

Indice

1	Calcolo differenziale	3
1.1	Introduzione	4
1.2	Spazi	4
1.2.1	Spazi metrici	4
1.2.1.1	Distanza/metrica, spazio metrico	4
1.2.1.2	Intorni circolari, aperti, chiusi, punti di accumu- lazione	6
1.2.1.3	Apertura/chiusura di \emptyset e X	10
1.2.1.4	Sconnessione/connessione e insiemi aperti/chiusi	11
1.2.1.5	Apertura/chiusura di unione/intersezione di fa- miglie di aperti/chiusi	12
1.2.1.6	Funzioni $f : A \subseteq X \rightarrow Y$ continue, aperti e chiusi	14
1.2.1.7	Sottospazio metrico	16
1.2.2	Spazi normati	16
1.2.2.1	Spazio vettoriale e spazio normato	16
1.2.2.2	Spazio normato e spazio metrico	18
1.2.2.3	Sottospazio normato	19
1.2.3	Spazi con prodotto interno	19
1.2.3.1	Prodotto interno e spazi con	19
1.2.3.2	Spazio con prodotto interno euclideo è normato	19
1.2.3.3	Disuguaglianza di Cauchy-Schwartz e interpre- tazione geometrica del prodotto interno euclideo	20
1.2.3.4	Spazio con prodotto interno generico è normato	21
1.3	Funzioni $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ e curve in \mathbb{R}^m	22
1.3.1	Funzioni	22
1.3.2	Curve	23
1.3.2.1	Definizioni e primi esempi	23
1.3.2.2	Vettori tangenti	26
1.4	Funzioni $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$	27
1.4.1	Differenziabilità	27
1.4.1.1	Osservazioni preliminari	27
1.4.1.2	Derivate direzionali, derivate parziali	28
1.4.1.3	Derivate parziali valutate in un punto	30
1.4.1.4	Funzioni differenziabili	31

1.4.1.5	Differenziabilità e derivate direzionali	33
1.4.1.6	Gradiente, prodotto interno e valutazione del differenziale	35
1.4.1.7	Differenziabilità e continuità	38
1.4.1.8	Il teorema del differenziale totale	39
1.4.1.9	Operazioni tra funzioni differenziabili in un punto	40
1.4.1.10	Differenziali come combinazione lineare	40
1.4.1.11	Derivate successive	43
1.4.1.12	Derivazione di funzioni composte	45
1.4.2	Ottimizzazione libera	46
1.4.2.1	Massimi/minimi relativi	46
1.4.2.2	Setup di ottimizzazione	47
1.4.2.3	Identificazione dei punti critici di f : derivata prima di F	47
1.4.2.4	Matrici Hessiane	49
1.4.2.5	Classificazione dei punti critici	50
1.4.2.6	Classificazione dei punti critici	52
1.4.3	Esercizi su ottimizzazione libera	57
1.4.4	Il teorema di Dini	64
1.5	Funzioni $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$	69
1.5.1	Funzioni a valori vettoriali di più variabili	69
1.5.2	Funzioni differenziabili	70
1.6	Ottimizzazione vincolata	74
1.6.1	Teorema delle funzioni implicite	74
1.6.2	Varietà in \mathbb{R}^n	76
1.6.2.1	Definizione di varietà. Punti regolari	76
1.6.2.2	Varietà, spazi tangenti e spazi normali	82
1.6.3	Punti critici vincolati e moltiplicatori di Lagrange	84
1.6.4	Esempi	86

Capitolo 1

Calcolo differenziale

Osservazione 1 (Preparazione dell'esame). Sono importanti enunciati, definizioni (farsi esempi (es un insieme aperto semplice) e controesempi (non aperto)) e 5 dimostrazioni a scelta (essere in grado di fare).

Qualche esercizio: differenziale, derivata direzionale, calcolare derivata parziale, calcolo spazio tangente o varietà.

Da preparare, in aggiunta agli appunti L^AT_EX, alla fine del paragrafo 3, il teorema di Dini e il teorema delle funzioni implicite (solo l'enunciato, è il Dini generalizzato).

Ultima parte su varietà e su curve è abbastanza astratta (appunti sintetici).

Le dimostrazioni che si possono studiare sono quelle dove c'è scritto DIMOSTRAZIONE in maiuscolo (nel pdf) + la dimostrazione del teorema di dini (5 di cui chiesta 1):

- insieme chiuso e punti di accumulazione (1.2.1)
- sconnessione e sottoinsiemi sia aperti che chiusi (1.2.3)
- unioni/intersezioni famiglie aperti/chiusi (1.2.7)
- una funzione differenziabile in un punto ha tutte le derivate direzionali (1.4.2)
- una funzione differenziabile in un punto è continua in esso (1.4.6)
- i funzionali coordinata formano base canonica dello spazio dei funzionali lineari (base canonica spazio duale 1.4.9)
- differenziale come combinazione lineare dei funzionali coordinata 1.4.10
- condizione necessaria punti critici (1.4.14)
- teorema di Dini (1.4.24)
- teorema dei moltiplicatori di Lagrange (1.6.7)

4-5 giorni prima mandare una mail e chiedere appuntamento.

1.1 Introduzione

Osservazione 2. Nel corso di analisi 1, si sono studiate le funzioni $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ ossia le funzioni reali di variabile reale, per le quali input e output erano numeri reali. Possiamo però immaginare altre situazioni:

- funzioni $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ l'input è un singolo numero ma l'output è un insieme (vettore) di numeri reali; in tal caso si parla di funzioni di una variabile reale a valori vettoriali (o funzione a valori vettoriali di una variabile reale).
Nel caso di $m = 2, 3$ le funzioni $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ hanno il significato geometrico di curve nel piano o nello spazio tridimensionale.
- funzioni $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ nelle quali gli input sono più di uno: ad un gruppo di n ($n > 1$) variabili viene associato univocamente un numero reale. Parliamo in questo caso di funzioni di n variabili reali a valori reali (o funzione reale di n variabili reali).
- funzioni $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$

Osservazione 3. Il corso si concentra sulle funzioni $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, per le quali comunque ci serviranno le rimanenti.

Osservazione 4. Prima alcuni concetti di base.

1.2 Spazi

1.2.1 Spazi metrici

1.2.1.1 Distanza/mettrica, spazio metrico

Definizione 1.2.1 (Distanza/Metrica). Dato un insieme X , viene chiamata così una funzione $d : X \times X \rightarrow \mathbb{R}$ tale che, per $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$:

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$; $d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$, (Positiva o nulla);
2. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$, (Simmetria);
3. $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$, (Disuguaglianza triangolare)

Definizione 1.2.2 (Spazio metrico). La coppia ordinata (X, d) composta dall'insieme X e dalla distanza su esso definita.

Esempio 1.2.1 (Spazio metrico (\mathbb{R}, d)). Se definiamo su \mathbb{R} la funzione

$$d(a, b) = \sqrt{(a - b)^2} = |a - b|$$

con $a, b \in \mathbb{R}$, questa è una metrica in quanto soddisfa i tre assiomi; i primi due sono banali da verificare, per il terzo si ha che

$$|a - b| \leq |a - c| + |c - b|$$

partendo dalla generica disuguaglianza triangolare

$$|x + y| \leq |x| + |y|$$

e sostituendo $x = a - c$ e $y = c - b$.

Pertanto la coppia (\mathbb{R}, d) è uno spazio metrico; e costituisce un caso particolare del generico spazio metrico euclideo seguente (quando $n = 1$).

Esempio 1.2.2 (Spazio metrico euclideo (\mathbb{R}^n, d)). Su \mathbb{R}^n si può definire¹ la funzione $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ seguente:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(y_1 - x_1)^2 + \cdots + (y_n - x_n)^2}.$$

dati $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$.

La funzione definita è una metrica, soddisfacendone la definizione e prende il nome di *metrica euclidea*: le prime due condizioni della definizione sono di banale verifica, per la terza si necessita di strumenti non ancora introdotti.

Pertanto (\mathbb{R}^n, d) con d metrica euclidea è uno spazio metrico, detto *spazio metrico euclideo*.

TODO: Completa e fai riferimento a questo esempio.

Esempio 1.2.3 (Spazio metrico (\mathbb{R}^2, d)). Se $n = 2$, ossia in \mathbb{R}^2 , si ha che \mathbf{x} è descritto da due coordinate (x_1, x_2) , una per ciascun asse (quello che classicamente sono x e y) e allo stesso modo avviene per \mathbf{y} , descritto da y_1, y_2 . La distanza tra i due punti è

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}.$$

con $(y_1 - x_1)^2$ il quadrato costruito sul segmento tra i due punti nella prima dimensione, $(y_2 - x_2)^2$ quello sulla seconda, e dunque la distanza altro non è che (essendo la somma sotto radice) la lunghezza dell'ipotenusa del triangolo rettangolo formato sui due punti per il teorema di Pitagora.

Osservazione 5. D' ora in avanti riguarderemo sempre \mathbb{R}^n come *spazio metrico euclideo*, vale a dire muniremo \mathbb{R}^n della metrica (o, distanza) euclidea.

Esempio 1.2.4 (Spazio metrico (X, d) con metrica della topologia discreta (spazio discreto)). Dato un insieme qualsiasi X , una funzione $d : X \times X \rightarrow \mathbb{R}$ che segnala la diversità degli elementi $\mathbf{x}, \mathbf{y} \in X$ forniti:

$$d(\mathbf{x}, \mathbf{y}) = [\mathbf{x} \neq \mathbf{y}] = \begin{cases} 0 & \text{se } \mathbf{x} = \mathbf{y} \\ 1 & \text{se } \mathbf{x} \neq \mathbf{y} \end{cases}$$

¹Formalmente possiamo porre la definizione di metrica euclidea poiché l'insieme delle n -ple ordinate di numeri reali:

$$\mathbb{R}^n = \{(x_1, \dots, x_n); x_i \in \mathbb{R}, i = 1, \dots, n\}.$$

possiede una struttura di spazio vettoriale (reale) di dimensione n , con le operazioni di somma e moltiplicazione per scalare definite come di consueto. Pertanto possiamo effettuare la differenza di due vettori appartenenti a (differenza sottintesa nella formula)

è una metrica. La verifica delle prime due condizioni della definizione è banale, la verifica di

$$d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$$

avviene come segue: se $d(\mathbf{x}, \mathbf{z}) = 0$ allora la disuguaglianza è certamente verificata (dato che i due addendi al secondo membro possono assumere valori 0 o 1); se viceversa $d(\mathbf{x}, \mathbf{z}) = 1$ (\mathbf{x} e \mathbf{z} differiscono) non si può avere contemporaneamente che $d(\mathbf{x}, \mathbf{y}) = 0$ e $d(\mathbf{y}, \mathbf{z}) = 0$ (ossia che, congiuntamente, $\mathbf{x} = \mathbf{y}$ e $\mathbf{y} = \mathbf{z}$) altrimenti si avrebbe $d(\mathbf{x}, \mathbf{z}) = 0$ (ossia sarebbe $\mathbf{x} = \mathbf{z}$ per la proprietà transitiva dell'uguaglianza). Pertanto deve essere che quanto meno $d(\mathbf{x}, \mathbf{y}) = 1$ o $d(\mathbf{y}, \mathbf{z}) = 1$ (o possono valere entrambi) e pertanto la disuguaglianza numerica è verificata.

Dunque la coppia (X, d) costituisce uno spazio metrico.

1.2.1.2 Intorni circolari, aperti, chiusi, punti di accumulazione

Osservazione 6. Nel seguito ci concentriamo sullo spazio metrico generico (X, d) . A titolo di esempio possiamo pensare a X come \mathbb{R}^n e a d come alla distanza euclidea.

Definizione 1.2.3 ($I(x_0, r)$: intorno aperto di centro x_0 e raggio r). Dato $x_0 \in X$ e $r \in \mathbb{R}^+$ ($r \in \mathbb{R} : r > 0$), è così chiamato il sottoinsieme di X composto dagli elementi x che distano meno di r da x_0 :

$$I(x_0, r) = \{x \in X : d(x, x_0) < r\}$$

Esempio 1.2.5 ($I(x_0, r)$ in (\mathbb{R}, d)). Si ha

$$|x - x_0| < r \iff \begin{cases} x - x_0 < r \\ x - x_0 > -r \end{cases} \iff -r < x - x_0 < r \iff x_0 - r < x < x_0 + r$$

pertanto $I(x_0, r) = (x_0 - r, x_0 + r)$, ossia è il segmento sulla retta reale di ampiezza $2r$ e centrato in x_0 .

Esempio 1.2.6 ($I(\mathbf{x}_0, r)$ in (\mathbb{R}^2, d) ed (\mathbb{R}^3, d)). In \mathbb{R}^2 si tratta dell'insieme di punti che formano il cerchio di raggio r e centro \mathbf{x} ; in \mathbb{R}^3 della sfera di raggio r e centro \mathbf{x} e così via.

Esempio 1.2.7 ($I(\mathbf{x}_0, r)$ nello spazio discreto (X, d)). Nello spazio metrico con metrica della topologia discreta tutti gli elementi stanno tra loro ad una distanza 0 o 1. Se $r \leq 1$ nessun elemento oltre \mathbf{x}_0 sta entro r ; viceversa se $r > 1$ tutti stanno entro r . Pertanto

$$I(\mathbf{x}_0, r) = \begin{cases} \{\mathbf{x}_0\} & \text{se } r \leq 1 \\ X & \text{se } r > 1 \end{cases}$$

Attenzione, va bene la definizione con $r \leq 1$ poiché l'intorno circolare è settato con la distanza strettamente minore di r , quindi strettamente minore di 1.

Definizione 1.2.4 (Insieme $A \subseteq X$ aperto). Insieme $A \subseteq X$ per il quale ogni elemento ha (almeno) un intorno sferico interamente contenuto in A :

$$\forall x \in A, \exists r \in \mathbb{R}^+ : I(x, r) \subseteq A$$

Esempio 1.2.8 (Esempio di aperto in $(\mathbb{R}, |\cdot|)$). In $(\mathbb{R}, |\cdot|)$, A è aperto se e solo se $\forall x \in A, \exists r \in \mathbb{R}^+$ tale che $(x - r, x + r) \subseteq A$.

Ad esempio (a, b) è aperto, infatti per qualsiasi punto x in esso contenuto si può costruire un insieme sferico con $r < \min\{d(x, a), d(x, b)\}$ che rispetta la definizione.

Esempio 1.2.9 (Controesempio di aperto in $(\mathbb{R}, |\cdot|)$). In $(\mathbb{R}, |\cdot|)$, $(a, b]$ invece non è aperto perché b appartiene all'insieme e un qualsiasi suo intorno sferico non è interamente contenuto in $(a, b]$.

Esempio 1.2.10 (Esempio di aperto in \mathbb{R}^2 euclideo). Sia $A = \{(x, y) \in \mathbb{R}^2 : x > 0\}$ (metà piano a destra dell'asse delle y). A è aperto in quanto per ogni punto (x_0, y_0) di A possiamo costruire un intorno con $r < d((0, y_0), (x_0, y_0))$ (distanza dall'asse delle y).

Esempio 1.2.11 (Controesempio di aperto in \mathbb{R}^2 euclideo). Se $A = \{(x, y) \in \mathbb{R}^2 : x \geq 0\}$ (metà piano a destra dell'asse delle y , questa inclusa), A non è aperto per i punti residenti sull'asse y . Gli intorno sferici (cerchi) non sono interamente contenuti in A .

Esempio 1.2.12 (Aperti nello spazio discreto (X, d)). Nello spazio metrico discreto abbiamo visto che:

$$I(\mathbf{x}_0, r) = \begin{cases} \{\mathbf{x}_0\} & \text{se } r \leq 1 \\ X & \text{se } r > 1 \end{cases}$$

Ne consegue, dato un $\mathbf{x}_0 \in A \subseteq X$ che con un r piccolo a piacere $I(\mathbf{x}_0, r) = \{\mathbf{x}_0\} \subseteq A$ per cui $\forall \mathbf{x} \in A, \exists I(\mathbf{x}, r) \subseteq A$ e dunque A è aperto.

Pertanto se (X, d) è uno spazio discreto, ogni $A \subseteq X$ è un aperto.

Definizione 1.2.5 (Punto di accumulazione per un insieme $A \subseteq X$). Considerato $A \subseteq X$, un punto $x \in X$ si dice di accumulazione per A se e solo se *qualsiasi* suo intorno sferico (escludendo x stesso) ha elementi in comune con A :

$$\forall r \in \mathbb{R}^+ : (I(x, r) \setminus \{x\}) \cap A \neq \emptyset$$

Esempio 1.2.13 ($0 \in \mathbb{R}$ è punto di accumulazione per $A = \{\frac{1}{n} : n \in \mathbb{Z}^+\}$?). Sia $A = \{\frac{1}{n} : n \in \mathbb{Z}^+\}$ l'insieme dei tick sempre più fitti nello spostarsi da 1 a 0 (es $1/2, 1/3$ etc), che include 1 ma non 0.

Sia $x_0 = 0 \in \mathbb{R}$: ci chiediamo se 0 sia punto di accumulazione per A . Se $I(0, r) = (-r, r)$, per n sufficientemente grande si ha che $\frac{1}{n} < r \iff \frac{1}{n} \in I(0, r)$, pertanto 0 è di accumulazione per A .

Esempio 1.2.14 (\mathbb{Z} ha punti di accumulazione in \mathbb{R} euclideo?). Sia $A = \mathbb{Z} \subset \mathbb{R}$. Ci chiediamo se \mathbb{Z} abbia punti di accumulazione in \mathbb{R} . Abbiamo due casi:

- se $x_0 \notin \mathbb{Z}$, ossia consideriamo un numero non intero si ha che, per un r piccolo a piacere l'intervallo circolare non include elementi di \mathbb{Z}

$$(I(x_0, r) \setminus \{x_0\}) \cap \mathbb{Z} = \emptyset$$

pertanto non può essere punto di accumulazione per \mathbb{Z} .

- se viceversa $x_0 \in \mathbb{Z}$ (consideriamo ora un intero), per un $r \leq 1$ si ha che $I(x_0, r) = \{x_0\}$ ossia l'intervallo circolare contiene solo il centro su cui è costruito, e applicando la definizione si ha

$$(I(x_0, r) \setminus \{x_0\}) \cap \mathbb{Z} = (\{x_0\} \setminus \{x_0\}) \cap \mathbb{Z} = \emptyset \cap \mathbb{Z} = \emptyset$$

pertanto non può essere punto di accumulazione per \mathbb{Z} .

Quindi si conclude che \mathbb{Z} non ha punti di accumulazione in \mathbb{R} .

Esempio 1.2.15 (Punti di accumulazione per un sottoinsieme dello spazio discreto (X, d)). Sia $A \subseteq X$; esistono punti di accumulazione per A ? Dato un $\mathbf{x}_0 \in X$ (sia esso interno ad A o meno) e sia $r \leq 1$ si ha che $I(\mathbf{x}_0, r) = \{\mathbf{x}_0\}$, pertanto:

$$(I(\mathbf{x}_0, r) \setminus \{\mathbf{x}_0\}) \cap A = (\{\mathbf{x}_0\} \setminus \{\mathbf{x}_0\}) \cap A = \emptyset \cap A = \emptyset$$

Perciò \mathbf{x}_0 non è punto di accumulazione per A e, in generale, non esistono punti di accumulazione (né interni né esterni) per un sottoinsieme dello spazio discreto.

Definizione 1.2.6 (Insieme $C \subseteq X$ chiuso). Un insieme $C \subseteq X$ si dice chiuso se e solo se (per definizione) il suo complementare $C^c = X \setminus C$ è aperto in X .

Osservazione 7. E viceversa un insieme è aperto se il complementare è chiuso.

Esempio 1.2.16 (Insiemi né chiusi né aperti (es in $(\mathbb{R}, |\cdot|)$ euclideo)). In generale, un insieme $B \subseteq X$ può essere né aperto né chiuso. Ad esempio in $(\mathbb{R}, |\cdot|)$ euclideo, l'intervallo $[0, 1)$ non è né aperto né chiuso:

- non è aperto perché per $0 \in [0, 1)$ è impossibile costruire un intorno circolare interamente contenuto in $[0, 1)$ (e quindi ai sensi della definizione non è aperto);
- non è chiuso poiché il complemento $\mathbb{R} \setminus [0, 1) = (-\infty, 0) \cup [1, +\infty)$ non è aperto. Infatti qui è impossibile costruire un intorno circolare per 1 che sia contenuto in $(-\infty, 0) \cup [1, +\infty)$

Teorema 1.2.1 (Insieme chiuso e punti di accumulazione). *Un insieme $C \subseteq \mathbb{R}^n$ è chiuso se e solo se contiene tutti i suoi punti di accumulazione (non ve ne sono al suo esterno).*

Dimostrazione. Effettuiamo la dimostrazione di entrambe le implicazioni di fatto dimostrandone le rispettive negazioni ossia:

1. nell'ipotesi che A non sia chiuso dimostriamo che non contiene tutti i suoi punti di accumulazione.

Sia A non chiuso, per cui A^c non è aperto: ergo esiste in quest'ultimo un suo punto con intorno non interamente contenuto in A^c :

$$\exists x_0 \in A^c, \forall r \in \mathbb{R}^+ : I(x_0, r) \not\subseteq A^c$$

Ma questo implica che questo intorno “esca da A^c ” ed abbia una intersezione con A ossia:

$$I(x_0, r) \cap A \neq \emptyset$$

Ora, se $x_0 \in A^c$, si ha che $x_0 \notin A$ pertanto possiamo riscrivere equivalentemente la precedente come:

$$(I(x_0, r) \setminus \{x_0\}) \cap A \neq \emptyset, \quad \forall r \in \mathbb{R}^+$$

Pertanto x_0 è punto di accumulazione per A con $x_0 \notin A$. Quindi non è vero che A contiene tutti i suoi punti di accumulazione.

2. nell'ipotesi di avere un x_0 di accumulazione che non appartenga ad un insieme A dimostriamo che questi non è chiuso.

Sia x_0 di accumulazione per A ma $x_0 \notin A$, ossia che $x_0 \in A^c$. Essendo punto di accumulazione si ha che qualsiasi intorno circolare ha punti in comune con A , ossia:

$$\forall r \in \mathbb{R}^+ : (I(x_0, r) \setminus \{x_0\}) \cap A \neq \emptyset$$

Ora dato che $x_0 \notin A$ (ossia $x_0 \in A^c$) possiamo togliere tranquillamente dall'intersezione di sopra, giungendo al fatto che valga, sotto le nostre ipotesi, quanto segue:

$$\forall r \in \mathbb{R}^+ : (I(x_0, r)) \cap A \neq \emptyset$$

Ma allora $\forall r \in \mathbb{R}^+$, se l'intorno ha punti di intersezione con A possiamo affermare che non sia un sottoinsieme interamente contenuto in A^c , ossia che valga:

$$I(x_0, r) \not\subseteq A^c$$

Pertanto $\exists x_0 \in A^c : \exists I(x_0, r) \not\subseteq A^c$. Ergo A^c non è aperto; e quindi A non è chiuso.

□

Osservazione 8. Per dimostrare che un insieme non è chiuso basta trovare un punto di accumulazione che non gli appartenga.

Esempio 1.2.17 (Applicazioni: $A = \{\frac{1}{n} : n \in \mathbb{Z}^+\}$ è chiuso?). Si ha che $A = \{\frac{1}{n} : n \in \mathbb{Z}^+\}$ non è chiuso, in quanto 0 è di accumulazione per A ma $0 \notin A$.

Esempio 1.2.18 (Applicazioni: sottoinsiemi chiusi nello spazio discreto (X, d)). Considerando lo spazio discreto (X, d) quali sono gli insiemi chiusi?

Ipotizziamo ad esempio che sia $X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$. Dire che il sottoinsieme $\{\mathbf{x}_1\}$ contiene tutti i suoi punti di accumulazione equivale a dire che non ve ne sono di esterni e sia \mathbf{x}_2 che \mathbf{x}_3 (che sono diversi da \mathbf{x}_1) non hanno intorno piccoli a piacere che includano \mathbf{x}_1 , quindi non sono punti di accumulazione e quindi $\{\mathbf{x}_1\}$ contiene “tutti” i suoi punti di accumulazione (o meglio non ve ne sono all'esterno come visto in precedenza).

Pertanto tutti i sottoinsiemi di X sono chiusi.

1.2.1.3 Apertura/chiusura di \emptyset e X

Proposizione 1.2.2 (Apertura/chiusura di \emptyset, X). *Dato uno spazio metrico (X, d) i sottoinsiemi X e \emptyset sono sia aperti che chiusi.*

NB: Tentativo mio

Dimostrazione. Partiamo dimostrando che X è sia aperto che chiuso:

- X è aperto perché ogni punto è interno e ha almeno un intorno sferico con elementi in comune con X
- X è chiuso perché non esiste altro punto al suo esterno che possa fungere da punto di accumulazione

Dimostriamo quanto concerne \emptyset a partire da questo. Dato che:

- X è aperto: \emptyset è chiuso essendo il suo complementare $X \setminus \emptyset = X$ aperto;
- X è chiuso: da definizione deriva che il suo complementare $X \setminus X = \emptyset$ è aperto.

□

Esempio 1.2.19 (Apertura/chiusura di \emptyset, \mathbb{R}). I sottoinsiemi \emptyset e \mathbb{R} in \mathbb{R} euclideo sono contemporaneamente sia aperti che chiusi.

NB: esempio mio

Esempio 1.2.20 (Apertura/chiusura di $X = [1, 3]$). Siano \emptyset e $X = [1, 3]$

- sicuramente $X = [1, 3]$ è chiuso, perché non esistono punti di accumulazione al di fuori di esso;
- tuttavia è anche aperto: l'intorno circolare (anche sia esso centrato su 1 o 3) non “guarda” all'esterno dell'insieme perché per definizione

$$I(x_0, r) = \{x \in X : d(x, x_0) < r\}$$

dove è l' $x \in X$ della definizione ad essere importante. Ad esempio in questo caso l'intorno circolare centrato in 1 e di raggio 0.5 è:

$$I(1, 0.5) = [1, 1.5] \subseteq [1, 3]$$

Quindi qualsiasi intorno circolare costruito su uno degli elementi dell'insieme è contenuto nell'insieme stesso e l'insieme è aperto.

1.2.1.4 Sconnessione/connessione e insiemi aperti/chiusi

Definizione 1.2.7 (Spazio metrico sconnesso). Lo spazio metrico (X, d) si dice sconnesso se esistono due sottoinsiemi (con determinate caratteristiche) formanti una partizione e nello specifico se:

- $\exists A_1, A_2 \subseteq X$, aperti
- $A_1, A_2 \neq \emptyset, X$
- si ha che:

$$X = A_1 \cup A_2 \quad A_1 \cap A_2 = \emptyset \quad (1.1)$$

Esempio 1.2.21 (Spazio metrico sconnesso: \mathbb{R}^2 senza circonferenza unitaria). Sia il piano cartesiano senza la circonferenza unitaria

$$X = \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

si ha che X è sconnesso, infatti:

- $X = A_1 \cup A_2$ con
- $A_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 > 1\}$, aperto, ossia l'insieme dei punti esterni alla circonferenza unitaria;
- $A_2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$, aperto, ossia l'insieme dei punti interni alla stessa.

Esempio 1.2.22 (Spazio metrico sconnesso: \mathbb{Q}). L'insieme dei razionali \mathbb{Q} con distanza euclidea è sconnesso. Se consideriamo $\pi \in \mathbb{R} \setminus \mathbb{Q}$ possiamo partizionare \mathbb{Q} in due sottoinsiemi:

$$\begin{aligned} \mathbb{Q}_{<\pi} &= \{q \in \mathbb{Q} : q < \pi\} \subseteq \mathbb{Q} \neq \emptyset, \mathbb{Q} \\ \mathbb{Q}_{>\pi} &= \{q \in \mathbb{Q} : q > \pi\} \subseteq \mathbb{Q} \neq \emptyset, \mathbb{Q} \end{aligned}$$

Ora si ha che $\mathbb{Q}_{<\pi}$ e $\mathbb{Q}_{>\pi}$ sono entrambi aperti, che $\mathbb{Q} = \mathbb{Q}_{<\pi} \cup \mathbb{Q}_{>\pi}$ e $\mathbb{Q}_{<\pi} \cap \mathbb{Q}_{>\pi} = \emptyset$.

Definizione 1.2.8 (Spazio metrico connesso). Non sconnesso.

Osservazione 9. Abbiamo visto che \emptyset e X sono sia aperti che chiusi. Il seguente risultato mostra un legame tra sconnessione e *altri* sottoinsiemi di X sia aperti che chiusi.

Teorema 1.2.3 (Sconnessione e sottoinsiemi sia aperti che chiusi). (X, d) è uno spazio metrico sconnesso $\iff \exists A \subseteq X, A \neq \emptyset, X$ con A sia aperto che chiuso.

Dimostrazione. **Dimostriamo:**

- ipotizziamo che (X, d) sia sconnesso: allora esistono $A_1, A_2 \subseteq X$ con $A_1, A_2 \neq \emptyset, X$, ed entrambi aperti tali che: $X = A_1 \cup A_2$ e $A_1 \cap A_2 = \emptyset$. Tuttavia $A_1 = A_2^c$ ed essendo A_1 aperto A_2 è chiuso. Pertanto abbiamo dimostrato che A_2 è sia aperto che chiuso (e stessa cosa si può fare per A_1 a parti inverse; quindi possiamo dimostrare che ve ne sono ben due di insiemi sia aperti che chiusi)
- ipotizzando invece che $\exists A \subset X$ ($A \neq \emptyset, X$) sia aperto che chiuso: possiamo allora porre $A_1 = A$ e $A_2 = A^c$. Si ha che A_1 è aperto (come A) e A_2 anche (perché il suo complementare A è anche chiuso); inoltre $X = A_1 \cup A_2$, $A_1 \cap A_2 = \emptyset$ e $A_1, A_2 \neq \emptyset, X$. Quindi (X, d) è sconnesso.

□

Corollario 1.2.4 (Connessione e sottoinsiemi sia aperti che chiusi). (X, d) è connesso $\iff \nexists A \neq \emptyset, X$ con A sia aperto che chiuso.

Dimostrazione. Negazione degli enunciati del precedente teorema. □

Teorema 1.2.5 (Teorema di connessione per \mathbb{R}^n euclideo). Lo spazio \mathbb{R}^n euclideo (con $n \in \mathbb{Z}^+$) è connesso.

Osservazione 10. In \mathbb{R}^n euclideo ($n \in \mathbb{Z}^+$), gli unici sottoinsiemi contemporaneamente aperti e chiusi sono \emptyset e \mathbb{R}^n .

Corollario 1.2.6. In \mathbb{R}^n euclideo, per i rimanenti $B \subset \mathbb{R}^n$, $B \neq \emptyset$:

- se B è aperto, allora B non è chiuso.
- se B è chiuso, allora B non è aperto.

1.2.1.5 Apertura/chiusura di unione/intersezione di famiglie di aperti/chiusi

Proposizione 1.2.7 (Apertura/chiusura di operazioni insiemistiche su famiglie di aperti/chiusi). In uno spazio metrico qualsiasi (X, d) sia $\{A_i \subseteq X : i \in \mathcal{I}\}$ una famiglia di sottoinsiemi di X (finita o infinita a seconda della cardinalità dell'insieme di indici \mathcal{I}):

- A_1) l'unione di una famiglia qualsiasi di insiemi aperti è un insieme aperto.
- A_2) l'intersezione di una famiglia finita di insiemi aperti è un insieme aperto.
- C_1) l'unione di una famiglia finita di insiemi chiusi è un insieme chiuso.
- C_2) l'intersezione di una famiglia qualsiasi di insiemi chiusi è un insieme chiuso.

