Luciano Battaia

Appunti per un corso di matematica

Teoria ed esercizi

Università Ca' Foscari di Venezia - Dipartimento di Economia

Appunti per un corso di matematica
Teoria ed esercizi

Luciano Battaia

Università Ca' Foscari di Venezia - Dipartimento di Economia

Versione 2.1 del 23 febbraio 2017

Se gli allievi non capiscono, il torto è dell'insegnante che non sa spiegare. Né vale addossare la responsabilità alle scuole inferiori.

Dobbiamo prendere gli allievi così come sono, richiamare ciò che essi hanno dimenticato, o studiato sotto altra nomenclatura.

Se l'insegnante tormenta i suoi alunni, e invece di cattivarsi il loro amore, eccita odio contro sè e la scienza che insegna, non solo il suo insegnamento sarà negativo, ma il dover convivere con tanti piccoli nemici sarà per lui un continuo tormento.

Giuseppe Peano $(1858 - 1932)$

Indice

Premessa

Questi appunti contengono lo schema di un corso di Matematica per il corso di laurea in Economia. Essi possono essere considerati la continuazione degli appunti per un precorso di Matematica, reperibili su www.batmath.it.

Gli studenti sono pregati di segnalare eventuali, inevitabili, errori all'indirizzo di posta elettronica batmath@gmail.com.

1 Insiemi e funzioni

Questo capitolo contiene sostanzialmente la riproposizione di alcuni concetti già introdotti negli appunti per il precorso, con alcune estensioni.

1.1 Insiemi limitati e illimitati di numeri reali

Attenzione: in tutto questo paragrafo gli insiemi considerati sono sottoinsiemi dell'insieme dei numeri reali.

Definizione 1.1. *Sia $A \subseteq \mathbb{R}$ un insieme. Un numero reale x si dice un* maggiorante *di A se*

$$x \geq a, \ \forall a \in A. \tag{1.1}$$

Un numero reale y si dice un minorante *di A se*

$$y \leq a, \ \forall a \in A. \tag{1.2}$$

Esempio 1.1. Sia $A =]-2, 8]$. Allora -5, $-\pi$, -2 sono minoranti; 8, 10, $\sqrt{89}$ sono maggioranti.

Esempio 1.2. Sia $A = \mathbb{N}$. Allora non esistono maggioranti, mentre tutti i numeri reali minori o ugauli a zero sono minoranti.

Esempio 1.3. Sia $A = \mathbb{Z}$. Allora non esistono né maggioranti, né minoranti.

Definizione 1.2. *Sia $A \subseteq \mathbb{R}$ un insieme. A dice* limitato superiormente *(o anche* limitato a destra*) se ha almeno un maggiorante. A dice* limitato inferiormente *(o anche* limitato a sinistra*) se ha almeno un minorante. Un insieme che sia limitato sia superiormente che inferiormente si dice* limitato. *Un insieme che non sia limitato superiormente o inferiormente si dice* illimitato superiormente *o* illimitato inferiormente. *Un insieme illimitato sia superiormente che inferiormente si dice* illimitato.

Esempio 1.4. $\mathbb{N}$ è limitato inferiormente, ma non superiormente.

Esempio 1.5. $\mathbb{Z}$ non è limitato né inferiormente né superiormente.

Esempio 1.6. $A =]2, 6[$ è limitato (sia superiormente che inferiormente).

Definizione 1.3. *Sia $A \subseteq \mathbb{R}$ un insieme. Un numero $M \in A$ si dice il* massimo *di A se $M \geq a$, $\forall a \in A$; un numero $m \in A$ si dice il* minimo *di A se $m \leq a$, $\forall a \in A$.*

Si noti che abbiamo detto "il" massimo e "il" minimo perché, come si può dimostrare se il massimo c'è, è unico; analogamente per il minimo. Si tenga ben presente che il massimo e il minimo, se ci sono, appartengono all'insieme.

Esempio 1.7. Se $A = [0,1]$, 1 è il massimo, mentre 0 è il minimo di A. Si scrive anche $1 = \max(A)$, $0 = \min(A)$.

Esempio 1.8. Se $A =]0,1[$, A non ha né massimo né minimo.

Esempio 1.9. Se $A =]0,1]$, $1 = \max(A)$, mentre A non ha minimo.

Esempio 1.10. $\min(\mathbb{N}) = 0$, mentre $\mathbb{N}$ non ha massimo.

Esempio 1.11. $\mathbb{Z}$ non ha né massimo né minimo.

Come è ovvio dalla definizione, e come mostrano gli esempi, un insieme illimitato superiormente non può avere massimo, uno illimitato inferiormente non può avere minimo, ma anche insiemi limitati possono non avere né massimo né minimo. Per ovviare a questo inconveniente si introducono i concetti di estremo superiore e inferiore che, in un certo senso, estendono quelli di massimo e minimo. L'introduzione di questi concetti è resa possibile dal fatto che si dimostra che se un insieme è limitato superiormente allora l'insieme dei suoi maggioranti ha sempre un minimo, se un insieme è limitato inferiormente allora l'insieme dei suoi minoranti ha sempre un massimo.

Definizione 1.4. *Sia $A \subseteq \mathbb{R}$ un insieme superiormente limitato. Allora il minimo dei maggioranti di A si chiama* estremo superiore *di A e si indica con* $\sup(A)$*. Sia $A \subseteq \mathbb{R}$ un insieme inferiormente limitato. Allora il massimo dei minoranti di A si chiama* estremo inferiore *di A e si indica con* $\inf(A)$*.*

Se A è superiormente illimitato si pone, per definizione, $\sup(A) = +\infty$*; se A è inferiormente illimitato si pone, per definizione,* $\inf(A) = -\infty$*.*

Possiamo quindi dire che ogni insieme di numeri reali ha sempre un estremo superiore (eventualmente $+\infty$) e un estremo inferiore (eventualmente $-\infty$). Per distinguere gli insiemi limitati dagli illimitati, per i primi parleremo di estremo superiore, oppure inferiore, *finiti*.

Esempio 1.12. $+\infty = \sup(\mathbb{N})$, $0 = \inf(\mathbb{N}) = \min(\mathbb{N})$.

Esempio 1.13. $+\infty = \sup(\mathbb{Z})$, $-\infty = \inf(\mathbb{Z})$.

Esempio 1.14. $1 = \sup(]0,1[)$, $0 = \inf(]0,1[)$.

Esempio 1.15. $1 = \sup([0,1]) = \max([0,1])$, $0 = \inf([0,1]) = \min([0,1])$.

Come mostrano gli esempi, e come è facile mostrare, se un insieme ha massimo, il massimo è anche estremo superiore, se un insieme ha minimo, il minimo è anche estremo inferiore.

Gli estremi superiore e inferiore, se finiti, godono delle due proprietà caratteristiche[1] elencate nel teorema che segue (invitando i più volenterosi a dimostrarlo).

Teorema 1.5. *Se $A \subseteq \mathbb{R}$ è un insieme superiormente limitato, allora* $\sup(A)$ *è caratterizzato dalle seguenti due proprietà.*

1. $\sup(A) \geq a$, $\forall a \in A$;
2. $\forall \varepsilon > 0$, $\exists a \in A$, $a > \sup(A) - \varepsilon$.

Se $A \subseteq \mathbb{R}$ è un insieme inferiormente limitato, allora $\inf(A)$ *è caratterizzato dalle seguenti due proprietà.*

1. $\inf(A) \leq a$, $\forall a \in A$;
2. $\forall \varepsilon > 0$, $\exists a \in A$, $a < \inf(A) + \varepsilon$.

[1]Si dice proprietà caratteristiche perché sono condizioni necessarie e sufficienti perché un numero sia il sup di insieme.

1.2 Insiemi limitati e illimitati nel piano

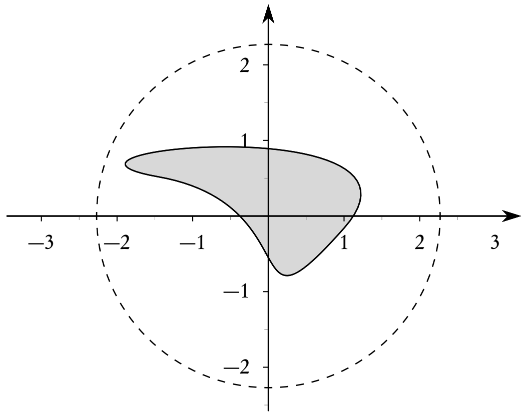
Anche per i sottoinsiemi del piano si può introdurre il concetto di insieme limitato e illimitato, ma la cosa è diversa dal caso degli insiemi sulla retta, perché sulla retta reale esiste un ordine (cioè nei numeri reali si può parlare di maggiore e di minore), mentre nel piano non esiste alcun ordine.

Definizione 1.6. *Sia $A \subseteq \mathbb{R}^2$ un sottoinsieme del piano. A si dice* limitato *se esiste un cerchio di centro l'origine e raggio r che lo contiene, altrimenti si dice* illimitato.

Come si vede si parla solo di insieme limitato o illimitato, non ha alcun senso il concetto di limitatezza superiore o inferiore, così come non hanno senso i concetti di maggiorante, minorante, massimo, minimo, estremo superiore o inferiore.

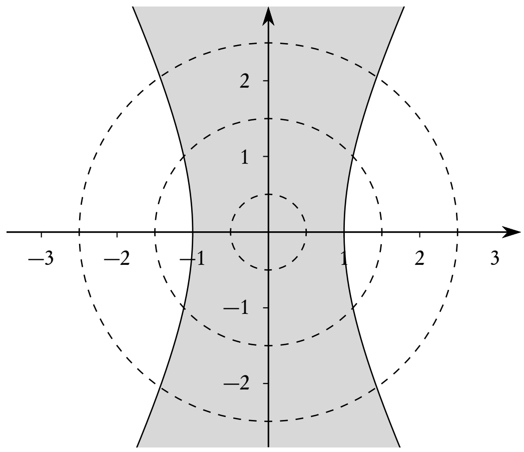
Esempio 1.16.

Un insieme limitato del piano e il cerchio che lo contiene



Esempio 1.17.

Un insieme illimitato del piano: nessun cerchio lo può contenere.



1.3 Un po' di topologia

Nel seguito avremo bisogno di utilizzare concetti simili per sottoinsiemi della retta e del piano. Per uniformare le definizioni conviene dare la seguente definizione.

Definizione 1.7. *Dato un punto P sulla retta (numero reale) o sul piano (coppia di numeri reali), diremo* palla *di centro P e raggio ε l'insieme di tutti i punti che hanno da P distanza minore di ε; diremo invece* palla chiusa *di centro P e raggio ε l'insieme di tutti i punti cha hanno da P distanza minore o uguale a ε.*

È evidente che sulla retta una palla è un intervallo (aperto o chiuso a seconda dei casi) che ha un numero c come centro e ε come semiampiezza; nel piano una palla è un cerchio di centro P e raggio ε, comprensivo o no della circonferenza di bordo, a seconda dei casi.

Definizione 1.8 (Intorno)**.** *Dato un punto P sulla retta o nel piano, si chiama* intorno *di P, e si indica con $I(P)$, o con I_P, una qualunque palla aperta a cui P appartiene.*

Ci interesseranno in maniera particolare gli *intorni circolari*, che sono gli intorni costituiti dalle palle aperte che hanno P al centro: per questi intorni useremo, di solito, il simbolo $I(P, \varepsilon)$. Il numero ε si chiama raggio dell'intorno.

Definizione 1.9 (Punto interno)**.** *Dato un insieme A, un punto P si dice* interno *ad A se esiste almeno un intorno di P tutto contenuto in A. È ovvio che un punto interno appartiene sempre all'insieme.*

Definizione 1.10 (Punto esterno)**.** *Dato un insieme A, un punto P si dice* esterno *ad A se esso è interno al complementare di A, cioè se esiste almeno un intorno di P tutto contenuto nel complementare di A. È ovvio che un punto esterno non può appartenere all'insieme.*

Definizione 1.11 (Punto isolato)**.** *Dato un insieme A, un punto P di A si dice* isolato *in A se esiste un intorno $I(P)$ di P tale che $I(P) \cap A = \{P\}$, cioè se esiste un intorno di P nel quale P è l'unico punto di A. È ovvio che un punto isolato appartiene sempre all'insieme.*

Definizione 1.12 (Punto di frontiera)**.** *Dato un insieme A, un punto P si dice* di frontiera *per A se per ogni intorno $I(P)$ di P si ha $I(P) \cap A \neq \emptyset$ e contemporaneamente $I(P) \cap \complement A \neq \emptyset$, cioè se in ogni intorno di P cade almeno un punto di A e un punto fuori da A. Un punto di frontiera può appartenere oppure no all'insieme.*

Definizione 1.13 (Punto di accumulazione)**.** *Dato un insieme A, un punto P si dice* di accumulazione *per A se in ogni intorno $I(P)$ di P cadono infiniti punti di A, cioè se l'insieme $I(P) \cap A$ contiene infiniti punti. Un punto di accumulazione può appartenere oppure no all'insieme.*

Di seguito alcuni esempi, prima con sottoinsiemi della retta, poi con sottoinsiemi del piano.

Esempio 1.18. In questo esempio sulla retta l'insieme A è così definito: $A = [0, 2[\cup \{5\}$.

- 1 è un punto interno, perché l'intorno $I(1) =]1/2, 3/2[$ è tutto contenuto in A. L'insieme di tutti i punti interni è $]0, 2[$.
- 7 è un punto esterno, perché l'intorno $I(5) =]6, 8[$ è tutto contenuto nel complementare di A. L'insieme di tutti i punti esterni è $]-\infty, 0[\cup]2, 5[\cup]5, +\infty[$.
- 5 è un punto isolato, anzi è l'unico punto isolato, perché l'intorno $I(5) =]4, 6[$, se intersecato con A, dà solo il punto 5 stesso.
- 0 è un punto di frontiera perché qualunque intorno di 0 contiene punti alla sua sinistra (che non stanno in A) e punti alla sua destra (e quelli immediatamente a destra di 0 stanno in A). Anche 2 è un punto di frontiera, per motivi simili. Si noti che 0 sta in A, mentre 2 non sta in A. Anche 5 è un punto di frontiera perché in ogni intorno di 5 cade un punto di A (5 stesso!) e punti del complementare di A (quelli immediatamente a sinistra e a destra di 5). 0, 2, 5 sono gli unici punti di frontiera.

– 1 è un punto di accumulazione, perché l'intorno $I(1) =]1/2, 3/2[$ contiene infiniti punti di A (anzi è costituito solo da punti di A). Anche 2 è punto di accumulazione, perché qualunque suo intorno contiene infiniti punti di A (quelli immediatamente a sinistra di 2 stesso). L'insieme di tutti i punti di accumulazione è $[0, 2]$.

Si noti che essere *interno* non è la stesso cosa di *appartenere*, essere *esterno* non è la stessa cosa di *non appartenere*. Valgono poi alcune proprietà che si possono desumere dagli esempi e che i più volenterosi sono invitati a provare.

– Un punto interno è sempre di accumulazione;
– un punto interno non può essere né isolato né di frontiera;
– un punto isolato è sempre di frontiera;
– un punto isolato non può essere di accumulazione, anzi, in un certo senso punto isolato è il contrario di punto di accumulazione.

Esempio 1.19. In questo esempio sul piano, l'insieme A è costituito dall'unione del cerchio di centro l'origine e raggio 1, comprensivo della semicirconferenza di bordo contenuta nel semipiano $y \geq 0$, del punto $P = (1, 1)$ e dei punti della retta r di equazione $x = 2$. Lasciamo al lettore, come utile esercizio, il compito di provare quanto affermato.

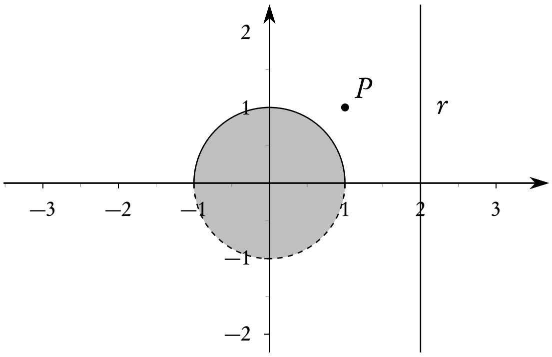


Figura 1.1 *Un insieme del piano*

– L'insieme dei punti interni è costituito dall'interno del cerchio di centro l'origine e raggio 1 (esclusa dunque la circonferenza di bordo).
– L'insieme dei punti esterni è costituito dai punti che stanno fuori dal cerchio chiuso di centro l'origine e raggio 1, con l'esclusione del punto P e dei punti della retta r.
– P è l'unico punto isolato.
– L'insieme dei punti di frontiera è costituito dai punti della circonferenza (non cerchio!) di centro l'origine e raggio 1, dal punto P e dai punti della retta r.
– L'insieme dei punti di accumulazione è costituito dal cerchio di centro l'origine e raggio 1 e dai punti della retta r.
– L'insieme A e un insieme illimitato del piano.

Definizione 1.14 (Insieme chiuso)**.** *Un insieme A si dice* chiuso *se contiene tutti i suoi punti di accumulazione.*

Definizione 1.15 (Insieme aperto)**.** *Un insieme A si dice* aperto *se il suo complementare è chiuso.*

Esempio 1.20. Le palle aperte sono insiemi aperti, le palle chiuse sono insiemi chiusi. In particolare gli intervalli aperti sono insiemi aperti, gli intervalli chiusi sono insiemi chiusi.

Esempio 1.21. L'insieme vuoto (sia come sottoinsieme di $\mathbb{R}$ che di $\mathbb{R}^2$) è sia aperto che chiuso. Analogamente tutto $\mathbb{R}$ (sulla retta) o tutto $\mathbb{R}^2$ (sul piano) sono aperti e chiusi. Questi sono gli unici insiemi contemporaneamente aperti e chiusi.

Esempio 1.22. Un intervallo del tipo $[a,b[$, oppure $]a,b]$ non è né aperto né chiuso.

Esempio 1.23. L'insieme del piano tracciato nella figura 1.1 non è né aperto né chiuso. Se a questo insieme aggiungo la semicirconferenza inferiore, diventa un insieme chiuso.

Esempio 1.24. L'insieme $\mathbb{N} \subset \mathbb{R}$ è chiuso. Analogamente l'insieme $\mathbb{Z} \subset \mathbb{R}$.

Seguono alcune proprietà la cui dimostrazione, come al solito, è lasciata per esercizio ai più volenterosi.

- Un insieme è aperto se e solo se tutti i suoi punti sono interni.
- Un insieme è chiuso se e solo contiene tutti i suoi punti di frontiera.
- Un insieme che abbia punti isolati non può essere aperto.
- Un insieme che abbia *solo* punti isolati è chiuso.
- Se A e B sono chiusi, anche $A \cup B$ e $A \cap B$ sono chiusi.
- Se A e B sono aperti, anche $A \cup B$ e $A \cap B$ sono aperti. Se però si passa ad unioni o intersezioni di infiniti insiemi ci possono essere delle sorprese. Senza entrare troppo nei dettagli, consideriamo per esempio gli insiemi
  $$]-1,1[,\ \left]-\frac{1}{2},\frac{1}{2}\right[,\ \left]-\frac{1}{3},\frac{1}{3}\right[,\ \left]-\frac{1}{4},\frac{1}{4}\right[,\ \dots,$$
  che sono tutti aperti. Facendo la loro intersezione resta solo il punto 0, che è un insieme chiuso, anzi un insieme costituito solo da un punto isolato.

1.4 Insiemi connessi. Insiemi convessi

Definizione 1.16 (Insieme connesso)**.** *Un insieme A (della retta o del piano) si dice* connesso *quando presi comunque due suoi punti P e Q esiste un arco di linea continua che li connette e tutto contenuto in A.*[2]

In $\mathbb{R}$ sono connessi tutti e soli gli intervalli, di qualunque tipo. In $\mathbb{R}^2$ le palle (aperte o chiuse) sono sempre connesse, ma ci sono anche insiemi connessi più complessi, come per esempio l'insieme costituito dai punti del primo e terzo quadrante, inclusi gli assi cartesiani.

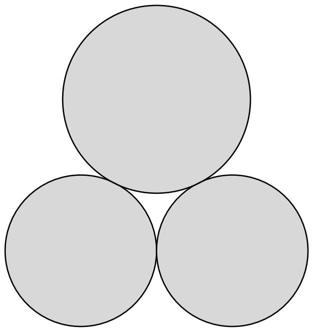
Definizione 1.17 (Insieme convesso)**.** *Un insieme A (della retta o del piano) si dice* convesso *quando presi comunque due suoi punti P e Q esiste un segmento che li connette e tutto contenuto in A.*

È evidente che un insieme convesso è sempre connesso, ma, almeno nel piano, il viceversa non è vero: ci sono insiemi connessi ma non convessi, come vedremo sugli esempi. In $\mathbb{R}$, invece, i due concetti coincidono: gli unici insiemi connessi o convessi sono gli intervalli, e la cosa è quasi ovvia.

[2] In realtà la definizione che qui abbiamo dato è quella di *connessione per archi*, mentre la definizione di connessione sarebbe più complessa. Per gli scopi del nostro corso, comunque, questa definizione "semplificata" è più che sufficiente.

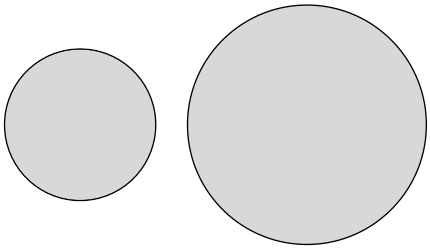
Esempio 1.25.

Un insieme connesso ma non convesso (le tre circonferenze bordo sono comprese nell'insieme).



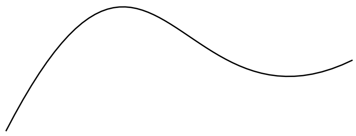
Esempio 1.26.

Un insieme non connesso (e quindi nemmeno convesso).



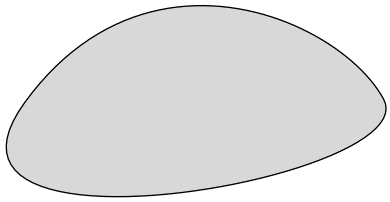
Esempio 1.27.

Un insieme connesso ma non convesso (una curva continua).



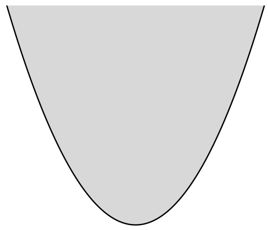
Esempio 1.28.

Un insieme connesso e convesso.



Esempio 1.29.

Un insieme connesso e convesso (si intende che l'insieme prosegue fino all'infinito, comprendendo tutta la parte interna alla parabola rappresentata).



1.5 Funzioni

Il concetto di funzione è di capitale importanza in matematica e in tutte le sue applicazioni e ne daremo qui una definizione formale.

Definizione 1.18. *Dati due insiemi A e B, si chiama* funzione *di A in B un sottoinsieme F del prodotto cartesiano $A \times B$ che goda della seguente proprietà: per ogni $x \in A$ esiste un unico $y \in B$ tale che $(x, y) \in F$.*

L'insieme A è detto dominio *della funzione, l'insieme B è detto* codominio.

Nella sostanza, per assegnare una funzione occorre assegnare i due insiemi A e B e precisare una regola che indichi, per ogni $x \in A$, qual è l'unico $y \in B$ tale che $(x, y) \in F$. Si potrebbe anche dire, in maniera più discorsiva, che una funzione è una legge che faccia corrispondere ad ogni x di un insieme A un unico y di un insieme B. Normalmente questa legge si indica con una lettera minuscola dell'alfabeto latino, per esempio f: se x è un elemento dell'insieme A e y è l'*unico* elemento di B tale che $(x, y) \in F$, si dice che y è *funzione* di x e si scrive $y = f(x)$ (leggi: "y uguale a effe di x"). La x di A prende anche il nome di *variabile indipendente*, la corrispondente y di B di *variabile dipendente*.

È molto importante ricordare che per assegnare una funzione occorre assegnare

- il dominio
- il codominio
- una legge o regola che indichi, per ogni x del dominio, quale sia l'unico y corrispondente del codominio.

Una (apparente) eccezione a questo fatto si ha nel caso delle funzioni elementari che hanno come dominio un sottoinsieme di $\mathbb{R}$ o $\mathbb{R}^2$ e come codominio $\mathbb{R}$. Si veda a questo proposito il successivo paragrafo 1.10 nella pagina 18.

La notazione più completa per le funzioni è la seguente:

$$f : A \to B,\ x \mapsto f(x),$$

ma spesso si scrive solo

$$x \mapsto f(x),$$

se gli insiemi A e B sono già stati precisati o sono chiari dal contesto. Si può anche dire semplicemente *la funzione* $y = f(x)$, anche se i puristi potrebbero storcere il naso.

Esempio 1.30. Se A e B sono l'insieme dei numeri reali, si può considerare la funzione che a ogni numero reale x fa corrispondere il suo quadrato. In questo caso si dovrebbe scrivere

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$$

ma si può scrivere anche semplicemente

$$x \mapsto x^2$$

oppure (e noi lo faremo sistematicamente)

$$y = x^2.$$

Per visualizzare le funzioni si usano spesso dei diagrammi a frecce, come quello che segue.

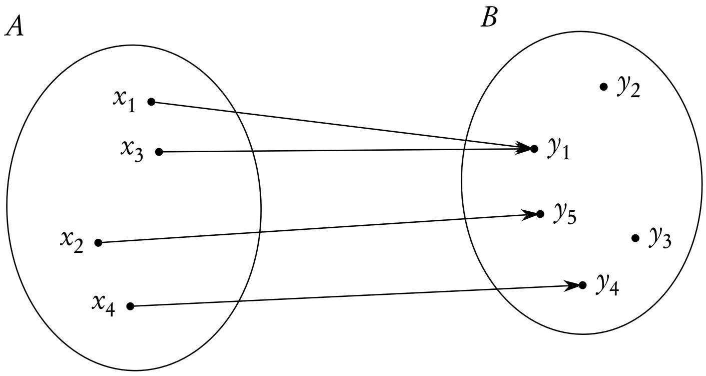


Figura 1.2 *Diagramma “a frecce” per visualizzare una funzione (tra insiemi finiti)*

Si noti che *è obbligatorio* che da *ogni* punto (elemento dell'insieme) A parta *esattamente* una freccia, mentre sui punti dell'insieme B possono anche arrivare *più* frecce, oppure *nessuna* freccia. Si potrebbe dire, usando un linguaggio figurato, che A è l'insieme degli arcieri, B l'insieme dei bersagli e che ogni arciere ha a disposizione nella propria faretra solo una freccia che è costretto a lanciare, mentre non ci sono limitazioni sui bersagli da colpire: ci possono essere bersagli colpiti da più frecce, e anche bersagli non colpiti da alcuna freccia.

Ha particolare interesse nelle applicazioni la determinazione del sottoinsieme del codominio costituito da tutti i punti dove arriva almeno una freccia, cioè, formalmente, l'insieme

$$I \subseteq B = \{\, y \in B \mid \exists x \in A, y = f(x) \,\}, \tag{1.3}$$

o anche, a parole, l'insieme degli y di B tali che esiste almeno un x di A, la cui immagine sia y. L'insieme I si chiama *insieme immagine*[(3)]. L'insieme immagine si indica anche con $f(A)$, proprio a significare il fatto che si tratta dell'insieme delle immagini di tutte le x di A. Se C è un sottoinsieme di A, si può considerare l'insieme delle immagini di tutte le x di C (che sarà naturalmente un sottoinsieme dell'insieme immagine). Questo insieme si indica con $f(C)$.

È chiaro che rappresentazioni grafiche come quella appena vista hanno senso solo se gli insiemi in questione sono finiti: in caso contrario si dovrebbero disegnare infinite frecce, cosa chiaramente impossibile.

Si usano anche altri tipi di rappresentazione per le funzioni. Per esempio se si considera la funzione che a ogni numero naturale compreso tra 1 e 5 fa corrispondere la sua metà (funzione che ha come dominio i numeri naturali citati e come codominio i numeri razionali), si può usare una tabella a doppia entrata, in cui nella prima colonna si scrivono i numeri naturali 1, 2, …, 5 e nella seconda colonna le *corrispondenti* metà di questi numeri.

[3]Purtroppo la nomenclatura a questo proposito non è uniforme e molti autori chiamano codominio l'insieme immagine, spesso senza assegnare alcun nome all'insieme B.

x	$x/2$
1	$1/2$
2	1
3	$3/2$
4	2
5	$5/2$

Tabella 1.1 *Rappresentazione "tabulare" di una funzione*

Un altro tipo di rappresentazione è quello dei diagrammi a torta, molto significativo in casi speciali. Consideriamo, ad esempio, un corso universitario dove si sono iscritti 120 alunni, provenienti da varie provincie, come nella tabella che segue:

Gorizia	Pordenone	Treviso	Trieste	Udine
5	70	15	10	20

Si comincerà con il calcolare le percentuali relative alle varie provincie:

Gorizia	Pordenone	Treviso	Trieste	Udine
4.17	58.33	12.5	8.33	16.67

Successivamente si calcoleranno le ampiezze delle "fette di torta" da utilizzare per ciascuna provincia, tenendo conto che la torta totale ha un'apertura di 360°:

Gorizia	Pordenone	Treviso	Trieste	Udine
15°	210°	45°	30°	60°

Il grafico è a questo punto immediato:

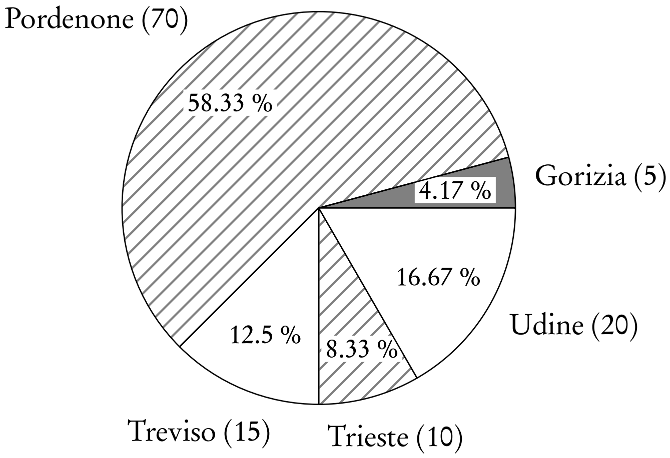


Figura 1.3 *Provenienza degli studenti del Corso ..., ripartiti per Provincia, diagramma "a torta"*

Ancora un'altra possibilità è quella di un diagramma a barre, che proponiamo qui di seguito, senza commenti.

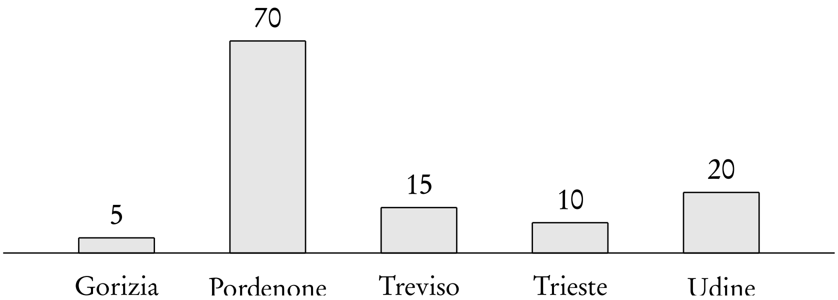


Figura 1.4 *Provenienza degli studenti del Corso ..., ripartiti per Provincia, diagramma "a barre"*

La rappresentazione più conveniente[(4)] nel caso delle funzioni tra due insiemi di numeri reali è però quella dei diagrammi o grafici cartesiani, in particolare nel caso in cui gli insiemi siano infiniti quando le rappresentazioni precedenti non sono utilizzabili. L'idea è di considerare un piano in cui si sia fissato un sistema di coordinate cartesiane (ortogonali per semplicità) Oxy e rappresentarvi tutte le coppie (x,y) in cui x è un punto (numero) del dominio della funzione e $y = f(x)$ è il *corrispondente* valore nel codominio della funzione. Riprendendo in esame l'esempio proposto nella tabella 1.1, dobbiamo rappresentare i punti

$$A=(1,1/2),\ B=(2,1),\ C=(3,3/2),\ D=(4,2),\ E=(5,5/2),$$

ottenendo il grafico che segue.

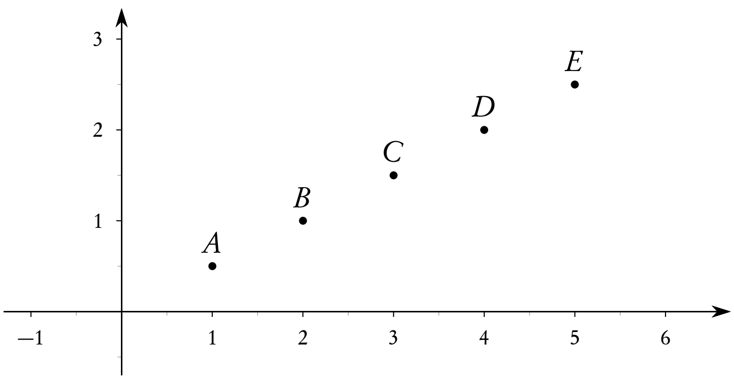


Figura 1.5 *Esempio di grafico cartesiano*

Il grafico della precedente figura 1.5 è in realtà un grafico a frecce "compattato": siccome i valori del dominio sono punti dell'asse x e quelli del codominio punti dell'asse y, possiamo sempre pensare di tracciare delle frecce che colleghino i punti del dominio con i corrispondenti del codominio, come quelle della figura 1.2, solo che è opportuno che le frecce "passino" per i punti A, B, ...:

[4]Vista la definizione di funzione come sottoinsieme del prodotto cartesiano di due insiemi, si può anzi ritenere questa la rappresentazione "standard" per le funzioni.

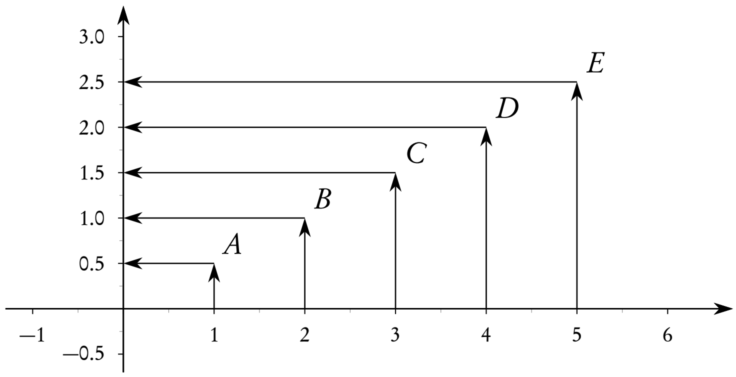


Figura 1.6 *Esempio di grafico cartesiano, con frecce*

Il grafico 1.5 "compatta" il grafico 1.6 nel senso che ne prende solo gli elementi essenziali, cioè gli "spigoli delle frecce": è evidente che dalla conoscenza degli spigoli si possono facilmente ricostruire le frecce.

Se si confronta la figura 1.5 con la tabella 1.1, ci si rende immediatamente conto dei notevoli vantaggi che il grafico presenta: da esso si può per esempio capire, "a colpo d'occhio", che al crescere di x nel dominio la corrispondente y del codominio cresce, e che tale crescita è *costante*. La cosa diventa ancora più significativa se si vuole considerare la funzione che a ogni numero reale x faccia corrispondere la sua metà: a differenza di quanto succedeva con la funzione rappresentata nella tabella 1.1, questa volta la x non varia più in un insieme finito e quindi una rappresentazione tabulare non ha alcun senso[5]. Un diagramma cartesiano è decisamente più significativo:

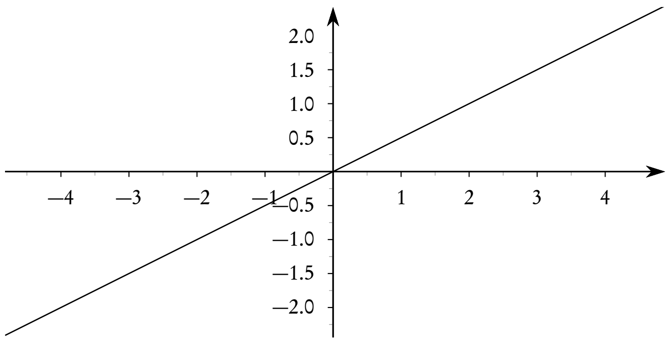


Figura 1.7 *Grafico della funzione* $y = {}^{x}/_{2}$

Naturalmente il diagramma 1.7 contiene anche i punti già rappresentati nel diagramma 1.5, ma contiene anche infiniti altri punti. Anche se non è chiaramente possibile rappresentare nel grafico *tutte* le coppie $(x,y) = (x, f(x))$ che visualizzano l'andamento della funzione, tuttavia la parte tracciata è sufficiente a rendere evidenti quasi tutte le proprietà che interessano.

[5] Si noti comunque che la regola (legge) che collega la x alla y è la stessa del caso precedente: per assegnare una funzione *non* è sufficiente assegnare la regola di calcolo, occorre anche fissare il dominio e il codominio.

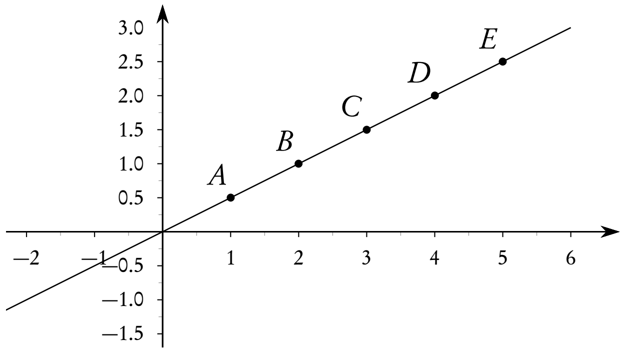


Figura 1.8 *Grafico della funzione $y = x/2$, con evidenziati alcuni punti*

Una buona parte del corso di Matematica sarà dedicata proprio allo studio di strategie adatte a evidenziare le caratteristiche essenziali di una funzione (avente come dominio e codominio sottoinsiemi dell'insieme dei numeri reali) e a tracciarne un grafico indicativo. Un grande aiuto in questo senso può essere fornito dai numerosi software dedicati allo scopo[6], ma, come al solito, bisogna tenere conto che il computer *è una macchina finita* e quindi non può risolvere tutti i problemi. A questo proposito proponiamo un esempio "estremo", precisamente il grafico della funzione

$$f(x) = \sin 1/x\,.$$

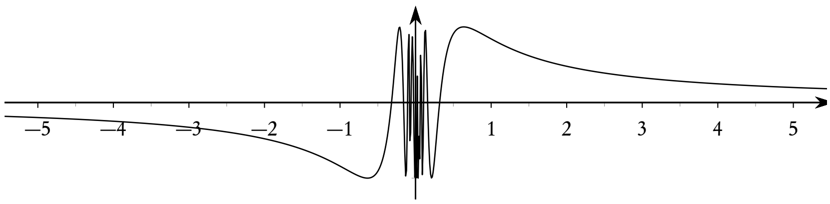


Figura 1.9 *Grafico di $f(x) = \sin 1/x$*

È chiaro che, per valori di x prossimi allo zero, questo grafico è poco significativo. Purtroppo nemmeno zoomate (in orizzontale) migliorano granché la situazione, come mostrano le due successive figure.

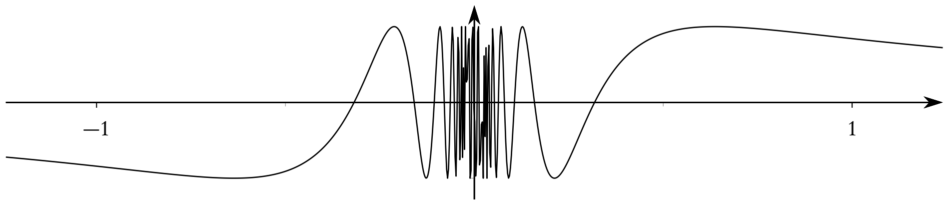


Figura 1.10 *Grafico di $f(x) = \sin 1/x$, con uno zoom sull'asse delle x*

[6]Tra i software commerciali segnaliamo *Mathematica* e *Maple*, due pacchetti estremamente sofisticati e complessi. Tra i software non commerciali segnaliamo *Maxima* (molto simile a *Mathematica*™ anche se non ne possiede tutte le potenzialità) e *Geogebra*. Riteniamo quest'ultimo particolarmente adatto per questo corso e segnaliamo che la maggior parte dei grafici contenuti in questo testo sono ottenuti proprio con *Geogebra*.

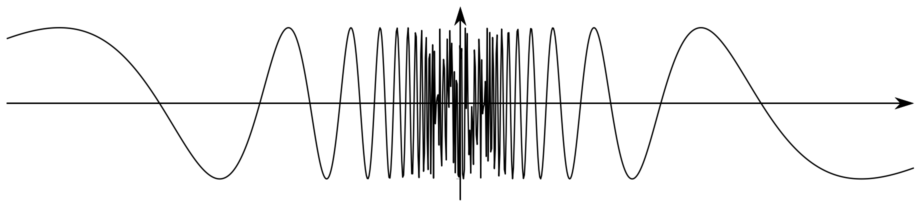


Figura 1.11 *Grafico di $f(x) = \sin 1/x$, con un ulteriore zoom sull'asse delle x*

Naturalmente non sempre le cose vanno così male (per fortuna!). Per la funzione $f(x) = x^3 - 3x^2$, per esempio, il grafico fornito da un software di calcolo è sufficientemente accurato da contenere con buona accuratezza le informazioni necessarie.

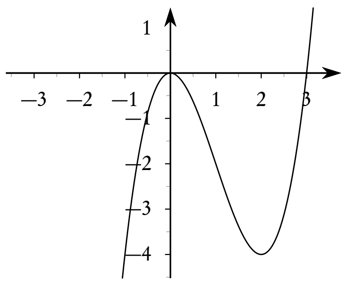


Figura 1.12 *Grafico di $f(x) = x^3 - 3x^2$*

Da questo grafico si vede subito che, al crescere della x da valori negativi fino allo 0, anche la corrispondente y cresce (e abbastanza rapidamente) fino a raggiungere il valore 0; successivamente se la x cresce da 0 a 2, la y decresce fino a raggiungere il valore −4, per poi aumentare di nuovo (e di nuovo abbastanza rapidamente) al crescere di x.

In tutti i grafici cartesiani che abbiamo fatto, tranne quelli delle figure 1.10 e 1.11, abbiamo usato la stessa unità di misura sui due assi: sistemi cartesiani siffatti sono detti *monometrici*. Di solito però nelle applicazioni la cosa non è possibile, e ne vedremo in seguito i motivi. È opportuno tenere presente che se un sistema cartesiano nel piano non è monometrico, le figure possono essere deformate. Per esempio i due grafici della figura seguente mostrano la circonferenza di centro l'origine e raggio 1, in due diversi sistemi di coordinate, di cui solo il primo è monometrico.

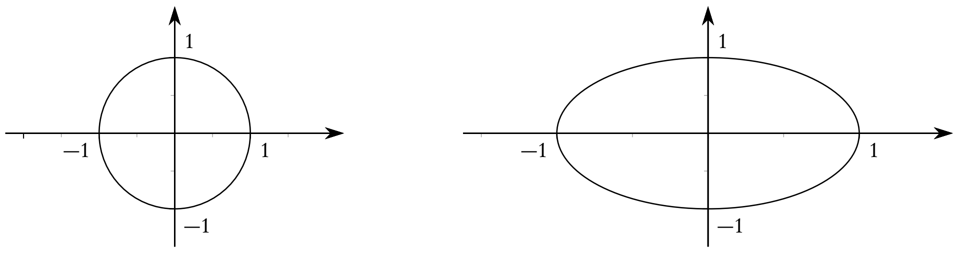


Figura 1.13 *Circonferenza di centro l'origine e raggio 1, in due diversi sistemi di coordinate, il primo monometrico, il secondo no*

1.6 Funzioni di due variabili - Introduzione

Un caso molto importante di funzioni con cui avremo a che fare nel seguito è quello delle funzioni in cui il dominio è un insieme di coppie di numeri reali (cioè un sottoinsieme di $\mathbb{R}^2$) e il codominio è l'insieme dei numeri reali: diremo brevemente *funzioni di due variabili*. Potremo usare una scrittura del tipo

$$z = f(x, y). \tag{1.4}$$

La rappresentazione grafica cartesiana di funzioni di questo tipo richiede un sistema di tre assi (che per noi saranno sempre mutuamente ortogonali): abbiamo bisogno infatti di una coppia di numeri per i punti del dominio, più un numero per i corrispondenti valori del codominio. Come vedremo, nelle situazioni che ci interesseranno, questi grafici avranno l'aspetto di superfici nello spazio. Riservandoci di approfondire a suo tempo l'argomento, proponiamo solo un grafico di esempio nella figura 1.14.

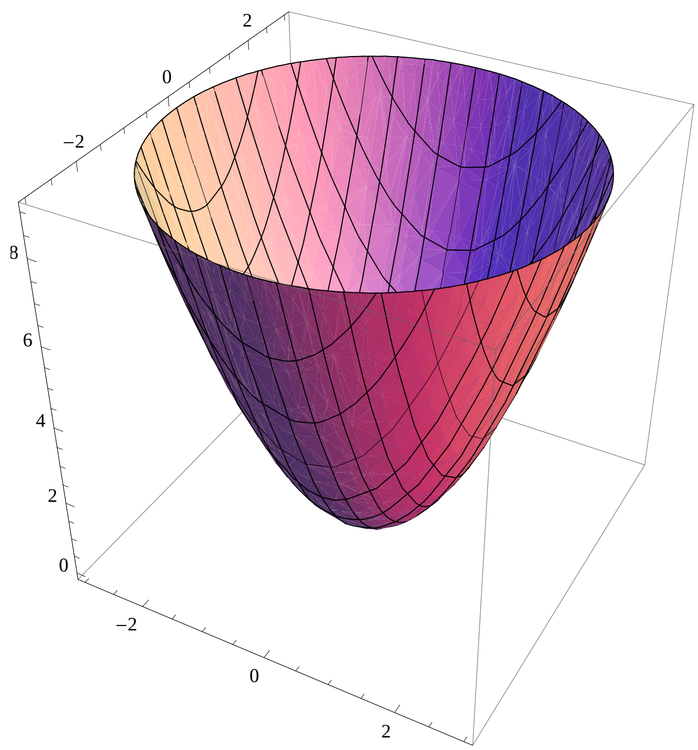


Figura 1.14 *Grafico della funzione* $z = x^2 + y^2$

1.7 Successioni

Il caso in cui il dominio di una funzione è l'insieme $\mathbb{N}$ dei naturali o un suo sottoinsieme è di particolare importanza: le funzioni di questo tipo prendono il nome di *successioni finite* se il dominio è un insieme finito o *infinite* se il dominio è un insieme infinito. Normalmente parlando di successioni ci riferiremo sempre a successioni infinite e saremo interessati solo a successioni il cui codominio è $\mathbb{R}$: parleremo di *successioni reali*.

Per le successioni la variabile indipendente si indica di solito con n, m invece che con x, e al posto di $f(x)$ si preferiscono scritture tipo a_n, b_n, ecc. (la variabile indipendente si scrive come pedice, anziché tra parentesi).

Esempio 1.31. La funzione data da

$$a_n = \frac{n}{n+1}, \quad n \in \mathbb{N},$$

è una successione reale. Anche di essa si può tracciare un grafico cartesiano, che sarà naturalmente costituito solo da "punti isolati", visto il particolare tipo di dominio.

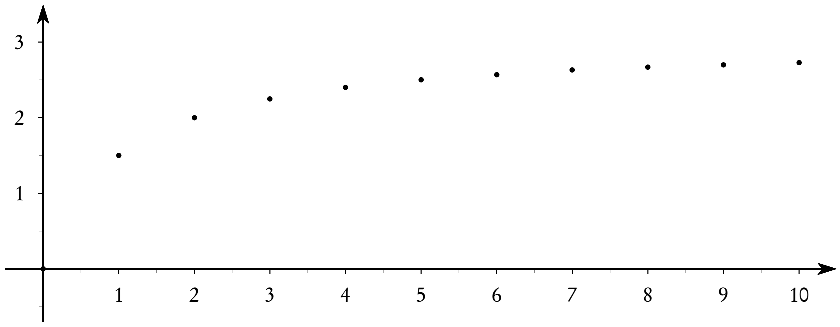


Figura 1.15 *Grafico della successione* $a_n = {}^{3n}/_{(n+1)}$

1.8 Operazioni sulle funzioni

In questo paragrafo il dominio delle funzioni è un sottoinsieme A di $\mathbb{R}$ o di $\mathbb{R}^2$, mentre il codominio è sempre $\mathbb{R}$: diremo brevemente funzioni reali di una o due variabili reali, ma spesso parleremo semplicemente di funzioni di una o due variabili, senza ulteriori precisazioni.

Date due funzioni f e g, esse si possono sempre sommare, sottrarre e moltiplicare; se la seconda è sempre diversa da zero, si possono anche dividere[7].

Esempio 1.32. Se $f(x) = |x|$ e $g(x) = x^2 + 1$, entrambe con dominio $\mathbb{R}$, la funzione somma di f e g è $|x| + x^2 + 1$, la differenza è $|x| - x^2 - 1$, il prodotto $|x|(x^2 + 1)$, il quoziente $|x|/(x^2+1)$.

Esempio 1.33. Se $f(x) = \mathrm{e}^x$ e $g(x) = x^2$, si può sempre fare la somma e il prodotto; per poter fare f/g si deve "restringere" il dominio, in modo da escludere il valore 0, che annullerebbe il denominatore.

Con opportune condizioni due funzioni f e g si possono anche *comporre*, cioè farle agire in successione: il risultato (in termini informatici diremmo l'output) della prima lo usiamo come input per la seconda, ottenendo alla fine il risultato voluto. Se per esempio la prima funzione è $f(x) = x^2$ e la seconda è $g(x) = \mathrm{e}^x$, allora la composta di f (prima funzione) e g (seconda funzione) è e^{x^2}. La funzione composta si indica con $g \circ f$ è il risultato finale della composizione si indica con $g(f(x))$; si presti

[7] Si noti che sommare, sottrarre, moltiplicare o dividere due funzioni equivale a sommare, sottrarre, moltiplicare o dividere le loro immagini, quindi è indispensabile che tali operazioni possano essere eseguite sul codominio.

particolare attenzione al fatto che la prima funzione è la più interna nella scrittura, la seconda è la più esterna.

Per poter fare la composizione si deve naturalmente richiedere che l’insieme immagine della prima sia contenuto nel dominio della seconda, visto che l’output della prima deve essere usato come input per la seconda. Se per esempio la prima funzione è $f(x)=x^2-1$ e la seconda è $g(x)=\sqrt{x}$, non si può fare a cuor leggero la composta, perché se, per esempio, $x=0$, la prima dà come risultato -1 che non può essere “dato in pasto” alla seconda funzione. In casi come questo, comunque, basterà restringere il dominio della prima funzione: nell’esempio basterà limitarsi a considerare solo $x\leq -1 \vee x\geq 1$.

Esempio 1.34. Siano date le due funzioni f e g, di $\mathbb{R}$ in $\mathbb{R}$, date dalle seguenti leggi:

$$f(x)=2x-3, \quad g(x)=5x+1.$$

Si possono allora considerare le funzioni composte

$$(f\circ g)(x)=f(g(x))=f(5x+1)=2(5x+1)-3=10x-1,$$

e

$$(g\circ f)=g(f(x))=g(2x-3)=5(2x-3)+1=10x-14.$$

Da questo esempio si deduce anche che $f\circ g$ e $g\circ f$ sono, in generale, diverse: nel fare la composta di due funzioni è importante l’ordine.

1.9 Funzioni elementari e funzioni definite “a pezzi”

Si chiamano *elementari* le funzioni (definite in un sottoinsieme di $\mathbb{R}$ o di $\mathbb{R}^2$) in cui sulla variabile, o sulle due variabili, si eseguono solo operazioni di somma, sottrazione, prodotto, quoziente, o dove sono coinvolte le funzioni potenza, radice, logaritmo, esponenziale, seno e coseno. Si tratta praticamente di tutte le funzioni con cui avremo a che fare nel nostro corso.

Il modo più semplice che abbiamo a disposizione (e di cui faremo largo uso) per costruire funzioni non elementari è quello della definizione "a pezzi” (*piecewise definition* nei software più comuni). Si tratta sostanzialmente di “unire” due funzioni (o meglio due grafici di funzioni) definite su due sottoinsiemi diversi (o ristrette a due insiemi diversi) di $\mathbb{R}$ (più raramente per noi di $\mathbb{R}^2$).

Esempio 1.35. Consideriamo le due funzioni $f(x)=\sqrt{x}$ e $g(x)=\sqrt{-x+1}$, la prima definita per $x\geq 0$, la seconda per $x\leq 1$. Di entrambe sappiamo già tracciare i grafici, riportati nelle figure 1.16 e 1.17.

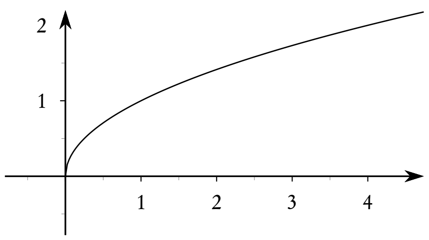


Figura 1.16 *La funzione* $\sqrt{x}$

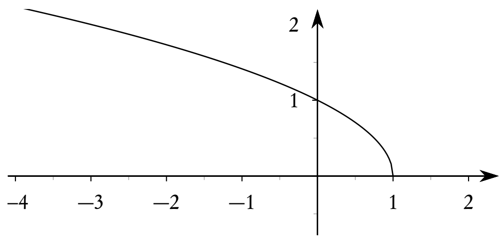
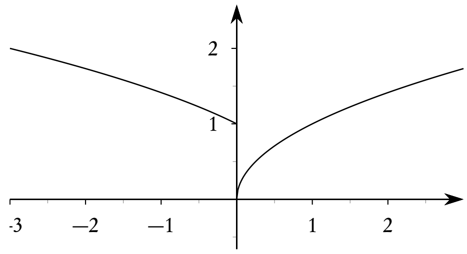


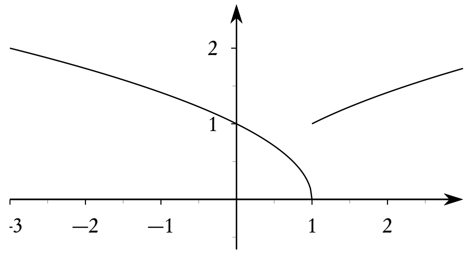
Figura 1.17 *La funzione* $\sqrt{-x-1}$

A partire da queste funzioni possiamo costruirne altre, come mostrato di seguito.

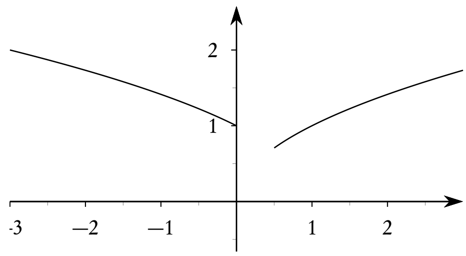
$$l(x) = \begin{cases} \sqrt{-x+1}, & \text{se} \quad x < 0 \\ \sqrt{x}, & \text{se} \quad x \geq 0 \end{cases} .$$



$$h(x) = \begin{cases} \sqrt{-x+1}, & \text{se} \quad x < 1 \\ \sqrt{x}, & \text{se} \quad x \geq 1 \end{cases} .$$



$$k(x) = \begin{cases} \sqrt{-x+1}, & \text{se} \quad x < 0 \\ \sqrt{x}, & \text{se} \quad x \geq 1 \end{cases} .$$



1.10 Dominio delle funzioni elementari

Nel dare la definizione di funzione abbiamo detto che per assegnare una funzione occorre assegnare un dominio, un codominio, e una legge che associ a ciascun punto del dominio un punto (e uno solo) del codominio. Nel caso delle funzioni elementari si sottintende sempre che il dominio sia il più grande sottoinsieme (di $\mathbb{R}$ o di $\mathbb{R}^2$) in cui le operazioni da eseguire sulla variabile o sulle variabili hanno senso. Per determinare il dominio di queste funzioni si deve dunque in generale risolvere un sistema di disequazioni che esplicitano le condizioni da imporre affinché le operazioni da eseguire siano lecite.

Esempio 1.36. Per trovare il dominio di $f(x) = \sqrt{x-1} + \ln(2-x)$, devo considerare il seguente sistema di disequazioni

$$\begin{cases} x - 1 \geq 0 \\ 2 - x > 0 \end{cases} ,$$

che traducono in formule le condizioni che il radicando della radice quadrata sia non negativo e che l'argomento del logaritmo sia strettamente positivo. Il dominio è $1 \leq x < 2$.

Esempio 1.37. Per trovare il dominio di $f(x,y)=\sqrt{x^2+y^2-1}$, devo risolvere la disequazione

$$x^2+y^2-1\geq 0$$

(equivalente a imporre la condizione che il radicando non sia negativo), che ha come soluzioni tutti i punti del piano non interni alla circonferenza di centro l'origine e raggio 1.

1.11 Funzioni crescenti e decrescenti

Attenzione: questi concetti sono validi *solo* per *funzioni di una variabile*.

Definizione 1.19 (Funzioni crescenti o decrescenti). *Una funzione $f(x)$, definita in un sottoinsieme $A\subseteq\mathbb{R}$ si dice* crescente *se, presi comunque x_1, x_2 nel dominio, l'essere $x_1<x_2$ implica che $f(x_1)\leq f(x_2)$ (se invece di $\leq$ si ha $<$, si dice* strettamente crescente*); la funzione si dice invece* decrescente *se, presi comunque x_1, x_2 nel dominio, l'essere $x_1<x_2$ implica che $f(x_1)\geq f(x_2)$ (se invece di $\geq$ si ha $>$, si dice* strettamente decrescente*).*

In pratica una funzione è crescente se al crescere di x cresce anche il corrispondente valore di $y=f(x)$, decrescente in caso contrario, cioè al crescere di x decresce il corrispondente valore di $y=f(x)$.

Le funzioni e^x, $\ln x$, $\sqrt{x}$ sono tutte crescenti (strettamente); la funzione e^{-x} è decrescente (strettamente); la funzione x^2 non è né crescente né decrescente.

Se una funzione non è né crescente né decrescente, può succedere che sia *crescente* o *decrescente a tratti*. Per esempio la funzione x^2 è decrescente per $x<0$, crescente per $x>0$; la funzione $1/x$ è decrescente sia per $x<0$ che per $x>0$ (ma, attenzione, non su tutto il suo dominio!). Le funzioni crescenti o decrescenti a tratti sono quelle che più comunemente ci capiterà di incontrare nel seguito.

1.12 Funzioni limitate e illimitate, massimi e minimi

Questi concetti si applicano sia alle funzioni di una variabile che a quelle di due variabili e sono concetti che riguardano l'insieme immagine di una funzione.

Definizione 1.20 (Funzioni limitate e illimitate). *Una funzione f di una o due variabili si dice* limitata *se il suo insieme immagine è limitato,* illimitata *se tale è il suo insieme immagine. Si usano anche gli aggettivi* superiormente *e* inferiormente, *esattamente come per i sottoinsiemi di $\mathbb{R}$.*

Le funzioni $\sin x$ e $\cos x$ sono limitate; la funzione $1/x$ è illimitata; la funzione e^x è superiormente illimitata; la funzione $-x^2$ è inferiormente illimitata.

Definizione 1.21 (Massimo e minimo per una funzione). *Il massimo e il minimo dell'insieme immagine di una funzione, se esistono, si chiamano rispettivamente* massimo assoluto *e* minimo assoluto *della funzione. L'estremo superiore e inferiore dell'insieme immagine di una funzione (che esistono sempre, eventualmente infiniti), si chiamano estremo superiore e inferiore della funzione. Spesso l'aggettivo* assoluto *si tralascia.*

La ricerca del massimo e del minimo assoluto di una funzione (se esistenti) è una delle più importanti applicazioni dell'analisi (problemi di *ottimizzazione*).

La funzione e^x non ha né max né min, il suo inf è 0, il suo sup $+ +\infty$. La funzione $\ln x$ non ha né max né min, il suo inf è $-\infty$, il suo sup $+ +\infty$. La funzione $\sin x$ ha 1 come max e -1 come min. La funzione x^2 ha 0 come min, mentre non ha max, il suo sup è $+\infty$.

Uno dei punti in corrispondenza dei quali la funzione assume il suo massimo o minimo si chiama *punto di massimo assoluto* o *punto di minimo assoluto*. La funzione $\sin x$ ha infiniti punti di massimo assoluto e di minimo assoluto,la funzione x^2 ha un solo punto di minimo assoluto (che è $x = 0$).

Definizione 1.22 (Massimi e minimi relativi)**.** *Un punto P del dominio di una funzione si chiama* punto di massimo relativo *se $\exists I(P)$ tale che, per ogni punto $Q \in I(P)$, si abbia $f(Q) \leq f(P)$; il punto si chiama* punto di minimo relativo *se, nelle stesse condizioni, si ha invece $f(Q) \geq f(P)$. Il valore $f(P)$ si chiama* massimo relativo *o* minimo relativo *rispettivamente. Mentre il massimo (assoluto) e il minimo (assoluto), se esistono, sono unici, i massimi e minimi relativi possono essere più d'uno (anche infiniti).*

Esempio 1.38. La funzione rappresentata nella figura 1.18 ha i punti x_1 e x_3 come punti di massimo relativo, x_2 e x_4 come punti di minimo relativo; non ha né massimo né minimo assoluti.

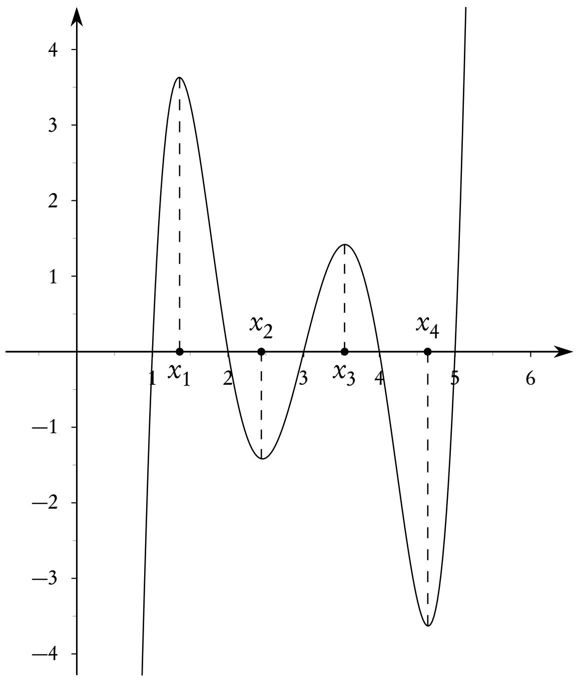


Figura 1.18 *Una funzione con due massimi e due minimi relativi*

Esempio 1.39. La funzione rappresentata nella figura 1.19 (si presuppone che il grafico mantenga indefinitamente l'andamento visualizzato) ha infiniti punti di massimo e di minimo relativo. Anche questa non ha né massimo né minimo assoluti.

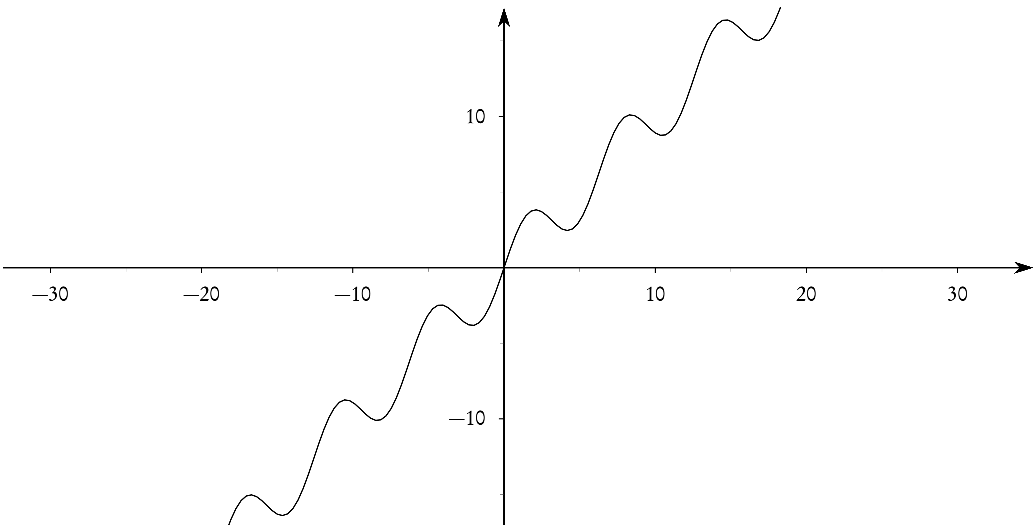


Figura 1.19 *Una funzione con infiniti massimi e minimi relativi*

Esempio 1.40. La situazione è identica nel caso di funzioni di due variabili, anche se, naturalmente, la produzione di un grafico è decisamente più complessa.

