Analisi Matematica II

Appunti delle lezioni tenute dal Prof. A. Fonda

Università di Trieste, CdL Ingegneria, a.a. 2021/2022

Nota. Questi appunti sono in continua evoluzione. Si prega di segnalare eventuali errori, per poterli correggere rapidamente.

1 Lo spazio $\mathbb{R}^N$

Consideriamo l'insieme $\mathbb{R}^N$, costituito dalle N-uple $(x_1, x_2, \dots, x_N)$, dove x_1, x_2, ..., x_N sono numeri reali. Indicheremo i suoi elementi con i simboli

$$\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{x}'', \dots$$

Cominciamo con l'introdurre un'operazione di addizione in $\mathbb{R}^N$: dati due elementi $\boldsymbol{x} = (x_1, x_2, \dots, x_N)$ e $\boldsymbol{x}' = (x'_1, x'_2, \dots, x'_N)$, si definisce $\boldsymbol{x} + \boldsymbol{x}'$ in questo modo:

$$\boldsymbol{x} + \boldsymbol{x}' = (x_1 + x'_1, x_2 + x'_2, \dots, x_N + x'_N)\,.$$

Valgono le seguenti proprietà:

a) (associativa) $(\boldsymbol{x} + \boldsymbol{x}') + \boldsymbol{x}'' = \boldsymbol{x} + (\boldsymbol{x}' + \boldsymbol{x}'')$;
b) esiste un "elemento neutro" $\mathbf{0} = (0, 0, \dots, 0)$: si ha $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x} = \mathbf{0} + \boldsymbol{x}$;
c) ogni elemento $\boldsymbol{x} = (x_1, x_2, \dots, x_N)$ ha un "opposto"
$(-\boldsymbol{x}) = (-x_1, -x_2, \dots, -x_N)$: si ha $\boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0} = (-\boldsymbol{x}) + \boldsymbol{x}$;
d) (commutativa) $\boldsymbol{x} + \boldsymbol{x}' = \boldsymbol{x}' + \boldsymbol{x}$.

Pertanto, $(\mathbb{R}^N, +)$ è un "gruppo abeliano". Normalmente, si usa scrivere $\boldsymbol{x} - \boldsymbol{x}'$ per indicare $\boldsymbol{x} + (-\boldsymbol{x}')$.

Definiamo ora la moltiplicazione di un elemento di $\mathbb{R}^N$ per un numero reale: considerati $\boldsymbol{x} = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$ e un numero reale $\alpha \in \mathbb{R}$, si definisce $\alpha\boldsymbol{x}$ in questo modo:

$$\alpha\boldsymbol{x} = (\alpha x_1, \alpha x_2, \dots, \alpha x_N)\,.$$

Valgono le seguenti proprietà:

a) $\alpha(\beta\boldsymbol{x}) = (\alpha\beta)\boldsymbol{x}$;
b) $(\alpha + \beta)\boldsymbol{x} = (\alpha\boldsymbol{x}) + (\beta\boldsymbol{x})$;
c) $\alpha(\boldsymbol{x} + \boldsymbol{x}') = (\alpha\boldsymbol{x}) + (\alpha\boldsymbol{x}')$;
d) $1\boldsymbol{x} = \boldsymbol{x}$.

Pertanto, con le operazioni introdotte, $\mathbb{R}^N$ è uno "spazio vettoriale". Chiameremo i suoi elementi "vettori"; i numeri reali, in questo ambito, verranno chiamati "scalari".

È utile introdurre il "prodotto scalare" tra due vettori: dati $\boldsymbol{x} = (x_1, x_2, \dots, x_N)$ e $\boldsymbol{x}' = (x'_1, x'_2, \dots, x'_N)$, si definisce il numero reale $\boldsymbol{x} \cdot \boldsymbol{x}'$ in questo modo:

$$\boldsymbol{x} \cdot \boldsymbol{x}' = \sum_{k=1}^{N} x_k x'_k \,.$$

Il prodotto scalare è spesso indicato con simboli diversi, quali ad esempio

$$\langle \boldsymbol{x} | \boldsymbol{x}' \rangle \,, \quad \langle \boldsymbol{x}, \boldsymbol{x}' \rangle \,, \quad (\boldsymbol{x} | \boldsymbol{x}') \,, \quad (\boldsymbol{x}, \boldsymbol{x}') \,.$$

Valgono le seguenti proprietà:
a) $\boldsymbol{x} \cdot \boldsymbol{x} \geq 0\,;$
b) $\boldsymbol{x} \cdot \boldsymbol{x} = 0 \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{0}\,;$
c) $(\boldsymbol{x} + \boldsymbol{x}') \cdot \boldsymbol{x}'' = (\boldsymbol{x} \cdot \boldsymbol{x}'') + (\boldsymbol{x}' \cdot \boldsymbol{x}'')\,;$
d) $(\alpha \boldsymbol{x}) \cdot \boldsymbol{x}' = \alpha(\boldsymbol{x} \cdot \boldsymbol{x}')\,;$
e) $\boldsymbol{x} \cdot \boldsymbol{x}' = \boldsymbol{x}' \cdot \boldsymbol{x}\,;$

Se $\boldsymbol{x} \cdot \boldsymbol{x}' = 0$, si dice che i due vettori $\boldsymbol{x}$ e $\boldsymbol{x}'$ sono "ortogonali".

1.1 Norma e distanza euclidea in $\mathbb{R}^N$

A partire dal prodotto scalare, possiamo definire la "norma" di un vettore $\boldsymbol{x} = (x_1, x_2, \dots, x_N)$:

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{\sum_{k=1}^{N} x_k^2} \,.$$

Valgono le seguenti proprietà:
a) $\|\boldsymbol{x}\| \geq 0\,;$
b) $\|\boldsymbol{x}\| = 0 \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{0}\,;$
c) $\|\alpha \boldsymbol{x}\| = |\alpha|\, \|\boldsymbol{x}\|\,;$
d) $\|\boldsymbol{x} + \boldsymbol{x}'\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{x}'\|\,.$

Per dimostrare la d), detta "disuguaglianza triangolare" per la norma, abbiamo bisogno della seguente **disuguaglianza di Schwarz**.

Teorema 1 *Presi due vettori $\boldsymbol{x}, \boldsymbol{x}'$, si ha*

$$|\boldsymbol{x} \cdot \boldsymbol{x}'| \leq \|\boldsymbol{x}\|\, \|\boldsymbol{x}'\| \,.$$

<u>Dimostrazione</u>. La disuguaglianza è sicuramente verificata se $\boldsymbol{x}' = \boldsymbol{0}$, essendo in tal caso $\boldsymbol{x} \cdot \boldsymbol{x}' = 0$ e $\|\boldsymbol{x}'\| = 0$. Supponiamo quindi $\boldsymbol{x}' \neq \boldsymbol{0}$. Per ogni $\alpha \in \mathbb{R}$, si ha

$$0 \leq \|\boldsymbol{x} - \alpha \boldsymbol{x}'\|^2 = (\boldsymbol{x} - \alpha \boldsymbol{x}') \cdot (\boldsymbol{x} - \alpha \boldsymbol{x}') = \|\boldsymbol{x}\|^2 - 2\alpha \boldsymbol{x} \cdot \boldsymbol{x}' + \alpha^2 \|\boldsymbol{x}'\|^2 \,.$$

Prendendo $\alpha = \frac{1}{\|\boldsymbol{x}'\|^2}\,\boldsymbol{x}\cdot\boldsymbol{x}'$, si ottiene

$$0 \le \|\boldsymbol{x}\|^2 - 2\frac{1}{\|\boldsymbol{x}'\|^2}(\boldsymbol{x}\cdot\boldsymbol{x}')^2 + \frac{1}{\|\boldsymbol{x}'\|^4}(\boldsymbol{x}\cdot\boldsymbol{x}')^2\|\boldsymbol{x}'\|^2 = \|\boldsymbol{x}\|^2 - \frac{1}{\|\boldsymbol{x}'\|^2}(\boldsymbol{x}\cdot\boldsymbol{x}')^2\,,$$

da cui la tesi.[1] ∎

Dimostriamo ora la proprietà d) della norma, usando la disuguaglianza di Schwarz:

$$\begin{aligned}\|\boldsymbol{x}+\boldsymbol{x}'\|^2 &= (\boldsymbol{x}+\boldsymbol{x}')\cdot(\boldsymbol{x}+\boldsymbol{x}')\\ &= \|\boldsymbol{x}\|^2 + 2\boldsymbol{x}\cdot\boldsymbol{x}' + \|\boldsymbol{x}'\|^2\\ &\le \|\boldsymbol{x}\|^2 + 2\|\boldsymbol{x}\|\,\|\boldsymbol{x}'\| + \|\boldsymbol{x}'\|^2\\ &= (\|\boldsymbol{x}\| + \|\boldsymbol{x}'\|)^2\,,\end{aligned}$$

da cui la disuguaglianza cercata.

Notiamo ancora la seguente **identità del parallelogramma**, di semplice verifica:

$$\|\boldsymbol{x}+\boldsymbol{x}'\|^2 + \|\boldsymbol{x}-\boldsymbol{x}'\|^2 = 2(\|\boldsymbol{x}\|^2 + \|\boldsymbol{x}'\|^2)\,.$$

Definiamo ora, a partire dalla norma, la "distanza euclidea" tra due vettori $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ e $\boldsymbol{x}' = (x_1', x_2', \ldots, x_N')$:

$$d(\boldsymbol{x},\boldsymbol{x}') = \|\boldsymbol{x}-\boldsymbol{x}'\| = \sqrt{\sum_{k=1}^{N}(x_k - x_k')^2}\,.$$

Valgono le seguenti proprietà:

a) $d(\boldsymbol{x},\boldsymbol{x}') \ge 0$;
b) $d(\boldsymbol{x},\boldsymbol{x}') = 0 \quad\Leftrightarrow\quad \boldsymbol{x} = \boldsymbol{x}'$;
c) $d(\boldsymbol{x},\boldsymbol{x}') = d(\boldsymbol{x}',\boldsymbol{x})$;
d) $d(\boldsymbol{x},\boldsymbol{x}'') \le d(\boldsymbol{x},\boldsymbol{x}') + d(\boldsymbol{x}',\boldsymbol{x}'')$.

Quest'ultima viene spesso chiamata "disuguaglianza triangolare"; la dimostriamo:

$$\begin{aligned}d(\boldsymbol{x},\boldsymbol{x}'') &= \|\boldsymbol{x}-\boldsymbol{x}''\|\\ &= \|(\boldsymbol{x}-\boldsymbol{x}') + (\boldsymbol{x}'-\boldsymbol{x}'')\|\\ &\le \|\boldsymbol{x}-\boldsymbol{x}'\| + \|\boldsymbol{x}'-\boldsymbol{x}''\|\\ &= d(\boldsymbol{x},\boldsymbol{x}') + d(\boldsymbol{x}',\boldsymbol{x}'')\,.\end{aligned}$$

2 Spazi metrici

Dato un insieme non vuoto E, una funzione $d : E\times E \to \mathbb{R}$ si chiama "distanza" (su E) se soddisfa alle seguenti proprietà:

[1] Lo stesso risultato si ottiene osservando che, se $a\alpha^2 + b\alpha + c \ge 0$ per ogni $\alpha \in \mathbb{R}$, con $a > 0$, allora $\Delta = b^2 - 4ac \le 0$.

a) $d(x,x') \geq 0\,;$
b) $d(x,x') = 0 \quad \Leftrightarrow \quad x = x'\,;$
c) $d(x,x') = d(x',x)\,;$
d) $d(x,x'') \leq d(x,x') + d(x',x'')$

(la disuguaglianza triangolare). L'insieme E, dotato della distanza d, si dice "spazio metrico". I suoi elementi verranno spesso chiamati "punti".

Abbiamo visto che $\mathbb{R}^N$, dotato della distanza euclidea, è uno spazio metrico (nel seguito, parlando dello spazio metrico $\mathbb{R}^N$, se non altrimenti specificato sottintenderemo che la distanza sia sempre quella euclidea). Nel caso $N = 1$, abbiamo la distanza usuale su $\mathbb{R}$: $d(\alpha, \beta) = |\alpha - \beta|$.

È però possibile considerare diverse distanze su uno stesso insieme. Ad esempio, presi due vettori $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ e $\boldsymbol{x'} = (x'_1, x'_2, \ldots, x'_N)$, la funzione

$$d_*(\boldsymbol{x}, \boldsymbol{x'}) = \sum_{k=1}^{N} |x_k - x'_k|$$

rappresenta anch'essa una distanza in $\mathbb{R}^N$. Lo stesso dicasi per la funzione

$$d_{**}(\boldsymbol{x}, \boldsymbol{x'}) = \max\{|x_k - x'_k| : k = 1, 2, \ldots, N\}\,.$$

Oppure, si può definire la seguente:

$$\hat{d}(\boldsymbol{x}, \boldsymbol{x'}) = \begin{cases} 0 & \text{se } \boldsymbol{x} = \boldsymbol{x'}\,, \\ 1 & \text{se } \boldsymbol{x} \neq \boldsymbol{x'}\,. \end{cases}$$

Anche questa è una distanza, per quanto strana possa sembrare.

Dati $x_0 \in E$ e un numero $\rho > 0$, definiamo la palla aperta di centro x_0 e raggio ρ:

$$B(x_0, \rho) = \{x \in E : d(x, x_0) < \rho\}\,;$$

analogamente definiamo la palla chiusa

$$\overline{B}(x_0, \rho) = \{x \in E : d(x, x_0) \leq \rho\}\,,$$

e la sfera

$$S(x_0, \rho) = \{x \in E : d(x, x_0) = \rho\}\,.$$

In $\mathbb{R}$, ogni intervallo $]a, b[$ è una palla aperta e ogni intervallo $[a, b]$ è una palla chiusa: si ha

$$]a, b[= B\left(\frac{a+b}{2}, \frac{b-a}{2}\right), \quad [a, b] = \overline{B}\left(\frac{a+b}{2}, \frac{b-a}{2}\right).$$

Una sfera in $\mathbb{R}$ è quindi costituita da due soli punti.

In $\mathbb{R}^2$, con la distanza euclidea, una palla è un cerchio: la palla aperta non comprende i punti della circonferenza esterna, la palla chiusa sì. Una sfera è semplicemente una circonferenza.

Se in $\mathbb{R}^2$ consideriamo la distanza d_* definita in precedenza, una palla sarà un quadrato, con i lati inclinati di 45 gradi, avente x_0 come punto centrale. Una sfera sarà il perimetro di tale quadrato. Se invece consideriamo la distanza d_{**}, la palla sarà ancora un quadrato, ma con i lati paralleli agli assi cartesiani.

Se invece prendiamo la distanza $\hat{d}$, su un qualsiasi insieme E, allora

$$B(x_0,\rho) = \begin{cases} \{x_0\} & \text{se } \rho \leq 1\,, \\ E & \text{se } \rho > 1\,, \end{cases} \qquad \overline{B}(x_0,\rho) = \begin{cases} \{x_0\} & \text{se } \rho < 1\,, \\ E & \text{se } \rho \geq 1\,, \end{cases}$$

per cui

$$S(x_0,\rho) = \begin{cases} E \setminus \{x_0\} & \text{se } \rho = 1\,, \\ \emptyset & \text{se } \rho \neq 1\,. \end{cases}$$

Un insieme $U \subseteq E$ si dice "intorno" di un punto x_0 se esiste un $\rho > 0$ tale che $B(x_0,\rho) \subseteq U$; in tal caso, il punto x_0 si dice "interno" ad U. L'insieme dei punti interni ad U si chiama "l'interno" di U e si denota con $\mathring{U}$. Chiaramente, si ha sempre $\mathring{U} \subseteq U$. Si dice che U è un "insieme aperto" se coincide con il suo interno, ossia se $\mathring{U} = U$.

Teorema 2 *Una palla aperta è un insieme aperto.*

Dimostrazione. Sia $B(x_0,\rho)$ la palla in questione; prendiamo un $x_1 \in B(x_0,\rho)$. Scelto $r > 0$ tale che $r \leq \rho - d(x_0,x_1)$, si ha che $B(x_1,r) \subseteq B(x_0,\rho)$; infatti, se $x \in B(x_1,r)$, allora

$$d(x,x_0) \leq d(x,x_1) + d(x_1,x_0) < r + d(x_1,x_0) \leq \rho\,,$$

per cui $x \in B(x_0,\rho)$. Abbiamo quindi dimostrato che ogni punto x_1 di $B(x_0,\rho)$ è interno a $B(x_0,\rho)$. ■

Consideriamo ora tre esempi particolari: nel primo, l'insieme U coincide con E; nel secondo, U è l'insieme vuoto; nel terzo, esso è costituito da un unico punto.

Ogni punto di E è interno all'insieme E stesso, in quanto ogni palla è per definizione contenuta in E. Quindi, l'interno di E coincide con tutto E, ossia $\mathring{E} = E$. Questo significa che E è un insieme aperto.

L'insieme vuoto non può avere punti interni. Quindi, l'interno di $\emptyset$, non avendo elementi, è vuoto. In altri termini, $\mathring{\emptyset} = \emptyset$, il che significa che $\emptyset$ è anch'esso un insieme aperto.

L'insieme $U = \{x_0\}$, costituito da un unico punto, in generale non è un insieme aperto (ad esempio in $\mathbb{R}^N$ con la distanza euclidea), ma può esserlo in casi particolari, quando x_0 è un "punto isolato" di E. Ad esempio, se si considera la distanza $\hat{d}$, oppure in $\mathbb{N}$, tutti i punti solo isolati.

Teorema 3 *L'interno di un insieme è un insieme aperto.*

Dimostrazione. Se $\mathring{U}$ è vuoto, la tesi è sicuramente vera. Supponiamo allora che $\mathring{U}$ sia non vuoto. Sia $x_1 \in \mathring{U}$. Allora esiste un $\rho > 0$ tale che $B(x_1, \rho) \subseteq U$. Sia $V = B(x_1, \rho)$, per cui $V \subseteq U$. Ne segue che $\mathring{V} \subseteq \mathring{U}$, quindi, essendo V un insieme aperto, $V \subseteq \mathring{U}$, ossia $B(x_1, \rho) \subseteq \mathring{U}$. Pertanto ogni punto x_1 di $\mathring{U}$ è interno a $\mathring{U}$. ■

Si può dimostrare la seguente implicazione:

$$U_1 \subseteq U_2 \quad \Rightarrow \quad \mathring{U}_1 \subseteq \mathring{U}_2 \,.$$

Da essa segue che $\mathring{U}$ è il più grande insieme aperto contenuto in U: se A è un aperto e $A \subseteq U$, allora $A \subseteq \mathring{U}$.

Diremo che il punto x_0 è "aderente" all'insieme U se per ogni $\rho > 0$ si ha che $B(x_0, \rho) \cap U \neq \emptyset$. L'insieme dei punti aderenti ad U si chiama "la chiusura" di U e si denota con $\overline{U}$. Chiaramente, si ha sempre $U \subseteq \overline{U}$. Si dice che U è un "insieme chiuso" se coincide con la sua chiusura, ossia se $U = \overline{U}$.

Teorema 4 *Una palla chiusa è un insieme chiuso.*

Dimostrazione. Sia $U = \overline{B}(x_0, \rho)$ la palla in questione; voglio dimostrare che $\overline{U} \subseteq U$. A tal fine vedremo che $\mathcal{C}U \subseteq \mathcal{C}\overline{U}$. [2] Prendiamo un $x_1 \in \mathcal{C}U$, ossia $x_1 \notin \overline{B}(x_0, \rho)$. Scelto $r > 0$ tale che $r \leq d(x_0, x_1) - \rho$, si ha che $B(x_1, r) \cap \overline{B}(x_0, \rho) = \emptyset$; infatti, se per assurdo esistesse un $x \in B(x_1, r) \cap \overline{B}(x_0, \rho)$, allora si avrebbe

$$d(x_0, x_1) \leq d(x_0, x) + d(x, x_1) < r + \rho \,,$$

in contrasto con la scelta fatta per r. Quindi, $x_1 \notin \overline{U}$, ossia $x_1 \in \mathcal{C}\overline{U}$. ■

Essendo E il l'insieme universo, ogni punto aderente ad E deve comunque appartenere ad E stesso. Quindi, la chiusura di E coincide con E, ossia $\overline{E} = E$. Questo significa che E è un insieme chiuso.

