

## Calcolo differenziale

Testo di riferimento: N. Fusco, P. Marcellini, C. Sbordone, *Analisi Matematica Due*, Liguori Editore.

### 1 Algebra lineare e geometria in $\mathbb{R}^n$

In questo capitolo rivediamo alcuni concetti di algebra lineare e geometria che ci servono per introdurre il calcolo differenziale per funzioni di più variabili reali a valori vettoriali. Per approfondimenti sulla teoria degli spazi vettoriali si rimanda al corso di Algebra lineare e Geometria.

#### 1.1 Prodotto scalare e norma euclidea

Sia  $n$  un intero positivo.  $\mathbb{R}^n$  è l'insieme dei vettori

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

dove  $x_1, \dots, x_n$  sono numeri reali;  $x_k$  è la  $k$ -esima *componente* del vettore  $x$ . La somma di due vettori  $x, y$  e il prodotto di un vettore  $x$  per uno scalare  $\lambda \in \mathbb{R}$  sono definiti in questo modo:

$$x + y = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda x = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}, \quad x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}.$$

Queste operazioni rendono  $\mathbb{R}^n$  uno *spazio vettoriale*.  $n$  è la sua dimensione. La *base canonica* è costituita dai vettori  $e_1, e_2, \dots, e_n$ , dove  $e_k$  è il vettore che ha la  $k$ -esima componente uguale a 1 e tutte le altre uguali a zero:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Ogni vettore  $x$  di componenti  $x_k$  si può scrivere come

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + x_2 e_2 + \dots + x_n e_n = \sum_{k=1}^n x_k e_k \quad (1)$$

( $x_k$  sono scalari e  $e_k$  sono vettori). Ogni vettore di  $\mathbb{R}^n$  può essere pensato anche come *punto* dello spazio. Spesso un vettore  $x$  si indica anche in riga,  $x = (x_1, x_2, \dots, x_n)$ , per comodità di scrittura. Il vettore nullo  $0 = (0, 0, \dots, 0)$  di  $\mathbb{R}^n$  è detto *origine*.

**Definizione 1.1.** Si chiama *prodotto scalare* di due vettori  $x, y \in \mathbb{R}^n$  il numero

$$x \cdot y = x_1 y_1 + \dots + x_n y_n = \sum_{k=1}^n x_k y_k \in \mathbb{R}.$$

Si noti che  $x \cdot y$  è uno scalare. Si indica in vari modi:

$$x \cdot y = x \circ y = (x, y) = (x|y) = \langle x, y \rangle = \langle x|y \rangle. \quad \square$$

**Proposizione 1.2.** *Il prodotto scalare soddisfa le seguenti proprietà:*

- (a)  $x \cdot x \geq 0$  per ogni  $x \in \mathbb{R}^n$ ;
- (b)  $x \cdot x = 0$  se e solo se  $x = 0$ ;
- (c)  $x \cdot y = y \cdot x$  per ogni  $x, y \in \mathbb{R}^n$ ;
- (d)  $(x + y) \cdot z = x \cdot z + y \cdot z$  per ogni  $x, y, z \in \mathbb{R}^n$ ;
- (e)  $(\lambda x) \cdot y = \lambda(x \cdot y)$  per ogni  $x, y \in \mathbb{R}^n$  ed ogni  $\lambda \in \mathbb{R}$ .

*Dimostrazione.* (a) e (b) valgono perché  $x \cdot x = x_1^2 + \dots + x_n^2$ , la somma dei quadrati delle componenti. (c), (d) e (e) valgono per le proprietà commutativa e associativa del prodotto e della somma di numeri reali, ragionando sulle componenti.  $\square$

**Definizione 1.3.** Si chiama *norma euclidea* di un vettore  $x \in \mathbb{R}^n$  il numero

$$|x| = \sqrt{x \cdot x} = \left( \sum_{k=1}^n x_k^2 \right)^{1/2} = \sqrt{x_1^2 + \dots + x_n^2}.$$

Si indica sia  $|x|$  che  $\|x\|$ , ed è un numero  $\geq 0$ .  $\square$

La norma euclidea  $|x|$  è la lunghezza del vettore  $x$ , cioè la distanza del punto  $x$  dall'origine (teorema di Pitagora). Si noti che la radice quadrata nella definizione di norma ha senso perché  $x \cdot x \geq 0$ .

**Lemma 1.4** (Disuguaglianza di Cauchy–Schwartz). *Per ogni  $x, y \in \mathbb{R}^n$  è*

$$|x \cdot y| \leq |x| |y|$$

( $|x \cdot y|$  è il valore assoluto del numero reale  $x \cdot y$ , mentre  $|x|$  e  $|y|$  sono le norme dei vettori  $x$  e  $y$ ).

*Dimostrazione.* Se  $y = 0$ ,  $x \cdot y = x \cdot 0 = 0$ ,  $|y| = |0| = 0$ , e la disuguaglianza è soddisfatta banalmente ( $0 \leq 0$ ). Se invece  $y \neq 0$ , consideriamo il vettore  $x + ty$ ,  $t \in \mathbb{R}$ , e, usando le proprietà del prodotto scalare, calcoliamo

$$\begin{aligned} |x + ty|^2 &= (x + ty) \cdot (x + ty) = x \cdot (x + ty) + ty \cdot (x + ty) \\ &= x \cdot x + x \cdot (ty) + ty \cdot x + (ty) \cdot (ty) \\ &= x \cdot x + 2tx \cdot y + t^2 y \cdot y \\ &= |x|^2 + 2tx \cdot y + |y|^2 t^2 =: p(t). \end{aligned}$$

Ora,  $|x + ty|^2 \geq 0$  per ogni  $t \in \mathbb{R}$ , quindi  $p(t)$  è un polinomio di secondo grado in  $t$  (perché  $|y|^2 \neq 0$ ), e  $p(t) \geq 0$  per ogni  $t \in \mathbb{R}$ . Perciò  $p(t)$  non può avere due radici reali distinte, e di conseguenza il suo discriminante deve essere  $\leq 0$ :

$$(x \cdot y)^2 - |x|^2 |y|^2 \leq 0.$$

Dunque  $(x \cdot y)^2 \leq |x|^2 |y|^2$ , da cui la tesi.  $\square$

La disuguaglianza di Cauchy-Schwartz riscritta in componenti è:

$$\left| \sum_{k=1}^n x_k y_k \right| \leq \left( \sum_{k=1}^n x_k^2 \right)^{1/2} \left( \sum_{k=1}^n y_k^2 \right)^{1/2}$$

per ogni  $x_k, y_k \in \mathbb{R}$ .

**Definizione 1.5.** Siano  $x \neq 0$  e  $y \neq 0$  due vettori di  $\mathbb{R}^n$ . Dalla disuguaglianza di Cauchy–Schwartz segue che

$$-1 \leq \frac{x \cdot y}{|x| |y|} \leq 1.$$

Di conseguenza esiste un unico  $\vartheta \in [0, \pi]$  tale che

$$\cos \vartheta = \frac{x \cdot y}{|x| |y|}.$$

$\vartheta$  si chiama *angolo* formato dai vettori  $x$  e  $y$ . Se  $\vartheta = \pi/2$ , cioè se  $x \cdot y = 0$ , i due vettori si dicono *ortogonali*; in simboli,  $x \perp y$ .  $\square$

**Definizione 1.6.** Un vettore  $x \in \mathbb{R}^n$  si dice *versore* se  $|x| = 1$ . □

Dato un vettore  $x \neq 0$ , il suo versore è il vettore

$$\hat{x} = \frac{1}{|x|} x = \frac{x}{|x|} = \left( \frac{x_1}{|x|}, \frac{x_2}{|x|}, \frac{x_3}{|x|}, \dots, \frac{x_n}{|x|} \right). \quad (2)$$

$\hat{x}$  è versore perché

$$|\hat{x}|^2 = \frac{x_1^2}{|x|^2} + \frac{x_2^2}{|x|^2} + \dots + \frac{x_n^2}{|x|^2} = \frac{x_1^2 + x_2^2 + \dots + x_n^2}{|x|^2} = \frac{|x|^2}{|x|^2} = 1.$$

L'angolo tra due vettori non nulli  $x$  e  $y$  è uguale all'angolo tra i corrispondenti versori  $\hat{x}$  e  $\hat{y}$ , perché (usando la proprietà (e) del prodotto scalare)

$$\frac{x \cdot y}{|x||y|} = \frac{1}{|x||y|} (x \cdot y) = \left( \frac{1}{|x|} x \right) \cdot \left( \frac{1}{|y|} y \right) = \hat{x} \cdot \hat{y} = \frac{\hat{x} \cdot \hat{y}}{|\hat{x}||\hat{y}|}.$$

**Osservazione 1.7.** Dato un vettore  $a = (a_1, \dots, a_n)$ , l'insieme dei vettori ortogonali ad  $a$  è

$$\{x \in \mathbb{R}^n : a \cdot x = 0\} = \{x \in \mathbb{R}^n : a_1 x_1 + \dots + a_n x_n = 0\},$$

cioè l'iperpiano (sottospazio vettoriale di codimensione 1) di equazione  $a_1 x_1 + \dots + a_n x_n = 0$ . □

**Proposizione 1.8.** *La norma euclidea soddisfa:*

- (i)  $|x| = 0$  se e solo se  $x = 0$ ;
- (ii)  $|\lambda x| = |\lambda| |x|$  per ogni  $x \in \mathbb{R}^n$  ed ogni  $\lambda \in \mathbb{R}$ ;
- (iii)  $|x + y| \leq |x| + |y|$  per ogni  $x, y \in \mathbb{R}^n$  (disuguaglianza triangolare).

*Dimostrazione.* (i) è la proprietà (b) del prodotto scalare.

(ii) vale perché, per la (e) del prodotto scalare,

$$|\lambda x|^2 = (\lambda x) \cdot (\lambda x) = \lambda^2 (x \cdot x) = \lambda^2 |x|^2,$$

la cui radice quadrata dà la tesi.

(iii) Per provare la disuguaglianza triangolare si applica la disuguaglianza di Cauchy-Schwartz:

$$\begin{aligned} |x + y|^2 &= (x + y) \cdot (x + y) \\ &= x \cdot x + x \cdot y + y \cdot x + y \cdot y \\ &= |x|^2 + 2x \cdot y + |y|^2 \\ &\leq |x|^2 + 2|x \cdot y| + |y|^2 \\ &\leq |x|^2 + 2|x||y| + |y|^2 \\ &= (|x| + |y|)^2. \end{aligned} \quad \square$$

Notiamo che, dalla (ii), si ha che

$$|-x| = |(-1)x| = |-1| |x| = |x| \quad \forall x \in \mathbb{R}^n, \quad (3)$$

un vettore  $x$  e il suo opposto  $-x$  hanno la stessa norma.

Infine una disuguaglianza di uso molto frequente: se  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ , per ciascuna componente  $x_k$  si ha

$$x_k^2 \leq x_1^2 + x_2^2 + \dots + x_n^2,$$

perché nella somma compare il termine  $x_k^2$  più tutti gli altri quadrati; di conseguenza

$$|x_k| \leq |x| \quad \forall x \in \mathbb{R}^n. \quad (4)$$

## 1.2 Funzioni lineari

**Definizione 1.9.** Una funzione  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  si dice *lineare* se valgono:

(i)  $f(x + y) = f(x) + f(y)$  per ogni  $x, y \in \mathbb{R}^n$ ;

(ii)  $f(\lambda x) = \lambda f(x)$  per ogni  $x \in \mathbb{R}^n$  ed ogni  $\lambda \in \mathbb{R}$ . □

Le funzioni lineari vengono anche dette *applicazioni lineari*.

**Proposizione 1.10.** Sia  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  una funzione lineare. Siano  $v_1, v_2, \dots, v_p \in \mathbb{R}^n$  un certo numero  $p$  di vettori, e  $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$  siano  $p$  scalari. Allora

$$f(\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_p v_p) = \lambda_1 f(v_1) + \lambda_2 f(v_2) + \dots + \lambda_p f(v_p) = \sum_{i=1}^p \lambda_i f(v_i).$$

*Dimostrazione.* Per induzione sul numero  $p$  di vettori e scalari.

Passo base: per  $p = 1$  la tesi è  $f(\lambda_1 v_1) = \lambda_1 f(v_1)$ , vera per la proprietà (ii) delle funzioni lineari.

Passo induttivo: supponiamo che la tesi sia dimostrata fino a  $p - 1$ , e proviamo che allora vale anche per  $p$ . Per la proprietà (i) delle funzioni lineari, applicata per  $x = \lambda_1 v_1 + \dots + \lambda_{p-1} v_{p-1}$  e  $y = \lambda_p v_p$ ,

$$f(\lambda_1 v_1 + \dots + \lambda_p v_p) = f(\lambda_1 v_1 + \dots + \lambda_{p-1} v_{p-1}) + f(\lambda_p v_p).$$

Per ipotesi induttiva, trattandosi di  $p - 1$  vettori e  $p - 1$  scalari,

$$f(\lambda_1 v_1 + \dots + \lambda_{p-1} v_{p-1}) = \lambda_1 f(v_1) + \dots + \lambda_{p-1} f(v_{p-1}),$$

mentre, per la (ii),  $f(\lambda_p v_p) = \lambda_p f(v_p)$ , e quindi

$$f(\lambda_1 v_1 + \dots + \lambda_p v_p) = \lambda_1 f(v_1) + \dots + \lambda_{p-1} f(v_{p-1}) + \lambda_p f(v_p),$$

che è la tesi per  $p$ . Per induzione, la tesi è provata per ogni  $p \in \mathbb{N}$ . □

La proposizione che segue dà una *caratterizzazione* delle funzioni lineari di  $\mathbb{R}^n$  in  $\mathbb{R}^m$ : le funzioni lineari sono, di fatto, i prodotti righe per colonne con matrici.

**Proposizione 1.11.** Una funzione  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  è lineare se e solo se è rappresentata da una matrice, cioè se e solo se esiste una matrice  $A$ , matrice a  $m$  righe e  $n$  colonne, tale che

$$f(x) = Ax \quad \forall x \in \mathbb{R}^n.$$

In tal caso, la  $k$ -esima colonna della matrice  $A$  è uguale al vettore  $f(e_k) \in \mathbb{R}^m$ , dove  $e_k = (0, \dots, 0, 1, 0, \dots, 0)$  è il  $k$ -esimo elemento della base canonica di  $\mathbb{R}^n$ .

*Dimostrazione.*  $\Rightarrow$ ) Sia  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  una funzione lineare. Se  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ , come osservato in (1), possiamo scrivere

$$x = \sum_{k=1}^n x_k e_k.$$

Per la proposizione 1.10,

$$f(x) = f(x_1 e_1 + \dots + x_n e_n) = x_1 f(e_1) + \dots + x_n f(e_n).$$

Ora,  $f(e_1)$  è un vettore di  $\mathbb{R}^m$ , cioè una colonna di  $m$  numeri reali,

$$f(e_1) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{m1} \end{pmatrix} \in \mathbb{R}^m,$$

per certi  $a_{jk} \in \mathbb{R}$ ,  $j = 1, \dots, m$ . Similmente  $f(e_2), f(e_3), \dots$  sono tutti vettori di  $\mathbb{R}^m$ ,

$$f(e_k) = \begin{pmatrix} a_{1k} \\ a_{2k} \\ a_{3k} \\ \vdots \\ a_{mk} \end{pmatrix} \in \mathbb{R}^m, \quad k = 1, \dots, n,$$

per certi  $a_{jk} \in \mathbb{R}$ . Quindi

$$\begin{aligned} f(x) &= x_1 f(e_1) + \dots + x_n f(e_n) \\ &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= Ax. \end{aligned}$$

Abbiamo provato che  $f(x) = Ax$  per ogni  $x \in \mathbb{R}^n$ , dove  $A$  è la matrice le cui colonne sono i vettori  $f(e_1), f(e_2), \dots, f(e_n)$ .

$\Leftrightarrow$  Supponiamo che  $f$  sia rappresentata da una certa matrice  $A$ ,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in \text{Mat}(m \times n),$$

cioè supponiamo che sia  $f(x) = Ax$  per ogni  $x \in \mathbb{R}^n$ . Dobbiamo mostrare che  $f$  è lineare. Svolgendo il prodotto riga per colonna, il vettore  $f(x + y) = A(x + y)$  è

$$\begin{aligned} A(x + y) &= \begin{pmatrix} a_{11}(x_1 + y_1) + a_{12}(x_2 + y_2) + \dots + a_{1n}(x_n + y_n) \\ a_{21}(x_1 + y_1) + a_{22}(x_2 + y_2) + \dots + a_{2n}(x_n + y_n) \\ \dots \\ a_{m1}(x_1 + y_1) + a_{m2}(x_2 + y_2) + \dots + a_{mn}(x_n + y_n) \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} + \begin{pmatrix} a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \\ a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n \\ \dots \\ a_{m1}y_1 + a_{m2}y_2 + \dots + a_{mn}y_n \end{pmatrix} \\ &= Ax + Ay, \end{aligned}$$

e la proprietà (i) delle funzioni lineari è soddisfatta. Analogamente,  $f(\lambda x) = A(\lambda x)$  è

$$\begin{aligned} A(\lambda x) &= \begin{pmatrix} a_{11} \lambda x_1 + a_{12} \lambda x_2 + \dots + a_{1n} \lambda x_n \\ a_{21} \lambda x_1 + a_{22} \lambda x_2 + \dots + a_{2n} \lambda x_n \\ \dots \\ a_{m1} \lambda x_1 + a_{m2} \lambda x_2 + \dots + a_{mn} \lambda x_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) \\ \lambda(a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n) \\ \dots \\ \lambda(a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n) \end{pmatrix} \\ &= \lambda \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} \\ &= \lambda Ax, \end{aligned}$$

e anche la (ii) è soddisfatta. Dunque  $f$  è lineare. Infine, svolgendo il prodotto riga per colonna, si vede che  $f(e_k) = Ae_k =$  la  $k$ -esima colonna della matrice  $A$ .  $\square$

**Osservazione 1.12.** Se  $m = 1$ , la matrice  $A$  si riduce ad una matrice riga,  $A = (a_1 \ a_2 \ \dots \ a_n)$ , e il prodotto righe per colonne  $Ax = a_1x_1 + \dots + a_nx_n$  è uguale al prodotto scalare di  $x$  per il vettore di componenti  $(a_1, a_2, \dots, a_n)$ .  $\square$

## 2 Elementi di topologia in $\mathbb{R}^n$

**Definizione 2.1.** Sia  $x_0 \in \mathbb{R}^n$ ,  $r > 0$ . Si chiama *palla centrata in  $x_0$  di raggio  $r$*  l'insieme

$$B(x_0, r) := \{x \in \mathbb{R}^n : |x - x_0| < r\}. \quad \square$$

$B(x_0, r)$  è l'insieme dei punti dello spazio  $\mathbb{R}^n$  che distano meno di  $r$  dal punto  $x_0$ . In dimensione  $n = 1$ ,  $B(x_0, r)$  è un segmento, l'intervallo aperto  $(x_0 - r, x_0 + r)$ . In dimensione  $n = 2$ ,  $B(x_0, r)$  è un disco, la regione interna alla circonferenza centrata in  $x_0$  e raggio  $r$ . In dimensione  $n = 3$ ,  $B(x_0, r)$  è una palla, la regione interna alla sfera centrata in  $x_0$  e raggio  $r$ .

**Definizione 2.2.** Un sottoinsieme  $D \subseteq \mathbb{R}^n$  si dice *aperto* se per ogni  $x \in D$  esiste  $r > 0$  tale che  $B(x, r) \subseteq D$ .

Un sottoinsieme  $D \subseteq \mathbb{R}^n$  si dice *chiuso* se il suo complementare  $\mathbb{R}^n \setminus D$  è aperto.

Per definizione,  $\mathbb{R}^n$  e l'insieme vuoto  $\emptyset$  sono aperti, e quindi sono anche chiusi (essendo l'uno il complementare dell'altro).

Un sottoinsieme  $D \subseteq \mathbb{R}^n$  si dice *limitato* se è contenuto in una palla, cioè se esiste  $R > 0$  tale che  $D \subseteq B(0, R)$ .

Un sottoinsieme  $D \subseteq \mathbb{R}^n$  si dice *compatto* se ha la seguente proprietà: per ogni successione  $\{x_k\}_k \subseteq D$ , esiste una sua sottosuccessione  $\{x_{k_j}\}_j$  e un punto  $\bar{x} \in D$  tali che  $x_{k_j}$  converge a  $\bar{x}$  per  $j \rightarrow \infty$ .  $\square$

**Definizione 2.3.** Sia  $D \subseteq \mathbb{R}^n$ . Un punto  $x_0 \in \mathbb{R}^n$  si dice *interno a  $D$*  se esiste  $r > 0$  tale che  $B(x_0, r) \subseteq D$ .

Un punto  $x_0 \in \mathbb{R}^n$  si dice *esterno a  $D$*  se è interno a  $\mathbb{R}^n \setminus D$ .

Un punto  $x_0 \in \mathbb{R}^n$  si dice *di frontiera per  $D$*  se non è né interno né esterno a  $D$ .

Un punto  $x_0 \in \mathbb{R}^n$  si dice *di accumulazione per  $D$*  se esiste una successione  $\{x_k\}_k \subseteq D \setminus \{x_0\}$  che converge a  $x_0$  (nota:  $x_k \neq x_0 \ \forall k$ ).

Si dice *interno di  $D$* , e si indica  $\overset{\circ}{D}$  oppure  $\text{int}(D)$ , l'insieme dei punti interni a  $D$ ,

$$\overset{\circ}{D} = \text{int}(D) = \{x \in \mathbb{R}^n : x \text{ è punto interno a } D\}.$$

Si dice *chiusura di  $D$* , e si indica  $\overline{D}$  oppure  $\text{cl}(D)$ , l'insieme

$$\overline{D} = \text{cl}(D) = \left\{ x \in \mathbb{R}^n : x \text{ è il limite di una successione } \{x_k\}_k \subseteq D \right\}.$$

Si dice *frontiera di D*, e si indica  $\partial D$  oppure  $\text{fr}(D)$ , l'insieme dei punti di frontiera di  $D$ ,

$$\partial D = \text{fr}(D) = \{x \in \mathbb{R}^n : x \text{ è punto di frontiera per } D\}. \quad \square$$

**Proposizione 2.4.** Per ogni  $D \subseteq \mathbb{R}^n$ ,

- $\text{int}(D) \subseteq D \subseteq \text{cl}(D)$ .
- $D$  è aperto se e solo se  $D = \text{int}(D)$ .
- $D$  è chiuso se e solo se  $D = \text{cl}(D)$ .
- $\text{cl}(D)$  è l'unione di  $D$  e dei suoi punti di accumulazione.

*Dimostrazione.* La dimostrazione non è in programma. Comunque non è troppo difficile, e si può provare a farla per esercizio.  $\square$

**Teorema 2.5** (Heine-Borel). Sia  $D \subseteq \mathbb{R}^n$ .  $D$  è compatto se e solo se  $D$  è chiuso e limitato.

*Dimostrazione.* La dimostrazione non è in programma. Si può provare a farla per esercizio. Si trova sul testo a pag. 109.  $\square$

### 3 Limiti e continuità

Consideriamo una successione  $\{x^{(j)}\}_j$  di vettori di  $\mathbb{R}^n$  (indichiamo in alto l'indice della successione perché in basso indichiamo le componenti di ciascun vettore). La successione è dunque  $\{x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}, x^{(5)}, x^{(6)}, \dots\}$ , e

$$x^{(1)} = \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} x_1^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix}, \quad \dots \quad x^{(j)} = \begin{pmatrix} x_1^{(j)} \\ \vdots \\ x_n^{(j)} \end{pmatrix}, \quad \dots, \quad j = 1, 2, 3, \dots$$

**Definizione 3.1.** Diciamo che una successione  $\{x^{(j)}\}_j$  in  $\mathbb{R}^n$  converge al vettore  $\bar{x}$  se la successione numerica  $\{|x^{(j)} - \bar{x}|\}_j$  converge a 0, cioè se: per ogni  $\varepsilon > 0$  esiste un intero  $\bar{j}$  tale che

$$|x^{(j)} - \bar{x}| < \varepsilon \quad \forall j \geq \bar{j}.$$

In simboli:  $\lim_{(j \rightarrow \infty)} x^{(j)} = \bar{x}$ , o anche  $x^{(j)} \rightarrow \bar{x}$  per  $j \rightarrow \infty$ .  $\square$

Notiamo che questa definizione e quella di convergenza di una successione numerica sono formalmente uguali: l'unica differenza è che in  $\mathbb{R}^n$  compare la norma di un vettore invece che il valore assoluto di un numero.

