Nota sulla convessità

L. P.

7 Marzo 2012

Una funzione $f(\boldsymbol{x})$, dove $\boldsymbol{x} = (x_1, \dots x_n)$, è detta **convessa** 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

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

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

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

Dimostrazione che (1)⇒(2) Prendendo la derivata della (1) rispetto ad α per $\alpha = 0$ si ottiene

$$\left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_0} \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_0) \leq f(\boldsymbol{x}_1) - f(\boldsymbol{x}_0), \tag{3}$$

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):

$$f(\boldsymbol{x}_0) \quad\leq\quad f(\bar{\boldsymbol{x}}) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}} \cdot (\boldsymbol{x}_0 - \boldsymbol{x})\,; \tag{4}$$

$$f(\boldsymbol{x}_1) \quad\leq\quad f(\bar{\boldsymbol{x}}) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}} \cdot (\boldsymbol{x}_1 - \boldsymbol{x})\,. \tag{5}$$

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

$$(1-\alpha) f(\boldsymbol{x}_0) + \alpha f(\boldsymbol{x}_1) \leq f(\bar{\boldsymbol{x}}) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}} \cdot \left[(1-\alpha)\boldsymbol{x}_0 + \alpha\boldsymbol{x}_1 - \bar{\boldsymbol{x}}\right], \tag{6}$$

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

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 ($k = 0, \dots, r$) 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):

$$f(\boldsymbol{x}_k) \leq f(\bar{\boldsymbol{x}}) + \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}} \cdot (\boldsymbol{x}_k - \boldsymbol{x})\,, \qquad \forall k. \tag{7}$$

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

$$\overline{f(\boldsymbol{x})} \leq f(\bar{\boldsymbol{x}}), \tag{8}$$

dove abbiamo definito

$$\overline{f(\boldsymbol{x})} = \sum_k p_k f(\boldsymbol{x}_k). \tag{9}$$