Derivate parziali del II ordine, matrice Hessiana, ecc.

Def. Sia $\Omega \underset{\text{aperto}}{\subseteq} \mathbb{R}^n$, $f \in C^1(\Omega, \mathbb{R})$. Se $\forall x \in \Omega$ $\forall i, j \in \{1, \dots, n\}$

$$\exists \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i} f\right)(x),$$

diciamo che f è parzialmente derivabile due volte in Ω e chiamiamo

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x) := \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i} f\right)(x)$$

la derivata parziale seconda di f rispetto a (x_j, x_i).

Se poi

$$x \longmapsto \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

è continua su Ω $\forall i, j \in \{1, \dots, n\}$, diciamo che f è di classe C^2 su Ω:

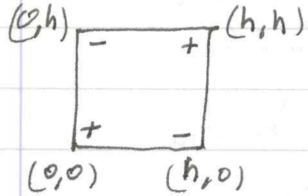
$$f \in C^2(\Omega, \mathbb{R}).$$

Teorema (di Schwarz). Se $\Omega \subseteq \mathbb{R}^n$ è aperto e $f \in C^2(\Omega, \mathbb{R})$, allora $\forall i, j \in \{1, \dots, n\}$

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

(Chiaramente, ciò è interessante se $i \neq j$!).

dim. Posso supporre $n = 2$, e basta mostrare che $\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial^2 f}{\partial y \partial x}(0,0)$ $((0,0) \in \Omega)$.



Esiste $\delta > 0$ tale che $[-\delta, \delta] \times [-\delta, \delta] \in \Omega$. Sia $h \in \mathbb{R}$ t.c. $|h| \leq \delta$ (suppongo $h > 0$) e sia

$$\Delta = f(h,h) - f(0,h) - f(h,0) + f(0,0)$$
$$= [f(h,h) - f(0,h)] - [f(h,0) - f(0,0)].$$

Siano $\phi_h(t) = f(t,h)$, $\phi_h : [-\delta, \delta] \to \mathbb{R}$

$\phi_0(t) = f(t,0)$, $\phi_0 : [-\delta, \delta] \to \mathbb{R}$

Per la formula di Taylor al II ordine con resto alla Lagrange e per la def. di $\partial_x, \partial_{xx}$:

$$f(h,h) - f(0,h) = \phi_h(h) - \phi_h(0) =$$
$$= [\phi_h(0) + \phi_h'(0)h + \tfrac{1}{2}\phi''(h')h^2] - \phi_h(0) \quad (\exists h' \in (0,h))$$
$$= \phi_h'(0)h + \frac{h^2}{2}\phi''(h') = \partial_x f(0,h)h + \frac{h^2}{2}\partial_{xx} f(h',h)$$

e $f(h,0) - f(0,0) = \dots = \partial_x f(0,0) + \frac{h^2}{2}\partial_{xx} f(h'',0)$ $(\exists h'' \in (0,h))$

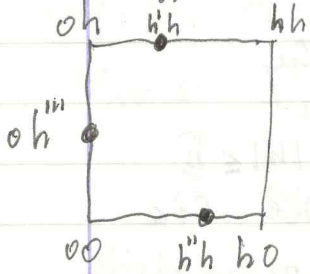
quindi

$$\Delta = [\partial_x f(0,h) - \partial_x f(0,0)]h + \frac{h^2}{2}[\partial_{xx} f(h',h) - \partial_{xx} f(h'',0)]$$
$$\overset{*}{=} \partial_y f(0,h''') \cdot h^2 + \frac{h^2}{2} \cdot [\partial_{xx} f(h',h) - \partial_{xx} f(h'',0)], \exists h''' \in (0,h)$$

$*$ Teorema di Lagrange per $\Psi(t) = \partial_x f(0,t)$

Dividendo per h^2, $\frac{\Delta}{h^2} =$

$$= \partial_{yx} f(0,h''') + \frac{1}{2}\left[\partial_{xx} f(h',h) - \partial_{xx} f(h'',0)\right]$$