Dimostrazione. Mostriamo A_1 e A_2 ; si ha poi che $A_1 \iff C_2$ e $A_2 \iff C_1$ sulla base delle leggi di DeMorgan:

$$\begin{aligned} \left(\bigcup_i A_i \right)^C &= \bigcap_i A_i^C \\ \left(\bigcap_{i=1}^n A_i \right)^C &= \bigcup_{i=1}^n A_i^C \end{aligned}$$

in quanto ipotizzando gli A_i aperti, si ha che A_i^C sono chiusi, mentre la prima è una unione/intersezione infinita e la seconda finita.

1. proviamo A_1): siano A_i aperti, proviamo che l'unione infinita $\bigcup_i A_i = A$ è un aperto. Si ha che

$$\forall i \in \mathcal{I}, \forall x \in A_i : \exists I(x, r_i) \subseteq A_i$$

Da questa discende che andando a fare l'unione di questi punti appartenenti ai vari A_i si abbia (intendendo con A l'unione infinita degli A_i):

$$\forall x \in A : \exists I(x, r_i) \subseteq A$$

cioè A è aperto.

2. Proviamo A_2): siano A_i aperti, proviamo che l'intersezione finita $\bigcap_{i=1}^n A_i = A$ è un aperto. Dal fatto che gli A_i siano aperti si ha per definizione che

$$\forall x \in A_i : \exists I(x, r_i) \subseteq A_i$$

Ora se x appartiene all'intersezione, appartiene a tutti gli A_i , $\forall i \in \mathcal{I}$ quindi per ipotesi possiamo costruire un intorno che appartenga a ciascuno degli A_i , ossia

$$\begin{aligned} \exists r_1 \in \mathbb{R}^+ : I(x, r_1) &\subseteq A_1 \\ \exists r_2 \in \mathbb{R}^+ : I(x, r_2) &\subseteq A_2 \\ \dots \\ \exists r_n \in \mathbb{R}^+ : I(x, r_n) &\subseteq A_n \end{aligned}$$

Ora se però prendiamo il minimo degli n r_i , ossia $\underline{r} = \min \{r_1, r_2, \dots, r_n\}$. Si ha che l'intorno costruito appartenga a tutti di A_i , quindi

$$I(x, \underline{r}) \subseteq \bigcap_{i=1}^n A_i$$

ossia $\bigcap_{i=1}^n A_i$ è aperto.

□

Osservazione 11. A_2 e C_1 sono *false* se si trascura l'ipotesi di *finitzza*. Vediamo due controesempi in \mathbb{R} euclideo.

Esempio 1.2.23 (Controesempio intersezione infinita di aperti). La seguente è una intersezione infinita di aperti (i segmenti attorno a 0 via via più piccoli):

$$\bigcap_{n \in \mathbb{Z}^+} \left(-\frac{1}{n}, \frac{1}{n} \right) = \{0\}$$

Il risultato è un chiuso, in quanto il suo complementare a \mathbb{R} , ossia $(-\infty, 0) \cup (0, \infty)$ è aperto. Pertanto essendo chiuso ed essendo diverso da \emptyset o \mathbb{R} è *non* aperto.

Esempio 1.2.24 (Controesempio unione infinita di chiusi). La seguente è una unione infinita di chiusi:

$$\bigcup_{n \in \mathbb{Z}^+} \left[-1 + \frac{1}{n}, 1 - \frac{1}{n} \right] = (-1, 1)$$

Il risultato è un aperto, in base alla definizione, ed essendo diverso da \emptyset o \mathbb{R} è *non* chiuso.

1.2.1.6 Funzioni $f : A \subseteq X \rightarrow Y$ continue, aperti e chiusi

Osservazione 12. Siano (X, d_X) , (Y, d_Y) spazi metrici; sia $f : A \subseteq X \rightarrow Y$ funzione.

Definizione 1.2.9 (Continuità in un punto \mathbf{x}_0 di una funzione f). Dato $\mathbf{x}_0 \in A$, si dice che la funzione f è *continua nel punto* \mathbf{x}_0 se e solo se

$$\forall \varepsilon \in \mathbb{R}^+, \exists \delta \in \mathbb{R}^+ : d_Y(f(\mathbf{x}), f(\mathbf{x}_0)) < \varepsilon, \quad \forall \mathbf{x} \in I(\mathbf{x}_0, \delta) \cap A. \quad (1.2)$$

Esempio 1.2.25 (Continuità in un punto \mathbf{x}_0 di una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$). Se $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione a valori reali, e supponendo (come faremo d'ora negli esempi su \mathbb{R}^n) che il dominio \mathbb{R}^n ed il codominio \mathbb{R} siano muniti delle rispettive metriche euclidee, allora f è continua in $\mathbf{x}_0 \in A$ se e solo se

$$\forall \varepsilon \in \mathbb{R}^+, \exists \delta \in \mathbb{R}^+ : \sqrt{(f(\mathbf{x}) - f(\mathbf{x}_0))^2} = |f(\mathbf{x}) - f(\mathbf{x}_0)| < \varepsilon, \quad \forall \mathbf{x} \in I(\mathbf{x}_0, \delta) \cap A$$

Definizione 1.2.10 (Continuità globale sul dominio A di una funzione f). La funzione f si dice *continua sul dominio* A se e solo se risulta continua *ogni punto* di A .

Definizione 1.2.11 (Immagine di una funzione f). Data una funzione $f : A \subseteq X \rightarrow Y$ tra spazi metrici (X, d_X) (Y, d_Y) , la sua immagine è l'insieme dei valori di Y assunti ossia $f[A] = \{f(x) : x \in X\} \subseteq Y$

Definizione 1.2.12 (Preimmagine, o fibra, di un sottoinsieme del codominio). Dato $C \subseteq Y$, la sua preimmagine rispetto alla funzione f è l'insieme:

$$f^{-1}[C] = \{x \in A : f(x) \in C\} \subseteq A \subseteq X.$$

Esempio 1.2.26 (Fibra di una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$). Nel caso di $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ la preimmagine (o, fibra) di $C \subseteq \mathbb{R}$ rispetto alla funzione f è l'insieme:

$$f^{-1}[C] = \{\mathbf{x} \in A : f(\mathbf{x}) \in C\} \subseteq A \subseteq \mathbb{R}^n$$

Osservazione 13. I seguenti risultati stanno a significare che una funzione continua trasforma aperti in aperti e chiusi in chiusi.

Teorema 1.2.8 (Caratterizzazione delle funzioni continue sul proprio dominio).

Data $f : A \subseteq X \rightarrow Y$, $A \neq \emptyset$, le seguenti affermazioni sono equivalenti:

- f è continua su A
- $\forall B \subseteq Y$, con B aperto, $\exists B_1 \subseteq X$, B_1 anch'esso aperto, tale che $f^{-1}[B] = A \cap B_1$.
Se poi A è anche aperto (o sia aperto che chiuso, es se $A = X$), anche $f^{-1}[B]$ è aperto in quanto intersezione di aperti;
- $\forall C \subseteq Y$, con C chiuso, $\exists C_1 \subseteq X$, C_1 anch'esso chiuso, tale che $f^{-1}[C] = A \cap C_1$.
Se poi A è anche chiuso (o sia aperto che chiuso, es se $A = X$), anche $f^{-1}[C]$ è chiuso in quanto intersezione di chiusi

Dimostrazione. Non fatta. □

Osservazione 14. Il teorema precedente ha applicazioni pratiche nel senso che permette di dire se una funzione è continua, anche in setup particolari, verificando “semplicemente” se trasforma aperti in aperti o chiusi in chiusi.

Esempio 1.2.27 (Quali funzioni \mathbb{R} discreto $\rightarrow \mathbb{R}$ euclideo sono continue?). Supponendo di avere l'insieme di $f : \mathbb{R} \rightarrow \mathbb{R}$ con dominio avente la metrica della topologia discreta e codominio l'euclidea, quali sono le funzioni continue? Sia B aperto in \mathbb{R} euclideo (codominio), da definizione della fibra si ha che

$$f^{-1}[B] = \{x \in \mathbb{R} \text{ discreto} : f(x) \in B\} \subseteq \mathbb{R} \text{ discreto}$$

Ossia la fibra è un sottoinsieme di uno spazio discreto: tali sottoinsiemi sono tutti aperti (cfr esempio 1.2.12), si ha che $f^{-1}[B]$ è aperto! Pertanto qualsiasi funzione di questa tipologia è continua.

Esempio 1.2.28 (Quali funzioni \mathbb{R} euclideo $\rightarrow \mathbb{R}$ discreto sono continue?). A setup di metrica opposto su dominio/codominio quali sono le funzioni continue? Sia f non costante, ossia per $x_1, x_2 \in \mathbb{R}$ euclideo si ha $f(x_1) = L_1 \neq L_2 = f(x_2)$.

- supponiamo f continua. Sia $B = \mathbb{R} \setminus \{L_1\} \subseteq Y$ con B aperto (per proposizione 1.2.7 è l'unione finita di aperti, ossia $(-\infty, L_1) \cup (L_1, \infty)$). Allora si ha che:
 - la fibra $f^{-1}[\mathbb{R} \setminus \{L_1\}]$ contiene x_2 ma non x_1 ; pertanto non è né l'insieme vuoto \emptyset né \mathbb{R} (che potrebbe esser sia aperto che chiuso);
 - essendo $\mathbb{R} \setminus \{L_1\}$ aperto ed essendo per ipotesi f continua si ha che $f^{-1}[\mathbb{R} \setminus \{L_1\}]$ è aperto.

Ora veniamo a considerare la parte rimanente del codominio ossia $\{L_1\}$. Questi è aperto, sempre essendo un sottoinsieme di spazio discreto per esempio 1.2.12; pertanto nell'ipotesi che f sia continua $f^{-1}[\{L_1\}]$ è anch'esso un aperto in \mathbb{R} euclideo.

Ma se sia $f^{-1}[\{L_1\}]$ che $f^{-1}[\mathbb{R} \setminus \{L_1\}]$ sono aperti formano una partizione di \mathbb{R} ossia potremmo scrivere \mathbb{R} come spazio sconnesso in quanto

$$\begin{aligned} f^{-1}[\mathbb{R} \setminus \{L_1\}] \cup f^{-1}[\{L_1\}] &= \mathbb{R} \\ f^{-1}[\mathbb{R} \setminus \{L_1\}] \cap f^{-1}[\{L_1\}] &= \emptyset \end{aligned}$$

con $f^{-1}[\mathbb{R} \setminus \{L_1\}]$, $f^{-1}[\{L_1\}]$ aperti $\neq \emptyset, \mathbb{R}$. Ma tutto ciò è assurdo per il teorema di connessione euclidea 1.2.5.

Pertanto partiti con f costante e ipotizzando f continua, abbiamo mandato in crash il sistema, quindi abbiamo provato

$$f \text{ non costante} \implies f \text{ non continua}$$

- provare il viceversa

$$f \text{ non continua} \implies f \text{ non costante}$$

è banale

Pertanto per le funzioni da \mathbb{R} euclideo a \mathbb{R} discreto si ha che:

$$f \text{ non costante} \iff f \text{ non continua}$$

e quindi

$$f \text{ costante} \iff f \text{ continua}$$

da cui le funzioni continue in questo setup sono le funzioni costanti

1.2.1.7 Sottospazio metrico

Definizione 1.2.13 (Sottospazio metrico (A, d_A) di (X, d)). Sia (X, d) spazio metrico e sia $A \subseteq X$, $A \neq \emptyset$; allora A eredita una struttura di spazio metrico per cui sia

$$d_A : A \times A \rightarrow \mathbb{R} : d_A(a, a') = d(a, a') \quad \forall a, a' \in A \quad (1.3)$$

A parole la distanza funziona allo stesso modo anche se settata operativamente sul sottoinsieme di elementi composto da A .

1.2.2 Spazi normati

1.2.2.1 Spazio vettoriale e spazio normato

Definizione 1.2.14 (Spazio vettoriale su \mathbb{R}^n). È l'insieme di elementi, chiamati *vettori*, definiti come tupla ordinata di reali $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$, sul quale sono definite le operazioni (interne all'insieme stesso):

1. addizione tra vettori: funzione $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita come

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

2. prodotto di vettore per scalare: funzione $\mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita come

$$c \cdot \mathbf{x} = (c \cdot x_1, c \cdot x_2, \dots, c \cdot x_n)$$

che soddisfano i seguenti assiomi/ *regole di calcolo* (qui $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e $a, b \in \mathbb{R}$):

(A₀) Proprietà commutativa² dell'addizione:

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \quad (1.4)$$

(A₁) Proprietà associativa dell'addizione:

$$(\mathbf{u} + \mathbf{x}) + \mathbf{y} = \mathbf{u} + (\mathbf{x} + \mathbf{y}) \quad (1.5)$$

(A₂) Esiste un vettore $\mathbf{0} = (0, 0, \dots, 0) \in X$, detto vettore nullo, neutro per l'addizione:

$$\mathbf{x} + \mathbf{0} = \mathbf{x} \quad (1.6)$$

(A₃) Esistenza dell'opposto di un vettore dato: esiste un vettore $\mathbf{y} \in X$ detto opposto di \mathbf{x} tale che:

$$\mathbf{x} + \mathbf{y} = \mathbf{0} \quad (1.7)$$

(M₁) Proprietà associativa del prodotto per uno scalare:

$$a(b\mathbf{x}) = (ab)\mathbf{x} \quad (1.8)$$

(M₂) Proprietà distributiva del prodotto per uno scalare rispetto alla somma di scalari:

$$(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x} \quad (1.9)$$

(M₃) Proprietà distributiva del prodotto per uno scalare rispetto alla somma di vettori:

$$\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y} \quad (1.10)$$

(M₄) Normalizzazione del prodotto per uno scalare:

$$1\mathbf{x} = \mathbf{x} \quad (1.11)$$

Definizione 1.2.15 (Norma). Sia X uno spazio vettoriale su \mathbb{R}^n ; una funzione $\|\cdot\| : X \rightarrow \mathbb{R}$ si dice norma se soddisfa i seguenti assiomi (considerando $\mathbf{x}, \mathbf{y} \in X$ e $\lambda \in \mathbb{R}$):

1. $\|\mathbf{x}\| \geq 0$; $\|\mathbf{x}\| = 0 \in \mathbb{R} \iff \mathbf{x} = \mathbf{0} \in X$;
2. $\|\lambda \cdot \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$;
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (disuguaglianza triangolare).

Definizione 1.2.16 (Spazio normato). La coppia $(X, \|\cdot\|)$ formata da uno spazio vettoriale ed una norma su esso definita.

²In alcune trattazioni (ad esempio Salce) questa proprietà è fatta derivare dalle altre e pertanto non costituisce assioma. Per semplicità qui la si considera come assioma

1.2.2.2 Spazio normato e spazio metrico

Proposizione 1.2.9 (Spazio normato è spazio metrico). *Sia $(X, \|\cdot\|)$ uno spazio normato. Se per $\mathbf{x}, \mathbf{y} \in X$ definiamo la funzione*

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \quad (1.12)$$

allora d è una distanza/metrica e (X, d) è spazio metrico.

Dimostrazione. Applicando le proprietà di una norma per mostrare la prima proprietà della distanza si ha:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \geq 0$$

ed inoltre ipotizzando di avere due elementi uguali la loro distanza (via norma) è nulla:

$$\mathbf{x} = \mathbf{y} \iff \mathbf{x} - \mathbf{y} = \mathbf{0} \xLeftrightarrow{(1)} \underbrace{\|\mathbf{x} - \mathbf{y}\|}_{d(\mathbf{x}, \mathbf{y})} = 0$$

dove in (1) siamo passati alle norme ad entrambi i membri.

Per dimostrare la seconda proprietà della distanza si usa la seconda della norma:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \|(-1)(\mathbf{y} - \mathbf{x})\| = |-1| \|\mathbf{y} - \mathbf{x}\| = d(\mathbf{y}, \mathbf{x})$$

Infine per la terza proprietà:

$$d(\mathbf{x}, \mathbf{z}) = \|\mathbf{x} - \mathbf{z}\| = \|(\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z})\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\| = d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$$

□

Osservazione 15. Quindi dalla 1.2.9 discende che:

- gli spazi normati sono un sottoinsieme degli spazi metrici e ogni concetto definito per gli spazi metrici rimane naturalmente definito per gli spazi normati;
- per gli spazi normati, si ha che la norma di un vettore $\|\mathbf{x}\|$ coincide con la distanza tra questo e il vettore nullo:

$$\|\mathbf{x}\| = d(\mathbf{x}, \mathbf{0})$$

dove abbiamo solo sostituito $\mathbf{0}$ al posto di \mathbf{y} nella precedente proposizione.

Esempio 1.2.29 (Norma euclidea e spazio normato $(\mathbb{R}^n, \|\cdot\|)$). Su \mathbb{R}^n la funzione $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ definita come

$$\|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

è norma (nello specifico è detta *norma euclidea*) in quanto soddisfa gli assiomi della definizione. Pertanto $(\mathbb{R}^n, \|\cdot\|)$ è spazio normato.

Esempio 1.2.30 (Distanza euclidea reforged). Data la norma euclidea nello spazio normato $(\mathbb{R}^n, \|\cdot\|)$, dati $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$, la distanza euclidea è definita come:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

1.2.2.3 Sottospazio normato

Definizione 1.2.17 (Sottospazio normato). Sia $(X, \|\cdot\|)$ uno spazio normato e sia $A \subseteq X$, $A \neq \emptyset$. Allora A eredita una struttura di spazio normato se anche esso è uno spazio vettoriale, ossia A è un sottospazio vettoriale, ossia è chiuso rispetto a somma di elementi e al prodotto di un elemento per una costante:

$$\begin{aligned} \mathbf{x}, \mathbf{y} \in A &\implies \mathbf{x} + \mathbf{y} \in A \\ a \in \mathbb{R}, \mathbf{x} \in A &\implies a\mathbf{x} \in A \end{aligned}$$

Esempio 1.2.31 ((meh) \mathbb{R}^2 è sottospazio normato di \mathbb{R}^n con norma euclidea). Si essendo \mathbb{R}^2 sottospazio.

1.2.3 Spazi con prodotto interno

1.2.3.1 Prodotto interno e spazi con

Definizione 1.2.18 (Prodotto interno). Dato uno spazio vettoriale X sul campo reale \mathbb{R} , si dice prodotto interno³ su X una funzione $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$ che rispetti i seguenti assiomi ($\mathbf{x}, \mathbf{x}', \mathbf{y}, \mathbf{y}' \in X$, $\lambda \in \mathbb{R}$):

1. è *definito positivo* ossia $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$; $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$
2. è *simmetrico*, ossia $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
3. è *bilineare*⁴ ossia:
 - $\langle \mathbf{x} + \mathbf{x}', \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}', \mathbf{y} \rangle$ e anche $\langle \mathbf{x}, \mathbf{y} + \mathbf{y}' \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y}' \rangle$
 - $\langle \lambda \cdot \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \lambda \cdot \mathbf{y} \rangle = \lambda \cdot \langle \mathbf{x}, \mathbf{y} \rangle$,

Definizione 1.2.19 (Spazio con prodotto interno). La coppia $(X, \langle \cdot, \cdot \rangle)$ composta dallo spazio vettoriale X e dal prodotto interno su esso definito.

1.2.3.2 Spazio con prodotto interno euclideo è normato

Esempio 1.2.32 (Prodotto interno standard/euclideo in \mathbb{R}^n). Su \mathbb{R}^n , la funzione definita come

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i = x_1 y_1 + \cdots + x_n y_n$$

con $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ è un prodotto interno, in quanto soddisfa la definizione.

Per Sylvester's Law of inertia, il prodotto interno euclideo è essenzialmente l'unico prodotto interno in \mathbb{R}^n

³In altre trattazioni, il prodotto interno viene indicato in altri modi; ai fini di questo corso per raffronto con altri è meglio abituarsi a leggere

$$\langle \mathbf{x}, \mathbf{y} \rangle = (\mathbf{x}|\mathbf{y}) = \mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}$$

⁴In parole povere funziona come una funzione lineare separatamente, se entrambe le componenti. Funzioni lineari, concetto di algebra lineare, sono richiamate sotto in definizione 1.4.5.

Proposizione 1.2.10 (\mathbb{R}^n con prodotto interno è spazio normato). *Se \mathbb{R}^n è spazio con prodotto interno è anche spazio normato (e dunque anche spazio metrico).*

Dimostrazione. Sia \mathbb{R}^n spazio con prodotto interno: considerato $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, la radice del prodotto interno del vettore con se stesso è una norma, e nello specifico la norma euclidea

$$\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_1^2 + \dots + x_n^2} = \|\mathbf{x}\|$$

Vi è pertanto un legame diretto tra prodotto interno di un vettore (con se stesso) e norma euclidea. \square

Osservazione 16. Dato che \mathbb{R}^n con prodotto interno è anche normato, allora è anche spazio metrico. La distanza si definirà applicando la precedente a ritroso come

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle}$$

1.2.3.3 Disuguaglianza di Cauchy-Schwartz e interpretazione geometrica del prodotto interno euclideo

Osservazione 17 (Prodotto interno in \mathbb{R}^n : interpretazione geometrica). Il prodotto interno euclideo in \mathbb{R}^n permette di dare una definizione trasparente della nozione di *angolo* tra due vettori non nulli di \mathbb{R}^n . Il punto di partenza è il seguente risultato.

Teorema 1.2.11 (Disuguaglianza di Cauchy-Schwarz). *Se $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$:*

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle} = \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

Proposizione 1.2.12 (Interpretazione geometrica del prodotto interno). *Siano $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, con $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$. Si ha che:*

$$\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

con θ angolo tra i vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, ossia il prodotto interno è proporzionale all'angolo che formano i due vettori

Dimostrazione. Riscriviamo la disuguaglianza di Cauchy-Schwartz come:

$$-\|\mathbf{x}\| \cdot \|\mathbf{y}\| \leq \langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

Dividendo tutti i membri per il positivo $\|\mathbf{x}\| \cdot \|\mathbf{y}\|$ abbiamo che:

$$-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \leq 1$$

Ora se consideriamo la funzione $\cos : [0, \pi] \rightarrow [-1, 1]$ in questo tratto $[0, \pi]$ è biunivoca e invertibile. Pertanto esisterà uno e un solo reale $\theta \in [0, \pi]$, appartenente al dominio, tale che il valore assunto dalla funzione

$$\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

Per definizione, il valore θ si dice *angolo* tra i vettori (non nulli) $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Per angolo si intende l'angolo non orientato (ossia il più piccolo, quello che sta entro i 180 gradi). \square

Osservazione 18. Si ha che

- $\langle \mathbf{x}, \mathbf{y} \rangle > 0 \iff 0 \leq \theta < \frac{\pi}{2}$ ossia se l'angolo θ è acuto
- $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \iff \theta = \frac{\pi}{2}$ ossia se l'angolo θ è retto e in vettori \mathbf{x}, \mathbf{y} sono ortogonali
- $\langle \mathbf{x}, \mathbf{y} \rangle < 0 \iff \frac{\pi}{2} < \theta \leq \pi$ ossia se l'angolo θ è ottuso

Corollario 1.2.13 (Prodotto interno: definizione alternativa). *Dall'equazione precedente otteniamo una definizione/specificazione alternativa del prodotto interno come:*

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cdot \cos \theta \quad (1.13)$$

Esempio 1.2.33 (Calcolo di prodotto interno e angolo di due vettori). Siano $\mathbf{x} = (x_1, 0), \mathbf{y} = (0, y_2)$ vettori in \mathbb{R}^2 , $x_1, y_2 \neq 0$. Allora

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 = \cos \left(\frac{\pi}{2} \right),$$

cioè i vettori \mathbf{x} e \mathbf{y} sono *ortogonali*.

1.2.3.4 Spazio con prodotto interno generico è normato

Osservazione 19. Generalizziamo ora il rapporto visto in precedenza per \mathbb{R}^n tra prodotto interno e norma

Proposizione 1.2.14 (Spazio (generico) con prodotto interno è normato). *Se $(X, \langle \cdot, \cdot \rangle)$ è uno spazio con prodotto interno e poniamo*

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

si ha che $\|\cdot\|$ è effettivamente una norma su X

Dimostrazione. Dato che $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ per definizione allora anche $\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \geq 0$; inoltre

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = 0 \iff \langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$$

Per la seconda proprietà della norma si ha che

$$\|\lambda \mathbf{x}\| = \sqrt{\langle \lambda \mathbf{x}, \lambda \mathbf{x} \rangle} = \sqrt{\lambda \langle \mathbf{x}, \lambda \mathbf{x} \rangle} = \sqrt{\lambda^2 \langle \mathbf{x}, \mathbf{x} \rangle} = |\lambda| \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = |\lambda| \|\mathbf{x}\|$$

Infine per la disuguaglianza triangolare, ossia:

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

consideriamo che $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$; se sviluppiamo il primo membro $\|\mathbf{x} + \mathbf{y}\|$ al quadrato (per eliminare la radice) si ha

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

Si noti che, se i vettori \mathbf{x} ed \mathbf{y} sono ortogonali, allora $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ e la precedente identità si riduce al teorema di Pitagora.

Concentrandoci sull'ultimo pezzo addendo, ossia su $2\langle \mathbf{x}, \mathbf{y} \rangle$, dalla disuguaglianza di Cauchy-Schwartz si ha (ignorando il valore assoluto, tanto la scrittura è valida comunque):

$$2\langle \mathbf{x}, \mathbf{y} \rangle \leq 2\|\mathbf{x}\|\|\mathbf{y}\|$$

Perciò:

$$\|\mathbf{x} + \mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$$

e quindi, come volevasi dimostrare (estraendo la radice da entrambi 2 reali positivi):

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

□

Osservazione 20 (Conclusione importante). Quindi ogni spazio con prodotto interno è, secondo la proposizione 1.2.14, uno spazio normato, e per proposizione 1.2.9 anche uno spazio metrico. Si ha allora

$$\text{spazi con prodotto interno} \subseteq \text{spazi normati} \subseteq \text{spazi metrici}$$

Pertanto tutti i concetti definiti nell'ambito degli spazi metrici (ad esempio convergenza, limitatezza, continuità) hanno senso per gli spazi con prodotto interno (e per gli spazi normati).

Osservazione 21. Perciò tutti i concetti definiti per \mathbb{R}^n come spazio metrico euclideo (ad es. *insiemi limitati, insiemi aperti, punti di accumulazione, insiemi chiusi, funzioni continue, successioni convergenti e di Cauchy, completezza*) mantengono invariato il loro significato quando riguardiamo \mathbb{R}^n come spazio con prodotto interno euclideo.

1.3 Funzioni $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ e curve in \mathbb{R}^m

1.3.1 Funzioni

Definizione 1.3.1 (Funzione parametrica in spazio vettoriale e componenti). Funzione del tipo $\mathbf{f} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ che può esser definita fornendo, per ciascun

elemento del vettore risultante, la *funzione componente* $f_i : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ da applicare all'input:

$$\mathbf{f}(t) = (f_1(t), \dots, f_m(t))$$

Osservazione 22 (Interpretazione grafica e notazione). Si pensa spesso a $t \in I$, il *parametro*, come al tempo, e si immagina la curva descritta dal moto del punto $\mathbf{f}(t)$ quanto t corre in I . Inoltre si usa il grassetto in \mathbf{f} a ricordare che come output la funzione fornisce un vettore.

Proposizione 1.3.1 (Limite di funzione parametrica in spazio vettoriale). *Il limite della funzione a valori vettoriali (se esiste) si calcola componente per componente; in simboli:*

$$\lim_{t \rightarrow t_0} (\mathbf{f}(t)) = \left(\lim_{t \rightarrow t_0} f_1(t), \dots, \lim_{t \rightarrow t_0} f_m(t) \right)$$

Osservazione 23. Il calcolo dei limiti per funzioni a valori vettoriali non introduce quindi nuove difficoltà rispetto al calcolo dei limiti per funzioni a valori reali; inoltre l'invarianza rispetto ai casi passati fa sì che proprietà riguardanti i limiti di funzioni vettoriali si possano enunciare e dimostrare in modo analogo al caso unidimensionale (ad esempio vale il *teorema di unicità del limite*, il *limite della somma* e del *prodotto per una costante*, ecc), e inoltre vale la seguente.

Definizione 1.3.2 (Funzione $\mathbf{f} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ continua). Si dice che $\mathbf{f} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ è continua in t_0 se $\lim_{t \rightarrow t_0} \mathbf{f}(t) = \mathbf{f}(t_0)$ ossia se e solo se lo sono tutte le sue (funzioni) componenti.

Proposizione 1.3.2 (Derivata di $\mathbf{f} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$). *La derivata di una funzione a valori vettoriali, se esiste, si calcola componente per componente:*

$$\mathbf{f}'(t_0) = (f'_1(t_0), \dots, f'_m(t_0))$$

Il risultato, detto vettore derivato, è il vettore delle derivate delle funzioni componenti.

Definizione 1.3.3 (Funzione $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ di classe $C^{(1)}(I)$). Funzione \mathbf{f} derivabile in tutto I e per la quale la funzione \mathbf{f}' è continua in I . Si scrive $\mathbf{f} \in C^{(1)}(I)$

Definizione 1.3.4 (Funzione $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ di classe $C^k(I)$). Funzione \mathbf{f} derivabile $k > 1$ volte in tutto I e per la quale la k -esima derivata è continua in I .

1.3.2 Curve

1.3.2.1 Definizioni e primi esempi

Definizione 1.3.5 (Curva in \mathbb{R}^m). Funzione $\mathbf{f} : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^m$ continua.

Osservazione 24. Quindi le peculiarità della curva sono il dominio ad intervallo chiuso e la continuità della funzione in essa

Definizione 1.3.6 (Sostegno della curva \mathbf{f} (immagine)). Data una curva $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$ viene chiamato così l'insieme $\mathbf{f}([a, b]) = \{\mathbf{f}(t) : t \in [a, b]\} \subset \mathbb{R}^m$.

Osservazione 25. Si tratta dell'insieme dei punti di \mathbb{R}^m percorsi dal punto mobile

Definizione 1.3.7 (Curva di classe $C^{(k)}$). Posto $\mathbf{f} \equiv (f_1, \dots, f_m)$, con $f_i : [a, b] \rightarrow \mathbb{R}$, la curva \mathbf{f} si dice di classe $C^{(k)}$ se e solo se tutte le sue componenti scalari f_i lo sono.

Definizione 1.3.8 (Curva chiusa). Curva $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$ per la quale $\mathbf{f}(a) = \mathbf{f}(b)$.

Definizione 1.3.9 (Curva aperta). Non chiusa.

Definizione 1.3.10 (Curva semplice). Curva $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$ con $\mathbf{f}(t_1) \neq \mathbf{f}(t_2)$ per ogni $t_1, t_2 \in [a, b]$.

Osservazione 26. In altre parole una curva semplice non ripassa mai dallo stesso punto: può essere chiusa ma non deve intrecciarsi

Definizione 1.3.11 (Curva regolare). Curva $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$ che congiuntamente è:

- di classe $C^{(1)}$;
- tale che $\mathbf{f}'(t) \neq \mathbf{0}$, $\forall t \in (a, b)$
- tale che se $\mathbf{f}(a) = \mathbf{f}(b)$ (ossia chiusa), allora $\mathbf{f}'(a) = \mathbf{f}'(b) \neq \mathbf{0}$

Esempio 1.3.1 (Esempi di curve e classificazione 1: circonferenza unitaria). La curva $\mathbf{f} : [0, 2\pi] \rightarrow \mathbb{R}^2$

$$\mathbf{f}(t) = (\cos t, \sin t)$$

è regolare e il sostegno della curva è la circonferenza unitaria. La curva è semplice chiusa.

