Corso di

Matematica per l'Economia

- Appunti -

Università Ca' Foscari di Venezia Dipartimento di Economia

Appunti per un corso di Matematica per l'Economia

Luciano Battaia

Versione del 15 febbraio 2016

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 sé 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 alcuni degli argomenti di un corso di "Matematica per l'Economia" per il corso di laurea in Economia.

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

1. Funzioni di più variabili

Ci occupiamo in questo capitolo delle funzioni reali di più variabili reali. Per favorire un più proficuo approccio a questo argomento cominceremo a considerare le funzioni di due variabili, per le quali esiste la possibilità di una rappresentazione grafica simile a quanto visto per il caso delle funzioni di una variabile. Passeremo poi al caso delle funzioni di tre o più variabili, per le quali questo tipo di rappresentazione grafica non è più possibile.

1.1. Introduzione illustrata alle funzioni di due variabili

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 naturali, come nella tabella 1.1.

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

Tabella 1.1. *Rappresentazione "tabulare" di una funzione di una variabile*

I dati di questa tabella possono essere riportati in un grafico cartesiano, come nella figura 1.1.

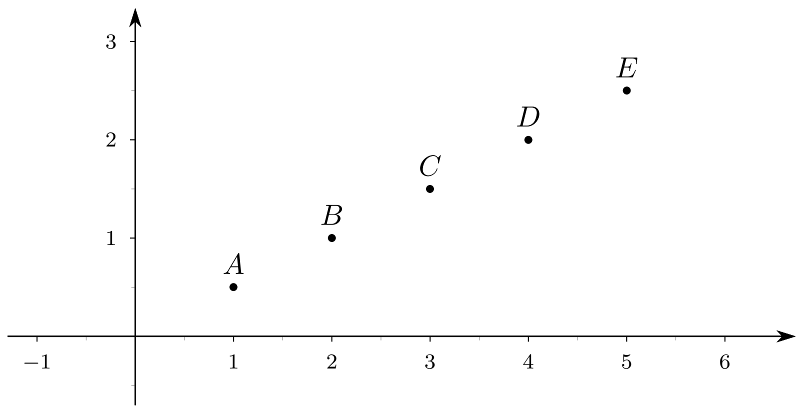


Figura 1.1. *Grafico cartesiano relativo alla tabella* 1.1

Ci interessa osservare che 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 1.2.

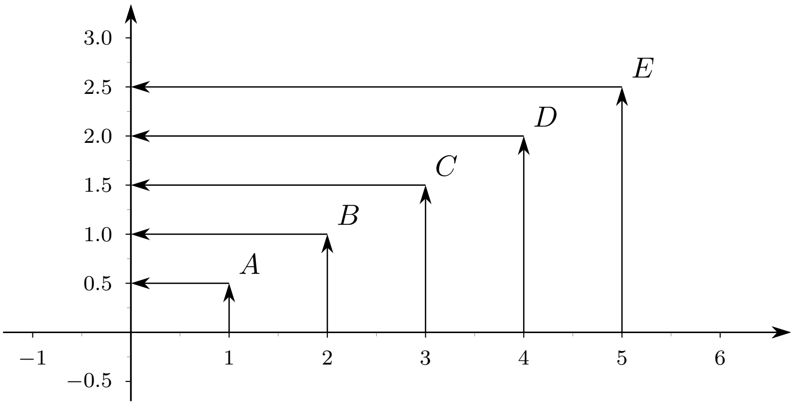


Figura 1.2. *Grafico cartesiano con frecce, relativo alla tabella 1.1*

Se si riportano nel grafico della figura 1.1 anche i punti corrispondenti ai valori di x che non compaiono nella tabella, si ottiene il risultato visualizzato nella figura 1.3: 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 studiando le funzioni di una variabile.

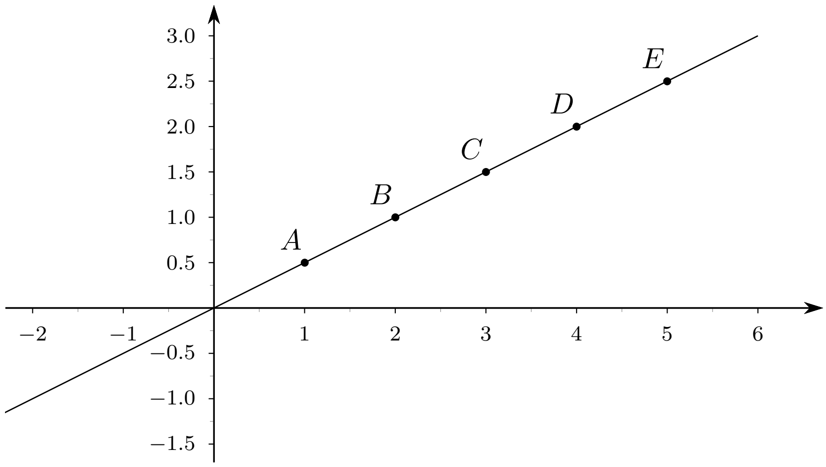


Figura 1.3. *Grafico della funzione $y = {}^{x}/_{2}$, comprendente i punti della figura 1.1*

Le funzioni reali di due variabili sono funzioni in cui il dominio è un sottoinsieme di $\mathbb{R}^2$ e il codominio è l'insieme $\mathbb{R}$ dei numeri reali. Se indichiamo con (x, y) un generico "punto" di $\mathbb{R}^2$, ovvero una generica coppia di numeri reali (è questa coppia che costituisce la variabile indipendente), e indichiamo con z il numero reale corrispondente a questa coppia, potremo usare una scrittura del tipo

$$z = f(x, y) \tag{1.1}$$

Se consideriamo per esempio la funzione di due variabili data dalla legge $f(x, y) = x + y$, potremo ancora costruire una tabella come la 1.1, 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).

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, $Oxyz$.

x	y	$x+y$
1	0	1
0	1	1
1	1	2
1	-1	0
3	0	3
2	1	3
-1	4	3
...	...	...

Tabella 1.2. *Rappresentazione "tabulare" di una funzione di due variabili*

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 1.4 per un singolo punto (x,y) del dominio.

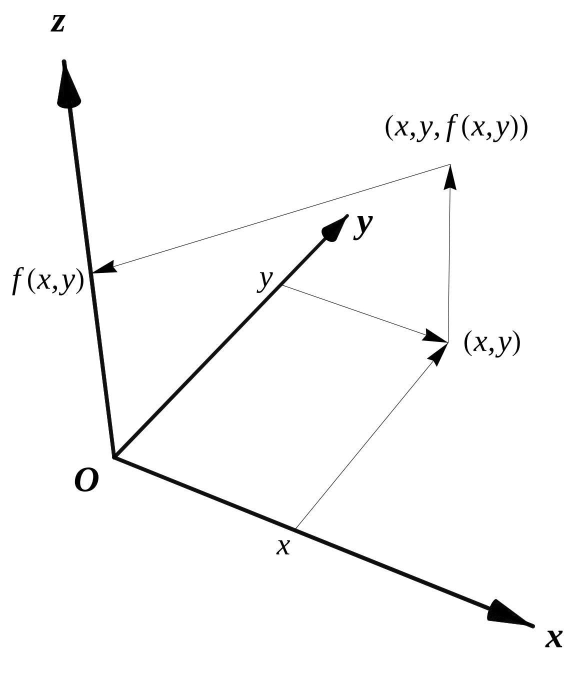


Figura 1.4. *Procedimento per tracciare il grafico di una funzione di due variabili*

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 1.5.

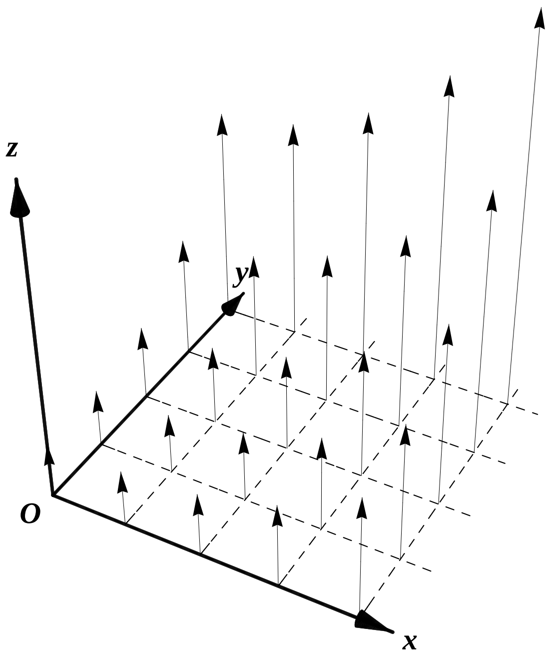


Figura 1.5. *Un "boschetto" di frecce*

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".

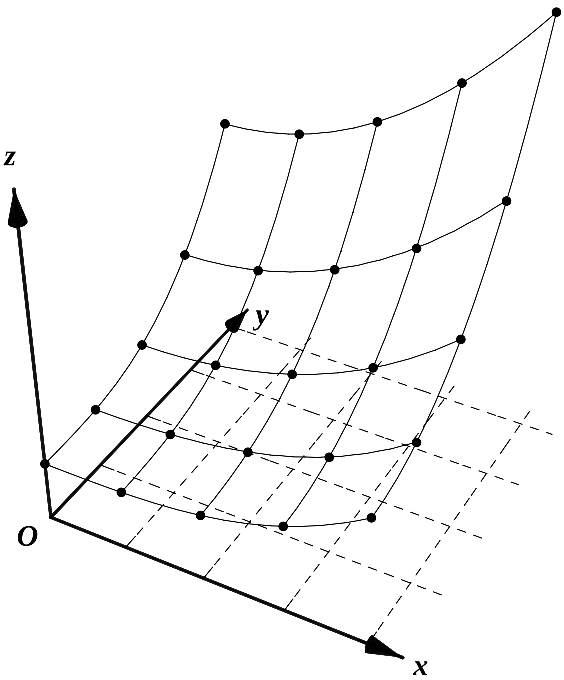


Figura 1.6. *Una superficie-grafico*

Per rendere più significativo il grafico si possono introdurre anche colorazioni come nella figura 1.7.

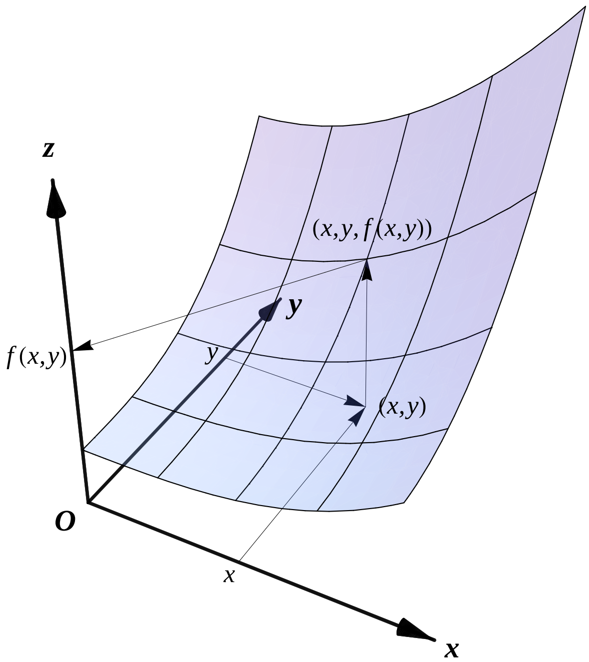


Figura 1.7. *Uso di colorazioni per le superfici-grafico*

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 di massimo e minimo (relativo o assoluto). Come suggerisce la figura 1.8, potremo usare l'appellativo *monte* e *cima* per riferirci ai massimi, l'appellativo *valle* e *fondovalle* per riferirci ai minimi.

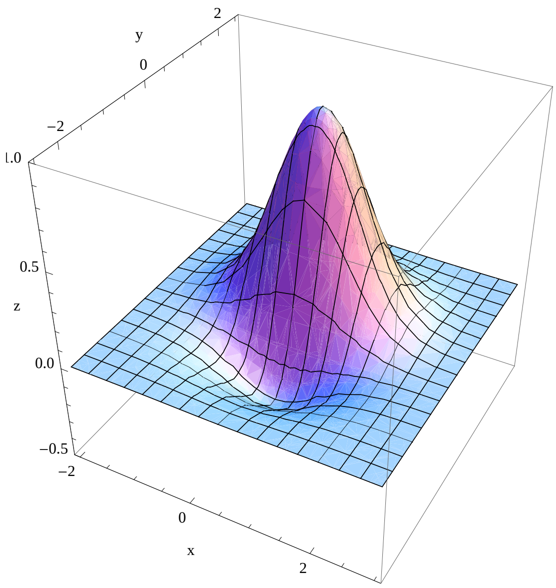


Figura 1.8. *Massimi e minimi in una funzione di due variabili*

La figura 1.9 mostra, come ulteriore esempio, una situazione in cui sono presenti due monti e una valle.

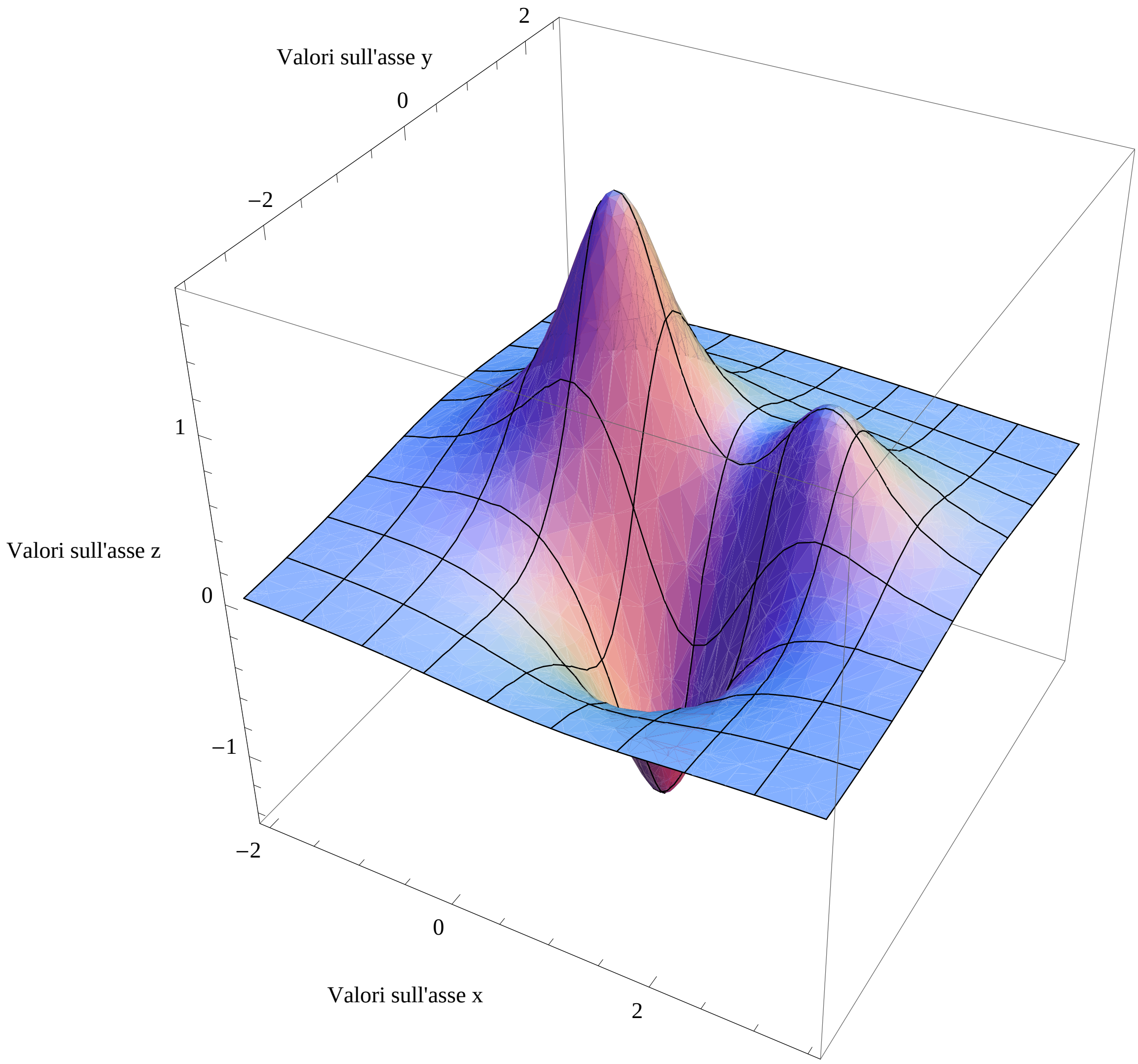


Figura 1.9. *Una funzione con due "monti" e una "valle"*

In queste figure non sono tracciati gli assi con le stesse convenzioni usate per le funzioni di una variabile, 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 1.10 mostra alcune di queste linee per la stessa superficie della figura 1.9.

