Analisi Matematica 1B - a.a. 2020–2021 - Lezione 14

Differenziabilità di funzioni composte

Damiano Foschi

(versione aggiornata il 29 aprile 2021)

Nelle precedenti lezioni abbiamo più che altro considerato la differenziabilità per funzioni scalari. Ora vogliamo estendere il concetto di differenziabilità anche per funzioni vettoriali. Vedremo anche come utilizzando la definizione di differenziabilità possiamo calcolare derivate parziali per composizioni di funzioni vettoriali.

1 Applicazioni lineari

Facciamo un piccolo ripasso sulle applicazioni lineari tra spazi $\mathbb{R}^n$ (rimandando al corso di Geometria e Algebra per maggiori dettagli).

Definizione 1.1. Siano V e W due spazi vettoriali sul campo $\mathbb{R}$. Una funzione $\boldsymbol{L}\colon V \to W$ si dice *applicazione lineare* quando valgono le seguenti proprietà:

$$\boldsymbol{L}(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{L}(\boldsymbol{x}) + \boldsymbol{L}(\boldsymbol{y}), \qquad \boldsymbol{L}(\lambda\boldsymbol{x}) = \lambda\boldsymbol{L}(\boldsymbol{x}), \qquad \forall \boldsymbol{x},\boldsymbol{y} \in V,\ \lambda \in \mathbb{R}.$$

C'è uno stretto legame tra matrici e applicazioni lineari tra spazi vettoriali di dimensione finita. Ad ogni matrice a valori reali $A = (a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ con m righe e n colonne, corrisponde un'applicazione lineare $\boldsymbol{L}_A\colon \mathbb{R}^n \to \mathbb{R}^m$ che al vettore (colonna) $\boldsymbol{x} \in \mathbb{R}^n$ associa il prodotto matriciale (riga per colonna) tra la matrice A e il vettore $\boldsymbol{x}$

$$\boldsymbol{L}_A(\boldsymbol{x}) := A\boldsymbol{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}.$$

Gli elementi del vettore $A\boldsymbol{x}$ sono dati dai prodotti scalari delle righe di A con $\boldsymbol{x}$. Se indichiamo con $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m$ i vettori formati dalle righe della matrice A abbiamo

$$\boldsymbol{v}_i := \big(a_{i1}, a_{i2}, \dots, a_{in}\big), \qquad A = \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_m \end{pmatrix}; \qquad A\boldsymbol{x} = \begin{pmatrix} \boldsymbol{v}_1 \cdot \boldsymbol{x} \\ \boldsymbol{v}_2 \cdot \boldsymbol{x} \\ \vdots \\ \boldsymbol{v}_m \cdot \boldsymbol{x} \end{pmatrix}.$$

Ogni applicazione lineare $\boldsymbol{L}\colon \mathbb{R}^n \to \mathbb{R}^m$ è della forma $\boldsymbol{L}(\boldsymbol{x}) = A\boldsymbol{x}$ per una opportuna matrice A con m righe e n colonne, le cui componenti a_{ij} si possono ottenere valutando $\boldsymbol{L}$ sui vettori delle basi canoniche,

$$a_{ij} = \widetilde{\boldsymbol{e}}_i \cdot \boldsymbol{L}(\boldsymbol{e}_j), \qquad i = 1, \dots, m,\ j = 1, \dots, n,$$

dove $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ sono i vettori della base canonica di $\mathbb{R}^n$ e $\widetilde{\boldsymbol{e}}_1, \dots, \widetilde{\boldsymbol{e}}_m$ sono i vettori della base canonica di $\mathbb{R}^m$. Il vettore $\boldsymbol{L}(\boldsymbol{e}_j)$ coincide con la j-esima colonna di A; inoltre ogni vettore $\boldsymbol{y} = \boldsymbol{L}(\boldsymbol{x})$ che è immagine tramite $\boldsymbol{L}$ di un vettore $\boldsymbol{x} = (x_1, \dots, x_n)$ è una combinazione lineare delle colonne di A,

$$\boldsymbol{L}(\boldsymbol{x}) = \boldsymbol{L}\Big(\sum_{j=1}^{n} x_j \boldsymbol{e}_j\Big) = \sum_{j=1}^{n} x_i \boldsymbol{L}(\boldsymbol{e}_j).$$

In particolare le applicazioni lineari $L\colon \mathbb{R}^n \to \mathbb{R}$ sono rappresentate da matrici con una sola riga, e quindi possono essere scritte nella forma $L(\boldsymbol{x}) = \boldsymbol{v} \cdot \boldsymbol{x}$ dove $\boldsymbol{v} = (v_1, \dots, v_n)$ è il vettore riga della matrice che rappresenta L, con $v_j = L(\boldsymbol{e}_j)$.

Esempio 1.2. La funzione $\boldsymbol{L}\colon \mathbb{R}^3 \to \mathbb{R}^2$ definita da $\boldsymbol{L}(x, y, z) = (x - 2y + 3z, 5y - 4x)$ è lineare ed è rappresentata dalla matrice $A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & 0 \end{pmatrix}$.

La funzione $L\colon \mathbb{R}^3 \to \mathbb{R}$ definita da $L(x, y, z) = 2x - 3z$ è lineare ed è rappresentata dal vettore $\boldsymbol{v} = (2, 0, -3)$.

Esercizio 1.3. Sia $\boldsymbol{L}\colon \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione lineare.

1. Spiega perché se $m > n$ l'applicazione $\boldsymbol{L}$ non può essere suriettiva.

2. Spiega perché se $m < n$ l'applicazione $\boldsymbol{L}$ non può essere iniettiva.

Una matrice A con m righe e n colonne è formata da mn elementi, e in un certo senso possiamo immaginarla come un vettore di $\mathbb{R}^{mn}$ (solo che le sue componenti invece di essere tutte il fila sono disposte su righe e colonne); possiamo definire dunque la norma euclidea di una matrice esattamente come abbiamo fatto per i vettori ponendo

$$\|A\| := \sqrt{\sum_{\substack{i=1,\dots,m \\ j=1,\dots,n}} a_{ij}^2} = \sqrt{\sum_{i=1}^{m} \|\boldsymbol{v}_i\|^2},$$

dove $\boldsymbol{v}_i$ è il vettore formato dagli elementi della i-esima riga di A. Se $\boldsymbol{L}$ è l'applicazione lineare corrispondente alla matrice A, applicando la disuguaglianza di Cauchy-Schwarz ai prodotti $\boldsymbol{v}_i \cdot \boldsymbol{x}$, otteniamo

$$\|\boldsymbol{L}(\boldsymbol{x})\|^2 = \sum_{i=1}^{m} (\boldsymbol{v}_i \cdot \boldsymbol{x})^2 \leqslant \sum_{i=1}^{m} \|\boldsymbol{v}_i\|^2 \, \|\boldsymbol{x}\|^2 = \|A\|^2 \, \|\boldsymbol{x}\|^2 \,,$$

e dunque

$$\|\boldsymbol{L}(\boldsymbol{x})\| \leqslant \|A\| \, \|\boldsymbol{x}\| \,.$$

In particolare ne segue che tutte le applicazioni lineari sono lipschitziane,

$$\|\boldsymbol{L}(\boldsymbol{x}_1) - \boldsymbol{L}(\boldsymbol{x}_2)\| = \|\boldsymbol{L}(\boldsymbol{x}_1 - \boldsymbol{x}_2)\| \leqslant \|A\| \, \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \,.$$

Ne segue che, essendo lipschitziane, le applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}^m$ sono sempre continue.