Esempio 1.3.2 (Esempi di curve e classificazione 2: (t^2, t^3)). La curva $\mathbf{f} : [-1, 1] \rightarrow \mathbb{R}^2$

$$\mathbf{f}(t) = (t^2, t^3)$$

non è regolare⁵, infatti per $t = 0$ si ha

$$\left(\frac{\partial t^2}{\partial t}(0), \frac{\partial t^3}{\partial t}(0) \right) = (0, 0)$$

La curva è semplice aperta.

Esempio 1.3.3 (Esempi di curve e classificazione 3). La curva $\mathbf{f} : [-1, 1] \rightarrow \mathbb{R}^2$ tale che

$$\mathbf{f}(t) = (t(t-1), t(t-1)(2t-1))$$

è regolare aperta, ma non semplice essendo $\mathbf{f}(0) = \mathbf{f}(1) = (0, 0)$

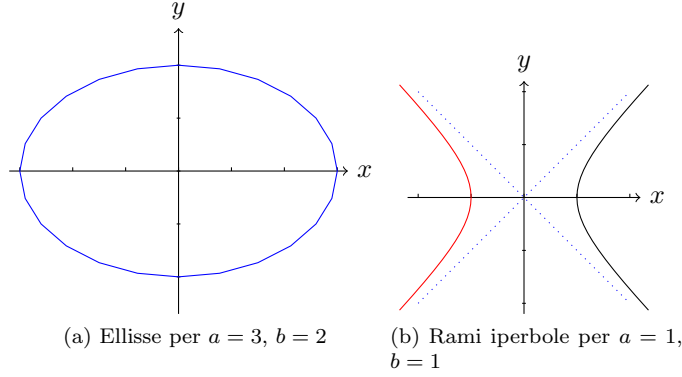


Figura 1.1: Ellisse ed iperbole

Esempio 1.3.4 (Esempi di curve e classificazione 4: ellisse). L'ellisse (figura 1.1):

$$\begin{cases} x = 3 \cos t \\ y = 2 \sin t \end{cases} \quad t \in [0, 2\pi]$$

è una curva chiusa, semplice ed è regolare, in quanto il vettore derivato $\mathbf{f}'(t) = (-3 \sin t, 2 \cos t) \neq (0, 0)$ per ogni t (quando una componente si annulla, l'altra è diversa da zero).

Esempio 1.3.5 (Esempi di curve e classificazione 5: rami di iperbole). Le equazioni

$$\begin{cases} x = a \cosh t \\ y = b \sinh t \end{cases} \quad t \in \mathbb{R}$$

con $a, b > 0$ descrivono il ramo di iperbole

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad x > 0$$

da cui l'origine del nome di *funzioni iperboliche* per $\cosh t$, $\sinh t$ (in figura 1.1 2, il ramo a destra, nero).

L'altro ramo (per $x < 0$, nella figura in rosso e a sinistra) è descritto

$$\begin{cases} x = -a \cosh t \\ y = b \sinh t \end{cases} \quad t \in \mathbb{R}$$

Ognuno dei due archi è una curva continua, aperta e direi sia regolare (di classe C^1) date le caratteristiche di \sinh e \cosh , non si ha mai che $\mathbf{f}'(t) = \mathbf{0}$, $\forall t \in \mathbb{R}$ e non si ha mai che $\mathbf{f}(a) = \mathbf{f}(b)$ che semplice.

⁵ma è regolare a tratti

Osservazione 27 (Ripasso). Il fatto che una funzione f sia biiettiva e continua non implica, in generale che l'inversa f^{-1} sia continua⁶

Definizione 1.3.12 (Omeomorfismo). Dati X, Y spazi metrici, applicazione $f : X \rightarrow Y$ continua e biiettiva, e per la quale anche f^{-1} è continua.

Definizione 1.3.13 (Curva semplice aperta). Una curva $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$ si dice *semplice aperta* se e solo se definisce un *omeomorfismo* da $[a, b]$ a $\mathbf{f}([a, b]) = \{\mathbf{f}(t) : t \in [a, b]\} \subseteq \mathbb{R}^m$.

Osservazione 28. Un ultimo risultato generale segue.

Proposizione 1.3.3 (Compatti ed omeomorfismi). *Dati X, Y spazi metrici con X compatto, se $f : X \rightarrow Y$ è continua e biiettiva, allora è un omeomorfismo (e la curva è semplice aperta).*

Osservazione 29. Poiché $[a, b]$ è un compatto, dalla precedente proposizione segue che la curva $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$, se continua, è anche semplice aperta (omeomorfismo) se e solo se è iniettiva (la suriettività si dà per buona ipotizzando la restrizione del codominio all'immagine).

1.3.2.2 Vettori tangenti

Osservazione 30. Nel seguito (nonché nella sezione sulle varietà), ci limiteremo a considerare curve regolari semplici aperte $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$.

Definizione 1.3.14 (Versore tangente). Data una curva $\mathbf{f} : [a, b] \rightarrow \mathbb{R}^m$ regolare semplice aperta e dato $t \in (a, b)$, si dice *versore tangente* alla curva nel punto $\mathbf{f}(t)$ il versore

$$\mathbf{T}(t) = \frac{\mathbf{f}'(t)}{\|\mathbf{f}'(t)\|} = \frac{(f'_1(t), \dots, f'_m(t))}{\sqrt{f'_1(t)^2 + \dots + f'_m(t)^2}}$$

Esempio 1.3.6 (Esempio di calcolo del versore tangente in un punto). L'ellisse $\mathbf{f}(t) = (3 \cos t, 2 \sin t)$, $t \in [0, 2\pi]$ ha versore tangente

$$\mathbf{T}(t) = \frac{(-3 \sin t, 2 \cos t)}{\sqrt{9 \sin^2 t + 4 \cos^2 t}}$$

Ad esempio, se $t = 0$, ossia in $\mathbf{f}(t) = (3, 0)$ si ha che il versore tangente è

$$\mathbf{T}(t) = \frac{(0, 2)}{\sqrt{4}} = (0, 1)$$

Definizione 1.3.15 (Vettore tangente alla curva nel punto $\mathbf{f}(t)$). Ogni vettore $\lambda \cdot \mathbf{T}(t)$ con $\lambda \in \mathbb{R}$, $\lambda \neq 0$

⁶Ad esempio Sia $X = [0, 2\pi[$ e $Y = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ con le metriche euclidee di \mathbb{R} e \mathbb{R}^2 , rispettivamente. L'applicazione $f : X \rightarrow Y$, $f(t) = (\cos t, \sin t)$ è evidentemente biiettiva e continua. Tuttavia la funzione inversa $f^{-1} : Y \rightarrow X$ non è continua nel punto $(1, 0) \in Y$; infatti se consideriamo una successione di punti $(y_n)_{n \in \mathbb{N}}$ della circonferenza Y che converge a $(1, 0)$ “dal semipiano di ordinata negativa” risulterà ovviamente $f^{-1}(y_n) > \pi$ per ogni $n \in \mathbb{N}$ e quindi la successione $(f^{-1}(y_n))_{n \in \mathbb{N}}$ non può convergere a $f^{-1}(1, 0) = 0$.

Definizione 1.3.16 (Retta tangente alla curva in un punto $\mathbf{f}(t)$). Dato $\lambda \in \mathbb{R}$, $\lambda \neq 0$, l'insieme di punti descritto da:

$$\mathbf{f}(t) + \lambda \mathbf{T}(t)$$

Esempio 1.3.7 (Calcolo retta tangente alla curva in un punto $\mathbf{f}(t)$: parabola). Considerando la curva (regolare, semplice aperta) $\mathbf{f} : [-1, 1] \rightarrow \mathbb{R}^2$, definita da $\mathbf{f}(t) = (t, t^2)$ (il grafico coincide con quello della parabola $y = x^2$ in $-1 \leq x \leq 1$). Sia $t = 1/2$. Il versore tangente a \mathbf{f} in t è il versore

$$\mathbf{T}(1/2) = \frac{(1, 2 \cdot \frac{1}{2})}{\sqrt{1+1}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

La retta tangente in $\mathbf{f}(1/2)$ è

$$\left\{ \left(\frac{1}{2}, \frac{1}{4} \right) + \lambda \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) : \lambda \in \mathbb{R} \right\}$$

che coincide con la retta $y = x - \frac{1}{4}$.

Esempio 1.3.8 (Calcolo retta tangente alla curva in un punto $\mathbf{f}(t)$: circonferenza unitaria). Data la curva $\mathbf{f} : [0, 2\pi] \rightarrow \mathbb{R}^2$, definita da $\mathbf{f}(t) = (\cos t, \sin t)$ (ossia la circonferenza goniometrica), per $t_0 = \frac{\pi}{2}$, la retta tangente in $\mathbf{f}(\pi/2)$ è l'insieme $\{(0, 1) + \lambda(-1, 0) : \lambda \in \mathbb{R}\}$, cioè la retta di equazioni cartesiane $y = 1$

1.4 Funzioni $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$

1.4.1 Differenziabilità

1.4.1.1 Osservazioni preliminari

Osservazione 31. In questa sezione, estenderemo le nostre conoscenze sul calcolo differenziale per funzioni $f : \mathbb{R} \rightarrow \mathbb{R}$ al caso di funzioni $f : \mathbb{R}^n \rightarrow \mathbb{R}$ in $n \geq 1$ variabili. I principali cambiamenti che intercorrono sono riportati nelle seguenti osservazioni.

Osservazione 32. Ricordiamo che in \mathbb{R} una funzione $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ con A aperto e $x \in A$ è derivabile in x se esiste ed è finito il limite

$$\lim_{t \rightarrow 0} \frac{f(x+t) - f(x)}{t}$$

e questo è detto derivata prima di f in x .

Osservazione 33. Il fatto che una funzione sia derivabile in un punto implica che sia continua nello stesso ma non vale il viceversa; possono esistere funzioni continue in un punto che non sono derivabili nello stesso (es il valore assoluto).

Osservazione 34 (Derivabilità e direzione). Se in \mathbb{R}^n si ha che $n > 1$, non ha più senso parlare di funzione derivabile in termini “assoluti”: nel caso di più variabili, il concetto di *derivabilità* deve essere sempre associato alla scelta di una *direzione*, o *versore*.

Nel caso $n = 1$ questo fenomeno non appare poiché in \mathbb{R} (dominio) abbiamo solo due direzioni, espresse dai versori $(1, 0)$ e $(-1, 0)$ e, di conseguenza, essendo l'uno l'opposto dell'altro, le due possibili definizioni di derivabilità (e, di *derivata*) di fatto coincidono, essendo la stessa a meno di un segno. (Si pensi alla pendenza della retta tangente che tende allo stesso se il punto che la definisce proviene da destra o sinistra).

Nel caso di \mathbb{R}^2 come dominio le direzioni (o versori) di derivazione sono i “raggi” della circonferenza unitaria.

Osservazione 35 (Derivabilità e differenziabilità 1). Il caso di funzioni in n variabili che ammettano derivate secondo *tutte le direzioni* non fornisce il vero equivalente della nozione di funzione derivabile in una variabile: ad esempio l'esistenza anche di tutte le derivate secondo ogni direzione in un punto *non implica* la *continuità* in tale punto (mentre nel caso di funzioni $f : \mathbb{R} \rightarrow \mathbb{R}$ la derivabilità in un punto implica la continuità nello stesso).

Osservazione 36 (Derivabilità e differenziabilità 2). Diviene perciò necessario introdurre un nuovo concetto che risulti il “vero” analogo del concetto di derivabilità nel caso di funzioni di più variabili: questo è appunto il concetto di funzione *differenziabile*, che costituirà il tema principale.

Osservazione 37 (Derivabilità e differenziabilità 3). Se $n = 1$, la tradizionale nozione di derivabilità e la nozione di differenziabilità *coincidono*: questo spiega la ragione per cui, al livello di funzioni di una variabile, non sia necessario introdurre il concetto di differenziabilità.

1.4.1.2 Derivate direzionali, derivate parziali

Definizione 1.4.1 (Versore (vettore di direzione)). Vettore $\mathbf{v} \in \mathbb{R}^n$ di norma euclidea unitaria, ossia $\|\mathbf{v}\| = 1$.

Definizione 1.4.2 (Retta passante per \mathbf{x}_0 di direzione \mathbf{v}). Dato $\mathbf{x}_0, \mathbf{v} \in \mathbb{R}^n$ con \mathbf{v} direzione, $t \in \mathbb{R}$, è l'insieme

$$r_{\mathbf{x}_0, \mathbf{v}} = \{\mathbf{x}' \in \mathbb{R}^n : \mathbf{x}' = \mathbf{x}_0 + t\mathbf{v}\}$$

Definizione 1.4.3 (Funzione derivabile in $\mathbf{x}_0 \in A \subseteq \mathbb{R}^n$ secondo la direzione \mathbf{v} e derivata direzionale). La funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto) è derivabile in $\mathbf{x}_0 \in A$ secondo la direzione $\mathbf{v} \in \mathbb{R}^n$ se e solo se esiste finito il limite

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} \quad (1.14)$$

chiamato *derivata direzionale* di f nel punto \mathbf{x}_0 secondo la direzione \mathbf{v} e indicato con $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$.

Osservazione importante 1 (Visualizzazione grafica derivata direzionale per $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ secondo \mathbf{v}). Ipotizziamo di avere f visualizzabile mediante un tendone, farne la derivata significa tagliarla dall'alto nella direzione \mathbf{v} dopodiché vedere la fetta tagliata dal lato e analizzarne la derivata nel punto di interesse.

Osservazione importante 2 (Numero di derivate direzionali). In generale, se $n > 1$, le derivate direzionali di una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (se esistono) sono infinite (tante quante possono essere i vettori di lunghezza unitaria in \mathbb{R}^n).

Definizione 1.4.4 (i -esima derivata parziale di f in \mathbf{x}_0). È la derivata direzionale di f nel punto \mathbf{x}_0 secondo la direzione \mathbf{e}_i (i -esimo vettore della base canonica): dato che $\mathbf{x}_0 + t\mathbf{e}_i = (x_1, \dots, x_i + t, \dots, x_n)$ si ha che

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{e}_i}(\mathbf{x}_0) &= \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_i) - f(\mathbf{x}_0)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(x_1, \dots, x_i + t, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{t} \\ &= \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \end{aligned}$$

Osservazione importante 3 (Metodo di calcolo i -esima derivata parziale). Perciò la derivata parziale i -esima in \mathbf{x}_0 si ottiene considerando le variabili $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ costanti (si pensi graficamente a quanto dentro parentesi in termini di due punti nello spazio, l'unica cosa che cambia è sulla dimensione i -esima), e derivando “in modo tradizionale” rispetto alla variabile x_i (per questo il nome di parziale).

Esercizio 1.4.1 (Calcolo derivate parziali per una $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$: 1). Calcolare le derivate parziali per la funzione

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

Soluzione. Le derivate parziali rispetto a x ed y saranno

$$\begin{aligned} \frac{\partial f}{\partial x} &= \sin xy + xy \cos xy + y^2 \cdot e^{xy^2} \\ \frac{\partial f}{\partial y} &= x^2 \cos xy + e^{xy^2} \cdot 2xy \end{aligned}$$

Esercizio 1.4.2 (Calcolo derivate parziali per una $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$: 1). brini is in da house. Calcolare le derivate parziali di

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

Soluzione. Allora le derivate parziali rispetto a x ed y saranno

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xy \cdot e^{x^2 y} + y^2 \\ \frac{\partial f}{\partial y} &= x^2 e^{x^2 y} + 2xy \end{aligned}$$

Osservazione 38 (Numero di derivate parziali). Le derivate parziali (se esistono) sono in numero *finito* (al massimo n), cioè la dimensione dello spazio dominio della funzione, o, equivalentemente, il numero delle variabili.

1.4.1.3 Derivate parziali valutate in un punto

Osservazione 39 (Notazioni alternative per la i -esima derivata parziale in un punto).

$$\frac{\partial f}{\partial \mathbf{e}_i}(\mathbf{x}_0) = \frac{\partial f}{\partial x_i}(\mathbf{x}_0) = \frac{\partial f(\mathbf{x}_0)}{\partial x_i} = D_i(f)(\mathbf{x}_0)$$

Proposizione 1.4.1 (Metodi di calcolo per derivate parziali valutate in un punto). *Data una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e volendo ottenere la i -esima derivata parziale si può procedere in due modi:*

1. *scrivere la funzione di una variabile che si ottiene sostituendo le variabili tenute costanti col il valore assunto nel punto, dopodiché calcolare la derivata normalmente rispetto all'incognita di interesse e infine sostituire col valore assunto nel punto;*
2. *calcolare la derivata considerando costanti le $n-1$ variabili non di interesse dopodiché effettuare la sostituzione dei valori del punto in blocco*

Il primo è da preferire. I vantaggi (ad esempio considerando $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$):

- *la funzione $f(x, y_0)$ potrebbe risultare derivabile per un certo y_0 fissato, mentre non lo è per y generico, quindi il primo metodo risulta praticabile in più casi;*
- *la funzione $f(x, y_0)$ potrebbe risultare molto semplice (ad esempio costante) e quindi banale da derivare, mentre $f(x, y)$ potrebbe non esserlo.*

Esempio 1.4.1 (Caso in cui entrambi i metodi funzionano). Sia $f(x, y) = x^2 y^3$. Vogliamo calcolare la derivata parziale rispetto a x in $(1, 2)$: abbiamo due strade:

1. *scrivere prima la funzione di una variabile che si ottiene sostituendo $y = 2$, quindi calcolando la derivata della funzione ottenuta di questa rispetto a x e valutando infine la derivata in $x = 1$:*

$$\frac{\partial f}{\partial x}(1, 2) = \frac{d}{dx}[8x^2]_{x=1} = (16x)_{x=1} = 16$$

2. *derivare genericamente rispetto a x (considerando y come una costante) e valutare sostituendo le coordinate del punto considerato:*

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial}{\partial x}(x^2 y^3) = (2xy^3)_{|(x,y)=(1,2)} = 16$$

Esempio 1.4.2 (Caso in cui solo il primo funziona). Vogliamo calcolare la derivata di $z = y\sqrt{x}$ rispetto a x nel punto $(0, 0)$. Si ha che

1. valutando parzialmente e poi calcolando la derivata si ha

$$\frac{\partial}{\partial x}(y\sqrt{x})|_{(x,y)=(0,0)} = \frac{d}{dx}(0\sqrt{x})|_{x=0} = \frac{d}{dx}(0) = 0$$

2. calcolando la derivata generica invece

$$\frac{\partial}{\partial x}(y\sqrt{x}) = \frac{y}{2\sqrt{x}}$$

ma valutando l'espressione trovata in $(0,0)$ si trova $0/0$ che non ha senso;

La derivata parziale rispetto a x nell'origine pertanto esiste e vale zero.

1.4.1.4 Funzioni differenziabili

Osservazione importante 4 (Derivate parziali in un punto e continuità). Come anticipato l'esistenza delle derivate parziali nel punto \mathbf{x}_0 non implica in generale la continuità nello stesso: una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ può avere derivate in $\mathbf{x}_0 \in A$ secondo tutte le direzioni, pur non essendo continua in \mathbf{x}_0 .

Esempio 1.4.3 (Funzione non continua in un punto \mathbf{x}_0 ma con tutte le derivate direzionali). Un esempio di $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. La funzione si presenta come un gradino (alto 1) nell'area sotto la retta $y = x^2$ e sopra l'asse delle x (esclusa)

$$f(x, y) = \begin{cases} 0 & \text{se } y \leq 0 \text{ oppure } y \geq x^2 \\ 1 & \text{se } 0 < y < x^2 \end{cases}$$

Ci interessa il punto $\mathbf{x}_0 = \mathbf{0} = (0,0)$. In corrispondenza di questo si ha che $f(\mathbf{x}_0) = 0$ (è tipo termopili tra due monti a destra e sinistra venendo da sud). Contemporaneamente

- la definizione 1.2.9 di continuità della funzione in \mathbf{x}_0 non regge poiché, non riusciamo a trovare un intorno (si pensi alla circonferenza unitaria) attorno a $\mathbf{0}$ dove per qualsiasi \mathbf{x} appartenentevi sia $d(f(\mathbf{x}), f(\mathbf{0})) < 1$. Pertanto f non è continua in $\mathbf{0}$;
- al contempo si ha che per qualsiasi direzione \mathbf{v} (si pensi al fascio di rette passante per l'origine), per $t \rightarrow 0$ $f(\mathbf{x}_0 + t\mathbf{v}) = 0$ (rimaniamo nella piana delle termopili per $t \rightarrow 0$, non saliamo su uno dei due monti se t è alto) e come detto precedentemente $f(\mathbf{x}_0) = 0$, dunque

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} = \lim_{t \rightarrow 0} \frac{0}{t} = 0$$

Ergo il limite esiste finito, quindi per qualsiasi direzione \mathbf{v} si ha che f ha derivate parziali finite (nulle).

Osservazione 40. Muoviamoci dunque verso la differenziabilità, introducendo il concetto di differenziale, che è una particolare funzione lineare.

Definizione 1.4.5 (Funzione lineare). Siano X e Y spazi vettoriali e sia $f : X \rightarrow Y$ una funzione tra X e Y . Tale funzione si dice *lineare* se:

$$f(\mathbf{x}_1 + \mathbf{x}_2) = f(\mathbf{x}_1) + f(\mathbf{x}_2) \quad \mathbf{x}_1, \mathbf{x}_2 \in X \quad (1.15)$$

$$f(\alpha \mathbf{x}) = \alpha f(\mathbf{x}) \quad \mathbf{x} \in X, \alpha \in \mathbb{C} \quad (1.16)$$

Definizione 1.4.6 (Funzionale). Dato uno spazio vettoriale X su un campo \mathbb{K} , un funzionale f è una funzione lineare del tipo $X \rightarrow \mathbb{K}$.

Esempio 1.4.4 (Esempio di funzionali). Sono funzionali le funzioni lineari del tipo $f : \mathbb{C}^3 \rightarrow \mathbb{C}$ o $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, ma non $h : \mathbb{R}^2 \rightarrow \mathbb{C}$.

Definizione 1.4.7 (Funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile in un punto \mathbf{x}_0). Una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto) si dice differenziabile in $\mathbf{x}_0 \in A$ se e solo se esiste un funzionale $L_{\mathbf{x}_0} : \mathbb{R}^n \rightarrow \mathbb{R}$ (o $df(\mathbf{x}_0)$), detto *differenziale di f nel punto \mathbf{x}_0* , tale che dato un $\mathbf{h} \in \mathbb{R}^n$

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Osservazione importante 5 (Interpretazione geometrica differenziabilità in un punto). Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile in $\mathbf{x}_0 \in A$, e sia $L_{\mathbf{x}_0}$ il suo differenziale in \mathbf{x}_0 .

Il numeratore che compare nel limite (riguardato come funzione dell'incremento vettoriale $\mathbf{h} \in \mathbb{R}^n$):

$$E_{\mathbf{x}_0}(\mathbf{h}) = f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(\mathbf{h}) = f(\mathbf{x}_0 + \mathbf{h}) - [f(\mathbf{x}_0) + L_{\mathbf{x}_0}(\mathbf{h})]$$

può essere interpretato come l'*errore* (anche esso funzione di \mathbf{h}) che si commette, nel punto $\mathbf{x}_0 + \mathbf{h}$, "approssimando" la funzione f con $f(\mathbf{x}_0) + L_{\mathbf{x}_0}(\mathbf{h})$.

Ad esempio per le funzioni $\mathbb{R} \rightarrow \mathbb{R}$ (figura 1.2) si ha che il differenziale nel punto x_0 per una variazione h è $L_x : h \rightarrow f'(x_0) \cdot h$, $h \in \mathbb{R}$ per cui $E(x_0, h)$ è lo spazio verticale tra $f(x_0 + h)$ e $f(x_0) + f'(x_0) \cdot h$ (ossia il valore assunto dalla funzione in $x + h$ e la sua approssimazione usando il differenziale).

D'altra parte il denominatore (opportunamente riscritto):

$$\|\mathbf{h}\| = \|(\mathbf{x}_0 + \mathbf{h}) - \mathbf{x}_0\|$$

fornisce la distanza del punto $\mathbf{x}_0 + \mathbf{h}$ dal punto \mathbf{x}_0 .

Dunque la condizione di differenziabilità di f in \mathbf{x}_0 può essere riscritta come segue: esiste $L_{\mathbf{x}_0} : \mathbb{R}^n \rightarrow \mathbb{R}$ lineare e $E_{\mathbf{x}_0} : \mathbb{R}^n \rightarrow \mathbb{R}$, funzione di \mathbf{h} in un intorno di \mathbf{x}_0 , tali che

$$f(\mathbf{x}_0 + \mathbf{h}) = (f(\mathbf{x}_0) + L_{\mathbf{x}_0}(\mathbf{h})) + E_{\mathbf{x}_0}(\mathbf{h})$$

con

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{E_{\mathbf{x}_0}(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

A parole: in prossimità di \mathbf{x}_0 è possibile approssimare f con il polinomio $f(\mathbf{x}_0) + L_{\mathbf{x}_0}$ commettendo un errore $E_{\mathbf{x}_0}$ (funzione di $\|\mathbf{h}\| = d(\mathbf{x}_0 + \mathbf{h}, \mathbf{x}_0)$) che, per $\mathbf{h} \rightarrow \mathbf{0}$, "tende a zero più velocemente" della norma di \mathbf{h} .

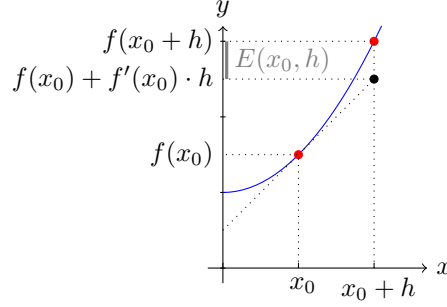


Figura 1.2: Visualizzazione grafica

Definizione 1.4.8 (Funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile in A). Una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto) si dice differenziabile in A se differenziabile in ogni punto di A .

1.4.1.5 Differenziabilità e derivate direzionali

Proposizione 1.4.2 (Differenziabilità e derivate direzionali). Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile in $\mathbf{x}_0 \in A$ (A un aperto). Allora f ha, in \mathbf{x}_0 , tutte le derivate direzionali $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$, per ogni vettore \mathbf{v} , e nello specifico:

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = L_{\mathbf{x}_0}(\mathbf{v})$$

Dimostrazione. Sia \mathbf{v} un vettore \mathbb{R}^n . Nell'ipotesi che f sia differenziabile specializziamo la condizione di differenziabilità al caso $\mathbf{h} = t\mathbf{v}$, con $t \in \mathbb{R}$, ottenendo il nostro punto di partenza, ossia cosa “vale” sotto ipotesi:

$$\lim_{t\mathbf{v} \rightarrow \mathbf{0}} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(t\mathbf{v})}{\|t\mathbf{v}\|} = 0$$

Considerato che $\|t\mathbf{v}\| = |t| \|\mathbf{v}\| = |t| \cdot 1 = |t|$ (sostituito), e che $t\mathbf{v} \rightarrow \mathbf{0} \iff t \rightarrow 0$, si giunge a:

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(t\mathbf{v})}{|t|} = 0$$

Applicando la definizione di limite, in un opportuno intorno di $t = 0$ si ha che:

$$\left| \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(t\mathbf{v})}{|t|} \right| < \varepsilon$$

Ma se questa vale, vale allora anche la seguente (dove al denominatore possiamo ignorare un valore atteso di t):

$$\left| \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(t\mathbf{v})}{t} \right| < \varepsilon$$

In altre parole la condizione dalla quale siamo partiti è *equivalente* alla condizione:

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(t\mathbf{v})}{t} = 0$$

Per linearità di $L_{\mathbf{x}_0}$, l'ultima condizione è equivalente alla condizione:

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - t \cdot L_{\mathbf{x}_0}(\mathbf{v})}{t} = 0$$

Per proprietà elementari dei limiti, ne segue (limite della somma è somma dei limiti)

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} = \lim_{t \rightarrow 0} \frac{t \cdot L_{\mathbf{x}_0}(\mathbf{v})}{t} = L_{\mathbf{x}_0}(\mathbf{v})$$

perciò

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$$

esiste finito (f è derivabile in \mathbf{x}_0 nella direzione \mathbf{v}) ed è uguale alla valutazione $L_{\mathbf{x}_0}(\mathbf{v})$ del differenziale $L_{\mathbf{x}_0}$ sul versore \mathbf{v} . In conclusione, non solo esiste la derivata di f nel punto \mathbf{x}_0 secondo la direzione \mathbf{v} , ma, di più, abbiamo:

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = L_{\mathbf{x}_0}(\mathbf{v})$$

□

Osservazione 41. In generale, come *modus operandi*: ci basta trovare una direzione per la quale la funzione non sia derivabile in un punto per dimostrare che sicuramente non è differenziabile.

Osservazione importante 6 (Funzione con tutte derivate direzionali in \mathbf{x}_0 è differenziabile?). Se $n > 1$, il “viceversa” di proposizione 1.4.2 è *falso*, ossia una funzione f può avere tutte le derivate direzionali in un punto \mathbf{x}_0 *ma non essere* differenziabile nello stesso.

Ad esempio in esempio 1.4.3 la funzione è non continua in $\mathbf{0}$, quindi per il teorema 1.4.6 è non differenziabile; tuttavia aveva tutte le derivate direzionali.

Osservazione 42. Il seguente risultato mostra, nuovamente, come derivabilità e differenziabilità coincidano in \mathbb{R}^1 .

Proposizione 1.4.3 (Derivabilità e differenziabilità in $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$). *Per le funzioni $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una funzione ammette derivata in un punto \mathbf{x}_0 (vale il grassetto) se e solo se è differenziabile in \mathbf{x}_0 e nel caso si ha che*

$$f'(\mathbf{x}_0) = L_{\mathbf{x}_0}(1)$$

Dimostrazione. Ipotizziamo che:

- f sia differenziabile in \mathbf{x}_0 ; allora per definizione esiste $L_{\mathbf{x}_0} : \mathbb{R} \rightarrow \mathbb{R}$ tale che

$$\lim_{h \rightarrow 0 \in \mathbb{R}} \frac{f(\mathbf{x}_0 + h) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(h)}{|h|} = 0$$

Effettuando lo stesso trick visto nella dimostrazione di proposizione 1.4.2 (vedere il limite entro valore assoluto per eliminare il doppio assoluto al denominatore) giungiamo a

$$\lim_{h \rightarrow 0} \frac{f(\mathbf{x}_0 + h) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(h)}{h} = 0$$

Dopodiché spezziamo il limite come fatto in precedenza

$$\lim_{h \rightarrow 0} \frac{f(\mathbf{x}_0 + h) - f(\mathbf{x}_0)}{h} = \lim_{h \rightarrow 0} \frac{L_{\mathbf{x}_0}(h)}{h} = \lim_{h \rightarrow 0} \frac{L_{\mathbf{x}_0}(1 \cdot h)}{h} \stackrel{(1)}{=} \lim_{h \rightarrow 0} \frac{\cancel{h} \cdot L_{\mathbf{x}_0}(1)}{\cancel{h}}$$

con (1) per linearità. Pertanto, essendo $L_{\mathbf{x}_0}(1)$ costante, prendendo primo e ultimo membro si conclude che

$$\exists f'(\mathbf{x}_0) = L_{\mathbf{x}_0}(1)$$

- sia derivabile; abbia derivata nel punto \mathbf{x}_0 pari a $f'(\mathbf{x}_0)$, costante. Qualora sia differenziabile, il differenziale per una variazione h sarà una cosa del genere $L_{\mathbf{x}_0}(h) = f'(\mathbf{x}_0) \cdot h$. Ci poniamo allora di verificare la definizione di differenziabilità al nostro caso. La funzione sarà differenziabile se vale:

$$\lim_{h \rightarrow 0 \in \mathbb{R}} \frac{f(\mathbf{x}_0 + h) - f(\mathbf{x}_0) - f'(\mathbf{x}_0) \cdot h}{|h|} = 0$$

Sempre mediante il solito trick del doppio valore assoluto arriviamo a

$$\lim_{h \rightarrow 0 \in \mathbb{R}} \frac{f(\mathbf{x}_0 + h) - f(\mathbf{x}_0) - f'(\mathbf{x}_0) \cdot h}{h} = 0$$

L'uguaglianza sarà vera (e la funzione differenziabile) se, applicando sempre le proprietà dei limiti, è vera la seguente, equivalente:

$$\lim_{h \rightarrow 0 \in \mathbb{R}} \frac{f(\mathbf{x}_0 + h) - f(\mathbf{x}_0)}{h} = \lim_{h \rightarrow 0 \in \mathbb{R}} \frac{f'(\mathbf{x}_0) \cdot \cancel{h}}{\cancel{h}} = f'(\mathbf{x}_0)$$

Ma l'uguaglianza è vera per definizione: il limite al primo membro è la definizione di derivata prima!