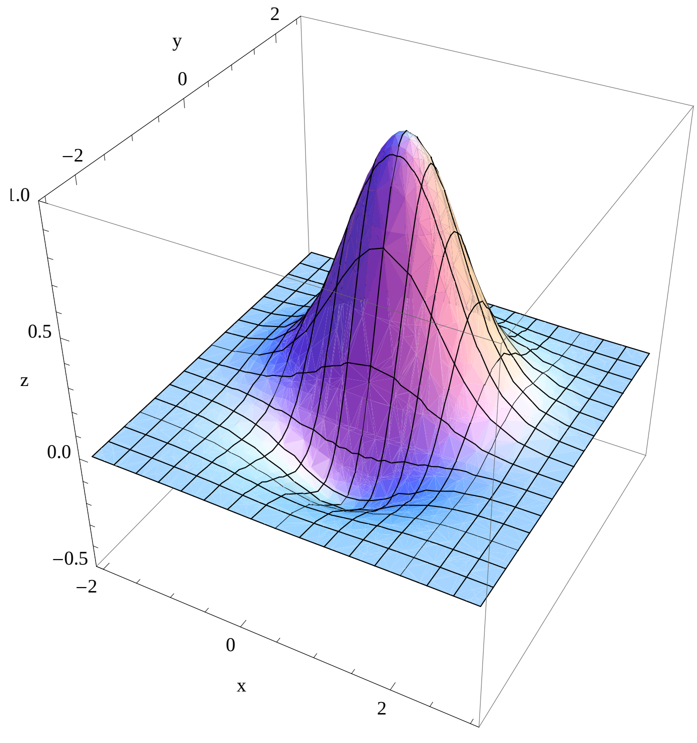


Figura 1.20 *Massimi e minimi in una funzione di due variabili*

1.13 Funzioni iniettive, suriettive, biiettive

Definizione 1.23. *Una funzione* $f: A \to B$ *si dice* iniettiva *se due punti diversi del dominio* P_1 *e* P_2 *hanno immagini diverse; una funzione si dice* suriettiva *se ogni punto del codominio è immagine di almeno un punto del dominio, ovvero se l'insieme immagine coincide con il codominio; una funzione che sia contemporaneamente iniettiva e suriettiva si dice* biiettiva *o* biunivoca.

Questi concetti hanno per noi interesse in particolare nel caso delle funzioni di una sola variabile. In questo caso è evidente che una funzione strettamente crescente o strettamente decrescente è iniettiva.

Esempio 1.41. Le funzioni e^x, $\ln x$, x^3, $\sqrt{x}$ sono tutte funzioni iniettive.

Esempio 1.42. Le funzioni, aventi come codominio $\mathbb{R}$, $\ln x$ e x^3 sono funzioni suriettive. Anche le funzioni e^x e $\sqrt{x}$ possono diventare suriettive se "restringiamo" il codominio rispettivamente agli $y > 0$ e agli $y \geq 0$.

Esempio 1.43. La funzione, di $\mathbb{R}$ in $\mathbb{R}$, x^3 è iniettiva e suriettiva, dunque biunivoca.

Esempio 1.44. La funzione x^2 non è iniettiva: i punti -1 e 1, per esempio, pur essendo diversi hanno la stessa immagine.

Esempio 1.45. La funzione, di $\mathbb{R}$ in $\mathbb{R}$, $\sin x$, non è iniettiva: esistono addirittura infiniti punti diversi del dominio che hanno la stessa immagine. La funzione non è nemmeno suriettiva, può tuttavia diventarlo se si restringe il codominio al solo intervallo $[-1,1]$.

Si noti, dagli esempi 1.42 e 1.45 che se si restringe il codominio di una funzione all'insieme immagine, ogni funzione può diventare suriettiva e questa restrizione non modifica nella sostanza l'andamento della funzione stessa: dunque è sempre possibile, senza grandi rinunce, rendere suriettiva una funzione. Non così per l'iniettività: a volte si restringe il dominio di una funzione per ottenere una funzione iniettiva, ma questa operazione modifica sostanzialmente la funzione stessa, in quanto porta a considerare solo una parte del suo grafico.

1.14 L'inversa di una funzione

Considerata una funzione $f: A \to B$ è spesso utile chiedersi se esiste una funzione $g: B \to A$ tale che se f assume in x il valore $f(x)$, allora g assume nel punto $f(x)$ il valore x, ovvero una funzione g che, agendo sul codominio di f, ne *annulli l'effetto*. In termini di frecce si può dire che la funzione g, se esiste, ha il compito di rispedire al mittente le frecce che hanno colpito i bersagli. È chiaro che ciò è possibile solo se si verificano le seguenti condizioni:

1. tutti i bersagli sono stati colpiti,
2. nessun bersaglio è stato colpito da più di una freccia.

La condizione 1. è dovuta al fatto che g, dovendo avere dominio B, deve agire su tutti i punti di B stesso; la condizione 2. è dovuta al fatto che g, per essere una funzione, deve associare ad ogni punto del suo dominio (cioè di B) un solo punto del suo codominio (cioè di A). Se si conviene di restringere il codominio di f alla sua immagine $f(A)$, si può naturalmente prescindere dalla prima condizione.

La funzione g, se esiste, è detta *inversa* di f e si indica[8] con la notazione f^{-1}.

[8] Purtroppo si tratta di una notazione leggermente infelice in quanto può ingenerare confusione con il simbolo di funzione reciproco, $1/f$. Tuttavia è la notazione ufficialmente adottata dalla quasi totalità degli autori.

La definizione formale è la seguente.

Definizione 1.24. *Data una funzione $f: A \to B$ individuata dal sottoinsieme F del prodotto cartesiano $A \times B$, essa si dice* invertibile *se l'insieme*

$$F^{-1} = \{ (y,x) \mid (x,y) \in F \} \tag{1.5}$$

definisce, a sua volta, una funzione.

Tenendo conto di quanto detto a proposito delle funzioni iniettive e delle funzioni suriettive, si può facilmente concludere che una funzione è invertibile se e solo se è contemporaneamente iniettiva e suriettiva. Se si restringe il codominio di f all'immagine si può prescindere dalla suriettività.

Si noti che, se f è una funzione invertibile ($f: A \to B$) e f^{-1} ne è l'inversa ($f^{-1}: B \to A$), si ha, contemporaneamente,

$$(f^{-1} \circ f)(x) = x, \quad \forall x \in A \quad \text{e} \quad (f \circ f^{-1})(y) = y, \quad \forall y \in B. \tag{1.6}$$

La funzione $f^{-1} \circ f$ fa corrispondere ad ogni x di A lo stesso x, cioè è la funzione *identità* di A; la funzione $f \circ f^{-1}$ fa corrispondere ad ogni y di B lo stesso y, cioè è la funzione *identità* di B. Se la funzione invertibile f è, come nella maggior parte dei casi che ci interesseranno, una funzione di $\mathbb{R}$ in $\mathbb{R}$, allora sia $f^{-1} \circ f$ che $f \circ f^{-1}$ sono la funzione identità di $\mathbb{R}$, ovvero la funzione che ha come grafico cartesiano la bisettrice del primo e terzo quadrante.

Esempio 1.46. Consideriamo la funzione reale di variabile reale data dalla legge $f(x) = 3x - 2$. Si tratta di una funzione iniettiva e suriettiva come si deduce subito dal suo grafico cartesiano (una retta non orizzontale) e come si può provare anche per calcolo diretto. Sia infatti y un qualunque numero reale; posto $y = 3x - 2$, si può ricavare $x = (y+2)/3$, ovvero esiste un e un solo x di $\mathbb{R}$ di cui y è immagine. Sempre dal grafico si può addirittura dedurre che la funzione è strettamente crescente, ma la cosa si può anche provare per calcolo diretto. Se infatti $x_1 < x_2$ sono due numeri reali, si ha $3x_1 < 3x_2$ e, successivamente, $3x_1 - 2 < 3x_2 - 2$, ovvero $y_1 < y_2$.

La funzione è dunque invertibile. Per ricavare esplicitamente la funzione inversa si può richiamare il calcolo già effettuato: posto $y = 3x - 2$, si può ricavare $x = (y+2)/3$. La funzione inversa è dunque quella che a y fa corrispondere $(y+2)/3$. In ragione del fatto che si usa indicare sempre con x la variabile indipendente si preferisce scrivere la funzione inversa come

$$f^{-1}: \mathbb{R} \to \mathbb{R}, \quad x \mapsto \frac{x+2}{3}.$$

Per un controllo della proprietà che lega una funzione e la sua inversa si ha:

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

e

$$(f^{-1} \circ f)(x) = f^{-1}(f(x)) = \frac{(3x-2)+2}{3} = x.$$

Esempio 1.47. Sia data la funzione esponenziale $f(x) = 2^x$, di dominio $\mathbb{R}$ e codominio $\mathbb{R}$. Possiamo restringere il codominio all'immagine senza alcun sostanziale sacrificio, ottenendo una funzione iniettiva e suriettiva di $\mathbb{R}$ in $\mathbb{R}^+$. La sua inversa è la funzione, di $\mathbb{R}^+$ in $\mathbb{R}$, definita da $f^{-1}(x) = \log_2 x$. Si ha, come è ben noto,

$$2^{\log_2 x} = x, \quad \forall x \in \mathbb{R}^+ \quad \text{e} \quad \log_2(2^x) = x, \quad \forall x \in \mathbb{R}.$$

Esempio 1.48. Sia data la funzione di $\mathbb{R}$ in $\mathbb{R}$ data da $f(x) = x^2$. La funzione non è né suriettiva né iniettiva. Come già detto più volte, è facile ottenere da essa una funzione suriettiva che mantenga nella sostanza le stesse caratteristiche. Per ottenere una funzione iniettiva occorre invece una modifica del dominio, considerandone solo una parte: si sceglie in genere l'insieme dei reali maggiori o uguali a zero. Si noti che questo modifica sostanzialmente la funzione: il grafico sarà ora solo una semiparabola contenuta nel primo quadrante.

Considerata allora la funzione

$$g: [0, +\infty[\to [0, +\infty[, \quad x \mapsto x^2,$$

essa è invertibile e la sua inversa si chiama funzione radice quadrata

$$g^{-1}: [0, +\infty[\to [0, +\infty[, \quad x \mapsto \sqrt{x}.$$

1.14.1 Il grafico di una funzione e della sua inversa

Poiché l'inversa di una funzione f è definita semplicemente come l'insieme

$$F^{-1} = \{ (y, x) \mid (x, y) \in F \}$$

essa semplicemente "scambia la x con la y": il suo grafico si otterrà allora da quello della funzione f con una simmetria rispetto alla bisettrice del primo e terzo quadrante. Le figure che seguono mostrano i grafici delle funzioni considerate negli esempi 1.46, 1.47 e 1.48 (in linea continua) e delle loro inverse (in linea tratteggiata).

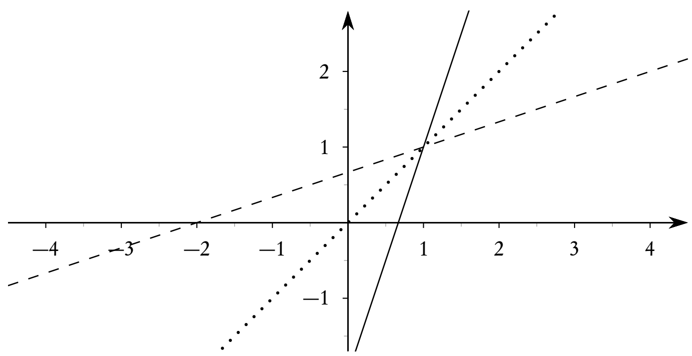


Figura 1.21 *La funzione* $f(x) = 3x - 2$ *e la sua inversa* $f^{-1}(x) = (x+2)/3$

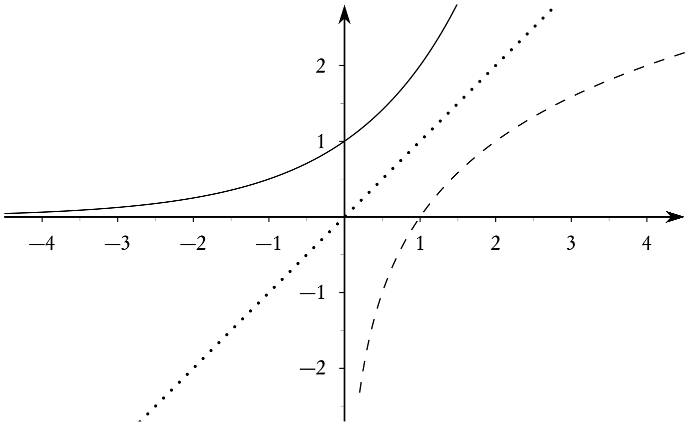


Figura 1.22 *La funzione* $f(x) = 2^x$ *e la sua inversa* $f^{-1}(x) = \log_2 x$

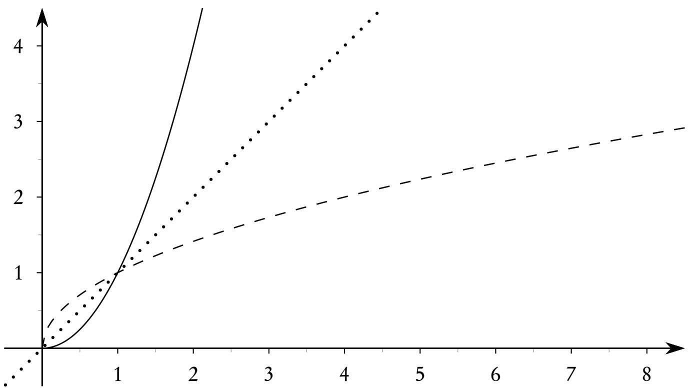


Figura 1.23 *Restrizione della funzione* $f(x) = x^2$ *e sua inversa* $f^{-1}(x) = \sqrt{x}$

1.15 Modelli lineari e non lineari

Le funzioni di $\mathbb{R}$ in $\mathbb{R}$ definite dalla regola $f(x) = ax$ oppure dalla regola $f(x) = ax + b$ giocano un ruolo molto importante nelle applicazioni e prendono il nome, rispettivamente, di *funzioni lineari* e *funzioni affini*. Le funzioni lineari hanno come grafico cartesiano una retta non verticale per l'origine, le funzioni affini una retta non verticale e non passante per l'origine.

Le funzioni lineari e affini possono essere utilizzate nella costruzione di un semplice modello di natura economica, come nell'esempio che segue. Un'impresa produce un bene sostenendo dei costi fissi di produzione (ad esempio collegati agli impianti) e dei costi che dipendono dalla quantità del bene prodotto. Dalla vendita del bene l'impresa ottiene un ricavo e dalla differenza fra ricavi e costi si può ottenere il profitto d'impresa.

Nell'ipotesi più semplice si può supporre che sia i costi variabili che il ricavo siano direttamente proporzionali alla quantità del bene prodotta. In questo caso la funzione che rappresenta il costo totale è una funzione affine della quantità , mentre la funzione ricavo è una funzione lineare.

Indicando con q la quantità prodotta del bene, con c_f il costo fisso d'impresa, con c e p, rispettivamente, il costo unitario di produzione e il prezzo unitario di vendita, si possono determinare la funzione di costo totale $C(q)$, di ricavo totale $R(q)$ e di profitto $\Pi(q)$ modo seguente:

$$C(q) = c_f + cq, \tag{1.7}$$

$$R(q) = pq, \tag{1.8}$$

$$\Pi(q) = R(q) - C(q) = pq - c_f - cq = (p-c)q - c_f. \tag{1.9}$$

Le funzioni lineari e affini trovano applicazione in moltissime situazioni in cui si debbano modellizzare situazioni concrete: basta pensare alle imposte come funzione del reddito o allo spazio percorso come funzione del tempo in un moto uniforme. Risulta importante notare che nei modelli lineari la crescita o la decrescita avvengono ad un tasso costante e questa ipotesi risulta spesso non realistica, per cui si deve ricorrere ad altri modelli.

Riprendendo l'esempio precedente dell'impresa che produce un bene sostenendo costi di produzione e costi fissi, in molti casi è realistico supporre che il profitto d'impresa cresca all'aumentare della produzione e, raggiunto un certo limite, decresca. Una funzione quadratica può spesso descrivere un comportamento di questo tipo. Detta q la quantità prodotta del bene, si può assumere che la funzione profitto sia la seguente

$$\Pi(q) = aq^2 + bq + c, \quad a < 0,\ b > 0,\ c < 0. \tag{1.10}$$

Avendo questa funzione come grafico una parabola con la concavità verso il basso, si capisce bene come modellizzi bene la situazione descritta. Naturalmente, scegliendo $a > 0$, si avrebbe un modello in cui ad una fase di decrescita segue una fase di crescita.

Altri modelli di uso frequente sono quelli esponenziali o logaritmici; tra questi si possono ricordare il fenomeno della crescita di una popolazione, il decadimento di una sostanza radioattiva, il valore di un investimento a capitalizzazione composta.

Ad esempio se si vuole modellizzare un fenomeno di crescita rispetto al tempo in cui vi sia un raddoppio in ogni intervallo unitario si può usare una funzione del tipo

$$f(t) = k2^t. \tag{1.11}$$

Se invece si vuole modellizzare fenomeni di decrescita si possono considerare funzioni esponenziali del tipo $f(t) = ka^t$, con base a dell'esponenziale minore di 1.

1.16 Esercizi

Esercizio 1.1. *Discutere le seguenti questioni in modo sintetico, ma esauriente.*

- *Si possono trovare due insiemi uno aperto e l'altro chiuso tali che la loro unione sia un insieme aperto?*
- *Si possono trovare due insiemi uno aperto e l'altro chiuso tali che la loro intersezione sia un insieme aperto?*
- *Si possono trovare due insiemi entrambi aperti tali che la loro unione sia un insieme aperto?*
- *Si possono trovare due insiemi entrambi chiusi tali che la loro unione sia un insieme chiuso?*

Esercizio 1.2. *Discutere le seguenti questioni in modo sintetico, ma esauriente.*

- *Se si costruisce una funzione definita a pezzi (in due pezzi) usando due funzioni limitate si ottiene sempre una funzione limitata.*
- *Se si costruisce una funzione definita a pezzi (in due pezzi) usando una funzione limitata e una funzione illimitata si può ottenere una funzione limitata.*
- *Se si costruisce una funzione definita a pezzi (in due pezzi) usando una funzione crescente e una decrescente si può ottenere una funzione crescente.*
- *Se si costruisce una funzione definita a pezzi (in due pezzi) usando due funzioni decrescenti si può ottenere una funzione crescente.*

Esercizio 1.3. *Determinare il dominio delle seguenti funzioni e tutte le caratteristiche di tale insieme (limitato, illimitato, aperto, chiuso, punti interni, esterni, ecc.)*

1. $f(x) = x + 1$;
2. $f(x) = \dfrac{x}{2-x}$;
3. $f(x) = \sqrt{x+1}$;
4. $f(x) = \sqrt{x} \cdot \dfrac{1}{1+x}$;
5. $f(x) = \sqrt{\dfrac{x}{2-x}}$;
6. $f(x) = \sqrt{(x-1)(1+x)}$;
7. $f(x) = \dfrac{\sqrt{x}}{\sqrt{2x-3}}$;
8. $f(x) = \sqrt{\dfrac{x}{2x-3}}$;
9. $f(x) = \dfrac{\sqrt{x^2-9}}{3-x}$;
10. $f(x) = \sqrt{2x}\sqrt{x+3}$;
11. $f(x) = \sqrt{x+1} - x + \sqrt{2-x}$.

Esercizio 1.4. *Determinare il dominio e le sue caratteristiche (aperto, ecc., punti interni, ecc.) per le seguenti funzioni di due variabili.*

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

3. $f(x,y) = \sqrt{x} \cdot \dfrac{1}{1+y}$;
4. $f(x,y) = \sqrt{\dfrac{y}{1+x}}$;
5. $f(x,y) = \sqrt{(x^2-1)(2-y)}$;
6. $f(x,y) = \sqrt{xy}$;
7. $f(x,y) = \dfrac{\sqrt{x^2-4}}{y+3}$;
8. $f(x,y) = \sqrt{2x}\sqrt{2y-3}$;
9. $f(x,y) = \sqrt{x+y-1} - \sqrt{x-y}$.

2 Matrici e sistemi lineari

2.1 Introduzione

Questo capitolo è dedicato ai sistemi lineari, un argomento che trova applicazioni in numerosi campi, non ultimo in campo economico. Per cominciare proponiamo due semplici esempi relativi a un argomento classico di geometria analitica, ovvero quello di trovare l'intersezione di due rette nel piano cartesiano e a un problema di composizione di cibo per animali.

Nel caso dell'intersezione di due rette, ricordiamo che ciascuna delle due rette è rappresentata da un'equazione di primo grado in due incognite: la determinazione del loro eventuale punto di intersezione si fa risolvendo un sistema del tipo

$$\begin{cases} ax + by + c = 0 \\ dx + ey + f = 0 \end{cases} .$$

Il sistema ha una soluzione se le rette sono incidenti, infinite soluzioni se sono coincidenti, nessuna soluzione se sono parallele.

In vista di quanto diremo è utile scrivere il sistema precedente nella forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} .$$

In questa scrittura si evidenziano

- la coppia delle incognite (x_1, x_2), che conviene scrivere in colonna per un motivo che apparirà chiaro nel seguito:
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix};$$
- la coppia dei termini noti (b_1, b_2), che, ancora una volta, conviene scrivere in colonna:
$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix};$$
- una tabella con due righe e due colonne costituita dai coefficienti delle incognite:
$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Passiamo ora ad esaminare un problema di composizione di cibo per animali. Un certo animale deve assumere giornalmente, per una dieta equilibrata, 80 grammi di proteine, 200 grammi di carboidrati, 50 grammi di grassi. Un negozio ha in magazzino quattro tipi di mangime da utilizzare, disponibili in sacchetti da 100 grammi l'uno, che chiameremo "unità di mangime". Questi mangimi forniscono, per ogni unità, le seguenti quantità di proteine, carboidrati e grassi.

	Proteine	Carboidrati	Grassi
Mangime 1	5	25	3
Manigme 2	6	30	4
Mangime 3	7	15	8
Mangime 4	13	10	7

Per determinare una miscela dei quattro mangimi che fornisca le quantità ritenute ottimali di ciascun componente, indichiamo con x_1, x_2, x_3, x_4 le unità necessarie dei vari mangimi.

Si dovrà risolvere il seguente sistema di equazioni.

$$\begin{cases} 5x_1 + 6x_2 + 7x_3 + 13x_4 = 80 \\ 25x_1 + 30x_2 + 15x_3 + 10x_4 = 200 \\ 3x_1 + 4x_2 + 8x_3 + 7x_4 = 50 \end{cases} .$$

Anche in questa scrittura si evidenziano

- la quaterna delle incognite (x_1, x_2, x_3, x_4), che conviene sempre scrivere in colonna:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix};$$

- la terna dei termini noti $(80, 200, 50)$, che, ancora una volta, conviene scrivere in colonna:

$$\begin{pmatrix} 80 \\ 200 \\ 50 \end{pmatrix};$$

- una tabella con tre righe e quattro colonne costituita dai coefficienti delle incognite:

$$\begin{pmatrix} 5 & 6 & 7 & 13 \\ 25 & 30 & 15 & 10 \\ 3 & 4 & 8 & 7 \end{pmatrix}.$$

Quello che ci interessa in questi esempi è il fatto che abbiamo evidenziato due liste ordinate di numeri reali, che abbiamo scritto in riga o in colonna, e una tabella di numeri reali con un certo numero di righe e un certo numero di colonne.

Esistono molte situazioni che inducono a considerare oggetti di questo tipo e che ci consentiranno di individuare una strategia per la risoluzione dei sistemi lineari. Le liste ordinate prendono il nome di *vettori*, le tabelle prendono il nome di *matrici*. Ci occuperemo ora in maniera abbastanza sistematica di questi oggetti.

2.2 Vettori

Definizione 2.1. *Si dice* vettore n-dimensionale*, o semplicemente* vettore*, una n-upla ordinata di numeri reali: $\vec{u} = (u_1, u_2, \dots, u_n)$. L'insieme dei vettori n-dimensionali si indica con $\mathbb{R}^n$: esso non è altro che il* *prodotto cartesiano di* $\mathbb{R}$ *per se stesso* n*-volte. Due vettori* $(u_1, u_2, \dots, u_n)$ *e* $(v_1, v_2, \dots, v_n)$ *sono uguali se e solo* $u_1 = v_1, \dots, u_n = v_n$. *Il vettore* $\vec{0} = (0, 0, \dots, 0)$ *(cioè l'*n*-upla costituita da tutti zeri), si chiama* vettore nullo *di* $\mathbb{R}^n$.

Definizione 2.2. *Dati due vettori* $\vec{u} = (u_1, u_2, \dots, u_n)$ *e* $\vec{v} = (v_1, v_2, \dots, v_n)$ *si chiama loro somma il vettore*

$$\vec{u} + \vec{v} = (u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n) = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n).$$

Definizione 2.3. *Dato un vettore* $\vec{u} = (u_1, u_2, \dots, u_n)$ *di* $\mathbb{R}^n$ *e un numero reale* k *si chiama* prodotto del vettore $\vec{u}$ per il numero reale k *il vettore*

$$k\vec{u} = (ku_1, ku_2, \dots, ku_n)$$

Nell'insieme $\mathbb{R}^n$ si possono dunque eseguire due operazioni, la somma, detta operazione interna perché sia i due addendi che il risultato sono vettori, e la moltiplicazione per un numero (si dice anche moltiplicazione per uno scalare), detta operazione esterna perché dei tre oggetti coinvolti (due in partenza e uno in arrivo) uno non è un vettore.

Per l'operazione di somma tra vettori valgono le usuali proprietà della somma tra numeri: commutativa, associativa, esistenza dell'elemento neutro (il vettore nullo) e dell'opposto (che si indica con $-\vec{u}$).

Per l'operazione di prodotto per un numero valgono alcune proprietà, simili (ma ovviamente non identiche perché qui moltiplico tra di loro oggetti diversi: un numero e un vettore) alle proprietà del prodotto fra numeri.

1. $(h+k)\vec{u} = h\vec{u} + k\vec{u}$, $\quad h(\vec{u} + \vec{v}) = h\vec{u} + h\vec{v}$ (proprietà distributive).
2. $h(k\vec{u}) = (hk)\vec{u}$.
3. $0\vec{u} = \vec{0}$, $\quad 1\vec{u} = \vec{u}$, $\quad -1\vec{u} = -\vec{u}$.

In molte situazioni conviene scrivere i vettori con gli elementi disposti su una colonna anziché su una riga:

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Nell'insieme dei vettori si introduce una ulteriore operazione, come precisato nella seguente definizione.

Definizione 2.4. *Dati due vettori* $\vec{u} = (u_1, u_2, \dots, u_n)$ *e* $\vec{v} = (v_1, v_2, \dots, v_n)$ *si chiama loro* prodotto scalare *il numero reale*

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

Si noti che anche questa operazione è esterna, perché dei tre oggetti coinvolti l'ultimo, cioè il risultato, non è un vettore.

Esempio 2.1. Una fabbrica produce tre oggetti, P_1, P_2, P_3, venduti al prezzo unitario di 3000, 2500, e 4000 rispettivamente, e ha tre stabilimenti, A, B, C, che producono le seguenti quantità dei tre oggetti:

$$\begin{array}{c|ccc} & A & B & C \\ \hline P_1 & 100 & 50 & 80 \\ P_2 & 60 & 200 & 30 \\ P_3 & 150 & 180 & 250 \end{array}\,.$$

Possiamo introdurre un "vettore di produzione" per ciascuno stabilimento:

$$\vec{x}_A = \begin{pmatrix} 100 \\ 60 \\ 150 \end{pmatrix}, \quad \vec{x}_B = \begin{pmatrix} 50 \\ 200 \\ 180 \end{pmatrix}, \quad \vec{x}_C = \begin{pmatrix} 80 \\ 30 \\ 250 \end{pmatrix},$$

e un "vettore prezzo unitario":

$$\vec{p} = \begin{pmatrix} 3000 \\ 2500 \\ 4000 \end{pmatrix}.$$

Il totale della produzione sarà

$$\vec{x}_T = \vec{x}_A + \vec{x}_B + \vec{x}_C = \begin{pmatrix} 230 \\ 290 \\ 580 \end{pmatrix}.$$

Il ricavo di ciascuno stabilimento sarà:

$$\vec{x}_A \cdot \vec{p} = 1050000, \quad \vec{x}_B \cdot \vec{p} = 1370000, \quad \vec{x}_C \cdot \vec{p} = 1315000.$$

Il ricavo totale si potrà ovviamente trovare o sommando i tre ricavi o facendo $\vec{x}_T \cdot \vec{p}$: si tratta di un semplice esempio di applicazione della proprietà distributiva del prodotto scalare.

2.2.1 Dipendenza ed indipendenza lineare

Definizione 2.5. *Siano $\vec{w}$ e $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_r$ vettori di $\mathbb{R}^n$. Si dice che il vettore $\vec{w}$ è* combinazione lineare *dei vettori $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_r$ se esistono r scalari (numeri reali) c_1, c_2, ..., c_r tali che*

$$\vec{w} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_r\vec{v}_r.$$

Esempio 2.2. Si considerino i vettori $\vec{v}_1 = (1,1,0)$, $\vec{v}_2 = (0,0,1)$ e $\vec{v}_3 = (3,3,-2)$. È facile verificare che

$$\vec{v}_3 = 3\vec{v}_1 - 2\vec{v}_2,$$

ovvero che $\vec{v}_3$ è combinazione lineare dei vettori $\vec{v}_1$ e $\vec{v}_2$ con coefficienti 3 e −2.

Esempio 2.3. Si considerino i vettori $\vec{v}_1 = (1,0,0)$, $\vec{v}_2 = (0,1,0)$ e $\vec{v}_3 = (0,0,1)$. In questo caso non è possibile scrivere nessuno dei tre vettori come combinazione lineare degli altri due: per esempio qualunque combinazione dei vettori $\vec{v}_2$ e $\vec{v}_3$ avrà uno zero come primo elemento e quindi non potrà mai dare il vettore $\vec{v}_1$.

Definizione 2.6. *I vettori $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_r$ di $\mathbb{R}^n$ si dicono* linearmente dipendenti *se è possibile scrivere almeno uno di essi come combinazione lineare degli altri,* linearmente indipendenti *in caso contrario.*

Vale il seguente teorema, che discende quasi immediatamente dalla definizione.

Teorema 2.7. *I vettori $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_r$ di $\mathbb{R}^n$ sono linearmente indipendenti se e solo se dall'uguaglianza*

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_r\vec{v}_r = \vec{0} = (0, 0, \ldots, 0)$$

segue $c_1 = c_2 = \cdots = c_r = 0$. Se invece l'uguaglianza precedente vale con almeno uno dei coefficienti diverso da zero, allora i vettori sono linearmente dipendenti.

Il problema di verificare se un insieme di vettori è o no indipendente conduce alla risoluzione di un sistema lineare, come si può vedere dal seguente esempio.

Esempio 2.4. Si considerino ad esempio i vettori $\vec{v}_1 = (1, -1, 2)$, $\vec{v}_2 = (2, 1, 3)$ e $\vec{v}_3 = (1, 0, 3)$ e supponiamo di voler stabilire se $\vec{v}_3$ è combinazione lineare di $\vec{v}_1$ e di $\vec{v}_2$. Si devono ricercare le soluzioni dell'equazione:

$$c_1 \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$$

nelle incognite c_1 e c_2. Tenendo conto che due vettori sono uguali se e solo se hanno gli stessi elementi, si ottiene:

$$\begin{cases} c_1 & + & 2c_2 & = & 1 \\ -c_1 & + & c_2 & = & 0 \\ 2c_1 & + & 3c_2 & = & 3 \end{cases}.$$

2.3 Matrici

Definizione 2.8. *Si dice* matrice *$m \times n$ (in $\mathbb{R}$) una tabella rettangolare costituita da numeri reali disposti su m righe e n colonne. Indicheremo una matrice con la scrittura*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

segnalando che alcuni usano le parentesi quadre al posto delle tonde.

Per indicare esplicitamente che la matrice ha m righe e n colonne scriveremo $A_{m \times n}$ o anche, se non ci sono possibilità di equivoco, A_{mn}. Gli $m \cdot n$ numeri reali che costituiscono una matrice si dicono i suoi *elementi*: l'elemento generico sarà indicato con a_{ij} dove i è l'*indice di riga* e j è l'*indice di colonna*.

Due matrici dello stesso tipo sono uguali se hanno tutti gli elementi corrispondenti uguali.

In sostanza una matrice è una fila di numeri, solo che i numeri invece di essere scritti in un'unica riga (in successione), sono scritti separandoli su varie righe. Come vedremo non ci sarà nessun problema a fare la somma di due matrici o il prodotto di una matrice per un numero. Il fatto però di aver scritto questi numeri in tabella ci consentirà di introdurre una ulteriore operazione, il prodotto tra matrici, che si rivelerà cruciale per le applicazioni.

Definizione 2.9. *Date due matrici dello stesso tipo, cioè entrambe $m \times n$, si dice somma delle due matrici la matrice $m \times n$ ottenuta sommando ordinatamente gli elementi delle due matrici. Si dice prodotto di una matrice per un numero la matrice ottenuta moltiplicando ciascun elemento della matrice per quel numero.*

Per l'operazione di somma tra matrici valgono le usuali proprietà della somma tra numeri reali.

1. Proprietà *commutativa*: se A e B sono due matrici qualunque (dello stesso tipo) $A+B=B+A$.
2. Proprietà *associativa*: se A, B e C sono tre matrici qualunque (dello stesso tipo) $(A+B)+C = A+(B+C)$. Questo consente di evitare l'uso di parentesi nello scrivere la somma di più matrici.
3. Esistenza dell'*elemento neutro*: si tratta della matrice A_{mn} i cui elementi sono tutti nulli, che si dice *la matrice nulla $m \times n$* e si indica con 0_{mn}, a volte anche semplicemente con 0 se non ci sono possibilità di equivoci: bisogna però prestare attenzione a non confondere la matrice nulla con il numero 0.
4. Esistenza dell'*opposto*: data una matrice A, esiste una matrice B dello stesso tipo tale che $A+B=0$. La matrice B ha come elementi gli opposti degli elementi della matrice A e si indica con $-A$.

Per l'operazione di prodotto per un numero valgono le seguenti proprietà.

1. $(h+k)A = hA + kA$, $\quad h(A+B) = hA + hB$ (proprietà distributive).
2. $h(kA) = (hk)A$.
3. $0A = 0$, $\quad 1A = A$, $\quad -1A = -A$.

Il fatto che valgano tutte queste proprietà si esprime dicendo che l'insieme di tutte le matrici $m \times n$ è uno spazio vettoriale, di dimensione $m \times n$.

E ora un po' di nomenclatura relativa alle matrici.

- Un matrice si dice *quadrata* se il numero di righe è uguale al numero di colonne. In questo caso il numero comune di righe e colonne si dice l'*ordine* della matrice.
- Data una matrice A_{mn} diremo *trasposta* di A la matrice A^{T}_{nm}, ottenuta scambiando in A le righe con le colonne.

  Esempio 2.5. La trasposta di $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 8 \end{pmatrix}$ è la $A^{\mathrm{T}} = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 8 \end{pmatrix}$.

- Una matrice quadrata si dice simmetrica se coincide con la sua trasposta.

  Esempio 2.6. La matrice $A = \begin{pmatrix} 1 & 3 & 0 & 2 \\ 3 & 2 & 3 & 1 \\ 0 & 3 & 0 & -1 \\ 2 & 1 & -1 & 1 \end{pmatrix}$ è simmetrica.

- In una matrice quadrata gli elementi aventi lo stesso numero di riga e di colonna costituiscono la *diagonale principale*

  Esempio 2.7. Nella matrice seguente gli elementi incasellati sono quelli della diagonale principale.

$$A = \begin{pmatrix} \boxed{1} & 3 & 0 & 2 \\ 3 & \boxed{2} & 3 & 1 \\ 0 & 3 & \boxed{0} & -1 \\ 2 & 1 & -1 & \boxed{1} \end{pmatrix}$$

– Una matrice quadrata si dice *diagonale* se tutti gli elementi fuori dalla diagonale principale sono nulli.

Esempio 2.8. La matrice $A = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}$ è diagonale.

– La matrice quadrata diagonale in cui tutti gli elementi della diagonale principale valgono 1 si dice *matrice unità* $n \times n$ e si indica con I_n.

Le matrici a una sola riga (e n colonne) si dicono anche *vettori-riga*, quelle a una sola colonna (e n righe) e si dicono anche *vettori-colonna*. Spesso i vettori colonna si indicano con scritture del tipo $\vec{u}$, $\vec{v}$, ecc.

2.4 Il prodotto tra matrici

Come già anticipato, il fatto di avere scritto gli $m \times n$ numeri che compongono una matrice in una tabella anzichè in una riga, ci consente di introdurre una nuova operazione nell'insieme delle matrici, precisamente il *prodotto righe per colonne*. La definizione e le proprietà di questa operazione sono un po' più complesse e delicate che non quelle delle due operazioni precedenti e le tratteremo con un po' più di dettaglio.

Conviene subito precisare che, mentre l'operazione di somma si esegue tra matrici dello stesso tipo (cioè con lo stesso numero di righe e di colonne), ciò non accade per il prodotto (righe per colonne) di matrici. La condizione indispensabile perché questa operazione si possa eseguire è che *il numero di colonne della prima matrice sia uguale al numero di righe della seconda*.

Precisamente date una matrice $A_{m\times n}$ e una matrice $B_{n\times p}$, il prodotto, che definiremo tra poco, tra A e B, *nell'ordine*, è un matrice $C_{m\times p}$:

$$C_{m\times p} = A_{m\times n} \cdot B_{n\times p} .$$

Anche se in questa formula abbiamo usato il punto per indicare la moltiplicazione, normalmente non useremo alcun simbolo, esattamente come si fa con il prodotto di numeri, quando non ci sono equivoci.

È ovvio che per questa operazione non ha nemmeno senso chiedersi, in generale, se vale la proprietà commutativa: infatti di solito se si può fare il prodotto AB, non si può fare quello BA. Per esempio il prodotto $A_{2\times 3}B_{3\times 4}$ si può fare e produce una matrice $C_{2\times 4}$, mentre il prodotto inverso non si può fare in quanto non è soddisfatta la regola generale sul numero di righe della prima matrice e il numero di colonne della seconda. Alcune volte ha senso "invertire l'ordine" nel prodotto, come nel caso $A_{2\times 3}B_{3\times 2}$, che produce una matrice $C_{2\times 2}$; il prodotto in ordine inverso produce invece una matrice completamente diversa, in quanto si tratta di una matrice $D_{3\times 3}$. Se per caso le due matrici sono quadrate e dello stesso ordine, allora il prodotto, in qualunque ordine, dà una matrice ancora dello stesso ordine: tuttavia, come vedremo, anche in questo caso uno scambio di ordine può cambiare il risultato.

L'algoritmo da usare per eseguire il prodotto è un po' noioso da esprimere a parole e cercheremo di sintetizzarlo in modo da ottenere una formulazione facilmente utilizzabile, ragionando su un esempio.

Supponiamo di dover moltiplicare una matrice $A_{2\times 3}$ per una matrice $B_{3\times 4}$: il prodotto dovrà essere una matrice $C_{2\times 4}$. Potremo simbolicamente rappresentare la situazione con la seguente figura

$$\begin{pmatrix} \square & \square & \square \\ \square & \square & \square \end{pmatrix} \cdot \begin{pmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{pmatrix} = \begin{pmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \end{pmatrix}.$$

Ciascuno dei box da riempire della matrice C ha un ben preciso numero di riga e di colonna, come evidenziato nella figura che segue:

$$\begin{pmatrix} \square & \square & \square \\ \square & \square & \square \end{pmatrix} \cdot \begin{pmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{pmatrix} = \begin{pmatrix} \boxed{1,1} & \boxed{1,2} & \boxed{1,3} & \boxed{1,4} \\ \boxed{2,1} & \boxed{2,2} & \boxed{2,3} & \boxed{2,4} \end{pmatrix}.$$

Consideriamo allora uno dei box da riempire, per esempio il box $\boxed{2,3}$. Basterà moltiplicare ciascun elemento della 2ª riga della prima matrice (che ha 3 elementi), per *il corrispondente* elemento della 3ª colonna della seconda matrice (che ha sempre 3 elementi) e poi sommare i prodotti così ottenuti: il risultato andrà ad occupare il box $\boxed{2,3}$ della matrice prodotto. La figura che segue illustra il procedimento.

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ \boxed{2,1} & \boxed{2,2} & \boxed{2,3} \end{pmatrix} \cdot \begin{pmatrix} \cdot & \cdot & \boxed{1,3} & \cdot \\ \cdot & \cdot & \boxed{2,3} & \cdot \\ \cdot & \cdot & \boxed{3,3} & \cdot \end{pmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \boxed{2,3} & \cdot \end{pmatrix}.$$

Un esempio numerico chiarirà ancora meglio come si opera.

$$\begin{pmatrix} 1 & -2 & 4 \\ \boxed{-3} & \boxed{-1} & \boxed{5} \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & \boxed{-2} & 3 \\ 1 & -1 & \boxed{3} & 4 \\ 5 & 2 & \boxed{1} & 3 \end{pmatrix} = \begin{pmatrix} 20 & 11 & -4 & 7 \\ 18 & 8 & \boxed{8} & 2 \end{pmatrix}$$

Il numero 8, al posto $(2,3)$ della matrice prodotto, è stato ottenuto[1] con il seguente calcolo:

$$(-3)(-2) + (-1)(3) + (5)(1) = 8\,.$$

Lo studente può provare, per esercizio, ad ottenere gli altri valori indicati.

Come ulteriore esempio consideriamo il prodotto seguente:

$$\begin{pmatrix} 1 & 2 & -2 & 4 \\ -3 & 2 & -1 & 5 \\ 1 & 0 & 3 & -1 \\ -1 & 1 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & -2 & 3 \\ 1 & -1 & 3 & 4 \\ 2 & 1 & -1 & 1 \\ 5 & 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 20 & 5 & 10 & 21 \\ 19 & 4 & 18 & 13 \\ 3 & 2 & -6 & 3 \\ 3 & 0 & 3 & 3 \end{pmatrix}.$$

[1] Tutti i software di calcolo simbolico consentono di ottenere facilmente il prodotto di matrici, anche molto grandi. In effetti il calcolo riportato nell'esempio è stato ottenuto con *Mathematica*, compresa la formattazione necessaria per la stampa con LaTeX.

Oltre a controllare la correttezza del risultato[2], è utile verificare che scambiando l'ordine il risultato è completamente diverso, anche se in questo caso lo scambio di ordine è perfettamente legittimo, trattandosi di matrici quadrate entrambe 4×4.

$$\begin{pmatrix} 2 & 1 & -2 & 3 \\ 1 & -1 & 3 & 4 \\ 2 & 1 & -1 & 1 \\ 5 & 2 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -2 & 4 \\ -3 & 2 & -1 & 5 \\ 1 & 0 & 3 & -1 \\ -1 & 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} -6 & 9 & -5 & 15 \\ 3 & 4 & 16 & -4 \\ -3 & 7 & -6 & 14 \\ -3 & 17 & -3 & 29 \end{pmatrix}.$$

Naturalmente il fatto che, *di solito*, non sia consentito scambiare l'ordine nel prodotto, non significa che la cosa non sia mai possibile. Un esempio è fornito dal calcolo che segue.

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 \\ 0 & 7 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix},$$

ma anche

$$\begin{pmatrix} 5 & 0 \\ 0 & 7 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}.$$

Un altra cosa che distingue il prodotto di matrici dall'ordinario prodotto di numeri è il fatto che, per il prodotto di matrici, *non vale la legge dell'annullamento del prodotto*: può benissimo darsi che il prodotto di due matrici sia la matrice nulla (cioè formata da tutti zeri), senza che le due matrici fattore siano nulle. Un esempio è il seguente:

$$\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Si verifica subito che si ha anche

$$\begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

cioè che il prodotto rimane la matrice nulla, anche scambiando l'ordine. Visto che però il prodotto di due matrici non è commutativo, non deve stupire il fatto che, in un prodotto di due matrici A e B, possa succedere che $AB = 0$, mentre $BA \neq 0$, come mostra l'esempio che segue.

$$\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

mentre

$$\begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 4 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ci sono naturalmente anche cose che il prodotto di matrici ha in comune con l'ordinario prodotto di numeri e sono, precisamente, le seguenti.

1. La *proprietà associativa*, ovvero $(AB)C = A(BC)$, che consente di evitare l'uso di parentesi.

[2]Sempre ottenuto con *Mathematica*, ma è meglio non fidarsi!

2. La *proprietà distributiva* rispetto alla somma, ovvero $(A+B)C = AC+BC$ e $A(B+C) = AB+AC$. Si noti che, non essendo il prodotto commutativo, le due scritture sono diverse. Anche con il prodotto di matrici si adotta la convenzione in uso con il prodotto di numeri, ovvero che il prodotto ha precedenza sulla somma: scrivendo $AB + C$ intendiamo $(AB) + C$.
3. L'esistenza, nel caso ci si limiti ad operare *solo* ed *esclusivamente* con matrici quadrate dello stesso ordine n, dell'elemento neutro, cioè di una particolare matrice che non ha alcuna influenza nel prodotto, come il numero 1 nel prodotto tra numeri ("il prodotto per 1 non occorre scriverlo!"). Questa matrice è la I_n che, non a caso, abbiamo chiamato *matrice unità*.

Si tengano ben presenti le condizioni per potere effettuare le operazioni indicate: la somma di due matrici si può fare solo tra matrici dello stesso tipo (e produce ancora una matrice dello stesso tipo); il prodotto tra matrici si può fare solo se il numero di colonne della prima matrice è uguale al numero di righe della seconda (e produce, in generale, una matrice diversa da entrambe). Operando nell'ambito delle matrici quadrate di ordine n, non c'è, naturalmente, alcun problema: le somme e i prodotti si possono sempre eseguire e tutti i risultati sono ancora matrici dello stesso ordine (ma attenzione a non scambiare i fattori in un prodotto!).

Le proprietà indicate si possono dimostrare rigorosamente, ma ci accontentiamo di una verifica in casi concreti, invitando ancora una volta lo studente, come utile esercizio, a fare tutti i calcoli necessari.

Esempio 2.9.

$$\left[\begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 5 \\ 2 & -6 \end{pmatrix}\right] \cdot \begin{pmatrix} 3 & 2 \\ -4 & 5 \end{pmatrix} =$$

$$= \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix} \cdot \left[\begin{pmatrix} 1 & 5 \\ 2 & -6 \end{pmatrix} \cdot \begin{pmatrix} 3 & 2 \\ -4 & 5 \end{pmatrix}\right] = \begin{pmatrix} 56 & -24 \\ 137 & -131 \end{pmatrix}.$$

Esempio 2.10.

$$\left[\begin{pmatrix} -3 & 1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 3 & -4 \\ -2 & 6 \end{pmatrix}\right] \cdot \begin{pmatrix} 1 & -5 \\ 4 & 3 \end{pmatrix} =$$

$$= \left[\begin{pmatrix} -3 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -5 \\ 4 & 3 \end{pmatrix}\right] + \left[\begin{pmatrix} 3 & -4 \\ -2 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & -5 \\ 4 & 3 \end{pmatrix}\right] = \begin{pmatrix} -12 & -9 \\ 31 & 29 \end{pmatrix}.$$

Esempio 2.11.

$$\begin{pmatrix} 1 & -2 & 3 & -5 \\ 5 & 3 & 7 & -2 \\ -3 & 6 & 8 & 1 \\ 4 & -9 & 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 & 3 & -5 \\ 5 & 3 & 7 & -2 \\ -3 & 6 & 8 & 1 \\ 4 & -9 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 & -5 \\ 5 & 3 & 7 & -2 \\ -3 & 6 & 8 & 1 \\ 4 & -9 & 3 & 2 \end{pmatrix}.$$

Un'altra sostanziale differenza tra il prodotto di matrici e il prodotto di numeri è legato al concetto di reciproco. È ben noto che, dato un numero a diverso da 0, esiste uno e un solo numero b diverso da zero tale che $a \cdot b = 1$. Il numero b si indica con a^{-1} o, anche, con $1/a$ e si chiama il reciproco di a. Ci si può porre l'analogo problema con le matrici. Intanto ci si scontra con la difficoltà legata al fatto che il prodotto di matrici non è sempre definito, ma questo si risolve facilmente, limitandosi a considerare solo l'insieme delle matrici quadrate di ordine n, con n fissato. Nonostante questo, data una matrice $A_{n\times n}$, non è affatto detto che esista una matrice $B_{n\times n}$, tale che $AB = I_n$. Come al solito vediamo alcuni esempi per capire il problema.

Esempio 2.12. Sia data la matrice seguente

$$A = \begin{pmatrix} 1 & -2 \\ 3 & 4 \end{pmatrix},$$

e poniamoci il problema della ricerca di una matrice $B_{2\times 2}$ tale che

$$AB = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

La matrice B dovrà essere del tipo

$$B = \begin{pmatrix} x & y \\ z & t \end{pmatrix},$$

con valori delle incognite x, y, z, t da determinare. Se calcoliamo il prodotto AB notiamo che deve essere:

$$\begin{pmatrix} 1 & -2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} x-2z & y-2t \\ 3x+4z & 4t+3y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

da qui otteniamo un sistema di 4 equazioni in 4 incognite (di primo grado!), di immediata risoluzione:

$$\begin{cases} x-2z=1 \\ y-2t=0 \\ 3x+4z=0 \\ 4t+3y=1 \end{cases} \quad \Rightarrow \quad \begin{cases} x = \dfrac{2}{5} \\ y = \dfrac{1}{5} \\ z = -\dfrac{3}{10} \\ t = \dfrac{1}{10} \end{cases} \quad \Rightarrow \quad B = \begin{pmatrix} \dfrac{2}{5} & \dfrac{1}{5} \\ -\dfrac{3}{10} & \dfrac{1}{10} \end{pmatrix}$$

Si noti che il sistema è di facile risoluzione perché può essere spezzato in due separati sistemi di due equazioni in due incognite. Dunque in questo caso è stato possibile trovare una matrice B tale che $AB = I_2$. Si può verificare che vale anche $BA = I_2$.

Purtroppo la cosa non è sempre possibile, come mostra il seguente esempio.

Esempio 2.13. Sia data la matrice seguente

$$A = \begin{pmatrix} 1 & 2 & -1 \\ -2 & 3 & 1 \\ -4 & 13 & 1 \end{pmatrix},$$

e cerchiamo, se esiste, una matrice B tale che $AB = I_3$. Usando la stessa strategia di prima, il problema si potrà scrivere nella forma seguente:

$$\begin{pmatrix} 1 & 2 & -1 \\ -2 & 3 & 1 \\ -4 & 13 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{pmatrix} =$$

$$= \begin{pmatrix} x_1 + 2x_4 - x_7 & x_2 + 2x_5 - x_8 & x_3 + 2x_6 - x_9 \\ -2x_1 + 3x_4 + x_7 & -2x_2 + 3x_5 + x_8 & -2x_3 + 3x_6 + x_9 \\ -4x_1 + 13x_4 + x_7 & -4x_2 + 13x_5 + x_8 & -4x_3 + 13x_6 + x_9 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si ottiene un sistema di 9 equazioni in 9 incognite, che si può spezzare in 3 sistemi di 3 equazioni in 3 incognite. Non è difficile provare che il sistema non ha alcuna soluzione, da cui si conclude che, questa volta, non esiste una matrice B che soddisfi i requisiti richiesti.

Questi esempi ci conducono a dare la seguente definizione.

Definizione 2.10. *Una matrice $A_{n\times n}$ si dice* invertibile, *se esiste una matrice $B_{n\times n}$ tale che $AB = BA = I_n$. In questo caso la matrice B si chiama* inversa *della matrice A e si indica con A^{-1}, in modo analogo a quanto si fa con i numeri.*

Si prova poi che una tal matrice se esiste è unica. Il fatto che l'inversa non esista sempre ha come conseguenza che *non* si definisce una operazione di divisione tra matrici.

2.5 Il determinante di una matrice quadrata

Consideriamo una matrice quadrata $A_{n\times n}$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

e scegliamo a caso un suo elemento. Se sopprimiamo la riga e la colonna che "passano per quell'elemento", otterremo una nuova matrice $B_{n-1\times n-1}$. Per esempio sia

$$A = \begin{pmatrix} 1 & 3 & 2 & -1 \\ 4 & 3 & 6 & 1 \\ -2 & 1 & 5 & 8 \\ 7 & -2 & 1 & 4 \end{pmatrix},$$

e consideriamo l'elemento $a_{23} = 6$. Se sopprimiamo la riga e la colonna che passano per quell'elemento, otterremo la nuova matrice B, a tre righe e tre colonne:

$$B = \begin{pmatrix} 1 & 3 & -1 \\ -2 & 1 & 8 \\ 7 & -2 & 4 \end{pmatrix}.$$

Questa matrice si chiama *matrice complementare* dell'elemento a_{23} e si indica con il simbolo A_{23}. In generale, se indichiamo con a_{ij} l'elemento, la matrice complementare si indicherà con A_{ij}.

Introdurremo ora un numero di fondamentale importanza, associato ad ogni matrice quadrata, e precisamente il *determinante* della matrice, indicato con

$$\det(A), \quad \text{oppure} \quad |A| .$$

Come al solito preferiamo una definizione operativa ad una formale, e senza nemmeno preoccuparci di dare la più corta possibile, ma la più immediatamente utilizzabile. Partiamo allora da una matrice $A_{1\times 1}$, cioè un numero reale: se $A = a$, porremo, per definizione,

$$\det(A) = a ,$$

cioè il determinante di una matrice 1×1 è il numero stesso (qui è meglio non usare il secondo simbolo per il determinante perché si può fare confusione con il simbolo di valore assoluto).

Consideriamo poi una matrice 2×2:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

In questo caso porremo, sempre per definizione,

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} .$$

Passiamo poi a una matrice 3×3:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Scegliamo una riga o una colonna qualsiasi e scriviamo, per ciascuno degli elementi di quella riga o colonna, la matrice complementare: si tratta di 3 matrici 2×2, di cui già sappiamo calcolare il determinante. Supponendo di avere scelto, per esempio, la 3 riga, il determinante della matrice $A_{3\times 3}$ sarà, per definizione,

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= (-1)^{3+1} a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} + (-1)^{3+2} a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + (-1)^{3+3} a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} .$$

Per una matrice 4×4 ripetiamo la strategia: scegliamo una riga o una colonna qualsiasi, calcoliamo le matrici complementari di ciascuno dei 4 elementi, ottenendo 4 matrici di ordine 3, di cui sappiamo calcolare il determinante e procediamo come abbiamo fatto con la matrice 3×3. Supponendo, per esempio, di avere scelto la seconda colonna, procederemo come mostra la formula seguente.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = (-1)^{1+2} a_{12} |A_{12}| + (-1)^{2+2} a_{22} |A_{22}| + (-1)^{3+2} a_{32} |A_{32}| + (-1)^{4+2} a_{42} |A_{42}|$$

Naturalmente dovremmo provare che il risultato non dipende dalla riga o colonna scelta, ma…ci fidiamo! In ogni caso, come al solito, è possibile verificarlo su esempi concreti.

È chiaro che si tratta di un calcolo lungo e noioso, anche se esistono numerose strategie per velocizzarlo. Esula dagli scopi di questo corso lo studio di queste strategie: segnaliamo solo che tutti i software di calcolo simbolico (compreso Geogebra) possono eseguire facilmente questo calcolo, che del resto è molto elementare, richiede solo di fare somme e prodotti di numeri (anche se in quantità industriale!).

Esempio 2.14. Nella matrice che segue, per trovare il determinante, scegliamo la prima riga, ottenendo:

$$\begin{vmatrix} 1 & 3 & -1 \\ -2 & 1 & -3 \\ 2 & -2 & 4 \end{vmatrix} =$$

$$= (-1)^{1+1} \cdot 1 \cdot \begin{vmatrix} 1 & -3 \\ -2 & 4 \end{vmatrix} + (-1)^{1+2} \cdot 3 \cdot \begin{vmatrix} -2 & -3 \\ 2 & 4 \end{vmatrix} + (-1)^{1+3} \cdot (-1) \cdot \begin{vmatrix} -2 & 1 \\ 2 & -2 \end{vmatrix} =$$

$$= 1 \cdot 1 \cdot (4-6) - 1 \cdot 3 \cdot (-8+6) + 1 \cdot (-1) \cdot (4-2) = 2 .$$

È un utile esercizio ripetere lo stesso calcolo scegliendo un'altra riga o colonna.

Per completezza segnaliamo che il numero

$$(-1)^{i+j} \det(A_{ij})$$

cioè il determinante della matrice complementare dell'elemento a_{ij}, moltiplicato per $(-1)^{i+j}$, si chiama *complemento algebrico* dell'elemento stesso.

Se A e B sono matrici quadrate di ordine n valgono le seguenti proprietà.

- $\det(A^{\mathrm{T}}) = \det(A)$.
- Se A ha una riga o una colonna di zeri allora $\det(A) = 0$.
- Se A ha due righe o due colonne uguali o proporzionali allora $\det(A) = 0$.
- Se gli elementi di una riga o di una colonna sono moltiplicati per uno scalare k, allora anche $\det(A)$ è moltiplicato per k.
- $\det(kA) = k^n \det(A)$.
- $\det(AB) = \det(A)\det(B)$ (teorema di Binet).

Osserviamo che se A è una matrice quadrata di ordine n, allora le sue righe e le sue colonne sono vettori di $\mathbb{R}^n$. Il seguente teorema fornisce un importante risultato sul legame tra determinante di una matrice di ordine n e dipendenza o indipendenza lineare di un insieme di n vettori.

Teorema 2.11. *Il determinante di una matrice quadrata di ordine n è nullo se e solo se gli n vettori riga o gli n vettori colonna sono linearmente dipendenti.*

Esempio 2.15. Si considerino i seguenti quattro vettori $v_1, v_2, v_3, v_4 \in \mathbb{R}^4$.

$$\vec{v}_1 = \begin{pmatrix} 4 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} -2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}.$$

Per stabilire se essi sono linearmente dipendenti o indipendenti si considera la matrice quadrata A di ordine $n = 4$ ottenuta prendendo come colonne i vettori dati:

$$A = \begin{pmatrix} 4 & 2 & -2 & 1 \\ -1 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Si calcola il determinante della matrice, sviluppando prima secondo la terza riga della matrice A e poi di nuovo secondo la terza della matrice rimasta:

$$\det(A) = -2 \det \begin{pmatrix} 4 & 2 & -2 \\ -1 & 1 & 1/2 \\ 0 & 1 & 0 \end{pmatrix} = 2 \det \begin{pmatrix} 4 & -2 \\ -1 & 1/2 \end{pmatrix} = 0$$

Per il teorema precedente si può concludere che i vettori sono linearmente dipendenti.

Si noti che questo teorema può essere utilizzato solo per stabilire la dipendenza o indipendenza di n vettori di $\mathbb{R}^n$. Si veda però il successivo teorema 2.13 nella pagina 45.

2.5.1 Il calcolo della matrice inversa

La conoscenza del determinante di una matrice ci permette di stabilire a priori se la matrice stessa è invertibile oppure no. Precisamente si dimostra il seguente teorema.

Teorema 2.12. *Una matrice quadrata è invertibile se e solo se il suo determinante è diverso da zero. Una matrice a determinante nullo si dice anche* singolare, *e il nome è proprio legato al fatto che la matrice in questo caso non è invertibile.*

Il calcolo pratico della matrice inversa (se non singolare!) si può fare con la seguente regola.

1. Si calcola il determinante di A, $|A|$;
2. si scrive la matrice trasposta, di A: A^{T};
3. si sostituisce, in tale matrice trasposta, ciascun elemento con il rispettivo complemento algebrico;
4. si divide la matrice così ottenuta per $|A|$, ottenendo l'inversa della matrice A.

Esempio 2.16. Sia

$$A = \begin{pmatrix} 6 & 8 \\ 7 & 4 \end{pmatrix}.$$

Si ha

$$\det(A) = 6 \cdot 4 - 8 \cdot 7 = -32 \quad , \quad A^{\mathrm{T}} = \begin{pmatrix} 6 & 7 \\ 8 & 4 \end{pmatrix}.$$

Sostituiamo in quest'ultima matrice ogni elemento con il suo complemento algebrico e dividiamo poi per -32:

$$\begin{pmatrix} 4 & -8 \\ -7 & 6 \end{pmatrix} \quad , \qquad A^{-1} = \begin{pmatrix} \frac{4}{-32} & \frac{-8}{-32} \\ \frac{-7}{-32} & \frac{6}{-32} \end{pmatrix} = \begin{pmatrix} -\frac{1}{8} & \frac{1}{4} \\ \frac{7}{32} & -\frac{3}{16} \end{pmatrix}.$$

A questo punto è facile verificare che

$$\begin{pmatrix} 6 & 8 \\ 7 & 4 \end{pmatrix} \cdot \begin{pmatrix} -\frac{1}{8} & \frac{1}{4} \\ \frac{7}{32} & -\frac{3}{16} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

e che

$$\begin{pmatrix} -\frac{1}{8} & \frac{1}{4} \\ \frac{7}{32} & -\frac{3}{16} \end{pmatrix} \cdot \begin{pmatrix} 6 & 8 \\ 7 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

2.6 Il rango di una matrice

Data una matrice qualunque $A_{m\times n}$, dunque anche non quadrata, scegliamo k righe e k colonne qualsiasi: gli elementi comuni a queste k righe e k colonne costituiscono una matrice quadrata, che si chiama *matrice estratta* dalla matrice A; di questa matrice si può calcolare il determinante che chiameremo *minore*, o anche minore estratto, di ordine k.

Per esempio dalla matrice $A_{2\times 3}$

$$A = \begin{pmatrix} 1 & 3 & -1 \\ -2 & 1 & -3 \end{pmatrix},$$

si possono estrarre 6 minori di ordine 1

$$1, 3, -1, -2, 1, -3,$$

e tre minori di ordine 2

$$\begin{vmatrix} 1 & 3 \\ -2 & 1 \end{vmatrix} = 7, \quad \begin{vmatrix} 1 & -1 \\ -2 & -3 \end{vmatrix} = -5, \quad \begin{vmatrix} 3 & -1 \\ 1 & -3 \end{vmatrix} = -8, \quad .$$

Il massimo intero positivo r per cui esiste un minore di ordine r diverso da zero si chiama *rango* della matrice.

Esempio 2.17. Il rango della matrice

$$A = \begin{pmatrix} 1 & 4 & 3 & -1 \\ -2 & -1 & 1 & -3 \\ 2 & 1 & -1 & 4 \end{pmatrix}$$

è 3, perché, per esempio, il minore di ordine 3 ottenuto con la prima, la terza e la quarta colonna vale 7, come è facile provare:

$$\begin{vmatrix} 1 & 3 & -1 \\ -2 & 1 & -3 \\ 2 & -1 & 4 \end{vmatrix} = 7\,.$$

Esempio 2.18. Il rango della matrice

$$A = \begin{pmatrix} 1 & 3 & -1 & 2 \\ 0 & 7 & -5 & 5 \\ -2 & 1 & -3 & 1 \\ 2 & -8 & 8 & -6 \end{pmatrix},$$

è 2, perché l'unico minore di ordine 4 (cioè il determinante della matrice stessa) vale 0 (provarlo per esercizio), tutti i minori di ordine 3 valgono sempre 0 (anche questo è da provare per esercizio), mentre per esempio il minore di ordine 2 ottenuto "intersecando" le prime due righe e le prime due colonne vale (è immediato!) 7.

Il seguente teorema è utile per verificare la dipendenza o indipendenza lineare di m vettori di $\mathbb{R}^n$.

Teorema 2.13. *Data una matrice $A_{m\times n}$, il rango di A è il massimo numero di vettori riga o di vettori colonna linearmente indipendenti.*

Esempio 2.19. Si considerino i due vettori di $\mathbb{R}^3$

$$\vec{v}_1 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Li disponiamo in una matrice a due colonne e tre righe:

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 0 \\ 4 & 1 \end{pmatrix}.$$

Questa matrice ha rango 2 perché

$$\det\begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix} = -3 \neq 0.$$

Dunque le sue due colonne (i due vettori dati) sono linearmente indipendenti. Nulla cambia se i vettori si dispongono in riga anziché in colonna.

Si noti che da questo teorema si può concludere che m vettori di $\mathbb{R}^n$, con $m > n$ sono sempre dipendenti: infatti disponendo i vettori in colonna si ha una matrice con n righe e m colonne che non può avere rango maggiore di n, visto che $m > n$.