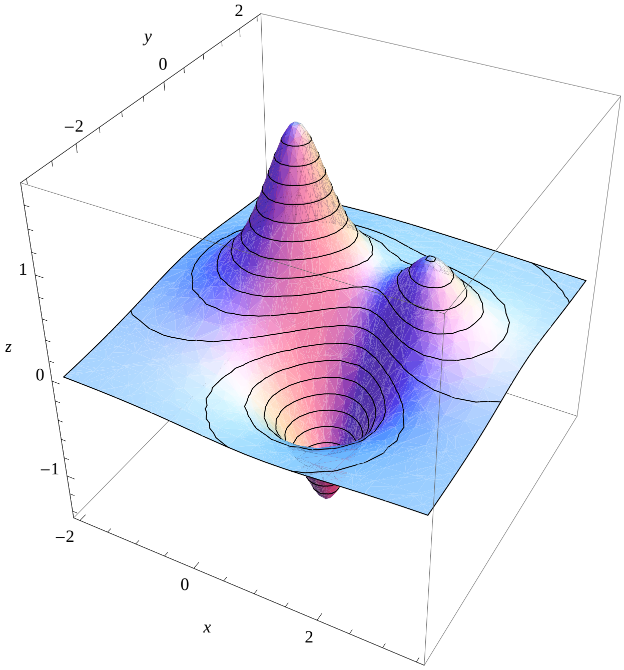


Figura 1.10. *Linee di livello*

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 1.9 nella figura 1.11.

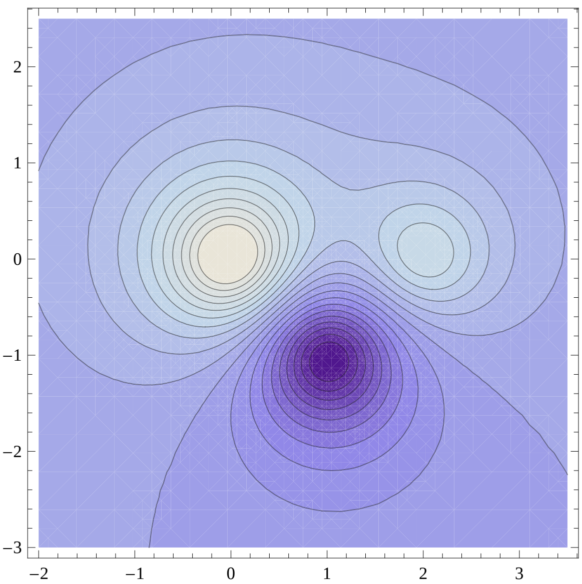


Figura 1.11. *Linee di livello raccolte sul piano Oxy*

La figura 1.12 mostra come si ottiene una delle linee di livello mediante intersezione della superficie con un piano orizzontale.

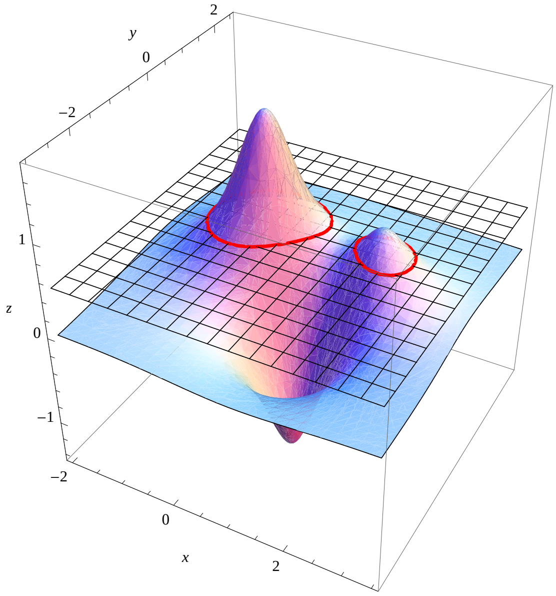


Figura 1.12. *Sezione di una superficie con un piano orizzontale*

Ritornando alla piastrellatura della figura 1.9, 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.

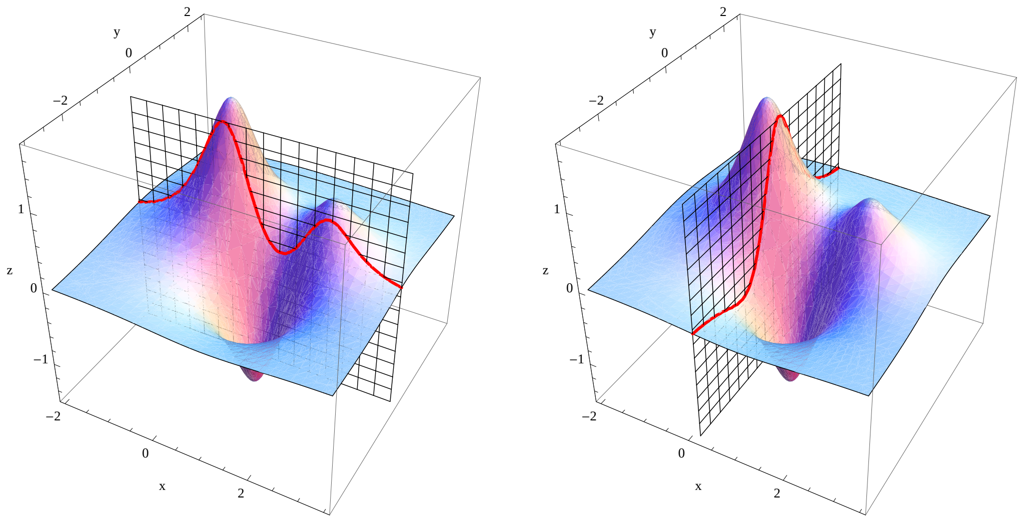


Figura 1.13. *Sezione di una superficie con piani verticali paralleli a Oxz e a Oyz*

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 1.14 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.

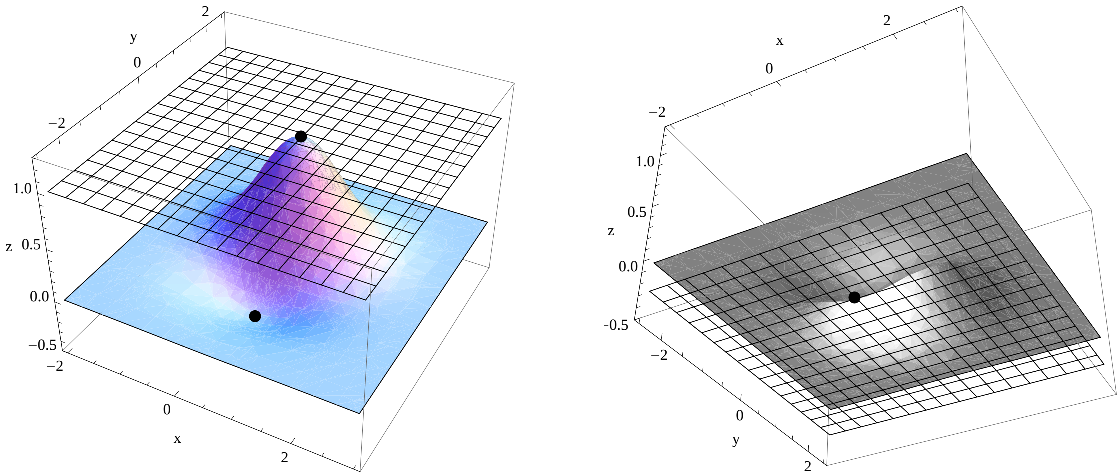


Figura 1.14. *Piani tangenti in un punto di massimo e in un punto di minimo*

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.

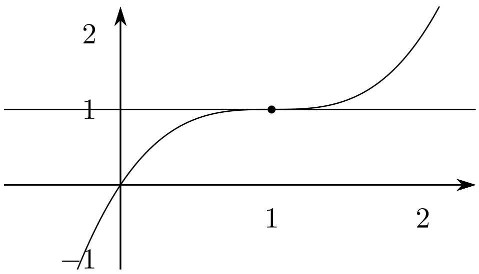


Figura 1.15. *Un flesso a tangente orizzontale*

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 1.16 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 1.17 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.

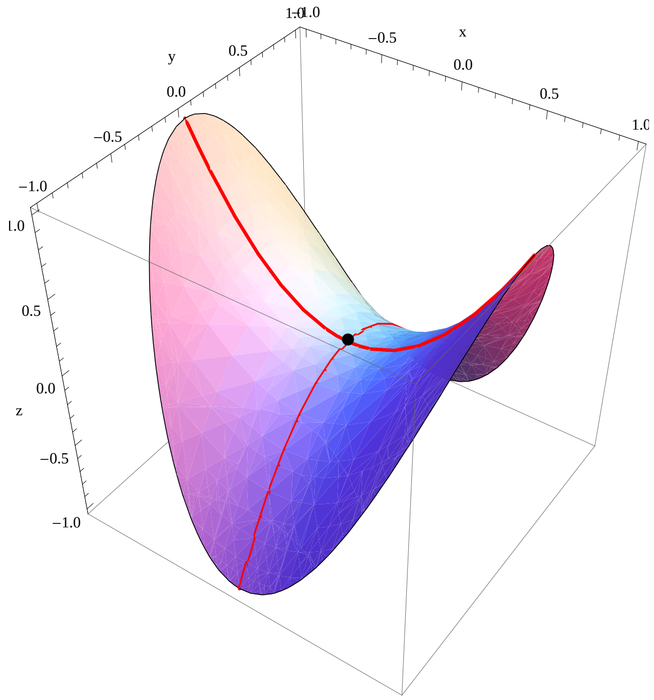


Figura 1.16. *Una "sella di cavallo"*

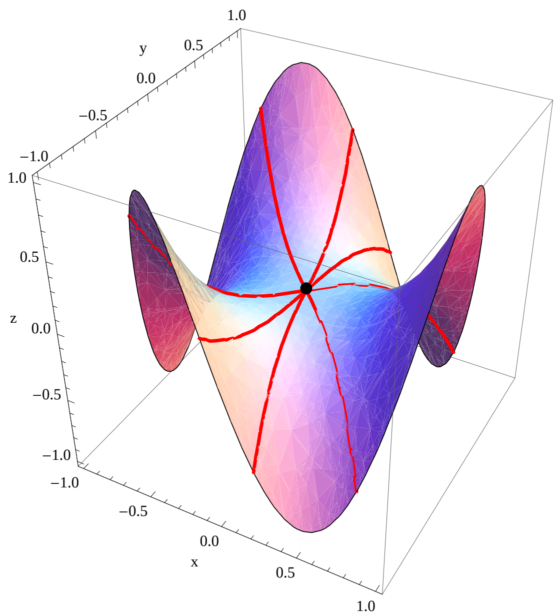


Figura 1.17. *Una "sella di scimmia"*

1.2. 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 $2x + 3y - z = 0$.

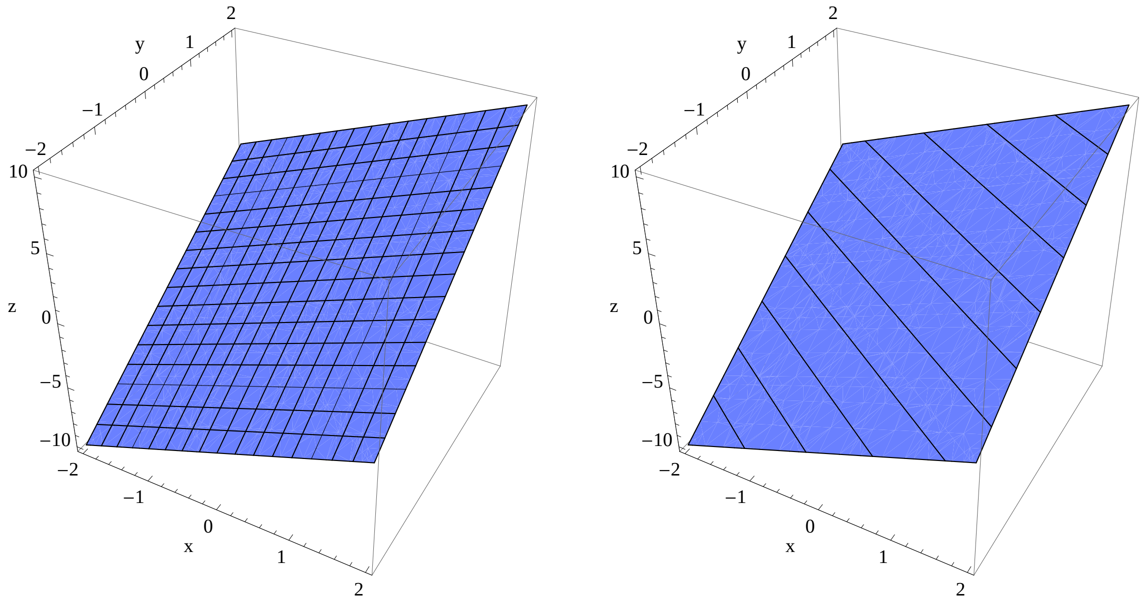


Figura 1.18. *Piano* $z = 2x + 3y$

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.

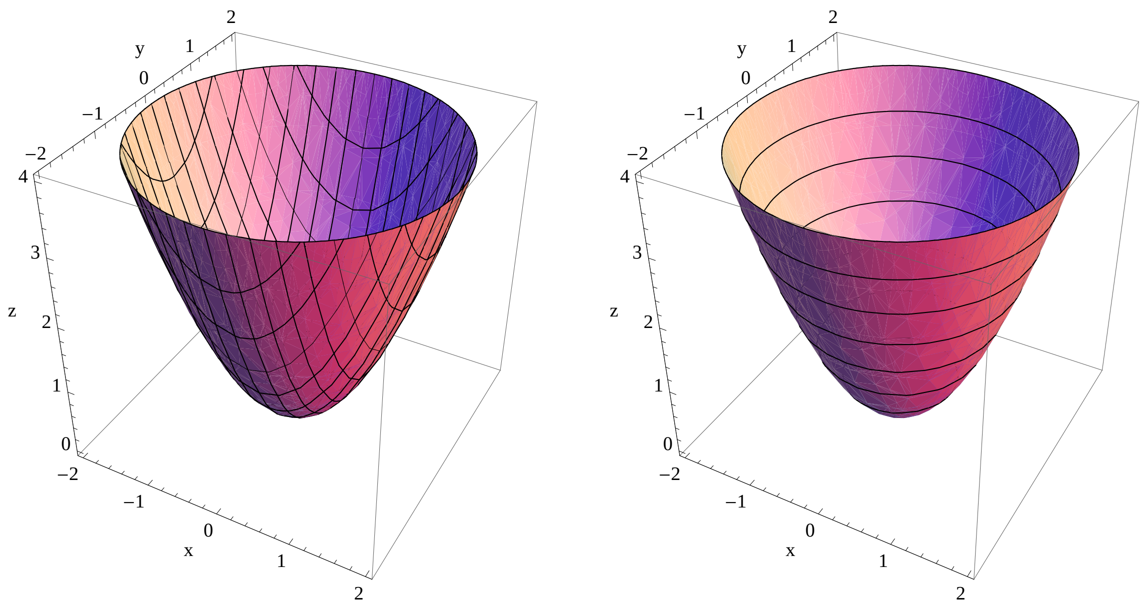


Figura 1.19. *Paraboloide* $z = x^2 + y^2$

3. Paraboloide a sezione ellittica: $z = 3x^2 + y^2$. Superficie simile a quella della figura 1.19, ma con curve di livello a sezione ellittica con centro sull'asse z.

