19/10/26

Metodi iterativi

Data $A \in \mathbb{R}^{n \times n}$ non sing, dato $\underline{b} \in \mathbb{R}^{n \times 1}$? $\underline{x} \in \mathbb{R}^{n \times 1} : A\underline{x} = \underline{b}$

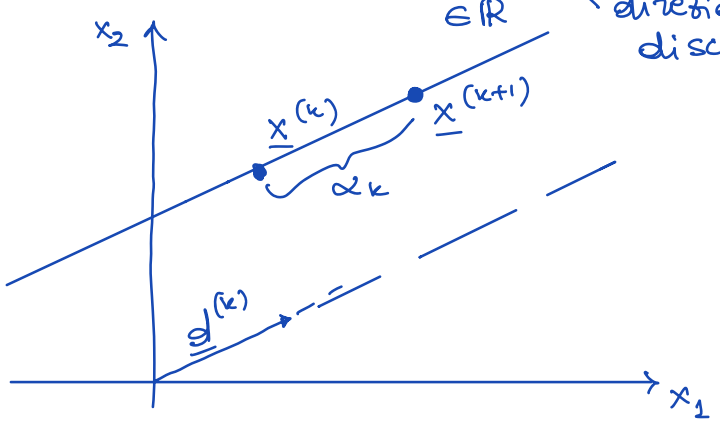
Dato un vettore $\underline{x}^{(0)} \in \mathbb{R}^{n \times 1}$

costruisco una succ $\underline{x}^{(k)}$ con $k \geq 0$: $\lim_{k \to \infty} \underline{x}^{(k)} = \underline{x}$

Vediamo i <u>metodi di tipo gradiente</u>

si costruisce $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha_k \cdot \underline{d}^{(k)}$

$\alpha_k \in \mathbb{R}$; $\underline{d}^{(k)}$ ↖ direzione di discesa



per calcolare $\underline{x}^{(k+1)}$ devo capire come costruire $\underline{d}^{(k)}$, α_k

Metodo del gradiente

$A \in \mathbb{R}^{n \times n}$ sdp , costruisco $\Phi(\underline{x})$, $\Phi : \mathbb{R}^n \to \mathbb{R}$

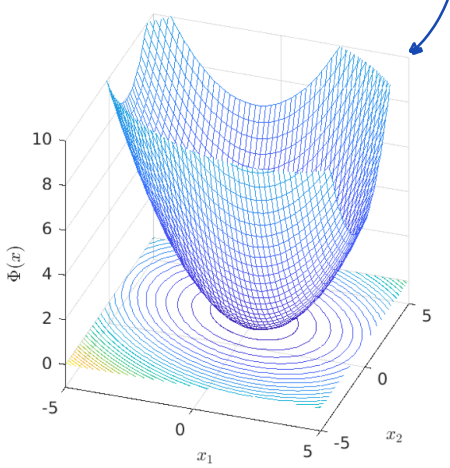
$\underline{x} \mapsto \Phi(\underline{x})$

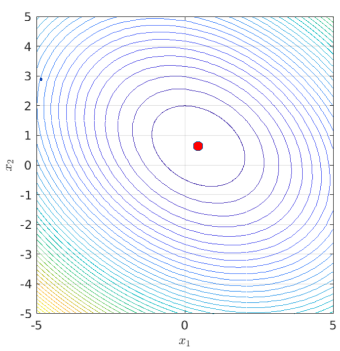
$$\Phi(\underline{x}) = \frac{1}{2} \underline{x}^T A \underline{x} - \underline{x}^T \underline{b}$$

