OTTIMIZZAZIONE NON VINCOLATA

$D = \mathbb{R}^n \longrightarrow$ (P) $\min\{f(x) : x \in \mathbb{R}^n\}$ $\quad (f : \mathbb{R}^n \to \mathbb{R})$

Condizioni di ottimalità

Teo (cond. necessarie). Sia $\bar{x} \in \mathbb{R}^n$ un punto di minimo locale di (P).

a) Se f è differenziabile in $\bar{x}$, allora $\nabla f(\bar{x}) = 0$

b) se f è differenziabile 2 volte in $\bar{x}$, allora $\nabla^2 f(\bar{x})$ è semidefinita positiva

dim Sia $\varepsilon > 0$ tale che $f(\bar{x}) = \min\{f(x) : x \in B(\bar{x}, \varepsilon)\}$, e siano $d \in \mathbb{R}^n$ con $\|d\|_2 = 1$ e $t \leq \varepsilon$:

a) $$0 \leq f(\bar{x} + td) - f(\bar{x}) = t\,\nabla f(\bar{x})^T d + r(td)$$

da cui $0 \leq \nabla f(\bar{x})^T d + r(td)/t \xrightarrow[t \to 0]{} \nabla f(\bar{x})^T d$, ovvero $\nabla f(\bar{x})^T d \geq 0$

Analogamente, utilizzando la direzione $(-d)$ si ottiene $\nabla f(\bar{x})^T d \leq 0$

Quindi $\nabla f(\bar{x})^T d = 0 \quad \forall d \in \mathbb{R}^n$ (con $\|d\|_2 = 1$), da cui $\nabla f(\bar{x}) = 0$ $\left(\begin{array}{c}\text{considerare} \\ d = \nabla f(\bar{x})\end{array}\right)$

b) $$0 \leq f(\bar{x} + td) - f(\bar{x}) = t\,\nabla f(\bar{x})^T d + \frac{1}{2} t^2 d^T \nabla^2 f(\bar{x})\, d + r(td)$$
$$= \frac{1}{2} t^2 d^T \nabla^2 f(\bar{x})\, d + r(td)$$

da cui $0 \leq d^T \nabla^2 f(\bar{x}) d + r(td)/2t^2 \longrightarrow d^T \nabla^2 f(\bar{x}) d$, ovvero $d^T \nabla^2 f(\bar{x}) d \geq 0$

Quindi $d^T \nabla^2 f(\bar{x}) d \geq 0 \quad \forall d \in \mathbb{R}^n$ (con $\|d\|_2 = 1$): $\nabla^2 f(\bar{x})$ è semidefinita positiva ■

Teo (caso convesso) Sia f convessa e differenziabile su $\mathbb{R}^n$. Allora

$$\bar{x} \in \mathbb{R}^n \text{ è un punto di minimo di (P)} \iff \nabla f(\bar{x}) = 0$$

dim $\Rightarrow$) segue immediatamente dal teorema precedente

$\Leftarrow$) f convessa $\rightarrow f(y) \geq f(\bar{x}) + \nabla f(\bar{x})^T (y - \bar{x}) \quad \forall y \in \mathbb{R}^n$

$\nabla f(\bar{x}) = 0 \rightarrow f(y) \geq f(\bar{x}) \quad \forall y \in \mathbb{R}^n \rightarrow \bar{x}$ punto di minimo di (P) ∎

<u>Oss</u> Tutte le funzioni convesse differenziabili 2 volte verificano la condizione necessaria b) in ogni punto.

<u>Esempio</u> $n=2 \quad f(x_1, x_2) = (x_2 - x_1^2)(x_2 - 4x_1^2)$

$$\nabla f(x) = \begin{pmatrix} 16x_1^3 - 10x_1x_2 \\ 2x_2 - 5x_1^2 \end{pmatrix}$$

$$\nabla f(x) = 0 \iff \begin{cases} 2x_2 = 5x_1^2 \\ 16x_1^3 = 10x_1x_2 \end{cases} \iff \begin{cases} x_2 = 5/2\, x_1^2 \\ 16x_1^3 = 25x_1^3 \end{cases} \iff x_1 = x_2 = 0 \leftarrow \bar{x} = (0,0)$$

$$\nabla^2 f(x) = \begin{pmatrix} 48x_1^2 - 10x_2 & -10x_1 \\ -10x_1 & 2 \end{pmatrix} : \quad \nabla^2 f(\bar{x}) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \text{ è semidefinita positiva}$$

(f non è convessa: $\nabla^2 f(10,11) = \begin{pmatrix} -10 & 0 \\ 0 & 2 \end{pmatrix}$ non è semidefinita)

$\bar{x} = (0,0)$ non è un punto di minimo locale per f:

$P = \{x \in \mathbb{R}^2 \mid x_2 = 2x_1^2\} \qquad f(x_1, 2x_1^2) = x_1^2(-2x_1^2) = -2x_1^4 < 0$ se $x_1 \neq 0$

f è negativa su $P \setminus \{\bar{x}\}$: $\quad x_k = (1/k, 2/k^2) \quad x_k \to \bar{x}$ e $f(x_k) = -\frac{2}{k^4} < 0$

<u>Def</u> $\bar{x} \in \mathbb{R}$ si dice PUNTO STAZIONARIO per f se $\nabla f(\bar{x}) = 0$.