Notiamo che non esiste alcun punto aderente all'insieme $\emptyset$. Infatti, qualsiasi sia il punto x_0, per ogni $\rho > 0$ si ha che $B(x_0, \rho) \cap \emptyset = \emptyset$. Quindi, la chiusura di $\emptyset$, non avendo elementi, è vuota. In altri termini, $\overline{\emptyset} = \emptyset$, il che significa che $\emptyset$ è un insieme chiuso.

L'insieme $U = \{x_0\}$, costituito da un unico punto, è sempre un insieme chiuso. Infatti, preso un $x_1 \notin U$, scegliendo $\rho > 0$ tale che $\rho < d(x_0, x_1)$ si ha che $B(x_1, \rho) \cap U = \emptyset$, per cui x_1 non è aderente ad U.

Teorema 5 *La chiusura di un insieme è un insieme chiuso.*

Dimostrazione. Poniamo $V = \overline{U}$. Se $V = E$, la tesi è verificata. Supponiamo quindi che sia $V \neq E$. Sia $x_1 \notin V$. Allora esiste un $\rho > 0$ tale che $B(x_1, \rho) \cap U = \emptyset$. Vediamo che anche $B(x_1, \rho) \cap V = \emptyset$. Infatti, se per assurdo ci fosse un $x \in B(x_1, \rho) \cap V$, essendo $B(x_1, \rho)$ un insieme aperto, esisterebbe un $r > 0$ tale che $B(x, r) \subseteq B(x_1, \rho)$. Siccome $x \in V = \overline{U}$, dovrebbe essere $B(x, r) \cap U \neq \emptyset$ e quindi anche $B(x_1, \rho) \cap U \neq \emptyset$, in contraddizione con quanto sopra. Quindi, nessun punto x_1 al di fuori di V può essere aderente a V. In altri termini, V contiene tutti i punti ad esso aderenti, pertanto è chiuso. ■

[2]Denotiamo con $\mathcal{C}U$ il complementare di U in E, ossia l'insieme $E \setminus U$.

Si può dimostrare che

$$U_1 \subseteq U_2 \quad \Rightarrow \quad \overline{U}_1 \subseteq \overline{U}_2 \,.$$

Da questa implicazione segue che $\overline{U}$ è il più piccolo insieme chiuso che contiene U: se C è un chiuso e $C \supseteq U$, allora $C \supseteq \overline{U}$.

Si può dimostrare che l'unione e l'intersezione di due insiemi aperti [chiusi] sono insiemi aperti [chiusi]. Lo stesso vale per un numero finito di insiemi aperti [chiusi]: lo si dimostra per induzione. Se invece si considera un numero infinito di insiemi, la cosa cambia. L'unione di un numero infinito di insiemi aperti è un insieme aperto, l'intersezione in generale non lo è. Ad esempio, in $\mathbb{R}$, prendendo gli aperti

$$A_n = \left] -\frac{1}{n+1}\,, \frac{1}{n+1} \right[\,,$$

con $n \in \mathbb{N}$, la loro intersezione è $\{0\}$, che non è un aperto. Analogamente, l'intersezione di un numero infinito di insiemi chiusi è un insieme chiuso, mentre l'unione in generale non lo è. Ad esempio, considerando i chiusi

$$C_n = \left[-1 + \frac{1}{n+1}\,, 1 - \frac{1}{n+1} \right]\,,$$

con $n \in \mathbb{N}$, la loro unione è l'intervallo $]-1, 1[$, che non è un chiuso.

Cercheremo ora di capire le analogie incontrate tra le nozioni di interno e chiusura di un insieme, e quelle di insieme aperto e chiuso.

Teorema 6 *Valgono le seguenti relazioni:*

$$\overline{\mathcal{C}U} = \mathcal{C}\mathring{U}\,, \qquad (\mathcal{C}\mathring{U}) = \mathcal{C}\overline{U}\,.$$

<u>Dimostrazione</u>. Vediamo la prima uguaglianza. Se $U = E$, allora $\mathcal{C}U = \emptyset$, per cui $\overline{\mathcal{C}U} = \emptyset$; d'altra parte, $\mathring{U} = E$, per cui $\mathcal{C}\mathring{U} = \emptyset$. L'uguaglianza è così verificata in questo caso. Supponiamo ora che sia $U \neq E$, per cui $\mathcal{C}U \neq \emptyset$. Si ha:

$$\begin{aligned} x \in \overline{\mathcal{C}U} \quad &\Leftrightarrow \quad \forall \rho > 0 \quad B(x,\rho) \cap \mathcal{C}U \neq \emptyset \\ &\Leftrightarrow \quad \forall \rho > 0 \quad B(x,\rho) \not\subseteq U \\ &\Leftrightarrow \quad x \notin \mathring{U} \\ &\Leftrightarrow \quad x \in \mathcal{C}\mathring{U}. \end{aligned}$$

Questo dimostra la prima uguaglianza. Possiamo ora usarla per dedurne la seguente:

$$\mathcal{C}(\mathcal{C}\mathring{U}) = \overline{\mathcal{C}(\mathcal{C}U)} = \overline{U}\,.$$

Passando ai complementari, si ottiene la seconda uguaglianza. ■

Abbiamo quindi che

$$\overline{U} = \mathcal{C}(\overset{\circ}{\mathcal{C}U})\,, \qquad \overset{\circ}{U} = \mathcal{C}(\overline{\mathcal{C}U})\,.$$

Come immediato corollario, abbiamo il seguente.

Teorema 7 *Un insieme è aperto [chiuso] se e solo se il suo complementare è chiuso [aperto].*

<u>Dimostrazione</u>. Se U è aperto, $U = \overset{\circ}{U}$ e quindi

$$\overline{\mathcal{C}U} = \mathcal{C}\overset{\circ}{U} = \mathcal{C}U\,,$$

per cui $\mathcal{C}U$ è chiuso.

Se U è chiuso, $U = \overline{U}$ e quindi

$$(\overset{\circ}{\mathcal{C}U}) = \mathcal{C}\overline{U} = \mathcal{C}U\,,$$

per cui $\mathcal{C}U$ è aperto. ∎

Si definisce la "frontiera" di un insieme U come differenza tra la sua chiusura e il suo interno:

$$\partial U = \overline{U} \setminus \overset{\circ}{U}\,.$$

Ad esempio, in $\mathbb{R}$ abbiamo:

$$\overline{\mathbb{Q}} = \mathbb{R}\,, \quad \overset{\circ}{\mathbb{Q}} = \emptyset\,, \quad \partial\mathbb{Q} = \mathbb{R}\,.$$

È bene essere prudenti su alcune conclusioni che possono esserci suggerite dalla nostra intuizione basata sulla distanza euclidea. Ad esempio, le uguaglianze

$$\overline{B(\boldsymbol{x}_0,\rho)} = \overline{B}(\boldsymbol{x}_0,\rho)\,, \quad \partial B(\boldsymbol{x}_0,\rho) = S(\boldsymbol{x}_0,\rho)\,.$$

valgono sicuramente in $\mathbb{R}^N$ con la distanza euclidea, ma possono non valere in altri casi. Prendiamo ad esempio la distanza $\hat{d}$ considerata sopra. Allora $B(\boldsymbol{x}_0,1) = \{\boldsymbol{x}_0\}$, che è un insieme chiuso, e $\overline{B}(\boldsymbol{x}_0,1) = E$ per cui $\overline{B(\boldsymbol{x}_0,1)} \neq \overline{B}(\boldsymbol{x}_0,1)$. Inoltre, $\partial B(\boldsymbol{x}_0,1) = \emptyset$, mentre $S(\boldsymbol{x}_0,1) = E \setminus \{\boldsymbol{x}_0\}$, per cui $\partial B(\boldsymbol{x}_0,1) \neq S(\boldsymbol{x}_0,1)$.

Nota. Nel seguito, qualora non specificato altrimenti, quando parleremo di $\mathbb{R}^N$ come spazio metrico o normato sarà sempre sottinteso che la distanza e la norma su di esso considerate siano quelle euclidee.

3 Funzioni continue

Siano E ed F due spazi metrici, con le loro distanze d_E e d_F, rispettivamente. Sia x_0 un punto di E e $f : E \to F$ una funzione.

Definizione 1 *Diremo che f è "continua" in x_0 se, comunque preso un numero positivo ε, è possibile trovare un numero positivo δ tale che, se x è un qualsiasi elemento del dominio E che disti da x_0 per meno di δ, allora $f(x)$ dista da $f(x_0)$ per meno di ε. In simboli:*

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \; \forall x \in E \quad d_E(x, x_0) < \delta \; \Rightarrow \; d_F(f(x), f(x_0)) < \varepsilon \, .$$

In questa formulazione, spesso la scrittura "$\forall x \in E$" verrà sottintesa.

Osserviamo che una o entrambe le disuguaglianze

$$d_E(x, x_0) < \delta \, , \qquad d_F(f(x), f(x_0)) < \varepsilon$$

possono essere sostituite rispettivamente da

$$d_E(x, x_0) \le \delta \, , \qquad d_F(f(x), f(x_0)) \le \varepsilon$$

ottenendo definizioni che sono tutte tra loro equivalenti. Questo è dovuto al fatto, da un lato, che ε è un *qualunque* numero positivo e, dall'altro lato, che se l'implicazione della definizione vale per un certo numero positivo δ, essa vale a maggior ragione prendendo al posto di quel δ un qualsiasi numero positivo più piccolo.

Una rilettura della definizione di continuità ci mostra che f è continua in x_0 se e solo se:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \; f(B(x_0, \delta)) \subseteq B(f(x_0), \varepsilon) \, .$$

Inoltre, è del tutto equivalente considerare una palla chiusa al posto di una palla aperta; risulta inoltre utile la seguente formulazione equivalente, per cui f è continua in x_0 se e solo se:

per ogni intorno V di $f(x_0)$ esiste un intorno U di x_0 tale che $f(U) \subseteq V$.

Nel caso in cui la funzione f sia continua in ogni punto x_0 del dominio E, diremo che "f è continua su E", o semplicemente "f è continua".

Vediamo ora alcuni esempi.

Esempio 1. La funzione costante: per un certo $\bar{c} \in F$, si ha che $f(x) = \bar{c}$, per ogni $x \in E$. Essendo $d_F(f(x), f(x_0)) = d_F(\bar{c}, \bar{c}) = 0$ per ogni $x \in E$, tale funzione è chiaramente continua (ogni scelta di $\delta > 0$ va bene).

Esempio 2. Siano $E = \mathbb{R}^N$ e $F = \mathbb{R}^N$. Fissato un numero $\alpha \in \mathbb{R}$, consideriamo la funzione $f : \mathbb{R}^N \to \mathbb{R}^N$ definita da $f(\boldsymbol{x}) = \alpha \boldsymbol{x}$. Vediamo che è continua. Infatti, se $\alpha = 0$, si tratta della funzione costante con valore $\mathbf{0}$, e sappiamo che tale funzione è continua. Sia ora $\alpha \neq 0$. Allora, fissato $\varepsilon > 0$, essendo

$$\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| = \|\alpha\boldsymbol{x} - \alpha\boldsymbol{x}_0\| = \|\alpha(\boldsymbol{x} - \boldsymbol{x}_0)\| = |\alpha| \, \|\boldsymbol{x} - \boldsymbol{x}_0\| \,,$$

basta prendere $\delta = \frac{\varepsilon}{|\alpha|}$ per avere l'implicazione

$$\|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta \;\Rightarrow\; \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| < \varepsilon \,.$$

Esempio 3. Siano $E = \mathbb{R}^N$ e $F = \mathbb{R}$. Vediamo che la funzione $f : \mathbb{R}^N \to \mathbb{R}$ definita da $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ è continua su $\mathbb{R}^N$. Questo seguirà facilmente dalla disuguaglianza

$$\Big| \|\boldsymbol{x}\| - \|\boldsymbol{x}'\| \Big| \leq \|\boldsymbol{x} - \boldsymbol{x}'\| \,,$$

che ora dimostriamo. Si ha:

$$\begin{aligned} \|\boldsymbol{x}\| &= \|(\boldsymbol{x} - \boldsymbol{x}') + \boldsymbol{x}'\| \leq \|\boldsymbol{x} - \boldsymbol{x}'\| + \|\boldsymbol{x}'\| \,, \\ \|\boldsymbol{x}'\| &= \|(\boldsymbol{x}' - \boldsymbol{x}) + \boldsymbol{x}\| \leq \|\boldsymbol{x}' - \boldsymbol{x}\| + \|\boldsymbol{x}\| \,. \end{aligned}$$

Essendo $\|\boldsymbol{x} - \boldsymbol{x}'\| = \|\boldsymbol{x}' - \boldsymbol{x}\|$, si ha che

$$\|\boldsymbol{x}\| - \|\boldsymbol{x}'\| \leq \|\boldsymbol{x} - \boldsymbol{x}'\| \qquad \text{e} \qquad \|\boldsymbol{x}'\| - \|\boldsymbol{x}\| \leq \|\boldsymbol{x} - \boldsymbol{x}'\| \,,$$

da cui la disuguaglianza cercata. A questo punto, considerato un $\boldsymbol{x}_0 \in \mathbb{R}^N$ e fissato un $\varepsilon > 0$, basta prendere $\delta = \varepsilon$ per avere che

$$\|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta \;\Rightarrow\; \Big| \|\boldsymbol{x}\| - \|\boldsymbol{x}_0\| \Big| < \varepsilon \,.$$

I tre teoremi seguenti hanno dimostrazione analoga a quella vista nel corso di Analisi 1, in cui si supponeva che E fosse un sottoinsieme di $\mathbb{R}$.

Teorema 8 *Se $f, g : E \to \mathbb{R}$ sono continue in x_0 e $\alpha \in \mathbb{R}$, anche αf, $f + g$, $f - g$, $f \cdot g$ sono continue in x_0.*

Teorema 9 (della permanenza del segno) *Sia $g : E \to \mathbb{R}$ continua in x_0. Se $g(x_0) > 0$, allora esiste un intorno U di x_0 per cui $g(x) > 0$ per ogni $x \in U$. Se invece $g(x_0) < 0$, allora esiste un intorno U di x_0 per cui $g(x) < 0$ per ogni $x \in U$.*

Teorema 10 *Se $f, g : E \to \mathbb{R}$ sono continue in x_0 e $g(x_0) \neq 0$, anche $\frac{f}{g}$ è continua in x_0.*

Vediamo ora come si comporta una funzione composta di due funzioni continue.

Teorema 11 *Siano E, F, G tre spazi metrici, $f : E \to F$ continua in x_0 e $g : F \to G$ continua in $f(x_0)$; allora $g \circ f$ è continua in x_0.*

Dimostrazione. Fissato un intorno W di $[g \circ f](x_0) = g(f(x_0))$, per la continuità di g in $f(x_0)$ esiste un intorno V di $f(x_0)$ tale che $g(V) \subseteq W$. Allora, per la continuità di f in x_0, esiste un intorno U di x_0 tale che $f(U) \subseteq V$. Ne segue che $[g \circ f](U) \subseteq W$. ■

Consideriamo ora, per ogni $k = 1, 2, \ldots, D$, la funzione "k-esima proiezione" $p_k : \mathbb{R}^D \to \mathbb{R}$ definita da

$$p_k(x_1, x_2, \ldots, x_D) = x_k \,.$$

Teorema 12 *Le funzioni p_k sono continue.*

Dimostrazione. Consideriamo un punto $\boldsymbol{x}_0 = (x_1^0, x_2^0, \ldots, x_D^0) \in \mathbb{R}^D$ e fissiamo $\varepsilon > 0$. Notiamo che, per ogni $\boldsymbol{x} = (x_1, x_2, \ldots, x_D) \in \mathbb{R}^D$, si ha

$$|x_k - x_k^0| \leq \sqrt{\sum_{j=1}^{D} (x_j - x_j^0)^2} = d(\boldsymbol{x}, \boldsymbol{x}_0) \,,$$

per cui, prendendo $\delta = \varepsilon$, si ha:

$$d(\boldsymbol{x}, \boldsymbol{x}_0) < \delta \;\Rightarrow\; |p_k(\boldsymbol{x}) - p_k(\boldsymbol{x}_0)| = |x_k - x_k^0| < \varepsilon \,,$$

il che dimostra che p_k è continua in $\boldsymbol{x}_0$. ■

Supponiamo ora $f : E \to \mathbb{R}^M$. Consideriamo le "componenti" della funzione f definite da $f_k = p_k \circ f : E \to \mathbb{R}$, con $k = 1, 2, \ldots, M$, per cui si ha

$$f(x) = (f_1(x), f_2(x), \ldots, f_M(x)) \,.$$

Teorema 13 *La funzione f è continua in x_0 se e solo se lo sono tutte le sue componenti.*

Dimostrazione. Se f è continua in x_0, lo sono anche le f_k in quanto composte di funzioni continue. Viceversa, supponiamo che le componenti di f siano tutte continue in x_0. Fissato $\varepsilon > 0$, per ogni $k = 1, 2, \ldots, M$ esiste un $\delta_k > 0$ tale che

$$d(x, x_0) < \delta_k \;\Rightarrow\; |f_k(x) - f_k(x_0)| < \varepsilon \,.$$

Posto $\delta = \min\{\delta_1, \delta_2, \ldots, \delta_M\}$, si ha

$$d(x, x_0) < \delta \;\Rightarrow\; d(f(x), f(x_0)) = \sqrt{\sum_{j=1}^{M} (f_j(x) - f_j(x_0))^2} < \sqrt{M}\varepsilon \,,$$

il che, per l'arbitrarietà di ε, completa la dimostrazione. ■

Teorema 14 *Ogni applicazione lineare $\ell : \mathbb{R}^N \to \mathbb{R}^M$ è continua.*

Dimostrazione. Osserviamo che, essendo le proiezioni p_k lineari, le componenti $\ell_k = p_k \circ \ell$ dell'applicazione lineare ℓ sono anch'esse lineari. Consideriamo la base canonica $[\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_N]$ di $\mathbb{R}^N$, con

$$\begin{aligned} \boldsymbol{e}_1 &= (1, 0, 0, \dots, 0)\,, \\ \boldsymbol{e}_2 &= (0, 1, 0, \dots, 0)\,, \\ &\vdots \\ \boldsymbol{e}_N &= (0, 0, 0, \dots, 1)\,. \end{aligned}$$

Ogni vettore $\boldsymbol{x} = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$ si piò scrivere come

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_N\boldsymbol{e}_N = p_1(\boldsymbol{x})\boldsymbol{e}_1 + p_2(\boldsymbol{x})\boldsymbol{e}_2 + \ldots + p_N(\boldsymbol{x})\boldsymbol{e}_N\,.$$

Quindi, per ogni $k \in \{1, 2, \dots, M\}$,

$$\ell_k(\boldsymbol{x}) = p_1(\boldsymbol{x})\ell_k(\boldsymbol{e}_1) + p_2(\boldsymbol{x})\ell_k(\boldsymbol{e}_2) + \ldots + p_N(\boldsymbol{x})\ell_k(\boldsymbol{e}_N)\,,$$

per cui ℓ_k risulta essere combinazione lineare delle proiezioni $p_1, p_2, \dots, p_N$. Essendo queste ultime continue, anche ℓ_k è continua. Avendo tutte le componenti continue, ℓ è pertanto continua. ■

4 La nozione di limite

Consideriamo due spazi metrici E, F, un punto x_0 di E e una funzione

$$f : E \to F\,, \qquad \text{oppure} \qquad f : E \setminus \{x_0\} \to F\,,$$

non necessariamente definita in x_0. Supporremo inoltre che x_0 sia un "punto di accumulazione" di E: ogni intorno di x_0 contiene infiniti punti di E.

