1 Preliminari su spazi vettoriali

Ricordiamoci che dati due spazi vettoriali U e V su $\mathbb{R}$ una funzione lineare $T : U \to V$ è caratterizzata dalla seguente proprietà:

a $T(u+u') = Tu + Tu'$ per ogni $(u,u') \in U^2$

b $T\lambda u = \lambda T u$ per ogni $\lambda \in \mathbb{R}$ e per ogni $u \in U$.

Denoto con $\mathcal{L}(U,V)$ l'insieme delle funzioni lineari da U a V. Notiamo che $\mathcal{L}(U,V)$ è a sua volta uno spazio vettoriale su $\mathbb{R}$.

Più in generale, dati spazi vettoriali $U_1, ..., U_n$ e V, una mappa n–lineare $T : U_1 \times ... \times U_n \to V$ è caratterizzata dalle seguenti proprietà:

a' per ogni $i \in \{1,...,n\}$ e per ogni scelta di vettori, si ha

$$T[(\mathbf{x}_1,...,\mathbf{x}_{i-1},\mathbf{x}_i+\mathbf{x}_i',\mathbf{x}_{i+1},...,\mathbf{x}_n)] = T[(\mathbf{x}_1,...,\mathbf{x}_{i-1},\mathbf{x}_i,\mathbf{x}_{i+1},...,\mathbf{x}_n)] + T[(\mathbf{x}_1,...,\mathbf{x}_{i-1},\mathbf{x}_i',\mathbf{x}_{i+1},...,\mathbf{x}_n)]$$

b' Per ogni $\lambda \in \mathbb{R}$, per ogni $i \in \{1,...,n\}$ e per ogni scelta di vettori, si ha

$$T[(\mathbf{x}_1,...,\mathbf{x}_{i-1},\lambda\mathbf{x}_i,\mathbf{x}_{i+1},...,\mathbf{x}_n)] = \lambda T[(\mathbf{x}_1,...,\mathbf{x}_{i-1},\mathbf{x}_i,\mathbf{x}_{i+1},...,\mathbf{x}_n)].$$

Denotiamo con $\mathcal{L}^n(U_1\times...\times U_n,V)$ l'insieme di queste mappe n–lineari. Se $U = U_1 = ... = U_n$ scriviamo $\mathcal{L}^n(U,V) = \mathcal{L}^n(U^n,V)$. Notiamo che $\mathcal{L}^n(U_1\times...\times U_n,V)$ è a sua volta uno spazio vettoriale su $\mathbb{R}$ dove per $T,S \in \mathcal{L}^n(U_1\times...\times U_n,V)$ e $\lambda,\mu \in \mathbb{R}$

$$(\lambda T + \mu S)(\mathbf{x}_1,...,\mathbf{x}_n) := \lambda T(\mathbf{x}_1,...,\mathbf{x}_n) + \mu S(\mathbf{x}_1,...,\mathbf{x}_n).$$

Lemma 1.1. *Esiste un naturale isomorfismo*

$$\mathcal{L}^n(U_1\times...\times U_n,V) \ni T \to \mathrm{t} \in \mathcal{L}(U_1,\mathcal{L}^{n-1}(U_2\times...\times U_n,V)) \tag{1.1}$$

definito da

$$\mathrm{t}(\mathbf{x}_1)(\mathbf{x}_2,...,\mathbf{x}_n) := T(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n) \tag{1.2}$$

Dim. Per prima cosa, verifichiamo che la mappa (1.1)–(1.2) è ben definita. Infatti, abbiamo una funzione

$$U_1 \ni \mathbf{x}_1 \to T(\mathbf{x}_1,\cdot,...,\cdot) \in \mathcal{L}^{n-1}(U_2\times...\times U_n,V) \tag{1.3}$$

perchè chiaramente

$$U_2\times...\times U_n \ni (\mathbf{x}_2,...,\mathbf{x}_n) \to T(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n) \in V \text{ , per } \mathbf{x}_1 \text{ fissato,}$$

è $(n-1)$–lineare. Inoltre, in termini di $\mathbf{x}_1$, (1.3) è lineare, perché

$$
\begin{aligned}
&T(\lambda\mathbf{x}_1+\mu\mathbf{x}'_1,\cdot,...,\cdot)(\mathbf{x}_2,...,\mathbf{x}_n)=T(\lambda\mathbf{x}_1+\mu\mathbf{x}'_1,\mathbf{x}_2,...,\mathbf{x}_n)\\
&\;=\lambda T(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n)+\mu T(\mathbf{x}'_1,\mathbf{x}_2,...,\mathbf{x}_n)=\big(\lambda T(\mathbf{x}_1,\cdot,...,\cdot)+\mu T(\mathbf{x}'_1,\cdot,...,\cdot)\big)\,(\mathbf{x}_2,...,\mathbf{x}_n).
\end{aligned}
$$

Quindi abbiamo verificato che (1.1)–(1.2) definisce una funzione lineare. Per verificare che si tratta di un isomorfismo, basta definire l'inversa. Presa $\mathrm{t}\in\mathcal{L}(U_1,\mathcal{L}^{n-1}(U_2\times...\times U_n,V))$ definiamo

$$
T(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n):=\mathrm{t}(\mathbf{x}_1)(\mathbf{x}_2,...,\mathbf{x}_n). \tag{1.4}
$$

E' facile verificare che $T\in\mathcal{L}^n(U_1\times...\times U_n,V)$. Ovviamente le mappe (1.1) e (1.4) sono inverse l'una dell'altra. □

Osservazione 1.2. Si noti che se $T:\mathbb{R}^{d_1}\times...\times\mathbb{R}^{d_n}\to\mathbb{R}^N$ è multilineare, per

$$
\mathbf{x}_j=\begin{pmatrix}x_{j1}\\ \vdots\\ x_{jd_j}\end{pmatrix}=\sum_{l=1}^{d_j}x_{jl}e_{jl}\text{ dove} \tag{1.5}
$$

$$
e_{jl}:=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ 0\\ \vdots\\ 0\end{pmatrix}\begin{matrix}\\ \\ \\ l\text{–esima}\\ \\ \\ \\ d_j\text{–esima}\end{matrix} \tag{1.6}
$$

abbiamo

$$
T[(\mathbf{x}_1,...,\mathbf{x}_n)]=\sum_{l_1=1}^{d_1}...\sum_{l_n=1}^{d_n}T[(e_{1l_1},...,e_{nl_n})]x_{1l_1}...x_{nl_n}. \tag{1.7}
$$

Si noti che (1.7) è un polinomio a coefficienti $T[(e_{1l_1},...,e_{nl_n})]\in\mathbb{R}^N$

Ricordiamoci che per ogni spazio vettoriale U su $\mathbb{R}$ si definisce duale di U lo spazio $U^*=\mathcal{L}(U,\mathbb{R})$. In generale U e U^* sono spazi diversi. Quando $U=\mathbb{R}^d$ esiste però una naturale identificazione di U e U^*, che enunciamo senza dimostrazione.

Lemma 1.3. *In $\mathbb{R}^d$ consideriamo il prodotto scalare $\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$*

$$
(\mathbf{x},\mathbf{y})=\sum_{i=1}^{d}x_iy_i
$$

Allora per ogni $\mathbf{x} \in \mathbb{R}^d$*, la funzione* $\mathbb{R}^d \ni \mathbf{y} \to (\mathbf{x}, \mathbf{y}) \in \mathbb{R}$ *è una funzione lineare, e quindi appartiene al duale* $(\mathbb{R}^d)^*$*. Denoto questo elemento del duale* $(\mathbb{R}^d)^*$ *con* $(\mathbf{x}, \cdot)$*. Risulta che la corrispondenza*

$$\mathbb{R}^d \ni \mathbf{x} \to (\mathbf{x}, \cdot) \in (\mathbb{R}^d)^* \tag{1.8}$$

è un isomorphismo lineare tra $\mathbb{R}^d$ *ed il suo duale* $(\mathbb{R}^d)^*$*.*

□

In questa parte del corso, considereremo funzioni tra spazi euclidei.

2 Spazi Euclidei

Indichiamo con $\mathbb{R}$ l'insime dei numeri reali, ossia la retta reale. Con $\mathbb{R}^2$ il piano, con $\mathbb{R}^3$ lo spazio tridimensionale. In generale $\mathbb{R}^d$ è l'insieme delle $d-$uple ordinate di numeri reali:

$$\mathbb{R}^d = \{(x_1, \cdots, x_d) \colon x_1 \in \mathbb{R}, \cdots, x_d \in \mathbb{R}\}.$$

In $\mathbb{R}^d$ la distanza, o metrica, euclidea tra due punti $\mathbf{x} = (x_1, \cdots, x_d)$ e $\mathbf{y} = (y_1, \cdots, y_d)$ è il numero

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{n=1}^{d} (x_n - y_n)^2}. \tag{2.1}$$

In $\mathbb{R}$, la distanza si riduce a

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|.$$

Nel seguito, utilizzeremo la metrica Euclidea (2.1), salvo indicazione esplicita, tuttavia si possono fare altri esempi.

Definizione 2.1 (Norme)**.** Sia X uno spazio vettoriale su $\mathbb{R}$. Una mappa $\|\cdot\| : X \to [0, +\infty)$ è una *norma* su X se soddisfa le seguenti proprietà:

1. $\|\mathbf{x}\| = 0 \Longleftrightarrow \mathbf{x} = 0$
2. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ per ogni coppia $\mathbf{x}, \mathbf{y} \in X$
3. $\|\lambda \mathbf{x}\| = |\lambda| \, \|\mathbf{x}\|$ per ogni $\lambda \in \mathbb{R}$ e $\mathbf{x} \in X$.

   Uno spazio vettoriale X con una norma $\|\cdot\|$ è detto uno spazio normato.

Esercizio 2.2. Verificare se in uno spazio normato $(X, \|\cdot\|)$ poniamo $d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$ questo definisce una metrica su X.

Vale il seguente.

Definizione 2.3. Abbiamo le seguenti definizioni

i Dato un insieme X e due metriche d_1 e d_2 su di esso, queste si dicono equivalenti se esiste una costante $C \geq 1$ tale che

$$\frac{1}{C} d_1(x, y) \leq d_2(x, y) \leq C d_1(x, y) \text{ per ogni coppia } x, y \in X.$$

ii Dato uno spazio vettoriale X con due norme $\|\cdot\|_1$ e $\|\cdot\|_2$, queste si dicono equivalenti se se esiste una costante $C \geq 1$ tale che

$$\frac{1}{C} \|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq C \|\mathbf{x}\|_1 \text{ per ogni } \mathbf{x} \in X.$$

Esercizio 2.4. Verificare se in uno spazio spazio vettoriale X ci sono due norme $\|\cdot\|_1$ e $\|\cdot\|_2$, queste sono equivalenti se e solo se le corrispondenti metriche sono equivalenti.

Osservazione 2.5. Le seguenti, sono tre norme su $\mathbb{R}^d$, tra loro equivalenti:

1. la norma euclidea $|\mathbf{x}| := \sqrt{\sum_{n=1}^{d} x_n^2}$, la cui associata metrica è la (2.1);
2. $|\mathbf{x}|_1 := \sum_{n=1}^{d} |x_n|$;
3. $|\mathbf{x}|_\infty := \sup\{|x_n| : n = 1, ..., d\}$.

Più in generale, le seguenti norme su $\mathbb{R}^d$, sono equivalenti alle precedenti:

$$|\mathbf{x}|_p := \left(\sum_{n=1}^{d} |x_n|^p \right)^{\frac{1}{p}} \text{ per } p \in (1, \infty). \tag{2.2}$$

Infatti, per prima cosa, vale

$$|\mathbf{x}|_q \leq |\mathbf{x}|_p \text{ per } p < q. \tag{2.3}$$

Considero solo il caso $1 \leq p < q < \infty$ (il caso $q = +\infty$ è il più semplice). Risulta che (2.3) è equivalente a

$$\left(\sum_{n=1}^{d} |x_n|^q \right)^{\frac{p}{q}} \leq \sum_{n=1}^{d} |x_n|^p \text{ per } p < q. \tag{2.4}$$

Consideriamo la funzione $f(t) = |t|^{\frac{p}{q}}$. Siccome $0 < \frac{p}{q} < 1$ è una funzione concava $f(t) \geq 0$ con $f(0) = 0$. Si può dimostrare allora che

$$f(t_1 + ... + t_d) \leq f(t_1) + ... + f(t_d) \text{ per ogni scelta di numeri } t_1 \geq 0, ..., t_d \geq 0.$$

Allora abbiamo quanto segue, che dimostra (2.4),

$$\left(\sum_{n=1}^{d} |x_n|^q\right)^{\frac{p}{q}} = f\left(\sum_{n=1}^{d} |x_n|^q\right) \le \sum_{n=1}^{d} f\left(|x_n|^q\right) = \sum_{n=1}^{d} |x_n|^p.$$

D'altra parte, siccome $|x_n| \le |\mathbf{x}|_\infty$, per $p > q$ abbiamo

$$\left(\sum_{n=1}^{d} |x_n|^q\right)^{\frac{1}{q}} \le \left(\sum_{n=1}^{d} |\mathbf{x}|_\infty^q\right)^{\frac{1}{q}} \le \sqrt[q]{d}\ |\mathbf{x}|_\infty \le \sqrt[q]{d}\ |\mathbf{x}|_p. \tag{2.5}$$

Esercizio 2.6. Consideriamo un operatore lineare $L \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$. Ricordiamo che esiste una naturale identificazione con lo spazio delle matrici $N \times d$, data da

$$L\mathbf{x} = \begin{pmatrix} L_{11} & \dots & L_{1d} \\ \vdots & \vdots & \vdots \\ L_{N1} & \dots & L_{Nd} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, \tag{2.6}$$

dove a destra ho il prodotto riga colonna. Dimostrare che se definiamo

$$|L| := \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{d} L_{ij}^2} \tag{2.7}$$

resta definita una norma nello spazio $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$.

Osservazione 2.7. La norma (2.7) non è quella usata di solito. Infatti di solito, anche in ambito infinito dimensionale, si definisce come norma $|L|$ di L la minima costante $|L|$ per la quale la proposizione (2.8) è vera. In dimensione finita le due norme sono equivalenti. La norma (2.7) in ambito infinito dimensionale corrisponde alla norma di Hilbert–Schmidt. In ambito infinito dimensionale, le norme non sono equivalenti. Si rinvia ad un corso di Analisi Funzionale alla Magistrale.

Esempio 2.8. Consideriamo un operatore lineare $L \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$. Dimostrare che utilizzando (2.7) si ha la disuguaglianza

$$|L\mathbf{x}| \le |L|\ |\mathbf{x}| \text{ per ogni } \mathbf{x} \in \mathbb{R}^d. \tag{2.8}$$

dove $|L\mathbf{x}|$ e $|\mathbf{x}|$ sono le norme euclidee.
Infatti

$$\begin{aligned} |L\mathbf{x}|^2 &= \sum_{i=1}^{N} \left(\sum_{j=1}^{d} L_{ij} x_j\right)^2 \le \sum_{i=1}^{N} \left(\sum_{j=1}^{d} L_{ij}^2\right) \sum_{j'=1}^{d} x_{j'}^2 \\ &= \sum_{i=1}^{N} \sum_{j=1}^{d} L_{ij}^2 |\mathbf{x}|^2 = |L|^2\ |\mathbf{x}|^2 \end{aligned}$$

Esempio 2.9. Dati $B \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$ e $A \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$ abbiamo

$$|AB| \leq |A|\ |B|. \tag{2.9}$$

Infatti, posto $C = AB$, abbiamo

$$c_{ij} = \sum_{k=1}^{N} a_{ik} b_{kj}$$

che implica

$$c_{ij}^2 = \left| \sum_{k=1}^{N} a_{ik} b_{kj} \right|^2 \leq \sum_{k=1}^{N} a_{ik}^2 \sum_{k'=1}^{N} b_{k'j}^2$$

Allora

$$\sum_{i=1}^{M} \sum_{j=1}^{d} c_{ij}^2 = \sum_{i=1}^{M} \sum_{j=1}^{d} \left| \sum_{k=1}^{N} a_{ik} b_{kj} \right|^2 \leq \sum_{i=1}^{M} \sum_{j=1}^{d} \left(\sum_{k=1}^{N} a_{ik}^2 \right) \left(\sum_{k'=1}^{N} b_{k'j}^2 \right) = \left(\sum_{i=1}^{M} \sum_{k=1}^{N} a_{ik}^2 \right) \left(\sum_{j=1}^{d} \sum_{k'=1}^{N} b_{k'j}^2 \right)$$

dando il risultato desiderato.

Esempio 2.10. Sia $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$. Procedendo per induzione, da (2.9) si ricava abbiamo

$$|A^n| \leq |A|^n \text{ per ogni } n = 2, 3, ... \tag{2.10}$$

Supponiamo $|A| < 1$. Allora risulta che la matrice $1 + A$ è invertibile, ed infatti è la somma della seguente serie (serie di Neumann):

$$(1 + A)^{-1} = \sum_{n=0}^{\infty} (-1)^n A^n. \tag{2.11}$$

Inoltre abbiamo

$$|(1 + A)^{-1}| = \left| \sum_{n=0}^{\infty} (-1)^n A^n \right| \leq \sum_{n=0}^{\infty} |A|^n = \frac{1}{1 - |A|} \tag{2.12}$$

Quindi abbiamo

$$|(1 + A)^{-1}| \leq 2 \text{ se } |A| \leq \frac{1}{2}. \tag{2.13}$$

Nel seguito, utilizzeremo la metrica Euclidea (2.1) e la corrispondente norma, salvo esplicita indicazione contraria.

Dato $\mathbf{x}_0 \in \mathbb{R}^d$ e dato un numero $r > 0$, la palla "aperta" di centro $\mathbf{x}_0$ e raggio r, che denoterò con $D(\mathbf{x}_0, r)$, è l'insieme dei punti la cui distanza da $\mathbf{x}_0$ è strettamente minore di r. In formule

$$D(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbb{R}^d \colon d(\mathbf{x}_0, \mathbf{x}) < r\}.$$

La palla "chiusa" di centro $\mathbf{x}_0$ e raggio r, che denoterò con $\overline{D(\mathbf{x}_0, r)}$, è l'insieme dei punti la cui distanza da $\mathbf{x}_0$ è minore o uguale di r. In formule

$$\overline{D(\mathbf{x}_0, r)} = \{\mathbf{x} \in \mathbb{R}^d \colon d(\mathbf{x}_0, \mathbf{x}) \leq r\}.$$

Quando $\mathbf{x}_0 = 0$ e $r = 1$, si parla di palla unitaria, aperta o chiusa.

Esercizio 2.11. Prendendo la dimensione $d = 2$, disegnare le palle unitarie per le metriche nell'esercizio 2.5.

Un sottoinsieme A di $\mathbb{R}^d$ si dice aperto se dato un qualsiasi punto $\mathbf{x}_0 \in A$ esiste una palla aperta di centro $\mathbf{x}_0$ contenuta in A, ossia se esiste $r > 0$ tale che $d(\mathbf{x}_0, \mathbf{x}) < r \Rightarrow \mathbf{x} \in A$. Si include tra gli aperti anche l'insieme vuoto.

Esempio 2.12. Una palla aperta è un aperto. Il sottoinsieme del piano definito da $x > 0$ ed $y > x^3$ è un aperto (lo si dimostrerà dopo). Eccetera.

Un sottoinsieme C di $\mathbb{R}^d$ si dice chiuso, se l'insieme complementare $\complement C$ (ossia l'insieme dei punti di $\mathbb{R}^d$ non in C) é aperto.

Definizione 2.13. Sia $X \subseteq \mathbb{R}^d$ un sottoinsieme non vuoto. Un punto $\mathbf{x}_0 \in \mathbb{R}^d$ si dice di accumulazione per X se per ogni $\epsilon > 0$ esiste $\mathbf{x} \in X$ tale che $0 < |\mathbf{x}_0 - \mathbf{x}| < \epsilon$. Denoterò con X' l'insieme dei punti di accumulazione di X.

Esercizio 2.14. Sia $X \subseteq \mathbb{R}^d$ un sottoinsieme non vuoto. Verificare che le seguenti due proposizioni sono equivalenti.

1. X è un insieme chiuso di $\mathbb{R}^d$.
2. $X \supseteq X'$.

Esercizio 2.15. Sia $X \subseteq \mathbb{R}^d$ un sottoinsieme non vuoto. Verificare che $\overline{X} = X \cup X'$.

Esercizio 2.16. Verificare che la palla chiusa $\overline{D(\mathbf{x}_0, r)}$ è la chiusura della palla aperta $D(\mathbf{x}_0, r)$.

3 Limiti

Definizione 3.1. Sia $\{\mathbf{x}_n\}_{n \in \mathbb{N}}$ una successione di elementi di $\mathbb{R}^d$. Diciamo che un punto $\mathbf{L} \in \mathbb{R}^d$ è il limite della successione in $\mathbb{R}^d$, e scriviamo $\lim_{n \to +\infty} \mathbf{x}_n = \mathbf{L}$, se

$$\text{per ogni } \epsilon > 0 \text{ esiste } n_\epsilon \in \mathbb{R} \text{ tale che } n > n_\epsilon \Longrightarrow |\mathbf{x}_n - \mathbf{L}| < \epsilon. \tag{3.1}$$

Osserviamo che

$$\mathbf{x}_n = \begin{pmatrix} x_n^1 \\ \vdots \\ x_n^d \end{pmatrix} \text{ e } \mathbf{L} = \begin{pmatrix} L^1 \\ \vdots \\ L^d \end{pmatrix}.$$

Ciò significa che assegnare $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ una successione in $\mathbb{R}^d$ è equivalente all'assegnare d successioni in $\mathbb{R}$ (prese in ordine). Notiamo che per ogni $i_0 \in \{1, ..., d\}$.

$$|x_n^{i_0} - L^{i_0}| = \sqrt{(x_n^{i_0} - L^{i_0})^2} \le \sqrt{\sum_{j=1}^{d}(x_n^j - L^j)^2} = |\mathbf{x}_n - \mathbf{L}| \tag{3.2}$$

Quindi concludiamo che (3.1) implica che per ogni i_0 e per ogni $\epsilon > 0$ si ha $n > n_\epsilon \Longrightarrow |x_n^{i_0} - L^{i_0}| < \epsilon$. Pertanto

$$\lim_{n\to+\infty} \mathbf{x}_n = \mathbf{L} \Longrightarrow \lim_{n\to+\infty} x_n^{i_0} = L^{i_0} \text{ per ogni } i_0 = 1, ..., d. \tag{3.3}$$

Vale anche l'implicazione opposta. Infatti, se $\lim\limits_{n\to+\infty} x_n^{i_0} = L^{i_0}$ per ogni $i_0 = 1, ..., d$, è facile concludere che per ogni $\epsilon > 0$ esiste $n_\epsilon \in \mathbb{R}$ tale che $n > n_\epsilon \Longrightarrow |x_n^{i_0} - L^{i_0}| < \epsilon$ per ogni i_0. Ma allora,

$$n > n_\epsilon \Longrightarrow |\mathbf{x}_n - \mathbf{L}| = \sqrt{\sum_{j=1}^{d}(x_n^j - L^j)^2} < \sqrt{\sum_{j=1}^{d}\epsilon^2} = \epsilon\sqrt{d}.$$

Quindi abbiamo dimostrato che

$$\text{per ogni } \epsilon > 0 \text{ esiste } m_\epsilon \in \mathbb{R} \text{ tale che } n > m_\epsilon \Longrightarrow |\mathbf{x}_n - \mathbf{L}| < \epsilon\sqrt{d}. \tag{3.4}$$

E' facile concludere che questo implica la Proposizione (3.1). E quindi abbiamo dimostrato che in (3.7) vale anche l'implicazione $\Longleftarrow$.

Definizione 3.2. Sia $X \subseteq \mathbb{R}^d$, $f : X \to \mathbb{R}^N$, $\mathbf{x}_0$ un punto di accumulazione di X e $\mathbf{L} \in \mathbb{R}^N$. Scriviamo che $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{L}$ se

$$\text{per ogni } \epsilon > 0 \text{ esiste } \delta_\epsilon \in \mathbb{R}_+ \text{ tale che } 0 < |\mathbf{x} - \mathbf{x}_0| < \delta_\epsilon \Longrightarrow |f(\mathbf{x}) - \mathbf{L}| < \epsilon. \tag{3.5}$$

Nel caso particolare in cui $\mathbf{x}_0 \in X$, se $L = f(\mathbf{x}_0)$ diciamo che f è continua nel punto $\mathbf{x}_0$. In particolare, se $\mathbf{x}_0 \in X \cap X'$, f è continua in $\mathbf{x}_0$ se

$$\text{per ogni } \epsilon > 0 \text{ esiste } \delta_\epsilon \in \mathbb{R}_+ \text{ tale che } |\mathbf{x} - \mathbf{x}_0| < \delta_\epsilon \Longrightarrow |f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon. \tag{3.6}$$

Notare che in un punto $x_0 \in X\backslash X'$, cioè in un punto $x_0 \in X$ che è isolato, la proposizione (3.6) è sempre vera. Cioè tutte le funzioni sono continue in tutti i punti isolati del loro dominio di definizione.

Una funzione $f : X \to \mathbb{R}^N$ che è continua in tutti i punti di X si dice continua in X. Denoterò il relativo insieme con $C^0(X, \mathbb{R}^N)$. Notare che $C^0(X, \mathbb{R}^N)$ è uno spazio vettoriale (siete in grado di dimostrarlo?).

Notiamo che

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_N(\mathbf{x}) \end{pmatrix} \text{ e } \mathbf{L} = \begin{pmatrix} L_1 \\ \vdots \\ L_N \end{pmatrix}.$$

In analogia con quanto visto sopra per ogni $i_0 \in \{1, ..., d\}$.

$$|f_{i_0}(\mathbf{x}) - L_{i_0}| = \sqrt{(f_{i_0}(\mathbf{x}) - L_{i_0})^2} \leq \sqrt{\sum_{j=1}^{d} (f_j(\mathbf{x}) - L_j)^2} = |f(\mathbf{x}) - \mathbf{L}|.$$

Quindi concludiamo che (3.5) implica che per ogni i_0 e per ogni $\epsilon > 0$, $0 < |\mathbf{x} - \mathbf{x}_0| < \delta_\epsilon \Longrightarrow |f_{i_0}(\mathbf{x}) - L_{i_0}| < \epsilon$.. Pertanto

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{L} \Longrightarrow \lim_{\mathbf{x} \to \mathbf{x}_0} f_{i_0}(\mathbf{x}) = L_{i_0} \text{ per ogni } i_0 \tag{3.7}$$

Anche qui, come sopra, vale $\Longleftarrow$. Infatti, se $\lim\limits_{\mathbf{x} \to \mathbf{x}_0} f_{i_0}(\mathbf{x}) = L_{i_0}$ per ogni i_0, è facile verificare che

$$\text{per ogni } \epsilon > 0 \text{ esiste } \delta'_\epsilon \in \mathbb{R}_+ \text{ tale che } 0 < |\mathbf{x} - \mathbf{x}_0| < \delta'_\epsilon \Longrightarrow |f_{i_0}(\mathbf{x}) - L_{i_0}| < \epsilon \text{ per ogni } i_0.$$

Allora

$$0 < |\mathbf{x} - \mathbf{x}_0| < \delta'_\epsilon \Longrightarrow |f(\mathbf{x}) - \mathbf{L}| = \sqrt{\sum_{j=1}^{d} (f_j(\mathbf{x}) - L_j)^2} < \sqrt{\sum_{j=1}^{d} \epsilon^2} = \epsilon\sqrt{d}.$$

Facile concludere che questo implica (3.5).

Concludiamo che nello studio dei limiti di funzioni ci possiamo concentrare sul caso di funzioni (a valori) scalari, cioè con $N = 1$.

Osservazione 3.3. Notiamo per $d = 1$, considerato che l'analisi $f : X \subseteq \mathbb{R} \to \mathbb{R}^N$ si riduce spesso all'analisi di funzioni N funzioni $f : X \subseteq \mathbb{R} \to \mathbb{R}$, in sostanza non ci sono particolari novità rispetto ad Analisi 1. Tuttavia più avanti considereremo qualche caso di funzioni $f : X \subseteq \mathbb{R} \to \mathbb{R}^N$.

Esempio 3.4. Una classe importante di funzioni continue è quella delle funzioni Lipschitziane. Una funzione $X \subseteq \mathbb{R}^d$, $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ si dice Lipschitziana con costante di Lipschitz $K > 0$ se

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq K|\mathbf{x} - \mathbf{y}| \text{ per ogni coppia } \mathbf{x}, \mathbf{y} \in X\ . \tag{3.8}$$

E' immediato che si tratta di funzioni continue: si può prendere $\delta_\epsilon = K^{-1}\epsilon$ per ogni $\epsilon > 0$.

Esempio 3.5. Sia $X \subseteq \mathbb{R}^d$ e per $i_0 \in 1, ..., d$ sia $\pi_{i_0} : X \to \mathbb{R}$ definito da $\pi_{i_0}(\mathbf{x}) := x_{i_0}$. Questa è la proiezione i_0–esima. Si noti che, come in (3.2),

$$|\pi_{i_0}(\mathbf{x}) - \pi_{i_0}(\mathbf{y})| = |x_{i_0} - y_{i_0}| \leq |\mathbf{x} - \mathbf{y}| \text{ per ogni coppia } \mathbf{x}, \mathbf{y} \in X \ .$$

Pertanto si tratta di una funzione Lipschitziana con costante di Lipschitz $K = 1$, e quindi è continua.

Esempio 3.6. Quindi la funzione $(x, y) \to y$ e la funzione $(x, y) \to x$ nel piano sono continue. Siccome $(x, y) \to x^3$ può essere interpretata come la composizione della proiezione e di $x \to x^3$, è continua. Pertanto l'insieme $x > 0$ è aperto in $\mathbb{R}^2$ e l'insieme $y > x^3$ è aperto in $\mathbb{R}^2$. La loro intersezione è un aperto in $\mathbb{R}^2$, confermando Esempio 2.12.

Esempio 3.7. Data una mappa lineare $L : \mathbb{R}^d \to \mathbb{R}^N$, da

$$|L\mathbf{x}| \leq |L| \ |\mathbf{x}| \text{ per ogni } \mathbf{x} \in \mathbb{R}^d$$

è facile concludere che L è Lipschitziana con costante di Lipschitz $|L|$ definita in (2.7).

Si noti che, nel precedente Esempio 3.5, si ha $N = 1$

Esempio 3.8. Fissiamo $\mathbf{v} \in \mathbb{R}^N$ dove

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}$$

e consideriamo la mappa $F : \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N) \to \mathbb{R}^d$ definita da (prodotto riga colonna)

$$\mathcal{L}(\mathbb{R}^d, \mathbb{R}^N) \ni L \to {}^t\mathbf{v}L \in \mathbb{R}^d \text{ dove } {}^t\mathbf{v} = (v_1, ..., v_N). \tag{3.9}$$

Risulta che F è un operatore lineare da $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$ a $\mathbb{R}^d$ con norma (2.7) data da $|F| = |\mathbf{v}|$.

Esempio 3.9. Fissiamo $\mathbf{u} \in \mathbb{R}^d$ dove

$$\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_d \end{pmatrix}$$

e consideriamo la mappa $G : \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N) \to \mathbb{R}^N$ definita da (prodotto riga colonna)

$$\mathcal{L}(\mathbb{R}^d, \mathbb{R}^N) \ni L \to L\mathbf{u} \in \mathbb{R}^N. \tag{3.10}$$

Risulta che G è un operatore lineare da $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$ a $\mathbb{R}^N$ con norma (2.7) data da $|G| = |\mathbf{u}|$.