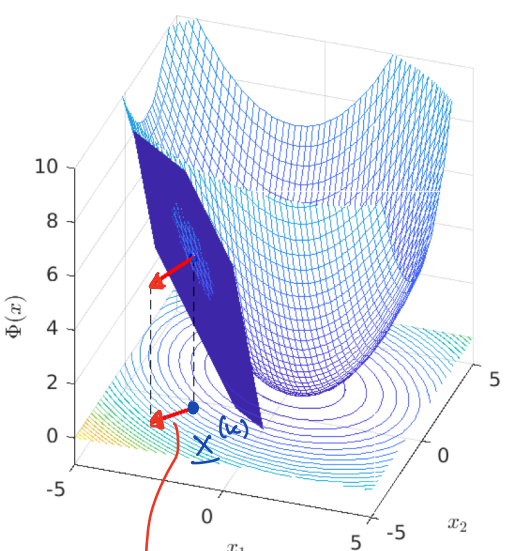
se $n=2$, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$ e

$$\underline{\Phi}(\underline{x}) = \frac{1}{2}\left(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2\right) - x_1b_1 - x_2b_2$$

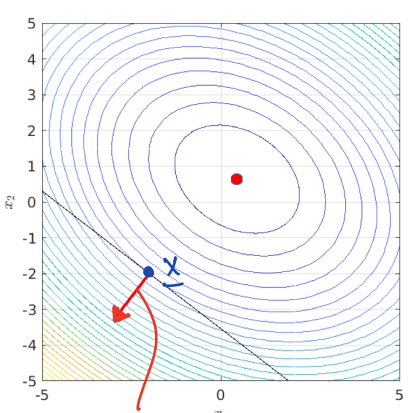
é um paraboloide







$\nabla\underline{\Phi}(\underline{x}^{(k)})$



$\nabla\underline{\Phi}(\underline{x})$

↳ gradiente di Φ valutato nel punto $x^{(k)}$

$$\nabla \Phi(x) = \begin{bmatrix} \frac{\partial \Phi}{\partial x_1}(x) \\ \frac{\partial \Phi}{\partial x_2}(x) \end{bmatrix} \in \mathbb{R}^{n\times 1}$$ è punto nella direzione di max crescita a partire da x

∇ = operatore nabla

Se x è il vertice del paraboloide $\Rightarrow \nabla \Phi(x) = 0$

Se $n=2$

$$\Phi(x) = \frac{1}{2}\left(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2\right) - x_1b_1 - x_2b_2$$

$$\nabla \Phi(x) = \begin{bmatrix} \frac{\partial \Phi}{\partial x_1}(x) \\ \frac{\partial \Phi}{\partial x_2}(x) \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 - b_1 \\ a_{12}x_1 + a_{22}x_2 - b_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{12}x_1 + a_{22}x_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$= Ax - b.$$

La relazione $\nabla \Phi(x) = Ax - b$ vale in realtà $\forall n \geq 2$

Nel punto di minimo del paraboloide

$$\nabla \Phi(x) = 0 \iff Ax - b = 0 \iff Ax = b$$

Il punto di minimo del paraboloide è la sol di $Ax = b$

Per calcolare la soluzione $\underline{x}$ di $A\underline{x} = \underline{b}$,
cerco il punto di minimo del paraboloide con la regola

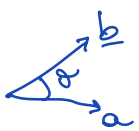
$$\begin{cases} \underline{x}^{(0)} \text{ dato} \\ \underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha_k \underline{d}^{(k)} \end{cases} \qquad k \geq 0$$

$\underline{d}^{(k)}$: direz di discesa perché devo convergere al pto di min.

Def $\underline{d}^{(k)}$ è **direz di discesa** per Φ se:

1) $\left(\nabla\Phi(\underline{x}^{(k)})\right)^T \underline{d}^{(k)} < 0$ quando $\nabla\Phi(\underline{x}^{(k)}) \neq \underline{0}$

2) $\underline{d}^{(k)} = 0$ quando $\nabla\Phi(\underline{x}^{(k)}) = \underline{0}$.