La composizione di due applicazioni lineari, quando è definita, è ancora un'applicazione lineare. Se $\boldsymbol{L} \colon \mathbb{R}^n \to \mathbb{R}^m$ è l'applicazione lineare $\boldsymbol{L}(\boldsymbol{x}) = A\boldsymbol{x}$, rappresentata dalla matrice A con m righe e n colonne, e $\boldsymbol{M} \colon \mathbb{R}^m \to \mathbb{R}^k$ è l'applicazione lineare $\boldsymbol{M}(\boldsymbol{y}) = B\boldsymbol{y}$, rappresentata dalla matrice A con k righe e m colonne, allora la funzione composta $\boldsymbol{M} \circ \boldsymbol{L} \colon \mathbb{R}^n \to \mathbb{R}^k$ è l'applicazione lineare $\boldsymbol{M}(\boldsymbol{L}(\boldsymbol{x})) = BA\boldsymbol{x}$, rappresentata dalla matrice prodotto BA con k righe e n colonne.

2 Differenziabilità per funzioni vettoriali

Per funzioni a valori vettoriali il concetto di differenziabilità rimane quello della possibilità di avere un'approssimazione locale del primo ordine, che significa che gli incrementi della funzione sono descritti a livello infinitesimale da una applicazione lineare degli incrementi delle variabili.

Definizione 2.1. Una funzione $\boldsymbol{f}(\boldsymbol{x})$ definita in un intorno del punto $\boldsymbol{p} \in \mathbb{R}^n$ a valori in $\mathbb{R}^m$ si dice *differenziabile* nel punto $\boldsymbol{p}$ quando esiste una matrice A con m righe e n colonne tale che

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{p}) + A(\boldsymbol{x} - \boldsymbol{p}) + o\big(\|\boldsymbol{x} - \boldsymbol{p}\|\big), \quad \text{per } \boldsymbol{x} \to \boldsymbol{p}. \tag{1}$$

La matrice A si dice *matrice Jacobiana* della funzione $\boldsymbol{f}$ nel punto $\boldsymbol{p}$, e la indichiamo con la notazione $J_{\boldsymbol{f}}(\boldsymbol{p}) = A$.

Utilizzando la definizione del simbolo o-piccolo, la scrittura (1) è equivalente al limite vettoriale

$$\lim_{\boldsymbol{x} \to \boldsymbol{p}} \frac{\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{p}) - A(\boldsymbol{x} - \boldsymbol{p})}{\|\boldsymbol{x} - \boldsymbol{p}\|} = \boldsymbol{0}. \tag{2}$$

Proposizione 2.2. *Sia $\boldsymbol{f}(\boldsymbol{x})$ una funzione definita in un intorno del punto $\boldsymbol{p} \in \mathbb{R}^n$ a valori in $\mathbb{R}^m$, e siano $f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})$ le sue m componenti scalari. Allora $\boldsymbol{f}$ è differenziabile nel punto $\boldsymbol{p}$ se e solo se tutte le sue componenti scalari f_k sono differenziabili nel* *punto* $\boldsymbol{p}$*, e in tal caso la matrice Jacobiana di* $\boldsymbol{f}$ *nel punto* $\boldsymbol{p}$ *è la matrice le cui righe sono formate dai vettori gradienti delle componenti scalari* f_k*,*

$$J_{\boldsymbol{f}}(\boldsymbol{p}) = \begin{pmatrix} \nabla f_1(\boldsymbol{p}) \\ \nabla f_2(\boldsymbol{p}) \\ \vdots \\ \nabla f_m(\boldsymbol{p}) \end{pmatrix}.$$

Dimostrazione. Segue dal fatto che il limite (2) a valori vettoriali che definisce la differenziabilità per $\boldsymbol{f}$ è equivalente agli m limiti scalari

$$\lim_{\boldsymbol{x}\to\boldsymbol{p}} \frac{f_k(\boldsymbol{x}) - f_k(\boldsymbol{p}) - \boldsymbol{v}_k \cdot (\boldsymbol{x} - \boldsymbol{p})}{\|\boldsymbol{x} - \boldsymbol{p}\|} = 0, \qquad k = 1, \dots, m.$$

che definiscono la differenziabilità di ciascuna componente f_k, dove il vettore $\boldsymbol{v}_k$ è la k-esima riga della matrice A. □

Esempio 2.3. La funzione $\boldsymbol{f}(x, y) := (x + y^2, y\sin(x), \mathrm{e}^{x^2-y})$ è composta da tre componenti scalari che dipendono da due variabili,

$$f_1(x, y) = x + y^2, \quad f_2(x, y) = y \sin x, \quad f_3(x, y) = \mathrm{e}^{x^2-y}.$$

I loro gradienti sono

$$\nabla f_1(x, y) = (1, 2y), \quad \nabla f_2(x, y) = (y \cos x, \sin x), \quad \nabla f_3(x, y) = (2x\mathrm{e}^{x^2-y}, -\mathrm{e}^{x^2-y}).$$

La matrice Jacobiana di $\boldsymbol{f}$ è data da

$$J_{\boldsymbol{f}}(x, y) = \begin{pmatrix} 1, & 2y \\ y\cos x, & \sin x \\ 2x\mathrm{e}^{x^2-y}, & -\mathrm{e}^{x^2-y} \end{pmatrix}.$$

Data la funzione vettoriale di più variabili

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_m(x_1, x_2, \dots, x_n) \end{pmatrix}$$

possiamo anche descrivere la condizione di differenziabilità in un punto $\boldsymbol{x}$ usando un vettore $\boldsymbol{h}$ per gli incrementi (infinitesimali),

$$\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{h}) = \boldsymbol{f}(\boldsymbol{x}) + J_f(\boldsymbol{x})\boldsymbol{h} + o\big(\|\boldsymbol{h}\|\big), \quad \text{per } \boldsymbol{h} \to \boldsymbol{0}.$$

La matrice Jacobiana di f non è altro che la matrice formata da tutte le derivate parziali prime delle varie componenti della funzione,

$$J_{\boldsymbol{f}}(\boldsymbol{x}) = \big(\partial_{x_j} f_i(\boldsymbol{x})\big)_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} \partial_{x_1} f_1(\boldsymbol{x}) & \partial_{x_2} f_1(\boldsymbol{x}) & \dots & \partial_{x_n} f_1(\boldsymbol{x}) \\ \partial_{x_1} f_2(\boldsymbol{x}) & \partial_{x_2} f_2(\boldsymbol{x}) & \dots & \partial_{x_n} f_2(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} f_n(\boldsymbol{x}) & \partial_{x_2} f_n(\boldsymbol{x}) & \dots & \partial_{x_n} f_n(\boldsymbol{x}). \end{pmatrix}$$

L'applicazione lineare rappresentata dalla matrice jacobiana $J_{\boldsymbol{f}}(\boldsymbol{x})$ si dice *differenziale* di $\boldsymbol{f}$ nel punto $\boldsymbol{x}$,

$$\mathrm{d}\boldsymbol{f}(\boldsymbol{x})\colon \mathbb{R}^n \to \mathbb{R}^m, \qquad \mathrm{d}\boldsymbol{f}(\boldsymbol{x})(\boldsymbol{h}) = J_{\boldsymbol{f}}(\boldsymbol{x})\boldsymbol{h}.$$

Esempio 2.4. Consideriamo la funzione $\boldsymbol{f}(x,y,z) := \big(\sqrt{x+y}, \log(y+z)\big)$ e determiniamo la sua approssimazione del primo ordine intorno al punto $(1,2,3)$. Per mettere in evidenza gli incrementi delle variabili rispetto al punto, scriviamo

$$x = 1+a, \qquad y = 2+b, \qquad z = 3+c,$$

dove (a,b,c) è il vettore degli incrementi. Cerchiamo di ricondurci con opportune sostituzioni alle approssimazioni del primo ordine per funzioni di una variabile di radici quadrate e logaritmi; abbiamo infatti che

$$\sqrt{1+t} = 1 + \frac{1}{2}t + o(t), \qquad \log(1+t) = t + o(t), \qquad \text{per } t \to 0.$$