La proposizione che segue mostra che questa nozione di convergenza equivale alla convergenza "componente per componente":

**Proposizione 3.2.** Sia  $\{x^{(j)}\}_j$  una successione in  $\mathbb{R}^n$  e  $\bar{x} \in \mathbb{R}^n$ . Si ha:

$$x^{(j)} \rightarrow \bar{x} \quad \text{se e solo se} \quad x_1^{(j)} \rightarrow \bar{x}_1, \quad x_2^{(j)} \rightarrow \bar{x}_2, \quad \dots, \quad x_n^{(j)} \rightarrow \bar{x}_n$$

per  $j \rightarrow \infty$ .

*Dimostrazione.*  $\Rightarrow$ ) Sia  $k \in \{1, \dots, n\}$ . Dalla (4) si ha

$$0 \leq |x_k^{(j)} - \bar{x}_k| \leq |x^{(j)} - \bar{x}|;$$

per ipotesi  $|x^{(j)} - \bar{x}| \rightarrow 0$  per  $j \rightarrow \infty$ , quindi (teorema dei carabinieri) anche  $|x_k^{(j)} - \bar{x}_k| \rightarrow 0$  per  $j \rightarrow \infty$ .

$\Leftarrow$ ) Per definizione di norma euclidea,

$$|x^{(j)} - \bar{x}| = \left( \sum_{k=1}^n |x_k^{(j)} - \bar{x}_k|^2 \right)^{1/2}.$$

Per ipotesi, ciascun termine  $|x_k^{(j)} - \bar{x}_k|^2 \rightarrow 0$  per  $j \rightarrow \infty$ . La radice quadrata è una funzione continua, dunque

$$\begin{aligned} \lim_{j \rightarrow \infty} |x^{(j)} - \bar{x}| &= \lim_{j \rightarrow \infty} \sqrt{\sum_{k=1}^n |x_k^{(j)} - \bar{x}_k|^2} \\ &= \sqrt{\lim_{j \rightarrow \infty} \sum_{k=1}^n |x_k^{(j)} - \bar{x}_k|^2} \\ &= \sqrt{\sum_{k=1}^n \left( \lim_{j \rightarrow \infty} |x_k^{(j)} - \bar{x}_k|^2 \right)} \\ &= \sqrt{\sum_{k=1}^n 0} \\ &= 0. \end{aligned} \quad \square$$

Sia  $D \subseteq \mathbb{R}^n$  e  $f: D \rightarrow \mathbb{R}^m$  una funzione. Si dice che  $f$  è una funzione di  $n$  variabili reali perché gli elementi del dominio  $D$  sono vettori  $x = (x_1, \dots, x_n)$  con  $n$  componenti. Si dice che  $f$  è una funzione a valori vettoriali perché il suo codominio è lo spazio  $\mathbb{R}^m$ : per ogni  $x \in D$ ,  $f(x)$  è un vettore con  $m$  componenti,

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}.$$

Ciascuna  $f_k$  è una funzione a valori in  $\mathbb{R}$ ,  $f_k: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ .

**Definizione 3.3.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Diciamo che  $f$  è continua nel punto  $\bar{x} \in D$  se per ogni  $\varepsilon > 0$  esiste  $\delta > 0$  tale che

$$|f(x) - f(\bar{x})| < \varepsilon \quad \forall x \in D \text{ tale che } |x - \bar{x}| < \delta.$$

In simboli si indica:

$$\lim_{|x - \bar{x}| \rightarrow 0} |f(x) - f(\bar{x})| = 0,$$

oppure  $\lim_{|x - \bar{x}| \rightarrow 0} f(x) = f(\bar{x})$ , o anche  $f(x) \rightarrow f(\bar{x})$  per  $x \rightarrow \bar{x}$ .

$f$  si dice continua in  $D$  se è continua in ogni punto  $x \in D$ . □

Notiamo che la definizione di continuità si scrive come quella studiata in Analisi 1, pur di sostituire valori assoluti di numeri con norme di vettori.

In modo analogo a quanto visto per le successioni di vettori, la continuità di una funzione a valori vettoriali equivale alla continuità di tutte le sue funzioni componenti:

**Proposizione 3.4.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $\bar{x} \in D$ .  $f$  è continua in  $\bar{x}$  se e solo se  $f_k$  è continua in  $\bar{x}$ , per ogni  $k = 1, \dots, m$ .

*Dimostrazione.*  $\Rightarrow$ ) Sia  $k \in \{1, \dots, m\}$ , e sia  $\varepsilon > 0$ . Dalla (4) si ha

$$|f_k(x) - f_k(\bar{x})| \leq |f(x) - f(\bar{x})|;$$

per ipotesi  $f$  è continua in  $\bar{x}$ , dunque esiste  $\delta > 0$  tale che  $|f(x) - f(\bar{x})| < \varepsilon$  non appena  $|x - \bar{x}| < \delta$ , e di conseguenza

$$|f_k(x) - f_k(\bar{x})| \leq |f(x) - f(\bar{x})| < \varepsilon \quad \forall x \in D \text{ tale che } |x - \bar{x}| < \delta,$$

cioè  $f_k$  è continua in  $\bar{x}$ .

$\Leftarrow$ ) Sia  $\varepsilon > 0$ . Per ipotesi, ciascuna funzione componente  $f_k$  è continua in  $\bar{x}$ , quindi ammette un suo  $\delta_k > 0$  tale che

$$|f_k(x) - f_k(\bar{x})| < \varepsilon / \sqrt{m} \quad \forall x \in D \text{ tale che } |x - \bar{x}| < \delta_k.$$

Sia  $\delta := \min\{\delta_1, \delta_2, \dots, \delta_m\}$ . Allora per  $|x - \bar{x}| < \delta$  tutte le disuguaglianze sono soddisfatte, e perciò, dalla definizione di norma euclidea,

$$\begin{aligned} |f(x) - f(\bar{x})| &= \left( \sum_{k=1}^m |f_k(x) - f_k(\bar{x})|^2 \right)^{1/2} \\ &\leq \left[ \sum_{k=1}^m \left( \frac{\varepsilon}{\sqrt{m}} \right)^2 \right]^{1/2} \\ &= \left( \frac{\varepsilon^2}{m} + \frac{\varepsilon^2}{m} + \dots + \frac{\varepsilon^2}{m} \right)^{1/2} \\ &= \varepsilon \end{aligned}$$

per ogni  $x \in D$  tale che  $|x - \bar{x}| < \delta$ . □

Fra le funzioni  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ , ci sono tre casi particolarmente interessanti:

- $n = 1$ , funzione vettoriale di una sola variabile reale,  $f: D \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ . La funzione è detta una *curva* in  $\mathbb{R}^m$ . Un esempio di curva piana,  $m = 2$ , è la curva

$$f: [0, 2\pi] \subset \mathbb{R} \rightarrow \mathbb{R}^2, \quad f(t) = \begin{pmatrix} x_0 + r \cos t \\ y_0 + r \sin t \end{pmatrix},$$

che percorre in senso antiorario una circonferenza di centro nel punto  $(x_0, y_0) \in \mathbb{R}^2$  e raggio  $r > 0$ . Un altro esempio di curva in  $\mathbb{R}^m$  è

$$f: [0, 1] \subset \mathbb{R} \rightarrow \mathbb{R}^m, \quad f(t) = x + t(y - x),$$

il segmento (porzione di retta) in  $\mathbb{R}^m$  di estremi nei punti  $x$  e  $y$  di  $\mathbb{R}^m$ . (Spesso la variabile del dominio è indicata con la lettera  $t$ , pensando a  $f$  come alla legge oraria che descrive il moto di una particella che si muove nello spazio  $\mathbb{R}^m$ : all'istante di tempo  $t$  la particella si trova nella posizione  $f(t) \in \mathbb{R}^m$ ).

- $m = 1$ , funzioni scalari, cioè a valori reali. Sono quelle di cui ci occuperemo più in dettaglio. Ad ogni punto  $x \in D \subseteq \mathbb{R}^n$  è associato il numero  $f(x) \in \mathbb{R}$ .
- $n = m$ , dominio e codominio hanno la stessa dimensione. La funzione si dice *campo*. (Pensando ad un campo di forze: nel punto  $x$  dello spazio  $\mathbb{R}^n$  agisce una forza espressa dal vettore  $f(x) \in \mathbb{R}^n$ . È questo il caso del campo gravitazionale, del campo elettromagnetico, ecc.).

Come quanto visto in dimensione 1 nel corso di Analisi 1, la continuità è equivalente alla “continuità per successioni”, e la dimostrazione è identica a quella del caso uni-dimensionale:

**Teorema 3.5** (Teorema ponte). *Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  e  $x_0 \in D$ . Allora  $f$  è continua in  $x_0$  se e solo se  $\lim_{(k \rightarrow \infty)} f(x_k) = f(x_0)$  per ogni successione  $\{x_k\} \subseteq D$  che converge a  $x_0$ .*

*Dimostrazione.*  $\Rightarrow$ ) Sia  $\varepsilon > 0$ . Per ipotesi, esiste  $\delta > 0$  tale che  $|f(x) - f(x_0)| < \varepsilon$  per ogni  $x \in D$  tale che  $|x - x_0| < \delta$ . Sia  $\{x_k\}$  una successione in  $D$  che converge a  $x_0$ . Per definizione di limite di una successione, esiste un intero  $\bar{k}$  tale che  $|x_k - x_0| < \delta$  per ogni  $k \geq \bar{k}$ . Dunque

$$|f(x_k) - f(x_0)| < \varepsilon \quad \forall k \geq \bar{k}.$$

Questo prova che  $f(x_k) \rightarrow f(x_0)$  per  $k \rightarrow \infty$ .

$\Leftarrow$ ) Sia  $\varepsilon > 0$ . Supponiamo, per assurdo, che  $f$  non sia continua in  $x_0$ . Allora non è vero che per ogni  $\varepsilon > 0$  esiste  $\delta > 0$  come richiesto dalla definizione di continuità, e dunque c'è un certo  $\varepsilon_0 > 0$  che non ammette alcun  $\delta > 0$  come richiesto dalla definizione. In altre parole, per ogni  $\delta > 0$  la disuguaglianza

$$|f(x) - f(x_0)| < \varepsilon_0 \quad \forall x \in D \text{ t.c. } |x - x_0| < \delta$$

è falsa, cioè esiste un certo punto  $x$  (dipendente da  $\delta$ ) tale che  $|f(x) - f(x_0)| \geq \varepsilon_0$ , nonostante  $|x - x_0|$  sia  $< \delta$ . Questo significa che per ogni intero  $k$ , scegliendo  $\delta = 1/k$ , esiste un punto  $x_k \in D$  tale che

$$|f(x_k) - f(x_0)| \geq \varepsilon_0, \quad |x_k - x_0| < 1/k.$$

In questo modo abbiamo costruito una successione in  $D$  che converge a  $x_0$  tale che  $f(x_k)$  non converge a  $f(x_0)$ , andando in contraddizione con l'ipotesi. Di conseguenza,  $f$  non può che essere continua in  $x_0$ . □

## 4 Funzioni lipschitziane

**Definizione 4.1.** Una funzione  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  si dice *lipschitziana* sul dominio  $D$  se esiste un numero  $L > 0$  tale che

$$|f(x) - f(y)| \leq L|x - y| \quad \forall x, y \in D. \quad \square$$

Prima di proseguire con funzioni di più variabili, vediamo due esempi semplici in dimensione 1.

**Esempio 4.2.** La funzione  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = \sin x$  è lipschitziana su  $\mathbb{R}$ , perché, se  $x, y \in \mathbb{R}$ , per il teorema di Lagrange esiste un punto  $\xi$  interno al segmento di estremi  $x$  e  $y$  tale che

$$f(x) - f(y) = f'(\xi)(x - y),$$

da cui

$$|f(x) - f(y)| = |f'(\xi)||x - y| = |\cos \xi||x - y| \leq |x - y|,$$

essendo  $|\cos \xi| \leq 1$ . Quindi  $f$  è lipschitziana, con  $L = 1$ . □

**Esempio 4.3.** La funzione  $f: [0, \infty) \subset \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = \sqrt{x}$  non è lipschitziana su  $[0, \infty)$ . Infatti, se lo fosse, esisterebbe una costante  $L > 0$  per cui  $|f(x) - f(y)| \leq L|x - y|$  per ogni  $x, y \geq 0$ . In particolare, per  $y = 0$  e  $x > 0$ , sarebbe  $|f(x) - f(0)| \leq L|x - 0|$ , cioè

$$\sqrt{x} \leq Lx \quad \forall x > 0,$$

da cui  $1 \leq L\sqrt{x}$  per ogni  $x > 0$ , assurdo (per esempio, prendere  $x = 1/(4L^2)$ , o farne il limite per  $x \rightarrow 0$ ). □

**Proposizione 4.4.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  lipschitziana in  $D$ . Allora  $f$  è continua in  $D$ .

*Dimostrazione.* Per ipotesi, esiste  $L > 0$  tale che  $|f(x) - f(y)| \leq L|x - y|$  per ogni  $x, y \in D$ . Sia  $x_0 \in D$ , e sia  $\varepsilon > 0$ . Scegliendo  $\delta = \varepsilon/L$ , si ha che

$$|f(x) - f(x_0)| \leq L|x - x_0| < L\delta = \varepsilon \quad \forall x \in D \text{ tale che } |x - x_0| < \delta.$$

Questo prova che  $f$  è continua in  $x_0$ , per ogni  $x_0 \in D$ . □

**Proposizione 4.5.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ .  $f$  è lipschitziana se e solo se ogni sua componente  $f_k$  è lipschitziana.

*Dimostrazione.*  $\Rightarrow$ ) Sia  $f$  lipschitziana con costante  $L > 0$ . Per ogni  $k \in \{1, \dots, m\}$ , dalla (4) si ha

$$|f_k(x) - f_k(y)| \leq |f(x) - f(y)| \leq L|x - y|,$$

quindi anche  $f_k$  è lipschitziana.

$\Leftarrow$ ) Supponiamo che ogni funzione componente  $f_k$  sia lipschitziana, ciascuna con la sua costante  $L_k > 0$ . Dalla definizione di norma euclidea,

$$\begin{aligned} |f(x) - f(y)| &= \left( \sum_{k=1}^m |f_k(x) - f_k(y)|^2 \right)^{1/2} \\ &\leq \left( \sum_{k=1}^m L_k^2 |x - y|^2 \right)^{1/2} \\ &= \left( |x - y|^2 \sum_{k=1}^m L_k^2 \right)^{1/2} \\ &= L|x - y|, \end{aligned}$$

dove  $L = (L_1^2 + \dots + L_m^2)^{1/2}$ . □

## 4.1 Alcuni esempi importanti di funzioni lipschitziane

**Proposizione 4.6.** *Le seguenti funzioni sono lipschitziane.*

- (a) *Traslazione in direzione  $\xi$ .  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $f(x) = x + \xi$ , dove  $\xi \in \mathbb{R}^n$ .*
- (b) *Omotetia di rapporto  $\lambda$ .  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $f(x) = \lambda x$ , dove  $\lambda \in \mathbb{R}$ .*
- (c) *Norma.  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = |x|$ .*
- (d) *Prodotto scalare per un vettore  $\xi$ .  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = x \cdot \xi$ , dove  $\xi \in \mathbb{R}^n$ .*
- (e)  *$k$ -esima proiezione.  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = x_k$ , dove  $k = 1, \dots, n$ .*
- (f) *Funzioni lineari, cioè prodotto righe per colonne per una matrice.  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $f(x) = Ax$ , dove  $A \in \text{Mat}(m \times n)$ .*

*Dimostrazione.* (a)  $|f(x) - f(y)| = |x + \xi - y - \xi| = |x - y|$ , lipschitziana con  $L = 1$ .

(b)  $|f(x) - f(y)| = |\lambda x - \lambda y| = |\lambda(x - y)| = |\lambda||x - y|$ , lipschitziana con  $L = |\lambda|$ .

(c) Scriviamo  $x$  come  $x - y + y$  e applichiamo la disuguaglianza triangolare:

$$|x| = |x - y + y| \leq |x - y| + |y|,$$

da cui  $|x| - |y| \leq |x - y|$ . Ora scriviamo  $y$  come  $y - x + x$  e applichiamo di nuovo la disuguaglianza triangolare:

$$|y| = |y - x + x| \leq |y - x| + |x| = |x - y| + |x|,$$

da cui  $|x| - |y| \geq -|x - y|$  (notiamo che  $|y - x| = |-(x - y)| = |x - y|$ , come osservato in (3)). Abbiamo dunque

$$-|x - y| \leq |x| - |y| \leq |x - y|,$$

cioè

$$||x| - |y|| \leq |x - y|.$$

Questo significa che  $|f(x) - f(y)| \leq |x - y|$ , e la norma è una funzione lipschitziana con costante  $L = 1$ .

(d) Dalla disuguaglianza di Cauchy-Schwartz,  $|f(x) - f(y)| = |\xi \cdot x - \xi \cdot y| = |\xi \cdot (x - y)| \leq |\xi| |x - y|$ , lipschitziana con  $L = |\xi|$ .

(e)  $x_k - y_k$  è la  $k$ -esima componente del vettore  $x - y$ , dunque, dall'osservazione (4),  $|f(x) - f(y)| = |x_k - y_k| \leq |x - y|$ , lipschitziana con  $L = 1$ .

(f) Sia  $f(x) = Ax$ , con  $A$  matrice a  $m$  righe e  $n$  colonne,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}.$$

Nella proposizione 1.11 abbiamo dimostrato che  $f$  è lineare. Quindi, per dimostrare che  $f$  è lipschitziana, basta provare che esiste  $L > 0$  tale che

$$|Ax| \leq L|x| \quad \forall x \in \mathbb{R}^n. \tag{5}$$

Infatti, una volta che la (5) è stata dimostrata, per la linearità di  $f$  si ha

$$|f(x) - f(y)| = |Ax - Ay| = |A(x - y)| \leq L|x - y| \quad \forall x, y \in \mathbb{R}^n,$$

cioè  $f$  è lipschitziana. Dunque dimostriamo (5). Sia  $x \in \mathbb{R}^n$ . Indichiamo  $r_j$  la  $j$ -esima riga di  $A$ ,

$$r_j = (a_{j1} \ a_{j2} \ a_{j3} \ \dots \ a_{jn}).$$

Per definizione di prodotto righe per colonne,

$$Ax = \begin{pmatrix} r_1x \\ r_2x \\ \vdots \\ r_mx \end{pmatrix},$$

dove

$$r_jx = (a_{j1} \ a_{j2} \ \dots \ a_{jn}) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n = \begin{pmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = v_j \cdot x,$$

indicando  $v_j$  il vettore di componenti  $(a_{j1}, a_{j2}, \dots, a_{jn})$  ( $v_j$  è il vettore che si ottiene semplicemente “mettendo in piedi” la riga  $r_j$ ). Per la disuguaglianza di Cauchy-Schwartz,

$$|r_jx| = |v_j \cdot x| \leq |v_j||x|.$$

Allora, per definizione di norma euclidea,

$$\begin{aligned} |Ax| &= \sqrt{(r_1x)^2 + (r_2x)^2 + \dots + (r_mx)^2} \\ &\leq \sqrt{|v_1|^2|x|^2 + |v_2|^2|x|^2 + \dots + |v_m|^2|x|^2} \\ &= \sqrt{|x|^2(|v_1|^2 + |v_2|^2 + \dots + |v_m|^2)} \\ &= L|x|, \quad L = \sqrt{|v_1|^2 + |v_2|^2 + \dots + |v_m|^2}. \end{aligned}$$

Abbiamo dimostrato che vale (5), con

$$L = \sqrt{|v_1|^2 + \dots + |v_m|^2} = \left( \sum_{j=1}^m |v_j|^2 \right)^{1/2} = \left( \sum_{j=1}^m \sum_{k=1}^n |a_{jk}|^2 \right)^{1/2}.$$

Quindi (f) è dimostrata. □

Osserviamo che l'omotetia di rapporto  $\lambda$  è rappresentata dalla matrice quadrata diagonale

$$A = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix},$$

la funzione  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = \xi \cdot x$  è rappresentata dalla matrice riga

$$A = (\xi_1 \ \xi_2 \ \dots \ \xi_n),$$

e la  $k$ -esima proiezione  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = x_k$  è del tipo (d) con  $\xi = e_k$ , ed è rappresentata dalla matrice riga

$$A = (0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0),$$

1 nella  $k$ -esima colonna, 0 in tutte le altre. Dunque (b),(d),(e) sono casi particolari di (f).

Il numero  $L$  che compare nella dimostrazione di (f) definisce la norma della matrice  $A$ :

**Definizione 4.7.** Sia  $A \in \text{Mat}(m \times n)$  una matrice a  $m$  righe e  $n$  colonne, di coefficienti  $a_{i,j}$ . Si dice *norma della matrice  $A$*  il numero

$$|A| = \left( \sum_{j=1}^m \sum_{k=1}^n |a_{jk}|^2 \right)^{1/2}. \quad \square$$

Nella dimostrazione di (f) abbiamo quindi provato che

$$|Ax| \leq |A||x| \quad \forall x \in \mathbb{R}^n, \quad \forall A \in \text{Mat}(m \times n). \quad (6)$$

## 5 Ancora sulla continuità

**Proposizione 5.1** (Composizione di funzioni continue). *Consideriamo le funzioni*

$$f: A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad g: B \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^p.$$

Definiamo

$$D = \{x \in A : f(x) \in B\},$$

e supponiamo che  $x_0 \in D$ . Il punto  $y_0 = f(x_0)$  appartiene a  $B$  per definizione di  $D$ . Sia  $F$  la funzione composta

$$F: D \rightarrow \mathbb{R}^p, \quad F(x) = g(f(x)).$$

- 1) Se  $f$  è continua in  $x_0$  e  $g$  è continua in  $y_0$ , allora  $F$  è continua in  $x_0$ .
- 2) Se  $f$  è continua in  $D$  e  $g$  è continua in  $B$ , allora  $F$  è continua in  $D$ .

*Dimostrazione.* 1) Sia  $\varepsilon > 0$ . Per ipotesi,  $g$  è continua in  $y_0$ , dunque esiste  $\delta_1 > 0$  tale che

$$|g(y) - g(y_0)| < \varepsilon \quad \forall y \in B \text{ t.c. } |y - y_0| < \delta_1.$$

Per ipotesi,  $f$  è continua in  $x_0$ , quindi (definizione di continuità con  $\delta_1$  nel ruolo di  $\varepsilon$ ) esiste  $\delta > 0$  tale che

$$|f(x) - f(x_0)| < \delta_1 \quad \forall x \in A \text{ t.c. } |x - x_0| < \delta.$$

Dunque per ogni  $x \in D$  tale che  $|x - x_0| < \delta$  si ha  $|g(f(x)) - g(f(x_0))| < \varepsilon$ , e questo prova che  $F$  è continua in  $x_0$ .

2) Sia  $f$  continua in  $D$  e  $g$  continua in  $B$ . Allora quanto provato in 1) vale per ogni punto  $x_0 \in D$ .  $\square$

Come conseguenza della proposizione 5.1, se  $f: A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  è continua in  $A$ , allora

- $\cos(f(x))$ ,  $\sin(f(x))$ ,  $e^{f(x)}$  sono continue in  $A$ ,
- $1/f(x)$  è continua su  $D = \{x \in A : f(x) \neq 0\}$ ,
- $\log(f(x))$  è continua su  $D = \{x \in A : f(x) > 0\}$ ,
- $\sqrt{f(x)}$  è continua su  $D = \{x \in A : f(x) \geq 0\}$ ,
- $\operatorname{tg} f(x)$  è continua su  $D = \{x \in A : f(x) \neq \frac{\pi}{2} + k\pi \ \forall k \in \mathbb{Z}\}$ ,

eccetera.

Inoltre, abbiamo provato che la norma e le funzioni lineari sono lipschitziane, e dunque continue. Di conseguenza, se  $f: A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  è continua, anche la funzione composta

$$|f|: A \rightarrow \mathbb{R}, \quad x \mapsto |f(x)|$$

è continua, e similmente  $x \mapsto Mf(x)$  e  $x \mapsto f(Mx)$  sono continue, per ogni matrice  $M$  (che abbia il corretto numero di righe e di colonne).

Un'altra conseguenza della proposizione 5.1 è il seguente “criterio di non continuità”, molto utile negli esercizi.

**Proposizione 5.2.** *Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $x_0 \in D$ . 1) Se esiste un intervallo  $(-a, a) \subseteq \mathbb{R}$  e una curva continua*

$$\varphi: (-a, a) \rightarrow D$$

*tale che  $\varphi(0) = x_0$ , e  $f(\varphi(t))$  non converge a  $f(\varphi(0)) = f(x_0)$  per  $t \rightarrow 0$ , allora  $f$  non è continua in  $x_0$ .*

*2) Se esistono due curve continue  $\varphi, \psi: (-a, a) \rightarrow D$  tali che  $\varphi(0) = \psi(0) = x_0$ , e*

$$\lim_{t \rightarrow 0} f(\varphi(t)) \neq \lim_{t \rightarrow 0} f(\psi(t)),$$

*allora  $f$  non è continua in  $x_0$ .*

*Dimostrazione.* 1)  $\varphi$  è continua per ipotesi. Se  $f$  fosse continua in  $x_0$ , per la proposizione 5.1 la composizione  $F(t) = f(\varphi(t))$  sarebbe continua in  $t = 0$ , cioè

$$\lim_{t \rightarrow 0} F(t) = F(0).$$

Ma per ipotesi questo limite non vale. Dunque  $f$  non è continua in  $x_0$ .