Definizione 2 *Si dice che $l \in F$ è il "limite di f in x_0", o anche "limite di $f(x)$ per x che tende a x_0" e si scrive*

$$l = \lim_{x \to x_0} f(x)\,,$$

se

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \ \forall x \in E \quad 0 < d(x, x_0) < \delta \ \Rightarrow \ d(f(x), l) < \varepsilon\,,$$

o equivalentemente,

$$\forall V, \text{ intorno di } l \quad \exists U, \text{ intorno di } x_0 : \quad f(U \setminus \{x_0\}) \subseteq V.$$

Talvolta si scrive anche $f(x) \to l$ *per* $x \to x_0$.

Per cominciare, verifichiamo l'unicità del limite.

Teorema 15 *Se esiste, il limite di* f *in* x_0 *è unico.*

Dimostrazione. Supponiamo per assurdo che ce ne siano due diversi, l e l'. Prendiamo $\varepsilon = \frac{1}{2}d(l,l')$. Allora esiste un $\delta > 0$ tale che

$$0 < d(x,x_0) < \delta \;\Rightarrow\; d(f(x),l) < \varepsilon\,,$$

ed esiste un $\delta' > 0$ tale che

$$0 < d(x,x_0) < \delta' \;\Rightarrow\; d(f(x),l') < \varepsilon\,.$$

Sia $x \neq x_0$ tale che $d(x,x_0) < \delta$ e $d(x,x_0) < \delta'$ (tale x esiste perché x_0 è di accumulazione). Allora

$$d(l',l) \leq d(l,f(x)) + d(f(x),l') < 2\varepsilon = d(l',l)\,,$$

una contraddizione. ■

Il seguente teorema evidenzia il legame stretto che intercorre tra i concetti di limite e di continuità.

Teorema 16 *Considerata la funzione* $f : E \to F$, *si ha che*

$$f \text{ è continua in } x_0 \quad \Leftrightarrow \quad \lim_{x\to x_0} f(x) = f(x_0)\,.$$

Iniziamo a vedere le proprietà dei limiti che vengono direttamente ereditate dalle funzioni continue. Nei teoremi seguenti, con relativo corollario, le funzioni f e g sono definite su E o su $E \setminus \{x_0\}$, indifferentemente.

Teorema 17 (della permanenza del segno) *Sia* $g : E\setminus\{x_0\} \to \mathbb{R}$ *tale che*

$$\lim_{x\to x_0} g(x) > 0\,,$$

allora esiste un $\delta > 0$ *tale che*

$$0 < d(x,x_0) < \delta \;\Rightarrow\; g(x) > 0\,.$$

Analogamente, se

$$\lim_{x\to x_0} g(x) < 0\,,$$

allora esiste un $\delta > 0$ *tale che*

$$0 < d(x,x_0) < \delta \;\Rightarrow\; g(x) < 0\,.$$

Corollario 1 *Se* $g(x) \leq 0$ *per ogni* x *in un intorno di* x_0*, allora, qualora il limite esista, si ha*

$$\lim_{x \to x_0} g(x) \leq 0 \,.$$

Analogamente, se $g(x) \geq 0$ *per ogni* x *in un intorno di* x_0*, allora, qualora il limite esista, si ha*

$$\lim_{x \to x_0} g(x) \geq 0 \,.$$

Vediamo ora come si comporta il limite nei confronti delle operazioni in $\mathbb{R}$.

Teorema 18 *Siano* $f, g : E \setminus \{x_0\} \to \mathbb{R}$ *tali che*

$$l_1 = \lim_{x \to x_0} f(x) \,, \quad l_2 = \lim_{x \to x_0} g(x) \,.$$

Allora

$$\lim_{x \to x_0} [f(x)g(x)] = l_1 l_2 \,;$$

inoltre, se $l_2 \neq 0$*,*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{l_1}{l_2} \,.$$

Consideriamo ora una funzione composta $g \circ f$. Abbiamo due possibili situazioni.

Teorema 19 *Sia* $f : E \to F$*, oppure* $f : E \setminus \{x_0\} \to F$*, tale che*

$$\lim_{x \to x_0} f(x) = l \,.$$

Se $g : F \to G$ *è continua in* l*, allora*

$$\lim_{x \to x_0} g(f(x)) = g(l) \,.$$

In altri termini,

$$\lim_{x \to x_0} g(f(x)) = g(\lim_{x \to x_0} f(x)) \,.$$

<u>Dimostrazione</u>. Riguardando la definizione di limite, si ha che $\tilde{f} : E \to \mathbb{R}$ ivi definita è continua in x_0 e g è continua in $l = \tilde{f}(x_0)$. Pertanto, $g \circ \tilde{f}$ è continua in x_0, da cui

$$\lim_{x \to x_0} g(f(x)) = \lim_{x \to x_0} g(\tilde{f}(x)) = g(\tilde{f}(x_0)) = g(l) \,.$$

∎

Teorema 20 *Sia $f : E \to F$, oppure $f : E \setminus \{x_0\} \to F$, tale che*

$$\lim_{x \to x_0} f(x) = l\,.$$

Supponiamo che l sia un punto di accumulazione di F e che la funzione

$$g : F \to G\,, \qquad \textit{oppure} \qquad g : F \setminus \{l\} \to G\,,$$

non necessariamente definita in l, sia tale che

$$\lim_{y \to l} g(y) = L\,.$$

Se $f(x) \neq l$ per ogni $x \in E \setminus \{x_0\}$, allora

$$\lim_{x \to x_0} g(f(x)) = L\,.$$

Dimostrazione. Consideriamo nuovamente la funzione $\tilde{f} : E \to F$, continua in x_0 con $\tilde{f}(x_0) = l$. Analogamente, consideriamo la funzione $\tilde{g} : F \to G$ così definita:

$$\tilde{g}(y) = \begin{cases} g(y) & \text{se } y \neq l\,, \\ L & \text{se } y = l\,. \end{cases}$$

Essa è continua in l con $\tilde{g}(l) = L$. Consideriamo la funzione composta $\tilde{g} \circ \tilde{f}$, che per quanto sopra è continua in x_0 con $\tilde{g}(\tilde{f}(x_0)) = \tilde{g}(l) = L$. Essendo $f(x) \neq l$ per ogni x, si ha che, per $x \in E \setminus \{x_0\}$,

$$g(f(x)) = \tilde{g}(f(x)) = \tilde{g}(\tilde{f}(x))\,,$$

e pertanto,

$$\lim_{x \to x_0} g(f(x)) = \lim_{x \to x_0} \tilde{g}(\tilde{f}(x)) = \tilde{g}(\tilde{f}(x_0)) = L\,.$$

■

Nota. La conclusione del teorema precedente si riassume con la formula

$$\lim_{x \to x_0} g(f(x)) = \lim_{y \to \lim\limits_{x \to x_0} f(x)} g(y)\,.$$

Spesso si dice che si è operato il "cambio di variabile $y = f(x)$". Riguardando inoltre le ipotesi dello stesso teorema, si vede subito che è sufficiente richiedere che sia $f(x) \neq l$ per gli x tali che $0 < d(x, x_0) < \delta$. Ciò è dovuto al fatto che la nozione di limite è, in un certo senso, di tipo "locale". Questa osservazione vale in generale e verrà spesso usata in seguito.

Esempio. Sia ora $f : \mathbb{R}^2 \to \mathbb{R}$ definita da $f(x, y) = x^2 + y^2$ e $g : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ definita da $g(z) = z \sin(1/z)$. Il Teorema 20 qui può essere applicato per concludere che

$$\lim_{(x,y) \to (0,0)} g(f(x, y)) = \lim_{z \to \lim\limits_{(x,y) \to (0,0)} f(x,y)} g(z) = \lim_{z \to 0} g(z) = 0\,.$$

Siamo infine interessati a studiare il limite di una funzione $f : E \setminus \{x_0\} \to \mathbb{R}^M$, dove x_0 è un punto di accumulazione di uno spazio metrico E. Consideriamo le sue componenti $f_k : E \setminus \{x_0\} \to \mathbb{R}$ di f, con $k = 1, 2, \dots, M$, per cui si ha:

$$f(x) = (f_1(x), f_2(x), \dots, f_M(x)) .$$

Teorema 21 *Il limite* $\lim_{x \to x_0} f(x) = \boldsymbol{l} \in \mathbb{R}^M$ *esiste se e solo se esistono i limiti* $\lim_{x \to x_0} f_k(x) = l_k \in \mathbb{R}$, *per ogni* $k = 1, 2, \dots, M$. *In tal caso, si ha* $\boldsymbol{l} = (l_1, l_2, \dots, l_M)$. *Vale quindi la formula*

$$\lim_{x \to x_0} f(x) = \left(\lim_{x \to x_0} f_1(x), \lim_{x \to x_0} f_2(x), \dots, \lim_{x \to x_0} f_M(x)\right) .$$

Dimostrazione. Segue direttamente dal teorema sulla continuità delle componenti di una funzione continua. ■

Finora abbiamo considerato due spazi metrici E, F, un punto x_0 di accumulazione per E e una funzione $f : E \to F$, oppure $f : E \setminus \{x_0\} \to F$. Siccome l'eventuale valore di f in x_0 è ininfluente ai fini dell'esistenza o meno del limite, nonchè del suo effettivo valore, da ora in poi per semplicità considereremo solo il caso $f : E \setminus \{x_0\} \to F$.

Si può verificare che tutte le considerazioni fatte continuano a valere per una funzione $f : \widehat{E} \setminus \{x_0\} \to F$, con $\widehat{E} \subseteq E$, purché x_0 sia di accumulazione per $\widehat{E}$: ogni intorno di x_0 deve contenere infiniti punti di $\widehat{E}$.

Sia ora $f : E \setminus \{x_0\} \to F$, e sia $\widehat{E} \subseteq E$. Possiamo considerare la restrizione di f a $\widehat{E} \setminus \{x_0\}$: è la funzione $\hat{f} : \widehat{E} \setminus \{x_0\} \to F$ i cui valori coincidono con quelli di f: si ha $\hat{f}(x) = f(x)$ per ogni $x \in \widehat{E} \setminus \{x_0\}$. Talvolta si scrive $\hat{f} = f|_{\widehat{E}}$.

Teorema 22 *Se esiste il limite di* f *in* x_0 *e* x_0 *è di accumulazione anche per* $\widehat{E}$, *allora esiste anche il limite di* $\hat{f}$ *in* x_0 *e ha lo stesso valore:*

$$\lim_{x \to x_0} \hat{f}(x) = \lim_{x \to x_0} f(x) .$$

Dimostrazione. Segue immediatamente dalla definizione di $\hat{f}$. ■

Il teorema precedente viene spesso usato per stabilire la non esistenza del limite per la funzione f: a tal scopo, è sufficiente trovare due diverse restrizioni lungo le quali i valori del limite differiscono.

Esempio 1. La funzione $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$, definita da

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

non ha limite per $(x, y) \to (0, 0)$, come si vede considerando le restrizioni alle due rette $\{(x, y) : x = 0\}$ e $\{(x, y) : x = y\}$.

Esempio 2. Più sorprendente è la funzione definita da

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

per la quale le restrizioni a tutte le rette passanti per $(0,0)$ hanno limite 0, ma la restrizione alla parabola $\{(x,y) : y = x^2\}$ vale costantemente $\frac{1}{2}$.

Esempio 3. Dimostriamo invece che

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y^2}{x^2 + y^2} = 0\,.$$

Fissiamo un $\varepsilon > 0$. Dopo aver verificato che

$$\frac{x^2 y^2}{x^2 + y^2} \leq \frac{1}{2}\,(x^2 + y^2)\,,$$

risulta naturale prendere $\delta = \sqrt{2\varepsilon}$, per avere che

$$d((x,y),(0,0)) < \delta \;\Rightarrow\; \left|\frac{x^2 y^2}{x^2 + y^2} - 0\right| < \varepsilon\,.$$

Consideriamo ora il caso in cui E è uno spazio metrico qualunque ed $F = \mathbb{R}$. Risulterà talvolta utile il seguente "teorema dei due carabinieri".

Teorema 23 *Supponiamo di avere due funzioni f_1, f_2 per cui*

$$\lim_{x\to x_0} f_1(x) = \lim_{x\to x_0} f_2(x) = l\,.$$

Se $f : E \setminus \{x_0\} \to \mathbb{R}$ è tale che, per ogni x,

$$f_1(x) \leq f(x) \leq f_2(x)\,,$$

allora

$$\lim_{x\to x_0} f(x) = l\,.$$

Dimostrazione. Fissato $\varepsilon > 0$, esistono $\delta_1 > 0$ e $\delta_2 > 0$ tali che

$$\begin{aligned} 0 < d(x,x_0) < \delta_1 \quad &\Rightarrow \quad l - \varepsilon < f_1(x) < l + \varepsilon\,,\\ 0 < d(x,x_0) < \delta_2 \quad &\Rightarrow \quad l - \varepsilon < f_2(x) < l + \varepsilon\,. \end{aligned}$$

Se $\delta = \min\{\delta_1, \delta_2\}$, allora

$$0 < d(x,x_0) < \delta \quad \Rightarrow \quad l - \varepsilon < f_1(x) \leq f(x) \leq f_2(x) < l + \varepsilon\,,$$

il che dimostra la tesi. ■

5 La retta ampliata

Consideriamo la funzione $\varphi : \mathbb{R} \to]-1,1[$, definita da

$$\varphi(x) = \frac{x}{1+|x|}\,.$$

Possiamo definire una nuova distanza su $\mathbb{R}$:

$$\tilde{d}(x,x') = |\varphi(x) - \varphi(x')|\,.$$

Si può vedere che ogni intorno per la nuova distanza è anche intorno per la vecchia distanza, e viceversa.

Introduciamo ora il nuovo insieme $\widetilde{\mathbb{R}}$, definito come unione di $\mathbb{R}$ e di due nuovi elementi, che indicheremo con $-\infty$ e $+\infty$:

$$\widetilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}\,.$$

L'insieme $\widetilde{\mathbb{R}}$ risulta totalmente ordinato se si mantiene l'ordine esistente tra coppie di numeri reali e si pone inoltre, per ogni $x \in \mathbb{R}$,

$$-\infty < x < +\infty\,.$$

Consideriamo la funzione $\tilde{\varphi} : \widetilde{\mathbb{R}} \to [-1,1]$, definita da

$$\tilde{\varphi}(x) = \begin{cases} -1 & \text{se } x = -\infty\,, \\ \varphi(x) & \text{se } x \in \mathbb{R}\,, \\ 1 & \text{se } x = +\infty\,. \end{cases}$$

Definiamo, per $x, x' \in \widetilde{\mathbb{R}}$,

$$\tilde{d}(x,x') = |\tilde{\varphi}(x) - \tilde{\varphi}(x')|\,.$$

Si verifica facilmente che $\tilde{d}$ è una distanza su $\widetilde{\mathbb{R}}$. In questo modo, $\widetilde{\mathbb{R}}$ risulta uno spazio metrico.

Si può verificare che un intorno di $+\infty$ è un insieme che contiene, oltre al punto $+\infty$, un intervallo del tipo $]\alpha, +\infty[$, per un certo $\alpha \in \mathbb{R}$.

Analogamente, un intorno di $-\infty$ è un insieme che contiene, oltre a $-\infty$, un intervallo del tipo $]-\infty, \beta[$, per un certo $\beta \in \mathbb{R}$.

Vediamo ora come si traduce la definizione di limite in alcuni casi in cui compaiono gli elementi $+\infty$ o $-\infty$. Ad esempio, sia $E \subseteq \mathbb{R}$, F uno spazio metrico e $f : E \to F$ una funzione. Considerando E come sottoinsieme di $\widetilde{\mathbb{R}}$, si ha che $+\infty$ è punto di accumulazione per E se e solo se E non è limitato superiormente. In tal caso, si ha:

$$\begin{aligned} \lim_{x\to+\infty} f(x) = l \in F \quad &\Leftrightarrow \quad \forall V \text{ intorno di } l \;\; \exists U \text{ intorno di } +\infty : \\ &\qquad\qquad f(U \cap E) \subseteq V \\ &\Leftrightarrow \quad \forall \varepsilon > 0 \;\; \exists \alpha \in \mathbb{R} : \quad x > \alpha \;\Rightarrow\; d(f(x), l) < \varepsilon\,. \end{aligned}$$

Analogamente, se E non è limitato inferiormente, si ha:

$$\lim_{x \to -\infty} f(x) = l \in F \quad \Leftrightarrow \quad \forall \varepsilon > 0 \quad \exists \beta \in \mathbb{R}: \quad x < \beta \ \Rightarrow \ d(f(x), l) < \varepsilon\,.$$

Si noti che

$$\lim_{x \to +\infty} f(x) = l \quad \Leftrightarrow \quad \lim_{x \to -\infty} f(-x) = l\,.$$

Vediamo ora il caso in cui E sia uno spazio metrico ed $F = \mathbb{R}$, considerato come sottoinsieme di $\widetilde{\mathbb{R}}$. Supponiamo che x_0 sia di accumulazione per E e consideriamo una funzione $f : E \to \mathbb{R}$, o $f : E \setminus \{x_0\} \to \mathbb{R}$. Si ha:

$$\begin{aligned}\lim_{x \to x_0} f(x) = +\infty \quad &\Leftrightarrow \quad \forall V \text{ intorno di } +\infty \ \exists U \text{ intorno di } x_0 : \\ &\qquad f(U \setminus \{x_0\}) \subseteq V \\ &\Leftrightarrow \quad \forall \alpha \in \mathbb{R} \quad \exists \delta > 0 : \quad 0 < d(x, x_0) < \delta \ \Rightarrow \ f(x) > \alpha\,;\end{aligned}$$

analogamente,

$$\lim_{x \to x_0} f(x) = -\infty \quad \Leftrightarrow \quad \forall \beta \in \mathbb{R} \quad \exists \delta > 0 : \quad 0 < d(x, x_0) < \delta \ \Rightarrow \ f(x) < \beta\,.$$

Si noti che

$$\lim_{x \to x_0} f(x) = +\infty \quad \Leftrightarrow \quad \lim_{x \to x_0} (-f(x)) = -\infty\,.$$

Le situazioni considerate in precedenza possono talvolta presentarsi assieme. Ad esempio, se $E \subseteq \mathbb{R}$ non è limitato superiormente ed $F = \mathbb{R}$, si avrà

$$\begin{aligned}\lim_{x \to +\infty} f(x) = +\infty \quad &\Leftrightarrow \quad \forall V \text{ intorno di } +\infty \ \exists U \text{ intorno di } +\infty : \\ &\qquad f(U \cap E) \subseteq V \\ &\Leftrightarrow \quad \forall \alpha \in \mathbb{R} \quad \exists \alpha' \in \mathbb{R} : \quad x > \alpha' \ \Rightarrow \ f(x) > \alpha\,;\end{aligned}$$

analogamente,

$$\lim_{x \to +\infty} f(x) = -\infty \quad \Leftrightarrow \quad \forall \beta \in \mathbb{R} \quad \exists \alpha \in \mathbb{R} : \quad x > \alpha \ \Rightarrow \ f(x) < \beta\,.$$

Se invece $E \subseteq \mathbb{R}$ non è limitato inferiormente ed $F = \mathbb{R}$, si avrà

$$\lim_{x \to -\infty} f(x) = +\infty \quad \Leftrightarrow \quad \forall \alpha \in \mathbb{R} \quad \exists \beta \in \mathbb{R} : \quad x < \beta \ \Rightarrow \ f(x) > \alpha\,;$$

analogamente,

$$\lim_{x \to -\infty} f(x) = -\infty \quad \Leftrightarrow \quad \forall \beta \in \mathbb{R} \quad \exists \beta' \in \mathbb{R} : \quad x < \beta' \ \Rightarrow \ f(x) < \beta\,.$$

Vediamo ad esempio il caso di una successione $(a_n)_n$ in uno spazio metrico F. Abbiamo quindi una funzione $f : \mathbb{N} \to F$ definita da $f(n) = a_n$. Considerando $\mathbb{N}$ come sottoinsieme di $\widetilde{\mathbb{R}}$, si vede che l'unico punto di accumulazione è $+\infty$. Adattando la definizione di limite a questo caso, possiamo scrivere:

$$\lim_{n\to+\infty} a_n = l \in F \quad \Leftrightarrow \quad \forall \varepsilon > 0 \quad \exists \bar{n} \in \mathbb{N} : \quad n \geq \bar{n} \;\Rightarrow\; d(a_n, l) < \varepsilon \,.$$

Pertanto, spesso il limite di una successione si denota semplicemente con $\lim\limits_n a_n$, sottintendendo che $n \to +\infty$.