Per la prima componente di $\boldsymbol{f}$, quando $(a,b,c) \to (0,0,0)$ abbiamo

$$\begin{aligned} f_1(1+a, 2+b, 3+c) &:= \sqrt{3+a+b} = \sqrt{3}\cdot\sqrt{1+\frac{a+b}{3}} = \\ &= \sqrt{3}\Big(1 + \frac{1}{2}\cdot\frac{a+b}{3} + o\big(|a|+|b|\big)\Big) = \sqrt{3} + \frac{1}{2\sqrt{3}}a + \frac{1}{2\sqrt{3}}b + o\big(\|(a,b,c)\|\big). \end{aligned}$$

Per la seconda componente di $\boldsymbol{f}$, quando $(a,b,c) \to (0,0,0)$ abbiamo

$$\begin{aligned} f_2(1+a, 2+b, 3+c) &:= \log(5+b+c) = \log(5) + \log\Big(1 + \frac{b+c}{5}\Big) = \\ &= \log(5) + \frac{b+c}{5} + o\big(|b|+|c|\big) = \log(5) + \frac{1}{5}b + \frac{1}{5}c + o\big(\|(a,b,c)\|\big). \end{aligned}$$

Mettendo insieme le due cose troviamo

$$\begin{aligned} \boldsymbol{f}(1+a, 2+b, 3+c) &= \begin{pmatrix} \sqrt{3} + \frac{1}{2\sqrt{3}}a + \frac{1}{2\sqrt{3}}b \\ \log(5) + \frac{1}{5}b + \frac{1}{5}c \end{pmatrix} + o\big(\|(a,b,c)\|\big) = \\ &= \begin{pmatrix} \sqrt{3} \\ \log 5 \end{pmatrix} + \begin{pmatrix} \frac{1}{2\sqrt{3}}, & \frac{1}{2\sqrt{3}}, & 0 \\ 0, & \frac{1}{5}, & \frac{1}{5} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} + o\big(\|(a,b,c)\|\big). \end{aligned}$$

Dunque la matrice Jacobiana di $\boldsymbol{f}$ nel punto $(1,2,3)$ è

$$J_{\boldsymbol{f}}(1,2,3) = \begin{pmatrix} \frac{1}{2\sqrt{3}}, & \frac{1}{2\sqrt{3}}, & 0 \\ 0, & \frac{1}{5}, & \frac{1}{5} \end{pmatrix}.$$

Osservazione 2.5. Per l'applicazione lineare $\boldsymbol{L}(\boldsymbol{x}) = A\boldsymbol{x}$ rappresentata dalla matrice A abbiamo

$$\boldsymbol{L}(\boldsymbol{x}+\boldsymbol{h}) = \boldsymbol{L}(\boldsymbol{x}) + \boldsymbol{L}(\boldsymbol{h}) = \boldsymbol{L}(\boldsymbol{x}) + A\boldsymbol{h} + \boldsymbol{0}, \quad \forall \boldsymbol{h},$$

e dunque ne consegue che $J_{\boldsymbol{L}}(\boldsymbol{x}) = A$ per ogni $\boldsymbol{x}$.

Per funzioni scalari la matrice Jacobiana coincide con il vettore gradiente. Se $f\colon A \to \mathbb{R}$, con $A \subseteq \mathbb{R}^n$, abbiamo

$$\nabla f(\boldsymbol{x}) = J_f(\boldsymbol{x}) = \big(\partial_{x_1} f(\boldsymbol{x}), \partial_{x_2} f(\boldsymbol{x}), \dots, \partial_{x_n} f(\boldsymbol{x})\big).$$

Per funzioni vettoriali di una sola variabile la matrice Jacobiana è un vettore (colonna) formato dalle derivate delle singoli componenti, che chiameremo *vettore derivata*. Se $\boldsymbol{f}\colon I \to \mathbb{R}^n$, con $I \subseteq \mathbb{R}$, abbiamo

$$\boldsymbol{f}'(t) = J_{\boldsymbol{f}}(t) = \begin{pmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{pmatrix}.$$

Esercizio 2.6. Determina la matrice jacobiana per le seguenti funzioni nei punti indicati:

1. funzione $f(x, y) := (x\sqrt{y}, y \log x)$, punto $(\mathrm{e}, 4)$;
2. funzione $f(x, y, z) := (x, xy, xyz)$, punto $(2, 1, 0)$;
3. funzione $f(x, y) := (x + y, x - y, xy, x/y)$, punto $(-1, 1)$;
4. funzione $f(x, y, z, t) := (1 + 2x + 3y + 4z, 5x + 6y + 7z + 8t)$, punto $(4, 3, 2, 1)$;
5. funzione $f(u, v, w) := \arctan(u + w \log v)$, punto $(0, 1, 2)$;
6. funzione $f(t) := (t \cos t, t \sin t, t)$, punto π.

---

3 Differenziabilità di funzioni composte

Date due funzioni vettoriali $\boldsymbol{f}(\boldsymbol{x})$ e $\boldsymbol{g}(\boldsymbol{y})$ possiamo considerare la funzione composta

$$(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x}))$$

solo quando il numero di componenti della funzione interna $\boldsymbol{f}$ coincide con il numero di componenti della variabile $\boldsymbol{y}$ da cui dipende la funzione esterna $\boldsymbol{g}$.

Definizione 3.1. Siano $n, m, k \in \mathbb{N}$. Siano $\boldsymbol{f}\colon A \to \mathbb{R}^m$ definita sul dominio $A \subseteq \mathbb{R}^n$ e $\boldsymbol{g}\colon B \to \mathbb{R}^k$ definita sul dominio $B \subseteq \mathbb{R}^m$. Sia $\widetilde{A}$ l'insieme dei punti del dominio di $\boldsymbol{f}$ la cui immagine tramite $\boldsymbol{f}$ appartiene al dominio di $\boldsymbol{g}$,

$$\widetilde{A} := \{\boldsymbol{x} \in A\colon \boldsymbol{f}(\boldsymbol{x}) \in B\} \subseteq A \subseteq \mathbb{R}^n.$$

Quando $\widetilde{A}$ non è vuoto possiamo definire la *funzione composta* $\boldsymbol{g} \circ \boldsymbol{f}$ ponendo

$$\boldsymbol{g} \circ \boldsymbol{f}\colon \widetilde{A} \to \mathbb{R}^k, \qquad (\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) := \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})).$$

Teorema 3.2. *Supponiamo che*

- $\boldsymbol{f}$ *è differenziabile nel punto* $\boldsymbol{x}$*;*
- $\boldsymbol{g}$ *è differenziabile nel punto* $\boldsymbol{f}(\boldsymbol{x})$*;*
- *la funzione composta* $\boldsymbol{g} \circ \boldsymbol{f}$ *è definita in un intorno di* $\boldsymbol{x}$*.*

Allora abbiamo che $\boldsymbol{g} \circ \boldsymbol{f}$ *è differenziabile nel punto* $\boldsymbol{x}$ *e la sua matrice Jacobiana è data dal prodotto della Jacobiana di* $\boldsymbol{g}$ *in* $\boldsymbol{f}(\boldsymbol{x})$ *con la Jacobiana di* $\boldsymbol{f}$ *in* $\boldsymbol{x}$*,*

$$J_{\boldsymbol{g}\circ\boldsymbol{f}}(\boldsymbol{x}) = J_{\boldsymbol{g}}(\boldsymbol{f}(\boldsymbol{x}))J_{\boldsymbol{f}}(\boldsymbol{x}). \tag{3}$$

Nel caso in cui le funzioni siano due funzioni scalari di una variabile la formula (3) coincide con la nota formula per la derivata di funzioni composte scalari

$$D[g(f(x))] = g'(f(x))f'(x).$$

Anche la dimostrazione procede in maniera simile a quella del caso scalare.

