RICAPITOLANDO

CONDIZIONI DI OTTIMALITÀ per PROBLEMI DI OTTIMIZZAZIONE CONTINUA NON VINCOLATA

$$\min_{x \in \mathbb{R}^n} f(x) \qquad (P)$$

Se $f(x)$ è coerciva $\Rightarrow$ (P) ammette minimo locale

CONDIZIONI NECESSARIE

- $f \in C^1(\mathbb{R}^n)$, x^* è minimo locale $\Rightarrow$ $\nabla f(x^*) = 0$
- $f \in C^2(\mathbb{R}^n)$, x^* è minimo locale $\Rightarrow$ $\nabla f(x^*) = 0$, $d^T \nabla^2 f(x^*) d \geq 0 \quad \forall d \in \mathbb{R}^n$

CONDIZIONE SUFFICIENTE

- $f \in C^2(\mathbb{R}^n)$, $\nabla f(x^*) = 0$, $d^T \nabla^2 f(x^*) d > 0 \quad \forall d \in \mathbb{R}^n, d \neq 0$ $\Rightarrow$ x^* è minimo locale stretto di f su $\mathbb{R}^n$

Queste condizioni possono essere scritte anche per i problemi di massimizzazione.

Per $\max_{x \in \mathbb{R}^n} f(x) = -\min_{x \in \mathbb{R}^n} -f(x)$

- x^* massimo locale $\Rightarrow$ $\nabla f(x^*) = 0$
- x^* massimo locale $\Rightarrow$ $\nabla f(x^*) = 0$, $d^T \nabla^2 f(x^*) d \leq 0 \quad \forall d \in \mathbb{R}^n$
- $\nabla f(x^*) = 0$, $d^T \nabla^2 f(x^*) d < 0 \quad \forall d \in \mathbb{R}^n;\ d \neq 0$ $\Rightarrow$ x^* è massimo locale per f su $\mathbb{R}^n$

Def: PUNTO STAZIONARIO

Sia $f: \mathbb{R}^n \to \mathbb{R}$; $f \in C^1(\mathbb{R}^n)$

$x^* \in \mathbb{R}^n$ e' detto PUNTO STAZIONARIO

se $\nabla f(x^*) = 0$

CASO $\mathbb{R}^2$ $\min_{x \in \mathbb{R}^2} f(x)$ (P) COME RICONOSCERE SE $\nabla^2 f(x^*)$ e' DEFINITO POSITIVO / NEGATIVO

- $\det(\nabla^2 f(x^*)) > 0$ e $\frac{\partial^2 f}{\partial x_1^2}(x^*) > 0 \Rightarrow \nabla^2 f(x^*)$ e' definito positivo
- $\det(\nabla^2 f(x^*)) < 0$ e $\frac{\partial^2 f}{\partial x_i^2}(x^*) < 0 \Rightarrow \nabla^2 f(x^*)$ e' definito negativo
- $\det(\nabla^2 f(x^*)) = 0 \Rightarrow \nabla^2 f(x^*)$ e' SEMIDEFINITO
- $\det(\nabla^2 f(x^*)) < 0 \Rightarrow \nabla^2 f(x^*)$ e' INDEFINITO (Se $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ e' indefinito x^* e' un punto di sella)

Es:

$f(x_1, x_2) = x_1^2 - x_2^2$ $\quad \nabla f(x) = \begin{pmatrix} 2x_1 \\ -2x_2 \end{pmatrix}$ $\quad \nabla^2 f(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$

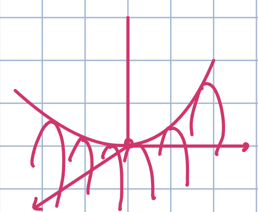
$\Downarrow$

$\det(\nabla^2 f(x)) < 0$

$x^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ e' P. STAZIONARIO

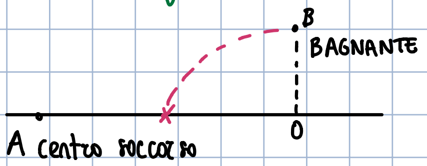
$\nabla f(x^*) = 0$, $\nabla^2 f(x^*)$ INDEFINITO

$\Rightarrow x^*$ e' PUNTO DI SELLA



Nel caso di problemi particolarmente strutturati, possiamo calcolare direttamente la soluzione del problema applicando le condizioni di ottimalità.