□

1.4.1.6 Gradiente, prodotto interno e valutazione del differenziale

Osservazione 43. Ricordiamo sempre che un qualsiasi vettore può essere espresso come sommatoria del prodotto delle sue componenti, ordinatamente, per gli elementi della base canonica

$$\mathbf{v} = (v_1, \dots, v_n) = \sum_{i=1}^n v_i \cdot \mathbf{e}_i$$

(\mathbf{e}_i vettore tutto nullo ad eccezione della i -esima componente, pari a 1).

Osservazione 44. Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto) differenziabile in $\mathbf{x}_0 \in A$ e $L_{\mathbf{x}_0}$ il differenziale di f nel punto \mathbf{x}_0 . Sappiamo che f ha tutte le derivate direzionali in $\mathbf{x}_0 \in A$ e quindi tutte le n derivate parziali. Possiamo allora calcolarne il gradiente.

Definizione 1.4.9 (Gradiente di f in un punto \mathbf{x}_0). Si dice *gradiente* di f in \mathbf{x}_0 il vettore avente per componenti le derivate parziali di f nel punto \mathbf{x}_0

$$\begin{aligned}\nabla f(\mathbf{x}_0) &= \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right) \in \mathbb{R}^n \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \cdot \mathbf{e}_i\end{aligned}$$

Proposizione 1.4.4 (Differenziale e gradiente). *Il differenziale di f nel punto $\mathbf{x}_0 \in A$ valutato per una variazione $\mathbf{h} \in \mathbb{R}^n$ è prodotto interno tra gradiente nel punto e variazione stessa:*

$$df(\mathbf{x}_0)(\mathbf{h}) = L_{\mathbf{x}_0}(\mathbf{h}) = \langle \nabla f(\mathbf{x}_0), \mathbf{h} \rangle$$

Dimostrazione. Si ha che:

$$L_{\mathbf{x}_0}(\mathbf{h}) = L_{\mathbf{x}_0} \left(\sum_{i=1}^n h_i \cdot \mathbf{e}_i \right) \stackrel{(1)}{=} \sum_{i=1}^n h_i \cdot L_{\mathbf{x}_0}(\mathbf{e}_i) \stackrel{(2)}{=} \sum_{i=1}^n h_i \cdot \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \stackrel{(3)}{=} \langle \nabla f(\mathbf{x}_0), \mathbf{h} \rangle$$

con (1) per linearità della funzione, (2) per proposizione 1.4.2 e (3) per la definizione di prodotto interno. \square

Osservazione 45. Pertanto il differenziale $L_{\mathbf{x}_0}$ di f in \mathbf{x}_0 per una variazione \mathbf{v} può essere visto come il funzionale lineare che associa ad ogni vettore $\mathbf{v} \in \mathbb{R}^n$ il suo prodotto interno col vettore gradiente.

Corollario 1.4.5 (Derivata direzionale e gradiente). *Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile in $\mathbf{x}_0 \in A$ (A aperto). Allora per ogni versore $\mathbf{v} \in \mathbb{R}^n$*

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle$$

Dimostrazione. Sfruttando ordinatamente 1.4.2 e 1.4.4

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = L_{\mathbf{x}_0}(\mathbf{v}) = \langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle$$

\square

Esempio 1.4.5 (Derivate parziali, gradiente in un punto, differenziale (non valutato e per variazione): funzione polinomiale). Data $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definita da

$$f(x, y, z) = 2 + x + xz + x^2z + y^3 + 3xy^2z$$

calcoliamo la valutazione del differenziale nel punto $\mathbf{x}_0 = (1, 1, 2)$ per la variazione di vettore $\mathbf{v} = (2, 1, 3)$. Le derivate parziali sono

$$\begin{cases} \frac{\partial f}{\partial x} = 1 + z + 2xz + 3y^2z \\ \frac{\partial f}{\partial y} = 3y^2 + 6xyz \\ \frac{\partial f}{\partial z} = x + x^2 + 3xy^2 \end{cases}$$

che valutate in \mathbf{x}_0 portano a $\frac{\partial f}{\partial x}(\mathbf{x}_0) = 13$, $\frac{\partial f}{\partial y}(\mathbf{x}_0) = 15$ e $\frac{\partial f}{\partial z}(\mathbf{x}_0) = 5$; ossia il gradiente di f in \mathbf{x}_0 è

$$\nabla f(\mathbf{x}_0) = \begin{bmatrix} 13 \\ 15 \\ 5 \end{bmatrix}$$

Quindi il differenziale in \mathbf{x}_0 è il funzionale lineare $df(\mathbf{x}_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ definito da

$$df(\mathbf{x}_0) = 13dx + 15dy + 5dz$$

Nel caso del vettore $\mathbf{h} = (2, 1, 3)$, si ha $dx(\mathbf{h}) = 2$, $dy(\mathbf{h}) = 1$, $dz(\mathbf{h}) = 3$, perciò

$$L_{\mathbf{x}_0}(\mathbf{h}) = df(\mathbf{x}_0)(\mathbf{h}) = 13 \cdot 2 + 15 \cdot 1 + 5 \cdot 3 = 56$$

Esempio 1.4.6 (Derivate parziali, gradiente in un punto, differenziale (non valutato e per variazione): funzione razionale). Sia $f : A \rightarrow \mathbb{R}$ con $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \neq 1\}$ definita come segue:

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

Le derivate parziali sono

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{(3y + 2x)(1 + x^2 + y^2) - (2xy + x^2) \cdot 2x}{(1 + x^2 + y^2)^2} = \frac{3y + 2x + 2xy^2 - 3x^2y + 3y^3}{1 + 2x^2 + 2y^2 + 2x^2y^2 + x^4 + y^4} \\ \frac{\partial f}{\partial y} &= \frac{3x + 3x^3 - 3xy^2 - 2x^2y}{1 + 2x^2 + 2y^2 + 2x^2y^2 + x^4 + y^4} \end{aligned}$$

Calcoliamo i gradienti di f nei punti $(0, 0)$, $(1, 2)$, $(2, 0)$:

$$\nabla f(0, 0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \nabla f(1, 2) = \begin{bmatrix} 17/18 \\ 11/36 \end{bmatrix} \quad \nabla f(2, 0) = \begin{bmatrix} 4/25 \\ 6/5 \end{bmatrix}$$

I differenziali in tali punti sono dunque i seguenti funzionali lineari:

$$\begin{aligned} df(0, 0) &= 0 \cdot dx + 0 \cdot dy \\ df(1, 2) &= \frac{17}{18}dx + \frac{11}{36}dy \\ df(2, 0) &= \frac{4}{25}dx + \frac{6}{5}dy \end{aligned}$$

Considerata la variazione $\mathbf{v} = (-1, 1)$, la valutazione del differenziale $df(1, 2)$ per \mathbf{v} è lo scalare

$$(df(1, 2))(-1, 1) = \frac{17}{18} \cdot (-1) + \frac{11}{36} \cdot 1 = -\frac{23}{36}$$

Esempio 1.4.7 (Esercizio fatto in classe). Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = 3x^2y + xy^3$. Calcolare $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$ con $\mathbf{x}_0 = (1, 1)$ e $\mathbf{v} = \{1/2, \sqrt{3}/2\}$. Si ha che

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= 6xy + y^3|_{(1,1)} = 6 + 1 = 7 \\ \frac{\partial f}{\partial y}(x, y) &= 3x^2 + 3xy^2|_{(1,1)} = 3 + 3 = 6\end{aligned}$$

Perciò $\nabla f(\mathbf{x}_0) = (7, 6)$. Il teorema poi implica che

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \langle (7, 6), (1/2, \sqrt{3}/2) \rangle = 7/2 + 3\sqrt{3}$$

1.4.1.7 Differenziabilità e continuità

Proposizione 1.4.6 (Differenziabilità in un punto \mathbf{x}_0 e continuità). *Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile in $\mathbf{x}_0 \in A$ (A aperto). Allora f è continua in \mathbf{x}_0 .*

Dimostrazione. Poiché f è differenziabile in \mathbf{x}_0 , ricordiamo che si ha

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + L_{\mathbf{x}_0}(\mathbf{h}) + E_{\mathbf{x}_0}(\mathbf{h})$$

con

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{E_{\mathbf{x}_0}(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Ora

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = L_{\mathbf{x}_0}(\mathbf{h}) + E_{\mathbf{x}_0}(\mathbf{h}) = \langle \nabla f(\mathbf{x}_0), \mathbf{h} \rangle + E_{\mathbf{x}_0}(\mathbf{h})$$

da cui sfruttando la disuguaglianza triangolare su primo e terzo membro

$$|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)| = |\langle \nabla f(\mathbf{x}_0), \mathbf{h} \rangle + E_{\mathbf{x}_0}(\mathbf{h})| \leq |\langle \nabla f(\mathbf{x}_0), \mathbf{h} \rangle| + |E_{\mathbf{x}_0}(\mathbf{h})|$$

Dalla disuguaglianza di Cauchy-Schwarz, abbiamo

$$|\langle \nabla f(\mathbf{x}_0), \mathbf{h} \rangle| \leq \|\nabla f(\mathbf{x}_0)\| \cdot \|\mathbf{h}\|$$

ne segue, sfruttando le ultime due, che

$$|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)| \leq \|\nabla f(\mathbf{x}_0)\| \cdot \|\mathbf{h}\| + |E_{\mathbf{x}_0}(\mathbf{h})|$$

Ovviamente $|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)| \geq 0$, ma si ha anche che $|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)|$ è maggiorata dalla somma di due funzioni (ossia $\|\nabla f(\mathbf{x}_0)\| \cdot \|\mathbf{h}\|$ e $|E_{\mathbf{x}_0}(\mathbf{h})|$) che tendono entrambe a 0 per $\mathbf{h} \rightarrow \mathbf{0}$ in quanto:

- il caso di $\|\nabla f(\mathbf{x}_0)\| \cdot \|\mathbf{h}\|$ è ovvio, per definizione
- per $|E_{\mathbf{x}_0}(\mathbf{h})|$ si ragiona come segue:

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{E_{\mathbf{x}_0}(\mathbf{h})}{\|\mathbf{h}\|} = 0 \implies \lim_{\mathbf{h} \rightarrow \mathbf{0}} E_{\mathbf{x}_0}(\mathbf{h}) = 0 \iff \lim_{\mathbf{h} \rightarrow \mathbf{0}} |E_{\mathbf{x}_0}(\mathbf{h})| = 0$$

Pertanto $|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)| \rightarrow 0$ per $\mathbf{h} \rightarrow \mathbf{0}$, ossia considerando esempio 1.2.25 f è continua in \mathbf{x}_0 . \square

1.4.1.8 Il teorema del differenziale totale

Osservazione 46. Ci interessa fornire una condizione sufficiente per sancire la differenziabilità di una funzione in un punto.

Teorema 1.4.7 (Teorema del differenziale totale). *Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, (A aperto). Supponiamo che f abbia tutte le derivate parziali $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ in un intorno $I(\mathbf{x}_0, \delta)$ del punto $\mathbf{x}_0 \in A$ e che queste siano continue in \mathbf{x}_0 . Allora f è differenziabile in \mathbf{x}_0 .*

Osservazione 47. Alcune applicazioni riguardano l'individuare le funzioni differenziabili. Vale che:

- una funzione polinomiale da \mathbb{R}^n a \mathbb{R} (ossia definita utilizzando solamente somma algebrica e prodotto) è differenziabile in ogni punto di \mathbb{R}^n
- una funzione razionale (rapporto tra funzioni polinomiali) con denominatore che non si annulla per qualsiasi valore del dominio è differenziabile nello stesso

Esempio 1.4.8 (Controesempio del teorema del differenziale totale). Esempio di funzione non continua che dimostriamo esser non differenziabile.

Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita come segue (una pianura a 0 separata dalla croce formata dagli assi a 1):

$$f(x, y) = \begin{cases} 1 & \text{se } xy = 0 \\ 0 & \text{se } xy \neq 0 \end{cases}$$

Ci interessa il punto $\mathbf{x}_0 = \mathbf{0} = (0, 0)$, in corrispondenza del quale $f(\mathbf{x}_0) = 1$. Ora chiaramente f non è continua in \mathbf{x}_0 dato che non possiamo rendere piccola a piacere scegliendo un opportuno intorno la distanza tra il valore assunto dalla funzione in $\mathbf{0}$ e nei punti della circonferenza/intorno: per la maggior parte degli elementi tale differenza varrà 1 (differenza tra la croce e la piana sottostante). Di questa funzione esistono alcune derivate direzionali, non tutte, per cui in base alla proposizione 1.4.2 non può essere differenziabile. Ad esempio:

- sì ok si hanno le derivate direzionali sugli assi, es avendo $\mathbf{e}_1 = (1, 0)^T$

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}_1) - f(\mathbf{x}_0)}{t} = \frac{1 - 1}{t} = 0/t = 0$$

e la stessa cosa avviene per $\mathbf{e}_2 = (0, 1)^T$

- tuttavia se adottiamo come vettore una direzione a 45 gradi, ossia $\mathbf{v} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ si ha che

$$\begin{aligned} \lim_{t \rightarrow 0^+} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} &= \frac{0 - 1}{t} = -1/t = -\infty \\ \lim_{t \rightarrow 0^-} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} &= \frac{0 - 1}{t} = -1/t = +\infty \end{aligned}$$

I due limiti non coincidono ergo il limite finito non esiste

1.4.1.9 Operazioni tra funzioni differenziabili in un punto

Osservazione 48. Qui sostituiamo, per comodità di notazione

$$L_{\mathbf{x}_0} = df(\mathbf{x}_0)$$

(per lasciare il riferimento alla funzione che ci serve nel prosieguo)

Proposizione 1.4.8. *Siano $f, g : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabili in $\mathbf{x}_0 \in A$ (A aperto). Allora :*

- $f + g$ è differenziabile in \mathbf{x}_0 e si ha:

$$d(f + g)(\mathbf{x}_0) = df(\mathbf{x}_0) + dg(\mathbf{x}_0)$$

- $f \cdot g$ è differenziabile in \mathbf{x}_0 e si ha:

$$d(f \cdot g)(\mathbf{x}_0) = \underbrace{df(\mathbf{x}_0)}_{f. \text{ lin}} \cdot \underbrace{g(\mathbf{x}_0)}_{\text{cost.}} + \underbrace{dg(\mathbf{x}_0)}_{f. \text{ lin}} \cdot \underbrace{f(\mathbf{x}_0)}_{\text{cost.}}$$

- per ogni $\lambda \in \mathbb{R}$, λf è differenziabile in \mathbf{x}_0 e si ha:

$$d(\lambda f)(\mathbf{x}_0) = \lambda \cdot df(\mathbf{x}_0)$$

Osservazione 49. Si tratta di regole speculari a quelle di derivate incontrate in analisi 1. (volendo vedi lez 091 per la roba $\mathbb{R} \rightarrow \mathbb{R}$ ma anche no)

1.4.1.10 Differenziali come combinazione lineare

Osservazione 50. Nell'ipotesi di avere una funzione differenziabile, si poniamo qui l'obiettivo di scrivere un differenziale in modo "intrinseco". Sappiamo già che

$$df(\mathbf{x}_0)(\mathbf{v}) = \langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \cdot v_i$$

Definizione 1.4.10 (Spazio duale di \mathbb{R}^n). L'insieme dei funzionali lineari aventi dominio \mathbb{R}^n , ossia

$$(\mathbb{R}^n)^* = \{\varphi : \mathbb{R}^n \rightarrow \mathbb{R} : \varphi \text{ è lineare}\}$$

munito dell'operazione di somma tra funzioni che gode di questa definizione:

$$(\varphi_1 + \varphi_2)(\mathbf{v}) = \varphi_1(\mathbf{v}) + \varphi_2(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbb{R}^n$$

e di moltiplicazione esterna per scalare:

$$(\lambda \varphi)(\mathbf{v}) = \lambda \varphi(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbb{R}^n, \quad \forall \lambda \in \mathbb{R}$$

Osservazione 51 ($(\mathbb{R}^n)^*$ come spazio vettoriale). $(\mathbb{R}^n)^*$ è uno spazio vettoriale (si può verificare che ne rispetta tutti i requisiti visti in algebra lineare). Per la cronaca il vettore nullo di questo spazio è la funzione $\mathbf{0}(\mathbf{v}) = 0$, per qualsiasi $\mathbf{v} \in \mathbb{R}^n$.

Osservazione 52 (Idea su differenziali come combinazione lineare). I differenziali sono funzionali lineari, e, quindi, elementi dello spazio duale $(\mathbb{R}^n)^*$; pertanto se riusciamo a costruire una “base canonica” per lo spazio duale $(\mathbb{R}^n)^*$, avremo una “scrittura canonica” anche per i differenziali.

Definizione 1.4.11 (Funzionale coordinata i -esimo). Dato un $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$ è il funzionale $dx_i : \mathbb{R}^n \rightarrow \mathbb{R}$, definito da

$$dx_i(\mathbf{v}) = v_i$$

Osservazione 53. Ossia un funzionale coordinata dx_i applicato ad un generico vettore, restituisce la sua i -esima componente.

Osservazione 54. Chiaramente se li applichiamo ad elementi della base canonica di \mathbb{R}^n indicata con $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ si ha che

$$dx_i(\mathbf{e}_j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Teorema 1.4.9 (“Base canonica” dello spazio duale $(\mathbb{R}^n)^*$). *L’insieme dei funzionali coordinata*

$$\{dx_1, \dots, dx_n\}$$

è base per lo spazio duale $(\mathbb{R}^n)^$, che pertanto è uno spazio vettoriale di dimensione finita n .*

Dimostrazione. Dobbiamo provare che $\{dx_1, \dots, dx_n\}$ è un sistema di generatori e che è un insieme linearmente indipendente nello spazio vettoriale $(\mathbb{R}^n)^*$:

- per l’insieme di generatori: sia $\varphi \in (\mathbb{R}^n)^*$, e sia $\mathbf{v} = (v_1, \dots, v_n) = \sum_{i=1}^n v_i \mathbf{e}_i \in \mathbb{R}^n$. Allora

$$\varphi(\mathbf{v}) = \varphi\left(\sum_{i=1}^n v_i \mathbf{e}_i\right) = \sum_{i=1}^n v_i \varphi(\mathbf{e}_i) = \sum_{i=1}^n dx_i(\mathbf{v}) \varphi(\mathbf{e}_i) \quad \forall \mathbf{v} \in \mathbb{R}^n$$

Questa famiglia (infinita, vale per qualsiasi vettore $\mathbf{v} \in \mathbb{R}^n$) di identità tra valutazioni implica la seguente identità nello spazio duale $(\mathbb{R}^n)^*$:

$$\varphi = \sum_{i=1}^n dx_i \cdot \varphi(\mathbf{e}_i)$$

Ossia la funzione (valutata in qualsiasi punto \mathbf{v}) è effettivamente una combinazione lineare che fa uso delle funzioni dx_i , pesandole per la funzione applicata agli elementi \mathbf{e}_i ;

- per l'indipendenza lineare: dato un generico vettore $\mathbf{v} = (v_1, \dots, v_n)$ dobbiamo provare che una combinazione lineare degli elementi della base (i funzionali coordinata applicati al vettore stesso) restituisce il vettore nullo (ossia la funzione che restituisce sempre zero) se e solo se i coefficienti impiegati nella combinazione sono tutti nulli. Ossia che valga la condizione:

$$\sum_{i=1}^n c_i \cdot dx_i(\mathbf{v}) = 0 \iff c_1 = \dots = c_n = 0$$

ovvero

$$c_1 v_1 + \dots + c_n v_n = 0 \iff c_1 = \dots = c_n = 0$$

Se ipotizziamo che $c_1 = \dots = c_n = 0$, allora è algebricamente immediato provare che $c_1 v_1 + \dots + c_n v_n = 0$.

Viceversa, ipotizzando per che un generico vettore $\mathbf{v} = (v_1, \dots, v_n)$ sia $\sum_{i=1}^n c_i dx_i(\mathbf{v}) = 0$, allora questa deve valere anche se sostituiamo $\mathbf{e}_1, \dots, \mathbf{e}_n$ al posto di \mathbf{v} . Ora

$$\sum_{i=1}^n c_i dx_i(\mathbf{e}_1) = c_1$$

$$\sum_{i=1}^n c_i dx_i(\mathbf{e}_2) = c_2$$

...

$$\sum_{i=1}^n c_i dx_i(\mathbf{e}_n) = c_n$$

pertanto deve essere $c_1 = \dots = c_n = 0$.

□

Corollario 1.4.10 (Differenziale come combinazione lineare funzionali coordinata). *Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto), differenziabile in $\mathbf{x}_0 \in A$. Denotiamo con $L_{\mathbf{x}_0}$ il differenziale di f in $\mathbf{x}_0 \in A$.*

Allora il differenziale per una variazione generica \mathbf{h} si scrive (in modo unico nello spazio duale $(\mathbb{R}^n)^$) come combinazione lineare di funzionali coordinata dx_i (applicati a \mathbf{h}) e coefficienti delle derivate parziali:*

$$L_{\mathbf{x}_0}(\mathbf{h}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \cdot dx_i(\mathbf{h}) \in (\mathbb{R}^n)^* \quad (1.17)$$

Dimostrazione. Segue immediatamente dalla equivalenza tra le identità (puntuali, sulle valutazioni):

$$L_{\mathbf{x}_0}(\mathbf{h}) = \langle \nabla f(\mathbf{x}_0), \mathbf{h} \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \cdot h_i = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \cdot dx_i(\mathbf{h}), \quad \forall \mathbf{h} \in \mathbb{R}^n$$

□

1.4.1.11 Derivate successive

Definizione 1.4.12 (Derivate parziali successive, pure e miste). Una derivata parziale rispetto ad una variabile può essere derivata ulteriormente; quello che si ottiene è una *derivata successiva*, che può essere *pura* (se si deriva rispetto alla medesima variabile per la quale si è derivato in passato) o *mista* (se si cambia variabile).

Definizione 1.4.13 (Ordine di una derivata successiva). Il numero di volte per le quali complessivamente (considerando tutte le variabili di derivazione) si è derivato per giungere lì.

Osservazione 55. A livello di notazione si indica con

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

la derivata (di secondo ordine) della funzione f , ordinatamente prima rispetto alla variabile x e poi rispetto alla variabile y .

Esempio 1.4.9. Dato

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

Si ha

$$\begin{aligned} \frac{\partial f}{\partial x} &= 4xy + 3 & \frac{\partial f}{\partial y} &= 2x^2 \\ \frac{\partial^2 f}{\partial y \partial x} &= 4x & \frac{\partial^2 f}{\partial x \partial y} &= 4x \end{aligned}$$

Notiamo che, in questo caso:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$$

Osservazione 56. In generale non necessariamente scambiando la sequenza della variabile di derivazione si ottiene il medesimo risultato e può benissimo essere che

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \neq \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right)$$

Esempio 1.4.10. Consideriamo la funzione

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = xy + |y|$$

Si noti che:

- la derivata parziale $\frac{\partial f}{\partial x} = y$ è definita su tutti i punti del dominio \mathbb{R}^2 , è ovunque continua (come funzione delle variabili x e y) ed è *derivabile* rispetto alla variabile y . Più precisamente, si ha:

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = 1$$

in ogni punto del dominio \mathbb{R}^2

- la derivata parziale $\frac{\partial f}{\partial y}$ rispetto alla variabile y *non esiste*⁷ nei punti per cui $y = 0$; quindi, in questi punti non può esistere la derivata successiva $\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right)$. Ergo le due derivate non coincidono.

Osservazione 57. Tuttavia, sotto opportune ipotesi (analoghe a quelle del Teorema del Differenziale Totale), sussiste il seguente fondamentale risultato

Teorema 1.4.11 (di Schwartz, o, sulla scambiabilità della sequenza di derivazione parziale). *Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto) e $\mathbf{x}_0 = (x_1, \dots, x_n) \in A$. Se, per $i, j = 1, 2, \dots, n$, le derivate successive*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}, \quad \frac{\partial^2 f}{\partial x_j \partial x_i}$$

esistono in un intorno di \mathbf{x}_0 e sono continue in \mathbf{x}_0 allora in prossimità di \mathbf{x}_0 coincidono

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Osservazione 58. Se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione polinomiale, tutte le derivate parziali successive esistono e sono polinomiali, quindi continue. Perciò si può sempre scambiare la sequenza di derivazione.

Esempio 1.4.11. Sia $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definita da

$$f(x, y, z) = x^2 y z + x^2 z^2 + 2 x y z^2$$

Verifichiamo che effettivamente:

$$\frac{\partial^3 f}{\partial x \partial y \partial z} = \frac{\partial^3 f}{\partial z \partial x \partial y}$$

Si ha che

$$\begin{aligned} \frac{\partial^3 f}{\partial x \partial y \partial z} &= \frac{\partial^2}{\partial x \partial y} (x^2 y + 2x^2 z + 4xyz) = \frac{\partial}{\partial x} (x^2 + 4xz) = 2x + 4z \\ \frac{\partial^3 f}{\partial z \partial x \partial y} &= \frac{\partial^2}{\partial z \partial x} (x^2 z + 2xz^2) = \frac{\partial}{\partial z} (2xz + 2z^2) = 2x + 4z \end{aligned}$$

e come preventivato dall'osservazione di cui sopra le due derivate coincidono.

Definizione 1.4.14 (Funzione di classe $C^k(A)$). Funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, con A aperto, le cui derivate parziali fino all'ordine k esistono e sono continue in A

Osservazione 59. Questa definizione torna utile nel caso di funzioni:

- funzioni $C^1(A)$ che sono differenziabili in A per il teorema del differenziale totale

⁷Sarebbe $x + \operatorname{sgn}(y)$, ma si pensi alla funzione valore atteso che presenta uno spigolo per $y = 0$

- funzioni $C^2(A)$ per sancire la simmetricità della matrice Hessiana. Se le derivate parziali sino all'ordine 2 esistono e sono continue si può applicare il teorema di Schwartz.

Definizione 1.4.15 (Funzione di classe C^∞). Se per ogni $k \in \mathbb{N}$, f è di classe C^k in A , diciamo che f è di classe C^∞ in A

1.4.1.12 Derivazione di funzioni composte

Osservazione 60. Nel caso di funzioni $f, g : \mathbb{R} \rightarrow \mathbb{R}$ si è visto come la funzione $h(x) = f \circ g = f(g(x))$ abbia derivata

$$h'(x) = f'(g(x)) \cdot g'(x)$$

Ora desideriamo generalizzare alla composizione di una funzione $\mathbf{r} : \mathbb{R} \rightarrow \mathbb{R}^n$ e $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Vogliamo determinare derivabilità e derivata di una funzione composta:

$$g : \mathbb{R} \rightarrow \mathbb{R}, \quad g = f \circ \mathbf{r}$$

Proposizione 1.4.12 (Derivazione di funzioni composte). *Sia:*

- una prima funzione $\mathbf{r} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ (I aperto) definita come:

$$\mathbf{r}(t) = (r_1(t), \dots, r_n(t))$$

con $r_i : \mathbb{R} \rightarrow \mathbb{R}$;

- una seconda funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto);
- sia l'immagine di \mathbf{r} contenuta in A

$$\mathbf{r}(I) = \{(r_1(t), \dots, r_n(t)) : t \in I\} \subseteq A$$

In tal modo è possibile definire la funzione composta $g = f \circ \mathbf{r} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ come:

$$g(t) = f(\mathbf{r}(t)) = f(r_1(t), \dots, r_n(t))$$

- sia $t_0 \in I$ tale che:
 - $r_i(t)$ è derivabile in t_0 , per ogni $i = 1, \dots, n$
 - f è differenziabile in $\mathbf{r}(t_0) \in A$.

Allora la funzione composta $g = f \circ \mathbf{r} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ è derivabile in t_0 e risulta:

$$\begin{aligned} g'(t_0) &= \langle \nabla f(\mathbf{r}(t_0)), \mathbf{r}'(t_0) \rangle \\ &= \langle \nabla f(\mathbf{r}(t_0)), (r'_1(t_0), \dots, r'_n(t_0)) \rangle \\ &= \sum_{i=1}^n \frac{\partial f(\mathbf{r}(t_0))}{\partial x_i} \cdot r'_i(t_0) \end{aligned} \tag{1.18}$$

Dimostrazione. Brini (appunti extended) pag 74 pdf □

Osservazione 61. Si nota che la formula, soprattutto quando espressa come $\nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$ non è molto dissimile da quella di due funzioni $\mathbb{R} \rightarrow \mathbb{R}$, ossia $f'(g(x)) \cdot g'(x)$

Esempio 1.4.12. Siano $\mathbf{r} : \mathbb{R} \rightarrow \mathbb{R}^2$ e $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definite come

$$\begin{aligned} \mathbf{r}(t) &= (r_1(t), r_2(t)) = (t, t^2) \\ f(x, y) &= x^2 + y^3 \end{aligned}$$

Allora

$$g(t) = f(\mathbf{r}(t)) = f(r_1(t), r_2(t)) = f(t, t^2) = t^2 + t^6, \quad t \in \mathbb{R}$$

Ovviamente si ha

$$g'(t) = 2t + 6t^5$$

Applicando (a titolo di esempio) la formula 1.18, si giunge al medesimo risultato. Alcuni elementi necessari:

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 3y^2, \quad r'_1(t) = 1, \quad r'_2(t) = 2t$$

Inoltre valutando le derivate parziali nel punto $\mathbf{r}(t)$ (ossia sostituendo $x = t$ e $y = t^2$ nelle derivate parziali precedentemente ottenute) si ha:

$$\frac{\partial f}{\partial x}(\mathbf{r}(t)) = 2t, \quad \frac{\partial f}{\partial y}(\mathbf{r}(t)) = 3(t^2)^2 = 3t^4$$

Risulta quindi:

$$\begin{aligned} g'(t) &= \nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = \frac{\partial f}{\partial x}(\mathbf{r}(t)) \cdot r'_1(t) + \frac{\partial f}{\partial y}(\mathbf{r}(t)) \cdot r'_2(t) \\ &= 2t \cdot 1 + 3t^4 \cdot 2t = 2t + 6t^5 \end{aligned}$$

1.4.2 Ottimizzazione libera

1.4.2.1 Massimi/minimi relativi

Osservazione 62. Sia funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e $\mathbf{x}_0 \in A$.