Dimostrazione. Per l'ipotesi di differenziabilità di $\boldsymbol{f}$ in $\boldsymbol{x}$, per piccoli incrementi $\boldsymbol{h} \in \mathbb{R}^n$ abbiamo

$$\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) = \boldsymbol{f}(\boldsymbol{x}) + A\boldsymbol{h} + \boldsymbol{\varphi}(\boldsymbol{h}), \tag{4}$$

dove $A = J_{\boldsymbol{f}}(\boldsymbol{x})$ e $\boldsymbol{\varphi}$ è una funzione definita in un intorno di $\boldsymbol{0} \in \mathbb{R}^n$ con $\boldsymbol{\varphi}(\boldsymbol{0}) = \boldsymbol{0}$ e $\boldsymbol{\varphi}(\boldsymbol{h}) = o\left(\|\boldsymbol{h}\|\right)$ per $\boldsymbol{h} \to \boldsymbol{0}$. Sia $\boldsymbol{y} := \boldsymbol{f}(\boldsymbol{x})$. Per l'ipotesi di differenziabilità di $\boldsymbol{g}$ in $\boldsymbol{y}$, per piccoli incrementi $\boldsymbol{k} \in \mathbb{R}^m$ abbiamo

$$\boldsymbol{g}(\boldsymbol{y}+\boldsymbol{k}) = \boldsymbol{g}(\boldsymbol{y}) + B\boldsymbol{k} + \boldsymbol{\psi}(\boldsymbol{k}), \tag{5}$$

dove $B = J_{\boldsymbol{g}}(\boldsymbol{y})$ e $\boldsymbol{\psi}$ è una funzione definita in un intorno di $\boldsymbol{0} \in \mathbb{R}^m$ con $\boldsymbol{\psi}(\boldsymbol{0}) = \boldsymbol{0}$ e $\boldsymbol{\psi}(\boldsymbol{k}) = o\left(\|\boldsymbol{k}\|\right)$ per $\boldsymbol{k} \to \boldsymbol{0}$. Combiniamo (4) con (5) tramite le sostituzioni $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ e $\boldsymbol{k} = A\boldsymbol{h} + \boldsymbol{\varphi}(\boldsymbol{h})$, osservando che $\boldsymbol{x}$ è un punto fissato, mentre l'incremento $\boldsymbol{k}$ può essere reso piccolo scegliendo $\boldsymbol{h}$ sufficientemente piccolo, e otteniamo

$$\begin{aligned} \boldsymbol{g}\big(\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h})\big) &= \boldsymbol{g}\big(\boldsymbol{f}(\boldsymbol{x})+\boldsymbol{k}\big) = \boldsymbol{g}\big(\boldsymbol{f}(\boldsymbol{x})\big) + B\big(A\boldsymbol{h}+\boldsymbol{\varphi}(\boldsymbol{h})\big) + \boldsymbol{\psi}\left(A\boldsymbol{h}+\boldsymbol{\varphi}(\boldsymbol{h})\right) = \\ &= \boldsymbol{g}\big(\boldsymbol{f}(\boldsymbol{x})\big) + (BA)\boldsymbol{h} + B\boldsymbol{\varphi}(\boldsymbol{h}) + \boldsymbol{\psi}\left(A\boldsymbol{h}+\boldsymbol{\varphi}(\boldsymbol{h})\right) = \boldsymbol{g}\big(\boldsymbol{f}(\boldsymbol{x})\big) + (BA)\boldsymbol{h} + \boldsymbol{\omega}(\boldsymbol{h}), \end{aligned} \tag{6}$$

dove $\boldsymbol{\omega}(\boldsymbol{h}) := B\boldsymbol{\varphi}(\boldsymbol{h}) + \boldsymbol{\psi}\left(A\boldsymbol{h}+\boldsymbol{\varphi}(\boldsymbol{h})\right)$.

Verifichiamo che $\boldsymbol{\omega}(\boldsymbol{h}) = o\left(\|\boldsymbol{h}\|\right)$ per $\boldsymbol{h} \to \boldsymbol{0}$. Siccome $\boldsymbol{\varphi}(\boldsymbol{h}) = o\left(\|\boldsymbol{h}\|\right)$ per $\boldsymbol{h} \to \boldsymbol{0}$, abbiamo che

$$\forall \varepsilon_1 > 0,\, \exists \delta_1 > 0 : \|\boldsymbol{h}\| \leqslant \delta_1 \implies \|\boldsymbol{\varphi}(\boldsymbol{h})\| \leqslant \varepsilon_1 \|\boldsymbol{h}\|\,;$$

analogamente, siccome $\boldsymbol{\psi}(\boldsymbol{k}) = o\left(\|\boldsymbol{k}\|\right)$ per $\boldsymbol{k} \to \boldsymbol{0}$, abbiamo che

$$\forall \varepsilon_2 > 0,\, \exists \delta_2 > 0 : \|\boldsymbol{k}\| \leqslant \delta_2 \implies \|\boldsymbol{\psi}(\boldsymbol{k})\| \leqslant \varepsilon_2 \|\boldsymbol{k}\|\,.$$

Dato $\varepsilon > 0$ scegliamo ora

$$\varepsilon_1 := \min\left\{\frac{\varepsilon}{3\,\|B\|}, \sqrt{\frac{\varepsilon}{3}}\right\}, \qquad \varepsilon_2 := \min\left\{\frac{\varepsilon}{3\,\|A\|}, \sqrt{\frac{\varepsilon}{3}}\right\}, \qquad \delta := \min\left\{\delta_1, \frac{\delta_2}{\|A\| + \varepsilon_1}\right\}.$$

Preso $\boldsymbol{h}$ tale che $\|\boldsymbol{h}\| \leqslant \delta$, poniamo $\boldsymbol{k} := A\boldsymbol{h} + \boldsymbol{\varphi}(\boldsymbol{h})$ e abbiamo

$$\|\boldsymbol{k}\| \leqslant \|A\boldsymbol{h}\| + \|\boldsymbol{\varphi}(\boldsymbol{h})\| \leqslant (\|A\| + \varepsilon_1)\,\|\boldsymbol{h}\| \leqslant (\|A\| + \varepsilon_1)\,\delta \leqslant \delta_2,$$

dunque

$$\begin{aligned}\|\boldsymbol{\omega}(\boldsymbol{h})\| &\leqslant \|B\boldsymbol{\varphi}(\boldsymbol{h})\| + \|\boldsymbol{\psi}(\boldsymbol{k})\| \leqslant \|B\|\,\|\boldsymbol{\varphi}(\boldsymbol{h})\| + \varepsilon_2\,\|\boldsymbol{k}\| \leqslant \\ &\leqslant \|B\|\,\varepsilon_1\,\|\boldsymbol{h}\| + \varepsilon_2\big(\|A\| + \varepsilon_1\big)\,\|\boldsymbol{h}\| = \big(\|B\|\,\varepsilon_1 + \|A\|\,\varepsilon_2 + \varepsilon_1\varepsilon_2\big)\,\|\boldsymbol{h}\| \leqslant \varepsilon\,\|\boldsymbol{h}\|\,.\end{aligned}$$

Abbiamo così provato che per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che quando $\|\boldsymbol{h}\| \leqslant \delta$ si ha $\|\boldsymbol{\omega}(\boldsymbol{h})\| \leqslant \varepsilon\,\|\boldsymbol{h}\|$, ma ciò significa che $\boldsymbol{\omega}(\boldsymbol{h}) = o(\|\boldsymbol{h}\|)$ per $\boldsymbol{h} \to \boldsymbol{0}$. Dunque tornando all'uguaglianza (6) otteniamo

$$(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x} + \boldsymbol{h}) = (\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) + (BA)\boldsymbol{h} + o(\|\boldsymbol{h}\|), \quad \text{per } \boldsymbol{h} \to \boldsymbol{0},$$

da cui ricaviamo che $\boldsymbol{g} \circ \boldsymbol{f}$ è differenziabile nel punto $\boldsymbol{x}$ e la matrice Jacobiana è data da $J_{\boldsymbol{g}\circ\boldsymbol{f}}(\boldsymbol{x}) = BA$. □

