AM210-2013-14: Tracce delle lezioni- IV Settimana

DERIVATE PARZIALI E DIFFERENZIABILITÁ

Funzioni $C^1(O)$. Una funzione é $C^1(O)$ se le sue derivate parziali esistono e sono continue in O; $f = (f_1, \ldots, f_m)$ é di classe $C^1(O)$ sse lo sono le f_i.

Proposizione 1. Sia $f \in C^1(O, \mathbf{R}^m)$. Allora f **é differenziabile**.

Prova. Basta mostrarlo per funzioni scalari. Per semplicitá, lo dimostriamo per funzioni di due variabili. Da $f_x,\ f_y$ continue segue

$$\forall \epsilon > 0\ \exists \delta = \delta_\epsilon > 0 : \ |\frac{\partial f}{\partial x}(w) - \frac{\partial f}{\partial x}(v)| + |\frac{\partial f}{\partial y}(w) - \frac{\partial f}{\partial y}(v)| \ \leq\ \epsilon \quad \forall\, w, v \in B_\delta(u) \subset O$$

Applicando il Teorema Fondamentale del Calcolo a $\frac{d}{d\tau} f(\tau, y+t) \ = \ \frac{\partial f}{\partial x}(\tau, y+t)$ e a $\frac{d}{d\tau} f(x, \tau) \ = \ \frac{\partial f}{\partial y}(x, \tau)$ otteniamo

$$\left| f(x+s, y+t)\ -\ f(x,y)\ -\ \left[\frac{\partial f}{\partial x}(x,y)s\ +\ \frac{\partial f}{\partial y}(x,y)t \right] \right| =$$

$$\left| f(x+s, y+t) - f(x, y+t) - \frac{\partial f}{\partial x}(x,y)s + f(x, y+t) - f(x,y) - \frac{\partial f}{\partial y}(x,y)t \right| \leq$$

$$\left| \int_x^{x+s} \left[\frac{\partial f}{\partial x}(\tau, y+t) - \frac{\partial f}{\partial x}(x,y) \right] d\tau \right| + \left| \int_y^{y+t} \left[\frac{\partial f}{\partial y}(x,\tau) - \frac{\partial f}{\partial y}(x,y) \right] d\tau \right|$$

$$\leq\ \epsilon\ (|s| + |t|) \qquad \text{se} \quad s^2 + t^2 \leq \delta_\epsilon^2$$

ESEMPI E CONTROESEMPI

(i) $\quad f(x,y) \ = \ \frac{xy}{x^2+y^2}$, se $x^2 + y^2 \neq 0, \quad f(0,0) = 0$

é di classe $C^1(\mathbf{R}^2 \setminus (0,0))$ ed ha anche derivate parziali, nulle, in $(0,0)$. Ma $\frac{\partial f}{\partial h}$ non esiste per alcun $h \neq e_i, i = 1, 2$, perché $t \to f(tx, ty)$ non é continua in $t = 0$ (a meno che non sia $xy = 0$). In particolare, f non é continua in zero:

una funzione puó avere derivate parziali in un punto senza essere continua in quel punto.

(ii) $f(x,y) = \frac{x^2y}{x^4+y^2}, \quad f(0,0) = 0.$

Siccome $f(x,0) = 0 \ \forall x$, é $f_x(x,0) = 0 \ \forall x$. Allo stesso modo, $f_y(0,y) = 0 \quad \forall y$. In particolare, f ha derivate parziali, nulle, in $(0,0)$. Poi, se $x^2 + y^2 > 0$, é $\frac{f(tx,ty)}{t} = \frac{x^2y}{t^2x^4+y^2} \to_{t\to 0} \frac{x^2}{y}$ e quindi anche nell'origine esistono le derivate in tutte le direzioni, ma non sono tutte nulle (come dovrebbero essere se la funzione fosse differenziabile). Siccome $f(x,x^2) \equiv \frac{1}{2}$, vediamo che

una funzione puó essere derivabile, in un punto, lungo tutte le direzioni senza essere continua in quel punto.