Teorema 24 *La funzione f è continua in x_0 se e solo se, presa una successione $(a_n)_n$ in E, si ha*

$$\lim_n a_n = x_0 \quad \Rightarrow \quad \lim_n f(a_n) = f(x_0)\,.$$

Dimostrazione. Supponiamo che f sia continua in x_0, e sia $(a_n)_n$ una successione in E tale che $\lim_n a_n = x_0$. Per il Teorema 19 sul limite di una funzione composta,

$$\lim_n f(a_n) = f(\lim_n a_n) = f(x_0)\,,$$

cosicchè una delle due implicazioni è dimostrata.

Ragioniamo ora per contrapposizione, e supponiamo che f non sia continua in x_0. Questo significa che esiste un $\varepsilon > 0$ tale che, per ogni $\delta > 0$, esiste almeno un $x \in E$ per cui $d(x, x_0) < \delta$ e $d(f(x), f(x_0)) \geq \varepsilon$. Prendendo $\delta = \frac{1}{n+1}$, per ogni $n \in \mathbb{N}$ esiste pertanto un a_n in E tale che $d(a_n, x_0) < \frac{1}{n+1}$ e $d(f(a_n), f(x_0)) \geq \varepsilon$. Ne segue che $\lim_n a_n = x_0$, ma sicuramente non può essere che $\lim_n f(a_n) = f(x_0)$. ■

Consideriamo ora il caso in cui x_0 sia un punto di accumulazione e $f : E \setminus \{x_0\} \to F$ una funzione. Come immediata conseguenza del teorema precedente, otteniamo il seguente

Corollario 2 *Avremo che*

$$\lim_{x\to x_0} f(x) = l$$

se e solo se, presa una successione $(a_n)_n$ in $E \setminus \{x_0\}$, si ha

$$\lim_n a_n = x_0 \quad \Rightarrow \quad \lim_n f(a_n) = l\,.$$

Sia ora U un sottoinsieme dello spazio metrico E. Possiamo caratterizzare la nozione di punto aderente a U facendo uso delle successioni.

Teorema 25 *Un punto $x \in E$ è aderente a U se e solo se esiste una successione $(a_n)_n$ in U tale che $\lim_n a_n = x$.*

Dimostrazione. Se x è aderente a U, allora l'intersezione $B\left(x, \frac{1}{n+1}\right) \cap U$ è non vuota, per ogni $n \in \mathbb{N}$, per cui posso sceglierne un elemento, che chiamo a_n. In questo modo, ho costruito una successione $(a_n)_n$ in U, ed è facile vedere che essa ha limite x. Una delle due implicazioni è così dimostrata.

Supponiamo ora che esista una successione $(a_n)_n$ in U tale che $\lim_n a_n = x$. Allora, fissato $\rho > 0$, esiste un $\bar{n} \in \mathbb{N}$ tale che

$$n \geq \bar{n} \;\Rightarrow\; d(a_n, x) < \rho\,,$$

ossia $a_n \in B(x, \rho)$. Quindi, $B(x, \rho) \cap U$ è non vuoto, e questo dimostra che x è aderente a U. ■

6 Insiemi compatti

Data che sia una successione $(a_n)_n$, una sua "sottosuccessione" si ottiene selezionando una successione strettamente crescente di indici $(n_k)_k$ e considerando la funzione composta

$$k \mapsto n_k \mapsto a_{n_k}\,.$$

Teorema 26 *Se una successione ha limite, allora tutte le sue sottosuccessioni hanno lo stesso limite.*

Dimostrazione. Essendo gli indici n_k in $\mathbb{N}$, dalla $n_{k+1} > n_k$ si deduce che $n_{k+1} \geq n_k + 1$ e, per induzione, che $n_k \geq k$, per ogni k. Ne segue che $\lim_k n_k = +\infty$. Pertanto,

$$\lim_{k\to+\infty} a_{n_k} = \lim_{n \to \lim\limits_{k\to+\infty} n_k} a_n = \lim_{n\to+\infty} a_n\,.$$ ■

In uno spazio metrico E, diremo che un sottoinsieme U è "compatto" se ogni successione $(a_n)_n$ in U possiede una sottosuccessione $(a_{n_k})_k$ che ha limite in U. Il Teorema di Bolzano–Weierstrass afferma quindi che, se $E = \mathbb{R}$, gli intervalli del tipo $U = [a, b]$ sono compatti.

Teorema 27 *Ogni insieme compatto di uno spazio metrico E è chiuso e limitato.*

Dimostrazione. Supponiamo dapprima che U sia compatto. Preso un $x \in \overline{U}$, esiste una successione $(a_n)_n$ in U tale che $\lim_n a_n = x$. Essendo U compatto, esiste una sottosuccessione $(a_{n_k})_k$ che ha limite in U. Ma, essendo una sottosuccessione, deve essere $\lim_k a_{n_k} = x$, per cui $x \in U$. Quindi, ogni punto aderente di U appartiene ad U, per cui U è chiuso.

Fissiamo ora un $x_0 \in U$ qualsiasi e dimostriamo che, se $n \in \mathbb{N}$ è sufficientemente grande, allora $U \subseteq B(x_0, n)$. Per assurdo, se così non fosse, potrei trovare una successione $(a_n)_n$ in U tale che $d(a_n, x_0) \geq n$, per ogni $n \in \mathbb{N}$. Ma, essendo U compatto, esiste una sottosuccessione $(a_{n_k})_k$ che ha un certo limite $\bar{x} \in U$. Usando la disuguaglianza triangolare, si ha che

$$|d(a_{n_k}, x_0) - d(\bar{x}, x_0)| \leq d(a_{n_k}, \bar{x})\,,$$

da cui segue che $\lim_k d(a_{n_k}, x_0) = d(\bar{x}, x_0)$, mentre dovrebbe essere

$$\lim_k d(a_{n_k}, x_0) = +\infty\,,$$

una contraddizione. Pertanto, U deve essere limitato. ■

Ora vorremmo focalizzare la nostra attenzione sui sottoinsiemi compatti di $\mathbb{R}^M$.

Teorema 28 *Un sottoinsieme $\mathbb{R}^M$ è compatto* se e solo se *è chiuso e limitato.*

Dimostrazione. Sappiamo già che ogni compatto è chiuso e limitato. Supponiamo ora che U sia un sottoinsieme chiuso e limitato di $\mathbb{R}^M$. Supporremo per semplicità $M = 2$. Allora U è contenuto in un rettangolo

$$I = [a_1, b_1] \times [a_2, b_2] .$$

Sia $(\boldsymbol{a}_n)_n$ una successione in U. Si ha $\boldsymbol{a}_n = (a_n^1, a_n^2)$, con $a_n^1 \in [a_1, b_1]$ e $a_n^2 \in [a_2, b_2]$. Per la proprietà di Bolzano–Weierstrass, esiste una sottosuccessione $(a_{n_k}^1)_k$ che ha un limite $l_1 \in [a_1, b_1]$. Consideriamo la sottosuccessione $(a_{n_k}^2)_k$, con gli stessi indici di quella appena trovata. Per la proprietà di Bolzano–Weierstrass, esiste una sottosuccessione $(a_{n_{k_j}}^2)_j$ che ha un limite $l_2 \in [a_2, b_2]$. Osservando che

$$d\Big(\boldsymbol{a}_{n_{k_j}}, (l_1, l_2)\Big) = \sqrt{(a_{n_{k_j}}^1 - l_1)^2 + (a_{n_{k_j}}^2 - l_2)^2} ,$$

se ne deduce che

$$\lim_j \boldsymbol{a}_{n_{k_j}} = (l_1, l_2) .$$

Il punto $\boldsymbol{l} = (l_1, l_2)$ è aderente ad U. Essendo U chiuso, $\boldsymbol{l}$ appartiene ad U. ■

Nel seguito, diremo che una funzione $f : U \to \mathbb{R}$ è “limitata superiormente” (o “limitata inferiormente”) se lo è la sua immagine $f(U)$. Diremo che f è “limitata” se è sia limitata superiormente che inferiormente. Diremo che “f ha massimo” (o “f ha minimo”) se $f(U)$ ce l’ha. Nel caso in cui f abbia massimo, chiameremo “punto di massimo” ogni $\bar{x}$ per cui $f(\bar{x}) = \max f(U)$; analoga definizione per “punto di minimo”.

Teorema 29 (di Weierstrass) *Se U è un insieme compatto e $f : U \to \mathbb{R}$ è una funzione continua, allora f ha massimo e minimo.*

Dimostrazione. Sia $s = \sup f(U)$. Dimostreremo che esiste un punto di massimo, ossia un $\bar{x} \in U$ tale che $f(\bar{x}) = s$.

Notiamo che è possibile trovare una successione $(y_n)_n$ in $f(U)$ tale che $\lim_n y_n = s$: se $s \in \mathbb{R}$, per ogni $n \geq 1$ possiamo trovare un $y_n \in f(U)$ per cui $s - \frac{1}{n} < y_n \leq s$; se invece $s = +\infty$, per ogni n esiste un $y_n \in f(U)$ tale che $y_n > n$.

In corrispondenza, possiamo trovare una successione $(x_n)_n$ in U tale che $f(x_n) = y_n$. Essendo U compatto, esiste una sottosuccessione $(x_{n_k})_k$ che ha un limite $\bar{x} \in U$. Siccome $\lim_n y_n = s$ e $y_{n_k} = f(x_{n_k})$, la sottosuccessione $(y_{n_k})_k$ ha anch’essa limite s. Allora, per la continuità di f,

$$f(\bar{x}) = f(\lim_k x_{n_k}) = \lim_k f(x_{n_k}) = \lim_k y_{n_k} = s .$$

Il teorema è così dimostrato, per quanto riguarda l’esistenza del massimo. Per il minimo, si procede in modo analogo (oppure, si considera la funzione continua $g = -f$ e si usa il fatto che g ha massimo). ■

7 Calcolo differenziale: funzioni da $\mathbb{R}^N$ a $\mathbb{R}$

In questa sezione, E sarà un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di E e $f : E \to \mathbb{R}$ una funzione. Vogliamo estendere il concetto di derivata già introdotto nel caso $N = 1$. Iniziamo con il fissare una "direzione", ossia un vettore $\boldsymbol{v} \in \mathbb{R}^N$ tale che $\|\boldsymbol{v}\| = 1$ (detto anche "versore"). Chiamiamo, se esiste, "derivata direzionale" di f in $\boldsymbol{x}_0$ nella direzione $\boldsymbol{v}$ il seguente limite

$$\lim_{t \to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t} \,,$$

che verrà indicato con il simbolo

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) \,.$$

Se $\boldsymbol{v}$ coincide con un elemento $\boldsymbol{e}_k$ della base canonica $[\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N]$ di $\mathbb{R}^N$, la derivata direzionale si chiamerà "derivata parziale" k-esima di f in $\boldsymbol{x}_0$ e si indicherà con

$$\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) \,.$$

Se $\boldsymbol{x}_0 = (x_1^0, x_2^0, \ldots, x_N^0)$, si ha quindi:

$$\begin{aligned} \frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) &= \lim_{t \to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{e}_k) - f(\boldsymbol{x}_0)}{t} \\ &= \lim_{t \to 0} \frac{f(x_1^0, x_2^0, \ldots, x_k^0 + t, \ldots, x_N^0) - f(x_1^0, x_2^0, \ldots, x_k^0, \ldots, x_N^0)}{t} \,, \end{aligned}$$

per cui si usa parlare di "derivata rispetto alla k-esima variabile".

Esistono delle funzioni che, pur avendo derivate direzionali in tutte le possibili direzioni, non sono continue. Ad esempio, la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita da

$$f(x, y) = \begin{cases} \dfrac{x^4 y^2}{(x^4 + y^2)^2} & \text{se } (x, y) \neq (0, 0) \,, \\[2ex] 0 & \text{se } (x, y) = (0, 0) \,, \end{cases}$$

ha tutte le derivate direzionali nulle in $\boldsymbol{x}_0 = (0, 0)$, ma non è continua in tale punto, come si vede considerando la restrizione alla parabola $\{(x, y) \in \mathbb{R}^2 : y = x^2\}$. Questo fatto ci porta a cercare una generalizzazione più appropriata del concetto di derivata.

Definizione 3 *Diremo che la funzione f è "differenziabile" in $\boldsymbol{x}_0$ se esiste una applicazione lineare $\ell : \mathbb{R}^N \to \mathbb{R}$ per cui si possa scrivere*

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x}) \,,$$

dove r è una funzione tale che

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0 \,.$$

Se f è differenziabile in $\boldsymbol{x}_0$, l'applicazione lineare ℓ si chiama "differenziale" di f in $\boldsymbol{x}_0$ e si indica con il simbolo

$$df(\boldsymbol{x}_0)\,.$$

Teorema 30 *Se f è differenziabile in $\boldsymbol{x}_0$, allora f è continua in $\boldsymbol{x}_0$.*

<u>Dimostrazione</u>. Sappiamo che l'applicazione $\ell = df(\boldsymbol{x}_0)$, essendo lineare, è continua e $\ell(\mathbf{0}) = 0$. Ne segue che

$$\begin{aligned}\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) &= \lim_{\boldsymbol{x}\to\boldsymbol{x}_0}[f(\boldsymbol{x}_0) + \ell(\boldsymbol{x}-\boldsymbol{x}_0) + r(\boldsymbol{x})]\\ &= f(\boldsymbol{x}_0) + \ell(\mathbf{0}) + \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} r(\boldsymbol{x})\\ &= f(\boldsymbol{x}_0) + \lim_{\boldsymbol{x}\to\boldsymbol{x}_0}\frac{r(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}\lim_{\boldsymbol{x}\to\boldsymbol{x}_0}\|\boldsymbol{x}-\boldsymbol{x}_0\|\\ &= f(\boldsymbol{x}_0)\,.\end{aligned}$$

■

Seguendo un'abitudine consolidata per le applicazioni lineari, si usa spesso scrivere $df(\boldsymbol{x}_0)\boldsymbol{h}$ invece di $df(\boldsymbol{x}_0)(\boldsymbol{h})$.

Teorema 31 *Se f è differenziabile in $\boldsymbol{x}_0$, allora esistono tutte le derivate direzionali di f in $\boldsymbol{x}_0$: per ogni direzione $\boldsymbol{v}\in\mathbb{R}^N$ si ha*

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = df(\boldsymbol{x}_0)\boldsymbol{v}\,.$$

<u>Dimostrazione</u>. Usando la definizione di differenziale, abbiamo

$$\begin{aligned}\lim_{t\to 0}\frac{f(\boldsymbol{x}_0+t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t} &= \lim_{t\to 0}\frac{df(\boldsymbol{x}_0)(t\boldsymbol{v}) + r(\boldsymbol{x}_0+t\boldsymbol{v})}{t}\\ &= \lim_{t\to 0}\frac{t\,df(\boldsymbol{x}_0)\boldsymbol{v} + r(\boldsymbol{x}_0+t\boldsymbol{v})}{t}\\ &= df(\boldsymbol{x}_0)\boldsymbol{v} + \lim_{t\to 0}\frac{r(\boldsymbol{x}_0+t\boldsymbol{v})}{t}\,;\end{aligned}$$

d'altra parte, essendo $\|\boldsymbol{v}\| = 1$, si ha

$$\lim_{t\to 0}\left|\frac{r(\boldsymbol{x}_0+t\boldsymbol{v})}{t}\right| = \lim_{t\to 0}\frac{|r(\boldsymbol{x}_0+t\boldsymbol{v})|}{\|(\boldsymbol{x}_0+t\boldsymbol{v})-\boldsymbol{x}_0\|} = \lim_{\boldsymbol{x}\to\boldsymbol{x}_0}\frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} = 0\,,$$

da cui la tesi. ■

In particolare, se $\boldsymbol{v}$ coincide con un elemento $\boldsymbol{e}_k$ della base canonica $[\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N]$, si ha:

$$\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) = df(\boldsymbol{x}_0)\boldsymbol{e}_k\,.$$

Scrivendo il vettore $\boldsymbol{h} \in \mathbb{R}^N$ come $\boldsymbol{h} = h_1\boldsymbol{e}_1 + h_2\boldsymbol{e}_2 + \ldots + h_N\boldsymbol{e}_N$, abbiamo

$$\begin{aligned} df(\boldsymbol{x}_0)\boldsymbol{h} &= h_1 df(\boldsymbol{x}_0)\boldsymbol{e}_1 + h_2 df(\boldsymbol{x}_0)\boldsymbol{e}_2 + \ldots + h_N df(\boldsymbol{x}_0)\boldsymbol{e}_N \\ &= h_1\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0) + h_2\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0) + \ldots + h_N\frac{\partial f}{\partial x_N}(\boldsymbol{x}_0)\,, \end{aligned}$$

ossia

$$df(\boldsymbol{x}_0)\boldsymbol{h} = \sum_{k=1}^N \frac{\partial f}{\partial x_k}(\boldsymbol{x}_0)h_k\,.$$

Introducendo il vettore "gradiente" di f in $\boldsymbol{x}_0$

$$\nabla f(\boldsymbol{x}_0) = \left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0), \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0), \ldots, \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0)\right)\,,$$

si può scrivere

$$df(\boldsymbol{x}_0)\boldsymbol{h} = \nabla f(\boldsymbol{x}_0)\cdot\boldsymbol{h}\,.$$

Possiamo quindi scrivere

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + r(\boldsymbol{x})\,,$$

con

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0}\frac{r(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} = 0\,.$$

Analizziamo con maggiore attenzione il caso $N = 2$. Come di consueto, invece di usare la noitazione (x_1, x_2), gli elementi di $\mathbb{R}^2$ verranno denotati con (x, y). Fissato quindi il punto $\boldsymbol{x}_0 = (x_0, y_0)$, possiamo scrivere

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) + r(x,y)\,,$$

con

$$\lim_{(x,y)\to(x_0,y_0)}\frac{r(x,y)}{\sqrt{(x-x_0)^2+(y-y_0)^2}} = 0\,.$$

Ricordando che il grafico di f è l'insieme

$$G_f = \{(x,y,z)\in\mathbb{R}^3 : z = f(x,y)\}\,,$$

chiameremo "piano tangente" al grafico di f nel punto $(x_0, y_0, f(x_0, y_0))$ l'insieme

$$\left\{(x,y,z)\in\mathbb{R}^3 : z = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)\right\}.$$

8 Funzioni di classe $\mathcal{C}^1$

Il seguente risultato è noto come "teorema del differenziale totale".