$\underline{a}\cdot\underline{b} > 0$

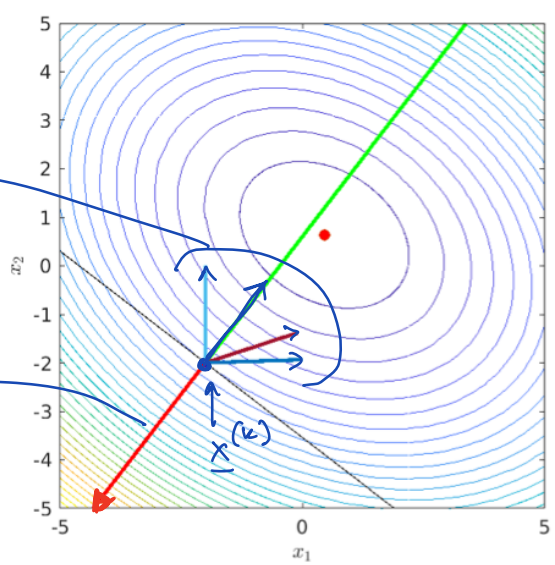


$\underline{a}\cdot\underline{b} < 0$

$\underline{a}\cdot\underline{b} = |a|\cdot|b|\cos\theta$

sono tutte direzioni di discesa

$\nabla\Phi(\underline{x}^{(k)})$



Il metodo del gradiente sceglie $\underline{d}^{(k)} = -\nabla \Phi(\underline{x}^{(k)})$.

È vero che $\underline{d}^{(k)} = -\nabla \Phi(\underline{x}^{(k)})$ è una direzione di discesa?

Sì perché $\left(\underline{d}^{(k)}\right)^T \cdot \nabla \Phi(x^{(k)}) = -\|\nabla \Phi(\underline{x}^{(k)})\|^2 < 0$

$\left(\underline{a}^T \underline{a} = \underline{a} \cdot \underline{a} = \|\underline{a}\|^2\right)$

Osserviamo che

$$\underline{d}^{(k)} = -\nabla \Phi(\underline{x}^{(k)}) = b - A\underline{x}^{(k)}$$ (vedere i conti sopra)

Inoltre, definiamo il residuo al passo k

$$\underline{r}^{(k)} = \underline{b} - A\underline{x}^{(k)}.$$

(se $\underline{x}$ = sol esatta del sistema $\Rightarrow \underline{b} - A\underline{x} = \underline{r} = \underline{0}$, cioè il residuo è nullo)

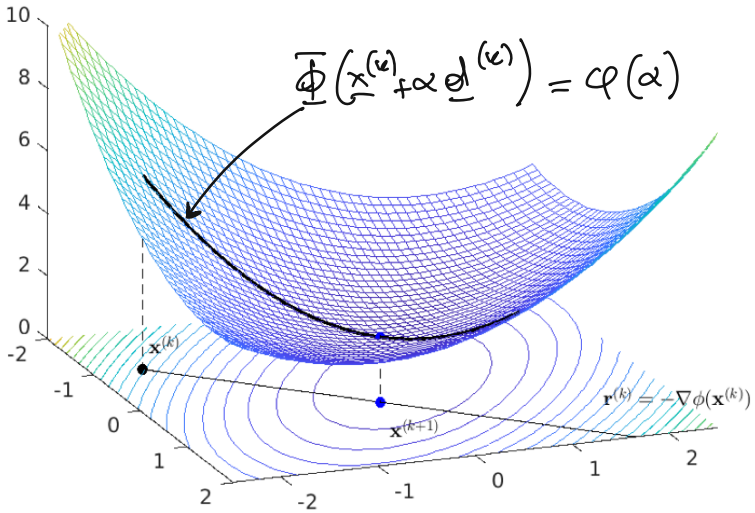
Allora la direzione di discesa del gradiente al passo k è:

$$\underline{d}^{(k)} = -\nabla \Phi(\underline{x}^{(k)}) = \underline{r}^{(k)}.$$

Ora vediamo come costruire il passo $\alpha_k \in \mathbb{R}$.

Scelgo α_k : $\Phi(\underline{x}^{(k)} + \alpha_k \underline{d}^{(k)}) = \min_{\alpha \in \mathbb{R}} \Phi(\underline{x}^{(k)} + \alpha \underline{d}^{(k)})$

ricaivo $\Phi(\underline{x}^{(k)} + \alpha \underline{d}^{(k)}) = \varphi(\alpha)$.