(iii) $f(x,y) = \frac{x^3y}{x^6+y^2}$, se $x^2 + y^2 \neq 0$, $f(0,0) = 0$

é di classe $C^1(\mathbf{R}^2 \setminus (0,0))$ ed ha derivate parziali, nulle, anche in $(0,0)$. Inoltre, in $(0,0)$, f **é derivabile in tutte le direzioni**: $\frac{\partial f}{\partial h}(0,0) = \lim_t \frac{f(tx,ty)}{t} = \lim_t \frac{tx^3y}{t^4x^6+y^2} = 0$. Tuttavia f **non é continua** in $(0,0)$, perché $f(x,x^3) \equiv \frac{1}{2}$. Dunque

una funzione puó avere derivate nulle lungo tutte le direzioni in un punto senza essere differenziabile in quel punto.

(iv) $f(x,y) = \frac{x^3y\sqrt{x^2+y^2}}{x^6+y^2}$ se $x^2 + y^2 \neq 0$, $f(0,0) = 0$

ha derivata nulla in tutte le direzioni, ed é anche continua, in $(0,0)$, ma non é differenziabile in $(0,0)$, perché $\frac{f(x,y)-f(0,0)}{\sqrt{x^2+y^2}} = \frac{x^3y}{x^6+y^2}$ non va a zero al tendere di $x^2 + y^2$ a zero (vale infatti $\frac{1}{2}$ lungo la cubica $y = x^3$). Dunque

una funzione continua puó avere derivate nulle in tutte le direzioni in un punto senza essere differenziabile in quel punto.

CONSIDERAZIONI GEOMETRICHE

Sia $f \in C^1(O, \mathbf{R})$, $(0,0) \in O \subset \mathbf{R}^2$, $f(0,0) = 0$. Allora, se $(x,y) \in O$,

$$f(x,y) = f_x(0,0)x + f_y(0,0)y + \circ(\sqrt{x^2 + y^2})$$

Ovvero, a meno di un errore di ordine superiore al primo, i valori di f, vicino all'origine, sono quelli della funzione lineare

$$z = df(0,0)(x,y) = f_x(0,0)x + f_y(0,0)y \qquad (*)$$

In linguaggio geometrico: il grafico di $df(0,0)$, ovvero il piano per l'origine in $\mathbf{R}^3$,

$$T_f(0,0) := \{(x, y, f_x(0,0)x + f_y(0,0)y)\}$$

é il piano tangente al grafico di f nel punto $(0,0,f(0,0))$.
Tale piano interseca i piani coordinati Oxz, Oyz, lungo le rette di equazione (in Oxz e Oyz rispettivamente) $z = f_x(0,0)x$, $z = f_y(0,0)y$, **rette tangenti** al grafico (in Oxz, Oyz) delle funzioni $z = f(x,0)$, $z = f(0,y)$, restrizioni della $z = f(x,y)$ rispettivamente ad Ox, Oy. I vettori tangenti a tali due curve, dati da $(1,0,f_x(0,0)$, $(0,1,f_y(0,0)y)$, generano lo spazio tangente $T_f(0,0) =$

$$\{(x,y,f_x(0,0)x + f_y(0,0)y)\} = \{x(1,0,f_x(0,0) + y(0,1,f_y(0,0)) : \quad x,y \in \mathbf{R}\}$$

Notiamo che il generico vettore $(x,y,f_x(0,0)x+f_y(0,0)y)$ in $T_f(0,0)$ é il vettore tangente alla curva $t \to (tx,ty,f(tx,ty))$ (giacente nel piano generato dall'asse Oz e dalla retta $(tx,ty,0)$ ed intersezione del grafico di f con tale piano). Dunque, **i vettori del piano tangente sono vettori tangenti al grafico di f.**
Il piano T_f si puó anche descrivere come il piano dei vettori ortogonali al vettore

$$(-f_x,-f_y,1) = (1,0,f_x(0,0) \wedge (0,1,f_y(0,0)y)$$

che é vettore normale al grafico.
Piú in generale, se $f \in C^1(O,\mathbf{R}), x_0 \in O \subset \mathbf{R}^n$, *l'iperpiano* in $\mathbf{R}^{n+1}$

$$T_f(x_0) := \{(x, f(x_0)+ < \nabla f(x_0), x - x_0 >) : \quad x \in \mathbf{R}^n\}$$