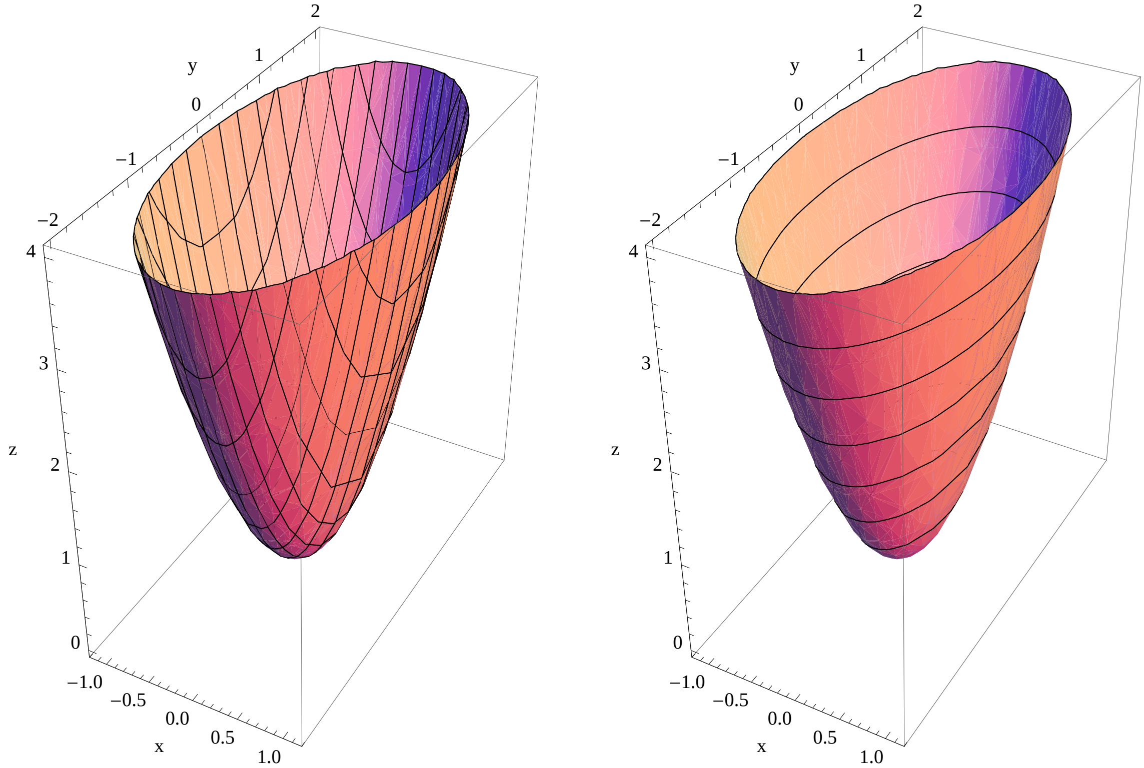


Figura 1.20. *Paraboloide a sezione ellittica:* $z = 3x^2 + y^2$

4. La sella $z = x^2 - y^2$.

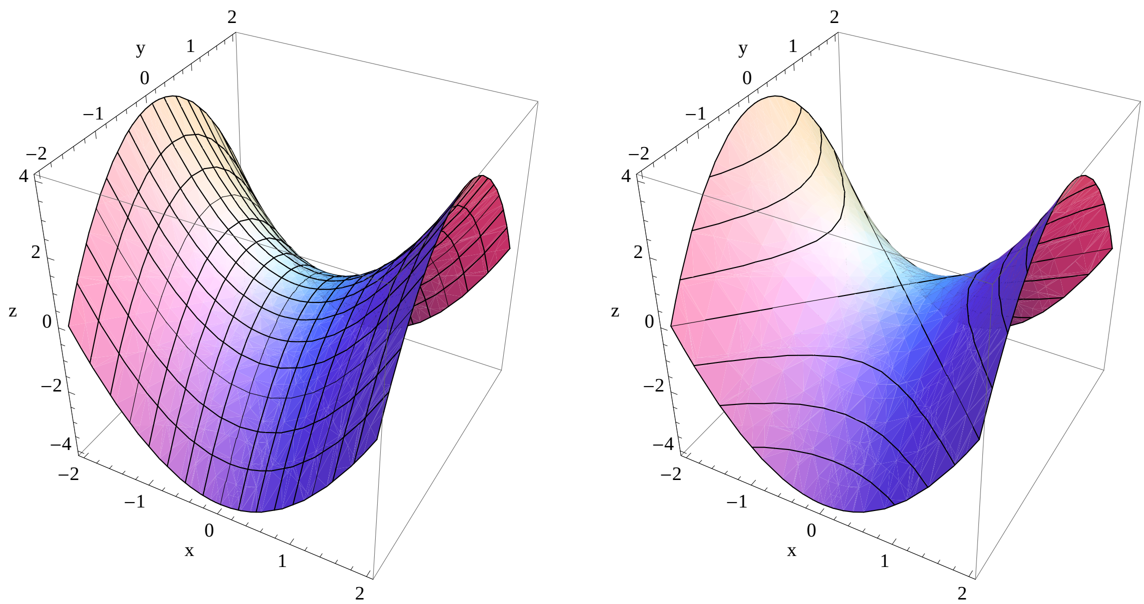


Figura 1.21. *La sella* $z = x^2 - y^2$

5. La superficie $z = x^2$. Si tratta della superficie ottenuta traslando la parabola $z = x^2$ lungo l'asse delle x.

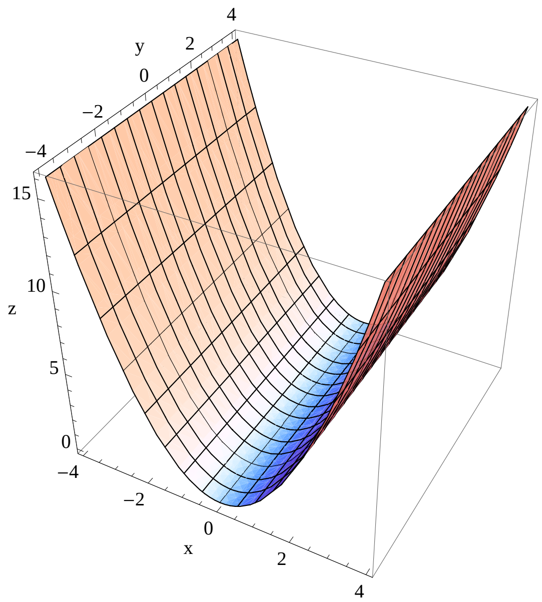


Figura 1.22. *La superficie* $z = x^2$

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.

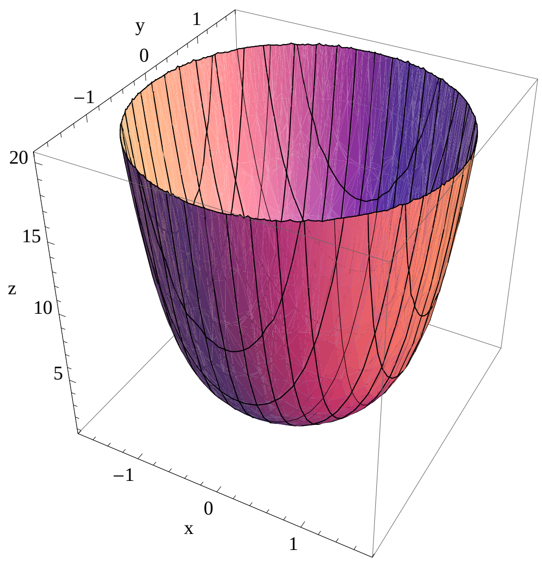


Figura 1.23. *La superficie* $z = \mathrm{e}^{x^2+y^2}$

1.3. Funzioni reali di tre o più variabili

Il passaggio dalle funzioni reali di una variabile a quelle di due variabili ha comportato alcune difficoltà, come per esempio la rinuncia alla possibilità di parlare di funzioni crescenti e decrescenti[1], tuttavia le loro caratteristiche importanti possono ancora essere illustrate utilizzando un grafico cartesiano. Quando si passa a trattare le funzioni reali di tre o più variabili le cose si complicano in quanto per tabulare una funzione, per esempio, di tre variabili abbiamo bisogno di *quattro* colonne: *tre* per le variabili indipendenti e *una* per la variabile dipendente. Una rappresentazione grafica necessiterebbe dunque di uno spazio a quattro dimensioni, che non è possibile visualizzare mediante illustrazioni, anche se la trattazione matematica è perfettamente legittima. Le cose si complicano ulteriormente aumentando il numero di variabili.

Nonostante questo è ancora possibile parlare di massimi e minimi (relativi o assoluti), ovvero di problemi di ottimizzazione che sono quelli che interessano maggiormente l'economia.

Per le funzioni di tre o più variabili possiamo usare notazioni simili alla 1.1, usata per il caso di due variabili. Avremo naturalmente necessità di introdurre nuove lettere per le variabili indipendenti, che aumentano di numero. È più utile però utilizzare la x con un indice per indicare le variabili indipendenti e la y per indicare la variabile dipendente:

$$y = f(x_1, x_2), \quad y = f(x_1, x_2, x_3), \quad y = f(x_1, x_2, x_3, x_4), \quad \dots \tag{1.2}$$

1.4. Richiami sui vettori

.

Definizione 1.1. *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 1.2. *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 1.3. *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] Ulteriori difficoltà si incontreranno, come vedremo, con il concetto di derivata.

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 1.4. *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 1.1. 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:

	A	B	C
P_1	100	50	80
P_2	60	200	30
P_3	150	180	250

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.

Ci interesserà anche nel seguito il concetto di *norma* di un vettore.

Definizione 1.5. *Dato un vettore $\vec{u} = (u_1, u_2, \dots, u_n)$, chiameremo* norma *di $\vec{u}$ il numero reale positivo*

$$\|\vec{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}. \tag{1.3}$$

Diremo poi versore *un vettore di norma* 1.

Al posto di norma si può anche usare la dicitura *modulo*, anche se non bisogna fare confusione con il modulo di un numero reale.

Utilizzando le notazioni vettoriali potremo anche uniformare le scritture delle funzioni:

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

Naturalmente dovrà essere chiaro dal contesto se

$$\vec{x} = (x_1, x_2) \quad \text{oppure} \quad \vec{x} = (x_1, x_2, x_3), \quad \text{ecc.}$$

A volte sarà utile chiamare punti gli elementi di $\mathbb{R}^n$, ed indicarli con P, Q, ...:

$$P = (x_1, x_2), \quad P = (x_1, x_2, x_3), \quad \dots \tag{1.5}$$

Potremo anche scrivere $f(P)$ al posto di $f(\vec{x})$.

Nel caso $n = 1$ i punti sono semplicemente numeri reali.

1.5. Operazioni sulle funzioni

Come per le funzioni di una variabile, anche le funzioni di più variabili si possono sommare, sottrarre, moltiplicare e dividere (quest'ultima operazione con le solite limitazioni sul denominatore): in effetti sommare, sottrarre, moltiplicare o dividere due funzioni reali significa eseguire queste operazioni sul codominio, che, per i casi che abbiamo considerato noi, rimane sempre $\mathbb{R}$.

Ci interessa esaminare in dettaglio un'altra operazione importante sulle funzioni e precisamente la *composizione* di due funzioni: date due funzioni f e g si tratta di farle agire in successione, ovvero usare il risultato della prima (in termini informatici diremmo l'output) come input per la seconda, ottenendo alla fine il risultato voluto. Se la prima funzione è f e la seconda funzione è g allora la composta si indica con $g \circ f$ e il risultato (output) finale si indica con $g(f(x))$: si presti particolare attenzione al fatto che la prima funzione è la più interna nella scrittura, la seconda la più esterna. Per rendere ancora più chiara la successione delle operazioni si potranno anche usare scritture del tipo

$$x \xrightarrow{\ f\ } f(x) \xrightarrow{\ g\ } g(f(x)) \quad \text{oppure} \quad A \xrightarrow{\ f\ } B \subseteq C \xrightarrow{\ g\ } D, \tag{1.6}$$

dove A e B sono il dominio e codominio di f, C e D il dominio e codominio di g.

Per poter eseguire questa operazione l'insieme immagine della prima funzione deve essere contenuto nel dominio della seconda, altrimenti la composizione non si può fare. Già nel caso di funzioni di una variabile si presentava questo problema, che comunque poteva di norma essere risolto con opportune restrizioni. Per esempio se la prima funzione è $f(x) = x^2 - 1$ e la seconda è $g(x) = \sqrt{x}$, la composta non si può fare a cuor leggero perché se, per esempio, $x = 0$, la prima dà come risultato -1, che non può essere "dato in pasto" alla seconda. Il problema può essere risolto semplicemente restringendo il dominio di f agli $x \leq -1 \vee x \geq 1$.

Ben diverso il caso quando si passa a considerare funzioni reali di più variabili. Se per esempio f e g sono funzioni di due variabili ($f, g \colon \mathbb{R}^2 \to \mathbb{R}$) non si potrà in alcun modo considerare la composta: il risultato di ciascuna delle due è un unico numero reale, che non può far parte del dominio dell'altra, che è in ogni caso costituito da coppie di numeri reali. Se invece f è una funzione di due variabili ($f\colon \mathbb{R}^2 \to \mathbb{R}$) e g una funzione di una variabile ($g\colon \mathbb{R} \to \mathbb{R}$) si potrà (salvo opportune restrizioni) considerare la funzione $g \circ f$, mentre non avrà alcun senso $f \circ g$: questo risulta chiaro da una scrittura come quella della (1.6).

$$(x_1, x_2) \xrightarrow{\quad f \quad} f(x_1, x_2) \xrightarrow{\quad g \quad} g(f(x, y)) \quad \text{oppure} \quad \mathbb{R}^2 \xrightarrow{\quad f \quad} \mathbb{R} \xrightarrow{\quad g \quad} \mathbb{R}.$$

Esempio 1.2. Siano $f(x_1, x_2) = x_1 + 3x_1x_2 - x_2^3$ e $g(x) = \sin x$. Allora

$$(g \circ f)(x_1, x_2) = g(f(x_1, x_2)) = \sin(x_1 + 3x_1x_2 - x_2^3).$$

La costruzione della funzione composta non richiede alcuna necessità di restrizioni sul dominio di f.

Esempio 1.3. Siano $f(x_1, x_2) = x_1^2 + y_1^2 - 1$ e $g(x) = \sqrt{x}$. Allora

$$(g \circ f)(x_1, x_2) = g(f(x_1, x_2)) = \sqrt{x_1^2 + y_1^2 - 1},$$

ma la costruzione della funzione composta richiede una restrizione sul dominio di f: occorre che

$$x_1^2 + y_1^2 - 1 \geq 0,$$

ovvero occorre considerare solo i punti esterni alla circonferenza di centro l'origine e raggio 1, con l'aggiunta dei punti della stessa circonferenza.