$\Phi(\underline{x})$ è un paraboloide $\Rightarrow$ $\varphi(\alpha)$ è una parabola

? α_k : $\Phi\left(\underline{x}^{(k)}+\alpha_k \underline{d}^{(k)}\right) = \min_{\alpha\in\mathbb{R}} \Phi\left(\underline{x}^{(k)}+\alpha \underline{d}^{(k)}\right)$

equivale a

? α_k : $\varphi(\alpha_k) = \min_{\alpha\in\mathbb{R}} \varphi(\alpha)$

Ciò equivale a dire che α_k è il punto di minimo della parabola $\varphi(\alpha)$, ovvero l'ascissa del vertice.

26/10/26

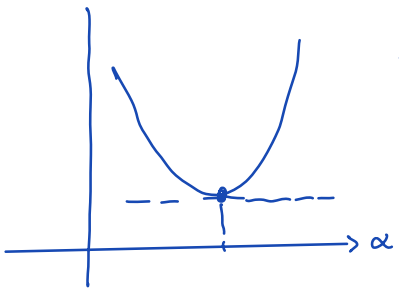
$$\varphi(\alpha) = \Phi\left(\underline{x}^{(k)}+\alpha \underline{d}^{(k)}\right) =$$

$$= \frac{1}{2}\left(\underline{x}^{(k)}+\alpha \underline{d}^{(k)}\right)^T A \left(\underline{x}^{(k)}+\alpha d^{(k)}\right) - \left(x^{(k)}+\alpha \underline{d}^{(k)}\right)^T \underline{b}$$

$$= \ldots = a\alpha^2 + b\alpha + c$$

parabola $a>0$ con concavità verso l'alto

calcolare il pto di min di $\varphi(\alpha)$ equivale a calcolare la radice di $\varphi'(\alpha) = 0$



trovo

$$\alpha_k = \frac{(\underline{d}^{(k)})^T \underline{r}^{(k)}}{(A\underline{d}^{(k)})^T \underline{d}^{(k)}} = \frac{(\underline{d}^{(k)})^T \underline{r}^{(k)}}{(\underline{d}^{(k)})^T A \underline{d}^{(k)}} \qquad \left(\underline{r}^{(k)} = \underline{b} - A\underline{x}^{(k)}\right)$$

passo di avanzamento lungo la direzione $\underline{d}^{(k)}$

posso calcolare $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha_k \underline{d}^{(k)}$

itero per $k \geq 0$ fino a quando

$$\|\underline{r}^{(k)}\| < \varepsilon \qquad \varepsilon = \text{tolleranza data}$$

Algoritmo del metodo del gradiente

Dati (input) : A, $\underline{b}$, $\underline{x}^{(0)}$, ε, k_{max}

Output : $\underline{x}$, k, $\|r^{(k)}\|$

$k = 0$

$\underline{r}^{(0)} = \underline{b} - A\underline{x}^{(0)}$, $\underline{d}^{(0)} = \underline{r}^{(0)} \left(= -\nabla \Phi(\underline{x}^{(0)})\right)$

$norma = \|\underline{r}^{(0)}\|$

while $k < k_{max}$ & $norma \geq \varepsilon$

$$\alpha_k = \frac{(\underline{d}^{(k)})^T \underline{r}^{(k)}}{(\underline{d}^{(k)})^T A \underline{d}^{(k)}} \qquad \text{(calcolo il passo)} \qquad \rightarrow n^2 \text{ op}$$

salvare in una variabile $\underline{v} = A\underline{d}^{(k)}$ e poi riutilizzo

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

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

($\underline{b} - A\underline{x}^{(k)} = \underline{r}^{(k)}$)

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

```
    k = k+1
    normar = || r^(k+1) ||
end
```