2.6.1 Il teorema di Kronecker per il calcolo del rango

Oltre al concetto di minore, per il calcolo del rango di una matrice è utile il concetto di *minore orlato*, che compare nel teorema di Kronecker, detto anche teorema degli “orlati”, di cui parleremo tra poco. Questo teorema può semplificare il calcolo del rango di una matrice.

Se M_p è un minore di ordine p estratto da una matrice A, si dice *orlato* di M_p un minore di ordine M_{p+1}, che contiene il minore M_p.

Teorema 2.14. *Una matrice A ha rango p se e solo se esiste almeno un suo minore M_p di ordine p non nullo e avente tutti gli orlati nulli.*

Se riprendiamo in esame l'esempio 2.18, per determinare il rango possiamo osservare che il minore di ordine 2

$$M_2 = \begin{vmatrix} 1 & 3 \\ 0 & 7 \end{vmatrix}$$

vale 7 e tutti i suoi quattro orlati di ordine 3 sono nulli, dunque il rango è 2. I quattro orlati di ordine 3 sono

$$\begin{vmatrix} 1 & 3 & 2 \\ 0 & 7 & 5 \\ 2 & -8 & -6 \end{vmatrix}, \quad \begin{vmatrix} 1 & 3 & -1 \\ 0 & 7 & -5 \\ -2 & 1 & -3 \end{vmatrix}, \quad \begin{vmatrix} 1 & 3 & 2 \\ 0 & 7 & 5 \\ -2 & 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & 3 & -1 \\ 0 & 7 & -5 \\ 2 & -8 & 8 \end{vmatrix}.$$

Come si vede ci sono decisamente meno calcoli da fare.

Per utilizzare questo teorema si può procedere secondo il seguente schema. Data una matrice A

1. si cerca un minore di ordine 2 che sia diverso da zero: se non ne esistono il rango di A è 1, altrimenti si passa al punto 2;
2. si calcolano i suoi orlati di ordine 3: se sono tutti nulli il rango di A è 2, altrimenti si esamina uno che sia non nullo e si procede con il punto 3;
3. si calcolano i suoi orlati di ordine 4: se sono tutti nulli il rango di A è 3, altrimenti si esamina uno che non sia nullo e si procede con lo stesso schema.

2.7 I sistemi lineari

Si chiama *sistema lineare* un sistema di m equazioni in n incognite, *di primo grado*. Questo tipo di sistemi ha grande importanza in numerose applicazioni, anche economiche. Ci occuperemo, seppure per sommi capi del problema della loro risoluzione.

In generale un sistema del tipo considerato si può scrivere nella forma seguente.

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \right. . \tag{2.1}$$

La matrice costruita con tutti i coefficienti delle incognite si dice matrice dei coefficienti o anche *matrice incompleta*.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{2.2}$$

La matrice a m righe e una colonna, contenente i termini noti si chiama *vettore dei termini noti*.

$$\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \tag{2.3}$$

Si considera poi una matrice a n righe e una colonna, contenente tutte le incognite, detta *vettore delle incognite*.

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{2.4}$$

A questo punto il sistema (2.2) si può scrivere nella seguente forma compatta:

$$A\vec{x} = \vec{b}\,, \tag{2.5}$$

che ha lo stesso identico aspetto delle equazioni di primo grado in una incognita

$$ax = b\,, \tag{2.6}$$

solo che ora gli oggetti A, $\vec{x}$ e $\vec{b}$ non sono più numeri, ma matrici.

Si chiama *soluzione* del sistema (2.2) una vettore numerico a n componenti

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \tag{2.7}$$

tale che sostituendo ordinatamente le sue componenti al posto delle incognite, il sistema risulti soddisfatto.

Si può dimostrare che un sistema di questo tipo può avere solo tre comportamenti rispetto alla sua risolubilità:

1. non avere soluzioni, nel qual caso si dice *incompatibile*;
2. avere una sola soluzione, nel qual caso si dice *determinato*;

3. avere infinite soluzioni, nel qual caso si dice *indeterminato*.

Un sistema che abbia soluzioni (una o infinite) si dice *compatibile*.

Per valutare la risolubilità del sistema occorre considerare, oltre alla matrice incompleta A, anche la *matrice completa*, che si indica con $A|b$ e si ottiene aggiungendo, a destra, alle colonne di A la colonna dei termini noti.

$$A|b = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right). \tag{2.8}$$

La separazione dell'ultima colonna con una barra verticale consente di visualizzare immediatamente nella stessa scrittura sia la matrice incompleta che quella completa. In sostanza parlando di sistemi lineari, ci si può limitare a scrivere direttamente la matrice completa, nel modo indicato: da questa scrittura, volendo, è immediato ricavare la scrittura tradizionale con le equazioni e le incognite. Negli esempi che seguono faremo sempre così.

La risolubilità o meno del sistema lineare si può decidere sulla base del seguente teorema.

Teorema 2.15 (Teorema di Rouché-Capelli). *Il sistema lineare* (2.1) *di m equazioni in n incognite ha soluzioni se e solo se la matrice incompleta e quella completa hanno lo stesso rango. Il rango comune delle due matrici, quando il sistema è compatibile, si chiama anche* rango del sistema.

Dimostrazione. La dimostrazione del teorema di Rouché-Capelli è interessante perché, nella sostanza,è una semplice applicazione dei concetti di dipendenza e indipendenza lineare di vettori.

Prima di procedere osserviamo che dire che una sistema ha soluzioni equivale a dire che il vettore $\vec{b}$ dei termini noti è combinazione lineare dei vettori colonna della matrice incompleta A, i coefficienti della combinazione essendo proprio i valori $x_1, x_2, \dots, x_n$ che costituiscono il vettore $\vec{x}$ delle soluzioni.

Indichiamo con $r(A)$ e $r(A|B)$ il rango di A e di $A|b$ rispettivamente. Se $r(A) = r(A|b) = r$ significa che esistono r colonne della matrice A linearmente indipendenti, indichiamole con $\vec{c}_1, \vec{c}_2, \dots, \vec{c}_r$. Ogni altra colonna di A è combinazione lineare di queste r colonne. Anche il rango di $A|b$ è r e quindi il vettore $\vec{b}$ è combinazione lineare delle stesse r colonne di A. A maggior ragione $\vec{b}$ è combinazione lineare di tutte le colonne di A, ma questo significa che il sistema ha soluzioni.

Supponiamo viceversa che il sistema ammetta soluzioni ovvero che il vettore $\vec{b}$ sia combinazione lineare delle colonne di A. Se A ha rango r, $\vec{b}$ deve essere combinazione anche solo delle colonne $\vec{c}_1$, $\vec{c}_2$, ..., $\vec{c}_r$, che sono il massimo numero di colonne linearmente indipendenti. Consideriamo ora il rango di $A|b$. Sicuramente $r(A|b) \geq r(A)$ (perché $(A|b)$ ha più colonne di A). Se fosse $r(A|b) > r(A)$ dovrebbero esistere $r+1$ colonne di $A|b$ linearmente indipendenti e fra queste dovrebbe comparire il vettore $\vec{b}$, ma questo è assurdo perché, come già detto, $\vec{b}$ è combinazione lineare delle colonne $\vec{c}_1$, $\vec{c}_2$, ..., $\vec{c}_r$ di A. Dunque $r(A) = r(A|b)$. □

Esempio 2.20. Il sistema lineare di matrice completa

$$A|b = \left(\begin{array}{cc|c} 3 & 4 & 7 \\ 5 & 6 & 11 \\ 2 & 7 & 4 \end{array}\right)$$

non ha soluzioni, perché la matrice completa ha determinante

$$\det(A|b) = \begin{vmatrix} 3 & 4 & 7 \\ 5 & 6 & 11 \\ 2 & 7 & 4 \end{vmatrix} = 10 \neq 0,$$

e quindi rango 3, mentre quella incompleta potrebbe avere al massimo rango 2, in quanto ha solo due colonne.

Esempio 2.21. Il sistema lineare di matrice completa

$$A|b = \left(\begin{array}{cc|c} 1 & 3 & 2 \\ 2 & 4 & 5 \\ 1 & 5 & 1 \end{array}\right)$$

è compatibile. Infatti si ha $\det(A|b) = 0$, mentre, per esempio,

$$\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = 4 - 6 = -2 \neq 0$$

e quindi le matrici incompleta e, a maggior ragione quella completa, hanno rango 2. Si noti che, ovviamente, il rango della matrice completa non può essere più piccolo di quello della matrice incompleta, in quanto ogni minore della incompleta è amche minore della completa.

Come mostra l'ultimo esempio sopra riportato, per la compatibilità del sistema occorre trovare un minore non nullo di ordine massimo possibile che sia contemporaneamente minore della matrice incompleta e di quella completa. Fatto questo la risoluzione del sistema procede con i passi indicati di seguito (e che saranno meglio chiariti, al solito, con un esempio).

1. Sopprimi tutte le eventuali equazioni corrispondenti alle righe i cui coefficienti non intervengono nel minore trovato.
2. Trasporta a secondo membro tutti gli eventuali termini contenenti le incognite corrispondenti a colonne che non intervengono nel calcolo del minore trovato: queste incognite saranno considerate come parametri, cioè rimarranno completamente arbitrarie.
3. Il sistema residuo è ora un sistema "quadrato", ovvero con lo stesso numero di equazioni e di incognite e con matrice dei coefficienti a determinante non nullo: risolvilo usando il teorema di Cramer che segue.

Se il sistema è compatibile e non c'è nessuna incognita da considerare come parametro, allora esso ha una sola soluzione; se è compatibile e ci sono incognite da considerare come parametro, allora il sistema ha infinite soluzioni e, se i parametri sono k, si usa dire che ha ∞^k soluzioni.

Teorema 2.16 (Teorema di Cramer). *Un sistema quadrato $A\vec{x} = \vec{b}$ di n equazioni in n incognite, di matrice dei coefficienti A, ammette per ogni $\vec{b} \in \mathbb{R}^n$ un'unica soluzione se e solo se la matrice A ha ha determinante* $\det(A) \neq 0$. *In tal caso, considerate le matrici A_i, ottenute sostituendo alla i-esima colonna di A la colonna dei termini noti, l'unica soluzione del sistema è il vettore $\vec{x}$ di componenti*

$$x_1 = \frac{\det(A_1)}{\det(A)}, \quad x_2 = \frac{\det(A_2)}{\det(A)}, \quad \dots, \quad x_n = \frac{\det(A_n)}{\det(A)}.$$

Prima di procedere alla dimostrazione vediamo un esempio.

Esempio 2.22. Il sistema di matrice completa

$$A|b = \left(\begin{array}{ccc|c} 3 & 2 & 2 & 13 \\ 2 & 4 & 3 & 19 \\ 4 & 5 & 2 & 20 \end{array}\right)$$

soddisfa le condizioni poste dal teorema di Cramer. Per le tre matrici A_i si ha:

$$\det(A_1) = \begin{vmatrix} 13 & 2 & 2 \\ 19 & 4 & 3 \\ 20 & 5 & 2 \end{vmatrix} = -17 \ \det(A_2) = \begin{vmatrix} 3 & 13 & 2 \\ 2 & 19 & 3 \\ 4 & 20 & 2 \end{vmatrix} = -34 \ \det(A_2) = \begin{vmatrix} 3 & 2 & 13 \\ 2 & 4 & 19 \\ 4 & 5 & 20 \end{vmatrix} = -51 \,.$$

La soluzione del sistema è allora data da

$$x_1 = \frac{-17}{-17} = 1, \quad x_2 = \frac{-34}{-17} = 2, \quad x_1 = \frac{-51}{-17} = 3\,,$$

cioè dal vettore colonna

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Si noti che l'applicazione del teorema di Rouché-Capelli porta sempre a concludere che il sistema è compatibile e ha una sola soluzione perché il rango di A è 3 e il rango di $A|b$ deve necessariamente essere ancora 3 (ci sono solo tre righe).

Dimostrazione. Supponiamo che la matrice A abbia determinante $|A| \neq 0$: allora esiste la matrice inversa A^{-1} e la soluzione del sistema $\vec{x}$ è data da

$$\vec{x} = A^{-1}\vec{b}.$$

Infatti se si premoltiplica $A\vec{x} = \vec{b}$ per A^{-1} si ottiene

$$A^{-1}A\vec{x} = A^{-1}\vec{b},$$

da cui, sviluppando il primo membro, utilizzando la proprietà associativa dell'operazione di moltiplicazione fra matrici ed il fatto che $I_n\vec{x} = \vec{x}$, si ha

$$A^{-1}(A\vec{x}) = (A^{-1}A)\vec{x} = I_n\vec{x} = \vec{x}$$

cioè

$$\vec{x} = A^{-1}\vec{b}.$$

Quindi, se la matrice dei coefficienti A è invertibile, il sistema ammette un'unica soluzione e questa è data da $\vec{x} = A^{-1}\vec{b}$.

Poiché

$$A^{-1}=\frac{1}{|A|}\begin{pmatrix}A_{11} & A_{21} & \dots & A_{n1}\\ A_{12} & A_{22} & \dots & A_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ A_{1n} & A_{2n} & \dots & A_{nn}\end{pmatrix}$$

segue che

$$\vec{x}=\frac{1}{|A|}\begin{pmatrix}b_1A_{11}+b_2A_{21}+\ldots+b_nA_{n1}\\ b_1A_{12}+b_2A_{22}+\ldots+b_nA_{n2}\\ \vdots\\ b_1A_{1n}+b_2A_{2n}+\ldots+b_nA_{nn}\end{pmatrix}$$

cioè

$$x_i=\frac{1}{|A|}\left(b_1A_{1i}+b_2A_{2i}+\ldots+b_nA_{ni}\right).$$

L'espressione tra parentesi rappresenta il determinante della matrice

$$\begin{pmatrix}a_{11} & a_{12} & \dots & b_1 & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & b_2 & \dots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \dots & b_n & \dots & a_{nn}\end{pmatrix}$$

ottenuta da A sostituendo il vettore $\vec{b}$ alla colonna i-esima.

Viceversa, se il sistema $A\vec{x}=\vec{b}$ ammette una ed una sola soluzione, qualunque sia il vettore $\vec{b}$, nel caso particolare in cui $\vec{b}=\vec{0}$, risolvere il sistema significa determinare quel vettore $\vec{x}$ tale che

$$A\vec{x}=\vec{0}.$$

Per ipotesi la soluzione $\vec{x}$ è unica: quindi $\vec{x}=\vec{0}$. In altre parole, $\vec{0}$ è l'unico vettore $\vec{x}$ tale che

$$x_1\vec{c}_1+x_2\vec{c}_2+\cdots+x_n\vec{c}_n=\vec{0}.$$

Ma allora le colonne di A sono linearmente indipendenti e quindi, ricordando una delle proprietà del determinante, $|A|\neq 0$. □

Un sistema di equazioni lineari si dice *sistema di Cramer* se la matrice dei coefficienti A è una matrice quadrata di ordine n (il numero di equazioni è uguale al numero n delle incognite) con determinante diverso da zero (quindi A invertibile).

Esempio 2.23. Il sistema lineare di matrice completa

$$A|b=\left(\begin{array}{ccc|c}3 & 2 & -1 & 1\\ 4 & 3 & -1 & 2\\ -2 & -1 & 1 & 0\end{array}\right)$$

ha rango 2. Infatti i quattro minori della matrice completa sono tutti nulli (compreso quindi quello costituito solo dalla matrice incompleta)

$$\begin{vmatrix} 3 & 2 & -1 \\ 4 & 3 & -1 \\ -2 & -1 & 1 \end{vmatrix} = 0, \begin{vmatrix} 3 & 2 & 1 \\ 4 & 3 & 2 \\ -2 & -1 & 0 \end{vmatrix} = 0 \begin{vmatrix} 3 & -1 & 1 \\ 4 & -1 & 2 \\ -2 & 1 & 0 \end{vmatrix} = 0 \begin{vmatrix} 2 & -1 & 1 \\ 3 & -1 & 2 \\ -1 & 1 & 0 \end{vmatrix} = 0,$$

mentre la matrice estratta prendendo le prime due righe e le prime due colonne (sottomatrice comune della completa e della incompleta), ha determinante (minore) $9-8=1\neq 0$.

Trascuriamo quindi la terza equazione e portiamo i termini contenenti la terza incognita a secondo membro.

$$\left(\begin{array}{cc|c} 3 & 2 & 1+x_3 \\ 4 & 3 & 2+x_3 \end{array}\right)$$

Questo sistema si risolve con Cramer ottenendo

$$x_1 = \frac{\begin{vmatrix} 1+x_3 & 2 \\ 2+x_3 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 2 \\ 4 & 3 \end{vmatrix}} = x_3 - 1, \qquad x_2 = \frac{\begin{vmatrix} 3 & 1+x_3 \\ 4 & 2+x_3 \end{vmatrix}}{\begin{vmatrix} 3 & 2 \\ 4 & 3 \end{vmatrix}} 2 - x_3,$$

mentre x_3 rimane completamente arbitrario (parametro) e possiamo indicarlo con t. I vettori soluzione sono allora

$$\begin{pmatrix} t-1 \\ 2-t \\ t \end{pmatrix},$$

che sono infinite, al variare di t. Poiché è rimasto un solo parametro, il sistema ha ∞^1 soluzioni.

Esempio 2.24. Una industria fabbrica tre prodotti, P_1, P_2, P_3. Ciascuno di essi subisce un ciclo lavorativo in tre diversi reparti, A, B, C, con i tempi seguenti, in ore:

	A	B	C
P_1	2	1	1
P_2	5	3	2
P_3	3	2	2

Se in una lavorazione il carico orario dei 3 reparti è stato di 104 ore, 64 ore e 55 ore rispettivamente per i reparti A, B e C, si chiede quante sono le quantità fabbricate di ciascun prodotto.

Dette x_1, x_2, x_3 queste quantità si deve avere:

$$\begin{cases} 2x_1 + 5x_2 + 3x * 3 = 104 \\ x_1 + 3x_2 + 2x_3 = 64 \\ x_1 + 2x_2 + 2x_3 = 55 \end{cases}.$$

Si tratta dunque di risolvere un sistema lineare di matrice completa

$$\left(\begin{array}{ccc|c} 2 & 5 & 3 & 104 \\ 1 & 3 & 2 & 64 \\ 1 & 2 & 2 & 55 \end{array}\right).$$

Poiché il determinante della matrice incompleta vale 1, e quindi è diverso da zero, il sistema ha rango 3 e quindi ha una sola soluzione che si può trovare facilmente con la regola di Cramer, ottenendo

$$x_1 = 7, \quad x_2 = 9, \quad x_3 = 15.$$

È però utile la seguente osservazione. Posto

$$A = \begin{pmatrix} 2 & 5 & 3 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 104 \\ 64 \\ 55 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

il sistema si può scrivere in forma compatta

$$A\vec{x} = \vec{b}.$$

La matrice A ha determinante diverso da zero e quindi è invertibile, con

$$A^{-1} = \begin{pmatrix} 2 & -4 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Dalla precedente uguaglianza si può dunque ricavare

$$\vec{x} = A^{-1}\vec{b} = \begin{pmatrix} 2 & -4 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 104 \\ 64 \\ 55 \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \\ 15 \end{pmatrix},$$

in completo accordo con il risultato prima trovato.

In sostanza la risoluzione di un sistema quadrato compatibile, con determinante della matrice incompleta, non nullo si può fare, formalmente, esattamente come per un'equazione di primo grado in un'incognita:

$$ax = b, \quad a \neq 0, \quad \Rightarrow \quad x = a^{-1}b.$$

2.8 Il modello input-output di Leontief

Per comprendere e governare l'economia di un paese si deve fornire un modello basato sui diversi settori di questa economia. Wassily Leontief propose, ancora negli anni 30 del secolo scorso, un modello input-output che porta il suo nome. A causa delle sue numerose ricerche in campo economico, al professor Leontief fu attribuito, nel 1973, il premio Nobel. Faremo un breve cenno al modello di Leontief, che usa estesamente i sistemi lineari (anche molto grandi con centinaia di equazioni e di incognite: lo stesso Leontief sviluppò un modello dell'economia degli Stati Uniti con ben 500 settori!).

Consideriamo dunque un sistema economico composto da n settori, o industrie, interdipendenti: $S_1, S_2, \dots, S_n$. Ciò significa che ogni settore consuma prodotti dei diversi settori, compreso il suo (per esempio una fabbrica che produce automobili può "consumare" automobili da essa stessa prodotte per esempio per motivi di trasporto).

Se non ci sono scambi commerciali con l'esterno dei settori interessati il sistema economico si dice *chiuso*, altrimenti si dice *aperto* (ed è quest'ultima l'opzione più comune).

Indichiamo con x_i il livello di produzione, cioè la quantità di bene prodotta, del settore S_i e con x_{ij} la quantità del bene prodotta da S_i e consumata dal settore S_j. Indichiamo infine con d_i la domanda esterna del bene prodotto dal settore S_i, con $d_i = 0$ se il sistema è chiuso.

Perché l'economia sia equilibrata si deve avere

$$x_i = x_{i1} + x_{i2} + \cdots + x_{in} + d_i, \quad i = 1, 2, \ldots, n. \tag{2.9}$$

Si fa normalmente l'ipotesi, del tutto plausibile, che il settore S_j consumi una quantità del bene prodotto dal settore S_i proporzionale al volume x_j della sua propria produzione. Se indichiamo con m_{ij} la costante di proporzionalità si deve dunque avere

$$x_{ij} = m_{ij} x_j. \tag{2.10}$$

Il coefficiente m_{ij} è detto *coefficiente tecnico di produzione* e, in base a (2.10), rappresenta la quantità del bene prodotta da S_i necessaria per produrre una unità del bene prodotto da S_j.

Con questa ipotesi le n condizioni (2.9) danno luogo al seguente sistema lineare.

$$\left\{ \begin{array}{ccccccccccc} x_1 & = & m_{11}x_1 & + & m_{12}x_2 & + & \ldots & + & m_{1n}x_n & + & d_1 \\ x_2 & = & m_{21}x_1 & + & m_{22}x_2 & + & \ldots & + & m_{2n}x_n & + & d_2 \\ \ldots & = & \ldots & + & \ldots & + & \ldots & + & \ldots & + & \ldots \\ x_n & = & m_{n1}x_1 & + & m_{n2}x_2 & + & \ldots & + & m_{nn}x_n & + & d_n \end{array} \right. . \tag{2.11}$$

Indicata con A la matrice contenente tutti gli m_{ij}, detta *matrice di input-output* del sistema, con $\vec{d}$ il vettore dei termini noti e, al solito, con $\vec{x}$ il vettore delle incognite, il sistema (2.11) si può scrivere in forma compatta

$$\vec{x} = A\vec{x} + \vec{d} \quad \Rightarrow \quad (I - A)\vec{x} = \vec{d}, \tag{2.12}$$

ove I è la matrice unità. Si tratta dunque di un sistema lineare di n equazioni in n incognite, con matrice dei coefficienti $I - A$. Se il sistema economico è chiuso $\vec{d} = 0$ e il sistema è omogeneo, se è aperto $\vec{d} \neq 0$ e il sistema non è omogeneo.

Proponiamo di seguito due esempi concreti, uno di sistema chiuso e uno di sistema aperto, con pochi settori interdipendenti.

Esempio 2.25 (Un sistema chiuso di Leontief). Si consideri un sistema economico basato su tre settori, per esempio carbone, energia elettrica e acciaio. Ciascun settore ha bisogno degli altri per funzionare e consuma una parte del suo stesso prodotto. La tabella seguente fornisce sulle colonne l'output totale di ciascun settore (riferito a una unità di bene prodotto) e sulle righe l'input di ciascun settore per ogni bene (abbiamo supposto, a titolo d'esempio, che il settore carbone non utilizzi carbone per funzionare).

	Carbone	Elettricità	Acciaio
Carbone	0.0	0.4	0.6
Elettricità	0.6	0.1	0.2
Acciaio	0.4	0.5	0.2

Si noti come la somma di ogni riga e di ogni colonna è 1.

Il sistema (2.12) diventa, tenendo conto che $d_i = 0$,

$$\begin{cases} x_1 - 0.4x_2 - 0.6x_3 = 0 \\ -0.6x_1 + 0.9x_2 - 0.2x_3 = 0 \\ -0.4x_1 - 0.5x_2 + 0.8x_3 = 0 \end{cases}.$$

Il determinante della matrice incompleta è 0 e si trova subito che il suo rango è 2, per esempio considerando le prime due righe e le prime due colonne; poiché in questo caso di sistema omogeneo la matrice completa e quella incompleta hanno necessariamente lo stesso rango, il sistema ha rango 2 e quindi ha infinite soluzioni: possiamo lasciare x_3 arbitraria, chiamandola per esempio t. Si ottiene

$$\vec{x} = \begin{pmatrix} 0.939394\,t \\ 0.848485\,t \\ t \end{pmatrix}.$$

A ogni valore non negativo di $t = x_3$ corrisponde un valore per x_1 e x_2.

Esempio 2.26 (Un sistema aperto di Leontief). Riconsideriamo il sistema (2.12):

$$(I - A)\vec{x} = \vec{d}.$$

Se la matrice $I - A$ è invertibile (cosa che non è sempre vera), questo sistema può essere risolto come visto nell'esempio 2.24:

$$\vec{x} = (I - a)^{-1}\vec{d}.$$

Se inoltre la matrice $(I - A)^{-1}$ ha tutti i suoi elementi non negativi, allora le componenti del vettore $\vec{x}$ sono non negative e sono dunque accettabili per questo modello. In questo caso si dice che la matrice A è *vitale* o *produttiva*.

Consideriamo un sistema economico composto da tre settori: un settore, S_1, per l'estrazione e raffinazione di petrolio, un settore, S_2, di produzione di elettricità e uno, S_3, di fabbricazione di automobili. La matrice di input-output sia la seguente

$$A = \begin{pmatrix} 0.1 & 0.25 & 0.2 \\ 0.3 & 0.4 & 0.5 \\ 0.1 & 0.15 & 0.1 \end{pmatrix},$$

mentre il vettore della domanda esterna sia dato da

$$\begin{pmatrix} 5000 \\ 7500 \\ 12500 \end{pmatrix}.$$

Si trova, con metodi ormai usuali,

$$I - A = \begin{pmatrix} 0.9 & -0.25 & -0.2 \\ -0.3 & 0.6 & -0.5 \\ -0.1 & -0.15 & 0.9 \end{pmatrix}, \quad \text{e} \quad (I - A)^{-1} = \begin{pmatrix} 1.46457 & 0.80315 & 0.771654 \\ 1.00787 & 2.48819 & 1.6063 \\ 0.330709 & 0.503937 & 1.46457 \end{pmatrix}.$$

È ora facile trovare la soluzione del sistema

$$\vec{x} = (I - A)^{-1}\vec{d} = \begin{pmatrix} 22992.1 \\ 43779.5 \\ 23740.2 \end{pmatrix}.$$

3 Limiti e continuità per funzioni di una variabile

3.1 Considerazioni introduttive

Prima di iniziare la vera e propria trattazione dell'importante concetto di limite per le funzioni reali, consideriamo alcuni esempi per capire il senso delle definizioni formali che daremo, segnalando che, vista la natura e gli scopi di questo corso privilegeremo sempre gli aspetti più propriamente "applicativi", naturalmente non rinunciando al rigore necessario.

Esempio 3.1. Consideriamo la funzione definita da

$$f(x) = \frac{\sin x}{x}$$

il cui dominio naturale è $\mathbb{R} \setminus \{0\}$, cioè l'insieme dei reali non nulli, e poniamoci il seguente problema, che avrà notevole importanza nel seguito: *considerato che non è possibile calcolare il valore di f in corrispondenza di $x = 0$, perché 0 non appartiene al dominio della funzione, è almeno possibile valutare il comportamento di f se la x, a partire da un valore* "prossimo *a*" 0, *si avvicina* "quanto più è possibile" *a* 0[(1)]?

Per rispondere a questa domanda usiamo un foglio di calcolo e calcoliamo i valori di x (in radianti), $sin(x)$ e $\sin x / x$, a partire, per esempio, da $x = 1$ fino ad avvicinarci a $x = 0$. Abbiamo riportato i risultati di questo calcolo nella tabella 3.1.

Come si può notare, il rapporto presente nell'ultima colonna si avvicina sempre più a 1 (si noti, nelle ultime 3 righe, che i valori di x e di $\sin x$ nella tabella sono identici, ma questo fatto è una conseguenza degli errori di arrotondamento del foglio di calcolo: in realtà $\sin x$ è sempre minore di x, come risulta evidente dalla definizione che ne abbiamo dato).

Una ulteriore verifica di questo fatto può essere ottenuta se tracciamo il grafico della funzione f in questione: si noti che, in corrispondenza del punto $(0, 1)$, non viene tracciato alcun punto sul grafico.

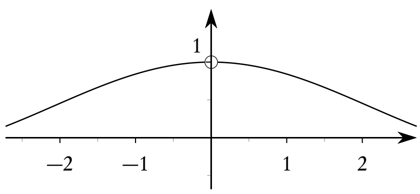


Figura 3.1 *Grafico della funzione* $\sin x / x$, *nei pressi di* 0

[1] Abbiamo usato le virgolette su "*prossimo*" e su "*quanto più è possibile*", perché sono termini che dovranno essere precisati, in quanto non hanno un significato univoco

x	$\sin x$	$\sin x/x$
1,000000000000000	0,841470984807896	0,841470984807896
0,500000000000000	0,479425538604203	0,958851077208406
0,400000000000000	0,389418342308650	0,973545855771626
0,300000000000000	0,295520206661340	0,985067355537799
0,200000000000000	0,198669330795061	0,993346653975306
0,100000000000000	0,099833416646828	0,998334166468282
0,050000000000000	0,049979169270678	0,999583385413567
0,040000000000000	0,039989334186634	0,999733354665854
0,030000000000000	0,029995500202496	0,999850006749855
0,020000000000000	0,019998666693333	0,999933334666654
0,010000000000000	0,009999833334167	0,999983333416667
0,005000000000000	0,004999979166693	0,999995833338542
0,004000000000000	0,003999989333342	0,999997333335467
0,003000000000000	0,002999995500002	0,999998500000675
0,002000000000000	0,001999998666667	0,999999333333466
0,001000000000000	0,000999999833333	0,999999833333342
0,000500000000000	0,000499999979167	0,999999958333334
0,000100000000000	0,000099999999833	0,999999998333333
0,000050000000000	0,000049999999979	0,999999999583333
0,000010000000000	0,000010000000000	0,999999999983333
0,000005000000000	0,000005000000000	0,999999999995833
0,000001000000000	0,000001000000000	0,999999999999833

Tabella 3.1 *Valori di x, $\sin x$, $\sin x/x$, per x variabile da 1 a "quasi 0"*

Potremo riassumere questo comportamento dicendo che, per x che "tende" a 0, $\sin x/x$ "tende' a 1, in formule

$$x \to 0 \Rightarrow \frac{\sin x}{x} \to 1,$$

o, come scriveremo di solito,

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

Esempio 3.2. Come secondo esempio consideriamo la funzione

$$f(x) = \frac{\sqrt{x}-1}{x-1},$$

il cui dominio naturale è $[0,1[\cup]1,+\infty[$. Anche qui, usando un foglio di calcolo, determiniamo (se c'è!!) il valore a cui tende questo rapporto, al tendere di x a 1 (valore non compreso nel dominio), a partire, per esempio, da $x=2$. Otteniamo la tabella 3.2.

Anche in questo caso possiamo usare un grafico per un ulteriore controllo numerico del risultato.

x	$\sqrt{x}-1$	$x-1$	$(\sqrt{x}-1)/(x-1)$
2,000000000000000	0,414213562373095	1,000000000000000	0,414213562373095
1,500000000000000	0,224744871391589	0,500000000000000	0,449489742783178
1,400000000000000	0,183215956619923	0,400000000000000	0,458039891549808
1,300000000000000	0,140175425099138	0,300000000000000	0,467251416997127
1,200000000000000	0,095445115010332	0,200000000000000	0,477225575051661
1,100000000000000	0,048808848170152	0,100000000000000	0,488088481701516
1,050000000000000	0,024695076595960	0,050000000000000	0,493901531919198
1,040000000000000	0,019803902718557	0,040000000000000	0,495097567963925
1,030000000000000	0,014889156509222	0,030000000000000	0,496305216974065
1,020000000000000	0,009950493836208	0,020000000000000	0,497524691810391
1,010000000000000	0,004987562112089	0,010000000000000	0,498756211208895
1,005000000000000	0,002496882788171	0,005000000000000	0,499376557634223
1,004000000000000	0,001998003990028	0,004000000000000	0,499500997506952
1,003000000000000	0,001498876684342	0,003000000000000	0,499625561447503
1,002000000000000	0,000999500499376	0,002000000000000	0,499750249687958
1,001000000000000	0,000499875062461	0,001000000000000	0,499875062461019
1,000500000000000	0,000249968757810	0,000500000000000	0,499937515620330
1,000100000000000	0,000049998750062	0,000100000000000	0,499987500624021
1,000050000000000	0,000024999687508	0,000050000000000	0,499993750157135
1,000010000000000	0,000004999987500	0,000010000000000	0,499998749999897
1,000005000000000	0,000002499996875	0,000005000000000	0,499999375011051
1,000001000000000	0,000000499999875	0,000001000000000	0,499999875099910

Tabella 3.2 *Valori di x, $\sqrt{x}-1$, $x-1$, $(\sqrt{x}-1)/(x-1)x$, per x variabile da 2 a "quasi 1"*

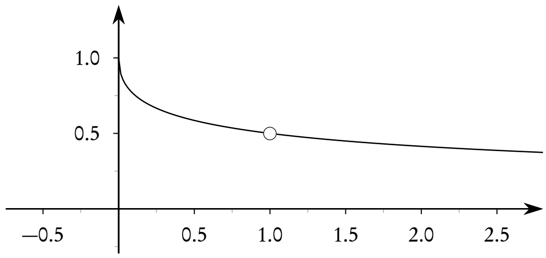


Figura 3.2 *Grafico della funzione $f(x)=(\sqrt{x}-1)/(x-1)$*

Come prima, notiamo che in corrispondenza di $x=1$ non viene tracciato alcun punto sul grafico della funzione.

Potremo scrivere ora:

$$\lim_{x\to 1}\frac{\sqrt{x}-1}{x-1}=\frac{1}{2},$$

naturalmente riservandoci di precisare meglio il senso della scrittura.

Esempio 3.3. Il terzo esempio è relativo alla funzione

$$f(x)=\frac{\sin x}{x^2},$$

per la quale cerchiamo di capire, sempre usando il solito foglio di calcolo, che cosa succede quando x si avvicina a `0` (punto non compreso nel dominio naturale). Otteniamo la tabella 3.3.

x	$\sin x$	x^2	$\sin x/x^2$
1,000000000000000	0,841470984807896	1,000000000000000	0,841470984807896
0,500000000000000	0,479425538604203	0,250000000000000	1,917702154416810
0,400000000000000	0,389418342308650	0,160000000000000	2,433864639429070
0,300000000000000	0,295520206661340	0,090000000000000	3,283557851792660
0,200000000000000	0,198669330795061	0,040000000000000	4,966733269876530
0,100000000000000	0,099833416646828	0,010000000000000	9,983341664682810
0,050000000000000	0,049979169270678	0,002500000000000	19,991667708271300
0,040000000000000	0,039989334186634	0,001600000000000	24,993333866646300
0,030000000000000	0,029995500202496	0,000900000000000	33,328333558328500
0,020000000000000	0,019998666693333	0,000400000000000	49,996666733332700
0,010000000000000	0,009999833334167	0,000100000000000	99,998333341666600
0,005000000000000	0,004999979166693	0,000025000000000	199,999166667708000
0,004000000000000	0,003999989333342	0,000016000000000	249,999333333867000
0,003000000000000	0,002999995500002	0,000009000000000	333,332833333558000
0,002000000000000	0,001999998666667	0,000004000000000	499,999666666733000
0,001000000000000	0,000999999833333	0,000001000000000	999,999833333342000
0,000500000000000	0,000499999979167	0,000000250000000	1999,999916666670000
0,000100000000000	0,000099999999833	0,000000010000000	9999,999983333340000
0,000050000000000	0,000049999999979	0,000000002500000	19999,999991666700000
0,000010000000000	0,000010000000000	0,000000000100000	99999,999998333300000
0,000005000000000	0,000005000000000	0,000000000025000	199999,999999167000000
0,000001000000000	0,000001000000000	0,000000000001000	999999,999999833000000

Tabella 3.3 *Valori di x, $\sin x$, x^2, $\sin x/x^2$, per x variabile da 1 a “quasi 0”*

Questa volta notiamo che, quanto più x è prossimo a 0, tanto più i valori della funzione crescono. Riservandoci una precisazione ulteriore, possiamo dire che se x tende a 0, $f(x)$ diventa enormemente grande, ovvero “tende a $+\infty$”, in formule

$$\lim_{x\to 0} \frac{\sin x}{x^2} = +\infty\,.$$

Il solito grafico conferma l'impressione numerica.

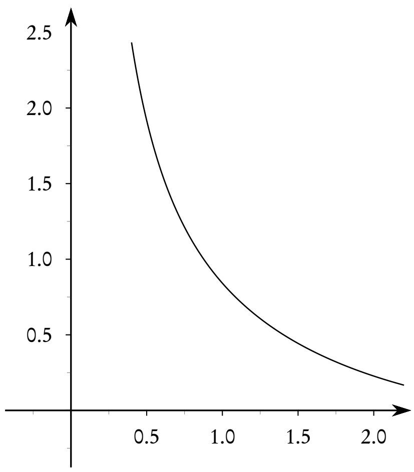


Figura 3.3 *Grafico della funzione $\sin x/x^2$*

Esempio 3.4. Concludiamo questa carrellata di numeri e grafici con la funzione

$$f(x) = \frac{\sin x}{\sqrt{x}},$$

per la quale siamo ancora interessati a scoprire che cosa succede quando x si avvicina a 0 (punto non compreso nel dominio naturale). La relativa tabella è la 3.4.

x	$\sin x$	$\sqrt{x}$	$\sin x / \sqrt{x}$
1,000000000000000	0,841470984807896	1,000000000000000	0,841470984807896
0,500000000000000	0,479425538604203	0,707106781186548	0,678010098842090
0,400000000000000	0,389418342308650	0,632455532033676	0,615724462171224
0,300000000000000	0,295520206661340	0,547722557505166	0,539543611290014
0,200000000000000	0,198669330795061	0,447213595499958	0,444238128702149
0,100000000000000	0,099833416646828	0,316227766016838	0,315700983200547
0,050000000000000	0,049979169270678	0,223606797749979	0,223513639896411
0,040000000000000	0,039989334186634	0,200000000000000	0,199946670933171
0,030000000000000	0,029995500202496	0,173205080756888	0,173179101163883
0,020000000000000	0,019998666693333	0,141421356237310	0,141411928335454
0,010000000000000	0,009999833334167	0,100000000000000	0,099998333341667
0,005000000000000	0,004999979166693	0,070710678118655	0,070710383491198
0,004000000000000	0,003999989333342	0,063245553203368	0,063245384548694
0,003000000000000	0,002999995500002	0,054772255750517	0,054772173592170
0,002000000000000	0,001999998666667	0,044721359549996	0,044721329735762
0,001000000000000	0,000999999833333	0,031622776601684	0,031622771331221
0,000500000000000	0,000499999979167	0,022360679774998	0,022360678843303
0,000100000000000	0,000099999999833	0,010000000000000	0,009999999983333
0,000050000000000	0,000049999999979	0,007071067811865	0,007071067808919
0,000010000000000	0,000010000000000	0,003162277660168	0,003162277660116
0,000005000000000	0,000005000000000	0,002236067977500	0,002236067977490
0,000001000000000	0,000001000000000	0,001000000000000	0,001000000000000

Tabella 3.4 *Valori di x, $\sin x$, $\sqrt{x}$, $\sin x/\sqrt{x}$, per x variabile da 1 a "quasi 0"*

Questa volta il foglio di calcolo è un po' più impreciso (sempre a causa di problemi di arrotondamento), ma, magari anche facendo uso del solito grafico, ci assumiamo la responsabilità di affermare che

$$\lim_{x \to 0} \frac{\sin x}{\sqrt{x}} = 0.$$

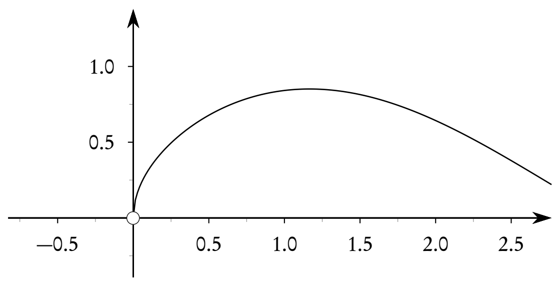


Figura 3.4 *Grafico della funzione $f(x) = \sin x/\sqrt{x}$*

È interessante osservare che le quattro situazioni che abbiamo esaminato negli esempi sono molto simili: in tutti i quattro casi abbiamo considerato rapporti di due quantità, variabili al variare di x, che per x tendente a un opportuno valore si presentavano sempre con l'aspetto di frazioni aventi numeratore e denominatore "*infinitamente piccolo*": eppure il risultato finale è stato completamente diverso nei quattro casi.

Purtroppo non è sempre possibile decidere il comportamento di una funzione in prossimità di un dato punto usando un foglio di calcolo (o simili tecniche) e, nemmeno, usando un grafico. Un esempio classico in tal senso è dato dalla funzione

$$f(x) = \sin \frac{1}{x} .$$

La tabella 3.5, costruita al solito modo, questa volta non ci è di grande aiuto.

x	$1/x$	$\sin 1/x$
1,000000000000000	1,000000000000000	0,841470984807897
0,500000000000000	2,000000000000000	0,909297426825682
0,400000000000000	2,500000000000000	0,598472144103957
0,300000000000000	3,333333333333330	−0,190567962875485
0,200000000000000	5,000000000000000	−0,958924274663138
0,100000000000000	10,000000000000000	−0,544021110889370
0,050000000000000	20,000000000000000	0,912945250727628
0,040000000000000	25,000000000000000	−0,132351750097773
0,030000000000000	33,333333333333300	0,940529576628763
0,020000000000000	50,000000000000000	−0,262374853703929
0,010000000000000	100,000000000000000	−0,506365641109759
0,005000000000000	200,000000000000000	−0,873297297213995
0,004000000000000	250,000000000000000	−0,970528019541805
0,003000000000000	333,333333333333000	0,318846344358746
0,002000000000000	500,000000000000000	−0,467771805322476
0,001000000000000	1000,000000000000000	0,826879540532003
0,000500000000000	2000,000000000000000	0,930039504416137
0,000100000000000	10000,000000000000000	−0,305614388888252
0,000050000000000	20000,000000000000000	0,581984761994295
0,000010000000000	100000,000000000000000	0,035748797986561
0,000005000000000	200000,000000000000000	−0,071451895241553
0,000001000000000	1000000,000000000000000	−0,349993502171308

Tabella 3.5 *Valori di x, $1/x$, $\sin 1/x$, per x variabile da 1 a "quasi 0"*

Né ci conforta molto l'esame del grafico, che abbiamo già visto nella figura 1.9 della pagina 13, e nelle due figure subito successive.

Abbiamo bisogno di una teoria e di tecniche di calcolo più raffinate per risolvere il problema che ha, come già accennato, grande importanza applicativa. Purtroppo le cose non sono molto semplici, e qui ci limiteremo solo agli aspetti essenziali.

3.2 La retta reale estesa

Per velocizzare la trattazione del problema del calcolo dei limiti e per semplificare molte scritture è utile *ampliare* l'insieme dei numeri reali, aggiungendo due ulteriori elementi che chiameremo, anche se in maniera impropria, ancora "punti". Attenzione però: non useremo *mai* per questi due elementi la dicitura "numero", in quanto, come vedremo, il loro comportamento nei confronti delle operazioni elementari è *alquanto strano*.

Definizione 3.1. *Chiameremo* retta reale estesa*, l'insieme*

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

cioè l'insieme dei numeri reali a cui siano stati aggiunti altri due elementi, o punti, detti rispettivamente −infinito e +infinito, per i quali stabiliremo le regole di seguito elencate per quanto riguarda l'ordine e le operazioni fondamentali.

Nella retta reale estesa a volte (ma non sempre!!) potremo attribuire un segno anche allo zero, con delle regole che vedremo in seguito: se saremo interessati a questa scelta indicheremo con 0^+ uno "*zero positivo*", con 0^- uno "*zero negativo*".

Ordine nella retta reale estesa

Per ogni numero reale a, si pone, per definizione,

$$-\infty < a < +\infty,$$

ovvero $-\infty$ precede tutti i i numeri reali (è una specie di "primo elemento"), mentre $+\infty$ segue tutti i numeri reali (è una specie di "ultimo elemento").

Operazioni nella retta reale estesa

Le operazioni elementari in uso tra i numeri reali possono essere estese, entro certi limiti, ad operazioni coinvolgenti anche i nuovi simboli di $\pm\infty$, nel modo indicato qui di seguito. Segnaliamo che, scrivendo ∞, intendiamo riferirci indifferentemente al simbolo $+\infty$ o $-\infty$. Tutte le volte che serve ed è possibile, si deve inoltre applicare la usuale "regola dei segni" per quanto riguarda il prodotto e il quoziente.

1. Per ogni numero reale a, $a \pm (+\infty) = \pm\infty$.
2. Per ogni numero reale a, $a \pm (-\infty) = \mp\infty$.
3. $(+\infty) + (+\infty) = +\infty$.
4. $(-\infty) + (-\infty) = -\infty$.
5. Per ogni numero reale a *diverso* da 0, $a \cdot (\infty) = \infty$ (con la regola dei segni).
6. $(\infty) \cdot (\infty) = \infty$ (con la regola dei segni).
7. Per ogni numero reale a, anche 0, $a/\infty = 0$.
8. Per ogni numero reale a *diverso* da 0, $a/0 = \infty$ (con la regola dei segni, se applicabile).
9. Per ogni numeri reale a (anche 0), $\infty/a = \infty$ (con la regola dei segni, se applicabile).

Osserviamo che *non* abbiamo definito le operazioni nei casi seguenti:

1. Somma di $(+\infty)$ e $(-\infty)$ (e analoghe che si ottengono usando le regole dei segni: diremo brevemente che abbiamo escluso dalle regole di calcolo il caso $\infty - \infty$).

2. Prodotto tra 0 e ∞: diremo brevemente che abbiamo escluso dalle regole di calcolo il caso $0 \cdot \infty$.
3. Quoziente tra 0 e 0: diremo brevemente che abbiamo escluso dalle regole di calcolo il caso $0/0$.
4. Quoziente tra ∞ e ∞: diremo brevemente che abbiamo escluso dalle regole di calcolo il caso ∞/∞.

Chiameremo queste situazioni *forme di indecisione* o anche (ma la nomenclatura ci pare oltremodo brutta e inadeguata) *forme indeterminate*[2].

Intorni dell'infinito

Anche per i due nuovi oggetti aggiunti alla retta reale si introduce il concetto di intorno. Precisamente si dice *intorno di* $+\infty$ un qualunque intervallo aperto superiormente illimitato: $I_{+\infty} =]a, +\infty[$, oppure $]-\infty, +\infty[$; si dice invece *intorno di* $-\infty$ un qualunque intervallo aperto inferiormente illimitato: $I_{+\infty} =]-\infty, a[$, oppure $]-\infty, +\infty[$.

3.3 La definizione di limite

Siamo ora pronti per dare una definizione il più possibile formale e rigorosa del concetto di limite per una funzione reale. Come al solito accompagneremo questa definizione con esempi grafici esplicativi. Segnaliamo subito una difficoltà nella definizione: gli esempi che abbiamo fornito implicavano un *movimento* della x verso un dato valore x_0; ebbene, un tale concetto non è formalizzabile in maniera chiara e univoca, e nella definizione che daremo ogni idea di movimento è sparita. È questo il risultato di una lunga discussione tra i matematici all'inizio dell'analisi, discussione che ha condotto alla definizione formale che segue.

Definizione 3.2. *Sia data una funzione f, di dominio D, e sia x_0 un punto di accumulazione per D (non essendo escluso che x_0 possa essere uno dei due simboli di infinito). Diremo che l (anche qui non essendo escluso che l possa essere uno dei due simboli di infinito) è il* limite *di $f(x)$ per x tendente a x_0, e scriveremo*

$$\lim_{x \to x_0} f(x) = l \tag{3.1}$$

se, scelto un arbitrario intorno I_l di l, è possibile trovare in corrispondenza un opportuno intorno I_{x_0} di x_0, in modo tale che i valori della funzione calcolati in I_{x_0}, tranne *x_0 stesso, cadano in I_l.*

Detto in termini meno formali: vale la formula (3.1) se, considerato un segmento arbitrario contenente punti situati nei pressi di l, è possibile trovare un segmento che contenga il punto x_0, in modo tale che *tutte* le frecce che partono da questi punti cadano solo su punti tra quelli precedentemente scelti, con la clausola che non interessa sapere che cosa succede in corrispondenza del punto x_0.

Quando l è un numero reale non è restrittivo, e di solito lo si fa sempre, limitarsi a considerare solo intorni circolari di l stesso. È inoltre evidente che ci si può limitare a considerare solo "intorni piccoli" di l. Tenendo conto di questo la definizione di limite può anche essere riformulata come segue

[2] Le definizioni che abbiamo dato relative alle operazioni in $\widetilde{\mathbb{R}}$ servono in realtà a calcolare i limiti: vedremo che solo quando si hanno situazioni del tipo delle quattro chiamate forme di indecisione, il calcolo è, in generale, complesso. Per questo sarebbe forse meglio chiamare queste situazioni "*forme difficili*", e non forme di indecisione, ma la tradizione ha il suo peso... In ogni caso è bene tenere presente fin da subito che *non c'è nulla di indeterminato*, solo che in questi casi è complicato decidere che cosa succede.

Definizione 3.3. *Sia data una funzione f, di dominio D, e sia x_0 un punto di accumulazione per D (non essendo escluso che x_0 possa essere uno dei due simboli di infinito). Diremo che $l \in \mathbb{R}$ è il* limite *di $f(x)$ per x tendente a x_0, e scriveremo*

$$\lim_{x \to x_0} f(x) = l \tag{3.2}$$

se, scelto un numero $\varepsilon > 0$ arbitrariamente piccolo, è possibile trovare in corrispondenza un opportuno intorno I_{x_0} di x_0, in modo tale che per i valori della funzione calcolati in I_{x_0}, tranne *x_0 stesso si abbia:*

$$l - \varepsilon < f(x) < l + \varepsilon,$$

disuguaglianza che può anche essere scritta nella forma compatta

$$|f(x) - l| < \varepsilon.$$

Esempio 3.5. Si provi, usando la definizione con l'ε, che

$$\lim_{x \to 1} x + 1 = 2.$$

Si tratta di verificare se la doppia disequazione

$$2 - \varepsilon < x + 1 < 2 + \varepsilon$$

è o no verificata in un intorno di 1. La risoluzione è immediata e si ottiene

$$1 - \varepsilon < x < 1 + \varepsilon,$$

che costituisce un intorno (circolare, ma la cosa non è molto importante) di 1.

Esempio 3.6.

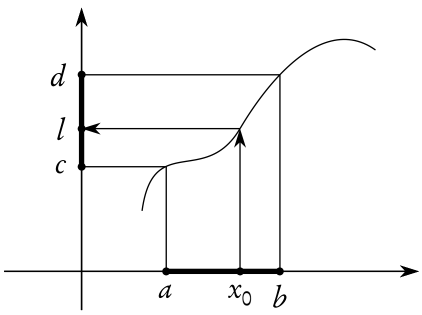


Figura 3.5 *Tutte le frecce che partono dal segmento* $]a, b[$ *cadono nel segmento* $]c, d[$

Esempio 3.7.

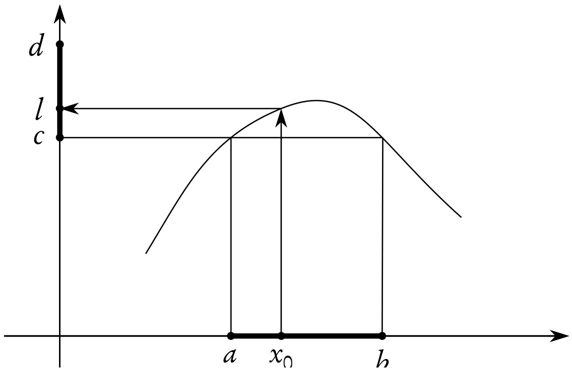


Figura 3.6 *Anche in questo caso le frecce che partono dal segmento* $]a, b[$ *cadono nel segmento* $]c, d[$

Esempio 3.8.

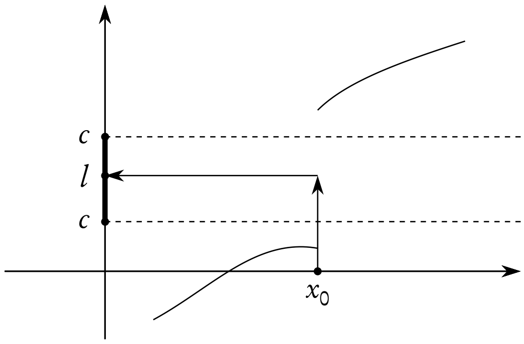


Figura 3.7 *In questo caso non è possibile trovare un segmento, circondante* x_0*, con le proprietà richieste*

Negli esempi relativi alle figure 3.5, 3.6 e 3.7, si è potuto verificare la validità della definizione di limite, usando i grafici delle funzioni. Tutto sarebbe ovviamente più complesso se i grafici non fossero disponibili, e nel seguito faremo qualche cenno sulle tecniche da usare, limitatamente ai casi di nostro interesse.

3.4 Tre teoremi fondamentali sui limiti

Enunciamo tre teoremi fondamentali sui limiti, di cui daremo solo una dimostrazione grafica: non è comunque difficile tradurre in un discorso formale e rigoroso quanto diremo.

Teorema 3.4 (Unicità del limite)**.** *Se una funzione ha un limite* l*, per* x *tendente a* x_0*, tale limite è unico.*

Dimostrazione. Si supponga che ci siano due limiti diversi, l_1 e l_2 e si esamini la figura che segue.

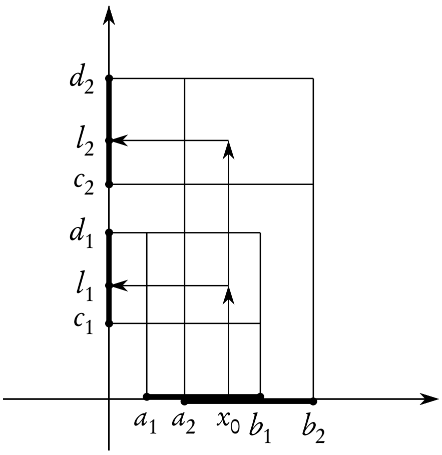


Figura 3.8 *Unicità del limite*

È chiaro che le frecce lanciate dai punti tra a_2 e b_1 dovrebbero cadere *contemporaneamente* in $]c_1, d_1[$ e $]c_2, d_2[$, cosa palesemente impossibile. □

Teorema 3.5 (Permanenza del segno)**.** *Se una funzione ha un limite positivo, per x tendente a x_0, la funzione è positiva in un intorno del punto x_0. Discorso complementare se il limite è negativo.*

Dimostrazione. Supponiamo che il limite l sia positivo, ed esaminiamo la figura che segue, che non ha bisogno di commenti.

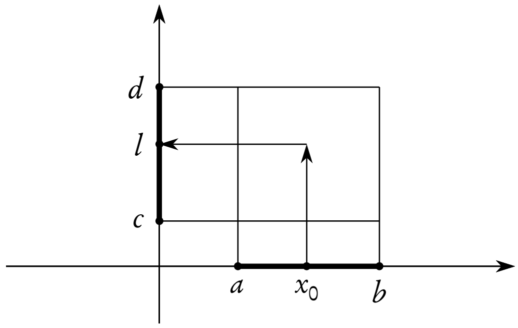


Figura 3.9 *Permanenza del segno*

□

Osserviamo esplicitamente che se

$$\lim_{x \to x_0} f(x) = 0,$$

allora la funzione può essere positiva nei pressi di x_0, e allora diremo che il limite è "zero positivo", o 0^+, negativa nei pressi di x_0, e allora diremo che il limite è "zero negativo", o 0^-, o infine può cambiare di segno nei pressi di x_0 e allora diremo che il limite è 0, senza precisare "positivo" o "negativo". I tre grafici della figura seguente illustrano tre situazioni possibili.

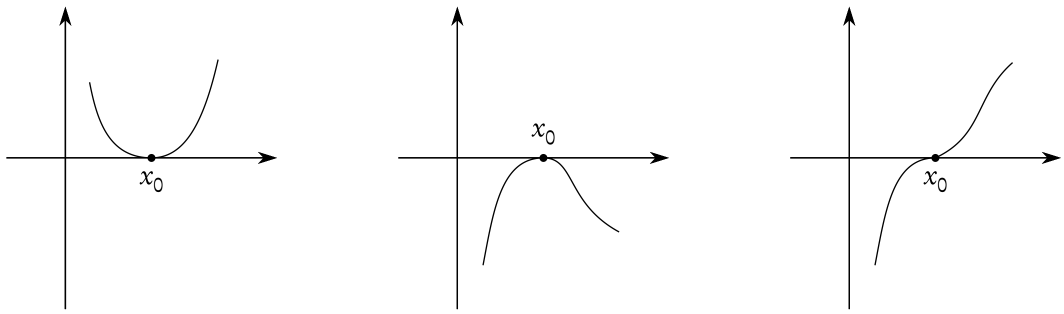


Figura 3.10 0^+, 0^-, 0 *"senza segno"*

Limite destro e limite sinistro

Molto spesso, nei problemi di limite, è utile limitarsi a considerare solo le x del dominio di una funzione che si trovano "a destra" di x_0, oppure "a sinistra", di x_0. Parleremo allora di *limite destro* e *limite sinistro*, e scriveremo

$$\lim_{x\to x_0^+} f(x) = l \quad \text{oppure} \quad \lim_{x\to x_0^-} f(x) = l\,.$$

Se riesaminiamo il terzo grafico della precedente figura 3.10 alla luce di questa definizione, potremo dire che la funzione ha limite 0^- a sinistra e 0^+ a destra. Purtroppo però, quando una funzione ha limite 0 "senza segno", non è affatto detto che il limite sia 0^- a sinistra e 0^+ a destra o viceversa: le situazioni possono essere molto più complesse. Proponiamo solo un esempio grafico, senza commenti, relativo alla funzione

$$f(x) = x \sin\frac{1}{x}\,.$$

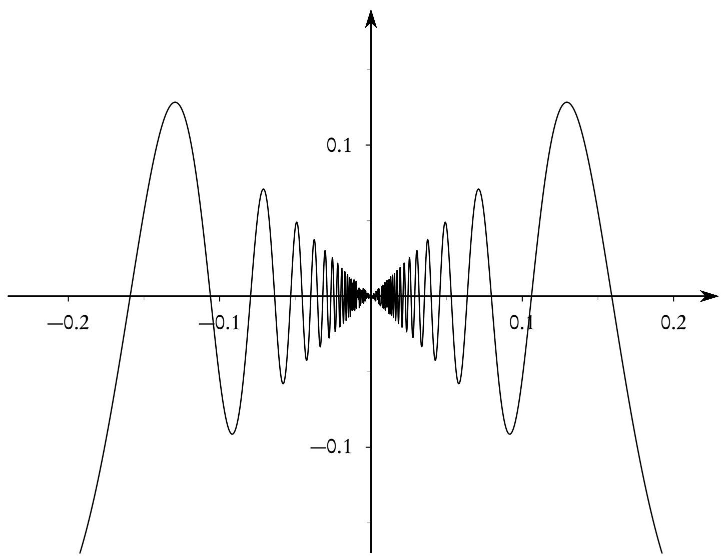


Figura 3.11 *La funzione* $x \sin\frac{1}{x}$

Teorema 3.6 (Del confronto o dei due carabinieri). *Se due funzioni f e g hanno lo stesso limite l per x tendente a x_0, anche una funzione h che sia compresa tra le due ha lo stesso limite.*

Dimostrazione. È sufficiente esaminare la figura che segue.

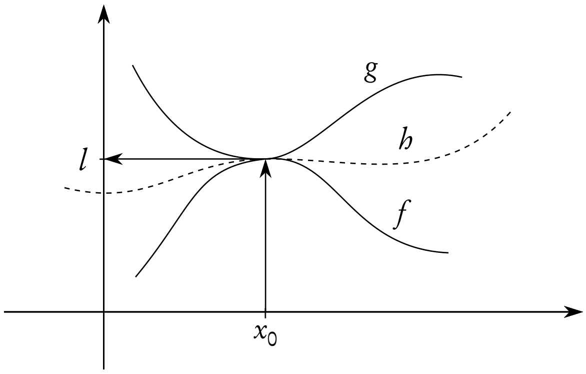


Figura 3.12 *Il teorema dei due carabinieri*

□

3.5 Limiti di successioni

Le successioni sono particolari funzioni (a valori reali nei casi che ci interessano). Dal punto di vista formale non ci sarebbe quindi nulla da aggiungere per quanto riguarda il concetto di limite. Tuttavia, visto il particolare tipo di dominio di una successione (un sottoinsieme dei naturali), conviene fare alcune precisazioni.

Innanzitutto osserviamo che per parlare di limite è indispensabile in concetto di punto di accumulazione, altrimenti non avrebbe alcun senso parlare di "avvicinarsi quanto più è possibile" ad un punto x_0. Poiché nella definizione di punto di accumulazione per un insieme si richiede che in ogni intorno di un punto cadano infiniti punti dell'insieme, per le successioni finite non esistono punti di accumulazione, per le successioni infinite l'unico "punto" di accumulazione è $+\infty$. Dunque per le successioni infinite ha senso solo il concetto di

$$\lim_{n\to+\infty} a_n, \tag{3.3}$$

tanto che si può anche scrivere semplicemente

$$\lim_n a_n, \quad \text{o, addirittura,} \quad \lim a_n. \tag{3.4}$$

Per completezza riscriviamo la definizione di limite adattandola a questo caso speciale.

Definizione 3.7. *Sia data una successione (infinita) a_n di numeri reali. Si dice che*

$$\lim_{n\to+\infty} a_n = l$$

se, scelto un arbitrario intorno I_l di l è trovare in corrispondenza un opportuno numero naturale M, tale che per ogni $n > M$ il corrispondente valore a_n della successione cada nell'intorno scelto I_l di l.

Anche qui, se $l \in \mathbb{R}$, si potrà limitarsi a considerare intorno circolari e "arbitrariamente piccoli" di l.

Esempio 3.9. Si provi, usando la definizione, che

$$\lim \frac{n+1}{n} = 1.$$

Fissato $\varepsilon > 0$ si tratta di vedere se la doppia disequaglianza

$$1-\varepsilon < \frac{n+1}{n} < 1+\varepsilon$$

è verificata per tutti i valori di n maggiori di un opportuno naturale M. La disuguaglianza di sinistra è sempre verificata. Per quella di destra si ha

$$\frac{n+1}{n} < 1+\varepsilon \Rightarrow n+1 < n+n\varepsilon \Rightarrow n > \frac{1}{\varepsilon}.$$

Per concludere basterà prendere per M il primo naturale successivo a $1/\varepsilon$.

3.6 Funzioni continue

Definizione 3.8. *Sia data una funzione f, di dominio D, e sia x_0 un punto di accumulazione per D, appartenente a D. La funzione f si dice* continua *in x_0 se*

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

Se invece x_0 è isolato per D, la funzione è per definizione continua in x_0.

È come dire che una funzione è continua se il calcolo del limite (quando ha senso) si può fare semplicemente sostituendo x_0 al posto di x nell'espressione della funzione: una bella facilitazione, se si riesce a scoprire a priori quali sono le funzioni continue!

Si noti che per le successioni il problema della continuità non si pone: il dominio delle successioni è sempre costituito solo da punti isolati.

Si dimostra, non senza qualche difficoltà, che tutte le funzioni elementari che abbiamo considerato sono continue in tutti i punti del loro dominio.

È parimenti possibile dimostrare che anche le altre funzioni elementari che non abbiamo considerato sono continue in tutti i punti del loro dominio: si tratta di tutte le funzioni polinomiali, razionali fratte, contenenti radicali, potenze con esponente di vario tipo, esponenziali, logaritmiche, trigonometriche, e quelle che si ottengono per somma, sottrazione, prodotto, quoziente e composizione di queste in tutti i modi possibili. Per ottenere funzioni non continue, al livello del nostro corso, bisogna ricorrere alle funzioni definite a pezzi, come la funzione, detta funzione segno, o signum, definita come segue:

$$\operatorname{sgn}(x) = \begin{cases} -1, & \text{se} \quad x < 0 \\ 0, & \text{se} \quad x = 0 \\ 1, & \text{se} \quad x > 0 \end{cases}, \tag{3.5}$$

il cui grafico è riportato nella figura che segue.