Esempio 3.10. Consideriamo una mappa n–lineare $T : \mathbb{R}^{d_1} \times ... \times \mathbb{R}^{d_n} \to \mathbb{R}^N$. Allora come in (1.7) abbiamo

$$T[(\mathbf{x}_1, ..., \mathbf{x}_n)] = \sum_{l_1=1}^{d_1} ... \sum_{l_n=1}^{d_n} T[(e_{1l_1}, ..., e_{nl_n})] x_{1l_1} ... x_{nl_n}.$$

In $\mathbb{R}^N$ abbiamo

$$
\begin{aligned}
|T[(\mathbf{x}_1, ..., \mathbf{x}_n)]| &\leq \sum_{l_1=1}^{d_1} ... \sum_{l_n=1}^{d_n} |T[(e_{1l_1}, ..., e_{nl_n})]| \ |x_{1l_1} ... x_{nl_n}| \\
&\leq C \sqrt{\sum_{l_1=1}^{d_1} ... \sum_{l_n=1}^{d_n} |x_{1l_1} ... x_{nl_n}|^2} = C \sqrt{\sum_{l_1=1}^{d_1} |x_{1l_1}|^2 ... \sum_{l_n=1}^{d_n} |x_{nl_n}|^2} = C|\mathbf{x}_1| ... |\mathbf{x}_d| \qquad (3.11) \\
\text{per } C &:= \sqrt{\sum_{l_1=1}^{d_1} ... \sum_{l_n=1}^{d_n} |T[(e_{1l_1}, ..., e_{nl_n})]|^2} \qquad (3.12)
\end{aligned}
$$

Notare che se definisco

$$
|(\mathbf{x}_1, ..., \mathbf{x}_n)| = \sqrt{\sum_{j=1}^{n} |\mathbf{x}_j|^2} = \sqrt{\sum_{j=1}^{n} \sum_{l_j=j}^{d_j} x_{jl_j}^2}
$$

ottengo la metrica euclidea in $\mathbb{R}^{d_1+...+d_n} = \mathbb{R}^{d_1} \times ... \times \mathbb{R}^{d_n}$.
Notiamo che

$$
|T[(\mathbf{x}_1, ..., \mathbf{x}_n)]| \leq C|\mathbf{x}_1| ... |\mathbf{x}_n| \leq C|(\mathbf{x}_1, ..., \mathbf{x}_n)|^n. \qquad (3.13)
$$

Chiaramente si tratta di una funzione continua che per qualsiasi $n \geq 2$.

Esempio 3.11. Ci chiediamo quale sia

$$
\lim_{(x,y)\to(0,0)} \frac{x^3 - y^3}{x^2 + y^2} =?
$$

Posto $r = \sqrt{x^2 + y^2}$, notare che il precedente limite coincide con

$$
\lim_{r\to 0} \frac{x^3 - y^3}{x^2 + y^2}.
$$

Qui può essere conveniente considerare le coordinate polari

$$
\begin{aligned}
x &= r\cos\theta \\
y &= r\sin\theta.
\end{aligned}
$$

Qui $r \geq 0$ è la distanza di (x, y) dall'origine, mentre θ è un angolo, per esempio in $[0, 2\pi]$. Notare per inciso che $\mathbb{R}^2$ nel piano (r, θ) diventa $[0, +\infty) \times [0, 2\pi]$ e la palla chiusa in $\mathbb{R}^2$ di centro l'origine e raggio R diviene il rettangolo $[0, R] \times [0, 2\pi]$.

Ora scriviamo

$$
\frac{x^3 - y^3}{x^2 + y^2} = \frac{r^3 \cos^3\theta - r^3 \sin^3\theta}{r^2} = r\left(\cos^3\theta - \sin^3\theta\right)
$$

ed osserviamo che all'avvicinarsi di (x, y) all'origine, ossia quando $r \to 0$, abbiamo

$$\left| r\left(\cos^3\theta - \sin^3\theta\right)\right| \leq 2r \to 0.$$

Ossia il limite esiste ed è 0.

Esempio 3.12. Ci chiediamo quale sia

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

Dobbiamo guardare a

$$\frac{xy}{x^2+y^2} = \frac{r^2\cos\theta\sin\theta}{r^2} = \cos\theta\sin\theta.$$

Come si vede, per $r \to 0$ quest'ultima quantità ha, a seconda dell'angolo con cui ci avviciniamo all'origine, valori differenti. Quindi concludiamo che

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} \quad \text{non esiste.}$$

Esempio 3.13. Ci chiediamo quale sia

$$\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{x^2+y^2} = \lim_{r\to 0} \frac{\sin(xy)}{x^2+y^2} =?$$

Ora possiamo porre

$$\frac{\sin(xy)}{x^2+y^2} = \frac{\sin(xy)}{xy}\frac{xy}{x^2+y^2}.$$

Per la regola del prodotto, se il limite esiste

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

Siccome $\lim_{(x,y)\to(0,0)} xy = 0$ abbiamo per via dei limiti notevoli

$$\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{xy} = 1.$$

Allora il limite esiste se e solo se esiste il limite nel precedente esempio. Quindi

$$\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{x^2+y^2} \quad \text{non esiste.}$$

Esempio 3.14. Ci chiediamo quale sia

$$\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^4+y^2} = \lim_{r\to 0} \frac{x^2y}{x^4+y^2} = L =?$$

Se spediamo (x, y) verso $(0, 0)$ mantenendoci lungo la generica retta per l'origine $y = mx$, posto

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

abbiamo

$$f(x, mx) = \frac{mx^3}{x^4 + m^2x^2} = \frac{mx}{x^2 + m^2} \xrightarrow{x \to 0} \frac{0}{m^2} = 0$$

e pertanto, se il limite L esiste si ha $L = 0$. Tuttavia, se ci mettiamo sulla parabola $y = x^2$,

$$f(x, x^2) = \frac{x^4}{x^4 + x^4} = \frac{1}{2} \xrightarrow{x \to 0} \frac{1}{2} \neq 0.$$

Ossia L non esiste.

Esempio 3.15. Consideriamo la funzione

$$f(x, y) = \begin{cases} \frac{\exp\left(-\frac{1}{x^2}\right) y}{\exp\left(-\frac{2}{x^2}\right) + y^2} \text{ se } x \neq 0 \\ 0 \text{ se } x = 0. \end{cases} \tag{3.14}$$

Si nota che se ci manteniamo sulla curva $x^m = (y/c)^n$, con m ed n primi tra loro, cioè $y = cx^{m/n}$, con $x > 0$ se n è pari, allora

$$f\left(x, cx^{m/n}\right) = c\frac{\exp\left(-\frac{1}{x^2}\right) x^{m/n}}{\exp\left(-\frac{2}{x^2}\right) + c^2 x^{2m/n}} = c\frac{\exp\left(-\frac{1}{x^2}\right) x^{-m/n}}{\exp\left(-\frac{2}{x^2}\right) x^{-2m/n} + c^2} \xrightarrow{x \to 0} 0.$$

Tuttavia, per $y = \exp\left(-\frac{1}{x^2}\right)$ (notare che $\lim_{x \to 0} \exp\left(-\frac{1}{x^2}\right) = 0$) abbiamo

$$f\left(x, \exp\left(-\frac{1}{x^2}\right)\right) = \frac{\exp\left(-\frac{1}{x^2}\right) \exp\left(-\frac{1}{x^2}\right)}{\exp\left(-\frac{2}{x^2}\right) + \exp\left(-\frac{2}{x^2}\right)} \equiv \frac{1}{2}.$$

Di conseguenza, $\lim_{(x,y) \to (0,0)} f(x, y)$ non esiste e quindi in particolare f non è continua nel punto $(0, 0)$.

Gli ultimi due esempi sono tratti da [1].

Esempio 3.16. Consideriamo la funzione

$$f(x, y) = \begin{cases} x \sin\left(\frac{1}{y}\right) + y \sin\left(\frac{1}{x}\right) \text{ se } xy \neq 0 \\ 0 \text{ se } xy = 0. \end{cases}$$

Notiamo che

$$0 \leq |f(x, y)| \leq |x| + |y|$$

e pertanto, dal Teorema dei Carabinieri, siccome $\lim_{(x,y) \to (0,0)} (|x| + |y|) = 0$, segue che $\lim_{(x,y) \to (0,0)} f(x, y) = 0$. Si noti che i due seguenti limiti in generale non esistono:

$$\lim_{x \to 0} \lim_{y \to 0} f(x, y) \text{ e } \lim_{y \to 0} \lim_{x \to 0} f(x, y).$$

Definizione 3.17. Sia $X \subseteq \mathbb{R}^d$, un insieme illimitato e consideriamo una funzione $f : X \to \mathbb{R}^N$ ed $\mathbf{L} \in \mathbb{R}^N$. Scriviamo che $\lim_{\mathbf{x}\to\infty} f(\mathbf{x}) = \mathbf{L}$ se

$$\text{per ogni } \epsilon > 0 \text{ esiste } r_\epsilon \in \mathbb{R}_+ \text{ tale che } |\mathbf{x}| > r_\epsilon \text{ e } \mathbf{x} \in X \Longrightarrow |f(\mathbf{x}) - \mathbf{L}| < \epsilon. \tag{3.15}$$

Nel caso particolare $N = 1$, scriviamo che $\lim_{\mathbf{x}\to\infty} f(\mathbf{x}) = +\infty$ se

$$\text{per ogni } M \in \mathbb{R} \text{ esiste } r_M \in \mathbb{R}_+ \text{ tale che } |\mathbf{x}| > r_M \text{ e } \mathbf{x} \in X \Longrightarrow f(\mathbf{x}) > M \tag{3.16}$$

oppure scriviamo che $\lim_{\mathbf{x}\to\infty} f(\mathbf{x}) = -\infty$ se

$$\text{per ogni } M \in \mathbb{R} \text{ esiste } r_M \in \mathbb{R}_+ \text{ tale che } |\mathbf{x}| > r_M \text{ e } \mathbf{x} \in X \Longrightarrow f(\mathbf{x}) < M \tag{3.17}$$

In tutti gli esempi di questa sezione $\lim_{\mathbf{x}\to\infty} f(\mathbf{x})$ non esiste.

4 Teorema di Bolzano–Weierstrass

Theorem 4.1 (Bolzano–Weierstrass)**.** *Un sottoinsieme X di $\mathbb{R}^d$ è compatto per successioni se e solo se è chiuso e limitato.*

Dim. $\Longrightarrow$ Incominciamo col supporre che X è compatto per successioni. Ciò significa che

$$\forall \text{ successione } \{\mathbf{x}_n\}_{n\in\mathbb{N}} \text{ in } X\ \exists\ \{\mathbf{x}_{n_k}\}_{k\in\mathbb{N}} \text{ ed } \overline{x} \in X \text{ con } \mathbf{x}_{n_k} \xrightarrow{k\to+\infty} \overline{x}. \tag{4.1}$$

Se X non fosse chiuso, esisterebbe un punto $y \in \mathbb{R}^d \backslash X$ di accumulazione di X. Pertanto, dalla Definizione 2.13, per ogni $n \in \mathbb{N}$ esisterebbe un $\mathbf{x}_n \in X$ tale che $0 < |\mathbf{y} - \mathbf{x}_n| < 1/n$. Dalla Definizione 3.1 questo implicherebbe che $\lim_{n\to+\infty} \mathbf{x}_n = \mathbf{y}$. Questo a sua volta implica che, per ogni sottosuccessione $\{\mathbf{x}_{n_k}\}_{k\in\mathbb{N}}$ di questa particolare successione, si ha $\lim_{k\to+\infty} \mathbf{x}_{n_k} = \mathbf{y}$ (sapete dimostrarlo?). Ma questo contraddice la proposizione (4.1). Siccome quest'ultima è vera per ipotesi, concludiamo che non può esistere un punto $y \in \mathbb{R}^d \backslash X$ di accumulazione di X, e che pertanto X è un insieme chiuso.
Sempre assumendo che X è compatto per successioni, dimostriamo che è limitato, cioè dimostriamo che

$$\exists\ r_0 > 0 \text{ t.c. } |\mathbf{x}| \le r_0 \text{ per ogni } \mathbf{x} \in X. \tag{4.2}$$

Supponiamo che X sia illimitato. Questo equivale a scrivere che

$$\forall\ n \in \mathbb{N}\ \exists\ \mathbf{x}_n \in X \text{ t.c. } |\mathbf{x}_n| \ge n. \tag{4.3}$$

Resta determinata una successione $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ in X. Si noti che (4.3) implica

$$\lim_{n\to+\infty} |\mathbf{x}_n| = +\infty. \tag{4.4}$$

Siccome, per ipotesi, X è compatto per successioni, è vera la proposizione (4.1). In particolare, per una opportuna sottosuccessione $\{\mathbf{x}_{n_k}\}_{k\in\mathbb{N}}$ con $\mathbf{x}_{n_k} \xrightarrow{k\to+\infty} \overline{\mathbf{x}} \in X$, per la continuità della funzione $|\cdot| : \mathbb{R}^d \to \mathbb{R}$, segue che

$$\lim_{k\to+\infty} |\mathbf{x}_{n_k}| = |\overline{\mathbf{x}}|. \tag{4.5}$$

Ma, siccome (4.4) implica

$$\lim_{k\to+\infty} |\mathbf{x}_{n_k}| = +\infty,$$

per l'unicità del limite, otteniamo la contraddizione $+\infty > |\overline{\mathbf{x}}| = +\infty$, cioè un assurdo. Quindi (4.3) non può essere vera. Pertanto (4.3) è vera.

Dim. $\Longleftarrow$ Supponiamo ora che $X \subseteq \mathbb{R}^d$ sia chiuso e limitato (e non vuoto). In particolare vale (4.3). Questo implica che per ogni $\mathbf{x} \in X$ e per ogni sua componente x_{j_0}, si ha

$$|x_{j_0}| = \sqrt{x_{j_0}^2} \leq \sqrt{\sum_{j=1}^{d} x_j^2} \leq r_0.$$

Consideriamo ora una qualsiasi successione $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ in X. Considerata la successione $\{x_n^1\}_{n\in\mathbb{N}}$ delle prime coordinate, siccome questa successione è contenuta nell'intervallo $[-r_0, r_0]$, per il Teorema di Bolzano Weierstrass in dimensione $d = 1$, segue che, per $i = 1$

$$\text{esistono una sottosuccessione } \{x_{n_k}^i\}_{k\in\mathbb{N}} \text{ e un } y_i \in [-r_0, r_0] \text{ con } \lim_{k\to+\infty} x_{n_k}^i = y_i. \tag{4.6}$$

Non è restrittivo supporre che (4.6) sia vera direttamente per la successione $\{x_n^i\}_{n\in\mathbb{N}}$: basta sostituire $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ con $\{\mathbf{x}_{n_k}\}_{k\in\mathbb{N}}$. Che può allora assumere che (4.6) è vera per $i = 1, 2$. Non è restrittivo supporre che (4.6) sia vera per $i = 1, 2$ direttamente per la successione $\{x_n^i\}_{n\in\mathbb{N}}$. Ripetendo la procedura, possiamo concludere che esiste una sottosucessione $\{\mathbf{x}_{n_k}\}_{k\in\mathbb{N}}$ tale che (4.6) è vera per ogni $i = 1, ..., d$. Posto allora $\mathbf{y} = (y_1, ..., y_d)$, sege allora, si veda §3, che $\lim_{k\to+\infty} \mathbf{x}_{n_k} = \mathbf{y}$. Siccome X è chiuso per ipotesi, segue che $\mathbf{y}$. Abbiamo quindi dimostrato che X soddisfa (4.1).

□

Si noti che l'implicazione $\Longrightarrow$ è vera molto più in generale.

Esercizio 4.2. Sia (X, d) uno spazio metrico e $K \subseteq X$ un sottospazio compatto per successioni. Dimostrare che K è chiuso e limitato.

Esempio 4.3. L'implicazione $\Longleftarrow$ in generale è invece falsa. Ad esempio, se consideriamo l'insieme

$$BC^0(\mathbb{R}) := \{f \in C^0(\mathbb{R}) : \ f \text{ è limitata in } \mathbb{R}\} \tag{4.7}$$

che ha la norma

$$\|f\|_\infty = \sup\{|f(x)| : x \in \mathbb{R}\} \tag{4.8}$$

per la quale $BC^0(\mathbb{R})$ è uno spazio metrico completo, risulta che se ad esmpio consideriamo la funzione

$$f(x) = \begin{cases} 1 - 2|x| \text{ se } |x| \leq 1/2 \\ 0 \text{ se } |x| \geq 1/2 \end{cases}$$

e se consideriamo la successione di funzioni $\{f(\cdot - n)\}_{n\in\mathbb{N}}$, risulta che non esiste alcuna sottosuccessione di $\{f(\cdot - n)\}_{n\in\mathbb{N}}$ che sia convergente in $BC^0(\mathbb{R})$. D'altra parte $\{f(\cdot - n)\}_{n\in\mathbb{N}}$ è una successione limitata perché

$$\|f(\cdot - n)\|_\infty = \sup\{|f(x-n)| : x \in \mathbb{R}\} = \sup\{|f(x)| : x \in \mathbb{R}\} = \|f\|_\infty.$$

Sareste in grado di generare altri esempi? Sareste in grado di generare esempi in $BC^0(I)$ in un generico intervallo I?

Esercizio 4.4. Data una successione $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}^d$, scriviamo che

$$\lim_{n\to+\infty} \mathbf{x}_n = \infty \tag{4.9}$$

se vale quanto segue:

$$\forall\, r \in \mathbb{R} \quad \exists n_r \in \mathbb{R} \text{ t.c.} n > n_r \Longrightarrow |\mathbf{x}_n| > r. \tag{4.10}$$

Dimostrare allora quanto segue:

Per ogni successione $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}^d$ esistono una sottosuccessione $\{\mathbf{x}_{n_k}\}_{k\in\mathbb{N}}$ ed un $\mathbf{y} \in \mathbb{R}^d \cup \{\infty\}$ tali che

$$\lim_{k\to+\infty} \mathbf{x}_{n_k} = \mathbf{y}. \tag{4.11}$$

Enunciamo, senza dimostrazione, il seguente.

Theorem 4.5 (Weierstrass). *Sia $f \in C^0(X,\mathbb{R})$ con X uno spazio metrico non vuoto e compatto per successioni. Risulta che f ammette sia punti di massimo assoluto che punti di minimo assoluto.*

Ci sono molti modi per dimostrarlo. In particolare, si potrebbe ripetere il tipo di dimostrazione utilizzata in Analisi 1. □

Esercizio 4.6. Sia X un sottoinsieme chiuso e illimitato di $\mathbb{R}^d$. Dimostrare che se $f \in C^0(X,\mathbb{R})$ soddisfa $\lim\limits_{\mathbf{x}\to\infty} f(\mathbf{x}) = +\infty$ allora f ha punti di minimo assoluto. Se invece $f \in C^0(X,\mathbb{R})$ soddisfa $\lim\limits_{\mathbf{x}\to\infty} f(\mathbf{x}) = -\infty$ allora f ha punti di massimo assoluto.

Esercizio 4.7. Sia X un sottoinsieme chiuso e illimitato di $\mathbb{R}^d$. Sia $f \in C^0(X, \mathbb{R})$ con $\lim_{\mathbf{x}\to\infty} f(\mathbf{x}) = L \in \mathbb{R}$. Dimostrare che se esiste $x_0 \in X$ con $f(x_0) < L$ allora f ha punti di minimo assoluto. Se esiste $x_0 \in X$ con $f(x_0) > L$ allora f ha punti di massimo assoluto.

Osservazione 4.8. Il Teorema di Weierstrass è assolutamente cruciale in Matematica, ad esempio nella ricerca di soluzioni di Equazioni alle Derivate Parziali, dove spesso le soluzioni più importanti sono punti di massimo o di minimo di funzioni (ad esempio l'Energia) definite in spazi di funzioni quali, ad esempio $BC^0(\mathbb{R})$, ed a valori in $\mathbb{R}$. Per esempio, in modelli di rilevanza fisica, gli stati stabili di un sistema sono quelli ad energia minima, e spesso vengono chiamati Stati Fondamentali. Quindi, il problema dell'esistenza di Stati Fondamentali è equivalente al problema di dimostrare che esistono punti di minimo assoluto dell' Energia. Solo che Bolzano–Weierstrass ed il Teorema di Weierstrass non valgono in questi spazi di funzioni. Eppure la dimostrazione dell'esistenza di minimo assoluto non può che essere quella, essenzialmente, che avete studiato in Analisi 1. Ed infatti c'è spesso il modo di introdurre altre *topologie* (le topologie deboli) in questi spazi, per le quali invece Bolzano–Weierstrass vale ed in certi casi importanti si può dimostrare il Teorema di Weierstrass.

La ricerca di punti di massimo e di minimo di funzioni scalari è il nostro prossimo scopo.

5 Derivate

Definizione 5.1 (Derivata parziale)**.** Consideriamo una funzione $f : X \to \mathbb{R}^N$ con $X \subseteq \mathbb{R}^d$. Supponiamo che $\mathbf{x}^0 \in \mathring{X}$(:=parte interna di X). Allora diciamo che se il seguente limite esiste ed è finito,

$$\lim_{x_i \to x_i^0} \frac{f(x_1^0, ..., x_{i-1}^0, x_i, x_{i+1}^0, ..., x_d^0) - f(\mathbf{x}^0)}{x_i - x_i^0}. \tag{5.1}$$

esso è la i–esima derivata parziale di f in $\mathbf{x}^0$, e viene dotato con questi simboli

$$\frac{\partial}{\partial x_i} f(\mathbf{x}^0) = \partial_{x_i} f(\mathbf{x}^0) = \partial_i f(\mathbf{x}^0) = f_{x_i}(\mathbf{x}^0).$$

Esercizio 5.2. Mostrare che se

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_N(\mathbf{x}) \end{pmatrix}$$

allora $\frac{\partial}{\partial x_i} f(\mathbf{x}^0)$ esiste se e solo se $\frac{\partial}{\partial x_i} f_j(\mathbf{x}^0)$ esiste per tutti i $j = 1, ..., d$ e che in quel caso

$$\frac{\partial}{\partial x_i} f(\mathbf{x}^0) = \begin{pmatrix} \frac{\partial}{\partial x_i} f_1(\mathbf{x}^0) \\ \vdots \\ \frac{\partial}{\partial x_i} f_N(\mathbf{x}^0) \end{pmatrix}.$$

Esempio 5.3. Per $f(x,y) = \dfrac{\sin(xy)+y}{x^2+y^2}$ abbiamo

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

Definizione 5.4 (Gradiente)**.** Consideriamo una funzione $f : X \to \mathbb{R}$ con $X \subseteq \mathbb{R}^d$. Supponiamo in un punto $\mathbf{x}^0 \in \mathring{X}$ esistano tutte le derivate parziali $\partial_i f(\mathbf{x}^0)$ per $i = 1, ..., d$. Il gradiente di f nel punto $\mathbf{x}^0$ è il vettore

$$\nabla f(\mathbf{x}^0) = \mathrm{grad} f(\mathbf{x}^0) = \nabla f(\mathbf{x}^0) = (\partial_1 f(\mathbf{x}^0), ..., \partial_d f(\mathbf{x}^0)). \tag{5.2}$$

Più in generale, con una funzione $f : X \to \mathbb{R}^N$ con $X \subseteq \mathbb{R}^d$ ed $N \geq 2$, se in un punto $\mathbf{x}^0 \in \mathring{X}$ esistano tutte le derivate parziali

$$\partial_i f(\mathbf{x}^0) = \begin{pmatrix} \partial_i f_1(\mathbf{x}^0) \\ \vdots \\ \partial_i f_N(\mathbf{x}^0) \end{pmatrix} \text{ per } i = 1, ..., d, \tag{5.3}$$

invece del gradiente si parla della matrice Jacobiana

$$Df(\mathbf{x}^0) = Jf(\mathbf{x}^0) = \begin{pmatrix} \partial_1 f_1(\mathbf{x}^0) & \dots & \partial_d f_1(\mathbf{x}^0)) \\ \vdots & \vdots & \vdots \\ \partial_1 f_N(\mathbf{x}^0) & \dots & \partial_d f_N(\mathbf{x}^0) \end{pmatrix}. \tag{5.4}$$

Molto importante, quando $d = N$ è il determinante $\det Df(\mathbf{x}^0)$, che qualche volta viene chiamato Jacobiano.

Definizione 5.5. Data $f : X \to \mathbb{R}^N$ con $X \subseteq \mathbb{R}^d$ per ogni $\mathbf{y}_0 \in \mathbb{R}^N$ il corrispondente insieme di livello è

$$f^{-1}(\mathbf{y}_0) := \{\mathbf{x} \in X : f(\mathbf{x}) = \mathbf{y}_0\}. \tag{5.5}$$

Quando $N = 1$ e $d = 2$ si parla anche di curva di livello, anche se non è detto che (5.5) sia una curva.

Esempio 5.6. Consideriamo l'esempio banale $f(x,y) = x^2 + y^2$. L'insieme di definizione è tutto $\mathbb{R}^2$.

- Per $c < 0$ l'insieme di livello $x^2 + y^2 = c$ è vuoto
- L'insieme di livello $x^2 + y^2 = 0$ è esattamente l'origine.
- Per $c > 0$ l'insieme di livello $x^2 + y^2 = c$ è esattamente il cerchio di raggio $\sqrt{c}$ e centro l'origine.

Ovunque

$$\nabla f(x,y) = f_x(x,y)\overrightarrow{i} + f_x(x,y)\overrightarrow{j} = 2\left(x\overrightarrow{i} + y\overrightarrow{j}\right)$$

La cosa interessante, se si disegnano alcune curve di livello, è che ovunque il gradiente $\nabla f(x,y)$ è perpendicolare alla curva di livello per (x,y). Vedremo che questo fatto è vero in generale.

Esercizio 5.7. Verificare le seguenti affermazioni usando la definizione di derivata parziale:

- La funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita come $f(x,y) = x$ ammette derivate parziali ovunquee valgono
$$\frac{\partial f}{\partial x}(x,y) = 1\,, \quad \frac{\partial f}{\partial y}(x,y) = 0\,.$$
- La funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita come $f(x,y) = \sin(xy)$ ammette derivate parziali per ogni $(x,y) \in \mathbb{R}^2$ e valgono
$$\frac{\partial f}{\partial x}(x,y) = \cos(xy)y\,, \quad \frac{\partial f}{\partial y}(x,y) = \cos(xy)x\,.$$
- La funzione $f : (-1,+\infty) \times \mathbb{R} \to \mathbb{R}$ definita come $f(x,y) = \log(1+x)e^y$ ammette derivate parziali per ogni (x,y) del dominio e valgono
$$\frac{\partial f}{\partial x}(x,y) = \frac{e^y}{1+x}\,, \quad \frac{\partial f}{\partial y}(x,y) = \log(1+x)e^y\,.$$

Rispetto ad Analisi 1, ci sono delle novità, essenzialmente di natura geometrica, e non saranno le uniche.

Esempio 5.8. Consideriamo la funzione

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} \text{ se } (x,y) \neq (0,0) \\ 0 \text{ se } (x,y) = (0,0). \end{cases} \tag{5.6}$$

Abbiamo visto prima che $\lim_{(x,y)\to(0,0)} f(x,y)$ non esiste, quindi f non è continua in $(0,0)$. Eppure $f_x(0,0) = 0$ e $f_y(0,0) = 0$ esistono.

Un altro esempio, ancora più eloquente, è il seguente.

Esempio 5.9. Consideriamo, per costanti $a \neq b$, la funzione

$$f(x,y) = \begin{cases} a \text{ se } xy \neq 0 \\ b \text{ se } xy = 0. \end{cases}$$

Ovviamente, esibendo un salto in prossimità degli assi coordinati, sugli assi la funzione è discontinua. Eppure $f_x(0,0) = 0$ e $f_y(0,0) = 0$ esistono.

Una generalizzazione della nozione di derivata parziale è la seguente.

Definizione 5.10 (Derivata direzionale). Consideriamo una funzione $f : X \to \mathbb{R}^N$ con $X \subseteq \mathbb{R}^d$. Supponiamo che $\mathbf{x}^0 \in \mathring{X}$. Sia inoltre $\mathbf{v} \in \mathbb{R}^d$ un vettore. Diciamo che f ammette derivata direzionale nella direzione di $\mathbf{v}$ nel punto $\mathbf{x}^0$ se esiste a ed è finito,

$$\frac{d}{dt} f\left(\mathbf{x}^0 + t\mathbf{v}\right)\Big|_{t=0} = \lim_{t\to 0} \frac{f\left(\mathbf{x}^0 + t\mathbf{v}\right) - f(\mathbf{x}^0)}{t}. \tag{5.7}$$

Esempio 5.11. Rifacendoci all'Esempio 3.15, abbiamo

$$\lim_{x\to 0} \frac{f(x, mx)}{x} = m \lim_{x\to 0} \frac{\exp\left(-\frac{1}{x^2}\right)}{\exp\left(-\frac{2}{x^2}\right) + m^2 x^2} = 0$$

per ogni $m \in \mathbb{R}$. Quindi, le derivate direzionali in $(0,0)$ esistono, sono tutte nulle, con f non continua in $(0,0)$.

Esempio 5.12. Consideriamo la seguente funzione $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x,y) = \begin{cases} \left(\frac{x^2 y}{x^4 + y^2}\right)^2 & (x,y) \neq (0,0)\,, \\ 0 & (x,y) = (0,0)\,. \end{cases}$$

Per esercizio, verificare che le derivate direzionali sono nulle nell'origine. Consideriamo ora il versore $\boldsymbol{\nu} = (\cos\theta, \sin\theta)$ con $\theta \in [0, 2\pi)$. Possiamo calcolare

$$\begin{aligned} &\lim_{t\to 0} \frac{f(0 + t\boldsymbol{\nu}) - f(0)}{t} \\ =&\lim_{t\to 0} \frac{f(t\cos\theta, t\sin\theta) - f(0,0)}{t} \\ =&\lim_{t\to 0} \frac{1}{t} \frac{t^6 \cos^4\theta \sin^2\theta}{t^4(t^2\cos^4\theta + \sin^2\theta)^2} \\ =&\lim_{t\to 0} t \cdot \frac{\cos^4\theta \sin^2\theta}{(t^2\cos^4\theta + \sin^2\theta)^2} = 0 \end{aligned}$$

dove il termine nell'ultima frazione è limitato una volta fissato θ. Ne consegue che ogni derivata direzionale esiste ed è nulla nell'origine. Tuttavia questo non basta a garantire la continuità di f. Infatti f non è continua valendo

$$f(s, as^2) = \frac{a^2}{(1 + a^2)^2}$$

per ogni $a \in \mathbb{R}$ e $s \in \mathbb{R}$. In particolare la restrizione di f sulla parabola $y = x^2$ è la funzione costante $\frac{1}{4}$ e quindi esistono infiniti punti in un intorno dell'origine che assumono valori distanti da zero.

6 Il Differenziale

Definizione 6.1 (Differenziale)**.** Sia data una funzione $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ e un punto $\mathbf{x}^0 \in \mathring{X}$. Diremo che f è differenziabile in $\mathbf{x}^0$ se esiste un'applicazione lineare $L : \mathbb{R}^d \to \mathbb{R}^N$, cioè una matrice $N \times d$, tale che

$$\lim_{\substack{\mathbf{h}\to 0\\ \mathbf{h}\in\mathbb{R}^d}} \frac{f(\mathbf{x}^0+\mathbf{h}) - f(\mathbf{x}^0) - L\mathbf{h}}{|\mathbf{h}|} = 0 \in \mathbb{R}^N . \tag{6.1}$$

L'applicazione L si dice differenziale di f in $\mathbf{x}^0$ (verrà identificata meglio tra poco).

