Claudio Saccon (*)
Ingegneria Aerospaziale
Lezioni di Analisi Matematica 2 e Complementi

Lezione 19 07/11/2023

email: claudio.sacconCHIOCCIOLAunipi.it
web: http://pagine.dm.unipi.it/csblog1/

Ricevimento su appuntamento da concordare per email

Proprietà dei diff / matrici Jacobiane . $f: A \to \mathbb{R}^M$ $A \subset \mathbb{R}^N$ aperto $x_0 \in A$

Ricordiamo che se f è $C^1(A)$ la matrice Jacobiana, è

$$J_f(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_N} \\ \vdots & & \vdots \\ \frac{\partial f_M}{\partial x_1} & \cdots & \frac{\partial f_M}{\partial x_N} \end{bmatrix}$$ è $M \times N$ (M righe N colonne)

$$\left(J_f(x_0)\right)_{ij} = \frac{\partial f_i}{\partial x_j} \qquad i = 1 \ldots M \quad j = 1 \ldots N$$

Proprietà (a) LINEARITÀ : $d(\lambda f + \mu g)(x_0) = \lambda\, df(x_0) + \mu\, dg(x_0)$

oppure $J_{\lambda f + \mu g}(x_0) = \lambda J_f(x_0) + \mu J_g(x_0)$

NO DIM.

(b) (b1) "differenziale del prodotto" : $f: A \to \mathbb{R}^M$ $g: A \to \mathbb{R}$

dunque posso fare $h(x) = g(x) f(x)$ $h: A \to \mathbb{R}^M$

È chiaro che $$\frac{\partial h_i}{\partial x_j} = \frac{\partial}{\partial x_j}\left(g f_i\right) = \frac{\partial g}{\partial x_j} f_i + g \frac{\partial f_i}{\partial x_j}$$

In termini di matrici Jacobiane :

$$J_{gf}(x_0) = J_g \otimes f(x_0) + g\, J_f(x_0)$$ dove ho usato:

$$[a_1 \cdots a_N] \otimes \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix} = \begin{bmatrix} a_1 b_1 & \cdots & a_N b_1 \\ \vdots & & \vdots \\ a_1 b_M & \cdots & a_N b_M \end{bmatrix}$$

(b2) (prodotto scalare) $f, g: A \to \mathbb{R}^M$ $f \cdot g: A \to \mathbb{R}$

$$\circledast \quad \frac{\partial}{\partial x_i} f \cdot g = \frac{\partial}{\partial x_i} \sum_{j=1}^{M} f_j g_j = \sum_{j=1}^{M} \frac{\partial f_j}{\partial x_i} g_j + \sum_{j=1}^{M} f_j \frac{\partial g_j}{\partial x_i}$$

in termini di matrici Jacobiane ho:

$$J_{f\cdot g}(x_0) = \underbrace{g^t J_f(x_0) + f^t J_g(x_0)}_{\text{Prodotti matriciali}} \qquad g = \begin{bmatrix} g_1 \\ \vdots \\ g_M \end{bmatrix} \quad f = \begin{bmatrix} f_1 \\ \vdots \\ f_M \end{bmatrix}$$

(si capisce la formula perché $f \cdot g = f^t g = g^t f$)

La formula si capisce meglio se introduciamo il <u>GRADIENTE</u>

<u>Def.</u> Se $f : A \to \underline{\mathbb{R}}$ (f C^1) chiamo gradiente di f in x_0 il vettore

$$\nabla f(x_0) = J_f(x_0)^t = \begin{bmatrix} \frac{\partial f(x_0)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x_0)}{\partial x_N} \end{bmatrix} \qquad (\nabla f(x_0) \in \mathbb{R}^N)$$

Il gradiente è quel vettore $\nabla f(x_0)$ tale che

$$f'(x_0)(\vec{v}) = \nabla f(x_0) \cdot \vec{v}$$

(Il gradiente $\nabla f(x_0)$ è quel vettore $\vec{w}$ tale che $df(x_0)\vec{v} = \vec{w} \cdot \vec{v}$ $\forall \vec{v} \in \mathbb{R}^N$)