Teorema 32 *Se f possiede le derivate parziali in un intorno di $\boldsymbol{x}_0$ ed esse sono continue in $\boldsymbol{x}_0$, allora f è differenziabile in $\boldsymbol{x}_0$.*

<u>Dimostrazione</u>. Supporremo per semplicità di notazioni $N = 2$. Definiamo l'applicazione lineare $\ell : \mathbb{R}^2 \to \mathbb{R}$ che ad ogni vettore $\boldsymbol{h} = (h_1, h_2)$ associa

$$\ell(\boldsymbol{h}) = \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)h_1 + \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)h_2 \,.$$

Vedremo che ℓ è proprio il differenziale di f in $\boldsymbol{x}_0$. Intanto, è lineare, come si vede immediatamente. Inoltre, scrivendo $\boldsymbol{x}_0 = (x_1^0, x_2^0)$ e $\boldsymbol{x} = (x_1, x_2)$, per il Teorema di Lagrange si ha

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{x}_0) &= (f(x_1, x_2) - f(x_1^0, x_2)) + (f(x_1^0, x_2) - f(x_1^0, x_2^0)) \\ &= \frac{\partial f}{\partial x_1}(\xi_1, x_2)(x_1 - x_1^0) + \frac{\partial f}{\partial x_2}(x_1^0, \xi_2)(x_2 - x_2^0)\,, \end{aligned}$$

per un certo $\xi_1 \in [x_1^0, x_1]$ e un certo $\xi_2 \in [x_2^0, x_2]$. Quindi,

$$\begin{aligned} r(\boldsymbol{x}) &= f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - \ell(\boldsymbol{x} - \boldsymbol{x}_0) \\ &= \left[\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(x_1^0, x_2^0)\right](x_1 - x_1^0) + \\ &\quad + \left[\frac{\partial f}{\partial x_2}(x_1^0, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, x_2^0)\right](x_2 - x_2^0)\,, \end{aligned}$$

ed essendo $|x_1 - x_1^0| \le \|\boldsymbol{x} - \boldsymbol{x}_0\|$ e $|x_2 - x_2^0| \le \|\boldsymbol{x} - \boldsymbol{x}_0\|$,

$$\frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \le \left|\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(x_1^0, x_2^0)\right| + \left|\frac{\partial f}{\partial x_2}(x_1^0, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, x_2^0)\right| \,.$$

Facendo tendere $\boldsymbol{x}$ a $\boldsymbol{x}_0$, si ha che $(\xi_1, x_2) \to (x_1^0, x_2^0)$ e $(x_1^0, \xi_2) \to (x_1^0, x_2^0)$ per cui, essendo $\frac{\partial f}{\partial x_1}$ e $\frac{\partial f}{\partial x_2}$ continue in $\boldsymbol{x}_0 = (x_1^0, x_2^0)$, si ha

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0\,,$$

da cui la tesi. ■

Diremo che la funzione f è di classe $\mathcal{C}^1$ su E se f possiede le derivate parziali ed esse sono continue su tutto E. Dal teorema precedente segue che una funzione di classe $\mathcal{C}^1$ è "differenziabile su E", ossia in ogni punto di E.

9 Derivate parziali successive

Supponiamo, per semplicità, $N = 2$. Consideriamo E, un insieme aperto di $\mathbb{R}^2$ e una funzione $f : E \to \mathbb{R}$ che abbia le derivate parziali $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ in tutti i punti di E. Se esse posseggono a loro volta derivate parziali in un punto $\boldsymbol{x}_0$, queste si dicono "derivate parziali seconde" della f in $\boldsymbol{x}_0$ e si denotano con i simboli

$$\frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_1}\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\,, \quad \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_2}\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\,,$$
$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_1}\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\,, \quad \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_2}\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\,.$$

Teorema 33 (di Schwarz) *Se esistono le derivate parziali seconde* $\frac{\partial^2 f}{\partial x_2 \partial x_1}$, $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ *in un intorno di* $\boldsymbol{x}_0$ *ed esse sono continue in* $\boldsymbol{x}_0$, *allora*

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0)\,.$$

Dimostrazione. Sia $\rho > 0$ tale che $B(\boldsymbol{x}_0, \rho) \subseteq E$. Scriviamo $\boldsymbol{x}_0 = (x_1^0, x_2^0)$ e prendiamo un $\boldsymbol{x} = (x_1, x_2) \in B(\boldsymbol{x}_0, \rho)$ tale che $x_1 \neq x_1^0$ e $x_2 \neq x_2^0$. Possiamo allora definire

$$g(x_1, x_2) = \frac{f(x_1, x_2) - f(x_1, x_2^0)}{x_2 - x_2^0}\,, \quad h(x_1, x_2) = \frac{f(x_1, x_2) - f(x_1^0, x_2)}{x_1 - x_1^0}\,.$$

Si verifica che vale l'uguaglianza

$$\frac{g(x_1, x_2) - g(x_1^0, x_2)}{x_1 - x_1^0} = \frac{h(x_1, x_2) - h(x_1, x_2^0)}{x_2 - x_2^0}\,.$$

Per il Teorema di Lagrange, esiste un $\xi_1 \in]x_1^0, x_1[$ tale che

$$\frac{g(x_1, x_2) - g(x_1^0, x_2)}{x_1 - x_1^0} = \frac{\partial g}{\partial x_1}(\xi_1, x_2) = \frac{\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(\xi_1, x_2^0)}{x_2 - x_2^0}\,,$$

ed esiste un $\xi_2 \in]x_2^0, x_2[$ tale che

$$\frac{h(x_1, x_2) - h(x_1, x_2^0)}{x_2 - x_2^0} = \frac{\partial h}{\partial x_2}(x_1, \xi_2) = \frac{\frac{\partial f}{\partial x_2}(x_1, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, \xi_2)}{x_1 - x_1^0}\,.$$

Di nuovo per il Teorema di Lagrange, esiste un $\eta_2 \in]x_2^0, x_2[$ tale che

$$\frac{\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(\xi_1, x_2^0)}{x_2 - x_2^0} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \eta_2)\,,$$

ed esiste un $\eta_1 \in]x_1^0, x_1[$ tale che

$$\frac{\frac{\partial f}{\partial x_2}(x_1, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, \xi_2)}{x_1 - x_1^0} = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\eta_1, \xi_2)\,.$$

Quindi,

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \eta_2) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\eta_1, \xi_2)\,.$$

Facendo tendere $\boldsymbol{x} = (x_1, x_2)$ a $\boldsymbol{x}_0 = (x_1^0, x_2^0)$, si ha che sia (ξ_1, η_2) che (η_1, ξ_2) tendono a $\boldsymbol{x}_0$, e per la continuità delle derivate seconde miste si ha la tesi. ■

Diremo che la funzione f è di classe $\mathcal{C}^2$ su E se f possiede tutte le derivate parziali seconde ed esse sono continue su tutto E. Dal teorema precedente segue che se una funzione di classe $\mathcal{C}^2$, le derivate parziali "miste" sono uguali.

È utile definire la "matrice hessiana" di f nel punto $\boldsymbol{x}_0$:

$$Hf(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) \end{pmatrix};$$

se f è di classe $\mathcal{C}^2$, si tratta di una matrice simmetrica.

Quanto sopra si può estendere senza difficoltà alle funzioni di N variabili, con N qualunque. Se f è di classe $\mathcal{C}^2$, la matrice hessiana risulta allora una matrice simmetrica del tipo $N \times N$:

$$Hf(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) & \dots & \frac{\partial^2 f}{\partial x_N \partial x_1}(\boldsymbol{x}_0) \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) & \dots & \frac{\partial^2 f}{\partial x_N \partial x_2}(\boldsymbol{x}_0) \\ \\ \vdots & \vdots & \dots & \vdots \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_N}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_N}(\boldsymbol{x}_0) & \dots & \frac{\partial^2 f}{\partial x_N^2}(\boldsymbol{x}_0) \end{pmatrix}.$$

Procedendo per induzione, si possono definire le derivate parziali n-esime di una funzione. Si dice che la funzione f è di classe $\mathcal{C}^n$ su E se f possiede tutte le derivate parziali n-esime ed esse sono continue su tutto E.

10 La formula di Taylor

Supponiamo ora che $f : E \to \mathbb{R}$ sia una funzione di classe $\mathcal{C}^{n+1}$, per un certo $n \geq 1$.

Consideriamo come sopra, per semplicità, il caso $N = 2$. Introduciamo le seguenti notazioni:

$$D_{x_1} = \frac{\partial}{\partial x_1}\,, \quad D_{x_2} = \frac{\partial}{\partial x_2}\,,$$

$$D_{x_1}^2 = \frac{\partial^2}{\partial x_1^2}\,, \quad D_{x_1} D_{x_2} = \frac{\partial^2}{\partial x_1 \partial x_2}\,, \quad D_{x_2}^2 = \frac{\partial^2}{\partial x_2^2}\,,$$

e così via, per le derivate parziali successive. Si noti che, per un vettore $\boldsymbol{h} = (h_1, h_2) \in \mathbb{R}^2$, si ha

$$df(\boldsymbol{x}_0)\boldsymbol{h} = h_1 D_{x_1} f(\boldsymbol{x}_0) + h_2 D_{x_2} f(\boldsymbol{x}_0)\,,$$

che risulterà conveniente scrivere

$$df(\boldsymbol{x}_0)\boldsymbol{h} = [h_1 D_{x_1} + h_2 D_{x_2}] f(\boldsymbol{x}_0)\,.$$

In questo modo, possiamo pensare che f viene trasformata dall'operatore $[h_1 D_{x_1} + h_2 D_{x_2}]$ nella nuova funzione $[h_1 D_{x_1} + h_2 D_{x_2}]f = h_1 D_{x_1} f + h_2 D_{x_2} f$.

Dati due punti $\boldsymbol{x}_0$ e $\boldsymbol{x}$ in $\mathbb{R}^N$, si definisce il "segmento" che li congiunge:

$$[\boldsymbol{x}_0, \boldsymbol{x}] = \{\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0) : t \in [0,1]\}\,;$$

analogamente, scriveremo

$$]\boldsymbol{x}_0, \boldsymbol{x}[= \{\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0) : t \in]0,1[\,\}\,.$$

Supponiamo ora che $[\boldsymbol{x}_0, \boldsymbol{x}]$ sia un segmento contenuto in E e consideriamo la funzione $\phi : [0,1] \to \mathbb{R}$ definita da

$$\phi(t) = f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\,.$$

Dimostriamo che ϕ è derivabile $n+1$ volte su $[0,1]$. Per $t \in [0,1]$, essendo f differenziabile in $\boldsymbol{u}_0 = \boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0)$, si ha

$$f(\boldsymbol{u}) = f(\boldsymbol{u}_0) + df(\boldsymbol{u}_0)(\boldsymbol{u} - \boldsymbol{u}_0) + r(\boldsymbol{u})\,,$$

con

$$\lim_{\boldsymbol{u} \to \boldsymbol{u}_0} \frac{r(\boldsymbol{u})}{\|\boldsymbol{u} - \boldsymbol{u}_0\|} = 0\,.$$

Quindi,

$$\begin{aligned}
\lim_{s \to t} \frac{\phi(s) - \phi(t)}{s - t} &= \lim_{s \to t} \frac{f(\boldsymbol{x}_0 + s(\boldsymbol{x} - \boldsymbol{x}_0)) - f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))}{s - t} \\
&= \lim_{s \to t} \frac{df(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))((s-t)(\boldsymbol{x} - \boldsymbol{x}_0)) + r(\boldsymbol{x}_0 + s(\boldsymbol{x} - \boldsymbol{x}_0))}{s - t} \\
&= df(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))(\boldsymbol{x} - \boldsymbol{x}_0) + \lim_{s \to t} \frac{r(\boldsymbol{x}_0 + s(\boldsymbol{x} - \boldsymbol{x}_0))}{s - t}\,,
\end{aligned}$$

ed essendo

$$\lim_{s \to t} \left| \frac{r(\boldsymbol{x}_0 + s(\boldsymbol{x} - \boldsymbol{x}_0))}{s - t} \right| = \lim_{\boldsymbol{u} \to \boldsymbol{u}_0} \frac{|r(\boldsymbol{u})|}{\|\boldsymbol{u} - \boldsymbol{u}_0\|} \|\boldsymbol{x} - \boldsymbol{x}_0\| = 0\,,$$

si ha

$$\phi'(t) = \lim_{s \to t} \frac{\phi(s) - \phi(t)}{s - t} = df(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))(\boldsymbol{x} - \boldsymbol{x}_0)\,.$$

Con le nuove notazioni, ponendo $\boldsymbol{x}-\boldsymbol{x}_0=\boldsymbol{h}=(h_1,h_2)$, abbiamo

$$\phi'(t)=[h_1D_{x_1}+h_2D_{x_2}]f(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))=g(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\,,$$

dove g è la nuova funzione $[h_1D_{x_1}+h_2D_{x_2}]f$. Possiamo allora iterare il procedimento, e calcolare la derivata seconda di ϕ:

$$\begin{aligned}\phi''(t)&=[h_1D_{x_1}+h_2D_{x_2}]g(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\\&=[h_1D_{x_1}+h_2D_{x_2}][h_1D_{x_1}+h_2D_{x_2}]f(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\,.\end{aligned}$$

Per brevità, scriveremo

$$\phi''(t)=[h_1D_{x_1}+h_2D_{x_2}]^2f(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\,.$$

Notiamo che, usando la linearità delle derivate parziali e l'uguaglianza delle derivate miste (Teorema di Schwarz), si ha

$$\begin{aligned}[h_1D_{x_1}+h_2D_{x_2}]^2f&=h_1^2D_{x_1}^2f+2h_1h_2D_{x_1}D_{x_2}f+h_2^2D_{x_2}^2f\\&=[h_1^2D_{x_1}^2+2h_1h_2D_{x_1}D_{x_2}+h_2^2D_{x_2}^2]f\,.\end{aligned}$$

Osserviamo che l'espressione

$$[h_1D_{x_1}+h_2D_{x_2}]^2=[h_1^2D_{x_1}^2+2h_1h_2D_{x_1}D_{x_2}+h_2^2D_{x_2}^2]$$

si ottiene formalmente come il quadrato di un binomio. Procedendo in questo modo, si può dimostrare per induzione che, per $k=1,2,\dots,n+1$, la formula della derivata k-esima di ϕ è

$$\phi^{(k)}(t)=[h_1D_{x_1}+h_2D_{x_2}]^kf(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\,,$$

e che, usando formalmente la formula del binomio di Newton

$$(a_1+a_2)^k=\sum_{j=0}^{k}\binom{k}{j}a_1^{k-j}a_2^j,$$

si ha

$$[h_1D_{x_1}+h_2D_{x_2}]^k=\left[\sum_{j=0}^{k}\binom{k}{j}h_1^{k-j}h_2^jD_{x_1}^{k-j}D_{x_2}^j\right]$$

(in questa formula, i simboli $D_{x_1}^0$ e $D_{x_2}^0$ vanno interpretati come l'operatore identità).

Per poter scrivere agevolmente la formula di Taylor, introduciamo la notazione

$$d^kf(\boldsymbol{x}_0)\boldsymbol{h}^k=[h_1D_{x_1}+h_2D_{x_2}]^kf(\boldsymbol{x}_0)\,.$$

Teorema 34 *Sia $f : E \to \mathbb{R}$ di classe $\mathcal{C}^{n+1}$ e $[\boldsymbol{x}_0, \boldsymbol{x}]$ un segmento contenuto in E. Allora esiste un $\boldsymbol{\xi} \in]\boldsymbol{x}_0, \boldsymbol{x}[$ tale che*

$$f(\boldsymbol{x}) = p_n(\boldsymbol{x}) + r_n(\boldsymbol{x}) ,$$

dove

$$p_n(\boldsymbol{x}) = f(\boldsymbol{x}_0)+df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)+\frac{1}{2!}d^2f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^2+\ldots+\frac{1}{n!}d^nf(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^n$$

è il "polinomio di Taylor di grado n associato alla funzione f nel punto $\boldsymbol{x}_0$" e

$$r_n(\boldsymbol{x}) = \frac{1}{(n+1)!}d^{n+1}f(\boldsymbol{\xi})(\boldsymbol{x}-\boldsymbol{x}_0)^{n+1}$$

è il "resto di Lagrange".

Dimostrazione. Per la formula di Taylor applicata alla funzione ϕ, si ha

$$\phi(t) = \phi(0) + \phi'(0)t + \frac{1}{2!}\phi''(0)t^2 + \ldots + \frac{1}{n!}\phi^{(n)}(0)t^n + \frac{1}{(n+1)!}\phi^{(n+1)}(\xi)t^{n+1} ,$$

per un certo $\xi \in]0, t[$. La formula cercata si ottiene prendendo $t = 1$ e sostituendo i valori delle derivate di ϕ trovati sopra. ■

Il polinomio di Taylor si può anche scrivere nella forma compatta

$$p_n(\boldsymbol{x}) = \sum_{k=0}^{n} \frac{1}{k!}d^kf(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^k ,$$

con la convenzione che $d^0f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^0$, il primo addendo della somma, sia $f(\boldsymbol{x}_0)$. Si ha quindi

$$\begin{aligned} p_n(\boldsymbol{x}) &= \sum_{k=0}^{n} \frac{1}{k!}[(x_1 - x_1^0)D_{x_1} + (x_2 - x_2^0)D_{x_2}]^k f(\boldsymbol{x}_0) \\ &= \sum_{k=0}^{n} \frac{1}{k!} \left(\sum_{j=0}^{k} \binom{k}{j} \frac{\partial^k f}{\partial^{k-j}x_1 \partial^j x_2}(\boldsymbol{x}_0)\ (x_1 - x_1^0)^{k-j}(x_2 - x_2^0)^j \right) . \end{aligned}$$

Può essere utile la seguente espressione per il polinomio di secondo grado:

$$p_2(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x}-\boldsymbol{x}_0) + \tfrac{1}{2}\Big(Hf(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\Big) \cdot (\boldsymbol{x}-\boldsymbol{x}_0) .$$

Il teorema sopra dimostrato resta valido per qualsiasi dimensione N, pur di interpretare correttamente le notazioni: ad esempio, per un vettore $\boldsymbol{h} = (h_1, h_2, \ldots, h_N)$, si dovrà leggere

$$d^kf(\boldsymbol{x}_0)\boldsymbol{h}^k = [h_1D_{x_1} + h_2D_{x_2} + \ldots + h_ND_{x_N}]^k f(\boldsymbol{x}_0) .$$

In questo caso, volendo esplicitare il polinomio di Taylor, sarà utile utilizzare la formula di Leibniz

$$(a_1 + a_2 + \ldots + a_N)^k = \sum_{m_1+m_2+\ldots+m_N=k} \frac{k!}{m_1!\,m_2!\cdots m_N!}\, a_1^{m_1}a_2^{m_2}\cdots a_N^{m_N} .$$

11 La ricerca di massimi e minimi

Come sopra, consideriamo un insieme aperto $E \subseteq \mathbb{R}^N$ e una funzione $f : E \to \mathbb{R}$. Diremo che $\boldsymbol{x}_0 \in E$ è un "punto di massimo locale" per la funzione f se esiste un intorno U di x_0 contenuto in E per cui x_0 è punto di massimo della restrizione di f a U. Equivalentemente, se

$$\exists \delta > 0 : \ \forall \boldsymbol{x} \in E \quad d(\boldsymbol{x}, \boldsymbol{x}_0) < \delta \ \Rightarrow \ f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0) \,.$$

Analogamente per "punto di minimo locale".

Teorema 35 (di Fermat) *Se $\boldsymbol{x}_0$ è un punto di massimo o di minimo locale e f è differenziabile in $\boldsymbol{x}_0$, allora $\nabla f(\boldsymbol{x}_0) = 0$.*

<u>Dimostrazione</u>. Se $\boldsymbol{x}_0$ è punto di massimo locale, per ogni direzione $\boldsymbol{v} \in \mathbb{R}^N$ avremo che

$$\frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t} \quad \begin{cases} \geq 0 & \text{se } t < 0 \,, \\ \leq 0 & \text{se } t > 0 \,. \end{cases}$$

Siccome f è differenziabile in $\boldsymbol{x}_0$, ne deduciamo che

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \lim_{t \to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t} = 0 \,.$$

In particolare, sono nulle tutte le derivate parziali, per cui $\nabla f(\boldsymbol{x}_0) = 0$. Nel caso in cui $\boldsymbol{x}_0$ sia un punto di minimo locale, si procede in modo analogo. ■

Un punto il cui il gradiente si annulli è detto "punto stazionario". Naturalmente un tale punto potrebbe non essere nè di massimo nè di minimo.