Osservazione 6.2. Si noti che (6.1) si può riscrivere equivalentemente scrivendo che si ha

$$f(\mathbf{x}^0+\mathbf{h}) - f(\mathbf{x}^0) = L\mathbf{h} + o(\mathbf{h}) \text{ dove } \lim_{\substack{\mathbf{h}\to 0\\ \mathbf{h}\in\mathbb{R}^d}} \frac{o(\mathbf{h})}{|\mathbf{h}|} = 0. \tag{6.2}$$

Esercizio 6.3. Data una funzione $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ e un punto $\mathbf{x}^0 \in \mathring{X}$, sia $N > 1$ e si scriva

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

Si dimostri che f è differenziabile in punto $\mathbf{x}^0$ se e solo se ciascuna delle funzioni f_j è differenziabile in $\mathbf{x}^0$. Inoltre, si dimostri che si ha

$$Df(\mathbf{x}^0)\mathbf{u} = \begin{pmatrix} Df_1(\mathbf{x}^0)\mathbf{u} \\ \vdots \\ Df_N(\mathbf{x}^0)\mathbf{u} \end{pmatrix} \text{ per tutti gli } \mathbf{u} \in \mathbb{R}^d. \tag{6.3}$$

Esercizio 6.4. Sia $f : \mathbb{R}^d \to \mathbb{R}^N$ definita da

$$f(\mathbf{x}) = A\mathbf{x} + \mathbf{y}^0 \tag{6.4}$$

dove $A : \mathbb{R}^d \to \mathbb{R}^N$ è un operatore lineare e $\mathbf{y}^0 \in \mathbb{R}^N$. Dimostrare che f è dappertutto differenziabile con $Df(\mathbf{x}) = A$ per ogni $\mathbf{x} \in \mathbb{R}^d$.

Esercizio 6.5. Consideriamo una mappa n–lineare $T : \mathbb{R}^{d_1} \times ... \times \mathbb{R}^{d_n} \to \mathbb{R}^N$ dove $n \geq 2$. Verificare che $DT(0) = 0$. Più in generale, calcolare, $DT(\mathbf{x}_1, ..., \mathbf{x}_n)$.

Definizione 6.6 (Spazio Tangente)**.** Data una funzione $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ e un punto $\mathbf{x}^0 \in \mathring{X}$ dove f è differenziabile con differenziale $L : \mathbb{R}^d \to \mathbb{R}^N$. Consideriamo il grafico di f

$$\Gamma_f = \{(\mathbf{x}, f(\mathbf{x})) \in \mathbb{R}^d \times \mathbb{R}^N : \mathbf{x} \in X\}. \tag{6.5}$$

Lo spazio tangente a Γ_f nel punto $(\mathbf{x}^0, \mathbf{y}^0)$ dove $\mathbf{y}^0 = f(\mathbf{x}^0)$ è lo spazio affine

$$T_{(\mathbf{x}^0,\mathbf{y}^0)}\Gamma_f := \{(\mathbf{x}, \mathbf{y}^0 + L\mathbf{v}) \in \mathbb{R}^d \times \mathbb{R}^N : \mathbf{v} \in \mathbb{R}^d\}. \tag{6.6}$$

Il seguente lemma chiarisce che la differenzabilità è una nozione diversa rispetto a quella della semplice esistenza delle derivate parziali.

Lemma 6.7. *Sia* $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ *differenziabile in un punto* $\mathbf{x}^0 \in \mathring{X}$. *Allora* f *è continua in* $\mathbf{x}^0$.

Dim. Abbiamo

$$f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) = \frac{f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) - L\mathbf{h}}{|\mathbf{h}|}|\mathbf{h}| + L\mathbf{h}.$$

Ora abbiamo (continuità in 0 segue da Esempio 3.7) che

$$\lim_{\substack{\mathbf{h}\to 0\\ \mathbf{h}\in\mathbb{R}^d}} L\mathbf{h} = L0 = 0.$$

D'altra parte, per la regola del prodotto nei limiti,

$$\lim_{\mathbf{h}\to 0} \frac{f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) - L\mathbf{h}}{|\mathbf{h}|}|\mathbf{h}| = \lim_{\mathbf{h}\to 0} \frac{f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) - L\mathbf{h}}{|\mathbf{h}|} \lim_{\mathbf{h}\to 0} |\mathbf{h}| = 0\ 0 = 0.$$

Di conseguenza, si ha quanto segue, cioè la continuità,

$$\lim_{\substack{\mathbf{h}\to 0\\ \mathbf{h}\in\mathbb{R}^d}} f(\mathbf{x}^0 + \mathbf{h}) = f(\mathbf{x}^0).$$

□

La differenziabilità di f in $\mathbf{x}^0$ è una nozione più forte di quella dell'esistenza delle derivate parziali. Infatti la differenziabilità in $\mathbf{x}^0$ implica l'esistenza della matrice Jacobiana in $\mathbf{x}^0$.

Lemma 6.8. *Sia* $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ *differenziabile in un punto* $\mathbf{x}^0 \in \mathring{X}$. *Allora il differenziale di* f *in* $\mathbf{x}^0$ *è la matrice Jacobiana in* $\mathbf{x}^0$*:*

$$L = \begin{pmatrix} \partial_1 f_1(\mathbf{x}^0) & \dots & \partial_d f_1(\mathbf{x}^0)) \\ \vdots & \vdots & \vdots \\ \partial_1 f_N(\mathbf{x}^0) & \dots & \partial_d f_N(\mathbf{x}^0) \end{pmatrix}.$$

Dim. Consideriamo la base canonica in $\mathbb{R}^d$ formata, al variare di $i \in \{1, ..., d\}$, dai versori

$$e_i = {}^t(0, ..., 0, \underset{i}{1}, 0, ..., 0). \tag{6.7}$$

Allora (6.1) implica

$$\lim_{h\to 0^+} \frac{f(\mathbf{x}^0 + he_i) - f(\mathbf{x}^0) - Lhe_i}{h} = 0$$

cioè

$$\lim_{h\to 0^+} \frac{f(\mathbf{x}^0 + he_i) - f(\mathbf{x}^0)}{h} = Le_i \tag{6.8}$$

Similmente

$$\lim_{h\to 0^-} \frac{f(\mathbf{x}^0 + he_i) - f(\mathbf{x}^0) - Lhe_i}{|h|} = 0 \Longleftrightarrow \lim_{h\to 0^-} \frac{f(\mathbf{x}^0 + he_i) - f(\mathbf{x}^0) - hLe_i}{-h} = 0$$
$$\Longleftrightarrow \lim_{h\to 0^-} \frac{f(\mathbf{x}^0 + he_i) - f(\mathbf{x}^0)}{-h} = -Le_i$$

implica

$$\lim_{h\to 0^-} \frac{f(\mathbf{x}^0 + he_i) - f(\mathbf{x}^0)}{h} = Le_i. \tag{6.9}$$

Ma allora (6.8)–(6.9) sono equivalenti a

$$\lim_{h\to 0} \frac{f(\mathbf{x}^0 + he_i) - f(\mathbf{x}^0)}{h} = Le_i. \tag{6.10}$$

Segue che $\partial_i f(\mathbf{x}^0) = Le_i$. Questa formula implica che la i–esima colonna di L (che è uguale a Le_i) coincide con (5.3). In particolare, siccome questo vale per ogni $i \in \{1, ..., d\}$, seque che la matrice Jacobiana (5.4) è ben definita e coincide con L. □

Osservazione 6.9. Si noti che per $d = N = 1$ l'esistenza della derivata $f'(x_0)$ è equivalente alla differenziabilità. Infatti, sappiamo che la derivata $f'(x_0)$, se esiste, è ciò che rende vero il sequente limite,

$$\lim_{x\to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0.$$

Ma questa è proprio la definizione di differenziabilità.

Abbiamo quanto segue.

Theorem 6.10 (Regola della Catena)**.** *Siano $f : X \subseteq \mathbb{R}^d \to Y \subseteq \mathbb{R}^N$ e $g : Y \subseteq \mathbb{R}^N \to \mathbb{R}^M$, $\mathbf{x}^0 \in \mathring{X}$, $\mathbf{y}^0 \in \mathring{Y}$ con $\mathbf{y}^0 = f(\mathbf{x}^0)$. Supponiamo che f sia differenziabile in $\mathbf{x}^0$ e che g sia differenziabile in $\mathbf{y}^0$. Allora la composizione $g \circ f$ è differenziabile in $\mathbf{x}_0$ con differenziale*

$$D(g \circ f)(\mathbf{x}_0) = Dg(\mathbf{y}_0)Df(\mathbf{x}_0) \ \textit{(prodotto riga colonna).} \tag{6.11}$$

Dim. Abbiamo

$$\begin{aligned}
&f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) = Df(\mathbf{x}_0)\mathbf{h} + o(\mathbf{h}) \text{ e} \\
&g\left(f\left(\mathbf{x}^0 + \mathbf{h}\right)\right) - g\left(f\left(\mathbf{x}^0\right)\right) = g\left(f\left(\mathbf{x}^0 + \mathbf{h}\right) - f(\mathbf{x}^0) + f(\mathbf{x}^0)\right) - g\left(f\left(\mathbf{x}^0\right)\right) \\
&= Dg(\mathbf{y}_0)\left(f\left(\mathbf{x}^0 + \mathbf{h}\right) - f(\mathbf{x}^0)\right) + o\left(f\left(\mathbf{x}^0 + \mathbf{h}\right) - f(\mathbf{x}^0)\right) \\
&= Dg(\mathbf{y}_0)\left(Df(\mathbf{x}_0)\mathbf{h} + o(\mathbf{h})\right) + o\left(Df(\mathbf{x}_0)\mathbf{h} + o(\mathbf{h})\right) \\
&= Dg(\mathbf{y}_0)Df(\mathbf{x}_0)\mathbf{h} + Dg(\mathbf{y}_0)o(\mathbf{h}) + o\left(Df(\mathbf{x}_0)\mathbf{h} + o(\mathbf{h})\right).
\end{aligned}$$

Siccome risulta, come dimostreremo subito, che

$$Dg(\mathbf{y}_0)o(\mathbf{h}) + o\left(Df(\mathbf{x}_0)\mathbf{h} + o(\mathbf{h})\right) = o(\mathbf{h}), \tag{6.12}$$

abbiamo ottenuto

$$g\left(f\left(\mathbf{x}^0 + \mathbf{h}\right)\right) - g\left(f\left(\mathbf{x}^0\right)\right) = Dg(\mathbf{y}_0)Df(\mathbf{x}_0)\mathbf{h} + o(\mathbf{h}),$$

che dà il risultato desiderato.

Ci resta da dimostrare (6.12) [1]. Per prima cosa abbiamo

$$\frac{1}{|\mathbf{h}|}|Dg(\mathbf{y}_0)o(\mathbf{h})| \leq |Dg(\mathbf{y}_0)| \, \frac{|o(\mathbf{h})|}{|\mathbf{h}|} \xrightarrow{\mathbf{h}\to 0} 0$$

che dimostra $Dg(\mathbf{y}_0)o(\mathbf{h}) = o(\mathbf{h})$. Per finire, ci resta da dimostrare,

$$o_1\left(Df(\mathbf{x}_0)\mathbf{h} + o_2(\mathbf{h})\right) = o_3(\mathbf{h}), \tag{6.13}$$

dove ho messo gli indici per distinguere le tre o piccole.

(6.13) è equivalente a dimostrare che

$$\forall\ \epsilon > 0\ \exists\ \delta_\epsilon > 0 \text{ t.c. } |\mathbf{h}| \leq \delta_\epsilon \Longrightarrow |o_1\left(Df(\mathbf{x}_0)\mathbf{h} + o_2(\mathbf{h})\right)| \leq \epsilon|\mathbf{h}|. \tag{6.14}$$

Per o_j con $j = 1, 2$ abbiamo

$$\forall\ \epsilon > 0\ \exists\ \delta_\epsilon^{(j)} > 0 \text{ t.c. } |\mathbf{v}| \leq \delta_\epsilon^{(j)} \Longrightarrow |o_j\left(\mathbf{v}\right)| \leq \epsilon|\mathbf{v}|,$$

dove possiamo sempre assumere $\delta_\epsilon^{(1)} \leq 1$. Per

$$|\mathbf{h}| < \delta^{(2)}_{2^{-1}\delta_\epsilon^{(1)}} \tag{6.15}$$

abbiamo

$$\begin{aligned}
&|Df(\mathbf{x}_0)\mathbf{h} + o_2(\mathbf{h})| \leq |Df(\mathbf{x}_0)\mathbf{h}| + |o_2(\mathbf{h})| \leq |Df(\mathbf{x}_0)| \ |\mathbf{h}| + |o_2(\mathbf{h})| \\
&\leq |Df(\mathbf{x}_0)| \ |\mathbf{h}| + \frac{\delta_\epsilon^{(1)}}{2} |\mathbf{h}| \,.
\end{aligned}$$

Se oltre a (6.15) chiediamo anche

$$|\mathbf{h}| \leq \min\left\{1, \frac{\delta_\epsilon^{(1)}}{2\,|Df(\mathbf{x}_0)|}\right\} \tag{6.16}$$

[1] Piccola pillola di *saggezza*: in Analisi Matematica le luci della ribalta, ad esempio nelle presentazioni ai convegni, illuminano i termini principali, ma il lavoro vero è nei resti.

concludiamo

$$|Df(\mathbf{x}_0)\mathbf{h} + o_2(\mathbf{h})| \le |Df(\mathbf{x}_0)| \ |\mathbf{h}| + \frac{\delta_\epsilon^{(1)}}{2} |\mathbf{h}| \le \delta_\epsilon^{(1)}$$

e questo implica

$$\begin{aligned}|o_1 (Df(\mathbf{x}_0)\mathbf{h} + o_2(\mathbf{h}))| \le \epsilon |Df(\mathbf{x}_0)\mathbf{h} + o_2(\mathbf{h})| &\le \epsilon \left(|Df(\mathbf{x}_0)| + \frac{\delta_\epsilon^{(1)}}{2}\right) |\mathbf{h}| \\ &\le \epsilon (|Df(\mathbf{x}_0)| + 1) |\mathbf{h}| .\end{aligned}$$

Abbiamo cioè dimostrato che se per ogni $\epsilon > 0$ consideriamo un

$$0 < \delta_\epsilon^{(3)} < \min \left\{ \delta^{(2)}_{2^{-1}\delta_\epsilon^{(1)}}, 1, \frac{\delta_\epsilon^{(1)}}{2 |Df(\mathbf{x}_0)| + 1} \right\}$$

allora risulta

$$|\mathbf{h}| \le \delta_\epsilon^{(3)} \Longrightarrow |o_1 (Df(\mathbf{x}_0)\mathbf{h} + o_2(\mathbf{h})) | \le \epsilon (|Df(\mathbf{x}_0)| + 1) |\mathbf{h}| .$$

Un δ_ϵ che renda vera la proposizione (6.14) si ottiene ponendo

$$\delta_\epsilon := \delta^{(3)}_{\frac{\epsilon}{|Df(\mathbf{x}_0)|+1}} .$$

□

Corollario 6.11. *Siano $g : Y \subseteq \mathbb{R}^d \to \mathbb{R}^N$ differenziabile in un punto $\mathbf{y}^0 \in \mathring{Y}$ e $\mathbf{v} \in \mathbb{R}^d$ un vettore non nullo. Allora g ammette derivata direzionale nella direzione di $\mathbf{v}$ nel punto $\mathbf{x}^0$ e vale l'uguaglianza*

$$\left.\frac{d}{dt} g \left(\mathbf{y}^0 + t\mathbf{v}\right)\right|_{t=0} = Dg(\mathbf{y}_0)\mathbf{v}. \tag{6.17}$$

Dim. $\mathbf{y}^0 \in \mathring{Y}$ significa che esiste un $r_0 > 0$ tale che $D(\mathbf{y}^0, r_0) \subseteq Y$. Se poniamo $f(t) := \mathbf{y}^0 + t\mathbf{v}$, risulta che

$$|f(t) - \mathbf{y}^0| = |t\mathbf{v}| < r_0 \text{ per } |t| < \frac{r_0}{|\mathbf{v}|}.$$

Pertanto abbiamo $f : \left(-\frac{r_0}{|\mathbf{v}|}, \frac{r_0}{|\mathbf{v}|}\right) \to X$. Questa funzione è differenziabile in $t = 0$ con differenziale dato da

$$\left.\frac{d}{dt} f(t)\right|_{t=0} = \left(\mathbf{y}^0 + t\mathbf{v}\right)\big|_{t=0} = \mathbf{v}.$$

D'altra parte, dal Teorema 6.10 anche la composizione $t \to g\left(\mathbf{y}^0 + t\mathbf{v}\right) = g \circ f(t)$ è differenziabile in $t = 0$ con differenziale dato esattamente dalla formula (6.17). □

A questo punto, ci serve un criterio di differenziabilità. Il seguente criterio, basato sul Teorema di Lagrange in Analisi 1, è fondamentale.

Theorem 6.12. *Sia data una funzione $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ e $\mathbf{x}^0 \in \mathring{X}$. Se f ammette derivate parziali rispetto a tutte le variabili in un intorno di $\mathbf{x}^0$ e se queste sono tutte continue in $\mathbf{x}^0$ allora f è differenziabile in $\mathbf{x}^0$.*

Dim. Si può dimostrare il teorema ragionando sulle componenti f_k, quindi senza perdere in generalità si può porre $N = 1$. Si tratta allora di dimostrare che

$$f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) - \nabla f(\mathbf{x}^0)\mathbf{h} = o(|\mathbf{h}|). \tag{6.18}$$

Abbiamo

$$\begin{aligned}
&f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) = \\
&f(x_1^0 + h_1, x_2^0 + h_2, ..., x_d^0 + h_d) - f(x_1^0, x_2^0 + h_2, ..., x_d^0 + h_d) + f(x_1^0, x_2^0 + h_2, ..., x_d^0 + h_d) - f(x_1^0, x_2^0, ..., x_d^0) \\
&= \sum_{j=1}^{d} \big(f(x_1, ..., x_{j-1}^0, x_j^0 + h_j, ..., x_d^0 + h_d) - f(x_1^0, ..., x_j^0, x_{j+1}^0 + h_{j+1}, ..., x_d^0 + h_d)\big) \\
&= \sum_{j=1}^{d} \partial_j f(x_1^0, ..., x_{j-1}^0, x_j^0 + t_j h_j, ..., x_d^0 + h_d) h_j \text{ dove } t_j \in (0, 1).
\end{aligned}$$

Di conseguenza,

$$\begin{aligned}
f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) - \nabla f(\mathbf{x}^0)\mathbf{h} &= f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) - \sum_{j=1}^{d} \partial_j f(x_1^0, ..., x_d^0) h_j \\
= \sum_{j=1}^{d} \underbrace{\left[\partial_j f(x_1^0, ..., x_{j-1}^0, x_j^0 + t_j h_j, ..., x_d^0 + h_d) - \partial_j f(x_1^0, ..., x_d^0)\right]}_{o(1)} h_j,
\end{aligned}$$

dove ho degli $o(1)$ per la continuità delle derivate parziali in $\mathbf{x}^0$.
Quindi

$$f(\mathbf{x}^0 + \mathbf{h}) - f(\mathbf{x}^0) - \nabla f(\mathbf{x}^0)\mathbf{h} = \sum_{j=1}^{d} o(1) h_j.$$

Il termine a destra è un $o(|\mathbf{h}|)$ poiché

$$\frac{1}{|\mathbf{h}|}\left|\sum_{j=1}^{d} o(1) h_j\right| \leq \sum_{j=1}^{d} |o(1)| \frac{|h_j|}{|\mathbf{h}|} \leq \sum_{j=1}^{d} |o(1)| = o(1).$$

Quindi la

$$\lim_{\mathbf{h}\to 0} \frac{1}{|\mathbf{h}|}\left|\sum_{j=1}^{d} o(1) h_j\right| = 0$$

segue dal teorema dei carabinieri e (6.18). □

Definizione 6.13. Sia data una funzione $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ con X aperto. Essa si dice di classe C^1 in X se esistono tutte le derivate parziali in X e queste sono tutte continue in X. Con tale affermazione chiediamo che per ogni $k = 1, \ldots, d$ la funzione

$$\partial_k f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N , \quad \text{tale che } \mathbf{x} \mapsto \partial_k f(\mathbf{x})$$

sia continua in X. L'insieme delle funzioni $X \to \mathbb{R}^N$ che sono C^1 in X viene denotato con $C^1(X, \mathbb{R}^N)$. Per $N = 1$ si scrive $C^1(X)$.

Osservazione 6.14. Se f è di classe C^1 in X allora f è differenziabile in X, quindi f è continua in X.

Esercizio 6.15. Dimostrare che per un $X \subseteq \mathbb{R}^d$ aperto, una funzione f appartiene a $C^1(X, \mathbb{R}^N)$ se e solo se è differenziabile su tutto X e se $Df \in C^0(X, \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N))$.

Esercizio 6.16. Dimostrare che la funzione

$$F(x) = \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

è differenziabile in $x = 0$ ma non è di classe C^1 in un intorno di zero. Analogamente, dimostrare che, la funzione $G : \mathbb{R}^2 \to \mathbb{R}$ tale che $G(x, y) = F(x)$, è differenziabile in $(0, 0)$ ma non è di classe C^1 in un intorno di $(0, 0)$.

Definizione 6.17. Sia data una funzione $f : X \subseteq \mathbb{R}^d \to \mathbb{R}^N$ e sia $\mathbf{x}^0 \in \mathring{X}$ un punto dove f è differenziabile. Allora diciamo che $\mathbf{x}^0$ è un punto critico di f se $Df(\mathbf{x}^0) = 0$.

Theorem 6.18 (Fermat)**.** *Sia data una funzione $f : X \subseteq \mathbb{R}^d \to \mathbb{R}$ e sia $\mathbf{x}^0 \in \mathring{X}$ un punto dove f è differenziabile. Supponiamo inoltre che $\mathbf{x}^0$ sia un punto di massimo o di minimo locale per f. Allora $Df(\mathbf{x}^0) = 0$.*

Dim. Assumeremo il caso $d = 1$, che è già noto. $\mathbf{x}^0 \in \mathring{X}$ significa che esiste un $r_0 > 0$ tale che $D(\mathbf{x}^0, r_0) \subseteq X$. Consideriamo ora un generico $\mathbf{v} \in \mathbb{R}^d$ non nullo. Abbiamo già visto che la funzione

$$\left(-\frac{r_0}{|\mathbf{v}|}, \frac{r_0}{|\mathbf{v}|} \right) \to f\left(\mathbf{x}^0 + t\mathbf{v}\right) \in \mathbb{R} \tag{6.19}$$

è ben definita ed è derivabile in $t = 0$ con derivata

$$\left.\frac{d}{dt} f\left(\mathbf{x}^0 + t\mathbf{v}\right)\right|_{t=0} = Df(\mathbf{x}^0)\mathbf{v}. \tag{6.20}$$

Siccome è facile concludere che il punto $t = 0$ è di massimo o di minimo locale per la funzione in (6.19), dal Teorema di Fermat nel caso $d = 1$ possiamo concludere che la derivata in (6.20) è nulla:

$$Df(\mathbf{x}^0)\mathbf{v} = 0 \text{ per ogni } \mathbf{v} \in \mathbb{R}^d.$$

Questo implica $Df(\mathbf{x}^0) = 0$. $\square$

Esempio 6.19. Consideriamo la funzione $f(x,y) = xe^{-x^2-y^2}$.

Qui $f(x,y)$ è differenziabile visto che

$$f_x(x,y) = (1-2x^2)e^{-x^2-y^2}$$
$$f_y(x,y) = -2xye^{-x^2-y^2}$$

sono funzioni continue. Notare che la funzione soddisfa $f(-x,y) = -f(x,y)$ e $f(x,-y) = f(x,y)$, quindi quando disegnamo gli insiemi di livello dovremmo ottenere una figura simmetrica rispetto agli assi coordinati. Notare che $f(x,y) > 0$ se $x > 0$, $f(x,y) = 0$ se $x = 0$ (quindi l'asse y è l'insieme di livello 0) e $f(x,y) < 0$ se $x < 0$.

Notare che

$$\lim_{(x,y)\to\infty} f(x,y) = 0.$$

Ossia quando ci allontaniamo verso l'infinito, ci avviciniamo sempre più all'altitudine del livello del mare. Siccome la funzione è sia positiva che negativa, da qualche parte c'è un qualche punto di massima altezza ed un qualche punto di minima altezza, cioè punti di massimo e di minimo assoluto. Per trovarli poniamo

$$(1-2x^2)e^{-x^2-y^2} = 0$$
$$-2xye^{-x^2-y^2} = 0$$

ossia

$$1-2x^2 = 0$$
$$xy = 0.$$

Dalla prima equazione ricaviamo che le soluzioni del sistema sono della forma (x_0, y_0) con $x_0 = \pm\frac{1}{\sqrt{2}}$. Dalla seconda equazione ricaviamo che $y_0 = 0$. In effetti le soluzioni del sistema sono i due punti critici $\left(\pm\frac{1}{\sqrt{2}}, 0\right)$. Siccome c'e' almeno un punto di minimo assoluto ed almeno un punto di massimo assoluto, che non coincidono tra loro, uno di questi punti critici è il punto di massimo assoluto mentre l'altro è il punto di minimo assoluto. Siccome $f\left(\frac{1}{\sqrt{2}}, 0\right) > 0 > f\left(-\frac{1}{\sqrt{2}}, 0\right)$ evidentemente $\left(\frac{1}{\sqrt{2}}, 0\right)$ è il punto di massimo assoluto mentre $\left(-\frac{1}{\sqrt{2}}, 0\right)$ è il punto di minimo assoluto.

Esempio 6.20. Sia $f(x,y) = \dfrac{x+y}{1+x^2+y^2}$

Qui $f(x,y)$ è differenziabile visto che

$$f_x(x,y) = \frac{(1+x^2+y^2)-(x+y)2x}{(1+x^2+y^2)^2} = \frac{1-x^2-2xy+y^2}{(1+x^2+y^2)^2}$$
$$f_y(x,y) = \frac{(1+x^2+y^2)-(x+y)2y}{(1+x^2+y^2)^2} = \frac{1+x^2-2xy-y^2}{(1+x^2+y^2)^2}$$

Notare che $f(x,y) > 0$ se $x+y > 0$, $f(x,y) = 0$ se $x+y = 0$ (quindi la diagonale del secondo e quarto quadrante é l'insieme di livello 0) e $f(x,y) < 0$ se $x+y < 0$.

Notare che

$$\lim_{(x,y)\to\infty} f(x,y) = 0.$$

Perciò esistono punti di massimo e di minimo assoluto. Per trovarli poniamo

$$\begin{aligned} 1 - x^2 - 2xy + y^2 &= 0 \\ 1 + x^2 - 2xy - y^2 &= 0. \end{aligned}$$

Sommando, sottraendo e semplificando osserviamo che il sistema si riduce a

$$\begin{aligned} xy &= \frac{1}{2} \\ x^2 - y^2 &= 0. \end{aligned}$$

Dalla prima equazione ricaviamo che i punti critici (x_0, y_0) sono tali che x_0 e y_0 sono non nulli ed hanno lo stesso segno, mentre dalla seconda equazione abbiamo $|x_0| = |y_0|$. Quindi, mettendo le due cose insieme, abbiamo $x_0 = y_0$. Sostituendo nella prima equazione otteniamo $x^2 = 1/2$, da cui si ottengono i due punti critici $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ e $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$. Siccome c'e' almeno un punto di minimo assoluto ed almeno un punto di massimo assoluto, che non coincidono tra loro, uno di questi punti critici è il punto di massimo assoluto mentre l'altro è il punto di minimo assoluto. Siccome $f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) > 0 > f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ evidentemente $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ è il punto di massimo assoluto mentre $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ è il punto di minimo assoluto.

Esempio 6.21. Sia $f(x,y) = (y + x^2)e^{-x^2-y^2}$

Qui $f(x,y)$ è differenziabile con derivate parziali

$$\begin{aligned} f_x(x,y) &= (2x - 2x(y+x^2))e^{-x^2-y^2} = 2x(1 - x^2 - y)e^{-x^2-y^2} \\ f_y(x,y) &= (1 - 2x^2y - 2y^2)e^{-x^2-y^2}. \end{aligned}$$

L'insieme di livello 0 è la parabola $y + x^2 = 0$ ed inoltre

$$\lim_{(x,y)\to\infty} f(x,y) = 0.$$

Siccome la funzione è sia positiva che negativa, da qualche parte c'è un punto di massimo assoluto ed un punto di minimo assoluto. Per trovarli poniamo

$$\begin{aligned} x(1 - x^2 - y) &= 0 \\ 1 - 2x^2y - 2y^2 &= 0. \end{aligned}$$

Guardiamo alla prima equazione $x(1-x^2-y)=0$ e poniamo $x=0$. Sostituendo nella seconda otteniamo $y=\pm\frac{1}{\sqrt{2}}$, e quindi abbiamo i due punti critici

$$\left(0,\frac{1}{\sqrt{2}}\right), \quad \left(0,-\frac{1}{\sqrt{2}}\right). \tag{6.21}$$

Cerchiamo punti critici con $x\neq 0$. Dunque consideriamo il sistema

$$\begin{aligned}&1-x^2-y=0\\&1-2x^2y-2y^2=1-2(x^2+y)y=0.\end{aligned}$$

Il sistema è equivalente a

$$\begin{aligned}&1-x^2-y=0\\&1-2y=0.\end{aligned}$$

Pertanto otteniamo i punti critici

$$\left(\frac{1}{\sqrt{2}},\frac{1}{2}\right), \quad \left(-\frac{1}{\sqrt{2}},\frac{1}{2}\right). \tag{6.22}$$

Quindi abbiamo in totale 4 punti critici. Calcoliamo

$$\begin{aligned}f\left(0,-\frac{1}{\sqrt{2}}\right)&=-\frac{1}{\sqrt{2}}e^{-\frac{1}{2}}\\f\left(0,\frac{1}{\sqrt{2}}\right)&=\frac{1}{\sqrt{2}}e^{-\frac{1}{2}}\\f\left(\pm\frac{1}{\sqrt{2}},\frac{1}{2}\right)&=e^{-\frac{1}{4}-\frac{1}{2}}.\end{aligned}$$

Necessariamente $\left(0,-\frac{1}{\sqrt{2}}\right)$ è il punto di minimo assoluto. Inoltre,

$$f\left(0,\frac{1}{\sqrt{2}}\right)<f\left(\pm\frac{1}{\sqrt{2}},\frac{1}{2}\right)$$

poiché

$$\frac{1}{\sqrt{2}}<e^{-\frac{1}{4}}\Leftrightarrow\sqrt{2}>e^{\frac{1}{4}}\Leftrightarrow 4>e=2,7...$$

Quindi $\left(\pm\frac{1}{\sqrt{2}},\frac{1}{2}\right)$ sono i punti di massimo assoluto. Senza fare nemmeno un conto, semplicemente visualizzando come possano essere fatte le curve di livello, si capisce che $\left(0,\frac{1}{\sqrt{2}}\right)$ non può essere ne' un massimo ne' un minimo locale, ma presumibilmente è una sella (il punto più basso di un passo di montagna che unisce le due cime), come del resto verificheremo dopo, con uno strumento in più.

Esempio 6.22. Sia $f(x,y) = x^4 + y^4 - 2x^2y^2 + 2x^2$

$f(x,y)$ è differenziabile visto che

$$f_x(x,y) = 4x^3 - 4xy^2 + 4x$$
$$f_y(x,y) = 4y^3 - 4x^2y$$

sono funzioni continue.

Conviene sempre controllare se vi sia

$$\lim_{(x,y)\to\infty} f(x,y) = ?$$

Osserviamo che

$$f(x,y) = x^4 + y^4 - 2x^2y^2 + 2x^2 = (x^2 - y^2)^2 + 2x^2. \tag{6.23}$$

Da qui risulta $f(0,0) = 0$ e $f(x,y) \geq 0$ con $f(x,y) > 0$ per $(x,y) \neq (0,0)$. Quindi la (6.23) implica che $(0,0)$ è l'unico minimo assoluto. La (6.23) implica anche

$$\lim_{(x,y)\to\infty} f(x,y) = +\infty, \tag{6.24}$$

che ci dice che la funzione non ha punti di massimo assoluto mentre la (6.23) ci dice che $(0,0)$ è l'unico minimo assoluto. Per cercare eventuali altri punti critici risolviamo

$$x(x^2 - y^2 + 1) = 0$$
$$y(y^2 - x^2) = 0$$

Guardiamo alla prima equazione, $x(x^2 - y^2 + 1) = 0$. Se poniamo $x = 0$, e se sostituiamo nella $y(y^2 - x^2) = 0$, otteniamo $y^3 = 0$. Pertanto $x = 0$ implica $y = 0$. Se poi guardiamo alla $y(y^2 - x^2) = 0$ e poniamo $y = 0$, sostituendo in $x(x^2 - y^2 + 1) = 0$ otteniamo $x(x^2 + 1) = 0$ che implica $x = 0$. In conclusione, $(0,0)$ è una soluzione del sistema, come del resto sapevamo già. Eventuali altri punti critici dovranno essere tali che $x \neq 0$ ed $y \neq 0$ e dovranno risolvere

$$x^2 - y^2 = -1$$
$$x^2 - y^2 = 0.$$

Ovviamente quest'ultimo sistema non ha soluzioni e pertanto l'unico punto critico di $f(x,y)$ è $(0,0)$.