Esempio 3.3. Considera le funzioni

$$\begin{aligned}\boldsymbol{F}(x,y) &:= \left(x + y, x\log(y), \cos(xy), \sqrt{x + y^2}\right), \\ \boldsymbol{G}(a,b,c,d) &:= \left(b\arctan(c), a + b^2 + c^3, \frac{d^2}{a+b}\right).\end{aligned}$$

Determiniamo la matrice Jacobiana della funzione composta $\boldsymbol{H}(x,y) := \boldsymbol{G}(\boldsymbol{F}(x,y))$ nel punto $\boldsymbol{p} := (\pi, 1)$. Calcoliamo prima il valore di $\boldsymbol{F}$ in $\boldsymbol{p}$,

$$\boldsymbol{q} := \boldsymbol{F}(\boldsymbol{p}) = \boldsymbol{F}(\pi,1) = \left(\pi + 1, \pi\log(1), \cos(\pi\,1), \sqrt{\pi + 1^2}\right) = \left(1 + \pi, 0, -1, \sqrt{1+\pi}\right).$$

Determiniamo la matrice Jacobiana di $\boldsymbol{F}$ calcolando tutte le sue derivate parziali,

$$J_{\boldsymbol{F}}(x,y) = \begin{pmatrix} 1 & 1 \\ \log(y) & \frac{x}{y} \\ -y\sin(xy) & -x\sin(xy) \\ \frac{1}{2\sqrt{x+y^2}} & \frac{y}{\sqrt{x+y^2}} \end{pmatrix}, \quad J_{\boldsymbol{F}}(\boldsymbol{p}) = \begin{pmatrix} 1 & 1 \\ 0 & \pi \\ 0 & 0 \\ \frac{1}{2\sqrt{1+\pi}} & \frac{1}{\sqrt{1+\pi}} \end{pmatrix}.$$

Determiniamo la matrice Jacobiana di $\boldsymbol{G}$ calcolando tutte le sue derivate parziali,

$$J_{\boldsymbol{G}}(a,b,c,d) = \begin{pmatrix} 0 & \arctan(c) & \frac{b}{1+c^2} & 0 \\ 1 & 2b & 3c^2 & 0 \\ -\frac{d^2}{(a+b^2)^2} & -\frac{d^2}{(a+b^2)^2} & 0 & \frac{2d}{a+b} \end{pmatrix}, \quad J_{\boldsymbol{G}}(\boldsymbol{q}) = \begin{pmatrix} 0 & -\frac{\pi}{4} & 0 & 0 \\ 1 & 0 & 3 & 0 \\ -\frac{1}{1+\pi} & -\frac{1}{1+\pi} & 0 & \frac{2}{\sqrt{1+\pi}} \end{pmatrix}.$$

Otteniamo la matrce Jacobiana della funzione composta come prodotto delle matrici Jacobiane delle sue componenti,

$$J_{\boldsymbol{H}}(\boldsymbol{p}) = J_{\boldsymbol{g}}(\boldsymbol{q})J_{\boldsymbol{f}}(\boldsymbol{p}) = \begin{pmatrix} 0 & -\frac{\pi}{4} & 0 & 0 \\ 1 & 0 & 3 & 0 \\ -\frac{1}{1+\pi} & -\frac{1}{1+\pi} & 0 & \frac{2}{\sqrt{1+\pi}} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & \pi \\ 0 & 0 \\ \frac{1}{2\sqrt{1+\pi}} & \frac{1}{\sqrt{1+\pi}} \end{pmatrix} = \begin{pmatrix} 0 & -\frac{\pi^2}{4} \\ 1 & 1 \\ 0 & \frac{1-\pi}{1+\pi} \end{pmatrix}.$$

Osservazione 3.4. Consideriamo il caso particolare della composizione di una funzione scalare con una funzione vettoriale di una variabile. Sia $\boldsymbol{f}\colon I \to A$ una funzione vettoriale di una variabile definita sull'intervallo $I \subseteq \mathbb{R}$ a valori nel dominio $A \subseteq \mathbb{R}^n$, e sia $g\colon A \to \mathbb{R}$ una funzione scalare di più variabili definita su A. Supponiamo che $\boldsymbol{f}$ e g siano differenziabili. La funzione composta $h\colon I \to \mathbb{R}$, definita da $h(t) := g(\boldsymbol{f}(t))$, è una funzione scalare di una variabile. Possiamo pensare ad $\boldsymbol{f}$ come una curva che ha sostegno nel dominio del campo scalarre g, la funzione composta $h = g \circ \boldsymbol{f}$ non fa altro che leggere i valori del campo g lungo la curva descritta da $\boldsymbol{f}$.

La derivata della funzione composta h nel punto t, per la formula (3), è data dal prodotto di un vettore riga per un vettore colonna, che possiamo interpretare come il prodotto scalare di due vettori,

$$\begin{aligned} h'(t) &:= J_h(t) = J_g(\boldsymbol{f}(t))J_{\boldsymbol{f}}(t) = \\ &= \left(\frac{\partial g}{\partial x_1}(\boldsymbol{f}(t)), \dots, \frac{\partial g}{\partial x_n}(\boldsymbol{f}(t))\right) \begin{pmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{pmatrix} = \sum_{j=1}^{n} \frac{\partial g}{\partial x_j}(\boldsymbol{f}(t))\, f_j'(t) = \\ &= \nabla g(\boldsymbol{f}(t))\cdot \boldsymbol{f}'(t) = D_{\boldsymbol{f}'(t)} g(\boldsymbol{f}(t)), \end{aligned}$$

e coincide con la derivata direzionale di g nel punto $\boldsymbol{f}(t)$ rispetto alla direzione $\boldsymbol{f}'(t)$.

Se poniamo $\boldsymbol{x} = \boldsymbol{f}(t)$ e $y = g(\boldsymbol{x})$ possiamo scrivere tale formula alla maniera dei fisici,

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \sum_{j=1}^{n} \frac{\partial y}{\partial x_j}\frac{\mathrm{d}x_j}{\mathrm{d}t} = \frac{\partial y}{\partial x_1}\frac{\mathrm{d}x_1}{\mathrm{d}t} + \frac{\partial y}{\partial x_2}\frac{\mathrm{d}x_2}{\mathrm{d}t} + \dots + \frac{\partial y}{\partial x_n}\frac{\mathrm{d}x_n}{\mathrm{d}t},$$

questa formula viene anche chiamata *regola della catena*, e dice che la derivata *totale* di y rispetto a t si ottiene sommando le derivate parziali di y rispetto alle variabili x_j, ciascuna moltiplicata per la corrispondente derivata di x_j rispetto a t.

Esempio 3.5. Supponiamo di voler studiare il comportamento del campo scalare F definito da

$$F(x, y, z) := x^2 y + z^3,$$

lungo la curva parametrizzata da

$$\boldsymbol{\gamma}(t) := (x(t), y(t), z(t)) = (t^2, \cos t, \log t), \qquad t > 0.$$

Andiamo a considerare la composizione

$$f(t) := F(\boldsymbol{\gamma}(t)) = F(x(t), y(t), z(t)).$$

La derivata di f è data da

$$f'(t) = \nabla F(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{\gamma}'(t) = \frac{\partial F}{\partial x}\, x' + \frac{\partial F}{\partial y}\, y' + \frac{\partial F}{\partial z}\, z' = 2xy\, x' + x^2\, y' + 3z^2\, z' =$$
$$= 2t^2(\cos t)2t + t^4(-\sin t) + 3(\log t)^2 \frac{1}{t} = 4t^3 \cos t - t^4 \sin t + \frac{3}{t}(\log t)^2.$$

Esercizio 3.6. Calcola il valore di $g'(1)$ dove g è la funzione definita da

$$g(t) := F\big(F(t,t^2), F(t^3,t^4)\big), \qquad F(x,y) := x - \log(xy^2).$$

Esercizio 3.7. Considera la funzione

$$\boldsymbol{F}(x,y) := \left(\frac{\cos(x)}{x+y}, \frac{\sin(x)}{x-y}\right).$$

Sia $\boldsymbol{G}(x,y) = \boldsymbol{F}(\boldsymbol{F}(x,y))$ la funzione composta di $\boldsymbol{F}$ con se stessa.

