

Note del corso Complementi di Matematica

Fabio Bagagiolo*

October 8, 2009

1 Introduzione

1.1 Syllabus

Queste sono le note della prima parte del corso “Complementi di Analisi Matematica” per il secondo anno dei corsi professionalizzanti della Laurea Triennale in Ingegneria Civile, Ingegneria per l’Ambiente e il Territorio. In esse si introducono alcuni dei concetti classici che sono argomento dei secondi corsi di Analisi Matematica, ovvero lo studio di funzioni definite su domini di più variabili (continuità, derivabilità parziale, differenziabilità e studio dei loro massimi e minimi) e il calcolo dei loro integrali su domini del piano e dello spazio tridimensionale, in particolare: il calcolo dei volumi.

Lo scopo del corso “Complementi di Analisi Matematica” e quindi anche quello di queste note, è quello di fornire alla studentessa e allo studente gli strumenti necessari per poter formalizzare matematicamente in modo corretto e per poter affrontare in modo utile alla loro soluzione alcuni dei problemi che la professione futura potrà loro riservare.

Va da sé comunque che, per poter utilizzare al meglio uno strumento matematico, è necessario almeno saperlo inquadrare nel suo contesto teorico ed astratto. Pertanto in queste note, pur non puntando ad un apprendimento avanzato della teoria matematica sottostante, non si rinuncerà mai ad un linguaggio e ad una esposizione formalmente rigorosi. Il linguaggio formale e rigoroso della matematica (che poi è proprio alla base dell’universalità e della versatilità degli strumenti matematici) è quello degli insiemi e delle funzioni. Pertanto questa Introduzione continua con brevi richiami su tali concetti¹.

1.2 Sui concetti di insieme e di funzione.

Qui si richiama un po’ della teoria generale degli insiemi e delle funzioni. Si intende che la studentessa e lo studente abbiano già incontrato tali concetti e che, soprattutto, siano già familiari con gli esempi classici di insiemi numerici \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} , \mathbb{C} .

*Dipartimento di Matematica, Università di Trento, Povo, bagagiol@science.unitn.it

¹I richiami qui di teoria degli insiemi sono senz’altro utili anche per la seconda parte del corso, Teoria della Probabilità, non riportata in queste note

1.2.1 Il concetto di insieme.

Quello di insieme è un concetto primitivo, cioè non può essere spiegato senza usare la parola “insieme” o un suo sinonimo.

Definizione 1.1 Un insieme è una “raccolta” o una “collezione” di oggetti.

Solitamente indicheremo gli insiemi con lettere latine maiuscole e useremo invece le lettere latine minuscole per gli oggetti di cui essi sono composti. Inoltre tali oggetti verranno chiamati elementi.

Sia quindi A un insieme e a un suo elemento; allora scriveremo

$$a \in A,$$

che si legge a è un elemento dell'insieme A , oppure a appartiene ad A , oppure a sta in A .

Se invece a non è un elemento di A , allora scriveremo

$$a \notin A.$$

L'insieme privo di elementi si chiama insieme vuoto e si indica con il simbolo \emptyset .

Osservazione 1.2 È abbastanza evidente che, dati un elemento a ed un insieme A qualunque, allora accade una ed una sola delle seguenti cose: $a \in A$ o $a \notin A$. Cioè o a è un elemento di A , oppure a non è un elemento di A ; in particolare ambedue le cose non possono accadere contemporaneamente, né tantomeno può accadere che nessuna delle due cose sia vera.

Se a, b, c sono tutti gli elementi di un insieme A , allora scriveremo anche

$$A = \{a, b, c\}.$$

Se A è l'insieme di *tutti e soli* gli elementi che godono di una certa proprietà \mathcal{P} , allora useremo una delle seguenti notazioni

$$\begin{aligned} A &= \{a \mid \mathcal{P}\}, \\ A &= \{a : \mathcal{P}\}, \\ A &= \{a \text{ t.c. } \mathcal{P}\}, \end{aligned}$$

che si leggono A è l'insieme degli elementi a tali che per essi vale la proprietà \mathcal{P} (o tali che godono della proprietà \mathcal{P}). In questo caso la proprietà \mathcal{P} è caratterizzante per gli elementi di A (e quindi per l'insieme A stesso). Infatti, se $a \in A$, allora a gode della proprietà \mathcal{P} ; viceversa, se a gode della proprietà \mathcal{P} , allora $a \in A$.

Se A consiste esattamente di quegli elementi di un altro insieme C che godono della proprietà \mathcal{P} , allora scriveremo

$$A = \{a \in C \mid \mathcal{P}\}.$$

Esempio 1. Sia A l'insieme delle facoltà dell'Università di Trento, allora possiamo scrivere indifferentemente

$$A = \left\{ \begin{array}{l} \text{scienze MFN, ingegneria, lettere,} \\ \text{sociologia, economia, giurisprudenza, scienze cognitive} \end{array} \right\},$$

$$A = \left\{ a \mid a \text{ è una facoltà dell'Università di Trento} \right\}.$$

Notiamo che in questo caso la proprietà caratterizzante è “essere una facoltà dell'Università di Trento”.

Definizione 1.3 *Se A e B sono due insiemi tali che ogni elemento di B è anche elemento di A , allora si dice che B è un sottoinsieme di A .*

Se B è un sottoinsieme di A , allora si scrive

$$B \subseteq A,$$

che si legge B è un sottoinsieme di A , oppure B è incluso in A oppure A include B . La proprietà di B di essere un sottoinsieme di A si può anche esprimere nel seguente modo

$$b \in B \implies b \in A, \tag{1.1}$$

che si legge “ b appartiene a B ” implica “ B appartiene ad A ”, oppure se b è un elemento di B , allora b è anche un elemento di A .

Osservazione 1.4 *Nella prima possibile lettura di (1.1) sopra elencata, le virgolette stanno a significare che è un'implicazione tra affermazioni; cioè se l'affermazione “ b appartiene a B ” è vera, allora necessariamente è vera anche l'affermazione “ b appartiene ad A ”.*

Ovviamente, questa implicazione non dice nulla su quali proprietà possano avere gli elementi che non stanno in B , in particolare non impedisce il fatto che elementi che non stanno in B stiano comunque in A .

L'implicazione (1.1) può essere anche scritta in modo equivalente come

$$b \notin A \implies b \notin B, \tag{1.2}$$

che si legge se b non è un elemento di A , allora non è nemmeno un elemento di B . Come si fa a provare rigorosamente che (1.1) è equivalente a (1.2)? La procedura consiste in due passi: 1) si suppone che l'implicazione (1.1) sia vera e si dimostra che allora è vera anche l'implicazione (1.2); 2) si suppone che l'implicazione (1.2) è vera e si dimostra che allora è vera anche l'implicazione (1.1). Sviluppiamo questa procedura.

1) Supponiamo (1.1) vera. Per provare che (1.2) è vera, dobbiamo provare che ogniqualvolta b non è un elemento di A allora non è nemmeno un elemento di B . Proviamo questo per assurdo. Questo significa che supponiamo (1.2) falsa e mostriamo che questo ci porta ad un assurdo, da cui ne deduciamo che (1.2) deve essere necessariamente vera. Supponiamo quindi (1.2) falsa. Questo significa che ci deve essere almeno un caso in cui l'implicazione in (1.2) non è verificata. Questo significa che esiste un elemento b che non è un elemento di A ma che invece è un elemento di B . Quest'ultima affermazione si scrive in simboli

$$\exists b \text{ tale che } b \notin A \text{ e } b \in B, \text{ oppure } \exists b \in B \mid b \notin A.$$

Ma se b appartiene a B , allora essendo (1.1) vera per ipotesi, deve necessariamente essere vero anche $b \in A$. Questo è un assurdo perché abbiamo così un elemento che contemporaneamente sta in A e non sta in A (vedi Osservazione 1.2).

2) La dimostrazione è del tutto analoga a quella del punto 1) ed è lasciata per esercizio.

Dalla definizione di sottoinsieme, si ha immediatamente che, qualunque sia l'insieme A , A è sottoinsieme di se stesso.

Definizione 1.5 *Siano A e B due insiemi. Se B è un sottoinsieme di A ed esiste un elemento di $a \in A$ tale che $a \notin B$, allora B si dice un sottoinsieme proprio di A .*

Se B è un sottoinsieme proprio di A , allora scriviamo

$$B \subset A, \text{ oppure } B \subsetneq A.$$

Esempio 2. Sia B l'insieme delle facoltà “di collina” dell'Università di Trento. Allora

$$B = \{ \text{scienze MFN, ingegneria} \},$$

ed è un sottoinsieme proprio dell'insieme A dell'Esempio 1. Infatti, tutti gli elementi di B sono anche elementi di A , ma esistono elementi di A che non sono elementi di B (ad esempio *economia*). Possiamo anche scrivere

$$B = \left\{ a \in A \mid a \text{ è una facoltà “di collina”} \right\},$$

Definizione 1.6 *Due insiemi coincidono (ovvero sono lo stesso insieme) se hanno esattamente gli stessi elementi.*

Se due insiemi A e B coincidono, scriveremo

$$A = B.$$

Proposizione 1.7 (*Doppia inclusione*). *Due insiemi A e B coincidono se e soltanto se A è un sottoinsieme di B e B è un sottoinsieme di A . In simboli*

$$A = B \iff A \subseteq B \text{ e } B \subseteq A. \tag{1.3}$$

Dimostrazione. Per dimostrare un “se e solo se” dobbiamo provare due implicazione (la freccia in (1.3) è infatti una “doppia freccia”). Dobbiamo cioè provare che

$$\begin{aligned} i) \quad & A = B \implies A \subseteq B \text{ e } B \subseteq A, \\ ii) \quad & A \subseteq B \text{ e } B \subseteq A \implies A = B. \end{aligned}$$

i) Supponiamo vero che $A = B$ e dimostriamo che allora A è sottoinsieme di B e B è sottoinsieme di A .

Per definizione, $A = B$ significa che A e B hanno esattamente gli stessi elementi. Questo significa se a è elemento di A allora è anche un elemento di B e, viceversa, se b è un elemento di B allora è anche un elemento di A . Ma quanto appena detto è esattamente la definizione del fatto che $A \subseteq B$ e che $B \subseteq A$.

ii) Supponiamo che A sia sottoinsieme di B e che B sia sottoinsieme di A e proviamo che A e B coincidono.

Per provare che A e B coincidono ragioniamo per assurdo. Supponiamo che essi non coincidano

$$A \neq B.$$

Questo significa che almeno una delle seguenti due affermazioni è vera:

- a) esiste un elemento di A che non è un elemento di B ;
- b) esiste un elemento di B che non è un elemento di A .

Ma, per definizione di sottoinsieme, questo significa che almeno una delle due affermazioni seguenti è vera:

- a') A non è un sottoinsieme di B ;
- b') B non è un sottoinsieme di A .

E questo è un assurdo in quanto in contraddizione con l'ipotesi iniziale. □

La Proposizione 1.7, fornisce un metodo per provare che due insiemi sono uguali: bisogna provare che uno è sottoinsieme dell'altro e viceversa.

È evidente che, dato un insieme A , l'unico sottoinsieme non proprio di A è A stesso. Inoltre, presi una qualunque proprietà \mathcal{P} ed un qualunque insieme A , l'insieme

$$B = \left\{ a \in A \mid \mathcal{P} \right\},$$

è evidentemente un sottoinsieme di A . In particolare, può succedere che nessun elemento di A soddisfi alla proprietà \mathcal{P} e quindi che B sia l'insieme vuoto. Ne segue che l'insieme vuoto è un sottoinsieme di qualunque insieme.

Esempio 3. Se A è l'insieme delle facoltà dell'Università di Trento, definiamo

$$C = \left\{ a \in A \mid a \text{ è una facoltà "di mare"} \right\}.$$

È evidente che $C = \emptyset \subset A$.

1.2.2 Operazioni con gli insiemi.

Definizione 1.8 Dati due insiemi A e B , si definisce insieme unione di A e B (o semplicemente unione di A e B) e si denota con $A \cup B$, l'insieme degli elementi che stanno in A o in B (cioè in almeno uno dei due). In simboli

$$A \cup B = \{c \mid c \in A \text{ o } c \in B\}.$$

Si definisce insieme intersezione di A e B (o semplicemente intersezione di A e B) e si denota con $A \cap B$, l'insieme degli elementi che stanno in A e in B (cioè contemporaneamente in tutti e due). In simboli

$$A \cap B = \{c \mid c \in A \text{ e } c \in B\}.$$

Si definisce insieme differenza di A da B (o semplicemente differenza di A da B) e si denota con $A \setminus B$, l'insieme degli elementi di A che non stanno in B . In simboli

$$A \setminus B = \{a \in A \mid a \notin B\}.$$

Se $B \subseteq A$ allora si dice complementare di B in A l'insieme $A \setminus B$ e si denota con $C_A B$ (o semplicemente con $C B$ se non c'è ambiguità sullo "spazio ambiente" A).

Esempio 4. Siano $A = \{1, 3, 6, 7, 8, 9\}$ e $B = \{2, 3, 5, 6, 10\}$. Allora si ha

$$A \cup B = \{1, 2, 3, 5, 6, 7, 8, 9, 10\},$$

$$A \cap B = \{3, 6\},$$

$$A \setminus B = \{1, 7, 8, 9\}.$$

Proposizione 1.9 Dati tre insiemi qualunque A , B e C , si ha

$$i) \quad A \cap B \subseteq A, \quad A \cap B \subseteq B, \quad A \setminus B \subseteq A,$$

$$ii) \quad A \subseteq A \cup B, \quad B \subseteq A \cup B,$$

$$iii) \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

$$iv) \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

dove le parentesi hanno l'usuale significato di effettuare prima le operazioni dentro le parentesi e poi quelle esterne.

Dimostrazione. Lasciata per esercizio (ricorrere alle definizioni e, per dimostrare le uguaglianze, usare la doppia inclusione). \square

Osservazione 1.10 In generale, $A \setminus B \neq B \setminus A$. Nell'Esempio 4 infatti, $B \setminus A = \{2, 5, 10\}$, ben diverso da $A \setminus B$.

Osservazione 1.11 *Le proprietà iii) e iv) nella Proposizione 1.9 si chiamano, rispettivamente, proprietà distributiva dell'intersezione rispetto all'unione (rispetto all'intersezione).*

Osservazione 1.12 *Si noti l'uso delle congiunzioni “e” ed “o” fatto nelle argomentazioni precedenti.*

La congiunzione “e” è sempre forte, cioè significa che devono verificarsi contemporaneamente ambedue le proprietà che essa sta congiungendo. Spesso tale congiunzione viene sostituita da una virgola. Per esempio, l'insieme

$$\{c \mid c \in A, c \in B\},$$

è l'insieme degli elementi che stanno in A e in B contemporaneamente (è cioè $A \cap B$).

La congiunzione “o” è di solito debole, cioè significa che deve verificarsi almeno una delle due proprietà che essa sta congiungendo. Questo però non esclude che possano verificarsi ambedue contemporaneamente. Non è quindi una congiunzione esclusiva “o una o l'altra, ma mai tutte e due assieme”.

In matematica quindi la congiunzione “o” è quasi sempre debole. Nei pochi casi in cui essa dovesse essere forte, questo verrà esplicitamente detto, per esempio tramite l'uso della doppia congiunzione “o ... o ...”, “o ... oppure ...”. Questo ad esempio è già accaduto nell'Osservazione 1.2.

Definizione 1.13 *Dati due insiemi A e B , si definisce prodotto cartesiano di A per B e lo si denota con $A \times B$, l'insieme i cui elementi sono le coppie del tipo (a, b) dove il primo elemento sta in A e il secondo sta in B . In simboli*

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Se almeno uno tra A e B è l'insieme vuoto, allora si pone $A \times B = \emptyset$

Esempio 5. Siano $A = \{1, 2, 5\}$ e $B = \{0, 2\}$. Allora si ha

$$A \times B = \{(1, 0), (1, 2), (2, 0), (2, 2), (5, 0), (5, 2)\}.$$

Notiamo che $A \times B$ è in generale diverso da $B \times A$. Infatti nel nostro esempio si ha

$$B \times A = \{(0, 1), (0, 2), (0, 5), (2, 1), (2, 2), (2, 5)\} \neq A \times B.$$

Dati (a_1, b_1) e (a_2, b_2) in $A \times B$, essi sono lo stesso elemento se e soltanto se coincidono “componente per componente”, cioè se $a_1 = a_2$ e $b_1 = b_2$.

1.2.3 Il concetto di funzione tra insiemi.

Definizione 1.14 *Siano A e B due insiemi. Una funzione di A in B (o una mappa, o una applicazione) è una “legge” che ad ogni elemento di A associa uno ed uno solo elemento di B .*

Possiamo quindi dire che una funzione di A in B è una sorta di “scatola nera” o di “macchinetta” in cui possiamo introdurre un qualunque elemento $a \in A$ ed essa ci restituisce un elemento $b \in B$ (che in generale dipende da quale elemento di A abbiamo introdotto). La cosa fondamentale è che possiamo introdurre (sebbene uno alla volta) tutti gli elementi di A e, per ogni immissione, viene mandato fuori un solo elemento di B .

Cosa vieta la definizione di funzione? 1) vieta che sia possibile inserire nella macchinetta più di un elemento di A per volta; 2) vieta che ci sia un elemento di A che non possiamo introdurre nella macchinetta; 3) vieta che, una volta introdotto un elemento di A , la macchinetta restituisca più di un elemento di B ; 4) vieta che, una volta introdotto un elemento di A , la macchinetta non restituisca alcun elemento di B .

Che cosa non vieta la definizione di funzione? 1) non vieta che esistano due elementi distinti di A per i quali, se introdotti nella macchinetta, viene restituito il medesimo elemento di B ; 2) non vieta che esista un elemento di B che non viene mai restituito dalla macchinetta, qualunque sia l'elemento di A introdotto.

Le funzioni si indicano usualmente con lettere minuscole (latine o greche) di solito f , g , φ , ψ . Una funzione f di A in B , si indica in simboli nel seguente modo

$$\begin{aligned} f: A &\rightarrow B \\ a &\mapsto f(a), \end{aligned}$$

e $f(a)$ indica l'elemento di B che, tramite la funzione f , viene associato all'elemento $a \in A$. L'insieme A si chiama dominio della funzione f , mentre l'insieme B si chiama codominio della funzione f . L'elemento $f(a) \in B$ si chiama immagine di a tramite f . Il seguente sottoinsieme del codominio (denotato con $f(A)$)

$$f(A) = \{b \in B \mid \exists a \in A, f(a) = b\} \subseteq B,$$

si chiama immagine di A tramite f o anche immagine di f . Esso consta di tutti gli elementi di B che, tramite f , provengono da elementi di A . Alle volte l'immagine di f è anche denotata con il simbolo $\text{Im} f$ (da non confondersi con la parte immaginaria di un numero complesso).

Possiamo dire che una funzione è un “oggetto matematico” individuato da tre elementi: il dominio, il codominio e la legge.