Usando il fatto che $0 \le h', h'', h''' \le h$ e che $\partial_{xx} f \in C^0(\Omega, \mathbb{R})$,

$$\exists \lim_{h \to 0} \frac{\Delta}{h^2} = \partial_{yx} f(0,0) + \frac{1}{2}\left[\partial_{xx} f(0,0) - \partial_{xx} f(0,0)\right] = \partial_{yx} f(0,0).$$

Riordinando i termini,

$$\Delta = [f(h,h) - f(h,0)] - [f(0,h) - f(0,0)]$$

e lo stesso ragionamento a figura rovesciata mi dà:

$$\exists \lim_{h \to 0} \frac{\Delta}{h^2} = \partial_{xy} f(0,0),$$

quindi

$$\partial_{yx} f(0,0) = \partial_{xy} f(0,0)$$

per unicità del limite $\blacksquare$

<u>Def.</u> Sia $f \in C^2(\Omega, \mathbb{R})$, $\Omega \subseteq \mathbb{R}^n$ aperto. La <u>matrice Hessiana</u> di f in $x \in \Omega$ è $\mathrm{Hess} f(x) \in \mathcal{M}_{n \times n}$,

$$\mathrm{Hess} f(x) = \begin{bmatrix} \partial_{x_1 x_1} f(x) & \partial_{x_1 x_2} f(x) & \cdots & \partial_{x_1 x_n} f(x) \\ \partial_{x_2 x_1} f(x) & \partial_{x_2 x_2} f(x) & \cdots & \vdots \\ \vdots & & & \vdots \\ \partial_{x_n x_1} f(x) & \cdots & \cdots & \partial_{x_n x_n} f(x) \end{bmatrix}$$

- $\mathrm{Hess} f(x) = \mathrm{Hess} f(x)^T$ è una matrice simmetrica, per il T. di Schwarz.

<u>Forme quadratiche su $\mathbb{R}^n$: I parte.</u>

<u>Def.</u> Sia $A = A^T$, $A \in \mathcal{M}_{n \times n}$. La <u>forma quadratica</u> $Q = Q_A$ associata ad A è la funzione $Q_A : \mathbb{R}^n \longrightarrow \mathbb{R}$

$$x \longmapsto Q_A(x) = x^T A x$$

- Se $A = [a_{ij}]_{i,j=1}^n$, allora

$$Q_A(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j,$$

infatti, $x^T A x = x \cdot (Ax) = \sum_{i=1}^n x_i (Ax)_i$

$$= \sum_{i=1}^n x_i \left(\sum_{j=1}^n a_{ij} x_j\right) = \sum_{i,j=1}^n a_{ij} x_i x_j$$

- $Q_A(\lambda x) = \lambda^2 Q_A(x) \quad \forall \lambda \in \mathbb{R} \ \forall x \in \mathbb{R}^n$
- $Q_{A+B}(x) = Q_A(x) + Q_B(x) \quad \forall A, B \in \mathcal{M}_{n \times n}^{simm} \ \forall x \in \mathbb{R}^n$

poiché $x^T(A+B)x = (x^T A + x^T B)x = x^T A x + x^T B x$.

- $Q_A(x+y) = Q_A(x) + 2\, x^T A y + Q_A(y)$, $\forall x,y \in \mathbb{R}^n$ $\forall A \in M^{simm}_{n\times n}$

poiché
$(x+y)^T A (x+y) = x^T A (x+y) + y^T A (x+y) =$
$= x^T A x + x^T A y + y^T A x + y^T A y$ e
$y^T A x = (y^T A x)^T = x^T A^T y^{TT} = x^T A y$ ∎

