Derivata di funzioni composte: tips & tricks

Luciano Battaia(*)

1 gennaio 2010

In questo breve articolo propongo alcuni suggerimenti di carattere esclusivamente pratico da utilizzare nel calcolo della derivata di funzioni composte, nel caso di funzioni di $\mathbb{R}^n$ in $\mathbb{R}^m$.

Sia $f\colon A\subset\mathbb{R}^n\to\mathbb{R}^m$ una funzione di n variabili, a valori in $\mathbb{R}^m$, funzione che potremo sempre pensare costuituita da una m-upla di funzioni di $\mathbb{R}^n$ in $\mathbb{R}$, cioè di funzioni a valori reali. Se indichiamo con $\vec{x}=(x_1,\,x_2,\,\dots,\,x_n)$ i punti[1] di $\mathbb{R}^n$ e con $\vec{y}=(y_1,\,y_2,\,\dots,\,y_m)$ i punti di $\mathbb{R}^m$, potremo scrivere la funzione come :

$$f\colon\vec{x}\mapsto\vec{y}=f(\vec{x}),\quad\text{ovvero}\quad f\colon\vec{x}\mapsto\begin{cases} y_1=f_1(x_1,\,x_2,\,\dots,\,x_n)\\ y_2=f_2(x_1,\,x_2,\,\dots,\,x_n)\\ \quad\vdots\\ y_m=f_m(x_1,\,x_2,\,\dots,\,x_n)\end{cases}.\tag{1}$$

Come è abitudine, potremo anche parlare della funzione f scrivendo $\vec{y}=f(\vec{x})$, e dire che $\vec{x}$ è la variabile indipendente, mentre $\vec{y}$ è la variabile dipendente.

Nel seguito, per semplicità, ci limiteremo a considerare solo funzioni definite in tutto lo spazio: l'adattamento a funzioni definite in un sottoinsieme è elementare ed è lasciato al lettore.

Consideriamo due funzioni f e g; sappiamo che è possibile considerare la composta $g\circ f$ se l'immagine della prima funzione (quella più interna, cioè la f) è un sottoinsieme del dominio della seconda (quella più esterna, cioè la g). Quando si considerano funzioni composte è opportuno riservare il nome $\vec{x}$ per la variabile indipendente della prima funzione e il nome $\vec{y}$ per la variabile dipendente della seconda, usando un diverso nome, per esempio $\vec{z}$ per la variabile intermedia, che è la variabile dipendente della prima funzione e la variabile indipendente della seconda.

$$\vec{x}\xrightarrow{\ f\ }\vec{z}\xrightarrow{\ g\ }\vec{y}\ :\tag{2}$$

detta $h=g\circ f$ la funzione composta, avremo naturalmente

$$\vec{x}\xrightarrow{\ h\ }\vec{y},\tag{3}$$

e potremo scrivere $\vec{y}=g(\vec{z})$, $z=f(\vec{x})$, $\vec{y}=h(\vec{x})=g\big(f(\vec{x})\big)$.

Quindi la funzione composta si potrà costruire solo se

$$f\colon\mathbb{R}^n\to\mathbb{R}^p\quad\text{e}\quad g\colon\mathbb{R}^p\to\mathbb{R}^m\,,$$

*http://www.batmath.it

[1] Normalmente i punti di $\mathbb{R}^n$ o di $\mathbb{R}^m$ sono rappresentati come vettori colonna; qui li scriviamo come vettori riga per comodità tipografica.

con p valore comune delle dimensioni del codominio di f e del dominio di g. La composta sarà una funzione

$$h\colon \mathbb{R}^n \to \mathbb{R}^m\,. \tag{4}$$

Il problema di cui vogliamo occuparci è il calcolo delle derivate (in generale parziali) della funzione composta, passando attraverso le derivate delle due funzioni componenti, tenendo anche conto del fatto che ciascuna delle funzioni è in genere composta da più funzioni a valori reali. In particolare si avrà:

— $f = (f_1,\, f_2,\, \dots,\, f_p)$, e ciascuna della delle f_i avrà n derivate parziali

$$\frac{\partial f_i}{\partial x_1}\,,\ \frac{\partial f_i}{\partial x_2}\,,\ \dots\,,\ \frac{\partial f_i}{\partial x_n}\,;$$

— $g = (g_1,\, g_2,\, \dots,\, g_m)$, e ciascuna delle g_j avrà p derivate parziali

$$\frac{\partial g_j}{\partial z_1}\,,\ \frac{\partial g_j}{\partial z_2}\,,\ \dots\,,\ \frac{\partial g_j}{\partial z_p}\,;$$

— $h = (h_1,\, h_2,\, \dots,\, h_m)$, e ciascuna della delle h_j avrà n derivate parziali

$$\frac{\partial h_j}{\partial x_1}\,,\ \frac{\partial h_j}{\partial x_2}\,,\ \dots\,,\ \frac{\partial h_j}{\partial x_n}\,.$$