Esempio (pag. 19 dispense)



corre sulla sabbia $4\ m/s$
nuota $1.5\ m/s$
$|OA| = 200\ m$
$|OB| = 100\ m$
$|AX| = 200 - x$
$|BX| = \sqrt{x^2 + 100^2}$

Il tempo t impiegato dal soccorritore se inizia a nuotare dalla posizione x è dato da:

$$\frac{|AX|}{4} + \frac{|BX|}{1.5} = \frac{200 - x}{4} + \frac{\sqrt{x^2 + 100^2}}{1.5}$$

Il tempo t è funzione di x:

$$f(x) = \frac{200 - x}{4} + \frac{\sqrt{x^2 + 100^2}}{1.5}$$

$$f'(x) = -\frac{1}{4} + \frac{1}{1.5} \frac{2x}{2\sqrt{x^2 + 100^2}}$$

$f'(x) = 0 \Rightarrow x^* = 40.45$, $f''(x^*) > 0 \Rightarrow x^*$ è il minimo di $f(x)$

ALGORITMI PER PROBLEMI DI OTTIMIZZAZIONE NON VINCOLATA

$$(P) \min_{x \in \mathbb{R}^n} f(x)$$

IPOTESI

$f: \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$

$\exists\, x^0 \in \mathbb{R}^n$ c.t. $\mathcal{L}_{x^0} = \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$ è compatto

Gli algoritmi per la soluzione di (P)

consentono in generale di determinare punti stazionari di P,

ovvero l'insieme:

$$\Omega = \{x \in \mathbb{R}^n \mid \nabla f(x) = 0\}$$

Gli algoritmi che consideriamo generano successioni $\{x^k\} \subseteq \mathbb{R}^n$, t.c $\lim_{\substack{k\to+\infty \\ k\in K}} x^k = \hat{x}$ con $\hat{x}$ stazionario

che ammette sottosuccessioni che convergono a punti stazionari

Def: DIREZIONI DI DISCESA

Si definisce insieme delle DIREZIONI di DISCESA di f in un punto $\bar{x} \in \mathbb{R}^n$, l'insieme

e' l'insieme vuoto

$$D(\bar{x}) = \{ d \in \mathbb{R}^n \mid \exists\, \delta > 0 \text{ t.c } f(\bar{x}+\alpha d) < f(\bar{x}),\ \forall \alpha \in (0,\delta] \}$$

TEOREMA

Sia $f \in C^1(\mathbb{R}^n)$

Se $\bar{x} \in \mathbb{R}^n$ non e' un punto stazionario per (P) allora l'insieme

$$D_s(\bar{x}) = \{ d \in \mathbb{R}^n \mid \nabla f(\bar{x})^T d < 0 \}$$

e' non vuoto e vale $D_s(\bar{x}) \subseteq D(\bar{x})$

DIMOSTRAZIONE:

- Dimostriamo per prima cosa che $D_s(\bar{x}) \neq \emptyset$

$d = -\nabla f(\bar{x}) \neq 0$ (poiche' $\bar{x}$ non e' stazionario per ipotesi)

$\nabla f(\bar{x})^T d = -\nabla f(\bar{x})^T \nabla f(\bar{x}) = -\|\nabla f(\bar{x})\|^2 < 0$

$\Rightarrow -\nabla f(\bar{x}) \in D_s(\bar{x})$ e quindi $D_s(\bar{x}) \neq \emptyset$

- Dimostriamo che $D_s(\bar{x}) \subseteq D(\bar{x})$

Sia $d \in D_s(\bar{x})$. $f(\bar{x}+\alpha d) = f(\bar{x}) + \nabla f(\bar{x})^T(\alpha d) + R_1(\alpha\|d\|) = f(\bar{x}) + \alpha\left(\underbrace{\nabla f(\bar{x})^T d}_{<0} + \underbrace{\frac{R_1(\alpha\|d\|)}{\alpha}}_{\to 0}\right)$

poiche' $d \in D_s(\bar{x})$

per α sufficientemente piccoli $\nabla f(\bar{x})^T d + \frac{R_1(\alpha\|d\|)}{\alpha} < 0$

Quindi $\exists\, \delta > 0$ t.c

$f(\bar{x}+\alpha d) < f(\bar{x}) \quad \forall \alpha \in (0,\delta)$ ovvero $d \in D(\bar{x})$

Tra gli algoritmi di ottimizzazione per problemi non vincolati ci concentriamo e studiamo il METODO DEL GRADIENTE

La generica iterazione del metodo del gradiente è del tipo

$$x^{k+1} = x^k + \alpha^k d^k$$

con $d^k = -\nabla f(x^k)$

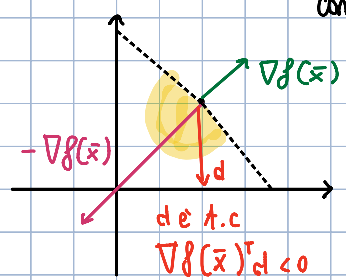
... L'antigradiente $-\nabla f(x^k)$ è la direzione di massima discesa di f in x^k (x^k non stazionario)

Sia $d = -\nabla f(\bar{x})$ con $\bar{x}$ non stazionario

Sia $d \in \mathbb{R}^n$

$$\nabla f(\bar{x}) d = \|\nabla f(\bar{x})\| \|d\| \cos\theta \geq \overbrace{- \|\nabla f(\bar{x})\| \cdot \|d\|}^{\text{si ha quando } \cos\theta = -1 \iff \theta = \pi}$$

con θ angolo compreso tra $\nabla f(\bar{x})$ e d



$$D_s(\bar{x}) = \{ d \mid \nabla f(\bar{x})^T d < 0 \}$$

$$\Downarrow$$

l'angolo tra $\nabla f(\bar{x})$ e d deve essere t.c $\cos\theta < 0$