Derivate parziali successive

Ordine di derivazione

Teorema 1 (Teorema di Schwarz). *Siano Ω un insieme aperto di $\mathbb{R}^d$ e $u : \Omega \to \mathbb{R}$ una funzione derivabile due volte su Ω. Se per una coppia di indici $1 \le i \le d$ e $1 \le j \le d$ le derivate parziali $\partial_{ij}u$ e $\partial_{ji}u$ sono continue nel punto $x_0 \in \Omega$, allora*

$$\partial_{ij}u(x_0) = \partial_{ji}u(x_0).$$

Dimostrazione: Supponiamo che $d = 2$ e $x_0 = (x, y)$. Per ogni $h, k \in \mathbb{R}$, consideriamo l'espressione

$$u(x+h, y+k) - u(x+h, y) - u(x, y+k) + u(x, y).$$

Fissati y e k, la funzione

$$f(t) = u(t, y+k) - u(t, y)$$

è derivabile in t ed esiste $h' \in (0, h)$ tale che

$$\begin{aligned} u(x+h, y+k) - u(x+h, y) - u(x, y+k) + u(x, y) &= f(x+h) - f(x) \\ &= hf'(x+h') \\ &= h\Big(\partial_x u(x+h', y+k) - \partial_x u(x+h', y)\Big). \end{aligned}$$

Utilizzando di nuovo il teorema di Lagrange, si ha

$$\partial_x u(x+h', y+k) - \partial_x u(x+h', y) = k\,\partial_y\partial_x u(x+h', y+k'),$$

per un qualche $k' \in (0, k)$. Quindi

$$\lim_{(h,k)\to(0,0)} \frac{u(x+h, y+k) - u(x+h, y) - u(x, y+k) + u(x, y)}{hk} = \partial_y\partial_x u(x, y).$$

Ora, fissiamo x e k e consideriamo la funzione

$$g(s) = u(x+h, s) - u(x, s).$$

Ora, per ogni k, esiste $k'' \in (0, k)$ tale che

$$\begin{aligned} u(x+h, y+k) - u(x+h, y) - u(x, y+k) + u(x, y) &= g(y+k) - g(y) \\ &= k\,g'(y+k'') \\ &= k\Big(\partial_y u(x+h, y+k'') - \partial_y u(x, y+k'')\Big). \end{aligned}$$

Come sopra, esiste $h'' \in (0, h)$ tale che

$$\partial_y u(x+h, y+k'') - \partial_y u(x, y+k'') = h\,\partial_x\partial_y u(x+h'', y+k''),$$

Quindi, per la continuità di $\partial_x\partial_y u$, si ha

$$\lim_{(h,k)\to(0,0)} \frac{u(x+h, y+k) - u(x+h, y) - u(x, y+k) + u(x, y)}{hk} = \partial_x\partial_y u(x, y).$$

Sviluppo di Taylor al secondo ordine

Teorema 2 (Teorema di Taylor - ordine 2)**.** *Siano* Ω *un insieme aperto di* $\mathbb{R}^d$ *e sia* $F : \Omega \to \mathbb{R}$ *una funzione in* $C^2(\Omega)$*. Allora*

$$F(X) = F(X_0) + \sum_{i=1}^{d} \partial_i F(X_0)\big(X^i - X_0^i\big) + \frac{1}{2}\sum_{i=1}^{d} \partial_{ij} F(X_0)\big(X^i - X_0^i\big)\big(X^j - X_0^j\big) + o\big(|X - X_0|^2\big).$$

Dimostrazione: Supponiamo per semplicità che $d = 2$ e la funzione F dipende dalle variabili x e y. Inoltre, senza perdità di generalità, possiamo supporre che $(x_0, y_0) = (0, 0)$. Ora, siccome le funzioni

$$\partial_x F : \Omega \to \mathbb{R} \qquad \text{e} \qquad \partial_y F : \Omega \to \mathbb{R}$$

sono di classe C^1, abbiamo

$$\partial_x F(x, y) = \partial_x F(0, 0) + (x, y) \cdot \nabla(\partial_x F)(0, 0) + \varepsilon_1(x, y),$$

$$\partial_y F(x, y) = \partial_y F(0, 0) + (x, y) \cdot \nabla(\partial_y F)(0, 0) + \varepsilon_2(x, y),$$

dove

$$\lim_{(x,y)\to(0,0)} \frac{\varepsilon_1(x, y)}{\sqrt{x^2 + y^2}} = 0 \qquad \text{e} \qquad \lim_{(x,y)\to(0,0)} \frac{\varepsilon_2(x, y)}{\sqrt{x^2 + y^2}} = 0.$$