2) Se  $f$  fosse continua in  $x_0$ , le composizioni  $F(t) = f(\varphi(t))$  e  $G(t) = f(\psi(t))$  sarebbero continue, e quindi avremmo

$$\begin{aligned} \lim_{t \rightarrow 0} f(\varphi(t)) &= \lim_{t \rightarrow 0} F(t) = F(0) = f(\varphi(0)) = f(x_0), \\ \lim_{t \rightarrow 0} f(\psi(t)) &= \lim_{t \rightarrow 0} G(t) = G(0) = f(\psi(0)) = f(x_0). \end{aligned}$$

Ma per ipotesi i due limiti sono diversi, dunque  $f$  non è continua in  $x_0$ . □

Le più semplici curve in  $\mathbb{R}^n$  passanti per un punto  $x_0$  e tali che  $\varphi(0) = x_0$  sono le rette:

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}^n, \quad \varphi(t) = x_0 + tv,$$

$v \in \mathbb{R}^n$ ,  $v \neq 0$ . Per questo, solitamente, andando alla ricerca di una o più curve per applicare il criterio di non continuità della proposizione (5.2) si parte dalle rette. Molto spesso le rette soddisfano le ipotesi del criterio (vedi per esempio l'esercizio 4(c)); altrimenti si cercano altre curve (vedi per esempio l'esercizio 3).

**Proposizione 5.3.** *Siano  $f$  e  $g$  due funzioni continue in  $D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Consideriamo la funzione somma*

$$s: D \rightarrow \mathbb{R}^m, \quad s(x) = f(x) + g(x),$$

e la funzione prodotto scalare

$$p: D \rightarrow \mathbb{R}, \quad p(x) = f(x) \cdot g(x).$$

Se  $f$  e  $g$  sono continue in  $x_0 \in D$ , allora anche  $s$  e  $p$  sono continue in  $x_0$ .

*Dimostrazione.*

$$\begin{aligned} 0 \leq |s(x) - s(x_0)| &= |f(x) + g(x) - f(x_0) - g(x_0)| \\ &\leq |f(x) - f(x_0)| + |g(x) - g(x_0)|. \end{aligned}$$

Entrambi i termini  $|f(x) - f(x_0)|$  e  $|g(x) - g(x_0)| \rightarrow 0$  per  $|x - x_0| \rightarrow 0$ , quindi anche

$$|s(x) - s(x_0)| \rightarrow 0 \quad \text{per } |x - x_0| \rightarrow 0,$$

cioè  $s$  è continua in  $x_0$ .

Usando la disuguaglianza triangolare e poi quella di Cauchy-Schwartz,

$$\begin{aligned} 0 \leq |p(x) - p(x_0)| &= |f(x) \cdot g(x) - f(x_0) \cdot g(x_0)| \\ &= |f(x) \cdot g(x) - f(x_0) \cdot g(x) + f(x_0) \cdot g(x) - f(x_0) \cdot g(x_0)| \\ &= \left| (f(x) - f(x_0)) \cdot g(x) + f(x_0) \cdot (g(x) - g(x_0)) \right| \\ &\leq \left| (f(x) - f(x_0)) \cdot g(x) \right| + \left| f(x_0) \cdot (g(x) - g(x_0)) \right| \\ &\leq |f(x) - f(x_0)| |g(x)| + |f(x_0)| |g(x) - g(x_0)|. \end{aligned}$$

Quando  $|x - x_0| \rightarrow 0$ ,  $g(x) \rightarrow g(x_0)$ , la norma è una funzione continua<sup>1</sup>, e dunque  $|g(x)| \rightarrow |g(x_0)|$ . Dunque entrambi i termini  $|f(x) - f(x_0)| |g(x)|$  e  $|f(x_0)| |g(x) - g(x_0)|$  sono il prodotto di una funzione limitata e di una infinitesima per  $|x - x_0| \rightarrow 0$ , e quindi il loro limite è 0. Di conseguenza

$$|p(x) - p(x_0)| \rightarrow 0 \quad \text{per } |x - x_0| \rightarrow 0,$$

cioè  $p$  è continua in  $x_0$ . □

<sup>1</sup>Nella proposizione 4.6(c) abbiamo dimostrato anche di più: la norma è lipschitziana.

**Osservazione 5.4.** La proposizione 5.3 applicata a funzioni scalari dà la continuità del prodotto di funzioni continue, dal momento che il prodotto scalare in  $\mathbb{R}$  è il classico prodotto tra numeri reali.  $\square$

Un'altra osservazione utile negli esercizi per provare la continuità di una funzione in un punto:

**Proposizione 5.5.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ . Supponiamo che esistano  $C > 0, R > 0, \alpha > 0$  tali che

$$|f(x) - f(x_0)| \leq C|x - x_0|^\alpha \quad \forall x \in D \text{ t.c. } |x - x_0| < R. \quad (7)$$

Allora  $f$  è continua in  $x_0$ .

*Dimostrazione.* Sia  $\varepsilon > 0$ . Scegliamo

$$\delta = \min\{R, (\varepsilon/C)^{1/\alpha}\}.$$

$\delta > 0$ , e per ogni  $x \in D$  tale che  $|x - x_0| < \delta$  si ha:

$$|f(x) - f(x_0)| \leq C|x - x_0|^\alpha < C\delta^\alpha \leq \varepsilon.$$

Dunque  $f$  è continua in  $x_0$ .  $\square$

Per provare la continuità di una funzione di più variabili, spesso si cerca di ricondursi ad una disuguaglianza del tipo (7). Si noti che una simile condizione è sufficiente, ma non necessaria per avere la continuità: per esempio, la funzione

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 1/\log|x| & \text{per } x \neq 0, \\ 0 & \text{per } x = 0 \end{cases}$$

è continua in  $x = 0$  ma

$$\frac{|f(x) - f(0)|}{|x - 0|^\alpha} = \frac{1}{|\log|x|| |x|^\alpha} \rightarrow \infty \quad \text{per } x \rightarrow 0,$$

per ogni  $\alpha > 0$ , quindi non esiste  $C > 0$  che soddisfi la (7).

Ricapitolando, per gli esercizi ribadiamo che:

- per dimostrare che  $f$  è continua in un punto, è sufficiente provare che la definizione di continuità è soddisfatta, oppure che vale una disuguaglianza del tipo (7);
- per dimostrare che  $f$  non è continua in un punto, è sufficiente trovare una curva come descritto nella proposizione 5.2 per cui la composizione  $f(\varphi(t))$  non ha limite, oppure due curve per le quali  $f(\varphi(t))$  e  $f(\psi(t))$  hanno limiti diversi;
- se la composizione con tutte le rette o con tutte le parabole o con tutte le curve di una certa famiglia soddisfa la condizione  $f(\varphi(t)) \rightarrow f(x_0)$  per  $t \rightarrow 0$ , **non si può trarre nessuna conclusione**:  $f$  potrebbe essere sia continua che non continua nel punto  $x_0$ . In tal caso, bisogna proseguire l'analisi, o trovando una o più curve che permettano l'applicazione del criterio di non continuità (proposizione 5.2), o dimostrando che la definizione di continuità (o una disuguaglianza tipo (7)) è soddisfatta.

Concludiamo con un importante teorema per funzioni scalari.

**Teorema 5.6** (Teorema di Weierstrass). Sia  $D \subseteq \mathbb{R}^n$  un insieme compatto, e sia  $f: D \rightarrow \mathbb{R}$  una funzione scalare continua. Allora esistono  $y, z \in D$  tali che

$$f(y) \leq f(x) \leq f(z) \quad \forall x \in D,$$

cioè  $f$  ha massimo e minimo in  $D$ .

*Dimostrazione.* Consideriamo

$$\alpha := \sup_{x \in D} f(x).$$

$\alpha$  può essere un numero reale oppure  $\alpha = \infty$ . Dimostriamo che, in ogni caso, esiste una successione  $\{x_k\}$  in  $D$  tale che  $f(x_k) \rightarrow \alpha$  per  $k \rightarrow \infty$ . Se  $\alpha = \infty$ , per definizione di “sup” per ogni  $k \in \mathbb{N}$  esiste  $x_k \in D$  tale che  $f(x_k) \geq k$ , e dunque  $\lim_{(k \rightarrow \infty)} f(x_k) = \infty = \alpha$ . Se invece  $\alpha \in \mathbb{R}$ , per definizione di “sup” per ogni  $k \in \mathbb{N}$  esiste  $x_k \in D$  tale che  $f(x_k) \geq \alpha - 1/k$ , e dunque  $\lim_{(k \rightarrow \infty)} f(x_k) = \alpha$ . Abbiamo provato l’esistenza della successione voluta in entrambi i casi.

$\{x_k\}$  è successione in  $D$ . Per definizione di compattezza, esiste un punto  $z \in D$  e una sottosuccessione  $\{x_{k_j}\}$  tale che

$$x_{k_j} \rightarrow z \quad \text{per } j \rightarrow \infty.$$

$f$  è continua, quindi, per il teorema ponte,

$$f(x_{k_j}) \rightarrow f(z) \quad \text{per } j \rightarrow \infty.$$

D’altra parte,  $f(x_k) \rightarrow \alpha$  per  $k \rightarrow \infty$ , e quindi anche la sua sottosuccessione:  $f(x_{k_j}) \rightarrow \alpha$  per  $j \rightarrow \infty$ . Per l’unicità del limite,  $f(z) = \alpha$ . Poiché  $\alpha$  è il sup,

$$f(x) \leq \alpha = f(z) \quad \forall x \in D,$$

e inoltre  $\alpha < \infty$  perché  $f(z) \in \mathbb{R}$ .

L’esistenza del minimo si prova in modo analogo. □

## 6 Derivate parziali

**Definizione 6.1.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione scalare,  $D$  un aperto di  $\mathbb{R}^n$  e  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n) \in D$ . Diciamo che esiste la *derivata parziale di  $f$  rispetto a  $x_1$  nel punto  $\bar{x}$*  se la funzione reale di *una* variabile reale

$$t \mapsto f(t, \bar{x}_2, \dots, \bar{x}_n)$$

è derivabile in  $t = \bar{x}_1$ , cioè se esiste il limite del rapporto incrementale

$$\begin{aligned} & \lim_{\substack{t \rightarrow \bar{x}_1 \\ t \in \mathbb{R} \setminus \{\bar{x}_1\}}} \frac{f(t, \bar{x}_2, \dots, \bar{x}_n) - f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)}{t - \bar{x}_1} \\ &= \lim_{\substack{s \rightarrow 0 \\ s \in \mathbb{R} \setminus \{0\}}} \frac{f(\bar{x}_1 + s, \bar{x}_2, \dots, \bar{x}_n) - f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)}{s} \\ &= \lim_{\substack{s \rightarrow 0 \\ s \in \mathbb{R} \setminus \{0\}}} \frac{f(\bar{x} + se_1) - f(\bar{x})}{s}. \end{aligned}$$

In tal caso, questo limite si dice derivata parziale di  $f$  in  $\bar{x}$  rispetto a  $x_1$  e si indica con uno di questi simboli:

$$\partial_{x_1} f(\bar{x}) = \frac{\partial}{\partial x_1} f(\bar{x}) = \frac{\partial f(\bar{x})}{\partial x_1} = D_{x_1} f(\bar{x}) = f_{x_1}(\bar{x}).$$

In modo analogo si definisce la derivata parziale di  $f$  rispetto a  $x_k$  nel punto  $\bar{x}$ ,  $k = 1, \dots, n$  come il limite, se esiste, del rapporto incrementale

$$\begin{aligned} & \lim_{\substack{s \rightarrow 0 \\ s \in \mathbb{R} \setminus \{0\}}} \frac{f(\bar{x} + se_k) - f(\bar{x})}{s} \\ &= \lim_{\substack{s \rightarrow 0 \\ s \in \mathbb{R} \setminus \{0\}}} \frac{f(\bar{x}_1, \dots, \bar{x}_k + s, \dots, \bar{x}_n) - f(\bar{x}_1, \dots, \bar{x}_k, \dots, \bar{x}_n)}{s}, \end{aligned}$$

cioè la derivata della funzione

$$t \mapsto f(\bar{x}_1, \dots, \bar{x}_{k-1}, t, \bar{x}_{k+1}, \dots, \bar{x}_n)$$

in  $t = \bar{x}_k$ , e si indica

$$\partial_{x_k} f(\bar{x}) = \frac{\partial}{\partial x_k} f(\bar{x}) = \frac{\partial f(\bar{x})}{\partial x_k} = D_{x_k} f(\bar{x}) = f_{x_k}(\bar{x}). \quad \square$$

**Definizione 6.2.** Supponiamo che esistano tutte le derivate parziali  $\partial_{x_1}f(x), \dots, \partial_{x_n}f(x)$  di  $f$  nel punto  $x$ . Si chiama *gradiente di  $f$  in  $x$*  il vettore che ha per componenti le derivate parziali di  $f$  in  $x$ , e si indica

$$\nabla f(x) = \text{grad } f(x) = \begin{pmatrix} \partial_{x_1}f(x) \\ \partial_{x_2}f(x) \\ \vdots \\ \partial_{x_n}f(x) \end{pmatrix} \in \mathbb{R}^n.$$

Il simbolo  $\nabla$  si chiama “nabla”. □

Se tutte le derivate parziali esistono in ogni punto  $x \in D$ , la funzione  $x \mapsto \partial_{x_k}f(x)$  è una funzione scalare  $D \rightarrow \mathbb{R}$ , mentre la funzione  $x \mapsto \nabla f(x)$  è un campo vettoriale,  $D \rightarrow \mathbb{R}^n$ .

**Definizione 6.3.** Sia  $D$  un aperto di  $\mathbb{R}^n$  e  $f: D \rightarrow \mathbb{R}$  una funzione scalare che ha tutte le derivate parziali in ogni punto  $x \in D$ . Allora la funzione

$$F: D \rightarrow \mathbb{R}^n, \quad F(x) = \nabla f(x)$$

è detta un *campo conservativo*, ed  $f$  è detta un *potenziale* del campo  $F$ . □

**Osservazione 6.4.** Per funzioni di  $\mathbb{R} \rightarrow \mathbb{R}$ , l’esistenza della derivata in un punto  $x$  implica la continuità in  $x$ . Per le derivate parziali questo non è vero, come mostra l’esempio seguente.

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x_1, x_2) = \begin{cases} 1 & \text{se } x_1x_2 \neq 0, \\ 0 & \text{se } x_1x_2 = 0. \end{cases}$$

La funzione  $t \mapsto f(t, 0)$  è la funzione costantemente nulla, che è derivabile e ha derivata zero. In altre parole: il rapporto incrementale in  $x = 0$  rispetto alla variabile  $x_1$  è

$$\frac{f(t, 0) - f(0, 0)}{t - 0} = \frac{0}{t} = 0 \quad \forall t \in \mathbb{R}, t \neq 0,$$

quindi esiste il suo limite per  $t \rightarrow 0$  e vale 0. Similmente per  $x_2$ , quindi esistono le derivate parziali nell’origine  $\partial_{x_1}f(0)$  e  $\partial_{x_2}f(0)$ , e valgono 0. D’altra parte la funzione non è continua nell’origine: la composizione con la curva continua  $\varphi(t) = (t, t)$  (che è la retta di equazione  $x_2 = x_1$ ) dà

$$\lim_{t \rightarrow 0} f(\varphi(t)) = \lim_{t \rightarrow 0} f(t, t) = 1,$$

mentre la composizione con la curva continua  $\psi(t) = (t, 0)$  (che è la retta di equazione  $x_2 = 0$ , cioè l’asse delle  $x_1$ ) dà

$$\lim_{t \rightarrow 0} f(\psi(t)) = \lim_{t \rightarrow 0} f(t, 0) = 0,$$

dunque per la proposizione 5.2  $f$  non è continua nell’origine. □

**Definizione 6.5.** Sia  $D$  un aperto di  $\mathbb{R}^n$ ,  $f: D \rightarrow \mathbb{R}^m$  una funzione vettoriale di componenti  $(f_1, \dots, f_m)$ . Supponiamo che in un punto  $x \in D$  esistano tutte le derivate parziali di tutte le funzioni componenti  $f_j$  (ogni funzione componente  $f_j$  è una funzione scalare, quindi ha senso la definizione di derivata parziale vista sopra). Si chiama *matrice jacobiana* o *jacobiano di  $f$  in  $x$*  la matrice

$$J_f(x) = \begin{pmatrix} \partial_{x_1}f_1(x) & \partial_{x_2}f_1(x) & \partial_{x_3}f_1(x) & \dots & \partial_{x_n}f_1(x) \\ \partial_{x_1}f_2(x) & \partial_{x_2}f_2(x) & \partial_{x_3}f_2(x) & \dots & \partial_{x_n}f_2(x) \\ \partial_{x_1}f_3(x) & \partial_{x_2}f_3(x) & \partial_{x_3}f_3(x) & \dots & \partial_{x_n}f_3(x) \\ \dots & \dots & \dots & \dots & \dots \\ \partial_{x_1}f_m(x) & \partial_{x_2}f_m(x) & \partial_{x_3}f_m(x) & \dots & \partial_{x_n}f_m(x) \end{pmatrix}.$$

La  $j$ -esima riga di  $J_f(x)$  è il vettore  $\nabla f_j(x)$  “scritto in riga”. La  $k$ -esima colonna di  $J_f(x)$  è la derivata parziale rispetto a  $x_k$  del vettore  $f(x) = (f_1(x), \dots, f_m(x))$ . □

**Osservazione 6.6.** • Se  $m = n$ ,  $J_f(x)$  è una matrice quadrata.

- Se  $m = 1$ ,  $f$  è una funzione scalare, e  $J_f(x)$  è la matrice riga

$$J_f(x) = (\partial_{x_1}f(x) \quad \partial_{x_2}f(x) \quad \partial_{x_3}f(x) \quad \dots \quad \partial_{x_n}f(x)),$$

cioè il vettore  $\nabla f(x)$  scritto in riga.

- Se  $n = 1$ ,  $f$  è una curva in  $\mathbb{R}^m$ ,  $x \in \mathbb{R}$  è l'unica variabile, la derivata parziale  $\partial_x f_j(x)$  si indica  $f'_j(x)$ , trattandosi di funzione reale di una variabile reale, e  $J_f(x)$  è la matrice colonna

$$J_f(x) = \begin{pmatrix} f'_1(x) \\ f'_2(x) \\ \vdots \\ f'_m(x) \end{pmatrix}. \quad \square$$

## 7 Il differenziale

Nel corso di Analisi 1 è stato provato che per funzioni di una variabile reale la derivabilità implica la continuità. Per funzioni di più variabili reali, invece, l'esistenza delle derivate parziali non implica la continuità (osservazione 6.4), e quindi, da questo punto di vista, l'esistenza delle derivate parziali non è la nozione che estende il concetto di derivabilità alle funzioni di più variabili. Tale estensione è rappresentata dalla nozione di *differenziabilità*. Prima di dare la definizione, facciamo un passo indietro e rivediamo la definizione di derivabilità per funzioni di una variabile reale:  $f: D \subseteq \mathbb{R} \rightarrow \mathbb{R}$  è derivabile nel punto  $x \in D$  se esiste un numero, indicato con  $f'(x)$ , tale che

$$\lim_{\substack{|h| \rightarrow 0 \\ h \in \mathbb{R} \setminus \{0\}}} \frac{f(x+h) - f(x)}{h} = f'(x),$$

cioè

$$\lim_{\substack{|h| \rightarrow 0 \\ h \in \mathbb{R} \setminus \{0\}}} \left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| = 0,$$

cioè

$$\lim_{\substack{|h| \rightarrow 0 \\ h \in \mathbb{R} \setminus \{0\}}} \left| \frac{f(x+h) - f(x) - f'(x)h}{h} \right| = 0,$$

cioè

$$\lim_{\substack{|h| \rightarrow 0 \\ h \in \mathbb{R} \setminus \{0\}}} \frac{|f(x+h) - f(x) - f'(x)h|}{|h|} = 0,$$

cioè, per definizione di "o piccolo",

$$f(x+h) - f(x) - f'(x)h = o(|h|) \quad \text{per } |h| \rightarrow 0,$$

che è la formula di Taylor di ordine 1 con resto in forma di Peano. Questa è la definizione che estendiamo dal caso scalare a quello vettoriale:

**Definizione 7.1.** Sia  $D$  un aperto di  $\mathbb{R}^n$  e  $f: D \rightarrow \mathbb{R}^m$ .  $f$  si dice *differenziabile nel punto*  $x \in D$  se esiste una funzione lineare di  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ , cioè se esiste una matrice  $A \in \text{Mat}(m \times n)$  tale che

$$\lim_{\substack{|h| \rightarrow 0 \\ h \in \mathbb{R}^n \setminus \{0\}}} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0. \quad (8)$$

In tal caso la funzione lineare  $h \mapsto Ah$  si chiama *differenziale di  $f$  in  $x$*  e si indica

$$df(x): \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad h \mapsto df(x)[h] = Ah.$$

In altre parole:  $f$  è differenziabile nel punto  $x$  se esiste una funzione lineare  $df(x): \mathbb{R}^n \rightarrow \mathbb{R}^m$  tale che

$$f(x+h) = f(x) + df(x)[h] + o(|h|) \quad \text{per } |h| \rightarrow 0. \quad \square$$

Quest'ultima uguaglianza è lo sviluppo di Taylor di ordine 1 per  $f$  intorno al punto  $x$  con resto in forma di Peano. Il differenziale di  $f$  in  $x$  applicato al vettore  $h$ ,  $df(x)[h]$ , è l'equivalente vettoriale del prodotto  $f'(x)h$  che compare nella formula di Taylor nel caso scalare. In effetti, applicando la definizione di differenziale ad una funzione reale di una variabile reale  $f: D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ , si vede che il differenziale di  $f$  in  $x$  è la funzione lineare "moltiplicazione per il numero reale  $f'(x)$ ",

$$df(x): \mathbb{R} \rightarrow \mathbb{R}, \quad df(x)[h] = f'(x)h.$$

**Proposizione 7.2.** *Supponiamo che  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  sia differenziabile in  $x$ . Allora esistono tutte le derivate parziali di  $f$  in  $x$ , e*

$$df(x)[h] = J_f(x)h \quad \forall h \in \mathbb{R}^n,$$

*cioè il differenziale  $df(x)$  è la funzione lineare rappresentata dalla matrice jacobiana  $J_f(x)$ .*

*Dimostrazione.* Sia  $f$  differenziabile in  $x$ , e sia

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

la matrice a  $m$  righe e  $n$  colonne che rappresenta il differenziale. Consideriamo il vettore  $h = se_k = (0, \dots, 0, s, 0, \dots, 0)$ , dove  $s \in \mathbb{R}$ ,  $s \neq 0$ , e  $e_k$  è il  $k$ -esimo vettore della base canonica. Allora

$$|h| = |s|, \quad Ah = sAe_k = s \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix},$$

e  $|h| \rightarrow 0$  se  $s \rightarrow 0$ . Il limite (8) (che vale per ipotesi) applicato per  $h = se_k$  dà

$$\lim_{|s| \rightarrow 0} \frac{|f(x + se_k) - f(x) - sAe_k|}{|s|} = 0.$$

Ora,  $f(x + se_k) - f(x) - sAe_k$  è il vettore

$$f(x + se_k) - f(x) - sAe_k = \begin{pmatrix} f_1(x + se_k) - f_1(x) - sa_{1k} \\ f_2(x + se_k) - f_2(x) - sa_{2k} \\ f_3(x + se_k) - f_3(x) - sa_{3k} \\ \vdots \\ f_m(x + se_k) - f_m(x) - sa_{mk} \end{pmatrix},$$

e, per la (4),

$$|f_j(x + se_k) - f_j(x) - sa_{jk}| \leq |f(x + se_k) - f(x) - sAe_k|$$

per ogni  $j = 1, \dots, m$ . Di conseguenza

$$0 \leq \frac{|f_j(x + se_k) - f_j(x) - sa_{jk}|}{|s|} \leq \frac{|f(x + se_k) - f(x) - sAe_k|}{|s|}$$

e dunque (carabinieri)

$$\lim_{s \rightarrow 0} \frac{|f_j(x + se_k) - f_j(x) - sa_{jk}|}{|s|} = 0.$$

Poi

$$\frac{|f_j(x + se_k) - f_j(x) - sa_{jk}|}{|s|} = \left| \frac{f_j(x + se_k) - f_j(x) - sa_{jk}}{s} \right| = \left| \frac{f_j(x + se_k) - f_j(x)}{s} - a_{jk} \right|,$$

quindi

$$\lim_{s \rightarrow 0} \left| \frac{f_j(x + se_k) - f_j(x)}{s} - a_{jk} \right| = 0,$$

cioè

$$\lim_{s \rightarrow 0} \frac{f_j(x + se_k) - f_j(x)}{s} = a_{jk},$$

e perciò esiste la derivata parziale di  $f_j$  rispetto a  $x_k$  nel punto  $x$ , e vale  $a_{jk}$ . Dal momento che questo argomento vale per ogni  $k = 1, \dots, n$  ed ogni  $j = 1, \dots, m$ ,  $A$  è proprio la matrice jacobiana  $J_f(x)$ .  $\square$

**Osservazione 7.3.** Se  $f$  è una funzione scalare, il differenziale ha una rappresentazione particolarmente semplice:

$$df(x)[h] = \nabla f(x) \cdot h.$$

Infatti (osservazione 6.6)  $J_f(x)$  è la matrice riga  $(\partial_{x_1} f(x) \quad \partial_{x_2} f(x) \quad \dots \quad \partial_{x_n} f(x))$ .  $\square$

**Proposizione 7.4.**  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  è differenziabile in  $x \in D$  se e solo se ogni funzione componente  $f_j: D \rightarrow \mathbb{R}$  è differenziabile in  $x$ , per ogni  $j = 1, \dots, m$ .