Mostreremo ora come la formula di Taylor possa essere usata per stabilire un criterio affinchè un punto stazionario sia di massimo, o di minimo. Iniziamo con una definizione. Diremo che una matrice $\mathbb{A}$ simmetrica $N \times N$ è *definita positiva* se

$$[\mathbb{A}\boldsymbol{h}] \cdot \boldsymbol{h} > 0 \,, \qquad \text{per ogni } \boldsymbol{h} \in \mathbb{R}^N \setminus \{\boldsymbol{0}\} \,.$$

Diremo che $\mathbb{A}$ è *definita negativa* se vale la disuguaglianza opposta, ossia se $-\mathbb{A}$ è definita positiva.

Teorema 36 *Se $\boldsymbol{x}_0$ è un punto stazionario e f è di classe $\mathcal{C}^2$, con matrice hessiana $Hf(\boldsymbol{x}_0)$ definita positiva, allora $\boldsymbol{x}_0$ è un punto di minimo locale. Se invece $Hf(\boldsymbol{x}_0)$ è definita negativa, allora $\boldsymbol{x}_0$ è un punto di massimo locale.*

<u>Dimostrazione</u>. Per la formula di Taylor, per $\boldsymbol{x} \neq \boldsymbol{x}_0$ in un intorno di $\boldsymbol{x}_0$ esiste un $\boldsymbol{\xi} \in]\boldsymbol{x}_0, \boldsymbol{x}[$ per cui

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \tfrac{1}{2}\Big(Hf(\boldsymbol{\xi})(\boldsymbol{x} - \boldsymbol{x}_0)\Big) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) \,.$$

Se $\mathbb{A} = Hf(\boldsymbol{x}_0)$ è definita positiva, esiste un $c > 0$ tale che, per ogni $\boldsymbol{v} \in \mathbb{R}^N$ con $\|\boldsymbol{v}\| = 1$,

$$[\mathbb{A}\boldsymbol{v}] \cdot \boldsymbol{v} \geq c \,.$$

(Abbiamo qui usato il Teorema di Weierstrass, e il fatto che la sfera $\{v \in \mathbb{R}^N : \|\boldsymbol{v}\| = 1\}$ è un insieme compatto.) Quindi

$$\Big(Hf(\boldsymbol{x}_0)\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}\Big)\cdot\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \geq c\,.$$

Per la continuità delle derivate seconde, se $\boldsymbol{x}$ è sufficientemente vicino a $\boldsymbol{x}_0$,

$$\Big(Hf(\boldsymbol{\xi})\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}\Big)\cdot\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \geq \tfrac{1}{2}c > 0\,.$$

(Lo si vede per assurdo, usando di nuovo laa compattezza della sfera.) Essendo $\nabla f(\boldsymbol{x}_0) = \mathbf{0}$, per tali $\boldsymbol{x}$ abbiamo che

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + \tfrac{1}{2}\Big(Hf(\boldsymbol{\xi})(\boldsymbol{x}-\boldsymbol{x}_0)\Big)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) \\ &\geq f(\boldsymbol{x}_0) + \tfrac{1}{2}c\|\boldsymbol{x}-\boldsymbol{x}_0\|^2 > f(\boldsymbol{x}_0)\,, \end{aligned}$$

per cui $\boldsymbol{x}_0$ è un punto di minimo locale.

La dimostrazione della seconda affermazione è analoga. ■

Enunciamo ora (senza dimostrazione) due criteri utili a stabilire quando una matrice $\mathbb{A}$ simmetrica $N \times N$ è definita positiva o negativa. Ricordiamo che gli autovalori di una matrice simmetrica sono tutti reali.

Primo criterio. *La martrice $\mathbb{A}$ è definita positiva se tutti i suoi autovalori sono positivi. Essa è definita negativa se tutti i suoi autovalori sono negativi.*

Secondo criterio. *La martrice $\mathbb{A} = (a_{ij})_{ij}$ è definita positiva se*

$$\begin{aligned} & a_{11} > 0\,, \\ & \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0\,, \\ & \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} > 0\,,\ \dots \\ & \det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{pmatrix} > 0\,. \end{aligned}$$

Essa è definita negativa se i determinanti scritti sopra hanno segno alternato: quelli delle sottomatrici con un numero dispari di righe e di colonne sono negativi, mentre quelli delle sottomatrici con un numero pari di righe e di colonne sono positivi.

12 Il teorema della funzione implicita - primo enunciato

Il seguente risultato porta il nome di Ulisse Dini.

Teorema 37 *Siano $\Omega \subseteq \mathbb{R} \times \mathbb{R}$ un aperto, $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$ e (x_0, y_0) un punto di Ω per cui si abbia:*

$$g(x_0, y_0) = 0 \qquad \frac{\partial g}{\partial y}(x_0, y_0) \neq 0\,.$$

Allora esistono un intorno aperto U di x_0, un intorno aperto V di y_0 e una funzione $\eta : U \to V$ di classe $\mathcal{C}^1$ tali che $U \times V \subseteq \Omega$ e, presi $x \in U$ e $y \in V$, si ha:

$$g(x, y) = 0 \quad \Leftrightarrow \quad y = \eta(x)\,.$$

Inoltre, la funzione η è di classe $\mathcal{C}^1$ e vale la formula

$$\eta'(x) = -\frac{\frac{\partial g}{\partial x}(x, \eta(x))}{\frac{\partial g}{\partial y}(x, \eta(x))}\,.$$

La funzione η risulta definita "implicitamente" dall'equazione $g(x, y) = 0$; il suo grafico è l'insieme

$$Gr(\eta) = \{(x, y) \in U \times V : g(x, y) = 0\}\,.$$

Dimostrazione. Supponiamo ad esempio $\frac{\partial g}{\partial y}(x_0, y_0) > 0$. Per la proprietà di permanenza del segno, esiste un $\delta > 0$ tale che, se $|x - x_0| \leq \delta$ e $|y - y_0| \leq \delta$, allora $\frac{\partial g}{\partial y}(x, y) > 0$. Quindi, per ogni $x \in [x_0 - \delta, x_0 + \delta]$, la funzione $g(x, \cdot)$ è strettamente crescente su $[y_0 - \delta, y_0 + \delta]$. Essendo $g(x_0, y_0) = 0$, avremo che

$$g(x_0, y_0 - \delta) < 0 < g(x_0, y_0 + \delta)\,.$$

Per la permanenza del segno, esiste un $\delta' > 0$ tale che, se $x \in [x_0 - \delta', x_0 + \delta']$, allora

$$g(x, y_0 - \delta) < 0 < g(x, y_0 + \delta)\,.$$

Definiamo $U =]x_0 - \delta', x_0 + \delta'[$ e $V =]y_0 - \delta, y_0 + \delta[$. Quindi, per ogni $x \in U$, siccome $g(x, \cdot)$ è strettamente crescente, esiste uno ed un solo $y \in]y_0 - \delta, y_0 + \delta[$ per cui $g(x, y) = 0$; chiamo $\eta(x)$ tale y. Resta così definita una funzione $\eta : U \to V$ tale che, presi $x \in U$ e $y \in V$, si ha:

$$g(x, y) = 0 \quad \Leftrightarrow \quad y = \eta(x)\,.$$

Per vedere che η è continua, fissiamo ora un $\bar{x} \in U$ e dimostriamo la continuità in $\bar{x}$. Preso un $x \in U$ e considerata la funzione $\gamma : [0, 1] \to \mathbb{R} \times \mathbb{R}$ definita da

$$\gamma(t) = (\bar{x} + t(x - \bar{x}), \eta(\bar{x}) + t(\eta(x) - \eta(\bar{x}))),$$

applicando il Teorema di Lagrange alla funzione $g \circ \gamma$ si ha che esiste un $\xi \in]0,1[$ per cui

$$g(x,\eta(x)) - g(\bar{x},\eta(\bar{x})) = \frac{\partial g}{\partial x}(\gamma(\xi))(x-\bar{x}) + \frac{\partial g}{\partial y}(\gamma(\xi))(\eta(x)-\eta(\bar{x}))\,.$$

Essendo $g(x,\eta(x)) = g(\bar{x},\eta(\bar{x})) = 0$, si ha che

$$|\eta(x)-\eta(\bar{x})| = \left|\frac{\frac{\partial g}{\partial x}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))}\right| |x-\bar{x}|\,.$$

Siccome le derivate parziali di g sono continue e $\frac{\partial g}{\partial y}$ è non nulla sul compatto $\overline{U} \times \overline{V}$, si ha che $|\frac{\partial g}{\partial x}(\gamma(\xi))(\frac{\partial g}{\partial y}(\gamma(\xi)))^{-1}|$ è limitato superiormente e ne segue la continuità di η in $\bar{x}$. Resta da vedere la derivabilità: procedendo come sopra si ha che

$$\frac{\eta(x)-\eta(\bar{x})}{x-\bar{x}} = -\frac{\frac{\partial g}{\partial x}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))}\,,$$

con $\gamma(\xi)$ appartenente al segmento che congiunge $(\bar{x},\eta(\bar{x}))$ con $(x,\eta(x))$. Se x tende a $\bar{x}$, si ha che $\gamma(\xi)$ tende a $(\bar{x},\eta(\bar{x}))$ e quindi

$$\eta'(\bar{x}) = \lim_{x\to\bar{x}} \frac{\eta(x)-\eta(\bar{x})}{x-\bar{x}} = -\frac{\frac{\partial g}{\partial x}(\bar{x},\eta(\bar{x}))}{\frac{\partial g}{\partial y}(\bar{x},\eta(\bar{x}))}\,.$$

Ne segue che η è di classe $\mathcal{C}^1$. ■

Vale naturalmente anche il seguente enunciato simmetrico rispetto al precedente.

Teorema 38 *Siano $\Omega \subseteq \mathbb{R}\times\mathbb{R}$ un aperto, $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$ e (x_0,y_0) un punto di Ω per cui si abbia:*

$$g(x_0,y_0) = 0 \qquad \frac{\partial g}{\partial x}(x_0,y_0) \neq 0\,.$$

Allora esistono un intorno aperto U di x_0, un intorno aperto V di y_0 e una funzione $\eta : V \to U$ di classe $\mathcal{C}^1$ tali che $U \times V \subseteq \Omega$ e, presi $x \in U$ e $y \in V$, si ha:

$$g(x,y) = 0 \quad \Leftrightarrow \quad x = \eta(y)\,.$$

Inoltre, la funzione η è di classe $\mathcal{C}^1$ e vale la formula

$$\eta'(y) = -\frac{\frac{\partial g}{\partial y}(y,\eta(y))}{\frac{\partial g}{\partial x}(y,\eta(y))}\,.$$

13 Il differenziale di una funzione a valori vettoriali

Sia E un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di E e $f : E \to \mathbb{R}^M$ una funzione.

Definizione 4 *Diremo che la funzione f è "differenziabile" in $\boldsymbol{x}_0$ se esiste una applicazione lineare $\ell : \mathbb{R}^N \to \mathbb{R}^M$ per cui si possa scrivere*

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x}) \,,$$

dove r è una funzione tale che

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \boldsymbol{0} \,.$$

Se f è differenziabile in $\boldsymbol{x}_0$, l'applicazione lineare ℓ si chiama "differenziale" di f in $\boldsymbol{x}_0$ e si indica con il simbolo

$$df(\boldsymbol{x}_0) \,.$$

Siano $f_1, f_2, \ldots, f_M$ le componenti di f, per cui

$$f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_M(\boldsymbol{x})) \,.$$

Teorema 39 *La funzione f è differenziabile in $\boldsymbol{x}_0$ se e solo se lo sono tutte le sue componenti. In tal caso, per ogni vettore $\boldsymbol{h} \in \mathbb{R}^N$ si ha*

$$df(\boldsymbol{x}_0)\boldsymbol{h} = (df_1(\boldsymbol{x}_0)\boldsymbol{h}, df_2(\boldsymbol{x}_0)\boldsymbol{h}, \ldots, df_M(\boldsymbol{x}_0)\boldsymbol{h}) \,.$$

<u>Dimostrazione</u>. Considerando le componenti nell'equazione

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x}) \,,$$

possiamo scrivere

$$f_j(\boldsymbol{x}) = f_j(\boldsymbol{x}_0) + \ell_j(\boldsymbol{x} - \boldsymbol{x}_0) + r_j(\boldsymbol{x}) \,,$$

con $j = 1, 2, \ldots, M$, e sappiamo che

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \boldsymbol{0} \quad \Leftrightarrow \quad \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r_j(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0 \ \text{ per ogni } j = 1, 2, \ldots, M \,,$$

da cui la tesi. ■

Il teorema precedente permette di ricondurre lo studio del differenziale di una funzione a valori vettoriali a quello delle sue componenti, che sono funzioni a valori scalari.

È utile considerare la matrice associata all'applicazione lineare $\ell = df(\boldsymbol{x}_0)$, data da

$$\begin{pmatrix} \ell_1(\boldsymbol{e}_1) & \ell_1(\boldsymbol{e}_2) & \dots & \ell_1(\boldsymbol{e}_N) \\ \ell_2(\boldsymbol{e}_1) & \ell_2(\boldsymbol{e}_2) & \dots & \ell_2(\boldsymbol{e}_N) \\ \vdots & \vdots & & \vdots \\ \ell_M(\boldsymbol{e}_1) & \ell_M(\boldsymbol{e}_2) & \dots & \ell_M(\boldsymbol{e}_N) \end{pmatrix},$$

dove $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots \boldsymbol{e}_N$ sono i vettori della base canonica di $\mathbb{R}^N$. Tale matrice si chiama "matrice jacobiana" associata alla funzione f nel punto $\boldsymbol{x}_0$ e si denota con $Jf(\boldsymbol{x}_0)$. Ricordando che

$$\frac{\partial f_j}{\partial x_k}(\boldsymbol{x}_0) = df_j(\boldsymbol{x}_0)\boldsymbol{e}_k ,$$

con $j = 1, 2, \dots, M$ e $k = 1, 2, \dots, N$, si ottiene la matrice

$$Jf(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_1}{\partial x_N}(\boldsymbol{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_2}{\partial x_N}(\boldsymbol{x}_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_M}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_M}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_M}{\partial x_N}(\boldsymbol{x}_0) \end{pmatrix}.$$

Studiamo ora la differenziabilità di una funzione composta.

Teorema 40 *Se $f : E \to \mathbb{R}^M$ è differenziabile in $\boldsymbol{x}_0$, E' è un aperto di $\mathbb{R}^M$ contenente $f(E)$ e $g : E' \to \mathbb{R}^L$ è differenziabile in $f(\boldsymbol{x}_0)$, allora $g \circ f$ è differenziabile in $\boldsymbol{x}_0$, e si ha*

$$d(g \circ f)(\boldsymbol{x}_0) = dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0) .$$

Dimostrazione. Ponendo $\boldsymbol{y}_0 = f(\boldsymbol{x}_0)$, si ha

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + df(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + r_1(\boldsymbol{x}) , \quad g(\boldsymbol{y}) = g(\boldsymbol{y}_0) + dg(\boldsymbol{y}_0)(\boldsymbol{y} - \boldsymbol{y}_0) + r_2(\boldsymbol{y}) ,$$

con

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r_1(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \boldsymbol{0} , \quad \lim_{\boldsymbol{y} \to \boldsymbol{y}_0} \frac{r_2(\boldsymbol{y})}{\|\boldsymbol{y} - \boldsymbol{y}_0\|} = \boldsymbol{0} .$$

Introduciamo la funzione $R_2 : E' \to \mathbb{R}^L$ così definita:

$$R_2(\boldsymbol{y}) = \begin{cases} \dfrac{r_2(\boldsymbol{y})}{\|\boldsymbol{y} - \boldsymbol{y}_0\|} & \text{se } \boldsymbol{y} \neq \boldsymbol{y}_0 , \\ \boldsymbol{0} & \text{se } \boldsymbol{y} = \boldsymbol{y}_0 . \end{cases}$$

Si noti che R_2 è continua in $\boldsymbol{y}_0$. Allora

$$\begin{aligned} g(f(\boldsymbol{x})) &= g(f(\boldsymbol{x}_0)) + dg(f(\boldsymbol{x}_0))[f(\boldsymbol{x}) - f(\boldsymbol{x}_0)] + r_2(f(\boldsymbol{x})) \\ &= g(f(\boldsymbol{x}_0)) + dg(f(\boldsymbol{x}_0))[df(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + r_1(\boldsymbol{x})] + r_2(f(\boldsymbol{x})) \\ &= g(f(\boldsymbol{x}_0)) + [dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0)](\boldsymbol{x} - \boldsymbol{x}_0) + r_3(\boldsymbol{x}) , \end{aligned}$$

dove

$$\begin{aligned}r_3(\boldsymbol{x}) &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + r_2(f(\boldsymbol{x}))\\ &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| R_2(f(\boldsymbol{x}))\\ &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + \|df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0) + r_1(\boldsymbol{x})\| R_2(f(\boldsymbol{x}))\,.\end{aligned}$$

Quindi,

$$\begin{aligned}\frac{\|r_3(\boldsymbol{x})\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \le{}& \left\| dg(f(\boldsymbol{x}_0))\left(\frac{r_1(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}\right)\right\| + \\ &+ \left(\left\| df(\boldsymbol{x}_0)\left(\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}\right)\right\| + \frac{\|r_1(\boldsymbol{x})\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}\right)\|R_2(f(\boldsymbol{x}))\|\,.\end{aligned}$$

Se $\boldsymbol{x} \to \boldsymbol{x}_0$, il primo addendo tende a 0, poiché $dg(f(\boldsymbol{x}_0))$ è continua; f è continua in $\boldsymbol{x}_0$ e R_2 è continua in $\boldsymbol{y}_0 = f(\boldsymbol{x}_0)$ con $R_2(\boldsymbol{y}_0)) = \boldsymbol{0}$, per cui $\|R_2(f(\boldsymbol{x}))\|$ tende a 0; $df(\boldsymbol{x}_0)$, essendo continua, è limitata sull'insieme compatto $\bar{B}(\boldsymbol{0}, 1)$. Quindi, si ha che

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{\|r_3(\boldsymbol{x})\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} = 0\,.$$

Ne segue che $g \circ f$ è differenziabile in x_0 con differenziale $dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0)$. ■

Come noto, la matrice associata alla composizione di due applicazioni lineari è il prodotto delle due matrici corrispondenti. Dal teorema precedente abbiamo quindi la seguente formula per le matrici jacobiane:

$$J(g\circ f)(\boldsymbol{x}_0) = Jg(f(\boldsymbol{x}_0)) \cdot Jf(\boldsymbol{x}_0)\,,$$

ossia

$$\begin{aligned}&\begin{pmatrix} \frac{\partial (g\circ f)_1}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial (g\circ f)_1}{\partial x_N}(\boldsymbol{x}_0)\\ \vdots & \cdots & \vdots \\ \frac{\partial (g\circ f)_L}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial (g\circ f)_L}{\partial x_N}(\boldsymbol{x}_0)\end{pmatrix} = \\ &\qquad = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(\boldsymbol{x}_0)) & \cdots & \frac{\partial g_1}{\partial y_M}(f(\boldsymbol{x}_0))\\ \vdots & \cdots & \vdots \\ \frac{\partial g_L}{\partial y_1}(f(\boldsymbol{x}_0)) & \cdots & \frac{\partial g_L}{\partial y_M}(f(\boldsymbol{x}_0))\end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_1}{\partial x_N}(\boldsymbol{x}_0)\\ \vdots & \cdots & \vdots \\ \frac{\partial f_M}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_M}{\partial x_N}(\boldsymbol{x}_0)\end{pmatrix}.\end{aligned}$$

Ne segue la formula per le derivate parziali:

$$\begin{aligned}&\frac{\partial (g\circ f)_i}{\partial x_k}(\boldsymbol{x}_0) = \\ &= \frac{\partial g_i}{\partial y_1}(f(\boldsymbol{x}_0))\frac{\partial f_1}{\partial x_k}(\boldsymbol{x}_0) + \frac{\partial g_i}{\partial y_2}(f(\boldsymbol{x}_0))\frac{\partial f_2}{\partial x_k}(\boldsymbol{x}_0) + \ldots + \frac{\partial g_i}{\partial y_M}(f(\boldsymbol{x}_0))\frac{\partial f_M}{\partial x_k}(\boldsymbol{x}_0)\\ &= \sum_{j=1}^{M} \frac{\partial g_i}{\partial y_j}(f(\boldsymbol{x}_0))\frac{\partial f_j}{\partial x_k}(\boldsymbol{x}_0)\,,\end{aligned}$$

dove $i = 1, 2, \ldots, L$ e $k = 1, 2, \ldots, N$.