7 Derivate parziali di ordine superiore

Da qui in poi consideriamo $f : A \subseteq \mathbb{R}^d \to \mathbb{R}$ definita su A aperto. Nel caso di funzioni con codominio $\mathbb{R}^N$ si ragiona analogamente sulle componenti.

Supponiamo che, per un certo indice $i \in \{1, \dots, N\}$, esista per ogni $\mathbf{x} \in A$ la derivata parziale $\dfrac{\partial f}{\partial x_i}$. Quindi possiamo considerare la funzione

$$\frac{\partial f}{\partial x_i} : A \subseteq \mathbb{R}^d \to \mathbb{R}\,.$$

Di questa funzione potrei calcolare le derivate parziali di indice $j \in \{1, \dots, d\}$ nei punti di A, ovvero $\dfrac{\partial}{\partial x_j}\left(\dfrac{\partial f}{\partial x_i}\right)$.

Definizione 7.1. Si dice derivata parziale del secondo ordine di f in $\mathbf{x}^0 \in A$ rispetto alle variabili x_i e x_j, la derivata parziale j-esima della funzione $\frac{\partial f}{\partial x_i}$ e le denoteremo con

$$\text{der. parz. del sec. ord. } \textit{mista}: \qquad \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j \partial x_i}\,, \qquad i \neq j$$

$$\text{der. parz. del sec. ord. } \textit{pura}: \qquad \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_i^2}$$

Iterando questo procedimento possiamo definire, dove possibile, derivate parziali di ordine k di f in $\mathbf{x}^0 \in A$ rispetto alle variabili $x_{i_1}, x_{i_2}, \dots, x_{i_k}$:

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_2} \partial x_{i_1}}\,.$$

Esempio 7.2. Consideriamo la funzione $f(x, y) = y^2\ \cos x$ e calcoliamo le derivate parziali fino al terzo ordine.

$$\frac{\partial f}{\partial x}(x,y) = -y^2\ \sin x\,, \qquad \frac{\partial f}{\partial y}(x,y) = 2y\ \cos x\,,$$

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = -y^2\ \cos x\,, \qquad \frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = -2y\ \sin x\,,$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = -2y\ \sin x\,, \qquad \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = 2\ \cos x\,,$$

$$\frac{\partial^3 f}{\partial x^3}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x}\right) = y^2 \sin x\,, \qquad \frac{\partial^3 f}{\partial y^3}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y}\right) = 0\,,$$

$$\frac{\partial^3 f}{\partial x \partial x \partial y}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial}{\partial x}\left(-2y\ \sin x\right) = -2y \cos x$$

$$\frac{\partial^3 f}{\partial x \partial y \partial x}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial}{\partial x}\left(-2y\ \sin x\right) = -2y \cos x$$

$$\frac{\partial^3 f}{\partial y \partial x \partial x}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x^2}\right) = \frac{\partial}{\partial y}\left(-y^2\ \cos x\right) = -2y \cos x$$

$$
\begin{aligned}
\frac{\partial^3 f}{\partial x \partial y \partial y}(x,y) &= \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y^2}\right) = \frac{\partial}{\partial x}(2\ \cos x) = -2\ \sin x \\
\frac{\partial^3 f}{\partial y \partial x \partial y}(x,y) &= \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial}{\partial y}(-2y\ \sin x) = -2\ \sin x \\
\frac{\partial^3 f}{\partial y \partial y \partial x}(x,y) &= \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial}{\partial y}(-2y\ \sin x) = -2\ \sin x
\end{aligned}
$$

Notiamo che nel precedente esempio le derivate miste assumono lo stesso valore indipendentemente dall'ordine con cui facciamo le derivate. Tuttavia importa quante volte deriviamo rispetto ad una variabile piuttosto che rispetto ad un'altra. **Questa non è una regola sempre valida** come illustra il prossimo esempio.

Esempio 7.3. La funzione

$$
f(x,y) = \begin{cases} xy\ \frac{x^4 - y^2}{x^4 + y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}
$$

ammette le derivate parziali

$$
\frac{\partial f}{\partial x}(x,y) = \begin{cases} y\ \frac{x^4 - y^2}{x^4 + y^2} + \frac{8x^4 y^3}{(x^4+y^2)^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}
$$

$$
\frac{\partial f}{\partial y}(x,y) = \begin{cases} x\ \frac{x^4 - y^2}{x^4 + y^2} - \frac{4x^5 y^2}{(x^4+y^2)^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}
$$

Verificare i calcoli per esercizio, specialmente nell'origine usando il comportamento di f lungo gli assi. A questo punto possiamo calcolare le derivate seconde nell'origine e trovare:

$$
\frac{\partial^2 f}{\partial x^2}(0,0) = 0\,, \frac{\partial^2 f}{\partial y^2}(0,0) = 0\,, \frac{\partial^2 f}{\partial y \partial x}(0,0) = -1\,, \frac{\partial^2 f}{\partial x \partial y}(0,0) = 1\,.
$$

I calcoli si fanno sempre considerando le restrizioni lungo gli assi, ma stavolta dobbiamo considerare le funzioni $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$. Le prime due affermazioni seguono dal fatto che $\frac{\partial f}{\partial x}(x,0) = 0$ per ogni $x \in \mathbb{R}$ e $\frac{\partial f}{\partial y}(0,y) = 0$ per ogni $y \in \mathbb{R}$. Le seconde invece dal fatto che $\frac{\partial f}{\partial x}(0,y) = -y$ per ogni $y \in \mathbb{R}$ e $\frac{\partial f}{\partial y}(x,0) = x$ per ogni $x \in \mathbb{R}$. Mostro i dettagli dell'ultimo caso, scrivere esplicitamente gli altri per esercizio:

$$
\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) = \lim_{h \to 0} \frac{\frac{\partial f}{\partial y}(h,0) - \frac{\partial f}{\partial y}(0,0)}{h} = \lim_{h \to 0} \frac{h - 0}{h} = 1\,.
$$

Esempio 7.4. Più in generale, consideriamo una funzione $\phi : \mathbb{R}^2 \backslash \{(0,0)\} \to \mathbb{R}$ che ammetta derivate prime ovunque in $\mathbb{R}^2 \backslash \{(0,0)\}$. Supponiamo inoltre che

$$
\lim_{x \to 0} \phi(x,0) = a \text{ e } \lim_{y \to 0} \phi(0,y) = b \text{ con } a \neq b.
$$

Si noti che la funzione $\phi(x,y) = \dfrac{x^4 - y^2}{x^4 + y^2}$ dell'Esempio 7.3 soddisfa le condizioni, ma infatti anche, $\phi(x,y) = \dfrac{x^2 - y^2}{x^2 + y^2}$, $\phi(x,y) = \dfrac{x^4 - y^4}{x^4 + y^4}$ ecc., sono possibili esempi.
Come in Esempio 7.3, definiamo

$$f(x,y) = \begin{cases} xy\ \phi(x,y) & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

E' immediato, da $f(x,0) \equiv 0$ e $f(0,y) \equiv 0$, che $\partial_x f(x,0) \equiv 0$ e $\partial_y f(0,y) \equiv 0$. Abbiamo, inoltre,

$$\begin{aligned} \partial_y f(x,0) &= x\ \phi(x,0) + x\ y|_{y=0}\ \partial_y \phi(x,0) = x\ \phi(x,0) \\ \partial_x f(0,y) &= y\ \phi(0,y) + y\ x|_{x=0}\ \partial_x \phi(0,y) = y\ \phi(0,y). \end{aligned}$$

Pertanto

$$\begin{aligned} \partial_x \partial_y f(0,0) &= \lim_{x\to 0} \frac{\partial_y f(x,0) - \partial_y f(0,0)}{x} = \lim_{x\to 0} \frac{x\ \phi(x,0) - 0}{x} = \lim_{x\to 0} \phi(x,0) = a \\ \partial_y \partial_x f(0,0) &= \lim_{y\to 0} \frac{\partial_x f(0,y) - \partial_x f(0,0)}{x} = \lim_{y\to 0} \frac{y\ \phi(0,y) - 0}{y} = \lim_{y\to 0} \phi(0,y) = b. \end{aligned}$$

Da qui si vede che $\partial_x \partial_y f(0,0) = a \neq b = \partial_y \partial_x f(0,0)$.

In realtà, i casi in cui le derivate miste dipendono dall'ordine di derivazione, sono "patologici". C'è tutta una classe di teoremi, che possiamo archiviare sotto il titolo di Teorema di Schwartz, che dimostrano che, in molti casi importanti, il valore della derivata non dipende dall'ordine di differenziazione. Sotto consideriamo un caso particolare di Teorema di Schwartz. Un'altra versione la si può trovare in Giusti [2, pagina 15].

Definizione 7.5 (Doppia differenziabilità)**.** Consideriamo una funzione $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$, dove A è un aperto. Supponiamo che f sia differenziabile su tutto A e consideriamo la funzione $Df : A \subseteq \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$, dove $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$ è lo spazio degli operatori lineari definiti in $\mathbb{R}^d$ a valori in $\mathbb{R}^N$, che noi identifichiamo con lo spazio delle matrici $N \times d$. Sia $\mathbf{x}^0 \in A$. Diremo che f è due volte differenziabile in $\mathbf{x}^0$ se Df è differenziabile in $\mathbf{x}^0$ e scriveremo $D^2 f(\mathbf{x}^0) = D(Df)(\mathbf{x}^0)$.

Osservazione 7.6. Ricordiamo che per $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$ si ha

$$Df(\mathbf{x}) = \begin{pmatrix} \partial_1 f_1(\mathbf{x}) & \dots & \partial_d f_1(\mathbf{x})) \\ \vdots & \vdots & \vdots \\ \partial_1 f_N(\mathbf{x}) & \dots & \partial_d f_N(\mathbf{x}) \end{pmatrix}$$

dove $\partial_i f_j(\mathbf{x})$ sono le componenti della funzione a valori vettoriali $Df(\mathbf{x})$. Allora, la differenziabilità di Df in $\mathbf{x}^0$ è equivalente alla differenziabilità di tutte le componenti $\partial_i f_j$. Inoltre, se Df è differenziabile in $\mathbf{x}^0$, allora esistono tutte le derivate parziali $\partial_k \partial_i f_j$ (o, equivalentemente, $\partial_k \partial_i f$) in $\mathbf{x}^0$. Inoltre, se tutte le derivate parziali $\partial_k \partial_i f_j$ sono definite in un intorno di $\mathbf{x}^0$ e sono continue in $\mathbf{x}^0$, allora Df è differenziabile in $\mathbf{x}^0$.

Definizione 7.7 (Matrice Hessiana)**.** Consideriamo una funzione $f : A \subseteq \mathbb{R}^d \to \mathbb{R}$, dove A è un aperto. Supponiamo che f ammetta tutte le derivate parziali su tutto A. Allora resta definito il gradiente $\nabla f : A \subseteq \mathbb{R}^d \to \mathbb{R}^d$

$$\nabla f(\mathbf{x}) = (\partial_1 f(\mathbf{x}), \partial_2 f(\mathbf{x}), ..., \partial_d f(\mathbf{x})).$$

Supponiamo che tutte le derivate parziali di ∇f siano a loro volta definite in un punto $\mathbf{x}^0 \in A$. Questo equivale a dire che tutte le derivate parziali di $\partial_j f$ siano a loro volta definite in $\mathbf{x}^0$ per ogni $j = 1, ..., d$. Resta pertanto definita la seguente matrice, detta la Matrice Hessiana di f in $\mathbf{x}^0$:

$$Hf(\mathbf{x}^0) = (\partial_i \partial_j f(\mathbf{x}_0))_{i,j=1}^d = \begin{pmatrix} \partial_1^2 f(\mathbf{x}^0) & \partial_1\partial_2 f(\mathbf{x}^0) & \cdots & \partial_1\partial_d f(\mathbf{x}^0) \\ \partial_2\partial_1 f(\mathbf{x}^0) & \partial_2^2 f(\mathbf{x}^0) & \cdots & \partial_2\partial_d f(\mathbf{x}^0) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_d\partial_1 f(\mathbf{x}^0) & \partial_d\partial_2 f(\mathbf{x}^0) & \cdots & \partial_d^2 f(\mathbf{x}^0) \end{pmatrix}. \tag{7.1}$$

Esercizio 7.8. Verificare che se $f : A \subseteq \mathbb{R}^d \to \mathbb{R}$ è differenziabile due volte in un punto $\mathbf{x}^0 \in A$ allora la matrice Hessiana (7.1) è definita in $\mathbf{x}_0$.

Esempio 7.9. Sia $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$ differenziabile ovunque in A (quindi resta definita la funzione $Df : A \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N))$ e due volte differenziabile in un punto $\mathbf{x}^0 \in A$. Consideriamo un vettore $\mathbf{v} \in \mathbb{R}^d$ e la funzione

$$A \ni \mathbf{x} \to g(\mathbf{x}) := Df(\mathbf{x})\mathbf{v} \in \mathbb{R}^N \tag{7.2}$$

Risulta che (qui ricorda che $D^2 f(\mathbf{x}^0) \in \mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N))$ e che $D^2 f(\mathbf{x}^0)\mathbf{u} \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N))$

$$g \text{ è differenziabile in } \mathbf{x}^0 \text{ con } Dg(\mathbf{x}^0)\mathbf{u} = \left(D^2 f(\mathbf{x}^0)\mathbf{u}\right)\mathbf{v} \text{ per ogni } \mathbf{u} \in \mathbb{R}^d. \tag{7.3}$$

Per dimostrare (7.3) osserviamo che $g = G \circ Df$ con G la funzione in (3.10) data da

$$\mathcal{L}(\mathbb{R}^d, \mathbb{R}^N) \ni L \to L\mathbf{v} \in \mathbb{R}^N.$$

G è differenziabile ovunque con $DG = G$ ovunque. Inoltre, dalla regola della catena, otteniamo

$$\begin{aligned} Dg(\mathbf{x}^0)\mathbf{u} &= GD^2 f(\mathbf{x}^0)\mathbf{u} = G\left(D^2 f(\mathbf{x}^0)\mathbf{u}\right) = \left(D^2 f(\mathbf{x}^0)\mathbf{u}\right)\mathbf{v} && (7.4) \\ &=: D^2 f(\mathbf{x}^0)(\mathbf{u}, \mathbf{v}). && (7.5) \end{aligned}$$

Vedremo infatti in Teorema 7.10 che $\left(D^2 f(\mathbf{x}^0)\mathbf{u}\right)\mathbf{v} = \left(D^2 f(\mathbf{x}^0)\mathbf{v}\right)\mathbf{u}$ per ogni coppia di vettori $\mathbf{u}$ e $\mathbf{v}$.

Theorem 7.10 (Teorema di Schwartz)**.** *Sia $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$, differenziabile in A con $Df \in C^0(A, \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N))$ e differenziabile due volte in un punto $\mathbf{x}^0 \in A$. Allora la funzione bilineare $D^2 f(\mathbf{x}_0) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^N$ è simmetrica, cioè*

$$\left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v} = \left(D^2 f(\mathbf{x}_0)\mathbf{v}\right)\mathbf{u} \text{ per ogni coppia } (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^d \times \mathbb{R}^d. \tag{7.6}$$

Dim. Abbiamo $D(\mathbf{x}_0, r) \subseteq A$ per un opportuno $r > 0$. Consideriamo ora

$$\mathbf{u}, \mathbf{v} \in \mathbb{R}^d \text{ con } |\mathbf{u}| < r/2 \text{ e } |\mathbf{v}| < r/2 \tag{7.7}$$

e consideriamo la funzione

$$[-1, 1] \ni t \to f(\mathbf{x}_0 + \mathbf{u} + t\mathbf{v}) \in \mathbb{R}^N.$$

Introduciamo la funzione

$$g(t) := f(\mathbf{x}_0 + \mathbf{u} + t\mathbf{v}) - f(\mathbf{x}_0 + t\mathbf{v})$$

Abbiamo

$$\begin{aligned} \left|g(1) - g(0) - g'(0)\right| &= \left|\int_0^1 \big(g'(t) - g'(0)\big)\, dt\right| \le \int_0^1 |g'(t) - g'(0)| dt \\ &\le \sup_{t \in [0,1]} |g'(t) - g'(0)|. \end{aligned} \tag{7.8}$$

Dalla regola della catena, Teorema 6.10, abbiamo

$$\begin{aligned} g'(t) &= Df(\mathbf{x}_0 + \mathbf{u} + t\mathbf{v})\mathbf{v} - Df(\mathbf{x}_0 + t\mathbf{v})\mathbf{v} \\ &= (Df(\mathbf{x}_0 + \mathbf{u} + t\mathbf{v}) - Df(\mathbf{x}_0))\,\mathbf{v} - (Df(\mathbf{x}_0 + t\mathbf{v}) - Df(\mathbf{x}_0))\,\mathbf{v} \\ &= \big(Df(\mathbf{x}_0 + \mathbf{u} + t\mathbf{v}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0)(\mathbf{u} + t\mathbf{v})\big)\,\mathbf{v} \\ &- \big(Df(\mathbf{x}_0 + t\mathbf{v}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0) t\mathbf{v}\big)\,\mathbf{v} \\ &+ \underbrace{\big(D^2 f(\mathbf{x}_0)(\mathbf{u} + t\mathbf{v})\big)\,\mathbf{v} - \big(D^2 f(\mathbf{x}_0) t\mathbf{v}\big)\,\mathbf{v}}_{(D^2 f(\mathbf{x}_0)\mathbf{u})\mathbf{v}}. \end{aligned}$$

Concludendo

$$\begin{aligned} g'(t) - \big(D^2 f(\mathbf{x}_0)\mathbf{u}\big)\,\mathbf{v} &= \big(Df(\mathbf{x}_0 + \mathbf{u} + t\mathbf{v}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0)(\mathbf{u} + t\mathbf{v})\big)\,\mathbf{v} \\ &- \big(Df(\mathbf{x}_0 + t\mathbf{v}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0) t\mathbf{v}\big)\,\mathbf{v}. \end{aligned} \tag{7.9}$$

Ricordiamoci che la doppia differenziabilità in $\mathbf{x}^0$ implica che

$$\frac{Df(\mathbf{x}_0 + \mathbf{h}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0)\mathbf{h}}{|\mathbf{h}|} = o(1) \Longleftrightarrow |Df(\mathbf{x}_0 + \mathbf{h}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0)\mathbf{h}| \le |o(1)|\ |\mathbf{h}|.$$

In particolare, abbiamo

$$\forall\ \epsilon > 0 \text{ esiste } \delta_\epsilon > 0 \text{ t.c. } |\mathbf{h}| < \delta_\epsilon \Rightarrow |Df(\mathbf{x}_0 + \mathbf{h}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0)\mathbf{h}| < \epsilon |\mathbf{h}|.$$

Allora, per $0 < r < \delta_\epsilon$ in (7.7), per (7.9) concludiamo

$$\begin{aligned} &|g'(t) - \big(D^2 f(\mathbf{x}_0)\mathbf{u}\big)\,\mathbf{v}| \\ &\le \big|Df(\mathbf{x}_0 + \mathbf{u} + t\mathbf{v}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0)(\mathbf{u} + t\mathbf{v})\big|\, |\mathbf{v}| \\ &+ \big|Df(\mathbf{x}_0 + t\mathbf{v}) - Df(\mathbf{x}_0) - D^2 f(\mathbf{x}_0) t\mathbf{v}\big|\, |\mathbf{v}| \\ &< \epsilon |\mathbf{u} + t\mathbf{v}||\mathbf{v}| + \epsilon |t\mathbf{v}||\mathbf{v}| \le \epsilon |\mathbf{v}|\, (|\mathbf{u} + t\mathbf{v}| + t|\mathbf{v}|) \\ &\le \epsilon |\mathbf{v}|\, (|\mathbf{u}| + t|\mathbf{v}| + t|\mathbf{v}|) \le 2\epsilon |\mathbf{v}|(|\mathbf{u}| + |\mathbf{v}|). \end{aligned}$$

Pertanto, da (7.8) concludiamo

$$\begin{aligned}
&\left|g(1) - g(0) - \left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v}\right| \le \left|g(1) - g(0) - g'(0)\right| + \left|g'(0) - \left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v}\right| \\
&\le \sup_{t\in[0,1]} |g'(t) - g'(0)| + 2\epsilon|\mathbf{v}|(|\mathbf{u}| + |\mathbf{v}|) \\
&\le \sup_{t\in[0,1]} \left(\left|g'(t) - \left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v}\right| + \left|g'(0) - \left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v}\right|\right) + 2\epsilon|\mathbf{v}|(|\mathbf{u}| + |\mathbf{v}|) \\
&\le 6\epsilon|\mathbf{v}|(|\mathbf{u}| + |\mathbf{v}|).
\end{aligned}$$

In conclusione abbiamo dimostrato

$$\begin{aligned}
&\left|\left[\underbrace{f(\mathbf{x}_0 + \mathbf{u} + \mathbf{v}) - f(\mathbf{x}_0 + \mathbf{v})}_{g(1)} - \underbrace{f(\mathbf{x}_0 + \mathbf{u}) + f(\mathbf{x}_0)}_{g(0)}\right] - \left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v}\right| \\
&\le 6\epsilon|\mathbf{v}|(|\mathbf{u}| + |\mathbf{v}|).
\end{aligned}$$

Ma siccome la quantità nella parentesi quadrata è simmetrica in $(\mathbf{u}, \mathbf{v})$, scambiando di ruolo $\mathbf{u}$ e $\mathbf{v}$ resta egualmente dimostrato

$$\left|[f(\mathbf{x}_0 + \mathbf{u} + \mathbf{v}) - f(\mathbf{x}_0 + \mathbf{v}) - f(\mathbf{x}_0 + \mathbf{u}) + f(\mathbf{x}_0)] - \left(D^2 f(\mathbf{x}_0)\mathbf{v}\right)\mathbf{u}\right| \le 6\epsilon|\mathbf{u}|(|\mathbf{u}| + |\mathbf{v}|).$$

Dalla disuguaglianza triangolare, le ultime due disuguaglianze implicano

$$\begin{aligned}
&\left|\left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v} - \left(D^2 f(\mathbf{x}_0)\mathbf{v}\right)\mathbf{u}\right| \\
&= \left|-\left(g(1) - g(0) - \left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v}\right) + \left(g(1) - g(0) - \left(D^2 f(\mathbf{x}_0)\mathbf{v}\right)\right)\mathbf{u}\right| \\
&\le \left|g(1) - g(0) - \left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v}\right| + \left|g(1) - g(0) - \left(D^2 f(\mathbf{x}_0)\mathbf{v}\right)\mathbf{u}\right| \\
&\le 6\epsilon|\mathbf{v}|(|\mathbf{u}| + |\mathbf{v}|) + 6\epsilon|\mathbf{u}|(|\mathbf{u}| + |\mathbf{v}|) \\
&= 6\epsilon(|\mathbf{u}| + |\mathbf{v}|)^2 \ \forall\ (\mathbf{u}, \mathbf{v}) \text{ che soddisfa (7.7) per un } 0 < r < \delta_\epsilon.
\end{aligned}$$

Quindi, in conclusione

$$\left|\left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v} - \left(D^2 f(\mathbf{x}_0)\mathbf{v}\right)\mathbf{u}\right| \le 6\epsilon(|\mathbf{u}| + |\mathbf{v}|)^2 \ \forall\ (\mathbf{u}, \mathbf{v}) \text{ che soddisfa (7.7) per un } 0 < r < \delta_\epsilon.$$

Siccome l'ultima disuguaglianza coinvolge in entrambi i membri funzioni omogenee di ordine 2, la disuguaglianza è vera per tutte le coppie $(\mathbf{u}, \mathbf{v})$. Pertanto concludiamo che

$$\left|\left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v} - \left(D^2 f(\mathbf{x}_0)\mathbf{v}\right)\mathbf{u}\right| \le 6\epsilon(|\mathbf{u}| + |\mathbf{v}|)^2 \ \forall\ (\mathbf{u}, \mathbf{v}) \text{ e } \forall\ \epsilon > 0.$$

Ma allora questo implica che il termine a sinistra è uguale a 0, concludiamo

$$\left(D^2 f(\mathbf{x}_0)\mathbf{u}\right)\mathbf{v} = \left(D^2 f(\mathbf{x}_0)\mathbf{v}\right)\mathbf{u} \ \forall\ (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^d \times \mathbb{R}^d.$$

$\square$

Corollario 7.11. *Sia $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$ differenziabile in A e differenziabile due volte in un punto $\mathbf{x}^0 \in A$. Allora*

$$\partial_i \partial_j f(\mathbf{x}^0) = \partial_j \partial_i f(\mathbf{x}^0) \text{ per ogni coppia di indici.} \tag{7.10}$$

Dim. Consideriamo $Df : A \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$. Per la regola della catena (6.17)

$$\frac{d}{dt} f(\mathbf{x} + t\mathbf{v})\Big|_{t=0} = Df(\mathbf{x})\mathbf{v} = \sum_{j=1}^{d} \partial_j f(\mathbf{x}) v_j.$$

Sempre per la regola della catena,

$$\frac{d}{dt} Df(\mathbf{x}^0 + t\mathbf{u})\mathbf{v}\Big|_{t=0} = \left(D^2 f(\mathbf{x}^0)\mathbf{u}\right)\mathbf{v} = \sum_{j=1}^{d} (D\partial_j f(\mathbf{x}^0)\mathbf{u}) v_j =$$
$$\sum_{i,j=1}^{d} \partial_i \partial_j f(\mathbf{x}^0) u_i v_j.$$

Scambiando di ruolo u e v abbiamo

$$\frac{d}{dt} Df(\mathbf{x}^0 + t\mathbf{v})\mathbf{u}\Big|_{t=0} = \left(D^2 f(\mathbf{x}^0)\mathbf{v}\right)\mathbf{u} = \sum_{j=1}^{d} (D\partial_j f(\mathbf{x}^0)\mathbf{v}) u_j =$$
$$\sum_{i,j=1}^{d} \partial_i \partial_j f(\mathbf{x}^0) v_i u_j.$$

Siccome da (7.6) abbiamo $\left(D^2 f(\mathbf{x}^0)\mathbf{u}\right)\mathbf{v} = \left(D^2 f(\mathbf{x}^0)\mathbf{v}\right)\mathbf{u}$ segue

$$\sum_{i,j=1}^{d} \partial_i \partial_j f(\mathbf{x}^0) u_i v_j = \partial_i \partial_j f(\mathbf{x}^0) v_i u_j \text{ per ogni coppia } (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Scegliendo $\mathbf{u} = e_a$ e $\mathbf{v} = e_b$, ricaviamo $\partial_a \partial_b f(\mathbf{x}^0) = \partial_b \partial_a f(\mathbf{x}^0)$, per ogni scelta di $a, b \in \{1, ..., d\}$. □

Una condizione sufficiente per garantire la doppia differenziabilità di una funzione in un punto è la seguente.

Definizione 7.12 (Differenziabilità di ordine k)**.** Consideriamo una funzione $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$, dove A è un aperto. Supponiamo di avere definito la nozione di differenziabilità di ordine $k-1$, e sia il differenziale di ordine $k-1$ definito su tutto A, $D^{k-1}f : A \to \mathcal{L}^{k-1}(\mathbb{R}^d, \mathbb{R}^N)$. Se quest'ultima funzione è a sua volta differenziabile in un punto $\mathbf{x}^0$, poniamo

$$D^k f(\mathbf{x}^0) := D(D^{k-1}f)(\mathbf{x}^0) \in \mathcal{L}(\mathbb{R}^d, \mathcal{L}^{k-1}(\mathbb{R}^d, \mathbb{R}^N)) \sim \mathcal{L}^k(\mathbb{R}^d, \mathbb{R}^N). \tag{7.11}$$

Chiamiamo $D^k f(\mathbf{x}^0)$ il differenziale di ordine k di f nel punto $\mathbf{x}^0$ e diciamo che f è differenziabile fino all'ordine k nel punto $\mathbf{x}^0$.

Lemma 7.13. *Sia $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$ differenziabile $k-1$ volte ovunque in A e k volte differenziabile in un punto $\mathbf{x}^0 \in A$. Allora risulta che f ammette in x^0 tutte le derivate parziali di ordine k e che*

$$D^k f(\mathbf{x}^0)(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k) = \sum_{i_1,...,i_k=1}^{d} \partial_{i_1}...\partial_{i_k} f(\mathbf{x}^0) v_{1i_1}...v_{ki_k}. \tag{7.12}$$

In particolare, vale

$$\partial_{i_1}...\partial_{i_k} f(\mathbf{x}_0) = D^k f(\mathbf{x}^0)(e_{i_1}, e_{i_2}, ..., e_{i_k}) \text{ per ogni scelta di indici.} \tag{7.13}$$

Dim. Consideriamo degli arbitrari vettori $\mathbf{v}_2, ..., \mathbf{v}_k \in \mathbb{R}^d$ e la funzione

$$A \ni \mathbf{x} \to g(\mathbf{x}) := D^{k-1} f(\mathbf{x})(\mathbf{v}_2, ..., \mathbf{v}_k) \in \mathbb{R}^N. \tag{7.14}$$

Supponiamo per induzione che

$$D^{k-1} f(\mathbf{x})(\mathbf{v}_2, ..., \mathbf{v}_k) = \sum_{i_2,...,i_k=1}^{d} \partial_{i_2}...\partial_{i_k} f(\mathbf{x}) v_{2i_2}...v_{ki_k} \text{ dove} \tag{7.15}$$

$$\mathbf{v}_i = \begin{pmatrix} v_{i1} \\ \vdots \\ v_{id} \end{pmatrix}.$$

Si noti che (7.15) implica subito

$$D^{k-1} f(\mathbf{x})(e_{i_2}, ..., e_{i_k}) = \partial_{i_2}...\partial_{i_k} f(\mathbf{x}). \tag{7.16}$$

Vogliamo dimostrare che

$$g \text{ è differenziabile in } \mathbf{x}^0 \text{ con } Dg(\mathbf{x}^0)\mathbf{u} = \left(D^k f(\mathbf{x}^0)\mathbf{u}\right)(\mathbf{v}_2, ..., \mathbf{v}_k) \text{ per ogni } \mathbf{u} \in \mathbb{R}^d. \tag{7.17}$$

Per dimostrare (7.17) osserviamo che $g = G \circ D^{k-1} f$ con G la funzione data da

$$G : \mathcal{L}^{k-1}(\mathbb{R}^d, \mathbb{R}^N) \ni L \to L(\mathbf{v}_2, ..., \mathbf{v}_k) \in \mathbb{R}^N.$$

G è una funzione lineare, pertanto differenziabile ovunque e con ovunque $DG = G$. Dalla regola della catena, otteniamo

$$\begin{aligned} &Dg(\mathbf{x}^0)\mathbf{u} = D\left(G \circ D^{k-1} f(\mathbf{x}^0)\right)\mathbf{u} = G\left(DD^{k-1} f(\mathbf{x}^0)\mathbf{u}\right) \\ &= \left(DD^{k-1} f(\mathbf{x}^0)\mathbf{u}\right)(\mathbf{v}_2, ..., \mathbf{v}_k) = \left(D^k f(\mathbf{x}^0)\mathbf{u}\right)(\mathbf{v}_2, ..., \mathbf{v}_k) =: D^k f(\mathbf{x}^0)(\mathbf{u}, \mathbf{v}_2, ..., \mathbf{v}_k) \end{aligned} \tag{7.18}$$

(In Teorema 7.15 verificheremo che è una funzione k–lineare simmetrica).