CON LA DEF DI GRADIENTE LA FORMULA SUL PROD. SCAL. DIVENTA:

Se $h(x) = \underbrace{f(x) \cdot g(x)}_{\text{PROD. SCAL. IN } \mathbb{R}^M}$ $\quad (f, g : A \to \mathbb{R}^M \quad h : A \to \mathbb{R})$

$$\boxed{\underbrace{\nabla h(x_0)}_{\text{Vettore in } \mathbb{R}^N} \left(= \nabla (f \cdot g)(x_0) \right) = \underbrace{J_f(x_0)^t}_{N \times M} \underbrace{g(x_0)}_{\in \mathbb{R}^M} + \underbrace{J_g(x_0)^t}_{N \times M} \underbrace{f(x_0)}_{\in \mathbb{R}^M}}$$

(è una rilettura della formula (*) scritta sopra)

<u>Esempio</u> Se $f : A \to \mathbb{R}^M$ allora posso considerare

$g(x) = \|f(x)\|^2$ $\quad$ ($h(x) = \|f(x)\|$ $\quad (h = \sqrt{g}$) dopo)

Suppongo che f sia C^1. $g, h : A \to \mathbb{R}$ VOGLIO CALCOLARNE I GRADIENTI:

$$g(x) = f(x) \cdot f(x)$$ posso applicare le formule

$$\nabla g(x) = 2\, J_f(x)^t f(x)$$ ← vero per ogni x

Caso semplice $f(x) = x$ (dunque $M = N$)

in questo caso $J_f = I$ (matrice identica) ← ??

· MODO OVVIO: $f(x) = x$ significa che $f(x)_i = x_i$

$$\frac{\partial f_i}{\partial x_j} = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}$$ dunque $J_f(x) = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = I$

· OSSERVAZIONE GENERALE Se $f(x) = Lx$ dove $L : \mathbb{R}^N \to \mathbb{R}^M$ è lineare allora $df(x) = L$ $\forall x$, nel senso che

$$\forall x, \forall \vec{v} \quad df(x)(\vec{v}) = L\vec{v}$$!!

ATTENZIONE $df : A \to \{$APPLICAZIONI LINEARI DA $\mathbb{R}^N \to \mathbb{R}^M\}$

Se $M = 1$ è def. $\nabla f : A \to \mathbb{R}^N$

Continuando l'esempio $f(x) = x \Rightarrow g(x) = x \cdot x = \|x\|^2$

ho come gradiente $\nabla g(x) = 2 I \cdot x = 2x$.

Naturalmente lo posso dimostrare direttamente:

$$\frac{\partial}{\partial x_j} \|x\|^2 = \frac{\partial}{\partial x_j} \sum_{i=1}^N x_i^2 = \sum_{i=1}^N \frac{\partial}{\partial x_j} x_i^2 = \sum_{i=1}^N 2x_i \frac{\partial x_i}{\partial x_j} =$$

$$\frac{\partial x_i}{\partial x_j} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

$2x_j$ DUNQUE $\frac{\partial}{\partial x_j}\|x\|^2 = 2x_j \iff \nabla g(x) = 2\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = 2x$

Più in generale se $f(x) = Lx$ dove L è lineare, allora

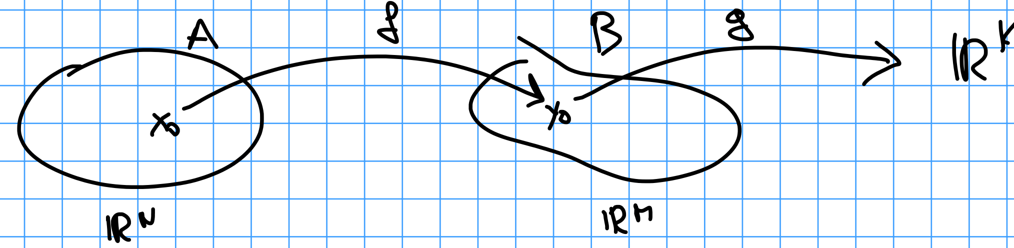
$$g(x) = \|Lx\|^2 \Rightarrow \nabla g = 2 \underbrace{J_f(x)}_{L} \cdot \underbrace{f(x)}_{Lx} = 2 L\, Lx = 2L^2 x$$