In generale, per i tipi di funzioni a cui siamo interessati in questo corso, potremo solo considerare composizioni in cui la funzione più interna sia una $f\colon \mathbb{R}^n \to \mathbb{R}$ e la funzione più esterna sia una $g\colon \mathbb{R} \to \mathbb{R}$:

$$\mathbb{R}^n \xrightarrow{\quad f \quad} \mathbb{R} \xrightarrow{\quad g \quad} \mathbb{R}, \quad \Rightarrow \quad \mathbb{R}^n \xrightarrow{\quad g\circ f \quad} \mathbb{R}. \tag{1.7}$$

L'effettiva costruibilità della funzione composta può comportare restrizioni sul dominio di f.

1.6. Insiemi limitati e illimitati in n dimensioni

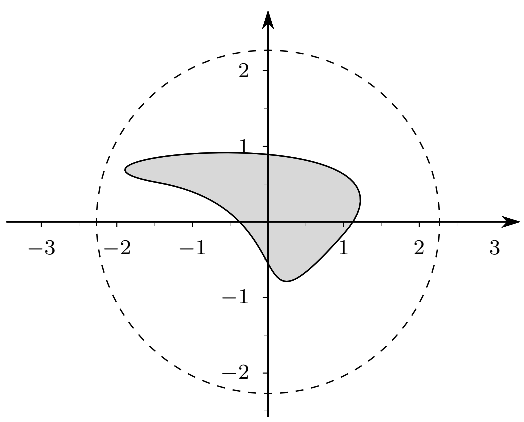
Anche per i sottoinsiemi del piano, cioè di $\mathbb{R}^2$, si può introdurre il concetto di insieme limitato e illimitato, ma la cosa è diversa dal caso degli insiemi sulla retta, perché sulla retta reale esiste un ordine (cioè nei numeri reali si può parlare di maggiore e di minore), mentre nel piano 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.

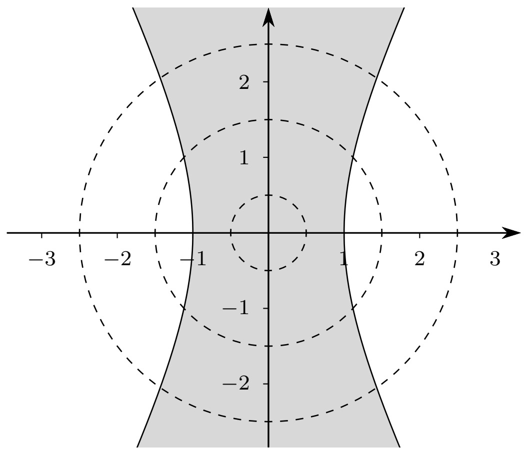
Esempio 1.4.

Un insieme limitato del piano e il cerchio che lo contiene



Esempio 1.5.

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



L'estensione del concetto di insieme limitato e illimitato a $\mathbb{R}^n$ è molto semplice: basta sostituire al cerchio di centro l'origine e raggio r, una palla di centro l'origine e raggio r.

Definizione 1.7. *Sia $A \subseteq \mathbb{R}^n$ un sottoinsieme di $\mathbb{R}^n$. A si dice* limitato *se esiste una palla di centro l'origine e raggio r che lo contiene, altrimenti si dice* illimitato.

1.7. Un po' di topologia

Per poter parlare di limiti e continuità per funzioni di più variabili occorre introdurre i concetti di distanza e di intorno. Introdurremo anche alcuni altri concetti, connessi a quello di intorno, che ci saranno utili nel seguito.

Definizione 1.8. *Dati due punti*

$$P = (x_1, x_2, \dots, x_n) \quad e \quad Q = (x'_1, x'_2, \dots, x'_n)$$

di $\mathbb{R}^n$, diremo loro distanza *il numero reale positivo*

$$d(P,Q) = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + \cdots + (x_n - x'_n)^2}. \tag{1.8}$$

Si tratta nella sostanza della generalizzazione della usuale distanza tra due punti determinata con il teorema di Pitagora.

Definizione 1.9. *Dato un punto P di $\mathbb{R}^n$, diremo* palla *o* palla aperta *di centro P e raggio r l'insieme di tutti i punti che hanno da P distanza minore di r; diremo* palla chiusa *di centro P e raggio r l'insieme di tutti i punti che hanno da P distanza minore o uguale a r.*

Se $n = 1$ una palla è semplicemente un intervallo che ha centro su P, se $n = 2$ è un cerchio di centro P, se $n = 3$ una sfera di centro P, se $n > 3$ non ne possiamo dare un'interpretazione geometrica: potremo parlare di *ipersfera* di centro P.

Definizione 1.10. *Dato un punto P di $\mathbb{R}^n$ diremo* intorno *di P una qualunque palla aperta cui P appartiene. Indicheremo con $I(P)$ un generico intorno di un punto P.*

Definizione 1.11. *Dato un sottoinsieme A di $\mathbb{R}^n$, un punto P di $\mathbb{R}^n$ si dice di* accumulazione *per A se in ogni intorno di P cadono infiniti punti di A.*

Esattamente come succedeva nel caso di $\mathbb{R}$, un punto di accumulazione può appartenere o no all'insieme A.

Esempio 1.6. Sia $A \subset \mathbb{R}^2$, $A = \{ (x, y) \in \mathbb{R}^2 \mid (x, y) \neq (0, 0) \}$ (A è l'insieme dei punti del piano diversi dall'origine). Allora tutti i punti del piano (compresa l'origine che non appartiene all'insieme) sono di accumulazione per A. Questo insieme è il dominio, per esempio, della funzione

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

Esempio 1.7. Sia $A \subset \mathbb{R}^2$, $A = \{ (x, y) \in \mathbb{R}^2 \mid x \neq 0 \wedge y \neq 0 \}$ (A è l'insieme dei punti del piano fuori dagli assi). Allora tutti i punti di A sono di accumulazione per A (e appartengono all'insieme), ma anche tutti i punti sui due assi sono di accumulazione (e questi non appartengono all'insieme). Questo insieme è il dominio, per esempio, della funzione

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

Aggiungiamo ora alcune altre definizioni che estendono al caso di $\mathbb{R}^n$ analoghi concetti già visti sulla retta reale. Alcuni degli esempi proposti sono relativi ad $\mathbb{R}$, proprio per evidenziare le analogie tra i concetti su $\mathbb{R}$ e gli analoghi su $\mathbb{R}^n$.

Definizione 1.12 (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 1.13 (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 1.14 (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 1.15 (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.*

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

Esempio 1.8. 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 1.9. 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.

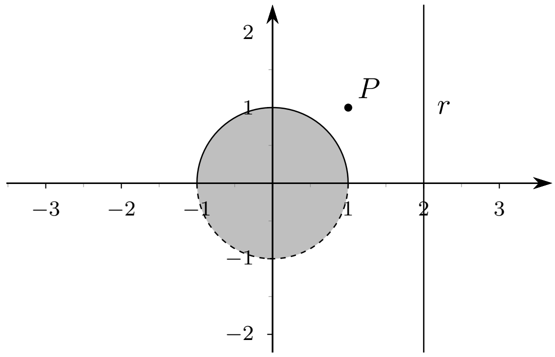


Figura 1.24. *Un insieme del piano*

- 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 dall'interno del cerchio di centro l'origine e raggio 1 e dai punti della retta r.
- L'insieme A e un insieme illimitato del piano.

Definizione 1.16 (Insieme chiuso)**.** *Un insieme A si dice* chiuso *se contiene tutti i suoi punti di accumulazione.*

Definizione 1.17 (Insieme aperto)**.** *Un insieme A si dice* aperto *se il suo complementare è chiuso.*

Esempio 1.10. 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 1.11. 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 1.12. Un intervallo del tipo $[a, b[$, oppure $]a, b]$ non è né aperto né chiuso.

Esempio 1.13. L'insieme del piano tracciato nella figura 1.24 non è né aperto né chiuso. Se a questo insieme aggiungo la semicirconferenza inferiore, diventa un insieme chiuso.

Esempio 1.14. 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.

1.8. Insiemi connessi. Insiemi convessi

Definizione 1.18 (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]

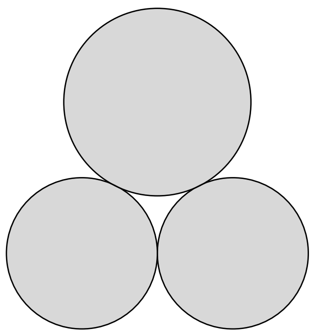
In $\mathbb{R}$ sono connessi tutti e soli gli intervalli, di qualunque tipo. In $\mathbb{R}^2$ le palle (aperte o chiuse) sono sempre connesse, ma ci sono anche insiemi connessi più complessi, come per esempio l'insieme costituito dai punti del primo e terzo quadrante, inclusi gli assi cartesiani.

Definizione 1.19 (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.

Esempio 1.15.

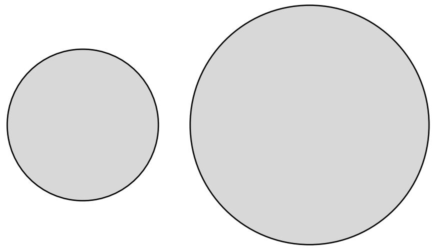
Un insieme connesso ma non convesso (le tre circonferenze bordo sono comprese nell'insieme).



[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.

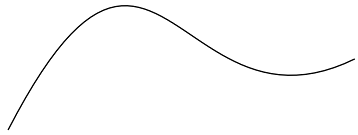
Esempio 1.16.

Un insieme non connesso (e quindi nemmeno convesso).



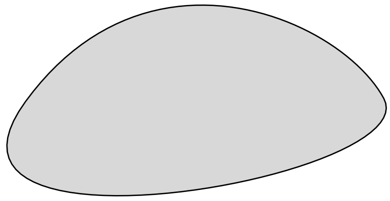
Esempio 1.17.

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



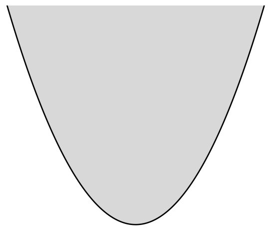
Esempio 1.18.

Un insieme connesso e convesso.



Esempio 1.19.

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



La definizione di insieme convesso si estende a sottoinsiemi di $\mathbb{R}^n$ previa la definizione di segmento in $\mathbb{R}^n$: naturalmente in $\mathbb{R}^2$ e $\mathbb{R}^3$ si otterrà il concetto già noto dalla geometria elementare. Non è difficile provare che se consideriamo una funzione $f\colon \mathbb{R} \to \mathbb{R}^2$ del tipo[3]

$$\begin{cases} x = at + b \\ y = ct + d \end{cases}, \quad t \in \mathbb{R}$$

il suo insieme immagine, cioè l'insieme dei punti del piano che hanno le coordinate (x, y) che siottengono dalle precedenti equazioni al variare di t in $\mathbb{R}$, è una retta. Infatti ricavando t da una delle due equazioni e sostituendo nell'altra si ottiene un'equazione di primo grado nelle due variabili x ed y che, come sappiamo, rappresenta una retta del piano. Se invece di far variare t su tutto $\mathbb{R}$ lo facciamo variare solo su un intervallo $[\alpha, \beta]$, otterremo invece di tutta la retta solo un segmento, di estremi P e Q, dove $P = (a\alpha + b, c\alpha + d)$ e $Q = (a\beta + b, c\beta + d)$. Un segmento di $\mathbb{R}^n$ sarà allora l'immagine di una funzione $f\colon \mathbb{R} \to \mathbb{R}^n$, in cui tutte le componenti sono al massimo di primo grado nella variabile t.

Definizione 1.20 (Insieme convesso di $\mathbb{R}^n$)**.** *Un insieme $A \subseteq \mathbb{R}^n$ si dice* convesso *quando presi comunque due suoi punti P e Q esiste un segmento che li connette e tutto contenuto in A.*

1.9. Cenno su limiti e continuità

Possiamo ora estendere la definizione di limite che è stata data per funzioni di una variabile anche al caso di $\mathbb{R}^n$, quasi con le stesse parole. Unica differenza importante è che in $\mathbb{R}^n$ non si

[3] Come sarà precisato più avanti nel capitolo 4, una funzione $f\colon \mathbb{R} \to \mathbb{R}^2$ è una coppia di funzioni di $\mathbb{R}$ in $\mathbb{R}$.

possono introdurre i concetti di $+\infty$ e $-\infty$: si può parlare solo genericamente di punti all'∞ (senza segno), e si può chiamare *intorno di* ∞ *in* $\mathbb{R}^n$ l'esterno di una qualunque palla centrata sull'origine. Con questa precisazione si può ripetere quasi pari pari la definizione data per funzioni di una variabile.

Definizione 1.21 (Limite in più variabili). *Sia data una funzione $f(P)$, di dominio D, e sia P_0 un punto di accumulazione per D (non essendo escluso che P_0 possa essere l'infinito). Diremo che l (non essendo escluso che l possa essere uno dei due simboli di infinito[4]) è il* limite *di $f(P)$ per P tendente a P_0, e scriveremo*

$$\lim_{P \to P_0} f(P) = l \tag{1.9}$$

se, scelto un arbitrario intorno I_l di l, è possibile trovare in corrispondenza un opportuno intorno $I(P_0)$ di P_0, in modo tale che i valori della funzione calcolati in $I(P_0)$, tranne *P_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 più variabili hanno dominio in $\mathbb{R}^n$, 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 1.22 (Continuità in più variabili). *Sia data una funzione $f(P)$, di dominio D, e sia P_0 un punto di accumulazione per D,* appartenente a *D. La funzione f si dice* continua *in P_0 se*

$$\lim_{P \to P_0} f(P) = f(P_0)\,.$$

Anche qui è come dire che una funzione è continua se il calcolo del limite si può fare semplicemente sostituendo P_0 al posto di P 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 più variabili è estremamente complesso e non alla portata di questo corso, per cui non ce ne occuperemo.

1.10. Rette, piani, iperpiani

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 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

[4] Il valore l del limite appartiene alla retta reale estesa, in quanto la funzione f ha come codominio $\mathbb{R}$.

figura 1.25 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.

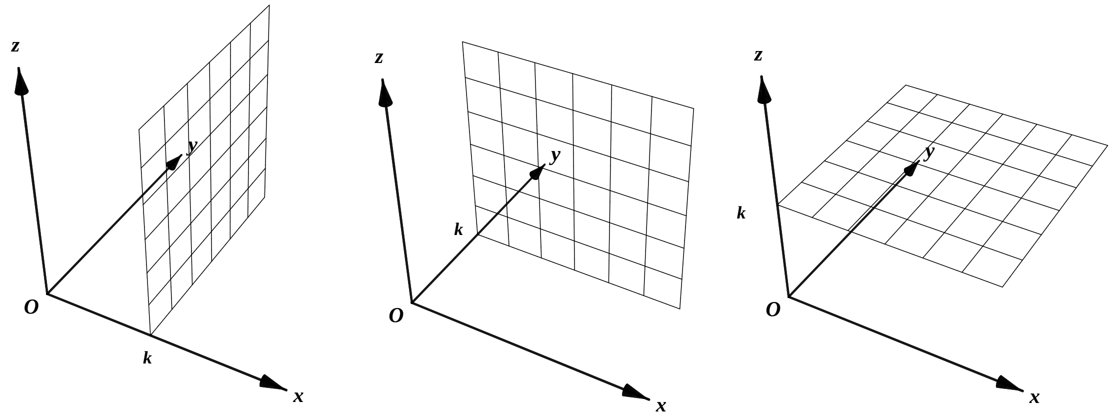


Figura 1.25. *Piani $x = k$, $y = k$, $z = k$, rispettivamente*

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

$$z = mx + ny + q \,, \tag{1.10}$$

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 $x = 0$.