*Dimostrazione.* Ricordiamo che la  $j$ -esima riga della matrice jacobiana è il gradiente di  $f_j$  scritto in riga, e quindi

$$df(x)[h] = \begin{pmatrix} \nabla f_1(x) \cdot h \\ \nabla f_2(x) \cdot h \\ \vdots \\ \nabla f_m(x) \cdot h \end{pmatrix} = \begin{pmatrix} df_1(x)[h] \\ df_2(x)[h] \\ \vdots \\ df_m(x)[h] \end{pmatrix}.$$

Perciò  $f(x+h) - f(x) - df(x)[h]$  è il vettore che ha  $f_j(x+h) - f_j(x) - df_j(x)[h]$  come  $j$ -esima componente,

$$f(x+h) - f(x) - df(x)[h] = \begin{pmatrix} f_1(x+h) - f_1(x) - df_1(x)[h] \\ f_2(x+h) - f_2(x) - df_2(x)[h] \\ \vdots \\ f_m(x+h) - f_m(x) - df_m(x)[h] \end{pmatrix}.$$

$\Rightarrow$ ) Sia  $f$  differenziabile in  $x$ . Per la (4),

$$0 \leq \frac{|f_j(x+h) - f_j(x) - df_j(x)[h]|}{|h|} \leq \frac{|f(x+h) - f(x) - df(x)[h]|}{|h|},$$

per ogni  $j = 1, \dots, m$ . Il termine a destra  $\rightarrow 0$  per  $|h| \rightarrow 0$  per ipotesi, dunque anche il termine in mezzo  $\rightarrow 0$ , cioè  $f_j$  è differenziabile in  $x$ , per ogni  $j = 1, \dots, m$ .

$\Leftarrow$ ) Supponiamo che  $f_1, \dots, f_m$  siano differenziabili in  $x$ . Per definizione di norma euclidea,

$$\begin{aligned} & \frac{|f(x+h) - f(x) - df(x)[h]|}{|h|} \\ &= \frac{\sqrt{|f_1(x+h) - f_1(x) - df_1(x)[h]|^2 + \dots + |f_m(x+h) - f_m(x) - df_m(x)[h]|^2}}{|h|} \\ &= \sqrt{\frac{|f_1(x+h) - f_1(x) - df_1(x)[h]|^2 + \dots + |f_m(x+h) - f_m(x) - df_m(x)[h]|^2}{|h|^2}} \\ &= \sqrt{\left(\frac{|f_1(x+h) - f_1(x) - df_1(x)[h]|}{|h|}\right)^2 + \dots + \left(\frac{|f_m(x+h) - f_m(x) - df_m(x)[h]|}{|h|}\right)^2}, \end{aligned}$$

e questo è infinitesimo per  $|h| \rightarrow 0$  perché, per ipotesi, ciascun termine in parentesi  $\rightarrow 0$ .  $\square$

**Proposizione 7.5.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  differenziabile in  $x \in D$ . Allora  $f$  è continua in  $x$ .

*Dimostrazione.* Il differenziale  $df(x)$  è una funzione lineare, dunque lipschitziana,

$$|df(x)[h]| \leq L|h| \quad \forall h \in \mathbb{R}^n,$$

con  $L = |J_f(x)|$  (vedi proposizione 7.2 e la disuguaglianza (6)). Per ipotesi,  $f(x+h) - f(x) - df(x)[h] = o(|h|)$  per  $|h| \rightarrow 0$ , dunque, per disuguaglianza triangolare,

$$\begin{aligned} |f(x+h) - f(x)| &= |f(x+h) - f(x) - df(x)[h] + df(x)[h]| \\ &\leq |f(x+h) - f(x) - df(x)[h]| + |df(x)[h]| \\ &\leq o(|h|) + L|h| \rightarrow 0 \end{aligned}$$

per  $h \rightarrow 0$ . □

Abbiamo provato che la differenziabilità implica l'esistenza di tutte le derivate parziali. Per contro, può capitare che esistano tutte le derivate parziali senza che ci sia differenziabilità. Un esempio di questo fenomeno è la funzione descritta nell'osservazione 6.4: le derivate in 0 esistono entrambe, ma la funzione non è differenziabile, perché, se lo fosse, sarebbe anche continua in 0 grazie alla proposizione 7.5, mentre invece, come si è visto nell'osservazione 6.4, la funzione non è continua in 0.

Tuttavia, l'esistenza delle derivate parziali in tutti i punti e la loro continuità bastano per implicare la differenziabilità:

**Teorema 7.6** (Teorema del differenziale). *Sia  $D \subseteq \mathbb{R}^n$  un aperto,  $f: D \rightarrow \mathbb{R}^m$  una funzione che ha tutte le derivate parziali in tutti i punti di  $D$ . Supponiamo che tutte le derivate parziali siano continue in un punto  $x \in D$ . Allora  $f$  è differenziabile in  $x$ .*

*Dimostrazione.* Dimostriamo il teorema per  $n = 3$ . La dimostrazione generale per  $n$  qualsiasi è più laboriosa da scrivere, ma identica come ragionamento.

1) Dapprima dimostriamo il teorema nel caso in cui  $f$  sia una funzione scalare,  $f: D \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$ . Scriviamo la differenza  $f(x+h) - f(x)$  aggiungendo e togliendo termini in cui cambiamo soltanto una componente nel passare da  $x+h$  a  $x$ :

$$\begin{aligned} f(x+h) - f(x) &= f(x_1+h_1, x_2+h_2, x_3+h_3) - f(x_1, x_2, x_3) \\ &= f(x_1+h_1, x_2+h_2, x_3+h_3) - f(x_1, x_2+h_2, x_3+h_3) \\ &\quad + f(x_1, x_2+h_2, x_3+h_3) - f(x_1, x_2, x_3+h_3) \\ &\quad + f(x_1, x_2, x_3+h_3) - f(x_1, x_2, x_3). \end{aligned}$$

I primi due termini sono

$$f(x_1+h_1, x_2+h_2, x_3+h_3) - f(x_1, x_2+h_2, x_3+h_3) = \varphi(x_1+h_1) - \varphi(x_1),$$

dove  $\varphi$  è la funzione

$$\varphi(t) = f(t, x_2+h_2, x_3+h_3),$$

funzione reale di una variabile reale. Se  $h_1 \neq 0$ , applichiamo il teorema di Lagrange, ricordando che la derivata di  $\varphi$  è, per definizione, la derivata parziale di  $f$  rispetto a  $x_1$ : esiste un numero reale  $\xi_1$  nell'intervallo aperto di estremi  $x_1$  e  $x_1+h_1$  tale che

$$\varphi(x_1+h_1) - \varphi(x_1) = \varphi'(\xi_1)h_1,$$

cioè

$$f(x_1+h_1, x_2+h_2, x_3+h_3) - f(x_1, x_2+h_2, x_3+h_3) = \partial_{x_1} f(\xi_1, x_2+h_2, x_3+h_3) h_1.$$

Se  $h_1 = 0$ , quest'ultima uguaglianza vale con  $\xi_1 = x_1$ . Quindi, in ogni caso, esiste un numero reale  $\xi_1$  nell'intervallo chiuso di estremi  $x_1$  e  $x_1+h_1$  che soddisfa l'ultima uguaglianza. Applicando lo stesso ragionamento alla funzione

$$\psi(t) = f(x_1, t, x_3+h_3),$$

si ha che esiste un numero reale  $\xi_2$  nell'intervallo chiuso di estremi  $x_2$  e  $x_2+h_2$  tale che

$$\begin{aligned} f(x_1, x_2+h_2, x_3+h_3) - f(x_1, x_2, x_3+h_3) &= \psi(x_2+h_2) - \psi(x_2) \\ &= \psi'(\xi_2) h_2 \\ &= \partial_{x_2} f(x_1, \xi_2, x_3+h_3) h_2. \end{aligned}$$

Applicando lo stesso ragionamento alla funzione

$$\rho(t) = f(x_1, x_2, t),$$

si ha che esiste un numero reale  $\xi_3$  nell'intervallo chiuso di estremi  $x_3$  e  $x_3 + h_3$  tale che

$$\begin{aligned} f(x_1, x_2, x_3 + h_3) - f(x_1, x_2, x_3) &= \rho(x_3 + h_3) - \rho(x_3) \\ &= \rho'(\xi_3) h_3 \\ &= \partial_{x_3} f(x_1, x_2, \xi_3) h_3. \end{aligned}$$

Di conseguenza,

$$\begin{aligned} f(x + h) - f(x) &= \partial_{x_1} f(\xi_1, x_2 + h_2, x_3 + h_3) h_1 \\ &\quad + \partial_{x_2} f(x_1, \xi_2, x_3 + h_3) h_2 \\ &\quad + \partial_{x_3} f(x_1, x_2, \xi_3) h_3. \end{aligned}$$

Per quanto riguarda il differenziale, si ha (osservazione 7.3)

$$df(x)[h] = \partial_{x_1} f(x) h_1 + \partial_{x_2} f(x) h_2 + \partial_{x_3} f(x) h_3.$$

Perciò

$$f(x + h) - f(x) - df(x)[h] = b_1 h_1 + b_2 h_2 + b_3 h_3,$$

dove

$$\begin{aligned} b_1 &:= \partial_{x_1} f(\xi_1, x_2 + h_2, x_3 + h_3) - \partial_{x_1} f(x), \\ b_2 &:= \partial_{x_2} f(x_1, \xi_2, x_3 + h_3) - \partial_{x_2} f(x), \\ b_3 &:= \partial_{x_3} f(x_1, x_2, \xi_3) - \partial_{x_3} f(x). \end{aligned}$$

Di conseguenza, per  $h \neq 0$ , usando la disuguaglianza triangolare e la (4),

$$\begin{aligned} \frac{|f(x + h) - f(x) - df(x)[h]|}{|h|} &= \frac{|b_1 h_1 + b_2 h_2 + b_3 h_3|}{|h|} \\ &\leq \frac{|b_1| |h_1| + |b_2| |h_2| + |b_3| |h_3|}{|h|} \\ &= |b_1| \frac{|h_1|}{|h|} + |b_2| \frac{|h_2|}{|h|} + |b_3| \frac{|h_3|}{|h|} \\ &\leq |b_1| + |b_2| + |b_3|. \end{aligned}$$

Quando  $|h| \rightarrow 0$ ,  $\xi_k \rightarrow x_k$ ,  $k = 1, 2, 3$ , perché

$$|\xi_k - x_k| \leq |h_k| \leq |h| \rightarrow 0, \quad k = 1, 2, 3,$$

e quindi i punti

$$\begin{pmatrix} \xi_1 \\ x_2 + h_2 \\ x_3 + h_3 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ \xi_2 \\ x_3 + h_3 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \\ \xi_3 \end{pmatrix} \rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x$$

per  $|h| \rightarrow 0$ . Per ipotesi, le derivate parziali di  $f$  sono continue in  $x$ , dunque  $b_1, b_2, b_3 \rightarrow 0$  per  $|h| \rightarrow 0$ , e quindi anche

$$\frac{|f(x + h) - f(x) - df(x)[h]|}{|h|} \rightarrow 0 \quad \text{per } |h| \rightarrow 0,$$

cioè  $f$  è differenziabile in  $x$ .

2) Dimostriamo ora il teorema per  $f$  vettoriale,  $f: D \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}^m$ , con  $f = (f_1, \dots, f_m)$ . Ciascuna componente  $f_j$  è una funzione scalare che ha tutte le derivate parziali continue per ipotesi, quindi, per quanto dimostrato al punto 1), è differenziabile. Dunque, per la proposizione 7.4,  $f$  è differenziabile.  $\square$

In dimensione 1, se  $f$  e  $g$  sono funzioni reali di variabile reale che sono derivabili, anche la loro composizione  $F(x) = g(f(x))$  è derivabile, e vale la “regola della catena” per la derivata della funzione composta:

$$F'(x) = g'(f(x))f'(x).$$

La proposizione che segue estende questa formula al caso vettoriale.

**Proposizione 7.7.** *Sia  $D$  un sottoinsieme aperto di  $\mathbb{R}^n$  e  $B$  un sottoinsieme aperto di  $\mathbb{R}^m$ . Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $g: B \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^p$ , e supponiamo che  $f(x) \in B$  per ogni  $x \in D$ . Sia  $\bar{x} \in D$ ,  $\bar{y} = f(\bar{x}) \in B$ . Supponiamo che  $f$  sia differenziabile in  $\bar{x}$  e  $g$  sia differenziabile in  $\bar{y}$ . Allora la funzione composta*

$$F: D \rightarrow \mathbb{R}^p, \quad F(x) = g(f(x))$$

è differenziabile in  $\bar{x}$ , il suo differenziale in  $\bar{x}$  è la funzione lineare  $dF(\bar{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^p$ ,

$$dF(\bar{x})[h] = dg(\bar{y})[df(\bar{x})[h]] = J_g(\bar{y})J_f(\bar{x})h \quad \forall h \in \mathbb{R}^n,$$

e la sua matrice jacobiana è dunque il prodotto delle matrici jacobiane di  $g$  e di  $f$  (usuale prodotto righe per colonne),

$$J_F(\bar{x}) = J_g(\bar{y})J_f(\bar{x}). \quad (9)$$

Di conseguenza, dalla regola del prodotto di matrici righe per colonne, la derivata parziale rispetto a  $x_k$  della funzione componente  $F_j$  nel punto  $\bar{x}$  è

$$\partial_{x_k} F_j(\bar{x}) = \sum_{i=1}^m \partial_{y_i} g_j(\bar{y}) \partial_{x_k} f_i(\bar{x}), \quad (10)$$

per ogni  $k = 1, \dots, n$  e ogni  $j = 1, \dots, p$ .

*Dimostrazione.* La dimostrazione non è in programma. □

Vediamo come si ricava la formula (10) dalla (9) e dalla regola del prodotto righe per colonne:  $\partial_{x_k} F_j(\bar{x})$  è l'elemento di posto riga  $j$ , colonna  $k$  della matrice  $J_F(\bar{x})$ , e quindi è il prodotto della riga  $j$  della matrice  $J_g(\bar{y})$  per la colonna  $k$  della matrice  $J_f(\bar{x})$ . Ora, la  $j$ -esima riga di  $J_g(\bar{y})$  è

$$(\partial_{y_1} g_j(\bar{y}) \quad \partial_{y_2} g_j(\bar{y}) \quad \partial_{y_3} g_j(\bar{y}) \quad \dots \quad \partial_{y_m} g_j(\bar{y})),$$

cioè il gradiente  $\nabla g_j(\bar{y})$  della funzione componente  $g_j$  nel punto  $\bar{y}$  scritto in riga, mentre la  $k$ -esima colonna di  $J_f(\bar{x})$  è

$$\begin{pmatrix} \partial_{x_k} f_1(\bar{x}) \\ \partial_{x_k} f_2(\bar{x}) \\ \partial_{x_k} f_3(\bar{x}) \\ \vdots \\ \partial_{x_k} f_m(\bar{x}) \end{pmatrix},$$

cioè la derivata rispetto a  $x_k$  del vettore  $(f_1(x), \dots, f_m(x))$  in  $\bar{x}$ . Perciò

$$\partial_{x_k} F_j(\bar{x}) = (\partial_{y_1} g_j(\bar{y}) \quad \dots \quad \partial_{y_m} g_j(\bar{y})) \begin{pmatrix} \partial_{x_k} f_1(\bar{x}) \\ \vdots \\ \partial_{x_k} f_m(\bar{x}) \end{pmatrix} = \sum_{i=1}^m \partial_{y_i} g_j(\bar{y}) \partial_{x_k} f_i(\bar{x}),$$

che è la (10).

Vediamo alcuni casi particolari.

- Se  $p = 1$ ,  $g$  è una funzione scalare,  $g: B \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ , allora anche  $F$  è una funzione scalare,  $F: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , e (10) diventa

$$\partial_{x_k} F(\bar{x}) = \sum_{i=1}^m \partial_{y_i} g(\bar{y}) \partial_{x_k} f_i(\bar{x}).$$

- Se  $n = 1$ ,  $f$  è una curva,  $f: D \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ , al posto del vettore  $x = (x_1, \dots, x_n)$  c'è una sola variabile reale  $x \in \mathbb{R}$ , e la derivata parziale  $\partial_x$  si indica semplicemente con  $'$ . Dunque  $F$  è funzione di un'unica variabile reale,  $F: D \subseteq \mathbb{R} \rightarrow \mathbb{R}^p$ ,  $F$  è una curva in  $\mathbb{R}^p$ , e la (10) diventa

$$F'_j(\bar{x}) = \sum_{i=1}^m \partial_{y_i} g_j(\bar{y}) f'_i(\bar{x}).$$

- Se poi  $n = p = 1$ ,  $f$  è una curva e  $g$  è scalare, allora  $F: D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ ,  $F$  è una funzione reale di una variabile reale (il tipo di funzioni studiate in Analisi 1), e

$$F'(\bar{x}) = \sum_{i=1}^m \partial_{y_i} g(\bar{y}) f'_i(\bar{x}).$$

- Infine, se  $n = m = p = 1$ , la somma di indice  $i = 1, \dots, m$  ha un unico termine, anche al posto del vettore  $y = (y_1, \dots, y_m)$  c'è un'unica variabile reale  $y \in \mathbb{R}$ , anche la derivata parziale  $\partial_y$  si indica semplicemente con  $'$ , e si ritrova la regola della catena vista in Analisi 1:

$$F'(\bar{x}) = g'(\bar{y}) f'(\bar{x})$$

(ricordiamo che, per ipotesi,  $\bar{y} = f(\bar{x})$ ).

## 8 Derivate direzionali

Da questo capitolo in poi limitiamo il nostro studio alle funzioni scalari.

**Osservazione 8.1.** Sia  $D \subseteq \mathbb{R}^n$ ,  $f: D \rightarrow \mathbb{R}$  una funzione scalare. Sia  $x$  un punto interno di  $D$ ,  $x \in \text{int}D$ . Allora, per definizione, esiste  $r > 0$  tale che  $B(x, r) \subseteq D$ . Sia  $h \in \mathbb{R}^n$ ,  $h \neq 0$ . Sia

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}^n, \quad \varphi(s) = x + sh.$$

I punti  $\varphi(s)$ ,  $s \in \mathbb{R}$ , formano la retta che passa per il punto  $x$  e ha direzione data dal vettore  $h$ . Un punto  $\varphi(s) = x + sh$  sta nella palla  $B(x, r)$  se  $|x + sh - x| = |s||h| < r$ . Poniamo  $\delta = r/|h|$ . Allora, per ogni  $s \in (-\delta, \delta)$ , il punto  $x + sh$  appartiene a  $B(x, r)$ , e quindi a  $D$ . Di conseguenza sull'intervallo  $(-\delta, \delta)$  è definita la funzione composta

$$F: (-\delta, \delta) \rightarrow \mathbb{R}, \quad F(s) = f(\varphi(s)) = f(x + sh).$$

Se l'ipotesi è che  $D$  sia aperto e  $x \in D$ , la situazione è identica: infatti, se  $D$  è aperto,  $D = \text{int}D$ .

Quando si dice che la funzione composta  $F(s) = f(x + sh)$  è definita "per  $s$  sufficientemente piccolo", si intende l'osservazione appena fatta.  $\square$

**Definizione 8.2.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione scalare e  $x \in \text{int}D$ . Diciamo che esiste la *derivata direzionale di  $f$  nella direzione  $v$  nel punto  $x$*  se la funzione reale di una variabile reale

$$s \mapsto f(x + sv)$$

è derivabile in  $s = 0$ , cioè se esiste il limite del rapporto incrementale

$$\lim_{\substack{s \rightarrow 0 \\ s \in \mathbb{R} \setminus \{0\}}} \frac{f(x + sv) - f(x)}{s}.$$

In tal caso, questo limite si dice derivata direzionale di  $f$  in  $x$  nella direzione  $v$  e si indica con uno di questi simboli:

$$\partial_v f(x) = \frac{\partial}{\partial v} f(x) = \frac{\partial f(x)}{\partial v} = D_v f(x). \quad \square$$

Osserviamo che, scegliendo  $v = e_k$ , la derivata direzionale nella direzione  $v = e_k$  è la derivata parziale rispetto a  $x_k$  (vedi la definizione di derivata parziale).

**Proposizione 8.3.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione scalare e  $x \in \text{int}D$ . Se  $f$  è differenziabile in  $x$ , allora esistono tutte le derivate direzionali di  $f$  in  $x$ , e

$$\partial_v f(x) = \nabla f(x) \cdot v = \sum_{i=1}^n \partial_{x_i} f(x) v_i$$

per ogni  $v \in \mathbb{R}^n$ ,  $v \neq 0$ .

*Dimostrazione.* Come visto nell'osservazione 8.1, esiste un intervallo  $(-\delta, \delta) \subset \mathbb{R}$ ,  $\delta > 0$ , in cui è definita la funzione composta

$$F: (-\delta, \delta) \rightarrow \mathbb{R}, \quad F(s) = f(\varphi(s)) = f(x + sv).$$

Per ipotesi,  $f$  è differenziabile nel punto  $x = \varphi(0)$ .  $\varphi$  è differenziabile in ogni  $s$ , e

$$\varphi(s) = \begin{pmatrix} \varphi_1(s) \\ \varphi_2(s) \\ \vdots \\ \varphi_n(s) \end{pmatrix} = \begin{pmatrix} x_1 + sv_1 \\ x_2 + sv_2 \\ \vdots \\ x_n + sv_n \end{pmatrix}, \quad \varphi'(s) = \begin{pmatrix} \varphi_1'(s) \\ \varphi_2'(s) \\ \vdots \\ \varphi_n'(s) \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v.$$

Dalla proposizione 7.7 segue che  $F$  è differenziabile in  $s = 0$ , e, applicando la formula per il differenziale della funzione composta (dimensionalmente parlando, siamo nel caso  $\mathbb{R} \rightarrow \mathbb{R}^n \rightarrow \mathbb{R}$ ),

$$F'(0) = \sum_{i=1}^n \partial_{x_i} f(x) \varphi_i'(0) = \sum_{i=1}^n \partial_{x_i} f(x) v_i = \nabla f(x) \cdot v.$$

Per definizione di derivata direzionale  $F'(0) = \partial_v f(x)$ , e la tesi è provata.  $\square$

**Osservazione 8.4.** Sia  $v$  un versore di  $\mathbb{R}^n$ , cioè  $|v| = 1$ , e supponiamo che  $\nabla f(x) \neq 0$ . Per la proposizione precedente e per la definizione di angolo tra due vettori,

$$\partial_v f(x) = \nabla f(x) \cdot v = |\nabla f(x)| |v| \cos \alpha = |\nabla f(x)| \cos \alpha,$$

dove  $\alpha \in [0, \pi]$  è l'angolo formato dai vettori  $v$  e  $\nabla f(x)$ . Quindi

$$-|\nabla f(x)| \leq \partial_v f(x) \leq |\nabla f(x)| \quad \forall v \in \mathbb{R}^n, |v| = 1,$$

e  $\partial_v f(x)$  è massima quando  $\cos \alpha = 1$ , cioè  $\alpha = 0$ , cioè quando  $v$  e  $\nabla f(x)$  sono due vettori con la stessa direzione e lo stesso verso, cioè  $v = \lambda \nabla f(x)$  per un certo  $\lambda > 0$ ;  $\partial_v f(x)$  è minima quando  $\cos \alpha = -1$ , cioè  $\alpha = \pi$ , cioè quando  $v$  e  $\nabla f(x)$  sono due vettori con la stessa direzione e il verso opposto, cioè  $v = \lambda \nabla f(x)$  per un certo  $\lambda < 0$ ;  $\partial_v f(x)$  non è né massima né minima quando  $\cos \alpha \in (-1, 1)$ , cioè  $\alpha \in (0, \pi)$ , cioè quando  $v$  e  $\nabla f(x)$  sono due vettori di direzione diversa, cioè quando non sono proporzionali.