Definizione 1.4.16 (Massimo relativo). \mathbf{x}_0 è punto di massimo relativo per f se e solo se esiste un intorno $I(\mathbf{x}_0, \delta) \subseteq \mathbb{R}^n$ tale che

$$f(\mathbf{x}_0) \geq f(\mathbf{x}), \quad \forall \mathbf{x} \in I(\mathbf{x}_0, \delta) \cap A$$

Definizione 1.4.17 (Minimo relativo). \mathbf{x}_0 è punto di minimo relativo per f se e solo se esiste un intorno $I(\mathbf{x}_0, \delta) \subseteq \mathbb{R}^n$ tale che

$$f(\mathbf{x}_0) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in I(\mathbf{x}_0, \delta) \cap A$$

1.4.2.2 Setup di ottimizzazione

Osservazione importante 7 (Setup ottimizzazione). Consideriamo congiuntamente le funzioni:

1. \mathbf{r} una funzione del tipo $\mathbf{r}(t) : (-\delta, \delta) \subset \mathbb{R} \rightarrow \mathbb{R}^n$, tale che, dato un punto $\mathbf{x}_0 \in A \subseteq \mathbb{R}^n$ e un versore direzione \mathbf{v} ($\|\mathbf{v}\| = 1$), restituisce

$$\mathbf{r}(t) = \mathbf{x}_0 + t\mathbf{v} = (r_1(t), \dots, r_n(t)) = (x_1 + tv_1, \dots, x_n + tv_n)$$

Ipotizzando che l'immagine $\mathbf{r}((-\delta, \delta)) \subseteq A$, la funzione \mathbf{r} pertanto descrive, dato un $t \in (-\delta, \delta)$, un punto che si muove in A tra $\mathbf{x}_0 - \delta\mathbf{v}$ e $\mathbf{x}_0 + \delta\mathbf{v}$, con traiettoria centrata in \mathbf{x}_0 .

Si noti che $r_i(t)$ è derivabile rispetto a t per ogni $i = 1, \dots, n$ e si ha:

$$\begin{aligned} r'_i(t) &= (x_i + tv_i)' = v_i \\ \mathbf{r}'(t) &= (v_1, \dots, v_n) = \mathbf{v} \end{aligned}$$

2. una seconda funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (con A aperto e lo stesso A della funzione di cui al punto precedente), di nostro interesse (per la quale vogliamo ricercare massimi/minimi etc)

Possiamo allora definire la funzione composta $F_{\mathbf{x}_0, \mathbf{v}}(t) = f \circ \mathbf{r} : (-\delta, \delta) \subset \mathbb{R} \rightarrow \mathbb{R}$ come

$$\begin{aligned} F_{\mathbf{x}_0, \mathbf{v}}(t) &= (f \circ \mathbf{r})(t) = f(\mathbf{r}(t)) = f(\mathbf{x}_0 + t\mathbf{v}) \\ &= f(x_1 + tv_1, \dots, x_n + tv_n) \end{aligned}$$

Si pensi al caso di una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ nello spazio tridimensionale: questa funzione composta F (che è del tipo $\mathbb{R} \rightarrow \mathbb{R}$ classico) descrive l'andamento della funzione f guardandone la fetta (tagliata dall'alto, ma guardata di lato) descritto dal punto mobile $\mathbf{x}_0 + t\mathbf{v}$ nell'intorno di \mathbf{x}_0 (ricordando che il punto si muove tra $\mathbf{x}_0 - \delta\mathbf{v}$ e $\mathbf{x}_0 + \delta\mathbf{v}$).

Osservazione 63. Aggiungendo ipotesi su f a questo setup si permette l'ottimizzazione della funzione f piccolo, ossia la ricerca dei punti di massimo/minimo etc.

1.4.2.3 Identificazione dei punti critici di f : derivata prima di F

Osservazione importante 8 (Idea alla base dell'identificazione di punti critici). A livello di intuizione grafica, un punto \mathbf{x}_0 sarà *critico* per f (massimo, minimo, flesso) se critico anche per la fetta descritta da $F_{\mathbf{x}_0, \mathbf{v}}$ (nello specifico se 0 è critico per quest'ultima dato che con $t = 0$ ci troviamo in \mathbf{x}_0); ci interessa pertanto trovare la derivata prima di $F_{\mathbf{x}_0, \mathbf{v}}$ (la pendenza della tangente sulla fetta considerata) e trovare per quali valori si annulla (per verificare per \mathbf{x}_0 dobbiamo verificare se $F'_{\mathbf{x}_0, \mathbf{v}}(0) = 0$).

Poi, affinché un punto \mathbf{x}_0 sia critico per f , 0 deve esserlo per tutte le F costruibili facendo variare il vettore direzione \mathbf{v} . In altre parole, ad esempio, \mathbf{x}_0 è massimo per f solo se lo è per tutte le sue fette costruibili/possibili passanti per \mathbf{x}_0 .

Osservazione 64. Si pone dunque il calcolo della derivata della funzione composta $F_{\mathbf{x}_0, \mathbf{v}}$, il metodo è presentato in 1.4.12; il nostro setup descritto nella precedente sezione soddisfa tutti i requirements della proposizione ad eccezione dell'ipotesi di differenziabilità di f . Qui ipotizziamo che f sia differenziabile in tutto A per semplicità.

Proposizione 1.4.13 (Derivata prima di $F_{\mathbf{x}_0, \mathbf{v}}$). *Sia f differenziabile in A : la derivata prima di $F_{\mathbf{x}_0, \mathbf{v}}$ secondo t è*

$$F'_{\mathbf{x}_0, \mathbf{v}}(t) = \langle \nabla f(\mathbf{x}_0 + t\mathbf{v}), \mathbf{v} \rangle$$

Dimostrazione. Se f è differenziabile, la derivata prima di F si ottiene con l'applicazione della regola di derivazione delle funzioni composte (proposizione 1.4.12) come segue

$$\begin{aligned} F'_{\mathbf{x}_0, \mathbf{v}}(t) &= \langle \nabla f(\mathbf{r}(t)), \mathbf{r}'(t) \rangle = \sum_{i=1}^n \frac{\partial f(\mathbf{r}(t))}{\partial x_i} \cdot r'_i(t) \\ &= \sum_{i=1}^n \frac{\partial f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i} v_i \\ &= \langle \nabla f(\mathbf{x}_0 + t\mathbf{v}), \mathbf{v} \rangle \end{aligned} \tag{1.19}$$

□

Osservazione 65. Per verificare la differenziabilità di f possiamo applicare il teorema del differenziale totale e poi procedere a ottimizzazione.

Osservazione 66. Nel prossimo risultato richiediamo che la $F' = 0$ per tutti i possibili \mathbf{v} .

Proposizione 1.4.14 (Condizione necessaria punti critici). *Siano $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto) ed $\mathbf{x}_0 \in A$ punto di massimo (minimo) relativo per essa. Se f è differenziabile in \mathbf{x}_0 , allora il differenziale di f in \mathbf{x}_0 è nullo per qualsiasi $\mathbf{v} \in \mathbb{R}^n$:*

$$\langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle = 0, \quad \forall \mathbf{v} \in \mathbb{R}^n$$

Dimostrazione. Ipotizziamo \mathbf{x}_0 come punto di massimo e \mathbf{v} una direzione. La funzione (di una variabile) $F_{\mathbf{x}_0, \mathbf{v}}(t) = f(\mathbf{x}_0 + t\mathbf{v})$ ha per definizione massimo in $t = 0$. Allora, guardando alla funzione F , dato che il massimo è in 0 deve essere $F'_{\mathbf{x}_0, \mathbf{v}}(0) = 0$. Applicando la proposizione precedente, si deve avere

$$F'_{\mathbf{x}_0, \mathbf{v}}(0) = \langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle = 0$$

Quindi per ogni direzione $\mathbf{v} \in \mathbb{R}^n$ si ha $\langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle = 0$ ossia implica che il differenziale è identicamente nullo. □

Definizione 1.4.18 (Punti critici della funzione f). I punti $\mathbf{x}_0 \in A$ del dominio di una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ per cui il differenziale risulta identicamente nullo (ossia per qualsiasi variazione $\mathbf{v} \in \mathbb{R}^n$)

Proposizione 1.4.15 (Il teorema di Luca). *Il differenziale risulta identicamente nullo (per qualsiasi direzione \mathbf{v}) se e solo se il suo gradiente è il vettore nullo, ossia $\nabla f(\mathbf{x}_0) = \mathbf{0}$:*

$$\langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle = 0 \iff \nabla f(\mathbf{x}_0) = \mathbf{0}$$

Osservazione 67. Il precedente risultato - indiscussa pietra miliare nella storia dell'analisi - porta alla seguente considerazione.

Osservazione importante 9 (Procedura ottimizzazione). Per cercare i punti di massimo o minimo locale nei quali una funzione è derivabile occorre:

1. determinare tutti i punti critici, ossia i punti $\mathbf{x}_0 = (x_1, \dots, x_n)$ che risolvono il sistema di n equazioni in n incognite

$$\begin{cases} \frac{\partial f(\mathbf{x}_0)}{\partial x_1} = 0 \\ \dots \\ \frac{\partial f(\mathbf{x}_0)}{\partial x_n} = 0 \end{cases} \quad (1.20)$$

2. studiare la natura dei punti critici, ovvero stabilire per ciascuno di essi se è punto di minimo, massimo o nessuna delle due cose (es sella). Un modo per farlo (analogamente al caso univariato) è di guardare alla derivata seconda di f , che come si vedrà in seguito a 1.21, nel generico punto \mathbf{x}_0 è $\mathbf{v}^T \mathbf{H}_{f(\mathbf{x}_0)} \mathbf{v}$.

1.4.2.4 Matrici Hessiane

Definizione 1.4.19 (Matrice Hessiana della funzione f in \mathbf{x}_0). Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto) di classe $C^{(2)}(A)$. Fissato un punto $\mathbf{x}_0 \in A$ viene chiamata così la matrice quadrata $n \times n$:

$$\mathbf{H}_{f(\mathbf{x}_0)} = \left(\frac{\partial^2 f(\mathbf{x}_0)}{\partial x_i \partial x_j} \right)_{i,j=1,\dots,n} = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_1^2} & \dots & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_2 \partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_2 \partial x_n} \\ \dots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_n^2} \end{bmatrix}$$

Osservazione importante 10 (Calcolabilità e simmetria dell'Hessiana). Di base l'Hessiana può essere definita anche solo se *esistono* le derivate parziali seconde; se poi queste sono *continue* (ossia f è di classe $C^2(A)$, come di fatto si è assunto qui) allora vale il teorema di Schwartz e la matrice Hessiana è *simmetrica*.

Esempio 1.4.13 (Esempio di calcolo di una Hessiana in un punto). Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

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

Per calcolare l'Hessiana in $\mathbf{x}_0 = (-1, 3)$:

$$H_f(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x \partial x}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{x}_0) \\ \frac{\partial^2 f}{\partial y \partial x}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial y \partial y}(\mathbf{x}_0) \end{bmatrix} = \begin{bmatrix} 4y & 4x \\ 4x & -18y \end{bmatrix}_{x=-1, y=3} \stackrel{(1)}{=} \begin{bmatrix} 12 & -4 \\ -4 & -54 \end{bmatrix}$$

dove in (1) abbiamo valutato la matrice ottenuta dalle derivazioni in $\mathbf{x}_0 = (-1, 3)$.

1.4.2.5 Classificazione dei punti critici

Osservazione importante 11. Sia ora \mathbf{x}_0 punto critico per f e sia massimo (risp. minimo). Allora deve essere che 0 sia massimo (rispettivamente minimo) per $F_{\mathbf{x}_0, \mathbf{v}}$ ossia valga

$$\begin{aligned} F''_{\mathbf{x}_0, \mathbf{v}}(0) &\leq 0 && \text{se massimo} \\ F''_{\mathbf{x}_0, \mathbf{v}}(0) &\geq 0 && \text{se minimo} \end{aligned}$$

Si pone dunque il problema ora di calcolare la derivata seconda.

Osservazione 68. Continuiamo a effettuare ipotesi su f . Sia ora di classe $C^{(2)}(A)$ per cui vale il teorema di Schwartz e la matrice Hessiana ($n \times n$) è simmetrica. Da algebra lineare dal fatto che sia simmetrica deriva che sia **diagonalizzabile** (ossia ha tutti gli n autovalori).

Proposizione 1.4.16 (Derivata seconda di $F_{\mathbf{x}_0, \mathbf{v}}$). *Se f è di classe $C^{(2)}(A)$ allora F ammette derivata seconda e si ha che*

$$F''_{\mathbf{x}_0, \mathbf{v}}(t) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_j} v_i v_j \quad (1.21)$$

Dimostrazione. La derivata seconda è la derivata prima della derivata prima; essendo f di classe $C^{(2)}$, le derivate parziali $\frac{\partial f}{\partial x_i}$ sono di classe $C^{(1)}$ e, quindi derivabili; derivando nuovamente 1.19 rispetto non più a i ma a j (anchora applicando la regola della funzione composta per cui il risultato è il prodotto della derivazione della funzione con argomento non derivato, da qui il $\frac{\partial}{\partial}$ aggiuntivo,

per la derivata dell'argomento, in questo caso v_j). Si ha

$$\begin{aligned}
 F''_{\mathbf{x}_0, \mathbf{v}}(t) &= (F'_{\mathbf{x}_0, \mathbf{v}}(t))' \\
 &= \left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{v}) \cdot v_i \right)' \\
 &= \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{v}) \right)' \cdot v_i \\
 &= \sum_{i=1}^n \left(\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j}(\mathbf{x}_0 + t\mathbf{v}) \right) \cdot v_j \right) \cdot v_i \\
 &= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_j} v_i v_j
 \end{aligned}$$

□

Osservazione 69. Per esprimerla equivalentemente in notazione matriciali sfruttiamo l'Hessiana (contenente le derivate seconde della funzione f)

$$F''_{\mathbf{x}_0, \mathbf{v}}(t) = \langle \mathbf{v} \times \mathbf{H}_{f(\mathbf{x}_0 + t\mathbf{v})}, \mathbf{v} \rangle = \mathbf{v}^T \mathbf{H}_{f(\mathbf{x}_0 + t\mathbf{v})} \mathbf{v}$$

Mostriamo la seconda notazione. Partiamo dalla post-moltiplicazione dell'Hessiana per il vettore \mathbf{v} . Ricordando che l'Hessiana è la matrice

$$\mathbf{H}_{f(\mathbf{x}_0 + t\mathbf{v})} = \left(\frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_j} \right)_{i,j=1,\dots,n} = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_2 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_n^2} \end{bmatrix}$$

Focalizziamoci sulla i -esima riga

$$\begin{bmatrix} \cdots & \cdots & \vdots \\ \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_n} \\ \cdots & \cdots & \vdots \end{bmatrix}$$

Moltiplicandola per il vettore colonna \mathbf{v} si avrà la sommatoria

$$\frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_1} \cdot v_1 + \cdots + \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_n} \cdot v_n = \sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_j} \cdot v_j$$

che comparirà come elemento i -esimo del vettore colonna $\mathbf{H}\mathbf{v}$. Nello svolgimento della pre-moltiplicazione per \mathbf{v}^T (riga), questo verrà moltiplicato per l' i -esimo elemento di \mathbf{v} , ottenendo

$$\sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_j} \cdot v_j \cdot v_i$$

Si concluderà sommando tutti questi i termini (e riordinando i \mathbf{v} per estetica)

$$\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}_0 + t\mathbf{v})}{\partial x_i \partial x_j} \cdot v_i \cdot v_j$$

che appunto coincide con $F''_{\mathbf{x}_0, \mathbf{v}}(t)$

Lemma 1.4.17 (Derivata seconda $F''_{\mathbf{x}_0, \mathbf{v}}(0)$). *Il precedente risultato ci permette di dire che per $t = 0$ (ossia nel punto \mathbf{x}_0):*

$$F''_{\mathbf{x}_0, \mathbf{v}}(0) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_i \partial x_j} v_i v_j = \mathbf{v}^T \mathbf{H}_{f(\mathbf{x}_0)} \mathbf{v}$$

Proposizione 1.4.18 (Massimi/minimi: condizione necessaria). *Sia data una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (con A aperto) di classe $C^{(2)}$. Se $\mathbf{x}_0 \in A$, punto critico per f , è punto di minimo (rispettivamente massimo) relativo per f , allora per ogni vettore $\mathbf{v} \in \mathbb{R}^n$ risulta:*

$$\langle \mathbf{v} \times \mathbf{H}_{f(\mathbf{x}_0)}, \mathbf{v} \rangle \geq 0, \quad (\text{rispettivamente } \langle \mathbf{v} \times \mathbf{H}_{f(\mathbf{x}_0)}, \mathbf{v} \rangle \leq 0)$$

Dimostrazione. commentata, salta □

Teorema 1.4.19 (Massimi/minimi: condizione sufficiente). *Sia data una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto) di classe $C^{(2)}$, con \mathbf{x}_0 punto critico per f . Se, per ogni vettore $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v} \neq \mathbf{0}$ risulta*

$$\langle \mathbf{v} \times \mathbf{H}_{f(\mathbf{x}_0)}, \mathbf{v} \rangle > 0, \quad \text{rispettivamente } \langle \mathbf{v} \times \mathbf{H}_{f(\mathbf{x}_0)}, \mathbf{v} \rangle < 0$$

allora \mathbf{x}_0 è punto di minimo (risp. massimo) relativo per f .

Dimostrazione. Omessa (vedi pag 106-111 briniext, nel caso) □

Osservazione 70. La differenza tra le due sta nell'ammettere l'uguale o meno, prestar attenzione.

1.4.2.6 Classificazione dei punti critici

Osservazione 71 (Segno della forma quadratica (FQ)). Dalle condizioni precedenti risulta chiaro che ci interessa il segno di $\mathbf{v}^T \mathbf{H}_{f(\mathbf{x}_0)} \mathbf{v}$ al variare di $\mathbf{v} \in \mathbb{R}^n$. Tale scrittura è anche detta *forma quadratica* e il suo segno dipende dalle caratteristiche algebriche di $\mathbf{H}_{f(\mathbf{x}_0)}$, che deve esser studiata.

Definizione 1.4.20 (Segno della FQ ed Hessiana). La matrice $\mathbf{H}_{f(\mathbf{x}_0)}$ si dice:

- *definita* positiva (risp. negativa) se $\mathbf{v}^T \mathbf{H}_{f(\mathbf{x}_0)} \mathbf{v} > 0$ (risp. < 0), $\forall \mathbf{v} \neq \mathbf{0}$
- *semidefinita* positiva (risp. negativa) se:
 1. $\mathbf{v}^T \mathbf{H}_{f(\mathbf{x}_0)} \mathbf{v} \geq 0$ (risp. ≤ 0), $\forall \mathbf{v} \neq \mathbf{0}$
 2. $\exists \mathbf{v} \neq \mathbf{0} : \mathbf{v}^T \mathbf{H}_{f(\mathbf{x}_0)} \mathbf{v} = 0$
- *non semidefinita* (o indefinita) se $\exists \mathbf{v}_1, \mathbf{v}_2$ tali che $\mathbf{v}_1^T \mathbf{H}_{f(\mathbf{x}_0)} \mathbf{v}_1 > 0$ e $\mathbf{v}_2^T \mathbf{H}_{f(\mathbf{x}_0)} \mathbf{v}_2 < 0$

Teorema 1.4.20 (Hessiana e classificazione dei punti critici). Siano $f \in C^2(A)$ e $\mathbf{x}_0 \in A$ un punto critico per f . Se la Hessiana è:

1. *definita positiva (negativa)*, allora \mathbf{x}_0 è punto di minimo (massimo) locale forte
2. *non semidefinita (indefinita)*, allora \mathbf{x}_0 è punto di sella

Osservazione 72. Questo è un po' la stessa cosa di condizione sufficiente per massimi/minimi di cui sopra.

Osservazione 73. Ora lo studio dell'Hessiana per determinarne le caratteristiche può seguire due strade: quella del calcolo degli autovalori/determinante della matrice stessa, oppure il calcolo dei determinanti delle sottomatrici principali k -esime. Partiamo col primo.

Teorema 1.4.21 (Hessiana ed autovalori/determinanti). La matrice $\mathbf{H}_{f(\mathbf{x})}$ è

- *definita positiva (negativa)* se e solo se tutti i suoi autovalori sono positivi (negativi);
- *semidefinita positiva (negativa)* se e solo se tutti i suoi autovalori sono ≥ 0 (≤ 0) e almeno uno di essi è nullo (o anche se il determinante, prodotto degli autovalori, è nullo);
- *non semidefinita (indefinita)* se ha almeno un autovalore positivo e uno negativo

Osservazione 74. Alternativamente al calcolo degli autovalori, per classificare un punto critico si può far uso dei determinanti delle n sottomatrici principali k -esime.

Definizione 1.4.21 (Sottomatrice principale k -esima). Indicata con \mathbf{H}_k si ottiene selezionando le prime k righe e k colonne della matrice Hessiana $\mathbf{H}_{f(\mathbf{x})}$.

Esempio 1.4.14. Se per compattezza di notazione indichiamo l'Hessiana con

$$\mathbf{H}_{f(\mathbf{x})} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \dots & \dots & \dots & \dots \\ h_{n1} & h_{n2} & \dots & h_{nn} \end{bmatrix}$$

Allora

$$\mathbf{H}_1 = [h_{11}] \quad \mathbf{H}_2 = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \quad , \dots, \quad \mathbf{H}_n = \mathbf{H}_{f(\mathbf{x})}$$

Definizione 1.4.22 (Minori principali). I determinanti di queste matrici.

Teorema 1.4.22 (Hessiane definite e minori principali). *Si ha che $\mathbf{H}_{f(\mathbf{x})}$ è:*

1. *definita positiva se e solo se $\det \mathbf{H}_k > 0$ per ogni $k = 1, \dots, n$*
2. *definita negativa se e solo se $(-1)^k \det \mathbf{H}_k > 0$ per ogni $k = 1, \dots, n$ (ovvero se $h_{11} < 0, \det \mathbf{H}_2 > 0, \det \mathbf{H}_3 < 0, \dots$)*

Osservazione 75. Sintetizziamo brevemente una possibile strategia per la ricerca di estremi liberi di una funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (A aperto):

1. si isolano i punti in cui f non è regolare (ad esempio non derivabile una o due volte). Questi punti vanno esaminati a parte con uno studio in generale diretto ed adattato al caso concreto
2. tolti i punti anomali trovati al passo 1, si determinano gli eventuali punti critici di f risolvendo il sistema 1.20
3. si esamina la natura di ogni punto stazionario studiando la matrice Hessiana. Se questa è definita o non semidefinita, si conclude usando il teorema 1.4.20. Se è nulla o è semidefinita, in generale si ricorre a uno studio diretto del segno dell'incremento $\Delta f(\mathbf{x}_0)$ in un intorno di \mathbf{x}_0

Osservazione 76. Nello specifico caso bidimensionale, le considerazioni precedenti conducono alla seguente regola.

Proposizione 1.4.23. *Sian $f \in C^2(A)$, A aperto in \mathbb{R}^2 , $(x_0, y_0) \in A$ un punto critico per f e*

$$\mathbf{H}_{f(x_0, y_0)} = \begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{xy}(x_0, y_0) & f_{yy}(x_0, y_0) \end{bmatrix}$$

la matrice hessiana di f nel punto critico. Allora:

1. *se $\det \mathbf{H}_{f(x_0, y_0)} > 0$ e*
 - *$f_{xx}(x_0, y_0) > 0$ allora (x_0, y_0) è di minimo locale forte*
 - *$f_{xx}(x_0, y_0) < 0$ allora (x_0, y_0) è di massimo locale forte*

(stiamo applicando la regola dei minori principali con $\det \mathbf{H}_1 = f_{xx}(x_0, y_0)$ e $\det \mathbf{H}_2 = \det \mathbf{H}_{f(x_0, y_0)}$)
2. *se $\det \mathbf{H}_{f(x_0, y_0)} < 0$ allora (x_0, y_0) è punto di sella (determinante prodotto degli autovalori, di segno discorde dunque)*
3. *se $\det \mathbf{H}_{f(x_0, y_0)} = 0$ occorre un'analisi ulteriore*

Osservazione 77. Vediamo alcuni esempi e osservazioni su come si utilizzano in pratica le idee precedenti

Esempio 1.4.15. Sia $f(x, y) = x^2 + y^2$, che graficamente si presenta come un cono rivolto all'insù. Ha un unico punto critico $\mathbf{x}_0 = (0, 0)$ e inoltre è minimo. Mostriamolo.

$\mathbf{x}_0 = 0$ è critico $\iff \nabla f(x, y) = (0, 0) \in \mathbb{R}^2$. Il gradiente è

$$\begin{bmatrix} \frac{\partial f}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff x, y = 0$$

Ossia esiste un punto critico ed è $\mathbf{x} = (0, 0)$. Studiamo la Hessiana

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

che è diagonale pertanto ha autovalori sulla diagonale ossia $\lambda_1 = \lambda_2 = 2 > 0$. L'Hessiana pertanto è definita positiva pertanto \mathbf{x}_0 è minimo.

Esempio 1.4.16. La funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = -x^2 - y^2 + 1$ è una coppa rovesciata con massimo in $\mathbf{x}_0 = (0, 0)$. Mostriamolo. Eventuali punti critici annullano il gradiente

$$\begin{bmatrix} \frac{\partial f}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) \end{bmatrix} = \begin{bmatrix} -2x \\ -2y \end{bmatrix} = 0 \iff x = y = 0$$

quindi si $\mathbf{x}_0 = (0, 0)$ effettivamente è punto critico. Calcoliamo l'Hessiana in corrispondenza

$$\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

Si ha che la matrice è definita negativa (tutti gli autovalori, sulla diagonale sono negativi) quindi si tratta di un punto di massimo.

Esempio 1.4.17. Sia $f(x, y) = x^2 - y^2$. Non riesco ad immaginare l'immagine, ma comunque calcoliamo il gradiente facendo le derivate per andare alla ricerca di eventuali punti critici

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = 2x \\ \frac{\partial f}{\partial y}(x, y) = -2y \end{cases}$$

Si ha che $\nabla f = \mathbf{0} \iff \mathbf{x}_0 = 0$. Studiamo l'hessiana nel punto che è

$$\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

Ancora matrice diagonale, gli autovalori si trovano sulla principale e sono discordi quindi la matrice è non semidefinita, ergo \mathbf{x}_0 è un punto di sella.

Per la cronaca il grafico è quello di una patatina pringles.

Esempio 1.4.18. Cerchiamo i punti critici di

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

Occorre risolvere il sistema

$$\begin{cases} \frac{\partial f}{\partial x} = 6x - 3x^2y = 0 \\ \frac{\partial f}{\partial y} = 2y - x^3 = 0 \end{cases}$$

La prima equazione da $x = 0$ oppure $y = 2/x$. Sostituendo queste nella seconda si trovano le soluzioni

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

Poiché f è differenziabile in tutti i punti del piano, per il teorema di Fermat⁸ non possono esservi altri punti di massimo e minimo. Questi sono tutti e soli i punti critici di f .

Si tratta ora di decidere, per ciascuno di questi tre punti, se sia effettivamente un punto di estremo oppure no. La matrice hessiana di f è

$$\mathbf{H}_{f(x,y)} = \begin{bmatrix} 6 - 6xy & -3x^2 \\ -3x^2 & 2 \end{bmatrix}$$

Nel punto:

- $(0, 0)$ abbiamo

$$\mathbf{H}_{f(0,0)} = \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix}$$

La matrice ha autovalori 6 e 2 per cui è definita positiva: $(0, 0)$ è punto di minimo locale forte. Non può essere di minimo globale poiché, per esempio, $f(x, 1) \rightarrow -\infty$ se $x \rightarrow +\infty$;

- $(\sqrt{2}, \sqrt{2})$ abbiamo

$$\mathbf{H}_{f(\sqrt{2}, \sqrt{2})} = \begin{bmatrix} -6 & -6 \\ -6 & 2 \end{bmatrix}$$

La matrice ha determinante negativo e perciò è indefinita: $(\sqrt{2}, \sqrt{2})$ è un punto di sella;

- $(-\sqrt{2}, -\sqrt{2})$ abbiamo

$$\mathbf{H}_{f(-\sqrt{2}, -\sqrt{2})} = \begin{bmatrix} -6 & -6 \\ -6 & 2 \end{bmatrix}$$

per cui anche $(-\sqrt{2}, -\sqrt{2})$ è punto di sella.

Esempio 1.4.19. Cerchiamo gli estremi liberi della funzione di tre variabili

$$f(x, y, z) = 3x^2 + 2y^2 + z^2 - 2xz + 2x + 2y + 1$$

⁸Sia $y = f(x)$ una funzione $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$. Se $x_0 \in A$ è un punto estremante per f , e la funzione è derivabile in quel punto, allora si ha che $f'(x_0) = 0$.

I punti critici risolvono il sistema

$$\begin{cases} f_x(x, y, z) = 6x + 2 - 2z = 0 \\ f_y(x, y, z) = 4y + 2 = 0 \\ f_z(x, y, z) = 2z - 2x = 0 \end{cases} \quad \text{ossia} \quad \begin{cases} z = x \\ y = -1/2 \\ 2x + 1 = 0 \end{cases}$$

Si trova l'unico punto critico

$$\mathbf{x}_0 = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2} \right)$$

Calcoliamo l'Hessiana di f in \mathbf{x}_0

$$\mathbf{H}_{f(\mathbf{x}_0)} = \begin{bmatrix} 6 & 0 & -2 \\ 0 & 4 & 0 \\ -2 & 0 & 2 \end{bmatrix}$$

Per studiarne il segno, usiamo il teorema 1.4.22:

$$\det \mathbf{H}_f = 0 \cdot \dots + 4(6 \cdot 2 - (-2 \cdot (-2))) + 0 \cdot \dots = 32 > 0,$$

$$\det \begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix} = 24 > 0,$$

$$\det[6] = 6 > 0$$

Poiché i minori principali sono positivi, la forma quadratica è definita positiva. Il punto \mathbf{x}_0 è di minimo locale forte

1.4.3 Esercizi su ottimizzazione libera

Osservazione 78. Sia data una funzione f di due variabili x e y ; nel seguito, per brevità di notazione denoteremo con $f_x, f_y, f_{xx}, f_{xy}, f_{yx}, f_{yy}$ le derivate $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial y^2}$ nel “generico” punto (x, y) del dominio (se esistono).

Osservazione 79. Dato un punto \mathbf{x} del dominio in cui tali derivate esistano, scriveremo

$$f_x(\mathbf{x}), f_y(\mathbf{x}), f_{xx}(\mathbf{x}), f_{xy}(\mathbf{x}), f_{yx}(\mathbf{x}), f_{yy}(\mathbf{x})$$

per le *valutazioni* di dette derivate nel punto \mathbf{x} .

Esempio 1.4.20. Si consideri la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

$$f(x, y) = x^3 - 3x^2 - 9x + \frac{1}{y^2 + 1}$$

1. la funzione f è differenziabile su \mathbb{R}^2 ?
2. si scriva il gradiente di f nel punto $\mathbf{x}_0 = (0, 0)$;
3. si calcoli la derivata direzionale di f nel punto $\mathbf{x}_0 = (0, 0)$, secondo la direzione $\mathbf{v} = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right)$;

4. si determinino gli eventuali punti di minimo relativo e di massimo relativo per f .