(c) "Derivata" della composizione. $f: A \to B$ $\quad g: B \to \mathbb{R}^K$

$A \subset \mathbb{R}^N \quad B \subset \mathbb{R}^M \quad x_0 \in A, \quad y_0 = f(x_0) \in B$

considero $h := g \circ f \quad \left(h(x) = g(f(x)) \right) \quad h: A \to \mathbb{R}^K$.

f e g sono C^1



ALLORA h è diff. e vale

$$\boxed{\underset{\uparrow}{dh(x_0)} = \underset{\uparrow}{dg(y_0)} \circ df(x_0)}$$

lineare da $\mathbb{R}^N \to \mathbb{R}^K$ $\quad$ lineare da $\mathbb{R}^M \to \mathbb{R}^K$ $\quad$ lineare da $\mathbb{R}^N \to \mathbb{R}^M$

In termini di matrice Jacobiana

PRODOTTO TRA MATRICI

$$J_{g\circ f}(x) = \underset{K \times N}{J_h(x)} = \underset{K \times M}{J_g(f(x))} \; \underset{M \times N}{J_f(x)}$$

Dim. Uso la def. di differenziabilità. Dato x_0 chiamo $L = df(x_0)$, $L_1 = dg(y_0)$. So che

$$(1) \quad \lim_{x \to x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} = 0_M = \begin{array}{c} \\ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \\ \\ \end{array} \begin{array}{l} \leftarrow 1 \\ \\ \\ \leftarrow M \end{array}$$

$$(2) \quad \lim_{y \to y_0} \frac{g(y) - g(y_0) - L_1(y - y_0)}{\|y - y_0\|} = 0_K$$

Dalla (2) metto $y = f(x)$ $\quad$ ($f(x) \to y_0$ se $x \to x_0$
f è diff $\Rightarrow$ f è continua)

$$\lim_{x \to x_0} \frac{g(f(x)) - g(f(x_0)) - L_1\overbrace{(f(x) - f(x_0))}^{\text{aggiungo e tolgo } L(x - x_0)}}{\|f(x) - f(x_0)\|} = 0$$

$\Updownarrow$

$$0 = \lim_{x \to x_0} \underbrace{\frac{h(x) - h(x_0) - L_1 \left(f(x) - f(x_0) - L(x - x_0) + L(x - x_0) \right)}{\|x - x_0\|} \cdot \frac{\|x - x_0\|}{\|f(x) - f(x_0)\|}}_{\star}$$

$$0 = \lim_{x \to x_0} \left(\frac{h(x) - h(x_0) - L_1 L (x - x_0)}{\|x - x_0\|} - L_1 \underbrace{\frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|}}_{\downarrow \ 0} \right) \underbrace{\frac{\|x - x_0\|}{\|f(x) - f(x_0)\|}}_{\downarrow\ ??}$$

$$\geq \frac{1}{M}$$

$$\left(\frac{\|x - x_0\|}{\|f(x) - f(x_0)\|} \right)^{-1} = \frac{\|f(x) - f(x_0) - L(x - x_0) + L(x - x_0)\|}{\|x - x_0\|} =$$

$$\left\| \underbrace{\frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|}}_{\downarrow\ 0} + \underbrace{L \left(\frac{x - x_0}{\|x - x_0\|} \right)}_{\left\| L \left(\frac{x - x_0}{\|x - x_0\|} \right) \right\| \leq \|L\|} \right\| \leftarrow \text{LIMITATA}$$

DUNQUE esiste una costante M tale che

$$\frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \leq M \implies \frac{\|x - x_0\|}{\|f(x) - f(x_0)\|} \geq \frac{1}{M}$$

