AM210 2012-13- V Settimana

DERIVAZIONE DI FUNZIONI COMPOSTE

1. 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

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

Infatti, dalla differenziabilitá di γ segue che

$$h(\tau) := \gamma(t+\tau) - \gamma(t) = \tau \dot\gamma(t) + \circ_\gamma(\tau) \quad \Rightarrow \quad \frac{h(\tau)}{\tau} \to_{\tau\to 0} \dot\gamma(t)$$

Ed allora, usando la differenziabilitá di f, troviamo che

$$\frac{f(\gamma(t+\tau)) - f(\gamma(t)}{\tau} = \frac{f(\gamma(t) + h(\tau)) - f(\gamma(t))}{\tau} = \frac{< \nabla f(\gamma(t), h(\tau) > + \circ_f(\|h\|)}{\tau}$$

$$\Rightarrow \quad \lim_{\tau \to 0} \frac{f(\gamma(t+\tau)) - f(\gamma(t)}{\tau} = < \nabla f(\gamma(t), \dot\gamma(t) >$$

perché $\frac{\circ_f(\|h\|)}{\tau} = \frac{\circ_f(\|h\|)}{\|h\|} \frac{\|h\|}{\tau} \to_{\tau \to 0} 0$.

ESEMPIO Sia $f \in C^1(\mathbf{R}^2)$ e sia $\gamma \in C^1((0,1), \mathbf{R}^2)$.

$$f(\gamma(t)) = 1 \, \forall t \in (0,1) \quad \Rightarrow \quad < \nabla f(\gamma(t)), \dot\gamma(t) >= 0 \quad \forall t \in (0,1)$$

Tale proprietá si descrive dicendo che 'se γ é contenuta su di una superfice di llivello di f allora in ogni $x = \gamma(t)$ si ha che: $\dot\gamma$ é ortogonale a $\nabla f(x)$.

2. Corollario. Se $g = (g_1(x_1, \dots, x_n), \dots, g_p(x_1, \dots, x_n))$ é funzione di classe C^1 in Ω aperto di $\mathbf{R}^n$, ed e_j é base canonica in $\mathbf{R}^n$, allora

$\gamma^j : t \to g(x + te_j)$ é cammino differenziabile e $\dot\gamma^j_l(0) = \frac{\partial g_l}{\partial x_j}(x)$, $l = 1, \dots, p$

Scriveremo $\frac{\partial g}{\partial x_j} = (\frac{\partial g_1}{\partial x_j}, \dots, \frac{\partial g_p}{\partial x_j})$. Per quanto sopra, se $f = f(y_1, \dots, y_p) \in C^1(\mathbf{R}^p, \mathbf{R})$, allora

$$\frac{\partial (f \circ g)}{\partial x_j} = \frac{d}{dt} f(g(x + te_j))_{t=0} = \sum_{l=1}^{p} \frac{\partial f}{\partial y_l}(g(x)) \frac{\partial g_l}{\partial x_j}(x) = < \nabla f(g(x)), \frac{\partial g}{\partial x_j}(x) >$$

3. Teorema.

Sia $g \in C^1(\Omega, \mathbf{R}^p)$, Ω aperto in $\mathbf{R}^n$. Sia $g(\Omega) \subset O$, O aperto in $\mathbf{R}^p$ aperto ed $f \in C^1(O, \mathbf{R}^m)$. Allora $f \circ g \in C^1(\Omega, \mathbf{R}^m)$ e

$$J_{f\circ g}(x) = J_f(g(x))\, J_g(x) \qquad \forall x \in \Omega$$

Infatti, come osservato in 2., l'elemento di posto i, j della matrice $J_{f\circ g}(x)$, dato da $\frac{\partial(f_i \circ g)}{\partial x_j}$ é uguale a $< \nabla f(g(x)), \frac{\partial g}{\partial x_j}(x) >$ che é appunto l'elemento di posto i, j della matrice prodotto $J_f(g(x))\, J_g(x)$, giacché la $i-$esima riga di $J_f(g(x))$ é $\nabla f_i(g(x))$ mentre la j-esima colonna di $J_g(x)$ é $\frac{\partial g}{\partial x_j}(x)$.

REGOLA DELLA CATENA

Vogliamo adesso dare una formulazione (e dimostrazione) intrinseca (cioé indipendente dalla scelta di una base) della formula di derivazione delle funzione composte, che, appunto in tale formulazione, prende il nome di 'regola della catena'. In tale riformulazione assumeremo, piú in generale, che le funzioni f, g siano differenziabili piuttosto che di classe C^1.

Sia $g : \Omega \to \mathbf{R}^p$, Ω aperto in $\mathbf{R}^n$. Sia $g(\Omega) \subset O$, O aperto in $\mathbf{R}^p$ ed $f : O \to \mathbf{R}^m$. Se g é differenziabile in x ed f é differenziabile in $g(x)$ allora $g \circ f$ é differenziabile in x e

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

Prova. Siccome f é differenziabile in $g(x)$ e g lo é in x, si ha che

$$f(g(x) + k) = f(g(x)) + df(g(x)k + \omega(k), \qquad g(x + h) = g(x) + dg(x)h + \circ(\|h\|)$$

ove $\|\omega(k)\| \leq \epsilon\|k\|$ se $\|k\| \leq \delta_\epsilon$. Dunque, posto $k(h) = g(x + h) - g(x)$, é

$$f\,(g(x + h)) = f(g(x) + k(h)) = f\,(g(x) + dg(x)h + \circ(\|h\|)) =$$

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

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

Ma $\|k(h)\| \;=\; \|dg(x)h + \circ(\|h\|)\| \;\leq\; \delta_\epsilon$ se $\|h\| \;\leq\; \delta'_\epsilon$, e quindi, per tali h, $\|\omega(k(h))\| \leq \epsilon\|k(h)\| = \epsilon\|dg(x)h + \circ(\|h\|)\| \leq C\epsilon\|h\|$. 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.

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. 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,

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

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

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

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

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