Ora se scegliamo

$$g(x) = D^{k-1}f(\mathbf{x})(e_{i_2}, ..., e_{i_k}) = \partial_{i_2}...\partial_{i_k}f(\mathbf{x}),$$

applicando (7.18) abbiamo

$$Dg(\mathbf{x}^0)\mathbf{u} = D\left(\partial_{i_2}...\partial_{i_k}f\right)(\mathbf{x}^0)\mathbf{u} = \sum_{i_1=1}^{d} \partial_{i_1}\partial_{i_2}...\partial_{i_k}f(\mathbf{x}^0)u_{i_1}$$

E pertanto, se ora ritorno alla $g(x)$ in (7.14), ho quanto segue, dove la prima uguaglianza è conseguenza di (7.18), mentre dalla seconda uguaglianza in poi, si usano proprietà di linearità della differenziazione, che vengono lasciate per esercizio,

$$\begin{aligned} D^k f(\mathbf{x}^0)(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k) &= Dg(\mathbf{x}^0)\mathbf{v}_1 = D\left(\sum_{i_2,...,i_k=1}^{d} \partial_{i_2}...\partial_{i_k}f(\mathbf{x})v_{2i_2}...v_{ki_k}\right)\Bigg|_{\mathbf{x}=\mathbf{x}^0} \mathbf{v}_1 \\ &= \sum_{i_2,...,i_k=1}^{d} D(\partial_{i_2}...\partial_{i_k}f(\mathbf{x})v_{2i_2}...v_{ki_k})|_{\mathbf{x}=\mathbf{x}_0}\,\mathbf{v}_1 \\ &= \sum_{i_1,...,i_k=1}^{d} \partial_{i_1}...\partial_{i_k}f(\mathbf{x}_0)v_{1i_1}...v_{ki_k}. \end{aligned}$$

Si noti in particolare che da (7.12) segue (7.13). □

Osservazione 7.14. Il Lemma 7.13 implica che, nelle ipotesi assegnate, esistono su A tutte le derivate parziali di ordine $k-1$ e che esse sono tutte differenziabili in $\mathbf{x}^0$.

Theorem 7.15 (Teorema di Schwartz)**.** *Sia $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$ con A aperto differenziabile $k-1$ volte in A con $D^{k-1}f \in C^0(A, \mathcal{L}^{k-1}(\mathbb{R}^d, \mathbb{R}^N))$ e k volte differenziabile in un punto $\mathbf{x}^0 \in A$. Allora la funzione k–lineare $D^k f(\mathbf{x}_0) : \underbrace{\mathbb{R}^d \times ... \times \mathbb{R}^d}_{k\ volte} \to \mathbb{R}^N$ è simmetrica, cioè*

$$D^k f(\mathbf{x}_0)(\mathbf{u}_1, ..., \mathbf{u}_k) = D^k f(\mathbf{x}_0)(\mathbf{u}_{\sigma_1}, ..., \mathbf{u}_{\sigma_k})\ \forall\ \textit{permutazione}\ (\sigma_1, ..., \sigma_k)\ \textit{di}\ (1, ..., k). \tag{7.19}$$

Dim. Il caso $k = 2$ è già noto. Supponiamo per induzione che il caso $k-1 \geq 2$ sia stato già dimostrato e consideriamo la funzione

$$g(\mathbf{x}) := D^{k-2}f(\mathbf{x})(\mathbf{u}_3, ..., \mathbf{u}_k) \text{ per } \mathbf{u}_3, ..., \mathbf{u}_k \text{ fissati.}$$

Da (7.12), abbiamo

$$Dg(\mathbf{x})\mathbf{u}_2 := D^{k-1}f(\mathbf{x})(\mathbf{u}_2, \mathbf{u}_3, ..., \mathbf{u}_k)$$

e, di nuovo da (7.12), abbiamo

$$D^2 g(\mathbf{x}^0)(\mathbf{u}_1, \mathbf{u}_2) = D^k f(\mathbf{x}^0)(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, ..., \mathbf{u}_k),$$

dove dal Teorema 7.10 abbiamo $D^2g(\mathbf{x}^0)(\mathbf{u}_1,\mathbf{u}_2) = D^2g(\mathbf{x}^0)(\mathbf{u}_2,\mathbf{u}_1)$, il che implica

$$D^k f(\mathbf{x}^0)(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,...,\mathbf{u}_k) = D^k f(\mathbf{x}^0)(\mathbf{u}_2,\mathbf{u}_1,\mathbf{u}_3,...,\mathbf{u}_k). \tag{7.20}$$

D'altra parte, per induzione

$$D^{k-1}f(\mathbf{x})(\mathbf{u}_2,\mathbf{u}_3,...,\mathbf{u}_k) = D^{k-1}f(\mathbf{x})(\mathbf{u}_{\sigma_2},...,\mathbf{u}_{\sigma_k}) \ \forall \text{ permutazione } (\sigma_2,...,\sigma_k) \text{ di } (2,...,k).$$

Pertanto

$$D^k f(\mathbf{x}^0)(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,...,\mathbf{u}_k) = D^k f(\mathbf{x}^0)(\mathbf{u}_1,\mathbf{u}_{\sigma_2},...,\mathbf{u}_{\sigma_k}) \ \forall \text{ permutazione } (\sigma_2,...,\sigma_k) \text{ di } (2,...,k). \tag{7.21}$$

Allora (7.19) segue da (7.20) e (7.21). □

Definizione 7.16. Consideriamo una funzione $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$, dove A è un aperto. Diremo che f è di classe C^k su A (con $k \in \mathbb{N}$) se tutte le derivate parziali di ordine k esistono e sono continue in A. In particolare denoteremo con $C^0(A,\mathbb{R}^N)$ lo spazio delle funzioni continue in A, con $C^k(A,\mathbb{R}^N)$ lo spazio delle funzioni di classe C^k e

$$C^\infty(A,\mathbb{R}^N) := \bigcap_{k\geq 1} C^k(A,\mathbb{R}^N)$$

l'insieme avente derivate parziali di qualsiasi ordine continue. In particolare

$$C^0(A,\mathbb{R}^N) \subsetneqq C^1(A,\mathbb{R}^N) \subsetneqq C^2(A,\mathbb{R}^N) \subsetneqq \cdots \subsetneqq C^k(A,\mathbb{R}^N) \subsetneqq \cdots \subsetneqq C^\infty(A,\mathbb{R}^N)\,.$$

Vale la seguente estensione dell'Osservazione 7.6

Theorem 7.17. *Sia data una funzione $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$ con A aperto e sia $\mathbf{x}^0 \in A$. Supponiamo che $f \in C^{k-1}(A,\mathbb{R}^N)$ e che tutte le derivate parziali di ordine k rispetto a tutte le variabili siano definite in un intorno di $\mathbf{x}^0$ e siano tutte continue in $\mathbf{x}^0$. Allora f è differenziabile k volte in $\mathbf{x}^0$.*

Dim.(Sketch) Come l'Osservazione 7.6, è una elementare conseguenza del Teorema 6.12. Intanto, grazie all'Osservazione 7.6, possiamo assumere il caso $k = 2$. Inoltre, per induzione, possiamo supporre che abbiamo $D^{k-1}f : A \to \mathcal{L}^{k-1}(\mathbb{R}^d,\mathbb{R}^N)$. Inoltre, da (7.12)–(7.13) abbiamo

$$D^{k-1}f(\mathbf{x})(\mathbf{v}_2,...,\mathbf{v}_k) = \sum_{i_2,...,i_k=1}^{d} \partial_{i_2}...\partial_{i_k} f(\mathbf{x}) v_{2i_1}...v_{ki_k}$$

$$\partial_{i_2}...\partial_{i_k} f(\mathbf{x}) = D^{k-1}f(\mathbf{x})(e_{i_2},...,e_{i_k}) \text{ per ogni scelta di indici.}$$

E' facile concludere che la funzione $D^{k-1}f$ è una funzioni a valori vettoriali con componenti

$$\partial_{i_2}...\partial_{i_k} f(\mathbf{x}) \text{ con } i_2,...,i_k = 1,...d. \tag{7.22}$$

Siccome per ipotesi ciascuna di queste funzioni ammette derivate parziali prime in un intorno di $\mathbf{x}^0$ che sono tutte continue in $\mathbf{x}^0$, segue dal Teorema 6.12 che tutte le funzioni in (7.22) sono differenziabili in $\mathbf{x}^0$. E quindi anche la funzione $D^{k-1}f$ è differenziabile in $\mathbf{x}^0$. Cioè f è differenziabile k volte in $\mathbf{x}^0$. □

Esercizio 7.18. Verificare che $f \in C^k(A, \mathbb{R}^N)$ se e solo se f è differenziabile fino all'ordine k con $D^k f \in C^k(A, \mathcal{L}^k(\mathbb{R}^d, \mathbb{R}^N))$.

Osservazione 7.19. Consideriamo $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$ con A aperto, $\mathbf{x}^0 \in A$, $D(\mathbf{x}^0, r_0) \subseteq A$ e sia $\mathbf{h} \in \mathbb{R}^d$ un vettore non nullo. Sia f differenziabile $k-1$ volte in $D(\mathbf{x}^0, r_0)$ e differenziabile k volte in $\mathbf{x}^0$. Sappiamo che $t \to f\left(\mathbf{x}^0 + t\mathbf{h}\right)$ è ben definita in $I := \left(-\frac{r_0}{|\mathbf{h}|}, \frac{r_0}{|\mathbf{h}|}\right)$. Abbiamo

$$\left(\frac{d}{dt}\right)^k f\left(\mathbf{x}^0 + t\mathbf{h}\right)\bigg|_{t=0} = D^k f(\mathbf{x}^0)\mathbf{h}^k \tag{7.23}$$

$$\text{dove } \mathbf{h}^k := \underbrace{(\mathbf{h}, ..., \mathbf{h})}_{k \text{ volte}} \tag{7.24}$$

Infatti, per $k = 1$ questo è discusso e dimostrato nel Corollario 6.11,ed è facile estendere per ogni $t \in I$, se f è differenziabile su tutto A. Supponiamo che (7.23) sia vera per $k-1$. E' facile estendere per ogni $t \in I$. Allora abbiamo

$$\begin{aligned}
&\frac{d}{dt}\left(\left(\frac{d}{dt}\right)^{k-1} f\left(\mathbf{x}^0 + t\mathbf{h}\right)\right)\bigg|_{t=0} = \frac{d}{dt}\left(D^{k-1} f(\mathbf{x}^0 + t\mathbf{h})\mathbf{h}^{k-1}\right)\bigg|_{t=0} \\
&= D\left(D^{k-1} f(\mathbf{x}^0)\mathbf{h}^{k-1}\right)\mathbf{h} = D^k f(\mathbf{x}^0)\mathbf{h}^k
\end{aligned}$$

dove nella penultima uguaglianza abbiamo applicato l'ipotesi che (7.23) sia vera per $k-1$ (in ogni t) e nell'ultima uguaglianza abbiamo applicato il Corollario 6.11 e la definizione in (7.18).

8 Espansioni di Taylor

Notazione 8.1. Scriveremo

$$\mathbb{N} = \{1, 2, 3, ...\} \text{ e } \mathbb{N}_0 = \{0, 1, 2, 3, ...\}.$$

Un multi–indice è un elemento

$$\begin{aligned}
&\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_d) \in \mathbb{N}_0^d \\
&|\boldsymbol{\alpha}| := \sum_{j=1}^{d} \alpha_j \\
&\boldsymbol{\alpha}! = \alpha_1! ... \alpha_d!.
\end{aligned}$$

Inoltre scriveremo

$$\partial^{\boldsymbol{\alpha}} = \partial_1^{\alpha_1} ... \partial_d^{\alpha_d}$$

e, per un vettore $\mathbf{h} \in \mathbb{R}^d$,

$$\mathbf{h}^{\boldsymbol{\alpha}} = h_1^{\alpha_1} ... h_d^{\alpha_d}.$$

Osservazione 8.2. Sia $A \in \mathcal{L}^k(\mathbb{R}^d, \mathbb{R}^N)$ simmetrica e consideriamo la funzione $f : \mathbb{R}^d \to \mathbb{R}^N$

$$f(\mathbf{x}) := \frac{1}{k!} A(\underbrace{\mathbf{x}, ..., \mathbf{x}}_{k \text{ volte}}).$$

E' facile verificare, ad esempio in termini delle coordinate, che $f \in C^\infty(\mathbb{R}^d, \mathbb{R}^N)$. Ora voglio verificare che

$$D^k f(\mathbf{x}) \equiv A. \tag{8.1}$$

La formula è vera per $k = 1$ e possiamo supporre, per induzione, che sia vera per $k - 1$. Abbiamo, come è facile dimostrare, ripetendo la regola di Leibnitz,

$$\begin{aligned}
&\frac{d}{dt} f(\mathbf{x} + t\mathbf{h}) = \frac{1}{k!} \frac{d}{dt} \left[A(\mathbf{x} + t\mathbf{h}, ..., \mathbf{x} + t\mathbf{h}) \right] \\
&= \frac{1}{k!} \sum_{i=1}^{k} A(\mathbf{x} + t\mathbf{h}, ..., \mathbf{x} + t\mathbf{h}, \underbrace{\mathbf{h}}_{i\text{–esimo}}, \mathbf{x} + t\mathbf{h}, ..., \mathbf{x} + t\mathbf{h}) \\
&= \frac{1}{(k-1)!} A(\mathbf{h}, \underbrace{\mathbf{x} + t\mathbf{h}, ..., \mathbf{x} + t\mathbf{h}}_{(k-1) \text{ volte}}).
\end{aligned}$$

Allora,

$$\begin{aligned}
&D^k f(\mathbf{x}) \mathbf{h}^k = \left(\left(\frac{d}{dt} \right)^k f(\mathbf{x} + t\mathbf{h}) \right) \Bigg|_{t=0} = \left(\left(\frac{d}{dt} \right)^{k-1} \frac{1}{(k-1)!} A(\mathbf{h}, \mathbf{x} + t\mathbf{h}, ..., \mathbf{x} + t\mathbf{h}) \right) \Bigg|_{t=0} \\
&= A\mathbf{h}^k,
\end{aligned}$$

dove nella prima uguaglianza abbiamo usato (7.23) e nell'ultima abbiamo usato (7.23) e l'ipotesi che (8.1) sia vera per $k - 1$. Ora abbiamo ottenuto che

$$B\mathbf{h}^k = 0 \text{ per ogni } \mathbf{h} \in \mathbb{R}^d \text{ dove } B := D^k f(\mathbf{x}) - A.$$

Sappiamo che $B \in \mathcal{L}^k(\mathbb{R}^d, \mathbb{R}^N)$ e che è simmetrica. Da qui si conclude che $B = 0$. Ciò segue dalla *Formula di Polarizzazione*

$$B(\mathbf{x}_1, ..., \mathbf{x}_k) = \frac{1}{2^k k!} \sum_{s_1, ..., s_k \in \{-1, 1\}} s_1 ... s_k B(s_1 \mathbf{x}_1 + ... + s_k \mathbf{x}_k)^k. \tag{8.2}$$

La formula (8.2) è vera per $k = 1$, riducendosi semplicemente a

$$B\mathbf{x} = \frac{B\mathbf{x} - B(-\mathbf{x})}{2}$$

mentre, per $k = 2$, abbiamo

$$B(\mathbf{x}, \mathbf{y}) = \frac{1}{4} B(\mathbf{x} + \mathbf{y})^2 - \frac{1}{4} B(\mathbf{x} - \mathbf{y})^2 = \frac{1}{8} B(\pm(\mathbf{x} + \mathbf{y}))^2 - \frac{1}{8} B(\pm(\mathbf{x} - \mathbf{y}))^2.$$

In generale, abbiamo

$$
\begin{aligned}
&\frac{1}{2^k k!} \sum_{s_1,...,s_k\in\{-1,1\}} s_1...s_k B(s_1\mathbf{x}_1 + ... + s_k\mathbf{x}_k)^k \\
&= \frac{1}{2^k k!} \sum_{s_1,...,s_k\in\{-1,1\}} s_1...s_k \sum_{|\boldsymbol{\alpha}|=k} \frac{k!}{\boldsymbol{\alpha}!} B(s_1\mathbf{x}_1)^{\alpha_1}...(s_k\mathbf{x}_k)^{\alpha_k} \\
&= \frac{1}{2^k} \sum_{s_1,...,s_k\in\{-1,1\}} \sum_{|\boldsymbol{\alpha}|=k} s_1^{\alpha_1+1}...s_k^{\alpha_k+1} \frac{1}{\boldsymbol{\alpha}!} B\mathbf{x}_1^{\alpha_1}...\mathbf{x}_k^{\alpha_k} \qquad (8.3)
\end{aligned}
$$

dove la prima uguaglianza è valida in analogia con la ben nota identità di Newton

$$
(a_1 + ... + a_d)^k = \sum_{|\boldsymbol{\beta}|=k} \frac{k!}{\boldsymbol{\beta}!} a_1^{\beta_1}...a_d^{\beta_d}. \tag{8.4}
$$

I termini in (8.3) dove c'è un esponente $\alpha_i + i$ dispari, si cancellano. Pertanto restano solo i termini dove tutti gli esponenti $\alpha_i + i$ sono pari. Il che significa, che contribuiscono alla somma in (8.3), solo i multiindici per i quali $\alpha_i \neq 0$ per ogni $i = 1, ..., k$ e $|\boldsymbol{\alpha}| = k$. Esiste un unico multiindice con questa proprietà, ed è

$$
\boldsymbol{\alpha} = (1, 1, ..., 1).
$$

Ottieniamo pertanto

$$
\begin{aligned}
&\frac{1}{2^k k!} \sum_{s_1,...,s_k\in\{-1,1\}} s_1...s_k B(s_1\mathbf{x}_1 + ... + s_k\mathbf{x}_k)^k \\
&= \frac{1}{2^k} \sum_{s_1,...,s_k\in\{-1,1\}} B\mathbf{x}_1...\mathbf{x}_k = B\mathbf{x}_1...\mathbf{x}_k,
\end{aligned}
$$

che dà il risultato desiderato, dimostrando (8.2).

Abbiamo il seguente teorema, sulla formula di Peano nelle espansione di Taylor.

Theorem 8.3. *Sia $f : A \subseteq \mathbb{R}^d \to \mathbb{R}^N$ con A aperto, $\mathbf{x}^0 \in A$ con $D(\mathbf{x}^0, r) \subseteq A$, sia $p \in \mathbb{N}$ e sia $f \in C^{p-1}(A, \mathbb{R}^N)$ e sia f differenziabile di ordine p in $\mathbf{x}^0$. Allora per $|\mathbf{h}| < r$ si ha*

$$
f(\mathbf{x}_0 + \mathbf{h}) = P_p(\mathbf{h}) + o(|\mathbf{h}|^p) \ \textit{dove} \tag{8.5}
$$

$$
P_p(\mathbf{h}) := \sum_{k=0}^{p} \frac{1}{k!} D^k f(\mathbf{x}_0)\mathbf{h}^k, \tag{8.6}
$$

dove $\mathbf{h}^k$ è definito in (7.24) e dove

$$
\lim_{\mathbf{h}\to 0} \frac{o(|\mathbf{h}|^p)}{|\mathbf{h}|^p} = 0. \tag{8.7}
$$

Dim. E' sufficiente considerare il caso $N = 1$. Considerando

$$g(\mathbf{h}) = f(\mathbf{x}_0 + \mathbf{h}) - P_p(\mathbf{h}) \Longrightarrow D^k g(0) = 0 \text{ per } k \leq p,$$

dove l'implicazione è una conseguenza dell'Osservazione 8.2. L'enunciato del teorema è una conseguenza di $g(\mathbf{h}) = o(|\mathbf{h}|^p)$, che dimostriamo ora.
Sappiamo che per $p = 1$ il risultato segue dalla definizione di differenziabilità, (6.2), che implica

$$g(\mathbf{h}) = o(|\mathbf{h}|).$$

Quindi, sia $p \geq 2$. Supponiamo per induzione che il risultato sia vero per $p - 1$. Allora

$$Dg(\mathbf{h}) = o(|\mathbf{h}|^{p-1}).$$

Abbiamo, per un $t_* \in (0, 1)$, il Teorema di Lagrange

$$g(\mathbf{h}) = g(t\mathbf{h})]_{t=0}^{t=1} = \frac{d}{dt} g(t\mathbf{h})\bigg|_{t_*} = Dg(t_*\mathbf{h})\mathbf{h} = o(|\mathbf{h}|^{p-1})\mathbf{h} = o(|\mathbf{h}|^p).$$

□

Vogliamo riscrivere in modo più semplice il polinomio di Taylor (8.6).

Proposizione 8.4. *Consideriamo le ipotesi in Teorema 8.3. Risulta che*

$$\sum_{k=0}^{p} \frac{1}{k!} D^k f(\mathbf{x}_0)\mathbf{h}^k = \sum_{|\boldsymbol{\alpha}| \leq p} \frac{1}{\boldsymbol{\alpha}!} \partial^{\boldsymbol{\alpha}} f(\mathbf{x}_0)\mathbf{h}^{\boldsymbol{\alpha}}. \tag{8.8}$$

Dim. In perfetta analogia con l' identità di Newton (8.4), abbiamo

$$\frac{1}{k!} D^k f(\mathbf{x}_0)\mathbf{h}^k = \frac{1}{k!} D^k f(\mathbf{x}_0) \underbrace{(\mathbf{h}, ..., \mathbf{h})}_{k \text{ volte}} = \frac{1}{k!} D^k f(\mathbf{x}_0) \left(h_1 e_1 + ... + h_d e_d, ..., h_1 e_1 + ... + h_d e_d\right)$$

$$= \frac{1}{k!} \sum_{|\boldsymbol{\alpha}|=k} \frac{k!}{\boldsymbol{\alpha}!} D^k f(\mathbf{x}_0) e_1^{\alpha_1} ... e_d^{\alpha_d} h_1^{\alpha_1} ... h_d^{\alpha_d}$$

$$= \sum_{|\boldsymbol{\alpha}|=k} \frac{1}{\boldsymbol{\alpha}!} D^k f(\mathbf{x}_0) \left(\underbrace{e_1, ...e_1}_{\alpha_1 \text{ volte}}, ..., \underbrace{e_d, ...e_d}_{\alpha_d \text{ volte}} \right) \mathbf{h}^{\boldsymbol{\alpha}}.$$

Infine, da (7.13), abbiamo

$$D^k f(\mathbf{x}_0) \left(\underbrace{e_1, ...e_1}_{\alpha_1 \text{ volte}}, ..., \underbrace{e_d, ...e_d}_{\alpha_d \text{ volte}} \right) = \partial_1^{\alpha_1} ... \partial_d^{\alpha_d} f(\mathbf{x}_0) = \partial^{\boldsymbol{\alpha}} f(\mathbf{x}_0).$$

□

9 Test delle derivate seconde per punti critici non degeneri

Definizione 9.1. Consideriamo ora una funzione $f : A \subseteq \mathbb{R}^d \to \mathbb{R}$, A aperto, $\mathbf{x}_0 \in A$, f differenziabile di ordine 2 in $\mathbf{x}_0$ ed, infine, $\mathbf{x}_0$ punto critico di f. Diciamo che $\mathbf{x}_0$ è punto critico non degenere di f se la matrice Hessiana

$$Hf(\mathbf{x}^0) = D^2 f(\mathbf{x}^0) = (\partial_i \partial_j f(\mathbf{x}_0))_{i,j=1}^d = \begin{pmatrix} \partial_1^2 f(\mathbf{x}^0) & \partial_1\partial_2 f(\mathbf{x}^0) & \cdots & \partial_1\partial_d f(\mathbf{x}^0) \\ \partial_2\partial_1 f(\mathbf{x}^0) & \partial_2^2 f(\mathbf{x}^0) & \cdots & \partial_2\partial_d f(\mathbf{x}^0) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_d\partial_1 f(\mathbf{x}^0) & \partial_d\partial_2 f(\mathbf{x}^0) & \cdots & \partial_d^2 f(\mathbf{x}^0) \end{pmatrix}.$$

è di rango d (cioè, è invertibile).

Osservazione 9.2. Ricordiamoci, dal Teorema 7.10, che, siccome f differenziabile è di ordine 2 in $\mathbf{x}_0$, il Hessiano $D^2 f(\mathbf{x}^0)$ è una matrice simmetrica. Ciò significa che esiste una matrice ortogonale A, tale che

$$^tAD^2 f(\mathbf{x}^0)A = \mathrm{diag}(\lambda_1, ..., \lambda_d) \tag{9.1}$$

dove gli $\lambda_j \in \mathbb{R}$ sono gli autovalori, ripetuti un numero di volte pari alla loro molteplicità. Si noti che se consideriamo la funzione

$$g(\mathbf{y}) := f(A\mathbf{y} + \mathbf{x}^0) \tag{9.2}$$

dalla regola della catena, per $x = A\mathbf{y} + \mathbf{x}^0$ abbiamo

$$Dg(\mathbf{y}) = Df(\mathbf{x})A$$

Notare che $Dg(0) = 0 \iff Df(\mathbf{x}^0) = 0$. Quindi 0 è un punto critico se e solo se $\mathbf{x}^0$ lo è. E' evidente che 0 è un punto di max/min locale per g se e solo se $\mathbf{x}^0$ è un punto di max/min locale per f.
Per $\mathbf{h} \in \mathbb{R}^d$ fissato,

$$Dg(\mathbf{y})\mathbf{h} = Df(A\mathbf{y} + \mathbf{x}^0)A\mathbf{h}$$

Ora

$$D^2 g(0)(\mathbf{h}, \mathbf{k}) = D^2 f(\mathbf{x}^0)(A\mathbf{h}, A\mathbf{k}).$$

Ricordiamoci che

$$D^2 g(0)(\mathbf{h}, \mathbf{k}) = {}^t\mathbf{k}D^2 g(0)\mathbf{h} \text{ e che } {}^t\mathbf{k}\,{}^tAD^2 f(\mathbf{x}^0)A = D^2 f(\mathbf{x}^0)(A\mathbf{h}, Ak).$$

Siccome sono uguali, segue

$$D^2 g(0) = {}^tAD^2 f(\mathbf{x}^0)A = \mathrm{diag}(\lambda_1, ..., \lambda_d) \tag{9.3}$$

Se applichiamo il Teorema 8.3 otteniamo, assumendo che 0 è un punto critico di g,

$$g(\mathbf{y}) = g(0) + \frac{1}{2}\sum_{j=1}^{d} \partial_j^2 g(0) y_j^2 + o(|\mathbf{y}|^2)$$
$$= g(0) + \frac{1}{2}\sum_{j=1}^{d} \lambda_j y_j^2 + o(|\mathbf{y}|^2). \tag{9.4}$$

Definizione 9.3. La matrice $D^2 f(\mathbf{x}^0)$ è strettamente positiva se $\lambda_j > 0$ per ogni j. E' positiva se $\lambda_j \geq 0$ per ogni j. E' strettamente negativa se $\lambda_j < 0$ per ogni j. E'negativa se $\lambda_j \leq 0$ per ogni j.

Lemma 9.4. *Sia $f : A \subseteq \mathbb{R}^d \to \mathbb{R}$, A aperto, $Df \in C^0(A, \mathcal{L}(\mathbb{R}^d, \mathbb{R}))$, $\mathbf{x}_0 \in A$, f differenziabile di ordine 2 in $\mathbf{x}_0$, $\mathbf{x}_0$ punto critico di f non degenere (cioè tale che $D^2 f(x_0) \in \mathcal{L}^2(\mathbb{R}^d, \mathbb{R}) \sim \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ è un isomorfismo) . Valgono le seguenti proprietà.*

a *Tutti gli autovalori sono $\lambda_j \neq 0$ per ogni $j = 1, ..., d$.*

b *Se $D^2 f(\mathbf{x}^0)$ è strettamente positiva, $\mathbf{x}_0$ è un punto di minimo locale.*

c *Se $D^2 f(\mathbf{x}^0)$ è strettamente negativa, $\mathbf{x}_0$ è un punto di massimo locale.*

d *Nel caso in cui esistano indici i e j con $\lambda_i < 0 < \lambda_j$, $\mathbf{x}_0$ non è ne' un punto di massimo ne' punto di minimo locale.*

Dim. Essendo $\mathbf{x}_0$ un punto critico, abbiamo $Df(\mathbf{x}^0) = 0$.
Consideriamo **a**. Siccome per ipotesi $D^2 f(\mathbf{x}^0)$ è di rango d, ed abbiamo

$$0 \neq \det D^2 f(\mathbf{x}^0) = \det D^2 g(0) = \lambda_1 ... \lambda_d,$$

segue che $\lambda_j \neq 0$ per ogni $j = 1, ..., d$.

Consideriamo **b**. E' equivalente dimostrare che 0 è un punto di minimo locale per g. Consideriamo pertanto (9.4). Sia $c_0 := \min\{\lambda_1, ..., \lambda_d\}$. Sappiamo che

$$\forall\ \epsilon > 0\ \exists\ \delta_\epsilon > 0 \text{ t.c. } |\mathbf{y}| < \delta_\epsilon \text{ implica } |o(|\mathbf{y}|^2)| < \epsilon |\mathbf{y}|^2. \tag{9.5}$$

Scegliamo allora $\epsilon = c_0/4$. Allora per $0 < |\mathbf{y}| < \delta_\epsilon$ abbiamo

$$g(\mathbf{y}) = g(0) + \frac{1}{2}\sum_{j=1}^{d} \lambda_j y_j^2 + o(|\mathbf{y}|^2) \geq g(0) + \frac{c_0}{2}|\mathbf{y}|^2 - \left|o(|\mathbf{y}|^2)\right|$$
$$> g(0) + \frac{c_0}{2}|\mathbf{y}|^2 - \frac{c_0}{4}|\mathbf{y}|^2 = g(0) + \frac{c_0}{4}|\mathbf{y}|^2 > g(0).$$

Questo conclude la dimostrazione di **b**. Si noti che **b** $\Longleftrightarrow$ **c**: basta moltiplicare f per -1.

Consideriamo **d**. E' equivalente dimostrare che 0 non è ne' un punto di massimo ne' punto di minimo locale per g. Non è restrittivo assumere $\lambda_1 > 0$ e $\lambda_2 < 0$. Poniamo $c_0 = \min\{|\lambda_1|, |\lambda_2|\}$

$$h_j(t) = g(te_j) \text{ per } j = 1, 2$$

In (9.4) otteniamo

$$h_j(t) = g(te_j) = g(0) + \frac{1}{2}\lambda_j t^2 + o(t^2).$$

Scegliamo $\epsilon = c_0/4$. Allora per $0 < |t| < \delta_\epsilon$, procedendo come sopra abbiamo

$$h_1(t) = g(0) + \frac{1}{2}\lambda_1 t^2 + o(t^2) \geq g(0) + \frac{c_0}{2}t^2 - |o(t^2)| > g(0) + \frac{c_0}{2}t^2 - \frac{c_0}{4}t^2 = g(0) + \frac{c_0}{4}t^2 > g(0)$$

garantisce che $t = 0$ è un punto di minimo locale per h_1 mentre

$$h_2(t) = g(0) + \frac{1}{2}\lambda_2 t^2 + o(t^2) \leq g(0) - \frac{c_0}{2}t^2 + |o(t^2)| < g(0) - \frac{c_0}{2}t^2 + \frac{c_0}{4}t^2 = g(0) - \frac{c_0}{4}t^2 < g(0)$$

garantisce che $t = 0$ è un punto di massimo locale per h_1. E' immediato che 0 non è ne' un punto di massimo ne' punto di minimo locale per g. □

Esempio 9.5. Consideriamo la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita come

$$f(x, y) = 3x^2 + 4y^2 - 2x + 2.$$

Poiché la funzione è di classe almeno C^2, i punti critici coincidono con i punti stazionari con gradiente nullo, quindi dobbiamo risolvere il sistema di equazioni

$$\nabla f(x, y) = (6x - 2, 8y) = (0, 0)$$

che presenta la soluzione $(1/3, 0)$. La matrice Hessiana risulta

$$H_f(x, y) = \begin{pmatrix} 6 & 0 \\ 0 & 8 \end{pmatrix}$$

per ogni $(x, y) \in \mathbb{R}^2$. In particolare notiamo immediatamente che la matrice $H_f(1/3, 0)$ ha tutti gli autovalori positivi. Quindi $(1/3, 0)$ è un punto di minimo locale con valore $f(1/3, 0) = 5/3$.