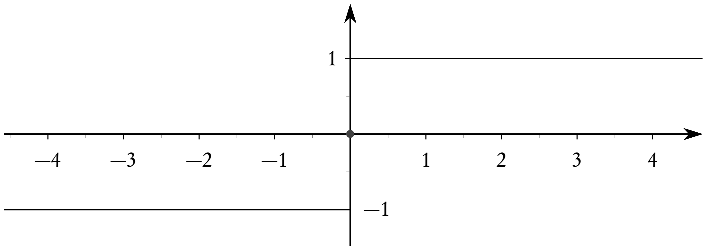


Figura 3.13 *La funzione segno*

Questa funzione non è continua nel punto 0 del suo dominio.

In primissima approssimazione si può dire che una funzione è continua se il suo grafico non presenta "strappi": l'affermazione andrebbe però precisata in dettaglio, ma ciò esula dagli scopi di questo corso.

3.7 I teoremi fondamentali sulle funzioni continue

Teorema 3.9 (Zeri di una funzione continua). *Sia $f : [a_0, b_0] \to \mathbb{R}$ una funzione continua, tale che*

$$f(a_0) \cdot f(b_0) < 0,$$

cioè tale che $f(a_0)$ e $f(b_0)$ abbiano segno opposto.

Allora esiste almeno un punto c di $[a_0, b_0]$ tale che $f(c) = 0$.

Dimostrazione. La dimostrazione si basa su una importante proprietà dei numeri reali di cui non abbiamo avuto la possibilità di occuparci, ma che possiamo considerare sufficientemente "intuitiva": Considerata una successione I_n di intervalli chiusi e limitati "inscatolati", ovvero tali che $I_{n+1} \subseteq I_n$, esiste almeno un numero reale c comune a tutti gli intervalli. Se poi l'ampiezza degli intervalli diventa arbitrariamente piccola al crescere di n allora tale numero c è unico.

Si considera allora il punto medio, m_0, di $[a_0, b_0]$: se $f(m_0) = 0$ abbiamo finito, se invece $f(m_0) \neq 0$, si considera l'intervallo $[a_1, b_1] = [a_0, m_0]$ se $f(a_0) \cdot f(m_0) < 0$, altrimenti $[a_1, b_1] = [m_0, b_0]$: si avrà in ogni caso $f(a_1) \cdot f(b_1) < 0$. Ripetendo il discorso su $[a_1, b_1]$, o si conclude con un punto c come richiesto o si procede. Dunque o si trova a un certo punto un c adatto e si conclude, o si ottiene una successione di intervalli inscatolati la cui ampiezza tende a 0. In questo secondo caso, sia c l'unico punto comune a tutti gli intervalli. Vogliamo provare che $f(c) = 0$. Se per caso fosse $f(c) \neq 0$, per il teorema della permanenza del segno esisterebbe un intorno di c dove f è diversa da 0. Ma ciò non può essere in quanto in questo intorno di c deve essere contenuto almeno uno degli intervalli precedentemente costruiti (in quanto la loro ampiezza tende a 0), e sugli estremi di questo intervallo la f ha valori di segno opposto. □

Il metodo usato in questa dimostrazione si chiama anche *metodo di bisezione* ed è anche un metodo pratico per trovare una approssimazione di uno zero, nel caso non sia possibile trovarlo con i metodi tradizionali dell'algebra.

Segnaliamo che, anche in casi semplici, è possibile che "occorrano tutti i passi" della dimostrazione, cioè che non si concluda in un numero finito di passi: basta considerare la funzione $f(x) = x^2 - 2$, nell'intervallo $[0,2]$. Si ha $f(0)=-2$ e $f(2)=2$: siamo dunque nelle condizioni di applicare il teorema, ma non potremo mai trovare la radice tra punti che si ottengono dividendo successivamente a metà gli intervalli via via costruiti, perché tutti questi punti sono razionali, mentre l'unico zero della funzione in $[0,2]$ è $\sqrt{2}$ che è irrazionale.

Il teorema vale anche se l'intervallo di definizione della funzione non è chiuso oppure non è limitato. Basta sostituire i valori $f(a_0)$ e $f(b_0)$ con i limiti da destra, sinistra, o all'infinito. La dimostrazione di questa "estensione" del teorema è una conseguenza del teorema di permanenza del segno e del teorema degli zeri che abbiamo provato. I volenterosi possono farla come utile esercizio.

È molto importante il fatto che la funzione deve essere continua in un intervallo: non basta un insieme qualunque. Per esempio la funzione $f(x)={}^1\!/x$ è continua in $[-1,0[\cup]0,1]$ e assume valori di segno opposto in -1 e 1, ma non si annulla mai.

Teorema 3.10 (Di connessione, o "Di tutti i valori", o "Dei valori intermedi"). *Sia f una funzione continua in un intervallo chiuso e limitato, $f:[a_0,b_0]\to\mathbb{R}$. Allora f assume tutti i valori compresi tra $f(a)$ e $f(b)$.*

Dimostrazione. Se $f(a)=f(b)$ non c'è nulla da provare, altrimenti basta prendere un γ tra $f(a)$ e $f(b)$ e applicare il teorema degli zeri alla funzione $g(x)=f(x)-\gamma$. □

Presi due punti c e d arbitrariamente tra a e b, il teorema si può applicare anche all'intervallo $[c,d]$: se ne deduce che la funzione assume tutti i valori compresi tra due suoi valori qualunque. Per questo si chiama "teorema di tutti i valori".

L'ultimo dei teoremi sulle funzioni continue a cui siamo interessati è il Teorema di Weierstrass, importantissimo per il problema della ricerca dei massimi e minimi e di cui parleremo nella successiva pagina 116.

3.8 Il calcolo dei limiti

Per il calcolo dei limiti delle funzioni continue, per x tendente a punti del dominio, non ci sono problemi, in quanto si può "eseguire una semplice sostituzione". Negli altri casi esistono numerose strategie e noi esamineremo in questo corso solo le più semplici.

Cominciamo con l'elencare alcuni risultati relativi alle funzioni elementari, nel caso di limiti per x tendente a punti non appartenenti al dominio; la quasi totalità di questi risultati sono intuitivi o di immediata verifica.

1. $\lim\limits_{x\to\pm\infty} x=\pm\infty$.
2. $\lim\limits_{x\to+\infty} a^x=\begin{cases}+\infty, & \text{se } a>1\\ 0, & \text{se } 0<a<1\end{cases}$.
3. $\lim\limits_{x\to-\infty} a^x=\begin{cases}0, & \text{se } a>1\\ +\infty, & \text{se } 0<a<1\end{cases}$.

   Poiché saremo interessati principalmente al caso che la base delle funzioni esponenziali sia il numero e, scriveremo brevemente questi risultati nel seguente modo:
   $$e^{+\infty}=+\infty \quad , \quad e^{-\infty}=0.$$

4. $\lim_{x\to+\infty} \log_a x = \begin{cases} +\infty, & \text{se} \quad a>1 \\ -\infty, & \text{se} \quad 0<a<1 \end{cases}$.

5. $\lim_{x\to 0^+} \log_a x = \begin{cases} -\infty, & \text{se} \quad a>1 \\ +\infty, & \text{se} \quad 0<a<1 \end{cases}$.

   Poiché saremo interessati principalmente al caso che la base delle funzioni logaritmo sia il numero e, scriveremo brevemente questi risultati nel seguente modo:

$$\ln(+\infty)=+\infty \quad , \quad \ln(0^+)=-\infty\,.$$

6. $\lim_{x\to\pm\infty} \sin x \nexists$.

7. $\lim_{x\to\pm\infty} \cos x \nexists$.

Successivamente riportiamo il risultato di alcuni teoremi che riguardano il calcolo di limiti importanti (*limiti notevoli*).

1. $\lim_{x\to 0} \dfrac{\sin x}{x} = 1$ (si tratta di un teorema che formalizza un risultato che avevamo già intravisto in uno degli esempi proposti all'inizio del capitolo).

2. $\lim_{x\to\pm\infty} \left(1+\dfrac{1}{x}\right)^x = \mathrm{e}$. A proposito di questo fondamentale limite si noti che esso si riferisce a una funzione esponenziale in cui sia la base che l'esponente sono variabili. Per trattare queste funzioni è, in generale, conveniente usare la seguente formula, conseguenza immediata della definizione di logaritmo:

$$\big(f(x)\big)^{g(x)} = \mathrm{e}^{g(x)\ln f(x)}, \tag{3.6}$$

   formula che permette di ottenere una funzione esponenziale vera e propria, cioè in cui solo l'esponente è variabile.

3. $\lim_{x\to 0} \dfrac{\mathrm{e}^x-1}{x} = 1\,.$

4. $\lim_{x\to 0} \dfrac{\ln(x+1)}{x} = 1\,.$

Successivamente segnaliamo l'uso delle regole di calcolo sulla retta reale estesa, regole di calcolo che sono state proprio definite nella previsione di un loro uso nel calcolo dei limiti: ognuna di quelle regole costituisce in realtà un apposito teorema sui limiti. Per esempio la regola

$$\forall a\in\mathbb{R},\ a+(+\infty)=+\infty$$

traduce in formule la tesi del seguente teorema: Date due funzioni f e g, tali che

$$\lim_{x\to x_0} f(x)=a\in\mathbb{R} \quad \text{e} \quad \lim_{x\to x_0} g(x)=+\infty\,,$$

si ha che

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

Chiameremo teoremi sull'*algebra dei limiti* l'insieme dei teoremi espressi mediante le regole di calcolo sulla retta reale estesa: in sostanza queste regole esprimono il fatto che i limiti si comportano bene rispetto alle operazioni fondamentali, tranne qualche caso…

Naturalmente i casi più interessanti saranno proprio quelli in cui quelle regole di calcolo non sono direttamente applicabili: in quei casi occorrerà applicare opportune strategie, di cui esamineremo solo alcuni esempi semplici. Segnaliamo comunque nuovamente che il problema del calcolo dei limiti è in generale un problema molto complesso che spesso richiede lunghe e faticose elaborazioni (non alla portata del nostro corso).

3.9 Ordini di infinito

Uno dei problemi che restano sospesi nel calcolo dei limiti è il caso in cui si presenti la situazione (forma di indecisione) $\infty - \infty$. Per esempio, nel calcolare

$$\lim_{x\to+\infty}(x^3 - x^2)$$

si ha proprio questo caso. Si può procedere nel seguente modo:

$$\lim_{x\to+\infty}(x^3 - x^2) = \lim_{x\to+\infty} x^3\left(1 - \frac{1}{x}\right) = +\infty\left(1 - \frac{1}{+\infty}\right) = +\infty(1-0) = +\infty.$$

Questo risultato si può interpretare nel seguente modo: la funzione x^3 tende all'infinito più rapidamente della funzione x^2 (o anche è un infinito più forte di x^2, o ancora è un *infinito di ordine superiore* rispetto a x^2) e quindi, dovendo fare la differenza tra un infinito più forte e una più debole, quello più debole non conta, può essere trascurato.

$$\lim_{x\to+\infty}(x^3 - x^2) = \lim_{x\to+\infty} x^3 = +\infty.$$

Anche se abbiamo espresso questo fatto con un linguaggio del tutto intuitivo, la cosa può essere resa rigorosa: gli infiniti più deboli *possono essere trascurati in una somma* (agli effetti del calcolo dei limiti, non in assoluto!).

Per poter applicare questo fatto occorre naturalmente avere una *scala* degli infiniti, in modo da sapere quali sono più forti e quali più deboli. Senza entrare nei dettagli, ai fini del nostro corso ci basterà sapere che, per $x \to +\infty$, i seguenti infiniti sono disposti in ordine crescente.

$$\ln x,\ \sqrt[3]{x},\ \sqrt{x},\ x,\ x^2,\ x^3,\ \dots,\ 2^x,\ \mathrm{e}^x,\ 3^x,\ \dots,\ \mathrm{e}^{x^2},\ \mathrm{e}^{x^3},\ \dots \tag{3.7}$$

È altresì facile mostrare che, in una somma, si possono trascurare, rispetto agli infiniti, anche le costanti, le funzioni che tendono a zero o a un valore finito, le funzioni che non hanno limite, ma rimangono limitate, come le funzioni seno e coseno per $x \to +\infty$.

Un'ulteriore applicazione del concetto di ordine di infinito si ha nel calcolo di limiti in cui si giunge alla forma di indecisione ∞/∞. In questo caso

– se il numeratore è un infinito di ordine superiore, allora il rapporto tende all'infinito;
– se il numeratore è un infinito di ordine inferiore, allora il rapporto tende a zero.

Si potrebbero fare discorsi simili per le quantità che tendono a zero, in quanto anche $0/0$ è una forma di indecisione, ma la cosa ha meno interesse (ed è sensibilmente più complessa) per il nostro corso e non ne parleremo.

Esempio 3.10. $\lim_{x\to+\infty} \frac{x^3-x^2+x-1}{3x^3+x} = \lim_{x\to+\infty} \frac{x^3}{3x^3} = \frac{1}{3}$.

Esempio 3.11. $\lim_{x\to+\infty} \frac{e^x-\ln x}{e^{x^2}-x^{33}} = \lim_{x\to+\infty} \frac{e^x}{e^{x^2}} = 0$.

3.10 Qualche esempio di calcolo dei limiti

Esempio 3.12. Calcolare

$$\lim_{x\to 0^+} \frac{\sin x}{x^2}.$$

Si ha

$$\lim_{x\to 0^+} \frac{\sin x}{x^2} = \lim_{x\to 0^+} \frac{\sin x}{x}\frac{1}{x} = 1\cdot\frac{1}{0^+} = 1\cdot(+\infty) = +\infty\,.$$

Esempio 3.13. Calcolare

$$\lim_{x\to 0^-} \frac{\sin x}{x^2}.$$

Si ha

$$\lim_{x\to 0^-} \frac{\sin x}{x^2} = \lim_{x\to 0^-} \frac{\sin x}{x}\frac{1}{x} = 1\cdot\frac{1}{0^-} = 1\cdot(-\infty) = -\infty\,.$$

Esempio 3.14. Calcolare

$$\lim_{x\to 1^+} \frac{1}{\ln x}.$$

Si ha

$$\ln 1^+ = 0^+ \Rightarrow \lim_{x\to 1^+} \frac{1}{\ln x} = \frac{1}{0^+} = +\infty\,.$$

Esempio 3.15. Calcolare

$$\lim_{x\to 1} \frac{x^3-1}{x-1}.$$

Si ha

$$\lim_{x\to 1} \frac{x^3-1}{x-1} = \frac{0}{0},$$

per cui i calcoli sulla retta reale estesa non possono essere usati. Osserviamo però che

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

da cui

$$\lim_{x\to 1} \frac{x^3-1}{x-1} = \lim_{x\to 1} \frac{\cancel{(x-1)}(x^2+x+1)}{\cancel{x-1}} = \lim_{x\to 1} x^2+x+1 = 3\,.$$

Esempio 3.16. Calcolare
$$\lim_{x\to 2^+} \frac{1-x}{(x-2)^2}\,.$$
Si ha
$$\lim_{x\to 2^+} \frac{1-x}{(x-2)^2} = \frac{-1}{0^+} = -\infty\,.$$

Esempio 3.17. Calcolare
$$\lim_{x\to 2^-} \frac{1-x}{(x-2)^2}\,.$$
Si ha
$$\lim_{x\to 2^-} \frac{1-x}{(x-2)^2} = \frac{-1}{0^+} = -\infty\,.$$

Esempio 3.18. Calcolare
$$\lim_{x\to -\infty} x\mathrm{e}^x\,.$$
La situazione è leggermente diversa da quelle esaminate prima, in quanto si giunge alla forma di indecisione $0\cdot\infty$. Si può però scrivere il limite come segue
$$\lim_{x\to -\infty} \frac{x}{\mathrm{e}^{-x}}\,.$$
Se ora si tiene conto che, per $x\to -\infty$, e^{-x} è esattamente come e^x per $x\to +\infty$, si conclude che il limite vale 0 (l'infinito del numeratore è più debole di quello del denominatore).

3.11 Esercizi

Esercizio 3.1. *Calcolare, se possibile, le seguenti espressioni sulla retta reale estesa.*

1. $-\infty+3((-\infty)(-2)+3-(-1)(+\infty))$;
2. $-\infty-(-2)(3+(+\infty)(-2)(-\infty))$;
3. $(-\infty)(-\infty)+2(3(-\infty)-2(-\infty))$;
4. $\dfrac{3(+\infty)-(-(+\infty)(-\infty)+3)}{1-(+\infty)-(+\infty+(-\infty)(+3))}$;
5. $\dfrac{-\infty(3+(+\infty-(-\infty)))}{1-(+\infty)(-\infty)(+3)}$;
6. $\dfrac{3(+\infty)-(3-(+\infty)(-\infty))}{1-(+\infty)-(+\infty+(-\infty)(+3))}$;
7. $\dfrac{2-(+\infty)(+\infty)(-\infty)+\infty}{(-2+(-\infty))(-\infty)(-3)}$.

Calcolare i limiti indicati.

Esercizio 3.2. $\lim\limits_{x\to+\infty} \dfrac{\sqrt{x}}{x+1}$.

Esercizio 3.3. $\lim\limits_{x\to 2} \dfrac{x+2}{x+1}$.

Esercizio 3.4. $\lim\limits_{x\to 1}(x^2+2)$.

Esercizio 3.5. $\lim\limits_{x\to+\infty}(x^2+2)$.

Esercizio 3.6. $\lim\limits_{x\to-\infty} \dfrac{x^2+2}{x+1}$.

Esercizio 3.7. $\lim\limits_{x\to-\infty} \dfrac{x-2}{x+1}$.

Esercizio 3.8. $\lim\limits_{x\to-\infty} \dfrac{2x}{x^3+1}$.

Esercizio 3.9. $\lim\limits_{x\to+\infty} \dfrac{2x+1}{3x-2}$.

Esercizio 3.10. $\lim\limits_{x\to+\infty} \sqrt{x+1}$.

Esercizio 3.11. $\lim\limits_{x\to+\infty} 2^x\sqrt{x^2-4}$.

Esercizio 3.12. $\lim\limits_{x\to-\infty} 2^x\sqrt{4-x}$.

Esercizio 3.13. $\lim\limits_{x\to+\infty} \dfrac{2^x}{x-1}$.

Esercizio 3.14. $\lim\limits_{x\to-\infty} \dfrac{2^x}{x-1}$.

Esercizio 3.15. $\lim\limits_{x\to+\infty} 2^x \dfrac{1}{\sqrt{x^2+1}}$.

Esercizio 3.16. $\lim\limits_{x\to+\infty} 2^x-(x^3-x+1)$.

Esercizio 3.17. $\lim\limits_{x\to-\infty} \dfrac{2^x}{x^4-x^2+2}$.

Esercizio 3.18. $\lim\limits_{x\to+\infty} \dfrac{\ln(x^2-x)}{x^2-1}$.

Esercizio 3.19. $\lim\limits_{x\to-\infty} \dfrac{x}{x^3+1}$.

Esercizio 3.20. $\lim\limits_{x\to-\infty} \dfrac{2^{x-2}}{4-x^2}$.

Esercizio 3.21. $\lim\limits_{x\to 2} e^{2-x}$.

Esercizio 3.22. $\lim\limits_{x\to +\infty} \sqrt{x-2}+x+2$.

Esercizio 3.23. $\lim\limits_{x\to +\infty} \dfrac{\sqrt{2x+1}}{x-1}$.

Esercizio 3.24. $\lim\limits_{x\to +\infty} \dfrac{e^{x^2}}{e^x}$.

Esercizio 3.25. $\lim\limits_{x\to +\infty} \dfrac{\ln(x^2)}{\ln x}$.

Esercizio 3.26. $\lim\limits_{x\to +\infty} \dfrac{x^4-2\ln x}{-\sin(2x)+x}$.

Esercizio 3.27. $\lim\limits_{x\to +\infty} \dfrac{x^2-\ln(x^3-3}{e^x-\sin x+x^2}$.

Esercizio 3.28. $\lim\limits_{x\to +\infty} \dfrac{x^3-\ln(x^3-1)+2}{e^x+\sin^2 x+x^3}$.

Esercizio 3.29. $\lim\limits_{x\to +\infty} \dfrac{e^{x^2}-x^3}{e^x+x^3}$.

Esercizio 3.30. $\lim\limits_{x\to -\infty} \dfrac{x^3-\ln(x^2)}{e^{-x}+\sin(x^2)}$.

Esercizio 3.31. $\lim\limits_{x\to -\infty} \dfrac{x^4-x^5}{e^{-x}-\sin x+2}$.

Esercizio 3.32. $\lim\limits_{x\to +\infty} \dfrac{\cos x+1}{\sqrt{x-1}}$.

Esercizio 3.33. $\lim\limits_{x\to +\infty} \dfrac{\cos(x-1)}{\sqrt{x^3-1}}$.

Esercizio 3.34. $\lim\limits_{x\to -\infty} \dfrac{\ln(1-x)}{\sqrt{x^2+1}}$.

Esercizio 3.35. $\lim\limits_{x\to 1} \dfrac{\ln(x^2)}{\ln x}$.

Esercizio 3.36. $\lim\limits_{x\to 0^+} \dfrac{e^{1/x}}{x}$.

Esercizio 3.37. $\lim\limits_{x\to -\infty} \dfrac{e^{1-x}}{x^2}$.

Esercizio 3.38. $\lim\limits_{x\to 0^+} \dfrac{\ln(1/x)}{x}$.

Esercizio 3.39. $\lim\limits_{x\to 0} \dfrac{\sin x}{1-x}$.

Esercizio 3.40. $\lim\limits_{x\to 0} \dfrac{\sin(x^2)}{x}$.

Esercizio 3.41. $\lim\limits_{x\to 0} \dfrac{\sin x}{\sqrt{x}}$.

Esercizio 3.42. $\lim\limits_{x\to 0} \dfrac{x}{1-e^x}$.

4 Introduzione alla matematica finanziaria

In questo capitolo proponiamo una breve introduzione ad alcuni argomenti di matematica finanziaria, con lo scopo principale di mostrare una applicazione delle nozioni e tecniche matematiche presentate.

4.1 Serie geometriche finite e infinite

Definizione 4.1. *Si chiama* serie geometrica finita *una successione di numeri reali*

$$a_1, a_2, \dots, a_n, \dots$$

tale che il rapporto tra ogni termine e il precedente sia costante

$$q = \frac{a_n}{a_{n-1}}, \quad \textit{ovvero} \quad a_n = q a_{n-1}. \tag{4.1}$$

Se la successione è finita la serie si chiama spesso progressione geometrica*. Il rapporto q definito nella* (4.1) *si chiama* ragione *della serie geometrica.*

Si noti che una ripetuta applicazione della (4.1) porta a concludere che

$$a_2 = a_1 q,\ a_3 = a_2 q = a_1 q^2,\ a_4 = a_3 q = a_1 q^3,\ \dots,\ a_n = a_{n-1} q = a_1 q^{n-1}. \tag{4.2}$$

Indicato semplicemente con a il primo termine, possiamo dunque scrivere una serie geometrica di ragione q come

$$a, aq, aq^2, aq^3, \dots, aq^{n-1}, \dots \tag{4.3}$$

Nel caso di una serie finita è facile calcolare la somma di tutti gli n termini della serie. Cominciamo col ricordare una formula per la scomposizione della differenza di due potenze: in analogia alla regola sulla scomposizione della differenza di due cubi si può provare che

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + a^2 b^{n-3} + a b^{n-2} + b^{n-1}). \tag{4.4}$$

Posto, nella (4.4), $a = 1$ e $b = q$ si ottiene, per $q \neq 1$,

$$1 - q^n = (1-q)(1 + q + q^2 + \cdots + q^{n-1}), \quad \text{ovvero} \quad 1 + q + q^2 + \cdots + q^{n-1} = \frac{1-q^n}{1-q}. \tag{4.5}$$

Si ha allora

$$a_1 + a_2 + \cdots + a_n = a + aq + \cdots + aq^{n-1} = a(1 + q + q^2 + \cdots + q^{n-1}) = a\frac{1-q^n}{1-q},\ q \neq 1, \tag{4.6}$$

mentre

$$a_1+a_2+\cdots+a_n=a+a+\cdots+a=n, \quad \text{se} \quad q=1. \tag{4.7}$$

Supponiamo ora che la serie sia infinita e vediamo che cosa si può affermare relativamente alla somma di tutti i suoi infiniti termini. Per $q=1$ chiaramente la somma diventa infinita, per gli altri valori di q tutto dipende dal comportamento della successione q^n. Si ha

1. se $q=-1$, q^n è una successione indeterminata;
2. se $q>1$, q^n è una successione divergente a $+\infty$;
3. se $q<-1$, q^n è una successione indeterminata (la successione dei termini di indice pari tende a $+\infty$, quella dei termini di indice dispari a $-\infty$;
4. se $|q|<1$, cioè $-1<q<1$, q^n è una successione convergente a 0.

Se ne deduce che la somma degli infiniti termini di una serie geometrica di ragione q è finita se e solo se $|q|<1$ e si ha precisamente

$$a_1+a_2+\cdots+a_n+\cdots=a+aq+\cdots+aq^{n-1}+\cdots=a\frac{1}{1-q} \tag{4.8}$$

Non deve meravigliare il fatto che la somma di infiniti addendi tutti positivi possa essere finita. Lo si può capire con il seguente semplice esempio.

Esempio 4.1. Si consideri sulla retta un segmento lungo 1 e gli si affianchi sulla destra un segmento lungo la metà, cioè 1/2, e successivamente un segmento lungo la metà della metà, cioè 1/4, e così via. Risulta evidente dalla figura 4.1 che il segmento complessivo (cioè la somma di tutti questi segmenti) avrà sempre lunghezza inferiore a 2 e raggiungerà 2 solo addizionando un'infinità di segmenti.

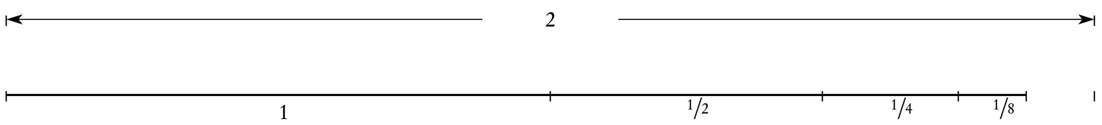


Figura 4.1 *Somma di infiniti segmenti*

4.2 Regimi finanziari

4.2.1 Considerazioni introduttive

Si definisce *operazione finanziaria* un'operazione che produce una variazione di capitale nel tempo. Il capitale iniziale, ad un certo tempo (o "epoca") x, indicato con C, viene chiamato *valore attuale*, il capitale maturato al tempo (o "epoca") successivo y, indicato con M, viene chiamato *montante*. Se due individui si scambiano i capitali C ed M, questi si dicono *finanziariamente equivalenti*. Si usa parlare di operazione con scadenzario $(C;M)/(x;y)$. Le operazioni finanziarie si dicono *certe* se gli importi sono disponibili con certezza: sono quelle di cui ci occuperemo brevemente.

In un'operazione di *investimento* avremo normalmente $M>C$ e la differenza positiva $M-C$ si chiama *interesse* e si indica con I:

$$M=C+I \quad \Rightarrow \quad \frac{M}{C}=\frac{C+I}{C}=1+\frac{I}{C}. \tag{4.9}$$

Si pone

$$\frac{M}{C} = r(x, y), \quad \frac{I}{C} = i(x, y), \tag{4.10}$$

e queste quantità sono chiamate, rispettivamente, *fattore di capitalizzazione* o di *montante* e *tasso d'interesse*, rispetto alle epoche x e y. Esse rappresentano rispettivamente il montante e l'interesse ottenuti investendo un capitale unitario dall'epoca x all'epoca y. La (4.9) si può riscrivere come

$$r(x, y) = 1 + i(x, y) \quad \Rightarrow \quad M = C \cdot r(x, y) = C \cdot (1 + i(x, y)). \tag{4.11}$$

L'operazione inversa dell'operazione di investimento si chiama operazione di *attualizzazione* o di *anticipazione*. In questo caso il capitale M disponibile all'epoca y viene attualizzato (riportato indietro nel tempo) all'epoca x. Anche in questo caso avremo $M > C$ e la differenza positiva $M - C$ si chiama *sconto* e si indica con D (da "discount"):

$$C = M - D \quad \Rightarrow \quad \frac{C}{M} = \frac{M - D}{M} = 1 - \frac{D}{M} \tag{4.12}$$

Si pone

$$\frac{C}{M} = v(x, y), \quad \frac{D}{M} = d(x, y), \tag{4.13}$$

e queste quantità sono chiamate, rispettivamente, *fattore di attualizzazione* o di *sconto* e *tasso di sconto*, rispetto alle epoche x e y. Esse rappresentano rispettivamente il valore attuale e lo sconto ottenuti attualizzando un capitale unitario dall'epoca y all'epoca precedente x. La (4.12) si può riscrivere come

$$v(x, y) = 1 - d(x, y) \quad \Rightarrow \quad C = M \cdot v(x, y) = M \cdot (1 - d(x, y)). \tag{4.14}$$

In base alle definizioni date si può dedurre che

$$r(x, y) = \frac{1}{v(x, y)}. \tag{4.15}$$

Si noti come le $M = C + I$ e $D = M - C$ siano in realtà la stessa formula, interpretata però in modo diverso: è una situazione frequente nelle varie formalizzazioni economiche che risponde a criteri di utilità.

Nel seguito indicheremo l'epoca iniziale con 0 e l'epoca finale con t (dunque t rappresenta la durata dell'investimento) e scriveremo

$$r(0, t) = r_t = r(t), \quad v(0, t) = v_t = v(t), \quad i(0, t) = i_t = i(t), \quad d(0, t) = d_t = d(t). \tag{4.16}$$

Nel caso in cui $t = 1$ useremo le notazioni ulteriormente semplificate

$$r(0, 1) = r, \quad v(0, 1) = v, \quad i(0, 1) = i, \quad d(0, 1) = d. \tag{4.17}$$

Se, come succede spesso nella pratica, l'unità dei tempi è l'anno[1] si parla di fattori o tassi *annui*.

[1] Abitualmente l'anno commerciale è composto da 12 mesi di 30 giorni ciascuno, per un totale di 360 giorni.

Riepiloghiamo le relazioni che esistono tra queste grandezze:

$$r=\frac{1}{v}=1+i=\frac{1}{1-d}; \tag{4.18}$$

$$v=1-d=\frac{1}{r}=\frac{1}{1+i}; \tag{4.19}$$

$$d=\frac{i}{1+i}=\frac{r-1}{r}=1-v; \tag{4.20}$$

$$i=r-1=\frac{d}{1-d}=\frac{1-v}{v}. \tag{4.21}$$

Esempio 4.2. Si debba corrispondere dopo un tempo $t=1$ un capitale di 1000 con il tasso effettivo i del 25%. Calcolare la somma da anticipare e il tasso effettivo di sconto.

Possiamo calcolare il fattore di attualizzazione v mediante la (4.19):

$$v=\frac{1}{1+i}=\frac{1}{1.25}=0.80.$$

Si ha dunque $C=1000\cdot v=800$ e $d=1-v=0.20=20\%$.

Osserviamo che normalmente i tassi sono espressi in percentuale e che $x\%$ significa $x/100$.

Possiamo ora dare la definizione di *regime finanziario* come un insieme di "regole" che consente di effettuare operazioni di capitalizzazione e attualizzazione e che quindi ci consente di confrontare importi disponibili in epoche diverse. Esamineremo in dettaglio solo due tra i più importanti regimi finanziari.

4.2.2 Regime dell'interesse semplice

Il più semplice regime finanziario è quello in cui l'interesse I è proporzionale sia al capitale iniziale C che alla durata t dell'impiego:

$$I=C\cdot i\cdot t, \tag{4.22}$$

dove i è il tasso di interesse per un periodo unitario (di solito il tasso annuo). Dunque il tasso di interesse nel periodo t è legato al tasso unitario dalla semplice legge $i(t)=i\cdot t$.

Di conseguenza la legge di capitalizzazione, il fattore di capitalizzazione e quello di attualizzazione risultano essere

$$M(t)=C(1+it),\quad r(t)=1+it,\quad v(t)=\frac{1}{1+it},\quad t\in[0,+\infty[. \tag{4.23}$$

Il grafico del fattore di capitalizzazione in funzione del tempo è una linea retta con pendenza o coefficiente angolare i: per questo motivo si parla di *legge lineare*.

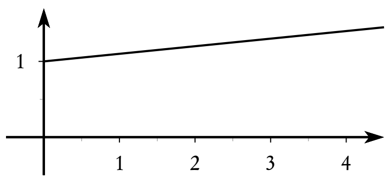


Figura 4.2 *Fattore di capitalizzazione nel regime dell'interesse semplice,* $i=10\%$

Si tenga presente che, nelle situazioni standard, i è molto piccolo e quindi la pendenza della retta è molto limitata.

Esempio 4.3. Calcolare il montante ottenuto da un investimento di capitale iniziale di 1000, al tasso annuo del 3% dopo 2 anni e 3 mesi nel regime dell'interesse semplice.

Intanto osserviamo che $i = 3\%$ equivale a $i = 0.03$ e che 2 anni e 3 mesi significa 2 anni e $1/4$, per un totale periodo $t = 9/4$ di anno. Si ha allora subito:

$$M = C(1+it) = 1000\left(1+0.03\cdot\frac{9}{4}\right) = 1067,5$$

Esempio 4.4. In quanto tempo un capitale di 530 genera un montante di 600 se investito in regime di capitalizzazione semplice al tasso annuo del 1%?

Da $M = C(1+it)$ ricaviamo

$$t = \frac{1}{i}\left(\frac{M}{C}-1\right).$$

Quindi

$$t = \frac{1}{0.01}\left(\frac{600}{530}-1\right) = 13,20755.$$

Il risultato finale è espresso in anni (13) e frazione di anni (0.20755). Per trasformare la frazione di anni in mesi basterà moltiplicarla per 12:

$$0.20755\cdot 12 = 2.4906,\quad \text{ovvero 2 mesi e una frazione di 0.4906 mese.}$$

Moltiplicando la frazione di mese per 30 otterremo i giorni:

$$0.4906\cdot 30 = 14.718,,\quad \text{ovvero 14 mesi e una frazione di 0.718 giorno.}$$

Si potrebbero ulteriormente ottenere le ore, i minuti, ecc.

4.2.3 Regime dell'interesse composto

Nel regime dell'interesse semplice gli interessi non vengono capitalizzati e quindi non generano ulteriore interesse. Nel regime dell'interesse composto, invece, gli interessi sono periodicamente capitalizzati e quindi generano nuovo interesse. La differenza tra questi due regimi è dunque evidente.

Se si investe un capitale C ad un interesse annuo i, alla fine dell'anno si avrà un montante $M_1 = C(1+i)$. Questo montante costituisce il capitale per il nuovo anno e dunque alla fine del secondo anno si avrà un montante

$$M_2 = M_1(1+i) = C(1+i)^2.$$

Se il numero di periodi (anni per noi) è intero, dopo t anni il montante sarà

$$M_t = C(1+i)^t. \tag{4.24}$$

Una legge di capitalizzazione di questo tipo si chiama, appunto, *regime dell'interesse composto* ed è espressa da una funzione di tipo esponenziale. Sulla base delle proprietà della funzione esponenziale, se $t = t_1 + t_2$, si ha

$$M_t = C(1+i)^t = C(1+i)^{t_1+t_2} = C(1+i)^{t_1}(1+i)^{t_2} \quad \Rightarrow \quad r(t_1+t_2) = r(t_1)r(t_2). \tag{4.25}$$

Un fattore di capitalizzazione che soddisfi la (4.25) si chiama *scindibile* e si dimostra che la legge di capitalizzazione degli interessi composti è l'unica scindibile.

La proprietà di scindibilità è molto importante nella matematica finanziaria e si può compiutamente formulare nel modo seguente.

Definizione 4.2. *Una legge di capitalizzazione si dice* scindibile *se il montante all'epoca t_3 di un capitale C investito all'epoca $t_1 < t_3$ non cambia se si disinveste il montante maturato ad un'epoca intermedia t_2 e lo si reinveste immediatamente lasciandolo investito sino all'epoca t_3.*

La legge di capitalizzazione (4.24) vale, in base alla proprietà di scindibilità, anche se il numero di periodi non è intero. Dunque

$$M_t = C(1+i)^t, \quad r(t) = (1+i)^t, \quad t \in [0, +\infty[. \tag{4.26}$$

Il grafico è un tipico grafico di funzione esponenziale.

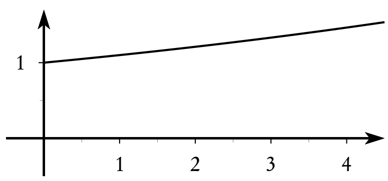


Figura 4.3 *Fattore di capitalizzazione nel regime dell'interesse composto, $i = 10\%$*

Confrontando le figure 4.2 e 4.3 è difficile apprezzare le differenze. Tutto diventa più evidente con tassi di interesse più elevati ed eseguendo un opportuno ingrandimento. La figura 4.4, con un interesse molto elevato (70%) per esaltare le differenze, mostra che per periodi inferiori all'anno l'interesse semplice è più conveniente, il contrario per periodi superiori all'anno.

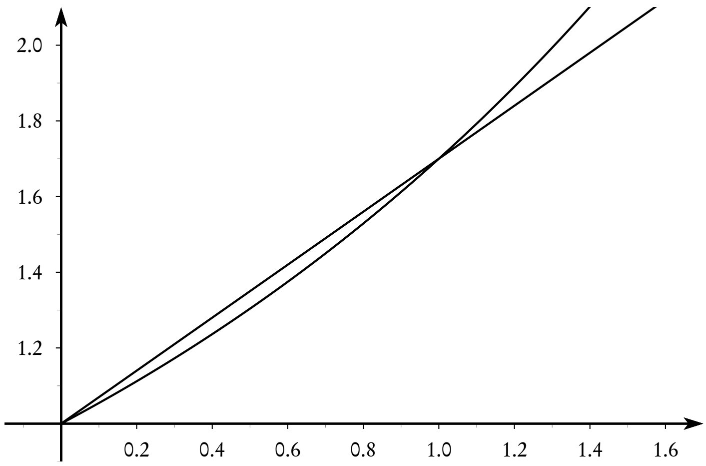


Figura 4.4 *Confronto fra i fattori di capitalizzazione nei regimi semplice e composto*

Esempio 4.5. Se si investe una somma di 1000 in regime di interesse composto al tasso annuo dell'1,32%, quanto tempo deve durare l'investimento affinché il montante arrivi a 1500?

Dalla (4.26) si trova

$$(1+i)^t=\frac{M}{C}\quad\Rightarrow\quad\ln(1+i)^t=\ln\frac{M}{C}\quad\Rightarrow\quad t=\frac{\ln(M/C)}{\ln(1+i)}.$$

Dunque

$$t=\frac{\ln 1.5}{\ln 1.0132}=30,919343=30 \text{ anni } 11 \text{ mesi e } 1 \text{ giorno.}$$

Esempio 4.6. Calcolare il valore attuale di 110 in un regime a capitalizzazione composta generati da un investimento a 2 anni e 5 mesi al tasso di interesse dello 0.5% annuo.

Intanto 2 anni e 5 mesi equivalgono a $^{29}/_{12}$ di anno. Dalla (4.26) si trova poi

$$C=M(1+i)^{-t}=110\cdot(1.005)^{-29/12}=108.68210.$$

4.3 Rendite

4.3.1 Considerazioni introduttive

Definizione 4.3. *Si chiama* rendita *una successione di capitali da riscuotere (o da pagare) a scadenze determinate.*

Esempi di rendita sono costituiti dal canone di affitto di un immobile, dalle rate di un mutuo, dallo stipendio di un dipendente, dalla pensione.

Si chiamano *rate* i singoli capitali esigibili o dovuti alle diverse scadenze.

Le rendite si dicono *certe* se le rate sono a priori fissate nel numero, nell'ammontare e nelle epoche di pagamento: gli esempi citati prima sono rendite certe; si dicono *aleatorie* se il numero, l'ammontare e le scadenze dipendono dal verificarsi di eventi dei quali non è noto a priori se e quando si verificheranno:

le pensioni di invalidità sono esempi di rendite aleatorie, in quanto saranno pagate solo se si verifica l'invalidità e dal momento in cui si verifica. Ci occuperemo qui brevemente solo delle rendite certe.

Le rendite si dicono *periodiche* se le rate sono tutte equiintervallate tra di loro, *aperiodiche* in caso contrario. Nel caso periodico (l'unico di cui ci occuperemo) l'intervallo costante tra una rata e la successiva è detto *periodo di competenza* o semplicemente *periodo*. Naturalmente la rendita si dirà *annua*, *semestrale*, *mensile*, ..., se il periodo è di un anno, di un semestre, di un mese, ...

Il pagamento delle rate può aversi all'inizio o alla fine di ciascun periodo: nel primo caso la rendita viene detta *anticipata*, nel secondo caso *posticipata*. Per esempio usualmente lo stipendio è pagato in rate posticipate, l'affitto al proprietario di un immobile è di solito pagato in rate anticipate. A seconda che il numero di rate sia finito o (almeno teoricamente) infinito, la rendita si chiama *temporanea* o *perpetua*.

Infine una rendita è detta *costante* se tutte le rate hanno lo stesso importo, *variabile* nel caso contrario. Le rendite costanti di importo unitario si dicono, appunto, *rendite unitarie*. Ci occuperemo solo delle rendite costanti.

4.3.2 Valore attuale e montante di una rendita

Data una rendita di qualunque tipo è importante nella pratica "valutare" la rendita stessa, ossia stabilire quale sia la somma che, in un determinato istante, è equivalente ad essa. Si parlerà, naturalmente, di *valore attuale della rendita* quando l'istante di valutazione è anteriore al primo periodo della rendita, di *montante della rendita* quando l'istante di valutazione è quello terminale della rendita stessa. Nel calcolo del valore attuale si parla di rendite *immediate* se l'istante di valutazione è l'istante iniziale del primo periodo, di rendite *differite* se l'istante di valutazione è anteriore all'inizio del primo periodo.

Il calcolo del valore attuale o del montante si farà naturalmente calcolando i valori attuali o i montanti delle singole rate e sommando gli importi così ottenuti. Per fare questo occorrerà scegliere uno specifico regime finanziario e, di norma, la scelta cade sul regime dell'interesse composto.

Ci occuperemo di seguito solo di alcune semplici situazioni, in particolare di rendite periodiche e costanti, temporanee o perpetue, nel regime di interesse composto. Inoltre possiamo supporre che le rendite siano unitarie: nel caso di rendite con rate non unitarie di valore R, basterà moltiplicare per R i risultati ottenuti per le rendite unitarie.

Rendita unitaria immediata, di durata n periodi

Se la rendita è posticipata, la prima rata sarà pagata 1 periodo esatto dopo l'istante di valutazione: il suo valore attuale sarà allora v. La seconda rata, da pagare dopo due periodi, avrà valore attuale v^2, l'ultima, da pagare dopo n periodi, avrà valore attuale v^n. Il valore attuale complessivo sarà allora

$$v + v^2 + v^3 + \cdots + v^n.$$

Questo valore si indica convenzionalmente con la scrittura

$$a_{\overline{n}|i} \tag{4.27}$$

che si legge "a figurato n al tasso i".

Tenendo conto delle formule relative alla somma di una progressione geometrica avremo

$$a_{\overline{n}|i} = v\,\frac{1-v^n}{1-v}. \tag{4.28}$$

Se teniamo conto del legame esistente tra v ed i, legame espresso dalla formula (4.19), otterremo

$$a_{\overline{n}|i}=\frac{1-(1+i)^{-n}}{i}. \tag{4.29}$$

Se invece la rendita è anticipata, basterà tenere conto che la prima rata non va affatto attualizzata, e che le successive vanno attualizzate per un anno di meno. Si avrà allora (si usa la notazione con la dieresi sopra la a per indicare che si tratta di rendita anticipata)

$$\ddot{a}_{\overline{n}|i}=1+v+v^2+\cdots+v^{n-1}=\frac{1-v^n}{1-v}=(1+i)\frac{1-(1+i)^{-n}}{i}=(1+i)a_{\overline{n}|i}. \tag{4.30}$$

Per quanto riguarda il calcolo del montante basterà ora capitalizzare anziché anticipare le singole rate. Nel caso della rendita posticipata la prima va capitalizzata per $n-1$ periodi, la seconda per $n-2$ periodi, l'ultima non va capitalizzata perché viene pagata nello stesso istante in cui si fa la valutazione. Si ha dunque, tenendo anche conto che il fattore di montante è $r=1+i$,

$$s_{\overline{n}|i}=r^{n-1}+r^{n-2}+\cdots+1=\frac{1-r^n}{1-r}=\frac{(1+i)^n-1}{i}=(1+i)^n a_{\overline{n}|i}, \tag{4.31}$$

ove abbiamo sempre utilizzato la scrittura convenzionale per il montante.

Se invece la rendita è anticipata, la prima rata andrà capitalizzata per n periodi, la seconda per $n-1$ periodi, l'ultima per 1 periodo. Avremo dunque

$$\ddot{s}_{\overline{n}|i}=r^n+r^{n-1}+\cdots+r=r\frac{1-r^n}{1-r}=(1+i)\frac{(1+i)^n-1}{i}=(1+i)s_{\overline{n}|i}=(1+i)^n\ddot{a}_{\overline{n}|i}. \tag{4.32}$$

Rendita unitaria differita di t periodi, di durata n periodi

Per ottenere le formule relative alle rendite differite rispetto all'istante di valutazione (la cosa ha interesse solo per il valore attuale), basterà tenere conto che ogni rata andrà anticipata per t periodi di più rispetto alle rendite immediate. Utilizzando i simboli autoesplicativi

$$_{t|}a_{\overline{n}|i} \quad \text{e} \quad _{t|}\ddot{a}_{\overline{n}|i}$$

per le rendite posticipate e anticipate rispettivamente si avrà dunque

$$_{t|}a_{\overline{n}|i}=v^{t+1}+v^{t+2}+\cdots+v^{t+n}=v^t\left(v+v^2+\cdots+v^n\right)=v^t a_{\overline{n}|i}, \tag{4.33}$$

$$_{t|}\ddot{a}_{\overline{n}|i}=v^t+v^{t+1}+\cdots+v^{t+n-1}=v^t\left(1+v+v^2*\cdots+v^{n-1}\right)=v^t\ddot{a}_{\overline{n}|i}. \tag{4.34}$$

Rendite unitarie perpetue

Per ottenere le formule relative alle rendite perpetue (naturalmente solo relativamente alle formule di attualizzazione: non ha senso parlare di montante in quanto non esiste un istante finale) basta fare il limite per $n \to +\infty$ nella formule relative alle rendite temporanee. Nelle nomenclatura si usa il simbolo ∞ al posto di n. Si ottengono facilmente le seguenti formule

$$a_{\overline{\infty}|i} = \frac{1}{i}. \tag{4.35}$$

$$\ddot{a}_{\overline{\infty}|i} = (1+i)\frac{1}{i} = 1 + \frac{1}{i}. \tag{4.36}$$

$$_{t|}a_{\overline{\infty}|i} = \frac{v^t}{i}. \tag{4.37}$$

Si noti in particolare che per le rendite perpetue quella anticipata, una volta pagata la prima rata che contribuisce al capitale solo con il suo valore nominale (1), non ha alcuna differenza con quella posticipata, come è giusto che sia.

Tutte le formule presentate sono di immediata applicazione. Proponiamo un semplice esempio applicativo.

Esempio 4.7. Calcolare il valore attuale e il montante di una rendita immediata posticipata annua con rata di 1200 e durata 15 anni nel regime di interesse composto e con il tasso di valutazione del 12% annuo.

Trasformiamo i in valore decimale: $i = 0.12$. Per ottenere i valori richiesti basta moltiplicare le rendite unitarie per il valore $R = 1200$. Si ottiene

$$Ra_{\overline{15}|i} = 1200\frac{1-(1+0.12)^{-15}}{0.12} = 8173.0374.$$

$$Rs_{\overline{15}|i} = (1+i)^{15} Ra_{\overline{15}|i} = 44735.6676.$$

Anche se non abbiamo scritto formule esplicite per il caso di rendite con rate variabili, mostriamo su un esempio come si possa comunque fare il calcolo semplicemente applicando le leggi di attualizzazione e capitalizzazione, nel regime di interesse opportuno, che in questo caso è quello composto.

Esempio 4.8. Data una rendita di 4 rate , rispettivamente di importo 1000, 1500, 1600, 2400 e di scadenze 1 anno, 1 anno e 4 mesi, 1 anno e 6 mesi, 3 anni a partire dal momento attuale, calcolarne il valore attuale e il montante supponendo un interesse del 9.5% annuo, in regime di interesse composto.

Per il valore attuale basterà attualizzare i singoli importi al tasso dello 0.095 per i periodi corretti e cioè

$$1, \quad \frac{16}{12}, \quad \frac{18}{12}, \quad 3.$$

Si ottiene

$$1000\frac{1}{1+0.095} + 1500\frac{1}{(1+0.095)^{16/12}} + 1600\frac{1}{(1+0.095)^{18/12}} + 2400\frac{1}{(1+0.095)^3} =$$
$$= 913.242 + 1329.043 + 1396.365 + 1826.969 = 5465.619.$$

Per il montante si può semplicemente capitalizzare a 3 anni il valore attuale ottenuto. Si ottiene

$$(1+0.095)^3\, 5465.619 = 7175.988.$$

5 Derivate per funzioni di una variabile

5.1 Tangenti a una circonferenza e tangenti a una curva

Prima di affrontare il problema della derivazione di una funzione, è opportuno richiamare il concetto di tangente a una circonferenza, evidenziandone i passi salienti. Ci sono sostanzialmente tre modi per definire la tangente a un circonferenza, di cui indicheremo con O il centro.

1ª *definizione* La tangente a una circonferenza in un suo punto P è la perpendicolare per P al raggio OP.

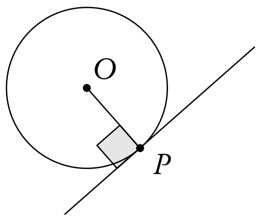


Figura 5.1 *Tangente a una circonferenza: la perpendicolare al raggio*

2ª *definizione* La tangente a una circonferenza in un punto P è l'unica retta per P avente in comune con la circonferenza solo il punto P.

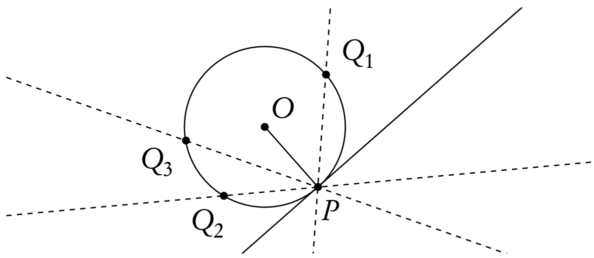


Figura 5.2 *Tangente a una circonferenza: l'unica retta avente un solo punto in comune con la circonferenza*

3ª *definizione* La tangente a una circonferenza in un suo punto P si ottiene da un "processo al limite": considerata la secante per P e un altro punto Q, se si fa tendere Q a P, la "posizione limite" è quella della tangente; si usa anche dire che la tangente è una "secante passante per due punti coincidenti", ma l'espressione, almeno enunciata così brutalmente, non ha senso, perché due punti coincidenti sono un unico punto e di rette per un punto ce ne sono infinite.

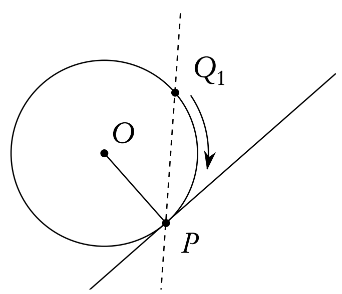


Figura 5.3 *Tangente a una circonferenza: la posizione limite della secante*

Ci possiamo ora porre il problema: è possibile estendere le definizioni precedenti a una curva generica che non sia una circonferenza? Sicuramente la prima definizione non può essere usata in generale, perché una curva generica non ha un centro e un raggio; purtroppo nemmeno la seconda definizione può essere usata. Se si esamina il grafico che segue, la retta t ha, intuitivamente, il diritto di essere chiamata tangente alla curva (grafico del seno) nel punto P, ma essa ha ben più di un punto in comune con la curva.

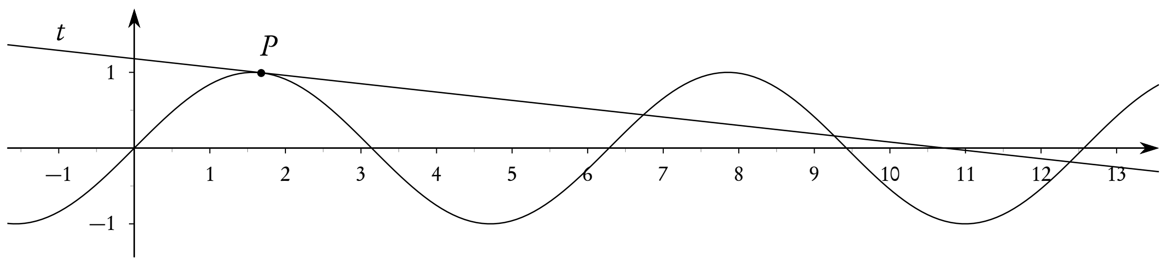


Figura 5.4 *Tangente alla funzione seno in un punto P*

La definizione che può invece essere estesa al caso generale è la terza, e questa generalizzazione conduce al concetto di derivata.

5.2 Derivata e tangente al grafico di una funzione

Sia data, nel piano cartesiano, una curva di equazione $y = f(x)$ (cioè il grafico di una funzione reale di variabile reale). Su questa curva fissiamo un punto $P(x_P, y_P) = (x_P, f(x_P))$. Il problema di cui vogliamo occuparci è il seguente: è possibile dare una definizione formalmente ineccepibile di tangente alla curva, sulla falsariga di quanto fatto nel caso della circonferenza (3^a definizione), e, in caso di risposta affermativa, è possibile costruire un algoritmo generale per trovare l'equazione di questa retta tangente (anzi, ci basterà il coefficiente angolare, perchè ovviamente la tangente, se esiste, passerà per P).

Ricordiamo che, se si hanno due punti $P(x_P, y_P)$ e $Q(x_Q, y_Q)$, aventi diversa ascissa, cioè non appartenenti a una retta verticale, la retta per i due punti ha coefficiente angolare

$$m_{PQ} = \frac{y_Q - y_P}{x_Q - x_P} = \frac{\Delta y}{\Delta x}.$$

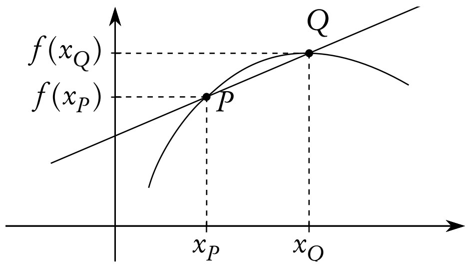


Figura 5.5 *Secante a una curva per due punti*

Se ora consideriamo una funzione di equazione $y = f(x)$ e un punto $P(x_P, f(x_P))$, per trovare la tangente in P possiamo procedere prendendo un secondo punto $Q(x_Q, f(x_Q))$ sul grafico e tracciando la retta per P e Q, detta *retta secante*, la quale avrà coefficiente angolare

$$m_{PQ} = \frac{y_Q - y_P}{x_Q - x_P} = \frac{f(x_Q) - f(x_P)}{x_Q - x_P} = \frac{\Delta f}{\Delta x}. \tag{5.1}$$

Definizione 5.1. *Il rapporto* $\Delta f/\Delta x$*, definito nella formula* (5.1) *si chiama* rapporto incrementale *della funzione* f *relativo al punto* x_P *e all'incremento* $h = x_Q - x_P = \Delta x$.

Se ora prendiamo il punto Q "sempre più vicino a P", la secante si avvicina sempre più a quella che, intuitivamente, ci pare la miglior candidata a essere definita retta *tangente*.

In generale indicheremo con x_0, o genericamente con x, l'ascissa del punto P e con $x_0 + h$ l'ascissa del punto Q. Il rapporto incrementale si scriverà allora

$$\frac{\Delta f}{\Delta x} = \frac{f(x_0 + h) - f(x_0)}{h}. \tag{5.2}$$

Definizione 5.2. *Data una funzione di equazione* $y = f(x)$*, definita in tutto un intorno* I_{x_0} *di un punto* x_0 *del dominio, si dice* derivata prima *della funzione* f *nel punto* x_0*, e si indica con*

$$f'(x_0), \ \textit{oppure} \ \mathrm{D}f(x_0),$$

il limite, se esiste ed è finito*, del rapporto incrementale, al tendere di* Q *a* P

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}. \tag{5.3}$$

In questo caso (cioè se il limite (5.3) esiste finito*), la funzione si dice* derivabile *in* x_0.

Esempio 5.1. Calcoliamo l'equazione della retta tangente al grafico della funzione $f(x) = \mathrm{e}^x$, nel punto di ascissa 1. Per il coefficiente angolare si ha

$$\lim_{h \to 0} \frac{\mathrm{e}^{1+h} - \mathrm{e}}{h} = \lim_{h \to 0} \frac{\mathrm{e}^1 \mathrm{e}^h - \mathrm{e}}{h} = \lim_{h \to 0} \frac{\mathrm{e}\,\mathrm{e}^h - \mathrm{e}}{h} = \lim_{h \to 0} \mathrm{e}\,\frac{\mathrm{e}^h - 1}{h} = \mathrm{e} \cdot 1 = \mathrm{e}.$$

Se teniamo conto che la retta deve passare per $P(1, \mathrm{e})$, otteniamo

$$y - \mathrm{e} = \mathrm{e}(x - 1) \Rightarrow y = \mathrm{e}\,x.$$

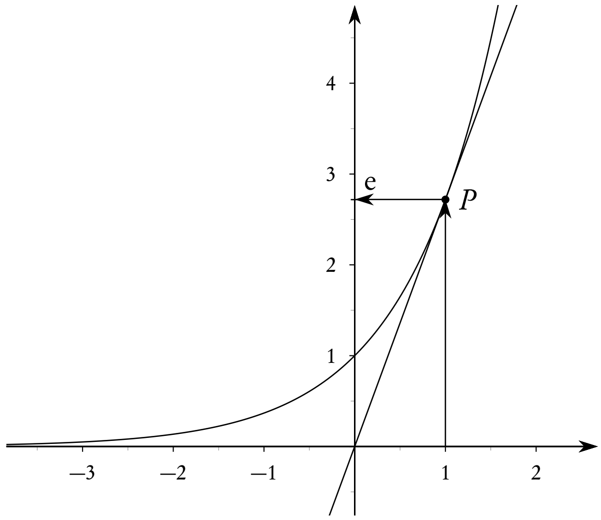


Figura 5.6 *Tangente alla funzione* e^x *nel punto di ascissa* 1

In generale non è indispensabile impostare il procedimento di calcolo del limite sulla base di un punto x_0 fissato: è possibile ricavare il coefficiente angolare della retta tangente a una curva di equazione $y = f(x)$ in un punto di ascissa x_0 qualsiasi. Così facendo si otterrà una espressione dipendente da x_0, e non più un singolo valore numerico. Questa espressione si chiama la *funzione derivata prima*, o semplicemente *funzione derivata* o a volte addirittura solo *derivata*. In generale, se non c'è possibilità di equivoco, in questi casi si scrive semplicemente x al posto di x_0.

Definizione 5.3. *Data una funzione f, definita in un intervallo I, si dice* funzione derivata prima *di f la funzione f' espressa da*

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}, \tag{5.4}$$

nei punti di ascissa $x \in I$ per i quali tale limite esiste finito.

Esempio 5.2. Riprendiamo in esame la funzione esponenziale dell'esempio precedente, e calcoliamo la derivata in un punto x generico:

$$\lim_{h\to 0} \frac{\mathrm{e}^{x+h} - \mathrm{e}^x}{h} = \lim_{h\to 0} \frac{\mathrm{e}^x \mathrm{e}^h - \mathrm{e}^x}{h} = \lim_{h\to 0} \mathrm{e}^x \, \frac{\mathrm{e}^h - \mathrm{e}^x}{h} = \mathrm{e}^x \cdot 1 = \mathrm{e}^x .$$

Naturalmente se al posto di x mettiamo il numero 1, come nell'esempio precedente, otteniamo nuovamente il numero e.

Esempio 5.3. Calcolare la derivata di $f(x) = x^3$. Si ha

$$\mathrm{D}(x^3) = \lim_{h\to 0} \frac{(x+h)^3 - x^3}{h} = \lim_{h\to 0} \frac{x^3 + 3x^2h + 3xh^2 + h^3 - x^3}{h} = \lim_{h\to 0} 3x^2 + 3xh + h^2 = 3x^2 .$$

Ribadiamo che questo risultato significa che

in corrispondenza all'ascissa	il grafico di $f(x) = x^3$ ammette retta tangente con coefficiente angolare
$\dots$	$\dots$
$x = -2$	$m = f'(-2) = 3(-2)^2 = 12$
$x = -1$	$m = f'(-1) = 3(-1)^2 = 3$
$x = 0$	$m = f'(0) = 3(0)^2 = 0$
$x = 1$	$m = f'(12) = 3(1)^2 = 3$
$x = 2$	$m = f'(2) = 3(2)^2 = 12$
$x = 3$	$m = f'(3) = 3(3)^2 = 27$
$\dots$	$\dots$

Anche per le derivate, trattandosi di un procedimento di limite, si potranno considerare separatamente il limite destro e il limite sinistro: si parlerà in questo caso di *derivata destra* e *derivata sinistra*.

Esempio 5.4. Sia data la funzione $f(x) = |x|$ e vediamo cosa succede per $x = 0$, separando il caso in cui per l'incremento h si ha $h < 0$ ($h \to 0^-$), da quello in cui si ha $h > 0$ ($h \to 0^+$):

$$\lim_{h\to 0^-} \frac{|0+h| - |0|}{h} = \lim_{h\to 0^-} \frac{|h|}{h} = \lim_{h\to 0^-} \frac{-h}{h} = -1\,,$$
$$\lim_{h\to 0^+} \frac{|0+h| - |0|}{h} = \lim_{h\to 0^+} \frac{|h|}{h} = \lim_{h\to 0^+} \frac{h}{h} = 1\,.$$

Questo risultato ha una evidente interpretazione grafica: la tangente sulla sinistra di 0 ha coefficiente angolare -1, sulla destra 1:

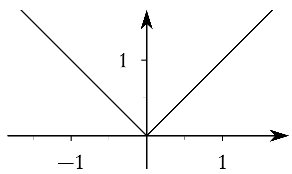


Figura 5.7 *La funzione valore assoluto*

In situazioni come quella appena vista si parla di *punto angoloso*: la derivata destra e quella sinistra sono entrambe finite, ma diverse.

La derivabilità di una funzione in un punto è legata alla continuità. Si dimostra infatti il seguente teorema:

Teorema 5.4. *Se una funzione è derivabile in un punto x_0 è anche continua in x_0.*

Non è vero il viceversa di questo teorema: una funzione può essere continua senza essere derivabile, come dimostra l'esempio della funzione valore assoluto.

Esempio 5.5. Nella definizione di derivata abbiamo chiesto che il limite del rapporto incrementale fosse finito: ci sono diversi motivi per fare questo (e la maggior parte di questi esulano dagli scopi di questo corso). Segnaliamo solo che si può presentare il caso che la tangente al grafico di una funzione può benissimo essere una retta verticale e, si sa, le rette verticali hanno la cattiva abitudine di non avere un coefficiente angolare. Come esempio consideriamo la funzione $f(x) = \sqrt[3]{x}$ e calcoliamo il limite del rapporto incrementale nell'origine.

$$\lim_{h\to 0} \frac{\sqrt[3]{0+h}-\sqrt[3]{0}}{h} = \lim_{h\to 0} \frac{\sqrt[3]{h}}{h} = \lim_{h\to 0} \frac{\sqrt[3]{h}}{\sqrt[3]{h^3}} = \lim_{h\to 0} \sqrt[3]{\frac{h}{h^3}} = \lim_{h\to 0} \frac{1}{\sqrt[3]{h^2}} = \frac{1}{0+} = +\infty$$

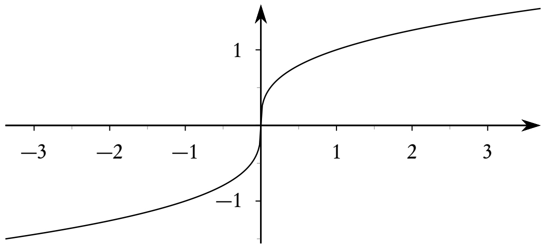


Figura 5.8 *La funzione* $f(x) = \sqrt[3]{x}$

Nella tabella 5.1 sono proposte, senza dimostrazione, le regole di derivazione più importanti, riguardanti la somma, il prodotto, ecc. di funzioni derivabili.