Definizione 1.15 *Una funzione $f: A \rightarrow B$ si dice iniettiva se a elementi diversi di A corrispondono elementi diversi di B , in simboli se*

$$a_1, a_2 \in A, a_1 \neq a_2 \implies f(a_1) \neq f(a_2).$$

Si dice suriettiva, se ogni elemento di B è immagine, tramite f , di un elemento di A (cioè se ogni elemento di B “proviene” da un elemento di A). In simboli

$$\forall b \in B \exists a \in A \mid b = f(a).$$

Si dice biiettiva se è sia iniettiva che suriettiva.

Esempio 8. Consideriamo la funzione

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto x^2, \end{aligned}$$

(verificare per esercizio che è una funzione). Questa funzione non è né iniettiva né suriettiva. Per verificare che non è iniettiva, basta “esibire un controesempio”, cioè trovare due elementi diversi di \mathbb{R} che hanno lo stesso quadrato. I numeri 2 e -2 soddisfano alla richiesta. Per verificare che non è suriettiva basta trovare un elemento di \mathbb{R} che non è il quadrato di alcun elemento di \mathbb{R} . Il numero -2 soddisfa alla richiesta.

Consideriamo la somma su \mathbb{Z} . Se vista come funzione di $\mathbb{Z} \times \mathbb{Z}$ in \mathbb{Z} , allora essa è suriettiva ma non iniettiva. Per mostrare che è suriettiva dobbiamo far vedere che, per ogni numero intero $m \in \mathbb{Z}$, esistono due numeri interi $m_1, m_2 \in \mathbb{Z}$ tali che $m = m_1 + m_2$ (cioè m proviene, tramite la funzione somma, dalla coppia (m_1, m_2)). Fissato $m \in \mathbb{Z}$, la coppia $(m - 1, m + 1)$ soddisfa alla richiesta. Per mostrare che non è iniettiva, bisogna determinare due coppie distinte di numeri interi che hanno la stessa somma. Le coppie $(1, -1)$ e $(-1, 1)$ soddisfano alla richiesta.

Consideriamo la seguente funzione (verificare per esercizio che è una funzione)

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \times \mathbb{R} \\ x &\mapsto (x, x). \end{aligned}$$

Questa funzione è iniettiva ma non suriettiva (verificarlo per esercizio).

Consideriamo la funzione (verificare per esercizio che è una funzione)

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto 3x - 2. \end{aligned}$$

Questa funzione è biiettiva. Proviamo che è iniettiva. Supponiamo che $y_1, y_2 \in f(\mathbb{R})$ e che $y_1 = y_2$. Il fatto che y_1 e y_2 stiano nell’immagine di f significa che esistono due numeri reali x_1 e x_2 tali che

$$y_1 = 3x_1 - 2, \quad y_2 = 3x_2 - 2.$$

Dall’uguaglianza $y_1 = y_2$ si ha che

$$3x_1 - 2 = 3x_2 - 2 \implies x_1 = x_2,$$

e quindi si ha l'iniettività². Proviamo che è suriettiva. Dobbiamo far vedere che, qualunque sia il numero reale y , esiste un numero reale x tale che $y = f(x)$ e cioè che $y = 3x - 2$. Basta prendere $x = (y + 2)/3 \in \mathbb{R}$.

Le funzioni biettive sono importanti perchè sono invertibili. Infatti supponiamo di avere una funzione biettiva $f : A \rightarrow B$. Per ogni elemento $b \in B$ denotiamo con $f^{-1}(b)$ quell'unico elemento $a \in A$ tale che $f(a) = b$. Abbiamo dato una "buona definizione"? oppure la procedura sopra descritta nella costruzione di $f^{-1}(b)$ non è sempre possibile o crea dell'ambiguità? Vediamo se è sempre possibile, cioè se qualunque sia $b \in B$ possiamo scegliere un elemento $a \in A$ tale che $f(a) = b$. Ma questa è proprio la richiesta di suriettività che stiamo supponendo essere verificata. Potrebbe però succedere che ci sia dell'ambiguità nella scelta di a , cioè che, dato $b \in B$ esistano più di un elemento di A che viene mandato in B . Ma questo non è vero, cioè esiste un solo elemento $a \in A$ tale che $f(a) = b$, in quanto stiamo supponendo che la funzione sia iniettiva. Quindi abbiamo un'unica possibile scelta per $f^{-1}(b) \in A$, cioè non c'è ambiguità. Possiamo pertanto definire la funzione di B in A , detta inversa di f ,

$$\begin{aligned} f^{-1} : B &\rightarrow A \\ b &\mapsto f^{-1}(b), \end{aligned}$$

e, con le argomentazioni fatte sopra, si vede che è effettivamente una funzione.

Perchè si chiama l'inversa di f ? perchè se prendiamo un qualunque elemento $b \in B$, ad esso ci applichiamo f^{-1} ottenendo $f^{-1}(b) \in A$ e a quest'ultimo applichiamo f ottenendo $f(f^{-1}(b)) \in B$, ebbene riotteniamo b , cioè

$$f(f^{-1}(b)) = b \quad \forall b \in B. \tag{1.4}$$

Allo stesso modo si ha

$$f^{-1}(f(a)) = a \quad \forall a \in A. \tag{1.5}$$

1.3 Su altri prerequisiti

Il concetto di insieme e quello di funzione sono alla base di tutta la matematica moderna (potremmo dire dal 1900 in poi) ed è quindi per questo che essi sono stati esplicitamente richiamati qui sopra. È ovvio che non sono gli unici concetti che servono per proseguire nella lettura di queste note. Altri requisiti essenziali sono: numeri reali, calcolo differenziale ed integrale per funzioni di una variabile reale a valori reali (argomento di Analisi I) e l'algebra lineare e gli spazi \mathbb{R}^n con le loro proprietà geometriche e algebriche (argomento di Geometria e Algebra). Questi prerequisiti non verranno qui richiamati.

²Una definizione equivalente dell'iniettività di $f : A \rightarrow B$ è la seguente (provare l'equivalenza per esercizio): $y_1, y_2 \in B, y_1 = f(x_1), y_2 = f(x_2), y_1 = y_2 \implies x_1 = x_2$

2 Calcolo Differenziale per Funzioni di più Variabili

2.1 Distanza, topologia e convergenza in \mathbb{R}^n

Ricordiamo che con il simbolo \mathbb{R}^n indichiamo il prodotto cartesiano di \mathbb{R} (insieme dei numeri reali) per se stesso n volte. Gli elementi x di \mathbb{R}^n sono quindi n -uple ordinate di numeri reali (stringhe di n numeri reali) $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $x_i \in \mathbb{R}$ per ogni $i = 1, \dots, n$. Gli elementi x di \mathbb{R}^n possono essere pensati sia come “punti” dello spazio n -dimensionale \mathbb{R}^n , aventi “coordinate cartesiane” (x_1, \dots, x_n) sia come “vettori” (freccie spiccate dall’origine con vertice nel punto x). Due elementi $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ sono uguali (cioè sono lo stesso elemento) se e soltanto se sono uguali “coordinata per coordinata” o “componente per componente”, cioè se $x_i = y_i$ per ogni $i = 1, \dots, n$. Su \mathbb{R}^n sono note le operazioni di “spazio vettoriale” (somma e moltiplicazione per uno scalare) e le loro proprietà.

Ricordiamo infine l’usuale identificazione (tramite le coordinate cartesiane) di \mathbb{R}^2 con il piano euclideo e di \mathbb{R}^3 con lo spazio euclideo tridimensionale (quello “fisico”).

Definizione 2.1 *Siano dati due punti $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ di \mathbb{R}^n , si dice distanza di x da y , e si indica con $\|x - y\|_{\mathbb{R}^n}$, la quantità non negativa*

$$\|x - y\|_{\mathbb{R}^n} := \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}.$$

Nei casi “fisici” $n = 2$ o $n = 3$ (piano o spazio 3D) la distanza appena definita altro non è che la lunghezza del segmento che congiunge il punto x con il punto y (Teorema di Pitagora). Se invece $n = 1$, allora si ritrova la usuale distanza di due punti sulla retta reale.

Definizione 2.2 *Dato un punto $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, si dice norma di x , e si indica con $\|x\|_{\mathbb{R}^n}$, la distanza di x dall’origine $O = (0, \dots, 0)$*

$$\|x\|_{\mathbb{R}^n} := \|x - O\|_{\mathbb{R}^n} = \sqrt{x_1^2 + \dots + x_n^2}.$$

Se $n = 1$, allora la norma altro non è che il valore assoluto in \mathbb{R} , che denotiamo con $|\cdot|$.

Notiamo che la distanza tra due punti coincide con la norma della loro differenza vettoriale (differenza componente per componente)

Proposizione 2.3 *Proprietà della distanza:*

- i) $\|x - x\|_{\mathbb{R}^n} = 0$ per ogni $x \in \mathbb{R}^n$ (un punto dista zero da se stesso),*
- ii) $\|x - y\|_{\mathbb{R}^n} = 0 \iff x = y$ (se due punti hanno distanza nulla tra loro, allora sono uguali),*
- iii) $\|x - y\|_{\mathbb{R}^n} = \|y - x\|_{\mathbb{R}^n}$ per ogni $x, y \in \mathbb{R}^n$ (la distanza di x da y è uguale alla distanza di y da x),*
- iv) $\|x - y\|_{\mathbb{R}^n} \leq \|x - z\|_{\mathbb{R}^n} + \|z - y\|_{\mathbb{R}^n}$ per ogni $x, y, z \in \mathbb{R}^n$ (disuguaglianza triangolare).*

Proprietà della norma

- a) $\|x\| \geq 0$ per ogni $x \in \mathbb{R}^n$,
- b) $\|x\|_{\mathbb{R}^n} = 0 \iff x = O$ (l'unico punto con norma nulla è l'origine),
- c) $\|x + y\|_{\mathbb{R}^n} \leq \|x\|_{\mathbb{R}^n} + \|y\|_{\mathbb{R}^n}$ per ogni $x, y \in \mathbb{R}^n$ (subadditività),
- d) $\|\lambda x\|_{\mathbb{R}^n} = |\lambda| \|x\|_{\mathbb{R}^n}$ per ogni $x \in \mathbb{R}^n, \lambda \in \mathbb{R}$ (positiva omogeneità).

Definizione 2.4 Dati $x \in \mathbb{R}^n$ e $r > 0$ si dice palla aperta centrata in x di raggio r , e si indica con $B_{\mathbb{R}^n}(x, r)$, il sottoinsieme

$$B_{\mathbb{R}^n}(x, r) := \left\{ y \in \mathbb{R}^n \mid \|x - y\|_{\mathbb{R}^n} < r \right\},$$

cioè l'insieme dei punti che distano meno di r da x .

Se $n = 1$ allora una palla centrata in x è un intervallo della retta reale; se $n = 2$ allora una palla è un disco (pieno) del piano; se $n = 3$ allora una palla è una sfera (piena) dello spazio $3D$.

Nel seguito, se non ci saranno ambiguità sulla dimensione (cioè sul numero naturale n), tralascieremo le notazioni con pedice $\|\cdot\|_{\mathbb{R}^n}$, $B_{\mathbb{R}^n}(x, r)$ per le più snelle $\|\cdot\|$, $B(x, r)$ senza pedice.

Definizione 2.5 Sia $A \subseteq \mathbb{R}^n$ un sottoinsieme.

Si dice interno di A , e si denota con $\text{int}(A)$ o con $\overset{\circ}{A}$, l'insieme dei punti $x \in A$ per cui esiste una palla aperta di raggio $r > 0$ in essi centrata e tutta contenuta in A

$$\text{int}(A) := \left\{ x \in A \mid \exists r > 0, B(x, r) \subseteq A \right\} \subseteq \mathbb{R}^n.$$

Si dice frontiera di A (o bordo di A), e si denota con ∂A o con $\text{fr}(A)$, l'insieme dei punti per i quali ogni palla in essi centrata interseca sia A che il suo complementare

$$\partial A = \left\{ x \in \mathbb{R}^n \mid \forall r > 0, B(x, r) \cap A \neq \emptyset, B(x, r) \cap (\mathbb{R}^n \setminus A) \neq \emptyset \right\},$$

cioè per ogni raggio $r > 0$ esiste un elemento $y \in A$ che dista da x meno di r ed esiste un elemento $z \notin A$ che dista da x meno di r .

Il sottoinsieme A si dice chiuso se contiene la sua frontiera $\partial A \subseteq A$.

Il sottoinsieme A si dice aperto se coincide con il suo interno $A = \text{int}(A)$.

Si dice chiusura di A , e si denota con \overline{A} , l'insieme formato da A unito la sua frontiera

$$\overline{A} := A \cup \partial A.$$

Il sottoinsieme A si dice limitato se è contenuto in una palla di centro l'origine e di raggio r opportuno, cioè se esiste $r > 0$ tale che

$$A \subseteq B(O, r).$$

Il sottoinsieme A si dice illimitato se non è limitato, cioè se non è contenuto in alcuna palla centrata nell'origine, cioè se, per ogni $r > 0$

$$A \not\subseteq B(O, r).$$

Osservazione 2.6 In generale i sottoinsiemi di \mathbb{R}^n non sono né chiusi né aperti, ma contengono pezzi della loro frontiera, ma non tutta. Un insieme è aperto se e soltanto se non contiene alcun pezzo della sua frontiera. Un insieme è chiuso se e soltanto se coincide con la sua chiusura. Esistono insiemi con interno vuoto (e cioè formati di sola frontiera) per esempio un segmento nel piano. L'unico insieme con frontiera vuota è tutto \mathbb{R}^n . Gli unici insiemi che sono sia aperti che chiusi sono tutto \mathbb{R}^n e (quasi per definizione) l'insieme vuoto $\emptyset \subset \mathbb{R}^n$. La chiusura di un insieme è un insieme chiuso, l'interno di un insieme è un insieme aperto. La palla aperta $B(x, r)$ definita in Definizione 2.4 è un insieme aperto nel senso della Definizione 2.5. La chiusura della palla aperta $B(x, r)$ è la cosiddetta palla chiusa

$$\bar{B}(x, r) = \left\{ y \in \mathbb{R}^n \mid \|x - y\| \leq r \right\},$$

che contiene il suo bordo

$$\partial\bar{B}(x, r) = \partial B(x, r) = \left\{ y \in \mathbb{R}^n \mid \|x - y\| = r \right\}.$$

Definizione 2.7 Una successione di punti in \mathbb{R}^n è una legge (funzione) che ad ogni numero naturale $k \in \mathbb{N}$ associa un punto³ x_k di \mathbb{R}^n (anche sempre lo stesso) e si indica con $\{x_k\}_{k \in \mathbb{N}}$ (o più semplicemente con $\{x_k\}_k$).

Una successione $\{x_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ converge ad un punto $x \in \mathbb{R}^n$ se essa “si avvicina arbitrariamente e definitivamente sempre di più ad x ” cioè se

$$\forall \varepsilon > 0 \exists \bar{k} \in \mathbb{N} \text{ tale che } k \geq \bar{k} \implies \|x_k - x\| \leq \varepsilon.$$

Diciamo invece che la successione diverge se “si allontana arbitrariamente e definitivamente sempre di più dall'origine” cioè se

$$\forall M > 0 \exists \bar{k} \in \mathbb{N} \text{ tale che } k \geq \bar{k} \implies \|x_k\| \geq M.$$

Nella definizione di convergenza qui sopra, l'avverbio “arbitrariamente” è formalizzato dalla scelta arbitraria di $\varepsilon > 0$ (cioè dal criterio di vicinanza a x), l'avverbio “definitivamente” è formalizzato dall'esistenza dell'indice \bar{k} per il quale, “da quel momento in poi”, tutti i punti distano da x al più ε , la locuzione “sempre di più” è formalizzata dal poter prendere ε sempre più piccolo (benché positivo). Analoghe considerazioni valgono per la definizione di divergenza.

Osservazione 2.8 Con riferimento all'Analisi Matematica unidimensionale (già nota), possiamo dire che la successione $\{x_k\}_k$ converge a x se e soltanto se la successione numerica $a_k = \|x_k - x\|$ converge a zero. Analogamente, dalla definizione di distanza (o da quella di norma), si ha subito che, posto $x_k = (x_1^k, \dots, x_n^k)$, $x = (x_1, \dots, x_n)$, la successione converge a x se e soltanto se le n successioni numeriche $\{x_i^k\}_k \subset \mathbb{R}$, $i = 1, \dots, n$, convergono rispettivamente a $x_i \in \mathbb{R}$, cioè se e soltanto se la successione converge “componente per componente”. Analogamente la successione diverge se e soltanto se la successione numerica $\|x_k\|$ diverge a $+\infty$.

³Notare che qui l'indice k non è un indice di componente del punto x , ma è l'indice della successione, cioè x_k è il k -esimo punto della successione.

Esempio 2.9 i) La successione

$$\left\{ x_n = \left(\frac{1}{n}, \frac{\sin(n)}{n} \right) \right\}_{n \in \mathbb{N} \setminus \{0\}} \subset \mathbb{R}^2$$

converge all'origine $O = (0, 0)$. Infatti, fissato $\varepsilon > 0$, prendiamo $\bar{n} \geq \frac{\sqrt{2}}{\varepsilon}$ e abbiamo, per ogni $n \geq \bar{n}$,

$$\|x_n - O\| = \left\| \left(\frac{1}{n}, \frac{\sin(n)}{n} \right) - (0, 0) \right\| = \sqrt{\frac{1}{n^2} + \frac{\sin(n)^2}{n^2}} \leq \sqrt{\frac{2}{n^2}} \leq \varepsilon.$$

ii) La successione

$$\{x_n = (\log(1 + n^2), e^{-n})\}_{n \in \mathbb{N}} \subset \mathbb{R}^2$$

diverge. Infatti, per ogni $M > 0$ prendiamo $\bar{n} \geq \sqrt{e^M - 1}$. Allora, per ogni $n \geq \bar{n}$ si ha

$$\|x_n\| = \sqrt{\log(1 + n^2)^2 + e^{-2n}} \geq \sqrt{\log(1 + n^2)^2} \geq M.$$

Notare che solo la prima componente di x_n va all'infinito, l'altra componente, addirittura, converge a zero.

iii) La successione

$$\left\{ x_n = \left((-1)^n, \frac{1}{2} - \frac{1}{n} \right) \right\}_{n \in \mathbb{N} \setminus \{0\}} \subset \mathbb{R}^2,$$

né converge (ad alcun punto) né diverge. Infatti, non converge ad alcun punto perché la prima componente non converge ad alcun valore reale: oscilla tra 1 e -1 (notare che invece la seconda componente converge a un mezzo). Non diverge perché sta sempre a distanza limitata dall'origine:

$$\|x_n\| = \sqrt{1 + \left(\frac{1}{2} - \frac{1}{n} \right)^2} \leq \sqrt{1 + \frac{1}{4}} = \frac{\sqrt{5}}{2} \quad \forall n \in \mathbb{N} \setminus \{0\}.$$

2.2 Esempi di funzioni da \mathbb{R}^2 in \mathbb{R} e loro grafici

2.2.1 Grafici

Data una funzione f definita su \mathbb{R}^2 a valori scalari in \mathbb{R}

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y),$$

il grafico di f è il sottoinsieme Γ_f di \mathbb{R}^3 così definito

$$\Gamma_f : \left\{ (x, y, z) \in \mathbb{R}^3 \mid z = f(x, y) \right\}.$$

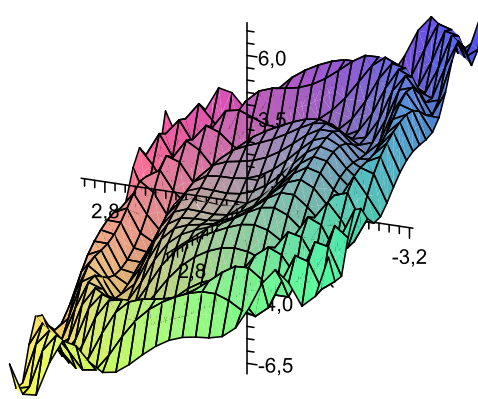


Figure 1: Grafico di $\sin(x^2 + y^2) - x - y$.

