Nota sulla convessità

L. P.

7 Marzo 2012

Una funzione $f(\boldsymbol{x})$, dove $\boldsymbol{x} = (x_1, \dots x_n)$, è detta **convessa**[1] se soddisfa la seguente condizione: dati qualunque due punti $\boldsymbol{x}_0$ e $\boldsymbol{x}_1$, e un qualunque numero α compreso fra 0 e 1, si ha

$$f((1-\alpha)\boldsymbol{x}_0 + \alpha\boldsymbol{x}_1) \leq (1-\alpha)f(\boldsymbol{x}_0) + \alpha f(\boldsymbol{x}_1) \tag{1}$$

Se la $f(\boldsymbol{x})$ è derivabile con continuità, questa proprietà è equivalente alla seguente: dato un punto qualunque $\boldsymbol{x}_0$, si ha

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_0} \cdot (\boldsymbol{x} - \boldsymbol{x}_0)\,. \tag{2}$$

Dimostrazione che (1)⇒(2) Per il teorema di Rolle, per $0 < \alpha < 1$ si ha

$$f((1-\alpha)\boldsymbol{x}_0 + \alpha\boldsymbol{x}_1) = f(\boldsymbol{x}_0) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}^*} \cdot \alpha\,(\boldsymbol{x}_1 - \boldsymbol{x}_0)\,, \tag{3}$$

dove

$$\boldsymbol{x}^* = (1-\alpha^*)\boldsymbol{x}_0 + \alpha^*\boldsymbol{x}_1, \tag{4}$$

con $0 < \alpha^* < \alpha$. Otteniamo così

$$\alpha\,[f(\boldsymbol{x}_1) - f(\boldsymbol{x}_0)] \geq \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}^*} \cdot \alpha\,(\boldsymbol{x}_1 - \boldsymbol{x}_0)\,. \tag{5}$$

Dividendo per $\alpha > 0$ e passando al limite per $\alpha \to 0$ otteniamo

$$f(\boldsymbol{x}_1) - f(\boldsymbol{x}_0) \geq \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_0} \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_0)\,, \tag{6}$$

da cui la conclusione.

Dimostrazione che (2)⇒(1) Si definisca $\bar{\boldsymbol{x}} = (1-\alpha)\boldsymbol{x}_0 + \alpha\boldsymbol{x}_1$. Si ha allora, per la (2):

$$f(\boldsymbol{x}_0) \;\geq\; f(\bar{\boldsymbol{x}}) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}} \cdot (\boldsymbol{x}_0 - \boldsymbol{x})\,; \tag{7}$$

$$f(\boldsymbol{x}_1) \;\geq\; f(\bar{\boldsymbol{x}}) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}} \cdot (\boldsymbol{x}_1 - \boldsymbol{x})\,. \tag{8}$$

Moltiplicando la prima per $(1-\alpha)$, la seconda per α e sommando, otteniamo

$$(1-\alpha)f(\boldsymbol{x}_0) + \alpha f(\boldsymbol{x}_1) \geq f(\bar{\boldsymbol{x}}) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}} \cdot [(1-\alpha)\boldsymbol{x}_0 + \alpha\boldsymbol{x}_1 - \bar{\boldsymbol{x}}]\,, \tag{9}$$

dove il secondo termine a secondo membro si annulla per la definizione di $\bar{\boldsymbol{x}}$.

[1] C'è spesso incertezza sulla terminologia. La convenzione più seguita è quella in cui una funzione con grafico a ∪ è detta **convessa**, e una con grafico a ∩ è detta **concava**.

Disuguaglianza di Jensen È da notare che il ragionamento dell'ultimo paragrafo può essere generalizzato al caso di più punti $\boldsymbol{x}_0$, ..., $\boldsymbol{x}_r$, cui sono assegnati i pesi (non negativi) p_k ($k = 0, \ldots, r$) che soddisfano la normalizzazione $\sum_k p_k = 1$. Definendo $\bar{\boldsymbol{x}} = \sum_k p_k \boldsymbol{x}_k$, si ha, per la (2):

$$f(\boldsymbol{x}_k) \geq f(\bar{\boldsymbol{x}}) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}} \cdot (\boldsymbol{x}_k - \boldsymbol{x}), \qquad \forall k. \tag{10}$$

Moltiplicando ciascuna di queste relazioni per p_k e sommando, si ottiene

$$\overline{f(\boldsymbol{x})} \geq f(\bar{\boldsymbol{x}}), \tag{11}$$

dove abbiamo definito

$$\overline{f(\boldsymbol{x})} = \sum_k p_k f(\boldsymbol{x}_k). \tag{12}$$