Funzione	***Derivata***
$k \cdot f(x)$	$k \cdot f'(x)$
$f(x)+g(x)$	$f'(x)+g'(x)$
$f(x)\cdot g(x)$	$f'(x)\cdot g(x)+f(x)\cdot g'(x)$
$\dfrac{1}{f(x)}$	$-\dfrac{f'(x)}{f^2(x)}$
$\dfrac{f(x)}{g(x)}$	$\dfrac{f'(x)\cdot g(x)-f(x)\cdot g'(x)}{g^2(x)}$
$f(g(x))$	$f'(g(x))\cdot g'(x)$

Tabella 5.1 *Regole di derivazione*

È opportuno evidenziare in dettaglio l'uso dell'ultima formula della tabella 5.1, che si riferisce alla regola di derivazione della composta di due funzioni. Come al solito ragioniamo su un esempio. Abbiamo già provato, in un esempio nella pagina 94 e nell'esempio successivo, che la derivata di $f(x) = \mathrm{e}^x$ è $f'(x) = \mathrm{e}^x$, e che la derivata di $g(x) = x^3$ è $g'(x) = 3x^2$. Se ora consideriamo la composta delle due funzioni

$$h(x) = f(g(x)) = \mathrm{e}^{g(x)} = \mathrm{e}^{x^3},$$

la derivata sarà
$$h'(x) = f'(g(x)) \cdot g'(x) = \mathrm{e}^{g(x)} \cdot g'(x) = \mathrm{e}^{x^3} 3x^2 = 3x^2 \mathrm{e}^{x^3}.$$

Nella tabella 5.2 proponiamo le regole per derivare le funzioni di uso più comune, ancora senza dimostrazione.

Funzione	***Derivata***
k	0
x^n, $n \in \mathbb{N} \setminus \{0\}$, $x \in \mathbb{R}$	nx^{n-1}
x^n, $n \in \mathbb{Z}$, $n < 0$, $x \in \mathbb{R} \setminus \{0\}$	nx^{n-1}
x^a, $a \in \mathbb{R}$, $x > 0$	ax^{a-1}
a^x	$a^x \cdot \ln a$
e^x	e^x
$\log_a \lvert x \rvert$	$\dfrac{1}{x} \cdot \log_a \mathrm{e} = \dfrac{1}{x \ln a}$
$\ln \lvert x \rvert$	$\dfrac{1}{x}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\operatorname{tg} x$	$1 + \operatorname{tg}^2 x = \dfrac{1}{\cos^2 x}$
$f^a(x)$	$nf^{a-1}(x) \cdot f'(x)$
$a^{f(x)}$	$a^{f(x)}(\ln a) f'(x)$
$\mathrm{e}^{f(x)}$	$\mathrm{e}^{f(x)} f'(x)$
$\log_a f(x)$	$\dfrac{f'(x)}{f(x)} \log_a \mathrm{e}$
$\ln f(x)$	$\dfrac{f'(x)}{f(x)}$
$\sin f(x)$	$\cos f(x) \cdot f'(x)$
$\cos f(x)$	$-\sin f(x) \cdot f'(x)$

Continua nella pagina successiva

Segue dalla pagina precedente

Funzione	***Derivata***
$\operatorname{tg} f(x)$	$\big(1+\operatorname{tg}^2 f(x)\big)f'(x)=\dfrac{f'(x)}{\cos^2 f(x)}$
$f(x)^{g(x)}=\mathrm{e}^{g(x)\ln f(x)}$	$\mathrm{e}^{g(x)\ln f(x)}\left(g'(x)\ln f(x)+g(x)\dfrac{f'(x)}{f(x)}\right)$

Tabella 5.2 *Derivate delle funzioni più comuni*

Si noti che la regola di derivazione delle potenze è sempre la stessa, quello che cambia è il dominio della funzione potenza. La regola che riguarda il caso in cui l'esponente sia un numero reale qualunque va bene anche per il caso dei radicali, e qui bisogna tenere conto che se l'indice è pari, la x deve essere maggiore di 0, se l'indice è dispari, la x può essere anche minore di 0. Esattamente come succede nel caso della radice cubica, se $x=0$ le funzioni radice non risultano derivabili[1]. Riportiamo in dettaglio le formule nei due casi che più ci interesseranno, cioè quello della radice quadrata e della radice cubica.

$$f(x)=\sqrt{x}\,,\quad f'(x)=\frac{1}{2\sqrt{x}}\quad ;\quad f(x)=\sqrt[3]{x}\,,\quad f'(x)=\frac{1}{3\sqrt[3]{x^2}}\,.$$

Esempio 5.6. Calcolare la derivata di

$$\sin(x^2+x)\,.$$

Si ha

$$\big(\sin(x^2+x)\big)'=(2x+1)\cos(x^2+x)\,.$$

Esempio 5.7. Calcolare la derivata di

$$(x^2+5)^{27}\,.$$

Si ha

$$\big((x^2+5)^{27}\big)'=27(x^2+5)^{26}(2x)=54x(x^2+5)^{26}\,.$$

Esempio 5.8. Calcolare la derivata di

$$\frac{x^2}{x^3+x^2-1}\,.$$

Si ha

$$\left(\frac{x^2}{x^3+x^2-1}\right)'=\frac{2x(x^3+x^2-1)-x^2(3x^2+2x)}{(x^3+x^2-1)^2}=\dots$$

Esempio 5.9. Calcolare la derivata di

$$\ln(\sin x^2)\,.$$

Si ha

$$\big(\ln(\sin x^2)\big)'=\frac{1}{\sin x^2}(\cos x^2)\,2x\,.$$

Si noti, in questo esempio, l'applicazione ripetuta della regola di derivazione delle funzioni composte.

[1] Si può osservare che il caso delle potenze è abbastanza complesso: prestare la massima attenzione!

5.3 La derivata della funzione inversa

Teorema 5.5 (Derivata della funzione inversa). *Sia f una funzione strettamente monotona definita in un intervallo e sia g la sua inversa. Se f è derivabile in un punto x_0 con derivata non nulla, posto $y_0 = f(x_0)$ e quindi $x_0 = g(y_0)$, la funzione inversa è derivabile in y_0 e si ha*

$$g'(y_0) = \frac{1}{f'(x_0)}. \tag{5.5}$$

Dimostrazione. Detto h l'incremento della variabile x e indicato con $k = f(x_0 + h) - f(x_0)$ il corrispondente incremento della variabile y, si ha

$$y_0 + k = f(x_0) + k = f(x_0 + h), \quad \text{da cui} \quad g(y_0 + k) = g(f(x_0 + h)) = x_0 + h = g(y_0) + h.$$

Osserviamo che, essendo f strettamente monotona, $k \neq 0$ se $h \neq 0$. Dunque

$$\frac{g(y_0 + k) - g(y_0)}{k} = \frac{h}{f(x_0 + h) - f(x_0)}.$$

Poiché la funzione inversa g è continua, si ha

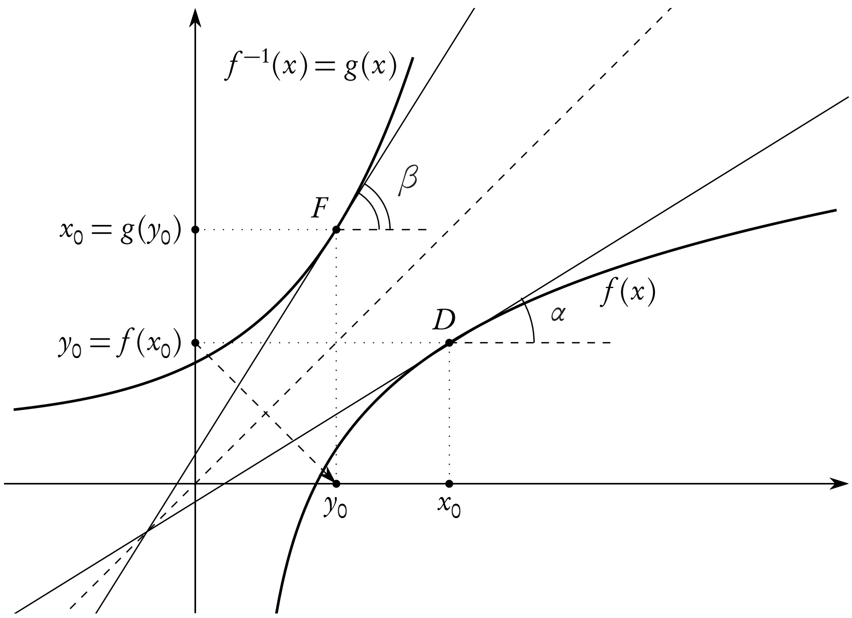
$$\lim_{h \to 0} h = \lim_{k \to 0} g(y_0 + k) - g(y_0) = 0.$$

Dunque

$$\lim_{k \to 0} \frac{g(y_0 + k) - g(y_0)}{k} = \lim_{h \to 0} \frac{h}{f(x_0 + h) - f(x_0)} = \frac{1}{f'(x_0)},$$

ovvero la tesi, essendo il primo membro proprio il limite del rapporto incrementale della funzione inversa g in y_0. □

Forniamo anche una significativa giustificazione per via grafica di questo teorema.



Il grafico proposto e le note proprietà di simmetria tra il grafico di una funzione e quello della sua inversa implicano anche che le tangenti al grafico di f e della sua inversa sono simmetrici rispetto alla bisettrice del primo e terzo quadrante. Tuttavia la tangente al grafico di f è calcolata nel punto x_0, quella al grafico dell'inversa nel punto y_0: ecco perché nella formula 5.5 compaiono due punti diversi, precisamente simmetrici rispetto alla bisettrice del primo e terzo quadrante.

Utilizziamo la formula 5.5 per ricavare le derivate di alcune funzioni elementari importanti. Cominciamo dalla funzione $f(x)=\ln x$. Come è noto la funzione $\ln x$ è l'inversa della funzione esponenziale $f(x)=e^x$. Poiché $D(e^x)=e^x$, posto $y=f(x)=e^x$, otteniamo $x=g(y)=\ln y$, da cui

$$g'(y_0)=\frac{1}{f'(x_0)}=\frac{1}{e^{x_0}}=\frac{1}{e^{\ln y_0}}=\frac{1}{y_0}.$$

Da qui ricaviamo la regola

$$f(x)=\ln x \quad \Rightarrow \quad f'(x)=\frac{1}{x}.$$

Sia ora $f(x)=\sin x$. La funzione non è biunivoca, e quindi non è invertibile, ma come sappiamo tale diventa se la restringiamo all'intervallo $[-\pi/2, \pi/2]$. L'inversa di questa restrizione è la funzione $\arcsin x$. Sapendo che $D(sin x)=\cos x$, posto $y=f(x)=\sin x$, otteniamo $x=g(y)=\arcsin y$, da cui

$$g'(y_0)=\frac{1}{f'(x_0)}=\frac{1}{\cos x_0}=\frac{1}{\cos\arcsin y_0}=\frac{1}{\sqrt{1-y_0^2}}.$$

Procedendo in maniera analoga per la funzione coseno (questa volta ristretta all'intervallo $[0,\pi]$) avremo $y=f(x)=\cos x$, da cui $x=g(y)=\arccos y$ e quindi

$$g'(y_0)=\frac{1}{f'(x_0)}=\frac{1}{-\sin x_0}=-\frac{1}{\sin\arccos y_0}=-\frac{1}{\sqrt{1-y_0^2}}.$$

Ne deduciamo le due regole

$$f(x)=\arcsin x \quad \Rightarrow \quad f'(x)=\frac{1}{\sqrt{1-x^2}} \qquad f(x)=\arccos x \quad \Rightarrow \quad f'(x)=-\frac{1}{\sqrt{1-x^2}}.$$

Si noti che entrambe le formule sono valide solo in $]-1,1[$, mentre per $x=\pm1$ le derivate diventano infinite, in accordo con il fatto che le tangenti al grafico di $\sin x$ e $\cos x$ sono orizzontali rispettivamente in $\pm\pi/2$ e in 0, π.

5.4 Derivate successive

Poiché la funzione derivata prima è a sua volta una funzione, ci si può chiedere se essa sia derivabile oppure no. Nei casi delle funzioni elementari che a noi interessano la risposta è affermativa e conduce al concetto di derivata seconda, terza, ecc., indicate con i simboli

$$f''(x)\Big(D^2(f(x))\Big), \quad f'''(x)\Big(D^3(f(x))\Big), f^{iv}(x)\Big(D^4(f(x))\Big), f^{(n)}(x)\Big(D^{(n)}(f(x))\Big)$$

Esempio 5.10. Calcolare la derivata 3ª di $f(x)=e^x$. Si ha, facilmente, $f'''(x)=e^x$.

Esempio 5.11. Calcolare le derivate prima, seconda, ecc., $(n+1)$-esima di $f(x)=x^n$. Si ha

$$f'(x)=nx^{n-1},\ f''(x)=n(n-1)x^{n-2},\ f'''(x)=n(n-1)(n-2)x^{n-3},$$
$$\ldots,\ f^{(n)}(x)=n(n-1)\cdots 1=n!,\ f^{n+1}(x)=0.$$

5.5 Polinomi di Taylor

Le funzioni polinomiali sono facilmente calcolabili in tutti i punti del loro dominio, in quanto coinvolgono solo operazioni elementari (somma e prodotto). Il calcolo esplicito delle altre funzioni è di solito impossibile "in termini esatti" e si ricorre ad opportune approssimazioni. Basta pensare, ad esempio, alla funzione $f(x)=\sqrt[4]{x}$: mentre è immediato che $\sqrt[4]{16}=32$, è lecito chiedersi come si possa calcolare $\sqrt[4]{17}$, ovvero qual è l'algoritmo che una calcolatrice (o un computer) usa per fornire un risultato del tipo

$$\sqrt[4]{17}\simeq 2.030543185\ldots$$

Stesso discorso per la quasi totalità delle altre funzioni, come seno, coseno, esponenziale, ecc.

Ci sono numerosi metodi per approssimare, con il dovuto grado di accuratezza, queste funzioni. Ci occuperemo qui dei *Polinomi di Taylor* che, anche se non costituiscono il metodo in assoluto più efficente, è quello più immediato da trattare, almeno per le funzioni che ci interessano. Inoltre i polinomi di Taylor sono teoricamente importanti, perché è proprio il loro uso che rende possibile la dimostrazione di un gran numero dei risultati dell'analisi, risultati che noi abbiamo dato senza giustificazione.

Definizione 5.6. *Data una funzione $f(x)$, definita in un intervallo e derivabile n volte in un punto x_0 dell'intervallo, si chiama* Polinomio di Taylor di ordine n della funzione f, relativo al punto[2] x_0, *e si indica con $T_{n,x_0}(x)$, il seguente polinomio*

$$T_{n,x_0}(x)=\frac{f(x_0)}{0!}+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \tag{5.6}$$

Il polinomio $T_{n,x_0}(x)$ ha le seguenti proprietà:

1. ha, in x_0, lo stesso valore della funzione: $T_{n,x_0}(x_0)=f(x_0)$;
2. ha, in x_0, la stessa derivata prima della funzione: $T'_{n,x_0}(x_0)=f'(x_0)$;
3. ha, in x_0, la stessa derivata seconda della funzione: $T''_{n,x_0}(x_0)=f''(x_0)$;
4. $\ldots$;
5. ha, in x_0, la stessa derivata n-esima della funzione: $T^{(n)}_{n,x_0}(x_0)=f^{(n)}(x_0)$;
6. è, *in un intorno di* x_0, il miglior polinomio di grado n approssimante $f(x)$.

Attenzione: le proprietà $1\ldots5$ elencate sopra valgono *solo ed esclusivamente* nel punto x_0, non negli altri punti (ma può succedere per certe funzioni che l'uguaglianza sia valida anche fuori dal punto x_0); l'ultima proprietà è valida in un intorno del punto x_0 e, purtroppo, non si può stabilire a priori quanto

[2] Su alcuni testi si usa il termine *Polinomio di Taylor* quando si considera un punto x_0 qualunque, *Polinomio di MacLaurin* quando il punto x_0 coincide con lo zero. I nomi sono legati ai matematici inglesi Brook Taylor (1685 — 1731) e Colin Maclaurin (1698 — 1746) che li introdussero. Noi useremo solo il primo nome.

grande sia questo intorno. Di solito all'aumentare dell'ordine l'approssimazione migliora sempre di più, ma questo non succede sempre: ci sono addirittura degli esempi in cui all'aumentare del grado l'approssimazione peggiora, altre situazioni in cui i polinomi di Taylor forniscono un'approssimazione che non ha alcun interesse applicativo.

Si noti che,

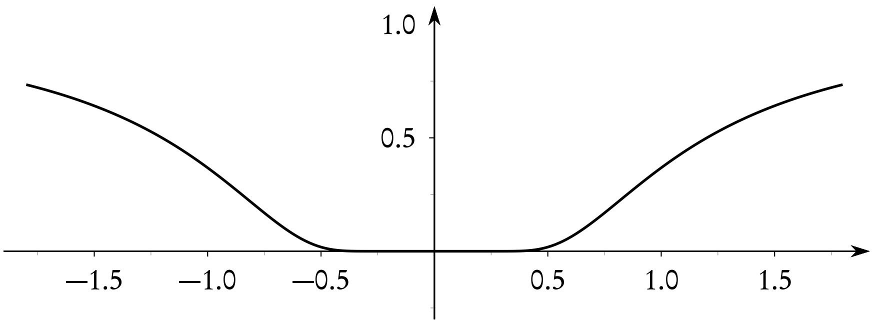
1. se $n=0$, il polinomio è di grado zero (cioè è una costante) e il suo grafico è la retta orizzontale passante per $\big(x_0, f(x_0)\big)$;
2. se $n=1$, il polinomio è di grado 1 e il suo grafico è quello della retta tangente al grafico della funzione nel punto $\big(x_0, f(x_0)\big)$;
3. se $n=2$, il polinomio è di grado 2 e il suo grafico è quello della parabola con asse verticale tangente al grafico della funzione nel punto $\big(x_0, f(x_0)\big)$.

Il primo problema che si presenta nella scrittura di un efficiente polinomio di Taylor è quello di scegliere opportunamente il punto x_0: occorre infatti che in x_0 siano calcolabili facilmente sia la funzione che le sue derivate. Molto spesso, in particolare nei casi di nostro interesse, questo avviene per il punto 0.

Esempio 5.12. Cominciamo con il proporre un esempio in cui le cose non vanno molto bene. Sia data la funzione

$$f(x)=\begin{cases} e^{-1/x^2}, & \text{se } x\neq 0 \\ 0, & \text{se } x=0 \end{cases} .$$

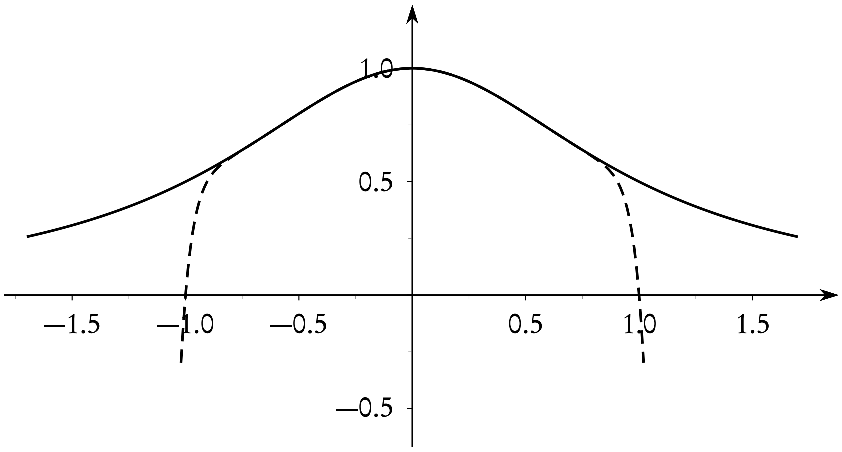
Si può provare (con un po' di pazienza!) che la funzione è derivabile in tutto $\mathbb{R}$ e che tutte le derivate, di qualsiasi ordine, sono nulle nell'origine. Se ne deduce che il polinomio di Taylor di ordine qualsiasi, di punto iniziale 0, è sempre il polinomio identicamente nullo ed è evidente che il polinomio identicamente nullo *non* costituisce una buona approssimazione "globale" della funzione. Il grafico che segue mostra solo che l'approssimazione è "ad occhio" accettabile nei pressi dell'origine (il polinomio di Taylor non è rappresentato in quanto coincide con l'asse x), ma ovviamente approssimare questa funzione, anche in un intorno piccolo dell'origine, con zero, è di poca utilità pratica.



Esempio 5.13. Come secondo esempio "patologico" proponiamo una situazione in cui all'aumentare del grado l'intorno in cui l'approssimazione vale non aumenta, se non in maniera insignificante. Il grafico che segue si riferisce alla funzione

$$f(x)=\frac{1}{1+x^2}$$

(con linea continua) e al suo polinomio di Taylor (in tratteggio) di ordine 22 nell'origine.



Per quanto si aumenti l'ordine, il polinomio non riesce a fornire un'approssimazione accettabile fuori dall'intervallo $[-1,1]$. Come utile esercizio si può provare che il polinomio di Taylor di ordine 8 e di punto iniziale 0 della funzione appena considerata è

$$T_{8,0}(x)=1-x^2+x^4-x^6+x^8\,.$$

Naturalmente dobbiamo fornire anche esempi in cui le cose vanno per il verso giusto: lo facciamo negli esempi che seguono, relativi alla funzione e^x e alla funzione $\sin(x)$: con tratto continuo la funzione, in tratteggio i vari polinomi di Taylor.

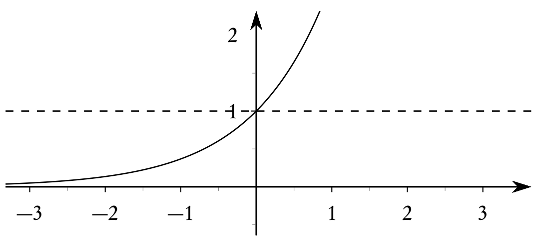


Figura 5.9 e^x *e* $T_{0,0}(x)$

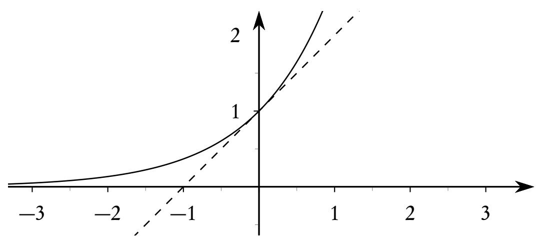


Figura 5.10 e^x *e* $T_{1,0}(x)$

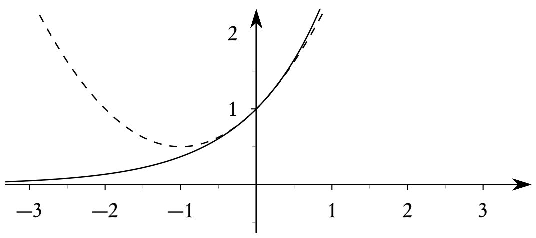


Figura 5.11 e^x *e* $T_{2,0}(x)$

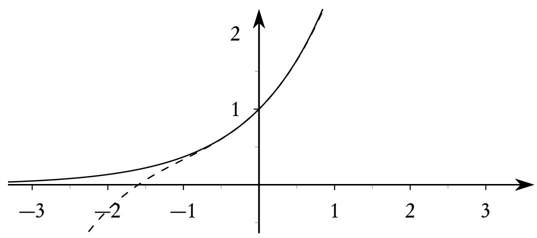


Figura 5.12 e^x *e* $T_{3,0}(x)$

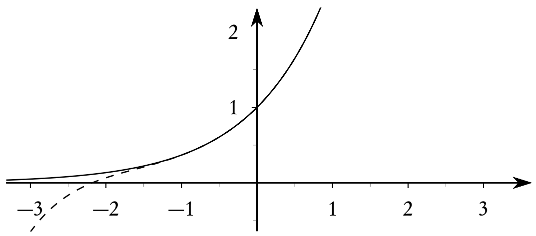


Figura 5.13 e^x *e* $T_{5,0}(x)$

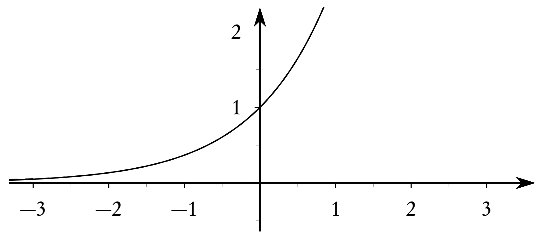


Figura 5.14 e^x *e* $T_{10,0}(x)$

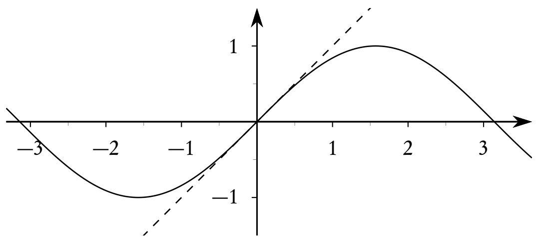


Figura 5.15 $\sin x$ *e* $T_{1,0}(x)$

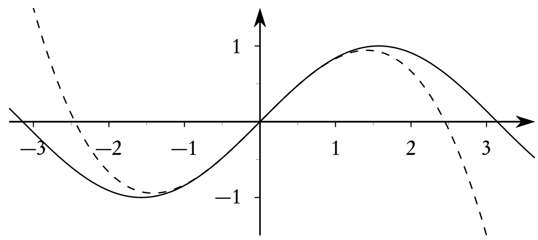


Figura 5.16 $\sin x$ *e* $T_{3,0}(x)$

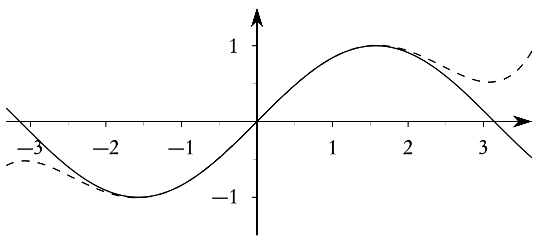


Figura 5.17 $\sin x$ *e* $T_{5,0}(x)$

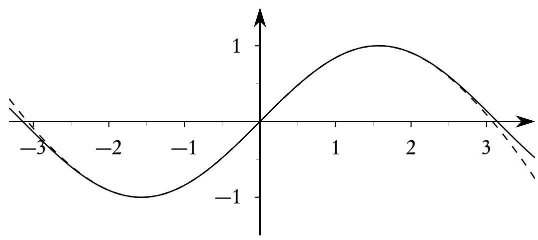


Figura 5.18 $\sin x$ *e* $T_{7,0}(x)$

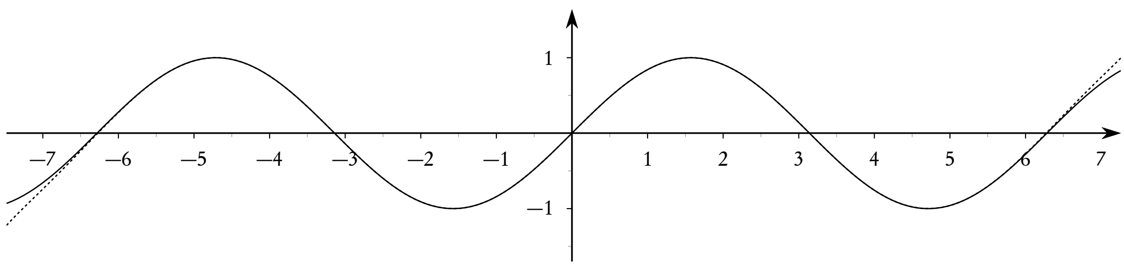


Figura 5.19 $\sin x$ *e* $T_{17,0}(x)$

Per concludere ritorniamo brevemente al problema che avevamo proposto all'inizio di questo paragrafo, e cioè il calcolo della radice quarta di 17. Consideriamo la funzione $f(x)=\sqrt[4]{x}$ e le sue prime quattro derivate, che calcoliamo nel punto $x_0=16$, che vogliamo usare come punto iniziale (o punto base) per la nostra formula di Taylor di ordine 4. Si ha quanto segue.

1. $f(x)=\sqrt[4]{x} \quad \Rightarrow \quad f(16)=2.$

2. $f'(x) = \dfrac{1}{\sqrt[4]{x^3}} \quad \Rightarrow \quad f'(16) = \dfrac{1}{32}.$

3. $f''(x) = -\dfrac{3}{16\sqrt[4]{x^7}} \quad \Rightarrow \quad f''(16) = -\dfrac{3}{2048}.$

4. $f'''(x) = \dfrac{21}{64\sqrt[4]{x^{11}}} \quad \Rightarrow \quad f'''(16) = \dfrac{21}{131\,072}.$

5. $f^{(4)}(x) = -\dfrac{231}{256\sqrt[4]{x^{15}}} \quad \Rightarrow \quad f^{(4)}(16) = -\dfrac{231}{8\,388\,608}.$

Si noti che abbiamo scelto come punto base 16 perché è "vicino" a 17, e perché tutte le derivate sono calcolabili esattamente in 16. Con questi valori possiamo scrivere il polinomio di Taylor di ordine 4, relativo al punto base 16:

$$T_{4,16}(x) = 2 + \frac{1}{32}(x-16+x) - \frac{3}{4096}(x-16)^2 + \frac{7}{262144}(x-16)^3 - \frac{77}{67108864}(x-16)^4$$

Si trova, naturalmente, $T_{4,16}(16)$=2, e poi, ed è questo che ci interessava,

$$T_{4,16}(17) \simeq 2.030543134\,,$$

risultato che coincide fino alla sesta cifra dopo la virgola con quello già scritto precedentemente: se avessimo preso un polinomio di grado più elevato, avremmo avuto un'approssimazione ancora migliore (anche se non è banale valutare quale sia il grado di approssimazione).

È molto importante osservare che questa approssimazione funziona bene "vicino a 16": se per esempio la usiamo per calcolare $\sqrt[4]{1}$, che vale ovviamente 1, otteniamo $\sqrt[4]{1} \simeq 1.218246266$, risultato chiaramente inaccettabile. Il tutto è confermato dal grafico della funzione $\sqrt[4]{x}$ e di $T_{4,16}(x)$: i due grafici sono quasi coincidenti nei pressi di 16, sono invece sensibilmente diversi nei pressi di 1.

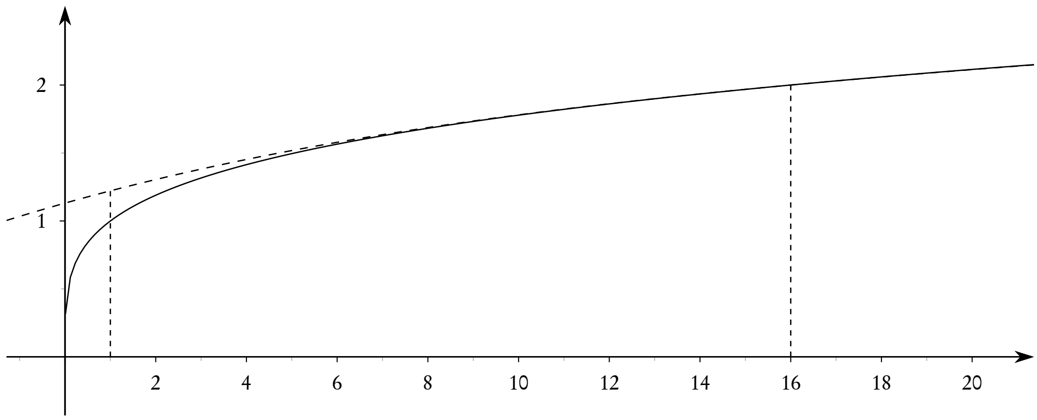


Figura 5.20 $\sqrt[4]{x}$ *e* $T_{4,16}(x)$

Lasciamo ai volenterosi di verificare che, per la stessa funzione,

$$T_{4,1}(x) = 1 + \frac{1}{4}(x-1) - \frac{3}{32}(x-1)^2 + \frac{7}{128}(x-1)^3 - \frac{77}{2048}(x-1)^4\,,$$

il cui grafico è rappresentato nella figura 5.21

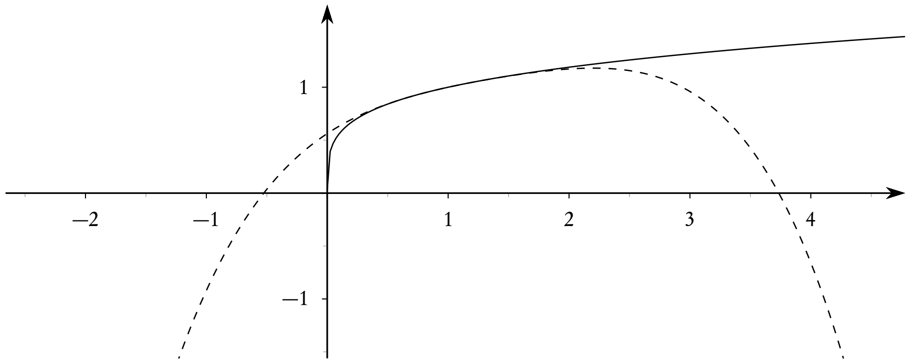


Figura 5.21 $\sqrt[4]{x}$ e $T_{4,1}(x)$

Ovviamente questa approssimazione funzione molto meglio della precedente vicino a 1, ma molto peggio vicino a 16; addirittura se la usassimo per calcolare $\sqrt[4]{16}$ otterremmo

$$\sqrt[4]{16} \simeq -1735.15478515625: \quad \text{un vero disastro!}$$

5.6 Polinomi di Taylor di alcune funzioni elementari

Per alcune funzioni elementari i polinomi di Taylor si possono scrivere per un ordine n qualunque (addirittura per $n \to +\infty$, ottenendo le serie di Taylor, di cui comunque non ci occupiamo in questo corso). Riportiamo le formule relative alle funzioni di uso più comune, dove abbiamo scelto come valore x_0 il numero 0, perché in questa situazione le formule diventano particolarmente semplici.

$\sin x$	$x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} + \dots + (-1)^{n-1}\dfrac{x^{2n-1}}{(2n-1)!}$
$\cos x$	$1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} + \dots + (-1)^{n}\dfrac{x^{2n}}{(2n)!}$
e^x	$1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \dots + \dfrac{x^n}{n!}$
$\ln(1+x)$	$x - \dfrac{x^2}{2} + \dfrac{x^3}{3} + \dots + (-1)^{n-1}\dfrac{x^n}{n}$

5.7 Esercizi

Esercizio 5.1. *Calcolare le derivate delle seguenti funzioni.*

1. $\sqrt{x} + 2x^5$;
2. $x^3 - 3x^2 - 2$;

3. $\ln x - x^4 + 4x^2 - x$;
4. $e^x - 2 - x^6$;
5. $e^x \cdot x^5$;
6. $-\sqrt{x}\left(x^3 - x^2 + x\right)$;
7. $e^x \ln x$;
8. $\ln(x)e^x\left(x^2 + x + 1\right)$;
9. $\dfrac{x-1}{x+3}$;
10. $\dfrac{x-x^2}{\ln x + x}$;
11. $\dfrac{x}{2^x + \sqrt{x}}$.

Esercizio 5.2. *Calcolare le derivate prime e seconde delle seguenti funzioni. Attenzione: in alcuni casi compaiono diverse lettere nelle funzioni qui di seguito; solo una però (quella esplicitamente indicata) è la variabile, le altre sono semplicemente dei parametri che vanno dunque trattati come costanti.*

1. $f(x) = \sqrt{1-4x} + 2x^5$;
2. $f(x) = x^3 - 3x^2 - 2$;
3. $f(x) = \ln(x^4 + 4x^2 - x)$;
4. $f(x) = e^{-x-2} - x^6$;
5. $f(x) = e^{5x}\, x^5$;
6. $f(x) = -\sqrt{x^3 - x^2 + x}$;
7. $f(x) = e^{x\ln(2x)}$;
8. $f(x) = \ln\left(e^3 x\left(x^2 + x + 1\right)\right)$;
9. $f(x) = x^2 k + y\ln(x^2 - hx)$;
10. $f(x) = e^{hx-k}$;
11. $f(y) = ky^3 + 2y^2 - hy$;
12. $f(y) = \sin(ky^2)$;
13. $f(x) = \sin(kx^2 - yx)$;
14. $f(y) = xy - x^2y^2$;
15. $f(y) = y\cos(ky) + xy$;

16. $f(z) = xyz^2$;
17. $f(z) = xz^3 + yz^2$;
18. $f(y) = x^3y - z^4 \ln\left(y^3 + k\right)$;
19. $f(x) = \mathrm{e}^{yx} + zx$.

Esercizio 5.3. *Per le seguenti funzioni determinare i polinomi di Taylor indicati.*

1. $f(x) = x^2\,; \quad T_{2,0}(x),\ T_{3,0}(x),\ T_{2,1}(x)$.
2. $f(x) = \ln x\,; \quad T_{2,1}(x),\ T_{3,1}(x),\ T_{2,\mathrm{e}^2}(x)$.
3. $f(x) = 2^x\,; \quad T_{3,0}(x),\ T_{3,1}(x)$.
4. $f(x) = x^2 + x^4 - 1\,; \quad T_{3,0}(x),\ T_{4,0}(x)$.

Esercizio 5.4. *Per le seguenti funzioni determinare i polinomi di Taylor indicati. Valutare se sarebbe stato possibile migliorare la scelta del punto base rispetto a quello indicato.*

1. $f(x) = \ln x\,; \quad T_{2,\mathrm{e}}(x)$.
2. $f(x) = \ln(x^2 + x)\,; \quad T_{3,1}(x)$.
3. $f(x) = x\mathrm{e}^x\,; \quad T_{1,e}(x)$.
4. $f(x) = x + \mathrm{e}^x\,; \quad T_{2,0}(x)$.
5. $f(x) = \sin x\,; \quad T_{3,1}(x)$.
6. $f(x) = \sin(x^2 + x)\,; \quad T_{2,0}(x)$.
7. $f(x) = x^2 + x\,; \quad T_{3,e}(x)$.
8. $f(x) = \sqrt{x^2 + 1}\,; \quad T_{2,0}(x)$.

6 Grafici di funzioni di una variabile

L'introduzione del concetto di derivata si rivela un importante successo per risolvere il problema di studiare le proprietà delle funzioni, fino a giungere al tracciamento di un grafico significativo. La parte dell'analisi che studia le proprietà delle funzioni che si possono ricavare sulla base delle loro derivate si chiama *calcolo differenziale*. Tale importante settore della matematica si basa su alcuni teoremi classici, di alcuni dei quali (quelli più importanti ai nostri fini) ci limiteremo a fornire gli enunciati e una giustificazione grafica. Si tratta dei cosiddetti *teoremi fondamentali del calcolo differenziale*: i teoremi di Rolle, Lagrange e Cauchy. Conseguenza di questi teoremi è la famosa *regola di l'Hôpital* per il calcolo di limiti in forma indeterminata.

6.1 I teoremi fondamentali del calcolo differenziale

Teorema 6.1 (Teorema di Lagrange). *Sia f una funzione definita in un intervallo chiuso $[a,b]$ e avente le seguenti proprietà:*

1. *f è continua in $[a,b]$ (compresi gli estremi!);*
2. *f è derivabile almeno in $]a,b[$ (potrebbe non essere derivabile negli estremi, per esempio potrebbe avere derivata infinita negli estremi).*

Allora esiste almeno un punto c interno *ad $[a,b]$ dove per la derivata prima della funzione si ha*

$$f'(c)=\frac{f(b)-f(a)}{b-a}.$$

Siccome il numero $(f(b)-f(a))/(b-a)$ è il coefficiente angolare della retta passante per i punti $A=(a,f(a))$ e $B=(b,f(b))$, mentre $f'(c)$ è il coefficiente angolare della tangente al grafico in un punto interno c, il teorema precedente si può interpretare geometricamente dicendo che esiste un punto interno al dominio dove la tangente è parallela alla secante passante per gli estremi. Si può vedere il grafico della figura 6.1 per rendersi ancora meglio conto del senso di questa interpretazione.

Dimostrazione. Si faccia riferimento alla figura 6.1.

Consideriamo i punti $A(a,f(a))$ e $B(b,f(b))$. La retta che passa per A e B ha coefficiente angolare

$$\frac{f(b)-f(a)}{b-a}$$

ed equazione

$$y=f(a)+\frac{f(b)-f(a)}{b-a}(x-a)\quad \big(=h(x)\big).$$

Consideriamo ora la funzione $g(x)=f(x)-h(x)$; essa ha le stesse proprietà di regolarità della funzione f e, in più, è tale che $g(a)=g(b)=0$. Alla funzione g possiamo applicare il teorema di

Weierstrass: esisteranno dunque due punti c e d in uno dei quali la funzione assume il suo massimo, mentre nell'altro assume il suo minimo. Se entrambi questi punti coincidessero con gli estremi di $[a,b]$, allora la funzione sarebbe costante in $[a,b]$ (con valore nullo della costante) e avrei $g'(x)=0$ su tutto $[a,b]$; se invece almeno uno dei due punti, diciamo c, è interno ad $]a,b[$, allora in esso deve essere $g'(c)=0$. In ogni caso in almeno un punto c di $]a,b[$ deve essere $g'(c)=0$. Poiché

$$g'(x)=f'(x)-\frac{f(b)-f(a)}{b-a},$$

si conclude subito con la tesi. □

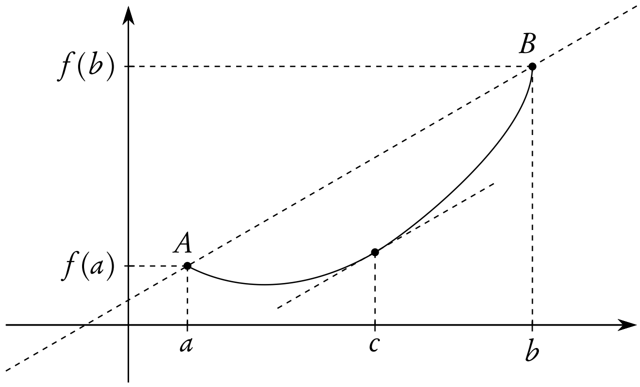


Figura 6.1 *Interpretazione geometrica del Teorema di Lagrange*

Esempio 6.1. Verifichiamo che la funzione $f(x)=x^2+x$ verifica, nell'intervallo $[-1,2]$, il Teorema di Lagrange e determiniamo il, oppure i, "punti di Lagrange".

La verifica delle ipotesi è immediata. Troviamo i punti di Lagrange. Si ha

$$f'(c)=2c+1;$$

mentre

$$\frac{f(b)-f(a)}{b-a}=\frac{f(2)-f(-1)}{2-(-1)}=\frac{(4+2)-(1-1)}{3}=2.$$

Deve dunque essere

$$2c+1=2 \quad\Rightarrow\quad c=\frac{1}{2},$$

che è interno all'intervallo del dominio. Si veda la figura seguente, dove abbiamo usato diverse unità di misura sui due assi.

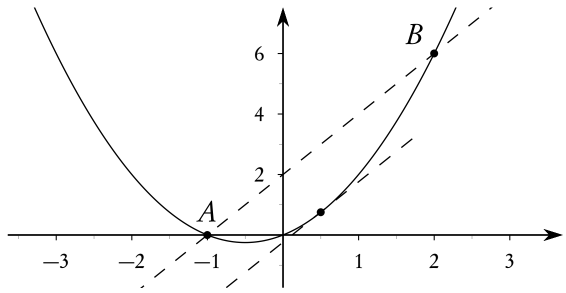


Figura 6.2 *Un esempio di applicazione del Teorema di Lagrange*

Teorema 6.2 (Teorema di Rolle)**.** *Sia f una funzione definita in un intervallo chiuso $[a,b]$ e avente le seguenti proprietà:*

1. *f è continua in $[a,b]$ (compresi gli estremi!);*
2. *f è derivabile almeno in $]a,b[$ (potrebbe non essere derivabile negli estremi, per esempio potrebbe avere derivata infinita negli estremi);*
3. *$f(a)=f(b)$ (le "quote" iniziale e finale del grafico sono identiche).*

Allora esiste almeno un punto c interno *ad $[a,b]$ dove la derivata prima della funzione si annulla.*

Dimostrazione. Il teorema di Rolle è un caso particolare del teorema di Lagrange in quanto la funzione f ha le stesse caratteristiche richieste per la validità del teorema di Lagrange e, in più, assume valori uguali agli estremi: da questo segue che la retta passante per i punti $A(a,f(a))$ e $B(b,f(b))$ ha coefficiente angolare nullo e questa è appunto la tesi del teorema di Rolle. □

Le figure che seguono danno una giustificazione grafica di questo risultato, se si tiene conto che avere derivata nulla significa avere tangente orizzontale. Il punto (o i punti di cui parla il teorema si possono chiamare "*punti di Rolle*")

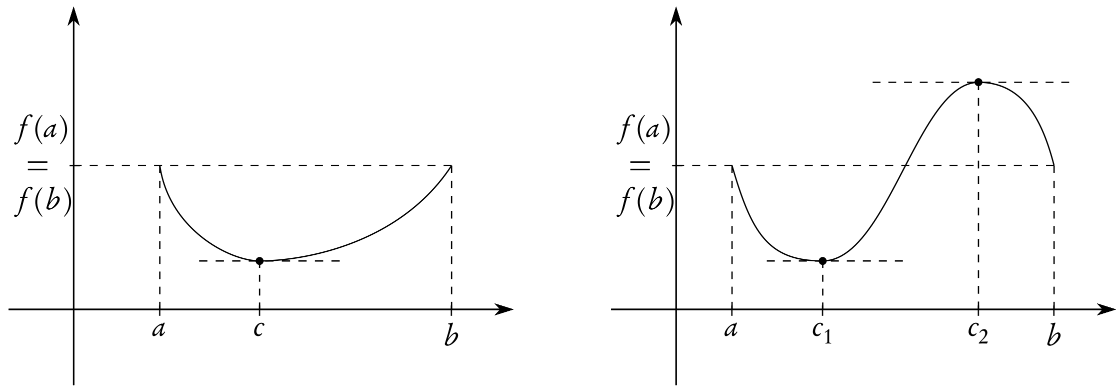


Figura 6.3 *Esempi di funzioni con uno e due "punti di Rolle"*

Approfittiamo di questo teorema per fare delle osservazioni su che cosa significhi *condizione sufficiente*. Le tre condizioni presenti nell'enunciato di questo teorema sono delle condizioni sufficienti per la validità dello stesso. Infatti se anche una o più di queste condizioni manca, non si può concludere con la tesi, come dimostrano i tre grafici che seguono: nel primo manca solo l'ipotesi di continuità in tutto $[a,b]$, nel secondo manca solo l'ipotesi di derivabilità all'interno di $[a,b]$, nel terzo manca solo l'ipotesi che le quote agli estremi siano uguali. In tutti e tre i casi non esiste alcun punto di Rolle.

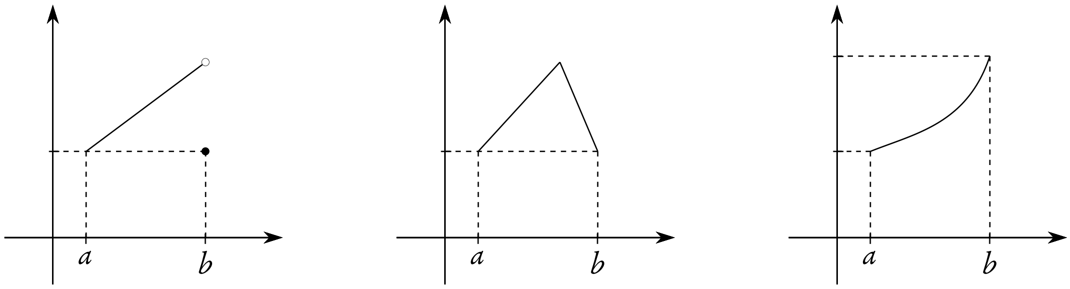


Figura 6.4 *Tre esempi di non applicabilità del Teorema di Rolle*

Tutto questo non significa affatto che se mancano una o più delle tre condizioni presenti nell'enunciato del teorema non esiste alcun punto di Rolle, come mostra il seguente grafico, in cui mancano addirittura tutte e tre le condizioni, ma dove esistono addirittura due punti in cui il grafico ha tangente orizzontale.

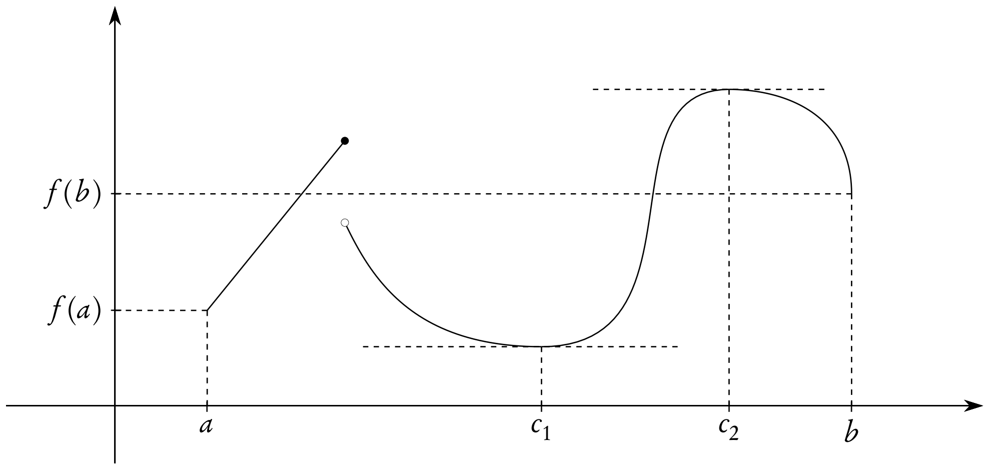


Figura 6.5 *Un esempio in cui esistono* punti di Rolle, nonostante le ipotesi non siano verificate

Il motivo di questo comportamento è da ricercarsi nel fatto che le condizioni per la validità del Teorema di Rolle *non sono necessarie*: anche se mancano, la tesi può essere ugualmente vera.

Esempio 6.2. Se una funzione soddisfa le ipotesi del teorema, almeno un punto di Rolle esiste sicuramente. Per trovarlo analiticamente basterà risolvere l'equazione, nell'incognita c, $f'(c)=0$.

Consideriamo la funzione, avente dominio l'intervallo $[0,1]$,

$$f(x)=5x^3-5x^2+1\,.$$

Essa soddisfa chiaramente le ipotesi del teorema di Rolle, per cui deve esistere almeno un punto c, *interno* al dominio, dove $f'(c)=0$. Per trovarlo basterà risolvere l'equazione

$$15c^2-10c=0 \quad\Rightarrow\quad c=0 \vee c=\frac{2}{3}.$$

Solo il punto $c={}^2/_3$ va bene, perché l'altro è situato proprio sulla frontiera del dominio.

Una generalizzazione del teorema di Lagrange è dovuta a Cauchy: riportiamo l'enunciato di questo teorema, per completezza.

Teorema 6.3 (Teorema di Cauchy)**.** *Siano f e g due funzioni continue in $[a,b]$ e derivabili almeno nei punti interni di $[a,b]$ e con $g'(x)\neq 0$ in tutto $]a,b[$. Allora esiste almeno un punto c interno ad $[a,b]$ tale che*

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(c)}{g'(c)}. \tag{6.1}$$

È ovvio che il teorema di Lagrange discende da questo teorema: basta porre $g(x)=x$.

Per le applicazioni sono importantissimi i seguenti tre corollari del teorema di Lagrange.

Primo corollario Se f è una funzione definita e continua in un intervallo $[a,b]$ e ha derivata >0 in $]a,b[$, allora f è *crescente* in $[a,b]$; se ha derivata <0 è invece *decrescente*. Per dimostrarlo basta osservare che se prendo due punti x_1 e x_2, con $x_1<x_2$, si ha

$$\frac{f(x_2)-f(x_1)}{x_2-x_1}=f'(c)>0 \quad\Rightarrow\quad f(x_2)>f(x_1),$$

il contrario se la derivata è negativa.

Secondo corollario Se f è una funzione definita e continua in un intervallo $[a,b]$ e ha derivata $=0$ in $]a,b[$, è costante in $[a,b]$. Per dimostrarlo basta prendere un punto x qualunque di $[a,b]$ e osservare che si ha

$$\frac{f(x)-f(a)}{x-a}=f'(c)=0 \quad\Rightarrow\quad f(x)=f(a),$$

ovvero che $f(x)$ si trova sempre alla stessa quota di $f(a)$.

Terzo corollario Se f e g sono due funzioni definite e continue in un intervallo $[a,b]$ e con la stessa derivata in $]a,b[$, allora la funzione $f-g$ è costante in $[a,b]$. Per dimostrarlo basta osservare che $f-g$ ha derivata nulla in $]a,b[$.

Teorema 6.4 (Teorema di l'Hôpital)**.** *Siano date due funzioni f e g definite e continue in un intorno di un punto c (eventualmente anche $\pm\infty$), derivabili almeno nei punti diversi da c, con $g'(x)$ sempre diversa da zero. Sia inoltre*

1. $\lim_{x\to c} f(x)=\lim_{x\to c} g(x)=0$,

oppure

2. $\lim_{x\to c} f(x)=\lim_{x\to c} g(x)=\infty$ *(con qualunque segno).*

Allora per calcolare il

$$\lim_{x \to c} \frac{f(x)}{g(x)}$$

che si presenta nella forma indeterminata $0/0$ *oppure* ∞/∞, *si può provare a calcolare il*

$$\lim_{x \to c} \frac{f'(x)}{g'(x)}.$$

Se quest'ultimo limite esiste (finito o no), esiste anche quello del rapporto delle due funzioni e i due limiti sono identici.

Esempio 6.3. Il

$$\lim_{x \to +\infty} \frac{e^x}{x}$$

si presenta nella forma indeterminata $+\infty/+\infty$. Si ha poi

$$\lim_{x \to +\infty} \frac{f'(x)}{g'(x)} = \lim_{x \to +\infty} \frac{e^x}{1} = +\infty$$

e quindi anche

$$\lim_{x \to +\infty} \frac{e^x}{x} = +\infty .$$

Si usa abbreviare questo procedimento nel modo seguente:

$$\lim_{x \to +\infty} \frac{e^x}{x} \overset{(H)}{=} \lim_{x \to +\infty} \frac{e^x}{1} = +\infty .$$

A questo punto si può "cancellare la (H) sopra l'uguale" e concludere che il limite cercato vale $+\infty$.

La regola di l'Hôpital si può applicare anche più volte in successione.

Esempio 6.4.

$$\lim_{x \to +\infty} \frac{e^x}{x^2} \overset{\cancel{(H)}}{=} \lim_{x \to +\infty} \frac{e^x}{2x} \overset{\cancel{(H)}}{=} \lim_{x \to +\infty} \frac{e^x}{2} = +\infty .$$

Esempio 6.5.

$$\lim_{x \to 0} \frac{\sin x - x}{x^3} \overset{\cancel{(H)}}{=} \lim_{x \to 0} \frac{\cos x - 1}{3x^2} \overset{\cancel{(H)}}{=} \lim_{x \to 0} \frac{-\sin x}{6x} = \lim_{x \to 0} \frac{-1}{6} \frac{\sin x}{x} = -\frac{1}{6} .$$

La regola di l'Hôpital non sempre funziona in maniera così semplice, come prova l'esempio che segue.

Esempio 6.6.

$$\lim_{x \to +\infty} \frac{e^{2x}}{e^x} \overset{(H)}{=} \lim_{x \to +\infty} \frac{2e^{2x}}{e^x} \overset{(H)}{=} \lim_{x \to +\infty} \frac{4e^{2x}}{e^x} \overset{(H)}{=}$$

ed è chiaro che proseguendo in questo modo non si ottiene nulla. Se però si osserva che

$$\frac{e^{2x}}{e^x} = e^{2x-x} = e^x ,$$

si conclude subito che...

Tuttavia non insistiamo oltre su questo tipo di difficoltà.

6.2 Massimi e minimi per una funzione

Abbiamo già dato la definizione di massimo e minimo relativo per una funzione (vedi le definizioni 1.21 e 1.22 del capitolo 1). La richiamiamo qui adattandola specificamente al caso che ci interessa di funzioni di una variabile.

Definizione 6.5. *Sia data una funzione f, definita in un insieme D. Un punto $x_0 \in D$ si dice* punto di massimo relativo *se esiste un intorno di x_0 tale che per tutti i punti dell'intorno si abbia che*

$$f(x) \leq f(x_0);$$

$x_0 \in D$ si dice invece punto di minimo relativo *se esiste un intorno di x_0 tale che per tutti i punti dell'intorno si abbia che*

$$f(x) \geq f(x_0).$$

Se le disuguaglianze valgono in senso stretto (senza gli uguali), allora i punti si chiamano di massimo o minimo relativo proprio.

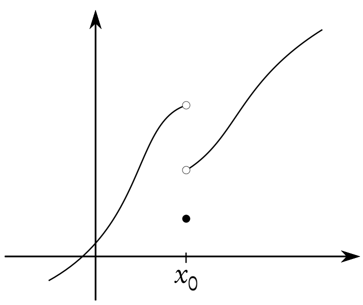
Il valore $f(x_0)$ si dice un (valore) massimo o minimo relativo *per la funzione.*

Se le disuguaglianze considerate valgono in tutto il dominio, si parla di punto di massimo, o minimo, assoluto *e di (valore) massimo o minimo* assoluto

La ricerca dei massimi e minimi relativi o assoluti per una funzione riveste grande importanza nelle applicazioni. Siamo particolarmente interessati a questa ricerca nel caso di funzioni derivabili, definite in un intervallo I. In questo caso valgono i seguenti risultati, che sono sostanzialmente delle conseguenze dei teoremi fondamentali che abbiamo considerato.

1. Se un punto x_0 è, per una funzione f, di massimo o minimo relativo *interno* ad I, allora $f'(x_0) = 0$.
2. Se una funzione è crescente a sinistra di x_0 e decrescente a destra di x_0, x_0 è di massimo relativo.
3. Se una funzione è decrescente a sinistra di x_0 e crescente a destra di x_0, x_0 è di minimo relativo.

L'esperienza ci aiuterà a trattare anche qualche caso più complesso di funzioni che non siano derivabili e/o continue in qualche punto. Segnaliamo comunque che in casi come questi possono anche presentarsi situazioni poco intuitive, come per esempio che una funzione sia crescente sia a sinistra che a destra di un punto e che in quel punto ci sia un minimo, o un massimo. Si vedano gli esempi grafici che seguono.



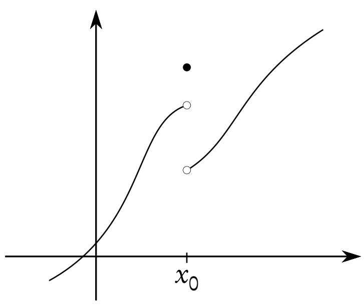


Figura 6.6 *Funzioni crescenti a sinistra e a destra di un punto, con minimo o massimo nel punto*

Sulla scorta dei due esempi proposti, il lettore è invitato provare a costruire graficamente altre situazioni "patologiche".

In situazioni standard potremo procedere come nell'esempio che segue.

Esempio 6.7. Per la funzione $f(x)=x^3-3x^2$ si ha $f'(x)=3x^2-6x$, da cui $f'(x)>0$ se $x<0$ oppure $x>2$, $f'(x)<0$ se $0<x<2$, $f'(x)=0$ se $x=0$ oppure $x=2$. Riporteremo questi risultati in un grafico come il seguente.

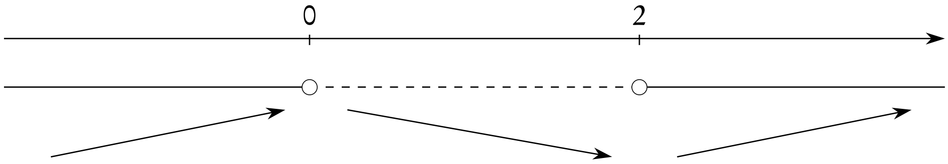


Figura 6.7 *Crescenza e decrescenza di una funzione*

Se teniamo conto che

$$\lim_{x\to+\infty} f(x)=+\infty\,, \quad \lim_{x\to-\infty} f(x)=-\infty\,, \quad f(0)=0\,, \quad f(2)=-4\,,$$

possiamo ben renderci conto dell'andamento grafico della funzione stessa:

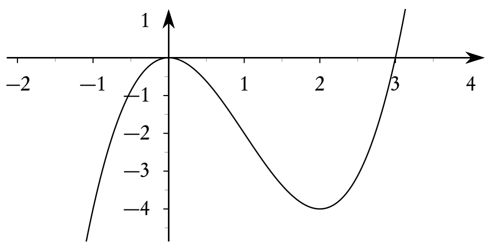


Figura 6.8 *Grafico della funzione* $f(x)=x^3-3x^2$

Un risultato di grande importanza per la ricerca dei massimi e minimi assoluti è costituito dal seguente teorema.

Teorema 6.6 (Teorema di Weierstrass). *Se $f(x)$ è una funzione definita e* continua *in un insieme chiuso e limitato, allora il massimo assoluto e il minimo assoluto esistono sicuramente.*

Questo teorema è importante perché se siamo sicuri che il massimo e minimo ci sono, gli sforzi per trovarli saranno sicuramente giustificati.

6.3 Funzioni convesse e concave

Definizione 6.7. *Una funzione f si dice* convessa *in un intervallo I se presi comunque due punti x_1 e x_2 di I e considerato il segmento di estremi $P_1=(x_1,f(x_1))$ e $P_2=(x_2,f(x_2))$, la parte del grafico di f coorrispondente all'intervallo $[x_1,x_2]$ sta tutta al di sotto di questo segmento.*

Definizione 6.8. *Una funzione f si dice* concava *in un intervallo I se presi comunque due punti x_1 e x_2 di I e considerato il segmento di estremi $P_1 = (x_1, f(x_1))$ e $P_2 = (x_2, f(x_2))$, la parte del grafico di f coorrispondente all'intervallo $[x_1, x_2]$ sta tutta al di sopra di questo segmento.*

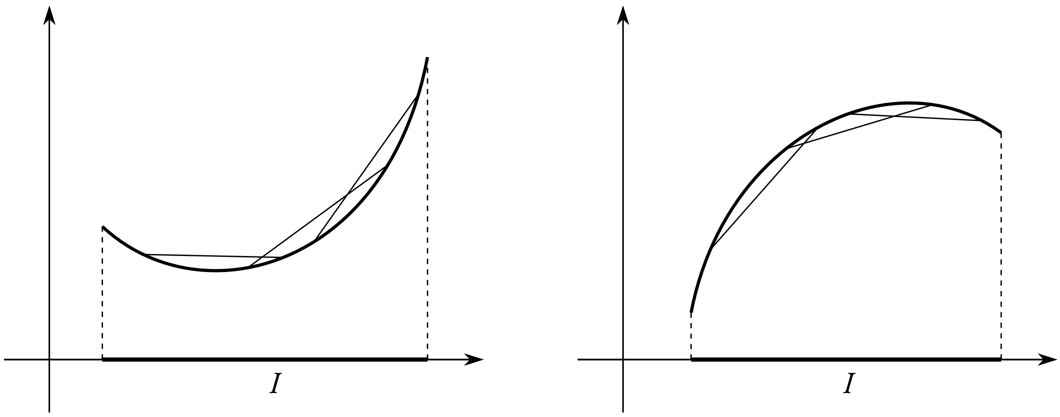


Figura 6.9 *Funzioni convesse e concave in un intervallo*

Per le funzioni derivabili due volte è possibile decidere se sono convesse o concave: si prova infatti che se una funzione ha $f''(x) > 0$ in un intervallo I, allora è convessa in I, se invece ha $f''(x) < 0$ in I, allora è concava in I.

Definizione 6.9. *Sia f una funzione derivabile in un intervallo I e x_0 un punto di I. Se esistono due intervalli del tipo $[x_1, x_0$ e del tipo $[x_0, x_2]$ tali che la funzione sia convessa nel primo e concava nel secondo, oppure concava nel primo e convessa nel secondo, allora il punto x_0 si dice* punto di flesso *o punto di inflessione per il grafico di f. La tangente al grafico nel punto $(x_0, f(x_0))$ si dice* tangente inflessionale.

È (abbastanza) evidente che se una funzione è derivabile due volte, la sua derivata seconda vale 0 in un punto di flesso.

La determinazione degli intervalli in cui una funzione è convessa oppure concava, e dei punti di flesso, migliora sensibilmente le informazioni sull'andamento grafico di una funzione, come mostra l'esempio che segue.

Esempio 6.8. Utilizzando le nozioni fin qui apprese, determiniamo i massimi, i minimi, i flessi e gli intervalli di crescenza, decrescenza, concavità, convessità della funzione

$$f(x) = \frac{x}{e^x},.$$

Vogliamo inoltre tracciare un abbozzo del grafico, tenendo anche conto delle ulteriori informazioni che possiamo ricavare dal calcolo di opportuni limiti, e magari determinando esplicitamente alcuni punti significativi per i quali il grafico deve passare.