- $Q_{\lambda A}(x) = \lambda \cdot Q_A(x)$ $\forall \lambda \in \mathbb{R}$ $\forall x \in \mathbb{R}^n$ $\forall A \in M^{simm}_{n\times n}$

poiché $Q_{\lambda A}(x) = \sum (\lambda a_{ij}) x_i x_j = \lambda \cdot \sum_{ij} a_{ij} x_i x_j$ ∎

Oltre a queste proprietà algebriche, abbiamo un'informazione quantitativa:

Lemma. $A \in M^{simm}_{n\times n}$, $x \in \mathbb{R}^n$
$\Rightarrow |x^T A x| \leq \|A\| \cdot \|x\|^2$.

dim. $|x^T A x| = |x \cdot (Ax)| \leq \|x\| \cdot \|Ax\|$
$\leq \|A\| \cdot \|x\|^2$ (uso Cauchy-Schwarz nella prima disuguaglianza). ∎

Oss. Invece che $\|A\|$, la norma di Hilbert-Schmidt, avrei potuto usare $\|A\|$, la norma operatoriale di A.

- $\nabla Q_A(x) = 2x^T A$ e $Hess(Q_A) = 2A$.

Infatti, $Q_A(x) = \sum_{ij} a_{ij} x_i x_j$

$\Rightarrow \partial_{x_i} Q_A(x) = \sum_{j\neq i} a_{ij} x_j + \sum_{j \neq i} a_{ji} x_j + 2 a_{ii} x_i$
$= 2 \sum_{j=1}^n a_{ij} x_j$ (perché $a_{ij} = a_{ji}$)
$= 2 (Ax)_i$, quindi
$\nabla Q_A(x) = 2\left((Ax)_1, \dots, (Ax)_n\right) = 2(Ax)^T$
$= 2\, x^T A^T = 2 x^T A$.

Per l'Hessiana, $\partial_{x_i x_j} Q_A(x) = 2 a_{ij}$ ∎

Def. Sia $A \in M^{simm}_{n\times n}$ e sia $Q_A = Q$ la corrispondente forma quadratica.

- Q è definita positiva sse
$\forall x \in \mathbb{R}^n \setminus \{0\} \Rightarrow Q(x) > 0$
- Q è definita negativa sse
$\forall x \in \mathbb{R}^n \setminus \{0\} \Rightarrow Q(x) < 0$
- Q è semidefinita positiva sse
$\forall x \in \mathbb{R}^n \Rightarrow Q(x) \geq 0$
- Q è semidefinita negativa sse
$\forall x \in \mathbb{R}^n \Rightarrow Q(x) \leq 0$
- Q è non definita sse $\exists x_1, x_2 \in \mathbb{R}^n$:
$Q(x_1) < 0 < Q(x_2)$

~~[illegible]~~

Formula di Taylor al II ordine.

Teorema. Siano $x_0 \in \Omega \xrightarrow{f} \mathbb{R}$, $\Omega \subseteq \mathbb{R}^n$ aperto; $f \in C^2(\mathbb{R}^n, \mathbb{R})$. Allora, se $x_0 + h \in \Omega$,

$$f(x_0+h) = f(x_0) + \nabla f(x_0) h + \frac{1}{2} h^T \cdot \mathrm{Hess} f(x_0) h + E_{f,x_0}(h), \text{ con}$$

$$\lim_{\substack{h \to 0 \\ \text{in } \mathbb{R}^n}} \frac{|E_{f,x_0}(h)|}{\|h\|^2} = 0$$

(cioè, $E_{f,x_0}(h) = o(\|h\|^2)$ $h \to 0$ in $\mathbb{R}^n$).

dim. ~~Ω aperto $\Rightarrow \exists \delta > 0$~~

Suppongo $x_0 = 0$. Ω aperto $\Rightarrow \exists \delta > 0$: $[-\delta,\delta]^n$ ~~$= \underbrace{[-\delta,\delta] \times \dots \times [-\delta,\delta]}_{\cancel{n \text{ volte}}}$~~ $\subseteq \Omega$.