Esso consiste nelle terne di \mathbb{R}^3 in cui le prime due coordinate (x, y) si pensano come elementi di \mathbb{R}^2 e la terza z è il valore che f assume in (x, y) ⁴

Se la funzione f è definita solo su di un sottoinsieme di \mathbb{R}^2 , diciamo $f : A \rightarrow \mathbb{R}$ con $A \subseteq \mathbb{R}^2$, allora il grafico di f è l'insieme delle terne che stanno “sopra” l'insieme A e la cui “quota” (terza coordinata) è il valore assunto da f (attenzione che questa quota può anche essere negativa se f assume valori negativi, e quindi la parola “sopra” può significare anche sotto: insomma, va presa in senso “algebrico”):

$$\Gamma_f = \left\{ (x, y, z) \in \mathbb{R}^3 \mid (x, y) \in A, z = f(x, y) \right\}.$$

Per esempio, il grafico della funzione

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

sopra il quadrato $[-\pi, \pi] \times [-\pi, \pi]$ è rappresentato in Figura 1.

2.2.2 Polinomi

Un polinomio in due variabili x, y di primo grado è una funzione

$$p : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto p(x, y) = ax + by + c.$$

⁴si pensi, per analogia, al grafico di una funzione di una variabile reale $g : \mathbb{R} \rightarrow \mathbb{R}$, che è quel sottoinsieme del piano, \mathbb{R}^2 , tale che ...

dove $a, b, c \in \mathbb{R}$ sono i coefficienti (fissati) del polinomio, con a o b diversi da zero.

Il polinomio è di secondo grado se è della forma

$$p(x, y) = \alpha x^2 + \beta y^2 + \gamma xy + \delta x + \epsilon y + \phi,$$

dove $\alpha, \beta, \gamma, \delta, \epsilon, \phi \in \mathbb{R}$ sono i coefficienti del polinomio, con α, β o γ diversi da zero.

Il polinomio è di terzo grado se è della forma

$$p(x, y) = Ax^3 + By^3 + Cx^2y + Dxy^2 + q(x, y)$$

dove q è un polinomio di primo o di secondo grado e $A, B, C, D \in \mathbb{R}$ con A, B, C o D diversi da zero.

In modo analogo si definiscono i polinomi di grado maggiore al terzo.

Ad esempio l'espressione

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

è un polinomio di terzo grado dove: $A = 1, B = 0, C = 0, D = -1$, e $q(x, y) = x^2 - y - 3$ è un polinomio di secondo grado con $\alpha = 1, \beta = 0, \gamma = 0, \delta = 0, \epsilon = -1$ e $\phi = -3$.

In modo analogo, infine, si definiscono i polinomi in n variabili.

2.2.3 Paraboloidi

Parabolide. Il parabolide è il grafico della funzione

$$f(x, y) = x^2 + y^2 = \|(x, y)\|^2.$$

Questa funzione è un polinomio di secondo grado omogeneo (non ci sono termini di grado inferiore al secondo), assume sempre valori maggiori o uguali a zero (zero solo nell'origine) ed è definita su tutto \mathbb{R}^2 . Il suo grafico (il parabolide) è rappresentato in Figura 2. Se "affettiamo" il parabolide con piani verticali otteniamo parabole, con piani orizzontali otteniamo invece circonferenze. È la "parabola" della televisione satellitare.

Parabolide iperbolico. Il parabolide iperbolico è il grafico della funzione

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

Questa funzione è un polinomio di secondo grado omogeneo, assume valori sia positivi che negativi ed è definita su tutto \mathbb{R}^2 . Il suo grafico (il parabolide iperbolico) è rappresentato in Figura 3. Se lo affettiamo con piani verticali otteniamo parabole (eventualmente degeneri: rette), con piani orizzontali otteniamo invece iperboli. È la sella del cavallo, o il passo di montagna (detto anche sella).

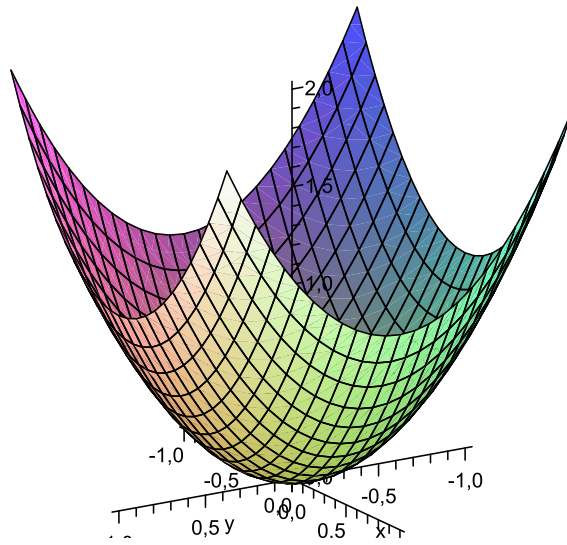


Figure 2: Paraboloide.

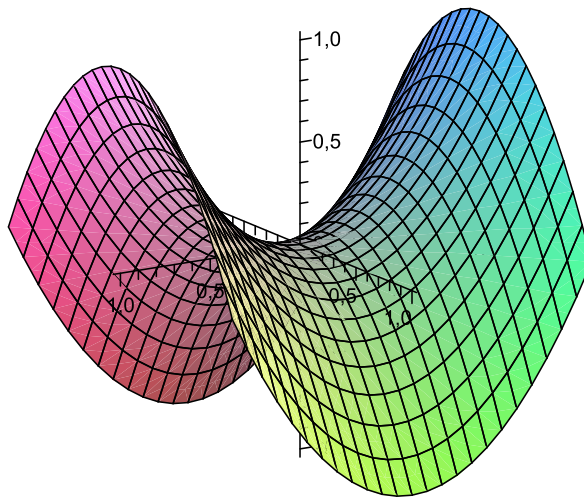


Figure 3: Paraboloide iperbolico.

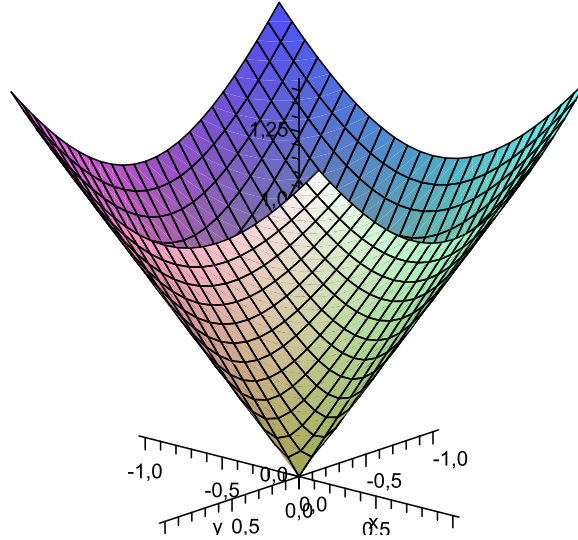


Figure 4: Cono.

2.2.4 Cono

Il cono è il grafico della funzione

$$f(x, y) = \sqrt{x^2 + y^2} = \|(x, y)\|.$$

Questa funzione non è un polinomio, è sempre maggiore o uguale a zero (zero solo nell'origine. N.B. Per definizione la “funzione radice quadrata” è positiva) ed è definita su tutto \mathbb{R}^2 . Il suo grafico (il cono) è rappresentato in Figura 4. Se lo affettiamo con piani verticali otteniamo rami di iperboli, con piani orizzontali otteniamo circonferenze.

2.2.5 Semisfera

La semisfera è il grafico di una funzione del tipo

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

Questa funzione non è un polinomio è sempre maggiore o uguale a zero ed è definita solamente sull'insieme (l'argomento della radice non può essere negativo)

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1 \right\},$$

che consiste nel cerchio di centro l'origine e raggio 1. Il grafico di f , rappresentato in Figura 5, è la semisfera superiore di centro l'origine e raggio 1.

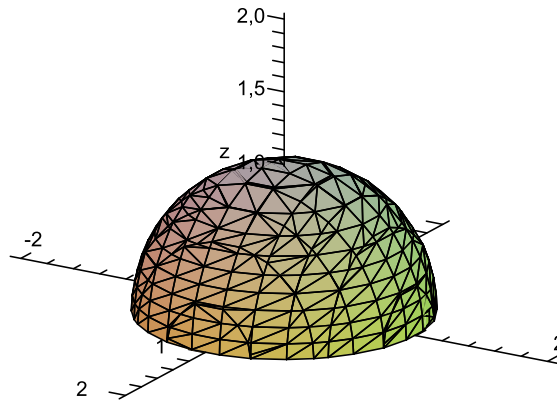


Figure 5: Semisfera.

2.3 Sottoinsiemi di \mathbb{R}^3 come luogo degli zeri di funzioni da \mathbb{R}^3 in \mathbb{R}

Nella sezione precedente, abbiamo visto importanti esempi di sottoinsiemi di \mathbb{R}^3 nella “veste” di grafici di funzioni da \mathbb{R}^2 in \mathbb{R} . Un altro modo di notevole di rappresentare (altri) insiemi di \mathbb{R}^3 è quello nella “veste” di luogo degli zeri di funzioni.

Sia data una funzione di tre variabili a valori reali

$$g : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad (x, y, z) \mapsto g(x, y, z).$$

Il luogo degli zeri di g è quel sottoinsieme di \mathbb{R}^3 , Z_g , eventualmente vuoto, così definito

$$Z_g = \left\{ (x, y, z) \in \mathbb{R}^3 \mid g(x, y, z) = 0 \right\},$$

cioè è l’insieme di tutti i punti di \mathbb{R}^3 in cui g si annulla, ovvero dei punti che soddisfano l’equazione $g(x, y, z) = 0$.

Ad esempio, il luogo degli zeri della funzione

$$g(x, y, z) = x^2 - \log(1 + y^2 + z^2) - x^2 z^2 - xy,$$

(all’interno del cubo $[-10, 10] \times [-10, 10] \times [-10, 10]$) è rappresentato in Figura 6.

2.3.1 Sfera

Consideriamo la funzione

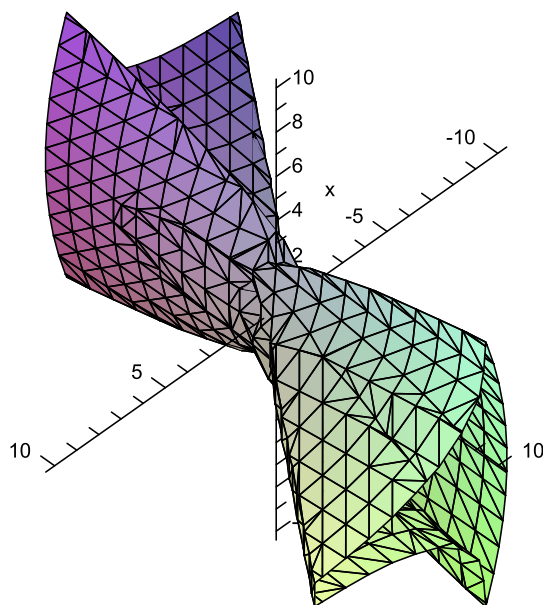


Figure 6: Luogo di zeri di $g(x, y, z) = x^2 - \log(1 + y^2 + z^2) - x^2z^2 - xy$.

$$g(x, y, z) = x^2 + y^2 + z^2 - 1 = \|(x, y, z)\|^2 - 1.$$

Questa funzione è un polinomio di secondo grado nelle tre variabili x, y, z , ed è definita su tutto \mathbb{R}^3 . Il suo luogo degli zeri è formato dai punti che soddisfano l'equazione

$$x^2 + y^2 + z^2 - 1 = 0, \text{ ovvero } x^2 + y^2 + z^2 = 1, \text{ ovvero } \|(x, y, z)\|^2 = 1, \text{ ovvero } \|(x, y, z)\| = 1$$

(N.B. la norma è, per definizione, positiva). Si tratta quindi del luogo dei punti di \mathbb{R}^3 che distano 1 dall'origine, ovvero è la sfera di centro l'origine e raggio 1 (vedi Figura 7).

2.3.2 Cilindro

Consideriamo la funzione

$$g(x, y, z) = x^2 + y^2 - 1.$$

È un polinomio di secondo grado ed è definita su tutto \mathbb{R}^3 . Notiamo che, anche se nell'espressione di g compaiono solo x e y , la nostra funzione è comunque funzione di tre variabili x, y, z , solamente che è costante nella terza variabile (cioè il suo valore dipende solo dalle prime due). Questo significa che quando cerchiamo il luogo degli zeri di g , ovvero quando risolviamo la seguente equazione nelle tre incognite x, y, z (cioè cerchiamo terne di numeri)

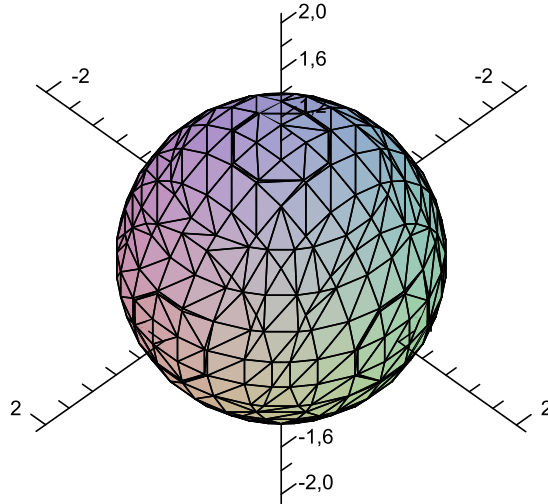


Figure 7: Sfera unitaria.

$$x^2 + y^2 - 1 = 0,$$

ad ogni coppia di valori (x, y) che soddisfa l'equazione possiamo aggiungere qualunque valore di z (qualunque “quota”). Ad esempio la terna $(1, 0, 0)$ soddisfa all'equazione, ma anche le terne $(1, 0, -1)$ e $(1, 0, 4)$ la soddisfano. Più in generale, tutte le terne del tipo $(1, 0, z)$ con qualunque $z \in \mathbb{R}$ sono soluzioni.

È facile ora convincersi che il luogo degli zeri di g , rappresentato in Figura 8, è il cilindro circolare retto indefinito (alto infinitamente in su e in giù), di sezione la circonferenza di raggio 1 e centro sull'asse z .

2.3.3 Ellissoide

Consideriamo la funzione

$$g(x, y, z) = \frac{x^2}{6} + \frac{y^2}{9} + \frac{z^2}{16} - 1.$$

Questa funzione è un polinomio di secondo grado ed è finita su tutto \mathbb{R}^3 . Il suo luogo degli zeri è un ellissoide con semiassi $\sqrt{6}, 3, 4$, rappresentato in Figura 9. Se a sezione circolare, cioè due semiassi uguali, è la palla da rugby.

2.3.4 Iperboloidi

Iperboloide a una falda. Consideriamo la funzione

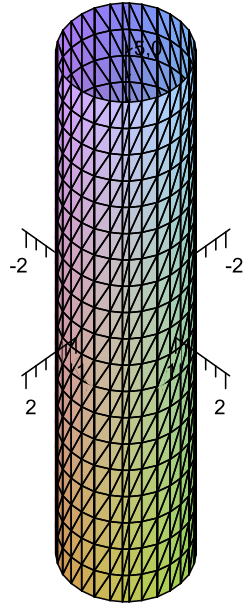


Figure 8: Cilindro.

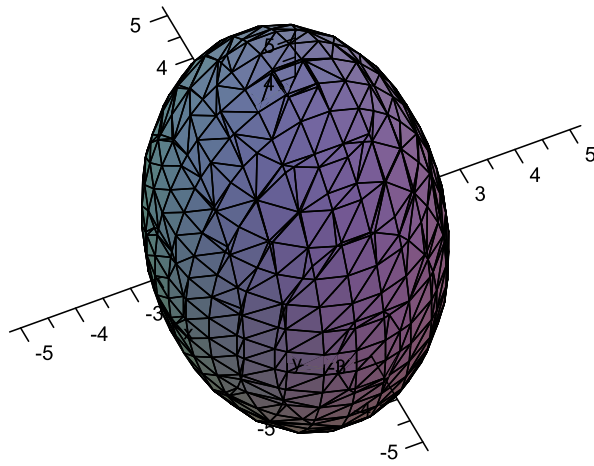


Figure 9: Ellissoide.

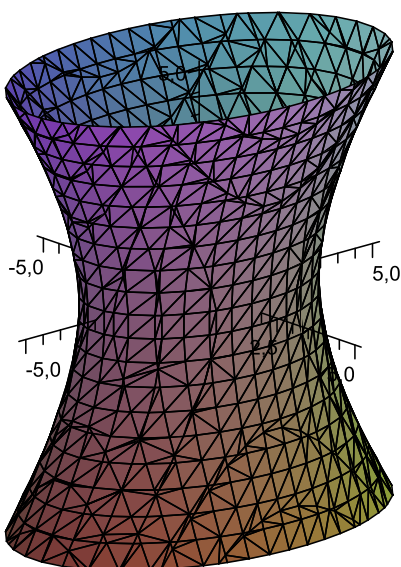


Figure 10: Iperboloide ad una falda.

$$g(x, y, z) = \frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{16} - 1.$$

È un polinomio di secondo grado, è definita su tutto \mathbb{R}^3 . Il suo luogo degli zeri è il cosiddetto iperboloide ad una falda (fatto di un pezzo solo) rappresentato in Figura 10. Le sezioni orizzontali sono ellissi, quelle verticali sono iperboli.

Iperboloide a due falde. Consideriamo la funzione

$$g(x, y, z) = -\frac{x^2}{4} - \frac{y^2}{9} + \frac{z^2}{16} - 1.$$

È un polinomio di secondo grado, è definita su tutto \mathbb{R}^3 . Il suo luogo degli zeri è il cosiddetto iperboloide a due falde (fatto di due pezzi) rappresentato in Figura 11. Le sezioni orizzontali sono ellissi, quelle verticali sono iperboli.

