AM2: Tracce delle lezioni- III Settimana

DIFFERENZIABILITÁ E MATRICE JACOBIANA

Siano O aperto in $\mathbf{R}^n$, $f : O \to \mathbf{R}^m$. Si dice che f é **differenziabile** in $x_0 \in O$ se esiste $L \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ (cioé L é trasformazione lineare di $\mathbf{R}^n$ in $\mathbf{R}^m$) tale che

$$f(x_0 + h) \;=\; f(x_0) + Lh + \circ(\|h\|) \qquad (*)$$

Tale L, *se esiste é unica.* Infatti: $\quad (L_1 - L_2)(h) = \circ(\|h\|) \quad \Rightarrow$
$t(L_1 - L_2)(h) = (L_1 - L_2)(th) = \circ(|t|) \quad \Rightarrow \quad (L_1 - L_2)(h) = 0 \quad \forall h.$

La trasformazione L in $(*)$ si chiama **differenziale** di f in x_0 e si indica $df(x_0)$.

La sua matrice rappresentativa $\quad (< Le_j, \hat{e}_i >)_{i,j}$ (nelle basi canoniche e_j, $\hat{e}_i$, $j = 1, \ldots, n$, $\quad i = 1, \ldots, m$) si chiama **matrice Jacobiana** e si indica $J_f(x_0)$.

Notiamo che, essendo L continua, f **é continua in** x_0.

n=1: Cammini differenziabili. Sia $\quad \gamma : (a, b) \to \mathbf{R}^m$, $\quad \gamma_i :=< \gamma, \hat{e}_i >$, cioé $\gamma = \sum_{i=1}^m \gamma_i \hat{e}_i$. Il 'cammino' γ risulta differenziabile in $t \in (a, b)$ se esiste un vettore $v = \sum_{i=1}^m v_i \hat{e}_i$ tale che $\quad \gamma(t + \tau) = \gamma(t) + \tau v + \circ(\tau)$, $\quad$ ovvero

$$\forall i = 1, \ldots, m \qquad \exists \quad \dot{\gamma}_i(t) := \lim_{\tau \to 0} \frac{\gamma_i(t + \tau) - \gamma_i(t)}{\tau} = v_i$$

Il vettore v, che si indica con $\dot{\gamma}$, é il **vettore tangente** al cammino γ in $\gamma(t)$. I cammini dotati di vettore tangente si chiamano differenziabili.

m=1: Derivate direzionali, derivate parziali. Sia $f : O \to \mathbf{R}$ differenziabile in $x \in O \subset \mathbf{R}^n$ aperto. Da $\frac{f(x+th)-f(x)}{t} = df(x)(h) + \circ(1) \quad \forall h \in \mathbf{R}^n$ e $|t| \le \delta_h$ segue:

$$\exists \quad \frac{\partial f}{\partial h}(x) := \frac{d}{dt} f(x+th)_{|t=0} = \lim_{t \to 0} \frac{f(x + th) - f(x)}{t} = df(x)h \qquad \forall h \in \mathbf{R}^n$$

Tale quantitá si chiama **derivata di f in x nella direzione h**. In particolare,

$$\exists \quad \frac{\partial f}{\partial x_j}(x) := \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t} = df(x)e_j$$

$\frac{\partial f}{\partial x_j}$ (che si scrive anche $f_{x_j}, \partial_j f, D_j f, \ldots$) si chiama **derivata parziale** di f fatta rispetto alla j-esima variabile. Il vettore

$$\nabla f(x) \;:=\; \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right)$$

si chiama **gradiente** di f in x, ed é il vettore che rappresenta $df(x)$:

$$df(x)h \ =< \ \nabla f(x) \ , \ h \ >$$

Se $\mathbf{m} > \mathbf{1}$, $\quad f$ differenziabile in $\quad x_0 \quad \Leftrightarrow \quad \frac{\|f(x_0+h) \ -[\ f(x_0)+df(x_0)h]\|}{\|h\|} \to_{\|h\|\to 0} 0$
$\Leftrightarrow \quad \frac{f_i(x_0+h) \ -[\ f_i(x_0)+<df(x_0)h,\hat{e}_i>]}{\|h\|} \to_{\|h\|\to 0} 0 \quad \forall i \quad \Leftrightarrow \quad f_i$ sono differenziabili con

$$\frac{\partial f_i}{\partial x_j} = df_i(x_0)e_j =< df(x_0)e_j, \hat{e}_i > \quad \text{ e quindi } \quad J_f(x_0) = \left(\frac{\partial f_i}{\partial x_j}\right)_{ij}$$

ovvero, $J_f(x_0)$ é la matrice che ha per righe i vettori $\nabla f_i(x_0)$.