Esaminando in dettaglio le equazioni di rette non verticali nel piano e piani non verticali nello spazio, possiamo concludere che essi sono il grafico di funzioni di primo grado in una o due incognite

$$y = mx + q \quad \text{oppure} \quad z = mx + ny + q.$$

Usando le notazioni che abbiamo indicato per le funzioni potremo scrivere queste funzioni in maniera similare:

$$y = m_1x_1 + q \quad \text{oppure} \quad y = m_1x_1 + m_2x_2 + q. \tag{1.11}$$

In questo modo è immediata la generalizzazione al caso di più di due variabili:

$$y = m_1x_1 + m_2x_2 + \cdots + m_nx_n + q. \tag{1.12}$$

Le funzioni del tipo 1.12 si chiamano *funzioni affini* e, se $q = 0$ (ovvero se manca il termine noto) si chiamano anche *funzioni lineari*. Come abbiamo visto, nel caso di una variabile hanno come grafico rette non verticali, nel caso di due variabili piani non verticali, nel caso di più di due variabili parleremo di *iperpiani* non verticali. Se manca il termine noto le rette, i piani o gli iperpiani *passano per l'origine*.

Le funzioni lineari e le funzioni affini hanno una grande importanza nelle applicazioni. Possiamo valutare questo fatto, a livello elementare, con un semplice esempio nel caso di una variabile.

Data la funzione $f(x) = \ln x$, la retta tangente al suo grafico nel punto di ascissa 1 è $y = x - 1$. Questa retta tangente ha, nei pressi del punto $x = 1$, un grafico che si discosta molto poco da quello della funzione logaritmo, tanto che, almeno in prima approssimazione, si può pensare di confondere il grafico del logaritmo con quello della tangente trovata: la funzione logaritmo, nei pressi del punto in esame, può essere linearizzata con la sua retta tangente: una bella comodità, visto che i calcoli con una funzione affine sono decisamente più semplici che non con una generica funzione.

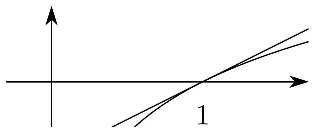


Figura 1.26. *Linearizzazione della funzione logaritmo in un intorno di* 1

Come è noto, per funzioni di una variabile questa possibilità è assicurata nel caso di funzioni derivabili. Vedremo che cosa succede per funzioni di più variabili.

1.11. Linee di livello e intersezioni con piani verticali

Nel caso, particolarmente importante, delle funzioni di due variabili per le quali è possibile una rappresentazione grafica cartesiana, hanno interesse anche le linee di livello, di cui abbiamo già parlato, e le intersezioni con piani verticali. Ne diamo ora una definizione formale.

Definizione 1.23 (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*

$$\begin{cases} z = f(x,y) \\ z = k \end{cases}, \tag{1.13}$$

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.

Esempio 1.20. 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 1.27.

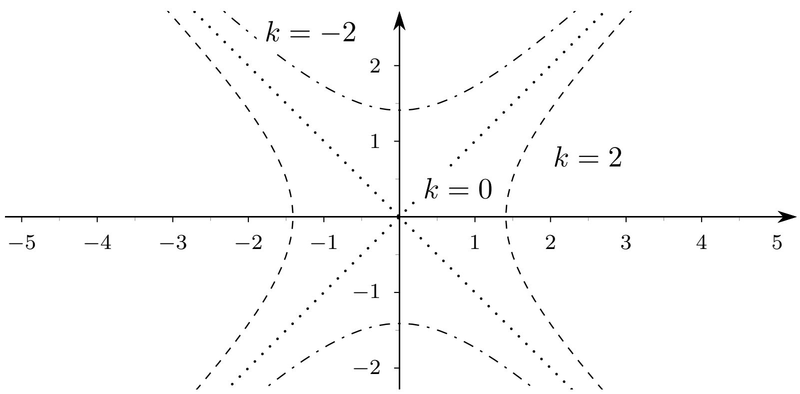


Figura 1.27. *Tre linee di livello per la funzione $f(x,y) = x^2 - y^2$*

Esempio 1.21. Data $f(x,y) = x - y^2$, la linea di livello 1 è la parabola della figura 1.28

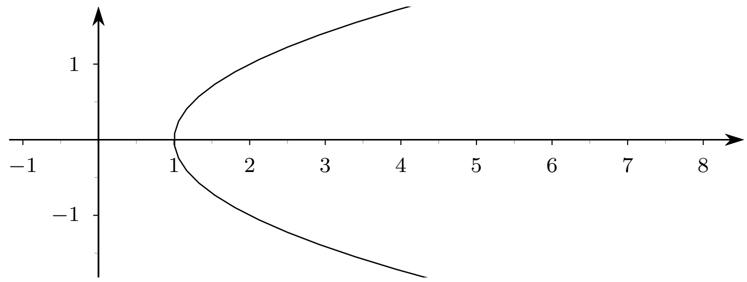


Figura 1.28. *Linea di livello* 1 *per la funzione* $f(x,y) = x - y^2$

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:

$$\left\{\begin{array}{l} z = f(x,y) \\ y = k \end{array}\right., \Rightarrow z = f(x,k) \qquad \left\{\begin{array}{l} z = f(x,y) \\ x = k \end{array}\right., \Rightarrow z = f(k,y). \tag{1.14}$$

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 1.22. Sia data la funzione $f(x,y) = x^3 - 4xy^2$, il cui grafico è rappresentato nella figura 1.29 (anche se in questo contesto il grafico è poco interessante).

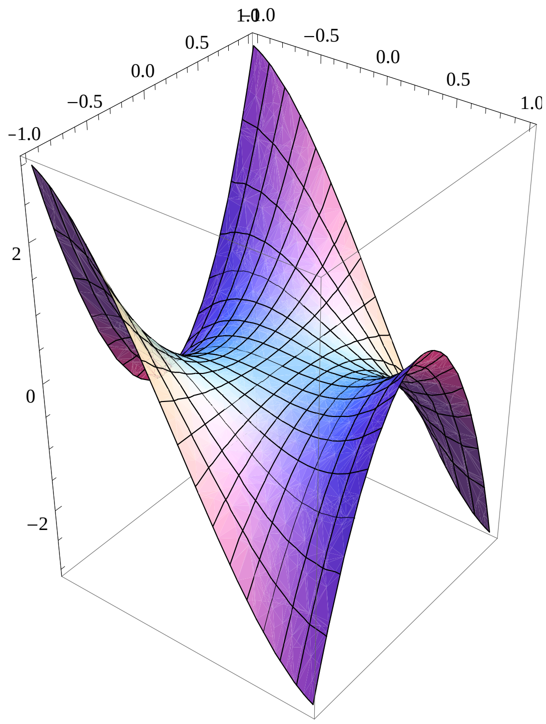


Figura 1.29. *Grafico della funzione* $f(x,y) = x^3 - 4xy^2$

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 1.30.

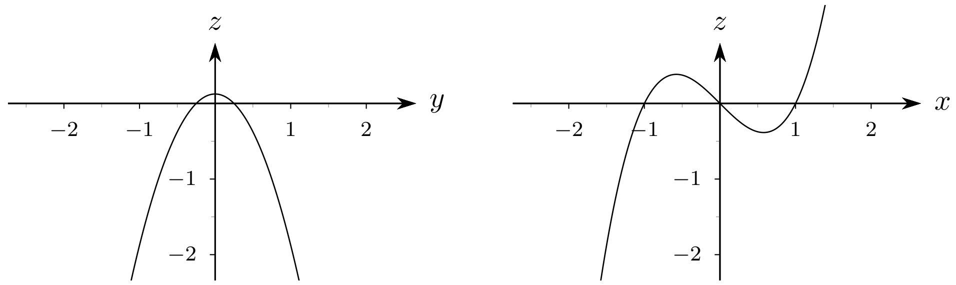


Figura 1.30. *Intersezioni della superficie* $z = x^3 - 4xy^2$ *con i piani* $x = 1/2$ *e* $y = 1/2$

La figura 1.31 mostra i piani sezionanti e le due curve direttamente sulla superficie.

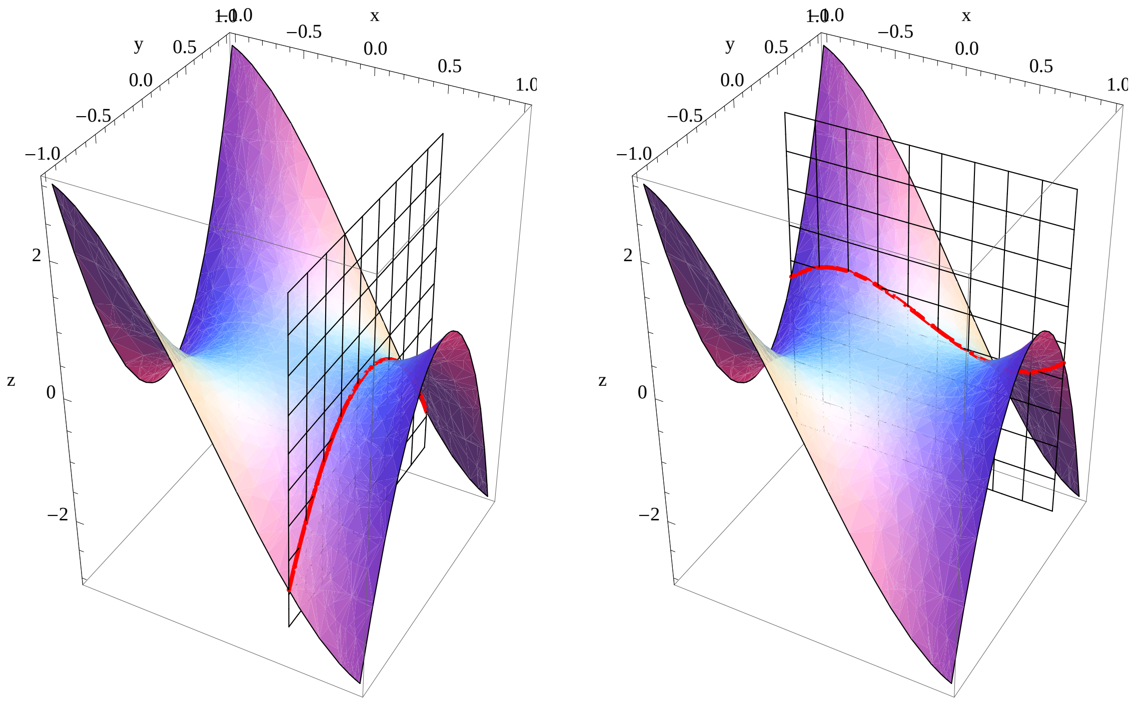


Figura 1.31. *Le intersezioni della figura 1.30, tracciate sulla superficie*

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.

2. Calcolo differenziale

Siamo ora interessati ad estendere il concetto di derivata ed in particolare quello di retta tangente e quindi di funzione affine approssimante (vedi le considerazioni a pagina 25), nonché i metodi per la ricerca di massimi e minimi al caso di funzioni di più variabili. Purtroppo le cose sono un po' più complesse rispetto al caso di una variabile e bisogna prestare la massima attenzione. Cominceremo trattando il caso delle funzioni di due variabili, più semplice in quanto, come più volte ricordato, è ancora possibile una rappresentazione grafica cartesiana.

2.1. Derivate parziali per funzioni di due variabili

Definizione 2.1 (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*

$$\lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}, \tag{2.1}$$

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*

$$f'_x(x_0, y_0) \quad \textit{oppure} \quad \frac{\partial f}{\partial x}(x_0, y_0). \tag{2.2}$$

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

$$\lim_{y \to y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}, \tag{2.3}$$

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*

$$f'_y(x_0, y_0) \quad \textit{oppure} \quad \frac{\partial f}{\partial y}(x_0, y_0). \tag{2.4}$$

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 2.1. 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 2.2. 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 2.3. Da $f(x,y)=\mathrm{e}^{x^2+y^2}$, si ottiene $f'_x(x,y)=2x\mathrm{e}^{x^2+y^2}$, $f'_y(x,y)=2y\mathrm{e}^{x^2+y^2}$.

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*[1], le ultime due si chiamano *derivate parziali seconde miste*.

Esempio 2.4. 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 2.2 (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 risulta dalle definizioni date, le derivate parziali sono le derivate della funzione ottenuta intersecando la superficie grafico di una funzione di due variabili con un piano verticale, rispettivamente parallelo all'asse x o all'asse y: come già osservato queste intersezioni, nei casi che ci interessano sono delle curve, grafico di funzioni di una sola variabile e quindi queste derivate parziali non sono altro che i coefficienti angolari delle rette tangenti a queste curve nel punto scelto. Le figure 2.1 e 2.2f mostrano la superficie grafico di una funzione di due variabili[2], le due curve sezione con due piani verticali paralleli all'asse x e all'asse y rispettivamente e, successivamente, le due rette tangenti a queste curve; nel grafico comprendente le rette tangenti abbiamo rappresentato solo una parte della superficie grafico, per maggiore chiarezza.

[1] Spesso l'appellativo "pure" si tralascia

[2] Per chi è interessato si tratta della funzione

$$f(x,y)=\left(y+\frac{1}{2}\right)\mathrm{e}^{-x^2-y^2+x},$$

mentre il punto scelto è il punto $(0.7, 0.3)$.

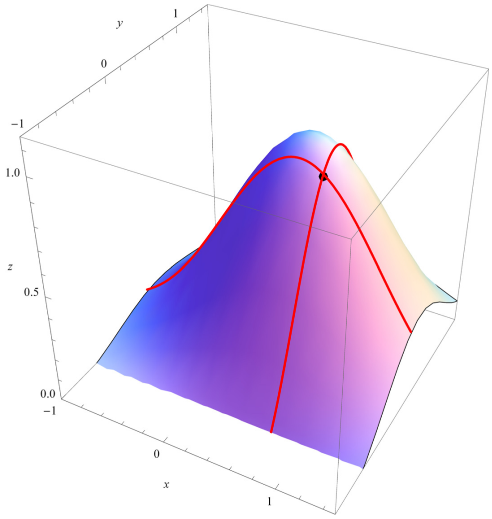


Figura 2.1. *Grafico di una funzione di due variabili e due curve sezione con piani verticali*

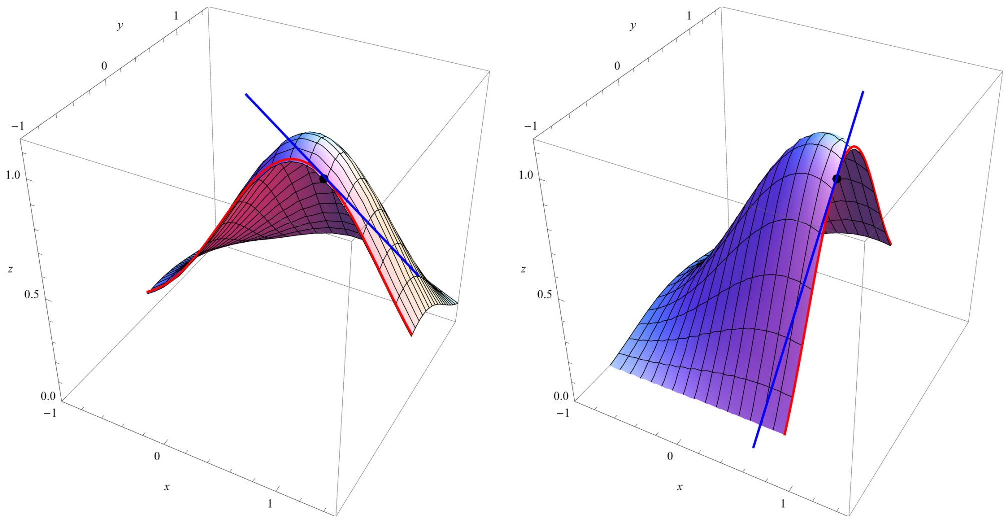


Figura 2.2. *Significato delle derivate parziali*

Nel caso delle funzioni di una variabile la derivata in un punto consente la determinazione della retta tangente al grafico della funzione e quindi la possibilità di approssimare, almeno in un intorno del punto scelto, la funzione con una funzione affine. Nel caso delle funzioni di due variabili l'esistenza delle derivate parziali consente, come già osservato, di determinare le rette tangenti alle curve intersezione tra la superficie grafico e il piano verticale parallelo al piano Oxz oppure Oyz. Purtroppo, in generale, questo non è sufficiente per garantire l'esistenza di un piano tangente[3] al grafico della superficie, cioè, anche se esistono le derivate parziali non è detto che sia possibile approssimare una funzione, in un intorno del punto scelto, con una funzione affine (che ha come grafico un piano). Naturalmente bisognerebbe dare un senso preciso al concetto di approssimazione e soprattutto di "grado di approssimazione", ma questo esula dal contesto di questo corso e ci accontenteremo di una nozione intuitiva (come del resto abbiamo fatto anche per le funzioni di una variabile).