~~Per $j = 1, 2, \dots, n$, definisco~~ $B(0,\delta) \subseteq \Omega$.

Sia $v \in \mathbb{R}^n$, $\|v\| = 1$. Definisco

$$\phi_v : (-\delta, \delta) \to \mathbb{R},$$
$$\phi_v(t) = f(tv).$$

Ho che $\phi_v'(t) = \nabla f(tv) \cdot v = \sum_{j=1}^n \partial_j f(tv) v_j$ (derivata di composizione) e

$$\phi_v''(t) = \sum_{i,j=1}^n \partial_{ij} f(tv) v_i v_j =$$
$$= v^T \cdot \mathrm{Hess} f(tv) v$$

Poichè ϕ_v'' è continua, $\phi_v \in C^2((-\delta,\delta), \mathbb{R})$.

~~Uso~~ Applico a ϕ_v la formula di Taylor al II ordine in una variabile con resto "alla Lagrange": $f(tv) =$

$$= \phi_v(t) = \phi_v(0) + \phi_v'(0) t + \frac{1}{2} \phi_v''(\Theta t)$$
con $\Theta = \Theta(v,t) \in [0,1]$

$$= f(0) + \nabla f(0) v \cdot t + \frac{t^2}{2} v^T \mathrm{Hess} f(\Theta t v) v$$

Sia $h \in B(0,\delta)$, $h \neq 0$: $h = \|h\| \cdot \frac{h}{\|h\|} = t v$, con $t \in \mathbb{R}$, $v \in \mathbb{R}^n$, $\|v\| = 1$

Allora, $f(h) = f(tv) =$

$$= f(0) + \nabla f(0) \cancel{t} tv + \frac{1}{2} (tv)^T \mathrm{Hess} f(\Theta t v)(tv)$$
$$= f(0) + \nabla f(0) h + \frac{1}{2} h^T \mathrm{Hess} f(\Theta \cdot h) \cdot h$$
con $\Theta = \Theta(v,t) = \Theta(h)$
$$= f(0) + \nabla f(0) h + \frac{1}{2} h^T \mathrm{Hess} f(0) h + E(h),$$

dove $E(h) = \frac{1}{2} h^T \operatorname{Hess} f(\Theta h)\, h - \frac{1}{2} h^T \operatorname{Hess} f(0)\, h$

$$= \frac{1}{2} h^T \left[\operatorname{Hess} f(\Theta h) - \operatorname{Hess} f(0)\right] \cdot h$$

è il termine d'errore, che stimo:

$$|E(h)| \le \frac{\|h\|^2}{2} \cdot \|\operatorname{Hess} f(\Theta h) - \operatorname{Hess} f(0)\|$$

$$= \frac{\|h\|^2}{2} \cdot \left(\sum_{ij} \left|\partial_{ij} f(\Theta h) - \partial_{ij} f(0)\right|^2 \right)^{1/2}$$

Per ipotesi, $\|\Theta h - 0\| = |\Theta| \cdot \|h\| \le \|h\|$,
e $\partial_{ij} f$ è continua su Ω,

quindi $$\frac{|E(h)|}{\|h\|^2} \le \frac{1}{2} \cdot \left(\sum_{ij} \left|\partial_{ij} f(\Theta h) - \partial_{ij} f(0)\right|^2 \right)^{\frac{1}{2}}$$

$$\le \varepsilon, \text{ se } \|h\| \le \delta = \delta(\varepsilon) \quad (\forall \varepsilon > 0\ \exists \delta > 0).$$

Cioè, $\dfrac{|E(h)|}{\|h\|^2} \xrightarrow[h \to 0 \text{ in } \mathbb{R}^n]{} 0$, come promesso ∎

Nota. In questa dimostrazione, come spesso avviene, la difficoltà consiste nell'avere stime indipendenti dalla direzione (il resto della dimostrazione è la riduzione del problema da più a una variabile).