Ora, fissiamo (x, y) in intorno di $(0, 0)$ e calcoliamo

$$\begin{aligned}
F(x, y) - F(0, 0) &= \int_0^1 (x, y) \cdot \nabla F(sx, sy)\, ds \\
&= \int_0^1 x\, \partial_x F(sx, sy)\, ds + \int_0^1 y\, \partial_y F(sx, sy)\, ds \\
&= \int_0^1 x\Big(\partial_x F(0, 0) + sx\partial_{xx} F(0, 0) + sy\partial_{yx} F(0, 0) + \varepsilon_1(sx, sy)\Big)\, ds \\
&\qquad + \int_0^1 y\Big(\partial_y F(0, 0) + sx\partial_{xy} F(0, 0) + sy\partial_{yy} F(0, 0) + \varepsilon_2(sx, sy)\Big)\, ds \\
&= (x, y) \cdot \nabla F(0, 0) + \frac{1}{2}(x, y) \begin{pmatrix} \partial_{xx} F(0, 0) & \partial_{yx} F(0, 0) \\ \partial_{xy} F(0, 0) & \partial_{yy} F(0, 0) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\
&\qquad + \int_0^1 \Big(x\varepsilon_1(sx, sy) + y\varepsilon_2(sx, sy)\Big)\, ds.
\end{aligned}$$

Infine, verificando che

$$\lim_{(x,y)\to(0,0)} \frac{1}{x^2 + y^2} \int_0^1 \Big(x\varepsilon_1(sx, sy) + y\varepsilon_2(sx, sy)\Big)\, ds = 0,$$

abbiamo la tesi. $\square$

Seconda dimostrazione del teorema: Sia $h = X - X_0$. Calcolare le derivate $g'(t)$ e $g''(t)$ della funzione

$$g(t) = F(X_0 + th).$$

Applicare il lemma seguente alla funzione g e usare Lemma 4 per dimostrare che

$$g''(\theta) - g''(0) = o(|h|^2).$$

Lemma 3. *Sia $g : [0,1] \to \mathbb{R}$ una funzione continua e derivabile su $[0,1]$, con derivata $g' : [0,1] \to \mathbb{R}$ continua e a sua volta derivabile in $[0,1]$ con derivata seconda $g'' : [0,1] \to \mathbb{R}$ continua.*
Allora, esiste $\theta \in [0,1]$ tale che

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(0) + (1-\theta)\Big(g''(\theta) - g''(0)\Big).$$

Dimostrazione: Considerare la funzione $\varphi(t) := g(t) + (1-t)g'(t)$ e usare il teorema fondamentale del calcolo integrale:

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)\,dt$$

per mostrare che

$$g(1) - g(0) - g'(0) = \int_0^1 (1-t)g''(t)\,dt.$$

Concludere usando l'uguaglianza

$$\frac{1}{2}g''(0) = \int_0^1 (1-t)g''(0)\,dt,$$

ed il teorema della media integrale. □

Lemma 4. *Siano $X = (x_1, \dots, x_n)$ e $Y = (y_1, \dots, y_n)$ due vettori in $\mathbb{R}^n$ e sia*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

una matrice $n \times n$ con coefficienti reali. Allora

$$\begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \le |X||Y| \left(\sum_{1 \le i,j \le n} a_{ij}^2 \right)^{1/2}$$

Dimostrazione: Per ogni $j = 1, \dots, n$ consideriamo i vettori

$$A_j = (a_{j1}, a_{j2} \dots, a_{jn}) \in \mathbb{R}^n.$$

Allora, per la disuguaglianza di Cauchy-Schwartz, abbiamo

$$\begin{aligned} \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= (y_1, y_2, \dots, y_n) \cdot \big(A_1 \cdot X,\, A_2 \cdot X, \dots,\, A_n \cdot X\big) \\ &\le |Y| \left(\sum_{j=1}^N (A_j \cdot X)^2 \right)^{1/2} \\ &\le |Y| \left(\sum_{j=1}^N |A_j|^2 |X|^2 \right)^{1/2} \le |Y|\,|X| \left(\sum_{j=1}^N |A_j|^2 \right)^{1/2}. \end{aligned}$$

il che conclude la dimostrazione del lemma. □

Condizioni necessarie e condizioni sufficienti per l'esistenza di massimi e minimi relativi

Definizione 5 (Massimi e minimi relativi)**.** *Sia Ω un sottoinsieme di $\mathbb{R}^d$ e sia $F : \Omega \to \mathbb{R}$ una funzione.*