- Determina la matrice jacobiana di $\boldsymbol{F}$ in un generico punto (x,y).
- Determina la matrice jacobiana di $\boldsymbol{G}$ nel punto $(0,1)$.

Esercizio 3.8. Considera le funzioni

$$f(x,y) := \log(1+x^2y), \quad \boldsymbol{G}(t) := (\arctan(1+t), \sin(3t)), \quad \boldsymbol{H}(x,y) := G(f(x,y)).$$

- Calcola la matrice jacobiana della funzione $\boldsymbol{H}$ nel punto $(1,0)$.
- In quali punti (x,y) si ha che la matrice jacobiana di $\boldsymbol{H}$ é invertibile?

Esercizio 3.9. Considera le funzioni

$$\boldsymbol{F}(x,y) := \big(x + x^2y,\, \sin(x+y^2)\big), \qquad \boldsymbol{G}(u,v) := \left(\frac{u}{1+v},\, \log(1+2u-3v)\right),$$
$$\boldsymbol{H}(x,y) := \boldsymbol{G}(\boldsymbol{F}(x,y)), \qquad \boldsymbol{K}(u,v) := \boldsymbol{F}(\boldsymbol{G}(u,v)).$$

Calcola le matrici jacobiane di H e K in $(0,0)$.

Esercizio 3.10. Considera le funzioni definite da

$$\boldsymbol{f}(x,y) := \left(x\cos(y), \frac{2y}{x^2+1}\right), \qquad g(s,t) := s - \int_0^t \mathrm{e}^{-r^2}\, \mathrm{d}r.$$

- Calcola tutte le derivate parziali di $\boldsymbol{f}$ e di g.
- Calcola il gradiente della funzione composta $g \circ \boldsymbol{f} \colon \mathbb{R}^2 \to \mathbb{R}$ nel punto $(0,0)$.

Esercizio 3.11. Considera le funzioni definite da

$$\boldsymbol{f}(t) := \big(t, t^2, t^3\big), \quad \boldsymbol{g}(x,y,z) := \big(\log(xyz), 1 - \mathrm{e}^{x+2y-3z}\big), \quad h(a,b) := \int_a^b \cos(\pi r^2)\, \mathrm{d}r.$$

- Calcola tutte le derivate parziali di $\boldsymbol{f}$, $\boldsymbol{g}$ e h.
- Sia k la funzione composta $k(t) := h(\boldsymbol{g}(\boldsymbol{f}(t)))$; calcola il valore della derivata $k'(1)$.

Esercizio 3.12. Considera le funzioni

$$F(x,y,z) := z\int_x^y \left(\frac{\log(t)}{t}\right)^2 \mathrm{d}t, \qquad \boldsymbol{G}(t) := (t, t^2, t^3),$$

definite per $t, x, y, z > 0$. Sia $h(t) := F(\boldsymbol{G}(t))$ la loro funzione composta.

- Calcola tutte le derivate prime di F e di $\boldsymbol{G}$.
- Calcola $h(1)$, $h'(1)$ e $h''(1)$.

---

4 Calcolo differenziale vettoriale

L'operazione di derivazione parziale $\frac{\partial}{\partial x_j}$, quando è ben definita, può essere intesa come un'applicazione lineare che trasforma funzioni in altre funzioni,

$$f \mapsto \frac{\partial f}{\partial x_j}.$$

Ad esempio possiamo vederla come una trasformazione lineare che prende funzioni di classe C^1 e restituisce funzioni continue,

$$\frac{\partial}{\partial x_j} : C^1 \to C^0.$$

Sia $k \in \mathbb{N}_0$. Indichiamo con la scrittura $C^k(A;B)$ l'insieme delle funzioni di classe C^k che hanno come dominio A e come codominio B. Essere di classe C^k significa possedere tutte le derivate fino all'ordine k ed esse sono tutte funzioni continue. Dato un sottoinsieme aperto $\Omega \subseteq \mathbb{R}^n$, l'operatore di derivazione parziale $\frac{\partial}{\partial x_j}$ agisce come operatore lineare che trasforma funzioni di $C^{k+1}(\Omega;\mathbb{R})$ in funzioni di $C^k(\Omega;\mathbb{R})$,

$$\frac{\partial}{\partial x_j} : C^{k+1}(\Omega;\mathbb{R}) \to C^k(\Omega;\mathbb{R}), \qquad j = 1, 2, \dots, n.$$

La funzione $\frac{\partial f}{\partial x_j}$ contiene informazioni su come $f(\boldsymbol{x})$ varia al variare della singola variabile x_j. Possiamo combinare queste derivate parziali in diversi modi per avere informazioni sulle modalità di variazione della funzione rispetto a tutte quante le variabili.

4.1 Gradiente

L'operatore gradiente associa ad una funzione scalare il campo vettoriale formato da tutte le sue derivate parziali.

Definizione 4.1. Sia $n \in \mathbb{N}$. Sia Ω un sottoinsieme aperto di $\mathbb{R}^n$. Definiamo l'operatore *gradiente* per funzioni di classe C^1:

$$\begin{aligned} \operatorname{grad}\colon C^1(\Omega;\mathbb{R}) &\to C^0(\Omega;\mathbb{R}^n), \\ \operatorname{grad} f(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) &:= \Big(\frac{\partial f}{\partial x_1}(\boldsymbol{x}), \dots, \frac{\partial f}{\partial x_n}(\boldsymbol{x})\Big), \quad \forall \boldsymbol{x} \in \Omega. \end{aligned}$$

Il gradiente è un operatore lineare e valgono le seguenti proprietà:

$$\nabla(f+g) = (\nabla f) + (\nabla g), \qquad \nabla(fg) = f\nabla g + g\nabla f.$$

Se indichiamo con $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ i vettori della base canonica di $\mathbb{R}^n$, per linearità possiamo pensare al gradiente come una somma vettoriale delle sue derivate parziali

$$\nabla f = \frac{\partial f}{\partial x_1}\boldsymbol{e}_1 + \dots + \frac{\partial f}{\partial x_n}\boldsymbol{e}_n = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\boldsymbol{e}_j.$$

A livello simbolico, l'operatore gradiente possiamo pensarlo come un vettore le cui componenti sono operatori di derivazione parziale,

$$\nabla = \Big(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\Big).$$

Per funzioni scalari di tre variabili, $f(x,y,z)$, abbiamo

$$\nabla = (\partial_x, \partial_y, \partial_z), \qquad \nabla f = \partial_x f\, \vec{\boldsymbol{i}} + \partial_y f\, \vec{\boldsymbol{j}} + \partial_z f\, \vec{\boldsymbol{k}}.$$

Abbiamo visto (nella lezione 12) che per funzioni differenziabili conoscere il gradiente permette di calcolare facilmente derivate direzionali rispetto a qualsiasi direzione $\boldsymbol{v}$:

$$D_{\boldsymbol{v}} f(\boldsymbol{p}) = \nabla f(\boldsymbol{p}) \cdot \boldsymbol{v}.$$

Dal punto di vista geometrico il gradiente ci fornisce informazioni sulle direzioni di massima crescita della funzione. Se $\nabla f(\boldsymbol{p}) \neq \boldsymbol{0}$ allora la direzione di massima crescita per la funzione $f(\boldsymbol{x})$ nel punto $\boldsymbol{p}$ è data dal versore

$$\boldsymbol{u} := \frac{1}{\|\nabla f(\boldsymbol{p})\|} \nabla f(\boldsymbol{p})$$

e la velocità di crescita in tale direzione è data dalla derivata direzionale

$$D_{\boldsymbol{u}} f(\boldsymbol{p}) = \nabla f(\boldsymbol{p}) \cdot \boldsymbol{u} = \|\nabla f(\boldsymbol{p})\|.$$

Esempio 4.2. Consideriamo il campo scalare $f(x, y, z) := x^3 + y^2 z$. Il suo gradiente è il campo vettoriale

$$\nabla f(x, y, z) = 3x^2 \vec{\boldsymbol{\imath}} + 2yz \vec{\boldsymbol{\jmath}} + y^2 \vec{\boldsymbol{k}}.$$

Esercizio 4.3. Sia $f(x, y)$ una funzione differenziabile nel punto $(x_\star, y_\star)$. Quali condizioni devono soddisfare due vettori $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^2$ affinché il gradiente $\nabla f(x_\star, y_\star)$ possa essere determinato conoscendo solo i valori delle derivate direzionali $D_{\boldsymbol{u}} f(x_\star, y_\star)$ e $D_{\boldsymbol{v}} f(x_\star, y_\star)$?