Il limite

$$\lim_{(x,y)\to\infty} f(x, y) = +\infty\,,$$

garantisce l'esistenza di un punto di minimo assoluto. Avendo trovato un solo minimo locale, questo è necessariamente anche il minimo assoluto.

Per dimostrare il limitebasta osservare che

$$f(x, y) = 3x^2 + 4y^2 - 2x + 2 = x^2 + y^2 + (2x^2 - 2x + 2 + 3y^2) \geq x^2 + y^2$$

e usare il teorema del confronto.

Esempio 9.6. Consideriamo la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita come

$$f(x, y) = \frac{1}{4}x^4 - \frac{1}{3}x^3 - x^2 + y^4 + y^2$$

(di classe C^2 su tutto il dominio). Il suo gradiente risulta

$$\nabla f(x, y) = \big(x^3 - x^2 - 2x\,,\, 4y^3 + 2y\big)$$

e si annulla nei punti che risolvono il sistema

$$\begin{cases} x^3 - x^2 - 2x = 0 \\ 4y^3 + 2y = 0 \end{cases} \qquad \Rightarrow \qquad \begin{cases} x(x^2 - x - 2) = 0 \\ y(4y^2 + 2) = 0 \end{cases}$$

che presenta le soluzioni $(0, 0), (-1, 0), (2, 0)$ (i punti critici). La matrice Hessiana di f è

$$H_f(x, y) = \begin{pmatrix} 3x^2 - 2x - 2 & 0 \\ 0 & 12y^2 + 2 \end{pmatrix}$$

Calcoliamo la matrice Hessiana nei tre punti critici

$$H_f(0, 0) = \begin{pmatrix} -2 & 0 \\ 0 & +2 \end{pmatrix},$$

$$H_f(-1, 0) = \begin{pmatrix} +3 & 0 \\ 0 & +2 \end{pmatrix},$$

$$H_f(2, 0) = \begin{pmatrix} +6 & 0 \\ 0 & +2 \end{pmatrix}.$$

Il punto $(0, 0)$ risulta un punto di sella, gli altri due punti di minimo locale stretto.
Si noti inoltre che

$$\lim_{(x,y)\to\infty} f(x, y) = +\infty\,,$$

garantisce l'esistenza di un punto di minimo assoluto. Siccome $f(-1, 0) = -\frac{5}{12}$ e $f(2, 0) = -\frac{8}{3}$, segue che $(2, 0)$ è il punto di minimo assoluto.

Esempio 9.7. Consideriamo $f(x, y) = (y + x^2)e^{-x^2-y^2}$.

Le derivate parziali le abbiamo calcolate e sappiamo che sono

$$f_x(x, y) = (2x - 2x(y + x^2))e^{-x^2-y^2} = 2x(1 - x^2 - y)e^{-x^2-y^2}$$
$$f_y(x, y) = (1 - 2x^2y - 2y^2)e^{-x^2-y^2}.$$

Ci ricordiamo di avere calcolato che i punti critici sono

$$(0, \frac{1}{\sqrt{2}})\,, \quad (0, -\frac{1}{\sqrt{2}})\,, \quad (\frac{1}{\sqrt{2}}, \frac{1}{2})\,, \quad (-\frac{1}{\sqrt{2}}, \frac{1}{2}).$$

Di questi abbiamo visto che $(0, -\frac{1}{\sqrt{2}})$ é il punto di minimo assoluto e che $(\pm\frac{1}{\sqrt{2}}, \frac{1}{2})$ sono i punti di massimo assoluto. Avevamo lasciato in sospeso l' asserzione che $(0, \frac{1}{\sqrt{2}})$ è una sella. Verifichiamolo.

Abbiamo

$$
\begin{aligned}
f_{xx}(x,y) &= (2-2y-6x^2)e^{-x^2-y^2} + (2x-2xy-2x^3)\frac{\partial}{\partial x}e^{-x^2-y^2} \\
f_{xy}(x,y) &= -2ye^{-x^2-y^2} + (2x-2xy-2x^3)\frac{\partial}{\partial y}e^{-x^2-y^2} \\
f_{yy}(x,y) &= (-2x^2-4y)e^{-x^2-y^2} + (1-2x^2y-2y^2)\frac{\partial}{\partial y}e^{-x^2-y^2}.
\end{aligned}
$$

ora, nel punto $(0, \frac{1}{\sqrt{2}})$ i secondi termini a destra si annullano e resta

$$
\begin{aligned}
f_{xx}(0, \frac{1}{\sqrt{2}}) &= (2-\sqrt{2})e^{-\frac{1}{2}} \\
f_{xy}(0, \frac{1}{\sqrt{2}}) &= -2\sqrt{2}e^{-\frac{1}{2}} \\
f_{yy}(0, \frac{1}{\sqrt{2}}) &= -4\sqrt{2}e^{-\frac{1}{2}}.
\end{aligned}
$$

Allora

$$
\det H_f(0, \frac{1}{\sqrt{2}}) = -\Big(4(2-\sqrt{2})+8\Big)\, e^{-1} < 0.
$$

Esempio 9.8. Consideriamo $f(x,y) = x^4 + y^4 - x^2 + y^2$.

Notare che la funzione è simmetrica per riflessione rispetto ad entrambi gli assi coordinati. In altre parole

$$
f(x,y) = f(-x,y) = f(x,-y) = f(-x,-y).
$$

Abbiamo

$$
\lim_{(x,y)\to\infty} f(x,y) = +\infty
$$

e pertanto ci sono minimo assoluti.

Le derivate sono

$$
\begin{aligned}
f_x(x,y) &= 4x^3 - 2x \\
f_y(x,y) &= 4y^3 + 2y.
\end{aligned}
$$

Cerchiamo i punti critici, cioé risolviamo il sistema

$$
\begin{aligned}
4x^3 - 2x &= 2x(2x^2-1) = 0 \\
4y^3 + 2y &= 2y(2y^2+1) = 0.
\end{aligned}
$$

Si vede che vi sono tre punti critici

$$
(0,0)\,, \quad (\pm\frac{1}{\sqrt{2}}, 0)
$$

con $f(\frac{1}{\sqrt{2}},0) = f(-\frac{1}{\sqrt{2}},0) < 0 = f(0,0)$. Quindi $(\pm\frac{1}{\sqrt{2}},0)$ sono necessariamente i punti di minimo assoluto. Verifichiamo che (0,0) é una sella. Abbiamo

$$\begin{aligned} f_{xx}(x,y) &= 12x^2 - 2 \\ f_{xy}(x,y) &= 0 \\ f_{yy}(x,y) &= 12y^2 + 2. \end{aligned}$$

Pertanto

$$\begin{aligned} f_{xx}(0,0) &= -2 \\ f_{xy}(0,0) &= 0 \\ f_{yy}(0,0) &= 2 \end{aligned}$$

e quindi $\det H_f(0,0) = -4 < 0$.

Esempio 9.9. Studiamo

$$f(x,y) = -\left(x^2-1\right)^2 - \left(x^2y - x - 1\right)^2.$$

Ovviamente $f(x,y) \le 0$ ovunque. Notare che

$$\lim_{(x,y)\to\infty} f(x,y) \text{ non esiste,}$$

perché ovviamente $f(x,y) \le -\left(x^2-1\right)^2 \xrightarrow{x\to\infty} -\infty$, ma sulla curva $x^2y = 1$, abbiamo $x \xrightarrow{y\to+\infty} 0$ e quindi $f(x,y) \xrightarrow{y\to+\infty} -1$.

L'equazione $f(x,y) = 0$ è equivalente al sistema

$$\begin{aligned} x^2 - 1 &= 0 \\ y - x - 1 &= 0 \end{aligned}$$

con soluzioni $(1,2)$ e $(-1,0)$, che evidentemente saranno anche punti critici. Verifichiamo se ci sono altri punti critici, calcolando

$$\begin{aligned} f_x &= -4\left(x^2-1\right)x - 2\left(x^2y - x - 1\right)(2xy-1) = 0 \\ f_y &= -2\left(x^2y - x - 1\right)x^2 = 0. \end{aligned}$$

La seconda equazione è soddisfatta per $x = 0$, che però non dà soluzioni della prima, oppure se $x \neq 0$ per $x^2y - x - 1 = 0$, per la quale la prima equazione diviene $-4\left(x^2-1\right)x = 0$, cioè $x^2 - 1 = 0$, da cui ho già ricavato $(1,2)$ e $(-1,0)$.

E' interessante in questo esempio che tra i due punti di massimo non c'è un passo di montagna con una sella, come in altri esempi.

Esempio 9.10 (baricentro). Dati dei punti $(x_1,y_1)...(x_n,y_n)$ trovare il punto che minimizza

$$f(x,y) = \sum_{j=1}^{n} \left[(x-x_j)^2 + (y-y_j)^2\right].$$

Osservare che

$$\lim_{(x,y)\to\infty} f(x,y) = +\infty$$

e pertanto ci sono minimo assoluti. Abbiamo

$$f_x(x,y) = 2\sum_{j=1}^{n}(x - x_j) = 0 \Leftrightarrow x = \frac{\sum_{j=1}^{n} x_j}{n}$$
$$f_y(x,y) = 2\sum_{j=1}^{n}(y - y_j) = 0 \Leftrightarrow y = \frac{\sum_{j=1}^{n} y_j}{n}.$$

ossia un unico punto critico (il baricentro) che è necessariamente il punto di minimo assoluto.

Esempio 9.11 (Metodo dei minimi quadrati: calcolo della retta di regressione). Consideriamo dei punti $(x_1, y_1)...(x_n, y_n)$ che supponiamo non allineati. Cerchiamo una retta che in qualche maniera meglio approssimi i punti.

La retta avrà equazione $y = ax + b$. Tra tutte le rette decidiamo di scegliere quella i cui coefficienti minimizzino

$$f(a,b) = \sum_{j=1}^{n}(ax_j + b - y_j)^2 = a^2P + b^2n + Q + 2Xab - 2Sa - 2bY$$

dove

$$P = \sum_{j=1}^{n} x_j^2 \quad , \quad X = \sum_{j=1}^{n} x_j \quad , \quad Y = \sum_{j=1}^{n} y_j$$
$$S = \sum_{j=1}^{n} x_j y_j \quad , \quad Q = \sum_{j=1}^{n} y_j^2.$$

Cerchiamo i punti critici.

$$f_a(a,b) = 2\sum_{j=1}^{n} x_j(ax_j + b - y_j) = 2(aP + bX - S) = 0$$
$$f_b(a,b) = 2\sum_{j=1}^{n}(ax_j + b - y_j) = 2(aX + nb - Y) = 0.$$

Quindi otteniamo il sistema

$$aP + bX = S$$
$$aX + nb = Y$$

ossia

$$\begin{pmatrix} P & X \\ X & n \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} S \\ Y \end{pmatrix}$$

che, se $nP - X^2 \neq 0$ ha soluzione

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{nP - X^2} \begin{pmatrix} n & -X \\ -X & P \end{pmatrix} \begin{pmatrix} S \\ Y \end{pmatrix}.$$

Verifichiamo che $nP - X^2 > 0$.

$$\begin{aligned} nP - X^2 &= \sum_{j=1}^{n} nx_j^2 - \left(\sum_{j=1}^{n} x_j \right)^2 = \sum_{j=1}^{n} nx_j^2 - \sum_{j,k=1}^{n} x_j x_k \\ &> \sum_{j=1}^{n} nx_j^2 - \sum_{j,k=1}^{n} \frac{x_j^2 + x_k^2}{2} = \sum_{j=1}^{n} nx_j^2 - \frac{1}{2} \sum_{j=1}^{n} nx_j^2 - \frac{1}{2} \sum_{k=1}^{n} nx_k^2 = 0. \end{aligned}$$

Qui l'ultima disuguaglianza è stretta, ossia abbiamo $2x_j x_k < x_j^2 + x_k^2$ per una qualche scelta di indici, perchè $2x_j x_k = x_j^2 + x_k^2$ implica $x_j = x_k$. Ma se questo fosse vero per tutti gli indici, avremmo $x_j = c$ per una fissata costante c, ossia i punti apparterrebbero ad una fissata retta verticale, il che è escluso per ipotesi. Infine notare che

$$2 \begin{pmatrix} P & X \\ X & n \end{pmatrix}$$

è la matrice Hessiana di f, e che il punto critico è un minimo locale. Infatti

$$\det \begin{pmatrix} P - \lambda & X \\ X & n - \lambda \end{pmatrix} = (P - \lambda)(n - \lambda) - X^2 = \lambda^2 - (P + n)\lambda - X^2$$

ha radici

$$\lambda_\pm = \frac{P + n}{2} \pm \frac{\sqrt{(P + n)^2 - 4nP + 4X^2}}{2} > 0.$$

Infine, risulta che è un minimo assoluto. Infatti $\lim\limits_{(a,b) \to \infty} f(a, b) = +\infty$ come vediamo subito, e quindi esiste un minimo assoluto. Per verificare il limite, poniamo

$$\mathbf{z} := \begin{pmatrix} a \\ b \end{pmatrix} \text{ e } f(\mathbf{z}) = f(a, b) = {}^t\mathbf{z} \begin{pmatrix} P & X \\ X & n \end{pmatrix} \mathbf{z} - 2(S, y)\mathbf{z}.$$

Sia ora A la matrice ortogonale tale che

$${}^tA \begin{pmatrix} P & X \\ X & n \end{pmatrix} A = \operatorname{diag}(\lambda_+, \lambda_-).$$

Allora, per $\mathbf{z} = A\mathbf{w}$, si ha

$$g(\mathbf{w}) := f(A\mathbf{w}) = \lambda_+ w_1^2 + \lambda_- w_2^2 - 2(S, y)A\mathbf{w} \geq \lambda_- |\mathbf{w}|^2 - 2|(S, y)|\ |A|\ |\mathbf{w}| \xrightarrow{\mathbf{w} \to \infty} +\infty.$$

10 Domini con bordo

Esempio 10.1. Troviamo punti di massimo e di minimo assoluto di

$$f(x, y) = x^2 + y^2 + 8y - 4x + 2$$

per (x, y) nel quadrato $Q = \{(x, y)\, t.c.\, x \in [0, 1],\, y \in [0, 1]\}$.

Prima di tutto cerchiamo punti di massimo e di minimo locali in $\mathring{Q}$, che sono punti critici. Poniamo

$$f_x = 2x - 4 = 0$$
$$f_y = 2y + 8 = 0.$$

Otteniamo cosi il punto critico $(2, -4)$ che è ovviamente fuori di Q. Pertanto all'interno di Q non abbiamo ne' punti di massimo ne' punti di minimo, ne' locali ne' assoluti.

Per trovare massimo e minimo assoluti bisogna restringere f sui lati di Q, Sul lato $I_1 = [0, 1] \times \{0\}$ abbiamo

$$f(x, 0) = x^2 - 4x + 2.$$

Abbiamo $f_x(x, 0) = 2(x - 2) < 0$, e quindi su I_1 il punto di massimo è $(0, 0)$ ed il punto di minimo è $(1, 0)$.

Sul lato $I_2 = \{1\} \times [0, 1]$ abbiamo

$$f(1, y) = y^2 + 8y - 1.$$

Abbiamo $f_y(1, y) = 2y + 8 > 0$ e quindi su I_2 il punto di massimo è $(1, 1)$ ed il punto di minimo è $(1, 0)$.

Sul lato $I_3 = [0, 1] \times \{1\}$ abbiamo $f_x(x, 1) = 2(x - 2) < 0$, e quindi su I_3 il punto di massimo è $(0, 1)$ ed il punto di minimo è $(1, 1)$.

Infine, sul lato $I_4 = \{0\} \times [0, 1]$, abbiamo $f(0, y) = y^2 + 8y + 2$, $f_y(0, y) = 2y + 8 > 0$, e quindi su I_4 il punto di massimo è $(0, 1)$ ed il punto di minimo è $(0, 0)$.

In conclusione, sul bordo $\partial Q = I_1 \cup I_2 \cup I_3 \cup I_4$, $(1, 0)$ è il punto di minimo assoluto mentre $(0, 1)$ è il punto di minimo assoluto. Infatti un calcolo elementare dà

$$f(0, 0) = 2\,,\; f(1, 0) = -1\,,\; f(0, 1) = 11\,,\; f(1, 1) = 8.$$

Concludiamo che $(1, 0)$ è il punto di minimo assoluto di f su Q e $(0, 1)$ è il punto di massimo assoluto di f su Q.

Esempio 10.2. Trovare massimo e minimo assoluto di $f(x, y) = x^2 - y^2 + y$ nella regione Ω

$$x^2 + (y^2)/2 \le 1$$

.

Iniziamo qui la risposta. Prima di tutto cerchiamo punti critici all'interno della regione, ossia per $x^2 + (y^2)/2 < 1$. Poniamo

$$\begin{aligned} f_x &= 2x = 0 \\ f_y &= 1 - 2y = 0. \end{aligned}$$

Otteniamo cosi il punto critico $(0, \frac{1}{2})$, che si trova in $\mathring{\Omega}$.

Il Hessiano è dato da

$$H_f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix},$$

quindi è una matrice costante con determinante uguale a $-4 < 0$ e quindi $(0, \frac{1}{2})$ è una sella. Pertanto i punti di massimo e di minimo si trovano sul bordo, ossia sulla ellisse di equazione $x^2 + (y^2)/2 = 1$.

Il bordo $\partial\Omega$ è l'ellisse ed è assimilabile ad un intervallo. Ci conviene *parametrizzare* l'ellisse. Specificamente, i punti dell'ellisse si possono scrivere nella forma

$$\begin{aligned} x &= \cos\vartheta \\ y &= \sqrt{2}\sin\vartheta \text{ per } \vartheta \in [0, 2\pi]. \end{aligned}$$

I valori della funzione ristretta su $\partial\Omega$ sono pertanto

$$h(\vartheta) = f(\cos\vartheta, \sqrt{2}\sin\vartheta) = \cos^2\theta - 2\sin^2\theta + \sqrt{2}\sin\theta = -3\sin^2\theta + \sqrt{2}\sin\theta + 1$$

Abbiamo

$$h'(\vartheta) = \left(-6\sin\theta + \sqrt{2}\right)\cos\vartheta$$

con punti critici per $\cos\vartheta = 0$, che sono esplicitamente noti, $\vartheta = \pi/2, 3\pi/2$, e per $\sin\theta = \dfrac{\sqrt{2}}{6}$.

$\cos\vartheta = 0$ corrisponde ai punti $(0, \pm\sqrt{2})$ dove abbiamo

$$\begin{aligned} f(0, \sqrt{2}) &= -2 + \sqrt{2} \\ f(0, -\sqrt{2}) &= -2 - \sqrt{2}. \end{aligned}$$

Per $\sin\theta = \dfrac{\sqrt{2}}{6}$ abbiamo

$$y = \sqrt{2}\sin\vartheta = \sqrt{2}\ \frac{\sqrt{2}}{6} = \frac{1}{3}$$

e quindi

$$x = \pm\sqrt{1 - \frac{y^2}{2}} = \pm\sqrt{1 - \frac{1}{18}} = \pm\sqrt{\frac{17}{18}}.$$

Abbiamo ottenuto gli ulteriori punti $\left(\pm\sqrt{\frac{17}{18}}, \frac{1}{3}\right)$. Abbiamo

$$f\left(\pm\sqrt{\frac{17}{18}}, \frac{1}{3}\right) = \frac{7}{6}$$

Pertanto $\left(\pm\sqrt{\dfrac{17}{18}}, \dfrac{1}{3}\right)$ sono i punti di massimo assoluto mentre $(0, -\sqrt{2})$ è il punto di minimo assoluto.

11 Il Teorema della Funzione Implicita

I precedenti esempi sollevano il problema di come si possa parametrizzare sottoinsiemi di spazi Euclidei. Uno strumento fondamentale è il Teorema della Funzione Implicita.

Prima di enunciare il Teorema della Funzione Implicita consideriamo il seguente lemma.

Lemma 11.1. *Supponiamo che $T \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$ sia un operatore invertibile e poniamo*

$$C_1 = \left|T^{-1}\right|. \tag{11.1}$$

Allora, se $B \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$, risulta

$$C_1|B - T| \leq 1/2 \Rightarrow \text{ esiste } B^{-1} \text{ e } \left|B^{-1}\right| \leq 2C_1. \tag{11.2}$$

Dim. Abbiamo

$$B = T + (B - T) = T\left(1 + T^{-1}(B - T)\right).$$

Notiamo da Esempio 2.10 che

$$\begin{aligned} &|T^{-1}(B - T)| \leq |T^{-1}|\ |B - T| = C_1|B - T| \leq 1/2 \Rightarrow \\ &\text{esiste } \left(1 + T^{-1}(B - T)\right)^{-1} \text{ e } \left|\left(1 + T^{-1}(B - T)\right)^{-1}\right| \leq 2. \end{aligned}$$

Siccome allora

$$B^{-1} = \left(1 + T^{-1}(B - T)\right)^{-1} T^{-1}$$

da (2.9) abbiamo

$$|B^{-1}| = |\left(1 + T^{-1}(B - T)\right)^{-1} T^{-1}| \leq 2C_1.$$

□

Esercizio 11.2. Sia $G \in C^0(X, \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N))$ una funzione continua da uno spazio metrico X a valori in $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$. Supponiamo che per ogni $x \in X$ risulti che l'operatore $G(x) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$ è invertibile. Dimostrare allora che

$$X \ni x \to (G(x))^{-1} \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N) \text{ è una funzione continua } X \to \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N). \tag{11.3}$$

Theorem 11.3. *Dati* $\mathbf{x}_0 \in \mathbb{R}^d$ *e* $\mathbf{y}_0 \in \mathbb{R}^N$*, un aperto* $A \subseteq \mathbb{R}^d \times \mathbb{R}^N$ *con* $(\mathbf{x}_0, \mathbf{y}_0) \in A$*. Consideriamo poi una funzione*

$$A \ni (\mathbf{x}, \mathbf{y}) \to F(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N \tag{11.4}$$

Supponiamo che $F(\mathbf{x}_0, \mathbf{y}_0) = 0$*, che sia* $F \in C^1(A, \mathbb{R}^N)$ *e che* $D_{\mathbf{y}} F(\mathbf{x}_0, \mathbf{y}_0) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$ *sia invertibile.*
Allora abbiamo quanto segue.

a *Esistono un intorno* $U = D(\mathbf{x}_0, r_1)$ *di* $\mathbf{x}_0$ *in* $\mathbb{R}^d$ *e un intorno* $V = D(\mathbf{y}_0, r_2)$ *di* $\mathbf{y}_0$ *in* $\mathbb{R}^N$ *con* $U \times V \subseteq A$*, ed una funzione* $\mathbf{x} \to f(\mathbf{x}) : U \to V$ *tali che*

$$F(\mathbf{x}, \mathbf{y}) = 0 \text{ con } (\mathbf{x}, \mathbf{y}) \in U \times V \Longleftrightarrow \mathbf{y} = f(\mathbf{x}) \text{ con } \mathbf{x} \in U. \tag{11.5}$$

b *Si ha* $f \in C^1(U, \mathbb{R}^N)$ *con*

$$Df(\mathbf{x}) = -\left(D_{\mathbf{y}} F(\mathbf{x}, f(\mathbf{x}))\right)^{-1} D_{\mathbf{x}} F(\mathbf{x}, f(\mathbf{x})) \text{ per ogni } \mathbf{x} \in U. \tag{11.6}$$

Osservazione 11.4. Ovviamente, l'enunciato deve comprendere i casi quando già in partenza

$$F(\mathbf{x}, \mathbf{y}) = \mathbf{y} - f(\mathbf{x}).$$

Notare che in questo caso,

$$D_{\mathbf{y}} F(\mathbf{x}, \mathbf{y}) = D_{\mathbf{y}} \left(\mathbf{y} - f(\mathbf{x})\right) = D_{\mathbf{y}} \left(\mathbf{y}\right) = 1 \text{ identità in } \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N),$$

che ovviamente è invertibile in $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$.

Dimostrazione di **a**. Siccome F è differenziabile in $(\mathbf{x}_0, \mathbf{y}_0)$, abbiamo

$$F(\mathbf{x}, \mathbf{y}) = D_{\mathbf{x}} F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0) + D_{\mathbf{y}} F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0) + \sigma(\mathbf{x}, \mathbf{y}) \tag{11.7}$$

con

$$\lim_{(\mathbf{x}, \mathbf{y}) \to (\mathbf{x}_0, \mathbf{y}_0)} \frac{\sigma(\mathbf{x}, \mathbf{y})}{|\mathbf{x} - \mathbf{x}_0| + |\mathbf{y} - \mathbf{y}_0|} = 0.$$

Rimpicciolendo A, se necessario, possiamo assumere che

$$|\sigma(\mathbf{x}, \mathbf{y})| < \varepsilon_0 \left(|\mathbf{x} - \mathbf{x}_0| + |\mathbf{y} - \mathbf{y}_0|\right) \text{ in } A,$$

dove $\varepsilon_0 > 0$ è un numero, che possiamo prendere piccolo a piacere, e che fisseremo sotto.
Introduciamo le costanti

$$C_1 = \left|\left(D_{\mathbf{y}} F(\mathbf{x}_0, \mathbf{y}_0)\right)^{-1}\right|$$
$$C_2 = \left|D_{\mathbf{x}} F(\mathbf{x}_0, \mathbf{y}_0)\right|.$$

Se ora un punto $(\mathbf{x}, \mathbf{y})$ soddisfa (11.5) in A, allora abbiamo

$$\mathbf{y} = \mathbf{y}_0 - (D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1} D_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0) - (D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1} \sigma(\mathbf{x}, \mathbf{y}). \tag{11.8}$$

Fissiamo $\mathbf{x}$ e consideriamo al variare di $\mathbf{y}$ la funzione

$$\Phi(\mathbf{y}) := \mathbf{y}_0 - (D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1} D_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{x} - \mathbf{x}_0) - (D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1} \sigma(\mathbf{x}, \mathbf{y}). \tag{11.9}$$

Abbiamo

$$\begin{aligned}|\Phi(\mathbf{y}) - \mathbf{y}_0| &\le \left|(D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1}\right| |D_{\mathbf{x}}F(\mathbf{x}_0, \mathbf{y}_0)| \, |\mathbf{x} - \mathbf{x}_0| + \left|(D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1}\right| \varepsilon_0 \left(|\mathbf{x} - \mathbf{x}_0| + |\mathbf{y} - \mathbf{y}_0|\right) \\ &\le C_1 C_2 |\mathbf{x} - \mathbf{x}_0| + C_2 \varepsilon_0 \left(|\mathbf{x} - \mathbf{x}_0| + |\mathbf{y} - \mathbf{y}_0|\right).\end{aligned}$$

Supponiamo ora che $|\mathbf{x} - \mathbf{x}_0| \le r_1$ e $|\mathbf{y} - \mathbf{y}_0| \le r_2$ per due numeri $r_1 > 0$ e $r_2 > 0$ da definire.
Allora

$$|\Phi(\mathbf{y}) - \mathbf{y}_0| \le (C_1 C_2 + C_2 \varepsilon_0) \, r_1 + C_2 \varepsilon_0 r_2.$$

Ora, scegliendo $\varepsilon_0 > 0$ ed $r_1 > 0$ in modo che

$$C_2 \varepsilon_0 < 1/4, \tag{11.10}$$

$$(C_1 C_2 + C_2 \varepsilon_0) \, r_1 < \frac{1}{4} r_2, \tag{11.11}$$

concludiamo che

$$|\Phi(\mathbf{y}) - \mathbf{y}_0| \le (C_1 C_2 + C_2 \varepsilon_0) \, r_1 + C_2 \varepsilon_0 r_2 < \frac{1}{2} r_2. \tag{11.12}$$

Questo significa, che

$$\text{per ogni } \mathbf{x} \in \overline{D(\mathbf{x}_0, r_1)} \text{ abbiamo } \Phi : \overline{D(\mathbf{y}_0, r_2)} \to \overline{D(\mathbf{y}_0, r_2/2)} \subseteq D(\mathbf{y}_0, r_2). \tag{11.13}$$

Finora abbiamo utilizzato unicamente la differenziabilità di F nel punto $(\mathbf{x}_0, \mathbf{y}_0)$.
Ora dimostriamo che possiamo fare in modo che Φ sia una contrazione. Per farlo, utilizzerema la continuità $D_{\mathbf{y}}F \in C^0(A, \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N))$. In particolare la continuità di $D_{\mathbf{y}}F(\mathbf{x}, \mathbf{y})$ nel punto $(\mathbf{x}_0, \mathbf{y}_0)$ ci garantisce che prendendo r_1 ed r_2 sufficientemente piccoli, per ogni preassegnato $\varepsilon_0 > 0$ abbiamo

$$|D_{\mathbf{y}}F\,(\mathbf{x}, \mathbf{y})) - D_{\mathbf{y}}F\,(\mathbf{x}_0, \mathbf{y}_0)| < \varepsilon_0 \text{ per ogni } x \in \overline{D(\mathbf{x}_0, r_1)} \text{ e } \mathbf{y} \in \overline{D(\mathbf{y}_0, r_2)} \tag{11.14}$$

In aggiunta a (11.10)–(11.16) chiediamo sia soddisfatta

$$C_1 \varepsilon_0 < 1/2. \tag{11.15}$$

Ora, abbiamo

$$|\Phi(\mathbf{y}_1) - \Phi(\mathbf{y}_2)| \le \left|(D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0))^{-1}\right| |\sigma(\mathbf{x}, \mathbf{y}_1) - \sigma(\mathbf{x}, \mathbf{y}_2)| = C_1 \left|\sigma(\mathbf{x}, \mathbf{y}_1) - \sigma(\mathbf{x}, \mathbf{y}_2)\right|. \tag{11.16}$$

Per stimare quest'ultima espressione torniamo a (11.7) e consideriamo

$$F(\mathbf{x},\mathbf{y}_2)=D_{\mathbf{x}}F(\mathbf{x}_0,\mathbf{y}_0)(\mathbf{x}-\mathbf{x}_0)+D_{\mathbf{y}}F(\mathbf{x}_0,\mathbf{y}_0)(\mathbf{y}_2-\mathbf{y}_0)+\sigma(\mathbf{x},\mathbf{y}_2)$$
$$F(\mathbf{x},\mathbf{y}_1)=D_{\mathbf{x}}F(\mathbf{x}_0,\mathbf{y}_0)(\mathbf{x}-\mathbf{x}_0)+D_{\mathbf{y}}F(\mathbf{x}_0,\mathbf{y}_0)(\mathbf{y}_1-\mathbf{y}_0)+\sigma(\mathbf{x},\mathbf{y}_1).$$

Facendo la differenza, otteniamo

$$\begin{aligned}\sigma(\mathbf{x},\mathbf{y}_2)-\sigma(\mathbf{x},\mathbf{y}_1)&=F(\mathbf{x},\mathbf{y}_2)-F(\mathbf{x},\mathbf{y}_1)-D_{\mathbf{y}}F(\mathbf{x}_0,\mathbf{y}_0)(\mathbf{y}_2-\mathbf{y}_1)\\&=F(\mathbf{x},t(\mathbf{y}_2-\mathbf{y}_1)+\mathbf{y}_1)]_{t=0}^{t=1}-D_{\mathbf{y}}F(\mathbf{x}_0,\mathbf{y}_0)(\mathbf{y}_2-\mathbf{y}_1).\end{aligned}$$