Rispettivamente:

1. le derivate parziali nel punto generico (x, y) sono

$$f_x = 3x^2 - 6x - 9, \quad f_y = \frac{-2y}{(y^2 + 1)^2}.$$

La funzione f è ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (vedi Teorema del Differenziale Totale)

2. si ha $\nabla f(0, 0) = (-9, 0)$

3. si ha

$$\frac{\partial f}{\partial \mathbf{v}}(0, 0) = \left\langle (-9, 0), \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right) \right\rangle = \frac{-9\sqrt{2}}{2} = -\frac{9}{\sqrt{2}}$$

4. per determinare gli eventuali punti di estremo relativo, dobbiamo prima trovare i punti critici:

$$\begin{cases} f_x = 3x^2 - 6x - 9 = 0 \\ f_y = \frac{-2y}{(y^2 + 1)^2} = 0 \end{cases}$$

dalla prima si hanno due soluzioni $x = 3 \vee x = -1$ dalla seconda $y = 0$. Abbiamo perciò due punti critici:

$$\mathbf{x}_1 = (3, 0), \quad \mathbf{x}_2 = (-1, 0)$$

Considerato che

$$f_{xx} = 6x - 6, \quad f_{yx} = f_{xy} = 0, \quad f_{yy} = \frac{6y^2 - 2}{(y^2 + 1)^3}$$

la generica Hessiana risulta

$$\mathbf{H}_{f(\mathbf{x})} = \begin{bmatrix} 6x - 6 & 0 \\ 0 & \frac{6y^2 - 2}{(y^2 + 1)^3} \end{bmatrix}$$

e valutandola nei punti critici, abbiamo

$$\mathbf{H}_{f(\mathbf{x}_1)} = \begin{bmatrix} 12 & 0 \\ 0 & -2 \end{bmatrix}, \quad \mathbf{H}_{f(\mathbf{x}_2)} = \begin{bmatrix} -12 & 0 \\ 0 & -2 \end{bmatrix}$$

Possiamo utilizzare la regola degli autovalori per concludere che la prima matrice è *non semidefinita*, per cui \mathbf{x}_1 è punto di *sella*, mentre la seconda matrice è *definita negativa*, per cui \mathbf{x}_2 è punto di *massimo*.

Esempio 1.4.21. Considerando la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) = (x^2 + 1)(y^3 - 3y)$$

1. f è differenziabile in ogni punto di \mathbb{R}^2 ?
2. si determini il gradiente di f in $\mathbf{x}_0 = (0, 0)$ e la derivata di f lungo la direzione $\mathbf{v} = (\frac{3}{5}, -\frac{4}{5})$ nel punto \mathbf{x}_0 ;
3. quale fra i punti \mathbf{x}_0 , $\mathbf{x}_1 = (0, 1)$ e $\mathbf{x}_2 = (0, -1)$ è un estremo relativo per f ?

Rispettivamente

1. le derivate parziali nel punto generico (x, y) sono

$$f_x = 2x(y^3 - 3y), \quad f_y = (x^2 + 1)(3y^2 - 3)$$

La funzione f è ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr Teorema del Differenziale Totale).

2. si ha $\nabla f(0, 0) = (0, -3)$, da cui

$$\frac{\partial f}{\partial \mathbf{v}}(0, 0) = \left\langle (0, -3), \left(\frac{3}{5}, -\frac{4}{5}\right) \right\rangle = \frac{12}{5}$$

3. Il punto $\mathbf{x}_0 = (0, 0)$ non è critico (basta sostituirlo nelle derivate parziali e vedere che $f_y \neq 0$ non soddisfacendo dunque le condizioni necessarie per un punto critico), quindi non è punto di estremo relativo. I punti $\mathbf{x}_2 = (0, 1)$ e $\mathbf{x}_3 = (0, -1)$ sono critici (per essi $f_x = f_y = 0$) e dobbiamo quindi esaminare la Hessiana valutata in questi.

Considerato che

$$f_{xx} = 2(y^3 - 3y), \quad f_{yx} = f_{xy} = 2x(3y^2 - 3), \quad f_{yy} = 6y(x^2 + 1)$$

l'Hessiana (nel punto generico (x, y)) è:

$$\mathbf{H}_{f(\mathbf{x})} = \begin{bmatrix} 2(y^3 - 3y) & 2x(3y^2 - 3) \\ 2x(3y^2 - 3) & 6y(x^2 + 1) \end{bmatrix}$$

e nei punti critici \mathbf{x}_1 e \mathbf{x}_2 si ha:

$$\mathbf{H}_{f(\mathbf{x}_1)} = \begin{bmatrix} -4 & 0 \\ 0 & 6 \end{bmatrix} \quad \mathbf{H}_{f(\mathbf{x}_2)} = \begin{bmatrix} 4 & 0 \\ 0 & -6 \end{bmatrix}$$

Entrambe le matrici sono *non semidefinite*, per cui \mathbf{x}_1 e \mathbf{x}_2 sono *punti di sella*.

Esempio 1.4.22. Considerando la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definita da

$$f(x, y) = (x^2 + 1)(y^3 + 3y^2 + 1)$$

1. la funzione f è differenziabile su \mathbb{R}^2 ?
2. si determini il gradiente di f in $\mathbf{x}_0 = (1, 1)$.
3. si determinino gli eventuali punti di massimo e minimo relativo per la funzione f .

Rispettivamente:

1. le derivate parziali nel punto generico (x, y) sono

$$f_x = 2x(y^3 + 3y^2 + 1), \quad f_y = (x^2 + 1)(3y^2 + 6y)$$

La funzione f è ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale).

2. si ha $\nabla f(1, 1) = (10, 18)$.
3. impostando

$$\begin{cases} f_x = 2x(y^3 + 3y^2 + 1) = 0 \\ f_y = (x^2 + 1)(3y^2 + 6y) = 0 \end{cases}$$

la derivata parziale f_y si annulla per $y = 0 \vee y = -2$. Poiché, per tali valori si ha $y^3 + 3y + 1 > 0$, per soddisfare il sistema deve essere $x = 0$. I punti critici sono pertanto:

$$\mathbf{x}_1 = (0, 0), \quad \mathbf{x}_2 = (0, -2)$$

Considerato che:

$$\begin{aligned} f_{xx} &= 2(y^3 + 3y^2 + 1) \\ f_{yx} &= f_{xy} = 0 \\ f_{yy} &= (6y + 6)(x^2 + 1) \end{aligned}$$

si ha

$$\mathbf{H}_{f(\mathbf{x})} = \begin{bmatrix} 2(y^3 + 3y^2 + 1) & 0 \\ 0 & (6y + 6)(x^2 + 1) \end{bmatrix}$$

e dunque

$$\mathbf{H}_f(\mathbf{x}_1) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}, \quad \mathbf{H}_f(\mathbf{x}_2) = \begin{bmatrix} 2(-8 + 12 + 1) & 0 \\ 0 & -6 \cdot 1 \end{bmatrix} = \begin{bmatrix} 10 & 0 \\ 0 & -6 \end{bmatrix}$$

Perciò \mathbf{x}_1 è di minimo, \mathbf{x}_2 è di sella.

Esempio 1.4.23. Data la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

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

- f è differenziabile in \mathbb{R}^2 ?

- si calcoli il gradiente di f nel punto $\mathbf{x}_0 = (1, 1)$.
- si calcoli la derivata direzionale di f in $\mathbf{x}_0 = (1, 1)$ secondo la direzione $\mathbf{v} = (\sqrt{2}/2, \sqrt{2}/2)$.
- la funzione f ammette punti di estremo relativo?

Rispettivamente:

1. le derivate parziali nel punto generico (x, y) sono

$$f_x = 2x + y, \quad f_y = x - 2y$$

La funzione f è ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale);

2. si ha $\nabla f(1, 1) = (3, -1)$;
3. l'unico punto critico è perciò il punto $\mathbf{x} = (0, 0)$. Gli elementi della matrice Hessiana (nel punto generico (x, y)) sono:

$$f_{xx} = 2 \quad f_{yx} = f_{xy} = 1, \quad f_{yy} = -2$$

Pertanto l'Hessiana è

$$\mathbf{H}_f(\mathbf{x}) = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$$

Il determinante della matrice Hessiana vale -5 ; poiché tale determinante fornisce il prodotto degli autovalori, gli autovalori hanno segno discorde e quindi la matrice è *non semidefinita*. Perciò $\mathbf{x}_0 = (0, 0)$ è di sella.

Esempio 1.4.24. Data la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

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

si determinino i punti critici di f . Quali tra questi sono punti di estremo relativo?

- le derivate parziali nel punto generico (x, y) sono

$$f_x = 9x^2 - 1, \quad f_y = 6y$$

La funzione f è ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale);

- i punti critici soddisfano

$$\begin{cases} 9x^2 - 1 = 0 \\ 6y = 0 \end{cases}$$

ossia sono

$$\mathbf{x}_1 = \left(\frac{1}{3}, 0\right), \quad \mathbf{x}_2 = \left(-\frac{1}{3}, 0\right)$$

- gli elementi della matrice Hessiana (nel punto generico (x, y)) sono:

$$f_{xx} = 18x, \quad f_{yx} = f_{xy} = 0, \quad f_{yy} = 6$$

per cui

$$\mathbf{H}_{f(\mathbf{x})} = \begin{bmatrix} 18x & 0 \\ 0 & 6 \end{bmatrix}$$

L'Hessiana valutata nei punti critici è

$$\mathbf{H}_{f(\mathbf{x}_1)} = \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}, \quad \mathbf{H}_{f(\mathbf{x}_2)} = \begin{bmatrix} -6 & 0 \\ 0 & 6 \end{bmatrix}$$

Perciò \mathbf{x}_1 è di minimo, \mathbf{x}_2 è di sella.

Esempio 1.4.25. Si consideri la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

$$f(x, y) = x \sin(y)$$

1. la funzione f è differenziabile su \mathbb{R}^2 ?
2. si calcoli il gradiente di f nel punto $\mathbf{x}_0 = (\pi, \pi)$
3. si calcoli la derivata direzionale di f nel punto \mathbf{x}_0 , secondo la direzione $\mathbf{v} = (-1/2, \sqrt{3}/2)$
4. si determinino i punti critici di f
5. si determinino gli eventuali punti di estremo relativo per f

Rispettivamente:

1. le derivate parziali nel punto generico (x, y) sono

$$f_x = \sin(y), \quad f_y = x \cos(y)$$

La funzione f è ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale).

2. il gradiente in $\mathbf{x}_0 = (\pi, \pi)$ è

$$\nabla f(\mathbf{x}_0) = (\sin(\pi), \pi \cos(\pi)) = (0, -\pi)$$

3. per la derivata direzionale in $\mathbf{x}_0 = (\pi, \pi)$ per $\mathbf{v} = (-1/2, \sqrt{3}/2)$

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{v}} = \nabla f(\mathbf{x}_0) \cdot \mathbf{v} = (0, -\pi) \cdot (-1/2, \sqrt{3}/2) = -\frac{3\pi}{2}$$

4. si ha

$$\begin{cases} \sin(y) = 0 \\ x \cos(y) = 0 \end{cases} \quad \begin{cases} y = \pi \cdot k \\ y = \frac{\pi}{2} + k\pi \vee x = 0 \end{cases}$$

i punti critici dunque debbono essere della forma $(0, k\pi)$ con $k \in \mathbb{Z}$;

5. gli elementi della matrice Hessiana (nel punto generico (x, y)) sono:

$$f_{xx} = 0, \quad f_{yx} = f_{xy} = \cos(y), \quad f_{yy} = -x \sin(y)$$

L'Hessiana

$$\mathbf{H}_{f(\mathbf{x})} = \begin{bmatrix} 0 & \cos(y) \\ \cos(y) & -x \sin(y) \end{bmatrix}$$

e per $\mathbf{x}_k = (0, k\pi)$ con $k = 1$ e $k = 2$, si ha

$$\mathbf{H}_{f(\mathbf{x}_2)} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \quad \mathbf{H}_{f(\mathbf{x}_1)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

con $\mathbf{x}_1, \mathbf{x}_2$ i primi due punti critici (che si ripeteranno infinitamente); in entrambi i casi essendo determinante negativo, è *non semidefinita*: perciò tutti i punti critici sono *punti di sella*.

Esempio 1.4.26. Verifichiamo che la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

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

ha un minimo relativo in $\mathbf{x}_0 = (0, 0)$. Infatti si ha

$$f_x = 3x^2 + 2x + 2xy$$

$$f_y = 2y + x^2$$

$$f_{xx} = 6x + 2 + 2y$$

$$f_{xy} = 2x$$

$$f_{yx} = 2x$$

$$f_{yy} = 2$$

Il differenziale in $\mathbf{x}_0 = (0, 0)$ è $df(0, 0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v} = 0, \forall \mathbf{v} \in \mathbb{R}$, pertanto $\mathbf{x}_0 = (0, 0)$ è punto critico. Inoltre la matrice Hessiana nel generico punto (x, y) è data da

$$\mathbf{H}_{f(x, y)} = \begin{bmatrix} 6x + 2y + 2 & 2x \\ 2x & 2 \end{bmatrix}$$

che in $\mathbf{x}_0 = (0, 0)$ è definita positiva; perciò questo è punto di minimo relativo.

Esempio 1.4.27. Verifichiamo che $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definita da

$$f(x, y) = 5 \cos xy + e^{(x-1)y} - (x-1)^2 + 2(x-1)y$$

ha un punto di massimo relativo in $(1, 0)$. Si ha:

$$f_x = -5y \sin xy + ye^{(x-1)y} - 2(x-1) + 2y$$

$$f_y = -5x \sin xy + (x-1)e^{(x-1)y} + 2(x-1)$$

In $(1, 0)$ risulta

$$f_x(1, 0) = 0, \quad f_y(1, 0) = 0$$

NB: Esercizio dagli appunti extended

NB: Esercizio dagli appunti extended

per cui $df(1,0) = 0$ e quindi $(1,0)$ è punto critico per f . Inoltre si ha

$$\begin{aligned} f_{xx} &= -5y^2 \cos xy + y^2 e^{(x-1)y} - 2 \\ f_{xy} = f_{yx} &= -5 \cos xy \cdot xy - 5 \sin xy + e^{(x-1)y} + y(x-1)e^{(x-1)y} + 2 \\ f_{yy} &= -5 \cos xy \cdot x^2 + (x-1)^2 e^{(x-1)y} \end{aligned}$$

Pertanto risulta

$$\mathbf{H}_{f(1,0)} = \begin{bmatrix} -2 & 3 \\ 3 & -5 \end{bmatrix}$$

che è definita negativa per il teorema 1.4.22. Ne consegue che $(1,0)$ è punto di massimo relativo

1.4.4 Il teorema di Dini

Osservazione 80. Questa parte costituisce una introduzione particolare al teorema delle funzioni implicite e può essere posticipata/letta prima di quello.

Osservazione 81 (Il problema). Consideriamo la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ data da

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

ossia il paraboloide volto all'insù con origine nel piano $z = -1$. Ci interessa sapere l'insieme di valori dove la funzione si annulla e se si può descrivere una relazione tra questi.

Il luogo degli zeri di f è l'insieme

$$Z_f = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 - 1 = 0\}$$

cioè la circonferenza unitaria centrata nell'origine.

Posto $X = \mathbb{R}$ per l'asse delle ascisse, $Y = \mathbb{R}$ per l'asse delle ordinate in \mathbb{R}^2 possiamo chiederci: è possibile esprimere l'insieme Z_f (la circonferenza) come *grafico* alternativamente di funzioni

$$\varphi : X \rightarrow Y, \quad \psi : Y \rightarrow X$$

(o comunque aventi dominio un sottoinsieme di X o Y rispettivamente)?

Osservazione 82 (Soluzione in forma globale). In questa forma il problema ha risposta *negativa* per entrambe le funzioni:

- supponiamo ad esempio che la circonferenza Z_f sia grafico di una funzione $\varphi : [-1, 1] \subset X \rightarrow Y$. Geometricamente è evidente che ciò è assurdo: per ogni $x \in (-1, 1)$, $\varphi(x)$ dovrebbe assumere 2 valori (pertanto non sarebbe una funzione) (figura 1.3, caso (a));
- specularmente avverrebbe con una ipotetica $\psi : [-1, 1] \subset Y \rightarrow X$ per $y \in (-1, 1)$. (figura 1.3 caso (b))

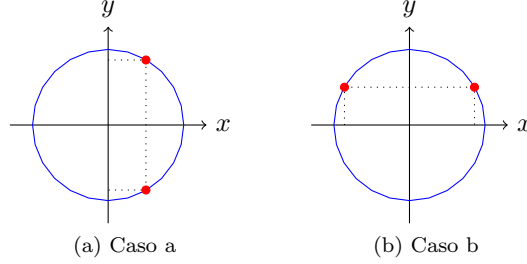


Figura 1.3: Due casi (senza soluzione in forma globale)

Osservazione 83 (Soluzione in forma locale). Possiamo riformulare il problema ma in versione *locale*: dato un punto $(x_0, y_0) \in Z_f$ è possibile determinare un intorno rettangolare $I_{(x_0, y_0)} = \{(x, y) \in \mathbb{R}^2 : x \in I(x_0, \delta_1), y \in I(y_0, \delta_2)\}$ all'interno del quale l'insieme $Z_f \cap I_{x_0, y_0}$ (ossia l'insieme degli zeri di questo intorno) sia il grafico alternativamente di una delle due funzioni?

$$\varphi : I(x_0, \delta_1) \subset \mathbb{R} \rightarrow I(y_0, \delta_2) \subset \mathbb{R}$$

$$\psi : I(y_0, \delta_2) \subset \mathbb{R} \rightarrow I(x_0, \delta_1) \subset \mathbb{R}$$

Così posto il problema *può* avere risposta *affermativa*.

Esempio 1.4.28. Se fissiamo per semplicità $\delta_1 = \delta_2 = \varepsilon < 1$:

1. in $\mathbf{x}_0 = (x_0, y_0) = (0, 1)$, si ha che $Z_f \cap I_{(0,1)}$ è la porzione di cerchio in alto, attorno a $(0, 1)$ (figura 1.4, a).

Si ha che $Z_f \cap I_{(0,1)}$ è il grafico di $\varphi : (-\varepsilon, \varepsilon) \rightarrow Y$ definita da

$$\varphi(x) = \sqrt{1 - x^2} = y \quad \text{localmente}$$

2. in $\mathbf{x}_1 = (0, -1)$ si ha che $Z_f \cap I_{(0,-1)}$ è uno spicchio di cerchio in basso attorno a $(0, -1)$ (figura 1.4, b), e che è il grafico di $\varphi : (-\varepsilon, \varepsilon) \rightarrow Y$ definita da

$$\varphi(x) = -\sqrt{1 - x^2} = y \quad \text{localmente}$$

3. per $\mathbf{x}_2 = (1, 0)$ il problema non può essere risolto seguendo lo stesso procedimento utilizzato nei due precedenti casi. Lo spicchio di grafico della funzione (figura 1.4, c) è quello attorno a $(1, 0)$ e, per quanto piccolo sia l'intorno $I_{(1,0)}$ la presunta funzione $\varphi(x) = y$ dovrebbe assumere due valori contemporaneamente (contro la definizione di funzione)

4. sempre per $\mathbf{x}_2 = (1, 0)$, l'insieme $Z_f \cap I_{(1,0)}$ considerato in precedenza può essere tuttavia riguardato come grafico di una funzione ψ dalle y alle x specificamente

$$\psi(y) = \sqrt{1 - y^2} = x \quad (\text{localmente})$$

Si pensi a $I_{(x_0, y_0)}$ come a un rettangolino che fa da intorno a (x_0, y_0) e che ha i suoi lati paralleli agli assi. Si pensi a $I_{(x_0, y_0)}$ come a un rettangolino che fa da intorno a (x_0, y_0) e che ha i suoi lati paralleli agli assi.

5. per $\mathbf{x}_3 = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ l'intorno rettangolare evidenzia uno spicchio attorno ai 135 gradi (figura 1.4 d) dove, a differenza dei casi precedenti dove mancava l'iniettività, si ha una biezione; in questo caso $Z_f \cap I\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ può essere riguardato *sia* come grafico di

$$\varphi(x) = \sqrt{1-x^2} = y$$

che come grafico di

$$\psi(y) = -\sqrt{1-y^2} = x$$

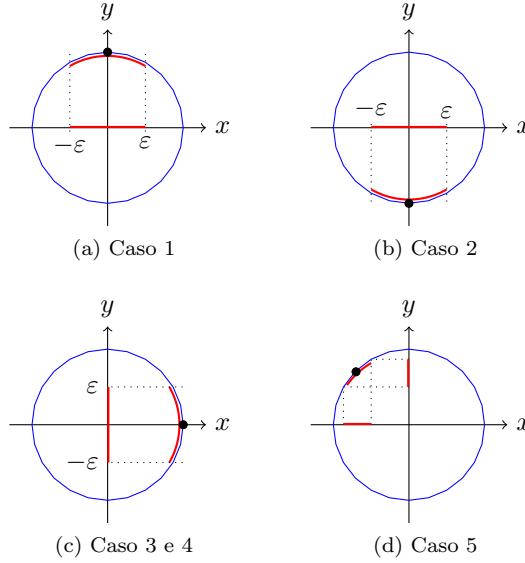


Figura 1.4: Quattro casi

Teorema 1.4.24 (Teorema di Dini (sulle funzioni implicite, caso speciale a 2 variabili)). *Sia:*

- $f : A \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ (A aperto) di classe $C^{(1)}(A)$. Nell'esempio sopra questo è il nostro paraboloide;
- sia (x_0, y_0) tale che $f(x_0, y_0) = 0$
- supponiamo che sia

$$\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$$

(ossia nell'esempio di sopra di non essere $2y = 0$ ossia di non trovarsi nei punti $(-1, 0)$ o $(1, 0)$)

Allora:

- esiste un intorno di $I(x_0, \delta_1)$, e un intorno di $I(y_0, \delta_2)$ tali che per ogni $x \in I(x_0, \delta_1)$ esiste uno e un solo $y \in I(y_0, \delta_2)$ coi quali si abbia che $f(x, y) = 0$.
Pertanto possiamo “costruire”, almeno in linea teorica essendo la definizione rispettata, una funzione del tipo $\varphi : I(x_0, \delta_1) \rightarrow I(y_0, \delta_2)$ con $y = \varphi(x)$

- detta funzione φ è di classe C^1
- vale la seguente

$$\varphi'(x) = - \frac{\frac{\partial f(x, y)}{\partial x}}{\frac{\partial f(x, y)}{\partial y}} \quad \forall x \in I(x_0, \delta_1)$$

Osservazione importante 12 (Lessico). Si usa dire che y è esplicitabile come funzione di x localmente e la funzione φ è definita implicitamente dalla “equazione” $f(x, y) = 0$.

Osservazione 84. Chiaramente la cosa è speculare e si può fare a variabili invertite, ossia se/partendo da:

$$\frac{\partial f}{\partial x}(x_0, y_0) \neq 0$$

e proseguendo specularmente.

Osservazione importante 13 (Forma di Z_f). Sia f di tipo Dini, ossia $\forall (x_0, y_0) \in Z_f$ si ha almeno una delle due condizioni:

$$\begin{aligned} \frac{\partial f}{\partial y}(x_0, y_0) &\neq 0 \\ \frac{\partial f}{\partial x}(x_0, y_0) &\neq 0 \end{aligned}$$

Che forma avrà $Z_f \subseteq \mathbb{R}^2$? Avrà una forma liscia: senza cuspidi/punti o senza nodi (non si incrocia).

Esempio 1.4.29. Tornando al caso $f(x, y) = x^2 + y^2 - 1$ l’equazione $x^2 + y^2 - 1 = 0$ definisce geometricamente la circonferenza unitaria centrata nell’origine.

1. per $\mathbf{x}_0 = (x_0, y_0) = (0, 1)$ poiché:

$$\frac{\partial f}{\partial y}(0, 1) = 2 \neq 0$$

l’equazione $x^2 + y^2 - 1 = 0$ definisce implicitamente in un intorno di $x_0 = 0$ la funzione

$$y = \varphi(x) = \sqrt{1 - x^2}$$

2. per $\mathbf{x}_1 = (0, -1)$

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

l'equazione $x^2 + y^2 - 1 = 0$ definisce implicitamente, in un intorno di $x_0 = 0$ la funzione

$$y = \varphi(x) = -\sqrt{1 - x^2}$$

3. per $\mathbf{x}_2 = (1, 0)$, l'equazione $x^2 + y^2 - 1 = 0$ *non definisce* implicitamente una funzione $y = \varphi(x)$, essendo

$$\frac{\partial f}{\partial y}(1, 0) = 0$$

Tuttavia, essendo

$$\frac{\partial f}{\partial x}(1, 0) = 2 \neq 0$$

l'equazione $x^2 + y^2 - 1 = 0$ *definisce implicitamente* in un intorno di $y_0 = 0$ la funzione

$$x = \psi(y) = \sqrt{1 - y^2}$$

Esempio 1.4.30. Considerando una nuova funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

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

Poiché $x^2 - y^2 - 2x + 2y = (x + y - 2)(x - y)$ il luogo degli zeri di f è geometricamente la coppia di rette $(x - y = 0) \vee (x + y - 2 = 0)$ (che si incrociano in $(1, 1)$, figura 1.5). Abbiamo

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= 2x - 2 \\ \frac{\partial f}{\partial y}(x, y) &= -2y + 2\end{aligned}$$

Ora:

1. poiché la derivata parziale rispetto ad y è diversa da zero per ogni $y \neq 1$, l'equazione $x^2 - y^2 - 2x + 2y = 0$ comunque si fissi un punto $\mathbf{x}_0 \neq (1, 1)$ definisce implicitamente una funzione

$$\varphi : x \mapsto y = \varphi(x)$$

da un intorno U di x_0 ad un intorno V di y_0 ;

2. poiché la derivata parziale rispetto ad x è diversa da zero per ogni $x \neq 1$, l'equazione $x^2 - y^2 - 2x + 2y = 0$ comunque si fissi un punto $\mathbf{x}_0 \neq (1, 1)$ definisce implicitamente una funzione

$$\psi : y \mapsto x = \psi(y)$$

da un intorno V di y_0 ad un intorno U di x_0 ;

3. nel punto $(1, 1)$ si ha

$$\frac{\partial f}{\partial x}(1, 1) = \frac{\partial f}{\partial y}(1, 1) = 0$$

Si noti che l'equazione $x^2 - y^2 - 2x + 2y = 0$ *non definisce* una funzione $\varphi : x \mapsto y$ né una funzione $\psi : y \mapsto x$ in alcun intorno di $(1, 1)$ (come ci si muove ci si muove, al ridosso dell'incrocio qualsiasi valore del dominio sarebbe associato a due valori del codominio)

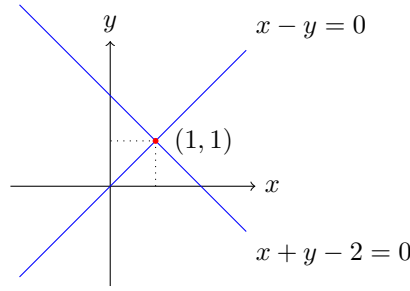


Figura 1.5: Luogo degli zeri di $f(x, y) = x^2 - y^2 - 2x + 2y$

1.5 Funzioni $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$

1.5.1 Funzioni a valori vettoriali di più variabili

Definizione 1.5.1 (Funzione a valori vettoriali e funzioni componenti). Funzione $\mathbf{f} : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$, (A aperto) definita da m funzioni $f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$ dette *componenti scalari*, ossia $\mathbf{f} \equiv (f_1, \dots, f_m)$.

Definizione 1.5.2 (Immagine di un vettore). Dato $\mathbf{x} = (x_1, \dots, x_n) \in A$, la sua *immagine* mediante \mathbf{f} è $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$; e si può scrivere come m -pla di funzioni componenti

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$$

Esempio 1.5.1. La funzione $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tale che per $\mathbf{x} = (x, y) \in \mathbb{R}^2$ risulti

$$\mathbf{f}(\mathbf{x}) = (2xy^2, 3x - 2xy, -x^2y)$$

si scriverà $\mathbf{f} \equiv (f_1, f_2, f_3)$ con

$$\begin{aligned} f_1 : \mathbb{R}^2 &\rightarrow \mathbb{R}, & f_1(x, y) &= 2xy^2 \\ f_2 : \mathbb{R}^2 &\rightarrow \mathbb{R}, & f_2(x, y) &= 3x - 2xy \\ f_3 : \mathbb{R}^2 &\rightarrow \mathbb{R}, & f_3(x, y) &= -x^2y \end{aligned}$$

Osservazione 85. Dato che in generale, per studiare una funzione a valori vettoriali si può ragionare componente per componente, e che le componenti sono funzioni reali di più variabili, la definizione del concetto di limite, continuità, e differenziabilità per una funzione di più variabili a valori vettoriali non introduce nuove difficoltà sostanziali rispetto a quanto visto in precedenza.

Definizione 1.5.3 (Limite). Sia $\mathbf{f} : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ definita almeno in un intorno del punto $\mathbf{x}_0 \in \mathbb{R}^n$, salvo al più \mathbf{x}_0 stesso e sia $\mathbf{L} \in \mathbb{R}^m$. Si dice che

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{L}$$

se

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \|\mathbf{f}(\mathbf{x}) - \mathbf{L}\| = 0$$

Osservazione 86. Analogamente a quanto visto nel capitolo delle funzioni vettoriali di variabile reale, si dimostra che il limite si calcola componente per componente, ossia

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})) = \left(\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_1(\mathbf{x}), \dots, \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_m(\mathbf{x}) \right)$$

Definizione 1.5.4 (Continuità in un punto). Si ha che \mathbf{f} è continua in \mathbf{x}_0 se

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0)$$

Definizione 1.5.5 (Continuità in un aperto). Continuità in tutti i punti che compongono l'aperto

Proposizione 1.5.1 (Condizione di continuità). *La funzione \mathbf{f} è continua (in un punto o in un aperto) se e solo se lo sono tutte le sue componenti.*

1.5.2 Funzioni differenziabili

Definizione 1.5.6 (Funzione differenziabile in un punto \mathbf{x}_0). Si dice che $\mathbf{f} : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ (A aperto) è *differenziabile nel punto* $\mathbf{x}_0 \in A$ se e solo se esiste un operatore lineare $\mathbf{L}_{\mathbf{x}_0} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ dipendente da \mathbf{x}_0 tale che risulti, per una variazione $\mathbf{h} \in \mathbb{R}^n$:

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) - \mathbf{L}_{\mathbf{x}_0}(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0} \in \mathbb{R}^m \quad (1.22)$$

ove, come sempre, $\|\mathbf{h}\|$ indica la norma euclidea del vettore $\mathbf{h} \in \mathbb{R}^n$.