4.2 Divergenza

L'operatore di divergenza associa ad un campo vettoriale il campo scalare definito dalla traccia della sua matrice Jacobiana. (La traccia di una matrice è la somma degli elementi sulla sua diagonale principale.)

Definizione 4.4. Sia $n \in \mathbb{N}$. Sia Ω un sottoinsieme aperto di $\mathbb{R}^n$. Definiamo l'operatore di *divergenza* per funzioni di classe C^1:

$$\operatorname{div}\colon C^1(\Omega; \mathbb{R}^n) \to C^0(\Omega; \mathbb{R}),$$
$$\operatorname{div} \boldsymbol{f}(\boldsymbol{x}) = \nabla \cdot \boldsymbol{f}(\boldsymbol{x}) := \operatorname{tr} J_{\boldsymbol{f}}(\boldsymbol{x}) = \sum_{j=1}^{n} \frac{\partial f_j}{\partial x_j}(\boldsymbol{x}) = \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}) + \cdots + \frac{\partial f_n}{\partial x_n}(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \Omega.$$

La divergenza è un operatore lineare e valgono le seguenti proprietà:

$$\nabla \cdot (\boldsymbol{f} + \boldsymbol{g}) = (\nabla \cdot \boldsymbol{f}) + (\nabla \cdot \boldsymbol{g}), \qquad \nabla \cdot (\varphi \boldsymbol{f}) = \varphi (\nabla \cdot \boldsymbol{f}) + (\nabla \varphi) \cdot \boldsymbol{f},$$

per ogni coppia di campi vettoriali $\boldsymbol{f}, \boldsymbol{g}$ e per ogni campo scalare φ.

A livello simbolico, l'operatore di divergenza applicato ad un campo vettoriale possiamo pensarlo come il prodotto scalare tra l'operatore gradiente e il campo vettoriale,

$$\operatorname{div} \boldsymbol{f} = \nabla \cdot \boldsymbol{f} = \frac{\partial}{\partial x_1} f_1 + \cdots + \frac{\partial}{\partial x_n} f_n.$$

Attenzione però all'ordine dei fattori: le scritture $\nabla \cdot \boldsymbol{f}$ e $\boldsymbol{f} \cdot \nabla$ non significano la stessa cosa: il primo è un campo scalare (la divergenza di $\boldsymbol{f}$), mentre il secondo è un operatore differenziale scalare. Se φ è un campo scalare abbiamo

$$(\nabla \cdot \boldsymbol{f}) \varphi = \frac{\partial f_1}{\partial x_1} \varphi + \cdots + \frac{\partial f_n}{\partial x_n} \varphi,$$
$$(\boldsymbol{f} \cdot \nabla) \varphi = f_1 \frac{\partial \varphi}{\partial x_1} + \cdots + f_n \frac{\partial \varphi}{\partial x_n}.$$

Per funzioni vettoriali scalari di tre variabili a valori in $\mathbb{R}^3$,

$$\boldsymbol{f}(x, y, z) = \big(f_1(x, y, z), f_2(x, y, z), f_3(x, y, z)\big), \tag{7}$$

abbiamo

$$\nabla\cdot\boldsymbol{f} = \partial_x f_1 + \partial_y f_2 + \partial_z f_3.$$

Dal punto di vista geometrico il valore della divergenza di un campo vettoriale in un punto è una misura di quanto i vettori del campo tendono ad "aprirsi" (ad allontanarsi tra loro).

Da un punto di vista fisico, se immaginiamo il vettore $\boldsymbol{f}(\boldsymbol{x})$ come la velocità della particella di un fluido che si trova nel punto $\boldsymbol{x}$ allora div $\boldsymbol{f}(\boldsymbol{x})$ indica la densità del flusso del fluido nel punto $\boldsymbol{x}$: immaginiamo di misurare la quantità di fluido che esce (o entra) attraverso la superficie di un piccolo intorno regolare del punto $\boldsymbol{x}$, il rapporto tra il flusso misurato e il volume dell'intorno tende (sotto ragionevoli ipotesi di regolarità) al valore della divergenza nel punto, quando l'intorno tende a concentrarsi nel punto.

Esempio 4.5. Consideriamo il campo vettoriale $\boldsymbol{f}(x, y, z) := (x^3, y^2 z, xz + y)$. La sua divergenza è il campo scalare

$$\nabla\cdot\boldsymbol{f}(x, y, z) = 3x^2 + 2yz + x.$$

Esercizio 4.6. Calcola la divergenza dei seguenti campi vettoriali:

1. $\boldsymbol{F}(x, y) = \Big(\log\sqrt{x^2 + y^2}, \arctan\frac{y}{x}\Big)$;
2. $\boldsymbol{F}(x, y, z) = yz\vec{\boldsymbol{i}} + xz\vec{\boldsymbol{j}} + xy\vec{\boldsymbol{k}}$;
3. $\boldsymbol{F}(x, y, z, t) = \nabla(xy^2z^3\mathrm{e}^t)$.

Esercizio 4.7. Sia $f(r)$ una funzione derivabile di una variabile. Dimostra che per $\boldsymbol{x} \in \mathbb{R}^n$ vale la formula

$$\nabla\cdot\big(f(\|\boldsymbol{x}\|)\,\boldsymbol{x}\big) = \|\boldsymbol{x}\|\, f'(\|\boldsymbol{x}\|) + n f(\|\boldsymbol{x}\|).$$

4.3 Rotore

L'operatore rotore associa ad un campo vettoriale su $\mathbb{R}^3$ un altro campo vettoriale, operando simbolicamente il prodotto vettoriale tra l'operatore gradiente e il campo vettoriale dato.

Definizione 4.8. Sia Ω un sottoinsieme aperto di $\mathbb{R}^3$. Definiamo l'operatore *rotore* per funzioni di classe C^1:

$$\mathrm{rot}\colon C^1(\Omega; \mathbb{R}^3) \to C^0(\Omega; \mathbb{R}^3),$$
$$\mathrm{rot}\,\boldsymbol{f}(\boldsymbol{x}) = \nabla\times\boldsymbol{f}(\boldsymbol{x}) := \Big(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\Big)\vec{\boldsymbol{i}} + \Big(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}\Big)\vec{\boldsymbol{j}} + \Big(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\Big)\vec{\boldsymbol{k}}, \quad \forall \boldsymbol{x} \in \Omega.$$

Il rotore è un operatore lineare e valgono le seguenti proprietà:

$$\nabla\times(\boldsymbol{f}+\boldsymbol{g}) = (\nabla\times\boldsymbol{f}) + (\nabla\times\boldsymbol{g}), \qquad \nabla\times(\varphi\boldsymbol{f}) = \varphi(\nabla\times\boldsymbol{f}) + (\nabla\varphi)\times\boldsymbol{f},$$

per ogni coppia di campi vettoriali $\boldsymbol{f}, \boldsymbol{g}$ e per ogni campo scalare φ.