Cominciamo con il calcolare le derivate prima e seconda di f.

$$f'(x) = \frac{1 \cdot e^x - x \cdot e^x}{(e^x)^2} = \frac{e^x(1-x)}{(e^x)^2} = \frac{\cancel{e^x}(1-x)}{(e^x)^{\cancel{2}}} = \frac{1-x}{e^x};$$

$$f''(x) = \frac{-1 \cdot e^x - (1-x)e^x}{(e^x)^2} = \frac{\cancel{e^x}(x-2)}{(e^x)^{\cancel{2}}} = \frac{x-2}{e^x}.$$

Se teniamo conto che e^x è una quantità sempre positiva, possiamo facilmente concludere che la derivata prima è positiva (funzione crescente) per $x < 1$, negativa (funzione decrescente) per $x > 1$, nulla (tangente orizzontale) per $x = 1$, dove avrà un punto di massimo (relativo) in quanto è prima crescente e poi decrescente (naturalmente la cosa è vera perchè la funzione non presenta "strappi" nel suo grafico!); per la derivata seconda si ha invece che è positiva (funzione convessa) per $x > 2$, negativa (funzione concava) per $x < 2$, nulla per $x = 2$, dove ha un punto di flesso in quanto a sinistra di 2 è concava, a destra è convessa. L'ordinata corrispondente all'ascissa 2 è $^2/_{e^2} \simeq 0.27$, mentre la derivata per $x = 2$ vale $-^1/_{e^2} \simeq -0.14$: questo ci consente di scrivere subito l'equazione della tangente inflessionale,

$$y - f(2) = f'(2)(x-2), \quad \text{ovvero} \quad y - \frac{2}{e^2} = \frac{-1}{e^2}(x-2)$$

Per raffinare ulteriormente le informazioni a nostra disposizione relative al grafico di f calcoliamo anche i limiti per x tendente a $-\infty$ e a $+\infty$ per sapere "da dove parte" e "dove arriva" il grafico stesso.

$$\lim_{x \to -\infty} \frac{x}{e^x} = \frac{-\infty}{0^+} = -\infty \quad \text{(Ricordare il grafico di } e^x\text{!!);}$$

$$\lim_{x \to \infty} \frac{x}{e^x} = \frac{+\infty}{+\infty} \overset{\cancel{(H)}}{=} \lim_{x \to \infty} \frac{1}{e^x} = \frac{1}{+\infty} = 0$$

Di solito è anche utile sapere se il grafico "sta sopra" oppure "sta sotto" all'asse x. Per fare questo basta controllare dove $f(x)$ è positiva e dove è negativa. In questo caso la cosa è immediata: $f(x)$ sta sopra all'asse x per $x > 0$, sta sotto all'asse x per $x < 0$, taglia l'asse delle x per $x = 0$.

Conviene riportare tutti questi risultati in un diagramma che ci consentirà di tracciare facilmente un grafico significativo della funzione stessa. Nella pratica è conveniente costruire questo diagramma un po' alla volta, man mano che si ottengono i vari risultati.

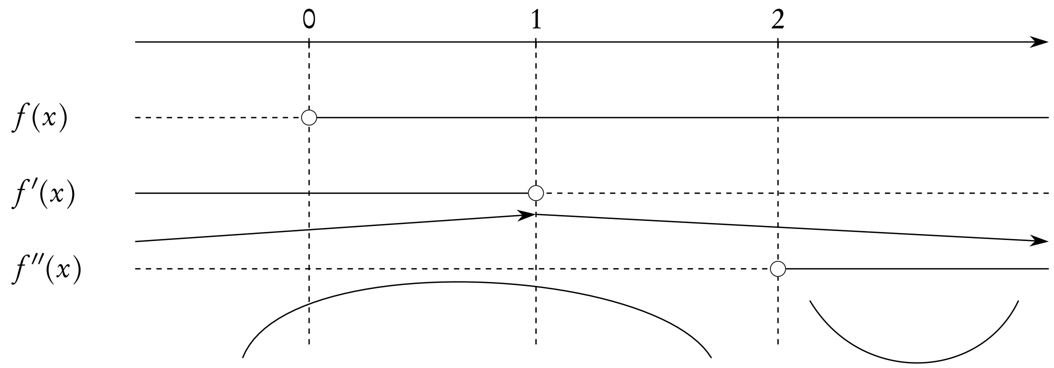


Figura 6.10 *Diagramma preparatorio allo studio della funzione* $f(x) = {}^x/_{e^x}$

Per tracciare un grafico il più corretto possibile ci serve ancora trovare l'ordinata del punto di massimo (massimo relativo ma anche assoluto in quanto la funzione non supera mai questo valore):

$$f(1)=\frac{1}{e}\simeq 0.37\,.$$

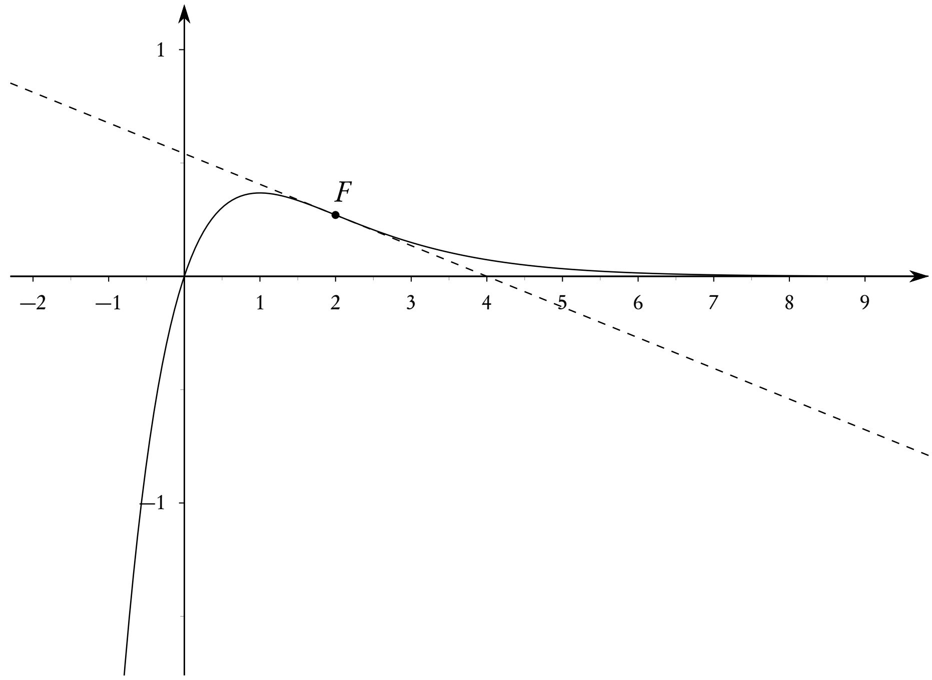


Figura 6.11 *Grafico di $f(x)={}^{x}/_{e^x}$ (Attenzione: unità di misura diverse sui due assi!)*

Da quanto detto sul significato della derivata seconda si può concludere che se in un punto x_0 la derivata prima di una funzione si annulla e se in quel punto la derivata seconda è positiva, allora il punto è di minimo relativo, se invece la derivata seconda è negativa, allora il punto è di massimo relativo.

6.4 Asintoti al grafico di una funzione

L'ultimo raffinamento che ci interesserà relativamente alla rappresentazione grafica di una funzione è quello della ricerca di particolari rette, dette *asintoti*, a cui il grafico stesso "si avvicina indefinitamente". Ci sono tre tipi possibili di asintoti: verticali, orizzontali e obliqui.

1. Una retta verticale (cioè del tipo $x=a$, con a numero reale) è un *asintoto verticale* per una funzione se

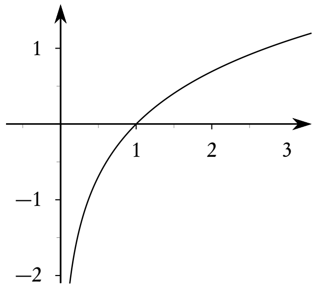
$$\lim_{x\to a^+} f(x)=\infty \quad \vee \quad \lim_{x\to a^-} f(x)=\infty \quad \vee \quad \lim_{x\to a} f(x)=\infty\,,$$

dove non ha importanza il segno di infinito.

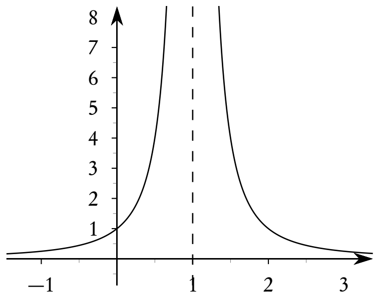
2. Una retta orizzontale (cioè del tipo $y = b$, con b numero reale) è un *asintoto orizzontale* per una funzione se
$$\lim_{x\to+\infty} f(x) = b \quad \vee \quad \lim_{x\to-\infty} f(x) = b\,.$$

3. Una retta obliqua (cioè del tipo $y = mx + q$, con $m \neq 0$) è un *asintoto obliquo* per una funzione se
$$\lim_{x\to+\infty} [f(x) - (mx+q)] = 0 \quad \vee \quad \lim_{x\to-\infty} [f(x) - (mx+q)] = 0\,.$$

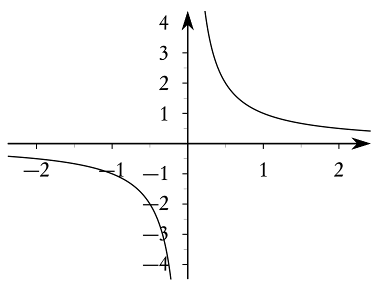
Esempio 6.9. La retta $x = 0$ è un asintoto verticale per la funzione $f(x) = \ln x$, infatti si ha
$$\lim_{x\to0^+} \ln x = -\infty\,.$$



Esempio 6.10. La retta $x = 1$ è un asintoto verticale per la funzione $f(x) = 1/(x-1)^2$, infatti si ha
$$\lim_{x\to1} \frac{1}{(x-1)^2} = +\infty\,.$$
La retta $y = 0$ è asintoto orizzontale per la stessa funzione, infatti si ha
$$\lim_{x\to\pm\infty} \frac{1}{(x-1)^2} = 0\,.$$

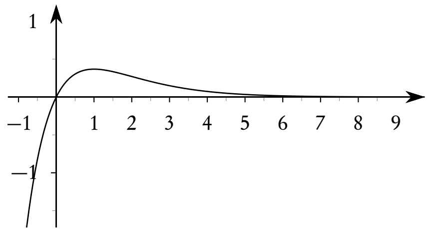


Esempio 6.11. La retta $x = 0$ è un asintoto verticale per la funzione $f(x) = 1/x$, infatti si ha
$$\lim_{x\to0^-} \frac{1}{x} = -\infty \wedge \lim_{x\to0^+} \frac{1}{x} = +\infty\,.$$
La retta $y = 0$ è asintoto orizzontale per la stessa funzione, infatti si ha
$$\lim_{x\to\pm\infty} \frac{1}{x} = 0\,.$$



Esempio 6.12. La retta $y = 0$ è asintoto orizzontale per la funzione $f(x) = {}^x/_{e^x}$, infatti si ha

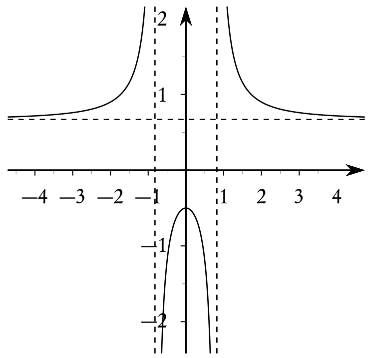
$$\lim_{x\to+\infty} \frac{x}{e^x} = 0\,.$$



Esempio 6.13. La retta $y = {}^2/_3$ è asintoto orizzontale per la funzione $f(x) = {}^{(2x^2+1)}/_{(3x^2-2)}$, infatti si ha

$$\lim_{x\to\pm\infty} \frac{2x^2+1}{3x^2-2} = \frac{2}{3}\,.$$

Le rette $x = \pm\sqrt{{}^2/_3}$ sono asintoti verticali per la stessa funzione, infatti…

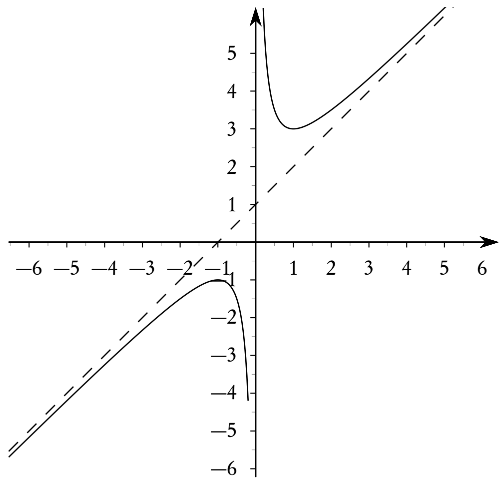


Esempio 6.14. La retta $x = 0$ è asintoto verticale per la funzione $f(x) = {}^{(x^2+x+1)}/_x$, infatti si ha

$$\lim_{x\to0^\pm} \frac{x^2+x+1}{x} = \pm\infty\,.$$

La retta $y = x + 1$ è asintoto obliquo per la stessa funzione, infatti si ha

$$\lim_{x\to\pm\infty} \left[\frac{(x^2+x+1)}{x} - (x+1)\right] = \lim_{x\to\pm\infty} \frac{1}{x} = 0\,.$$



Metodi pratici per la ricerca degli asintoti

Asintoti verticali Nei casi che ci interesseranno gli asintoti verticali si possono trovare (ma bisogna provare!) in presenza di funzioni fratte nei punti in cui si annulla il denominatore (punti che vanno esclusi dal dominio naturale), oppure in presenza di funzioni logaritmiche in punti in cui l'argomento del logaritmo si annulla.

Asintoti orizzontali Per ricercare questi asintoti basta fare i limiti per $x \to \pm\infty$ (di solito separatamente a $+\infty$ e a $-\infty$): se uno almeno di questi limiti è finito e vale b, allora $y = b$ è asintoto orizzontale.

Asintoti obliqui Si procede secondo il seguente schema:

1. Si calcola il limite, per $x \to +\infty$: se il limite è finito si ha un asintoto orizzontale e il gioco finisce; se il limite non c'è, il gioco finisce ugualmente.

2. Se il precedente limite è infinito si calcola il limite

$$\lim_{x\to+\infty} \frac{f(x)}{x}\ :$$

se questo limite è 0, infinito, o non esiste, tutto finisce.

3. Se il limite precedente è finito e diverso da zero, e lo indichiamo con m, si calcola il limite

$$\lim_{x\to+\infty} [f(x) - mx]\ :$$

se questo limite è infinito o non esiste, tutto finisce. Se invece esso è finito (anche zero), e lo indichiamo con q, allora la retta

$$y = mx + q$$

è un asintoto obliquo.

4. Si ripete il tutto per $x \to -\infty$.

Esempio 6.15. Nella funzione

$$f(x) = \frac{x^2-1}{x^2-x}$$

il denominatore si annulla per $x = 0$ e per $x = 1$. Si ha poi

$$\lim_{x\to 1} \frac{x^2-1}{x^2-x} = \lim_{x\to 0} \frac{(x-1)(x+1)}{x(x-1)} = \lim_{x\to 0} \frac{x+1}{x} = 2\,,$$

mentre

$$\lim_{x\to 0^\pm} \frac{x^2-1}{x^2-x} = \pm\infty\,.$$

Dunque solo $x = 0$ è asintoto verticale.

Si ha poi

$$\lim_{x\to\pm\infty} \frac{x^2-1}{x^2-x} = 1\,,$$

dunque $y = 1$ è asintoto orizzontale.

Esempio 6.16. Nella funzione

$$f(x) = \ln(x+2)$$

il dominio naturale è $x > -2$, e l'argomento del logaritmo si annulla per $x = -2$. Poiché si ha

$$\lim_{x\to -2^+} \ln(x+2) = -\infty,$$

la retta $x = -2$ è asintoto verticale. Si ha poi

$$\lim_{x\to+\infty} \ln(x+2) = +\infty\,;$$

dunque non ci sono asintoti orizzontali. Calcoliamo allora

$$\lim_{x\to+\infty}\frac{f(x)}{x}=\lim_{x\to+\infty}\frac{\ln(x+2)}{x}=\frac{+\infty}{+\infty}\overset{(H)}{=}\lim_{x\to+\infty}\frac{\frac{1}{x+2}}{1}=\lim_{x\to+\infty}\frac{1}{x+2}=0\,,$$

per cui non ci sono asintoti obliqui.

Esempio 6.17. Nella funzione

$$f(x)=\frac{x^2+1}{x-1}$$

il dominio naturale è $x\neq 1$. Poiché si ha

$$\lim_{x\to 1^{\pm}}\frac{x^2+1}{x-1}=\pm\infty\,,$$

la retta $x=1$ è asintoto verticale. Si ha poi

$$\lim_{x\to\pm\infty}\frac{x^2+1}{x-1}=\pm\infty\,,$$

per cui non ci sono asintoti orizzontali. Calcoliamo allora

$$\lim_{x\to\pm\infty}\frac{f(x)}{x}=\lim_{x\to\pm\infty}\frac{\frac{x^2+1}{x-1}}{x}=\lim_{x\to\pm\infty}\frac{x^2+1}{x^2-x}=\cdots=1\,(=m)$$

Possiamo procedere con il calcolo:

$$\lim_{x\to\pm\infty}[f(x)-mx]=\lim_{x\to\pm\infty}\left[\frac{x^2+1}{x-1}-1\cdot x\right]=\lim_{x\to\pm\infty}\frac{x^2+1-x^2+x}{x-1}=\cdots=1\,,$$

dunque la retta $y=x+1$ è asintoto obliquo (e la cosa funziona sia a $+\infty$ che a $-\infty$).

Il lettore è invitato a controllare questi risultati usando, per esempio, Geogebra.

6.5 Esercizi

Esercizio 6.1. *Si considerino i limiti seguenti; si verifichi se si può applicare la regola di l'Hôpital e, in caso affermativo, si calcoli il limite.*

1. $\displaystyle\lim_{x\to+\infty}\frac{x^2-1}{x^3+3}$;
2. $\displaystyle\lim_{x\to 1}\frac{\sin^2(x-1)}{x^3-1}$;
3. $\displaystyle\lim_{x\to 0}\frac{\sin^2(x-1)}{x^3-1}$;

4. $\lim_{x\to 2} \dfrac{\cos(x^2-4)}{x^3-8}$;
5. $\lim_{x\to +\infty} \dfrac{\cos(x^2-4)}{x^3-8}$;
6. $\lim_{x\to 2} \dfrac{e^{x^2-4}-1}{x^3-8}$;
7. $\lim_{x\to -\infty} \dfrac{e^{x^2-4}-1}{x^3-8}$;
8. $\lim_{x\to 0} \dfrac{e^x x^2}{\sin x}$;
9. $\lim_{x\to -\infty} \dfrac{e^x}{\sin x}$;
10. $\lim_{x\to 0} \dfrac{e^x}{\sin x}$;
11. $\lim_{x\to 0} \dfrac{e^x-1}{\sin x}$;
12. $\lim_{x\to +\infty} \dfrac{e^x}{x}$;
13. $\lim_{x\to +\infty} \dfrac{\sin x}{e^x}$

Esercizio 6.2. *Studiare le seguenti funzioni.*

1. $f(x)=\sqrt{x}+\ln x$;
2. $f(x)=\sqrt{x}-\ln\sqrt{x}$;
3. $f(x)=x-\ln(x^2-1)$;
4. $f(x)=x^4-x^3$;
5. $f(x)=x^3+x^4$;
6. $f(x)=x^6-x^4$;
7. $f(x)=x^4-1$;
8. $f(x)=x^3-x^2+x$.

Esercizio 6.3. *Studiare le seguenti funzioni ristrette all'intervallo I indicato, calcolando, in particolare, il massimo e minimo assoluti, se esistono.*

1. $f(x)=\dfrac{1}{x+1}$, $I=[0,1]$;
2. $f(x)=\dfrac{1}{x+1}$, $I=[-3,0]$;
3. $f(x)=\dfrac{x}{x+1}$, $I=[-3,-1]$;

4. $f(x) = \dfrac{x}{1-x}$, $I = [-3,3]$;
5. $f(x) = x + \dfrac{1}{x}$, $I = [3,4]$;
6. $f(x) = x + \sqrt{x-1}$, $I = [0,4]$;
7. $f(x) = \dfrac{1}{x^2} - 1$, $I = [-3,3]$;
8. $f(x) = \dfrac{1}{x^3}$, $I = [-3.3]$;
9. $f(x) = x^3 + x$, $I = [-10,0]$;
10. $f(x) = x^3 - x^2$, $I = [-1,10]$;
11. $f(x) = \ln(x^3 + x^2)$, $I = [0,10]$;
12. $f(x) = \sqrt{x^2+1}$, $I = [-10,10]$;
13. $f(x) = x^3 + x^2$, $I = [0,3]$;
14. $f(x) = x^5 + x$, $I = [0,1]$.

Esercizio 6.4. *Studiare le seguenti funzioni, tralasciando lo studio della derivata seconda se troppo complesso.*

1. $f(x) = \dfrac{1}{x^2+1}$;
2. $f(x) = \dfrac{1}{x^2-1}$;
3. $f(x) = \dfrac{1}{x^2+x}$;
4. $f(x) = \dfrac{x^2}{x^2+1}$;
5. $f(x) = \ln\dfrac{1}{x}$;
6. $f(x) = \ln\left(\dfrac{1}{x^2}+1\right)$;
7. $f(x) = e^{x^2}$;
8. $f(x) = xe^{1/x}$;
9. $f(x) = e^{1-x^2}$;
10. $f(x) = \dfrac{\ln x}{x}$;
11. $f(x) = \dfrac{e^x}{x}$.

Esercizio 6.5. *Determinare i punti di massimo e minimo relativo delle seguenti funzioni usando solo* $f' = 0$ *e* $f'' > 0$ *oppure* $f'' < 0$.

1. $f(x) = \ln(2x^2 - x)$;
2. $f(x) = e^{3x}(x^2 + x)$;
3. $f(x) = x^3 - x^4$;
4. $f(x) = e^{x^2-2x}$.

Esercizio 6.6. *Usando il metodo indicato nell'esercizio 6.5, dire se i punti indicati sono di massimo o minimo relativo per le funzioni seguenti.*

1. $f(x) = x\ln(x^2 - 4x^3)$; $x = 0$, $x = 2$;
2. $f(x) = e^{x^2-3x^3}$; $x = 0$, $x = 1$;
3. $f(x) = e^{4x^3-3x^4}$; $x = 0$, $x = -1$, $x = 1$;
4. $f(x) = x^2 e^{2x} - 2x^2 \ln x$; $x = 0$, $x = 3$.

Esercizio 6.7. *Tracciare il grafico delle seguenti funzioni, valutando in particolare la continuità e la derivabilità.*

1. $f(x) = \begin{cases} x, & \textit{se} \quad x \leq 0 \\ -2x, & \textit{se} \quad x > 0 \end{cases}$;
2. $f(x) = \begin{cases} x^2, & \textit{se} \quad x \leq 0 \\ -x^3, & \textit{se} \quad x > 0 \end{cases}$;
3. $f(x) = \begin{cases} \sqrt{x}, & \textit{se} \quad x \leq 0 \\ -x^2, & \textit{se} \quad x > 0 \end{cases}$;
4. $f(x) = \begin{cases} 2^x, & \textit{se} \quad x \leq -1 \\ \ln(x-2), & \textit{se} \quad x > -1 \end{cases}$;
5. $f(x) = \begin{cases} \sqrt{x}, & \textit{se} \quad x < 1 \\ \sqrt[3]{x} + 1, & \textit{se} \quad x \geq 1 \end{cases}$;
6. $f(x) = \begin{cases} 2^{-x} + 1, & \textit{se} \quad x < -1 \\ -\ln(x+2), & \textit{se} \quad x \geq -1 \end{cases}$;
7. $f(x) = \begin{cases} x, & \textit{se} \quad x \leq 1/2 \\ x^2 + 1/2, & \textit{se} \quad x > 1/2 \end{cases}$;
8. $f(x) = \begin{cases} x^2, & \textit{se} \quad x \leq 1 \\ (2/3)x^3, & \textit{se} \quad x > 1 \end{cases}$;
9. $f(x) = \begin{cases} 4\sqrt{x}, & \textit{se} \quad x \leq 1 \\ x^2, & \textit{se} \quad x > 1 \end{cases}$;
10. $f(x) = \begin{cases} 3^x, & \textit{se} \quad x \leq 2 \\ \ln(x^2 + 1), & \textit{se} \quad x > 2 \end{cases}$;

11. $f(x) = \begin{cases} 2\sqrt{x}, & se \quad x < 1 \\ 3\sqrt[3]{x}, & se \quad x \geq 1 \end{cases}$.

7 Integrali per funzioni di una variabile

7.1 Introduzione

Il calcolo integrale (per le funzioni reali di variabile reale), si occupa della risoluzione di due problemi, apparentemente indipendenti:

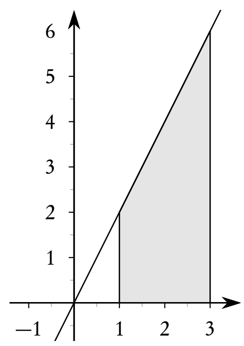
1. il calcolo dell'area di parti di piano qualsiasi (non solo dunque di poligoni o figure riconducibili a parti di cerchio);
2. la ricerca di funzioni che hanno una derivata assegnata.

Per quanto attiene ai contenuti di questo corso il primo problema sarà limitato al caso di figure piane "racchiuse" tra il grafico di funzioni di una variabile, che saranno sempre almeno continue; anche per il secondo problema ci limiteremo solo a trattare alcune situazioni molto semplici.

Il *Teorema fondamentale del calcolo integrale* stabilirà un legame molto stretto tra i due problemi indicati, almeno nei casi di funzioni continue a cui siamo interessati.

Come è ormai abitudine, prima di entrare nel vivo del discorso proponiamo alcuni esempi esplicativi.

Esempio. Consideriamo la funzione $f(x) = 2x$, ristretta all'intervallo $[1,3]$, il cui grafico è rappresentato nella figura qui a lato, e proponiamoci di calcolare l'area della regione evidenziata (compresa tra il grafico della funzione f, l'asse delle ascisse, e le due rette verticali $x = 1$ e $x = 3$).



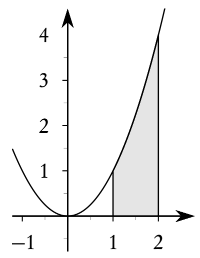
La regione evidenziata è un trapezio rettangolo, di base minore lunga 2, base maggiore lunga 6 e altezza lunga 2: l'area misurerà dunque 8.

È molto facile anche trovare una funzione che abbia f come derivata: si tratta, per esempio, della funzione $g(x) = x^2$.

La cosa interessante, e sarà proprio questo il contenuto del già citato Teorema fondamentale, è che l'area della regione evidenziata si può calcolare a partire dalla funzione g, facendo semplicemente $g(3) - g(1)$:

$$Area = 8 = 3^2 - 1^2 = g(3) - g(1).$$

Esempio. Consideriamo la funzione $f(x) = x^2$, ristretta all'intervallo $[1,2]$, il cui grafico è rappresentato nella figura qui a lato, e proponiamoci di calcolare l'area della regione evidenziata (compresa tra il grafico della funzione f, l'asse delle ascisse, e le due rette verticali $x = 1$ e $x = 2$).



La regione evidenziata non è più un trapezio rettangolo come nel caso precedente, anche se è molto simile a un trapezio rettangolo, con il lato obliquo sostituito da un arco di parabola: chiameremo questa regione un *trapezoide*. Il calcolo dell'area non sarà più elementare come prima, anche se Archimede già nel terzo secolo avanti Cristo era in grado di calcolarla, ottenendo il valore $7/3$.

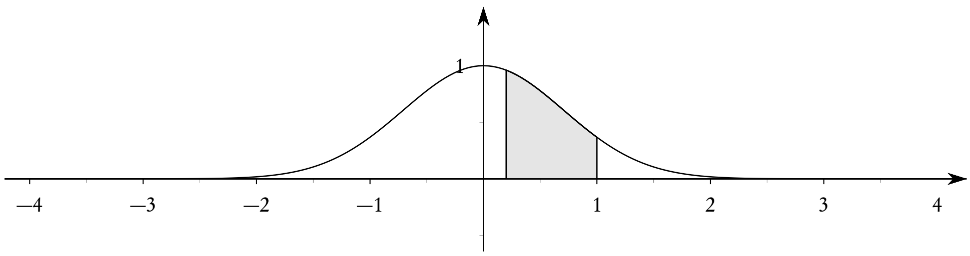
Anche questa volta è molto facile trovare una funzione che abbia f come derivata: si tratta, per esempio, della funzione $g(x)=x^3/3$.

Ancora una volta la cosa interessante è che il valore dell'area si può calcolare direttamente usando la funzione g e calcolando $g(2)-g(1)$:

$$Area=\frac{7}{3}=\frac{2^3}{3}-\frac{1^3}{3}=g(2)-g(1).$$

Purtroppo le cose non sono sempre così semplici, nemmeno per i software di calcolo simbolico. In genere questi software sono facilmente in grado di trovare valori approssimati, con il voluto grado di approssimazione, di aree del tipo che stiamo considerando. Il problema di trovare una funzione cha abbia una derivata assegnata è invece estremamente difficile, anche se è noto che esso ha, teoricamente, sempre soluzione nel caso che la funzione assegnata sia continua.

È interessante eseguire una prova specifica, con due software molto diffusi (Geogebra e Mathematica), su una funzione di grande importanza applicativa come la funzione di Gauss $f(x)=\mathrm{e}^{-x^2}$, per esempio sia calcolando l'area evidenziata nella figura seguente ($0.2<x<1$) che cercando una funzione che abbia f come derivata.



Il problema del calcolo (approssimato a 10 cifre decimali) dell'area è quasi immediato con entrambi i software: 0.5494591019.

Il problema del calcolo di una funzione g che abbia f come derivata è estremamente più complesso, tanto che il primo dei due software citati (Geogebra) non fornisce alcuna risposta (g "non definito") mentre il secondo (Mathematica) fornisce una risposta che a questo livello del corso non è interpretabile semplicemente:

$$\frac{1}{2}\sqrt{\pi}\,\mathrm{Erf}(x).$$

In ogni caso, anche se in quest'ultimo caso non abbiamo la possibilità di una verifica diretta, esiste sempre lo stesso legame tra l'area cercata e la funzione g (supposto che si sia in grado di calcolarla!).

7.2 Primitive per una funzione reale di variabile reale

Cominciamo a trattare con un po' più di dettaglio il problema del calcolo di una funzione avente assegnata derivata. Già sappiamo, in base a uno dei corollari del teorema di Lagrange, che se due funzioni definite su un intervallo I hanno la stessa derivata, allora esse differiscono per una costante. Se cerchiamo, per esempio, una funzione che abbia x^2 come derivata, oltre a $x^3/3$ andranno anche bene tutte le funzioni del tipo $x^3/3+c$, essendo c una costante arbitraria. Viceversa se ci viene chiesto di trovare *tutte* le funzioni che abbiano x^2 come derivata, sulla base dello stesso corollario potremo concludere che esse sono tutte e sole le funzioni

$$\frac{x^3}{3}+c\,.$$

La cosa è vera in generale: se è data una funzione $f(x)$ definita in un intervallo I (e che per noi sarà sempre continua in tutto l'intervallo), e se $g(x)$ è una funzione tale che

$$g'(x)=f(x),\ \forall x\in I,$$

allora tutte e sole le funzioni che hanno $f(x)$ come derivata sono date dalla formula

$$g(x)+c,$$

essendo c una costante arbitraria. Ci sono problemi per le funzioni che non siano definite su un intervallo, ma il loro esame esula dagli scopi di questo corso: trattando di questo tipo di problemi supporremo sempre che il dominio delle funzioni sia un intervallo; se così non fosse "restringeremo" il dominio a un intervallo.

Definizione 7.1. *Data una funzione f, definita in un intervallo I, chiameremo* primitiva *di f ogni funzione F definita e derivabile nello stesso intervallo e tale che*

$$F'(x)=f(x),\ \forall x\in I\,.$$

L'insieme di tutte *le primitive di una funzione f, in un intervallo I, si denota con il simbolo*

$$\int f(x)\,\mathrm{d}x$$

che si legge Integrale (indefinito) di $f(x)$ in dx.

La definizione data implica che, se $F(x)$ è una qualunque primitiva di $f(x)$, nell'intervallo I, il simbolo di integrale indefinito ha la seguente interpretazione:

$$\int f(x)\,\mathrm{d}x=\{\,F(x)+c\mid c\in\mathbb{R}\,\}\,.$$

Basterà dunque riuscire a trovare una sola primitiva di una funzione f per trovarle tutte (sempre nell'ipotesi, che non ripeteremo più, che il dominio sia un intervallo.

Si può dimostrare che tutte le funzioni continue in I hanno sempre primitive, ma, come mostra l'esempio della funzione di Gauss, la loro effettiva determinazione è in genere un problema arduo e, di solito, non risolubile con metodi elementari. Noi ci occuperemo della risoluzione di questo problema solo in casi molto semplici, segnalando che tutti i software di calcolo simbolico sono in grado di trovare primitive anche di funzioni complesse (ma non di tutte...).

Una prima osservazione che possiamo fare relativamente al calcolo di primitive è legata alle note proprietà delle derivate:

$$\big(kf(x)\big)'=kf'(x)\quad \text{e}\quad \big(f(x)+g(x)\big)'=f'(x)+g'(x),$$

da cui si ricava subito che

$$\int kf(x)\,\mathrm{d}x = k\int f(x)\,\mathrm{d}x \quad \text{e} \quad \int \big(f(x)+g(x)\big)\,\mathrm{d}x = \int f(x)\,\mathrm{d}x + \int g(x)\,\mathrm{d}x\,.$$

Queste formule si leggono a parole, rispettivamente, nel seguente modo: *le costanti possono essere "portate fuori" dal segno di integrale* e *l'integrale di una somma è uguale alla somma degli integrali*. Esse si chiamano, brevemente, proprietà di linearità dell'integrale.

Purtroppo l'analogia con le derivate si ferma qui: non esiste, per esempio, alcuna formula generale per calcolare integrali di prodotti o di quozienti di funzioni che si sappiano integrare, e neppure per calcolare l'integrale della composta di due funzioni che si sappiano integrare.

Una seconda strategia di calcolo si basa sulla lettura della tabella delle derivate "da destra a sinistra". Si costruisce in questo modo una nuova tabella, detta *Tabella delle primitive fondamentali*. Nella tabella 7.1 riportiamo le situazioni più comuni, unitamente a qualche altro caso di interesse applicativo.

Funzione	***Primitive***
k	$kx+c$
$x^\alpha,\ \alpha\in\mathbb{R}\setminus\{-1\}$	$\dfrac{x^{\alpha+1}}{\alpha+1}+c$
$x^{-1}=\dfrac{1}{x}$	$\ln\lvert x\rvert+c$
e^x	e^x+c
$\ln x$	$x\ln x - x + c$
$\sin x$	$-\cos x + c$
$\cos x$	$\sin x + c$
$\big(f(x)\big)^\alpha f'(x),\ \alpha\in\mathbb{R}\setminus\{-1\}$	$\dfrac{\big(f(x)\big)^{\alpha+1}}{\alpha+1}+c$
$\big(f(x)\big)^{-1} f'(x)=\dfrac{f'(x)}{f(x)}$	$\ln\lvert f(x)\rvert+c$
$f'(x)\mathrm{e}^{f(x)}$	$\mathrm{e}^{f(x)}+c$
$f'(x)\sin f(x)$	$-\cos f(x)+c$

Continua nella pagina successiva

Segue dalla pagina precedente

Funzione	***Primitive***
$f'(x)\cos f(x)$	$\sin f(x)+c$

Tabella 7.1 *Alcune primitive fondamentali*

Si noti, in particolare, che le potenze hanno un comportamento completamente diverso a seconda che l'esponente sia -1 o un numero reale diverso da -1.

7.3 Integrazione per parti

Per il caso dell'integrale di un prodotto di due funzioni vale la cosiddetta *formula di integrazione per parti*, utile in molte situazioni ma non sempre risolutiva. Supponiamo di dover calcolare l'integrale del prodotto di due funzioni

$$\int f(x)g(x)\,\mathrm{d}x\,,$$

e supponiamo di conoscere una qualunque primitiva di una delle due, per esempio di $f(x)$, primitiva che possiamo chiamare $F(x)$. Allora vale la seguente formula

$$\int f(x)g(x)\,\mathrm{d}x = F(x)g(x) - \int F(x)g'(x)\,\mathrm{d}x\,, \tag{7.1}$$

che si può leggere come segue: *l'integrale del prodotto di due funzioni è uguale a una primitiva della prima per la seconda, meno l'integrale del prodotto tra la primitiva della prima e la derivata della seconda.* Come si vede, non serve saper trovare l'integrale di entrambe, solo di una delle due, successivamente il problema è rinviato alla ricerca di un nuovo integrale (quello del prodotto tra la primitiva di una funzione e la derivata dell'altra): se questo integrale "residuo" è più facile di quello iniziale il metodo funziona, altrimenti no.

La formula di integrazione per parti è, come già detto, utile per calcolare in alcuni casi l'integrale del prodotto di due funzioni. La sua efficiente applicazione richiede alcuni accorgimenti, come risulta chiaro dai seguenti esempi.

Esempio 7.1. Per calcolare

$$\int x\mathrm{e}^x\,\mathrm{d}x$$

si può sia partire da una primitiva di x che da un primitiva di e^x, entrambe facilmente determinabili. Seguendo le due strategie e applicando la regola per parti si ottiene

$$\int x\mathrm{e}^x\,\mathrm{d}x = \frac{x^2}{2}\mathrm{e}^x - \int \frac{x^2}{2}\mathrm{e}^x\,\mathrm{d}x\,,$$

oppure
$$\int xe^x\,dx = xe^x - \int e^x\,dx.$$

Nel primo caso l'integrale "residuo" è più complesso di quello di partenza e quindi l'uso della formula per parti non è di alcun beneficio, mentre nel caso l'integrale residuo è elementare e si può facilmente concludere che
$$\int xe^x\,dx = xe^x - \int e^x\,dx = xe^x - e^x + c.$$

Esempio 7.2. Nell'integrale
$$\int \ln x\,dx$$
non si evidenzia immediatamente alcun prodotto di due funzioni e quindi, a prima vista, sembra che la formula per parti non possa essere applicata. Se però si osserva che
$$\int \ln x\,dx = \int 1\cdot \ln x\,dx$$
e che $F(x) = x$ è una primitiva di $f(x) = 1$, si ottiene, applicando la formula per parti,
$$\int \ln x\,dx = \int 1\cdot \ln x\,dx = x\ln x - \int x\frac{1}{x}\,dx = x\ln x - \int 1\,dx = x\ln x - x + c.$$

Esempio 7.3. Se consideriamo l'integrale
$$\int 2xe^{x^2}\,dx,$$
non conviene considerare la funzione integranda come il prodotto delle due funzioni $f(x) = 2x$ e $g(x) = e^{x^2}$, quanto piuttosto osservare che la funzione integranda è semplicemente la derivata di e^{x^2}. Si ha dunque, senza bisogno di applicare alcuna formula speciale,
$$\int 2xe^{x^2}\,dx = e^{x^2} + c.$$

Questo tipo di integrali è esplicitamente trattato nella tabella 7.1.

Si noti che la regola di integrazione per parti, che si riferisce nella sostanza all'integrale del prodotto di due funzioni, è profondamente diversa dalla regola di derivazione del prodotto di due funzioni. Infatti date due funzioni $f(x)$ e $g(x)$, se sappiamo calcolare la derivata di entrambe possiamo facilmente calcolare la derivata del prodotto secondo la nota regola:
$$D\big(f(x)g(x)\big) = f'(x)g(x) + f(x)g'(x).$$

Nel caso dell'integrale, invece, può benissimo succedere che si sappia calcolare l'integrale di due funzioni, ma che non ci sia alcuna strategia elementare per calcolare l'integrale del prodotto. Un esempio famoso è il seguente.

Esempio 7.4. Siano date le funzioni

$$f(x)=\frac{1}{x} \quad \text{e} \quad g(x)=\sin x.$$

Si ha facilmente

$$\int \frac{1}{x}\,\mathrm{d}x=\ln x+c \quad \text{e} \quad \int \sin x\,\mathrm{d}x=-\cos x+c.$$

Purtroppo non esiste invece alcuna strategia elementare per calcolare

$$\int \frac{1}{x}\cdot \sin x\,\mathrm{d}x=\int \frac{\sin x}{x}\,\mathrm{d}x.$$

Dunque l'integrazione indefinita è l'operazione inversa della derivazione, ma, come ogni operazione inversa, può essere tecnicamente molto più complessa.

7.4 Integrazione per sostituzione

La regola di derivazione delle funzioni composte ha come conseguenza una importante tecnica per il calcolo delle primitive, detta *integrazione per sostituzione*. La regola si esprime, in dettaglio, con il seguente teorema.

Teorema 7.2 (Integrazione per sostituzione). *Si debba calcolare*

$$\int f(x)\,\mathrm{d}x\,,$$

e si consideri una funzione $x=g(t)$, derivabile, con derivata continua e sempre diversa da zero; si indichi con $t=h(x)$ la sua funzione inversa[1]*. Si consideri poi l'integrale*

$$\int f(g(t))g'(t)\,\mathrm{d}t.$$

Se questo integrale è calcolabile e se $F(t)$ è una primitiva, si ha

$$\int f(x)\,\mathrm{d}x=F(h(x))+c.$$

In sostanza si procede nel seguente modo.
Dato l'integrale

$$\int f(x)\,\mathrm{d}x\,,$$

si considera una "sostituzione" (derivabile e con derivata sempre non nulla)

$$x=g(t)$$

[1]Se la funzione ha derivata sempre diversa da zero è sicuramente invertibile.

e la sua inversa

$$t = h(x).$$

Si cerca di calcolare l'integrale

$$\int f(g(t))g'(t)\,\mathrm{d}t\,,$$

che ovviamente deve essere più semplice di quello originale. Se si riesce a trovare una primitiva $F(t)$, l'integrale dato si ottiene sostituendo a t la sua espressione in funzione di x: $t = h(x)$.

Si tenga ben presente che nel nuovo integrale non si deve solo operare la sostituzione $x = g(t)$, ma anche moltiplicare per la derivata $g'(t)$ della funzione $g(t)$.

In molte situazioni è più facile partire da $t = h(x)$ e poi trovare l'inversa $x = g(t)$: l'importante è considerarle, fin dall'inizio, entrambe.

Esempio 7.5. Per calcolare

$$\int \sqrt{\frac{1+x}{1-x}}\,\mathrm{d}x,$$

possiamo considerare la sostituzione $x = g(t) = \sin t$ e la sua inversa[2] $t = h(x) = \arcsin x$. Si ha $g'(t) = \cos t$ e quindi si deve calcolare

$$\int \sqrt{\frac{1+\sin t}{1-\sin t}}\cos t\,\mathrm{d}t = \int \sqrt{\frac{(1+\sin t)^2}{1-\sin^2 t}}\cos t\,\mathrm{d}t = \int (1+\sin t)\,\mathrm{d}t = t - \cos t + c.$$

Per l'integrale dato si ha dunque

$$\int \sqrt{\frac{1+x}{1-x}}\,\mathrm{d}x = \arcsin x - \cos(\arcsin x) + c \ (= \arcsin x - \sqrt{1-x^2} + c).$$

Esempio 7.6. Per calcolare

$$\int \frac{x}{\sqrt{x-1}}\,\mathrm{d}x,$$

si può porre

$$\sqrt{x-1} = t = h(x), \quad \text{da cui} \quad x = 1+t^2 = g(t), \quad \text{e} \quad g'(t) = 2t.$$

Si deve dunque calcolare

$$\int \frac{1+t^2}{t}2t\,\mathrm{d}t = \int 2(1+t^2)\,\mathrm{d}t = 2t + \frac{2}{3}t^3 + c.$$

Quindi

$$\int \frac{x}{\sqrt{x-1}}\,\mathrm{d}x = 2\sqrt{x-1} - \frac{2}{3}\sqrt{(x-1)^3} + c.$$

[2]Anche se ci sono alcuni problemi nel considerare l'inversa della funzione seno, tutto funziona correttamente.

Si noti, in entrambi gli esempi proposti, come sia proprio il fattore $g'(t)$ che semplifica il calcolo.

Naturalmente il problema nell'uso di questa tecnica è quello di "indovinare" la sostituzione giusta, cosa spesso tutt'altro che facile.

Anche utilizzando le proprietà di linearità, le regole di integrazione per parti e sostituzione, e la tabella 7.1 non si riesce ad andare molto lontano nel calcolo; in ogni caso le situazioni di nostro interesse potranno essere risolte con queste poche regole.

Proponiamo di seguito qualche semplice esempio.

Esempio 7.7. $\displaystyle\int x^2+\sin x\,\mathrm{d}x=\frac{x^3}{3}-\cos x+c\,.$

Esempio 7.8. $\displaystyle\int \frac{x+1}{x^2}\,\mathrm{d}x=\int\left(\frac{1}{x}+\frac{1}{x^2}\right)\mathrm{d}x=\int\left(\frac{1}{x}+x^{-2}\right)\mathrm{d}x=\ln|x|+\frac{x^{-1}}{-1}+c=\ln|x|-\frac{1}{x}+c\,.$

Esempio 7.9. $\displaystyle\int 2\sin(2x)\,\mathrm{d}x=\int(2x)'\sin(2x)\,\mathrm{d}x=-\cos(2x)+c\,.$

Esempio 7.10. $\displaystyle\int\left(3x^4-2x^3+x-1\right)\mathrm{d}x=3\frac{x^5}{5}-2\frac{x^4}{4}+\frac{x^2}{2}-x+c=\frac{3x^5}{5}-\frac{x^4}{2}+\frac{x^2}{2}-x+c\,.$

Esempio 7.11. $\displaystyle\int\frac{2x}{x^2+1}\,\mathrm{d}x=\int\frac{(x^2+1)'}{x^2+1}\,\mathrm{d}x=\ln\left|x^2+1\right|+c=\ln(x^2+1)+c\,.$

Esempio 7.12. $\displaystyle\int\sqrt[7]{x^5}\,\mathrm{d}x=\int x^{5/7}\,\mathrm{d}x=\frac{x^{5/7+1}}{5/7+1}+c=\frac{7}{12}x^{12/7}+c=\frac{7}{12}\sqrt[7]{x^{12}}+c\,.$

7.5 Area di un trapezoide

Consideriamo una funzione f, definita e continua in un intervallo $[a,b]$ e *sempre positiva* in tutto l'intervallo. Siamo interessati a calcolare l'area della regione racchiusa tra il grafico della funzione f, l'asse della x, e le due rette verticali $x=a$ e $x=b$, regione che chiameremo, come già anticipato, un *trapezoide*.

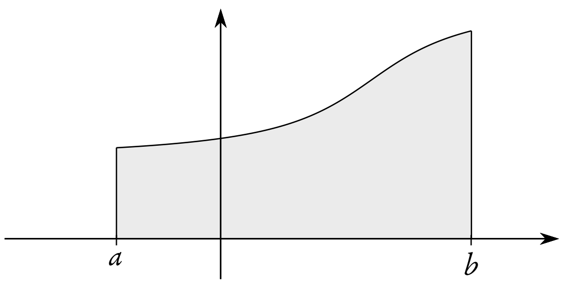


Figura 7.1 *Trapezoide relativo a una funzione positiva (e continua)*

Per valutare quest'area l'idea è quella di approssimarla, mediante dei "*plurirettangoli*", inscritti e circoscritti, ottenuti suddividendo l'intervallo $[a,b]$ in un certo numero di parti (uguali per semplicità) e procedendo come le figure 7.2 (relativa a un plurirettangolo inscritto) e 7.3 (relativa a un plurirettangolo circoscritto) indicano chiaramente.

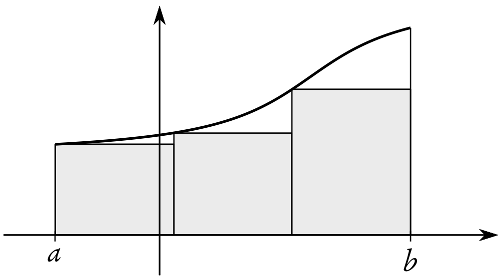


Figura 7.2 *Plurirettangolo inscritto*

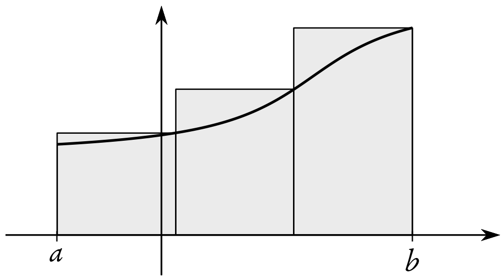


Figura 7.3 *Plurirettangolo circoscritto*

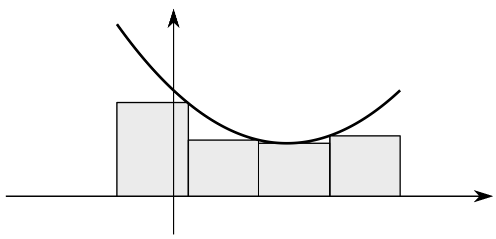
Si tratta di una generalizzazione del procedimento usato per "quadrare il cerchio", mediante una successione di poligoni inscritti e circoscritti. Anche in questo caso si dimostra (almeno nel caso delle funzioni continue che a noi interessano) che, se il numero di suddivisioni tende all'infinito (e quindi la loro ampiezza tende a zero), le aree dei due plurirettangoli inscritto e circoscritto tendono a un valore comune che si chiama *area del trapezoide* e si indica con il simbolo

$$\int_a^b f(x)\,\mathrm{d}x\,,$$

che si legge: *Integrale (definito) della funzione f tra a e b*.

L'origine di questo simbolo è legata alle considerazioni intuitive che seguono. L'area del trapezoide considerato si ottiene come somma delle aree di tanti rettangoli, che hanno come base la misura di $[a,b]$ divisa per il numero di suddivisioni, e come altezza il valore della funzione calcolato in un punto opportuno appartenente a ciascun intervallo della suddivisione. La misura della base si indicherà dunque con Δx (differenza tra l'ascissa dell'estremo destro e dell'estremo sinistro di ogni intervallo), mentre l'altezza sarà $f(x)$.

Nell'esempio proposto prima il punto in cui calcolare $f(x)$ è sempre o l'estremo sinistro o quello destro di ciascun intervallo della suddivisione, ma potrebbe anche essere un punto interno, come mostra la figura qui a lato, relativa a un plurirettangolo inscritto, in cui è stata considerata una suddivisione di $[a,b]$ in 4 parti.



L'area di un plurirettangolo inscritto o circoscritto sarà dunque esprimibile con una formula del tipo

$$\sum f(x)\Delta x\,,$$

dove la somma si intende estesa a tutti gli intervallini. Al tendere all'infinito del numero di suddivisioni, l'ampiezza di ciascuna tenderà a zero e viene indicata con $\mathrm{d}x$; il simbolo di sommatoria viene deformato in una s allungata: $\int$, per cui si ottiene proprio il simbolo già proposto.

7.6 Integrale definito

Se la funzione f anzichè essere sempre positiva fosse sempre negativa nell'intervallo $[a, b]$, la somma

$$\sum f(x)\Delta x\,, \tag{7.2}$$

avrebbe ancora senso, sarebbe negativa e corrisponderebbe all'opposto dell'area del plurirettangolo costruito con la stessa tecnica usata per le funzioni positive. Si vedano le figure seguenti.

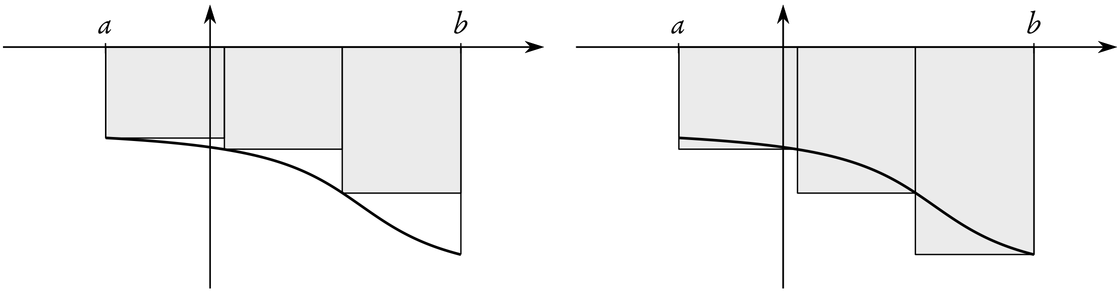


Figura 7.4 *Plurirettangolo inscritto* **Figura 7.5** *Plurirettangolo circoscritto*

È abbastanza evidente che si potrà ancora parlare di trapezoide e che il limite di una somma del tipo 7.2, al tendere all'infinito del numero di suddivisioni, sarà questa volta l'opposto dell'area del trapezoide: questo limite si indicherà ancora con lo stesso simbolo di prima

$$\int_a^b f(x)\,\mathrm{d}x\,,$$

e si userà la stessa nomenclatura.

Se una funzione è in parte positiva e in parte negativa la somma 7.2 avrà alcuni addendi positivi e alcuni negativi e il limite sarà chiaramente la differenza tra le aree di tutte le regioni soprastanti l'asse delle x e di tutte le regioni sottostanti l'asse stessa. In un caso come questo l'integrale potrebbe anche venire nullo, come dimostra chiaramente il caso della funzione seno, considerata nell'intervallo $[0, 2\pi]$, in cui, per questioni di simmetria, la parte soprastante l'asse x e quella sottostante sono chiaramente identiche.

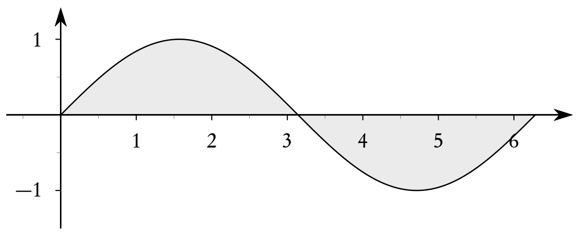


Figura 7.6 *Trapezoide relativo alla funzione* $\sin x$, *in* $[0, 2\pi]$

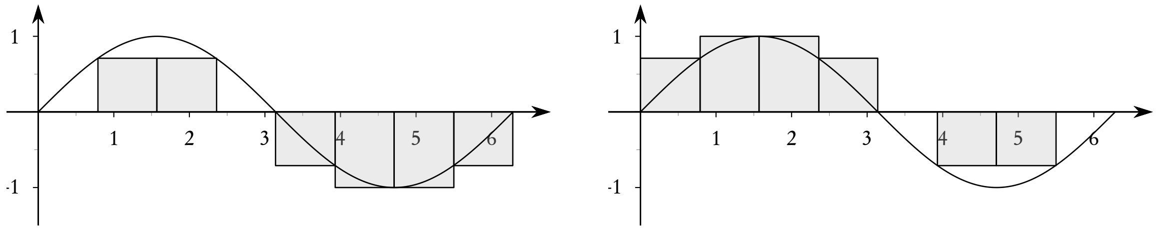


Figura 7.7 *Plurirettangoli relativi alla funzione* $\sin x$ *in* $[0, 2\pi]$, *con* 8 *suddivisioni*

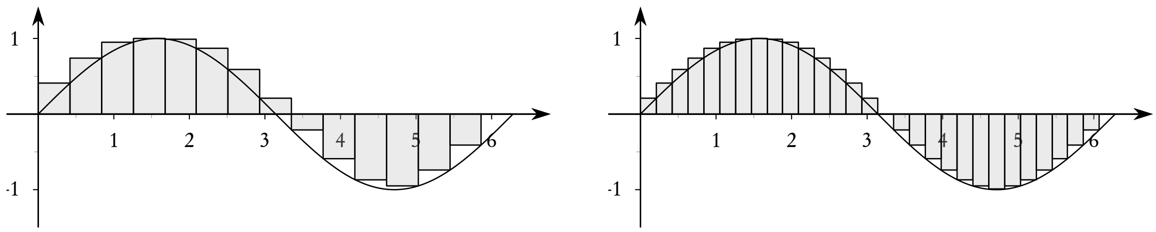


Figura 7.8 *Plurirettangoli relativi alla funzione* $\sin x$ *in* $[0, 2\pi]$, *con* 15 *e* 30 *suddivisioni*

C'è da fare un'osservazione abbastanza importante: tutto il processo di costruzione dei plurirettangoli ha senso solo perché le funzioni in esame sono limitate. La cosa è ovvia perché noi ci siamo limitati alle funzioni continue, ma se si vuole estendere il concetto a funzioni non continue bisogna prestare attenzione a questo fatto.

Siamo ora pronti per dare una definizione formalmente corretta del concetto di integrale definito.

Definizione 7.3. *Sia data una funzione f definita e continua in un intervallo chiuso e limitato $[a, b]$. Si consideri una suddivisione dell'intervallo $[a, b]$ in n parti uguali, ciascuna dunque di misura $\delta_i = (b-a)/n$ e si prenda, in ciascun sottointervallo, il massimo M_i e il minimo m_i della funzione. Le somme*

$$\sum_{i=1}^{n} m_i \delta_i \quad e \quad \sum_{i=1}^{n} M_i \delta_i$$

si chiamano, rispettivamente, Somma integrale inferiore *e* Somma integrale superiore *relative alla funzione f, all'intervallo $[a, b]$ e alla sua suddivisione in n parti.*

Si dimostra che, per le funzioni continue in un intervallo chiuso e limitato, le somme inferiori e superiori, al tendere di n all'infinito, tendono a un comune valore, che si chiama *Integrale definito di $f(x)$ tra a e b* e si indica con

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

Nel caso di funzioni positive l'integrale ha il significato di area del trapezoide individuato dalla funzione f sull'intervallo $[a, b]$, nel caso di funzioni negative ha il significato di opposto dell'area del trapezoide.

I numeri a e b si chiamano *estremi di integrazione*, la funzione f si chiama *funzione integranda*. È chiaro da quanto detto che, nel simbolo di integrale il nome della variabile non ha alcun interesse, tanto che alcuni lo omettono completamente; anche il $\mathrm{d}x$ serve soltanto a ricordare il processo di limite attraverso cui si è giunti alla definizione, e anche questo viene omesso da qualcuno. Si possono cioè considerare equivalenti i simboli seguenti:

$$\int_a^b f(x)\,\mathrm{d}x\,,\quad \int_a^b f(t)\,\mathrm{d}t\,,\quad \int_a^b f\,\mathrm{d}x\,,\quad \int_a^b f\,.$$

È altresì evidente che il numero a deve essere strettamente minore del numero b. Nelle applicazioni ha interesse considerare anche integrali tra c e d, con $c \geq d$. Per questo si pone, per definizione

$$\int_c^c f(x)\,\mathrm{d}x = 0\,,$$

e, se $c > d$,

$$\int_c^d f(x)\,\mathrm{d}x = -\int_d^c f(x)\,\mathrm{d}x\,.$$

Con queste definizioni il simbolo

$$\int_a^b f(x)\,\mathrm{d}x$$

ha senso qualunque sia l'ordine dei numeri a e b (naturalmente purché la funzione sia definita e continua in $[a,b]$ oppure $[b,a]$).

Una prima proprietà che si può dimostrare, molto utile nelle applicazioni, è quella relativa all'additività rispetto all'intervallo di integrazione.

Teorema 7.4. *Se a, b, c sono tre reali qualsiasi, si ha*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x\,.$$

La verifica grafica di questo teorema è ovvia per funzioni positive, nel caso che $a < c < b$, come mostra la figura che segue: il primo integrale è l'area del trapezoide totale, gli altri due le aree dei trapezoidi indicati con 1 e 2, ed è chiaro che l'area totale è la somma delle altre due. In ogni caso, facendo riferimento a opportuni grafici, non è difficile rendersi conto del perché la proprietà valga anche in casi più generali.

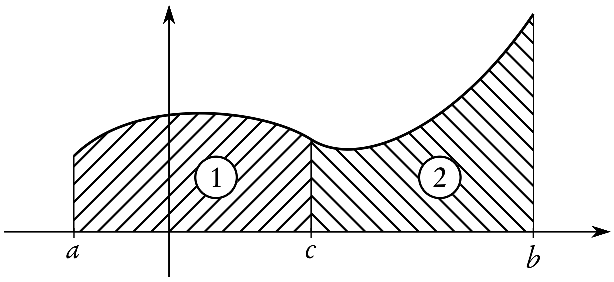


Figura 7.9 *Additività rispetto all'intervallo di integrazione*

Un importante risultato che si può dedurre come conseguenza delle definizioni che abbiamo dato è il metodo per il calcolo dell'area di una regione piana compresa tra i grafici di due funzioni, ed eventualmente due rette verticali, come nella figura che segue.

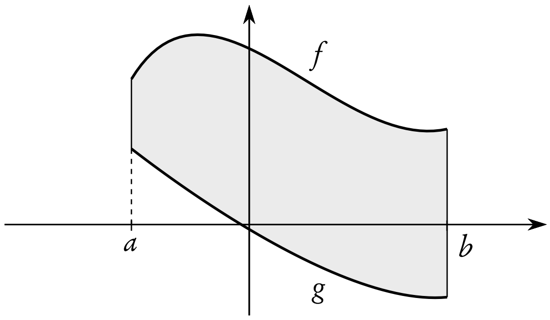


Figura 7.10 *Regione piana compresa tra i grafici di due funzioni*

L'area della regione in questione è *sempre* data dall'integrale, tra a e b ($a < b$), della differenza fra la funzione "più alta" e quella "più bassa"; nell'esempio della figura 7.10 si ha:

$$\int_a^b \big(f(x) - g(x)\big)\,\mathrm{d}x\,.$$

Può succedere che i due grafici si intersechino, come nella figura che segue.

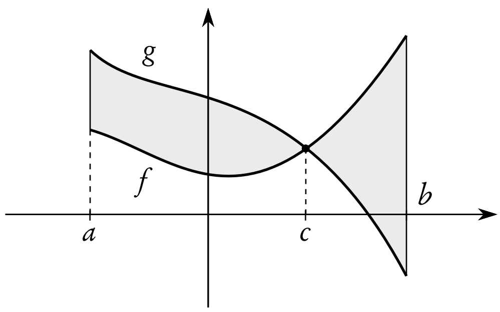


Figura 7.11 *Regione piana compresa tra i grafici di due funzioni*

In questo caso occorrerà spezzare il calcolo in due parti:

$$Area = \int_a^c \big(g(x) - f(x)\big)\,\mathrm{d}x + \int_c^b \big(f(x) - g(x)\big)\,\mathrm{d}x\,.$$

7.7 Il calcolo degli integrali definiti

Il calcolo effettivo degli integrali definiti usando la definizione è possibile solo in casi molto semplici e a prezzo di calcoli abbastanza complessi. Per fortuna viene in nostro aiuto il già citato teorema fondamentale del calcolo integrale che generalizza una proprietà che abbiamo già visto nei primi esempi di questo capitolo.

Teorema 7.5 (Teorema fondamentale del calcolo integrale)**.** *Sia data una funzione f definita e continua in un intervallo I. Sia inoltre F una primitiva di f in I. Se a e b sono due punti qualunque di I, si ha*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)\,.$$

E' chiaro il motivo per cui questo teorema si chiama *fondamentale*: in sostanza esso riconduce il calcolo dell'area di una regione piana (abbastanza) arbitraria a quello della ricerca di una primitiva, ricerca comunque non facile in generale, ma fattibile "a mano" in casi anche complessi.

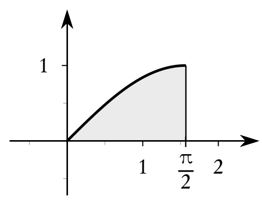
La formula precedente si usa scrivere, tradizionalmente, nel modo seguente:

$$\int_a^b f(x)\,\mathrm{d}x = \big[F(x)\big]_a^b\,.$$

Esempio. Calcolare l'area racchiusa tra il grafico della funzione $\sin x$, l'asse delle x, e le rette $x = 0$ e $x = \pi/2$.