Siccome dal Teorema Fondamentale del Calcolo e dalla Regola della Catena abbiamo

$$\begin{aligned}F(\mathbf{x},t(\mathbf{y}_2-\mathbf{y}_1)+\mathbf{y}_1)]_{t=0}^{t=1}&=\int_0^1\frac{d}{dt}\left(F(\mathbf{x},t(\mathbf{y}_2-\mathbf{y}_1)+\mathbf{y}_1)\right)dt\\&=\int_0^1 D_{\mathbf{y}}F\left(\mathbf{x},\mathbf{y}_1+t(\mathbf{y}_2-\mathbf{y}_1)\right)dt(\mathbf{y}_2-\mathbf{y}_1),\end{aligned}$$

dove, per la definizione dell'integrale, sfruttiamo $D_{\mathbf{y}}F\in C^0(A,\mathcal{L}(\mathbb{R}^N,\mathbb{R}^N))$, che ci garantisce che la funzione nell'integrale è continua in t, e che pertanto l'integrale è ben definito.
Dalla precedente formula, segue che

$$\sigma(\mathbf{x},\mathbf{y}_2)-\sigma(\mathbf{x},\mathbf{y}_1)=\int_0^1\left[D_{\mathbf{y}}F\left(\mathbf{x},\mathbf{y}_1+t(\mathbf{y}_2-\mathbf{y}_1)\right)-D_{\mathbf{y}}F\left(\mathbf{x}_0,\mathbf{y}_0\right)\right]dt(\mathbf{y}_2-\mathbf{y}_1).$$

Siccome l'insieme $\overline{D(\mathbf{x}_0,r_1)}\times\overline{D(\mathbf{y}_0,r_2)}$ è convesso, risulta

$$(\mathbf{x},\mathbf{y}_1+t(\mathbf{y}_2-\mathbf{y}_1))\in\overline{D(\mathbf{x}_0,r_1)}\times\overline{D(\mathbf{y}_0,r_2)}\text{ per ogni }t\in[0,1],$$

Allora dalla (11.14) segue

$$|D_{\mathbf{y}}F\left(\mathbf{x},\mathbf{y}_1+t(\mathbf{y}_2-\mathbf{y}_1)\right)-D_{\mathbf{y}}F\left(\mathbf{x}_0,\mathbf{y}_0\right)|<\varepsilon_0\text{ per ogni }t\in[0,1],\ x\in\overline{D(\mathbf{x}_0,r_1)}\text{ e }\mathbf{y}_1,\mathbf{y}_2\in\overline{D(\mathbf{y}_0,r_2)}.$$

Ma allora

$$\begin{aligned}&\left|\int_0^1\left[D_{\mathbf{y}}F\left(\mathbf{x},\mathbf{y}_1+t(\mathbf{y}_2-\mathbf{y}_1)\right)-D_{\mathbf{y}}F\left(\mathbf{x}_0,\mathbf{y}_0\right)\right]dt\right| \qquad (11.17)\\&\le\int_0^1\left|D_{\mathbf{y}}F\left(\mathbf{x},\mathbf{y}_1+t(\mathbf{y}_2-\mathbf{y}_1)\right)-D_{\mathbf{y}}F\left(\mathbf{x}_0,\mathbf{y}_0\right)\right|dt<\varepsilon_0\ \forall\ t\in[0,1],\ \mathbf{x}\in\overline{D(\mathbf{x}_0,r_1)}\text{ e }\mathbf{y}_1,\mathbf{y}_2\in\overline{D(\mathbf{y}_0,r_2)}.\end{aligned}$$

Segue

$$|\sigma(\mathbf{x},\mathbf{y}_2)-\sigma(\mathbf{x},\mathbf{y}_1)|\le\varepsilon_0|\mathbf{y}_2-\mathbf{y}_1|$$

Se inseriamo questo in (11.16) otteniamo

$$|\Phi(\mathbf{y}_1)-\Phi(\mathbf{y}_2)|\le C_1\left|\sigma(\mathbf{x},\mathbf{y}_1)-\sigma(\mathbf{x},\mathbf{y}_2)\right|\le C_1\varepsilon_0|\mathbf{y}_2-\mathbf{y}_1|.$$

Dalla condizione $C_1\varepsilon_0 < 1/2$ in (11.15), concludiamo

$$|\Phi(\mathbf{y}_1) - \Phi(\mathbf{y}_2)| \le \frac{1}{2}|\mathbf{y}_2 - \mathbf{y}_1| \text{ per ogni } \mathbf{x} \in \overline{D(\mathbf{x}_0, r_1)} \text{ e ogni } \mathbf{y}_1, \mathbf{y}_2 \in \overline{D(\mathbf{y}_0, r_2)}. \tag{11.18}$$

Pertanto possiamo concludere che la funzione in (11.13) data da (11.9) è una contrazione in $\overline{D(\mathbf{y}_0, r_2)}$. Ne segue che per ogni $|\mathbf{x}-\mathbf{x}_0| \le r_1$ esiste esattamente un punto fisso $|\mathbf{y}-\mathbf{y}_0| \le r_2$. Quindi abbiamo definito la funzione

$$f : \overline{D(\mathbf{x}_0, r_1)} \to \overline{D(\mathbf{y}_0, r_2)}. \tag{11.19}$$

con la proprietà che $F(\mathbf{x}, \mathbf{y}) = 0$ per $(\mathbf{x}, \mathbf{y}) \in D(\mathbf{x}_0, r_1) \times D(\mathbf{y}_0, r_2)$ se e solo se $\mathbf{y} = f(\mathbf{x})$.

Dimostrazione di **b**. Suddividiamo la dimostrazione in due parti distinte. Nella prima dimostriamo che f è continua, nella seconda che $f \in C^1$.

Dimostrazione che $f \in C^0(D(\mathbf{x}_0, r_1), \mathbb{R}^N)$. Fissiamo $\mathbf{x}_1 \in D(\mathbf{x}_0, r_1)$ e consideriamo $\mathbf{x} \in D(\mathbf{x}_0, r_1)$. Allora

$$\begin{aligned}
0 &= F(\mathbf{x}, f(\mathbf{x})) - F(\mathbf{x}_1, f(\mathbf{x}_1)) = F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1))]_{t=0}^{t=1} \\
&= \int_0^1 \frac{d}{dt}\left(F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1))\right) dt \\
&= \left(\int_0^1 D_{\mathbf{x}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt\right)(\mathbf{x} - \mathbf{x}_1) \\
&+ \left(\int_0^1 D_{\mathbf{y}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt\right)(f(\mathbf{x}) - f(\mathbf{x}_1)).
\end{aligned} \tag{11.20}$$

Notare, da (11.17), che

$$\left|\int_0^1 D_{\mathbf{y}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt - D_{\mathbf{y}} F(\mathbf{x}_0, \mathbf{y}_0)\right| < \varepsilon_0. \tag{11.21}$$

Allora, la condizione (11.15) e il Lemma 11.1 garantiscono

$$\left|\left(\int_0^1 D_{\mathbf{y}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt\right)^{-1}\right| \le 2C_1. \tag{11.22}$$

Da (11.20) ricaviamo

$$f(\mathbf{x}) - f(\mathbf{x}_1) = -\widetilde{G}(\mathbf{x}, \mathbf{x}_1)(\mathbf{x} - \mathbf{x}_1) \text{ dove} \tag{11.23}$$

$$\widetilde{G}(\mathbf{x}, \mathbf{x}_1) := \left(\int_0^1 D_{\mathbf{y}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt\right)^{-1} \int_0^1 D_{\mathbf{x}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt$$

dove

$\widetilde{G}(\mathbf{x}, \mathbf{x}_1) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$con

$$\left|\widetilde{G}(\mathbf{x}, \mathbf{x}_1)\right| \le \left|\left(\int_0^1 D_{\mathbf{y}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt\right)^{-1}\right| \left|\int_0^1 D_{\mathbf{x}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt\right|$$

$$\le 2C_1 \left|\int_0^1 D_{\mathbf{x}} F(t(\mathbf{x}, f(\mathbf{x})) + (1-t)(\mathbf{x}_1, f(\mathbf{x}_1)) dt\right|.$$

Poi usiamo

$$\left|\int_0^1 D_{\mathbf{x}}F(t(\mathbf{x},f(\mathbf{x}))+(1-t)(\mathbf{x}_1,f(\mathbf{x}_1))dt\right| \leq \int_0^1 |D_{\mathbf{x}}F(t(\mathbf{x},f(\mathbf{x}))+(1-t)(\mathbf{x}_1,f(\mathbf{x}_1))|dt$$
$$\leq \sup\{|D_{\mathbf{x}}F(\mathbf{x},\mathbf{y})| \text{ t.c. } (\mathbf{x},\mathbf{y}) \in \overline{D(\mathbf{x}_0,r_1)} \times \overline{D(\mathbf{y}_0,r_2)}\} =: L < +\infty.$$

In conclusione, abbiamo

$$\left|\widetilde{G}(\mathbf{x},\mathbf{x}_1)\right| \leq 2C_1 L \text{ per ogni } \mathbf{x},\mathbf{x}_1 \in D(\mathbf{x}_0,r_1).$$

Pertanto

$$|f(\mathbf{x})-f(\mathbf{x}_1)| \leq 2C_1 L\ |\mathbf{x}-\mathbf{x}_1| \text{ per ogni } \mathbf{x},\mathbf{x}_1 \in D(\mathbf{x}_0,r_1),$$

che dimostra che f è continua (infatty, Lipschitziana con costante di Lipschitz $2C_1L$) in $D(\mathbf{x}_0,r_1)$.

Dimostrazione della differenziabilità di f in $D(\mathbf{x}_0,r_1)$. Partendo dalla formula (11.23) ed utilizzando la formula (11.6), quest'ultima una formula ancora da dimostrare, ma che ci fornisce un candidato per il differenziale $Df(\mathbf{x}_1)$, abbiamo

$$\frac{f(\mathbf{x})-f(\mathbf{x}_1)+(D_{\mathbf{y}}F(\mathbf{x}_1,f(\mathbf{x}_1)))^{-1}D_{\mathbf{x}}F(\mathbf{x}_1,f(\mathbf{x}_1))(\mathbf{x}-\mathbf{x}_1)}{|\mathbf{x}-\mathbf{x}_1|} = \widetilde{\sigma}(\mathbf{x},\mathbf{x}_1) \text{ dove}$$
$$\widetilde{\sigma}(\mathbf{x},\mathbf{x}_1) = \Big((D_{\mathbf{y}}F(\mathbf{x}_1,f(\mathbf{x}_1)))^{-1}D_{\mathbf{x}}F(\mathbf{x}_1,f(\mathbf{x}_1)) - \widetilde{G}(\mathbf{x},\mathbf{x}_1)\Big)\frac{\mathbf{x}-\mathbf{x}_1}{|\mathbf{x}-\mathbf{x}_1|}.$$

Otteniamo che f è differenziabile in $\mathbf{x}_1$ con

$$Df(\mathbf{x}_1) = -(D_{\mathbf{y}}F(\mathbf{x}_1,f(\mathbf{x}_1)))^{-1}D_{\mathbf{x}}F(\mathbf{x}_1,f(\mathbf{x}_1)) \tag{11.24}$$

se riusciamo a dimostrare $\lim_{\mathbf{x}\to\mathbf{x}_1}\widetilde{\sigma}(\mathbf{x},\mathbf{x}_1)=0$. Quest'ultimo limite segue dal seguente limite, che dimostriamo sotto,

$$\lim_{\mathbf{x}\to\mathbf{x}_1}\widetilde{G}(\mathbf{x},\mathbf{x}_1) = (D_{\mathbf{y}}F(\mathbf{x}_1,f(\mathbf{x}_1)))^{-1}D_{\mathbf{x}}F(\mathbf{x}_1,f(\mathbf{x}_1)).$$

Dalla definizione di $\widetilde{G}$ in (11.23) e dalla regola del prodotto, abbiamo

$$\lim_{\mathbf{x}\to\mathbf{x}_1}\widetilde{G}(\mathbf{x},\mathbf{x}_1)$$
$$= \lim_{\mathbf{x}\to\mathbf{x}_1}\left(\int_0^1 D_{\mathbf{y}}F(t(\mathbf{x},f(\mathbf{x}))+(1-t)(\mathbf{x}_1,f(\mathbf{x}_1))dt\right)^{-1}\lim_{\mathbf{x}\to\mathbf{x}_1}\int_0^1 D_{\mathbf{x}}F(t(\mathbf{x},f(\mathbf{x}))+(1-t)(\mathbf{x}_1,f(\mathbf{x}_1))dt,$$

sempre che sia un caso definito. Siccome $f\in C^0(D(\mathbf{x}_0,r_1),\mathbb{R}^N)$, è elementare concludere che

$$\lim_{\mathbf{x}\to\mathbf{x}_1}\int_0^1 D_{\mathbf{x}}F(t(\mathbf{x},f(\mathbf{x}))+(1-t)(\mathbf{x}_1,f(\mathbf{x}_1))dt = D_{\mathbf{x}}F(\mathbf{x}_1,f(\mathbf{x}_1))$$

e, analogamente,

$$\lim_{\mathbf{x}\to\mathbf{x}_1}\int_0^1 D_{\mathbf{y}}F(t(\mathbf{x},f(\mathbf{x}))+(1-t)(\mathbf{x}_1,f(\mathbf{x}_1))dt = D_{\mathbf{y}}F(\mathbf{x}_1,f(\mathbf{x}_1)).$$

Segue dall'Esercizio 11.2 che quest'ultimo limite implica

$$\lim_{\mathbf{x}\to\mathbf{x}_1}\left(\int_0^1 D_{\mathbf{y}}F(t(\mathbf{x},f(\mathbf{x}))+(1-t)(\mathbf{x}_1,f(\mathbf{x}_1))dt\right)^{-1} = (D_{\mathbf{y}}F(\mathbf{x}_1,f(\mathbf{x}_1)))^{-1}.$$

Resta così dimostrata la formula (11.24), e cioè la (11.6). Il fatto che $f \in C^1(D(\mathbf{x}_0,r_1),\mathbb{R}^N)$, segue dalla formula (11.6).

□

Osservazione 11.5. Si noti che la dimostrazione è ispirata al *metodo delle tangenti* per la ricerca di soluzioni di equazioni $f(x)=0$, dove si imposta il problema in quello della ricerca di punti fissi della trasformazione $x \to x - \dfrac{f(x)}{f'(x)}$.

Esempio 11.6. L'insieme di livello $\{F=0\}$ dove $F(x,y)=y^2+x^4-x^2$ ha la forma di un otto orizzontale con estremi nei punti $(\pm 1,0)$. Notiamo che il gradiente $\nabla F(x,y)=(4x^3-2x,2y)$ si annulla nei punti $(\pm 1/\sqrt{2},0)$ e $(0,0)$. Il punto $(0,0)$ appartiene a $\{F=0\}$: in esso non possiamo applicare il teorema della funzione implicita. A posteriori ne capiamo il motivo: l'origine è il centro dell'otto che costituisce l'insieme ed esso, in un intorno dell'origine sufficientemente piccolo, risulta l'unione di due differenti archi di curva (in questo contesto non importa che in realtà *lontano dall'origine* essi si congiungono formando un'unica curva).

Nei punti $(x,y)\in\{F=0\}$ con $y\neq 0$ possiamo applicare il Teorema della Funzione Imlicita (*isolare* la y) e trovare $f:U\to V$, $f(x)=\pm\sqrt{x^2-x^4}$. Si noti che l'insieme U dato dal teorema è $(-1,0)$ oppure $(0,1)$ e non $(-1,1)$ (il pezzo di grafico in cui la funzione resta positiva, o negativa). Si noti inoltre che f non sarebbe derivabile in $x=0$, quindi non può essere $0\in U$. Riassumendo abbiamo le seguenti quattro alternative:

- se $(x,y)\in\{F=0\}$, $x>0$ e $y>0$ troviamo $\varphi:(0,1)\to(0,1)$ con $f(x)=\sqrt{x^2-x^4}$
- se $(x,y)\in\{F=0\}$, $x>0$ e $y<0$ troviamo $\varphi:(0,1)\to(-1,0)$ con $f(x)=-\sqrt{x^2-x^4}$
- se $(x,y)\in\{F=0\}$, $x<0$ e $y>0$ troviamo $\varphi:(-1,0)\to(0,1)$ con $f(x)=\sqrt{x^2-x^4}$
- se $(x,y)\in\{F=0\}$, $x<0$ e $y<0$ troviamo $\varphi:(-1,0)\to(-1,0)$ con $f(x)=-\sqrt{x^2-x^4}$

Se invece consideriamo l'insieme di livello $\{F=-1/4\}$, esso è costituito dai soli due punti $(\pm 1/\sqrt{2},0)$ e nulla più. Anche in questo caso notiamo che è *normale* che non si possa applicare il teorema della funzione implicita in questo caso.

Esempio 11.7. Sia $F : \mathbb{R}^3 \to \mathbb{R}$, con $F(x, y, z) = x^2 + y^2 + z^2 - 1$. Consideriamo il punto $P_0 = (0, 0, 1) \in \{F = 0\}$ (l'insieme di livello $\{F = 0\}$ è la sfera unitaria) Calcoliamo la matrice Jacobiana associata a F e troviamo

$$J_F(x, y, z) = \nabla F(x, y, z) = (\underbrace{2x, 2y}_{\frac{\partial F}{\partial \mathrm{x}}(P)} \mid \underbrace{2z}_{\frac{\partial F}{\partial \mathrm{y}}(P)})$$

Notiamo che valgono le ipotesi del teorema della funzione implicita per la nostra scelta di P_0, di conseguenza sappiamo che esiste un intorno $U \subseteq \mathbb{R}^2$ di $\mathrm{x}_0 = (0, 0)$ e un intorno $V \subseteq \mathbb{R}$ di $\mathrm{y}_0 = 1$ e una funzione $\varphi : U \to V$ tale che $z = f(x, y)$ e $F(x, y, f(x, y)) = 0$ per ogni $(x, y) \in U$. Inoltre vale

$$\nabla f(x, y) = -\left[\frac{\partial F}{\partial z}(x, y, f(x, y))\right]^{-1} \left(\frac{\partial F}{\partial x}(x, y, f(x, y)), \frac{\partial F}{\partial y}(x, y, f(x, y))\right)$$

da cui segue

$$\nabla f(x, y) = -\frac{1}{2f(x, y)}(2x, 2y) = \left(-\frac{x}{f(x, y)}, -\frac{y}{f(x, y)}\right)$$

in particolare per $(x, y) = \mathrm{x}_0 = (0, 0)$ troviamo $\nabla f(x, y) = (0, 0)$.

In questo caso la funzione f è facile da individuare, in quanto è facile isolare con semplici calcoli la variabile z, trovando

$$z = f(x, y) = \sqrt{1 - x^2 - y^2}\,.$$

Notiamo che l'aperto U deve essere necessariamente contenuto nella palla unitaria di $\mathbb{R}^2$ centrata nell'origine.

Osservazione 11.8. Un caso particolare del Teorema 11.3 ci fornisce il teorema della mappa aperta che ha il seguente enunciato:

Dati $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{R}^d$, un aperto $V_1 \subseteq \mathbb{R}^d$ con $\mathbf{y}_0 \in V_1$. Consideriamo poi una funzione

$$V_1 \ni \mathbf{y} \to g(\mathbf{y}) \in \mathbb{R}^d \text{ con } g(\mathbf{y}_0) = \mathbf{x}_0$$

Supponiamo che sia $g \in C^1(V_1, \mathbb{R}^d)$ e che $D_{\mathbf{y}}g(\mathbf{y}_0) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ sia invertibile.
Allora esistono un intorno V di $\mathbf{y}_0$ ed un intorno U di $\mathbf{x}_0$ tali che $g : V \to U$ è un diffeomorfismo C^1, cioè esiste una funzione inversa $f : U \to V$ che è in $C^1(U, \mathbb{R}^d)$.

Per dimostrare il teorema basta applicare il teorema della funzione implicita alla funzione $F(\mathbf{x}, \mathbf{y}) = \mathbf{x} - g(\mathbf{y})$ che è definita in $A = \mathbb{R}^d \times V_1$ ed a valori in $\mathbb{R}^d$. La funzione f è esattamente la funzione implicita.

12 Ricerca di massimi e minimi su spazi curvi: moltiplicatori di Lagrange

Osservazione 12.1. Consideriamo la situazione del teorema. Allora l'insieme $F(\mathbf{x}, \mathbf{y}) = 0$ vicino a $(\mathbf{x}_0, \mathbf{y}_0)$ può essere rappresentato com il grafico $\mathbf{y} = f(\mathbf{x})$. Lo spazio tangente a questo insieme nel punto $(\mathbf{x}_0, \mathbf{y}_0)$ è rappresentato dall'equazione

$$\mathbf{y} = D_{\mathbf{x}} f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \mathbf{y}_0$$

I vettori tangenti sono della forma $(\mathbf{u}, D_{\mathbf{x}} f(\mathbf{x}_0)\mathbf{u})$ al variare di $\mathbf{u} \in \mathbb{R}^d$.

Osserviamo che se considero l'identità $F(\mathbf{x}, f(\mathbf{x})) = 0$, allora

$$0 = D_{\mathbf{x}} \left(F(\mathbf{x}, f(\mathbf{x})) \right) = D_{\mathbf{x}} F(\mathbf{x}, f(\mathbf{x})) + D_{\mathbf{y}} F(\mathbf{x}, f(\mathbf{x})) D_{\mathbf{x}} f(\mathbf{x}).$$

Notare che quello nella formula è un operatore $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^N)$. Se lo applichiamo ad un vettore $\mathbf{u} \in \mathbb{R}^d$ otteniamo

$$DF(\mathbf{x}, f(\mathbf{x})) \left(\mathbf{u}, D_{\mathbf{x}} f(\mathbf{x})\mathbf{u} \right) = 0 \text{ per ogni } \mathbf{x} \text{ e per ogni } \mathbf{u} \in \mathbb{R}^d. \tag{12.1}$$

Cioè, per esempio nel punto $(\mathbf{x}_0, \mathbf{y}_0)$, $\ker DF(\mathbf{x}_0, \mathbf{y}_0)$ contiene i vettori tangenti. Più precisamente, il nucleo coincide esattamente con l'insieme dei vettori tangenti. Infatti, se

$$DF(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{u}, \mathbf{v}) = 0 \text{ per un } (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^d \times \mathbb{R}^N,$$

questo può essere riscritto come

$$D_{\mathbf{x}} F(\mathbf{x}_0, \mathbf{y}_0)\mathbf{u} + D_{\mathbf{y}} F(\mathbf{x}_0, \mathbf{y}_0)\mathbf{v} = 0 \Longleftrightarrow \mathbf{v} = -\left(D_{\mathbf{y}} F(\mathbf{x}_0, \mathbf{y}_0) \right)^{-1} D_{\mathbf{x}} F(\mathbf{x}_0, \mathbf{y}_0)\mathbf{u}.$$

Ma (11.6) ci dice che quest'ultima uguaglianza non è altro che $\mathbf{v} = Df(\mathbf{x}_0)\mathbf{u}$, cioè che $(\mathbf{u}, \mathbf{v})$ è tangente all'insieme $F = 0$ (vedere sotto).

Osservazione 12.2. Consideriamo la situazione del teorema e supponiamo che $\mathcal{C}$ sia l'insieme delle soluzioni di $F = 0$. Supponiamo di avere una funzione $G : A \to \mathbb{R}$. Consideriamo $G : \mathcal{C} \to \mathbb{R}$ e cerchiamo di volere capire se $(\mathbf{x}_0, \mathbf{y}_0)$ è un punto di massimo o di minimo per G su $\mathcal{C}$. Il Teorema della Funzione Implicita ci offre uno strumento per risolvere il problema. Infatti $\mathcal{C}$ è il grafico di $\mathbf{y} = f(\mathbf{x})$ e dunque $(\mathbf{x}_0, \mathbf{y}_0)$ è un punto di massimo o di minimo per G su $\mathcal{C}$ esattamente se $w(\mathbf{x}) := G(\mathbf{x}, f(\mathbf{x}))$ ha $\mathbf{x}_0$ come punto di massimo o di minimo. Notiamo che $w : U \to \mathbb{R}$ è una normale funzione definita in un aperto di $\mathbb{R}^d$ a valori in $\mathbb{R}$. Quindi possiamo applicare il Teorema di Fermat e concludere che perché $\mathbf{x}_0$ sia un punto di massimo o di minimo $G : \mathcal{C} \to \mathbb{R}$, cioè di w, deve essere un punto critico di w, ossia con $\mathbf{y}_0 = f(\mathbf{x}_0)$ abbiamo:

$$\begin{aligned} D_{\mathbf{x}} w(\mathbf{x}_0)\mathbf{u} &= D_{\mathbf{x}} G\left(\mathbf{x}_0, \mathbf{y}_0\right)\mathbf{u} + D_{\mathbf{y}} G\left(\mathbf{x}_0, \mathbf{y}_0\right) D_{\mathbf{x}} f(\mathbf{x}_0)\mathbf{u} \\ &= DG\left(\mathbf{x}_0, \mathbf{y}_0\right)\left(\mathbf{u}, D_{\mathbf{x}} f(\mathbf{x}_0)\mathbf{u}\right) = 0 \text{ per ogni } \mathbf{u} \in \mathbb{R}^d. \end{aligned} \tag{12.2}$$

Osservazione 12.3. Consideriamo la situazione di Osservazione 12.2. Ricapitolando, abbiamo quanto segue:

1. $(\mathbf{x}_0, \mathbf{y}_0) \in A \subseteq \mathbb{R}^d \times \mathbb{R}^N$ con $(\mathbf{x}_0, \mathbf{y}_0) \in A$ ed A aperto;
2. $A \ni (\mathbf{x}, \mathbf{y}) \to F(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N$;
3. $F \in C^1(A, \mathbb{R}^N)$ e $D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$ invertibile;
4. denoto con $\mathcal{C}$ sia l'insieme delle soluzioni di $F(\mathbf{x}, \mathbf{y}) = 0$, con $(\mathbf{x}_0, \mathbf{y}_0) \in \mathcal{C}$;
5. è assegnata $G \in C^1(A, \mathbb{R})$.

Abbiamo osservato che se $(\mathbf{x}_0, \mathbf{y}_0)$ è un punto di massimo o di minimo locale della restrizione $G|_{\mathcal{C}}$ allora si ha (12.2). Esaminiamone meglio il significato. Per prima cosa osserviamo, partendo da

$$F(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} F_1(\mathbf{x}, \mathbf{y}) \\ \vdots \\ F_N(\mathbf{x}, \mathbf{y}) \end{pmatrix}$$

che

$$DF(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \nabla F_1(\mathbf{x}, \mathbf{y}) \\ \vdots \\ \nabla F_N(\mathbf{x}, \mathbf{y}) \end{pmatrix} \text{ dove } \nabla F_j = (\partial_{x_1} F_j, ..., \partial_{x_d} F_j, \partial_{y_1} F_j, ..., \partial_{y_N} F_j) \, .$$

Notiamo che

$$\begin{aligned} &DF(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} D_{\mathbf{x}}F(\mathbf{x}, \mathbf{y}) & D_{\mathbf{y}}F(\mathbf{x}, \mathbf{y}) \end{pmatrix} \text{ con} \\ &D_{\mathbf{x}}F = \begin{pmatrix} \partial_{x_1} F_1 & \dots & \partial_{x_d} F_1 \\ \vdots & \vdots & \vdots \\ \vdots & & \\ \partial_{x_1} F_N & \dots & \partial_{x_d} F_N \end{pmatrix} \text{ e con } D_{\mathbf{y}}F = \begin{pmatrix} \partial_{y_1} F_1 & \dots & \partial_{y_N} F_1 \\ \vdots & \vdots & \vdots \\ \vdots & & \\ \partial_{y_1} F_N & \dots & \partial_{y_N} F_N \end{pmatrix} . \end{aligned}$$

Siccome $\det D_{\mathbf{y}}F \neq 0$, le righe che formano $D_{\mathbf{y}}F$ sono linearmente indipendenti, il che significa che i vettori ∇F_1,..., ∇F_N sono linearmente indipendenti. Notiamo che siccome nel punto $(\mathbf{x}_0, \mathbf{y}_0)$

$$\begin{aligned} 0 &= DF(\mathbf{x}_0, \mathbf{y}_0)(\mathbf{u}, \mathbf{v}) \\ &= \begin{pmatrix} \partial_{x_1} F_1 & \dots & \partial_{x_d} F_1 & \partial_{y_1} F_1 & \dots & \partial_{y_N} F_1 \\ \vdots & & & & & \\ \partial_{x_1} F_N & \dots & \partial_{x_d} F_N & \partial_{y_1} F_N & \dots & \partial_{y_N} F_N \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_d \\ v_1 \\ \vdots \\ v_N \end{pmatrix} \text{ per ogni } (\mathbf{u}, \mathbf{v}) \in T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C} \end{aligned}$$

abbiamo che i seguenti prodotti scalari sono nulli

$$\nabla F_j(\mathbf{x}_0, \mathbf{y}_0) \cdot (\mathbf{u}, \mathbf{v}) = 0 \text{ per ogni } (\mathbf{u}, \mathbf{v}) \in T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C} \text{ e } j = 1, ...N.$$

In modo simile

$$\nabla G(\mathbf{x}_0, \mathbf{y}_0) \cdot (\mathbf{u}, \mathbf{v}) = 0 \text{ per ogni } (\mathbf{u}, \mathbf{v}) \in T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C}.$$

Questo significa che i vettori

$$\nabla F_1(\mathbf{x}_0, \mathbf{y}_0), ..., \nabla F_N(\mathbf{x}_0, \mathbf{y}_0), \nabla G(\mathbf{x}_0, \mathbf{y}_0) \text{ appartengono a } \left(T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C}\right)^{\perp}$$

dove $\left(T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C}\right)^{\perp}$ è il sottospazio di $\mathbb{R}^d \times \mathbb{R}^N$ formato dai vettori che sono ortogonali allo spazio tangente $T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C}$. Ma siccome

$$\mathbb{R}^d \times \mathbb{R}^N = T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C} \bigoplus \left(T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C}\right)^{\perp}$$

e $\dim T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C} = d$, segue che $\dim \left(T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C}\right)^{\perp} = N$. Siccome ora i vettori

$$\nabla F_1(\mathbf{x}_0, \mathbf{y}_0), ..., \nabla F_N(\mathbf{x}_0, \mathbf{y}_0)$$

sono linearmente indipendenti, essi formano una base di $\left(T_{(\mathbf{x}_0, \mathbf{y}_0)}\mathcal{C}\right)^{\perp}$. Pertanto

$$\exists, \text{ unica, una scelta di scalari } \lambda_1, ..., \lambda_N \in \mathbb{R}, \text{ tale che } \nabla G(\mathbf{x}_0, \mathbf{y}_0) = \sum_{j=1}^{N} \lambda_j \nabla F_j(\mathbf{x}_0, \mathbf{y}_0). \tag{12.3}$$

Le Osservazioni 12.1–12.3 possono essere riassunte nel seguente teorema.

Theorem 12.4 (Moltiplicatori di Lagrange)**.** *Consideriamo due funzioni $F \in C^1(A, \mathbb{R}^N)$ e $G \in C^1(A, \mathbb{R})$ con $A \subseteq \mathbb{R}^{d+N}$ aperto, e $d > 0$. Sia*

$$F(\mathbf{z}) = \begin{pmatrix} F_1(\mathbf{z}) \\ \vdots \\ F_N(\mathbf{z}) \end{pmatrix}$$

Consideriamo $\mathbf{z}_0 \in A$ dove $F(\mathbf{z}_0) = 0$ e supponiamo che $DF(\mathbf{z}_0) : \mathbb{R}^{d+N} \to \mathbb{R}^N$ sia suriettiva. Infine denotiamo con $\mathcal{C}$ il sottoinsieme di A formato dalle soluzioni dell'equazione $F(\mathbf{z}) = 0$. Allora, se $\mathbf{z}_0$ è un punto di massimo o di minimo locale per la restrizione $G|_{\mathcal{C}} : \mathcal{C} \to \mathbb{R}$

$$\exists, \text{ unica, una scelta di scalari } \lambda_1, ..., \lambda_N \in \mathbb{R}, \text{ tale che } \nabla G(\mathbf{z}_0) = \sum_{j=1}^{N} \lambda_j \nabla F_j(\mathbf{z}_0). \tag{12.4}$$

I numeri $\lambda_1, ..., \lambda_N \in \mathbb{R}$ vengono chiamati moltiplicatori di Lagrange.