A livello simbolico, l'operatore rotore applicato ad un campo vettoriale possiamo pensarlo come il determinante di una matrice simbolica,

$$\operatorname{rot}\boldsymbol{f} = \nabla\times\boldsymbol{f} = \det\begin{pmatrix} \vec{\boldsymbol{i}} & \vec{\boldsymbol{j}} & \vec{\boldsymbol{k}} \\ \partial_x & \partial_y & \partial_z \\ f_1 & f_2 & f_3 \end{pmatrix}.$$

Dal punto di vista geometrico il vettore rotore di un campo vettoriale in un punto è una misura di quanto i vettori del campo tendano a "ruotare" intorno al punto.

Da un punto di vista fisico, se facciamo ruotare un corpo rigido che occupa lo spazio $\mathbb{R}^3$ intorno all'origine con velocità angolare costante $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$, allora il campo di velocità che descrive la velocità dei punti del corpo in ciascun punto dello spazio è descritto dal campo vettoriale dato da

$$\boldsymbol{v}(x,y,z) = \boldsymbol{\omega}\times(x,y,z) = (\omega_2 z - \omega_3 y)\vec{\boldsymbol{i}} + (\omega_3 x - \omega_1 z)\vec{\boldsymbol{j}} + (\omega_1 y - \omega_2 x)\vec{\boldsymbol{k}}.$$

In questo caso $\operatorname{rot}\boldsymbol{v}$ risulta essere proporzionale alla velocità angolare,

$$\begin{aligned}\operatorname{rot}\boldsymbol{v} = \nabla\times\boldsymbol{v} &= (\partial_y v_3 - \partial_z v_2)\,\vec{\boldsymbol{i}} + (\partial_z v_1 - \partial_x v_3)\,\vec{\boldsymbol{j}} + (\partial_x v_2 - \partial_y v_1)\,\vec{\boldsymbol{k}} = \\ &= (\omega_1 - (-\omega_1))\,\vec{\boldsymbol{i}} + (\omega_2 - (-\omega_2))\,\vec{\boldsymbol{j}} + (\omega_3 - (-\omega_3))\,\vec{\boldsymbol{k}} = 2\boldsymbol{\omega}.\end{aligned}$$

Quando il corpo non è rigido (ad esempio un fluido), la velocità angolare può variare da punto a punto ed essa può essere misurata tramite l'operatore rotore.

Esempio 4.9. Consideriamo il campo vettoriale $\boldsymbol{f}(x,y,z) := (x^3, y^2 z, xz + y)$. Il suo rotore è il campo vettoriale

$$\nabla\times\boldsymbol{f}(x,y,z) = (1-y^2)\vec{\boldsymbol{i}} + (0-z)\vec{\boldsymbol{j}} + (0-0)\vec{\boldsymbol{k}} = (1-y^2)\vec{\boldsymbol{i}} - z\vec{\boldsymbol{j}} = (1-y^2, -z, 0).$$

Osservazione 4.10. Se abbiamo un campo vettoriale dipendente da due variabili e a valori in $\mathbb{R}^2$,

$$\boldsymbol{g} \in C^1(\Omega; \mathbb{R}^2),$$

definito su Ω aperto in $\mathbb{R}^2$, possiamo estenderlo ad un campo dipendente da tre variabili a valori in $\mathbb{R}^3$ ponendo

$$\boldsymbol{G}(x,y,z) := \big(g_1(x,y), g_2(x,y), 0\big), \quad \forall (x,y,z) \in \Omega\times\mathbb{R}.$$

Se calcoliamo il rotore di $\boldsymbol{G}$ otteniamo

$$\operatorname{rot}\boldsymbol{G} = (\partial_y 0 - \partial_z g_2)\vec{\boldsymbol{i}} + (\partial_z g_1 - \partial_x 0)\vec{\boldsymbol{j}} + (\partial_x g_2 - \partial_y g_1)\vec{\boldsymbol{k}} = (\partial_x g_2 - \partial_y g_1)\vec{\boldsymbol{k}} = (0, 0, \partial_x g_2 - \partial_y g_1).$$

Osserviamo che si tratta sempre di un vettore parallelo all'asse z. L'informazione utile in questo calcolo è tutta contenuta nella terza componente, che è un campo scalare che dipende solo dalle variabili x e y, ed essa definisce il *rotore* (bidimensionale) del campo $\boldsymbol{g}$,

$$\operatorname{rot}\boldsymbol{g}(x,y) := \partial_x g_2(x,y) - \partial_y g_1(x,y).$$

Esercizio 4.11. Calcola il rotore dei seguenti campi vettoriali:

1. $\boldsymbol{F}(x, y) = \left(\log \sqrt{x^2 + y^2}, \arctan \frac{y}{x}\right)$;
2. $\boldsymbol{F}(x, y, z) = yz\vec{\boldsymbol{\imath}} + xz\vec{\boldsymbol{\jmath}} + xy\vec{\boldsymbol{k}}$;
3. $\boldsymbol{F}(x, y, z) = \nabla\times(z, x^2, y^3)$.

4.4 Composizione di operatori

Per campi di classe C^2 possiamo considerare la composizione di coppie di operatori differenziali in vari modi.

Sia $f(x, y, z)$ un campo scalare di classe C^2 e sia $\boldsymbol{F}(x, y, z) = (F_1, F_2, F_3)$ un campo vettoriale di classe C^2.

- Il rotore del gradiente di un campo scalare è sempre nullo, infatti per il teorema di Schwarz sull'uguaglianza delle derivate seconde miste abbiamo

$$\begin{aligned}\operatorname{rot}(\operatorname{grad} f) &= \nabla\times(\nabla f) = \\ &= \big(\partial_y\partial_z f - \partial_z\partial_y f\big)\vec{\boldsymbol{\imath}} + \big(\partial_z\partial_x f - \partial_x\partial_z f\big)\vec{\boldsymbol{\jmath}} + \big(\partial_x\partial_y f - \partial_y\partial_x f\big)\vec{\boldsymbol{k}} = \mathbf{0}.\end{aligned}$$

- La divergenza del rotore di un campo vettoriale è sempre nulla, infatti sempre per il teorema di Schwarz sull'uguaglianza delle derivate seconde miste abbiamo

$$\begin{aligned}\operatorname{div}(\operatorname{rot} \boldsymbol{F}) &= \nabla\cdot(\nabla\times\boldsymbol{F}) = \\ &= \partial_x(\partial_y F_3 - \partial_z F_2) + \partial_y(\partial_z F_1 - \partial_x F_3) + \partial_z(\partial_x F_2 - \partial_y F_1) = \\ &= \partial_y\partial_z F_1 - \partial_z\partial_y F_1 + \partial_z\partial_x F_2 - \partial_x\partial_z F_2 + \partial_x\partial_y F_3 - \partial_y\partial_x F_3 = 0.\end{aligned}$$

- La divergenza del gradiente di un campo scalare definisce un operatore differenziale del secondo ordine, detto *laplaciano*, esso coincide con la traccia della matrice Hessiana,

$$\Delta f := \operatorname{div}(\operatorname{grad} f) = \nabla\cdot(\nabla f) = \partial_{x_1}\partial_{x_1} f + \cdots + \partial_{x_n}\partial_{x_n} f = \sum_{j=1}^{n} \partial^2_{x_j} f = \operatorname{tr} H_f.$$

Esercizio 4.12. Sia $\boldsymbol{F}$ un campo vettoriale di classe C^2 su $\mathbb{R}^3$ a valori in $\mathbb{R}^3$. Calcola, cercando di semplicare al meglio il risultato il campo vettoriale $\boldsymbol{G}$ che si ottiene dalla differenza tra il rotore del rotore e il gradiente della divergenza,

$$\boldsymbol{G} := \nabla\times(\nabla\times\boldsymbol{F}) - \nabla(\nabla\cdot\boldsymbol{F}).$$

Esercizio 4.13. Dato il campo vettoriale

$$\boldsymbol{f}(x,y,z) := (x^2y, xyz, xz^2),$$

Calcola il campo vettoriale

$$\boldsymbol{g} := \nabla\times(\nabla\times\boldsymbol{f}) - \nabla(\nabla\cdot\boldsymbol{f}).$$