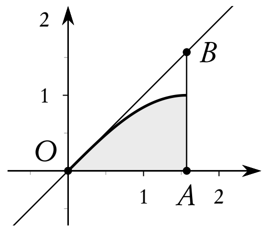
L'area richiesta è quella di un normale trapezoide, per cui si dovrà calcolare

$$\int_0^{\pi/2} \sin x\,\mathrm{d}x\,.$$



Poiché come è noto, una primitiva di $\sin x$ è $-\cos x$, applicando il teorema fondamentale si troverà subito

$$\int_0^{\pi/2} \sin x\,\mathrm{d}x = \big[-\cos x\big]_0^{\pi/2} = -\cos\frac{\pi}{2} - (-\cos 0) = 1\,.$$



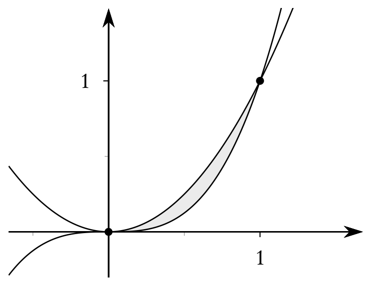
Si può fare un controllo "visivo" della bontà di questo risultato nel seguente modo. La funzione $\sin x$ ha come derivata $\cos x$, per cui, nell'origine, la tangente ha coefficiente angolare $\cos 0 = 1$ e conseguentemente equazione $y = x$ (la bisettrice del primo e terzo quadrante). Dunque l'area del trapezoide dovrà essere un po' più piccola di quella del triangolo OAB della figura a lato. L'area del triangolo è

$$\frac{1}{2}\cdot\frac{\pi}{2}\cdot\frac{\pi}{2}=\frac{\pi^2}{8}\simeq 1.2337\,,$$

per cui tutto torna. A voler essere sofisticati si poteva anche tracciare la tangente al grafico nel punto $C = (\pi/2, 0)$, determinarne l'intersezione D con la bisettrice e calcolare l'area del triangolo BCD, che va sottratta all'area di OAB per avere una stima più precisa dell'area cercata. Lasciamo al lettore questo (facile!) calcolo che fornisce il seguente risultato

$$Area(BCD)=\frac{1}{2}\left(\frac{\pi}{2}-1\right)^2\simeq 0.1629$$

Una migliore approssimazione dell'area del trapezoide sarà allora $1.2337 - 0.1629 = 1.0708$, che si avvicina sensibilmente al valore 1 trovato con l'integrale, confermando la bontà del risultato.



Esempio. Calcolare l'area della regione limitata racchiusa tra i grafici di x^2 e x^3 e appartenente al primo quadrante.

La figura qui a lato mostra che nella regione limitata di piano considerata si ha $0 \le x \le 1$, e che in questo tratto la funzione che sta più in alto è x^2, quella che sta più in basso è x^3. L'area richiesta sarà dunque

$$\int_0^1 (x^2-x^3)\,\mathrm{d}x\,.$$

Poiché

$$\int (x^2-x^3)\,\mathrm{d}x=\frac{x^3}{3}-\frac{x^4}{4}+c\,,$$

si avrà

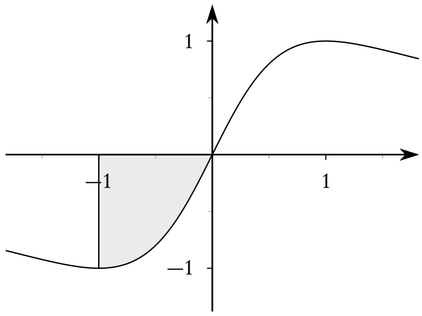
$$\int_0^1 (x^2-x^3)\,\mathrm{d}x=\left[\frac{x^3}{3}-\frac{x^4}{4}\right]_0^1=\left(\frac{1}{3}-\frac{1}{4}\right)-\left(\frac{0}{3}-\frac{0}{4}\right)=\frac{1}{12}\,.$$

Esempio. Calcolare l'area della parte di piano cartesiano individuato dalle condizioni.

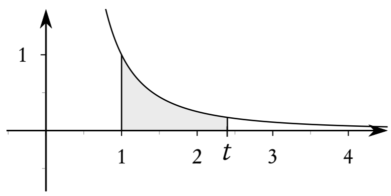
$$\begin{cases} 0\le y\le \dfrac{2x}{x^2+1} \\ -1\le x\le 0 \end{cases}$$

Basta osservare che, essendo la funzione sempre negativa nel tratto in questione, il trapezoide sta sotto l'asse delle ascisse, per cui la sua area sarà data da

$$-\int_{-1}^{0} \frac{2x}{x^2+1}\,dx = \int_{0}^{-1} \frac{2x}{x^2+1}\,dx = \left[\ln(x^2+1)\right]_0^{-1} = \ln 2 - \ln 1 = \ln 2 \simeq 0.69315\,.$$



Esempio. Calcolare l'area della regione limitata di piano compresa tra il grafico della funzione $f(x) = 1/x^2$, l'asse delle x, la retta $x = 1$ e la retta $x = t$, essendo t un numero reale strettamente maggiore di 1. Calcolare poi il limite di quest'area quanto $t \to +\infty$. Che interpretazione geometrica si può dare di questo risultato? Si tratta di un risultato intuitivamente evidente o "difficile da digerire"?



Con il teorema fondamentale il calcolo dell'area (si tratta di un normale trapezoide) è molto semplice:

$$\int_{1}^{t} \frac{1}{x^2}\,dx = \int_{1}^{t} x^{-2}\,dx = \left[\frac{x^{-1}}{-1}\right]_1^t = \frac{-1}{t} + 1\,.$$

Anche il limite per $t \to +\infty$ è immediato e fornisce come risultato il numero 1. Sulla base di quanto finora detto si può interpretare questo numero come l'*area* della regione illimitata di piano compresa tra il grafico della funzione, l'asse delle x e la retta $x = 1$. La cosa sorprendente è che quest'area risulta finita, pur riferendosi a una regione illimitata: questo fatto dà un chiaro significato geometrico al fatto che, da un certo punto in poi il grafico di $f(x) = 1/x^2$ è "talmente attaccato" all'asse x, da non lasciare praticamente scampo alcuno.

Se si ripete tutto il calcolo con la funzione $g(x) = 1/x$ al posto di $f(x)$ si ottiene:

$$\int_{1}^{t} \frac{1}{x}\,dx = \left[\ln|x|\right]_1^t = \ln|t| - \ln 1 = \ln t\,.$$

Adesso il limite per $t \to +\infty$ fornisce come risultato $+\infty$, e questa volta si può dire che la funzione $g(x) = 1/x$ è molto più "lontana" dall'asse x di quanto non sia la funzione $f(x) = 1/x^2$. La figura che segue illustra chiaramente questa situazione.

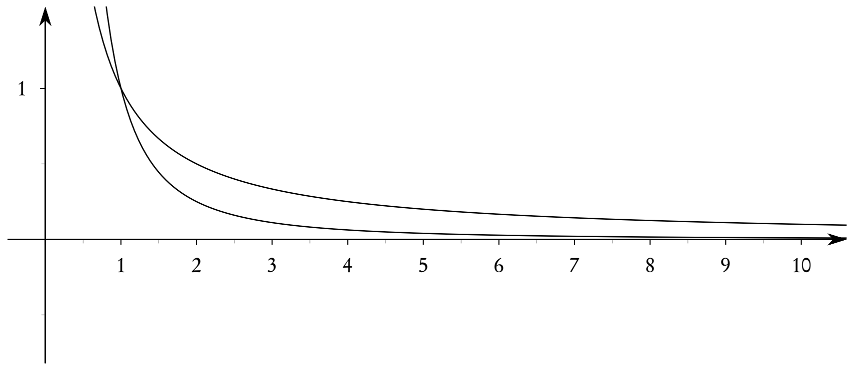


Figura 7.12 *Confronto fra i grafici di* $g(x) = 1/x$ *e* $f(x) = 1/x^2$

7.8 Integrali impropri

L'ultimo esempio del paragrafo precedente ci porta a considerare il problema di calcolare l'area di regioni illimitate del piano cartesiano. Per quanto ci riguarda saremo interessati a seguenti due tipi di problemi.

1. Calcolare l'area di regioni piane comprese tra l'asse delle ascisse e il grafico di una funzione avente un asintoto verticale in uno dei due estremi del suo intervallo di definizione; la funzione sarà naturalmente continua in $]a,b]$, oppure in $[a,b[$, a seconda che l'asintoto verticale sia sull'estremo sinistro o destro. Per esempio saremo interessati al problema di valutare l'area della regione compresa il grafico della funzione $f(x) = 1/x^2$ e l'asse delle x nell'intervallo $]0,1]$, che è l'esempio trattato precedentemente.
2. Calcolare l'area di regioni piane comprese tra il grafico di una funzione continua e l'asse delle x, in un intervallo del tipo $[a,+\infty[$, oppure $]-\infty,a]$.

In questi casi si parla di *integrali impropri*, e la tecnica è la stessa già applicata nell'esempio precedentemente considerato.

Per il primo caso supponiamo che l'asintoto verticale sia nell'estremo sinistro dell'intervallo di definizione della funzione. Allora si definisce:

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{t\to a^+} \int_t^b f(x)\,\mathrm{d}x\,. \tag{7.3}$$

Se invece l'asintoto si trova nell'estremo destro avremo

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{t\to b^-} \int_a^t f(x)\,\mathrm{d}x\,. \tag{7.4}$$

Per il secondo caso scriveremo (e i simboli sono ormai ovvi)

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \lim_{t\to +\infty} \int_a^t f(x)\,\mathrm{d}x \quad , \quad \int_{-\infty}^b f(x)\,\mathrm{d}x = \lim_{t\to -\infty} \int_t^b f(x)\,\mathrm{d}x\,. \tag{7.5}$$

Come mostra il già citato esempio queste aree possono essere finite o infinite.

7.9 Esercizi

Esercizio 7.1. *Calcolare i seguenti integrali indefiniti.*

1. $\displaystyle\int (x-\mathrm{e}^x)\,\mathrm{d}x\,;$
2. $\displaystyle\int \frac{2x-3}{x}\,\mathrm{d}x\,;$
3. $\displaystyle\int \frac{3}{x+1}\,\mathrm{d}x\,;$
4. $\displaystyle\int \left(\frac{1}{x}-\mathrm{e}^{3x}+1\right)\mathrm{d}x\,;$
5. $\displaystyle\int \frac{3x-2}{2x+1}\,\mathrm{d}x\,;$
6. $\displaystyle\int x\mathrm{e}^{3x^2}\,\mathrm{d}x\,;$
7. $\displaystyle\int 3x\mathrm{e}^{x}\,\mathrm{d}x\,;$
8. $\displaystyle\int x^2\mathrm{e}^{x}\,\mathrm{d}x\,;$
9. $\displaystyle\int x^2\ln x\,\mathrm{d}x\,;$
10. $\displaystyle\int \frac{x+2}{x}\,\mathrm{d}x\,;$
11. $\displaystyle\int x\sin x\,\mathrm{d}x\,;$
12. $\displaystyle\int 2x-3\sin(2x)\,\mathrm{d}x\,;$
13. $\displaystyle\int \left(\frac{2}{x}-3x+2x^2\right)\mathrm{d}x\,;$
14. $\displaystyle\int x^3\cos x\,\mathrm{d}x\,.$

Esercizio 7.2. *Determinare le somme integrali inferiori e superiori per le funzioni indicate, nell'intervallo a fianco segnato, e con il numero n di suddivisioni assegnato.*

1. $f(x)=x^2-1$, *su* $[-3,1]$, *con* $n=4$;

2. $f(x) = -x^3 - 2$, *su* $[-5,3]$, *con* $n = 2$;
3. $f(x) = \sqrt{-x}$, *su* $[-4,-3]$, *con* $n = 2$;
4. $f(x) = \ln(x-1)$, *su* $[2,4]$, *con* $n = 3$;
5. $f(x) = 2^{x-1}$, *su* $[-5,1]$, *con* $n = 3$;
6. $f(x) = x^2 - 1$, *su* $[0,2]$, *con* $n = 2$.

Esercizio 7.3. *Calcolare i seguenti integrali definiti.*

1. $\displaystyle\int_0^2 (2x-3)\,\mathrm{d}x$;
2. $\displaystyle\int_1^3 (2\ln x - 3x)\,\mathrm{d}x$;
3. $\displaystyle\int_0^1 2\frac{x}{x-2}\,\mathrm{d}x$;
4. $\displaystyle\int_{-1}^3 (x+x^4)\,\mathrm{d}x$.

Esercizio 7.4. *Per ciascuna delle funzioni seguenti, definite a pezzi, calcolare gli integrali indicati.*

1. $f(x) = \begin{cases} x, & \text{se } x \leq 0 \\ -2x-1, & \text{se } x > 0 \end{cases}$, $\displaystyle\int_0^2 f(x)\,\mathrm{d}x$ *e* $\displaystyle\int_{-2}^2 f(x)\,\mathrm{d}x$;
2. $f(x) = \begin{cases} x^2+x, & \text{se } x \leq 0 \\ -x^3, & \text{se } x > 0 \end{cases}$, $\displaystyle\int_{-1}^1 f(x)\,\mathrm{d}x$;
3. $f(x) = \begin{cases} \sqrt{x}-2x, & \text{se } x \leq 1 \\ -x^2, & \text{se } x > 1 \end{cases}$, $\displaystyle\int_0^2 f(x)\,\mathrm{d}x$;
4. $f(x) = \begin{cases} \mathrm{e}^x, & \text{se } x \leq 1 \\ \ln x, & \text{se } x > 1 \end{cases}$, $\displaystyle\int_{-1}^3 f(x)\,\mathrm{d}x$;
5. $f(x) = \begin{cases} \sqrt{x}, & \text{se } x \leq 1 \\ \sqrt[3]{x}, & \text{se } x > 1 \end{cases}$, $\displaystyle\int_0^3 f(x)\,\mathrm{d}x$;
6. $f(x) = \begin{cases} \mathrm{e}^{-x}+1, & \text{se } x < -2 \\ -\ln(-x)+2x, & \text{se } x \geq -2 \end{cases}$, $\displaystyle\int_{-3}^{-1} f(x)\,\mathrm{d}x$.

Esercizio 7.5. *Dire se esistono, ed eventualmente calcolarli, i seguenti integrali impropri.*

1. $\displaystyle\int_2^{+\infty} (2x-3)\,\mathrm{d}x$;
2. $\displaystyle\int_{-1}^3 \frac{2+x}{x+1}\,\mathrm{d}x$;

3. $\displaystyle\int_4^{+\infty} \frac{x+1}{x-3}\,\mathrm{d}x;$

4. $\displaystyle\int_4^{+\infty} \left(\frac{1}{x^2}-\frac{1}{x^3}\right)\mathrm{d}x;$

5. $\displaystyle\int_0^{2} \left(\frac{1}{x^2}-\frac{1}{x^3}\right)\mathrm{d}x;$

6. $\displaystyle\int_{-1}^{1} \frac{1}{x^2}\,\mathrm{d}x.$

8 Funzioni di due variabili

Abbiamo già parlato brevemente di funzioni di due variabili nel paragrafo 1.6 del capitolo 1. Ora torniamo sull'argomento per una trattazione più dettagliata, cominciando con una introduzione *visuale* al problema.

8.1 Introduzione illustrata

Richiamiamo alcuni concetti fondamentali relativi alla rappresentazione delle funzioni di una variabile, fissando l'attenzione su quanto sarà utile per affrontare con sicurezza il caso di due variabili. Se consideriamo la funzione che ad ogni numero reale x fa corrispondere la sua metà, possiamo costruire una *tabella a doppia entrata* in cui su una colonna mettiamo il valore di x (variabile indipendente) e sull'altra il corrispondente valore di $y = f(x)$ (variabile dipendente). Naturalmente potremo scrivere esplicitamente la tabella solo in corrispondenza a un numero finito di valori di x, per esempio per alcuni valori presi sui numeri naturali, come nella tabella 1.1 della pagina 10, tabella che qui riportiamo per comodità.

x	$x/2$
1	$1/2$
2	1
3	$3/2$
4	2
5	$5/2$

Tabella 8.1 *Rappresentazione "tabulare" di una funzione di una variabile*

I dati di questa tabella possono essere riportati in un grafico cartesiano, come nella figura 1.5 della pagina 11; riportiamo qui anche questa figura per comodità.

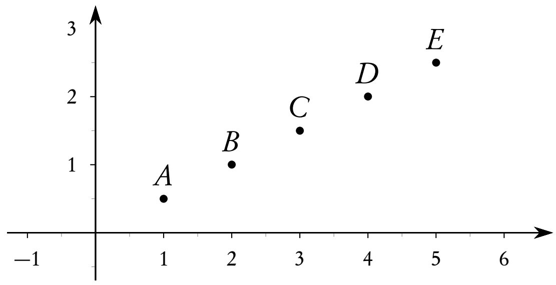


Figura 8.1 *Grafico cartesiano relativo alla tabella 8.1*

Come già accennato, questo grafico può essere desunto *compattando* un grafico "a frecce": da ogni punto x dell'asse delle ascisse facciamo partire una freccia verticale fino alla "quota" $f(x)$, cioè fino al punto $\big(x, f(x)\big)$; a partire da questa quota la freccia "piega" orizzontalmente fino a incontrare l'asse delle y esattamente in corrispondenza del valore $f(x)$, come nella figura seguente (già considerata nella pagina 12).

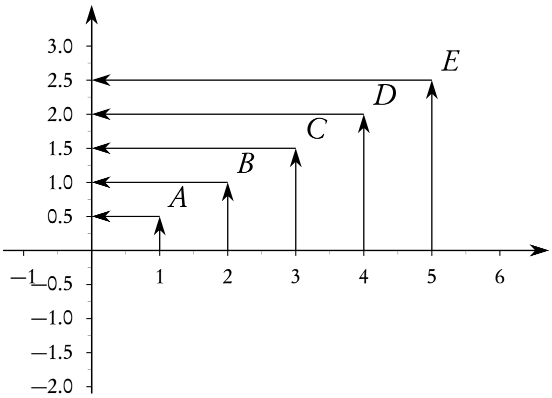


Figura 8.2 *Grafico cartesiano con frecce, relativo alla tabella 8.1*

Se si riportano nel grafico della figura 8.1 anche i punti corrispondenti ai valori di x che non compaiono nella tabella, si ottiene il risultato visualizzato nella figura seguente: i punti rappresentativi non si dispongono casualmente nel piano, ma su una linea, in questo caso su una linea retta, in casi più generali su una linea più complessa, come abbiamo già avuto modo di constatare.

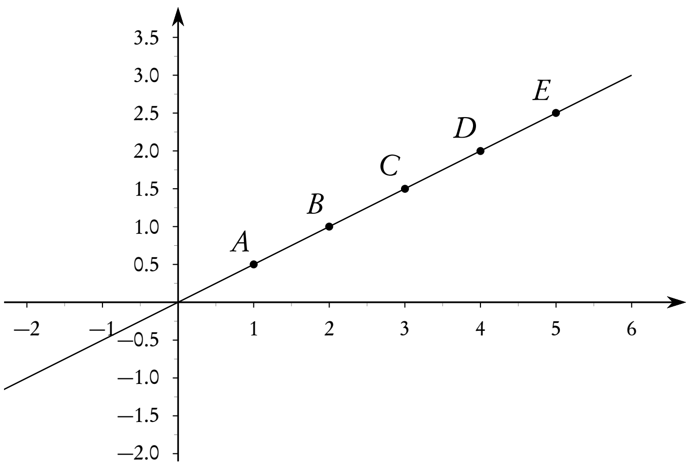


Figura 8.3 *Grafico della funzione $y = x/2$, comprendente i punti della figura 8.1*

Se consideriamo ora una funzione di due variabili, per esempio $f(x, y) = x + y$, potremo ancora costruire una tabella come la 8.1, ma dovremo utilizzare *tre* colonne: *due* per le variabili indipendenti e *una* per la variabile dipendente. Naturalmente anche qui la tabella potrà essere effettivamente costruita solo per alcune coppie di valori (x, y).

x	y	$x+y$
1	0	1
0	1	1
1	1	2
1	-1	0
2	-1	1
3	1	4
...	...	...

Tabella 8.2 *Rappresentazione "tabulare" di una funzione di due variabili*

Quello che si ottiene è un insieme di *terne* di numeri e le terne di numeri possono essere rappresentate nello spazio dove si sia introdotto un sistema di 3 assi cartesiani ortogonali, $Oxyz$.

Scegliamo, come è tradizione, di rappresentare le coppie (x,y) che stanno nel dominio di f sul piano Oxy. Da ciascuno di questi punti facciamo partire una freccia verticale fino alla "quota" $f(x,y)$, cioè fino al punto $\big(x,y,f(x,y)\big)$; a partire da questa quota la freccia "piega" orizzontalmente fino a incontrare l'asse z in corrispondenza al valore $f(x,y)$, come mostra la figura seguente per un singolo punto (x,y) del dominio.

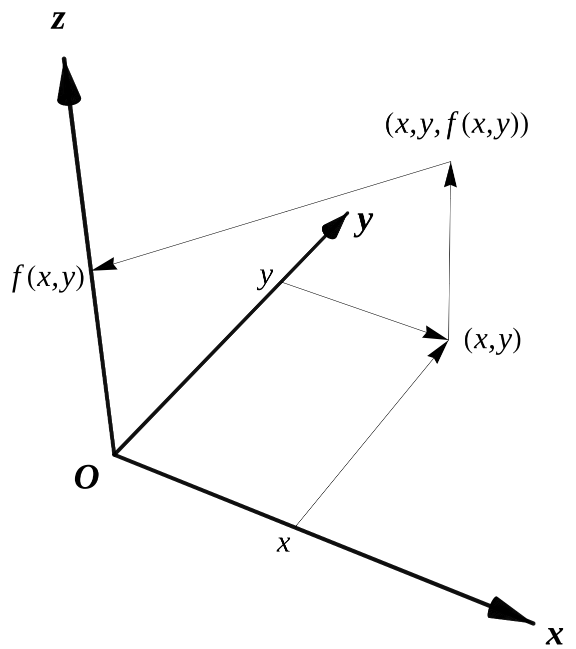


Figura 8.4 *Procedimento per tracciare il grafico di una funzione di due variabili*

Naturalmente, come già per le funzioni di una variabile, scegliamo alcuni punti nel dominio, per esempio quelli individuati da una griglia tracciata nel piano Oxy, e da ognuno innalziamo la freccia fino alla quota $f(x,y)$: ne viene un boschetto di frecce, come nella figura che segue.

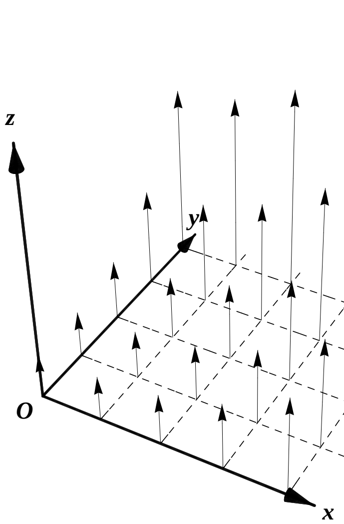


Figura 8.5 *Un "boschetto" di frecce*

Nei casi che interesseranno le punte delle frecce, cioè i punti di coordinate $\big(x, y, f(x, y)\big)$, non si distribuiscono a casaccio nello spazio, ma su una superficie, che possiamo evidenziare per esempio con una "piastrellatura".

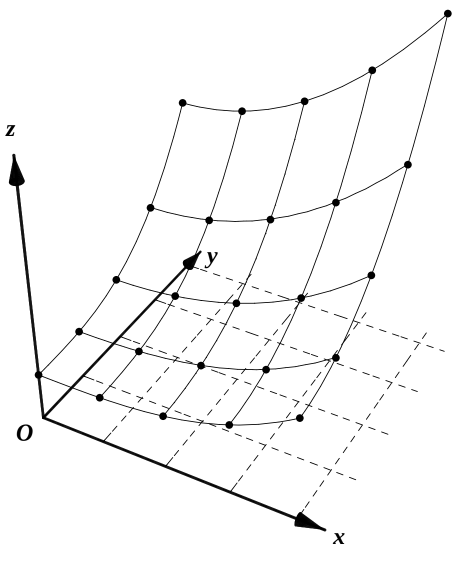


Figura 8.6 *Una superficie-grafico*

Per rendere più significativo il grafico si possono introdurre anche colorazioni come nella figura che segue.

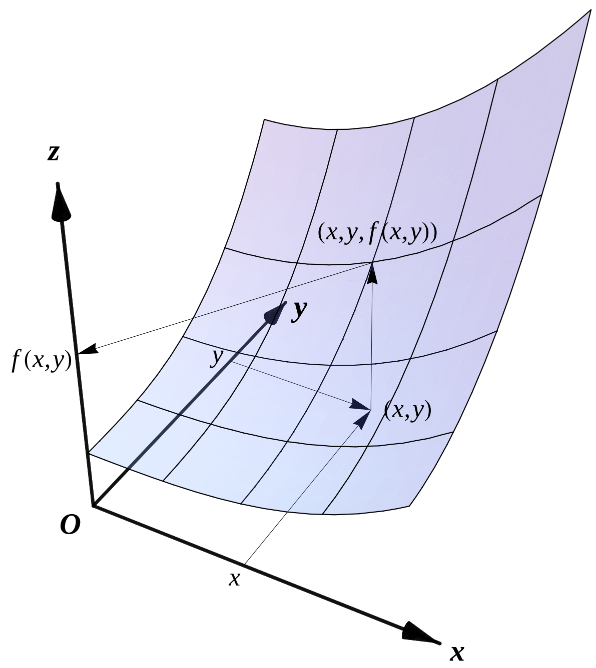


Figura 8.7 *Uso di colorazioni per le superfici-grafico*

Non tutte le caratteristiche che si evidenziano nel grafico delle funzioni di una variabile potranno essere trasferite ai grafici di funzioni di due variabili; per esempio non avrà alcun senso parlare di *crescenza* o *decrescenza*, mentre potremo ancora considerare (e la cosa sarà per noi della massima importanza) i concetti di massimo e minimo (relativo o assoluto). Come suggerisce la figura 1.20 della pagina 21, potremo usare l'appellativo *monte* e *cima* per riferirci ai massimi, l'appellativo *valle* e *fondovalle* per riferirci ai minimi.

La figura 8.8 mostra, come ulteriore esempio, una situazione in cui sono presenti due monti e una valle. In questa figura non sono tracciati gli assi, per non complicare il grafico: è una scelta che si fa normalmente nei grafici tridimensionali, dove si racchiude la parte di superficie che interessa in un "box", riportando sugli spigoli i valori delle variabili sui tre assi.

A volte, invece di tracciare sulla superficie una piastrellatura che riproduca la griglia del piano Oxy, conviene tracciare altre linee. Una delle scelte più comuni è quella delle *linee di livello*, o linee di quota: si tratta di evidenziare sulla superficie tutti i punti che si trovano a una determinata quota, punti che nelle situazioni comuni si distribuiscono su una linea che si può pensare ottenuta intersecando la superficie con un piano orizzontale (parallelo al piano Oxy). La figura 8.9 mostra alcune di queste linee per la stessa superficie della figura 8.8.

La considerazione delle linee di livello consente di costruire un rappresentazione grafica "bidimensionale" della stessa superficie: sarà sufficiente "raccogliere" tutte queste linee sul piano Oxy e magari usare colori via via più chiari per indicare le cime e via via più scuri per indicare le valli. Si tratta della convenzione che viene normalmente adottata nelle carte geografiche. Si può vedere questa rappresentazione per la stessa superficie della figura 8.8 nella figura 8.10.

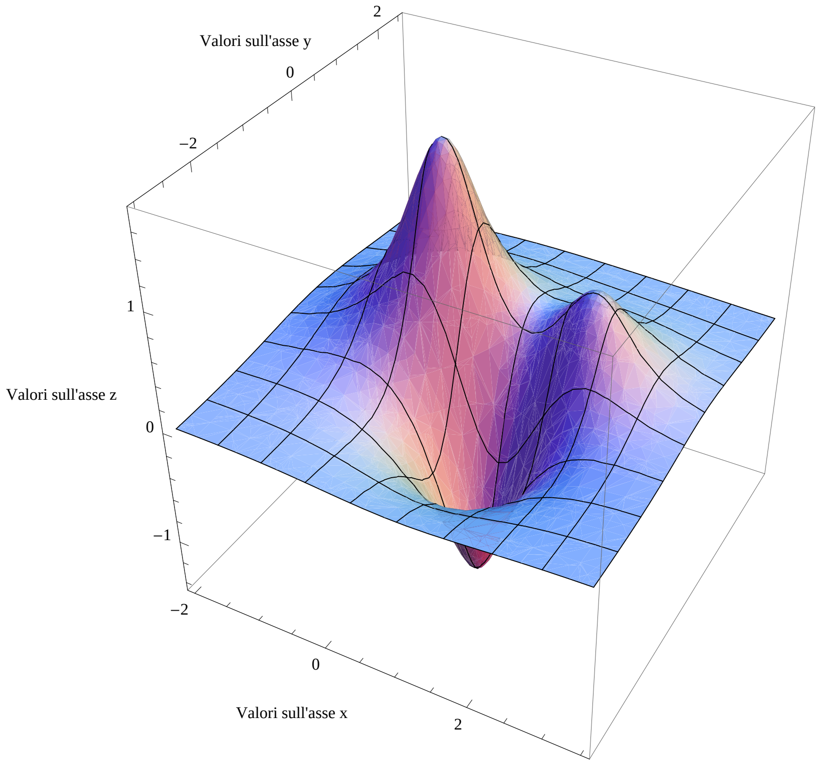


Figura 8.8 *Una funzione con due "monti" e una "valle"*

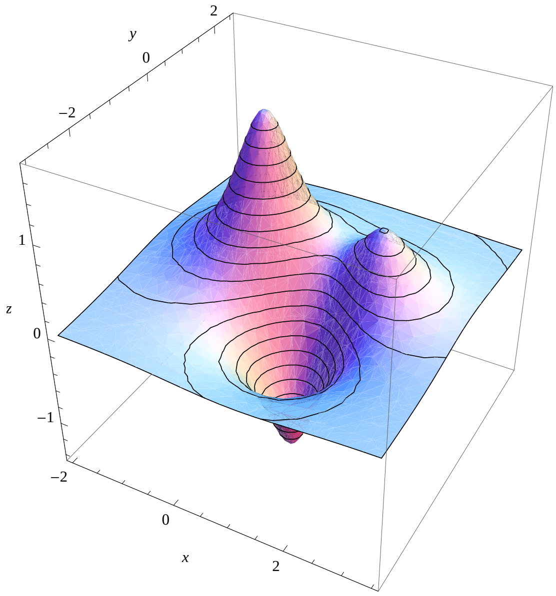


Figura 8.9 *Linee di livello*

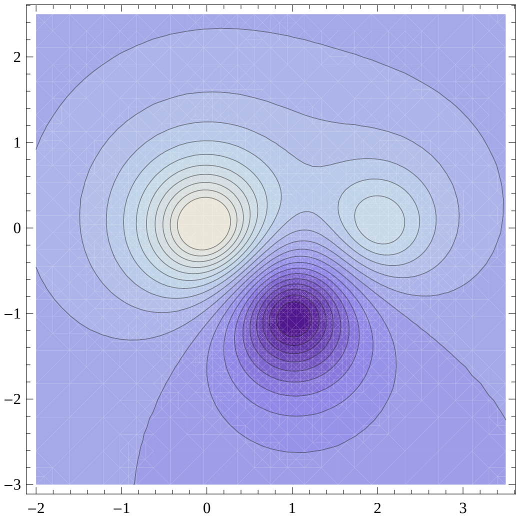


Figura 8.10 *Linee di livello raccolte sul piano Oxy*

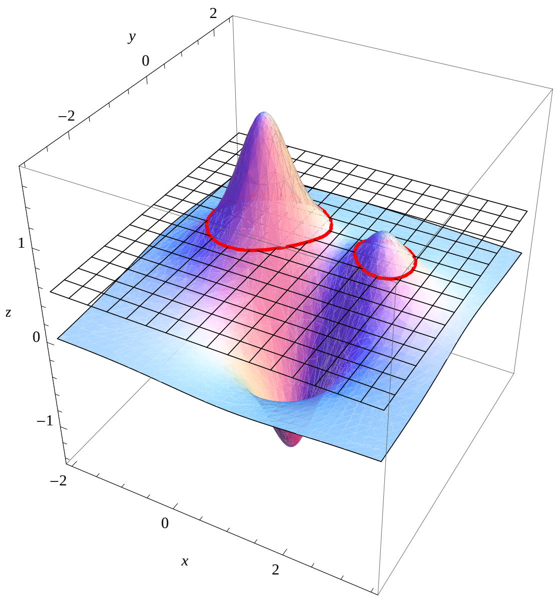


Figura 8.11 *Sezione di una superficie con un piano orizzontale*

La figura 8.11 mostra come si ottiene una delle linee di livello mediante intersezione della superficie con un piano orizzontale.

Ritornando alla piastrellatura della figura 8.8, possiamo osservare che le linee della piastrellatura non sono altro che le intersezioni della superficie con piani verticali paralleli o al piano Oxz o al piano Oyz.

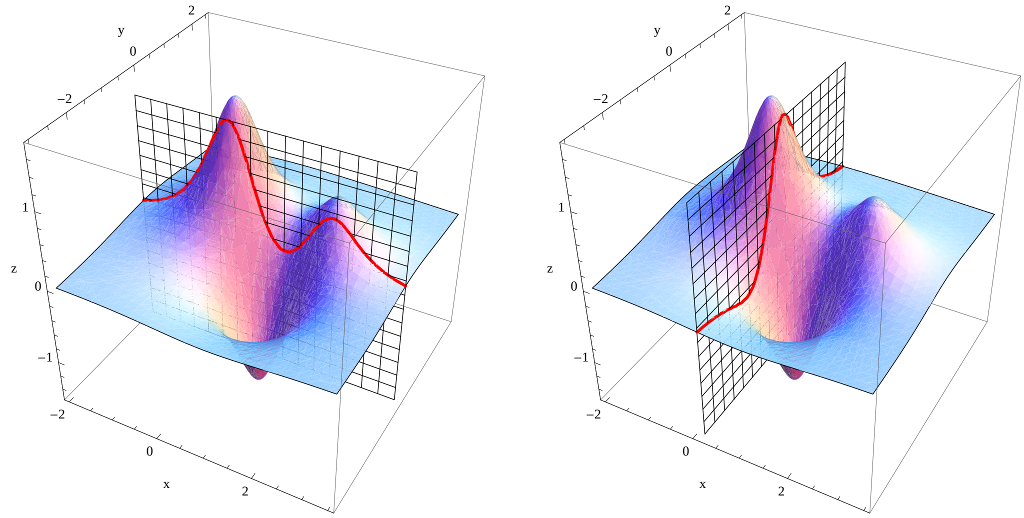


Figura 8.12 *Sezione di una superficie con piani verticali paralleli a Oxz e a Oyz*

Nel seguito saremo interessati a considerare anche questo tipo di sezioni.

Osserviamo anche esplicitamente che i massimi e minimi per funzioni di due variabili godono di proprietà grafiche simili a quelle delle funzioni di una variabile: per le funzioni di una variabile (opportunamente regolari e in particolare senza spigoli) nei massimi e minimi interni al dominio la retta tangente al grafico risultava orizzontale, ovvero parallela all'asse x; per le funzioni di due variabili (sempre opportunamente regolari) nei massimi e minimi interni al dominio sarà il piano tangente ad essere orizzontale, cioè parallelo al piano Oxy. Le immagini della figura 8.13 mostrano i piani tangenti in corrispondenza di un massimo e di un minimo; la prima immagine mostra la superficie vista dall'alto, la seconda vista dal basso, per evidenziare meglio i piani tangenti.

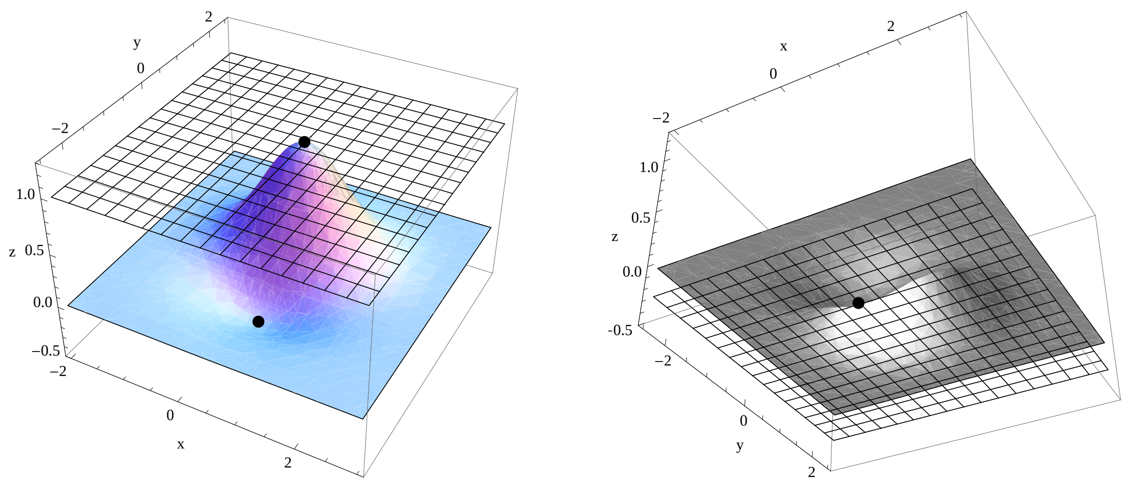


Figura 8.13 *Piani tangenti in un punto di massimo e in un punto di minimo*

Trattando le funzioni di una variabile, oltre ai massimi e minimi, abbiamo considerato anche i flessi a tangente orizzontale (come caso particolare di quelli a tangente obliqua). Non esiste nulla di simile per le funzioni di due variabili, nella quali però compare un fenomeno completamente nuovo: i *punti di sella*, dove, come vedremo, la situazione è decisamente più complessa che non con i flessi in una variabile.

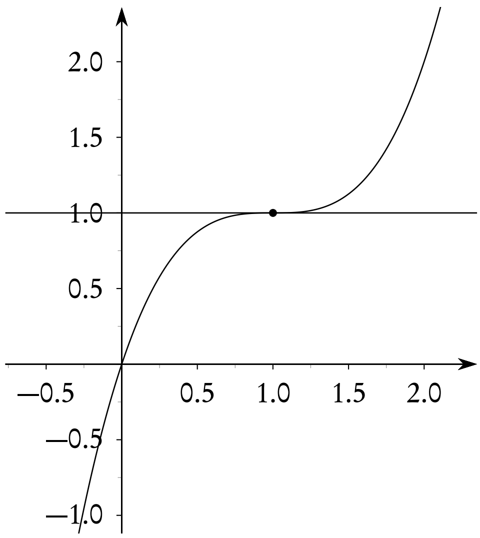


Figura 8.14 *Un flesso a tangente orizzontale*

Per le funzioni di una variabile l'idea fondamentale (per funzioni regolari) è che un punto di flesso (in particolare a tangente orizzontale) è un punto dove si ha un cambio di concavità.

Completamente diversa la situazione per funzioni di due variabili: si definisce punto di sella un punto in cui il piano tangente è orizzontale e in cui vale la seguente proprietà: se passiamo per il punto in certe direzioni il punto si presenta come un massimo, mentre in certe direzioni si presenta come un minimo.

Geograficamente un punto di sella corrisponde a un valico di montagna: per chi lo attraversa il valico è il punto più alto, per chi invece segue il crinale da una cima all'altra è il punto più basso.

Il nome *punto di sella* ricorda proprio la sella di un cavallo: il punto in cui il cavaliere è seduto è un massimo nella direzione destra-sinistra, è un minimo nella direzione avanti-dietro. Osserviamo anche che se su una normale sella di cavallo dovesse sedersi una scimmia, essa avrebbe difficoltà a sistemare la coda; esistono anche situazioni in cui la superficie ha un punto in cui potrebbe sedersi una scimmia, facendo posto sia alle gambe che alla coda (anche se non si conoscono cavalli su cui fissarla!), e si potrebbe parlare in questo caso di *selle di scimmia*.

La figura 8.15 mostra una sella nel senso ordinario del termine, con evidenziate due direzioni lungo le quali sulla superficie si ha un massimo e un minimo rispettivamente. La figura 8.16 mostra invece un punto a "sella di scimmia" su una superficie, e qui non si hanno direzioni lungo le quali si ha un massimo e direzioni lungo le quali si ha un minimo: dal punto di vista formale la situazione è ancora più complessa.

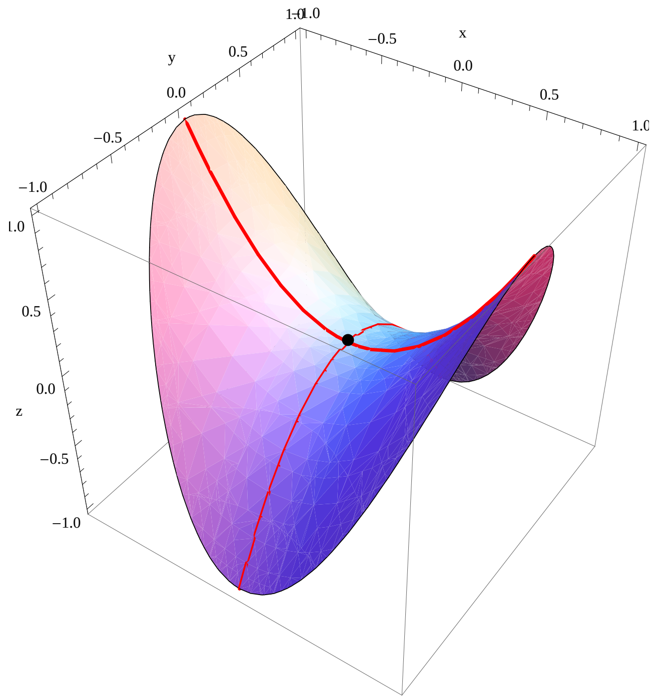


Figura 8.15 *Una "sella di cavallo"*

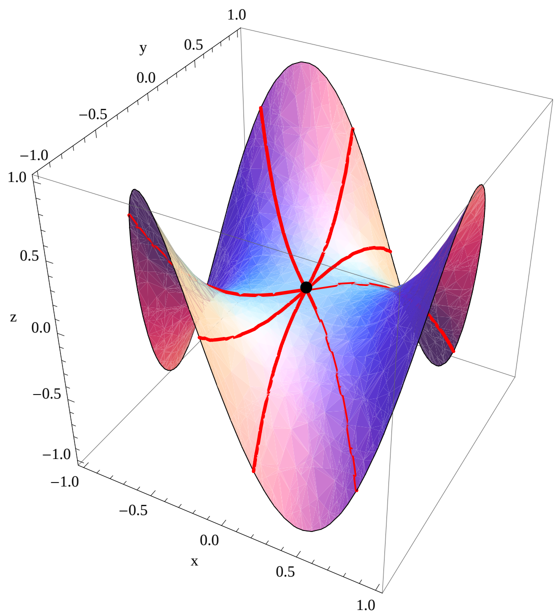


Figura 8.16 *Una "sella di scimmia"*

8.2 Qualche esempio significativo

Proponiamo alcuni esempi di grafici di funzioni di due variabili, che ci saranno utili nel seguito. Le figure rappresentano le superfici sia utilizzando una piastrellatura che curve di livello.

1. Piano $z = 2x + 3y$, o anche $2x + 3y - z = 0$.

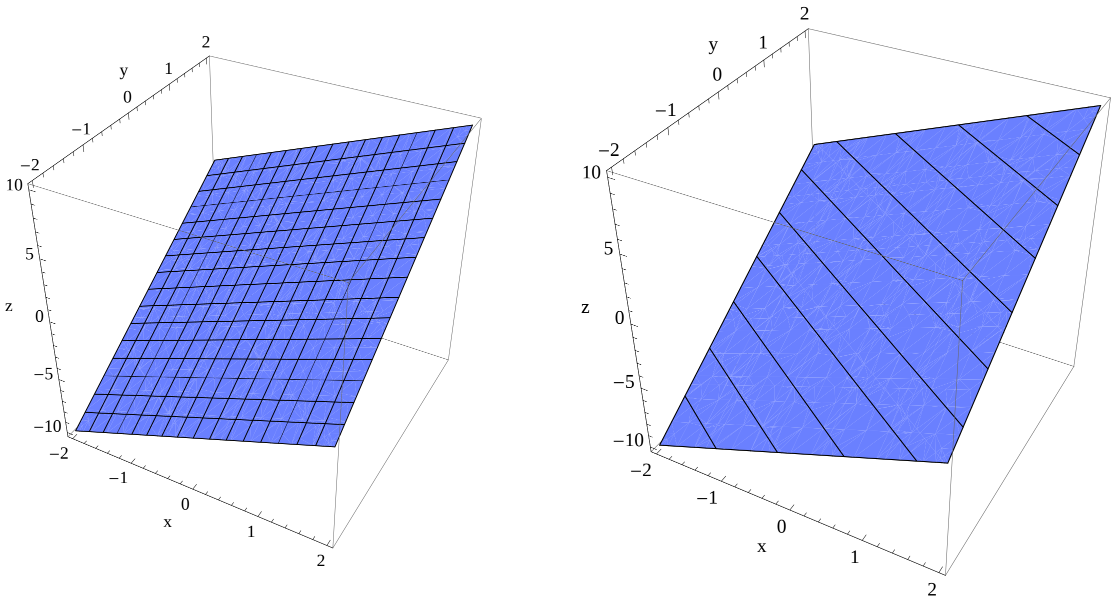


Figura 8.17 *Piano* $z = 2x + 3y$

2. Paraboloide $z = x^2 + y^2$. Si tratta della superficie ottenuta per rotazione della parabola $z = x^2$, attorno all'asse z. Le sue curve di livello sono circonferenze con centro sull'asse z.

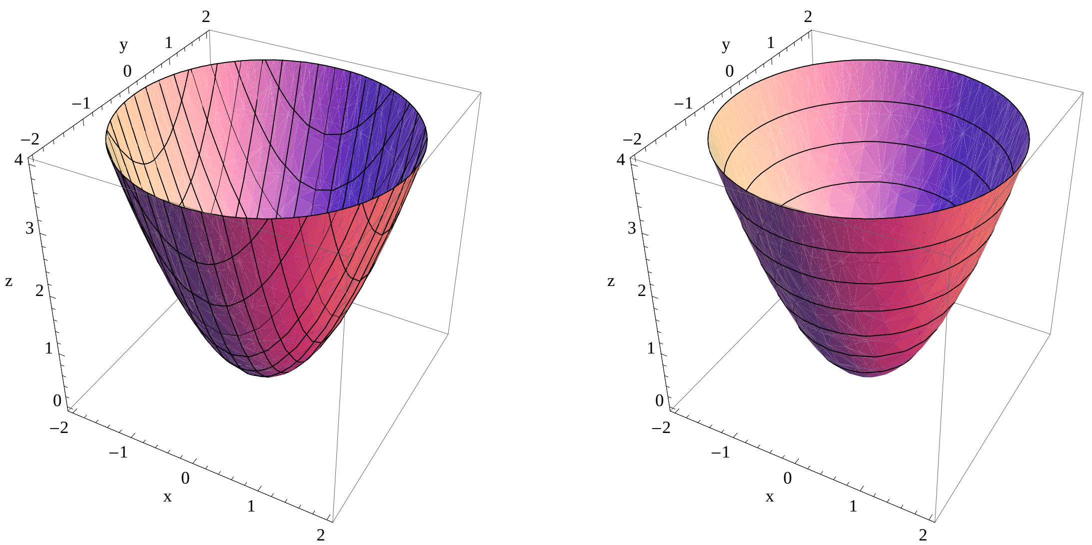


Figura 8.18 *Paraboloide* $z = x^2 + y^2$

3. Paraboloide a sezione ellittica: $z = 3x^2 + y^2$. Superficie simile a quella della figura 8.18, ma con curve di livello a sezione ellittica con centro sull'asse z.

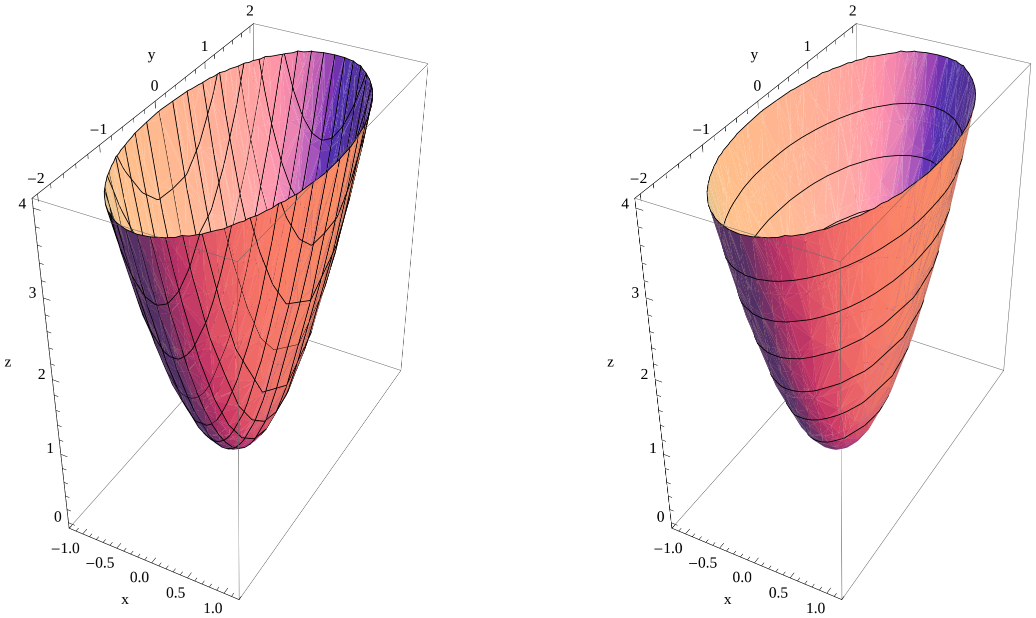


Figura 8.19 *Paraboloide a sezione ellittica:* $z = 3x^2 + y^2$

4. La sella $z = x^2 - y^2$.

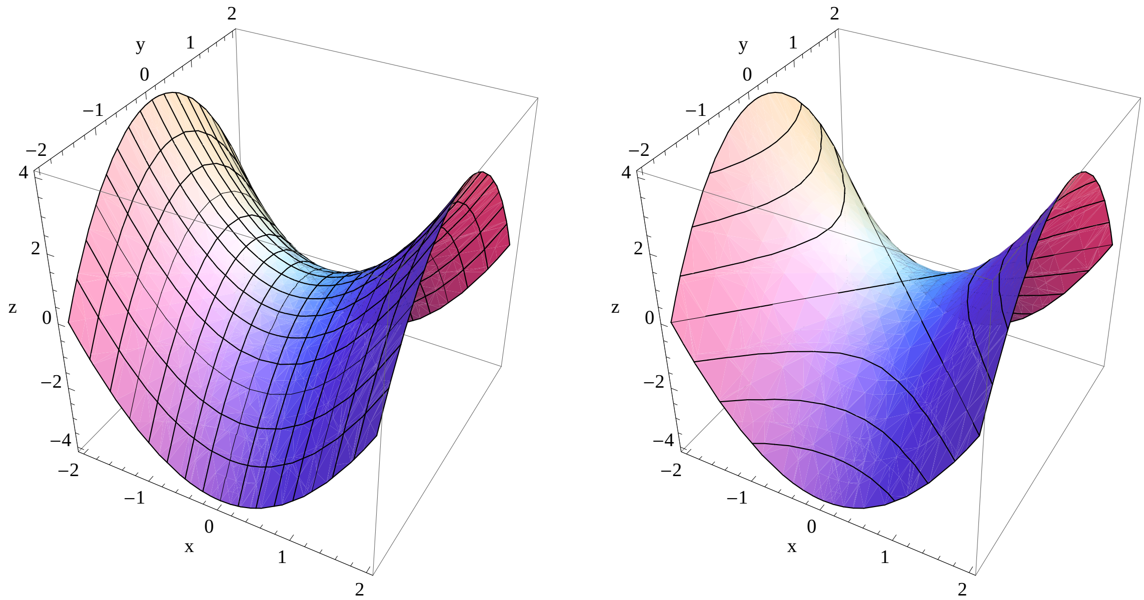


Figura 8.20 *La sella* $z = x^2 - y^2$

5. La superficie $z = x^2$. Si tratta della superficie ottenuta traslando la parabola $z = x^2$ lungo l'asse

delle x.

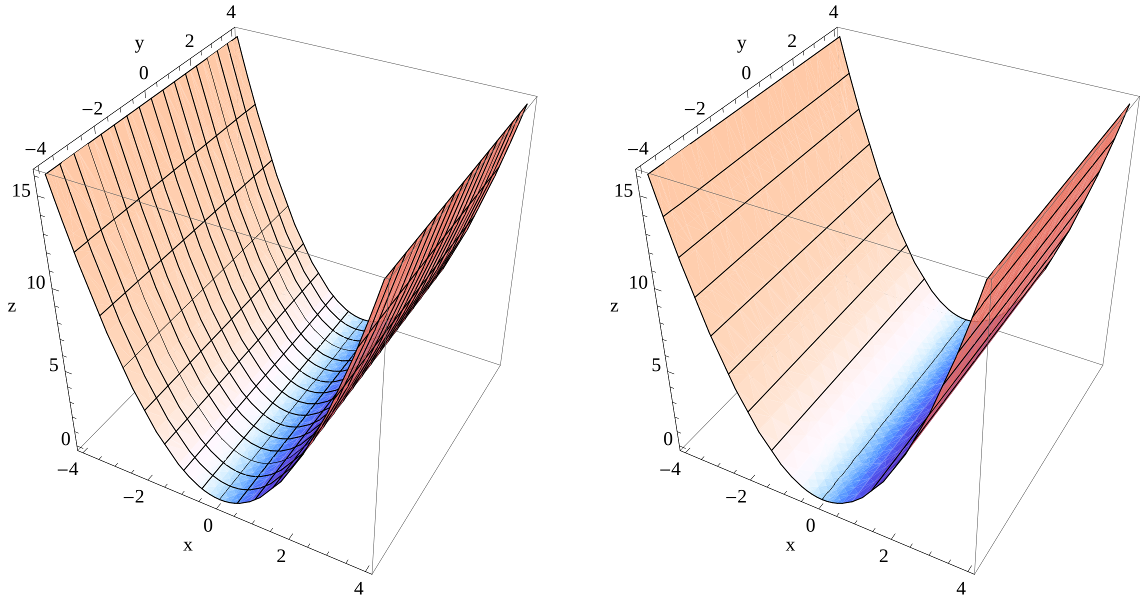


Figura 8.21 *La superficie* $z = x^2$

6. La superficie $z = \mathrm{e}^{x^2+y^2}$. Molto simile a un paraboloide, ma si osservi la grande differenza di unità di misura tra gli assi x e y da un lato e l'asse z dall'altro. Si noti anche che, in questo caso, il vertice si trova a quota 1 sull'asse z, mentre nel paraboloide si trova sull'origine.

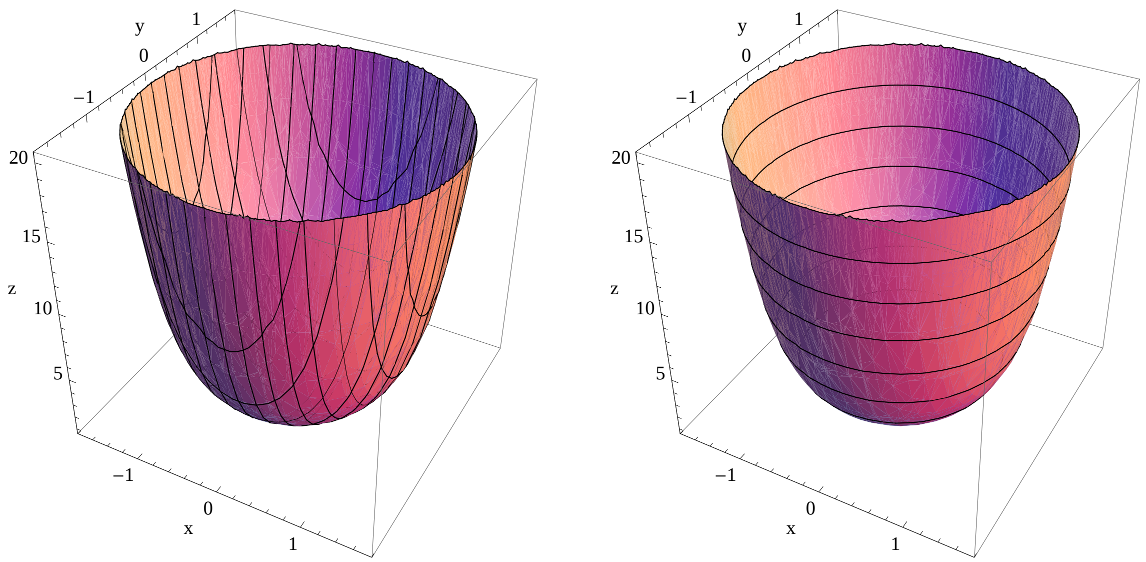


Figura 8.22 *La superficie* $z = \mathrm{e}^{x^2+y^2}$

8.3 Cenno su limiti e continuità

La definizione di limite che abbiamo dato per funzioni di una variabile (vedi la definizione 3.2 nella pagina 64) può essere estesa quasi con le stesse parole anche al caso di funzioni due variabili, in quanto basata solo sul concetto di intorno che abbiamo introdotto nella pagina 4 (definizione 1.8) anche per punti del piano. Unica differenza importante è che nel piano non si possono introdurre i concetti di $+\infty$ e $-\infty$: si può parlare solo genericamente di punti all'∞ (senza segno), e si può chiamare *intorno di* ∞ *nel piano* l'esterno di un qualunque disco centrato sull'origine. Con questa precisazione si può ripetere quasi pari pari la definizione 3.2.

Definizione 8.1 (Limite in due variabili)**.** *Sia data una funzione $f(x,y)$, di dominio D, e sia (x_0,y_0) un punto di accumulazione per D (non essendo escluso che (x_0,y_0) possa essere l'infinito). Diremo che l (non essendo escluso che l possa essere uno dei due simboli di infinito[1]) è il* limite *di $f(x,y)$ per (x,y) tendente a (x_0,y_0), e scriveremo*

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)=l \tag{8.1}$$

se, scelto un arbitrario intorno I_l di l, è possibile trovare in corrispondenza un opportuno intorno $I_{(x_0,y_0)}$ di (x_0,y_0), in modo tale che i valori della funzione calcolati in $I_{(x_0,y_0)}$, tranne *(x_0,y_0) stesso, cadano in I_l.*

Valgono tutti i teoremi sui limiti, opportunamente adattati e in particolare le regole di calcolo sulla retta reale estesa (ricordiamo che le funzioni di due variabili hanno dominio in $\mathbb{R}^2$, ma codominio in $\mathbb{R}$, esattamente come le funzioni di una variabile).

Si può anche introdurre il concetto di funzione continua con una definizione sostanzialmente identica a quella data per le funzioni di una variabile (definizione 3.8 nella pagina 70).

Definizione 8.2 (Continuità in due variabili)**.** *Sia data una funzione $f(x,y)$, di dominio D, e sia (x_0,y_0) un punto di accumulazione per D,* appartenente *a D. La funzione f si dice* continua *in (x_0,y_0) se*

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)=f(x_0,y_0).$$

Anche qui è come dire che una funzione è continua se il calcolo del limite si può fare semplicemente sostituendo (x_0,y_0) al posto di (x,y) nell'espressione della funzione: una bella facilitazione, se si riesce a scoprire a priori quali sono le funzioni continue! E anche qui si può dimostrare che tutte le funzioni costruite con *tecniche elementari* sono continue in tutti i punti del loro dominio.

Purtroppo al di fuori delle funzioni continue il calcolo dei limiti per funzioni di due variabili è estremamente complesso e non alla portata di questo corso, per cui non ce ne occuperemo.

8.4 Piani nello spazio

Ricordiamo che una retta non verticale nel piano ha equazione $y = mx + q$, dove q rappresenta l'ordinata (o quota) all'origine mentre m dà la pendenza o inclinazione della retta rispetto all'asse delle

[1]Il valore l del limite appartiene alla retta reale estesa, in quanto la funzione f ha come codominio $\mathbb{R}$.

x. Una retta verticale (parallela all'asse y) ha invece equazione $x = k$. Le rette orizzontali hanno equazioni del tipo $y = k$, e quindi hanno $m = 0$, cioè pendenza nulla, come è evidente. Per rendersi conto di questi fatti basta pensare che i punti appartenenti a rette verticali hanno tutti la stessa ascissa, mentre quelli appartenenti a rette orizzontali hanno tutti la stessa ordinata. Per memorizzare rapidamente queste proprietà si può osservare che nelle rette parallele all'asse x *manca la* x, in quelle parallele all'asse y *manca la* y.

Passando allo spazio possiamo cominciare a considerare le equazioni di piani paralleli a uno dei piani coordinati, ottenendo $x = k$ per i piani (verticali) paralleli al piano Oyz, $y = k$ per i piani (verticali) paralleli al piano Oxz e infine $z = k$ per i piani (orizzontali), paralleli al piano Oxy. Anche qui per rendersi conto di questi fatti basta tenere conto che se un piano è parallelo, per esempio, al piano Oxz, tutti i suoi punti hanno la stessa y; analogamente per gli altri casi. La figura 8.23 illustra queste tre situazioni. Ancora una volta per memorizzare rapidamente queste proprietà si può osservare che nei piani paralleli al piano Oxy *mancano la* x *e la* y, nei piani paralleli al piano Oxz *mancano la* x *e la* z, nei piani paralleli al piano Oyz *mancano la* y *e la* z.

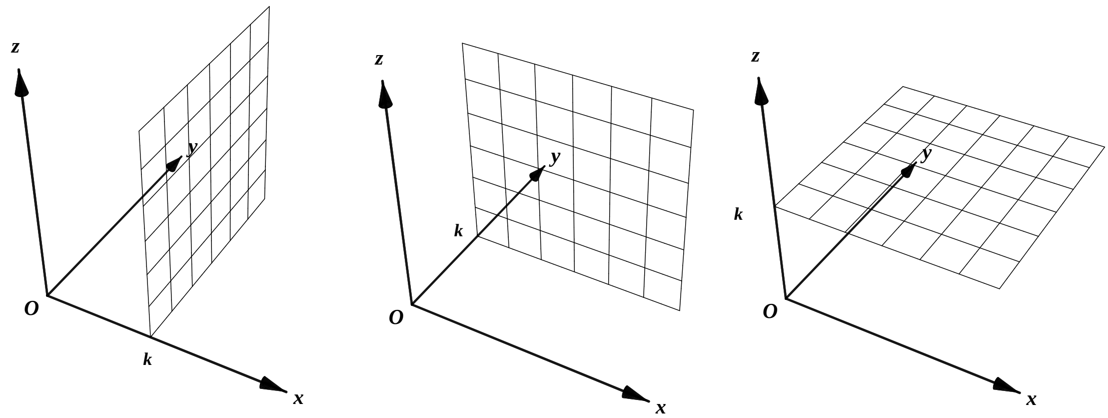


Figura 8.23 *Piani* $x = k$, $y = k$, $z = k$, *rispettivamente*

Passando ora a considerare piani non verticali, per ottenerne l'equazione possiamo considerare la generalizzazione dell'equazione di una retta non verticale; se teniamo conto che ora la variabile dipendente, cioè la quota, si indica abitualmente con z, otterremo una equazione del tipo

$$z = mx + ny + q\,, \tag{8.2}$$

dove q rappresenta la quota z all'origine. Per il significato di m ed n possiamo ragionare come segue (questo tipo di ragionamento ci sarà utile anche nel seguito). Se consideriamo un piano del tipo $z = mx + ny + q$ e lo intersechiamo con il piano $y = 0$ (cioé con il piano Oxz), otteniamo una retta del piano Oxz, di equazione $z = mx + q$. Dunque m rappresenta l'inclinazione di questa retta rispetto all'asse x. Analogamente si prova che n rappresenta l'inclinazione, rispetto all'asse y, della retta ottenuta intersecando $z = mx + ny + q$ con il piano $x = 0$.

Nelle due immagini della figura 8.24 è rappresentata questa situazione per il piano di equazione $z = x/2 + y + 1/2$.