Di conseguenza il gradiente di  $f$  in  $x$  indica la direzione e il verso in cui la derivata direzionale è massima, cioè la direzione e il verso "di massima pendenza", quella in cui la  $f$  cresce più rapidamente.  $\square$

## 9 Derivate parziali successive

Supponiamo che la funzione scalare  $f$  abbia derivata parziale rispetto a  $x_k$  in ogni punto del dominio. La derivata parziale è anch'essa una funzione scalare, e ha senso chiedersi se ha anch'essa derivate parziali.

**Definizione 9.1.** Sia  $D \subseteq \mathbb{R}^n$  aperto,  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione scalare che ha tutte le derivate parziali in ogni punto  $x \in D$ . Ciascuna derivata parziale  $\partial_{x_k} f$  è anch'essa una funzione scalare definita su  $D$ . Se a sua volta  $\partial_{x_k} f$  ha derivate parziali nel punto  $x$ , queste si chiamano *derivate parziali seconde* di  $f$  in  $x$ , e si indicano in vari modi,

$$\partial_{x_i x_k} f(x) = \frac{\partial^2}{\partial x_i \partial x_k} f(x) = \frac{\partial^2 f(x)}{\partial x_i \partial x_k} = \partial_{ik} f(x) = D_{x_i x_k} f(x) = D_{ik} f(x) = f_{x_i x_k}(x).$$

Se esistono tutte le derivate seconde di  $f$  in  $x$ , si chiama *hessiano* o *matrice hessiana* di  $f$  in  $x$  la matrice

$$H_f(x) = D^2f(x) = \begin{pmatrix} \partial_{11}f(x) & \partial_{12}f(x) & \partial_{13}f(x) & \dots & \partial_{1n}f(x) \\ \partial_{21}f(x) & \partial_{22}f(x) & \partial_{23}f(x) & \dots & \partial_{2n}f(x) \\ \partial_{31}f(x) & \partial_{32}f(x) & \partial_{33}f(x) & \dots & \partial_{3n}f(x) \\ \dots & \dots & \dots & \dots & \dots \\ \partial_{n1}f(x) & \partial_{n2}f(x) & \partial_{n3}f(x) & \dots & \partial_{nn}f(x) \end{pmatrix},$$

matrice quadrata  $n \times n$ . □

**Definizione 9.2.** Una funzione scalare  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  si dice *di classe  $C^2$*  sull'aperto  $D$ , e si indica  $f \in C^2(D)$ , se ha tutte le derivate seconde in ogni punto  $x \in D$  e sono tutte continue in  $D$ . □

In generale, la derivata parziale rispetto a  $x_i$  della funzione  $\partial_{x_k}f$  in  $x$  e la derivata parziale rispetto a  $x_k$  della funzione  $\partial_{x_i}f$  in  $x$  non coincidono, cioè in generale  $\partial_i\partial_kf \neq \partial_k\partial_if$ . Un esempio tipico è descritto nell'esercizio 6.

Tuttavia, se le derivate seconde sono continue, allora sono uguali:

**Teorema 9.3** (Teorema di Schwartz). *Sia  $D \subseteq \mathbb{R}^n$  aperto,  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione scalare che ha le derivate seconde in  $D$ . Se  $\partial_{x_i}\partial_{x_k}f$  e  $\partial_{x_k}\partial_{x_i}f$  sono continue nel punto  $x$ , allora*

$$\partial_{x_i}\partial_{x_k}f(x) = \partial_{x_k}\partial_{x_i}f(x).$$

*Dimostrazione.* Non in programma. □

Ricordiamo:

**Definizione 9.4.** Una matrice quadrata  $A = (a_{ij}) \in \text{Mat}(n \times n)$  si dice *simmetrica* se

$$a_{ij} = a_{ji} \quad \forall i, j = 1, \dots, n,$$

cioè se al posto (riga  $i$ , colonna  $j$ ) e al posto (riga  $j$ , colonna  $i$ ) si trova lo stesso valore. (Graficamente, una matrice quadrata è simmetrica se è speculare rispetto alla diagonale principale). □

Il teorema di Schwartz implica che, se  $f \in C^2(D)$ , per ogni  $x \in D$  la matrice hessiana  $H_f(x)$  è una matrice simmetrica.

Come conseguenza della formula per le derivate parziali delle funzioni composte (10) segue che la composizione di funzioni  $C^2$  è  $C^2$ :

**Proposizione 9.5.** *Siano  $D \subseteq \mathbb{R}^n$ ,  $B \subseteq \mathbb{R}^m$  aperti,  $f: D \rightarrow \mathbb{R}^m$ ,  $g: B \rightarrow \mathbb{R}^p$ , con  $f(x) \in B$  per ogni  $x \in D$ , cosicché la funzione composta  $F(x) = g(f(x))$  è definita per ogni  $x \in D$ . Se  $f$  e  $g$  sono di classe  $C^2$ , allora anche  $F$  è di classe  $C^2$ .*

*Dimostrazione.* Partiamo dalla formula della derivata parziale della funzione composta,

$$\partial_{x_k}F(x) = \sum_{i=1}^m \partial_{y_i}g(f(x)) \partial_{x_k}f_i(x),$$

e deriviamo rispetto a  $x_\ell$ :

$$\begin{aligned} \partial_{x_\ell}\partial_{x_k}F(x) &= \frac{\partial}{\partial x_\ell} \left( \sum_{i=1}^m \partial_{y_i}g(f(x)) \partial_{x_k}f_i(x) \right) \\ &= \sum_{i=1}^m \frac{\partial}{\partial x_\ell} \left( \partial_{y_i}g(f(x)) \partial_{x_k}f_i(x) \right) \\ &= \sum_{i=1}^m \left\{ \frac{\partial}{\partial x_\ell} \left( \partial_{y_i}g(f(x)) \right) \partial_{x_k}f_i(x) + \partial_{y_i}g(f(x)) \partial_{x_\ell}\partial_{x_k}f_i(x) \right\}. \end{aligned}$$

$\partial_{y_i}g(f(x))$  è la composizione della funzione  $\partial_{y_i}g$ , che è  $C^1$ , con  $f$ , che è  $C^1$  (di più, è addirittura  $C^2$ ), e quindi la composizione è di classe  $C^1$ , perciò  $\partial_{x_\ell}[\partial_{y_i}g(f(x))]$  è continua. Per gli altri termini,  $\partial_{x_\ell}\partial_{x_k}f_i$  è continua perché  $f$  è  $C^2$ , e anche  $\partial_{y_i}g(f(x))$  è continua (di più: è  $C^1$ ). Dunque tutti i termini sono somme e prodotti di funzioni continue, e quindi  $\partial_{x_\ell}\partial_{x_k}F$  esiste ed è una funzione continua. Questo vale per ogni indice  $\ell, k$ , quindi  $F$  è  $C^2$ . □

## 10 Formula di Taylor

Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x \in \text{int}D$ ,  $h \in \mathbb{R}^n$ ,  $h \neq 0$ . Consideriamo la composizione

$$F(s) = f(\varphi(s)) = f(x + sh),$$

funzione reale di variabile reale, definita in un certo intervallo  $(-\delta, \delta) \subset \mathbb{R}$ : vedi l'osservazione 8.1. Dalla formula per le derivate della funzione composta (vedi anche proposizione 8.3),

$$F'(s) = \nabla f(x + sh) \cdot h = \sum_{j=1}^n \partial_{x_j} f(x + sh) h_j.$$

Deriviamo un'altra volta rispetto alla variabile  $s$ , applicando a ciascun termine  $\partial_{x_j} f(x + sh)$  la formula di derivazione per le funzioni composte:

$$\begin{aligned} F''(s) &= \frac{d}{ds} \left( \sum_{j=1}^n \partial_{x_j} f(x + sh) h_j \right) \\ &= \sum_{j=1}^n \frac{d}{ds} \left( \partial_{x_j} f(x + sh) \right) h_j \\ &= \sum_{j=1}^n (\nabla \partial_{x_j} f(x + sh) \cdot h) h_j \\ &= \sum_{j=1}^n \left( \sum_{i=1}^n \partial_{x_i} \partial_{x_j} f(x + sh) h_i \right) h_j \\ &= \sum_{i,j=1}^n \partial_{x_i x_j} f(x + sh) h_i h_j. \end{aligned}$$

A questo punto facciamo un'osservazione sulle matrici, che ha interesse di per sé.

**Lemma 10.1.** *Sia  $A = (a_{ij}) \in \text{Mat}(n \times n)$  una matrice quadrata. Allora*

$$Ah \cdot h = \sum_{i,j=1}^n a_{ij} h_i h_j \quad \forall h = (h_1, \dots, h_n) \in \mathbb{R}^n.$$

*Dimostrazione.* Sia  $h \in \mathbb{R}^n$ , e indichiamo  $v = Ah$ . Il prodotto scalare  $Ah \cdot h$  è

$$Ah \cdot h = v \cdot h = v_1 h_1 + \dots + v_n h_n = \sum_{i=1}^n v_i h_i.$$

Dalla regola di prodotto righe per colonne, la  $i$ -esima componente  $v_i$  del vettore  $v = Ah$  è il prodotto della  $i$ -esima riga di  $A$  per  $h$ ,

$$v_i = (a_{i1} \quad a_{i2} \quad \dots \quad a_{in}) \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} = a_{i1} h_1 + \dots + a_{in} h_n = \sum_{j=1}^n a_{ij} h_j.$$

Dunque

$$Ah \cdot h = \sum_{i=1}^n v_i h_i = \sum_{i=1}^n \left( \sum_{j=1}^n a_{ij} h_j \right) h_i = \sum_{i,j=1}^n a_{ij} h_i h_j. \quad \square$$

Riprendiamo allora il calcolo della derivata seconda: per il lemma appena provato si ha

$$F''(s) = \sum_{i,j=1}^n \partial_{x_i x_j} f(x + sh) h_i h_j = D^2 f(x + sh) h \cdot h, \quad (11)$$

dove  $D^2 f(x + sh) = H_f(x + sh)$  è la matrice hessiana di  $f$  nel punto  $x + sh$ .

**Proposizione 10.2** (Formula di Taylor di ordine 1 con resto di Lagrange). Sia  $D \subseteq \mathbb{R}^n$  un aperto,  $f: D \rightarrow \mathbb{R}$  di classe  $C^1$  in  $D$ . Sia  $h \in \mathbb{R}^n$ ,  $h \neq 0$ , e supponiamo che tutto il segmento di estremi  $x$  e  $x+h$  sia contenuto in  $D$ , cioè

$$x + sh \in D \quad \forall s \in [0, 1].$$

Allora esiste  $\vartheta \in (0, 1)$  tale che

$$\begin{aligned} f(x+h) &= f(x) + \nabla f(x + \vartheta h) \cdot h \\ &= f(x) + \sum_{j=1}^n \partial_{x_j} f(x + \vartheta h) h_j. \end{aligned}$$

*Dimostrazione.* La funzione composta  $F(s) = f(x + sh)$  è una funzione reale di una variabile reale definita per  $s \in [0, 1]$ , e differenziabile in ogni  $s \in (0, 1)$ . Quindi, dal teorema di Lagrange (Analisi 1), esiste  $\vartheta \in (0, 1)$  tale che

$$F(1) - F(0) = F'(\vartheta)(1 - 0) = F'(\vartheta).$$

Poi  $F(1) = f(x+h)$ ,  $F(0) = f(x)$ , e  $F'(\vartheta) = \nabla f(x + \vartheta h) \cdot h$ . □

**Proposizione 10.3** (Formula di Taylor di ordine 2 con resto di Lagrange). Sia  $D \subseteq \mathbb{R}^n$  un aperto,  $f: D \rightarrow \mathbb{R}$  di classe  $C^2$  in  $D$ . Sia  $h \in \mathbb{R}^n$ ,  $h \neq 0$ , e supponiamo che tutto il segmento di estremi  $x$  e  $x+h$  sia contenuto in  $D$ , cioè  $x + sh \in D$  per ogni  $s \in [0, 1]$ . Allora esiste  $\vartheta \in [0, 1]$  tale che

$$\begin{aligned} f(x+h) &= f(x) + \nabla f(x) \cdot h + \frac{1}{2} D^2 f(x + \vartheta h) h \cdot h \\ &= f(x) + \sum_{j=1}^n \partial_{x_j} f(x) h_j + \frac{1}{2} \sum_{i,j=1}^n \partial_{x_i x_j} f(x + \vartheta h) h_i h_j. \end{aligned}$$

*Dimostrazione.* La funzione composta  $F(s) = f(x + sh)$  è di classe  $C^2$ , quindi, per la formula di Taylor di ordine 2 con resto di Lagrange per funzioni reali di una variabile reale, esiste  $\vartheta \in (0, 1)$  tale che

$$F(1) = F(0) + F'(0)(1 - 0) + \frac{1}{2} F''(\vartheta)(1 - 0)^2 = F(0) + F'(0) + \frac{1}{2} F''(\vartheta).$$

$F(1) = f(x+h)$ ,  $F(0) = f(x)$ ,  $F'(\vartheta) = \nabla f(x + \vartheta h) \cdot h$ , e, dalla (11),  $F''(\vartheta) = D^2 f(x + \vartheta h) h \cdot h$ . □

**Proposizione 10.4** (Formula di Taylor di ordine 1 con resto di Peano). Sia  $D \subseteq \mathbb{R}^n$  un aperto,  $f: D \rightarrow \mathbb{R}$  differenziabile nel punto  $x \in D$ . Allora

$$f(x+h) = f(x) + \nabla f(x) \cdot h + o(|h|) \quad \text{per } |h| \rightarrow 0.$$

*Dimostrazione.* La formula di Taylor di ordine 1 con resto di Peano è la definizione di differenziabilità. □

**Proposizione 10.5** (Formula di Taylor di ordine 2 con resto di Peano). Sia  $D \subseteq \mathbb{R}^n$  un aperto,  $f: D \rightarrow \mathbb{R}$  di classe  $C^2$  in  $D$ . Allora

$$f(x+h) = f(x) + \nabla f(x) \cdot h + \frac{1}{2} D^2 f(x) h \cdot h + o(|h|^2) \quad \text{per } |h| \rightarrow 0.$$

*Dimostrazione.*  $x \in D$ ,  $D$  è aperto, quindi esiste  $r > 0$  tale che  $B(x, r) \subseteq D$ , quindi  $x+h \in D$  per ogni  $h \in \mathbb{R}^n$  di norma  $|h| < r$  (e dunque anche tutto il segmento  $x + sh$ ,  $s \in [0, 1]$ , sta in  $D$ ). Dalla formula di Taylor di ordine 2 con resto di Lagrange, sappiamo che per ogni  $h$  di norma  $|h| < r$  esiste  $\vartheta_h \in (0, 1)$  (dipendente da  $h$ ) tale che

$$f(x+h) = f(x) + \nabla f(x) \cdot h + \frac{1}{2} D^2 f(x + \vartheta_h h) h \cdot h.$$

Dunque, sostituendo l'espressione appena scritta al posto di  $f(x+h)$  (nota che  $f(x)$  e  $\nabla f(x) \cdot h$  si elidono), si ha

$$\begin{aligned} \frac{|f(x+h) - f(x) - \nabla f(x) \cdot h - \frac{1}{2} D^2 f(x) h \cdot h|}{|h|^2} &= \frac{|D^2 f(x + \vartheta_h h) h \cdot h - D^2 f(x) h \cdot h|}{2|h|^2} \\ &= \frac{1}{2|h|^2} \left| \sum_{i,j=1}^n \left( \partial_{x_i x_j} f(x + \vartheta_h h) - \partial_{x_i x_j} f(x) \right) h_i h_j \right| \\ &\leq \frac{1}{2|h|^2} \sum_{i,j=1}^n |\partial_{x_i x_j} f(x + \vartheta_h h) - \partial_{x_i x_j} f(x)| |h_i| |h_j| \\ &= \frac{1}{2} \sum_{i,j=1}^n |\partial_{x_i x_j} f(x + \vartheta_h h) - \partial_{x_i x_j} f(x)| \frac{|h_i|}{|h|} \frac{|h_j|}{|h|} \\ &\leq \frac{1}{2} \sum_{i,j=1}^n |\partial_{x_i x_j} f(x + \vartheta_h h) - \partial_{x_i x_j} f(x)|, \end{aligned}$$

perché  $|h_i|/|h| \leq 1$  per ogni  $i = 1, \dots, n$ . Ora,  $\vartheta_h h \rightarrow 0$  per  $|h| \rightarrow 0$  (perché  $|\vartheta_h h| = |\vartheta_h| |h| \leq |h| \rightarrow 0$ ), quindi  $(x + \vartheta_h h) \rightarrow x$ , e, poiché ogni derivata seconda  $\partial_{x_i x_j} f$  è continua,

$$|\partial_{x_i x_j} f(x + \vartheta_h h) - \partial_{x_i x_j} f(x)| \rightarrow 0 \quad \text{per } |h| \rightarrow 0.$$

Visto che si tratta di una somma di un numero finito di quantità infinitesime,

$$\frac{|f(x+h) - f(x) - \nabla f(x) \cdot h - \frac{1}{2} D^2 f(x) h \cdot h|}{|h|^2} \leq \frac{1}{2} \sum_{i,j=1}^n |\partial_{x_i x_j} f(x + \vartheta_h h) - \partial_{x_i x_j} f(x)| \rightarrow 0 \quad \text{per } |h| \rightarrow 0,$$

cioè  $f(x+h) - f(x) - \nabla f(x) \cdot h - \frac{1}{2} D^2 f(x) h \cdot h = o(|h|^2)$  per  $|h| \rightarrow 0$ .  $\square$

Infine, è utile osservare che, se  $f$  è di classe  $C^1$ , vale anche il teorema fondamentale del calcolo integrale applicato alla funzione  $F(s)$ : con le stesse notazioni di prima,

$$f(x+h) - f(x) = F(1) - F(0) = \int_0^1 F'(s) ds = \int_0^1 \nabla f(x + sh) \cdot h ds.$$

## 11 Massimi e minimi

**Definizione 11.1.** Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione scalare. Un punto  $\bar{x} \in D$  si dice *punto di massimo relativo o locale* per  $f$  se esiste  $r > 0$  tale che

$$f(\bar{x}) \geq f(x) \quad \forall x \in D \text{ tale che } |x - \bar{x}| < r.$$

$\bar{x} \in D$  si dice *punto di minimo relativo o locale* per  $f$  se esiste  $r > 0$  tale che

$$f(\bar{x}) \leq f(x) \quad \forall x \in D \text{ tale che } |x - \bar{x}| < r.$$

$\bar{x} \in D$  si dice *punto di massimo assoluto o globale* per  $f$  in  $D$  se

$$f(\bar{x}) \geq f(x) \quad \forall x \in D.$$

$\bar{x} \in D$  si dice *punto di minimo assoluto o globale* per  $f$  in  $D$  se

$$f(\bar{x}) \leq f(x) \quad \forall x \in D.$$

Se nei punti  $x \neq \bar{x}$  la disuguaglianza è stretta,  $f(\bar{x}) > f(x)$  invece di  $f(\bar{x}) \geq f(x)$  (o  $<$  invece di  $\leq$ ), si dice che  $\bar{x}$  è punto di massimo (o di minimo), locale o globale, *in senso stretto*.  $\square$

La nozione di massimo globale è più forte di quella locale, nel senso che se  $\bar{x}$  è un punto di massimo globale, allora è anche di massimo locale (qualsiasi  $r > 0$  va bene nella definizione); idem per i punti di minimo.

Nel corso di Analisi 1 si è provato il seguente risultato:

**Proposizione 11.2** (Analisi 1). *Sia  $f: D \subseteq \mathbb{R} \rightarrow \mathbb{R}$  una funzione reale di variabile reale,  $x$  un punto di massimo o minimo locale per  $f$ . Supponiamo che  $x \in \text{int}D$ , cioè che esiste  $\delta > 0$  tale che  $(x - \delta, x + \delta) \subseteq D$ , e che  $f$  sia derivabile in  $x$ . Allora  $f'(x) = 0$ .*

Un risultato molto simile vale per le funzioni di più variabili reali, con il gradiente nel ruolo della derivata:

**Proposizione 11.3** (Condizione necessaria del primo ordine). *Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione scalare,  $x \in D$  un punto di massimo o di minimo locale. Se  $x \in \text{int}D$  ed esistono le derivate parziali di  $f$  in  $x$ , allora  $\nabla f(x) = 0$ .*

*Dimostrazione.* Per ipotesi,  $x$  è interno a  $D$ , cioè esiste  $r_1 > 0$  tale che  $B(x, r_1) \subseteq D$ . Supponiamo che  $x$  sia di massimo locale per  $f$ . Allora, per ipotesi, esiste  $r_2 > 0$  tale che  $f(x) \geq f(y)$  per ogni  $y \in D$  tale che  $|x - y| < r_2$ . Sia  $r = \min\{r_1, r_2\}$ , cosicché la palla  $B(x, r)$  è tutta contenuta in  $D$ , e  $f(x) \geq f(y)$  per ogni  $y \in B(x, r)$ . In questo modo possiamo definire la funzione composta

$$F: (-r, r) \rightarrow \mathbb{R}, \quad F(s) = f(x + se_1),$$

e si ha  $F(0) = f(x) \geq f(x + se_1) = F(s)$  per ogni  $s \in (-r, r)$ , cioè  $F$  ha un punto di massimo in  $s = 0$ . Per ipotesi,  $F$  è derivabile in  $s = 0$ , perché (definizione di derivata parziale)  $F'(0) = \partial_{x_1} f(x)$ . Dunque, per la proposizione 11.2 applicata a  $F$  in  $s = 0$ ,

$$\partial_{x_1} f(x) = F'(0) = 0.$$

Ripetendo lo stesso ragionamento per  $F(s) = f(x + se_k)$ , per ogni  $k = 1, 2, \dots, n$ , si ottiene che  $\partial_{x_k} f(x) = 0$  per ogni  $k$ , cioè  $\nabla f(x) = 0$ .

Se  $x$  è punto di minimo locale si procede nello stesso modo. □

**Osservazione 11.4.** L'ipotesi che  $x$  sia interno al dominio è davvero necessaria: si pensi, come esempio molto semplice, alla funzione di una variabile reale

$$f: [0, 1] \rightarrow \mathbb{R}, \quad f(x) = x;$$

il punto  $x = 0$  è di minimo (locale e globale nel dominio  $[0, 1]$ ), il punto  $x = 1$  è di massimo (locale e globale), ma la derivata in tali punti non è zero,  $f'(0) = f'(1) = 1 \neq 0$ .

Per funzioni di più variabili il gradiente in un punto di frontiera non è stato nemmeno definito: nella definizione di derivata parziale avevamo richiesto che il dominio fosse un aperto. Ricordiamo che l'interno di  $D$ ,  $\text{int}D$ , è sempre un aperto, qualunque sia l'insieme  $D$ . Se  $x \in \text{int}D$ , esiste una palla  $B(x, r) \subseteq D$ , quindi la definizione di derivata parziale, di derivata direzionale e di differenziale ha senso in ogni punto interno al dominio.

Osserviamo anche che ci possono essere punti di minimo o massimo in cui la derivata (o il gradiente) non esiste: si pensi alla funzione  $f(x) = |x|$  su  $\mathbb{R}$  o su  $\mathbb{R}^n$ , che ha  $x = 0$  come punto di minimo globale in senso stretto, senza che in  $x = 0$  esista la derivata (in  $\mathbb{R}$ ) o le derivate parziali (in  $\mathbb{R}^n$ ). □

**Definizione 11.5.** Un punto  $x$  interno al dominio di una funzione che ha derivate parziali si dice un *punto critico* o *punto stazionario* se  $\nabla f(x) = 0$ . □

Nella proposizione 11.3 abbiamo dunque provato che i punti di massimo o minimo, interni al dominio, in cui esiste il gradiente sono punti critici.

Ci proponiamo ora di studiare dei criteri per stabilire se un punto critico è un punto di massimo o un punto di minimo o un punto né di massimo né di minimo. A questo scopo introduciamo alcune definizioni e risultati sulle matrici quadrate (che verranno poi applicati alla matrice hessiana).