Dim. A meno di un cambio di coordinate, possiamo supporre che $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d \times \mathbb{R}^N$ con $D_y F(\mathbf{x}_0, \mathbf{y}_0)$ un isomorfismo in $\mathbb{R}^N$, dove $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0)$. Ora, se $(\mathbf{x}_0, \mathbf{y}_0)$ è un punto di massimo o di minimo locale per $G|_{\mathcal{C}}$, segue che abbiamo (12.3), e questo dà (12.4) e completa la dimostrazione. □

Osservazione 12.5. Il Teorema 12.4 è semplicemente una generalizzazione del Teorema di Fermat 6.18 al caso in cui una funzione scalere, invece di essere definita in un sottoinsieme aperto di uno spazio Euclideo, è definita su un aperto di uno spazio curvo (più precisamente, di un varietà differenziabile). In particolare, possiamo parafrasare il Teorema 12.4 dicendo che punti interni alla regione curva, se sono punti estremali debbono essere punti critici.

Osservazione 12.6. Operativamente, per trovare i punti estremali, si cercano i punti critici, risolvendo il sistema

$$\begin{cases} \partial_1 G(x_1, ..., x_d) = \sum_{j=1}^N \lambda_j \partial_1 F_j(x_1, ..., x_d) \\ \qquad \cdots \\ \partial_d G(x_1, ..., x_d) = \sum_{j=1}^N \lambda_j \partial_d F_j(x_1, ..., x_d) \\ \qquad F_1(x_1, ..., x_d) = 0 \\ \qquad \cdots \\ \qquad F_N(x_1, ..., x_d) = 0. \end{cases} \tag{12.5}$$

E' un sistema con incognite $(x_1, ..., x_d, \lambda_1, ..., \lambda_N) \in \mathbb{R}^{d+N}$ e con esattamente $d+N$ equazioni.

Osservazione 12.7. I moltiplicatori di Lagrange hanno un significato profondo. Ad esempio, supponiamo di dovere trovare trovare il massimo di $G(x, y)$ soggetta al vincolo $F(x, y) = b$ dove b varia in un certo intervallo I. Supponiamo che per ogni $b \in I$ ci sia esattamente un unico punto di massimo vincolato $(x(b), y(b))$. Allora abbiamo

$$\begin{aligned} &\partial_x G(x(b), y(b)) = \lambda(b) \partial_x F(x(b), y(b)) \\ &\partial_y G(x(b), y(b)) = \lambda(b) \partial_y F(x(b), y(b)) \\ &F(x(b), y(b)) = b \end{aligned}$$

ed abbiamo che $G(x(b), y(b))$ è il valore massimo di G sul vincolo $F(x, y) = b$. Ora abbiamo

$$\frac{d}{db} F(x(b), y(b)) = \partial_x F(x(b), y(b)) \frac{d}{db} x(b) + \partial_y F(x(b), y(b)) \frac{d}{db} y(b) = 1.$$

Moltiplicando per $\lambda(b)$ otteniamo

$$\begin{aligned} \lambda(b) &= \lambda(b) \partial_x F(x(b), y(b)) \frac{d}{db} x(b) + \lambda(b) \partial_y F(x(b), y(b)) \frac{d}{db} y(b) \\ &= \partial_x G(x(b), y(b)) \frac{d}{db} x(b) + \partial_y G(x(b), y(b)) \frac{d}{db} y(b) \\ &= \frac{d}{db} G(x(b), y(b)). \end{aligned}$$

Quindi concludiamo che

$$\frac{d}{db}G(x(b), y(b)) = \lambda(b), \tag{12.6}$$

che è un fatto notevole e che può essere generalizzato più in generale, si veda (12.9) sotto.

Osservazione 12.8. Più in generale, consideriamo il sistema

$$\begin{cases} \partial_1 G(x_1, ..., x_d) = \sum_{j=1}^{N} \lambda_j \partial_1 F_j(x_1, ..., x_d) \\ \dots \\ \partial_d G(x_1, ..., x_d) = \sum_{j=1}^{N} \lambda_j \partial_d F_j(x_1, ..., x_d) \\ F_1(x_1, ..., x_d) = b_1 \\ \dots \\ F_N(x_1, ..., x_d) = b_N \end{cases} \tag{12.7}$$

e supponiamo che per $\mathbf{b} \in \mathbb{R}^N$ il vettore con coordinate $b_1, ..., b_N$. G abbia un punto di massimo in $\mathbf{x}(\mathbf{b})$. I correspondenti moltiplicatori di Lagrange saranno funzioni $\lambda_j(\mathbf{b})$ di $\mathbf{b}$. Abbiamo

$$\begin{aligned} \partial_{b_k} F_j(\mathbf{x}(\mathbf{b})) &= \delta_{jk} \\ &= \sum_{\alpha=1}^{d} \partial_{x_\alpha} F_j(\mathbf{x}(\mathbf{b})) \partial_{b_k} x_\alpha(\mathbf{b}). \end{aligned}$$

Pertanto, moltiplicando per $\lambda_j(\mathbf{b})$

$$\sum_{\alpha=1}^{d} \lambda_j(\mathbf{b}) \partial_{x_\alpha} F_j(\mathbf{x}(\mathbf{b})) \partial_{b_k} x_\alpha(\mathbf{b}) = \lambda_j(\mathbf{b}) \delta_{jk}. \tag{12.8}$$

Sommando in j, abbiamo

$$\begin{aligned} &\text{dal termine a sinistra in (12.8)} = \sum_{j=1}^{N} \sum_{\alpha=1}^{d} \lambda_j(\mathbf{b}) \partial_{x_\alpha} F_j(\mathbf{x}(\mathbf{b})) \partial_{b_k} x_\alpha(\mathbf{b}) \\ &= \sum_{\alpha=1}^{d} \partial_{b_k} x_\alpha(\mathbf{b}) \sum_{j=1}^{N} \lambda_j(\mathbf{b}) \partial_{x_\alpha} F_j(\mathbf{x}(\mathbf{b})) = \sum_{\alpha=1}^{d} \partial_{b_k} x_\alpha(\mathbf{b}) \partial_{x_\alpha} G(\mathbf{x}(\mathbf{b})) \\ &= \partial_{b_k} \left(G(\mathbf{x}(\mathbf{b})) \right) \end{aligned}$$

mentre

$$\text{dal termine a destra in (12.8)} = \sum_{j=1}^{N} \lambda_j(\mathbf{b}) \delta_{jk} = \lambda_k(\mathbf{b}).$$

Quindi, confrontando le formule, concludiamo la seguente formula, che generalizza (12.6),

$$\partial_{b_k} \left(G(\mathbf{x}(\mathbf{b})) \right) = \lambda_k(\mathbf{b}). \tag{12.9}$$

Esempio 12.9. Trovare massimi e minimi di $f(x, y) = x^4 + y^4$ sul vincolo $x^6 + y^6 = 1$.

Notare subito che non è disponibile una ovvia rappresentazione parametrica in termini di funzioni elementari per il vincolo. Utilizziamo allora il metodo dei moltiplicatori considerando il sistema

$$\begin{aligned} &4x^3 = 6\lambda x^5 \\ &4y^3 = 6\lambda y^5 \\ &x^6 + y^6 = 1. \end{aligned}$$

Per $x = 0$ dall'ultima equazione otteniamo $y = \pm 1$ e dalla seconda $\lambda = \frac{2}{3}$. Per $y = 0$ dall'ultima equazione otteniamo $x = \pm 1$ e dalla prima $\lambda = \frac{2}{3}$. Pertanto

$$(\pm 1, 0)\,, \quad (0, \pm 1) \text{ sono punti critici.} \tag{12.10}$$

Ora vediamo se ci sono casi dove sia $x \neq 0$ che $y \neq 0$. Semplificando il sistema diviene

$$\begin{aligned} &2 = 2\lambda x^2 \\ &2 = 3\lambda y^2 \\ &x^6 + y^6 = 1 \end{aligned}$$

e risolvendo per λ nelle prime due si ottiene

$$\begin{aligned} &\lambda = \frac{2}{3x^2} \\ &\lambda = \frac{2}{3y^2} \\ &x^6 + y^6 = 1. \end{aligned}$$

Da qui si vede che necessariamente $x^2 = y^2$, ossia $x = \pm y$. Sostituendo nell'ultima si ottiene

$$2x^6 = 1$$

ossia $x = \frac{1}{2^{\frac{1}{6}}}$ e quindi si ottengono quattro valori

$$\pm(\frac{1}{2^{\frac{1}{6}}}, \pm\frac{1}{2^{\frac{1}{6}}}). \tag{12.11}$$

Per ciascuno di essi c'è un valore opportuno di λ che dà una soluzione del sistema. Abbiamo ottenuto 8 punti critici. Tra di essi cerchiamo i punti di minimo ed i punti di massimo. Abbiamo

$$f(\pm 1, 0) = f(0, \pm 1) = 1$$

e

$$f\left(\pm(\frac{1}{2^{\frac{1}{6}}}, \pm\frac{1}{2^{\frac{1}{6}}})\right) = \frac{1}{2^{\frac{2}{3}}} + \frac{1}{2^{\frac{2}{3}}} = 2^{\frac{1}{3}} > 1.$$

Quindi i punti in (12.10) sono i punti di minimo assoluto, mentre i punti in (12.11) sono i punti di massimo assoluto.

Esempio 12.10. Calcoliamo massimo e minimo della funzione $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$ sulla circonferenza unitaria $C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. Entrambi esistono per il teorema di Weierstrass in quanto la funzione f è continua su C che è chiuso e limitato, quindi compatto.

Notiamo facilmente che $C = \{(x, y) \in \mathbb{R}^2 \mid F(x, y) = 0\}$, con $F(x, y) = x^2 + y^2 - 1$, quindi

$$\nabla f(x, y) = (y, x)\,, \qquad \nabla F(x, y) = (2x, 2y)\,,$$

e notiamo che $\nabla F(x, y) \neq (0, 0)$ per ogni $(x, y) \in C$, quindi $J_F = \nabla F$ ha rango massimo su tutti i punti di C. Possiamo quindi utilizzare il teorema dei moltiplicatori di Lagrange e cercare quindi i valori λ, x, y che risolvono il seguente sistema.

$$\begin{cases} \nabla f(x, y) = \lambda \nabla F(x, y) \\ F(x, y) = 0 \end{cases} \iff \begin{cases} \frac{\partial f}{\partial x}(x, y) = \lambda \frac{\partial F}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) = \lambda \frac{\partial F}{\partial y}(x, y) \\ F(x, y) = 0 \end{cases} \iff \begin{cases} y = \lambda 2x \\ x = \lambda 2y \\ x^2 + y^2 = 1 \end{cases}$$

Sostituiamo la seconda equazione nella prima e troviamo

$$\begin{cases} y = 4\lambda^2 y \\ x = \lambda 2y \\ x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{cases} y(1 - 4\lambda^2) = 0 \\ x = \lambda 2y \\ x^2 + y^2 = 1 \end{cases}$$

Abbiamo due alternative: $y = 0$ oppure $\lambda = \pm 1/2$. Nel caso $y = 0$ troviamo

$$\begin{cases} y = 0 \\ x = 0 \\ 0 = 1 \end{cases}$$

che non porta a soluzioni. Consideriamo quindi gli altri due casi

$$\begin{cases} \lambda = \pm 1/2 \\ x = \pm y \\ 2x^2 = 1 \end{cases}$$

che porta alle quattro soluzioni

$$P_\pm = \left(\pm\frac{\sqrt{2}}{2}, \pm\frac{\sqrt{2}}{2} \right), \qquad Q_\pm = \left(\pm\frac{\sqrt{2}}{2}, \mp\frac{\sqrt{2}}{2} \right).$$

Sostituendo queste coordinate nella funzione f, troviamo che i punti $P_\pm$ sono punti di massimo con valore $1/2$, mentre i punti $Q_\pm$ sono punti di minimo con valore $-1/2$.

Esempio 12.11. Consideriamo la funzione $f : \mathbb{R}^2 \to \mathbb{R}$, definita come $f(x, y) = x^2 + y^2$ e il vincolo $C = \{F = 0\}$ con $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x^4 - x^2 + y^2$, e determiniamo il massimo di f su C (ovvero il punto più distante dall'origine).

Sia f che F sono differenziabili in tutto $\mathbb{R}^2$, e calcoliamo facilmente

$$\nabla f(x,y) = (2x, 2y)\,, \qquad \nabla F(x,y) = (4x^3 - 2x, 2y)$$

e notiamo che ∇F non ha rango massimo nei punti $O = (0,0) \in C$, $A_\pm = (\pm 1/\sqrt{2}\,, 0) \notin C$. L'origine è quindi un candidato punto di estremo con valore $F(0,0) = 0$.

Impostiamo ora il sistema suggerito dal teorema dei moltiplicatori di Lagrange:

$$\begin{cases} \nabla f(x,y) = \lambda \nabla F(x,y) \\ F(x,y) = 0 \end{cases} \iff \begin{cases} 2x = \lambda 2x(2x^2 - 1) \\ 2y = \lambda 2y \\ x^4 - x^2 + y^2 = 0 \end{cases} \iff \begin{cases} 2x\big(1 - \lambda(2x^2 - 1)\big) = 0 \\ 2y(1 - \lambda) = 0 \\ x^4 - x^2 + y^2 = 0 \end{cases}$$

Ci conviene iniziare col considerare la seconda equazione distinguendo i casi $y = 0$ e $\lambda = 1$.

$$[y = 0] \qquad \begin{cases} 2x\big(1 - \lambda(2x^2 - 1)\big) = 0 \\ y = 0 \\ x^4 - x^2 = 0 \end{cases} \Rightarrow \begin{cases} 2x\big(1 - \lambda(2x^2 - 1)\big) = 0 \\ y = 0 \\ x \in \{-1, 0, 1\} \end{cases}$$

Abbiamo quindi tre candidati: $O = (0,0)$ (già considerato prima), $P_\pm = (\pm 1, 0)$, con valore $f(P_\pm) = 1$.

$$[\lambda = 1] \qquad \begin{cases} 2x\big(1 - (2x^2 - 1)\big) = 0 \\ \lambda = 1 \\ x^4 - x^2 + y^2 = 0 \end{cases} \Rightarrow \begin{cases} 2x \cdot 2(1 - x^2) = 0 \\ \lambda = 1 \\ x^4 - x^2 + y^2 = 0 \end{cases} \Rightarrow \begin{cases} x \in \{-1, 0, 1\} \\ \lambda = 1 \\ y = 0 \end{cases}$$

che porta alle stesse soluzioni di poco fa. Quindi, dati i tre candidati O, P_+, P_- notiamo che il massimo di f su C è 1 raggiunto nei punti P_+, P_-.

Esempio 12.12. Consideriamo

$$\begin{aligned} &\text{massimizzare } f(x,y,z) = xyz \\ &\text{con vincolo } x + y + z = 1 \text{ per } x \geq 0 \quad y \geq 0 \text{ e } z \geq 0. \end{aligned}$$

Stiamo valutando f in un triangolo chiuso. f si annulla sul bordo ed è strettamente positiva all'interno. E' evidente che c'è un massimo assoluto all'interno del triangolo. Applicando il metodo dei moltiplicatori di Lagrange, scriviamo

$$\begin{aligned} yz &= \lambda \\ xz &= \lambda \\ xy &= \lambda \\ x + y + z &= 1. \end{aligned}$$

Si ricava subito che deve essere $x = y = z = 1/3$.

Esempio 12.13. Consideriamo

$$\begin{aligned}&\text{massimizzare } f(x,y,z) = xyz\\&\text{con vincolo } xy + xz + yz = 1 \text{ per } x \geq 0 \quad y \geq 0 \text{ e } z \geq 0.\end{aligned}$$

Ovviamente, $f > 0$ quando le variabili sono tutte non nulle, quindi il massimo, se c'è, è all'interno del vincolo. Notiamo che il vincolo è illimitato perchè ad esempio ho $z = \dfrac{1-xy}{x+y} \xrightarrow{(x,y)\to(0,0)} +\infty$. D'altra parte, se $(x,y,z) \to \infty$ sul vincolo, due delle variabili convergono a 0 mentre la terza diverge. Se ad esempio sono nella situazione $z = \dfrac{1-xy}{x+y} \xrightarrow{(x,y)\to(0,0)} +\infty$, allora $(x+y)z \longrightarrow 1$. Ma allora

$$0 < xyz < x(x+y)z \longrightarrow 0 \Longrightarrow \lim_{(x,y,z)\to\infty} f(x,y,z) = 0 \text{ sul vincolo.}$$

Siccome il vincolo è un chiuso, esiste almeno un punto di massimo e, come visto sopra, è interno. Applicando il metodo dei moltiplicatori di Lagrange, scriviamo

$$\begin{aligned}yz &= \lambda(y+z)\\xz &= \lambda(x+z)\\xy &= \lambda(x+y)\\xy + xz + yz &= 1.\end{aligned}$$

Confrontiamo le prime due, ricaviamo

$$\frac{y}{y+z} = \frac{x}{x+z}.$$

Siccome per fissato $z > 0$ la funzione $[0,+\infty) \ni t \to \frac{t}{t+z}$ è strettamente crescente, segue $x = y$. In modo simile, concludiamo $x = y = z$. Dal vincolo ricaviamo $3x^2 = 1 \Rightarrow x = 3^{-1/2}$. Quindi il punto di massimo assoluto è $(3^{-1/2}, 3^{-1/2}, 3^{-1/2})$.

13 Esempi di esercizi da vecchi esami (Ingegneria a Reggio Emilia)

Esercizio 13.1. Trovare punti di massimo e di minimo assoluto e relativo di

$$f(x) = \frac{x}{1+x^4+y^4}$$

Inoltre disegnare le curve di livello.

Esercizio 13.2. Trovare punti di massimo e di minimo assoluto di

$$f(x,y) = x^2 + y^2 + 8y - 4x + 2$$

per (x,y) nel quadrato $Q = \{(x,y)\, t.c.\, x \in [0,1],\, y \in [0,1]\}$.

Esercizio 13.3. Trovare punti critici (4 punti) di

$$f(x,y) = \frac{x+y}{1+x^2+y^2}.$$

Decidere quali di essi sono punti di massimo e quali punti di minimo (2 punti). Infine disegnare le curve di livello (2 punti).

Esercizio 13.4. Classificare i punti di massimo e di minimo assoluti di

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

e disegnare delle curve di livello che diano una buona descrizione qualitativa della funzione.

Per prima cosa abbiamo

$$\lim_{(x,y)\to\infty} f(x,y) = +\infty$$

quindi c' è sicuramente almeno un punto di minimo assoluto.

Poniamo

$$f_x = x^3 + x = 0$$
$$f_y = y^3 - y = 0$$

Ora le radici sono $(0,0)$, $(0,\pm 1)$. SI vede subito che $f_{xx}(0,0) = 2$, $f_{yy}(0,0) = -2$ e $f_{xy}(0,0) = 0$ e pertanto il det Hessiano$= -4$ e quindi $(0,0)$ è una sella e pertanto non è un minimo assoluto. Siccome $f(0,-1) = f(0,1)$ concludiamo che entrambi $(0,-1)$ & $(0,1)$ sono minimi assoluti.

Esercizio 13.5. Classificare i punti di massimo e di minimo assoluti di $f(x,y) = \frac{x^3}{3} - x^2 y + x + \frac{y^3}{3}$ nel quadrato $0 \le x \le 1$, $0 \le y \le 1$.

Esercizio 13.6. Classificare i punti critici di $f(x,y) = \frac{x^4}{4} + xy + \frac{y^4}{4}$ e trovare massimi e minimi relativi ed assoluti.

Per prima cosa abbiamo

$$\lim_{(x,y)\to\infty} f(x,y) = +\infty$$

quindi c'è sicuramente almeno un punto di minimo assoluto.

Poniamo

$$f_x = x^3 + y = 0$$
$$f_y = x + y^3 = 0$$

equivalente a

$$y^9 - y = y(y^8 - 1) = 0$$
$$x = -y^3$$

Otteniamo pertanto i punti critici $(0,0)$, $(1,-1)$ & $(-1,1)$. SI vede subito che $f_{xx}(0,0) = f_{yy}(0,0) = 0$ mentre $f_{xy}(0,0) = 1$ e pertanto il det Hessiano$= -1$ e quindi $(0,0)$ é una sella e pertanto non è un minimo assoluto. Siccome $f(1,-1) = f(-1,1)$ concludiamo che entrambi $(1,-1)$ & $(-1,1)$ sono minimi assoluti.

Esercizio 13.7. Trovare massimi e minimi assoluti di

$$f(x,y) = xe^{-x^2-y^2}$$

nel disco $x^2 + y^2 \leq 1$.

Esercizio 13.8. Trovare punti critici (4 punti) di

$$f(x,y) = \frac{x+y}{1+x^2+y^2}.$$

Decidere quali di essi sono punti di massimo e quali punti di minimo (3 punti). Disegnare (approssimativamente) un numero sufficiente di curve di livello per avere una comprensione qualitativa di $f(x,y)$ (2 punti).

Esercizio 13.9. Trovare i punti di massimo e minimo assoluto di $f(x,y) + x^2 - y^2 + 2y + x$ nel triangolo T definito da

$$x + y \leq 1\,, \quad x \geq 0\,, \quad y \geq 0.$$

Esercizio 13.10. Classificare i punti di massimo e di minimo assoluti di $f(x,y) = \frac{x^3}{3} + x^2y + x - \frac{y^3}{3}$ nel quadrato $0 \leq x \leq 1$, $0 \leq y \leq 1$.

Esercizio 13.11. Trovare punti critici di

$$f(x,y) = \frac{2x+y}{1+x^2+y^2}.$$

Decidere quali di essi sono punti di massimo e quali punti di minimo. Infine disegnare alcune curve di livello .

Per prima cosa abbiamo

$$\lim_{(x,y)\to\infty} f(x,y) = 0.$$

La retta $2x + y = 0$ è la curva di livello 0. Siccome $f(x,y)$ assume valori sia positivi che negativi, ci saranno sia punti di massimo che di minimo assoluto. Poniamo

$$f_x = \frac{2(1+x^2+y^2) - (2x+y)2x}{(1+x^2+y^2)^2} = \frac{2 - 2x^2 - 2xy + 2y^2}{(1+x^2+y^2)^2} = 0$$

$$f_y = \frac{(1+x^2+y^2) - (2x+y)2y}{(1+x^2+y^2)^2} = \frac{1 + x^2 - 4xy - y^2}{(1+x^2+y^2)^2} = 0$$

Il sistema è equivalente a

$$x^2 + xy - y^2 = 1$$
$$x^2 - 4xy - y^2 = -1.$$

Esso é equivalente a

$$x^2 - y^2 = 4/3$$
$$xy = 2/5.$$

Per sostituzione

$$x^2 - \frac{4}{25x^2} = \frac{4}{3} \Rightarrow x^4 - \frac{4}{3}x^2 - \frac{4}{25} = 0.$$

Per $z = x^2$ ho soluzioni

$$z_{1,2} = \frac{2}{3} \pm \sqrt{\frac{4}{9} + \frac{4}{25}}$$

e siccome $z = x^2$ ci interessa solo la soluzione col $+$, che chiamiamo z_1. Infine scopriamo ci sono esattamente due punti critici dati da $P_+ = (x_+, \frac{2}{5x_+})$ & $P_- = (x_-, \frac{2}{5x_-})$ dove $x_\pm = \pm\sqrt{z_1}$. Siccome sappiamo che ci sono almeno un punto di massimo assoluto ed almeno un punto di minimo assoluto, risulta che P_+ è il punto di massimo assoluto e P_- è il punto di minimo assoluto.

Esercizio 13.12. Classificare i punti critici di $f(x,y) = \frac{x^3}{3} + x^2y + x + \frac{y^3}{3}$ e trovare massimi e minimi relativi ed assoluti.

Esercizio 13.13. Trovare punti critici, distinguere quali sono punti di massimo e di minimo assoluto e relativo di

$$f(x,y) = x^6 + y^2 + x^2y$$

Inoltre disegnare alcune curve di livello.

SOLUZIONE. Per prima cosa $f(x,y) \to +\infty$ per $(x,y) \to \infty$, e quindi ci sono punti di minimo assoluto. I punti critici si ottengono risolvendo:

$$f_x = 6x^5 + 2xy = 2x(3x^4 + y) = 0$$
$$f_y = 2y + x^2 = 0.$$

Sostituendo $y = -\frac{x^2}{2}$ nella prima si ottiene

$$2x(3x^4 - \frac{x^2}{2}) = 6x^3(x^2 - \frac{1}{6}) = 0.$$

Le soluzioni sono

$$x = 0\,, \quad x = \pm\frac{1}{\sqrt{6}}.$$

Ponendo $y = -\frac{x^2}{2}$ si ottengono i punti critici.

Notare che $(0,0)$ non è ne' un massimo ne' un minimo locale. Per prima cosa osserviamo che $f(0,0)=0$. Se consideriamo la curva $y \to (|y|^\alpha, y)$, per un $\alpha > 0$ da definire, abbiamo

$$f(|y|^\alpha, y) = |y|^{6\alpha} + |y|^2 + |y|^{2\alpha} y.$$

Per $1/4 < \alpha < 1/2$ e per $|y| \ll 1$, risulta

$$f(|y|^\alpha, y) \approx |y|^{2\alpha} y$$

e quindi $f(|y|^\alpha, y) > 0$ per $y > 0$ e $f(|y|^\alpha, y) < 0$ per $y < 0$. Ossia $(0,0)$ non è ne' un massimo ne' un minimo locale. Siccome c'è necessariamente un minimo assoluto, necessariamente questi sono i due punti $(\pm\frac{1}{\sqrt{6}}, -\frac{1}{12})$ visto che

$$f(\frac{1}{\sqrt{6}}, -\frac{1}{12}) = f(-\frac{1}{\sqrt{6}}, -\frac{1}{12}).$$

Esercizio 13.14. Trovare punti critici, distinguere quali sono punti di massimo e di minimo assoluto e relativo di

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

Inoltre disegnare alcune curve di livello.

SOLUZIONE Per prima cosa $f(x,y) \to 0$ per $(x,y) \to \infty$, e siccome $f > 0$ dappertutto quindi ci sono punti di massimo assoluto.

$$f_x = \frac{4x^3 - y + x}{(x^4 + y^4 - xy + \frac{x^2}{2} + \frac{y^2}{2} + 1)^2} = 0$$
$$f_y = \frac{4y^3 - x + y}{(x^4 + y^4 - xy + \frac{x^2}{2} + \frac{y^2}{2} + 1)^2} = 0.$$

Quindi i punti critici risolvono

$$4x^3 - y + x = 0$$
$$4y^3 - x + y = 0.$$

Sommando e sottraendo otteniamo

$$4(x^3 + y^3) = 0$$
$$2(x-y) + 4(x^3 - y^3) = 2(x-y)(1 + x^2 - xy + y^2) = 0$$

la cui unica soluzione è $(0,0)$, necessariamente il punto di massimo globale.

Esercizio 13.15. Trovare punti di massimo e di minimo assoluti di $f(x,y) = \frac{x^3}{3} + \frac{y^2}{2}$ nell'insieme $x^2 + \frac{y^2}{4} \leq 1$.

Esercizio 13.16. Trovare massimi e minimi assoluti di $f(x,y) = x$ nel cardioide definito in coordinate polari da $0 \leq r \leq 1 - \cos\theta$, $\theta \in [0, 2\pi]$.

Esercizio 13.17. Trovare punti critici di

$$f(x,y) = -\frac{x^3}{3} + xy^2 + \frac{2}{3}y^3 - x - 2y.$$

Decidere quali di essi sono selle, punti di massimo e punti di minimo relativo od assoluto. In presenza di punti di massimo o minimo, dire se sono relativi o assoluti.

Abbiamo

$$\begin{aligned} f_x &= -x^2 + y^2 - 1 = 0 \\ f_y &= 2y^2 + 2xy - 2 = 0. \end{aligned}$$

Sostituendo nella seconda $y^2 = 1 + x^2$ otteniamo l'equivalente sistema

$$\begin{aligned} &(y+x)(y-x) = 1 \\ &(x+y)^2 = 1. \end{aligned}$$

Quindi

$$\begin{aligned} (y-x) &= 1 \\ (x+y) &= 1 \end{aligned}$$

che ha soluzione $(0,1)$, e

$$\begin{aligned} (y-x) &= -1 \\ (x+y) &= -1 \end{aligned}$$

che ha soluzione $(0,-1)$. Questi sono i punti critici....

Esercizio 13.18. Trovare punti critici (4 punti) di

$$f(x,y) = \frac{x+y}{1+x^2+y^2}.$$

Decidere quali di essi sono punti di massimo e quali punti di minimo (2 punti). Infine disegnare le curve di livello (2 punti).

Esercizio 13.19. Trovare massimo e minimo assoluto di $f(x,y) = x^2 - y^2 + y$ nella regione

$$x^2 + (y^2)/2 \leq 1$$

.

Esercizio 13.20. Trovare punti di massimo e minimo relativo ed assoluto di $f(x,y) = x^2 - y^2$ lungo la retta $x - 2y + 6 = 0$.

Esercizio 13.21. Trovare punti di massimo e minimo di $f(x,y) = x^2 + 3xy + y^2$ nel disco $x^2 + y^2 \leq 1$

Esercizio 13.22. Trovare massimi e minimi assoluti di

$$f(x,y) = (x-1)^2 - y^2$$

nella regione

$$x^2 + \frac{y^2}{4} \leq 1.$$

Esercizio 13.23. Trovare punti critici e classificarli, verificare se ci sono massimi e minimi relativi o assoluti, e infine disegnare un certo numero di curve di livello, per la funzione

$$f(x,y) = x^4 + y^4 + x^2 + y^2 + 8xy + 1.$$

Esercizio 13.24. Giustificare il fatto che la seguente funzione ha massimi e minimi assoluti nel piano, e trovarli.

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

Disegnare anche le curve di livello.

Esercizio 13.25. Trovare massimi e minimi relativi ed assoluti, se ci sono, di $f(x,y) = x^4 + y^4 + 2x^2y^2 - x^2 - y^2$ nel piano.

Esercizio 13.26. Trovare punti critici e classificarli, verificare se ci sono massimi e minimi relativi o assoluti, e infine disegnare un certo numero di curve di livello, per la funzione

$$f(x,y) = xye^{-x^2-y^2}.$$

Esercizio 13.27. Trovare punti critici e classificarli, verificare se ci sono massimi e minimi relativi o assoluti, e infine disegnare un certo numero di curve di livello, per la funzione

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

Esercizio 13.28. Trovare punti critici, eventuali punti di massimo e di minimo assoluti o relativi di

$$f(x,y) = x^3 + xy^2 + y^3 + x^2 + y^2.$$

Esercizio 13.29. Trovare punti critici e classificarli, verificare se ci sono massimi e minimi relativi o assoluti, e infine disegnare un certo numero di curve di livello, per la funzione

$$f(x,y) = x^4 + y^4 + x^2 + y^2 + 8xy + 1.$$

Esercizio 13.30. Trovare punti critici, dire se sono punti di massimo e di minimo assoluto, verificare se gli si può applicare il metodo della derivata seconda, se si può applicare il metodo della derivata seconda classificarli (massimi e minimi relativi o selle), per

$$f(x,y) = \frac{4}{3}x^3 + x^2y - x^2 + \frac{y^2}{2}.$$

Esercizio 13.31. Trovare massimi e minimi assoluti di $f(x,y) = x^2 - y^2$ sul vincolo $(x-1)^2 + y^2 = 1$.

Riferimenti bibliografici

[1] B.R. Gelbaum e J.M.H.Olmsted, *Counterexamples in Analysis*, Dover, (1992).

[2] E. Giusti,*Analisi Matematica 2*, Bollati Boringhieri (2003).