<u>Teo</u> (cond. sufficiente) Sia f differenziabile 2 volte in $\bar{x} \in \mathbb{R}^n$ e valga $\nabla f(\bar{x}) = 0$ ($\bar{x}$ punto stazionario). Se $\nabla^2 f(\bar{x})$ è definita positiva, allora $\bar{x}$ è un punto di minimo locale stretto di (P) ed inoltre esistono $\gamma > 0$ e $\delta > 0$ tali che

$$f(x) \geq f(\bar{x}) + \gamma \|x - \bar{x}\|_2^2 \qquad \forall x \in B(\bar{x}, \delta)$$

dim Sia $x \in \mathbb{R}^n$:

$$f(x) - f(\bar{x}) = \nabla f(\bar{x})^T (x-\bar{x}) + \tfrac{1}{2}(x-\bar{x})^T \nabla^2 f(\bar{x})(x-\bar{x}) + r(x-\bar{x}) \overset{\nabla f(\bar{x})=0}{=}$$

$$= \tfrac{1}{2}(x-\bar{x})^T \nabla^2 f(\bar{x})(x-\bar{x}) + r(x-\bar{x})$$

Sia $\lambda_{min} > 0$ il più piccolo autovalore di $\nabla^2 f(\bar{x})$ $\left(\rightarrow y^T \nabla^2 f(\bar{x}) y \geq \lambda_{min} \|y\|_2^2 \ \forall y \in \mathbb{R}^n\right)$

$$f(x) - f(\bar{x}) \geq \lambda_{min}/2 \, \|x-\bar{x}\|_2^2 + r(x-\bar{x})$$

$$[f(x) - f(\bar{x})]/\|x-\bar{x}\|_2^2 \geq \lambda_{min}/2 + r(x-\bar{x})/\|x-\bar{x}\|_2^2 \xrightarrow[x \to \bar{x}]{} \lambda_{min}/2$$

Sia $0 < \varepsilon < \lambda_{min}/2$: per la definizione di limite esiste $\delta > 0$ tale che

$$[f(x) - f(\bar{x})]/\|x-\bar{x}\|_2^2 \geq (\lambda_{min}/2 - \varepsilon) =: \gamma \quad \forall x \in B(\bar{x}, \delta)$$

da cui $f(x) \geq f(\bar{x}) + \gamma \|x-\bar{x}\|_2^2 \quad \forall x \in B(\bar{x}, \delta)$ ■

Oss Risultati analoghi per i punti di massimo:

$$\max\{f(x) : x \in \mathbb{R}^n\} = -\min\{(-f)(x) : x \in \mathbb{R}^n\} \rightarrow \text{(semi)definita positiva} \updownarrow \text{(semi)definita negativa}$$

Oss Ad eccezione di quelle costanti, le funzioni convesse [differenziabili] non ammettono punti di massimo locale/globale su $\mathbb{R}^n$:

$\bar{x}$ massimo di $f \Rightarrow \nabla f(\bar{x}) = 0$ $\qquad$ f convessa $+ \nabla f(\bar{x}) = 0 \Rightarrow \bar{x}$ minimo di f

ovvero massimo $\equiv$ minimo, cioè f è costante

Verifica alternativa

$\nabla f \equiv 0 \overset{\text{Taylor I ordine con resto}}{\Longrightarrow} f$ costante, $\quad$ sia $x \in \mathbb{R}^n$ tale che $\nabla f(x) \neq 0$:

$y(t) = t \nabla f(x) + x$

$$f \text{ convessa} \Rightarrow f(y(t)) \geq f(x) + \nabla f(x)^T (t \cdot \nabla f(x) + x - x) = f(x) + t \underbrace{\|\nabla f(x)\|_2^2}_{>0} \xrightarrow[t \to +\infty]{} +\infty$$

Oss Le funzioni quadratiche $\left(f(x) = \tfrac{1}{2} x^T Q x + b^T x + c\right)$ non convesse non ammettono punti di minimo locale in quanto $\nabla^2 f(x) \equiv Q$ non è semidefinita positiva.

Alcuni legami con l'analisi numerica

→ Caso delle funzioni quadratiche

$$f(x) = \frac{1}{2} x^T Q x + b^T x + c$$

$$\nabla f(x^*) = 0 \iff Q x^* = -b$$

ovvero

x^* punto stazionario per f $\iff$ x^* risolve il sistema lineare $Qx = -b$

→ In generale

$$\nabla f : \mathbb{R}^n \to \mathbb{R}^n \qquad \left(\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right) \to \frac{\partial f}{\partial x_i} : \mathbb{R}^n \to \mathbb{R} \right)$$

$$x \mapsto \nabla f(x)$$

∇f non lineare : $\nabla f(x) = 0$ è un sistema di equazioni non lineari

$$\downarrow$$

$$\begin{cases} \frac{\partial f}{\partial x_1}(x) = 0 \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) = 0 \end{cases}$$

→ Condizioni secondo ordine

Verificare se $\nabla^2 f(\bar{x})$ è (semi)definita positiva/negativa richiede il calcolo del più piccolo/grande autovalore di $\nabla^2 f(\bar{x})$

(nota: nel caso quadratico la matrice $\nabla^2 f(x)$ è nota una volta nota la funzione)

Calcolo ∇f, $\nabla^2 f$:

- differenze finite
- differenziazione automatica