## 11.1 Matrici definite, semidefinite e indefinite

**Definizione 11.6.** Una matrice quadrata  $A \in \text{Mat}(n \times n)$  si dice

- *definita positiva* se  $Ah \cdot h > 0$  per ogni  $h \in \mathbb{R}^n$ ,  $h \neq 0$ ,
- *definita negativa* se  $Ah \cdot h < 0$  per ogni  $h \in \mathbb{R}^n$ ,  $h \neq 0$ ,
- *semidefinita positiva* se  $Ah \cdot h \geq 0$  per ogni  $h \in \mathbb{R}^n$ ,
- *semidefinita negativa* se  $Ah \cdot h \leq 0$  per ogni  $h \in \mathbb{R}^n$ ,
- *indefinita* se non è semidefinita, cioè se esistono  $h, v \in \mathbb{R}^n$  tali che  $Ah \cdot h > 0$  e  $Av \cdot v < 0$ .  $\square$

Premettiamo al prossimo risultato un lemma che ha interesse di per sé.

**Lemma 11.7.** Sia  $K = \{x \in \mathbb{R}^n : |x| = 1\}$ . Allora  $K$  è chiuso e limitato (e dunque, per il teorema 2.5 di Heine-Borel,  $K$  è compatto).

*Dimostrazione.* Dimostriamo che  $K$  è chiuso. Per definizione di insieme chiuso, dobbiamo provare che il suo complementare

$$\mathcal{A} = \mathbb{R}^n \setminus K = \{x \in \mathbb{R}^n : |x| \neq 1\}$$

è aperto. Sia  $x \in \mathcal{A}$ , cioè  $|x| \neq 1$ . Può essere  $|x| > 1$  oppure  $|x| < 1$ .

Supponiamo che  $|x| > 1$ . Indichiamo  $r = (|x| - 1)/2$ , e mostriamo che la palla  $B(x, r)$  è contenuta in  $\mathcal{A}$ . Sia  $y \in B(x, r)$ . Scriviamo

$$x = y + x - y,$$

da cui, per disuguaglianza triangolare e dal fatto che  $|x - y| < r$ ,

$$|x| = |y + x - y| \leq |y| + |x - y| < |y| + r.$$

Dunque

$$|y| > |x| - r = |x| - \frac{|x| - 1}{2} = \frac{2|x| - |x| + 1}{2} = \frac{|x| + 1}{2} > \frac{1 + 1}{2} = 1,$$

cioè  $|y| > 1$ , e quindi  $y \in \mathcal{A}$ . Abbiamo dimostrato che ogni  $y$  nella palla  $B(x, r)$  appartiene ad  $\mathcal{A}$ , cioè  $B(x, r) \subseteq \mathcal{A}$ .

Supponiamo ora che  $|x| < 1$ . Indichiamo  $r = (1 - |x|)/2$ , e mostriamo che la palla  $B(x, r)$  è contenuta in  $\mathcal{A}$ . Sia  $y \in B(x, r)$ . Scriviamo

$$y = y - x + x,$$

da cui, per disuguaglianza triangolare e dal fatto che  $|x - y| < r$ ,

$$|y| = |y - x + x| \leq |y - x| + |x| < r + |x| = \frac{1 - |x|}{2} + |x| = \frac{1 + |x|}{2} < \frac{1 + 1}{2} = 1,$$

cioè  $|y| < 1$ , e quindi  $y \in \mathcal{A}$ . Abbiamo dimostrato che ogni  $y$  nella palla  $B(x, r)$  appartiene ad  $\mathcal{A}$ , cioè  $B(x, r) \subseteq \mathcal{A}$ .

In conclusione, abbiamo mostrato che per ogni punto  $x \in \mathcal{A}$  (sia per  $|x| > 1$  che per  $|x| < 1$ ) esiste un numero  $r > 0$  tale che  $B(x, r) \subseteq \mathcal{A}$ , e dunque  $\mathcal{A}$  è un aperto (per definizione di insieme aperto). Perciò  $K$  è chiuso (per definizione di insieme chiuso).

Infine,  $K$  è limitato perché è tutto contenuto nella palla centrata nell'origine e raggio, per esempio, 2,  $K \subset B(0, 2)$ .  $\square$

**Proposizione 11.8** (Caratterizzazione delle matrici definite). 1) Una matrice quadrata  $A \in \text{Mat}(n \times n)$  è definita positiva se e solo se esiste  $M > 0$  tale che

$$Ah \cdot h \geq M|h|^2 \quad \forall h \in \mathbb{R}^n.$$

2) Una matrice quadrata  $A \in \text{Mat}(n \times n)$  è definita negativa se e solo se esiste  $M > 0$  tale che

$$Ah \cdot h \leq -M|h|^2 \quad \forall h \in \mathbb{R}^n.$$

*Dimostrazione.* Dimostriamo 1).

⇒) Sia  $A$  definita positiva. Consideriamo la funzione

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(h) = Ah \cdot h.$$

$f$  è continua perché è il prodotto scalare tra le due funzioni continue  $h \mapsto h$  (la funzione identità) e  $h \mapsto Ah$  (funzione lineare rappresentata dalla matrice  $A$ , che è funzione lipschitziana e dunque continua). Dunque è continua sul compatto  $K = \{h \in \mathbb{R}^n : |h| = 1\}$ . Allora, per il teorema 5.6 di Weierstrass,  $f$  ha minimo in  $K$ , cioè esiste  $h_0 \in K$  tale che  $f(h) \geq f(h_0)$  per ogni  $h \in K$ . Indichiamo

$$f(h_0) = M.$$

Poiché  $h_0 \in K$ ,  $|h_0| = 1$ , quindi  $h_0 \neq 0$ , e dunque

$$M = f(h_0) = Ah_0 \cdot h_0 > 0$$

perché, per ipotesi,  $A$  è definita positiva. Abbiamo provato che

$$Ah \cdot h \geq M \quad \forall h \in \mathbb{R}^n, |h| = 1. \quad (12)$$

Sia ora  $h \in \mathbb{R}^n$ ,  $h \neq 0$ . Come in (2), scriviamo  $h$  come prodotto tra la sua norma e il suo versore,

$$h = |h|\hat{h}, \quad \hat{h} = \frac{h}{|h|}.$$

Dunque

$$Ah \cdot h = A(|h|\hat{h}) \cdot (|h|\hat{h}) = |h|A\hat{h} \cdot (|h|\hat{h}) = |h|^2 A\hat{h} \cdot \hat{h},$$

grazie alla proprietà (e) del prodotto scalare (vedi proposizione 1.2) e la proprietà (ii) delle funzioni lineari (vedi definizione 1.9 e proposizione 1.11). Poi  $|\hat{h}| = 1$ , quindi, per la (12),  $A\hat{h} \cdot \hat{h} \geq M$ , da cui

$$Ah \cdot h = |h|^2 A\hat{h} \cdot \hat{h} \geq |h|^2 M.$$

Questa disuguaglianza vale anche per  $h = 0$  (banalmente,  $0 \geq 0$ ), quindi vale per ogni  $h \in \mathbb{R}^n$ .

⇐) Se esiste  $M > 0$  tale che  $Ah \cdot h \geq M|h|^2$  per ogni  $h \in \mathbb{R}^n$ , allora, per  $h \neq 0$ , si ha

$$Ah \cdot h \geq M|h|^2 > 0,$$

e quindi  $A$  è definita positiva.

2) si prova in modo simile, prendendo il massimo di  $f$  su  $K$  invece del minimo e adattando i versi delle disuguaglianze. □

**Proposizione 11.9.** Se  $A = (a_{ij}) \in \text{Mat}(n \times n)$  è una matrice quadrata, allora

$$Ae_k \cdot e_k = a_{kk} \quad \forall k = 1, \dots, n. \quad (13)$$

Di conseguenza,

- se  $A$  è definita positiva, allora  $a_{kk} > 0$  per ogni  $k = 1, \dots, n$ ,
- se  $A$  è definita negativa, allora  $a_{kk} < 0$  per ogni  $k = 1, \dots, n$ ,
- se  $A$  è semidefinita positiva, allora  $a_{kk} \geq 0$  per ogni  $k = 1, \dots, n$ ,
- se  $A$  è semidefinita negativa, allora  $a_{kk} \leq 0$  per ogni  $k = 1, \dots, n$ .

*Dimostrazione.* Per provare (13), basta porre  $h = e_k = (0, \dots, 0, 1, 0, \dots, 0)$  nella formula del lemma 10.1, tenendo conto che  $h_i = 0$  per ogni indice  $i \neq k$ , mentre  $h_i = 1$  per  $i = k$ . Oppure, ricordando che  $Ae_k$  è la  $k$ -esima colonna della matrice  $A$  (vedi proposizione 1.11), si può calcolare

$$Ae_k \cdot e_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{k-1,k} \\ a_{kk} \\ a_{k+1,k} \\ \vdots \\ a_{nk} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_{kk}.$$

Di conseguenza, se  $A$  è definita positiva,  $Ah \cdot h > 0$  per  $h \neq 0$ , e per  $h = e_k$  (nota che  $e_k \neq 0$ ) si ha

$$Ah \cdot h = a_{kk} > 0.$$

Simile negli altri casi. □

Perciò una matrice che nella diagonale principale ha termini sia positivi che negativi non può che essere indefinita. Per stabilire se una matrice definita è definita positiva o negativa, basta conoscere il segno degli elementi della sua diagonale principale. Lo stesso per le matrici semidefinite.

Per le matrici simmetriche di dimensione 2 il segno del determinante determina il carattere della matrice:

**Proposizione 11.10.** *Sia  $A \in \text{Mat}(2 \times 2)$  una matrice simmetrica.*

- Se  $\det A > 0$ , allora  $A$  è definita;
- se  $\det A = 0$ , allora  $A$  è semidefinita;
- se  $\det A < 0$ , allora  $A$  è indefinita.

*Dimostrazione.* Consideriamo una matrice simmetrica  $2 \times 2$

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

con  $a, b, c \in \mathbb{R}$ ; il suo determinante è

$$\det A = ac - b^2,$$

e per ogni vettore  $h = (x, y) \in \mathbb{R}^2$  si ha

$$Ah \cdot h = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ bx + cy \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = (ax + by)x + (bx + cy)y = ax^2 + 2bxy + cy^2.$$

Caso 1: supponiamo che  $\det A > 0$ . Allora  $ac > b^2 \geq 0$ , dunque  $ac > 0$  e quindi  $a \neq 0$ . Perciò

$$p(t) = at^2 + 2bt + c$$

è un polinomio di 2° grado. Il suo discriminante è

$$\Delta = (2b)^2 - 4ac = -4(ac - b^2) = -4 \det A < 0,$$

dunque  $p(t)$  non ha radici reali, e

$$p(t) > 0 \quad \forall t \in \mathbb{R} \quad \text{se } a > 0, \quad p(t) < 0 \quad \forall t \in \mathbb{R} \quad \text{se } a < 0.$$

Caso 1.1: supponiamo che  $a > 0$ . Allora per ogni  $h = (x, y)$  con  $y \neq 0$  si ha

$$Ah \cdot h = ax^2 + 2bxy + cy^2 = y^2 \left[ a \left( \frac{x}{y} \right)^2 + 2b \left( \frac{x}{y} \right) + c \right] = y^2 p \left( \frac{x}{y} \right) > 0,$$

mentre per  $h = (x, y)$  con  $y = 0, x \neq 0$  si ha

$$Ah \cdot h = ax^2 > 0.$$

Dunque  $Ah \cdot h > 0$  per ogni  $h \neq 0$ , cioè  $A$  è definita positiva. Caso 1.2: supponiamo che  $a < 0$ . Allora

$$Ah \cdot h = y^2 p\left(\frac{x}{y}\right) < 0$$

per ogni  $h = (x, y)$  con  $y \neq 0$ , mentre per  $h = (x, 0), x \neq 0$  è

$$Ah \cdot h = ax^2 < 0,$$

e dunque  $Ah \cdot h < 0$  per ogni  $h \neq 0$ , cioè  $A$  è definita negativa. In ogni caso, sia per  $a > 0$  che per  $a < 0$ ,  $A$  è una matrice definita.

Caso 2: supponiamo che  $\det A = 0$ . Caso 2.1: supponiamo che  $a > 0$ . Allora  $p(t)$  è un polinomio di 2° grado, con discriminante  $\Delta = -4 \det A = 0$ , quindi ha due radici reali coincidenti. Dunque  $p(t) \geq 0$  per ogni  $t \in \mathbb{R}$  (perché  $a > 0$ ). Quindi per ogni  $h = (x, y)$  con  $y \neq 0$  si ha

$$Ah \cdot h = y^2 p\left(\frac{x}{y}\right) \geq 0,$$

mentre per  $h = (x, 0)$  si ha

$$Ah \cdot h = ax^2 \geq 0.$$

Dunque  $Ah \cdot h \geq 0$  per ogni  $h \in \mathbb{R}^2$ , e  $A$  è semidefinita positiva. Caso 2.2: supponiamo che  $a < 0$ . Come sopra,  $p(t)$  è un polinomio di 2° grado, con discriminante  $\Delta = 0$ , quindi ha due radici reali coincidenti.  $p(t) \leq 0$  per ogni  $t \in \mathbb{R}$  perché  $a < 0$ . Di conseguenza

$$Ah \cdot h = y^2 p\left(\frac{x}{y}\right) \leq 0$$

per ogni  $h = (x, y)$  con  $y \neq 0$ , mentre  $Ah \cdot h = ax^2 \leq 0$  per ogni  $h = (x, 0)$ . Dunque  $Ah \cdot h \leq 0$  per ogni  $h \in \mathbb{R}^2$ , e  $A$  è semidefinita negativa. Caso 2.3: supponiamo che  $a = 0$ . Allora  $b = 0$  perché  $ac = b^2$ . Dunque

$$Ah \cdot h = cy^2 \quad \forall h = (x, y) \in \mathbb{R}^2.$$

Se  $c \geq 0$ , si ha  $Ah \cdot h = cy^2 \geq 0$  per ogni  $h \in \mathbb{R}^2$ , e  $A$  è semidefinita positiva. Se  $c < 0$ , si ha  $Ah \cdot h = cy^2 \leq 0$  per ogni  $h \in \mathbb{R}^2$ , e  $A$  è semidefinita negativa. In ogni caso, qualunque siano i segni di  $a$  e di  $c$ ,  $A$  è semidefinita.

Caso 3: supponiamo che  $\det A < 0$ . Caso 3.1: supponiamo che  $a \neq 0$ . Allora  $p(t)$  è un polinomio di 2° grado, con discriminante  $\Delta = -4 \det A > 0$ , quindi ha due radici reali distinte. Scegliamo due numeri  $t$  ed  $s$ , uno interno all'intervallo che ha per estremi le due radici e l'altro fuori da tale intervallo, in modo che

$$p(t) < 0, \quad p(s) > 0.$$

Allora, posto  $h = (t, 1)$  e  $v = (s, 1)$ ,

$$Ah \cdot h = at^2 + 2bt + c = p(t) < 0, \quad Av \cdot v = as^2 + 2bs + c = p(s) > 0,$$

quindi  $A$  è una matrice indefinita. Caso 3.2: supponiamo che  $a = 0$ . Allora  $\det A = -b^2$ , e quindi  $b \neq 0$  (perché  $\det A < 0$ ). Perciò, per  $h = (x, 1)$ ,

$$Ah \cdot h = 2bx + c.$$

Scegliamo due numeri  $t$  e  $s$ , uno maggiore e uno minore di  $-c/2b$  (che è la radice del polinomio di 1° grado  $2bx + c$ ), in modo che

$$2bt + c < 0, \quad 2bs + c > 0.$$

Poniamo  $h = (t, 1), v = (s, 1)$ , così che

$$Ah \cdot h = 2bt + c < 0, \quad Av \cdot v = 2bs + c > 0,$$

quindi  $A$  è una matrice indefinita. In ogni caso, qualunque sia il segno di  $a$ ,  $A$  è indefinita.  $\square$

Torniamo all'analisi dei punti di massimo e minimo locali.

## 11.2 Condizioni del secondo ordine

Nel corso di Analisi 1 si è provato il seguente risultato:

**Proposizione 11.11** (Analisi 1). *Sia  $f: D \subseteq \mathbb{R} \rightarrow \mathbb{R}$  una funzione reale di variabile reale,  $x \in \text{int}D$ , cioè  $(x-r, x+r) \subseteq D$  per un certo  $r > 0$ , e  $f$  sia di classe  $C^2$  su  $(x-r, x+r)$ , cioè  $f$  è derivabile due volte con  $f''$  funzione continua nell'intervallo  $(x-r, x+r)$ .*

*Se  $x$  è un punto di minimo locale per  $f$ , allora  $f''(x) \geq 0$ .*

*Se  $x$  è un punto di massimo locale per  $f$ , allora  $f''(x) \leq 0$ .*

Osserviamo che, anche nel caso in cui  $x$  sia punto di minimo in senso stretto, la disuguaglianza  $f''(x) \geq 0$  non diventa, in generale,  $f''(x) > 0$ , ma deve rimanere com'è nell'enunciato. Idem per i punti di massimo in senso stretto. Un esempio di questo è la funzione

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = x^4;$$

$x = 0$  è un punto di minimo locale (e globale) in senso stretto ( $f(0) = 0 < x^4 = f(x)$  per ogni  $x \neq 0$ ), eppure  $f''(0) = 0$ : la derivata seconda nel punto di minimo stretto non è positiva.

Un risultato simile alla 11.11 vale per funzioni di più variabili reali, con la matrice hessiana nel ruolo della derivata seconda, matrici semidefinite positive invece di numeri reali  $\geq 0$ , matrici semidefinite negative invece di numeri reali  $\leq 0$ .

**Proposizione 11.12** (Condizione necessaria del 2° ordine). *Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x \in \text{int}D$ , e  $f$  di classe  $C^2$  su una palla  $B(x, r) \subseteq D$ .*

*Se  $x$  è un punto di minimo locale per  $f$ , allora la matrice hessiana  $D^2f(x)$  di  $f$  in  $x$  è semidefinita positiva.*

*Se  $x$  è un punto di massimo locale per  $f$ , allora  $D^2f(x)$  è semidefinita negativa.*

*Dimostrazione.* Sia  $h \in \mathbb{R}^n$ ,  $h \neq 0$ . La funzione composta  $F(s) = f(x+sh)$  è definita in un certo intervallo  $(-\delta, \delta)$ ,  $\delta > 0$ , per l'osservazione 8.1, ed è di classe  $C^2$  per la proposizione 9.5. Se  $x$  è punto di minimo locale per  $f$ , allora  $s = 0$  è punto di minimo locale per  $F$ , perché

$$F(0) = f(x) \leq f(x+sh) = F(s) \quad \forall s \in (-\delta, \delta),$$

e dunque, per la 11.11,  $F''(0) \geq 0$ . Poi  $F''(0) = D^2f(x)h \cdot h$  per la formula (11), quindi

$$D^2f(x)h \cdot h \geq 0,$$

e la matrice hessiana è semidefinita positiva. In modo del tutto simile, se  $x$  è di massimo locale per  $f$ , allora  $s = 0$  è di massimo locale per  $F$ ,  $F''(0) \leq 0$ , quindi  $D^2f(x)$  è semidefinita negativa.  $\square$

**Proposizione 11.13** (Condizione sufficiente del 2° ordine). *Sia  $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x \in \text{int}D$ , e  $f$  di classe  $C^2$  su una palla  $B(x, r) \subseteq D$ . Supponiamo che  $x$  sia un punto critico per  $f$ , cioè  $\nabla f(x) = 0$ .*

*Se la matrice hessiana  $D^2f(x)$  di  $f$  in  $x$  è definita positiva, allora  $x$  è un punto di minimo locale per  $f$  in senso stretto;*

*se  $D^2f(x)$  è definita negativa, allora  $x$  è un punto di massimo locale per  $f$  in senso stretto;*

*se  $D^2f(x)$  è indefinita, allora  $x$  è un punto critico per  $f$  che non è né di massimo né di minimo.*

*Infine, se  $D^2f(x)$  è semidefinita, non si può trarre nessuna conclusione.*

*Dimostrazione.* Supponiamo che  $D^2f(x)$  sia definita positiva. Allora, come provato nella proposizione 11.8, esiste  $M > 0$  tale che

$$D^2f(x)h \cdot h \geq M|h|^2 \quad \forall h \in \mathbb{R}^n. \quad (14)$$

Poiché  $f$  è di classe  $C^2$  su  $B(x, r)$ , vale la formula di Taylor di ordine 2 con il resto in forma di Peano, che, essendo  $\nabla f(x) = 0$ , si scrive

$$\lim_{\substack{|h| \rightarrow 0 \\ h \in \mathbb{R}^n \setminus \{0\}}} \frac{|f(x+h) - f(x) - \frac{1}{2}D^2f(x)h \cdot h|}{|h|^2} = 0.$$

Quindi (definizione di limite “con  $\varepsilon$  e  $\delta$ ”, scegliendo  $M/4$  nel ruolo di  $\varepsilon$ ) esiste  $\delta > 0$ ,  $\delta < r$ , tale che

$$\frac{|f(x+h) - f(x) - \frac{1}{2} D^2 f(x) h \cdot h|}{|h|^2} \leq \frac{M}{4} \quad \forall h \in \mathbb{R}^n, h \neq 0, |h| \leq \delta,$$

cioè  $|f(x+h) - f(x) - \frac{1}{2} D^2 f(x) h \cdot h| \leq \frac{M}{4} |h|^2$  per ogni  $h$  di norma  $|h| \leq \delta$ , da cui, in particolare,

$$f(x+h) - f(x) - \frac{1}{2} D^2 f(x) h \cdot h \geq -\frac{M}{4} |h|^2 \quad \forall h \in \mathbb{R}^n, |h| \leq \delta. \quad (15)$$

Allora, usando le disuguaglianze (14) e (15),

$$\begin{aligned} f(x+h) - f(x) &= \frac{1}{2} D^2 f(x) h \cdot h + \left( f(x+h) - f(x) - \frac{1}{2} D^2 f(x) h \cdot h \right) \\ &\geq \frac{1}{2} M |h|^2 - \frac{M}{4} |h|^2 \\ &= \frac{M}{4} |h|^2 > 0 \quad \forall h \in \mathbb{R}^n, h \neq 0, |h| \leq \delta, \end{aligned}$$

quindi  $f(x+h) > f(x)$  per ogni punto  $x+h \neq x$  nella palla  $B(x, \delta)$ . Dunque  $x$  è punto di minimo locale in senso stretto.

Se  $D^2 f(x)$  è definita negativa, la dimostrazione che  $x$  è un punto di massimo locale in senso stretto procede allo stesso modo, adattando segni e versi delle disuguaglianze.

Infine, supponiamo che  $D^2 f(x)$  sia indefinita. Se  $x$  fosse di massimo o minimo, per le condizioni necessarie del 2° ordine (proposizione 11.12)  $D^2 f(x)$  sarebbe una matrice semidefinita (positiva o negativa). Ma  $D^2 f(x)$  è indefinita, quindi  $x$  non è né di massimo né di minimo.  $\square$

## 12 Esercizi

Quella che segue non va considerata come una raccolta di esercizi (sono pochi!). A tale scopo si rimanda al volume di esercizi di P. Marcellini e C. Sbordone, *Esercitazioni di Matematica, 2° Volume, parte prima/seconda*, e in generale a tutti gli eserciziari che si trovano in biblioteca e in libreria e a esercizi e temi d'esame che si trovano in rete. Lo scopo di questo capitolo, invece, è quello di dare qualche esempio del modo di procedere nello svolgimento di esercizi di questa tipologia.

**Esercizio 1.** Sia  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = |x|$  la norma euclidea.

- Provare che  $f$  non ammette derivate parziali nell'origine.
- Provare che  $f$  è differenziabile in ogni punto  $x \neq 0$ .
- Calcolare, dove esiste, il differenziale di  $f$ .

**Esercizio 2.** Sia  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$f(x, y) = \begin{cases} \frac{1 - \cos(xy)}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0), \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

Studiarne

- la continuità,
- l'esistenza delle derivate parziali,
- la differenziabilità.
- Scrivere, se esiste, il differenziale in  $(0, 0)$  e in  $(\sqrt{\pi}, \sqrt{\pi})$ .

**Esercizio 3.** Stabilire se la funzione  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$f(x, y) = \begin{cases} \frac{x^3 y^2}{x^6 + y^4} & \text{per } (x, y) \neq (0, 0), \\ 0 & \text{per } (x, y) = (0, 0) \end{cases}$$

è continua.

**Esercizio 4.** (a) Stabilire se la funzione  $f: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ ,

$$f(x, y) = \frac{e^{xy^2} - 1}{x^2 + y^4}$$

si può estendere ad una funzione continua in  $\mathbb{R}^2$ , cioè stabilire se esiste un valore  $c \in \mathbb{R}$  tale che la funzione

$$\tilde{f}: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \tilde{f}(x, y) = \begin{cases} f(x, y) & \text{per } (x, y) \neq (0, 0), \\ c & \text{per } (x, y) = (0, 0) \end{cases}$$

sia continua in  $\mathbb{R}^2$ .