Interpretazione geometrica.
Caso **n=1**. Sia dapprima γ grafico cartesiano in $\mathbf{R}^2$: $\gamma(t) = (t, f(t))$, $t \in (a,b)$; $\gamma((a,b))$, é il grafico di f e differenziabilitá di γ equivale a derivabilitá di f. I punti $\gamma(t) + \tau\dot{\gamma}(t) = (t+\tau, f(t) + \tau f'(t))$, $\quad \tau \in \mathbf{R} \quad$ sono i punti della retta tangente a γ: $\dot{\gamma}$ é in tal senso vettore tangente in $\gamma(t)$ a γ.
Piú in generale, γ é differenziabile in t significa che $\gamma(t+\tau) - [\gamma(t) + \tau\dot{\gamma}(t)] = \circ(\tau)$, cioé la curva γ é approssimata, con un errore di ordine superiore al primo, dalla retta $\mathbf{R}\dot{\gamma}(t)$ traslata in $\gamma(t)$, che é quindi la retta tangente a γ in $\gamma(t)$.

Caso **m=1**. L'esistenza della derivata direzionale $\frac{\partial f}{\partial h}(x)$ equivale alla differenziabilitá del cammino $\gamma(t) = (x+th, f(x+th))$ risultando $\dot{\gamma}(t) = (h, \frac{\partial f}{\partial h}(x))$. La curva γ é l'intersezione del grafico di f con il piano (in $\mathbf{R}^3$ se $n = 2$) passante per $(x, f(x))$ e per la retta $x + \mathbf{R}h$; il vettore $(h, \frac{\partial f}{\partial h}(x))$, applicato in $(x, f(x))$ é vettore tangente a tale retta, e quindi al grafico di f, $\mathcal{G}_f$. Al variare di h si ottengono tutti i vettori tangenti al grafico. Siccome $\frac{\partial f}{\partial h}(x) =< \ \nabla f(x) \ , \ h \ >$, l'insieme di tali vettori,

$$\{(h, < \ \nabla f(x) \ , \ h \ >) : \quad h \in \mathbf{R}^n\}$$

descrive un piano che, traslato in $(x, f(x))$, é il piano tangente a $\mathcal{G}_f$ in $(x, f(x))$.
Ció si legge direttamente in $\quad f(x) = f(x_0) + < \nabla f(x_0), x - x_0 > + \circ(\|x - x_0\|)$ $(x := x_0 + h,)$: il 'piano' $\qquad T_f(x_0) := \{(x, f(x_0) + < \nabla f(x_0), x - x_0 \ >)\}$ approssima il grafico di f, in un intorno di $(x_0, f(x_0))$, con un errore di ordine superiore al primo: $\quad T_f(x_0)$é il **piano tangente a $\mathcal{G}_f$ in** $(x_0, f(x_0))$.
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 $\quad$ *curva di livello* $\{x : \ f(x) = f(x_0)\}$ $\quad$ e la corrispondente **'retta' tangente** in x_0, data appunto da $\quad < \nabla f(x_0), x - x_0 \ >= 0$. Il gradiente di f in x_0 é quindi *ortogonale* alle curve di livello ed 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**, $f = (f_1, \dots, f_m)$ parametrizza una $k-$superfice in $\mathbf{R}^m$. Le curve su tale superfice, date da $t \to f(t, x_2, \dots, x_n), \dots, t \to f(x_1, \dots, t)$ determinano k vettori tangenti alla k-superfice, le k colonne di J_f, che indichiamo $\frac{\partial f}{\partial x_j} = (\frac{\partial f_1}{\partial x_j}, \dots, \frac{\partial f_m}{\partial x_j})$.

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

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

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

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

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

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

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

Derivate parziali e differenziabilitá: esempi

(i) $f(x,y) \ = \ \frac{xy}{x^2+y^2}$, 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 in $(0,0)$:

$$\frac{\partial f}{\partial x} = \frac{y^3 - yx^2}{(x^2+y^2)^2}, \qquad \frac{\partial f}{\partial y} = \frac{x^3 - xy^2}{(x^2+y^2)^2}, \qquad \frac{\partial f}{\partial x}(0,0) \ = \ \frac{\partial f}{\partial y}(0,0) \ = \ 0$$

Tuttavia f non ha derivata in $(0,0)$ nella direzione h , 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. Dunque,

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

Notiamo che: $\frac{\partial f}{\partial x}(0,y) = \frac{1}{y}\ \ \forall y \neq 0$ e quindi $\frac{\partial f}{\partial x}$ non é continua in $(0,0)$.