Le derivate della funzione composta si possono calcolare passando attraverso le derivate delle funzioni componenti nell'ipotesi di differenziabilità delle funzioni componenti stesse (in punti corrispondenti) e, praticamente, per calcolare la derivata di una delle h, per esempio h_i, rispetto a una delle sue variabili, per esempio x_j ovvero

$$\frac{\partial h_i}{\partial x_j}\,, \tag{5}$$

si deve

— scrivere in un vettore riga le derivate parziali *solo* della g_i rispetto rispetto a *tutte* le sue p variabili:

$$\begin{pmatrix} \dfrac{\partial g_i}{\partial z_1} & \dfrac{\partial g_i}{\partial z_2} & \dots & \dfrac{\partial g_i}{\partial z_p} \end{pmatrix}\,; \tag{6}$$

— scrivere in un vettore colonna le derivate parziali di *tutte* le p componenti di f, $(f_1,\, f_2,\, \dots,\, f_p)$ *solo* rispetto a x_j:

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x_j} \\ \dfrac{\partial f_2}{\partial x_j} \\ \vdots \\ \dfrac{\partial f_p}{\partial x_j} \end{pmatrix} \tag{7}$$

— fare il prodotto scalare dei due vettori, tenendo ovviamente conto che la derivata della h_i va calcolata in $\vec{x}$, le derivate di della g_i vanno calcolate in $\vec{t} = f(\vec{x})$, mentre le derivate delle componenti di f vanno calcolate in $\vec{x}$.

Si ottiene:

$$\frac{\partial h_i}{\partial x_j} = \frac{\partial g_i}{\partial z_1}\frac{\partial f_1}{\partial x_j} + \frac{\partial g_i}{\partial z_2}\frac{\partial f_2}{\partial x_j} + \cdots + \frac{\partial g_i}{\partial z_p}\frac{\partial f_p}{\partial x_j}\,. \tag{8}$$

Naturalmente non è escluso che qualcuna delle derivate sia una derivata ordinaria, anziché parziale: basterà sostituire la ∂ con la d.

Si noti che p è il numero comune di colonne del primo vettore (vettore riga) e di righe del secondo vettore (vettore colonna), come deve essere nel prodotto righe per colonne di matrici, di cui il prodotto scalare è una caso particolare.

Esempio 1. Sia $f\colon \mathbb{R} \to \mathbb{R}^3$, $\vec{z} = \big(f_1(x), f_2(x), f_3(x)\big)$ e $g\colon \mathbb{R}^3 \to \mathbb{R}$, $y = g(\vec{z}) = g(z_1, z_2, z_3)$. Allora $h = g \circ f\colon \mathbb{R} \to \mathbb{R}$, $y = h(x)$ e

$$\frac{\mathrm{d}h}{\mathrm{d}x} = \begin{pmatrix} \dfrac{\partial g}{\partial z_1} & \dfrac{\partial g}{\partial z_2} & \dfrac{\partial g}{\partial z_3} \end{pmatrix} \begin{pmatrix} \dfrac{\mathrm{d}f_1}{\mathrm{d}x} \\ \dfrac{\mathrm{d}f_2}{\mathrm{d}x} \\ \dfrac{\mathrm{d}f_3}{\mathrm{d}x} \end{pmatrix} = \frac{\partial g}{\partial z_1}\frac{\mathrm{d}f_1}{\mathrm{d}x} + \frac{\partial g}{\partial z_2}\frac{\mathrm{d}f_2}{\mathrm{d}x} + \frac{\partial g}{\partial z_3}\frac{\mathrm{d}f_3}{\mathrm{d}x}\,,$$

dove abbiamo omesso di indicare il punto dove le derivate vanno calcolate. In un caso come questo spesso la variabile x viene indicata con t e la funzione f è detta una curva dello spazio.

Esempio 2. Sia $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\vec{z} = \big(f_1(x_1, x_2), f_2(x_1, x_2)\big)$ e $g\colon \mathbb{R}^2 \to \mathbb{R}$, $y = g(\vec{z}) = g(z_1, z_2)$. Allora $h = g \circ f\colon \mathbb{R}^2 \to \mathbb{R}$, $y = h(x_1, x_2)$ e

$$\frac{\partial h}{\partial x_1} = \begin{pmatrix} \dfrac{\partial g}{\partial z_1} & \dfrac{\partial g}{\partial z_2} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} \\ \dfrac{\partial f_2}{\partial x_1} \end{pmatrix} = \frac{\partial g}{\partial z_1}\frac{\partial f_1}{\partial x_1} + \frac{\partial g}{\partial z_2}\frac{\partial f_2}{\partial x_1}\,,$$

$$\frac{\partial h}{\partial x_2} = \begin{pmatrix} \dfrac{\partial g}{\partial z_1} & \dfrac{\partial g}{\partial z_2} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f_1}{\partial x_2} \\ \dfrac{\partial f_2}{\partial x_2} \end{pmatrix} = \frac{\partial g}{\partial z_1}\frac{\partial f_1}{\partial x_2} + \frac{\partial g}{\partial z_2}\frac{\partial f_2}{\partial x_2}\,.$$

Per rendersi conto a livello puramente intuitivo del senso delle formule indicate, esaminiamo in dettaglio l'ultima. Essa si può leggere pensando che h dipende da x_2 non direttamente, ma attraverso le due f, f_1 ed f_2, che sono state "sostituite" al posto delle due variabili della g: dunque per derivare la h rispetto a x_2 dovrò derivare la g rispetto alle sue due variabili e poi le due f, f_1 ed f_2, rispetto alla variabile interessata alla derivazione, cioè rispetto alla x_2.

Un po' di pratica renderà il tutto quasi automatico. . . !