(b) Stabilire se la funzione  $f: \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}$ ,

$$f(x, y, z) = \log \left( 2 + \frac{x^2 + y^2}{(x^2 + y^2 + z^2)^{1/2}} \right),$$

si può estendere ad una funzione continua in  $\mathbb{R}^3$ .

(c) La stessa domanda per la funzione  $f: \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}$ ,

$$f(x, y, z) = \frac{e^{x^2yz} - 1}{x^4 + y^4 + z^4}.$$

**Esercizio 5.** Trovare i punti di massimo e minimo, locali e globali, delle seguenti funzioni.

- (a)  $f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = y^2 \sin x;$       (b)  $f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = (1 + y^2) \sin x;$   
 (c)  $f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) = \cos \left( \frac{\pi}{2} \frac{|x|^2}{1 + |x|^2} \right);$       (d)  $f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) = \cos \left( \frac{\pi}{2} \frac{|x|}{1 + |x|} \right).$

**Esercizio 6.** Sia

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} \frac{x^3y - xy^3}{x^2 + y^2} & \text{per } (x, y) \neq (0, 0), \\ 0 & \text{per } (x, y) = (0, 0). \end{cases}$$

- (a) Provare che  $f$  è continua e differenziabile.  
 (b) Provare che  $f$  ha derivate seconde in  $(0, 0)$ , ma  $\partial_{xy}f(0, 0) \neq \partial_{yx}f(0, 0)$ .  
 (c) Trovare, se esistono, massimi e minimi, locali e globali, di  $f$ .

## 12.1 Svolgimento degli esercizi

**Esercizio 1.** (a) Per definizione, esiste la derivata parziale di  $f$  rispetto a  $x_k$  nel punto 0 se la funzione

$$\varphi(t) = f(0, \dots, 0, t, 0, \dots, 0), \quad t \in \mathbb{R},$$

è derivabile in  $t = 0$ . Ora,  $\varphi(t) = |t|$ , quindi  $\varphi$  non è derivabile in  $t = 0$ , perché, come noto, il limite destro e il limite sinistro del rapporto incrementale  $\frac{\varphi(t) - \varphi(0)}{t - 0}$  sono diversi. Dunque non esiste la derivata parziale di  $f$  rispetto a  $x_k$  in  $x = 0$ , per ogni  $k = 1, \dots, n$ .

(b) Sia  $x \neq 0$ . La funzione “radice quadrata di un numero non negativo”,  $[0, \infty) \rightarrow \mathbb{R}, s \mapsto \sqrt{s}$ , è derivabile in ogni punto  $s > 0$ , con derivata  $1/2\sqrt{s}$ . Quindi la funzione (di una variabile reale)

$$\varphi(t) = f(t, x_2, \dots, x_n) = \sqrt{t^2 + x_2^2 + \dots + x_n^2}$$

è derivabile nei punti  $t \in \mathbb{R}$  tali che  $t^2 + x_2^2 + \dots + x_n^2 > 0$ . La sua derivata è

$$\varphi'(t) = \frac{2t}{2\sqrt{t^2 + x_2^2 + \dots + x_n^2}},$$

e la derivata parziale di  $f$  rispetto a  $x_1$  nel punto  $x$  è, per definizione,  $\varphi'(t)$  calcolata per  $t = x_1$ , cioè

$$\partial_{x_1}f(x) = \varphi'(x_1) = \frac{x_1}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}} = \frac{x_1}{|x|}.$$

Notiamo che questa frazione ha senso perché  $x \neq 0$ . In modo analogo si ottiene l'esistenza della derivata parziale rispetto a  $x_k$  per ogni  $k = 1, \dots, n$ , con

$$\partial_{x_k} f(x) = \frac{x_k}{|x|}.$$

Proviamo che la derivata parziale  $\partial_{x_k} f: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$  è una funzione continua.  $\partial_{x_k} f$  è il prodotto di due funzioni: la funzione  $k$ -esima proiezione  $x \mapsto x_k$  e la funzione  $x \mapsto 1/|x|$ . La  $k$ -esima proiezione è lipschitziana (vedi proposizione 4.6(e)), e dunque continua, in tutto  $\mathbb{R}^n$ . La funzione  $x \mapsto 1/|x|$  è la composizione della funzione  $x \mapsto |x|$ , la norma, con la funzione  $t \mapsto 1/t$ , il reciproco di un numero reale. La norma è lipschitziana, e quindi continua, su tutto  $\mathbb{R}^n$  (vedi proposizione 4.6(c)), mentre la funzione  $t \mapsto 1/t$  è continua sul suo dominio  $\{t \in \mathbb{R} : t \neq 0\}$ . Di conseguenza (vedi la proposizione 5.1), la funzione composta  $x \mapsto 1/|x|$  è continua sul dominio  $D = \{x \in \mathbb{R}^n : |x| \neq 0\}$ . Infine, il prodotto di funzioni continue è continuo (proposizione 5.3 e osservazione 5.4), perciò  $x \mapsto x_k/|x|$  è continua su  $D$ .

Dal momento che tutte le derivate parziali di  $f$  sono continue in  $D$ ,  $f$  è differenziabile in ogni punto  $x \neq 0$  per il teorema del differenziale.

(c) Il differenziale di  $f$  nel punto  $x \neq 0$  applicato al vettore  $h \in \mathbb{R}^n$  è

$$df(x)[h] = \nabla f(x) \cdot h = \sum_{k=1}^n \partial_{x_k} f(x) h_k = \sum_{k=1}^n \frac{x_k}{|x|} h_k = \frac{x \cdot h}{|x|} = \left( \frac{x}{|x|} \right) \cdot h = \hat{x} \cdot h,$$

dove  $\hat{x} = x/|x|$  è il versore di  $x$ . Il differenziale della funzione norma euclidea nel punto  $x \neq 0$  è la funzione prodotto scalare per il versore  $\hat{x}$ .  $\square$

**Esercizio 2.** (a) La funzione  $\mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto xy$  è continua perché è il prodotto delle due proiezioni  $(x, y) \mapsto x$  e  $(x, y) \mapsto y$ , che sono funzioni continue. La funzione coseno è continua in  $\mathbb{R}$ , il denominatore  $x^2 + y^2$  è il quadrato della norma, quindi (con ragionamento simile a quello fatto nell'esercizio 1)  $f$  è continua in  $\mathbb{R}^2 \setminus \{(0, 0)\}$ .

Resta da studiare la continuità di  $f$  nel punto  $(0, 0)$ . Lo sviluppo di Taylor del coseno intorno a zero basterebbe per provare la continuità della  $f$  nell'origine; tuttavia possiamo ricordare che vale una disuguaglianza più esplicita che vale in tutto  $\mathbb{R}$  (Analisi 1), che ha un certo interesse di per sé, che qui ridimostriamo per completezza:

$$|1 - \cos t| \leq \frac{1}{2} t^2 \quad \forall t \in \mathbb{R}. \quad (16)$$

*Dimostrazione della (16).* Applichiamo la formula di Taylor di ordine 2 con resto in forma di Lagrange alla funzione  $g(t) = \cos t$ : fissato  $t \in \mathbb{R}$ , esiste  $\vartheta$  nell'intervallo aperto di estremi 0 e  $t$  tale che

$$g(t) = g(0) + g'(0)t + \frac{1}{2} g''(\vartheta) t^2.$$

$g(0) = \cos 0 = 1$ ,  $g'(0) = -\sin 0 = 0$ ,  $g''(\vartheta) = -\cos \vartheta$ , da cui

$$1 - \cos t = \frac{1}{2} t^2 \cos \vartheta,$$

e, visto che  $|\cos \vartheta| \leq 1$  qualunque sia  $\vartheta$ ,

$$|1 - \cos t| \leq \frac{1}{2} t^2. \quad \square$$

Torniamo alla  $f$  dell'esercizio. Essendo  $\frac{x^2}{x^2+y^2} \leq 1$ ,

$$\begin{aligned} |f(x, y) - f(0, 0)| &= \frac{|1 - \cos(xy)|}{x^2 + y^2} \\ &\leq \frac{\frac{1}{2} (xy)^2}{x^2 + y^2} = \frac{1}{2} \frac{x^2}{x^2 + y^2} y^2 \\ &\leq \frac{1}{2} y^2 \leq \frac{1}{2} |(x, y)|^2 \quad \forall (x, y) \in \mathbb{R}^2, \end{aligned}$$

quindi  $f(x, y) \rightarrow 0$  per  $|(x, y) - (0, 0)| \rightarrow 0$ , cioè  $f$  è continua in  $(0, 0)$ .<sup>2</sup> In conclusione,  $f$  è continua in tutto il suo dominio  $\mathbb{R}^2$ .

Se, invece della (16), avessimo usato lo sviluppo di Taylor del coseno  $\cos t = 1 - t^2/2 + O(t^4)$  per  $t \rightarrow 0$ , o il limite notevole  $(1 - \cos t)/t^2 \rightarrow 1/2$  per  $t \rightarrow 0$ , l'esercizio si sarebbe svolto essenzialmente nello stesso modo, con l'unica differenza che nell'usare la disuguaglianza  $|1 - \cos(xy)| \leq C(xy)^2$ ,  $C > 0$ , si deve specificare che essa vale per  $|(x, y)| < \delta$ , per un certo  $\delta > 0$ ; questo  $\delta$  verrebbe dal fatto che nel trattare il coseno si sta facendo un limite per  $t \rightarrow 0$ .

(b) Calcoliamo la derivata parziale rispetto a  $x$  della  $f$  in un punto  $(x, y) \neq (0, 0)$ , con le usuali regole di derivazione di Analisi 1 (qui  $x$  è la variabile, mentre  $y$  ha il ruolo di una costante):

$$\partial_x f(x, y) = \frac{y \sin(xy)(x^2 + y^2) - (1 - \cos(xy)) 2x}{(x^2 + y^2)^2} = \frac{y \sin(xy)}{x^2 + y^2} - \frac{(1 - \cos(xy)) 2x}{(x^2 + y^2)^2}.$$

In modo simile si calcola

$$\partial_y f(x, y) = \frac{x \sin(xy)}{x^2 + y^2} - \frac{(1 - \cos(xy)) 2y}{(x^2 + y^2)^2}.$$

Per calcolare, se esiste,  $\partial_x f(0, 0)$ , scriviamo il rapporto incrementale come richiede la definizione di derivata parziale,

$$\frac{f(x, 0) - f(0, 0)}{x - 0} = \frac{f(x, 0)}{x} = \frac{1 - \cos(x)}{x^2 + 0^2} \frac{1}{x} = \frac{0}{x^3} = 0 \quad \forall x \neq 0.$$

Questo rapporto incrementale ha limite per  $x \rightarrow 0, x \neq 0$ , e vale 0, dunque  $f$  ha derivata parziale rispetto a  $x$  nell'origine, e

$$\partial_x f(0, 0) = 0.$$

(Si arriva alla stessa conclusione — e in realtà è la stessa cosa — se si considera la funzione  $x \mapsto f(x, 0)$ , si osserva che è derivabile e se ne calcola la derivata in  $x = 0$ ). In modo del tutto analogo si prova che esiste e vale zero anche  $\partial_y f(0, 0)$ . In conclusione, le derivate parziali di  $f$  esistono in tutti i punti di  $\mathbb{R}^2$ .

(c) Essendo somme, prodotti e composizioni di funzioni continue, le derivate parziali di  $f$  sono funzioni continue in  $\mathbb{R}^2 \setminus \{(0, 0)\}$ , perciò, dal teorema del differenziale,  $f$  è differenziabile in tutti i punti  $(x, y) \neq (0, 0)$ . Resta da studiare la differenziabilità nell'origine.

*Primo modo.* Studiamo la continuità delle derivate parziali: se mostriamo che sono continue nell'origine, il teorema del differenziale implicherà la differenziabilità di  $f$  nell'origine. Abbiamo provato (punto (b)) che

$$\partial_x f(x, y) = \begin{cases} \frac{y \sin(xy)}{x^2 + y^2} - \frac{(1 - \cos(xy)) 2x}{(x^2 + y^2)^2} & \text{se } (x, y) \neq (0, 0), \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

Per disuguaglianza triangolare,

$$|\partial_x f(x, y) - \partial_x f(0, 0)| \leq \frac{|y| |\sin(xy)|}{x^2 + y^2} + \frac{|1 - \cos(xy)| 2|x|}{(x^2 + y^2)^2}.$$

Ricordiamo questa disuguaglianza:

$$|\sin t| \leq |t| \quad \forall t \in \mathbb{R}. \quad (17)$$

*Dimostrazione della (17).* Fissato  $t \in \mathbb{R}$ , dal Teorema di Lagrange esiste  $\vartheta$  nell'intervallo aperto di estremi 0 e  $t$  tale che

$$\sin t - \sin 0 = \sin' \vartheta (t - 0),$$

cioè  $\sin t = t \cos \vartheta$ , da cui

$$|\sin t| = |t| |\cos \vartheta| \leq |t|. \quad \square$$

<sup>2</sup>Notiamo che, di fatto, per provare la continuità abbiamo dimostrato che vale una disuguaglianza del tipo (7).

Torniamo alla stima di  $|\partial_x f(x, y) - \partial_x f(0, 0)|$ : applicando la (17) per il primo termine si ha

$$\frac{|y| |\sin(xy)|}{x^2 + y^2} \leq \frac{|y| |x| |y|}{x^2 + y^2} = \frac{y^2}{x^2 + y^2} |x| \leq |x| \leq |(x, y)|,$$

mentre applicando la (16) per il secondo termine

$$\frac{|1 - \cos(xy)| 2|x|}{(x^2 + y^2)^2} \leq \frac{x^2 y^2 |x|}{(x^2 + y^2)^2} = \frac{x^2}{x^2 + y^2} \frac{y^2}{x^2 + y^2} |x| \leq |x| \leq |(x, y)|.$$

Quindi

$$|\partial_x f(x, y) - \partial_x f(0, 0)| \leq \frac{|y| |\sin(xy)|}{x^2 + y^2} + \frac{|1 - \cos(xy)| 2|x|}{(x^2 + y^2)^2} \leq 2|(x, y)|,$$

e  $\partial_x f$  è continua in  $(0, 0)$ . Similmente si prova che anche  $\partial_y f(x, y) \rightarrow \partial_y f(0, 0) = 0$  per  $(x, y) \rightarrow (0, 0)$ . Per il teorema del differenziale,  $f$  è differenziabile in  $(0, 0)$ .

*Secondo modo.* Direttamente dalla definizione di differenziabilità:  $f$  è differenziabile in  $(0, 0)$  se esiste una funzione lineare  $df(0): \mathbb{R}^2 \rightarrow \mathbb{R}$  tale che

$$f(0 + h) - f(0) - df(0)[h] = o(|h|) \quad \text{per } |h| \rightarrow 0.$$

Il differenziale di  $f$  in  $0$ , se esiste, è la funzione nulla,

$$df(0)[h] = \nabla f(0) \cdot h = \partial_x f(0, 0)h_1 + \partial_y f(0, 0)h_2 = 0 \quad \forall h = (h_1, h_2) \in \mathbb{R}^2,$$

perché le derivate parziali di  $f$  nell'origine sono entrambe  $= 0$ . Applicando la (16), calcoliamo, per  $h = (h_1, h_2) \in \mathbb{R}^2, h \neq 0$ ,

$$\begin{aligned} \frac{|f(0 + h) - f(0) - df(0)[h]|}{|h|} &= \frac{|f(h) - 0 - 0|}{|h|} = \frac{|1 - \cos(h_1 h_2)|}{h_1^2 + h_2^2} \frac{1}{\sqrt{h_1^2 + h_2^2}} \\ &\leq \frac{\frac{1}{2}(h_1 h_2)^2}{h_1^2 + h_2^2} \frac{1}{\sqrt{h_1^2 + h_2^2}} = \frac{1}{2} \frac{h_1^2}{h_1^2 + h_2^2} \frac{|h_2|}{|h|} |h_2| \\ &\leq \frac{1}{2} |h_2| \leq \frac{1}{2} |h|, \end{aligned}$$

dove, come al solito, si è usato il fatto che  $\frac{|h_1|}{|h|}$  e  $\frac{|h_2|}{|h|}$  sono  $\leq 1$ . Perciò

$$\frac{|f(0 + h) - f(0) - df(0)[h]|}{|h|} \rightarrow 0 \quad \text{per } |h| \rightarrow 0, \quad h \in \mathbb{R}^2, \quad h \neq 0,$$

e la differenziabilità di  $f$  in  $(0, 0)$  è provata in modo diretto.

In conclusione,  $f$  è differenziabile in tutti i punti di  $\mathbb{R}^2$ .

(c) Il differenziale di  $f$  nel punto  $(x_0, y_0)$  è la funzione lineare

$$df(x_0, y_0): \mathbb{R}^2 \rightarrow \mathbb{R}, \quad df(x_0, y_0)[h] = \nabla f(x_0, y_0) \cdot h = \partial_x f(x_0, y_0) h_1 + \partial_y f(x_0, y_0) h_2.$$

In  $(0, 0)$  abbiamo già calcolato  $\partial_x f(0, 0) = \partial_y f(0, 0) = 0$ , quindi il differenziale di  $f$  in  $(0, 0)$  è la funzione nulla,  $df(0)[h] = 0$  per ogni  $h \in \mathbb{R}^2$ .

Sostituendo  $x = y = \sqrt{\pi}$  nella formula delle derivate parziali si calcola che

$$\partial_x f(\sqrt{\pi}, \sqrt{\pi}) = \partial_y f(\sqrt{\pi}, \sqrt{\pi}) = -\frac{1}{\pi^{3/2}},$$

quindi il differenziale di  $f$  in  $(\sqrt{\pi}, \sqrt{\pi})$  è la funzione lineare

$$df(\sqrt{\pi}, \sqrt{\pi})[h] = -\frac{h_1 + h_2}{\pi^{3/2}}. \quad \square$$

**Esercizio 3.**  $f$  è continua in ogni punto diverso dall'origine perché è fatta di somme, prodotti e composizioni di funzioni continue. Resta da studiare la continuità di  $f$  in  $(0, 0)$ . Sia numeratore che denominatore tendono a zero per  $(x, y) \rightarrow (0, 0)$ , e con le usuali maggiorazioni non riusciamo a provare che  $f(x, y) \rightarrow 0$  quando  $(x, y) \rightarrow (0, 0)$ . Questo ci fa venire il sospetto che  $f$  non sia continua in  $(0, 0)$ . Cerchiamo di dimostrarlo, applicando il criterio di non continuità basato sulla composizione di  $f$  con curve continue, e cominciamo con le curve più semplici, le rette passanti per l'origine. Consideriamo la retta di equazione  $y = 0$ , cioè l'asse delle  $x$ , che corrisponde alla curva

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \varphi(t) = (t, 0).$$

$\varphi$  è una funzione continua, e  $\varphi(0) = (0, 0)$ . La composizione è

$$f(\varphi(t)) = f(t, 0) = \frac{t^3 0^2}{t^6 + 0^4} = \frac{0}{t^6} = 0 \quad \forall t \neq 0,$$

e quindi  $\lim_{(t \rightarrow 0)} f(\varphi(t))$  effettivamente esiste ed è  $= 0$ . Lo stesso per l'asse delle  $y$ , retta di equazione  $x = 0$ , corrispondente a  $\varphi(t) = (0, t)$ . Consideriamo la retta passante per l'origine di equazione  $y = mx$ , cioè

$$\varphi(t) = (t, mt).$$

La composizione è

$$f(\varphi(t)) = f(t, mt) = \frac{t^3 (mt)^2}{t^6 + (mt)^4} = \frac{t^5 m^2}{t^4 (m^4 + t^2)} = \frac{m^2}{m^4 + t^2} t,$$

e quindi

$$\lim_{t \rightarrow 0} f(\varphi(t)) = 0$$

per ogni  $m$ . Quindi le rette non sono le curve adatte ad applicare il criterio della proposizione 5.2.

Se ci mettiamo a fare tentativi con parabole, cubiche, quartiche, e tutte le curve del tipo  $y = ax^n$ ,  $n = 1, 2, 3, \dots$ , ci viene sempre  $f(\varphi(t)) \rightarrow 0$  per  $t \rightarrow 0$ , come con le rette, e il criterio di non continuità non si applica.

Ma è la  $f$  stessa a suggerirci la curva giusta da provare: nella frazione i termini che compaiono sono  $x^3$ ,  $y^2$  e i loro quadrati. Proviamo dunque a considerare la curva  $y^2 = x^3$ , cioè

$$\varphi(t) = (t^2, t^3).$$

$\varphi$  è continua e  $\varphi(0) = (0, 0)$ . La composizione è

$$f(\varphi(t)) = f(t^2, t^3) = \frac{(t^2)^3 (t^3)^2}{(t^2)^6 + (t^3)^4} = \frac{t^6 t^6}{t^{12} + t^{12}} = \frac{t^{12}}{2t^{12}} = \frac{1}{2} \quad \forall t \neq 0,$$

e quindi

$$\lim_{t \rightarrow 0} f(\varphi(t)) = \frac{1}{2} \neq f(0, 0).$$

Abbiamo trovato una curva per cui si applica il criterio di non continuità, e di conseguenza  $f$  non è continua in  $(0, 0)$ . □

**Nota.** In questi esercizi stiamo usando la variabile  $t$  soltanto per rendere evidente che stiamo applicando il criterio di non continuità. Di solito, nel "test delle rette", quello che si fa concretamente è semplicemente riscrivere la  $f(x, y)$  con  $mx$  al posto di  $y$ . Di fatto è la stessa cosa: scrivendo  $x$  al posto di  $t$ , la composizione  $f(\varphi(t)) = f(t, mt) = f(x, mx)$  è quanto si ottiene sostituendo  $mx$  a  $y$  nell'espressione della  $f(x, y)$ . Questo vale non solo per le rette, ma per tutte le curve che hanno un'equazione del tipo  $y = (\text{funzione di } x)$  o  $x = (\text{funzione di } y)$ .

**Esercizio 4.** (a) Supponiamo che esista un'estensione continua  $\tilde{f}$ . Consideriamo la retta di equazione  $y = 0$ , cioè  $\varphi(t) = (t, 0)$ , e calcoliamo la composizione per  $t \neq 0$ ,

$$\tilde{f}(\varphi(t)) = \tilde{f}(t, 0) = f(t, 0) = \frac{e^{t^2} - 1}{t^2 + 0^4} = \frac{0}{t^2} = 0 \quad \forall t \neq 0,$$

quindi

$$\lim_{t \rightarrow 0} f(\varphi(t)) = \lim_{t \rightarrow 0} f(t, 0) = 0.$$

Ora consideriamo la curva suggerita dalla struttura dell'argomento dell'esponenziale, la parabola di equazione  $x = y^2$ ,

$$\psi(t) = (t^2, t),$$

che è continua e soddisfa  $\psi(0) = (0, 0)$ . La composizione è

$$f(\psi(t)) = \frac{e^{t^2 t^2} - 1}{(t^2)^2 + t^4} = \frac{e^{t^4} - 1}{2t^4},$$

quindi (limite notevole  $(e^s - 1)/s \rightarrow 1$  per  $s \rightarrow 0$ )

$$\lim_{t \rightarrow 0} f(\psi(t)) = \lim_{t \rightarrow 0} \frac{e^{t^4} - 1}{2t^4} = \frac{1}{2}.$$

Dal criterio 5.2,  $\tilde{f}$  non è continua in  $(0, 0)$ , assurdo. Di conseguenza  $f$  non ammette estensione continua.

(b)  $f$  è la composizione

$$f(x, y, z) = \log(g(x, y, z)), \quad g(x, y, z) = 2 + \frac{x^2 + y^2}{(x^2 + y^2 + z^2)^{1/2}}.$$

La funzione  $\log$  è continua nel suo dominio  $(0, \infty)$ . La funzione  $g$  è  $\geq 2$ , quindi  $g(x, y, z)$  appartiene al dominio di  $\log$  per ogni  $(x, y, z)$  nel dominio di  $g$ . Studiamo  $g$ :

$$|g(x, y, z) - 2| = \frac{x^2 + y^2}{(x^2 + y^2 + z^2)^{1/2}} \leq \frac{x^2 + y^2 + z^2}{(x^2 + y^2 + z^2)^{1/2}} = \sqrt{x^2 + y^2 + z^2} = |(x, y, z)|$$

e quindi

$$|g(x, y, z) - 2| \rightarrow 0 \quad \text{per } |(x, y, z)| \rightarrow 0.$$

Di conseguenza

$$f(x, y, z) = \log(g(x, y, z)) \rightarrow \log 2 \quad \text{per } (x, y, z) \rightarrow 0,$$

e la funzione

$$\tilde{f}: \mathbb{R}^3 \rightarrow \mathbb{R}, \quad \tilde{f}(x, y, z) = \begin{cases} f(x, y, z) & \text{per } (x, y, z) \neq (0, 0, 0), \\ \log 2 & \text{per } (x, y, z) = (0, 0, 0) \end{cases}$$

è continua, ed è l'estensione cercata.