Osservazione 87. Vediamo che qui, al numeratore:

- abbiamo una differenza tra vettori appartenenti a \mathbb{R}^m (anche la funzione $\mathbf{L}_{\mathbf{x}_0}$ restituisce ciò), mentre $\mathbf{x}_0, \mathbf{h} \in \mathbb{R}^n$
- nuovamente si vede come la funzione \mathbf{f} possa esser approssimata in un intorno di \mathbf{x}_0 come somma di vettore costante $\mathbf{f}(\mathbf{x}_0)$ e di una trasformazione lineare $\mathbf{L}_{\mathbf{x}_0}(\mathbf{h})$

Definizione 1.5.7 (Differenziale). L'operatore $\mathbf{L}_{\mathbf{x}_0}$:

- si dice ancora differenziale di \mathbf{f} in \mathbf{x}_0 e si denota brevemente anche con il simbolo $d\mathbf{f}(\mathbf{x}_0)$, mentre le sue “valutazioni”, per una variazione $\mathbf{h} \in \mathbb{R}^n$, si indicano con

$$\mathbf{L}_{\mathbf{x}_0}(\mathbf{h}) = d\mathbf{f}(\mathbf{x}_0)(\mathbf{h})$$

- essendo un vettore di m elementi, la j -esima componente (per $j = 1, \dots, m$) di $\mathbf{L}_{\mathbf{x}_0}$, può essere indicata con $L_{\mathbf{x}_0,j}$ (ossia $\mathbf{L}_{\mathbf{x}_0} \equiv (L_{\mathbf{x}_0,1}, \dots, L_{\mathbf{x}_0,m})$) e coincide con il differenziale della j -esima componente scalare f_j di \mathbf{f} . Ossia data una variazione $\mathbf{h} = (h_1, \dots, h_n) \in \mathbb{R}^n$, si ha:

$$L_{\mathbf{x}_0,j}(\mathbf{h}) = \langle \nabla f_j(\mathbf{x}_0), \mathbf{h} \rangle = \sum_{k=1}^n \frac{\partial f_j(\mathbf{x}_0)}{\partial x_k} \cdot h_k \quad (1.23)$$

Proposizione 1.5.2 (Differenziabilità e componenti). *La funzione a valori vettoriali $\mathbf{f} \equiv (f_1, \dots, f_m)$ è differenziabile in \mathbf{x}_0 se e solo se le sue componenti f_1, \dots, f_m sono tutte differenziabili in \mathbf{x}_0*

Dimostrazione. Possiamo spezzare il vettore risultante al numeratore di 1.22 nelle singole componenti e richiedere che ciascuna di esse sia 0. Ossia deve essere

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f_j(\mathbf{x}_0 + \mathbf{h}) - f_j(\mathbf{x}_0) - L_{\mathbf{x}_0,j}(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

per ciascun $j = 1, \dots, m$, ossia alla condizione di differenziabilità delle singole componenti. \square

Definizione 1.5.8 (Matrice jacobiana della funzione \mathbf{f} in \mathbf{x}_0). Data una funzione $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, si tratta della matrice $m \times n$ le cui righe sono i *gradienti* delle componenti scalari f_1, \dots, f_m ossia:

$$\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)} = \begin{bmatrix} \frac{\partial f_1(\mathbf{x}_0)}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x}_0)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x}_0)}{\partial x_1} & \dots & \frac{\partial f_m(\mathbf{x}_0)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla f_1(\mathbf{x}_0) \\ \dots \\ \nabla f_m(\mathbf{x}_0) \end{bmatrix} \quad (1.24)$$

Osservazione importante 14 (Differenziale come prodotto di Jacobiana per variazione). Alla luce della 1.23 e definita la Jacobiana come in 1.24, possiamo riscrivere il differenziale come:

$$\mathbf{L}_{\mathbf{x}_0}(\mathbf{h}) = d\mathbf{f}(\mathbf{x}_0)(\mathbf{h}) = \mathbf{J}_{\mathbf{f}(\mathbf{x}_0)} \mathbf{h} = \begin{bmatrix} L_{\mathbf{x}_0,1}(\mathbf{h}) \\ \dots \\ L_{\mathbf{x}_0,m}(\mathbf{h}) \end{bmatrix} = \begin{bmatrix} \langle \nabla f_1(\mathbf{x}_0), \mathbf{h} \rangle \\ \dots \\ \langle \nabla f_m(\mathbf{x}_0), \mathbf{h} \rangle \end{bmatrix}$$

Pertanto a differenza delle funzioni $\mathbb{R}^n \rightarrow \mathbb{R}$ dove la stima della variazione intercorsa nella funzione per la variazione \mathbf{h} è un singolo numero, qui la stima è un vettore.

Osservazione 88. Altresì ora la condizione di differenziabilità 1.22 si può riscrivere utilizzando il prodotto righe per colonne della Jacobiana con il vettore variazione $\mathbf{h} \in \mathbb{R}^n$, come segue

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) - \mathbf{J}_{\mathbf{f}(\mathbf{x}_0)} \mathbf{h}}{\|\mathbf{h}\|} = \mathbf{0} \in \mathbb{R}^m$$

Definizione 1.5.9 (Funzione di classe $C^{(k)}(A)$). Funzione $\mathbf{f} \equiv (f_1, \dots, f_m) : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ (con A aperto) per la quale risulta $f_j \in C^{(k)}(A)$ per ogni $j = 1, \dots, m$

Definizione 1.5.10 (Funzione di classe $C^{(\infty)}(A)$). Se $f_j \in C^{(\infty)}(A)$ per ogni $j = 1, \dots, m$.

Osservazione 89. Utilizzando gli analoghi risultati per funzioni a valori scalari, si deduce immediatamente

Proposizione 1.5.3 (Differenziabilità e continuità). Se

- $\mathbf{f} \in C^{(1)}(A)$, allora \mathbf{f} è differenziabile in A ;
- \mathbf{f} è differenziabile in $\mathbf{x}_0 \in A$, allora \mathbf{f} è continua in \mathbf{x}_0 ;
- \mathbf{f} è differenziabile in A , allora \mathbf{f} è continua in A ($\mathbf{f} \in C^{(0)}(A)$).

Osservazione 90. La prima affermazione dice che condizione sufficiente affinché una funzione risulti differenziabile in A è che tutti gli elementi della sua Jacobiana siano funzioni continue in A .

Questa è la versione “multidimensionale” del Teorema del differenziale totale e di fatto il criterio più applicato per garantire la differenziabilità di una funzione di più variabili a valori vettoriali, senza dover verificare direttamente la definizione.

Teorema 1.5.4 (Differenziabilità composizione di funzioni differenziabili). Sia:

- $\mathbf{g} : A \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^n$, A aperto
- $\mathbf{f} : B \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^p$, B aperto
- $\mathbf{g}(A) \subseteq B$

Possiamo allora definire la funzione composta

$$\begin{aligned} \mathbf{f} \circ \mathbf{g} : A &\subseteq \mathbb{R}^m \rightarrow \mathbb{R}^p \\ \mathbf{f} \circ \mathbf{g}(\mathbf{x}) &= \mathbf{f}(\mathbf{g}(\mathbf{x})) \end{aligned}$$

Ora facciamo qualche assunzione. Siano

1. $\mathbf{a} \in A$
2. $\mathbf{g}(\mathbf{a}) = \mathbf{b} \in B$

3. \mathbf{g} differenziabile in \mathbf{a}

4. \mathbf{f} differenziabile in $\mathbf{b} = \mathbf{g}(\mathbf{a})$

Allora si ha che: $\mathbf{f} \circ \mathbf{g}$ è differenziabile in \mathbf{a} e si ha:

$$d(\mathbf{f} \circ \mathbf{g})(\mathbf{a}) = d\mathbf{f}(\mathbf{g}(\mathbf{a})) \circ d\mathbf{g}(\mathbf{a}) = d\mathbf{f}(\mathbf{b}) \circ d\mathbf{g}(\mathbf{a}) \quad (1.25)$$

Dimostrazione. commentata □

Osservazione 91. In termini di matrici Jacobiane, la 1.25 assume una forma assai semplice come prodotto righe per colonne di $\mathbf{J}_{\mathbf{f}(\mathbf{b})}$ per $\mathbf{J}_{\mathbf{g}(\mathbf{a})}$

Corollario 1.5.5 (Jacobiana della composta). *La Jacobiana della composta si scrive come prodotto riga per colonna delle Jacobiane delle componenti \mathbf{f}, \mathbf{g} , valutate rispettivamente nei punti \mathbf{a} e \mathbf{b} :*

$$\underbrace{\mathbf{J}_{(\mathbf{f} \circ \mathbf{g})(\mathbf{a})}}_{p \times m} = \underbrace{\mathbf{J}_{\mathbf{f}(\mathbf{b})}}_{p \times n} \underbrace{\mathbf{J}_{\mathbf{g}(\mathbf{a})}}_{n \times m}$$

Esempio 1.5.2. Consideriamo le funzioni:

$$\begin{aligned} \mathbf{g} : \mathbb{R}^2 &\rightarrow \mathbb{R}^3, & \mathbf{g}(x, y) &\equiv (g_1(x, y), g_2(x, y), g_3(x, y)) = (x^2 - y, x + 1, x - y + 2) \\ \mathbf{f} : \mathbb{R}^3 &\rightarrow \mathbb{R}^3, & \mathbf{f}(x, y, z) &\equiv (f_1(x, y, z), f_2(x, y, z), f_3(x, y, z)) = (y, z - 1, xy) \end{aligned}$$

ovviamente sono differenziabili, in virtù delle osservazioni precedenti (sono funzioni polinomiali e quindi di classe $C^{(\infty)}$, nello specifico $\mathbf{g} \in C^{(\infty)}(\mathbb{R}^2)$, $\mathbf{f} \in C^{(\infty)}(\mathbb{R}^3)$). Sia $\mathbf{F} = \mathbf{f} \circ \mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$; risulta:

$$\begin{aligned} \mathbf{F}(x, y) &= \mathbf{f}(\mathbf{g}(x, y)) \\ &= \mathbf{f}(x^2 - y, x + 1, x - y + 2) \\ &= (x + 1, x - y + 1, x^3 + x^2 - xy - y) \\ &\equiv (F_1(x, y), F_2(x, y), F_3(x, y)) \end{aligned}$$

Posto $\mathbf{a} = (1, 2) \in \mathbb{R}^2$, si ha $\mathbf{b} = \mathbf{g}(\mathbf{a}) = (-1, 2, 1)$.

Per un generico punto $(x, y) \in \mathbb{R}^2$, la matrice Jacobiana di \mathbf{g} risulta:

$$\mathbf{J}_{\mathbf{g}(x, y)} = \begin{bmatrix} \frac{\partial g_1(x, y)}{\partial x} & \frac{\partial g_1(x, y)}{\partial y} \\ \frac{\partial g_2(x, y)}{\partial x} & \frac{\partial g_2(x, y)}{\partial y} \\ \frac{\partial g_3(x, y)}{\partial x} & \frac{\partial g_3(x, y)}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & -1 \\ 1 & 0 \\ 1 & -1 \end{bmatrix}$$

per cui, valutandola in $\mathbf{a} = (1, 2)$, si ha:

$$\mathbf{J}_{\mathbf{g}(1, 2)} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \\ 1 & -1 \end{bmatrix}$$

Per $(x, y, z) \in \mathbb{R}^3$, risulta

$$\mathbf{J}_{\mathbf{f}(x,y,z)} = \begin{bmatrix} \frac{\partial f_1(x,y,z)}{\partial x} & \frac{\partial f_1(x,y,z)}{\partial y} & \frac{\partial f_1(x,y,z)}{\partial z} \\ \frac{\partial f_2(x,y,z)}{\partial x} & \frac{\partial f_2(x,y,z)}{\partial y} & \frac{\partial f_2(x,y,z)}{\partial z} \\ \frac{\partial f_3(x,y,z)}{\partial x} & \frac{\partial f_3(x,y,z)}{\partial y} & \frac{\partial f_3(x,y,z)}{\partial z} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ y & x & 0 \end{bmatrix}$$

per cui, valutandola nel punto $\mathbf{b} = (-1, 2, 1)$, si ha:

$$\mathbf{J}_{\mathbf{f}(-1,2,1)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix}$$

Infine la Jacobiana della funzione composta \mathbf{F} per un generico $(x, y) \in \mathbb{R}^2$, risulta

$$\mathbf{J}_{\mathbf{F}(x,y)} = \begin{bmatrix} \frac{\partial F_1(x,y)}{\partial x} & \frac{\partial F_1(x,y)}{\partial y} \\ \frac{\partial F_2(x,y)}{\partial x} & \frac{\partial F_2(x,y)}{\partial y} \\ \frac{\partial F_3(x,y)}{\partial x} & \frac{\partial F_3(x,y)}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 3x^2 + 2x - y & -x - 1 \end{bmatrix}$$

per cui, valutando la matrice Jacobiana nel punto $\mathbf{a} = (1, 2)$, risulta

$$\mathbf{J}_{\mathbf{F}(1,2)} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 3 & -2 \end{bmatrix}$$

Come asserito dal Teorema, si ha:

$$\mathbf{J}_{\mathbf{f}(\mathbf{b})}\mathbf{J}_{\mathbf{g}(\mathbf{a})} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 3 & -2 \end{bmatrix} = \mathbf{J}_{\mathbf{F}(1,2)} = \mathbf{J}_{\mathbf{F}(\mathbf{a})}$$

1.6 Ottimizzazione vincolata

1.6.1 Teorema delle funzioni implicite

Teorema 1.6.1 (Teorema delle funzioni implicite (Dini generalizzato)). *Sia:*

- $\mathbf{f} : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{n-r}$ una applicazione di classe $C^{(1)}(A)$
- $Z_{\mathbf{f}} \subseteq \mathbb{R}^n$ il luogo degli zeri di \mathbf{f} , ossia l'insieme di valori di \mathbb{R}^n attraverso i quali \mathbf{f} produce il valore nullo $\mathbf{0} \in \mathbb{R}^{n-r}$:

$$Z_{\mathbf{f}} = \{(x_1, \dots, x_r, y_1, \dots, y_{n-r}) \in \mathbb{R}^n : \mathbf{f}(x_1, \dots, x_r, y_1, \dots, y_{n-r}) = \mathbf{0} \in \mathbb{R}^{n-r}\}$$

- considerando un singolo elemento appartenente a $Z_{\mathbf{f}}$, pensiamone le componenti spezzate in due e chiamiamole autonomamente come $\mathbf{x}_0 = (x_1, \dots, x_r) \in \mathbb{R}^r$ e $\mathbf{y}_0 = (x_1, \dots, x_{n-r}) \in \mathbb{R}^{n-r}$. Sarà sempre che $(\mathbf{x}_0, \mathbf{y}_0) \in \mathbb{R}^n$ e $\mathbf{f}((\mathbf{x}_0, \mathbf{y}_0)) = \mathbf{0}$ (facendo l'append, diciamo)

- la Jacobiana di \mathbf{f} nel punto $(\mathbf{x}_0, \mathbf{y}_0)$ è una matrice $(n-r) \times n$ composta da due parti: le prime r colonne (formanti una sottomatrice $(n-r) \times r$) riguardano le derivate parziali rispetto agli elementi di \mathbf{x}_0 le seconde $n-r$ (la sottomatrice quadrata $(n-r) \times (n-r)$ a destra) gli elementi di \mathbf{y}_0

$$J_{\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0, \mathbf{y}_0) & \dots & \frac{\partial f_1}{\partial x_r}(\mathbf{x}_0, \mathbf{y}_0) & \frac{\partial f_1}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \dots & \frac{\partial f_1}{\partial y_{n-r}}(\mathbf{x}_0, \mathbf{y}_0) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial f_{n-r}}{\partial x_1}(\mathbf{x}_0, \mathbf{y}_0) & \dots & \frac{\partial f_{n-r}}{\partial x_r}(\mathbf{x}_0, \mathbf{y}_0) & \frac{\partial f_{n-r}}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \dots & \frac{\partial f_{n-r}}{\partial y_{n-r}}(\mathbf{x}_0, \mathbf{y}_0) \end{bmatrix}$$

Se per ipotesi, focalizzandosi sulla sottomatrice quadrata $(n-r) \times (n-r)$ a destra delle \mathbf{y}_0 , si ha che è invertibile, quindi determinante non nullo

$$\det \begin{bmatrix} \frac{\partial f_1}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \dots & \frac{\partial f_1}{\partial y_{n-r}}(\mathbf{x}_0, \mathbf{y}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{n-r}}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \dots & \frac{\partial f_{n-r}}{\partial y_{n-r}}(\mathbf{x}_0, \mathbf{y}_0) \end{bmatrix} \neq 0$$

oppure avente rango $= n-r$, allora:

- esistono due intorni $I(\mathbf{x}_0, \delta_1) \subseteq \mathbb{R}^r$ e $I(\mathbf{y}_0, \delta_2) \subseteq \mathbb{R}^{n-r}$ tali che, per qualsiasi $\mathbf{x}' = (x'_1, \dots, x'_r) \in I(\mathbf{x}_0, \delta_1)$, esiste uno e un solo $\mathbf{y}' = (y'_1, \dots, y'_{n-r}) \in I(\mathbf{y}_0, \delta_2)$ per i quali $\mathbf{f}(\mathbf{x}', \mathbf{y}') = \mathbf{f}(x'_1, \dots, x'_r, y'_1, \dots, y'_{n-r}) = \mathbf{0} \in \mathbb{R}^{n-r}$, ossia con $(\mathbf{x}'_0, \mathbf{y}'_0) \in Z_{\mathbf{f}}$.
- Esiste pertanto la funzione $\varphi : I(\mathbf{x}_0, \delta_1) \subseteq \mathbb{R}^r \rightarrow \mathbb{R}^{n-r}$ definita come

$$\mathbf{y}' = \varphi(\mathbf{x}')$$

e scalarmente (guardando elemento per elemento del vettore)

$$\begin{aligned} y'_1 &= \varphi_1(x'_1, \dots, x'_r) \\ &\dots \\ y'_{n-r} &= \varphi_{n-r}(x'_1, \dots, x'_r) \end{aligned}$$

- inoltre $\varphi_1, \dots, \varphi_{n-r}$ sono ancora di classe C^1

Osservazione 92. Per $n=2, r=1$, si ottiene, come caso particolare, il teorema di Dini.

Esempio 1.6.1. Consideriamo (ponendo $n=3, r=1$) la funzione $\mathbf{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definita da

$$\begin{aligned} \mathbf{f}(x, y, z) &= (f_1(x, y, z), f_2(x, y, z)) \\ f_1(x, y, z) &= x^2 + y^2 + z^2 - 4 \\ f_2(x, y, z) &= z - 1 \end{aligned}$$

L'insieme

$$Z_{\mathbf{f}} = \{(x, y, z) \in \mathbb{R}^3 : \mathbf{f}(x, y, z) = (f_1(x, y, z), f_2(x, y, z)) = \mathbf{0}\}$$

è dato dall'intersezione degli insiemi dove si ha lo zero nelle due componenti:

$$\begin{cases} x^2 + y^2 + z^2 - 4 = 0 \\ z - 1 = 0 \end{cases} \quad \begin{cases} x^2 + y^2 + z^2 = 4 \\ z = 1 \end{cases} \quad \begin{cases} x^2 + y^2 = 3 \\ z = 1 \end{cases}$$

e consiste geometricamente nella circonferenza (in \mathbb{R}^3) ottenuta intersecando la sfera (centrata nell'origine) di raggio 2 (avente equazione $x^2 + y^2 + z^2 = 4$) con il piano parallelo al piano xy di altezza 1 (avente equazione $z = 1$). Quindi è la circonferenza di raggio $\sqrt{3}$ riportata, sul piano $z = 1$.

Fissiamo il punto $\mathbf{x}_0 = (1, \sqrt{2}, 1) \in Z_{\mathbf{f}}$ (come si può verificare sostituendo le coordinate nel sistema). La Jacobiana valutata rispetto alle sole y, z di \mathbf{f} è

$$\begin{bmatrix} \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{bmatrix} = \begin{bmatrix} 2y & 2z \\ 0 & 1 \end{bmatrix}$$

Se valutata in \mathbf{x}_0 , ha determinante

$$\det \begin{bmatrix} \frac{\partial f_1(\mathbf{x}_0)}{\partial y} & \frac{\partial f_1(\mathbf{x}_0)}{\partial z} \\ \frac{\partial f_2(\mathbf{x}_0)}{\partial y} & \frac{\partial f_2(\mathbf{x}_0)}{\partial z} \end{bmatrix} = \det \begin{bmatrix} 2\sqrt{2} & 2 \\ 0 & 1 \end{bmatrix} = 2\sqrt{2} \neq 0$$

Perciò a norma del teorema, possiamo esprimere localmente (intorno ad $\mathbf{x}_0 = (1, \sqrt{2}, 1)$) la circonferenza $Z_{\mathbf{f}}$ come grafico di una funzione $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^2$, ossia “esplicitare” (y, z) come funzione di x , ovvero scrivere

$$y = g_1(x), \quad z = g_2(x)$$

Geometricamente, stiamo dicendo che, se ragioniamo localmente, dato un punto della circonferenza $Z_{\mathbf{f}}$ vicino ad \mathbf{x}_0 , le sue coordinate y, z sono univocamente determinate una volta che sia nota la sua coordinata x . Nel nostro caso, si vede subito che risulta:

$$z = g_2(x) = 1 \tag{1.26}$$

$$y = g_1(x) = \sqrt{3 - x^2} \tag{1.27}$$

con 1.26 banale e 1.27 dovuta al fatto che, se $y = \sqrt{2}$ ci muoviamo nella parte positiva (rispetto a y), del piano avente $z = 1$.

In conclusione, l'equazione $\mathbf{f}(x, y, z) = \mathbf{0}$ determina (localmente, per $\mathbf{x}_0 = (1, \sqrt{2}, 1)$) la funzione $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^2$, definita come $\mathbf{g}(x) = (\sqrt{3 - x^2}, 1)$.

1.6.2 Varietà in \mathbb{R}^n

1.6.2.1 Definizione di varietà. Punti regolari

Osservazione 93. L'idea della varietà è (il modo moderno di) descrivere curve o punti nello spazio mediante equazioni, specificando che la definizione/descrizione è di carattere locale.

Definizione 1.6.1 (Varietà di \mathbb{R}^n di dimensione r e classe $C^{(k)}$). Sia $V \subseteq \mathbb{R}^n$: V è così chiamata se per ogni $\mathbf{x}_0 \in V$ esistono

1. un suo intorno $I(\mathbf{x}_0, \delta) \subseteq \mathbb{R}^n$;
2. una funzione $\mathbf{f} : I(\mathbf{x}_0, \delta) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{n-r}$ di classe $C^{(k)}$ ($k \geq 1$);

tali che:

- 1.

$$V \cap I(\mathbf{x}_0, \delta) = \{\mathbf{x} \in I(\mathbf{x}_0, \delta) \subseteq \mathbb{R}^n : \mathbf{f}(\mathbf{x}) = \mathbf{0} \in \mathbb{R}^{n-r}\}$$

Questa prima è una condizione “definente” la varietà: stiamo dicendo che un dato insieme di punti, se lo consideriamo localmente può essere descritto da un sistema di $n - r$ equazioni in n incognite, rappresentato compattamente mediante $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, o estesamente mediante il sistema

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ \dots \\ f_{n-r}(x_1, \dots, x_n) = 0 \end{cases}$$

2. la Jacobiana $(n - r) \times n$ di \mathbf{f} nel punto \mathbf{x}_0 :

$$\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)} = \begin{bmatrix} \frac{\partial f_1(\mathbf{x}_0)}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x}_0)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{n-r}(\mathbf{x}_0)}{\partial x_1} & \dots & \frac{\partial f_{n-r}(\mathbf{x}_0)}{\partial x_n} \end{bmatrix}$$

ha rango $n - r$.

Poiché la matrice Jacobiana di \mathbf{f} ha rango $n - r$, ciò significa che esisteranno $n - r$ indici $i_1 < \dots < i_{n-r}$ (associati ad altrettante variabili, che per semplicità supporremo essere le ultime, ossia $i_1 = r + 1, \dots, i_{n-r} = n$): tali che

$$\det \begin{bmatrix} \frac{\partial f_1(\mathbf{x}_0)}{\partial x_{r+1}} & \dots & \frac{\partial f_1(\mathbf{x}_0)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_{n-r}(\mathbf{x}_0)}{\partial x_{r+1}} & \dots & \frac{\partial f_{n-r}(\mathbf{x}_0)}{\partial x_n} \end{bmatrix} \neq 0$$

Questa seconda è una condizione di regolarità che richiediamo affinché valga il *teorema delle funzioni implicite* $\forall \mathbf{x}_0 \in V$; i punti $\mathbf{x}' \in V \cap I(\mathbf{x}_0, \delta)$ potranno essere espressi come punti in cui $n - r$ coordinate y'_1, \dots, y'_{n-r} variano come funzioni dipendenti di r coordinate x'_1, \dots, x'_r indipendenti, attraverso le funzioni $\varphi_1, \dots, \varphi_{n-r}$

$$\begin{aligned} y'_1 &= \varphi_1(x'_1, \dots, x'_r) \\ &\dots \\ y'_{n-r} &= \varphi_{n-r}(x'_1, \dots, x'_r) \end{aligned}$$

i punti di V saranno quelli rappresentabili *localmente* (in un opportuno intorno di \mathbf{x}_0) mediante una funzione $\varphi : \mathbb{R}^r \rightarrow \mathbb{R}^{n-r}$ come:

$$(x_1, \dots, x_r, \varphi_1(x_1, \dots, x_r), \dots, \varphi_{n-r}(x_1, \dots, x_r))$$

cioè al variare “liberamente” delle r coordinate x_1, \dots, x_r con le coordinate x_{r+1}, \dots, x_n dipendenti dalle x_1, \dots, x_r tramite il legame funzionale φ_i , ($i = r+1, \dots, n$) con φ_i di classe $C^{(k)}$ per ogni i .

Osservazione 94. Dire che abbiamo una varietà equivale a dire di avere un insieme di punti che ammetta una descrizione locale in termini di luogo di zeri. Si dice che $\mathbf{f} \equiv (f_1, \dots, f_{n-r})$ è la funzione che fornisce (“localmente”, vicino ad \mathbf{x}_0) le “equazioni” della varietà V .

Osservazione 95. Il fatto che i punti di V siano descritti localmente dal variare di r coordinate x_1, \dots, x_r , giustifica la terminologia “di dimensione r ” per la varietà V . In altre parole possiamo selezionare $n-r$ coordinate ed esprimerle come funzione di r lasciate libere.

Osservazione 96. Per la funzione \mathbf{f} siamo nella condizione analoga al teorema di Dini; es $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ si tratta di un sistema di una sola equazione in due variabili. Se la Jacobiana di f (che è il gradiente trasposto, composto da $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$) ha rango massimo (qui $n-r=1$), significa che almeno una delle derivate parziali è diversa da 0, quindi che possiamo applicare Dini.

Definizione 1.6.2 (Punti regolari e varietà). Dato un sottoinsieme $V \subseteq \mathbb{R}^n$, sono chiamati così i punti \mathbf{x}_0 soddisfacenti le precedenti condizioni. Di conseguenza, in questo linguaggio, una varietà è un sottoinsieme $V \subset \mathbb{R}^n$ i cui punti risultino tutti regolari.

Osservazione 97. Geometricamente la seconda condizione, quella della Jacobiana, fa sì che stiamo parlando di un oggetto liscio/regolare. Due cose non possono succedere: l’oggetto non ha ne cuspidi, ne nodi:

- una 1-varietà in \mathbb{R}^2 con nodi corrisponde al simbolo della lotta contro AIDS, mentre una cuspid è una cuspid;
- una 2-varietà in \mathbb{R}^3 con nodo è un foglio di carta arrotolato, mentre con una cuspid è un foglio piegato in mezzo tipo libro

Se regolare abbiamo tangenti, ossia uno *spazio tangente*

NB: Questo è un tentativo di dimostrazione

Esempio 1.6.2 (Circonferenza unitaria come varietà). L’insieme V ,

$$V = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 - 1 = 0\}$$

ossia la circonferenza unitaria centrata in $(0, 0)$, è una varietà di \mathbb{R}^2 di dimensione 1.

Infatti considerando $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da $f(x, y) = x^2 + y^2 - 1$, la Jacobiana di f in un generico punto è il gradiente trasposto

$$\mathbf{J}_f = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 2y \end{bmatrix}$$

e valutandola in un qualsivoglia $\mathbf{x}_0 = (x_0, y_0) \in V$

$$\mathbf{J}_{f(\mathbf{x}_0)} = [2x_0 \quad 2y_0]$$

Dunque si ha che $\mathbf{J}_{f(\mathbf{x}_0)} = (0, 0) \iff (x_0, y_0) = (0, 0)$ (e in tal punto il rango è nullo e non pari a 1); ma $(0, 0) \notin V$, e dunque in corrispondenza di tutti i punti di V la Jacobiana valutata ha rango 1, ossia tali punti sono regolari. Ossia V è una varietà.

Esempio 1.6.3 (Sfera unitaria come varietà). L'insieme

$$V = \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = x^2 + y^2 + z^2 - 1 = 0\}$$

ossia la sfera di raggio 1 centrata nell'origine $(0, 0, 0)$ è una varietà di \mathbb{R}^3 di dimensione 2 (tutti i suoi punti sono regolari).

Il procedimento è lo stesso di quello della circonferenza con solamente una modifica marginale alla Jacobiana nel punto generico

$$\mathbf{J}_{f(\mathbf{x}_0)} = [2x_0 \quad 2y_0 \quad 2z_0]$$

Questa ha rango 1 in tutti i punti $\mathbf{x}_0 \in V$ e dunque la varietà è di dimensione $3 - 1 = 2$.

Esempio 1.6.4 (Varietà: esempio in classe). Sia $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $F \equiv (F_1, F_2)$ con

$$\begin{aligned} F_1(x, y, z) &= x^2 + y^2 + z^2 - 4 \\ F_2(x, y, z) &= z - 1 \end{aligned}$$

da cui

$$V = \{(x, y, z) \in \mathbb{R}^3 : F(x, y, z) = \mathbf{0} \in \mathbb{R}^2\}$$

Ci chiediamo se sia regolare. La Jacobiana è

$$\mathbf{J}_f = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{bmatrix} = \begin{bmatrix} 2x & 2y & 2z \\ 0 & 0 & 1 \end{bmatrix}$$

Si può avere rango non massimo (2) solo se le righe si presentano come segue

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \vee \begin{bmatrix} 0 & 0 & k \\ 0 & 0 & 1 \end{bmatrix}$$

Nel primo caso ciò avviene se $\mathbf{x}_0 = (0, 0, 0)$, ma lo escludiamo non appartenendo alla varietà V ; il secondo si verifica se $x = 0, y = 0$ e per quanto riguarda z si abbia, dato $k = z/2$, che $(0, 0, k) \in V$, per averlo deve essere congiuntamente per le due componenti

$$\begin{cases} k^2 - 4 = 0 \\ k - 1 = 0 \end{cases} \quad \begin{cases} k = \pm 2 \\ k = -1 \end{cases}$$

NB: Lui dice di dimensione 1 ma dovrebbe essere un refuso ...

Dato che non vi è nessun valore che soddisfa entrambe le equazioni del sistema anche $(0, 0, k)$ non appartiene alla varietà. Pertanto il rango è 2 e V è varietà di \mathbb{R}^3 di dimensione 1.

Geometricamente consiste nella circonferenza determinata dall'incrocio tra la sfera centrata nell'origine e di raggio 2 e del piano $z = 1$.

Vogliamo descrivere localmente tale intersezione in prossimità di $\mathbf{x}_a = (\sqrt{3}, 0, 1)$ (appartenente alla varietà, come si può verificare): la jacobiana valutata nel punto è

$$\begin{bmatrix} 2\sqrt{3} & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

La colonna tutta a zero viene utilizzata per scegliere la variabile indipendente, in questo caso y , le altre sono dipendenti: scriviamo

$$z = \varphi_1(y) = 1$$

per quanto riguarda la $x = \varphi_2(y)$ invece, in prossimità di $(\sqrt{3}, 0, 1)$ (che sta sulla direzione est della circonferenza) possiamo scrivere

$$x^2 + y^2 + z^2 - 4 = 0$$

$$x^2 + y^2 + 1 - 4 = 0$$

$$x^2 = 3 - y^2$$

$$x = \sqrt{3 - y^2}$$

dove abbiamo tenuto la soluzione positiva (essendo di in prossimità di $\sqrt{3}$ per il punto).