passante per $(x_0,f(x_0))$ e parallelo all'iperpiano per l'origine $z =< \nabla f(x_0), h >$ (grafico di $df(x_0)$) é il **piano tangente al grafico di f, $\mathcal{G}_f$, in** $(x_0,f(x_0))$. L'equazione del piano tangente é quindi

$$z = f(x_0)+ < \nabla f(x_0), x - x_0 >$$

Sezionando $\mathcal{G}_f$ e $T_f(x_0)$ con il piano $z = f(x_0)$ e proiettando tali sezioni sul piano $z = 0$, otteniamo la *'superfice' di livello* $\Sigma := \{x : f(x) = f(x_0)\}$ e la corrispondente **'retta' tangente** in x_0, data appunto da $< \nabla f(x_0), x - x_0 >= 0$, retta per x_0 e ortogonale a $\nabla f(x_0)$.
Infine, ∇f é direzione di massima pendenza del grafico di f in $(x,f(x))$:

$$\frac{d}{dt} f(x+th) =< \nabla f(x+th), h > \qquad \text{e} \qquad \sup_{\|h\|=1} < \nabla f(x), h >= \|\nabla f(x)\|$$

Se **n=k$\leq$ m**, $X = X(t_1,\ldots,t_k) = (X_1(t_1,\ldots,t_k),\ldots,X_m(t_1,\ldots,t_k))$ parametrizza una $k-$superfice in $\mathbf{R}^m$. Le curve su tale superfice, date da

$$t \to X(t,t_2,\ldots,t_k),\ldots,t \to X(t_1,\ldots,t)$$

determinano k vettori tangenti alla k-superfice, le k colonne di J_X, che indichiamo $\frac{\partial X}{\partial t_j} = (\frac{\partial X_1}{\partial t_j},\ldots,\frac{\partial X_m}{\partial t_j})$.

Il TEOREMA DEL VALOR MEDIO

Sia $f \in C^1(O, \mathbf{R})$, O aperto convesso in $\mathbf{R}^n$. Allora

$$|f(u) - f(v)| \;\leq\; ||u - v|| \sup_{t\in[0,1]} \|\nabla f(v + t(u - v))\|$$

Infatti $$|f(u) - f(v)| \;=\; |\int_0^1 [\frac{d}{dt} f(v + t(u - v))] \; dt| \;=$$

$$|\int_0^1 < \nabla f(v + t(u - v)), u - v > \; dt| \;\leq\; ||u - v|| \sup_{t\in[0,1]} \|\nabla f(v + t(u - v))\|$$

Corollario 1 Sia $f \in C^1(O)$, O aperto connesso. Allora

$$\nabla f(u) \;=\; 0 \;\forall u \in O \quad \Rightarrow \quad f \equiv \text{ cost. } \quad \text{in} \quad O$$

Prova. Il teorema del valor medio implica che f é costante sui dischi contenuti in O:

$$x \in B_r(x_0) \subset O \quad \Rightarrow \quad \|f(x) - f(x_0)\| \leq \sup_{\|x-x_0\|<r} \|\nabla f(x)\| = 0$$

In particolare, se $x_0 \in O$, $\{x \in O : \; f(x) = f(x_0)\}$ é aperto. Ma, per la continuitá di f, é aperto anche $\{x \in O : \; f(x) \neq f(x_0)\}$. Siccome O é connesso, e

$$O = \{x \in O : \; f(x) = f(x_0)\} \cup \{x \in O : \; f(x) \neq f(x_0)\} \text{ (unione disgiunta)}$$

concludiamo che uno dei due aperti, ovviamente $\{x \in O : \; f(x) \neq f(x_0)\}$, é vuoto .

Corollario 2. Sia $f \in C^1(O, \mathbf{R}^m)$. Allora f é **localmente Lipschitziana** in O:

$$\forall \; \overline{B}_r(x_0) \subset O, \exists L = L(r, x_0) : \qquad \|f(x) - f(y)\| \leq L\|x - y\| \quad \forall \, x, y \in \overline{B}_r(x_0)$$

Prova. Intanto, dal Teorema del valor medio

$$|f_i(x) - f_i(y)| \leq \sup_{z\in\overline{B}_r(x_0)} \|\nabla f_i(z)\| \quad \|x - y\| \qquad \forall \quad x, y \in \overline{B}_r(x_0)$$