Senza entrare troppo nei dettagli diremo che una funzione è *differenziabile* in un punto se è possibile approssimarla opportunamente in un intorno del punto con una funzione affine, cioè se

[3] In realtà le cose sono ancora più complicate: se una funzione ammette le derivate parziali in un punto non è nemmeno detto che sia continua in quel punto!

in quel punto ammette piano tangente. Una condizione sufficiente, ma non necessaria, perché questo succeda è che le derivate parziali prime siano continue[4]. Questo succede in tutti i casi di nostro interesse. In questi casi l'equazione del piano tangente alla superficie grafico della funzione nel punto (x_0, y_0, z_0), con $z_0 = f(x_0, y_0)$ è:

$$z = f(x_0, y_0) + f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0)\,. \tag{2.5}$$

Esempio 2.5. 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 in $(1, -1, 0)$ è

$$z = 0 + 1(x - 1) - 2(y + 1) \;\Rightarrow\; z = x - 2y - 3\,. \tag{2.6}$$

Esempio 2.6. 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 2.3, dove sono rappresentate anche le due curve sezione.

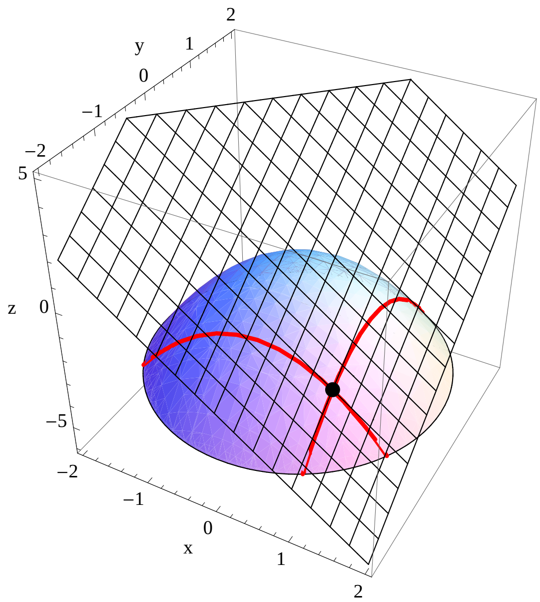


Figura 2.3. *Superficie* $z = -x^2 - y^2$ *e piano tangente in* $(1, -1)$

Per riunire in un certo senso le derivate parziali di una funzione di due variabili in un unico concetto si dà la seguente definizione.

Definizione 2.3 (Gradiente)**.** *Sia $f(x, y)$ una funzione di due variabili e sia $P_0 = (x_0, y_0)$ un punto del suo dominio dove esistono entrambe le derivate parziali. Si chiama* gradiente *della funzione f nel punto P_0 il vettore*

$$\nabla f(x_0, y_0) = \big(f'_x(x_0, y_0)\,,\; f'_y(x_0, y_0)\big)\,.$$

La scrittura ∇f si legge "nabla f" o anche "del f".

Questo concetto ci servirà anche tra poco per parlare di derivata direzionale.

[4]Si noti che, mentre l'esistenza delle derivate parziali coinvolge solo i valori della funzione sulla sezione con un piano verticale per il punto scelto, cioè su un intorno *unidimensionale* del punto, la continuità delle stesse coinvolge invece tutto un intorno *bidimensionale* del punto.

2.2. Derivate direzionali per funzioni di due variabili

Data una funzione di due variabili e un punto $P_0 = (x_0, y_0)$ del suo dominio abbiamo introdotto le derivate parziali considerando le sezioni verticali del suo grafico con piani paralleli all’asse x e all’asse y. Possiamo generalizzare questo procedimento considerando le sezioni con piani verticali passanti per il punto, ma non paralleli a nessuno dei due assi: otterremo ancora generalmente una curva tracciata sulla superficie grafico della funzione, e potremo ancora cercarne la tangente nel punto sul grafico corrispondente a P_0. Prima di dare la definizione formale vediamo una illustrazione riferita sempre alla stessa funzione e allo stesso punto considerato nelle figure 2.1 e 2.2. Anche in questo caso, nella illustrazione della tangente abbiamo visualizzato solo una parte della superficie grafico, per maggiore chiarezza.

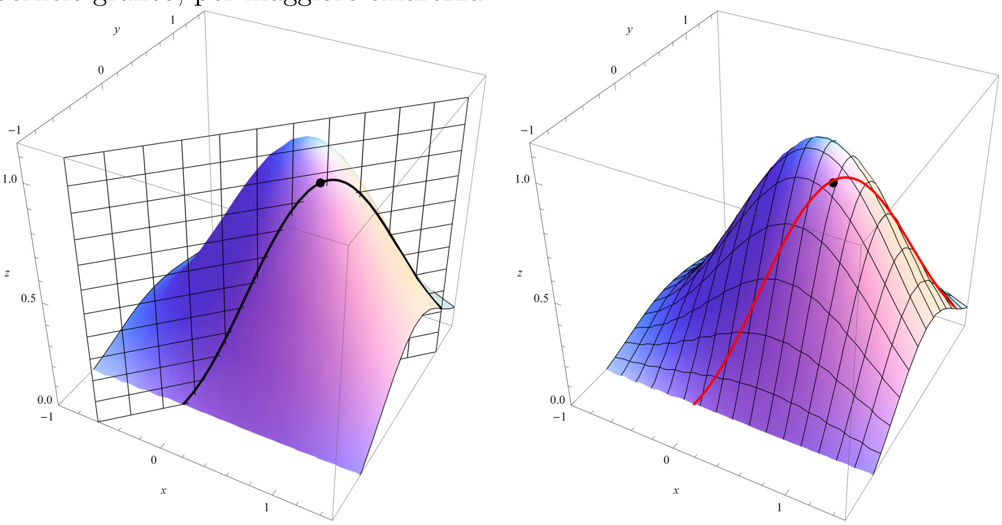


Figura 2.4. *Grafico di una funzione di due variabili e curva sezione con un piano verticale non parallelo agli assi x e y*

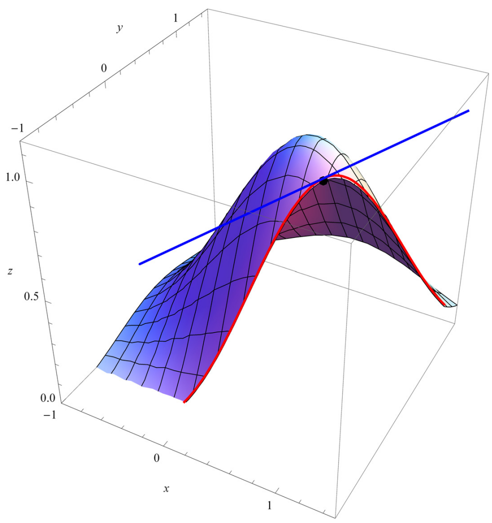


Figura 2.5. *Tangente alla curva sezione della figura 2.4*

Per introdurre in modo formale questo concetto consideriamo nel piano un punto $P_0 = (x_0, y_0)$ e un versore $\vec{v} = (v_1, v_2)$. È facile verificare (e comunque noi lo daremo per noto) che per ogni numero reale h il punto

$$P = (x_0 + hv_1, y_0 + hv_2)$$

appartiene alla retta passante per P_0 e parallela al vettore $\vec{v}$ e anzi, facendo variare h su tutto $\mathbb{R}$, si ottengono tutti i punti di questa retta. Si veda per un esempio la figura 2.6 dove

$$P_0 = (2, 1) \quad \text{e} \quad \vec{v} = \left(\frac{3}{5}, \frac{4}{5}\right).$$

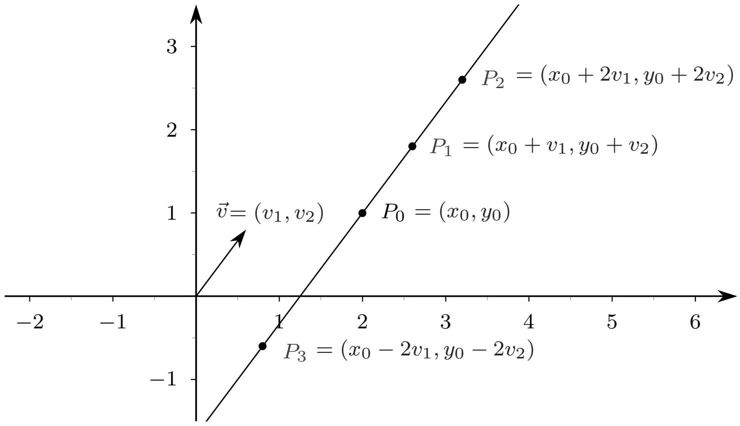


Figura 2.6. *Retta per un punto e parallela ad un versore*

Si noti che se $\vec{v} = (1, 0)$ la retta risulta parallela all'asse x, se $\vec{v} = (0, 1)$ la retta risulta parallela all'asse y.

Data ora una funzione $f(x, y)$, un punto $P_0 = (x_0, y_0)$ interno al suo dominio e un versore $\vec{v} = (v_1, v_2)$, possiamo considerare la funzione, della variabile h,

$$g(h) = f(x_0 + hv_1, y_0 + hv_2).$$

Al variare di h i valori di questa funzione forniscono le quote dei punti del grafico di f, in corrispondenza ai punti della retta passante per P_0 e parallela al versore $\vec{v}$. Naturalmente h dovrà essere tale che il punto $(x_0 + hv_1, y_0 + hv_2)$ non esca dal dominio della funzione f. Poiché saremo interessati al limite per $h \to 0$ e poiché il punto P_0 è interno al dominio non ci sono problemi. Si dà allora la seguente definizione

Definizione 2.4 (Derivata direzionale)**.** *Si chiama* derivata direzionale *della funzione f nel punto P_0 e secondo la direzione orientata di $\vec{v}$ il*

$$\lim_{h \to 0} \frac{g(h) - g(0)}{h} = \lim_{h \to 0} \frac{f(x_0 + hv_1, y_0 + hv_2) - f(x_0, y_0)}{h}, \tag{2.7}$$

se questo limite esiste finito. Questa derivata si indica con la scrittura

$$\frac{\partial f}{\partial \vec{v}}(x_0, y_0). \tag{2.8}$$

Se $\vec{v} = (1, 0)$ la derivata direzionale non è altro che la derivata parziale rispetto alla x, mentre se $\vec{v} = (0, 1)$ la derivata direzionale non è altro che la derivata parziale rispetto alla y. Il concetto di derivata direzionale dunque è una generalizzazione di quello di derivata parziale. L'esistenza della derivata direzionale secondo una data direzione garantisce la possibilità di tracciare la tangente alla curva intersezione del grafico di f con il piano verticale contenente la retta per P_0 e parallela a $\vec{v}$. Come abbiamo già avuto modo di segnalare, la sola esistenza delle derivate parziali non garantisce la possibilità di tracciare il piano tangente alla superficie grafico di una funzione.

Purtroppo nemmeno l'esistenza delle derivate direzionali secondo tutte le direzioni[5] garantisce la possibilità di tracciare questo piano tangente: anche se la cosa può sembrare sconcertante, sezionando il grafico di una funzione di due variabili con una piano verticale qualunque passante per un punto P_0 si può sempre ottenere una curva dotata di tangente, senza che vi sia un piano tangente alla superficie nel suo complesso. Non saremo comunque interessati a situazioni così patologiche. Segnaliamo invece che si dimostra che se una funzione è differenziabile in un punto essa ammette derivate direzionali in quel punto secondo ogni direzione orientata, quindi in particolare anche le derivate parziali.

Esempio 2.7. Sia $f(x,y) = x^2 - y^2$ una funzione di due variabili, $P_0 = (2,1)$ e $\vec{v} = (4/5\,,\,3/5)$. Si ha

$$g(h) = f(x_0 + hv_1, y_0 + hv_2) = f\left(2 + h\frac{4}{5}\,,\,1 + h\frac{4}{5}\right) = 3 + \frac{7}{25}h^2 + 2h, \quad g(0) = f(2,1) = 3.$$

Quindi

$$\lim_{h\to 0}\frac{g(h)-g(0)}{h} = \lim_{h\to 0}\left(\frac{7}{25}h + 2\right) = 2 = \frac{\partial f}{\partial \vec{v}}(2,1).$$

Per le derivate direzionali vale il seguente importante teorema, che facilita enormemente il calcolo delle derivate direzionali.

Teorema 2.5. *Sia $f(x,y)$ una funzione differenziabile*[6] *in un punto P_0 interno al suo dominio e sia $\vec{v}$ un versore. Allora la derivata direzionale di f nel punto P_0, secondo la direzione orientata di $\vec{v}$ è il prodotto scalare tra il gradiente della funzione nel punto e il versore $\vec{v}$:*

$$\frac{\partial f}{\partial \vec{v}}(x_0,y_0) = \nabla f(x_0,y_0)\cdot\vec{v} = f'_x(x_0,y_0)v_1 + f'_y(x_0,y_0)v_2. \tag{2.9}$$

Esempio 2.8. Riprendiamo in esame l'esempio 2.7. Essendo

$$f'_x = 2x \quad \text{e} \quad f'_y = -2y, \quad \text{da cui} \quad f'_x(2,1) = 4 \quad \text{e} \quad f'_y(2,1) = -2,$$

si ha

$$\frac{\partial f}{\partial \vec{v}}(2,1) = (4,-2)\cdot\left(\frac{4}{5}\,,\,\frac{3}{5}\right) = \frac{16}{5} - \frac{6}{5} = 2,$$

esattamente come avevamo ottenuto per calcolo diretto in base alla definizione.

2.3. Derivazione delle funzioni composte

Trattando le funzioni di una variabile abbiamo già considerato la regola di derivazione delle funzioni composte, regola particolarmente importante nelle applicazioni, in quanto la quasi totalità delle funzioni di uso comune è costruita componendo opportunamente le funzioni "elementari". Richiamiamo, per completezza, quella regola.

Teorema 2.6. *Siano f e g due funzioni derivabili di una variabile, tali che sia possibile costruire la funzione composta $h = g\circ f$, ovvero $h(x) = g(f(x))$. Si ha allora*

$$h'(x) = g'(f(x))f'(x). \tag{2.10}$$

[5] Come segnalato già trattando le derivate parziali, purtroppo le cose sono ancora più complicate: l'esistenza delle derivate direzionali secondo tutte le direzioni non garantisce nemmeno la continuità della funzione nel punto in esame!

[6] Ricordiamo che una funzione è differenziabile se il suo grafico ammette piano tangente in corrispondenza al punto P_0 e che una condizione sufficiente per questo è che le derivate parziali siano continue.

Siamo interessati alla estensione di questa regola al caso di funzioni composte di due variabili. L'unico caso che ci interessa è del tipo $h = g \circ f$ con $f\colon \mathbb{R}^2 \to \mathbb{R}$ e $g\colon \mathbb{R} \to \mathbb{R}$, ovvero $h(x_1, x_2) = g(f(x_1, x_2))$. In questo caso la formula (2.10) è immediatamente generalizzabile: basterà considerare una volta x_2 come costante (per ottenere h'_{x_1}) e una volta x_1 come costante (per ottenere h'_{x_2}). Si trova

$$h'_{x_1}(x_1, x_2) = g'(f(x_1, x_2))f'_{x_1}(x_1, x_2)\,, \quad h'_{x_2}(x_1, x_2) = g'(f(x_1, x_2))f'_{x_2}(x_1, x_2). \tag{2.11}$$

Anche se la formula (2.11) è una immediata estensione della (2.10), per la sua validità basta che la funzione g sia derivabile (g è funzione di una sola variabile), mentre occorre cha la funzione f sia differenziabile (f è funzione di due variabili e, come più volte osservato, l'esistenza delle derivate parziali non è sufficiente per la differenziabilità).