Osservazione 2.10 È facile convincersi che tutti gli esempi di questo paragrafo (luoghi di zeri) non sono rappresentabili come grafici di funzioni $f(x, y)$. Ad esempio la sfera non è il grafico di una funzione di tal tipo perchè, altrimenti, per esempio dovrebbe essere $f(0, 0) = 1$ e anche $f(0, 0) = -1$ (“sopra” il punto $(0, 0)$ ci sono i punti della sfera $(0, 0, 1)$ e $(0, 0, -1)$) e questo contraddice la definizione di funzione (ad ogni valore ne associa uno solo). Più in generale la sfera non è il grafico nemmeno di funzioni di (x, z) e di (y, z) (cioè grafici visti da altre angolazioni).

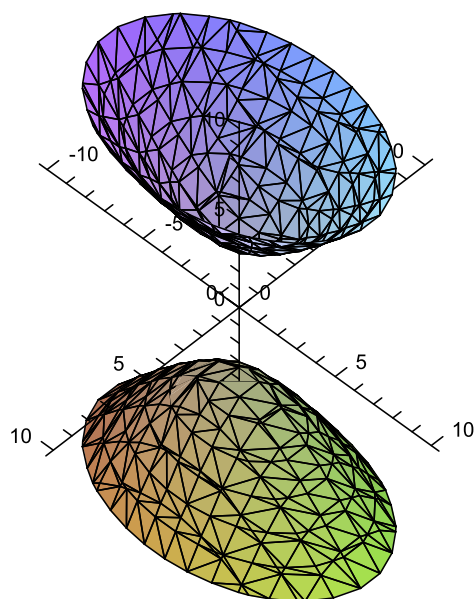


Figure 11: Iperboloide a due falde.

Quello che invece accade è che ogni grafico in \mathbb{R}^3 di funzione f di due variabili (x, y) è il luogo degli zeri di una funzione g di tre variabili (x, y, z) : per definizione di grafico, basta porre

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

2.4 Restrizione di una funzioni di più variabili ad un segmento

Un segmento (o una retta) di \mathbb{R}^n è un oggetto "unidimensionale", cioè può essere descritto facendo variare un solo parametro reale⁵. Ad esempio il segmento nel piano \mathbb{R}^2 di estremi $(1, 0)$ e $(3, 0)$ può essere descritto, in modo ovvio, nel seguente modo

$$\left\{ (t, 0) \in \mathbb{R}^2 \mid 1 \leq t \leq 3 \right\}.$$

È evidente che, al variare del parametro t tra i valori 1 e 3, la coppia $(t, 0)$ descrive il segmento in questione.

Rimanendo nel piano, dato un segmento di estremi $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$, allora, se P_1 e P_2 non sono sulla stessa retta verticale (cioè se $x_1 \neq x_2$), allora possiamo considerare l'equazione $y = mx + q$ della (unica) retta che passa per P_1 e per P_2 . È evidente allora che il segmento di vertici P_1 e P_2 si può descrivere come

$$\left\{ (x, mx + q) \mid x_1 \leq x \leq x_2 \right\}.$$

⁵Praticamente, pensando il parametro reale come il "tempo", si può interpretare tale descrizione del segmento come la legge oraria del moto di una particella che si muove lungo il segmento stesso.

Ad esempio se $P_1 = (2, 5)$ e $P_2 = (1, -2)$, la retta ha equazione $y = 7x - 9$ e quindi il nostro segmento è

$$\left\{ (x, 7x - 9) \mid 1 \leq x \leq 2 \right\}.$$

Se invece, ad esempio, abbiamo il segmento verticale di vertici $P_1 = (2, 45)$ e $P_2 = (2, 112)$, allora esso si può descrivere come

$$\left\{ (2, t) \mid 45 \leq t \leq 112 \right\}.$$

Più in generale, in qualunque \mathbb{R}^n , dati due punti $X = (x_1, \dots, x_n)$ e $Y = (y_1, \dots, y_n)$, il segmento da essi determinato $[X, Y]$, si può descrivere nel seguente modo

$$\left\{ tX + (1 - t)Y = (tx_1 + (1 - t)y_1, \dots, tx_n + (1 - t)y_n) \mid 0 \leq t \leq 1 \right\}.$$

Ancora in generale, se $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ è un vettore di norma unitaria, $\|v\|_{\mathbb{R}^n} = 1$, cioè un versore, e se $P_0 = (x_1, \dots, x_n)$ è un punto di \mathbb{R}^n , allora il segmento di lunghezza $\ell > 0$ spiccato da P nel verso e nella direzione di v è, in forma vettoriale

$$\left\{ P_0 + tv \mid 0 \leq t \leq \ell \right\},$$

o, in forma cartesiana,

$$\left\{ (x_1 + tv_1, \dots, x_n + tv_n) \mid 0 \leq t \leq \ell \right\}.$$

Sia ora $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione e sia $[X, Y]$ un segmento di \mathbb{R}^n . Allora possiamo considerare la restrizione di f al segmento, ovvero considerare solo quei valori che f assume sui punti del segmento. Per fare questo, rappresentiamo il segmento al variare di un parametro reale t (o x , o...) in un opportuno intervallo $[a, b]$ di valori (come abbiamo fatto giusto qui sopra), e indichiamo con $P(t)$ il punto del segmento corrispondente al valore t del parametro. La restrizione di f al segmento consiste quindi di una funzione di una variabile reale (il parametro $t \in [a, b]$) che, ad ogni valore della variabile, associa il valore reale assunto da f sul punto $P(t)$. Cioè abbiamo la funzione

$$\varphi : [a, b] \rightarrow \mathbb{R}, \quad t \mapsto f(P(t)).$$

Ad esempio, riprendendo il segmento nel piano di vertici $(2, 5)$ e $(1, -2)$, e considerando la funzione di due variabili

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

abbiamo che la restrizione di f al segmento è una funzione del solo parametro x così definita

$$\begin{aligned} \varphi : [1, 2] &\rightarrow \mathbb{R}, \\ x &\mapsto f(x, 7x - 9) = 3x - (7x - 9) + \sqrt{x^2 + (7x - 9)^2} = 9 - 4x + \sqrt{50x^2 - 126x + 81}. \end{aligned}$$

2.5 Funzioni di più variabili a valori scalari

Una funzione di più variabili reali a valori scalari (o funzione scalare) è una funzione definita su un sottoinsieme A di \mathbb{R}^n (eventualmente anche tutto \mathbb{R}^n) e che assume valori scalari (numeri reali), cioè:

$$f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}. \quad (2.6)$$

In questo paragrafo analizzeremo le proprietà analitiche di tali funzioni.

2.5.1 Continuità

Definizione 2.11 *La funzione (2.6) si dice continua in $\bar{x} \in A$ se, per ogni successione di punti in A , $\{x_k\} \subseteq A$, convergente a \bar{x} , si ha che la successione numerica $\{f(x_k)\} \subseteq \mathbb{R}$ converge a $f(\bar{x})$ ⁶ La funzione si dice continua in A (o semplicemente continua) se è continua in ogni punto di A .*

In generale una funzione non è continua. Ad ogni modo noi incontreremo sempre funzioni continue. Ad esempio, tutti i polinomi sono funzioni continue.

Somma di funzioni continue è continua, prodotto di funzioni continue è continua, quoziente di funzioni continue è continua nei punti in cui non si annulla il denominatore. Inoltre, se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ è continua e se anche $g : \mathbb{R} \rightarrow \mathbb{R}$ è continua, allora la composizione

$$g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto g(f(x))$$

è pure continua. Ad esempio se $f(x, y) = 2xy - 3x + 7$ e $g(s) = \sin(s)$, allora anche la funzione composizione

$$(x, y) \mapsto \sin(2xy - 3x + 7)$$

è continua. Ugualmente sono continue le funzioni⁷

$$e^{2xy-5y}$$

composizione di $f(x, y) = 2xy - 5y$ e $g(s) = e^s$;

$$\log(1 + x^2 + y^2)$$

composizione di $f(x, y) = 1 + x^2 + y^2 > 0$ e $g(s) = \log(s)$, $s > 0$;

$$\frac{x - y + xyz}{2 + z^2}$$

rappporto tra $f(x, y, z) = x - y + xyz$ e $h(x, y, z) = 2 + z^2 \neq 0$;

$$1 + x + y + e^{2xy}$$

somma di $f(x, y) = 1 + x + y$ e $h(x, y) = e^{2xy}$;

$$\cos(2x - y^3 + \sqrt{x^2 + z^2 + 3})$$

composizione di $f(x, y, z) = 2x - y^3 + \sqrt{x^2 + z^2 + 3}$ e $g(s) = \cos(s)$

⁶Una definizione equivalente è quella classica con gli ε e i δ : per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $\|x - \bar{x}\|_{\mathbb{R}^n} \leq \delta$ allora $|f(x) - f(\bar{x})| \leq \varepsilon$.

⁷È ovvio che l'elenco qui riportato è solo esemplificativo, ben lungi dal rappresentare tutte le funzioni continue che sono infinitamente di più.

2.5.2 Derivabilità

Iniziamo questo paragrafo con un "flashback". Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione di una variabile reale a valori reali. Sia poi $x_0 \in \mathbb{R}$. Dire che f è derivabile in x_0 e che la sua derivata è $f'(x_0)$ significa dire che x_0 è interno al dominio di f (cioè f è definita (almeno) in un intervallo aperto contenente x_0), e soprattutto che il seguente limite del rapporto incrementale

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

esiste finito e vale $f'(x_0) \in \mathbb{R}$. È noto che l'esistenza della derivata $f'(x_0)$ risolve simultaneamente due problemi: a) calcolare il "tasso di variazione" della funzione f (quanto "pende" il grafico di f nel punto $(x_0, f(x_0))$); b) determinare la retta tangente al grafico di f nel punto $(x_0, f(x_0))$: è la retta di equazione $y = f'(x_0)x + f(x_0) - f'(x_0)x_0$. In particolare, il punto b) risponde anche al problema di approssimare, "vicino" a x_0 , la f con una funzione lineare (il cui grafico è appunto la retta tangente).

Nel caso a più dimensioni, $n \geq 2$, questi due aspetti (tasso di variazione e approssimazione con una funzione lineare) sono separati, ovvero non sono simultaneamente risolti dall'esistenza del limite del rapporto incrementale, come per il caso unidimensionale.

Una prima questione che si pone, nel caso a più variabili, è cosa significhi "tasso di variazione". In una dimensione questo è ovvio: abbiamo un'unica direzione in cui muoversi (cioè far variare x): la linea stessa dei numeri reali. Già in dimensione 2 questo non è più vero: abbiamo infinite direzioni in cui muoverci a partire da un punto fissato. Quindi rispetto a quale direzione dobbiamo calcolare il tasso di variazione? È ragionevole pensare che ci sia un tasso di variazione per ogni possibile direzione.

Definizione 2.12 *Siano $A \subseteq \mathbb{R}^n$ un aperto, $f : A \rightarrow \mathbb{R}$ una funzione, $\bar{x} \in A$ un punto e $v \in \mathbb{R}^n$ un versore (norma unitaria) fissati. Si dice che f è derivabile in \bar{x} nella direzione v se esiste finito il limite del rapporto incrementale*

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

Il valore di tale limite (quando esiste) si denota con $\frac{\partial f}{\partial v}(\bar{x})$ e si chiama la derivata direzionale di f nel punto \bar{x} rispetto alla direzione v .

Se $v = e_i$, i -esimo elemento della base canonica di \mathbb{R}^n ($e_i = (0, \dots, 0, 1, 0, \dots, 0)$ con 1 al posto i - esimo), allora la derivata direzionale rispetto a e_i si chiama i -esima derivata parziale e si denota con $\frac{\partial f}{\partial x_i}(\bar{x})$.

Osservazione 2.13 *La derivata direzionale di f nella direzione v è quindi la derivata usuale della funzione di una sola variabile reale che consiste nella restrizione di f al segmento passante per \bar{x} e parallelo alla direzione di v .*

Nel caso di \mathbb{R}^2 , $e_1 = (1, 0)$ (rispettivamente: $e_2 = (0, 1)$) è la direzione della retta delle ascisse (rispettivamente: delle ordinate), cioè delle x (rispettivamente: delle y) e quindi spesso si usa, per indicare la prima derivata parziale (rispettivamente: la seconda derivata parziale) il simbolo $\frac{\partial f}{\partial x}$ (rispettivamente: $\frac{\partial^2 f}{\partial x^2}$). Analogamente in \mathbb{R}^3 si usano i simboli

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}.$$

Alle volte si usano anche i simboli f_x, f_y, f_z .

Esempio 2.14 Siano $f(x, y) = 2xy - x^2$, $(\bar{x}, \bar{y}) = (1, 2)$ e $v = (\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}})$. Allora si ha

$$\begin{aligned} \frac{\partial f}{\partial v}(1, 2) &= \lim_{t \rightarrow 0} \frac{f(1 + (t/\sqrt{10}), 2 + (3t/\sqrt{10})) - f(1, 2)}{t} = \lim_{s \rightarrow 0} \frac{5s^2 + 8s}{\sqrt{10}s} = \frac{8}{\sqrt{10}}; \\ \frac{\partial f}{\partial x}(1, 2) &= \lim_{t \rightarrow 0} \frac{f(1+t, 2) - f(1, 2)}{t} = \lim_{t \rightarrow 0} \frac{-t^2 + 2t}{t} = 2; \\ \frac{\partial f}{\partial y}(1, 2) &= \lim_{t \rightarrow 0} \frac{f(1, 2+t) - f(1, 2)}{t} = \lim_{t \rightarrow 0} \frac{2t}{t} = 2. \end{aligned}$$

Regola per il calcolo delle derivate. Finora abbiamo parlato di “derivata in un punto” che quindi è un numero reale. Ma, come per il caso unidimensionale, si può considerare anche la “funzione derivata (parziale, direzionale..)” (o più semplicemente la “derivata” (parziale, direzionale...)). Ad esempio, nel caso unidimensionale, la derivata della funzione $f(x) = 3x^2 - \sin(x)$ è la funzione $f'(x) = 6x - \cos(x)$, e sono ben note le regole di derivazione per ottenere la funzione derivata. Ebbene, per il calcolo delle derivate parziali, si applicano le stesse regole di derivazione, con l'accortezza di applicarle solo alla variabile rispetto a cui si sta derivando, e considerando le altre come costanti. Ad esempio, data la funzione

$$f(x, y, z) = 3xy - y^3 + \sin(x - y + z),$$

si ha

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y, z) &= 3y + \cos(x - y + z), \\ \frac{\partial f}{\partial y}(x, y, z) &= 3x - 3y^2 - \cos(x - y + z), \\ \frac{\partial f}{\partial z}(x, y, z) &= \cos(x - y + z). \end{aligned}$$

Riprendendo l'Esempio 2.14 si ha

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) = 2y - 2x &\implies \frac{\partial f}{\partial x}(1, 2) = 2, \\ \frac{\partial f}{\partial y}(x, y) = 2x &\implies \frac{\partial f}{\partial y}(1, 2) = 2. \end{aligned}$$

Definizione 2.15 Siano $A \subseteq \mathbb{R}^n$ aperto, $\bar{x} \in A$ un punto fissato $f : A \rightarrow \mathbb{R}$ una funzione. La funzione f si dice derivabile in \bar{x} se, in \bar{x} , esistono tutte le sue derivate parziali. Se f è derivabile in \bar{x} , allora si dice gradiente di f in \bar{x} , e si denota con $\nabla f(\bar{x})$, il vettore di \mathbb{R}^n formato dalle n derivate parziali di f in \bar{x} cioè

$$\nabla f(\bar{x}) = \left(\frac{\partial f}{\partial x_1}(\bar{x}), \dots, \frac{\partial f}{\partial x_n}(\bar{x}) \right).$$

Ad esempio, riprendendo l'Esempio 2.14, si ha

$$\nabla f(1, 2) = (2, 2).$$

2.5.3 Differenziabilità

Il concetto di "derivabilità" espresso nel precedente paragrafo, rende conto solo dell'aspetto "tasso di variazione" della funzione, anche se con la complicazione di doverne calcolare uno per ogni possibile direzione. Ora passiamo all'altro aspetto, quello di "approssimazione con una funzione lineare". Ricordiamo che, nel caso di una funzione di una variabile reale, questa approssimazione vicino ad un punto $x_0 \in \mathbb{R}$, da un punto di vista geometrico, consiste nel determinare la retta tangente al grafico di f nel punto $(x_0, f(x_0))$. Tale retta è il grafico della funzione $x \mapsto f'(x_0)x + f(x_0) - f'(x_0)x_0$.

Iniziamo considerando una funzione $f : A \rightarrow \mathbb{R}$ dove A è un aperto di \mathbb{R}^2 e sia $(x_0, y_0) \in A$ un punto fissato. In questo caso il grafico di f è un sottoinsieme di \mathbb{R}^3 (una cosiddetta "superficie") e quindi, al posto di retta tangente, dobbiamo parlare di "piano tangente" al grafico di f nel punto $(x_0, y_0, f(x_0, y_0)) \in \mathbb{R}^3$. È ovvio che tale piano tangente deve contenere le rette tangenti alle due "curvette" sul grafico di f corrispondenti ai punti $(x, y, f(x, y))$ dove (x, y) sono, rispettivamente, i punti di "segmentini" passanti per (x_0, y_0) e paralleli agli assi coordinati. Tali rette hanno quindi pendenza misurabile con $\frac{\partial f}{\partial x}(x_0, y_0)$ e $\frac{\partial f}{\partial y}(x_0, y_0)$, rispettivamente.

Si dimostra che, se tale piano tangente esiste, deve avere la seguente equazione in \mathbb{R}^3

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) - z + f(x_0, y_0) = 0, \quad (2.7)$$

Qualora esistano le derivate parziali di f in (x_0, y_0) , il piano di equazione (2.7) esiste sempre e passa per $(x_0, y_0, f(x_0, y_0))$. Ma non è detto che sia tangente al grafico di f . Questo avviene se, intorno a (x_0, y_0) , la differenza fra le "quote" $f(x, y)$ e

$$z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0)$$

è un infinitesimo di ordine superiore alla distanza di (x, y) da (x_0, y_0) (in cui le quote coincidono), ovvero se il seguente limite esiste e vale zero⁸

⁸Questo significa che, per ogni successione di punti (x_n, y_n) convergente a (x_0, y_0) , l'espressione a sinistra tende a zero (in \mathbb{R}). Lo stesso dicasi per limiti di funzioni scalari definite in \mathbb{R}^n , con n arbitrario.

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x,y) - f(x_0,y_0) - \nabla f(x_0,y_0) \cdot (x - x_0, y - y_0)}{\|(x,y) - (x_0,y_0)\|_{\mathbb{R}^2}} = 0$$

Diamo ora una definizione generale.