14 Il teorema della funzione implicita - caso generale

Vediamo come si generalizza il teorema della funzione implicita. Considereremo un insieme aperto Ω di $\mathbb{R}^M \times \mathbb{R}^N$ e una funzione $g : \Omega \to \mathbb{R}^N$, di classe $\mathcal{C}^1$. Quindi, g ha N componenti

$$g(\boldsymbol{x}, \boldsymbol{y}) = (g_1(\boldsymbol{x}, \boldsymbol{y}), \dots, g_N(\boldsymbol{x}, \boldsymbol{y}))\,.$$

Qui $\boldsymbol{x} = (x_1, \dots, x_M) \in \mathbb{R}^M$ e $\boldsymbol{y} = (y_1, \dots, y_N) \in \mathbb{R}^N$. Useremo la seguente notazione per le matrici jacobiane:

$$\frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x},\boldsymbol{y}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\boldsymbol{x},\boldsymbol{y}) \cdots \frac{\partial g_1}{\partial x_M}(\boldsymbol{x},\boldsymbol{y}) \\ \vdots \quad \dots \quad \vdots \\ \frac{\partial g_N}{\partial x_1}(\boldsymbol{x},\boldsymbol{y}) \cdots \frac{\partial g_N}{\partial x_M}(\boldsymbol{x},\boldsymbol{y}) \end{pmatrix}, \quad \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x},\boldsymbol{y}) = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\boldsymbol{x},\boldsymbol{y}) \cdots \frac{\partial g_1}{\partial y_N}(\boldsymbol{x},\boldsymbol{y}) \\ \vdots \quad \dots \quad \vdots \\ \frac{\partial g_N}{\partial y_1}(\boldsymbol{x},\boldsymbol{y}) \cdots \frac{\partial g_N}{\partial y_N}(\boldsymbol{x},\boldsymbol{y}) \end{pmatrix}.$$

Possiamo ora enunciare il Teorema di Dini in questo caso più generale.

Teorema 41 *Siano $\Omega \subseteq \mathbb{R}^M \times \mathbb{R}^N$ un aperto, $g : \Omega \to \mathbb{R}^N$ una funzione di classe $\mathcal{C}^1$ e $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ un punto di Ω per cui si abbia:*

$$g(\boldsymbol{x}_0, \boldsymbol{y}_0) = \boldsymbol{0}\,, \qquad \det \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0\,.$$

Allora esistono un intorno aperto U di $\boldsymbol{x}_0$, un intorno aperto V di $\boldsymbol{y}_0$ e una funzione $\eta : U \to V$ di classe $\mathcal{C}^1$ tali che $U \times V \subseteq \Omega$ e, presi $\boldsymbol{x} \in U$ e $\boldsymbol{y} \in V$, si ha:

$$g(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0} \quad \Leftrightarrow \quad \boldsymbol{y} = \eta(\boldsymbol{x})\,.$$

Inoltre, la funzione η è di classe $\mathcal{C}^1$ e vale la formula

$$J\eta(\boldsymbol{x}) = -\left(\frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}, \eta(\boldsymbol{x}))\right)^{-1} \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x}))\,.$$

Dimostrazione. Faremo la dimostrazione per induzione su N.

Nel caso $N = 1$ e $M \geq 2$, si procede in modo del tutto analogo a quanto già fatto nel caso $M = 1$. Basterà prendere, al posto dell'intervallo $[x_0 - \delta, x_0 + \delta]$, la palla chiusa $\overline{B}(\boldsymbol{x}_0, \delta)$, e similmente per gli intorni aperti di $\boldsymbol{x}_0$, per dimostrare l'esistenza e la continuità della funzione η. Resta da vedere la derivabilità: considerato $\bar{\boldsymbol{x}} = (\bar{x}_1, \dots, \bar{x}_M)$, prendiamo ora $\boldsymbol{x} = (\bar{x}_1 + h, \dots, \bar{x}_M)$; procedendo come in precedenza, si ha che

$$\frac{\eta(\bar{x}_1 + h, \dots, \bar{x}_M) - \eta(\bar{x}_1, \dots, \bar{x}_M)}{h} = -\frac{\frac{\partial g}{\partial x_1}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))}\,,$$

con $\gamma(\xi)$ appartenente al segmento che congiunge $(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))$ con $(\boldsymbol{x}, \eta(\boldsymbol{x}))$. Se h tende a 0, si ha che $\gamma(\xi)$ tende a $(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))$ e quindi

$$\frac{\partial \eta}{\partial x_1}(\bar{\boldsymbol{x}}) = \lim_{h \to 0} \frac{\eta(\bar{x}_1 + h, \dots, \bar{x}_M) - \eta(\bar{x}_1, \dots, \bar{x}_M)}{h} = -\frac{\frac{\partial g}{\partial x_1}(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))}{\frac{\partial g}{\partial y}(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))}\,.$$

Analogamente si calcolano le derivate parziali rispetto a $x_2, \dots, x_M$, per cui si vede che η è di classe $\mathcal{C}^1$ e

$$J\eta(\boldsymbol{x}) = -\frac{1}{\frac{\partial g}{\partial y}(\boldsymbol{x}, \eta(\boldsymbol{x}))} \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x}))\,.$$

Supponiamo ora l'enunciato valido fino a $N-1$, per un certo $N \geq 2$ (e $M \geq 1$ qualsiasi) e dimostriamo che vale anche per N. Useremo la notazione

$$\tilde{\boldsymbol{y}}_1 = (y_1, \dots, y_{N-1})\,,$$

per cui scriveremo $y = (\tilde{\boldsymbol{y}}_1, y_N)$. Siccome

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) & \cdots & \frac{\partial g_1}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_N}{\partial y_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) & \cdots & \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \end{pmatrix} \neq 0\,,$$

almeno uno degli elementi dell'ultima colonna è non nullo. Possiamo supporre senza perdita di generalità, eventualmente permutando le righe, che sia $\frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0$. Scrivendo $\boldsymbol{y}_0 = (\tilde{\boldsymbol{y}}_1^0, y_N^0)$, con $\tilde{\boldsymbol{y}}_1^0 = (y_1^0, \dots, y_{N-1}^0)$, sarà

$$g_N(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0, y_N^0) = 0\,, \qquad \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0, y_N^0) \neq 0\,.$$

Allora (caso unidimensionale) esistono un intorno aperto U_1 di $(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0)$, un intorno aperto V_N di y_N^0 e una funzione $\eta_1 : U_1 \to V_N$ di classe $\mathcal{C}^1$ tali che $U_1 \times V_N \subseteq \Omega$, per cui si abbia: se $(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \in U_1$ e $y_N \in V_N$,

$$g_N(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = 0 \quad \Leftrightarrow \quad y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)\,,$$

e

$$J\eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = -\frac{1}{\frac{\partial g_N}{\partial y_N}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1))} \frac{\partial g_N}{\partial(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1))\,.$$

Possiamo supporre U_1 della forma $\tilde{U} \times \tilde{V}_1$, con $\tilde{U}$ intorno aperto di $\boldsymbol{x}_0$ e $\tilde{V}_1$ intorno aperto di $\tilde{\boldsymbol{y}}_1^0$. Definiamo la funzione $\phi : \tilde{U} \times \tilde{V}_1 \to \mathbb{R}^{N-1}$, ponendo

$$\phi(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = (g_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)), \dots, g_{N-1}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)))\,.$$

Per brevità, scriveremo

$$g_{(1,\dots,N-1)}(\boldsymbol{x}, \boldsymbol{y}) = (g_1(\boldsymbol{x}, \boldsymbol{y}), \dots, g_{N-1}(\boldsymbol{x}, \boldsymbol{y}))\,.$$

Notiamo che $\phi(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = 0$ e che, essendo $\eta_1(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = y_N^0$,

$$\frac{\partial \phi}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = \frac{\partial g_{(1,\ldots,N-1)}}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) + \frac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0)\frac{\partial \eta_1}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0)\,. \qquad (*)$$

Inoltre, siccome $g_N(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)) = 0$, per ogni $(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \in U_1$, differenziando si ha:

$$0 = \frac{\partial g_N}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) + \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0)\frac{\partial \eta_1}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0)\,. \qquad (**)$$

Scriviamo

$$\det \frac{\partial \phi}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = \frac{1}{\frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0)} \det \left(\begin{array}{c|c} \dfrac{\partial \phi}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) & \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \\ \hline 0 & \dfrac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \end{array}\right),$$

avendo usato la notazione di matrice suddivisa a blocchi. Sostituendo le due uguaglianze $(*)$, $(**)$ e usando le proprietà dei determinanti, si ha:

$$\det \left(\begin{array}{c|c} \dfrac{\partial \phi}{\partial \tilde{y}_1}(x_0, \tilde{y}_1^0) & \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(x_0, y_0) \\ \hline 0 & \dfrac{\partial g_N}{\partial y_N}(x_0, y_0) \end{array}\right) =$$

$$= \det \left(\begin{array}{c|c} \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial \tilde{y}_1}(x_0, y_0) + \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(x_0, y_0)\dfrac{\partial \eta_1}{\partial \tilde{y}_1}(x_0, \tilde{y}_1^0) & \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(x_0, y_0) \\ \hline \dfrac{\partial g_N}{\partial \tilde{y}_1}(x_0, y_0) + \dfrac{\partial g_N}{\partial y_N}(x_0, y_0)\dfrac{\partial \eta_1}{\partial \tilde{y}_1}(x_0, \tilde{y}_1^0) & \dfrac{\partial g_N}{\partial y_N}(x_0, y_0) \end{array}\right)$$

$$= \det \left(\begin{array}{c|c} \dfrac{\partial g}{\partial \tilde{y}_1}(x_0, y_0) + \dfrac{\partial g}{\partial y_N}(x_0, y_0)\dfrac{\partial \eta_1}{\partial \tilde{y}_1}(x_0, \tilde{y}_1^0) & \dfrac{\partial g}{\partial y_N}(x_0, y_0) \end{array}\right)$$

$$= \det \left(\begin{array}{c|c} \dfrac{\partial g}{\partial \tilde{y}_1}(x_0, y_0) & \dfrac{\partial g}{\partial y_N}(x_0, y_0) \end{array}\right) = \det \frac{\partial g}{\partial y}(x_0, y_0) \neq 0\,.$$

Abbiamo quindi

$$\phi(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = \boldsymbol{0}\,, \qquad \det \frac{\partial \phi}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) \neq 0\,.$$

Per l'ipotesi induttiva, esistono un intorno aperto U di $\boldsymbol{x}_0$, un intorno aperto V_1 di $\tilde{\boldsymbol{y}}_1^0$ e una funzione $\eta_2 : U \to V_1$ di classe $\mathcal{C}^1$ tali che $U \times V_1 \subseteq \tilde{U} \times \tilde{V}_1$, per cui si abbia: per ogni $\boldsymbol{x} \in U$ e $\tilde{\boldsymbol{y}}_1 \in V_1$,

$$\phi(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = \boldsymbol{0} \quad \Leftrightarrow \quad \tilde{\boldsymbol{y}}_1 = \eta_2(\boldsymbol{x})\,.$$

In conclusione, per $\boldsymbol{x} \in U$ e $\boldsymbol{y} = (\tilde{\boldsymbol{y}}_1, y_N) \in V_1 \times V_2$, si ha:

$$g(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0} \quad \Leftrightarrow \quad \begin{cases} g_{(1,\ldots,N-1)}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = \boldsymbol{0} \\ g_N(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = 0 \end{cases}$$

$$
\begin{aligned}
&\Leftrightarrow \quad \begin{cases} g_{(1,\dots,N-1)}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = \mathbf{0} \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
&\Leftrightarrow \quad \begin{cases} \phi(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = \mathbf{0} \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
&\Leftrightarrow \quad \begin{cases} \tilde{\boldsymbol{y}}_1 = \eta_2(\boldsymbol{x}) \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
&\Leftrightarrow \quad \boldsymbol{y} = (\eta_2(\boldsymbol{x}), \eta_1(\boldsymbol{x}, \eta_2(\boldsymbol{x}))) \,.
\end{aligned}
$$

Ponendo $V = V_1 \times V_2$, resta pertanto definita la funzione $\eta : U \to V$:

$$\eta(\boldsymbol{x}) = (\eta_2(\boldsymbol{x}), \eta_1(\boldsymbol{x}, \eta_2(\boldsymbol{x}))) \,.$$

Tale funzione è di classe $\mathcal{C}^1$, siccome lo sono sia η_1 che η_2. Siccome $g(\boldsymbol{x}, \eta(\boldsymbol{x})) = 0$ per ogni $\boldsymbol{x} \in U$, se ne deduce che che

$$\frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x})) + \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}, \eta(\boldsymbol{x})) J\eta(\boldsymbol{x}) = \mathbf{0} \,,$$

da cui la formula per $J\eta(\boldsymbol{x})$. ■

Ed ecco l'enunciato simmetrico.

Teorema 42 *Siano $\Omega \subseteq \mathbb{R}^M \times \mathbb{R}^N$ un aperto, $g : \Omega \to \mathbb{R}^M$ una funzione di classe $\mathcal{C}^1$ e $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ un punto di Ω per cui si abbia:*

$$g(\boldsymbol{x}_0, \boldsymbol{y}_0) = \mathbf{0} \,, \qquad \det \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0 \,.$$

Allora esistono un intorno aperto U di $\boldsymbol{x}_0$, un intorno aperto V di $\boldsymbol{y}_0$ e una funzione $\eta : V \to U$ di classe $\mathcal{C}^1$ tali che $U \times V \subseteq \Omega$ e, presi $\boldsymbol{x} \in U$ e $\boldsymbol{y} \in V$, si ha:

$$g(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0} \quad \Leftrightarrow \quad \boldsymbol{x} = \eta(\boldsymbol{y}) \,.$$

Inoltre, la funzione η è di classe $\mathcal{C}^1$ e vale la formula

$$J\eta(\boldsymbol{y}) = -\left(\frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{y}, \eta(\boldsymbol{y})) \right)^{-1} \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{y}, \eta(\boldsymbol{y})) \,.$$

Vediamo ora un'importante conseguenza del teorema della funzione implicita.

Definizione 5 *Dati A e B, due aperti di $\mathbb{R}^N$, una funzione $\varphi : A \to B$ sè un "diffeomorfismo" se è di classe $\mathcal{C}^1$, biiettiva e la sua inversa $\varphi^{-1} : B \to A$ è anch'essa di classe $\mathcal{C}^1$.*

Enunciamo il **teorema di inversione locale**.

Teorema 43 *Siano A e B due aperti di $\mathbb{R}^N$ e $\varphi : A \to B$ una funzione di classe $\mathcal{C}^1$. Se per un certo $\boldsymbol{x}_0 \in A$ si ha che $\det J\varphi(\boldsymbol{x}_0) \neq 0$, allora esistono un intorno aperto U di $\boldsymbol{x}_0$ contenuto in A, e un intorno aperto V di $\varphi(\boldsymbol{x}_0)$ contenuto in B, tali che la restrizione $\varphi_{|U} : U \to V$ sia un diffeomorfismo.*

Dimostrazione. Consideriamo la funzione $g : A \times B \to \mathbb{R}^N$ definita da

$$g(\boldsymbol{x}, \boldsymbol{y}) = \varphi(\boldsymbol{x}) - \boldsymbol{y} .$$

Posto $\boldsymbol{y}_0 = \varphi(\boldsymbol{x}_0)$, si ha che

$$g(\boldsymbol{x}_0, \boldsymbol{y}_0) = \boldsymbol{0} , \qquad \text{e} \qquad \det \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}_0, \boldsymbol{y}_0) = \det J\varphi(\boldsymbol{x}_0) \neq 0 .$$

Per il teorema della funzione implicita, esistono un intorno aperto V di $\boldsymbol{y}_0$, un intorno aperto U di $\boldsymbol{x}_0$ e una funzione $\eta : V \to U$ di classe $\mathcal{C}^1$ tali che, presi $\boldsymbol{y} \in V$ e $\boldsymbol{x} \in U$, si ha:

$$\varphi(\boldsymbol{x}) = \boldsymbol{y} \quad \Leftrightarrow \quad g(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0} \quad \Leftrightarrow \quad \boldsymbol{x} = \eta(\boldsymbol{y}) .$$

Quindi, $\eta = \varphi_{|U}^{-1}$ e la dimostrazione è così completa. ■

15 Le M-superfici

Indichiamo con I un rettangolo di $\mathbb{R}^M$, dove $1 \leq M \leq N$.

Definizione 6 *Chiameremo* **M-superficie in $\mathbb{R}^N$** *una funzione* [3] $\sigma : I \to \mathbb{R}^N$ *di classe* $\mathcal{C}^1$. *Se* $M = 1$, σ *si dirà anche* **curva***; se* $M = 2$, *si dirà semplicemente* **superficie**. *L'insieme* $\sigma(I)$ *è detto* **supporto** *della M-superficie σ. Diremo che la M-superficie σ è* **regolare** *se, per ogni* $\boldsymbol{u} \in \mathring{I}$, *la matrice jacobiana* $\sigma'(\boldsymbol{u})$ *ha rango M.*

Consideriamo da vicino il caso $N = 3$. Una curva in $\mathbb{R}^3$ è una funzione $\sigma : [a, b] \to \mathbb{R}^3$, $\sigma = (\sigma_1, \sigma_2, \sigma_3)$. La curva è regolare se, per ogni $t \in]a, b[$, il vettore $\sigma'(t) = (\sigma_1'(t), \sigma_2'(t), \sigma_3'(t))$ è non nullo. In tal caso, si definisce il seguente **versore tangente** nel punto $\sigma(t)$:

$$\tau_\sigma(t) = \frac{\sigma'(t)}{||\sigma'(t)||} .$$

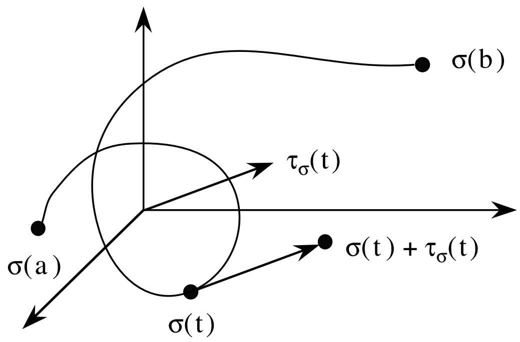
Esempio. La curva $\sigma : [0, 2\pi] \to \mathbb{R}^3$ definita da

$$\sigma(t) = (R\cos(2t), R\sin(2t), 0)$$

ha come supporto la circonferenza

$$\{(x, y, z) : x^2 + y^2 = R^2, z = 0\}$$

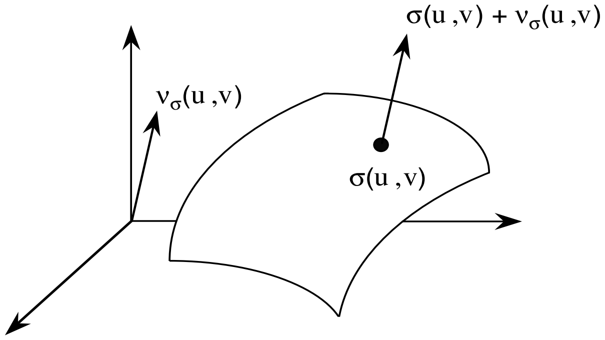
[3] Le derivate parziali di σ devono essere continue su tutto I e nei punti di frontiera vanno intese, se necessario, come derivate destre o sinistre. Equivalentemente, si potrebbe estendere σ ad una funzione di classe $\mathcal{C}^1$ definita su un aperto contenente I. In questa ottica, il dominio di σ potrebbe essere un insieme più generale, ad esempio la chiusura di un aperto limitato, affinché il differenziale risulti ben definito anche nei punti di frontiera. Considerazioni analoghe possono essere fatte per i domini delle forme differenziali considerate.