$\underline{x}$ ← l'ultima $\underline{x}^{(k+1)}$ trovata

k = n° iter effettuate

normar = norma dell'ultimo residuo

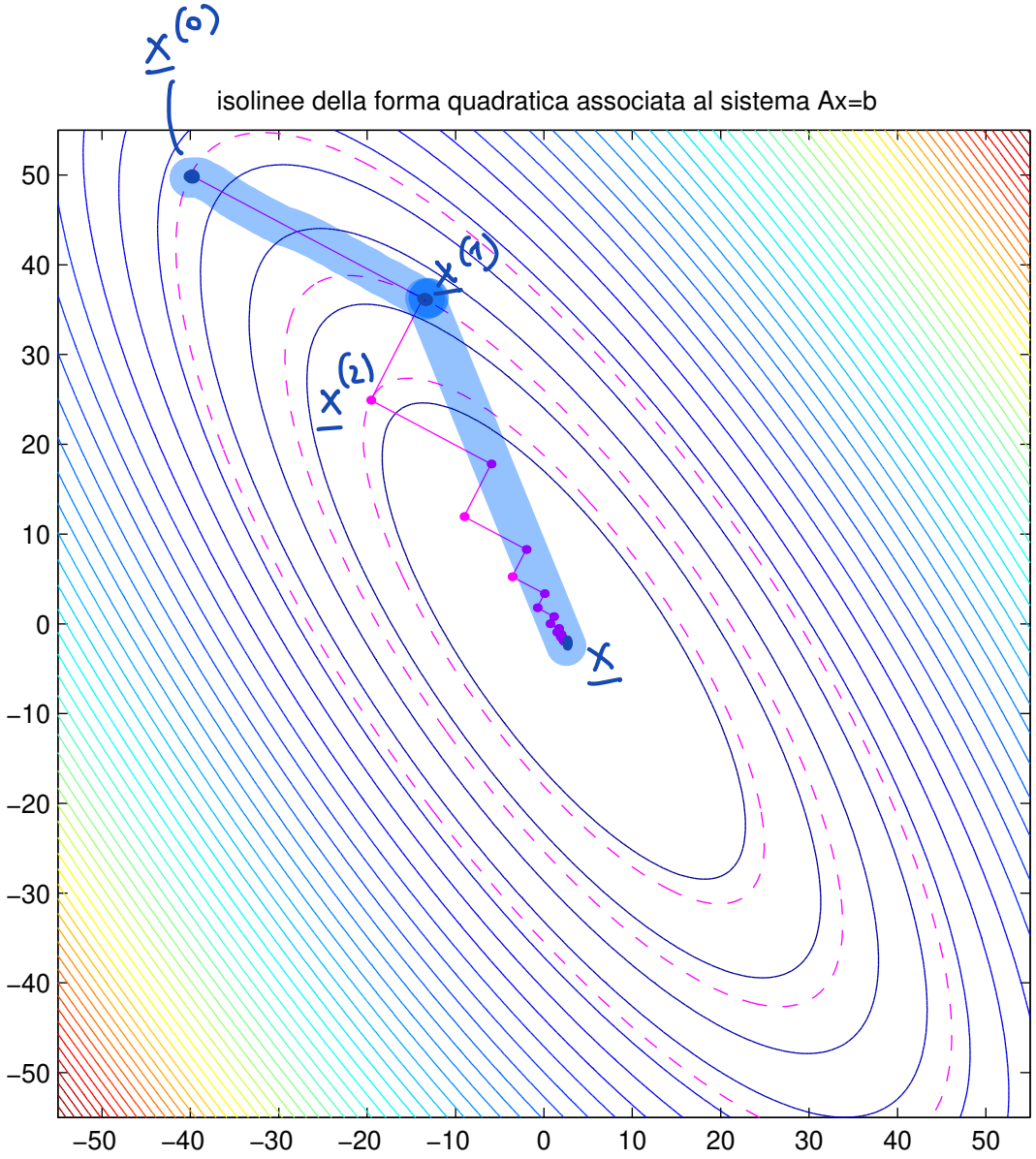
Costo computazionale ~ 1 p. matrice vettore ($\sim n^2$) per ogni iterazione.