Definizione 2.16 Siano $A \subseteq \mathbb{R}^n$ aperto, $\bar{x} \in A$ fissato e $f : A \rightarrow \mathbb{R}$ una funzione. Allora f si dice differenziabile in \bar{x} se essa è derivabile in \bar{x} (cioè se esistono tutte le derivate parziali) e se vale

$$\lim_{x \rightarrow \bar{x}} \frac{f(x) - f(\bar{x}) - \nabla f(\bar{x}) \cdot (x - \bar{x})}{\|x - \bar{x}\|_{\mathbb{R}^n}} = 0.$$

La funzione f si dice differenziabile se è differenziabile in ogni punto di A .

La seguente proposizione dà criteri per stabilire se una funzione è differenziabile e inoltre mette in luce alcune notevoli proprietà delle funzioni differenziabili.

Proposizione 2.17 *i) Se f ammette derivate parziali continue nell'intorno di un punto \bar{x} (cioè in una palla centrata in \bar{x}), allora f è differenziabile in \bar{x} .*

ii) I polinomi sono differenziabili; somma di differenziabili è differenziabile; prodotto di differenziabili è differenziabile; rapporto di differenziabili è differenziabile nei punti in cui non si annulla il denominatore; se f è differenziabile e $g : \mathbb{R} \rightarrow \mathbb{R}$ è derivabile, allora la composizione $g \circ f$ è differenziabile⁹.

iii) Una funzione differenziabile è continua e ammette derivate rispetto ad ogni direzione. Inoltre per ogni $v \in \mathbb{R}^n$ versore, se f è differenziabile in \bar{x} , vale la formula

$$\frac{\partial f}{\partial v}(\bar{x}) = \nabla f(\bar{x}) \cdot v$$

2.5.4 Derivate successive e formula di Taylor del secondo ordine

Data una funzione derivabile, possiamo chiederci se anche le sue derivate sono (ulteriormente) derivabili. Questo avviene, al solito, per i polinomi e per tutte le funzioni costruite sommando, moltiplicando, componendo ecc... come già detto nei paragrafi precedenti.

Ad esempio, la derivata parziale rispetto a x della derivata parziale rispetto a x di una funzione di due variabili (x, y) si denota (qualora esista) con

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

mentre la derivata rispetto a y della derivata rispetto a x è $\frac{\partial^2 f}{\partial y \partial x} (x, y)$.

⁹Ad esempio, sono differenziabili le funzioni $f(x, y) = e^{3x-y^2}$, $f(x, y) = \sin(x + y) - \cos(x + y)$, $f(x, y, z) = \sqrt{\sin^2(x + 3y - z^2) + 1}$ ecc...

Più in generale, data una funzione f di n variabili, allora, per ogni $m \in \mathbb{N}$ e per ogni $r_1, r_2, \dots, r_n \in \mathbb{N}$ tali che $r_1 + r_2 + \dots + r_n = m$, con il simbolo

$$\frac{\partial^m f}{\partial x_1^{r_1} \partial x_2^{r_2} \dots \partial x_n^{r_n}}(\bar{x})$$

si intende la derivata m -esima di f nel punto \bar{x} fatta r_1 volte rispetto alla prima componente x_1 , r_2 volte fatta rispetto alla seconda componente x_2, \dots, r_n volte fatta rispetto alla n -esima componente x_n .¹⁰

Ad esempio, consideriamo la funzione $f(x, y, z) = xyz + x - \log(1 + y^2 z^2)$. Allora si ha

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

Definizione 2.18 Una funzione $f : A \rightarrow \mathbb{R}$, con $A \subseteq \mathbb{R}^n$ aperto, si dice di classe C^m , con $m \in \mathbb{N} \setminus \{0\}$, se le derivate di f esistono tutte fino all'ordine m e sono continue. Si scrive anche $f \in C^m(A)$. Se esistono tutte le derivate di qualunque ordine e sono tutte continue, si scrive $f \in C^\infty(A)$.

Per un precedentemente enunciato teorema, è chiaro che le funzioni di classe C^m sono differenziabili.

Definizione 2.19 Siano $A \subseteq \mathbb{R}^n$ aperto, $f \in C^2(A)$ e $\bar{x} \in A$ fissato. Si chiama matrice Hessiana di f nel punto \bar{x} , e si denota con $H_f(\bar{x})$, la matrice quadrata $n \times n$

i cui coefficienti a_{ij} sono le derivate seconde $\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{x})$, cioè

$$H_f(\bar{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\bar{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\bar{x}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\bar{x}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\bar{x}) & \frac{\partial^2 f}{\partial x_2^2}(\bar{x}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(\bar{x}) \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\bar{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\bar{x}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\bar{x}) \end{pmatrix}.$$

¹⁰in realtà, fare prima la derivata rispetto ad una componente e poi quella rispetto ad un'altra componente e invece fare il viceversa, non danno, in generale lo stesso risultato. Vale però il Teorema di Schwarz che assicura che, sotto opportune ipotesi, scambiare l'ordine di derivazione rispetto alle singole variabili fa ottenere la medesima derivata. Al solito i polinomi e tutto il resto... godono di questa proprietà.

Chiamiamo matrice Hessiana di f , la matrice sopra definita, pensata come funzione di $x \in A$, cioè senza calcolare le derivate in uno specifico punto.

Poiché la funzione è di classe C^2 , allora, valendo il Teorema di Schwarz (vedi nota a piè di pagina), la matrice Hessiana è simmetrica.

Teorema 2.20 Siano $f \in C^2(A)$ e $\bar{x} \in A$ fissato, con $A \subseteq \mathbb{R}^n$ aperto. Allora vale la formula di Taylor al secondo ordine centrata in \bar{x} :

$$f(x) = f(\bar{x}) + \nabla f(\bar{x}) \cdot (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T H_f(\bar{x})(x - \bar{x}) + o(\|x - \bar{x}\|_{\mathbb{R}^n}^2), \quad x \rightarrow \bar{x}.$$

Nella formula di Taylor, il vettore $(x - \bar{x}) \in \mathbb{R}^n$ va inteso come vettore colonna e quindi il suo trasposto $(x - \bar{x})^T$ è un vettore riga. La scrittura $o(\|x - \bar{x}\|_{\mathbb{R}^n}^2)$, $x \rightarrow \bar{x}$ indica una quantità infinitesima quando x tende a \bar{x} di ordine superiore al secondo, cioè va a zero “più velocemente” di quanto vada a zero $\|x - \bar{x}\|^2$. La scrittura $(x - \bar{x})^T H_f(\bar{x})(x - \bar{x})$ indica la forma quadratica associata alla matrice Hessiana, applicata al punto $x - \bar{x}$.

2.5.5 Segno di una forma quadratica in \mathbb{R}^n

Definizione 2.21 Una forma quadratica in \mathbb{R}^n è un polinomio nelle n variabili $x = (x_1, \dots, x_n)$, che sia di secondo grado e omogeneo (cioè contiene solo addendi (monomi) di grado 2).

Ad esempio:

$$\begin{aligned} q(x, y) &= 2x^2 - 3xy && \text{è una forma quadratica su } \mathbb{R}^2, \\ q(x, y) &= 2x^2 - 3xy + x && \text{non è una forma quadratica,} \\ q(x, y, z) &= 3x^2 - xy + z^2 - yz && \text{è una forma quadratica su } \mathbb{R}^3, \\ q(x, y, z) &= 3x^2 - xy + z^2 - yz + 3 && \text{non è una forma quadratica.} \end{aligned}$$

Sia $q : \mathbb{R}^n \rightarrow \mathbb{R}$, $x \mapsto q(x)$, una forma quadratica su \mathbb{R}^n . Allora ad essa resta associata, in modo univoco, una matrice quadrata $n \times n$ simmetrica, A , tale che

$$q(x) = x^T A x, \quad \forall x \in \mathbb{R}^n,$$

dove x^T è il vettore riga, x è il vettore colonna e $x^T A x$ è l'usuale prodotto matriciale “righe per colonne”¹¹.

Ad esempio:

$$\begin{aligned} q(x, y) &= x^2 + 2xy - y^2 \longleftrightarrow (x, y) \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \\ q(x, y) &= -x^2 + 3xy \longleftrightarrow (x, y) \begin{pmatrix} -1 & \frac{3}{2} \\ \frac{3}{2} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \\ q(x, y, z) &= y^2 - 2xz + z^2 + 4yz \longleftrightarrow (x, y, z) \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned}$$

¹¹Se $A = (a_{ij})$ è tale matrice, allora a_{ii} è il coefficiente del termine x_i^2 , mentre, se $i \neq j$, a_{ij} e a_{ji} sono ambedue uguali a metà del coefficiente di $x_i x_j$.

È evidente che, con questa costruzione, vale anche il viceversa: ad ogni matrice simmetrica $n \times n$ resta univocamente associata una forma quadratica in \mathbb{R}^n . C'è quindi una naturale identificazione tra l'insieme delle forme quadratiche e quello delle matrici quadrate simmetriche.

Definizione 2.22 Una forma quadratica q su \mathbb{R}^n si dice

- i) definita positiva se $q(x) > 0$ per ogni $x \in \mathbb{R}^n$, $x \neq 0$,¹²
- ii) definita negativa se $q(x) < 0$ per ogni $x \in \mathbb{R}^n$, $x \neq 0$;
- iii) semidefinita positiva se $q(x) \geq 0$ per ogni $x \in \mathbb{R}^n$ e esiste $\bar{x} \neq 0$ tale che $q(\bar{x}) = 0$;
- iv) semidefinita negativa se $q(x) \leq 0$ per ogni $x \in \mathbb{R}^n$ e esiste $\bar{x} \neq 0$ tale che $q(\bar{x}) = 0$;
- v) non definita (o indefinita) se esistono $\bar{x}, \bar{y} \in \mathbb{R}^n$ tali che $q(\bar{x}) < 0 < q(\bar{y})$.

I seguenti esempi sono prototipi dei vari casi in \mathbb{R}^2 (verificarlo per esercizio)

$$\begin{aligned} q(x, y) &= x^2 + y^2 && \text{definita positiva,} \\ q(x, y) &= -x^2 - y^2 && \text{definita negativa,} \\ q(x, y) &= x^2 && \text{semidefinita positiva,} \\ q(x, y) &= -x^2 && \text{semidefinita negativa,} \\ q(x, y) &= x^2 - y^2 && \text{indefinita.} \end{aligned}$$

Il seguente risultato lega il “segno” della forma quadratica al segno degli autovalori della matrice simmetrica associata.

Proposizione 2.23 Sia q una forma quadratica su \mathbb{R}^n e sia A la matrice $n \times n$ simmetrica associata¹³. Allora si ha:

- i) q è definita positiva se e soltanto se tutti gli autovalori di A sono strettamente positivi;
- ii) q è definita negativa se e soltanto se tutti gli autovalori di A sono strettamente negativi;
- iii) q è semidefinita positiva se e soltanto se tutti gli autovalori di A sono non negativi (≥ 0) e però $\lambda = 0$ è autovalore di A ;
- iv) q è semidefinita negativa se e soltanto se tutti gli autovalori di A sono non positivi (≤ 0) e però $\lambda = 0$ è autovalore di A ;
- v) q è non definita se e soltanto se A ha autovalori strettamente positivi e autovalori strettamente negativi.

Data una matrice $n \times n$, chiamiamo “minori di nord-ovest” quei minori $r \times r$ formati dal blocco in alto a destra (nord-ovest, appunto), cioè $(a_{ij})_{i,j=1,\dots,r}$.

Proposizione 2.24 Data una matrice A simmetrica $n \times n$, allora essa: è definita positiva (ovvero tale è la forma quadratica associata) se e soltanto se tutti i suoi minori di nord-ovest hanno determinante strettamente positivo; è definita negativa (ovvero tale è la forma

¹²Notiamo che, per l'omogeneità, $q(0) = 0$ qualunque sia la forma quadratica (qui, lo zero argomento di q è lo zero vettore, il valore di $q(0)$ è lo zero numero reale!).

¹³Ricordiamo che una matrice simmetrica ha tutti gli autovalori reali.

quadratica ad essa associata) se soltanto se i minori di nordovest di ordine r dispari hanno determinante strettamente negativo e quelli di ordine r pari hanno determinante strettamente positivo.

Se $n = 2$, in particolare si ha

$$\begin{aligned} \det A > 0, a_{11} > 0 &\implies A \text{ definita positiva,} \\ \det A > 0, a_{11} < 0 &\implies A \text{ definita negativa,} \\ \det A = 0 &\implies A \text{ semidefinita (positiva o negativa),} \\ \det A < 0 &\implies A \text{ indefinita.} \end{aligned}$$

2.5.6 Massimi e minimi liberi

Definizione 2.25 Siano $A \subseteq \mathbb{R}^n$ aperto, $f : A \rightarrow \mathbb{R}$ una funzione e $\bar{x} \in A$ un punto fissato.

Il punto \bar{x} si dice un punto di massimo locale per f , e il valore $f(\bar{x}) \in \mathbb{R}$ si dice un massimo locale di f , se esiste una palla centrata in \bar{x} e tutta contenuta in A , $B(\bar{x}, r) \subseteq A$, $r > 0$, tale che

$$x \in B(\bar{x}, r) \implies f(x) \leq f(\bar{x}).$$

Il punto \bar{x} si dice un punto di minimo locale per f , e il valore $f(\bar{x}) \in \mathbb{R}$ si dice un minimo locale di f , se esiste una palla centrata in \bar{x} e tutta contenuta in A , $B(\bar{x}, r) \subseteq A$, $r > 0$, tale che

$$x \in B(\bar{x}, r) \implies f(x) \geq f(\bar{x}).$$

Il punto \bar{x} si dice un punto di massimo assoluto in A per f , e il valore $f(\bar{x}) \in \mathbb{R}$ si dice un massimo assoluto in A di f , se

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

Il punto \bar{x} si dice un punto di minimo assoluto in A per f , e il valore $f(\bar{x}) \in \mathbb{R}$ si dice un minimo assoluto in A di f , se

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

Osservazione 2.26 Si noti che, per esempio, nella definizione di massimo locale non si sta richiedendo che il valore massimo $f(\bar{x})$ non sia raggiunto anche in altri punti. In particolare non si esclude che la funzione sia costante (un unico valore raggiunto), nel qual caso tutti i punti sono sia di massimo che di minimo. Se tutte le disuguaglianze vengono sostituite con una disuguaglianza stretta (ad esempio $<$ al posto di \leq), allora si parla di massimo locale, minimo locale, massimo assoluto o minimo assoluto “stretto” (e in questo caso sÌ, è il più alto o il più basso mai raggiunto nella palla, o in tutto A .)

Notiamo infine che nella definizione di massimo o di minimo locale richiediamo che esista una palla centrata nel punto e tutta contenuta in A e quindi è naturale dare la

definizione per A aperto. Nella definizione di massimo e di minimo assoluti invece non è necessario che il dominio A sia aperto: la definizione può essere data per qualunque insieme A ¹⁴.

Teorema 2.27 Siano $A \subseteq \mathbb{R}^n$ un aperto e $f : A \rightarrow \mathbb{R}$ una funzione di classe C^1 . Allora, se $\bar{x} \in A$ è di massimo o di minimo locale per f , si ha

$$\nabla f(x) = 0 \in \mathbb{R}^n,$$

cioè tutte le derivate parziali prime di f in x si annullano.

Dimostrazione (intuitiva). Se, ad esempio, x è di massimo locale, allora, per definizione esiste una “palletta” centrata in x su cui $f(x)$ è il massimo. Ma allora $f(x)$ è anche il valore massimo raggiunto da f in un “segmentino” (contenuto nella palletta) passante per x e parallelo al primo asse coordinato (parallelo a $e_1 = (1, 0, \dots, 0)$). Se restringiamo f a tale segmentino, abbiamo una funzione φ di una sola variabile $t \in \mathbb{R}$ (il parametro che descrive il segmentino), e tale funzione deve avere un massimo nel valore del parametro che corrisponde al punto x . Quindi la derivata prima di φ deve annullarsi. Ma la derivata prima di φ altro non è che la derivata parziale di f rispetto a x_1 . \square

I punti in cui si annulla il gradiente vengono detti punti stazionari di f . Il teorema qui sopra dice che i punti di massimo o di minimo locale sono necessariamente punti stazionari. Ovviamente il viceversa non è vero, cioè possono esistere punti stazionari che non sono né di massimo né di minimo¹⁵.

Definizione 2.28 Un punto stazionario per f che non sia né di massimo locale né di minimo locale si dice un punto di sella (o una sella) per f . In particolare \bar{x} è un punto di sella se $\nabla f(\bar{x}) = 0$ e se, per ogni raggio $r > 0$ tale che $B(\bar{x}, r) \subseteq A$ si ha

$$\exists x, y \in B(\bar{x}, r) \text{ tali che } f(x) < f(\bar{x}) < f(y),$$

cioè mi posso avvicinare a \bar{x} quanto voglio con punti i cui valori di f sono strettamente più grandi di $f(\bar{x})$ e con punti i cui valori di f sono strettamente più piccoli di $f(\bar{x})$.

Teorema 2.29 Siano $A \subseteq \mathbb{R}^n$ aperto, $f : A \rightarrow \mathbb{R}$ di classe C^2 e $\bar{x} \in A$ un punto stazionario per f . Allora si ha:

- i) se la matrice Hessiana di f in \bar{x} è definita positiva, allora \bar{x} è un punto di minimo locale stretto per f ,
- ii) se la matrice Hessiana di f in \bar{x} è definita negativa, allora \bar{x} è un punto di massimo locale stretto per f ,

¹⁴In realtà anche per i massimi e i minimi locali si può dare una definizione con A non aperto, ma non addentriamoci nel merito.

¹⁵In analogia a quanto avviene in dimensione 1: l’annullarsi della derivata prima non implica il fatto di essere un punto di massimo o di minimo.

iii) se la matrice Hessiana di f in \bar{x} è indefinita, allora \bar{x} è un punto di sella.

Se invece la matrice Hessiana è semidefinita (positiva o negativa), allora il solo studio del segno della matrice non è sufficiente per determinare la natura del punto stazionario \bar{x} .

Dimostrazione (intuitiva). Se \bar{x} è stazionario, allora $\nabla f(\bar{x}) = 0$ e quindi la formula di Taylor al secondo ordine diventa, per $x \rightarrow \bar{x}$

$$f(x) - f(\bar{x}) = x^T H_f(\bar{x})x + o(\|x - \bar{x}\|^2).$$

È chiaro, ad esempio, che \bar{x} è di massimo locale stretto se e soltanto se, per punti $x \neq \bar{x}$ sufficientemente vicini, il membro di sinistra è strettamente negativo. Quindi anche il membro di destra deve essere strettamente negativo. Ma per x vicini a \bar{x} il segno del membro di destra è esattamente il segno della forma quadratica $x^T H_f(\bar{x})x$, in quanto il secondo addendo è di ordine superiore: è troppo piccolo per “mangiarsi” il segno della forma quadratica.