(c) Supponiamo che l'estensione continua  $\tilde{f}$  esista. Consideriamo la retta

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}^3, \quad \varphi(t) = (t, 0, 0),$$

che è continua e soddisfa  $\varphi(0) = (0, 0, 0)$ . La composizione è

$$\tilde{f}(\varphi(t)) = \tilde{f}(t, 0, 0) = f(t, 0, 0) = \frac{e^{t^2 0 0} - 1}{t^4 + 0^4 + 0^4} = \frac{0}{t^4} = 0 \quad \forall t \neq 0,$$

da cui

$$\lim_{t \rightarrow 0} f(\varphi(t)) = 0.$$

Ora consideriamo la retta

$$\psi: \mathbb{R} \rightarrow \mathbb{R}^3, \quad \psi(t) = (t, t, t),$$

che è continua e soddisfa  $\psi(0) = (0, 0, 0)$ . La composizione, per  $t \neq 0$ , dà

$$\tilde{f}(\psi(t)) = \tilde{f}(t, t, t) = f(t, t, t) = \frac{e^{t^2 t} - 1}{t^4 + t^4 + t^4} = \frac{e^{t^3} - 1}{3t^4} \quad \forall t \neq 0,$$

da cui (usando il limite notevole per l'esponenziale)

$$\lim_{t \rightarrow 0} \tilde{f}(\psi(t)) = \frac{1}{3}.$$

Questa volta il “test delle rette” è stato sufficiente per dimostrare che  $\tilde{f}$  non è continua in  $(0, 0, 0)$ . Siamo arrivati ad un assurdo, e quindi  $f$  non ammette nessuna estensione continua.  $\square$

**Esercizio 5.** (a) Innanzitutto, notiamo che per  $x = \pi/2$  si ha

$$f\left(\frac{\pi}{2}, y\right) = y^2 \sin\left(\frac{\pi}{2}\right) = y^2,$$

che  $\rightarrow \infty$  per  $y \rightarrow \pm\infty$ , quindi

$$\sup_{(x,y) \in \mathbb{R}^2} f(x, y) = \infty.$$

Questo toglie subito la speranza di trovare un punto di massimo globale: non ce ne sono, perché la funzione non è limitata superiormente. Similmente, per  $x = 3\pi/2$  si ha  $f(3\pi/2, y) = -y^2$ , quindi  $f$  non è limitata nemmeno inferiormente, e non esistono punti di minimo globale.

Cerchiamo ora punti di minimo e massimo locale. La funzione è differenziabile in ogni punto  $(x, y) \in \mathbb{R}^2$ , quindi i punti di minimo o massimo locale sono necessariamente punti critici (proposizione 11.3). Le derivate parziali sono

$$\partial_x f(x, y) = y^2 \cos x, \quad \partial_y f(x, y) = 2y \sin x.$$

$(x, y)$  è punto critico per  $f$  se

$$\begin{cases} y^2 \cos x = 0 \\ 2y \sin x = 0. \end{cases}$$

La prima equazione è soddisfatta se almeno uno fra  $y$  e  $\cos x$  è 0. Se  $y \neq 0$ , deve essere  $\cos x = 0$ ; ma allora  $\sin x = \pm 1 \neq 0$ , quindi la seconda equazione non è soddisfatta. Quindi non ci sono punti critici aventi  $y \neq 0$ . Se invece  $y = 0$ , entrambe sono soddisfatte. Perciò i punti critici di  $f$  sono tutti e soli i punti  $(x, 0)$ , quelli cioè che stanno sull'asse delle  $x$ . Ora rimane da capire quali, tra questi, sono di massimo, quali di minimo, quali né di massimo né di minimo.

Studiamo il segno di  $f$ . Innanzitutto,  $f$  è nulla su tutto la retta dei punti critici,  $f(x, 0) = 0$  per ogni  $x \in \mathbb{R}$ . Questo esclude immediatamente l'esistenza di massimi e minimi locali in senso stretto.

Se poi  $\bar{x}$  è un numero per cui  $\sin \bar{x} > 0$ , cioè se  $m\pi < \bar{x} < (m+1)\pi$  per un certo intero pari  $m$ , allora  $\sin x > 0$  per tutti gli  $x \in (m\pi, (m+1)\pi)$ , e quindi, in particolare, esiste  $r > 0$  tale che

$$\sin x > 0 \quad \forall x \in (\bar{x} - r, \bar{x} + r)$$

(basta prendere  $r$  più piccolo sia della distanza tra  $\bar{x}$  e  $m\pi$ , sia della distanza tra  $\bar{x}$  e  $(m+1)\pi$ ). Dunque per ogni punto  $(x, y)$  nella palla  $|(x, y) - (\bar{x}, 0)| < r$  si ha  $\sin x > 0$ ,  $y^2 \geq 0$ , quindi  $f(x, y) \geq 0$ . In altre parole,

$$f(\bar{x}, 0) = 0 \leq f(x, y) \quad \forall (x, y) \in \mathbb{R}^2 \text{ t.c. } |(x, y) - (\bar{x}, 0)| < r,$$

perciò  $(\bar{x}, 0)$  è un punto di minimo locale.

Nello stesso modo si prova che se  $\sin \bar{x} < 0$ , allora  $f(x, y) \leq 0$  per ogni punto  $(x, y)$  che dista meno di  $r$  dal punto  $(\bar{x}, 0)$ , e dunque  $(\bar{x}, 0)$  è punto di massimo locale.

Infine, se  $\sin \bar{x} = 0$ , cioè se  $\bar{x} = m\pi$  per un certo intero  $m$ , il punto  $(\bar{x}, 0)$  non è né di minimo né di massimo locale: infatti ogni palla centrata in  $(\bar{x}, 0)$ , raggio qualsiasi, contiene sempre un punto  $(x_1, y_1)$  in cui  $f > 0$  e un punto  $(x_2, y_2)$  in cui  $f < 0$ , e quindi  $f(\bar{x}, 0) = 0$  non è né il valore più grande né il valore più piccolo della palla, per nessuna palla. Per esempio, il punto  $(\bar{x} + \varepsilon, \varepsilon)$  e il punto  $(\bar{x} - \varepsilon, \varepsilon)$  stanno fuori dall'asse delle  $x$ , cosicché  $y^2 > 0$ , e stanno uno al di qua e l'altro al di là della retta verticale  $\{(m\pi, y) : y \in \mathbb{R}\}$ , quindi  $\sin x_1$  e  $\sin x_2$  sono uno positivo e l'altro negativo. Dato un qualsiasi raggio

$r > 0$ , basta scegliere  $0 < \varepsilon < r$  affinché i due punti in questione stiano nella palla di centro  $(\bar{x}, 0)$  e raggio  $r$ .

(b) La non esistenza di massimi e minimi globali è esattamente come per la funzione del punto (a).  $f$  è differenziabile ovunque, quindi i punti di massimo e minimo locali li cerchiamo tra i punti critici.  $(x, y)$  è punto critico se

$$\begin{cases} (1 + y^2) \cos x = 0, \\ 2y \sin x = 0. \end{cases}$$

La prima equazione è soddisfatta solo se  $\cos x = 0$ ; in tal caso,  $\sin x \neq 0$ , e la seconda è soddisfatta solo se  $y = 0$ . Quindi i punti critici di  $f$  sono i punti

$$\left(\frac{\pi}{2} + m\pi, 0\right), \quad m \in \mathbb{Z},$$

tutti sull'asse delle  $x$ , distanziati  $\pi$  l'uno dall'altro.

Se  $\bar{x} = \frac{\pi}{2} + m\pi$ ,  $f(\bar{x}, 0) = 0$ . Nei punti  $(\bar{x} + \varepsilon, \varepsilon)$  e  $(\bar{x} - \varepsilon, \varepsilon)$  la  $f$  è  $> 0$  nell'uno e  $< 0$  nell'altro. Quindi (come nel punto (a) di prima) questi punti critici non sono né di minimo né di massimo locale. In conclusione, non ci sono punti di massimo o di minimo, né globali né locali.

(c) L'argomento del coseno è  $= 0$  in  $x = 0$ , è  $> 0$  in ogni  $x \neq 0$ , ed è sempre  $< \pi/2$ , perché

$$0 < \frac{|x|^2}{1 + |x|^2} < 1 \quad \forall x \in \mathbb{R}^n, \quad x \neq 0.$$

Poiché  $\cos t \in (0, 1)$  per  $t \in (0, \pi/2)$ , e  $\cos 0 = 1$ , si ha

$$0 < f(x) < 1 \quad \forall x \in \mathbb{R}^n, \quad x \neq 0, \quad f(0) = 1.$$

Dunque  $f(0) > f(x)$  per ogni  $x \neq 0$ , e  $0$  è dunque un punto di massimo globale in senso stretto.

Al contrario, mostriamo che non esistono punti di minimo globale. Innanzitutto,  $f(x) > 0$  per ogni  $x \in \mathbb{R}^n$ ; quindi, se esistesse un punto  $\bar{x}$  di minimo globale, sarebbe  $f(\bar{x}) > 0$ . Dal momento che

$$\lim_{t \rightarrow \infty} \frac{t^2}{1 + t^2} = 1, \quad \lim_{t \rightarrow \infty} \cos\left(\frac{\pi}{2} \frac{t^2}{1 + t^2}\right) = 0,$$

per ogni  $\varepsilon > 0$  esiste  $t > 0$  tale che

$$\cos\left(\frac{\pi}{2} \frac{t^2}{1 + t^2}\right) < \varepsilon,$$

e quindi nel punto  $x = (t, 0, \dots, 0)$  si avrebbe

$$f(x) = \cos\left(\frac{\pi}{2} \frac{|x|^2}{1 + |x|^2}\right) = \cos\left(\frac{\pi}{2} \frac{t^2}{1 + t^2}\right) < \varepsilon.$$

Se applichiamo questo argomento con un  $\varepsilon$  tale che  $0 < \varepsilon < f(\bar{x})$  (per esempio  $\varepsilon = f(\bar{x})/2$ ), troviamo che esiste  $x \in \mathbb{R}^n$  tale che  $f(x) < \varepsilon$ , da cui

$$f(x) < f(\bar{x}),$$

assurdo perché  $\bar{x}$  è un punto di minimo globale. Dunque non ci sono punti di minimo globale. (In altre parole, abbiamo mostrato che

$$\inf_{x \in \mathbb{R}^n} f(x) = 0,$$

ma questo inf non è mai raggiunto, perché  $f(x) > 0$  per ogni  $x$ , e quindi non c'è minimo globale).

La funzione è differenziabile ovunque, quindi altri punti di minimo e massimo locali, se esistono, sono punti critici. La derivata parziale rispetto a  $x_k$  è

$$\partial_{x_k} f(x) = -\sin\left(\frac{\pi}{2} \frac{|x|^2}{1 + |x|^2}\right) \frac{\pi}{2} \frac{2x_k(1 + |x|^2) - |x|^2 2x_k}{(1 + |x|^2)^2} = -\sin\left(\frac{\pi}{2} \frac{|x|^2}{1 + |x|^2}\right) \frac{\pi x_k}{(1 + |x|^2)^2},$$

quindi  $\nabla f(x) = 0$  solo nel punto  $x = 0$  (per  $x \neq 0$ , l'argomento del seno sta nell'intervallo aperto  $(0, \pi/2)$  e quindi il seno è  $> 0$ ). In conclusione, esiste un unico punto critico,  $x = 0$ , che è anche punto di massimo

globale in senso stretto, e non ci sono altri punti né di minimo né di massimo, né locali né globali, né ci sono altri punti critici.

(d) L'esercizio è praticamente identico al punto (c), con l'unica differenza (che non deve passare inosservata!) che la funzione non ammette derivate parziali in  $x = 0$ . Al punto (c) il fatto che  $x = 0$  fosse di massimo globale in senso stretto era conseguenza di ragionamenti che non coinvolgevano le derivate parziali, e quindi valgono anche per (d). Inoltre  $f$  è differenziabile sull'insieme aperto  $\mathbb{R}^n \setminus \{0\}$ , dunque per cercare i punti di minimo e massimo locali di  $f$  in  $\mathbb{R}^n \setminus \{0\}$  si può studiare l'insieme dei punti critici, e dunque procede allo stesso modo; si trova, come prima, che nessun punto  $x \in \mathbb{R}^n \setminus \{0\}$  annulla il gradiente, e quindi non ci sono altri punti di massimo né di minimo. La funzione non ha nemmeno un punto critico (mentre, in (c), l'origine era punto critico).  $\square$

**Esercizio 6.** (a) In  $\mathbb{R}^2 \setminus \{(0,0)\}$   $f$  è differenziabile perché formata da prodotti, somme e composizioni di funzioni differenziabili. Proviamo che  $f$  è differenziabile anche nell'origine. Calcoliamo dapprima le derivate parziali. Per definizione, la derivata parziale di  $f$  rispetto a  $x$  nel punto  $(0,0)$  è la derivata, se esiste, della funzione

$$x \mapsto f(x, 0).$$

Poiché  $f(x, 0) = 0$  per ogni  $x \in \mathbb{R}$ , si tratta della funzione costantemente nulla, che è derivabile in ogni  $x \in \mathbb{R}$ , e in particolare in  $x = 0$ , con derivata nulla. Quindi  $\partial_x f(0,0)$  esiste ed è  $= 0$ . Con identico ragionamento si prova che anche  $\partial_y f(0,0)$  esiste ed è  $= 0$ . Dunque il gradiente di  $f$  nell'origine è il vettore  $\nabla f(0,0) = (0,0)$ . Quindi il differenziale in  $(0,0)$ , se esiste, è la funzione lineare  $h \mapsto \nabla f(0,0) \cdot h = 0$  per ogni  $h \in \mathbb{R}^2$ , la funzione nulla. Proviamo la differenziabilità di  $f$  in  $(0,0)$  direttamente dalla definizione: se  $0 = (0,0)$  e  $h = (h_1, h_2) \neq (0,0)$ , usando il fatto che  $|h_1| \leq |h|$  e  $|h_2| \leq |h|$ ,

$$\begin{aligned} \frac{|f(0+h) - f(0) - \nabla f(0) \cdot h|}{|h|} &= \frac{|f(h)|}{|h|} = \frac{|h_1^3 h_2 - h_1 h_2^3|}{|h|^2 |h|} \leq \frac{|h_1|^3 |h_2| + |h_1| |h_2|^3}{|h|^3} \\ &= \frac{|h_1|^3}{|h|^3} |h_2| + \frac{|h_2|^3}{|h|^3} |h_1| \leq |h_2| + |h_1| \leq 2|h|. \end{aligned}$$

Questa disuguaglianza prova che

$$\frac{|f(0+h) - f(0) - \nabla f(0) \cdot h|}{|h|} \rightarrow 0 \quad \text{per } |h| \rightarrow 0, \quad h \in \mathbb{R}^2, \quad h \neq 0,$$

cioè  $f$  è differenziabile in  $0$ . Quindi  $f$  è differenziabile in tutto il suo dominio  $\mathbb{R}^2$ . La continuità di  $f$  è conseguenza della sua differenziabilità (proposizione 7.5).

Un modo alternativo per provare la continuità e la differenziabilità di  $f$  è quello di calcolare le sue derivate parziali in tutti i punti del dominio e mostrare che sono continue, invocando il teorema 7.6 del differenziale. Tuttavia la via seguita qua richiede meno conti.

(b) La derivata seconda  $\partial_{xy} f(0,0) = \partial_x(\partial_y f)(0,0)$ , se esiste, è per definizione la derivata parziale rispetto a  $x$  della funzione  $\partial_y f$  nel punto  $(0,0)$ , cioè la derivata, se esiste, della funzione

$$F: x \mapsto F(x) = \partial_y f(x, 0)$$

nel punto  $x = 0$ , ossia  $\partial_x \partial_y f(0,0) = F'(0)$ .<sup>3</sup> Calcoliamo  $F(x) = \partial_y f(x, 0)$  per  $x \neq 0$ : in  $(x, y)$  è

$$\partial_y f(x, y) = \frac{(x^3 - 3xy^2)(x^2 + y^2) - (x^3y - xy^3)2y}{(x^2 + y^2)^2} = \frac{x^5 - 4x^3y^2 - xy^4}{(x^2 + y^2)^2},$$

quindi per  $y = 0$  si ha

$$\partial_y f(x, 0) = \frac{x^5}{x^4} = x.$$

In  $x = 0$ , invece, come abbiamo calcolato al punto (a),  $F(0) = \partial_y f(0,0) = 0$ . Quindi

$$F(x) = x \quad \forall x \in \mathbb{R},$$

<sup>3</sup>Ricordiamo che la derivata parziale di  $f(x_1, \dots, x_n)$  rispetto a  $x_1$  nel punto  $\bar{x}$  è la derivata della funzione  $t \mapsto f(t, \bar{x}_2, \dots, \bar{x}_n)$  in  $t = \bar{x}_1$ , bloccate tutte le componenti eccetto la prima. In  $\mathbb{R}^2$ ,  $t \mapsto f(t, \bar{y})$ , e qui  $\bar{y} = 0$ .

che è derivabile, e  $F'(x) = 1$  per ogni  $x$ , e in particolare in  $x = 0$ . Dunque

$$\partial_x(\partial_y f)(0, 0) = F'(0) = 1.$$

Analogamente, la derivata seconda  $\partial_{yx}f(0, 0) = \partial_y(\partial_x f)(0, 0)$ , se esiste, è per definizione la derivata parziale rispetto a  $y$  della funzione  $\partial_x f$  nel punto  $(0, 0)$ , cioè la derivata, se esiste, della funzione

$$G: y \mapsto G(y) = \partial_x f(0, y)$$

nel punto  $y = 0$ , ossia  $\partial_y \partial_x f(0, 0) = G'(0)$ . Calcoliamo  $G(y) = \partial_x f(0, y)$  per  $y \neq 0$ : in  $(x, y)$  è

$$\partial_x f(x, y) = \frac{(3x^2y - y^3)(x^2 + y^2) - (x^3y - xy^3)2x}{(x^2 + y^2)^2} = \frac{x^4y + 4x^2y^3 - y^5}{(x^2 + y^2)^2},$$

quindi per  $x = 0$  si ha

$$\partial_x f(0, y) = \frac{-y^5}{y^4} = -y.$$

In  $y = 0$ , invece, come calcolato al punto (a),  $G(0) = \partial_x f(0, 0) = 0$ . Quindi

$$G(y) = -y \quad \forall y \in \mathbb{R},$$

che è derivabile, e  $G'(y) = -1$  per ogni  $y$ , e in particolare in  $y = 0$ . Dunque

$$\partial_y(\partial_x f)(0, 0) = G'(0) = -1.$$

Questo prova che le due derivate miste sono diverse,

$$\partial_x(\partial_y f)(0, 0) = 1 \neq -1 = \partial_y(\partial_x f)(0, 0).$$

Ovviamente, calcolando i limiti dei rapporti incrementali (invece che calcolando le derivate con le regole di derivazione note da Analisi 1, come fatto qui) si ottiene lo stesso risultato.

(c) Visto che  $f$  è differenziabile e tutti i punti  $(x, y)$  sono interni al dominio  $\mathbb{R}^2$ , eventuali punti di massimo e di minimo locale sono necessariamente punti critici (proposizione 11.3). In altre parole, non ci sono punti di minimo o di massimo che non siano punti critici. Cerchiamo tutti i punti critici di  $f$ . Sappiamo già che  $\nabla f(0, 0) = (0, 0)$ , quindi l'origine è un punto critico. Fuori dall'origine, il gradiente in un punto  $(x, y) \neq (0, 0)$  è nullo se e solo se (dai conti fatti al punto (b))

$$\begin{cases} x(x^4 - 4x^2y^2 - y^4) = 0 & \text{(I),} \\ y(x^4 + 4x^2y^2 - y^4) = 0 & \text{(II).} \end{cases}$$

Se  $x \neq 0$  e  $y = 0$ , la (II) è soddisfatta ma la (I) no, perché  $x(x^4 - 4x^2y^2 - y^4) = x^5 \neq 0$ . Similmente, se  $x = 0$  e  $y \neq 0$ , la (I) è soddisfatta ma la (II) no, perché  $y(x^4 + 4x^2y^2 - y^4) = -y^5 \neq 0$ . Infine, se entrambi  $x \neq 0$  e  $y \neq 0$ , e  $(x, y)$  è una soluzione, allora

$$\begin{cases} x^4 - 4x^2y^2 - y^4 = 0, \\ x^4 + 4x^2y^2 - y^4 = 0. \end{cases}$$

Ma sottraendo si ottiene  $8x^2y^2 = 0$ , per cui almeno uno fra  $x$  e  $y$  deve essere nullo, assurdo. In conclusione,  $(0, 0)$  è l'unico punto critico di  $f$ .

Per capire se è punto di minimo o di massimo non possiamo usare i criteri del secondo ordine, perché manca l'ipotesi che  $f$  sia di classe  $C^2$  in una palla centrata nell'origine (se lo fosse, per il teorema di Schwartz sarebbe  $\partial_{xy}f(0, 0) = \partial_{yx}f(0, 0)$ , contro il punto (b)). Tuttavia si può ragionare in altro modo: lungo la retta di equazione  $y = mx$ , per  $x \neq 0$  si ha

$$f(x, mx) = \frac{x^3mx - x(mx)^3}{x^2 + (mx)^2} = \frac{m(1 - m^2)}{1 + m^2} x^2.$$

Il coefficiente  $m(1 - m^2)/(1 + m^2)$  può avere segni diversi a seconda del valore di  $m$ ; per esempio, se scegliamo  $m = 2$  risulta

$$f(x, 2x) = -\frac{6}{5}x^2 < 0 \quad \forall x \neq 0,$$

mentre per  $m = -2$  è

$$f(x, -2x) = \frac{6}{5}x^2 > 0 \quad \forall x \neq 0.$$

Dunque, dato  $r > 0$ , la palla di centro nell'origine e raggio  $r$  contiene sempre un punto del tipo  $(x, 2x)$  in cui  $f < 0$ , e un punto del tipo  $(x, -2x)$  in cui  $f > 0$  (basta scegliere  $x$  abbastanza piccolo, per esempio  $x = r/3$ ); quindi  $0 = f(0, 0)$  non è né il massimo né il minimo valore di  $f$  nella palla di raggio  $r$ . Poiché questo è vero per ogni  $r > 0$ , l'origine non è né un punto di massimo né un punto di minimo per  $f$ .

Punti di massimo o minimo globale non ce ne sono perché, se ci fossero, sarebbero anche punti di massimo o minimo locale; ma abbiamo provato che  $f$  non ha punti di minimo o massimo locale.

Un altro motivo per cui non ci sono punti di massimo o minimo globale è il fatto che  $f$  non è limitata: infatti

$$\lim_{x \rightarrow \infty} f(x, 2x) = -\infty, \quad \lim_{x \rightarrow \infty} f(x, -2x) = \infty,$$

e quindi

$$\sup_{(x,y) \in \mathbb{R}^2} f(x, y) = \infty, \quad \inf_{(x,y) \in \mathbb{R}^2} f(x, y) = -\infty,$$

mentre, se esistessero punti di massimo o minimo globale,  $f$  dovrebbe essere limitata. □

# Indice

<b>1</b>	<b>Algebra lineare e geometria in <math>\mathbb{R}^n</math></b>	<b>1</b>
1.1	Prodotto scalare e norma euclidea . . . . .	1
1.2	Funzioni lineari . . . . .	4
<b>2</b>	<b>Elementi di topologia in <math>\mathbb{R}^n</math></b>	<b>6</b>
<b>3</b>	<b>Limiti e continuità</b>	<b>7</b>
<b>4</b>	<b>Funzioni lipschitziane</b>	<b>10</b>
4.1	Alcuni esempi importanti di funzioni lipschitziane . . . . .	11
<b>5</b>	<b>Ancora sulla continuità</b>	<b>13</b>
<b>6</b>	<b>Derivate parziali</b>	<b>16</b>
<b>7</b>	<b>Il differenziale</b>	<b>18</b>
<b>8</b>	<b>Derivate direzionali</b>	<b>24</b>
<b>9</b>	<b>Derivate parziali successive</b>	<b>25</b>
<b>10</b>	<b>Formula di Taylor</b>	<b>27</b>
<b>11</b>	<b>Massimi e minimi</b>	<b>29</b>
11.1	Matrici definite, semidefinite e indefinite . . . . .	31
11.2	Condizioni del secondo ordine . . . . .	35
<b>12</b>	<b>Esercizi</b>	<b>36</b>
12.1	Svolgimento degli esercizi . . . . .	37