Quindi, presi x, y in $\overline{B}_r(x_0)$, risulta

$$\|f(x) - f(y)\|^2 = \sum_{i=1}^{m} |f_i(x) - f_i(y)|^2 \leq \quad \|x - y\|^2 \sum_{i=1}^{m}\sum_{j=1}^{n} \left[\sup_{z\in B_r(x_0)} \left|\frac{\partial f_i}{\partial x_j}(z)\right|\right]^2$$

REGOLA DELLA CATENA
(derivazione di funzioni composte)

Siano E, F, G spazi normati. Sia $x \in E$ e siano

$$B_r(x) \xrightarrow{g} B_\rho(g(x)) \subset F, \qquad B_\rho(g(x)) \xrightarrow{f} G$$

Se g é differenziabile in x e f é differenziabile in $g(x)$, allora $f \circ g$ é differenziabile in x e

$$d(f \circ g)(x) \ = \ df(g(x)) \circ dg(x) \qquad e \qquad J_{f \circ g}(x) \ = \ J_f(g(x))\, J_g(x)$$

Prova. Sia $h \in E$, $k(h) := g(x+h) - g(x) = dg(x)h + \circ(h)$, cosicché $\|k(h)\| \leq \|dg(x)h\| + \circ(\|h\|) \ \leq \ c\|h\|$ se $\|h\| \leq \delta$. Siccome f é differenziabile in $g(x)$, si ha che

$$f(g(x+h)) = f(g(x) + k(h)) = f(g(x)) + df(g(x))k + \omega(k) =$$

$$= f(g(x)) + [df(g(x)) \circ dg(x)]h + \circ(h)$$

perché $\|\omega(k)\| \leq \epsilon\|k\| \leq c\epsilon\|h\| \quad se \quad \|h\| << 1.$

L'affermazione sulle matrici Jacobiane segue dal fatto che se $L \in L(\mathbf{R}^n, \mathbf{R}^m)$, $U \in L(\mathbf{R}^m, \mathbf{R}^p)$, la matrice rappresentativa del prodotto di composizione $U \circ L$ é il prodotto delle matrici rappresentative

$$\mathbf{R}^n \xrightarrow{L} \mathbf{R}^m \xrightarrow{U} \mathbf{R}^p \quad \Rightarrow \quad \mathcal{A}_{U \circ L} = \mathcal{A}_U\, \mathcal{A}_L$$

ESEMPIO DI DERIVAZIONE DI FUNZIONI COMPOSTE
Derivazione lungo un cammino differenziabile

Siano $\gamma : (a,b) \to \mathbf{R}^n$, $f : \mathbf{R}^n \to \mathbf{R}$ di classe C^1. Allora

$$\frac{d}{dt} f(\gamma(t)) \ = \ \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\gamma(t))\, \frac{d\gamma_j}{dt}(t) \ = < \nabla f(\gamma(t)), \dot\gamma(t) >$$

É una conseguenza della regola della catena e del fatto che $J_f(x) = \nabla f$ (matrice riga) mentre $J_\gamma = \dot\gamma$ (matrice colonna).

COMPLEMENTI

1. Matrice rappresentativa del prodotto di composizione.

Se indichiamo con $\mathcal{A}_L, \mathcal{A}_U$ e $\mathcal{A}_{U\circ L}$ le matrici rappresentative di $L \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$, di $U \in \mathcal{L}(\mathbf{R}^m, \mathbf{R}^p)$ e di $U \circ L \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^p)$, allora

$$\mathcal{A}_{U\circ L} \ = \ \mathcal{A}_U\, \mathcal{A}_L \qquad (\textit{prodotto di matrici})$$

Siano $e_j, j = 1, \ldots, n$, $\hat{e}_i, i = 1, \ldots, m$, $\check{e}_l, l = 1, \ldots, p$ basi in $\mathbf{R}^n, \mathbf{R}^m, \mathbf{R}^p$. Siano $Le_j = \sum_{i=1}^m (Le_j)_i\, \hat{e}_i$, $\quad U\hat{e}_i = \sum_{l=1}^p (U\hat{e}_i)_l\, \check{e}_l$. Allora

$$\mathcal{A}_U \ = \ (\ (U\hat{e}_i)_l)_{i=1,\ldots,m,\, l=1,\ldots,p} \qquad\qquad \mathcal{A}_L \ = \ (\ (Le_j)_i)_{j=1,\ldots,n,\, i=1,\ldots,m}$$

$$\mathcal{A}_{U\circ L} \ = (\ ((U\circ L)e_j)_l)_{j=1,\ldots,n;\, l=1,\ldots,p}$$