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

É $\frac{f(tx,ty)}{t} = \frac{x^2 y}{t^2x^4+y^2} \to_{t\to 0} \frac{x^2}{y}$ e quindi $\frac{\partial f}{\partial h}(0,0) = \frac{h_1^2}{h_2}\quad \forall\, h=(h_1,h_2)\in \mathbf{R}^2$. 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) $f(x,y) = \frac{x^3 y}{x^6+y^2}$, 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, in un punto, derivate nulle lungo tutte le direzioni senza essere differenziabile in quel punto.

(iv) $f(x,y) = \frac{\sqrt{x^2+y^2}\,x^4y}{x^6+y^2}$ 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, in un punto, derivate nulle in tutte le direzioni senza essere differenziabile in quel punto.

REGOLA DELLA CATENA

Siano $f: O \to \mathbf{R}^m$ differenziabile in $x \in O \subset \mathbf{R}^n$, $g: f(O) \to \mathbf{R}^p$ differenziabile in $f(x)$. Allora $g\circ f$ é differenziabile in x e

$$d(g\circ f)(x) = dg(f(x)) \circ df(x)$$

$$J_{g\circ f}(x) = J_g(f(x))\, J_f(x) \qquad \text{(prodotto righe per colonne)}$$

Prova. $g\left(f(x+h)\right) = g\left(f(x) + df(x)h + \circ(\|h\|)\right) =$

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

ove $\quad \|\omega(k)\| \leq \epsilon\|k\| \quad$ se $\quad \|k\| \leq \delta_\epsilon. \qquad$ Dunque

$$g(f(x+h)) \;=\; g(f(x)) \;+\; dg(f(x)) \circ df(x)h + \;\circ(\|h\|) + \omega(df(x)h + \circ(\|h\|))$$

Ma siccome $\|df(x)h + \circ(\|h\|)\| \;\leq\; \delta_\epsilon$ se $\|h\| \leq \delta'_\epsilon$, abbiamo che, per tali h, $\|\omega(df(x)h + \circ(\|h\|)\| \leq \epsilon\|df(x)h + \circ(\|h\|)\| \leq C\epsilon\|h\|$. Da qui la tesi.
L'affermazione sulle matrici Jacobiane segue dal fatto che la matrice rappresentativa del prodotto di matrici é il prodotto delle matrici rappresentative (vedi sotto).

NOTA. 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

$$\mathcal{A}_{U\circ L} \;=\; \mathcal{A}_U\, \mathcal{A}_L$$

(prodotto di matrici). Infatti , indicate con $e_j, j = 1, \ldots, n$, $\hat{e}_i, i = 1, \ldots, m$, $\check{e}_l, l = 1, \ldots, p$ le basi (canoniche) in $\mathbf{R}^n, \mathbf{R}^m, \mathbf{R}^p$, si ha

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

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

D'altra parte, $\quad < (U \circ L)(e_j), \check{e}_l > \;=\; < U\,(\sum_{i=1}^m < L(e_j), \hat{e}_i > \hat{e}_i,)\,\check{e}_l > =$ $= \sum_{i=1}^m < Le_j, \hat{e}_i >< U(\hat{e}_i), \check{e}_l >$ é l'elemento di posto lj della matrice prodotto

$$(< U(\hat{e}_i), \check{e}_l >)_{i=1,\ldots,m, l=1,\ldots,p}\,(< Le_j, \hat{e}_i >)_{j=1,\ldots,n, i=1,\ldots,m}$$

Un esempio: 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,

$$g_\rho = f_x(\rho\cos\theta, \rho\sin\theta)\cos\theta + f_y(\rho\cos\theta, \rho\sin\theta)\sin\theta$$

$$g_\theta = -f_x(\rho\cos\theta, \rho\sin\theta)\rho\sin\theta + f_y(\rho\cos\theta, \rho\sin\theta)\rho\cos\theta$$

$$f_x(\rho\cos\theta, \rho\sin\theta) = g_\rho\cos\theta - \frac{1}{\rho}g_\theta\sin\theta, \quad f_x(\rho\cos\theta, \rho\sin\theta) = g_\rho\sin\theta + \frac{1}{\rho}g_\theta\cos\theta$$

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

Corollario: derivazione lungo un cammino.

Siano $\quad \gamma \in C^1(\mathbf{R}, \mathbf{R}^n), \quad f \in C^1(\mathbf{R}^n, \mathbf{R})$. Allora

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