Viceversa per $x_b = (0, \sqrt{3}, 1)$, la Jacobiana è

$$\begin{bmatrix} 0 & 2\sqrt{3} & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

per cui usiamo x come indipendente. Abbiamo sempre

$$z = \varphi_1(x) = 1$$

$$y = \varphi_2(x) = \sqrt{3 - x^2}$$

anche in quest'ultima prendendo la soluzione positiva della radice.

Esempio 1.6.5 (Iperbole). Si consideri l'insieme (figura 1.6)

$$V = \{(x, y) \in \mathbb{R}^2 : x^2 - y^2 = 1\}$$

Allora V è una varietà di dimensione 1 in \mathbb{R} . Poniamo $f : \mathbb{R}^2 \rightarrow \mathbb{R}^1 = \mathbb{R}^{2-1}$,

$$f(x, y) = x^2 - y^2 - 1$$

Se $\mathbf{x}_0 = (1, 0)$, risulta:

$$\mathbf{J}_{f(\mathbf{x}_0)} = \left(\frac{\partial f(\mathbf{x}_0)}{\partial x}, \frac{\partial f(\mathbf{x}_0)}{\partial y} \right) = (2, 0)$$

Dal teorema delle funzioni implicite, i punti di V , localmente intorno ad \mathbf{x}_0 (in un $I(\mathbf{x}_0, \delta)$), sono esprimibili esplicitando le coordinate x come funzione di y ; precisamente risulterà

$$x = g_1(y) = \sqrt{1 + y^2} \quad (\text{localmente})$$

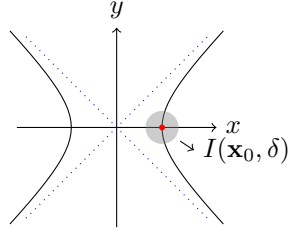


Figura 1.6: $x^2 - y^2 = 1$

Esempio 1.6.6 (Intersezione cerchio/retta). Considerando l'insieme (figura 1.7)

$$V = \{(x, y) \in \mathbb{R}^2 : (x^2 + y^2 - 1)(x + y) = 0\}$$

si ha che V non è una varietà. Si noti infatti che non esiste alcun intorno I di $\mathbf{x}_a = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, o analogamente di $\mathbf{x}_b = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$, per il quale i punti di $V \cap I$ soddisfano la condizione per cui la coordinata x è esplicitabile come funzione di y , o viceversa.

D'altronde potremmo scrivere

$$V = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$$

essendo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = (x^2 + y^2 - 1)(x + y)$ tuttavia

$$\mathbf{J}_{f(\mathbf{x}_a)} = \left(\frac{\partial f(\mathbf{x}_0)}{\partial x}, \frac{\partial f(\mathbf{x}_0)}{\partial y} \right) = (0, 0)$$

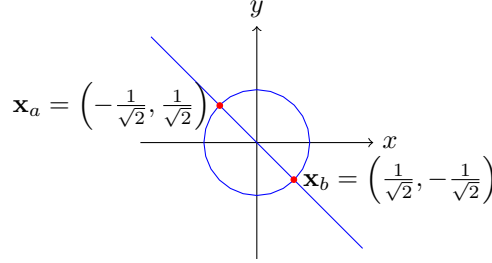
cioè la matrice jacobiana in \mathbf{x}_a ha rango 0.

Analogamente $\mathbf{J}_{f(\mathbf{x}_b)} = (0, 0)$, cioè $\mathbf{J}_{f(\mathbf{x}_b)}$ ha rango 0.

Esempio 1.6.7. Ancora nel caso

$$V = \{(x, y) \in \mathbb{R}^2 : (x^2 + y^2 - 1)(x + y) = 0\}$$

l'insieme $V' = V \setminus \{\mathbf{x}_a, \mathbf{x}_b\}$, con $\mathbf{x}_a = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ e $\mathbf{x}_b = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$, è effettivamente una varietà di \mathbb{R}^2 di dimensione 1.

Figura 1.7: $(x^2 + y^2 - 1)(x + y) = 0$

1.6.2.2 Varietà, spazi tangenti e spazi normali

Osservazione 98. Sia V una varietà di \mathbb{R}^n di dimensione r ($r \leq n$) e classe $C^{(k)}$ ($k \geq 1$).

Esempio 1.6.8. Ad esempio possiamo pensare alla sfera $x^2 + y^2 + z^2 = 1$ come 2-varietà nello spazio \mathbb{R}^3 . È liscia tipo mappamondo. Siamo interessati allo spazio tangente nel polo nord.

Definizione 1.6.3 (Vettore tangente ad una varietà). Un vettore $\mathbf{h} \in \mathbb{R}^n$ si dice *tangente* a V in \mathbf{x}_0 se e solo se esistono un reale positivo $\delta \in \mathbb{R}^+$ ed una curva regolare semplice aperta $\mathbf{r} : [-\delta, \delta] \rightarrow \mathbb{R}^n$ tale che

1. $\mathbf{r}([-\delta, \delta]) \subseteq V$, e $\mathbf{r}(0) = \mathbf{x}_0$;
2. $\mathbf{r}'(0) = \mathbf{h}$.

con, per ipotesi $\mathbf{h} \neq \mathbf{0}$ essendo la curva regolare.

Osservazione importante 15 (Interpretazione geometrica). Geometricamente, la condizione

1. significa che la curva \mathbf{r} “descrive una traiettoria” sulla varietà V “passante per \mathbf{x}_0 ” (in corrispondenza del valore $t = 0$)
2. significa che il vettore \mathbf{h} è proporzionale al “versore tangente” al sostegno di \mathbf{r} nel punto $\mathbf{r}(0) = \mathbf{x}_0$.

Esempio 1.6.9. Indicando con \mathbf{x}_0 il polo nord, da questi passano i meridiani (curve regolari) rappresentati \mathbf{r} (i quali risiedono sul mappamondo V e passano dal polo, interpretazione della prima condizione).

Per il vettore tangente al polo nord prendo la tangente al meridiano, nel polo nord. Se poi prendo nel polo nord tutti i tangenti ai vari meridiani si ha il piano tangente (che è uno spazio tangente).

Definizione 1.6.4 (Spazio tangente ad una varietà V nel punto \mathbf{x}_0). L'insieme

$$\mathbf{T}(\mathbf{x}_0) = \{\mathbf{h} \in \mathbb{R}^n : \mathbf{h} \text{ tangente a } V \text{ in } \mathbf{x}_0\} \cup \{\mathbf{0}\}$$

Osservazione 99. Questa definizione è chiara ma non operativa (occorrerebbe trovare tutte le curve passanti dal punto per generare lo spazio tangente.

Per $\mathbf{T}(\mathbf{x}_0)$ abbiamo bisogno di una definizione operativa, e la seguente lo è.

Teorema 1.6.2 (Spazio tangente ad una varietà V nel punto x_0 : teorema/definizione operativa). *Sia V una varietà di \mathbb{R}^n di dimensione r ($r < n$) e classe $C^{(k)}$, $k \geq 1$. Siano $\mathbf{x}_0 \in V$ ed $\mathbf{f} : I(\mathbf{x}_0, \delta) \subset \mathbb{R}^n \rightarrow \mathbb{R}^{n-r}$ di classe $C^{(k)}$, tali che:*

1. $V \cap I(\mathbf{x}_0, \delta) = \{\mathbf{x} \in I(\mathbf{x}_0, \delta) : \mathbf{f}(\mathbf{x}) = \mathbf{0}\}$.
2. la matrice Jacobiana $\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}$ ha rango $n - r$;

Allora

$$\mathbf{T}(\mathbf{x}_0) = \{\mathbf{h} \in \mathbb{R}^n : \mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}\mathbf{h} = \mathbf{0} \in \mathbb{R}^{n-r}\} = \text{Ker}(\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)})$$

ove Ker indica lo spazio nullo.

Dimostrazione. briniext pag 193 □

Osservazione 100. Dato che \mathbf{f} è (almeno) C^1 in \mathbf{x}_0 , è differenziabile ed ha la Jacobiana $\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}$. Si dimostra dunque che lo spazio tangente è il nocciolo della Jacobiana ⁹, ossia l'insieme di vettori \mathbf{h} che post-moltiplicati restituiscono il vettore nullo, ossia le soluzioni del sistema $\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}\mathbf{x} = \mathbf{0}$ (es da fare mediante eliminazione di Gauss).

Corollario 1.6.3. $\mathbf{T}(\mathbf{x}_0)$ è un sottospazio vettoriale di \mathbb{R}^n

Dimostrazione. Lo spazio nullo della Jacobiana è un sottospazio di \mathbb{R}^n , perché $\mathbf{0}$ vi appartiene e qualsiasi combinazione lineare di vettori, che attraverso la pre-moltiplicazione per Jacobiana risultano in $\mathbf{0}$, risulta anch'esso in $\mathbf{0}$ (essendo la pre-moltiplicazione per matrice Jacobiana un'applicazione lineare). □

Corollario 1.6.4. Si ha che $\dim \mathbf{T}(\mathbf{x}_0) = \dim (\text{Ker } \mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}) = r$

Dimostrazione. Dato dunque che la pre-moltiplicazione per matrice Jacobiana (ossia il differenziale) è una applicazione lineare tra spazi vettoriali (rispettivamente \mathbb{R}^n ed \mathbb{R}^m e dato che \mathbb{R}^n è finitamente generato, vale il teorema nullità+rango di algebra per cui:

$$n = \dim (\text{Ker } (\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)})) + \dim (\text{Im } (\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}))$$

Da cui

$$\begin{aligned} \dim (\text{Ker } \mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}) &= n - \dim (\text{Im } \mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}) \\ &= n - \text{rk } (\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}) \\ &= n - (n - r) \\ &= r \end{aligned}$$

Ossia basta una base di r elementi per generare lo spazio tangente. □

⁹Secondo Salce, per Brini è il nocciolo del differenziale $\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}\mathbf{h}$, ma basta intendersi.

Osservazione 101. Poiché la i -esima riga della matrice Jacobiana $\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}$ è il gradiente $\nabla f_i(\mathbf{x}_0)$ della i -esima componente scalare f_i della funzione \mathbf{f} ($i = 1, \dots, n - r$), il precedente risultato può essere riformulato come segue.

Corollario 1.6.5. *Lo spazio tangente $\mathbf{T}(\mathbf{x}_0)$ è lo spazio di tutti i vettori $\mathbf{h} \in \mathbb{R}^n$ tali che*

$$\langle \nabla f_i(\mathbf{x}_0), \mathbf{h} \rangle = 0, \quad \forall i = 1, \dots, n - r$$

cioè lo spazio dei vettori ortogonali ai gradienti delle $n - r$ componenti scalari di \mathbf{f} , valutati nel punto \mathbf{x}_0 .

Definizione 1.6.5 (Spazio normale alla varietà V nel punto \mathbf{x}_0). Sia V è una varietà di \mathbb{R}^n di dimensione r e classe $C^{(k)}$, $k \geq 1$. Si chiama così il *complemento ortogonale* (in \mathbb{R}^n) dello spazio tangente $\mathbf{T}(\mathbf{x}_0)$, cioè il sottospazio vettoriale:

$$\mathbf{N}(\mathbf{x}_0) = \mathbf{T}(\mathbf{x}_0)^\perp = \{\mathbf{u} \in \mathbb{R}^n : \langle \mathbf{u}, \mathbf{h} \rangle = 0, \text{ per ogni } \mathbf{h} \in \mathbf{T}(\mathbf{x}_0)\} \subset \mathbb{R}^n$$

Osservazione 102. Se V ha dimensione r , anche lo spazio tangente $\mathbf{T}(\mathbf{x}_0)$ ha dimensione r (come sottospazio vettoriale di \mathbb{R}^n): di conseguenza, dall'algebra lineare segue che lo spazio normale $\mathbf{N}(\mathbf{x}_0)$ ha dimensione $n - r$.

Corollario 1.6.6. *L'insieme $\{\nabla f_i(\mathbf{x}_0) : i = 1, \dots, n - r\}$ (composto dalle righe della matrice Jacobiana $\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}$) è una base dello spazio normale $\mathbf{N}(\mathbf{x}_0)$.*

Dimostrazione. Perché:

- lo spazio normale ha dimensione $n - r$;
- dal corollario 1.6.5 sappiamo che gli $n - r$ vettori $\nabla f_i(\mathbf{x}_0)$ (con $i = 1, \dots, n - r$), righe della Jacobiana, appartengono allo spazio normale $\mathbf{N}(\mathbf{x}_0)$;
- dato che la Jacobiana $\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)}$ ha rango massimo $n - r$, questi vettori sono linearmente indipendenti.

□

Osservazione 103. Pertanto per mezzo della Jacobiana riusciamo a determinare sia spazio normale (risolvendo un sistema) che lo spazio tangente (utilizzandone le righe per combinazioni lineari).

1.6.3 Punti critici vincolati e moltiplicatori di Lagrange

Osservazione importante 16 (Setup ottimizzazione con vincolo). Sia:

- una funzione $\Phi : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ di classe $C^{(1)}(A)$.
- V una varietà di \mathbb{R}^n di dimensione r ($r < n$) e classe $C^{(k)}$, $k \geq 1$, da intendersi come vincolo;
- $V \subset A \subseteq \mathbb{R}^n$ (A aperto);

Definizione 1.6.6 (Estremo relativo vincolato di Φ rispetto al “vincolo” V). Un massimo (risp. minimo) relativo della restrizione Φ/V di Φ a V : ossia esiste un intorno $I(\mathbf{x}_0, \delta)$ tale che $\Phi(\mathbf{x}_0) \geq \Phi(\mathbf{x})$, $\forall \mathbf{x} \in I(\mathbf{x}_0, \delta)$

Teorema 1.6.7 (Teorema dei moltiplicatori di Lagrange). *Se Φ ha estremo relativo vincolato in $\mathbf{x}_0 \in V$ (rispetto al vincolo V), allora esistono degli scalari $\lambda_1, \dots, \lambda_{n-r}$ tali che \mathbf{x}_0 sia punto critico per la funzione*

$$\Phi - \sum_{i=1}^{n-r} \lambda_i f_i$$

dove $\mathbf{f} : I(\mathbf{x}_0, \delta) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{n-r}$ è la funzione che fornisce (“localmente”, vicino \mathbf{x}_0) le “equazioni” della varietà V .

Dimostrazione. Sia \mathbf{h} un vettore tangente a V in \mathbf{x}_0 e sia $\mathbf{r} : [-\delta, \delta] \rightarrow \mathbb{R}^n$ una curva regolare semplice aperta tale che:

1. $\mathbf{r}([-\delta, \delta]) \subseteq V$ e $\mathbf{r}(0) = \mathbf{x}_0$;
2. $\mathbf{r}'(0) = \mathbf{h}$.

Posto $\omega(t) = \Phi(\mathbf{r}(t))$ (dunque $\omega : \mathbb{R} \rightarrow \mathbb{R}$), poiché $\mathbf{r}(t) \in V$ e Φ/V ha estremo relativo in \mathbf{x}_0 , ne segue che ω ha estremo relativo in $t = 0$, e quindi deve essere $\omega'(0) = 0$.

Risulta, d'altronde,

$$\omega'(0) \stackrel{(1)}{=} \sum_{i=1}^n \frac{\partial \Phi(\mathbf{r}(0))}{\partial x_i} \cdot \mathbf{r}'_i(0) = \left\langle \nabla \Phi(\mathbf{r}(0)), \mathbf{r}'(0) \right\rangle \stackrel{(2)}{=} \langle \nabla \Phi(\mathbf{x}_0), \mathbf{h} \rangle = 0$$

dove in (1) dove sfruttiamo la derivazione della funzione composta come in proposizione 1.4.12 e in (2) sostituiamo per le ipotesi poste sulla curva \mathbf{r} .

Si noti che quest'ultima relazione vale per ogni vettore tangente \mathbf{h} a V in \mathbf{x}_0 . Perciò $\nabla \Phi(\mathbf{x}_0) \in \mathbb{N}(x_0)$ ossia il gradiente appartiene allo spazio normale della varietà V nel punto \mathbf{x}_0 ; e, dato che questo è generato dai vettori della base $\nabla f_1(\mathbf{x}_0), \dots, \nabla f_{n-r}(\mathbf{x}_0)$, esistono degli scalari $\lambda_1, \dots, \lambda_{n-r}$ utili per generare $\nabla \Phi(\mathbf{x}_0)$ come combinazione lineare degli elementi della base:

$$\nabla \Phi(\mathbf{x}_0) = \sum_{i=1}^{n-r} \lambda_i \nabla f_i(\mathbf{x}_0)$$

Pertanto:¹⁰

$$\nabla \left(\Phi - \sum_{i=1}^{n-r} \lambda_i f_i \right) (\mathbf{x}_0) = \mathbf{0}$$

ossia \mathbf{x}_0 è punto critico per la funzione $\Phi - \sum_{i=1}^{n-r} \lambda_i f_i$ per il teorema di Luca. \square

¹⁰Qui sembra applichi qualcosa del genere $(f)'(a-b) = f'(a) - f'(b)$ al gradiente, ossia si può anche verificare che il gradiente mi pare sia una funzione lineare.

1.6.4 Esempi

Esempio 1.6.10. Sia $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\Phi(x, y) = x + y$ e sia

$$V = \{(x, y) \in \mathbb{R}^2 : f(x, y) = x^2 + y^2 - 1 = 0\}$$

la circonferenza unitaria centrata in $(0, 0)$, che come visto in esempio 1.6.2 è una varietà di \mathbb{R}^2 di dimensione 1.

I punti critici di V sono tutti e soli i punti per i quali $\nabla\Phi(x, y) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ è proporzionale a $\nabla f = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$; perciò deve aversi $x = y$ e quindi, dovendo tali punti appartenere a V si hanno due casi:

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \quad \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$$

Essendo V un compatto, in virtù del teorema di Weierstrass, questi due punti saranno rispettivamente uno punto di massimo (assoluto rispetto al vincolo V) e l'altro punto di minimo (assoluto rispetto al vincolo V).

Graficamente questi punti si trovano sull'intersezione tra il cilindro disegnato dalla circonferenza di cui sopra, per qualsiasi valore di z , e il piano $z = x + y$.

Esempio 1.6.11. Sia $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}$, $\Phi(x, y, z) = x + y + z^2$ e sia

$$V = \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = x^2 + y^2 + z^2 - 1 = 0\}$$

L'insieme V è una varietà di \mathbb{R}^3 di dimensione 2, come verificato in esempio 1.6.3.

I punti critici di V sono tutti e soli i punti per i quali $\nabla\Phi(x, y, z) = \begin{bmatrix} 1 \\ 1 \\ 2z \end{bmatrix}$ è

proporzionale a $\nabla f(x, y, z) = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}$; perciò deve risultare

$$\begin{bmatrix} 1 \\ 1 \\ 2z \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}$$

Ora:

- se $z \neq 0$ ne consegue che $\lambda = 1$ e, quindi, $x = \frac{1}{2}$, $y = \frac{1}{2}$; dovendo i punti appartenere a V , deve essere $z = \pm \frac{1}{\sqrt{2}}$.

Si hanno due punti critici

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

con moltiplicatore $\lambda = 1$.

- se $z = 0$, deve aversi $x^2 + y^2 = 1$, ma anche $x = y$; perciò $x = y = \pm \frac{1}{\sqrt{2}}$.
Si hanno altri due punti critici

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

con moltiplicatori $\lambda = \frac{\sqrt{2}}{2}$ e $\lambda = -\frac{\sqrt{2}}{2}$, rispettivamente.

Essendo V un compatto, in virtù del teorema di Weierstrass, la funzione Φ ammette punti di minimo (assoluto rispetto al vincolo V) e punti di massimo (assoluto rispetto al vincolo V).

Andando a valutare la funzione nei quattro punti critici trovati, si verifica immediatamente che il minimo assoluto si ottiene nel punto

$$\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right),$$

mentre il massimo assoluto si ottiene nei punti

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

Esempio 1.6.12. Sia $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}$, $\Phi(x, y, z) = x + y + z$ e sia $\mathbf{f} \equiv (f_1, f_2) : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definita da

$$\begin{cases} f_1(x, y, z) = x^2 + y^2 + z^2 - 4 \\ f_2(x, y, z) = x - 1 \end{cases}$$

Sia

$$V = \{(x, y, z) \in \mathbb{R}^3 : \mathbf{f}(x, y, z) = \mathbf{0}\} = \begin{cases} x^2 + y^2 + z^2 - 4 = 0 \\ x - 1 = 0 \end{cases}$$

L'insieme V è l'intersezione tra la sfera di raggio 2 descritta da f_1 e il piano $x = 1$ descritto da f_2 , ossia è una circonferenza, descritta dall'equazione;

$$y^2 + z^2 = 3$$

e posta sul piano $x = 1$. Detto insieme/circonferenza è dunque una varietà di \mathbb{R}^3 di dimensione 1 (tutti i suoi punti sono regolari, direi per quanto si è già visto nel caso della circonferenza unitaria classica, dove invece si aveva $z = 0$, invece che $x = 1$); volendo controllare, la Jacobiana

$$\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)} = \begin{bmatrix} 2x_0 & 2y_0 & 2z_0 \\ 1 & 0 & 0 \end{bmatrix}$$

ha rango 2 per qualsiasi $\mathbf{x}_0 \in V$, quindi la varietà V è di dimensione 1.

Si ha $\nabla \Phi(x, y, z) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\nabla f_1 = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}$, $\nabla f_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ e dobbiamo determinare i

punti di V tali che esistano $\lambda_1, \lambda_2 \in \mathbb{R}$ tali che

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Poiché il punto deve appartenere a V si ha $x = 1$. Inoltre da sopra deve essere $y = z$ per cui

$$y^2 + z^2 = 3 \iff 2y^2 = 3 \iff y = \pm \frac{\sqrt{3}}{\sqrt{2}} = z$$

quindi $y = z = \pm \frac{\sqrt{3}}{\sqrt{2}}$. In conclusione, i punti critici sono $\mathbf{x}_1 = \left(1, \frac{\sqrt{3}}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}}\right)$ e $\mathbf{x}_2 = \left(1, -\frac{\sqrt{3}}{\sqrt{2}}, -\frac{\sqrt{3}}{\sqrt{2}}\right)$.

Per andare a determinare i coefficienti λ_1, λ_2 , guardando alle coordinate in alto (di x), deve essere $\lambda_2 = 1 - 2\lambda_1$; determiniamo dunque prima λ_1 e poi λ_2 .

Per \mathbf{x}_1 si ha:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} 2 \cdot 1 \\ 2 \cdot \frac{\sqrt{3}}{\sqrt{2}} \\ 2 \cdot \frac{\sqrt{3}}{\sqrt{2}} \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

per cui si ha che $1 = \lambda_1 \cdot 2 \frac{\sqrt{3}}{\sqrt{2}}$ da cui $\lambda_1 = \frac{1}{\sqrt{2}\sqrt{3}}$ e pertanto $\lambda_2 = 1 - \frac{2}{\sqrt{2}\sqrt{3}} = 1 - \frac{\sqrt{2}}{\sqrt{3}}$. Similmente per \mathbf{x}_2 si ha

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} 2 \cdot 1 \\ 2 \cdot \left(-\frac{\sqrt{3}}{\sqrt{2}}\right) \\ 2 \cdot \left(-\frac{\sqrt{3}}{\sqrt{2}}\right) \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

da cui si conclude che $\lambda_1 = -\frac{1}{\sqrt{3}\sqrt{2}}$ e $\lambda_2 = 1 + \frac{2}{\sqrt{2}\sqrt{3}} = 1 + \frac{\sqrt{2}}{\sqrt{3}}$.

Essendo V un compatto, in virtù del teorema di Weierstrass, \mathbf{x}_1 e \mathbf{x}_2 saranno rispettivamente un punto di massimo (assoluto rispetto al vincolo V) e uno di minimo (assoluto rispetto al vincolo V).

Esempio 1.6.13. Sia $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\Phi(x, y) = xy$ e sia

$$V = \{(x, y) \in \mathbb{R}^2 : f(x, y) = x + 2y - 3 = 0\}$$

Risulta $\nabla \Phi(x, y) = \begin{bmatrix} y \\ x \end{bmatrix}$ e $\nabla f(x, y) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$; ne segue

$$y = \lambda, \quad x = 2\lambda$$

I punti critici sono perciò della forma $(x, y) = (2\lambda, \lambda)$; dovendo appartenere alla varietà V , si ha che $2\lambda + 2\lambda - 3 = 0$ e, quindi, $\lambda = \frac{3}{4}$.

In conclusione, l'unico punto critico è il punto $(\frac{3}{2}, \frac{3}{4})$, con moltiplicatore $\lambda = \frac{3}{4}$.

Esempio 1.6.14. Sia $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}$, $\Phi(x, y, z) = xy + z^2$ e sia $\mathbf{f} \equiv (f_1, f_2) : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definita da

$$\begin{cases} f_1(x, y, z) = x + y - 1 \\ f_2(x, y, z) = y^2 + z^2 - 1 \end{cases}$$

Sia

$$V = \{(x, y, z) \in \mathbb{R}^3 : \mathbf{f}(x, y, z) = \mathbf{0}\} = \begin{cases} x + y - 1 = 0 \\ y^2 + z^2 - 1 = 0 \end{cases}$$

Che è l'intersezione tra la retta $y = 1 - x$ sul piano $z = 0$ con il cilindro creato dalla circonferenza $y^2 + z^2 = 1$ (per qualsiasi valore di x).

L'insieme V è una varietà di \mathbb{R}^3 di dimensione 1: la Jacobiana per un generico $\mathbf{x}_0 = (x_0, y_0, z_0) \in V$ è

$$\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2y_0 & 2z_0 \end{bmatrix}$$

si ha che $\text{rango} = 1$ se $y_0 = 0 \wedge z_0 = 0$; ma ciò non può essere perché un punto $\mathbf{x}_0 = (x_0, 0, 0) \notin V$, perché non soddisferebbe la $y^2 + z^2 - 1 = 0$. Pertanto rango della matrice è 2, tutti i punti di V sono regolari e dunque V è una varietà di dimensione $3 - 2 = 1$.

Si ha $\nabla \Phi(x, y, z) = \begin{bmatrix} y \\ x \\ 2z \end{bmatrix}$, $\nabla f_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\nabla f_2 = \begin{bmatrix} 0 \\ 2y \\ 2z \end{bmatrix}$ e dobbiamo determinare i punti di V tali che esistano $\lambda_1, \lambda_2 \in \mathbb{R}$ tali che

$$\begin{bmatrix} y \\ x \\ 2z \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 2y \\ 2z \end{bmatrix}$$

Perciò $\lambda_1 = y$ e, se $z \neq 0$, $\lambda_2 = 1$, da cui $x = \lambda_1 + 2y = 3y$; dalla condizione $f_1(x, y, z) = 0$ evinciamo che $y = \frac{1}{4}$ e $x = \frac{3}{4}$.

Dalla condizione $f_2(x, y, z) = 0$ segue allora $z = \pm \frac{\sqrt{15}}{4}$.

In conclusione, vi sono due punti critici con $z \neq 0$, precisamente

$$\left(\frac{3}{4}, \frac{1}{4}, \frac{\sqrt{15}}{4} \right) \quad \left(\frac{3}{4}, \frac{1}{4}, -\frac{\sqrt{15}}{4} \right)$$

Se $z = 0$, allora $y = \pm 1$. Otteniamo allora altri due punti critici, precisamente $(2, -1, 0)$ con moltiplicatori $\lambda_1 = -1$ e $\lambda_2 = -\frac{3}{2}$, e $(0, 1, 0)$ con moltiplicatori $\lambda_1 = 1$ e $\lambda_2 = -\frac{1}{2}$.

Essendo V un compatto, in virtù del teorema di Weierstrass, la funzione Φ ammette punti di minimo (assoluto rispetto al vincolo V) e punti di massimo (assoluto rispetto al vincolo V).

Andando a valutare la funzione nei quattro punti critici trovati, si verifica immediatamente che il minimo assoluto si ottiene nel punto $(2, -1, 0)$, mentre il massimo assoluto si ottiene nei punti

$$\left(\frac{3}{4}, \frac{1}{4}, \frac{\sqrt{15}}{4} \right), \quad \left(\frac{3}{4}, \frac{1}{4}, -\frac{\sqrt{15}}{4} \right)$$

TODO: sviluppare per esteso qui

Esempio 1.6.15. Sia $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}$, $\Phi(x, y, z) = x^2 + y^2 + z^2$ e sia $\mathbf{f} \equiv (f_1, f_2) : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definita da

$$\begin{cases} f_1(x, y, z) = x^2 + y - 1 \\ f_2(x, y, z) = y + z - 1 \end{cases}$$

Sia

$$V = \{(x, y, z) \in \mathbb{R}^3 : \mathbf{f}(x, y, z) = \mathbf{0}\} = \begin{cases} x^2 + y - 1 = 0 \\ y + z - 1 = 0 \end{cases}$$

L'insieme V è una varietà di \mathbb{R}^3 di dimensione 1; la Jacobiana di un generico punto \mathbf{x}_0 è

$$\mathbf{J}_{\mathbf{f}(\mathbf{x}_0)} = \begin{bmatrix} 2x & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

che ha rango 2 (per qualsivoglia $\mathbf{x}_0 \in V$); quindi tutti i punti di V sono regolari e V è varietà di dimensione 1.

Si ha $\nabla \Phi(x, y, z) = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}$, $\nabla f_1 = \begin{bmatrix} 2x \\ 1 \\ 0 \end{bmatrix}$, $\nabla f_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ e dobbiamo determinare i punti di V tali che esistano $\lambda_1, \lambda_2 \in \mathbb{R}$ tali che

$$\begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} = \lambda_1 \begin{bmatrix} 2x \\ 1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Se $x \neq 0$, si ha chiaramente $\lambda_1 = 1$. Inoltre

$$2y = \lambda_1 + \lambda_2, \quad 2z = \lambda_2;$$

perciò

$$2y - 2z = 1, \quad y + z = 1,$$

da cui $z = \frac{1}{4}$ e $y = \frac{3}{4}$.

Poiché il punto deve appartenere a V , ne consegue che $x = \pm \frac{1}{2}$.

In conclusione, abbiamo due punti critici

$$\mathbf{x}_1 = \left(\frac{1}{2}, \frac{3}{4}, \frac{1}{4}\right), \quad \mathbf{x}_2 = \left(-\frac{1}{2}, \frac{3}{4}, \frac{1}{4}\right)$$

con moltiplicatori

$$\lambda_1 = 1, \quad \lambda_2 = \frac{1}{2}$$

Se $x = 0$, si ha $y = 1$ e $z = 0$. Perciò abbiamo un terzo punto critico

$$\mathbf{x}_3 = (0, 1, 0)$$

con moltiplicatori

$$\lambda_1 = 1, \quad \lambda_2 = 0.$$

Per finire i calcoli

$$\Phi(\mathbf{x}_1) = \frac{1}{4} + \frac{9}{16} + \frac{1}{16} = \frac{7}{8}$$

$$\Phi(\mathbf{x}_2) = \frac{1}{4} + \frac{9}{16} + \frac{1}{16} = \frac{7}{8}$$

$$\Phi(\mathbf{x}_3) = 1$$

Quindi \mathbf{x}_3 è massimo, $\mathbf{x}_1, \mathbf{x}_2$ minimi.