Esempio 2.9. Sia $f(x_1, x_2) = x_1^3 + x_2^2$ e $g(x) = \sin x$. Allora $h(x_1, x_2) = \sin(x_1^3 + x_2^2)$. Si ha, intanto,

$$f'_{x_1}(x_1, x_2) = 3x_1^2, \quad f'_{x_2}(x_1, x_2) = 2x_2, \quad g'(x) = \cos x.$$

Dunque

$$h'_{x_1}(x_1, x_2) = g'(f(x_1, x_2))f'_{x_1}(x_1, x_2) = \left(\cos\left(x_1^3 + x_2^2\right)\right) 3x_1^2,$$
$$h'_{x_2}(x_1, x_2) = g'(f(x_1, x_2))f'_{x_2}(x_1, x_2) = \left(\cos\left(x_1^3 + x_2^2\right)\right) 2x_2.$$

2.4. Funzioni implicite

2.4.1. Due esempi elementari

Consideriamo un'equazione lineare in due incognite

$$ax + by + c = 0\,, \quad \text{con } a,\ b \text{ e } c \text{ numeri reali qualunque,} \tag{2.12}$$

che sappiamo rappresentare una retta del piano cartesiano se $a^2 + b^2 \neq 0$, cioè se a e b non sono contemporaneamente nulli.

Consideriamo poi la funzione

$$f\colon \mathbb{R}^2 \to \mathbb{R} \quad , \quad f(x, y) = ax + by + c\,. \tag{2.13}$$

È immediato che si ha, per ogni $(x, y) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x, y) = a \quad , \quad \frac{\partial f}{\partial y}(x, y) = b\,, \tag{2.14}$$

cioè che il gradiente di f in qualunque punto (x, y) del piano vale sempre (a, b). Dunque la condizione affinché la (2.12) rappresenti una retta si può anche scrivere

$$\nabla f(x, y) \neq (0, 0). \tag{2.15}$$

Come ben sappiamo, se $a \neq 0$ dalla (2.12) si può ricavare la x come funzione della y (ovvero *esplicitare* la x):

$$x = -\frac{b}{a}\,y - \frac{c}{a} = h(y)\,; \tag{2.16}$$

se invece $b \neq 0$ sempre dalla (2.12) si può ricavare la y come funzione della x (ovvero *esplicitare* la y):

$$y = -\frac{a}{b}\,x - \frac{c}{b} = g(x)\,. \tag{2.17}$$

Detto in altri termini: se il gradiente di f è diverso da zero (non occorre qui precisare in quale punto in quanto il gradiente è costante), la retta può essere vista o come il grafico di una funzione della sola variabile x o come il grafico di una funzione della sola variabile y: precisamente se la derivata parziale rispetto *a una delle due* variabili è diversa da zero, allora si può vedere la retta come grafico di una funzione dell'*altra* variabile.

Si noti che se sostituiamo la x ricavata nella (2.16) nell'espressione della funzione f, la stessa, che è funzione della sola y, diventa identicamente nulla:

$$f(h(y), y) = a\,h(y) + by + c = a\left(-\frac{b}{a}\,y - \frac{c}{a}\right) + by + c = 0. \tag{2.18}$$

Analogamente se sostituiamo la y ricavata nella (2.17) nell'espressione della funzione f, la stessa, che è ora funzione della sola x, diventa identicamente nulla:

$$f(x, g(x)) = ax + b\,g(x) + c = ax + b\left(-\frac{a}{b}\,x - \frac{c}{b}\right) + c = 0. \tag{2.19}$$

Le (2.18) e (2.19) esprimono il fatto che tutte le soluzioni della (2.16) o della (2.17) sono anche soluzioni della (2.12).

Si può riassumere quanto detto nei seguenti termini.

Data la funzione di $\mathbb{R}^2$ in $\mathbb{R}$, $f(x, y) = ax + by + c$, se, per esempio, la derivata

$$\frac{\partial f}{\partial y}(x, y)$$

è diversa da zero (cioè se $b \neq 0$), allora dall'equazione $f(x, y) = 0$ si può ricavare la y come funzione derivabile di x, $y = g(x)$, in modo che

$$f(x, g(x)) \tag{2.20}$$

sia identicamente nulla su tutto $\mathbb{R}$. La funzione g costruita con questa tecnica si dice *funzione implicitamente definita* dall'equazione (2.12). Discorso simmetrico se invece è diversa da zero l'altra derivata parziale.

Se passiamo a considerare un'equazione come la

$$x^2 + y^2 - 1 = 0 \tag{2.21}$$

e la funzione di due variabili $f(x, y) = x^2 + y^2 - 1$, le cose si complicano un po', e dovremo fare qualche adattamento. Supponiamo, tanto per fissare le idee, di voler *esplicitare* la y. Per questo calcoliamo, in analogia a quanto fatto nel caso della retta, la derivata di f rispetto a y, che ora non sarà più costante, ma dipenderà dal punto (x, y):

$$\frac{\partial f}{\partial y}(x, y) = 2y\,.$$

Questa derivata è diversa da zero in tutti i punti dove la y è diversa da zero. Tenendo conto di come è fatto l'insieme, diciamolo $\mathcal{L}$, delle soluzioni dell'equazione (2.21) (la circonferenza di centro l'origine e raggio 1), si tratterà dei punti diversi da $(-1, 0)$ e $(1, 0)$. Fissiamo allora un punto, per esempio $(0, 1)$, sull'insieme $\mathcal{L}$, dove la derivata rispetto a y di f è diversa da zero. Si vede facilmente che dall'equazione (2.21) si potrà ricavare *una sola* funzione g della variabile x, definita in $[-1, 1]$, ivi *continua* e addirittura *derivabile* in $]-1, 1[$, il cui grafico *giaccia interamente* sull'insieme $\mathcal{L}$, cioè tale che

$$f(x, g(x)) = 0\,, \quad \forall x \in [-1, 1],$$

con la *ulteriore condizione* che $g(0) = 1$, cioè che il grafico passi per il punto $(0, 1)$ prima scelto. Si tratta precisamente della funzione

$$g(x) = \sqrt{1 - x^2}.$$

L'ultima condizione indicata è particolarmente importante perché, in sua mancanza, ci sarebbero due funzioni che soddisfano agli altri requisiti, precisamente g e $-g$.

Si noti che, a differenza del caso della retta, questa volta il grafico della funzione g non esaurisce tutto l'insieme delle soluzioni dell'equazione (2.21).

Si noti ancora, per concludere la trattazione di questi esempi elementari, che se avessimo scelto il punto $(1, 0)$ nel quale la derivata di f rispetto a y si annulla, *non* sarebbe stato possibile ricavare la y come funzione della x dall'equazione (2.21): basta tenere conto di come è fatto l'insieme $\mathcal{L}$. Si sarebbe però potuto procedere a ricavare la x come funzione della y, con considerazioni analoghe a quelle appena fatte.

2.4.2. Il teorema di Dini per le curve del piano

Passando ora ad una generica funzione $f(x, y)$, volendo considerare l'equazione $f(x, y) = 0$, dovremo fare ulteriori adattamenti, tra cui segnaliamo fin da subito i due seguenti, riservandoci di tornare di nuovo sull'argomento:

– non sarà in generale possibile trovare un'espressione analitica in termini di funzioni elementari per una funzione eventualmente definita implicitamente da un'equazione del tipo $f(x, y) = 0$, nemmeno se la funzione f stessa è espressa in termini di funzioni elementari;
– non sarà in generale nemmeno possibile precisare il dominio di una funzione eventualmente definita implicitamente da un'equazione del tipo $f(x, y) = 0$: saremo costretti ad accontentarci di un intorno (di cui non si potrà a priori stabilire l'ampiezza) di un opportuno punto.

Nonostante queste difficoltà, che sembrerebbero rendere addirittura inutile la trattazione che stiamo per fare, vedremo che sarà invece, in genere, possibile calcolare le derivate (anche se noi tratteremo in dettaglio la derivata prima), e quindi scrivere il polinomio di Taylor, di opportuno punto iniziale, della funzione eventualmente definita implicitamente da un'equazione del tipo $f(x, y) = 0$: questo basta in molte circostanze per risolvere i problemi che si presentano.

Veniamo ora ad enunciare il teorema fondamentale, dovuto a Ulisse Dini, enunciato che daremo in una forma dettagliata per evidenziare i fatti importanti, anche se si potrebbe darne un enunciato più compatto. Nell'enunciato di questo teorema si ritroverà la generalizzazione di quanto sopra detto nei due casi particolari trattati.

Teorema 2.7 (della funzione implicita, o di Dini, per le curve del piano). *Sia A un aperto non vuoto di $\mathbb{R}^2$, e f una funzione di A in $\mathbb{R}$, continua con le sue derivate prime. Sia poi (x_0, y_0) un punto di A ove $f(x_0, y_0) = 0$, cioè un punto soluzione dell'equazione $f(x, y) = 0$. Se*

$$\nabla f(x_0, y_0) \neq (0, 0)\,, \tag{2.22}$$

con, per esempio,

$$\frac{\partial f}{\partial y}(x_0, y_0) \neq 0\,, \tag{2.23}$$

allora esistono un intorno I di x_0 e una funzione $g\colon I \to \mathbb{R}$ tali che:

1. $\big(x, g(x)\big) \in A\,, \forall x \in I$;
2. $g(x_0) = y_0$;
3. $f\big(x, g(x)\big) = 0\,, \forall x \in I$;

4. *g è derivabile con derivata continua in I e si ha*

$$g'(x) = -\frac{f'_x(x, g(x))}{f'_y(x, g(x))}, \tag{2.24}$$

e dunque in particolare la derivata in x_0 è data da

$$g'(x_0) = -\frac{f'_x(x_0, y_0)}{f'_y(x_0, y_0)}; \tag{2.25}$$

5. *in un intorno di (x_0, y_0) l'insieme delle soluzioni dell'equazione $f(x, y) = 0$ coincide con il grafico della funzione g.*

Se, invece,

$$\frac{\partial f}{\partial x}(x_0, y_0) \neq 0,$$

allora esistono un intorno I di y_0 e una funzione $h\colon I \to \mathbb{R}$ tali che:

1. $\big(h(y), y\big) \in A, \forall y \in I$*;*
2. $h(y_0) = x_0$*;*
3. $f\big(h(y), y\big) = 0, \forall y \in I$*;*
4. *h è derivabile con derivata continua in I e si ha*

$$h'(y) = -\frac{f'_y(h(y), y)}{f'_x(h(y), y)} \tag{2.26}$$

e dunque in particolare la derivata in y_0 è data da

$$h'(y_0) = -\frac{f'_y(x_0, y_0)}{f'_x(x_0, y_0)}; \tag{2.27}$$

5. *in un intorno di (x_0, y_0) l'insieme delle soluzioni dell'equazione $f(x, y) = 0$ coincide con il grafico della funzione h.*

Esempio 2.10. Consideriamo l'equazione

$$f(x, y) = y^7 + 3y - 2x\mathrm{e}^{3x} = 0$$

e il punto $(0, 0)$ appartenente all'insieme delle sue soluzioni.

Avendosi

$$\frac{\partial f}{\partial x}(x, y) = -2\mathrm{e}^{3x}(1 + 3x), \quad \frac{\partial f}{\partial y}(x, y) = 7y^6 + 3,$$

possiamo notare che la derivata parziale rispetto a y è non nulla su tutto $\mathbb{R}^2$, in particolare in $(0, 0)$. Sarà pertanto possibile esplicitare y in funzione di x nell'intorno di $(0, 0)$. Sia $y = g(x)$ la funzione implicita. Si ha

$$g(0) = 0, \quad g'(0) = -\frac{f'_x(0, 0)}{f'_y(0, 0)} = \frac{2}{3}.$$

Dunque la retta tangente al grafico di g è:

$$y = \frac{2}{3}x.$$

Questo ci consente di approssimare il grafico della funzione implicita (che non riusciamo a scrivere esplicitamente!) con la retta tangente: purtroppo non è possibile stabilire a priori quanto ampio possa essere l'intorno di 0 in cui questa approssimazione è "buona". Tuttavia in molte questioni questa informazione è già sufficiente per le applicazioni. Ulteriori considerazioni, sempre basate sul teorema di Dini, e tecniche numeriche che però esulano dal contesto di questo corso permetterebbero addirittura di avere un'informazione abbastanza dettagliata del grafico della funzione implicita. Riportiamo, solo a titolo di curiosità, questo grafico.

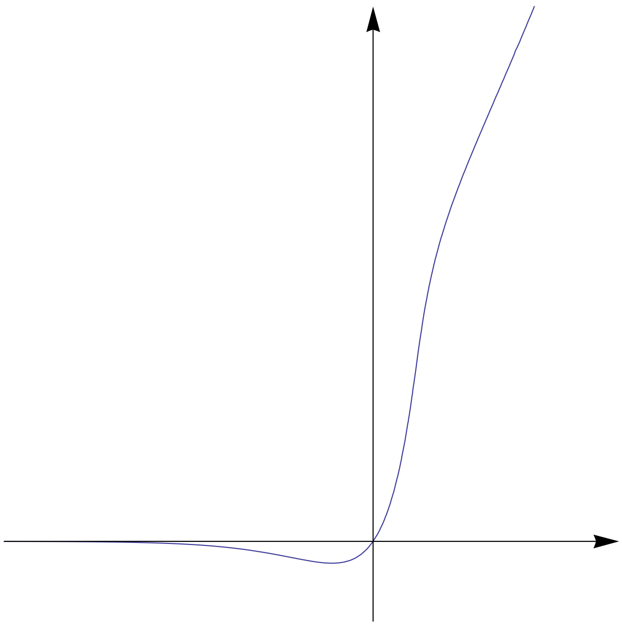


Figura 2.7. *Esempio di grafico di una funzione implicita*

Osservazione 2.8. Come abbiamo già sottolineato, e come prova l'esempio 2.10, in generale non è possibile scrivere esplicitamente una formula per la funzione $g(x)$ (nel primo caso) o $h(y)$ (nel secondo caso); per questo motivo le formule (2.24) e (2.26), che esprimono la derivata della funzione implicita in un punto generico, sono di scarsa utilità pratica, mentre le formule (2.25) e (2.27) permettono di calcolare effettivamente le derivate nel punto x_0 (o y_0 a seconda dei casi). Questo perché noi sappiamo quanto vale $g(x_0)$ (o $h(y_0)$). Tuttavia è importante la seguente considerazione, che ci limitiamo a fare solo per il primo caso (in cui si ricava la y come funzione implicita della x). Riesaminiamo la proprietà 3. della funzione implicita $g(x)$, che afferma che

3. $f(x, g(x)) = 0, \ \forall x \in I$.

Ora $f(x, y)$ è una funzione di due variabili e per ottenere $f(x, g(x))$ abbiamo sostituito alla variabile y la funzione $g(x)$ di cui il teorema di Dini afferma l'esistenza: questo equivale a dire che $f(x, g(x))$, che è funzione della sola variabile x, è stata ottenuta "componendo" $f(x, y)$ con $g(x)$. Si tratta di una composizione diversa da quella che noi abbiamo considerato nel paragrafo 1.5. Tuttavia anche in questo caso si può dimostrare una formula di derivazione sostanzialmente simile a quella che abbiamo proposto nel paragrafo 2.3. Precisamente vale il seguente teorema.

