Cauchy-Schwarz

$$|(u, v)| \leq \|u\|_2 \cdot \|v\|_2$$

$$-\|u\|_2 \|v\|_2 \leq (u, v) \leq \|u\|_2 \|v\|_2$$

Applicazione:

$$|(\nabla\phi(x), d)| = |\nabla\phi(x) \cdot d| \leq \|\nabla\phi(x)\|_2 \cdot \|d\|_2$$

$$-\|\nabla\phi(x)\|_2 \underbrace{\|d\|_2}_{1} \leq \nabla\phi(x) \cdot d \leq \|\nabla\phi(x)\|_2 \underbrace{\|d\|_2}_{1}$$

$$-\|\nabla\phi(x)\|_2 \leq \nabla\phi(x) \cdot d \leq \|\nabla\phi(x)\|_2$$

Scegliendo $d = \dfrac{\nabla\phi(x)}{\|\nabla\phi(x)\|_2}$ avente norma euclidea unitaria

$$\Downarrow$$

$$\nabla\phi(x) \cdot \frac{\nabla\phi(x)}{\|\nabla\phi(x)\|_2} = \|\nabla\phi(x)\|_2$$

si ottiene il massimo valore di $\nabla\phi(x) \cdot d$

$\Rightarrow$ direzione di massima discesa

Applicazione al problema della soluzione approssimata di sistemi lineari con metodi iterativi (metodi di discesa o del gradiente)

$$x_{k+1} = x_k + \alpha_k r_k \qquad \text{con } r_k = b - Ax_k$$

Come determinare α_k:

1) minimizzare $\phi(x_{k+1})$
2) minimizzare (r_{k+1}, r_{k+1})

1) Trovare α_k tale che

$$\phi(x_k+\alpha_k r_k) = \min_{\alpha\in\mathbb{R}} \phi(x_k+\alpha r_k)$$

$$\phi(x_k+\alpha r_k) = \frac{1}{2}\left(A(x_k+\alpha r_k), x_k+\alpha r_k\right) - (b, x_k+\alpha r_k) =$$

$$\frac{1}{2}\left[(Ax_k, x_k) + \alpha(Ar_k, x_k) + \alpha(Ax_k, r_k) + \alpha^2(Ar_k, r_k)\right]$$

$$-(b,x_k) - \alpha(b,r_k) =$$

$$\frac{1}{2}\left[(Ax_k, x_k) + 2\alpha(Ax_k, r_k) + \alpha^2(Ar_k, r_k)\right] - (b,x_k) - \alpha(b,r_k) =$$

$$= \frac{1}{2}(Ax_k, x_k) + \alpha(Ax_k, r_k) + \frac{1}{2}\alpha^2(Ar_k, r_k) - (b,x_k) - \alpha(b,r_k) =$$

$$= \frac{1}{2}\alpha^2(Ar_k, r_k) - \alpha(\underbrace{b-Ax_k}_{r_k}, r_k) + \frac{1}{2}(Ax_k, x_k) - (b, x_r) =$$

$$= E(\alpha)$$

Minimo : $\dfrac{dE(\alpha)}{d\alpha} = 0$

$$\alpha(Ar_k, r_k) - (r_k, r_k) = 0$$

$$\alpha = \frac{(r_k, r_k)}{(Ar_k, r_k)}$$

$$\Rightarrow \alpha_k = \frac{(r_k, r_k)}{(Ar_k, r_k)}$$

2) Trovare α_k tale che

$$\Big(\underbrace{b - A\underbrace{(x_k + \alpha_k r_k)}_{x_{k+1}}}_{r_{k+1}}, b - A(x_k + \alpha_k r_k)\Big) = \min_{\alpha \in \mathbb{R}} \big(b - A(x_k + \alpha r_k), b - A(x_k + \alpha r_k)\big)$$

$$(\underbrace{b - Ax_k}_{r_k} - \alpha A r_k, \underbrace{b - Ax_k}_{r_k} - \alpha A r_k) =$$

$$(r_k - \alpha A r_k, r_k - \alpha A r_k) =$$

$$(r_k, r_k) - \alpha (A r_k, r_k) - \alpha (r_k, A r_k) + \alpha^2 (A r_k, A r_k) =$$

$$\alpha^2 (A r_k, A r_k) - 2\alpha (A r_k, r_k) + (r_k, r_k) = F(\alpha)$$

Minimo: $\dfrac{dF(\alpha)}{d\alpha} = 0$

$$2\alpha (A r_k, A r_k) - 2 (A r_k, r_k) = 0$$

$$\alpha = \frac{(A r_k, r_k)}{(A r_k, A r_k)}$$

Algoritmo

- $r_k = b - A x_k$
- $x_{k+1} = x_k + \alpha_k r_k$
- $r_{k+1} = b - A x_{k+1} = b - A x_k - \alpha_k A r_k = r_k - \alpha_k A r_k$

Inizializzazione

x_0 dato, $r_0 = b - A x_0$

$k \geq 0$ $\quad z_k = A r_k$

$$\alpha_k = \frac{(r_k, r_k)}{(z_k, r_k)} \quad \text{oppure} \quad \alpha_k = \frac{(z_k, r_k)}{(z_k, z_k)}$$

$$x_{k+1} = x_k + \alpha_k r_k$$

$$r_{k+1} = r_k - \alpha_k z_k$$

Test: $\dfrac{\|r_{k+1}\|_2}{\|b\|_2} < \varepsilon$

Se $\alpha_k = \alpha \quad \forall k \Rightarrow \alpha = \dfrac{2}{\lambda_{max} + \lambda_{min}}$

$\lambda_{max} = \max \lambda(A)$

$\lambda_{min} = \min \lambda(A)$