Ma $\qquad (U \circ L)e_j \ = \ U\,(\sum_{i=1}^m (Le_j)_i\, \hat{e}_i) \ =$

$$\sum_{i=1}^m (Le_j)_i\, U\hat{e}_i = \sum_{i=1}^m \left[(Le_j)_i\ \sum_{l=1}^p (U\hat{e}_i)_l\, \check{e}_l \right] \ = \ \sum_{l=1}^p \left[\sum_{i=1}^m (Le_j)_i\, (U\hat{e}_i)_l \right] \check{e}_l$$

e quindi

$$((U\circ L)e_j)_l = \sum_{i=1}^m (Le_j)_i\, (U\hat{e}_i)_l$$

ovvero $((U \circ L)e_j)_l$ é l'elemento di posto lj della matrice prodotto $\mathcal{A}_U\, \mathcal{A}_L$.

2. Lipschitzianitá sui compatti

Se $f \in C^1(O, \mathbf{R}^m)$ allora f é Lipschitziana sui compatti di O :

$$\forall K \subset O \text{ compatto } \ \exists L = L(K) : \qquad \|f(x) - f(y)\| \leq L\|x - y\| \qquad \forall \quad x, y \in K$$

Prova. Siccome $\quad x, y \in K, \|x - y\| \geq r \quad \Rightarrow \quad \frac{\|f(x)-f(y)\|}{\|x-y\|} \ \leq \ \frac{2}{r} \sup\limits_{z\in K} \|f(z)\|$, basta provare che

$$\exists r > 0, \quad L > 0 : \quad x', x'' \in K, \quad \|x' - x''\| \leq r \quad \Rightarrow \quad \|f(x') - f(x'')\| \leq L\|x' - x''\|$$

Dalla compattezza di K segue che

$$\exists x_j \in K,\ r_j > 0,\ j = 1, \ldots, p \quad \text{tali che} \quad \overline{B}_{r_j}(x_j) \subset O \quad \text{e} \quad K \subset \bigcup_{j=1}^p B_{\frac{r_j}{2}}(x_j).$$

Per quanto sopra,

$$\exists L_j > 0 : \quad x', x'' \in K \cap \overline{B}_{r_j}(x_j) \quad \Rightarrow \quad \|f(x') - f(x'')\| \leq L_j \|x' - x''\|$$

Sia $0 < r \leq r_j \forall j$, $L := max L_j$. Se $x', x'' \in K, \|x' - x''\| \leq \frac{r}{2}$ e, diciamo, $x' \in B_{\frac{r_j}{2}}(x_j)$ si ha quindi che

$$\|x'' - x_j\| \leq \frac{r}{2} + \frac{r_j}{2} \leq r_j \quad \Rightarrow \quad \|f(x') - f(x'')\| \leq L \|x' - x''\|$$

3. Il gradiente in coordinate polari. .

Data $f \in C^1(\mathbf{R}^2, \mathbf{R})$, scriviamo f in coordinate polari: $g(\rho, \theta) := f(\rho \cos \theta, \rho \sin \theta)$. Allora,

$$g_\rho = f_x(\rho \cos \theta, \rho \sin \theta) \cos \theta + f_y(\rho \cos \theta, \rho \sin \theta) \sin \theta$$
$$g_\theta = -f_x(\rho \cos \theta, \rho \sin \theta) \rho \sin \theta + f_y(\rho \cos \theta, \rho \sin \theta) \rho \cos \theta$$
$$f_x(\rho \cos \theta, \rho \sin \theta) = g_\rho \cos \theta - \frac{1}{\rho} g_\theta \sin \theta$$
$$f_y(\rho \cos \theta, \rho \sin \theta) = g_\rho \sin \theta + \frac{1}{\rho} g_\theta \cos \theta$$
$$|\nabla f|^2(\rho \cos \theta, \rho \sin \theta) = g_\rho^2(\rho, \theta) + \frac{1}{\rho^2} g_\theta^2(\rho, \theta)$$