- *Diciamo che $X_0 \in \Omega$ è un punto di* MINIMO *relativo di F in Ω, se esiste $B_r(X_0)$ tale che*
$$F(X_0) \leq F(X) \qquad \textit{per ogni} \qquad X \in B_r(X_0) \cap \Omega\,.$$
- *Diciamo che $X_0 \in \Omega$ è un punto di* MASSIMO *relativo di F in Ω, se esiste $B_r(X_0)$ tale che*
$$F(X_0) \geq F(X) \qquad \textit{per ogni} \qquad X \in B_r(X_0) \cap \Omega\,.$$

Teorema 6 (Condizioni necessarie)**.** *Sia Ω un aperto di $\mathbb{R}^d$ e sia $F : \Omega \to \mathbb{R}$ una funzione di classe C^2 su Ω. Se F ha un minimo relativo nel punto $X_0 \in \Omega$, allora*
$$\nabla F(X_0) = 0 \qquad \textit{e} \qquad \nabla^2 F(X_0) \geq 0.$$
Se invece F ha un massimo relativo in X_0, allora
$$\nabla F(X_0) = 0 \qquad \textit{e} \qquad \nabla^2 F(X_0) \leq 0.$$

Dimostrazione: Per ogni vettore $X \in \mathbb{R}$ consideriamo la funzione $t \mapsto F(X_0 + tX)$. Se F ha un minimo relativo in X_0, allora la funzione (di una variabile reale) $t \mapsto F(X_0+tX)$ ha un minimo relativo in $t = 0$. Di conseguenza,
$$\frac{d}{dt}\Big|_{t=0} F(X_0 + tX) = 0 \qquad \text{e} \qquad \frac{d^2}{dt^2}\Big|_{t=0} F(X_0 + tX) \geq 0,$$
ovvero
$$X \cdot \nabla F(X_0) = 0 \qquad \text{e} \qquad X^t \nabla^2 F(X_0) X \geq 0.$$
Siccome il vettore X è arbitrario abbiamo che
$$X^t \nabla^2 F(X_0) X \geq 0 \qquad \text{per ogni} \qquad X \in \mathbb{R}^d.$$
Di conseguenza, la matrice hessiana $\nabla^2 F(X_0)$ è semi-definita positiva. Infine, scegliendo $X = \nabla F(X_0)$ nella condizione al primo ordine
$$X \cdot \nabla F(X_0) = 0,$$
otteniamo che
$$|\nabla F(X_0)|^2 = 0,$$
ovvero che X_0 è un punto critico per F. □

Teorema 7 (Condizioni sufficienti)**.** *Sia Ω un aperto di $\mathbb{R}^d$ e sia $F : \Omega \to \mathbb{R}$ una funzione di classe C^2 su Ω.*

- *Se il punto $X_0 \in \Omega$ è tale che*
$$\nabla F(X_0) = 0 \qquad \textit{e} \qquad \nabla^2 F(X_0) > 0,$$
allora $F :: \Omega \to \mathbb{R}$ ha un punto di minimo relativo in X_0.
- *Se invece il punto $X_0 \in \Omega$ è tale che*
$$\nabla F(X_0) = 0 \qquad \textit{e} \qquad \nabla^2 F(X_0) < 0,$$
allora $F : \Omega \to \mathbb{R}$ ha un punto di massimo relativo in X_0.

Dimostrazione: Supponiamo che

$$\nabla F(X_0) = 0 \qquad \text{e} \qquad \nabla^2 F(X_0) > 0.$$

Definiamo

$$M = \min_{\nu \in \partial B_1} \nu^t \nabla^2 F(X_0) \nu.$$

Siccome il minimo è realizzato in un vettore ν_{min}, abbiamo che $M > 0$. Inoltre, abbiamo

$$X^t \nabla^2 F(X_0) X \geq M|X|^2 \qquad \text{per ogni} \qquad X \in \mathbb{R}^d.$$

Ora, per il Teorema di Taylor, esiste una funzione $\varepsilon : \mathbb{R}^d \to \mathbb{R}$ tale che

$$F(X + X_0) = F(X_0) + \frac{1}{2} X^t \nabla^2 F(X_0) X + \varepsilon(X) \qquad \text{dove} \qquad \lim_{X \to 0} \frac{\varepsilon(X)}{|X|^2} = 0.$$

Di conseguenza,

$$F(X + X_0) - F(X_0) \geq |X|^2 \left(\frac{M}{2} + \frac{\varepsilon(X)}{|X|^2} \right) \geq 0,$$

per $|X|$ abbastanza piccolo. $\square$