RIASSUMENDO L'ESPRESSIONE $\star$ è del tipo

$$0 \leftarrow \left(\frac{h(x) - h(x_0) - L_1 L (x - x_0)}{\|x - x_0\|} + \text{infinitesimo} \right) \cdot (\text{qualcosa che è lontano da zero})$$

$$\Downarrow$$

$$0 \leftarrow \left(\frac{h(x) - h(x_0) - L_1 L_2 (x - x_0)}{\|x - x_0\|} + \text{infinitesimo} \right)$$

$$\Downarrow$$

$$\frac{h(x) - h(x_0) - L_1 L_2 (x - x_0)}{\|x - x_0\|} \to 0$$

cioè $\quad L_1 L_2 = d h(x_0)$ #

DUNQUE ho la proprietà delle matrici Jacobiane

$$\underline{J_h(x_0) = J_g(y_0)\, J_f(x_0)}$$

da questa formula ricavo le derivate parziali di $h = g \circ f$

Infatti se $i = 1 .. K$ $\quad j = 1 \dots N$ allora

$$\frac{\partial h_i}{\partial x_j}(x_0) = \left(J_h(x_0) \right)_{ij} = \left(J_g(y_0)\, J_f(x_0) \right)_{ij} = \sum_{k=1}^{M} \left(J_g(y_0) \right)_{ik} \left(J_f(x_0) \right)_{kj} =$$

$$\left(\text{se } C = A \cdot B \qquad c_{ij} = \sum a_{ik} b_{kj} \right)$$

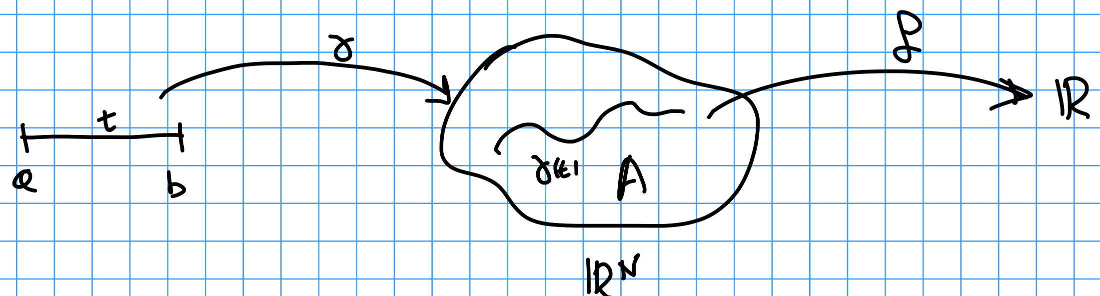
$$\sum_{k=1}^{M} \frac{\partial g_i}{\partial y_k}(y_0)\, \frac{\partial f_k}{\partial x_j}(x_0) .$$

RIASSUMENDO:

$$\boxed{\frac{\partial}{\partial x_j}(g \circ f)_i(x) = \sum_{k=1}^{M} \frac{\partial g_i}{\partial y_k}(f(x))\, \frac{\partial f_k}{\partial x_j}(x) \qquad \begin{array}{l} i = 1 \dots K \\ j = 1 \dots N \end{array}}$$

CASO SPECIALE SUPER IMPORTANTE

$f : A \to \mathbb{R}$ $\quad C^1$ $\qquad \gamma : [a\ b] \to A$ $\quad C^1$



è definita $h(t) = f(\gamma(t))$

$h : [a, b] \to \mathbb{R}$ $\qquad$ ("f lungo la curva γ")

$h'(t) = ??$ $\qquad$ Uso la formula con f al posto di g

e γ al posto di f

$$\frac{d}{dt}(f \circ \gamma)(t) = \sum_{k=1}^{N} \frac{\partial f}{\partial x_k}(\gamma(t))\, \frac{d\gamma_k}{dt}(t) = \qquad (i = j = 1)$$

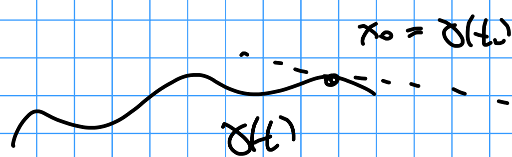
$$\sum_{h=1}^{N} \frac{\partial f}{\partial x_h}(\gamma(t))\,\gamma'_h(t) = \nabla f(\gamma(t)) \cdot \gamma'(t)$$

$$\boxed{\frac{d}{dt} f(\gamma(t)) = \nabla f(\gamma(t)) \cdot \gamma'(t)}$$

Osservo che l'ultima formula generalizza $f'(x_0)\vec{v} = \nabla f(x_0) \cdot \vec{v}$