Da quanto detto si deducono immediatamente i punti i) e ii). Per il punto iii) basta osservare che essere di sella, cioè avere punti vicini quanto si vuole in cui il membro di sinistra ha segno positivo e altri in cui ha segno negativo, è equivalente a dire, nel membro di destra, che la forma quadratica può assumere sia il segno positivo che quello negativo.

Con dei controesempi proviamo ora che, nel caso di matrice Hessiana semidefinita, qualsiasi cosa può accadere.

a) $f(x, y) = x^4 + y^4$, $H_f(0, 0) = O$ (matrice nulla, chiaramente semidefinita): è evidente che $(0, 0)$ è un punto di minimo locale (stretto);

b) $f(x, y) = -x^4 - y^4$, $H_f(0, 0) = O$ e $(0, 0)$ è un punto di massimo locale (stretto);

c) $f(x, y) = x^3 y^3$, $H_f(0, 0) = O$ e $(0, 0)$ è un punto di sella. \square

Esempio 2.30 Studiare la natura degli eventuali punti stazionari in \mathbb{R}^2 di

$$f(x, y) = (3 - x - y)(x - 3)(3 - y).$$

La funzione f ha massimo e/o minimo assoluti in \mathbb{R}^2 ?

Svolgimento. La prima parte dell'esercizio richiede di determinare gli (eventuali: possono anche non essercene) punti stazionari di f e di dire se essi sono punti di massimo locale, punti di minimo locale o punti di sella.

Notiamo innanzitutto che f è di classe C^∞ (è infatti un polinomio) e quindi ammette tutte le derivate, continue, di qualunque ordine.

I punti stazionari, per definizione, sono quelli che risolvono il sistema $\nabla f(x, y) = 0$ ovvero

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = 0, \\ \frac{\partial f}{\partial y}(x, y) = 0. \end{cases}$$

Tipicamente, questo tipo di sistemi non sono lineari e quindi non c'è un metodo standard, universale, per risolverli. Occorre lavorare “direttamente con le mani” caso per caso.

Nel nostro caso il sistema consiste in¹⁶

$$\begin{cases} (6 - 2x - y)(3 - y) = 0 \\ (x + 2y - 6)(x - 3) = 0, \end{cases}$$

che porge i seguenti punti stazionari $(3, 0)$, $(0, 3)$, $(3, 3)$, $(2, 2)$.

La matrice Hessiana di f , nel generico punto (x, y) è

$$H_f(x, y) = \begin{pmatrix} 2y - 6 & 2x + 2y - 9 \\ 2x + 2y - 9 & 2x - 6 \end{pmatrix},$$

da cui, nei quattro punti stazionari

$$\begin{aligned} H_f(3, 0) &= \begin{pmatrix} -6 & -3 \\ -3 & 0 \end{pmatrix}, & H_f(0, 3) &= \begin{pmatrix} 0 & -3 \\ -3 & -6 \end{pmatrix}, \\ H_f(3, 3) &= \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}, & H_f(2, 2) &= \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix}. \end{aligned}$$

Con un facile studio del segno delle matrici, si ottiene che $(3, 0)$, $(0, 3)$, $(3, 3)$ sono punti di sella e che $(2, 2)$ è un punto di massimo locale stretto.

Per rispondere alla domanda sui massimi e i minimi assoluti, proviamo, per esempio, a vedere come si comporta la funzione quando il punto va all'infinito lungo la retta $y = x$ bisettrice del primo e del terzo quadrante. Calcoliamo quindi i seguenti limiti

$$\begin{aligned} \lim_{x \rightarrow +\infty} f(x, x) &= \lim_{x \rightarrow +\infty} (2x - 3)(x - 3)^2 = +\infty, \\ \lim_{x \rightarrow -\infty} f(x, x) &= \lim_{x \rightarrow -\infty} (2x - 3)(x - 3)^2 = -\infty, \end{aligned}$$

da cui si deduce che f non ha massimo e minimo assoluti su \mathbb{R}^2 , in quanto assume valori arbitrariamente grandi e arbitrariamente piccoli (piccoli in senso algebrico)¹⁷.

Osservazione 2.31 *Notiamo che la funzione f dell'esempio ha un unico punto di massimo locale, che non è un punto di massimo assoluto, e inoltre essa non ha punti di minimo locale. Questo fatto, in dimensione 1, per una funzione continua è impossibile! (provare a disegnarne una per credere). Il grafico di f è rappresentato in Figura 12.*

¹⁶In questo particolare caso conviene forse non svolgere i prodotti (togliere le parentesi) e invece derivare usando le regole della derivazione del prodotto. Così da ottenere dei fattori già pronti per calcolarne gli zeri. Si può comunque, ovviamente, anche moltiplicare e poi calcolare le derivate.

¹⁷Del resto, essendo \mathbb{R}^2 aperto, se f avesse punti di massimo e di minimo assoluti, questi dovrebbero essere punti stazionari (in quanto punti interni). Ma già sappiamo che non ci sono punti stazionari di minimo. C'è un punto stazionario di massimo (il punto $(2, 2)$) e quindi, se un punto di massimo assoluto c'è, esso deve essere $(2, 2)$. Ma $f(2, 2) = 1$, mentre, ad esempio, $f(4, 4) = 5 > 1$ e 1 quindi non è un massimo assoluto

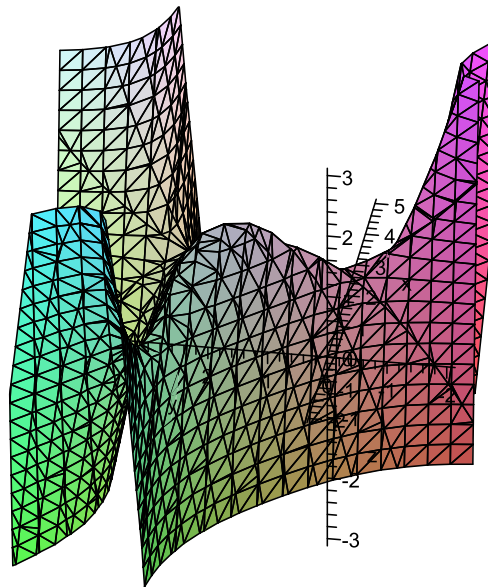


Figure 12: Funzione con un massimo e tre selle.

2.5.7 Massimi e minimi vincolati

Nel paragrafo precedente abbiamo studiato i massimi e i minimi liberi, cioè abbiamo dato dei criteri per stabilire la natura di un punto stazionario interno al dominio della funzione.

Il problema dei massimi e minimi assoluti in generale è un po' diverso. Innanzitutto bisogna essere sicuri che il massimo e il minimo di f esistano sul dominio (nell'esempio qui sopra, abbiamo visto che f non ha né massimo né minimo assoluti sul suo dominio che è \mathbb{R}^n (in generale questo avviene già in dimensione 1: una funzione, di solito, non ha né massimo né minimo assoluti.))

Il seguente teorema dà una condizione sufficiente affinché il massimo e il minimo assoluti di f esistano

Teorema 2.32 (*Weierstrass.*) *Sia $f : C \rightarrow \mathbb{R}^n$ una funzione continua con $C \subseteq \mathbb{R}^n$ chiuso e limitato (in una parola: compatto). Allora f ammette massimo e minimo assoluti su C , cioè esistono $\bar{x}, \bar{y} \in C$ tali che*

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

Se quindi abbiamo una funzione continua su un chiuso e limitato, siamo allora sicuri che il massimo e il minimo assoluti ci sono.¹⁸ Il problema è allora trovarli, e trovare i punti in cui essi sono raggiunti.

¹⁸Attenzione, il teorema di Weierstrass è solo una condizione sufficiente, ma non necessaria. Il massimo e il minimo assoluti possono ben esistere anche senza continuità della f e/o chiusura del dominio e/o limitatezza del dominio.

Supponiamo di voler cercare un punto di massimo assoluto (per il minimo si ragiona in modo analogo)¹⁹, forti del fatto che (ad esempio per il Teorema di Weierstrass) già sappiamo che esso esiste. Il ragionamento è il seguente: se un punto di massimo assoluto è interno al dominio, allora deve essere anche un punto di massimo locale (c'è tutta una palletta in cui è il valore più grande), quindi deve essere un punto stazionario. Se invece un punto di massimo assoluto sta sul bordo del dominio, allora non possiamo concludere che sicuramente è un punto stazionario (lì le derivate potrebbe benissimo non annullarsi). Se però ci dimentichiamo per il momento dell'interno del dominio e consideriamo solo i valori di f assunti sul bordo (cioè la "restrizione di f al bordo"), allora, a maggior ragione, questo punto deve essere anche di massimo per f sul bordo (cioè il valore di f è anche il più alto tra tutti quelli ottenuti da f sul bordo).

La strategia è la seguente: si cercano tutti i punti stazionari interni e si cercano i punti di massimo e di minimo per f ristretta al bordo. I punti di massimo e di minimo assoluti per f (che già sappiamo esistere) devono per forza essere tra questi. Questi punti così trovati sono (o almeno dovrebbero essere) sensibilmente di meno degli infiniti punti del dominio. Sperabilmente sono addirittura in numero finito. Si confrontano i valori di f su questi punti e si prendono quelli con valore più piccolo (minimo) e quelli con valore più grande (massimo).

Sorge però subito un problema. Come trovare il massimo e il minimo di f ristretta al bordo?

Se siamo nel piano e il dominio è una semplice figura geometrica (un quadrato, un rettangolo o un triangolo, cioè con bordo formato da una spezzata di segmenti) un metodo ovvio è quello di studiare la funzione f ristretta ai vari segmenti che compongono il bordo della figura. La funzione ristretta è funzione di una sola variabile il cui metodo di studio è ben noto²⁰. Similmente, sempre nel piano, se una parte del bordo del dominio è, ad esempio una porzione della circonferenza di centro l'origine, raggio $R > 0$, compreso tra gli angoli θ_1, θ_2 (in radianti), allora esso si descrive con la funzione, dipendente dall'angolo variabile θ

$$\varphi(\theta) = (R \cos \theta, R \sin \theta), \quad \theta \in [\theta_1, \theta_2].$$

La funzione f ristretta a questa parte di bordo, diventa allora la funzione della variabile reale θ

$$\psi(\theta) = f(\varphi(\theta)) = f(R \cos \theta, R \sin \theta), \quad \theta \in [\theta_1, \theta_2]$$

il cui studio, di solito, non comporta difficoltà.²¹ Se $\bar{\theta} \in [\theta_1, \theta_2]$ risulta essere di massimo assoluto per ψ , allora il punto $\varphi(\bar{\theta}) = (R \cos \bar{\theta}, R \sin \bar{\theta})$ è di massimo per f ristretta a questo pezzo di bordo.

¹⁹Attenzione, nessuno ha mai detto che il massimo, come il minimo, assoluto sia raggiunto in un solo punto.

²⁰Per questo caso rimandiamo senz'altro agli esercizi fatti in classe.

²¹Ad esempio, se $f(x, y) = 3x - xy + 4$, allora $\psi(\theta) = 3R \cos \theta - R^2 \cos \theta \sin \theta + 4$.

Un caso diverso, indipendentemente dalla dimensione n , è quando il bordo del dominio C è dato da un cosiddetto “vincolo di uguaglianza”, cioè dal luogo degli zeri di una funzione $g : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\partial C = \left\{ x \in \mathbb{R}^n \mid g(x) = 0 \right\}.$$

In questo caso, sotto opportune ipotesi, si può applicare il metodo dei moltiplicatori di Lagrange, il cui utilizzo si basa sul seguente teorema.

Teorema 2.33 *Siano $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ due funzioni di classe C^1 e consideriamo l'insieme*

$$S = \left\{ x \in \mathbb{R}^n \mid g(x) = 0 \right\}.$$

Allora se $\bar{x} \in S$ è un punto di massimo (o di minimo) assoluto per f su S , cioè

$$f(x) \leq f(\bar{x}) \quad \forall x \in S \quad (\text{o } f(x) \geq f(\bar{x}) \quad \forall x \in S),$$

(cioè \bar{x} è un punto di massimo (o minimo) per f vincolata a S), e se \bar{x} non è un punto singolare per S (cioè se $\nabla g(\bar{x}) \neq 0$), allora i due gradienti $\nabla g(\bar{x})$ e $\nabla f(\bar{x})$ sono paralleli. In altre parole, esiste un numero reale λ (detto moltiplicatore) tale che

$$\nabla f(\bar{x}) = \lambda \nabla g(\bar{x}), \quad \text{ovvero} \quad \frac{\partial f}{\partial x_i}(\bar{x}) = \lambda \frac{\partial g}{\partial x_i}(\bar{x}) \quad \forall i = 1, \dots, n.$$

Pertanto i punti di massimo e di minimo di f sul bordo del dominio $\partial C = \{g = 0\}$ vanno cercati tra gli eventuali punti singolari $\nabla g(x) = 0$ e tra quelli che, per un qualche $\lambda \in \mathbb{R}$, risolvono il sistema

$$\left\{ \begin{array}{l} \frac{\partial f}{\partial x_1}(x) = \lambda \frac{\partial g}{\partial x_1}(x), \\ \frac{\partial f}{\partial x_2}(x) = \lambda \frac{\partial g}{\partial x_2}(x), \\ \dots \\ \frac{\partial f}{\partial x_n}(x) = \lambda \frac{\partial g}{\partial x_n}(x), \\ g(x) = 0, \\ \nabla g(x) \neq 0. \end{array} \right. \quad (2.8)$$

I valori di f sugli eventuali punti così trovati, assieme agli eventuali punti singolari e agli eventuali stazionari interni, vanno tutti confrontati per determinare il massimo e il minimo assoluti di f sul dominio.

Osservazione 2.34 *Le soluzioni del sistema (2.8), in realtà sono $n + 1$ -uple di numeri reali $(x, \lambda) = (x_1, x_2, \dots, x_n, \lambda)$. A noi però del valore di λ interessa poco: ci basta l'esistenza e poi consideriamo solo le prime n variabili $x = (x_1, \dots, x_n)$.*

Esempio 2.35 Determinare, se esistono, i punti più vicini all'origine (e le loro distanze dall'origine)

$$S = \left\{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 + x + y + z = 0 \right\}.$$

*Svolgimento.*²² La distanza dall'origine in \mathbb{R}^3 è data da $\sqrt{x^2 + y^2 + z^2}$. Poiché la distanza è, per definizione, positiva, allora massimizzare/minimizzare la distanza o il suo quadrato è equivalente, nel senso che ambedue i procedimenti porgono i medesimi punti. Quindi dobbiamo studiare la seguente funzione su S (ricordandoci però che i valori delle distanze sono le radici quadrate dei valori di tale funzione)

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

L'insieme S è chiuso perché dato da un vincolo di uguaglianza con una funzione continua²³:

$$g(x, y, z) = x^2 + y^2 + z^2 + x + y + z, \quad S = \{g = 0\}.$$

È anche limitato perché si ha²⁴

$$0 = x^2 + y^2 + z^2 + x + y + z \geq \|(x, y, z)\|^2 - |x| - |y| - |z| \geq \|(x, y, z)\|^2 - \sqrt{3}\|(x, y, z)\|,$$

e quindi non possono esserci, in S , punti con norma “troppo” grande, in cui la disuguaglianza non potrebbe essere verificata. Quindi S è compatto e f , essendo continua, ha massimo e minimo assoluti su S .

Cerchiamo gli eventuali punti singolari di S (sui quali il metodo dei moltiplicatori di Lagrange non si applica).

$$\nabla g(x, y, z) = (2x + 1, 2y + 1, 2z + 1) = (0, 0, 0) \iff (x, y, z) = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right) \notin S.$$

Quindi non ci sono punti singolari su S ²⁵.

Studiamo il sistema

²²Qui è richiesto di studiare un problema solo sul vincolo $\{g = 0\}$ con il metodo dei moltiplicatori di Lagrange, e non anche nella parte di spazio eventualmente racchiusa dal vincolo stesso. Questo secondo caso consisterebbe nello studiare i massimi e i minimi sull'insieme $\{g \leq 0\}$ che è formato dalla parte interna $\{g < 0\}$ e dal bordo $\{g = 0\}$. In pratica, in questo esercizio, ci è richiesto solamente di studiare la funzione sul bordo.

²³In generale, se $g : \mathbb{R}^n \rightarrow \mathbb{R}$ è continua, allora i sottoinsiemi di \mathbb{R}^n del tipo $\{g \leq 0\}$ o $\{g = 0\}$ sono chiusi. Invece, in generale, gli insiemi del tipo $\{g < 0\}$ sono aperti.

²⁴Notare che, se $a, b, c > 0$ allora $a + b + c \leq \sqrt{3}\sqrt{a^2 + b^2 + c^2}$. Infatti, elevando al quadrato ambedue i membri e portando tutto a destra si ottiene la disuguaglianza $0 \leq 2a^2 + 2b^2 + 2c^2 - 2ab - 2ac - 2bc = (a - b)^2 + (a - c)^2 + (b - c)^2$, che è verificata, in quanto somma di quadrati.

²⁵Il punto trovato non sta in S perché non verifica l'equazione $g = 0$.

$$\begin{cases} 2x = \lambda(2x + 1) \\ 2y = \lambda(2y + 1) \\ 2z = \lambda(2z + 1) \\ x^2 + y^2 + z^2 + x + y + z = 0. \end{cases}$$

Raccogliendo la x , la y e la z nelle prime tre equazioni, si ha

$$\begin{cases} 2(1 - \lambda)x = \lambda \\ 2(1 - \lambda)y = \lambda \\ 2(1 - \lambda)z = \lambda \\ x^2 + y^2 + z^2 + x + y + z = 0. \end{cases}$$

Se $\lambda = 0$ allora otteniamo immediatamente il punto $(0, 0, 0)$. Se $\lambda = 1$, allora, dalle prime tre equazioni, si avrebbe $0 = 1$ e quindi se ne deduce che $\lambda = 1$ non è ammissibile. Quindi, per $\lambda \neq 1$ e $\lambda \neq 0$, possiamo dividere le prime tre equazioni per $2(1 - \lambda)$ e ottenere

$$x = y = z = \frac{\lambda}{2(1 - \lambda)} =: c \neq 0.$$

Ora, questo porge una soluzione se $g(c, c, c) = 0$ ovvero se (ricordare che $c \neq 0$)

$$3c^2 + 3c = 0 \implies c = -1,$$

al cui valore corrisponderebbe il valore (di poco interesse a questo punto) $\lambda = 2/3$.

Abbiamo ottenuto quindi un secondo punto $(-1, -1, -1) \in S$. Il massimo e il minimo assoluti di f sono necessariamente assunti tra questi due punti. Poiché $f(0, 0, 0) = 0$ e $f(-1, -1, -1) = 3$, ne segue che il punto su S di minima distanza dall'origine è l'origine stessa, con distanza nulla (!), e che il punto su S di massima distanza dall'origine è $(-1, -1, -1)$, con distanza $\sqrt{3}$.