8.5 Linee di livello e intersezioni con piani verticali

Definizione 8.3 (Linea di livello). *Data un funzione* $f(x,y)$ *una* linea di livello k, *che possiamo indicare con* l_k, *è l'insieme ottenuto come soluzione del sistema*

$$\begin{cases} z = f(x,y) \\ z = k \end{cases}, \tag{8.3}$$

ovvero è l'insieme (di solito una linea nel senso intuitivo del termine) intersezione tra la superficie grafico della funzione e il piano orizzontale a quota k. *Questa linea (essendo un'equazione in due variabili) va rappresentata sul piano* Oxy *(piano base), ma può anche essere tracciata direttamente sopra la superficie grafico della funzione.*

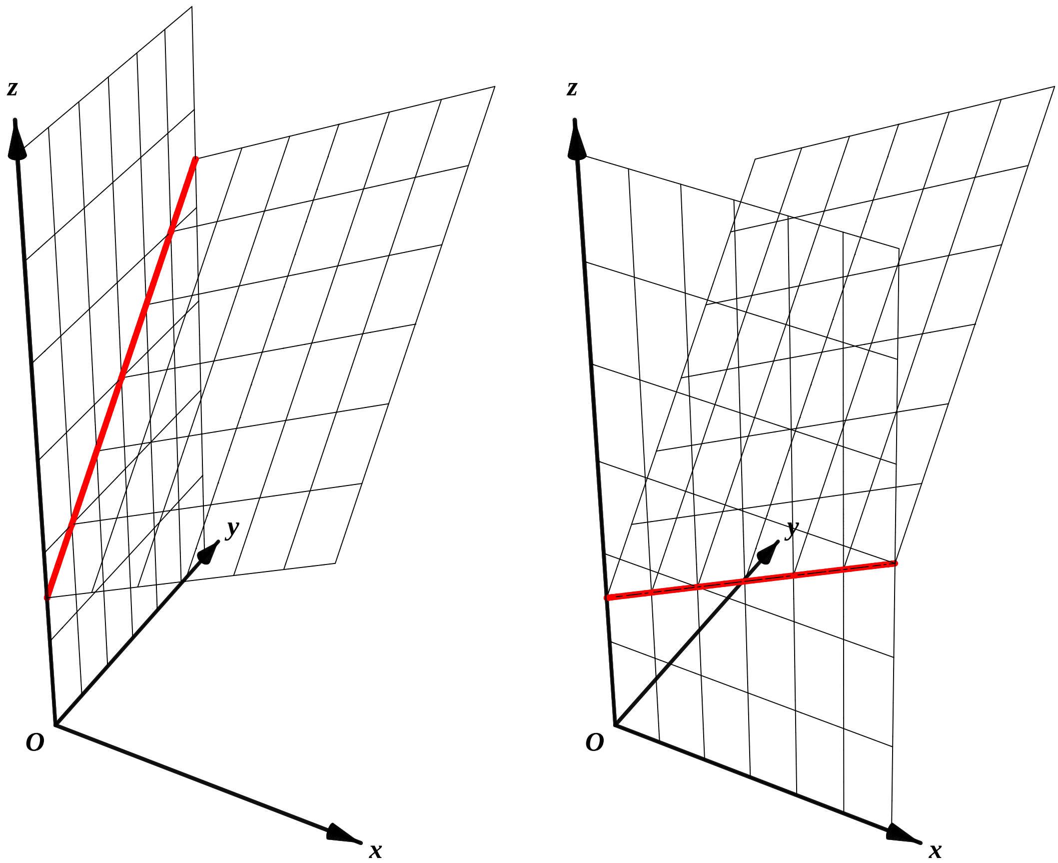


Figura 8.24 *Il piano di equazione* $z = x/2 + y + 1/2$ *e le due rette sezione con i piani coordinati verticali*

Esempio 8.1. Data $f(x,y) = x^2 - y^2$, la linea di livello k è $x^2 - y^2 = k$: se $k \neq 0$ si tratta di una iperbole, se $k = 0$, delle due rette $x = \pm y$. Tre di queste linee sono rappresentate nella figura 8.25.

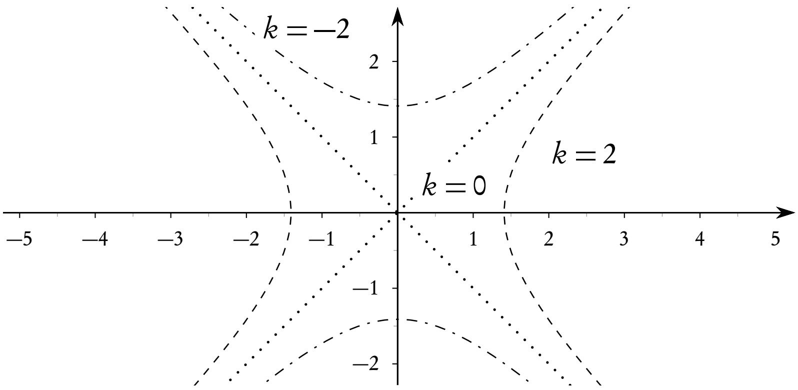


Figura 8.25 *Tre linee di livello per la funzione* $f(x,y)=x^2-y^2$

Esempio 8.2. Data $f(x,y)=x-y^2$, la linea di livello 1 è la parabola della figura 8.26

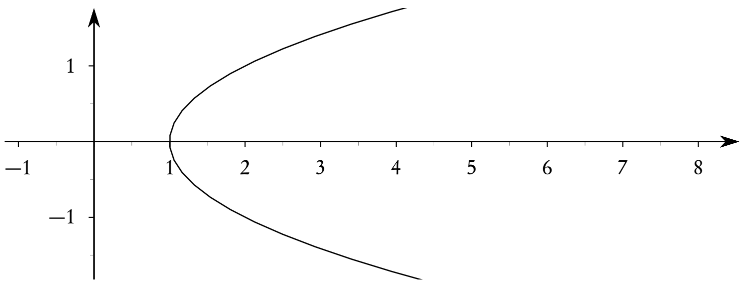


Figura 8.26 *Linea di livello* 1 *per la funzione* $f(x,y)=x-y^2$

Come vedremo, molto utili per studiare le proprietà delle funzioni di due variabili sono le linee intersezione della superfice-grafico della funzione con piani verticali paralleli ai piani coordinati, cioè del tipo $x=k$ e $y=k$. Queste linee si ottengono risolvendo uno dei seguenti due sistemi:

$$\begin{cases} z=f(x,y) \\ y=k \end{cases}, \Rightarrow z=f(x,k) \qquad \begin{cases} z=f(x,y) \\ x=k \end{cases}, \Rightarrow z=f(k,y). \tag{8.4}$$

Come è evidente nel primo caso si ottiene una funzione della variabile indipendente x, il cui grafico si potrà rappresentare in un piano Oxz, nel secondo caso si ottiene una funzione della variabile indipendente y, il cui grafico si potrà rappresentare in un piano Oyz. Naturalmente si potrà sempre immaginare queste curve anche tracciate direttamente sul grafico della funzione.

Esempio 8.3. Sia data la funzione $f(x,y)=x^3-4xy^2$, il cui grafico è rappresentato nella figura 8.27 (anche se in questo contesto il grafico è poco interessante).

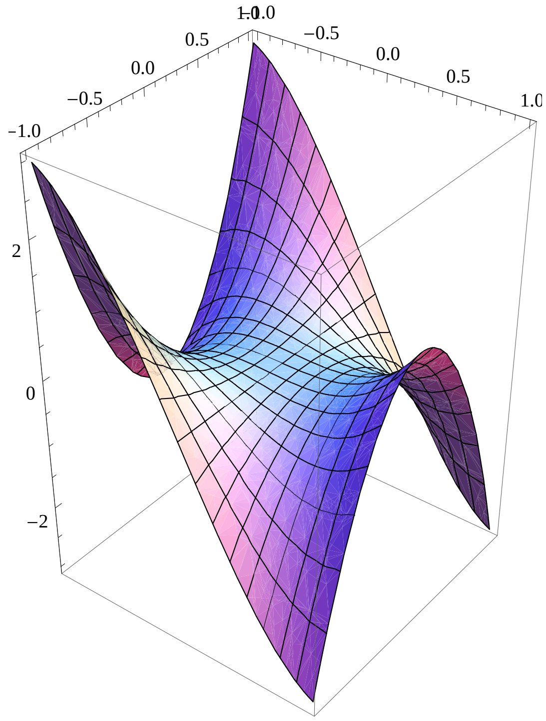


Figura 8.27 *Grafico della funzione* $f(x, y) = x^3 - 4xy^2$

L'intersezione con il piano $x = 1/2$ conduce alla funzione (della sola variabile y!) $z = 1/8 - 2y^2$, il cui grafico è (come è ben noto) una parabola (nel piano Oyz). L'intersezione con il piano $y = 1/2$ conduce alla funzione (della sola variabile x!) $z = x^3 - x$, il cui grafico (nel piano Oxz) possiamo tracciare con le note regole per studiare le funzioni di una variabile. Questi due grafici sono riportati nella figura 8.28.

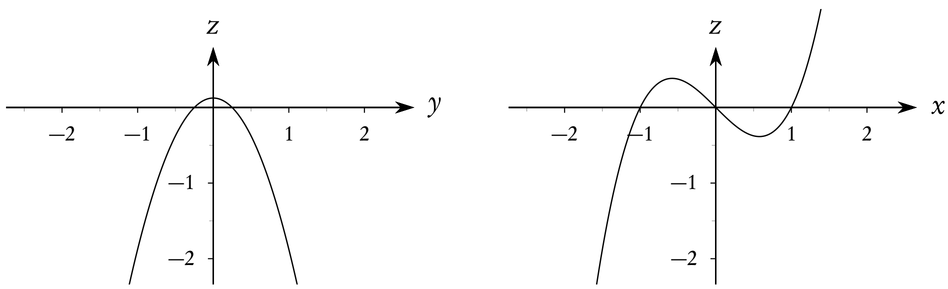


Figura 8.28 *Intersezioni della superficie* $z = x^3 - 4xy^2$ *con i piani* $x = 1/2$ *e* $y = 1/2$

La figura 8.29 mostra i piani sezionanti e le due curve direttamente sulla superficie.

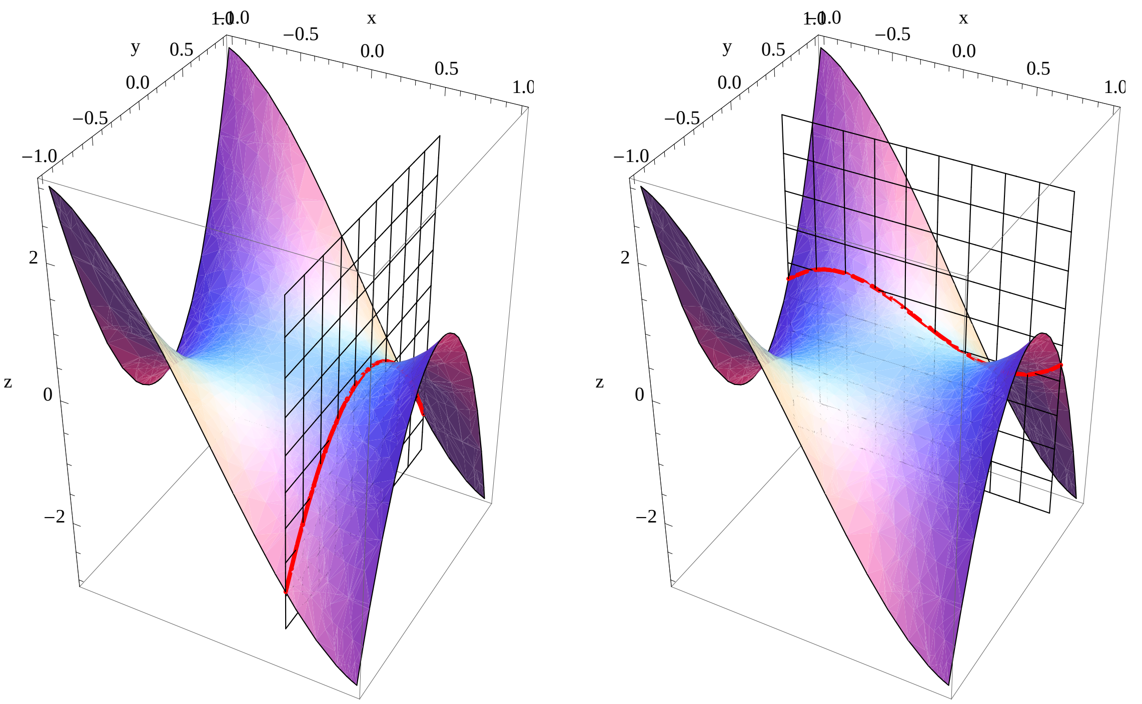


Figura 8.29 *Le intersezioni della figura 8.28, tracciate sulla superficie*

Le due funzioni ottenute per intersezione sono, come già notato, funzioni di una sola variabile e possono essere derivate, una o più volte, per valutare quando sono crescenti, decrescenti, concave, convesse, per trovare le rette tangenti, le eventuali formule di Taylor, ecc. Come vedremo queste derivate hanno interesse non solo per le curve intersezione, ma anche per la funzione di due variabili nel suo complesso.

8.6 Derivate parziali

Definizione 8.4 (Derivate parziali). *Data una funzione $z = f(x,y)$ e un punto (x_0,y_0) interno al suo dominio, possiamo considerare la funzione, della variabile x, $z = f(x,y_0) = g(x)$, ottenuta fissando y al valore y_0 e lasciando variare x, ovvero la funzione che si ottiene intersecando la superficie $z = f(x,y)$ con il piano verticale $y = y_0$. Possiamo ora considerare il*

$$\lim_{x\to x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0}, \tag{8.5}$$

ovvero il limite del rapporto incrementale della funzione $z = g(x)$. Se questo esiste ed è finito, esso si chiama derivata parziale prima rispetto a x della funzione f, nel punto (x_0,y_0) e si indica con

$$f'_x(x_0,y_0). \tag{8.6}$$

In maniera perfettamente analoga, possiamo considerare la funzione, della variabile y, $z = f(x_0,y) = h(y)$, ottenuta fissando x al valore x_0 e lasciando variare y, ovvero la funzione che si ottiene intersecando *la superficie* $z = f(x,y)$ *con il piano verticale* $x = x_0$. *Possiamo ora considerare il*

(8.7)
$$\lim_{y\to y_0} \frac{f(x_0,y)-f(x_0,y_0)}{y-y_0},$$

ovvero il limite del rapporto incrementale della funzione $z = h(y)$. *Se questo esiste ed è finito, esso si chiama derivata parziale prima rispetto a* y *della funzione* f, *nel punto* (x_0,y_0) *e si indica con*

(8.8)
$$f'_y(x_0,y_0).$$

In pratica il calcolo delle due derivate parziali in un punto generico (x,y) interno al dominio si fa pensando la funzione $f(x,y)$ come funzione di una sola delle due variabili e trattando l'altra come un parametro costante.

Esempio 8.4. Da $f(x,y) = x^2 + 4xy + 3xy^2$, si ottiene $f'_x(x,y) = 2x + 4y + 3y^2$, $f'_y(x,y) = 4x + 6xy$.

Esempio 8.5. Da $f(x,y) = \sin(x + x^2y)$, si ottiene $f'_x(x,y) = (1 + 2xy)\cos(x + x^2y)$, $f'_y(x,y) = x^2\cos(x + x^2y)$.

Esempio 8.6. Da $f(x,y) = e^{x^2+y^2}$, si ottiene $f'_x(x,y) = 2xe^{x^2+y^2}$, $f'_y(x,y) = 2ye^{x^2+y^2}$.

Come mostrano gli esempi proposti, le derivate parziali, calcolate in un generico punto, sono esse stesse funzioni di due variabili, e quindi posso riapplicare ad esse ancora la derivazione, ottenendo le derivate seconde; precisamente avendo ottenuto da *una* funzione *due* derivate parziali prime, da ciascuna otterrò *due* derivate parziali, per un totale di *quattro* derivate parziali seconde della funzione originaria:

- f''_{xx} sarà la derivata prima rispetto a x della f'_x;
- f''_{yy} sarà la derivata prima rispetto a y della f'_y;
- f''_{xy} sarà la derivata prima rispetto a y della f'_x;
- f''_{yx} sarà la derivata prima rispetto a x della f'_y.

Le prime due si chiamano *derivate parziali seconde pure*[2], le ultime due si chiamano *derivate parziali seconde miste*.

Esempio 8.7. Da $f(x,y) = x^2+4xy+3xy^2$, si ottiene, come già visto, $f'_x(x,y) = 2x+4y+3y^2$, $f'_y(x,y) = 4x + 6xy$ e, successivamente, $f''_{xx}(x,y) = 2$, $f''_{yy}(x,y) = 6x$, $f''_{xy}(x,y) = 4 + 6y$, $f''_{yx}(x,y) = 4 + 6y$.

Si potrebbe naturalmente proseguire ottenendo le derivate terze, e così via, ma non saremo interessati al loro uso. Osserviamo invece che, nell'esempio precedente, $f''_{xy}(x,y) = 4 + 6y = f''_{yx}(x,y)$. La cosa, anche se a prima vista sorprendente, non è casuale. Vale infatti il seguente notevole teorema.

Teorema 8.5 (Teorema di Schwartz). *Se le derivate seconde miste sono continue, allora esse sono uguali.*

Nei casi che ci interessano le cose andranno sempre nel senso previsto da questo teorema, ovvero le derivate seconde miste saranno sempre uguali.

Come abbiamo già avuto modo di constatare, la derivata prima per funzioni di una variabile permette il calcolo della pendenza della retta tangente al grafico della funzione e quindi la determinazione

[2]Spesso l'appellativo "pure" si tralascia

dell'equazione di questa tangente. Per le funzioni di due variabili le derivate parziali, in base a quanto abbiamo detto, serviranno a determinare le equazioni delle rette tangenti alle curve intersezione tra la superficie e il piano verticale parallelo al piano Oxz oppure Oyz. Esse però servono anche a determinare (almeno per funzioni abbastanza regolari) l'equazione del *piano tangente* alla superficie grafico della funzione di due variabili. Precisamente, data una funzione di due variabili $z = f(x,y)$ e un punto (x_0, y_0) del suo dominio, dove la funzione ammette derivate parziali prime *continue* (come succederà sempre nei nostri casi), l'equazione del piano tangente alla superficie grafico della funzione nel punto (x_0, y_0, z_0), con $z_0 = f(x_0, y_0)$ sarà:

$$z = f(x_0, y_0) + f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0). \tag{8.9}$$

Esempio 8.8. Riprendendo la funzione $f(x,y) = x^2 + 4xy + 3xy^2$ già trattata prima e considerato il punto $(1, -1)$, si ha $f(1,-1) = 0$, $f'_x(1,-1) = 1$, $f'_y(1,-1) = -2$, dunque l'equazione del piano tangente è

$$z = 0 + 1(x - 1) - 2(y + 1) \Rightarrow z = x - 2y - 3. \tag{8.10}$$

Esempio 8.9. Procedendo come nell'esempio precedente è facile provare che l'equazione del piano tangente al grafico di $z = -x^2 - y^2$, in corrispondenza al punto $(1, -1)$ è: $z = -2x + 2y + 2$. La situazione è rappresentata nella figura 8.30, dove sono rappresentate anche le due curve sezione.

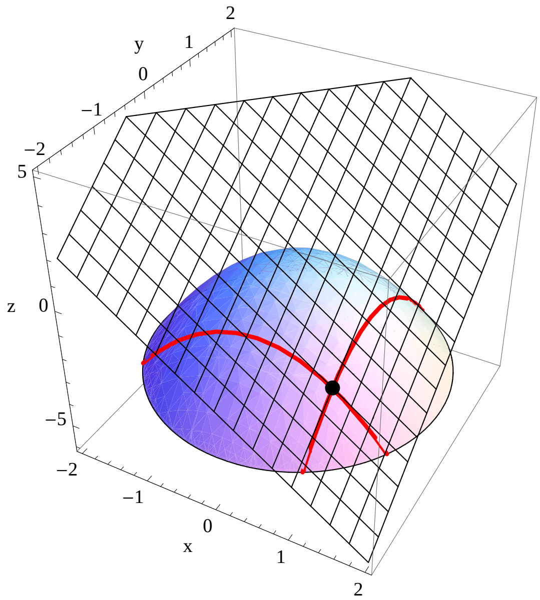


Figura 8.30 *Superficie* $z = -x^2 - y^2$ *e piano tangente in* $(1, -1)$

8.7 Ottimizzazione libera

I problemi principali a cui saremo interessati relativamente alle funzioni di due variabili sono i problemi di *ottimizzazione*, libera e vincolata, ovvero il problema della ricerca dei massimi e minimi nei punti interni al dominio della funzione, nei punti del bordo del dominio o, eventualmente, su un sottoinsieme del dominio. Nei casi che ci interessano questi problemi sono risolubili con lo studio delle derivate prime e seconde della funzione e, eventualmente di una nuova funzione (la funzione *Lagrangiana*) costruita a partire dalla funzione stessa in modo da tenere conto di eventuali ulteriori condizioni (vincoli).

Abbiamo già proposto grafici relativi a funzioni di due variabili in cui erano evidenziati massimi ("cime di monti") e minimi ("fondovalle"): per esempio le figure 1.20 (nella pagina 21), 8.8 (nella pagina 156) e 8.13 (nella pagina 158). Particolarmente significativa l'illustrazione 8.13 (nella pagina 158), in cui si evidenzia che il piano tangente alla superficie nei punti di massimo o di minimo (interni al dominio) è *orizzontale*, ovvero del tipo $z = k$. Si tratta di una situazione identica al caso delle funzioni di una variabile, dove, nei punti di massimo e minimo (interni al dominio) era la retta tangente ad essere orizzontale.

Se si tiene conto dell'equazione del piano tangente che abbiamo scritto nell'equazione (8.10), possiamo concludere che, in corrispondenza a un punto di massimo o minimo interno al dominio *entrambe* le derivate parziali saranno nulle, in perfetta analogia con il caso di una variabile dove si aveva l'annullamento della derivata prima.

Purtroppo (ancora come nel caso di funzioni di una variabile) l'annullarsi delle derivate *non garantisce* l'esistenza di un massimo o un minimo. Basta pensare ai punti di sella o alle selle di scimmia (vedi le figure 8.15 e 8.16).

Riproponiamo qui di seguito, per comodità, le stesse due figure con l'aggiunta del piano tangente.

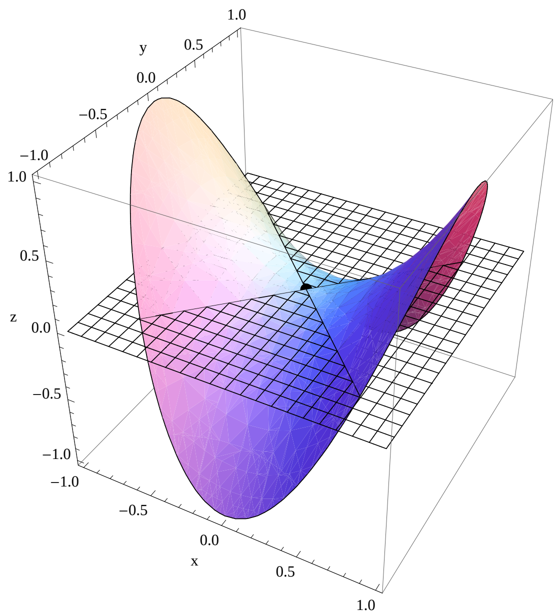


Figura 8.31 *Una "sella di cavallo" e il piano tangente*

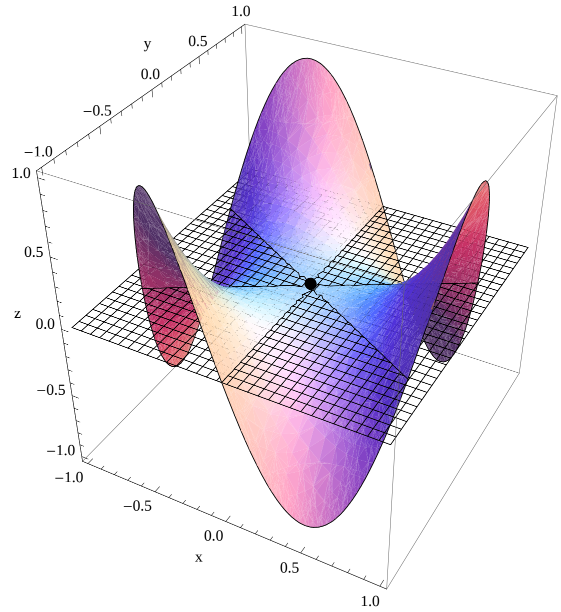


Figura 8.32 *Una "sella di scimmia" e il piano tangente*

La figura 8.32 mostra che la superficie ha un andamento "sfarfallante" rispetto al piano tangente nel punto dove esso risulta orizzontale. La situazione può essere anche più complessa, in quanto lo "sfarfallio" può essere ancora più accentuato, come mostra la figura 8.33.

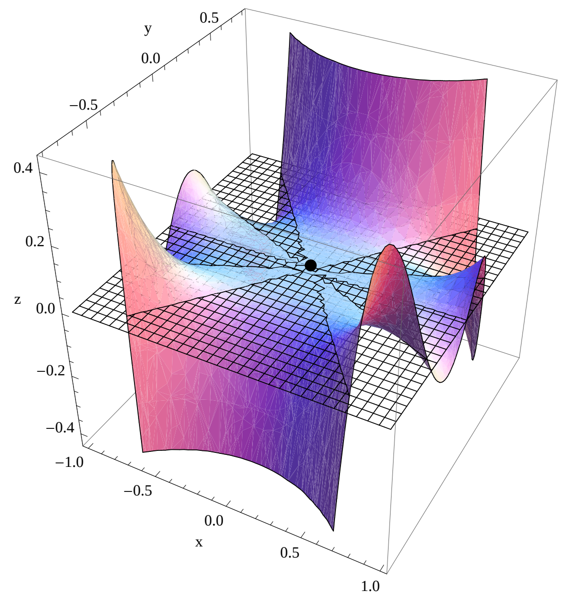


Figura 8.33 *Superficie con pronunciato "sfarfallio" rispetto al piano tangente orizzontale*

Quanto abbiamo detto si può riassumere nel seguente teorema.

Teorema 8.6 (Condizione necessaria per i massimi e minimi in due variabili)**.** *Se una funzione $f(x,y)$ dotata di derivate parziali ha, in corrispondenza a un punto (x_0,y_0)* interno *al dominio, un massimo o un minimo, allora necessariamente le derivate sono* contemporaneamente *nulle in (x_0,y_0).*

Un punto (interno al dominio) in cui le derivate parziali siano contemporaneamente nulle (senza che necessariamente sia un punto di minimo o di massimo) si chiama un *punto stazionario*, a volte anche *punto critico* per $f(x,y)$. Il teorema precedente si può allora riformulare dicendo che: condizione necessaria perché un punto (x_0,y_0) interno al dominio sia di massimo o di minimo per una funzione derivabile, è che esso sia un punto stazionario. La condizione *non* è in genere sufficiente.

Nel caso di una variabile per valutare se un punto (in cui la derivata prima si annulla) è di massimo di minimo (o di flesso), si può procedere a studiare la crescenza e decrescenza tramite il segno della derivata prima. Nulla di simile per le funzioni di due variabili, dove i concetti di funzione crescente e decrescente *non hanno alcun senso*. Per risolvere il problema ci viene in aiuto il teorema che segue, che dà una condizione *sufficiente* perché un punto stazionario sia di massimo o di minimo.

Teorema 8.7. *Sia data una funzione $f(x,y)$ dotata almeno di derivate seconde. Se (x_0,y_0) è un punto stazionario per f (interno al dominio), si calcolano, in (x_0,y_0), le quattro*[3] *derivate seconde e si costruisce la seguente tabella (matrice), detta* matrice hessiana,

$$\begin{pmatrix} f''_{xx}(x_0,y_0) & f''_{xy}(x_0,y_0) \\ f''_{yx}(x_0,y_0) & f''_{yy}(x_0,y_0) \end{pmatrix}. \tag{8.11}$$

Successivamente si calcola il seguente numero, detto determinante hessiano *o semplicemente* hessiano, *e indicato $H_f(x_0,y_0)$, o semplicemente con $H(x_0,y_0)$,*

$$f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)-f''_{xy}(x_0,y_0)f''_{yx}(x_0,y_0)=f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)-\left(f''_{xy}(x_0,y_0)\right)^2, \tag{8.12}$$

ottenuto facendo la differenza dei "prodotti in croce" degli elementi della precedente matrice. Ebbene:

- *Se $H(x_0,y_0)<0$, allora il punto (x_0,y_0) è un* punto di sella.
- *Se $H(x_0,y_0)>0$, allora si guarda uno dei due termini sulla diagonale principale della matrice (cioè $f''_{xx}(x_0,y_0)$ o $f''_{yy}(x_0,y_0)$):*
  - *se esso è >0 il punto è di* minimo *(relativo);*
  - *se esso è <0 il punto è di* massimo *(relativo).*
- *Se $H(x_0,y_0)=0$, allora* nulla si può concludere*: può succedere di tutto*[4].

Si tenga ben presente che se $H(x_0,y_0)>0$, allora $f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)-\left(f''_{xy}(x_0,y_0)\right)^2>0$, da cui $f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)>\left(f''_{xy}(x_0,y_0)\right)^2\geq 0$, per cui le due derivate seconde pure *devono* avere lo stesso segno e non possono annullarsi: è per questo che è indifferente considerare una o l'altra.

[3] In realtà ne bastano tre perché, nei casi che ci interessano, le due miste sono uguali.
[4] E occorrerebbe un'indagine approfondita che di solito esula dagli scopi di questo corso.

Esempio 8.10. Trovare i punti stazionari della funzione $f(x,y) = 2\ln(x^2+y^2+2) - xy$ e classificarli, usando la matrice hessiana.

Il primo passo consiste nel calcolare le derivate parziali prime e nel cercare i punti dove esse si annullano contemporaneamente (tecnicamente è questa la parte difficile perché si tratta di risolvere un sistema di due equazioni in due incognite, in genere non banale). Si ottiene:

$$\begin{cases} f'_x = \dfrac{4x}{x^2+y^2+2} - y = 0 \\ f'_y = \dfrac{4x}{x^2+y^2+2} - x = 0 \end{cases} .$$

Questo sistema è abbastanza bruttino, ma con un po' di pazienza si riesce a trovare che le sue soluzioni sono $(0,0)$, $(1,1)$ e $(-1,-1)$. Si hanno tre punti critici.

Si calcolano ora le derivate seconde e si scrivono le tre matrici hessiane, ottenendo, rispettivamente,

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1/2 & -3/2 \\ -3/2 & 1/2 \end{pmatrix}, \quad \begin{pmatrix} 1/2 & -3/2 \\ -3/2 & 1/2 \end{pmatrix}.$$

Nel primo punto si ha $H = 3 > 0$ e i termini sulla diagonale maggiore sono positivi: si tratta di un minimo relativo. Negli altri due punti si ha $H = -2 < 0$, quindi sono due punti di sella(5).

8.8 Ottimizzazione vincolata

Supponiamo ora di dover trovare i massimi e minimi non sui punti interni all'intero dominio, ma

1. sui punti interni a una parte del dominio;
2. oppure sul bordo del dominio;
3. oppure sulla restrizione della funzione a una curva tracciata all'interno del dominio(6).

Il primo dei casi elencati è semplice: si trovano i massimi e minimi sull'interno di *tutto* il dominio e poi si controlla se questi punti sono anche interni alla parte considerata. Il secondo dei casi è praticamente identico al terzo: nei casi che ci interessano il bordo del dominio è una curva. Pertanto il problema che rimane aperto è: come fare a trovare i massimi e minimi di una funzione di due variabili, se la consideriamo ristretta a una curva? È più propriamente in questo caso che parleremo di *massimi o minimi vincolati*. Per capire la difficoltà del problema utilizziamo, come al solito, un esempio grafico.

Riesaminiamo la figura 8.8 della pagina 156. Come già a suo tempo osservato, questa immagine evidenzia, per la funzione, la presenza di due massimi (relativi) e di un minimo (relativo), riferiti però all'intera superficie.

Supponiamo ora di voler considerare la restrizione della funzione all'insieme del piano Oxy individuato dall'equazione $(x-1)^2 + y^2 = 1$, cioè alla circonferenza di centro $(1,0)$ e raggio 1. In sostanza si tratta di questo: invece di considerare tutti i punti della superficie, consideriamo solo quelli relativi ai punti di questa circonferenza. Detto in altri termini: consideriamo la superficie cilindrica verticale ottenuta a partire da questa circonferenza e consideriamone l'intersezione con la superficie; si otterrà

[5] Attenzione: è casuale che i termini sulla diagonale principale siano uguali, mentre è naturale che lo siano quelli sulla diagonale secondaria (Teorema di Schwartz); è altresì casuale che la seconda e terza matrice hessiana siano uguali.

[6] Ci sono anche alcune altre situazioni, ma ci limiteremo a trattare queste tre

una curva dello spazio (almeno nei casi che ci interessano). Ebbene siamo interessati ai massimi e minimi che si ottengono se ci muoviamo *solo* su questa curva: è come dire che dobbiamo trovare i punti più alti e più bassi su una strada tracciata sulla superficie, la quale superficie comprende diverse montagne e vallate.

La figura 8.34 visualizza questa situazione. La figura 8.35, in cui è tracciata solo la strada, evidenzia ancora meglio il problema e ne palesa tutte le difficoltà: è chiaro che non potremo parlare di derivate parziali, di piano tangente, o cose simili. Un massimo e/o minimo su questa curva potrà non avere nulla a che fare con i massimi e/o minimi sull'intera superficie.

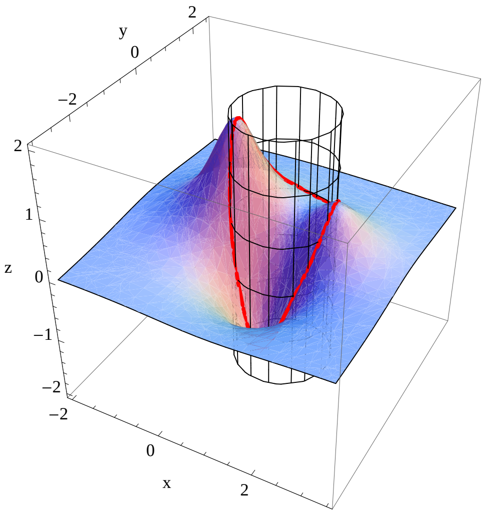


Figura 8.34 *Un problema di ottimizzazione vincolata*

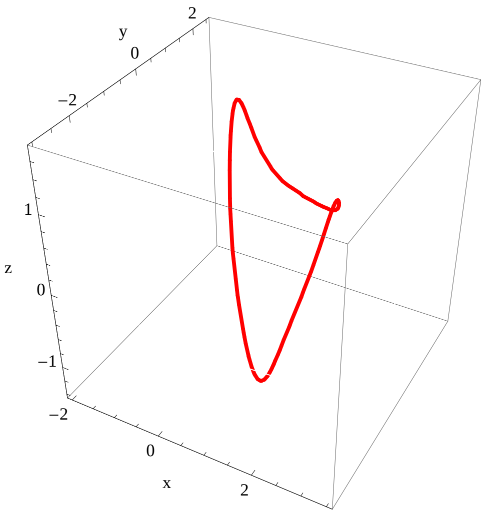


Figura 8.35 *Particolare di un problema di ottimizzazione vincolata*

Ebbene, è interessante il fatto che, nonostante le evidenti difficoltà, la determinazione di questi massimi e minimi può essere effettuata in una maniera molto simile a quanto fatto per i massimi e minimi liberi, seppure, naturalmente, con un opportuno adattamento.

Possiamo precisare il problema nel seguente modo (limitandoci, come al solito, ai casi di nostro interesse).

Sia data una funzione $f(x,y)$ definita in un dominio D, e si consideri una curva $\mathcal{C}$ tracciata nel dominio (potrebbe essere semplicemente il bordo del dominio). Questa curva, che sarà detta *vincolo* avrà generalmente una equazione del tipo $g(x,y)=0$[7]. Allora:

1. se da $g(x,y)=0$ si può esplicitare o la x o la y, la si sostituisce nella funzione f che diventa di una variabile, e si procede appunto come per le funzioni di una variabile;
2. se questo non è possibile (o è troppo complesso), si usa il metodo dei moltiplicatori di Lagrange, di cui parleremo tra poco.

Esempio 8.11. Se $z=f(x,y)=x^2-2y^2$ e il vincolo è $x-y+1=0$, dal vincolo si ricava $y=x+1$, che si sostituisce nella $f(x,y)$, ottenendo la funzione $z=-x^2-4x-2$, da cui si vede subito che si ha un massimo per $x=-2$, massimo che vale 2; naturalmente la y del massimo si ricaverà dal vincolo, ottenendo $y=-1$. La situazione è illustrata nella figura 8.36.

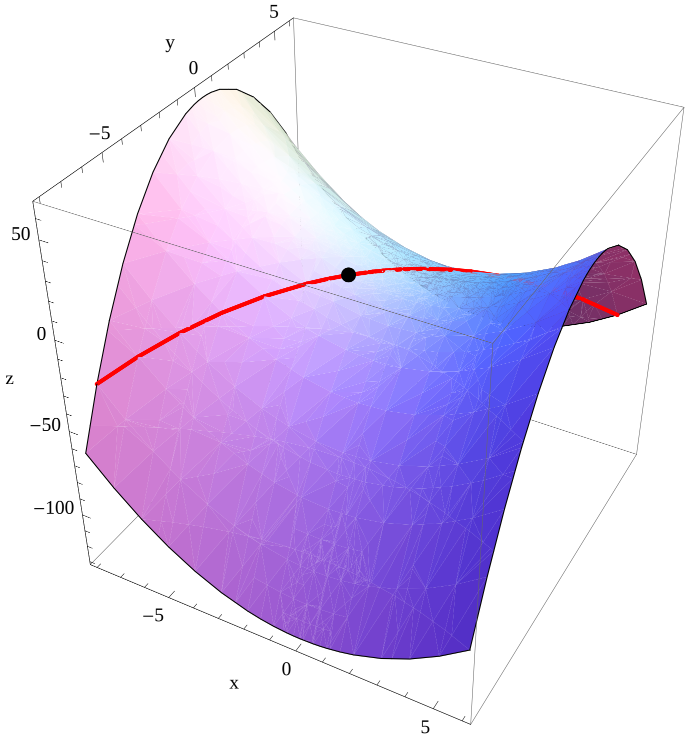


Figura 8.36 *Un massimo vincolato, con vincolo esplicitabile*

[7] Attenzione: ridurre sempre il vincolo alla forma indicata, ovvero $g(x,y)=0$.

E veniamo ora al metodo dei *moltiplicatori di Lagrange*, da applicare quando non si può esplicitare alcuna variabile[8].

Si procede nel seguente modo:

1. si costruisce la funzione lagrangiana (funzione di tre variabili)[9]: $L(x,y,\lambda)=f(x,y)+\lambda g(x,y)$;
2. si calcolano le derivate parziali rispetto alle tre variabili: L'_x, L'_y, L'_λ;
3. si risolve il sistema
$$\begin{cases} L'_x=0 \\ L'_y=0 \\ L'_\lambda=0 \end{cases};$$
4. se ci sono punti di massimo o minimo vincolato questi si trovano tra le coppie (x,y) estratte dalle terne (x,y,λ) che risolvono il sistema. Si noti come la situazione sia simile a quella dei massimi e minimi liberi: le coppie (x,y) estratte dalle terne (x,y,λ) che risolvono il sistema *non* è detto che siano punti di massimo o minimo vincolato, però gli eventuali punti di massimo o minimo vincolato vanno ricercati *solo* fra queste coppie (che in sostanza sono i punti stazionari della funzione lagrangiana. Si dice che i punti stazionari della lagrangiana sono gli unici "candidati" ad essere di massimo o minimo.

Nelle situazioni pratiche succede sempre che il vincolo è un insieme chiuso e limitato e la funzione è sufficientemente regolare, per cui (in base al teorema di Weierstrass, valido anche in due variabili) il massimo e minimo assoluto esistono sicuramente. Basterà allora trovare tutti i possibili candidati (che nei casi di nostro interesse sono in numero finito) usando il metodo di Lagrange e calcolare poi la funzione in tutti questi punti: il valore più alto corrisponderà al massimo, il più basso al minimo.

Esempio 8.12. Trovare il massimo e minimo assoluto di $f(x,y)=x+y$ sotto la condizione $x^2+y^2=1$.

Posto $L(x,y,\lambda)=x+y+\lambda(x^2+y^2-1)$, si ha

$$\begin{cases} L'_x(x,y,\lambda)=1+2\lambda x \\ L'_y(x,y,\lambda)=1+2\lambda y \\ L'_\lambda(x,y,\lambda)=x^2+y^2-1 \end{cases}$$

Uguagliando a zero le tre derivate parziali e risolvendo il sistema ottenuto, si trova:

$$\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right) \quad \text{e} \quad \left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)..$$

Estraendo[10] da queste terne le coppie dei valori (x,y) si trovano i seguenti due punti:

$$\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right) \quad \text{e} \quad \left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)..$$

[8]In realtà anche quando non si può esplicitare alcuna variabile si potrebbe evitare, almeno in certi casi, il metodo dei moltiplicatori, ma questo richiede lo studio della teoria delle curve, cosa che esula dal nostro programma.

[9]Alcuni testi prendono come funzione lagrangiana la $L(x,y,\lambda)=f(x,y)-\lambda g(x,y)$: non cambia assolutamente nulla, in quanto basta sostituire λ con $-\lambda$ per passare dall'una all'altra.

[10]Si noti che, in realtà, il valore di λ non ha interesse in questa questione, bisogna però che un valore di λ esista.

Poiché

$$f\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)=\sqrt{2} \quad \text{e} \quad f\left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)=-\sqrt{2},$$

se ne conclude che $\sqrt{2}$ è il massimo assoluto, mentre $-\sqrt{2}$ è il minimo assoluto. La figura 8.37 illustra la situazione.

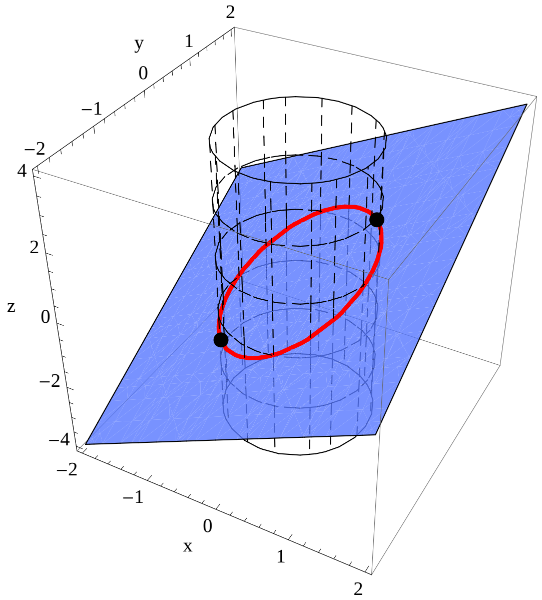


Figura 8.37 *Ottimizzazione vincolata, con il metodo dei moltiplicatori*

Esempio 8.13. Trovare il massimo e minimo assoluto di $f(x,y)=xy$ sotto la condizione $x^2+y^2=1$.

Posto $L(x,y,\lambda)=x+y+\lambda(x^2+y^2-1)$, calcolando le tre derivate parziali e uguagliandole a zero, si trova il seguente sistema:

$$\begin{cases} y-2\lambda x=0 \\ x-2\lambda y=0 \\ x^2+y^2-1=0 \end{cases},$$

che ha come soluzioni

$$A=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right),\ B=\left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right),\ C=\left(\frac{-\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right),\ D=\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right).$$

Poiché

$$f(A)=f(B)=\frac{1}{2} \quad \text{e} \quad f(C)=f(D)=-\frac{1}{2},$$

se ne deduce che il massimo vale $1/2$, mentre il minimo vale $-1/2$. La figura 8.38 illustra la situazione.

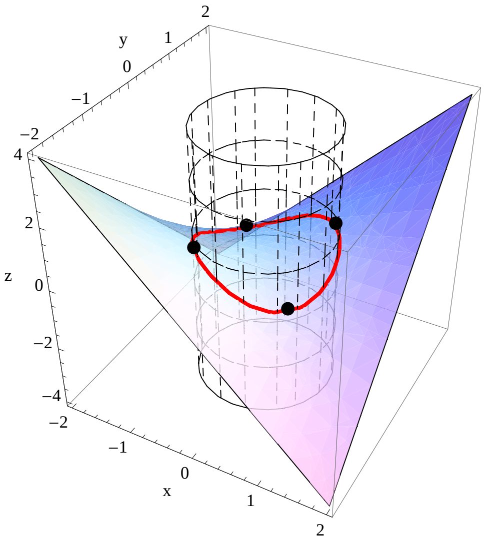


Figura 8.38 *Ancora una ottimizzazione vincolata, con il metodo dei moltiplicatori*

Esempio 8.14. Una ditta ha a disposizione 90€ per acquistare x oggetti di tipo A al prezzo di 3€ l'uno e y oggetti di tipo B al prezzo di 5€ l'uno. Per avere la massima utilità il prodotto xy[11] deve essere massimo. Quante macchine di ogni tipo deve acquistare?

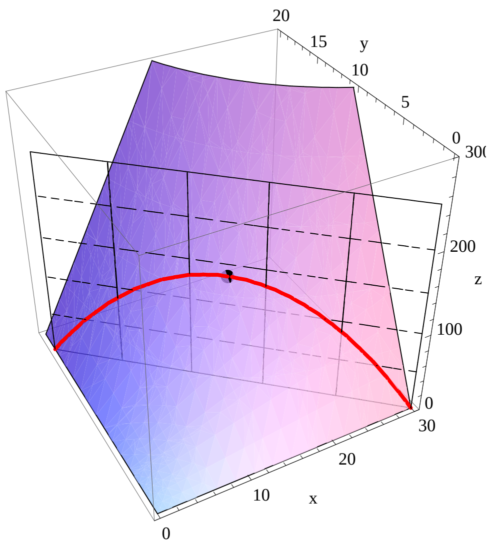


Figura 8.39 *Una ottimizzazione vincolata in un problema di economia*

Si deve massimizzare la funzione $f(x, y) = xy$ sottoposta al vincolo $3x + 5y = 90$. Procedendo come sopra si trova che l'unico punto stazionario della funzione Lagrangiana è $(15, 9, -3)$, cioè l'unico punto candidato ad essere di massimo (o minimo) vincolato è $(15, 9)$. In corrispondenza di questo punto la

[11]La funzione è detta proprio funzione di utilità.

funzione di utilità vale 135. Purtroppo il fatto che il vincolo sia una retta del piano, cioè un insieme non limitato, non ci permette di concludere agevolmente come negli esempi precedenti. Per il momento ci limitiamo ad usare il grafico della figura 8.39. Torneremo successivamente su questo punto.

Esiste una condizione sufficiente per i massimi e minimi vincolati e di questa faremo ora un breve cenno.

Sia data la funzione $z = f(x, y)$, opportunamente regolare, e il vincolo $g(x, y) = 0$, anch'esso opportunamente regolare. Consideriamo nuovamente la funzione lagrangiana $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ e supponiamo di averne determinato un punto stazionario (x_0, y_0, λ_0). In questo punto calcoliamo le derivate prime della funzione vincolo e le derivate seconde della funzione lagrangiana. Consideriamo poi la seguente matrice, detta *matrice hessiana orlata*:

$$\begin{pmatrix} 0 & g'_x & g'_y \\ g'_x & f''_{xx} + \lambda g''_{xx} & f''_{xy} + \lambda g''_{xy} \\ g'_y & f''_{yx} + \lambda g''_{yx} & f''_{yy} + \lambda g''_{yy} \end{pmatrix}. \tag{8.13}$$

Il determinante[12] di questa matrice è detto *hessiano orlato*. Se l'hessiano orlato è maggiore di zero, il punto (x_0, y_0) è di massimo per la funzione sul vincolo, se l'hessiano orlato è minore di zero il punto (x_0, y_0) è di minimo.

Nell'esempio precedente di massimo e minimo di un problema economico, l'hessiano orlato è

$$\begin{vmatrix} 0 & 3 & 5 \\ 3 & 0 & 1 \\ 5 & 1 & 0 \end{vmatrix} = 30 > 0,$$

da cui si conferma che il punto è di massimo, come già dedotto per via grafica.

8.9 Ottimizzazione globale su insiemi chiusi e limitati

L'ultimo problema di ottimizzazione di cui ci occupiamo è quello della ricerca del massimo e minimo assoluto di una funzione di due variabili su un insieme bidimensionale chiuso e limitato del piano, per esempio un cerchio, un quadrato, un triangolo, ecc. Se la funzione in esame è continua tale massimo e minimo assoluto esistono sicuramente per il teorema di Weierstrass. In generale si procede secondo quanto di seguito indicato.

1. si cercano gli eventuali punti stazionari interni al dominio annullando le derivate parziali prime;

[12] Il determinante di una matrice a 3 righe e 3 colonne come questa si fa con la regola di Sarrus, evidenziata nei passaggi seguenti.

$$\text{Dato } \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}, \text{ costruisci } \left|\begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix}\right|\begin{matrix} a & b \\ d & e \\ g & h \end{matrix}.$$

Successivamente calcola la somma dei prodotti sulle tre diagonali principali (dall'alto a sinistra al basso a destra) e delle tre diagonali secondarie (dall'alto a destra al basso a sinistra). Il determinante è la differenza di questi due numeri, ovvero

$$(aei + bfg + cdh) - (bdi + afh + crg).$$

2. si cercano (con uno dei metodi indicati) i punti stazionari sul bordo che, nei casi di nostro interesse, è costituito da una curva o dall'unione di più curve;
3. si calcola il valore della funzione su ciascuno dei punti trovati all'interno o sul bordo: il più grande valore corrisponderà al massimo, il più piccolo al minimo.

Chiariamo il tutto con un esempio.

Esempio 8.15. Trovare il massimo assoluto di

$$f(x,y)=x^2+y^2-xy+3x-7y$$

nell'insieme $T=\{(x,y\in\mathbb{R}^2 \mid 0\leq y\leq 4\wedge 0\leq x\leq y\}$, rappresentato nella figura 8.40.

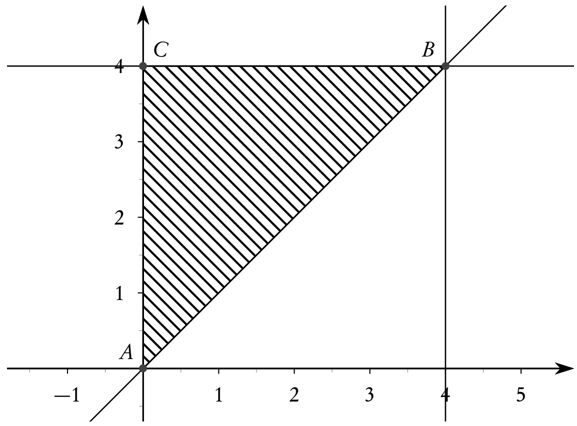


Figura 8.40 *Il triangolo* $T=\{(x,y\in\mathbb{R}^2 \mid 0\leq y\leq 4\wedge 0\leq x\leq y\}$

Cominciamo con il cercare eventuali punti stazionari per f all'interno di T, annullando le due derivate parziali.

$$\begin{cases} f'_x(x,y)=2x-y+3=0 \\ f'_y(x,y)=2y-x-7=0 \end{cases}.$$

Si trova come unica soluzione, interna al triangolo,

$$P\left(\frac{1}{3},\frac{11}{3}\right).$$

Vediamo allora cosa succede sul bordo di T, distinguendo 3 casi.

1. Sul segmento AB si ha $y=x$, con $0\leq x\leq 4$: la funzione f si riduce a $g_1(x)=x^2-4x$ che ha un minimo per $x=2$ e quindi due massimi agli estremi.
2. Sul segmento BC si ha $y=4$ e $0\leq x\leq 4$: la funzione si riduce a $g_2(x)=x^2-x-12$, che ha un minimo per $x=1/2$ e quindi due massimi agli estremi.
3. Sul segmento AC si ha $x=0$ e $0\leq y\leq 4$: la funzione si riduce a $h(y)=y^2-7y$, che ha un minimo per $y=7/2$ e quindi due massimi agli estremi.

Si devono dunque confrontare i valori di f nel punto P e nei tre punti A, B, C. Si ha

$$f\left(\frac{1}{3},\frac{11}{3}\right)=-\frac{111}{9},\ f(0,0)=0,\ f(0,4)=-12,\ f(4,4)=0.$$

Il massimo assoluto è dunque 0 e viene assunto nei punti O e B.

8.10 Esercizi

Esercizio 8.1. *Calcolare le derivate parziali prime e seconde delle seguenti funzioni.*

1. $f(x,y) = x^2y^2$.
2. $f(x,y) = xy - xy^2$.
3. $f(x,y) = e^x y$.
4. $f(x,y) = e^{xy} xy$.
5. $f(x,y) = y \ln x$.
6. $f(x,y) = \ln(xy$.
7. $f(x,y) = \dfrac{\ln x}{y}$.
8. $f(x,y) = e^{x+xy^2}$.

Esercizio 8.2. *Per le funzioni di seguito elencate dire se i punti indicati sono di massimo, minimo o sella (liberi); se possibile determinare se esistono altri punti di massimo, minimo, sella.*

1. $f(x,y) = x^2y$, $P(0,0)$, $Q(0,1)$.
2. $f(x,y) = xy - x^2y^2$, $P(0,0)$, $Q(1,1)$, $R(1,-1)$.
3. $f(x,y) = x \ln y$, $P(1,-1)$, $Q(0,1)$.
4. $f(x,y) = x^2e^y$, $P(0,0)$.
5. $f(x,y) = xye^{x+y}$, $P(1,0)$, $Q(-1,1)$, $R(1,-1)$.
6. $f(x,y) = x^2e^{3y-x}$, $P(1,0)$, $Q(0,0)$.
7. $f(x,y) = e^{xy}$, $P(0,0)$.
8. $f(x,y) = \ln(xy+1)$, $P(0,0)$, $Q(2,0)$.
9. $f(x,y) = \ln(x^2+y^2+2)$, $P(0,0)$.
10. $f(x,y) = e^{xy-x}$, $P(0,1)$.
11. $f(x,y) = x^2 - e^{y^2}$, $P(0,0)$.
12. $f(x,y) = 2x^2 - y$, $P(1,1)$.
13. $f(x,y) = x^2y^3$, $P(-1,1)$, $Q(0,0)$.
14. $f(x,y) = \ln x - y^2$, $P(1,-1)$.
15. $f(x,y) = e^{xy} - y$, $P(1,0)$, $Q(0,1)$.
16. $f(x,y) = e^{xy} - y^2$, $P(0,0)$.

17. $f(x,y) = e^{xy} - xy$, $P(0,1)$, $Q(1,0)$, $R(1,1)$.
18. $f(x,y) = x^3y - xy^3$, $P(\sqrt{2},1)$.
19. $f(x,y) = x^2y - 2xy + xy^2$, $P(0,2)$, $Q(2,0)$.

Esercizio 8.3. *Nei seguenti problemi di massimo e minimo vincolato, dove $f(x,y)$ è la funzione da studiare e il vincolo è indicato a fianco, scrivere la funzione lagrangiana e calcolare le sue derivate prime.*

1. $f(x,y) = x + y + 1$, $x^2 - y^3 = 0$.
2. $f(x,y) = x^2 - xy^3$, $xy - x^2y^2 = 0$.
3. $f(x,y) = \sqrt{x-y}$, $x - y + x^2 + y^2 = 2$.
4. $f(x,y) = e^{x-y+2}$, $\sqrt{x-y+xy} = 1$.
5. $f(x,y) = \ln(x-y)$, $x^2 - y^2 = 3$.

Esercizio 8.4. *Delle seguenti funzioni determinare se i punti indicati sono candidati ad essere di massimo o minimo vincolato sul vincolo indicato. Se possibile dire se si tratta di massimo o minimo e determinare il massimo e minimo assoluti.*

1. $f(x,y) = x + y + 1$ *sul vincolo* $x^2 + y^2 = 2$*;* $P(0,1)$, $Q(\sqrt{2},\sqrt{2})$.
2. $f(x,y) = x + y + 1$ *sul vincolo* $xy - 1 = 0$*;* $P(1,1)$, $Q(-1,-1)$.
3. $f(x,y) = x^2$ *sul vincolo* $x - 2y - 2 = 0$*;* $P(0,1)$, $Q(0,0)$.
4. $f(x,y) = x^2$ *sul vincolo* $x^2 + y^2 = 4$*;* $P(0,2)$, $Q(-2,0)$.
5. $f(x,y) = x^2$ *sul vincolo* $xy - x^2y^2 = 0$*;* $P(0,0)$, $Q(-1,0)$.
6. $f(x,y) = xy$ *sul vincolo* $x^2 + y^2 = 4$*;* $P(0,0)$, $Q(1,0)$, $R(-1,0)$.
7. $f(x,y) = x + y$ *sul vincolo* $(x-1)^2 + (y-1)^2 = 1$*;* $P(1,1)$, $Q((\sqrt{2}+1)/\sqrt{2}, (\sqrt{2}+1)/\sqrt{2})$.
8. $f(x,y) = x$ *sul vincolo* $x^2 - y^2 = 1$*;* $P(1,1)$, $Q(1,0)$, $R(-1,0)$.
9. $f(x,y) = x^2 + y^2$ *sul vincolo* $x^2 - y^2 = 1$*;* $P(1,1)$, $Q(1,0)$, $R(-1,0)$, $S(2,\sqrt{3})$.
10. $f(x,y) = x^2 - y^2$ *sul vincolo* $x^2 + y^2 = 1$*;* $P(0,0)$, $Q(1,0)$, $R(0,1)$.
11. $f(x,y) = x^2 - y^2$ *sul vincolo* $x^2 - y^2 = 1$*;* $P(0,0)$, $Q(1,0)$, $R(0,1)$.

Esercizio 8.5. *Come per l'esercizio 8.4; se il vincolo è esplicitabile utilizzare anche il relativo metodo per la ricerca dei massimi e minimi vincolati.*

1. $f(x,y) = x + y + 1$ *sul vincolo* $x^2 - y + 3 = 0$*;* $P(0,0)$, $Q(1,-2)$.
2. $f(x,y) = x^2 - y$ *sul vincolo* $x^3 - y = 0$*;* $P(0,0)$.
3. $f(x,y) = x - y^2$ *sul vincolo* $x - y^4 - 1 = 0$*;* $P(0,0)$.
4. $f(x,y) = x^2$ *sul vincolo* $y - x^2 = 0$*;* $P(0,0)$, $Q(1,1)$.
5. $f(x,y) = xy$ *sul vincolo* $x^2 - y^2 = 4$*;* $P(-2,0)$, $Q(2,0)$, $R)0,2)$.
6. $f(x,y) = x^2 - y^2$ *sul vincolo* $x^2 - y^2 = 4$*;* $P(-2,0)$, $Q(2,0)$, $R)0,2)$.

7. $f(x,y)=x^2+y^2$ *sul vincolo* $y-x-2=0$*;* $P(0,0)$.
8. $f(x,y)=x^3-y^2$ *sul vincolo* $x^3-y^3=0$*;* $P(0,0)$, $Q(1,1)$, $R(-1,-1)$.

Notazioni utilizzate

Le notazioni utilizzate in questo testo sono quelle di default nel sistema tipografico LaTeX 2_ε, notazioni che, nella maggior parte dei casi, concordano con quelle previste dalla normativa ISO 31 — 11.

Segnaliamo inoltre che, nella numerazione dei teoremi, definizioni, osservazioni, ecc., abbiamo scelto di usare una numerazione progressiva per capitolo. Altri testi usano invece numerazioni progressive separatamente per i teoremi, le definizioni, ecc. Si tratta naturalmente solo di una questione di gusto personale.

La scrittura di un testo contenente molta matematica è sempre un'impresa ardua e che richiede molto tempo e fatica. Un aiuto indispensabile è fornito da un sistema di composizione come quello che abbiamo adottato (e che costituisce ormai lo standard de facto per i testi scientifici). Per chi fosse interessato a conoscere LaTeX 2_ε segnaliamo che si tratta di un sistema di composizione tipografica di livello professionale e assolutamente gratuito. Tutte le informazioni utili si possono trovare sul sito ufficiale della comunità degli sviluppatori, `http://www.ctan.org/` e, in lingua italiana, sul sito degli Utilizzatori italiani di TeX e LaTeX, `http://www.guit.sssup.it/`. Alcuni manuali introduttivi e consigli per iniziare si trovano anche sul sito personale del docente, `http://www.batmath.it`.

Elenco delle notazioni

$\neg$	"non" (negazione logica).
$\vee$	"vel", o, oppure (disgiunzione logica).
$\wedge$	"et", e, e contemporaneamente (congiunzione logica).
$\Rightarrow$	"implica", se … allora … (implicazione logica).
$\Leftrightarrow$	"se e solo se" (equivalenza logica).
$\mathbb{N}$	Insieme dei numeri naturali: $\mathbb{N} = \{\, 0, 1, 2, \dots, n, \dots \,\}$.
$\mathbb{Z}$	Insieme dei numeri interi: $\mathbb{Z} = \{\, \dots, -2, -1, 0, 1, 2, \dots \,\}$.
$\mathbb{Q}$	Insieme dei numeri razionali: $\mathbb{Q} = \{\, m/n \mid m \in \mathbb{Z}, n \in \mathbb{N}, n \neq 0 \,\}$.
$\mathbb{R}$	Insieme dei numeri reali.
$\mathbb{C}$	Insieme dei numeri complessi.
$\mathbb{N}^+, \mathbb{Z}^+, \mathbb{Q}^+, \mathbb{R}^+$	Numeri naturali, interi, razionali, reali, maggiori di 0.
$A, B, \dots$	Notazione per gli insiemi.
$A \subseteq B$	A è un sottoinsieme di B.
$A \subset B$	A è un sottoinsieme proprio di B.
$B \supseteq A$	B è un soprainsieme di A.
$B \supset A$	B è un soprainsieme proprio di A.
$A \setminus B$	Differenza tra gli insiemi A e B.
$[a, b]$	$\{\, x \in \mathbb{R} \mid a \leq x \leq b \,\}$.
$]a, b[$	$\{\, x \in \mathbb{R} \mid a < x < b \,\}$.

Continua nella pagina successiva

Segue dalla pagina precedente

$]a,b]$	$\{ x \in \mathbb{R} \mid a < x \leq b \}$.
$[a,b[$	$\{ x \in \mathbb{R} \mid a \leq x < b \}$.
$[a,+\infty[$	$\{ x \in \mathbb{R} \mid x \geq a \}$.
$]a,+\infty[$	$\{ x \in \mathbb{R} \mid x > a \}$.
$]-\infty,a]$	$\{ x \in \mathbb{R} \mid x \leq a \}$.
$]-\infty,a[$	$\{ x \in \mathbb{R} \mid x < a \}$.
$f: D \to C\,,\ x \mapsto f(x)$	Notazione per le funzioni.
$\exp(x) = \mathrm{e}^x$	Notazione per la funzione esponenziale di base e.
$\ln(x)$	Logaritmo in base e di x.
$\log(x)$	Logaritmo in base 10 di x.

Osservazioni

- Per alcuni autori $\mathbb{N} = \{ 1, 2, \dots, n, \dots \}$, cioè l'insieme dei naturali non comprende lo zero.
- L'insieme dei numeri razionali è in realtà l'insieme delle frazioni, come più sopra definito, ma con una opportuna relazione che renda identiche due frazioni equivalenti. Inoltre nulla cambierebbe se si prendessero frazioni in cui anche il denominatore possa essere intero (naturalmente diverso da 0).
- La notazione utilizzata in questi appunti per gli insiemi non è l'unica possibile. Altri usano per esempio lettere maiuscole in grassetto: **A**, **B**, ... e questa scelta ha qualche indubbio vantaggio, in quanto anche i punti dello spazio sono abitualmente indicati con le lettere maiuscole corsive, con possibilità di confusione. In ogni caso tutto dovrebbe essere chiaro dal contesto.
- Molti usano $\subset$ per indicare i sottoinsiemi (propri o no) e $\subsetneq$, o $\subsetneqq$ per indicare i sottoinsiemi propri. Analoga osservazione per i soprainsiemi.
- Per indicare la differenza di due insiemi molti usano il simbolo $A - B$.
- Per quanto riguarda le notazioni sui logaritmi è da segnalare che la convenzione da noi scelta è quella in uso nella maggior parte dei software di calcolo e, quasi sempre, anche nelle calcolatrici tascabili. Altri adottano la notazione $\log(x)$ per indicare il logaritmo in base "e" e la notazione $\mathrm{Log}(x)$ o esplicitamente $\log_{10}(x)$ per indicare il logaritmo in base 10 del numero x.

Alfabeto greco

Riportiamo, per l'importanza che ha nelle notazioni matematiche, l'alfabeto greco con la relativa pronuncia.

alfa	α	A	nu (ni)	ν	N
beta	β	B	csi	ξ	Ξ
gamma	γ	Γ	omicron	o	O
delta	δ	Δ	pi	π	Π
epsilon	ε	E	ro	ϱ	R
zeta	ζ	Z	sigma	σ	Σ
eta	η	H	tau	τ	T
theta	ϑ	Θ	upsilon	υ	Υ
iota	ι	I	fi	φ	Φ
cappa	$\varkappa$	K	chi	χ	X
lambda	λ	Λ	psi	ψ	Ψ
mu (mi)	μ	M	omega	ω	Ω

Come si può notare, in molti casi le lettere greche maiuscole non differiscono nella grafia dalle lettere dell'alfabeto latino.

Segnaliamo anche una lettera dell'alfabeto ebraico di uso comune in matematica.

aleph $\aleph$

Indice analitico

Appunti per un corso di matematica
Teoria ed esercizi

Luciano Battaia

Università Ca' Foscari di Venezia - Dipartimento di Economia

Versione 2.1 del 23 febbraio 2017

Questa dispensa è rivolta agli studenti che frequentano il corso di matematica del primo anno di laurea in Economia e Commercio e affini.
La dispensa può essere considerata la continuazione di una analoga raccolta di appunti per un precorso di matematica.