(che viene percorsa due volte). Essendo $\sigma'(t) = (-2R\sin(2t), 2R\cos(2t), 0)$, si tratta di una curva regolare, e si ha:

$$\tau_\sigma(t) = (-\sin(2t), \cos(2t), 0)\,.$$

Una superficie in $\mathbb{R}^3$ è una funzione $\sigma : [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}^3$. La superficie è regolare se, per ogni $(u, v) \in\,]a_1, b_1[\, \times\,]a_2, b_2[$, i vettori $\frac{\partial\sigma}{\partial u}(u, v)$, $\frac{\partial\sigma}{\partial v}(u, v)$ sono linearmente indipendenti. In tal caso, essi individuano un piano, detto **piano tangente** alla superficie nel punto $\sigma(u, v)$, e si definisce il seguente **versore normale**:



$$\nu_\sigma(u, v) = \frac{\frac{\partial\sigma}{\partial u}(u, v) \times \frac{\partial\sigma}{\partial v}(u, v)}{||\frac{\partial\sigma}{\partial u}(u, v) \times \frac{\partial\sigma}{\partial v}(u, v)||}\,.$$

Esempi. 1. La superficie $\sigma : [0, \pi] \times [0, \pi] \to \mathbb{R}^3$ definita da

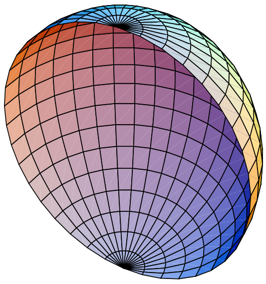
$$\sigma(\phi, \theta) = (R\sin\phi\cos\theta, R\sin\phi\sin\theta, R\cos\phi)$$

ha come supporto la semisfera

$$\{(x, y, z) : x^2 + y^2 + z^2 = R^2, y \geq 0\}\,.$$

Essendo

$$\frac{\partial\sigma}{\partial\phi}(\phi, \theta) = (R\cos\phi\cos\theta, R\cos\phi\sin\theta, -R\sin\phi)\,,$$
$$\frac{\partial\sigma}{\partial\theta}(\phi, \theta) = (-R\sin\phi\sin\theta, R\sin\phi\cos\theta, 0)\,,$$



si tratta di una superficie regolare, e si ha:

$$\nu_\sigma(\phi,\theta) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)\,.$$

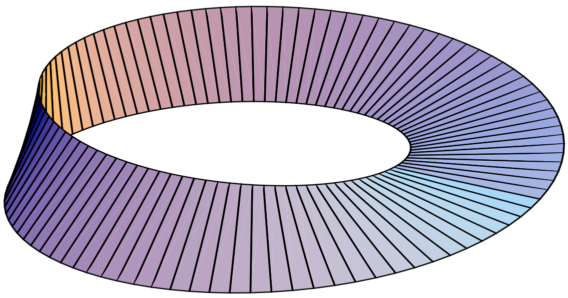
2. La superficie $\sigma : [r,R]\times[0,2\pi] \to \mathbb{R}^3$, con $0 \le r < R$, data da

$$\sigma(u,v) = (u\cos v, u\sin v, 0)\,,$$

ha come supporto un cerchio se $r=0$, una corona circolare se $r>0$. È una superficie regolare.

3. La superficie $\sigma : [r,R]\times[0,2\pi] \to \mathbb{R}^3$, con $0 < r < R$, definita da

$$\begin{aligned}\sigma(u,v) = \Bigg(\bigg(\frac{r+R}{2} + \Big(u - \frac{r+R}{2}\Big)\cos\Big(\frac{v}{2}\Big)\bigg)\cos v,\\ \bigg(\frac{r+R}{2} + \Big(u - \frac{r+R}{2}\Big)\cos\Big(\frac{v}{2}\Big)\bigg)\sin v,\\ \Big(u - \frac{r+R}{2}\Big)\sin\Big(\frac{v}{2}\Big)\Bigg)\,,\end{aligned}$$



ha comes supporto un nastro di Möbius. È anch'essa una superficie regolare.

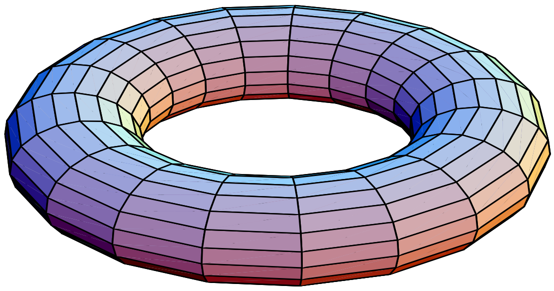
4. La superficie $\sigma : [0,2\pi]\times[0,2\pi] \to \mathbb{R}^3$ definita da

$$\sigma(u,v) = ((R + r\cos u)\cos v, (R + r\cos u)\sin v, r\sin u)$$

dove $0 < r < R$, ha come supporto l'anello toroidale o "toro"

$$\{(x,y,z) : (\sqrt{x^2+y^2} - R)^2 + z^2 = r^2\}\,.$$

Si può verificare che anche in questo caso si tratta di una superficie regolare.



Una 3-superficie in $\mathbb{R}^3$ si dice anche **volume**.

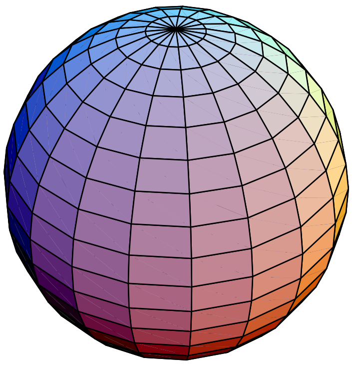
Esempio. La funzione $\sigma : [0, R] \times [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3$ definita da

$$\sigma(\rho, \phi, \theta) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi)$$

ha come supporto la palla chiusa

$$\{(x, y, z) : x^2 + y^2 + z^2 \leq R^2\} .$$

In questo caso, $\det \sigma'(\rho, \phi, \theta) = \rho^2 \sin \phi$ e pertanto si tratta di un volume regolare.



16 Analisi locale delle M-superfici

Sia $1 \leq M < N$. Identificando $\mathbb{R}^N$ con $\mathbb{R}^M \times \mathbb{R}^{N-M}$, ogni vettore $\boldsymbol{x} = (x_1, \ldots, x_N)$ di $\mathbb{R}^N$ si scriverà nella forma $\boldsymbol{x} = (\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}})$, con $\hat{\boldsymbol{x}} = (x_1, \ldots, x_M)$ e $\tilde{\boldsymbol{x}} = (x_{M+1}, \ldots, x_N)$.

Useremo inoltre la seguente notazione: dato $\hat{\boldsymbol{x}} = (x_1, \ldots, x_M) \in \mathbb{R}^M$ e $r > 0$,

$$B[\hat{\boldsymbol{x}}, r] = [x_1 - r, x_1 + r] \times \ldots \times [x_M - r, x_M + r] \subseteq \mathbb{R}^M .$$

Per semplicità, scriveremo $B[r]$ invece di $B[\mathbf{0}, r]$.

16.1 Le funzioni vincolo

Teorema 44 *Siano Ω un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di Ω e $g : \Omega \to \mathbb{R}^{N-M}$ una funzione di classe $\mathcal{C}^1$, tale che*

$$g(\boldsymbol{x}_0) = \mathbf{0}\,, \quad e \quad Jg(\boldsymbol{x}_0) \text{ ha rango } N - M.$$

Allora esistono un intorno U di $\boldsymbol{x}_0$ e una M-superficie regolare e iniettiva $\sigma : B[r] \to \mathbb{R}^N$, per un certo $r > 0$, tali che $\sigma(\mathbf{0}) = \boldsymbol{x}_0$ e

$$\{\boldsymbol{x} \in U : g(\boldsymbol{x}) = \mathbf{0}\} = \sigma(\mathbf{B[r]})\,.$$

<u>Dimostrazione</u>. Supponiamo, per esempio, che sia invertibile

$$\frac{\partial g}{\partial \tilde{\boldsymbol{x}}}(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial g_1}{\partial p_{M+1}}(\boldsymbol{x}_0) \cdots & \frac{\partial g_1}{\partial p_N}(\boldsymbol{x}_0) \\ \vdots \quad\quad \cdots & \vdots \\ \frac{\partial g_{N-M}}{\partial p_{M+1}}(\boldsymbol{x}_0) \cdots & \frac{\partial g_{N-M}}{\partial p_N}(\boldsymbol{x}_0) \end{pmatrix}.$$

Per il teorema della funzione implicita, esistono un intorno aperto $\hat{U}$ di $\hat{\boldsymbol{x}}_0$, un intorno aperto $\widetilde{U}$ di $\tilde{\boldsymbol{x}}_0$ e una funzione $\eta : \hat{U} \to \widetilde{U}$, tali che $\hat{U} \times \widetilde{U} \subseteq \Omega$ e, se $\hat{\boldsymbol{x}} \in \hat{U}$ e $\tilde{\boldsymbol{x}} \in \widetilde{U}$, si ha:

$$g(\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) = \mathbf{0} \quad \Leftrightarrow \quad \tilde{\boldsymbol{x}} = \eta(\hat{\boldsymbol{x}})\,.$$

Preso $r > 0$ tale che $B[\hat{\boldsymbol{x}}_0, r] \subseteq \hat{U}$, sia $U = B[\hat{\boldsymbol{x}}_0, r] \times \widetilde{U}$ e $\sigma : B[r] \to \mathbb{R}^N$ definita da $\sigma(\boldsymbol{u}) = (\boldsymbol{u} + \hat{\boldsymbol{x}}_0, \eta(\boldsymbol{u} + \hat{\boldsymbol{x}}_0))$. Si verifica che la matrice jacobiana $J\sigma(\boldsymbol{u})$ ha come sottomatrice la matrice $M \times M$ identità, per cui σ è regolare. Si vede facilmente che σ è iniettiva, in quanto lo è la prima componente $\boldsymbol{u} \mapsto \boldsymbol{u} + \hat{\boldsymbol{x}}_0$. Inoltre, se $\boldsymbol{x} = (\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) \in U$,

$$g(\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) = \mathbf{0} \quad \Leftrightarrow \quad \tilde{\boldsymbol{x}} = \eta(\hat{\boldsymbol{x}}) \quad \Leftrightarrow \quad (\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) = \sigma(\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}_{\mathbf{0}})\,,$$

da cui la tesi. Nel caso in cui la sottomatrice considerata non sia invertibile, basterà operare degli scambi nelle colonne della matrice $Jg(\boldsymbol{x}_0)$ per ricondursi alla situazione precedente. ∎

La M-superficie σ individuata dal teorema precedente è detta "M-parametrizzazione locale".

Vediamo tre casi di particolare interesse. Iniziamo con una curva in $\mathbb{R}^2$ (caso $M = 1$, $N = 2$).

Corollario 3 *Siano Ω un sottoinsieme aperto di $\mathbb{R}^2$, (x_0, y_0) un punto di Ω e $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$, tale che*

$$g(x_0, y_0) = 0 \quad e \quad \nabla g(x_0, y_0) \neq 0\,.$$

Allora esistono un intorno U di (x_0, y_0) e una curva regolare e iniettiva $\sigma : [-r, r] \to \mathbb{R}^2$, per un certo $r > 0$, tali che $\sigma(0) = (x_0, y_0)$ e

$$\{(x, y) \in U : g(x, y) = 0\} = \sigma([-r, r])\,.$$

Vediamo ora il caso di una superficie in $\mathbb{R}^3$ (caso $M = 2$, $N = 3$).

Corollario 4 *Siano Ω un sottoinsieme aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω e $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$, tale che*

$$g(x_0, y_0, z_0) = 0 \qquad e \qquad \nabla g(x_0, y_0, z_0) \neq 0\,.$$

Allora esistono un intorno U di (x_0, y_0, z_0) e una superficie regolare e iniettiva $\sigma : [-r, r] \times [-r, r] \to \mathbb{R}^3$, per un certo $r > 0$, tali che $\sigma(0, 0) = (x_0, y_0, z_0)$ e

$$\{(x, y, z) \in U : g(x, y, z) = 0\} = \sigma([-r, r] \times [-r, r])\,.$$

Infine, vediamo il caso di una curva in $\mathbb{R}^3$ (caso $M = 1$, $N = 3$).

Corollario 5 *Siano Ω un sottoinsieme aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω e $g_1, g_2 : \Omega \to \mathbb{R}$ due funzioni di classe $\mathcal{C}^1$, tali che*

$$g_1(x_0, y_0, z_0) = g_2(x_0, y_0, z_0) = 0 \qquad e \qquad \nabla g_1(x_0, y_0, z_0) \times \nabla g_2(x_0, y_0, z_0) \neq 0\,.$$

Allora esistono un intorno U di (x_0, y_0, z_0) e una curva regolare e iniettiva $\sigma : [-r, r] \to \mathbb{R}^3$, per un certo $r > 0$, tali che $\sigma(0) = (x_0, y_0, z_0)$ e

$$\{(x, y, z) \in U : g_1(x, y, z) = g_2(x, y, z) = 0\} = \sigma([-r, r])\,.$$

16.2 I moltiplicatori di Lagrange

Siano Ω un aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di Ω e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in $\boldsymbol{x}_0$. Vogliamo cercare eventuali punti di massimo o di minimo per la funzione ristretta a un "vincolo", che sarà descritto da un'altra funzione, in generale a valori vettoriali.

Teorema 45 *Sia $g : \Omega \to \mathbb{R}^{N-M}$ una funzione di classe $\mathcal{C}^1$ tale che*

$$g(\boldsymbol{x}_0) = \boldsymbol{0}\,, \quad e \quad Jg(\boldsymbol{x}_0) \text{ ha rango } N - M.$$

Posto

$$S = \{\boldsymbol{x} \in \Omega : g(\boldsymbol{x}) = \boldsymbol{0}\}\,,$$

se $\boldsymbol{x}_0$ è un punto di minimo o massimo locale per $f|_S$, allora esistono $(N - M)$ numeri reali $\lambda_1, \dots, \lambda_{N-M}$ tali che

$$\nabla f(\boldsymbol{x}_0) = \sum_{j=1}^{N-M} \lambda_j \nabla g_j(\boldsymbol{x}_0)\,.$$

I numeri $\lambda_1, \dots, \lambda_{N-M}$ si chiamano **moltiplicatori di Lagrange**.

Dimostrazione. Per il teorema precedente, esistono un intorno U di $\boldsymbol{x}_0$, un $r > 0$ e una M-superficie regolare $\sigma : B[r] \to \mathbb{R}^N$ tali che $\sigma(\mathbf{0}) = \boldsymbol{x}_0$ e

$$S \cap U = \sigma(B[r])\,.$$

Considerata la funzione $F : B[r] \to \mathbb{R}$, definita da $F(\boldsymbol{u}) = f(\sigma(\boldsymbol{u}))$, si ha che $\mathbf{0}$ è un punto di minimo o massimo locale per F. Quindi, $\nabla F(\mathbf{0}) = \mathbf{0}$, per cui

$$0 = JF(\mathbf{0}) = Jf(\boldsymbol{x}_0)J\sigma(\mathbf{0})\,.$$

Ne segue che

$$\nabla f(\boldsymbol{x}_0) \;\; \textit{è ortogonale a} \;\; \frac{\partial\sigma}{\partial u_1}(\mathbf{0}), \dots, \frac{\partial\sigma}{\partial u_M}(\mathbf{0})\,.$$

Inoltre, essendo $g(\sigma(\boldsymbol{u})) = \mathbf{0}$ per ogni $\boldsymbol{u} \in B[r]$, si ha che

$$Jg(\boldsymbol{x}_0)J\sigma(\mathbf{0}) = 0\,.$$

Quindi anche i vettori

$$\nabla g_1(\boldsymbol{x}_0), \dots, \nabla g_{N-M}(\boldsymbol{x}_0) \;\; \textit{sono tutti ortogonali a} \;\; \frac{\partial\sigma}{\partial u_1}(\mathbf{0}), \dots, \frac{\partial\sigma}{\partial u_M}(\mathbf{0})\,.$$

Siccome σ è regolare,

$$\textit{lo spazio vettoriale } \mathcal{T} \textit{ generato da } \frac{\partial\sigma}{\partial u_1}(\mathbf{0}), \dots, \frac{\partial\sigma}{\partial u_M}(\mathbf{0}) \textit{ ha dimensione } M\,.$$

Quindi lo spazio ortogonale $\mathcal{T}^\perp$ ha dimesnione $N - M$. E siccome, come abbiamo visto,

$$\nabla f(\boldsymbol{x}_0), \nabla g_1(\boldsymbol{x}_0), \dots, \nabla g_{N-M}(\boldsymbol{x}_0) \in \mathcal{T}^\perp\,,$$

questi vettori devono essere linearmente dipendenti. Quindi, essendo che i vettori $\nabla g_1(\boldsymbol{x}_0)$, ..., $\nabla g_{N-M}(\boldsymbol{x}_0)$ linearmente indipendenti, ne segue che $\nabla f(\boldsymbol{x}_0)$ si deve poter esprimere come una loro combinazione lineare. ∎

Vediamo anche qui tre casi particolari interessanti.

Corollario 6 *Siano Ω un aperto di $\mathbb{R}^2$, (x_0, y_0) un punto di Ω, $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$ tale che*

$$g(x_0, y_0) = 0 \quad \textit{e} \quad \nabla g(x_0, y_0) \neq 0,$$

e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in (x_0, y_0). Posto

$$S = \{(x, y) \in \Omega : g(x, y) = 0\}\,,$$

se (x_0, y_0) è un punto di minimo o massimo locale per $f|_S$, allora esiste un numero reale λ tale che

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)\,.$$

Corollario 7 *Siano Ω un aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω, $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$ tale che*

$$g(x_0, y_0, z_0) = 0 \qquad e \qquad \nabla g(x_0, y_0, z_0) \neq 0,$$

e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in (x_0, y_0, z_0). Posto

$$S = \{(x, y, z) \in \Omega : g(x, y, z) = 0\},$$

se (x_0, y_0, z_0) è un punto di minimo o massimo locale per $f|_S$, allora esiste un numero reale λ tale che

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0).$$

Corollario 8 *Siano Ω un aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω, $g_1, g_2 : \Omega \to \mathbb{R}$ due funzioni di classe $\mathcal{C}^1$ tali che*

$$g_1(x_0, y_0, z_0) = g_2(x_0, y_0, z_0) = 0 \qquad e \qquad \nabla g_1(x_0, y_0, z_0) \times \nabla g_2(x_0, y_0, z_0) \neq 0,$$

e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in (x_0, y_0, z_0). Posto

$$S = \{(x, y, z) \in U : g_1(x, y, z) = 0,\ g_2(x, y, z) = 0\},$$

se (x_0, y_0, z_0) è un punto di minimo o massimo locale per $f|_S$, allora esistono due numeri reali λ_1, λ_2 tali che

$$\nabla f(x_0, y_0, z_0) = \lambda_1 \nabla g_1(x_0, y_0, z_0) + \lambda_2 \nabla g_2(x_0, y_0, z_0).$$