2.6 Funzioni di più variabili a valori vettoriali

In questo paragrafo ci occupiamo di funzioni di più variabili il cui codominio però può avere dimensione maggiore di 1. In altre parole, ci occupiamo di funzioni

$$f : A \rightarrow \mathbb{R}^m, \quad x \mapsto f(x) = (f_1(x), f_2(x), \dots, f_m(x)),$$

dove $A \subseteq \mathbb{R}^n$ è un aperto. Le m funzioni

$$f_i : A \rightarrow \mathbb{R}, \quad x \mapsto f_i(x) \quad i = 1, 2, \dots, m$$

sono dette le funzioni componenti di f , sono funzioni di più variabili scalari (come quelle dei paragrafi precedenti) e indicano, per ogni $x \in A$, il valore della j -esima componente di $f(x) \in \mathbb{R}^m$. Ad esempio, data la funzione

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^4 \quad (x, y, z) \mapsto (x - y, 2x - yz, \cos(x + y + z), xyz), \quad (2.9)$$

si ha $f = (f_1, f_2, f_3, f_4)$ dove

$$f_1(x, y, z) = x - y, \quad f_2(x, y, z) = 2x - yz, \quad f_3(x, y, z) = \cos(x + y + z), \quad f_4(x, y, z) = xyz.$$

2.6.1 Continuità, derivabilità e differenziabilità.

Proposizione 2.36 *Siano $A \subseteq \mathbb{R}^n$ un aperto e $f : A \rightarrow \mathbb{R}^m$ una funzione. Allora f si dice continua in $x \in A$ (rispettivamente: derivabile, differenziabile) se tutte le sue m funzioni componenti sono continue (rispettivamente: derivabili, differenziabili) in x . Analogamente si dice che f è continua, derivabile o differenziabile in A se lo è in ogni punto $x \in A$.²⁶*

In particolare, la derivata (parziale o direzionale) di una funzione a valori vettoriali, è quel vettore che ha per componenti le derivate (parziali o direzionali) delle singole componenti. Ad esempio, riprendendo la funzione (2.9), si ha

$$\frac{\partial f}{\partial x}(x, y, z) = (1, 2, -\sin(x + y + z), yz) \in \mathbb{R}^4.$$

Definizione 2.37 *Sia $f : A \rightarrow \mathbb{R}^m$ derivabile in $x \in A$, con $A \subseteq \mathbb{R}^n$ aperto. Si dice matrice Jacobiana di f in x , e si denota con $J_f(x)$, la matrice $m \times n$ il cui generico coefficiente a_{ij} consiste nella derivata parziale di f_i rispetto a x_j , calcolata in x . Cioè*

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Osservazione 2.38 *Si noti la differenza tra matrice Hessiana e matrice Jacobiana. La prima è una matrice quadrata che coinvolge le derivate parziali seconde di una funzione scalare, la seconda è una matrice rettangolare (non necessariamente quadrata) che coinvolge le derivate parziali prime delle componenti di una funzione a valori vettoriali.*

Se $m = 1$ allora la funzione è scalare e la matrice Jacobiana è $1 \times n$, cioè è una riga e, ovviamente, coincide con il gradiente delle funzione.

2.6.2 Cambiamenti di variabili in \mathbb{R}^n

In questo paragrafo ci occupiamo solamente di funzioni a valori vettoriali in cui dominio e codominio hanno la medesima dimensione, n . Enunciamo il seguente teorema che dà una condizione sufficiente affinché, nell'intorno di un punto, una funzione sia invertibile²⁷.

²⁶In realtà, per ciascuno dei concetti di continuità, derivabilità, differenziabilità, si dà una definizione, tramite limiti e rapporti incrementali, indipendente dalle funzioni componenti. L'equivalenza con le medesime proprietà delle funzioni componenti è un risultato susseguente (di facile dimostrazione).

²⁷Per analogia, si pensi ad una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$. Se $f'(x_0) \neq 0$ allora, intorno a x_0 la funzione è strettamente crescente o strettamente decrescente (a seconda del segno di $f'(x_0)$) e quindi lì intorno è invertibile. Notiamo anche che in questo caso $n = 1$ e quindi la matrice Jacobiana è una matrice 1×1 il cui unico coefficiente è $f'(x_0)$ che ovviamente coincide anche con il suo determinante.

Teorema 2.39 (Teorema di inversione locale.) Siano $A \subseteq \mathbb{R}^n$ aperto, $\bar{x} \in A$ un punto e $f : A \rightarrow \mathbb{R}^n$ una funzione di classe C^1 , cioè con tutte le componenti di classe C^1 . Supponiamo che la matrice Jacobiana di f in \bar{x} , $J_f(\bar{x})$, sia non singolare, cioè che abbia determinante non nullo, cioè che sia invertibile (come matrice quadrata). Allora la funzione f è localmente invertibile attorno a \bar{x} , cioè esistono una palla U centrata in \bar{x} e una palla V centrata in $f(\bar{x})$ tale che la restrizione di f a U

$$f|_U : U \rightarrow V \quad x \mapsto f(x),$$

è biiettiva, ovvero invertibile. Inoltre l'inversa $f^{-1} : V \rightarrow U$ è pure lei di classe C^1 .

Si dice che f è un diffeomorfismo intorno a \bar{x} .

Definizione 2.40 Siano $A \subseteq \mathbb{R}^n$ aperto e $f : A \rightarrow \mathbb{R}^n$ una funzione di classe C^1 la cui matrice Jacobiana ha determinante non nullo in ogni punto di A , escluso al più un insieme di punti di misura (n -dimensionale) nulla²⁸.

Allora f si dice una trasformazione regolare di coordinate.

Osservazione 2.41 In virtù del teorema di inversione locale, possiamo dire che, a meno di un insieme formato da “pochi” punti, una trasformazione regolare di coordinate è localmente invertibile attorno ad ogni punto. E quindi tutto ciò che succede attorno al punto $\bar{x} \in A$ (ad esempio la rappresentazione mediante un particolare sistema di coordinate), avviene pari pari attorno al punto $f(\bar{x})$, purché reinterpretiamo il tutto (per esempio la rappresentazione mediante opportune coordinate) alla luce del “transfer” f (il cambiamento di coordinate, appunto).

Esempio 2.42 (Trasformazione lineare di coordinate.) Sia M una matrice quadrata $n \times n$ non singolare ($\det M \neq 0$). Consideriamo la funzione

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto Mx,$$

dove Mx indica l'usuale moltiplicazione matrice per vettore (righe per colonna). Un semplice conto mostra che $J_f(x) = M$ per ogni $x \in \mathbb{R}^n$, da cui f è una trasformazione regolare di coordinate.

Ad esempio, una rotazione nel piano di un angolo α in senso antiorario è una trasformazione (lineare) di coordinate data da

$$f(x, y) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

e il determinante è $\cos^2 \alpha + \sin^2 \alpha = 1 \neq 0$.

²⁸Ad esempio: un sottoinsieme del piano ha misura 2-dimensionale nulla se ha area nulla; un sottoinsieme dello spazio a tre dimensioni ha misura 3-dimensionale nulla se ha volume nullo. Solo questi due casi, $n = 2$ e $n = 3$, in realtà, ci interesseranno in seguito.

Esempio 2.43 (*Coordinate polari nel piano*) Rappresentiamo ogni punto (x, y) del piano \mathbb{R}^2 (escluso l'origine), tramite la sua distanza ρ dall'origine, $\rho = \sqrt{x^2 + y^2} \geq 0$ e tramite l'angolo $\theta \in [0, 2\pi[$ formato dalla retta delle ascisse positive e dal vettore (freccia spiccata dall'origine) (x, y) , contato in senso antiorario. Abbiamo quindi le relazioni

$$\begin{cases} x = \rho \cos \theta, \\ y = \rho \sin \theta. \end{cases}$$

L'origine $(0, 0)$, invece, la identifichiamo ponendo semplicemente $\rho = 0$, indipendentemente dall'angolo (che in realtà, in questo caso, non esiste.)

Abbiamo quindi una funzione

$$\varphi : [0, +\infty[\times [0, 2\pi[\rightarrow \mathbb{R}^2 \quad (\rho, \theta) \mapsto (\rho \cos \theta, \rho \sin \theta) = (x, y)$$

e notiamo che questa ha senso anche per $\rho = 0$ e vale $\varphi(0, \theta) = (0, 0)$ qualunque sia θ . La matrice Jacobiana di questa funzione (le cui variabili in entrata sono ρ e θ) è

$$J_\varphi(\rho, \theta) = \begin{pmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{pmatrix},$$

il cui determinante è $\rho \geq 0$ e, in particolare, esso è nullo solo sul segmento del piano $\rho\theta$: $\{(0, \theta) \mid \theta \in [0, 2\pi[\}$ che, ovviamente ha area nulla. Quindi φ è una trasformazione regolare di coordinate. Le coordinate (ρ, θ) si chiamano coordinate polari nel piano.

Ad esempio, nel piano cartesiano xy , il cerchio di centro l'origine e raggio 1, dato dalla disuguaglianza $x^2 + y^2 - 1 \leq 0$, è rappresentato in coordinate polari dall'insieme (cioè proviene tramite φ dall'insieme)

$$\left\{ (\rho, \theta) \mid 0 \leq \rho \leq 1, 0 \leq \theta \leq 2\pi \right\},$$

che è un rettangolo.

Esempio 2.44 (*Coordinate cilindriche nello spazio.*) In \mathbb{R}^3 rappresentiamo ogni punto (x, y, z) tramite la sua quota z e tramite le coordinate polari nel piano $z = 0$ del punto (proiezione) $(x, y, 0)$. Pertanto abbiamo una funzione

$$\varphi : [0, +\infty[\times [0, 2\pi[\times \mathbb{R} \rightarrow \mathbb{R}^3 \quad (\rho, \theta, z) \mapsto (\rho \cos \theta, \rho \sin \theta, z) = (x, y, z).$$

La matrice Jacobiana è

$$J_\varphi(\rho, \theta, z) = \begin{pmatrix} \cos \theta & -\rho \sin \theta & 0 \\ \sin \theta & \rho \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

il cui determinante è ancora uguale a $\rho \geq 0$ che, questa volta, si annulla solo sui punti dell'asse z che, ovviamente, ha volume nullo. Quindi φ è una trasformazione regolare di coordinate. Le coordinate (ρ, θ, z) si chiamano coordinate cilindriche in \mathbb{R}^3 .

Ad esempio, nello spazio cartesiano \mathbb{R}^3 , il cilindro (pieno) retto, di asse l'asse z , raggio 2 e compreso tra le quote $z = -1$, $z = 2$, descritto dalle disuguaglianze $x^2 + y^2 \leq 4$, $-1 \leq z \leq 2$, si rappresenta in coordinate cilindriche tramite l'insieme (cioè proviene tramite φ dall'insieme)

$$\left\{ (\rho, \theta, z) \mid 0 \leq \rho \leq 2, 0 \leq \theta \leq 2\pi, -1 \leq z \leq 2 \right\},$$

che è un parallelepipedo.

Esempio 2.45 (*Coordinate sferiche nello spazio.*) In \mathbb{R}^3 , rappresentiamo ogni punto (x, y, z) (escluso l'origine) tramite la sua distanza dall'origine $\rho = \sqrt{x^2 + y^2 + z^2} > 0$, l'angolo $\theta \in [0, 2\pi[$ formato nel piano $z = 0$ dalla retta delle ascisse positive e il vettore $(x, y, 0)$, contato in senso antiorario e l'angolo $\varphi \in [0, \pi]$ formato dall'asse delle z positive e il vettore (x, y, z) , contato a partire dall'alto.²⁹ Abbiamo quindi una funzione

$$\psi : [0, +\infty[\times [0, 2\pi[\times [0, \pi] \rightarrow \mathbb{R}^3, \quad (\rho, \theta, \varphi) \mapsto (\rho \sin \varphi \cos \theta, \rho \sin \varphi \sin \theta, \rho \cos \varphi) = (x, y, z),$$

e notiamo che essa ha senso anche per $\rho = 0$, avendosi $\psi(0, \theta, \varphi) = (0, 0, 0)$, qualunque siano θ e φ . La matrice Jacobiana di ψ è

$$J_\psi(\rho, \theta, \varphi) = \begin{pmatrix} \sin \varphi \cos \theta & -\rho \sin \varphi \sin \theta & \rho \cos \varphi \cos \theta \\ \sin \varphi \sin \theta & \rho \sin \varphi \cos \theta & \rho \cos \varphi \sin \theta \\ \cos \varphi & 0 & -\rho \sin \varphi \end{pmatrix},$$

il cui determinante è $-\rho^2 \sin \varphi$ che si annulla solo per $\rho = 0$ e per $\varphi = 0, \pi$, cioè su terne del tipo $(0, \theta, \varphi)$ (con θ, φ qualunque), $(\rho, \theta, 0)$, (ρ, θ, π) (con ρ, θ qualunque). È facile convincersi che questi punti formano un insieme di volume nullo. Quindi ψ è una trasformazione regolare di coordinate. Le coordinate (ρ, θ, φ) si chiamano coordinate sferiche in \mathbb{R}^3 .

Ad esempio, nello spazio cartesiano \mathbb{R}^3 , il pezzo di sfera (piena) di centro l'origine, raggio 3 e contenuta nel primo ottante $(x, y, z \geq 0)$, descritto dalle disequazioni $x^2 + y^2 + z^2 \leq 9$, $x, y, z \geq 0$, si rappresenta in coordinate sferiche con l'insieme (cioè proviene tramite ψ dall'insieme)

$$\left\{ (\rho, \theta, \varphi) \mid 0 \leq \rho \leq 3, 0 \leq \theta \leq \frac{\pi}{2}, 0 \leq \varphi \leq \frac{\pi}{2} \right\},$$

che è un parallelepipedo.

²⁹Con notazioni geografiche, θ sarebbe la longitudine e φ la latitudine, contata però dal polo nord anziché dall'equatore.

3 Calcolo Integrale per Funzioni di più Variabili

Ricordiamo il seguente fatto noto. Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione continua, dove $[a, b]$ è un intervallo chiuso e limitato di \mathbb{R} . Supponiamo $f(x) \geq 0$ per ogni $x \in [a, b]$. Qual è l'area della regione sottesa dal grafico di f sopra l'intervallo $[a, b]$? Tale regione di piano è l'insieme

$$\left\{ (x, y) \in \mathbb{R}^2 \mid x \in [a, b], 0 \leq y \leq f(x) \right\}.$$

È noto che l'area vale

$$\int_a^b f(x) dx,$$

l'integrale di f sull'intervallo $[a, b]$. È altresì noto che, in generale, l'integrale di una funzione continua a valori reali su un intervallo chiuso e limitato è l'area “con segno” sottesa dal grafico della funzione.

3.1 Integrale doppio di una funzione di due variabili.

Gli intervalli sono gli insiemi più semplici della retta reale³⁰. Non è sorprendente se diciamo che, nel piano, gli insiemi “più semplici” sono i rettangoli. Sia quindi $R = [a, b] \times [c, d] \subseteq \mathbb{R}^2$ un rettangolo del piano e sia $f : R \rightarrow \mathbb{R}$ una funzione continua. Se esiste un concetto di integrale di f sul rettangolo R , questo, per analogia con il caso unidimensionale, dovrebbe fornire il volume (con segno) della regione di spazio tridimensionale sottesa dal grafico di f sopra R . Supponiamo, per semplicità, $f(x, y) \geq 0$ per ogni $(x, y) \in R$. Allora la regione di spazio sottesa dal grafico è

$$V = \left\{ (x, y, z) \in \mathbb{R}^3 \mid (x, y) \in R, 0 \leq z \leq f(x, y) \right\}.$$

Quanto vale il volume di V ? Fissiamo un valore $\bar{x} \in [a, b]$ e consideriamo la restrizione di f al segmento (parallelo all'asse y) $\{(\bar{x}, y) \mid c \leq y \leq d\}$: $\varphi_{\bar{x}}(y) = f(\bar{x}, y)$ $y \in [c, d]$. Consideriamo poi la “lamina” $L(\bar{x})$ individuata, all'interno del solido V , da questo segmento come base e dai corrispondenti valori di f come vertice:

$$L(\bar{x}) = \left\{ (y, z) \in \mathbb{R}^2 \mid c \leq y \leq d, 0 \leq z \leq \varphi_{\bar{x}}(y) \right\}.$$

L'area di $L(\bar{x})$ è ovviamente data da

$$\int_c^d \varphi_{\bar{x}}(y) dy.$$

³⁰tra quelli su cui abbia senso fare dell'analisi...

È intuitivo che il volume di V deve essere uguale alla “somma” delle aree di tutte le lamine $L(x)$ al variare di $x \in [a, b]$. Ma queste aree sono infinite (più che numerabili!): una per ogni $x \in [a, b]$. Quindi la “somma” va sostituita con un integrale:

$$\text{Volume}(V) : \int_a^b \text{Area}(L(x)) dx = \int_a^b \left(\int_c^d \varphi_x(y) dy \right) dx = \int_a^b \left(\int_c^d f(x, y) dy \right) dx. \quad (3.10)$$

In modo del tutto analogo, si ottiene anche

$$\text{Volume}(V) = \int_c^d \left(\int_a^b f(x, y) dx \right) dy.$$

Nell'ultimo membro di (3.10), la notazione significa che prima si calcola l'integrale tra c e d della funzione della sola y , cioè con x fissato. Questa operazione porge una quantità che non dipende più da y (avendo integrato in y tra c e d , la variabile y è stata “saturata”), ma dipende solo da x : è l'area della lamina. Si esegue quindi l'integrale di tale quantità con x che varia tra a e b . Le regole di integrazione nei due integrali (quello “interno” e quello “esterno”) sono le usuali regole di integrazione di una funzione di una variabile, con l'accortezza, per esempio per l'integrale interno in (3.10), di integrare rispetto a y considerando x come una costante.

Esempio 3.1 Consideriamo il rettangolo $R = [0, 1] \times [1, 2]$ e la funzione $f(x, y) = xy + y^2$. Allora il volume della regione V di spazio sottesa dal grafico di f sopra il rettangolo R è

$$\begin{aligned} \text{Volume}(V) &= \int_0^1 \left(\int_1^2 (xy + y^2) dy \right) dx = \int_0^1 \left(\left(x \frac{y^2}{2} + \frac{y^3}{3} \right) \Big|_{y=1}^{y=2} \right) dx \\ &= \int_0^1 \left(\frac{3}{2}x + \frac{7}{3} \right) dx = \left(\frac{3}{4}x^2 + \frac{7}{3}x \right) \Big|_0^1 = \frac{37}{12}. \end{aligned}$$

Ovviamente (verificarlo per esercizio) vale anche

$$\int_1^2 \left(\int_0^1 (xy + y^2) dx \right) dy = \frac{37}{12}.$$

In generale, indipendentemente dal segno di f , la formula

$$\int_a^b \left(\int_c^d f(x, y) dy \right) dx \quad \left(= \int_c^d \left(\int_a^b f(x, y) dx \right) dy \right)$$

porge il volume con segno della regione sottesa dal grafico di f sopra il rettangolo $R = [a, b] \times [c, d]$. Normalmente non si scrivono le parentesi tra un integrale e l'altro, ma si scrive, ad esempio, semplicemente

$$\int_a^b \int_c^d f(x, y) dy dx.$$

Con una notazione standard, si scrive anche

$$\int \int_R f(x, y) dx dy,$$

che si legge integrale doppio di f su R (alle volte si scrive un solo simbolo di integrale $\int_R f(x, y) dx dy$).