Infatti se considero $\gamma(t) = x_0 + t\vec{v}$ allora $\gamma(0) = x_0$ e $\gamma'(t) = \vec{v}$, in particolare $\gamma'(0) = \vec{v}$. DUNQUE

$$f'(x_0)(\vec{v}) = \frac{d}{dt} f(x_0 + tv)\Big|_{t=0} = \nabla f(\underbrace{x_0}_{\gamma(0)}) \cdot \underbrace{\vec{v}}_{\gamma'(0)}$$



OSS. Dalla formula $\boxed{f'(x_0)(\vec{v}) = \vec{\nabla} f(x_0) \cdot \vec{v}}$ possiamo ricavare alcune "interpretazioni" del gradiente

DALLA FORMULA RICAVO CHE

$$-\|\vec{\nabla} f(x_0)\|\,\|\vec{v}\| \le f'(x_0)(\vec{v}) \le \|\vec{\nabla} f(x_0)\|\,\|\vec{v}\|$$

Se mi limito alle $\vec{v}$ con $\|\vec{v}\| = 1$ ho:

$$-\|\nabla f(x_0)\| \le f'(x_0)(\vec{v}) \le \|\nabla f(x_0)\| \qquad \forall \vec{v} \in S(0,1)$$

Se $\nabla f(x_0)$ $\Rightarrow$ $f'(x_0)(\vec{v}) = 0$ e basta...

Se $\nabla f(x_0) \neq 0$ posso considerare $\vec{v}_G = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$. La formula mi dà

$$f'(x_0)(\vec{v}_G) = \nabla f(x_0) \cdot \vec{v}_G = \nabla f(x_0) \cdot \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} = \|\nabla f(x_0)\|$$

DUNQUE $\quad \|\nabla f(x_0)\| = \max_{\vec{v} \in S(0,1)} f'(x_0)(\vec{v})$

e $\dfrac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$ è la direzione $\vec{v}$ su cui $f'(x_0)(\vec{v})$ è massimo

$\approx$ Il gradiente è un vettore che "punta nella direzione di massima salita" e il suo modulo è proprio il "tasso di salita massimo"

OSS. Sia $f: A \to \mathbb{R}$. Supponiamo che γ "percorra una linea di livello" $\Rightarrow f(\gamma(t)) = \text{costante}$

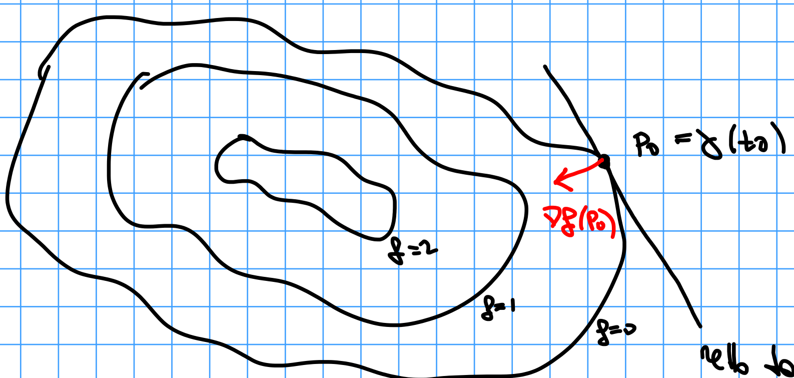
Ne segue che

$$\frac{d}{dt} f(\gamma(t)) = 0$$

$$\| $$

$$\nabla f(\gamma(t)) \cdot \gamma'(t)$$

$\leftarrow$ IL GRADIENTE È PERPENDICOLARE ALLE CURVE DI LIVELLO



retta tangente alla curva $- \; P_0 + t\vec{v}$ $\quad P_0 = \gamma(t_0)$, $\vec{v} = \gamma'(t_0)$