n° iterazioni $* n^2$ ~ costo comp. totale

Limite del metodo del gradiente

Durante le iterazioni $\underline{r}^{(k+1)} \perp \underline{r}^{(k)}$ ma $r^{(k+1)}$ non è detto che sia ortogonale a tutti i residui precedenti

$\underline{d}^{(k+1)} \quad \underline{d}^{(k)}$



Con il metodo del gradiente potrebbero servire molte iterazioni per arrivare a convergenza

Metodo del GRADIENTE CONIUGATO

$A \in \mathbb{R}^{n\times n}$ simm def pos, $\underline{b} \in \mathbb{R}^{n\times 1}$? $\underline{x} \in \mathbb{R}^{n\times 1}$ sol di $A\underline{x} = \underline{b}$

cerco $\underline{x}$ come punto di minimo di $\Phi(\underline{x}) = \frac{1}{2}\underline{x}^T A \underline{x} - \underline{x}^T \underline{b}$

Scelgo le direzioni di discesa:

$\underline{d}^{(0)} = \underline{r}^{(0)}$ (come per il gradiente)

per $k \geq 0$

$$\beta_k = \frac{(\underline{r}^{(k+1)})^T A \underline{d}^{(k)}}{(\underline{d}^{(k)})^T A \underline{d}^{(k)}} \in \mathbb{R}$$

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

(oss $\beta_k = 0 \Rightarrow \underline{d}^{(k+1)} = \underline{r}^{(k+1)}$ ritroviamo il grad)

garantisce: 1) $(\underline{d}^{(k+1)})^T A \underline{d}^{(j)} = 0 \quad \forall j \leq k$

($\underline{d}^{(k+1)}$ è A-coniugata con tutte le direzioni precedenti)

siccome A è sdp, questa proprietà garantisce che

$\underline{d}^{(0)}, \underline{d}^{(1)}, \dots, \underline{d}^{(k)}, \underline{d}^{(k+1)}$ sono vettori di $\mathbb{R}^n$ linearmente indip.

$\Rightarrow \{\underline{d}^{(0)}, \underline{d}^{(1)}, \dots, \underline{d}^{(n-1)}\}$ = base di $\mathbb{R}^n$ (dove cerco la soluz)

2) $(\underline{r}^{(k+1)})^T \underline{d}^{(j)} = 0 \quad \forall j \leq k$ il residuo al passo $k+1$ è ortogonale a tutte le direz di discesa precedenti.

Allora $(\underline{r}^{(n)})^T \underline{d}^{(j)} = 0 \quad \forall j \leq n-1$

cioè $\underline{r}^{(n)} \in \mathbb{R}^{n\times 1}$ ed è ortogonale alla base $\{\underline{d}^{(0)}, \dots, \underline{d}^{(n-1)}\}$ di $\mathbb{R}^{n\times 1}$

cioè $\underline{r}^{(n)} = \underline{0}$ cioè $\underline{b} - A\underline{x}^{(n)} = 0$

cioè $\underline{x}^{(n)} \equiv \underline{x}$ è la sol esatta

Il metodo del GC converge alla sol esatta del sistema $A\underline{x} = \underline{b}$ al + in n iterazioni, se n è la dim del sistema in aritmetica esatta -

La scelta di α_k è come nel metodo del grad.
L'unica variazione nell'alg. è la def di $\underline{d}^{(k+1)}$

gradiente
$\underline{d}^{(k+1)} = \underline{r}^{(k+1)}$ ⟵ GC $\begin{cases} \beta_k = \dots \\ \underline{d}^{(k+1)} = \underline{r}^{(k+1)} - \beta_k \underline{d}^{(k)} \end{cases}$

Generalizzazioni del metodo del GC per sistemi generici

- $A \in \mathbb{R}^{n \times n}$ non singolare

- GMRES (Generalized Minimal Residual)
- BiCGStab (Bi-conjugate gradient stabilized)

$\Big\}$ $\in$ Metodi di Krylov