3.1.1 Domini semplici nel piano

Nel precedente paragrafo abbiamo visto come integrare una funzione continua di due variabili su di un rettangolo del piano. Saper integrare solo sui rettangoli sarebbe forse un po' troppo poco. Abbiamo bisogno di una tecnica per poter integrare anche su insiemi più complessi. Un primo passo verso questo traguardo è l'integrazione sugli insiemi semplici.

Definizione 3.2 Un insieme $A \subseteq \mathbb{R}^2$ si dice semplice rispetto a y se

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid x \in [a, b], \varphi_1(x) \leq y \leq \varphi_2(x) \right\},$$

dove $[a, b]$ è un intervallo chiuso e limitato della retta reale e $\varphi_1, \varphi_2 : [a, b] \rightarrow \mathbb{R}$ sono due funzioni continue tali che $\varphi_1(x) \leq \varphi_2(x)$ per ogni $x \in [a, b]$.

Un insieme $A \subseteq \mathbb{R}^2$ si dice semplice rispetto a x se

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid y \in [c, d], \psi_1(y) \leq x \leq \psi_2(y) \right\},$$

dove $[c, d]$ è un intervallo chiuso e limitato della retta reale e $\psi_1, \psi_2 : [c, d] \rightarrow \mathbb{R}$ sono due funzioni continue tali che $\psi_1(y) \leq \psi_2(y)$ per ogni $y \in [c, d]$.

Un insieme $A \subseteq \mathbb{R}^2$ si dice semplice se è semplice rispetto a x o rispetto a y (almeno uno dei due).

Osservazione 3.3 La definizione “semplice rispetto a y ” dipende dal fatto che, fissato $\bar{x} \in [a, b]$ qualunque, la “fibra” di A sopra \bar{x} , cioè i punti del tipo $(\bar{x}, y) \in A$ è esattamente un segmento del piano, quello descritto da $\{(\bar{x}, y) \mid \varphi_1(\bar{x}) \leq y \leq \varphi_2(\bar{x})\}$. Analogamente per gli insiemi semplici rispetto a x .

È evidente che i rettangoli sono particolari insiemi semplici: sono semplici sia rispetto a x che rispetto a y .

La seguente proposizione segue da ragionamenti simili a quelli fatti nel caso del rettangolo.

Proposizione 3.4 Siano $A \subseteq \mathbb{R}^2$ un insieme semplice e $f : A \rightarrow \mathbb{R}$ una funzione continua. Allora f è integrabile su A (cioè il suo integrale esiste ed è finito³¹) e vale

³¹Non è un definizione proprio rigorosa, ma per i nostri scopi prendiamola per buona...

$$\int \int_A f(x, y) dx dy = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy dx, \quad (3.11)$$

se A è semplice rispetto a y , oppure

$$\int \int_A f(x, y) dx dy = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx dy,$$

se A è semplice rispetto a x .

Osservazione 3.5 Nella proposizione precedente, il significato dei simboli dovrebbe essere ormai ovvio: ad esempio nel secondo membro in (3.11), si intende che, per ogni x fissato si esegue prima l'integrale rispetto alla sola y che varia tra $\varphi_1(x)$ e $\varphi_2(x)$ e poi si integra in x che varia tra a e b . Questa procedura è indicata dall'ordine in cui si scrivono gli integrali (e il dx e il dy) e, come procedura, si chiama integrazioni successive.

Sempre, ad esempio, in (3.11), il primo membro si legge "integrale doppio di f su A ".

Esempio 3.6 Calcolare il seguente integrale doppio

$$\int \int_A (x + y) dx dy,$$

dove

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid y \leq -x^2 + 1, x \geq 0, y \geq 0 \right\}.$$

Svolgimento. L'insieme A consiste di quella porzione di piano nel primo quadrante e posta sotto la parabola di equazione $y = -x^2 + 1$. Esso è quindi un insieme semplice rispetto a y , con $x \in [0, 1]$ e $0 = \varphi_1(x) \leq y \leq \varphi_2(x) = -x^2 + 1$. Pertanto si ha

$$\begin{aligned} \int \int_A (x + y) dx dy &= \int_0^1 \int_0^{-x^2+1} (x + y) dy dx = \int_0^1 \left(\left(xy + \frac{y^2}{2} \right) \Big|_{y=0}^{y=-x^2+1} \right) dx \\ &= \int_0^1 \left(x(-x^2 + 1) + \frac{(-x^2 + 1)^2}{2} \right) dx = \int_0^1 \left(\frac{x^4}{2} - x^3 - x^2 + x + \frac{1}{2} \right) dx = \frac{31}{60}. \end{aligned}$$

3.1.2 Domini regolari del piano

Vediamo qui un'ulteriore categoria di insiemi del piano.

Definizione 3.7 Un insieme $A \subseteq \mathbb{R}^2$ si dice un dominio regolare se è unione finita di insiemi semplici che si intersecano a due a due al più sul bordo. Cioè

$$A = \bigcup_{i=1}^{i=N} A_i,$$

dove N è un numero naturale, gli insiemi A_i sono semplici e, se $i \neq j$,

$$\overline{A_i} \cap \overline{A_j} \subseteq \partial A_i \cap \partial A_j. \quad (3.12)$$

Proposizione 3.8 Sia $A \subseteq \mathbb{R}^2$ un dominio regolare e $f : A \rightarrow \mathbb{R}$ una funzione continua. Allora f è integrabile e

$$\int_A f(x, y) dx dy = \sum_{i=1}^{i=N} \int_{A_i} f(x, y) dx dy, \quad (3.13)$$

dove gli A_i sono i domini semplici come in (3.12)³².

Esempio 3.9 L'insieme

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 4, x \geq 0, y \geq -x + 1 \right\}$$

è un insieme regolare essendo uguale all'unione $A_1 \cup A_2$ dove

$$A_1 = \left\{ (x, y) \in \mathbb{R}^2 \mid x \in [0, 1], -x + 1 \leq y \leq \sqrt{4 - x^2} \right\},$$

$$A_2 = \left\{ (x, y) \in \mathbb{R}^2 \mid x \in [1, 2], 0 \leq y \leq \sqrt{4 - x^2} \right\}$$

sono insiemi semplici.

3.1.3 Cambiamento di variabili negli integrali doppi

Vediamo ora l'ultimo passo della generalizzazione dei domini su cui integrare.

Proposizione 3.10 Sia $\varphi : B \rightarrow \mathbb{R}^2$ una trasformazione regolare di coordinate in \mathbb{R}^2 . Siano poi $A = \varphi(B) \subseteq \mathbb{R}^2$ (immagine di $B \subseteq \mathbb{R}^2$ tramite φ) e $f : A \rightarrow \mathbb{R}$ una funzione integrabile. Denotiamo con (u, v) gli elementi di B (coordinate "nuove") e con (x, y) gli elementi di A (coordinate "vecchie"). Allora la funzione $(u, v) \mapsto f(\varphi(u, v)) |\det J_\varphi(u, v)|$, è integrabile su B e si ha

$$\int \int_A f(x, y) dx dy = \int \int_B f(\varphi(u, v)) |\det J_\varphi(u, v)| du dv. \quad (3.14)$$

La Proposizione 3.10 dice che integrare su A o integrare su B (quando $A = \varphi(B)$) sono di fatto la stessa cosa purché abbiamo l'accortezza di introdurre nell'integrale su B il fattore $|\det J_\varphi(u, v)|$ (valore assoluto del determinante della matrice Jacobiana di φ) detto Jacobiano. Infatti la trasformazione di coordinate φ può aver alterato le misure delle aree e degli angoli tra i segmenti e quindi bisogna tenerne conto introducendo lo Jacobiano che è un "fattore di riscaldamento delle aree"³³.

Il vantaggio della Proposizione 3.10 è evidente: anche se l'insieme A non è un dominio regolare, possiamo ugualmente integrare su A , tramite la formula (3.14), purché A sia l'immagine, tramite una trasformazione di coordinate, di un dominio regolare B .

³²Ovviamente la formula (3.13) è indipendente dalla particolare decomposizione di A in domini semplici che si intersecano solo sul bordo

³³Per analogia al caso unidimensionale, si ricordi che, se $\varphi : [c, d] \rightarrow [a, b]$ è derivabile e monotona (cioè è una trasformazione di coordinate in dimensione 1), allora $\int_a^b f(x) dx = \int_c^d f(\varphi(\xi)) \varphi'(\xi) d\xi$. Il termine $\varphi'(\xi)$ gioca il ruolo dello Jacobiano. Non entriamo nel dettaglio del perché in dimensione 1 non si prenda invece $|\varphi'(\xi)|$.

Esempio 3.11 Calcolare l'area della regione di piano³⁴³⁵

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid 1 \leq x^2 + y^2 \leq 2, x \leq y, -\sqrt{3}x \leq y \right\}$$

L'insieme A è parte della corona circolare di centro l'origine, raggio inferiore 1 e raggio superiore $\sqrt{2}$. Inoltre esso sta sopra le rette $y = x$ e $y = -\sqrt{3}x$. È quindi il pezzo di corona circolare con angolo θ compreso tra $\frac{\pi}{4}$ e $\frac{2}{3}\pi$. Tale insieme è in realtà un dominio regolare, in quanto si può scomporre nell'unione di domini semplici. È però più conveniente passare in coordinate polari. Infatti, in coordinate polari, (ρ, θ) , il nostro insieme è

$$\left\{ (\rho, \theta) \mid 1 \leq \rho \leq \sqrt{2}, \frac{\pi}{4} \leq \theta \leq \frac{2}{3}\pi \right\}.$$

Lo Jacobiano delle coordinate polari è $\rho \geq 0$ e quindi³⁶

$$\text{Area}(A) = \int_A dx dy = \int_{\frac{\pi}{4}}^{\frac{2}{3}\pi} \int_1^{\sqrt{2}} \rho d\rho d\theta = \frac{5}{12}\pi \frac{\rho^2}{2} \Big|_{\rho=1}^{\rho=\sqrt{2}} = \frac{5}{24}\pi.$$

Esempio 3.12 Calcolare

$$\int_A ((x + y - 3)^2 - \sin(y - 2x + 1)) dx dy,$$

dove

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid -1 \leq x + y \leq 1, 0 \leq y - 2x \leq \frac{\pi}{2} \right\}.$$

Svolgimento. Consideriamo il seguente cambiamento di variabili

$$u = x + y, \quad v = y - 2x. \quad (3.15)$$

È evidente che, nelle nuove variabili (u, v) , l'insieme A si scrive

$$A' = \left\{ (u, v) \mid -1 \leq u \leq 1, 0 \leq v \leq \frac{\pi}{2} \right\}.$$

Ora, per applicare la formula del cambiamento di variabili, abbiamo bisogno di una funzione (la trasformazione regolare di coordinate) φ che manda le coordinate nuove (u, v)

³⁴Generalmente, dato un insieme A nel piano, la sua area equivale all'integrale doppio su di esso della funzione costante $f \equiv 1$. Si pensi per analogia, che, in dimensione 1, l'integrale della funzione costante 1 sull'intervallo $[a, b]$ dà come risultato $b - a$ che è proprio la lunghezza dell'intervallo (la misura 1-dimensionale.) Provare per esercizio che l'area della regione di piano sottesa dal grafico della funzione continua $f : [a, b] \rightarrow \mathbb{R}$, che sappiamo essere $\int_a^b f(x) dx$, equivale (mediante integrazioni successive) all'integrale doppio della funzione 1 sull'insieme in questione.

³⁵Ovviamente, quando si integra la costante 1, il numero 1 non si scrive nell'integrale.

³⁶Risultato ovviamente atteso da ben note formule di geometria euclidea...

nelle coordinate vecchie (x, y) . Questa si ottiene semplicemente invertendo le relazioni in (3.15), ovvero esplicitando (x, y) in funzione di (u, v) . Si ha

$$(x, y) = \left(\frac{u - v}{3}, \frac{2u + v}{3} \right) = \varphi(u, v).$$

Il cambiamento φ è quindi lineare, associato alla matrice 2×2

$$\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

che ha determinante $\frac{1}{3} \neq 0$ (e quindi φ è effettivamente una trasformazione regolare di coordinate).

Quindi si ha

$$\begin{aligned} \int_A ((x + y - 3)^2 - \sin(y - 2x + 1)) \, dx dy &= \frac{1}{3} \int_{A'} ((u - 3)^2 - \sin(v + 1)) \, dudv \\ &= \frac{1}{3} \left(\frac{56}{3} + \cos\left(\frac{\pi}{2} + 1\right) - \cos(1) \right). \end{aligned}$$

3.2 Integrale triplo di una funzione a valori scalari

Anche per il caso di funzioni scalari di tre variabili si possono calcolare gli integrali su opportuni insiemi dello spazio tridimensionale. I rettangoli sono ovviamente rimpiazzati dai parallelepipedi, sui quali non è sorprendente che funzioni il metodo delle integrazioni successive. Per quanto riguarda i domini semplici in \mathbb{R}^3 , nel prossimo paragrafo diamo due definizioni che potrebbero ambedue corrispondere alla definizione di dominio semplice rispetto a z . Ci limiteremo a questi due casi, lasciando al lettore le analoghe definizioni, rispetto a x e rispetto a y .

3.2.1 Domini semplici e domini regolari nello spazio tridimensionale

Definizione 3.13 *Un insieme $A \subseteq \mathbb{R}^3$ si dice semplice per fili se esistono un dominio regolare $D \subseteq \mathbb{R}^2$ e due funzioni continue $h_1, h_2 : D \rightarrow \mathbb{R}$, con $h_1(x, y) \leq h_2(x, y)$ per ogni $(x, y) \in D$, tali che*

$$A = \left\{ (x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D, h_1(x, y) \leq z \leq h_2(x, y) \right\}. \quad (3.16)$$

Si dice invece che A è semplice per strati se esistono un intervallo $[a, b] \in \mathbb{R}$ e, per ogni $z \in [a, b]$, un dominio regolare $D_z \subseteq \mathbb{R}^2$ tali che

$$A = \left\{ (x, y, z) \in \mathbb{R}^3 \mid z \in [a, b], (x, y) \in D_z \right\}. \quad (3.17)$$

Diciamo che A è semplice se è semplice per fili o per strati. Infine diciamo che A è un dominio regolare di \mathbb{R}^3 se è unione finita di domini semplici che, a due a due, si intersecano al più sul bordo.

Una volta che sappiamo integrare sui domini semplici, sappiamo integrare anche sui domini regolari tramite l'analogo della formula (3.13).

Proposizione 3.14 *Sia A semplice per fili come in (3.16) e sia $f : A \rightarrow \mathbb{R}$ continua. Allora f è integrabile su A e vale (il membro di sinistra si dice "integrale triplo di f su A ")*

$$\int \int \int_A f(x, y, z) dx dy dz = \int \int_D \left(\int_{h_1(x, y)}^{h_2(x, y)} f(x, y, z) dz \right) dx dy \quad (3.18)$$

Sia A semplice per strati come in (3.17) e sia $f : A \rightarrow \mathbb{R}^3$ continua. Allora f è integrabile su A e vale

$$\int \int \int_A f(x, y, z) dx dy dz = \int_a^b \left(\int \int_{D_z} f(x, y, z) dx dy \right) dz \quad (3.19)$$

Osservazione 3.15 *La formula (3.18) consiste in una cosiddetta integrazione per fili: nell'integrale interno si integra in z , pensando x e y come costanti, con z che varia nell'intervallo (filo verticale) $[h_1(x, y), h_2(x, y)]$ e poi si esegue l'integrale doppio su D in x e y .*

Invece, la formula (3.19) consiste in una cosiddetta integrazione per strati: nell'integrale interno, una volta fissata la quota z , si integra, rispetto alle variabili x e y , sull'insieme (strato orizzontale) D_z , e poi si esegue l'integrale nella sola variabile $z \in [a, b]$.

Al solito le parentesi si possono omettere nelle notazioni, e si può scrivere l'integrale triplo anche con un solo simbolo di integrale.

Esempio 3.16 Calcolare il volume³⁷ della porzione di iperboloide

$$E = \left\{ (x, y, z) \in \mathbb{R}^3 \mid 0 \leq z \leq 1, x^2 + y^2 - z^2 - 1 \leq 0 \right\}.$$

Svolgimento. L'insieme E è un dominio semplice per strati con $z \in [0, 1]$ e

$$D_z = \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq z^2 + 1 \right\}$$

che, per ogni z fissato, consiste nel cerchio (pieno) in \mathbb{R}^2 di centro l'origine e raggio $\sqrt{z^2 + 1}$. In coordinate polari nel piano esso è l'insieme

$$D'_z = \left\{ (\rho, \theta) \mid 0 \leq \theta \leq 2\pi, 0 \leq \rho \leq \sqrt{z^2 + 1} \right\}.$$

Pertanto si ha, integrando per strati,

$$\begin{aligned} \text{Volume}(E) &= \int \int \int_E dx dy dz = \int_0^1 \left(\int \int_{D_z} dx dy \right) dz \\ &= \int_0^1 \left(\int_0^{2\pi} d\theta \int_0^{\sqrt{z^2+1}} \rho d\rho \right) dz = \frac{4}{3}\pi \end{aligned}$$

³⁷Come per gli insiemi del piano l'area è l'integrale doppio della funzione costante 1, così per gli insiemi di \mathbb{R}^3 il volume è l'integrale triplo della costante 1 (e ovviamente il numero 1 non si scrive nell'integrale).

3.2.2 Cambiamento di variabili negli integrali tripli

Come per gli integrali doppi, anche per gli integrali tripli vale l'analogo della formula (3.14) di cambiamento di variabili.

Esempio 3.17 Calcolare

$$\int_A (x - y + z) dx dy dz,$$

dove A è la semisfera (piena) di centro l'origine, raggio 1 e $z \geq 0$.

Svolgimento. In coordinate sferiche A è

$$\left\{ (\rho, \theta, \varphi) \mid 0 \leq \rho \leq 1, 0 \leq \theta \leq 2\pi, 0 \leq \varphi \leq \frac{\pi}{2} \right\}.$$

Il determinante della matrice Jacobiana delle coordinate sferiche è $-\rho^2 \sin \varphi$ che è sempre negativo ($\varphi \in [0, \pi]$ e quindi $\sin \varphi \geq 0$) e quindi il suo valore assoluto è $\rho^2 \sin \varphi$.
Ne segue

$$\int_A (x - y + z) = \int_0^{\frac{\pi}{2}} \int_0^{2\pi} \int_0^1 (\rho \sin \varphi \cos \theta - \rho \sin \varphi \sin \theta + \rho \cos \varphi) \rho^2 \sin \varphi \, d\rho d\theta d\varphi,$$

il cui calcolo si lascia per esercizio.