Teorema 2.9. *Sia $f(x, y)$ una funzione di due variabili dotata di derivate continue, e $g(x)$ una funzione di una variabile dotata di derivata continua. Allora la funzione composta $F(x) = f(x, g(x))$, della sola variabile x, è derivabile e si ha*

$$F'(x) = f'_x(x, g(x)) + f'_y(x, g(x)))\, g'(x). \tag{2.28}$$

Si noterà, anche in questa formula (2.28), la generalizzazione quasi immediata della (2.10), valida per funzioni di una variabile.

Tenendo conto di questo teorema si può ricavare la formula (2.24) che esprime la derivata della funzione implicita. Si ha infatti

$$F(x) = f(x, g(x)) = 0 \, \forall x \in I,$$

allora anche $F'(x) = 0 \, \forall x \in I$, ma

$$F'(x) = f'_x(x, g(x)) + f'_y(x, g(x)) \, g'(x) = 0, \tag{2.29}$$

da cui, appunto, la (2.24).

Esempio 2.11. Riprendiamo in esame la funzione

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

già considerata nell'equazione (2.21). Già sappiamo che, se fissiamo il punto $(0, 1)$ che è soluzione dell'equazione $f(x, y) = 0$, possiamo ricavare (questa volta in maniera esplicita) la funzione

$$g(x) = \sqrt{1 - x^2}$$

tale che $g(0) = 1$ e che

$$F(x) = f(x, g(x)) = f(x, \sqrt{1 - x^2}) = x^2 + \left(\sqrt{1 - x^2}\right)^2 - 1 = 0$$

Poiché $f'_x(x, y) = 2x$ e $f'_y(x, y) = 2y$, abbiamo, dalla (2.29)

$$F'(x) = f'_x(x, g(x)) + f'_y(x, g(x)) \, g'(x) = 2x + \left(2\sqrt{1 - x^2}\right) g'(x) = 0,$$

da cui

$$g'(x) = \frac{-x}{\sqrt{1 - x^2}} \, ,$$

come previsto. Si noti che abbiamo ricavato $g'(x)$ non direttamente dall'espressione di $g(x)$, ma usando il teorema di Dini.

Se riconsideriamo invece l'esempio 2.10, nel quale non è possibile ricavare un'espressione esplicita per $g(x)$, potremmo ancora scrivere, tenendo conto delle derivate di f che abbiamo già calcolato,

$$g'(x) = -\frac{-2e^{3x}(1 + 3x)}{7\big(g(x)\big)^6 + 3} \, ,$$

ma da qui non traiamo alcuna informazione sulla derivata di g in un generico punto x. Come visto però possiamo invece calcolare

$$g'(0) = \frac{2}{3} \, .$$

2.5. Funzioni di tre o più variabili

L'estensione dei concetti di derivata parziale e direzionale alle funzioni reali di più di due variabili reali non comporta particolari difficoltà, se si esclude il fatto che non è più possibile dare interpretazioni grafiche come quelle proposte per il caso di due variabili. Diamo di seguito le definizioni necessarie, che invitiamo a confrontare con quelle date nel caso di due variabili.

Definizione 2.10. *Data una funzione $y = f(x_1, x_2, \dots, x_n)$ e un punto $P_0 = (x_{1_0}, x_{2_0}, \dots, x_{n_0})$ interno al suo dominio, possiamo considerare la funzione, della variabile x_1, ottenuta fissando tutte le altre variabili al valore $x_{2_0}, \dots, x_{n_0}$ e lasciando variare solo la x_1: $g(x_1) = f(x_1, x_{2_0}, \dots, x_{n_0})$. Possiamo ora considerare il*

$$\lim_{x_1 \to x_{1_0}} \frac{f(x_1, x_{2_0}, \dots, x_{n_0}) - f(x_{1_0}, x_{2_0}, \dots, x_{n_0})}{x_1 - x_{1_0}}, \tag{2.30}$$

ovvero il limite del rapporto incrementale della funzione $g(x_1)$. Se questo limite esiste ed è finito, esso si chiama derivata parziale prima*, rispetto a x_1, della funzione f nel punto P_0 e si indica con*

$$f'_{x_1}(x_{1_0}, x_{2_0}, \dots, x_{n_0}) \quad \textit{oppure} \quad \frac{\partial f}{\partial x_1}(x_{1_0}, x_{2_0}, \dots, x_{n_0}). \tag{2.31}$$

In maniera perfettamente analoga ci si può comportare con le altre variabili, ottenendo le derivate parziali prime rispetto a ciascuna della variabili.

In pratica il calcolo delle derivate parziali in un punto generico $(x_1, x_2, \dots, x_n)$ interno al dominio di una funzione di n variabili si fa pensando la funzione come funzione di una sola variabile e trattando le altre come parametri costanti.

Esempio 2.12. Sia[7] $f(x, y, z) = x^2 + \sin y + x \cos z$. Allora

$$\begin{aligned} f'_x &= 2x + \cos z; \\ f'_y &= \cos y; \\ f'_z &= -x \sin z. \end{aligned}$$

Esempio 2.13. Sia $f(x_1, x_2, x_3, x_4) = x_1\sqrt{x_2} + \cos(x_2 + 3x_3) - x_1 \ln(1 + x_4^3)$. Allora

$$\begin{aligned} f'_{x_1} &= \sqrt{x_2} + \ln(1 + x_4^3); \\ f'_{x_2} &= x_1 \frac{1}{2\sqrt{x_2}} - \sin(x_2 + 3x_3); \\ f'_{x_3} &= -3 \sin(x_2 + 3x_3); \\ f'_{x_4} &= -x_1 \frac{1}{1 + x_4^3}\, 3x_4^2. \end{aligned}$$

Come mostrano gli esempi proposti, le derivate parziali, calcolate in un generico punto, sono esse stesse funzioni di più variabili, e quindi posso riapplicare ad esse ancora il processo di derivazione, ottenendo le derivate seconde: per ognuna delle n derivate prime si otterranno n derivate seconde, per un totale di n^2 derivate seconde.

Esempio 2.14. Riprendendo in esame la funzione dell'esempio 2.12, otteniamo quanto segue.

$$\begin{array}{lll} f''_{xx} = 2; & f''_{xy} = 0; & f''_{xz} = -\sin z; \\ f''_{yx} = 0; & f''_{yy} = -\sin y; & f''_{yz} = 0; \\ f''_{zx} = -\sin z & f''_{zy} = 0; & f''_{zz} = -x \cos z. \end{array}$$

Si usano ancora i termini di *derivate seconde pure* e *derivate seconde miste*, come nel caso delle funzioni di due variabili e vale ancora il teorema di Schwartz: se le derivate seconde miste sono continue, esse sono uguali. Se sistemiamo le derivate seconde in una matrice quadrata di

[7] Nel caso di tre variabili di solito si preferisce scrivere (x, y, z) al posto di (x_1, x_2, x_3).

ordine n, il teorema di Schwartz ha come conseguenza che questa matrice (nel caso dei derivate continue) è simmetrica.

Per la funzione degli esempi 2.12 e 2.14 questa matrice è la seguente.

$$\begin{pmatrix} 2 & 0 & -\sin z \\ 0 & -\sin y & 0 \\ -\sin z & 0 & -x\cos z \end{pmatrix}.$$

Come per le funzioni di due variabili si dà la definizione di gradiente.

Definizione 2.11. *Sia $f(x_1, x_2, \ldots, x_n)$ una funzione di n variabili e sia $P_0 = (x_{1_0}, x_{2_0}, \ldots, x_{n_0})$ un punto del suo dominio dove esistono tutte le derivate parziali. Si chiama* gradiente *della funzione f nel punto P_0 il vettore*

$$\nabla f(P_0) = \left(f'_{x_1}(P_0), f'_{x_2}(P_0), \ldots, f'_{x_n}(P_0), \right). \tag{2.32}$$

La scrittura ∇f si legge "nabla f" o anche "del f".

Sempre in analogia con le funzioni di due variabili si dà la definizione di derivata direzionale, naturalmente senza che sia più la possibilità di una visualizzazione grafica del significato.

Definizione 2.12. *Data una funzione $f(x_1, x_2, \ldots, x_n)$, un punto $P_0 =$ interno al suo dominio e un versore $\vec{v} = (v_1, v_2, \ldots, v_n)$, possiamo considerare la funzione, della sola variabile h*

$$g(h) = f(x_{1_0} + hv_1, x_{2_0} + hv_2, \ldots, x_{n_0} + hv_n).$$

Si chiama derivata direzionale *della funzione f, nel punto P_0 e secondo la direzione orientata di $\vec{v}$, il*

$$\lim_{h\to 0} \frac{g(h) - g(0)}{h} = \lim_{h\to 0} \frac{f(x_{1_0} + hv_1, x_{2_0} + hv_2, \ldots, x_{n_0} + hv_n) - f(x_{1_0}, x_{2_0}, \ldots, x_{n_0})}{h}, \tag{2.33}$$

se questo limite esiste finito. Questa derivata si indica con la scrittura

$$\frac{\partial f}{\partial \vec{v}}(P_0). \tag{2.34}$$

Se $\vec{v} = (1, 0, \ldots, 0)$ la derivata direzionale coincide con la derivata parziale rispetto alla prima variabile, e così via per le altre derivate parziali.

Anche nel caso di funzioni di più variabili vale un teorema analogo al teorema 2.5.

Teorema 2.13. *Sia $f(x_1, x_2, \ldots, x_n)$ una funzione differenziabile in un punto P_0 interno al suo dominio e sia $\vec{v}$ un versore. Allora la derivata direzionale di f nel punto P_0, secondo la direzione orientata di $\vec{v}$ è il prodotto scalare tra il gradiente della funzione nel punto e il versore $\vec{v}$:*

$$\frac{\partial f}{\partial \vec{v}}(P_0) = \nabla f(P_0) \cdot \vec{v} = f'_{x_1}(P_0)v_1 + f'_{x_2}(P_0)v_2 + \cdots + f'_{x_n}(P_0)v_n. \tag{2.35}$$

Come per le funzioni di due variabili siamo interessati all'approssimazione di una generica funzione con una funzione affine (che questa volta sarà un *iperpiano*). Ancora una volta né l'esistenza delle derivate parziali né quella delle derivate direzionali secondo ogni direzione orientata garantisce la possibilità di questa approssimazione. Una funzione che sia approssimabile, nell'intorno di un punto P_0, con una funzione affine (che diremo ancora "tangente", pur senza che questo termine abbia un preciso significato geometrico) sarà detta *differenziabile* e una condizione sufficiente per la differenziabilità è il fatto che tutte le derivate parziali prime siano continue. Questo succede in tutti i casi di nostro interesse. Per una funzione differenziabile l'iperpiano tangente ha la seguente equazione

$$y = f(P_0) + f'_{x_1}(P_0)(x_1 - x_{1_0}) + f'_{x_2}(P_0)(x_2 - x_{2_0}) + \cdots + f'_{x_n}(P_0)(x_n - x_{n_0}), \tag{2.36}$$

formula che costituisce la generalizzazione della (2.5).

Esempio 2.15. Se $f(x_1, x_2, x_3) = x_1^2 + x_1x_2 + 2x_3$, e $P_0 = (1, 1, 1)$, si ha $f(1, 1, 1) = 4$ e

$$\begin{gathered} f'_{x_1} = 2x_1 + x_2 \quad \Rightarrow \quad f'_{x_1}(1,1,1) = 3; \\ f'_{x_2} = x_1 \quad \Rightarrow \quad f'_{x_2}(1,1,1) = 1; \\ f'_{x_3} = 2 \quad \Rightarrow \quad f'_{x_3}(1,1,1) = 2. \end{gathered}$$

L'equazione dell'iperpiano tangente (funzione affine approssimante in un intorno di $(1, 1, 1)$) è

$$y = 4 + 3(x_1 - 1) + (x_2 - 1) + 2(x_3 - 1).$$

Anche per la regola di derivazione delle funzioni composte si possono fare considerazioni analoghe a quelle fatte nel caso di due variabili. Se $h = g \circ f$ con $: \mathbb{R}^n \to \mathbb{R}$ e $g\colon \mathbb{R} \to \mathbb{R}$, ovvero

$$h((x_1, x_2, \dots, x_n) = g\big(f(x_1, x_2, \dots, x_n)\big),$$

si ha

$$\begin{aligned} h'_{x_1}(x_1, x_2, \dots, x_n) &= g'\big((f(x_1, x_2, \dots, x_n)\big)\, f'_{x_1}(x_1, x_2, \dots, x_n), \\ h'_{x_2}(x_1, x_2, \dots, x_n) &= g'\big((f(x_1, x_2, \dots, x_n)\big)\, f'_{x_2}(x_1, x_2, \dots, x_n), \\ \dots &= \dots \\ h'_{x_n}(x_1, x_2, \dots, x_n) &= g'\big((f(x_1, x_2, \dots, x_n)\big)\, f'_{x_n}(x_1, x_2, \dots, x_n). \end{aligned}$$

Queste formule valgono, come già nel caso di due variabili, solo le la funzione f è differenziabile.

3. Ottimizzazione

3.1. Ottimizzazione libera in due variabili

I problemi principali a cui saremo interessati relativamente alle funzioni di più 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).

Cominceremo a trattare il caso delle funzioni di due variabili, in cui, come più volte accennato, è possibile e utile una significativa visualizzazione grafica delle situazioni che via via si presentano. Passeremo poi a trattare, seppure brevemente, il caso di funzioni di tre o più variabili.

È molto importante tenere presente che, trattando problemi di ottimizzazione per funzioni di più di una variabile, non si possono usare i termini "funzione crescente" o "funzione decrescente", in quanto il dominio è un sottoinsieme di $\mathbb{R}^n$, che non è un insieme ordinato. Il fatto che sia ancora possibile parlare di massimi e minimi è dovuto al fatto che il codominio di queste funzioni è ancora l'insieme $\mathbb{R}$ dei numeri reali che, invece, è un insieme ordinato.

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 1.8 (nella pagina 5), 1.9 (nella pagina 6) e 1.14 (nella pagina 9). Particolarmente significativa l'illustrazione 1.14 (nella pagina 9), 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 (2.6), 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 1.16 e 1.17).

Riproponiamo qui di seguito, per comodità, le stesse due figure con l'aggiunta del piano tangente.

La figura 3.2 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 3.3.

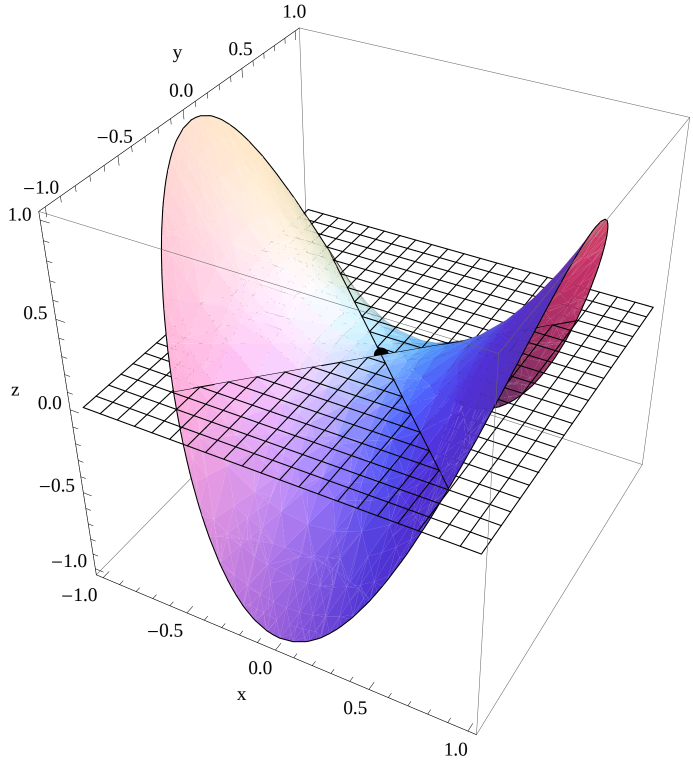


Figura 3.1. *Una "sella di cavallo" e il piano tangente*

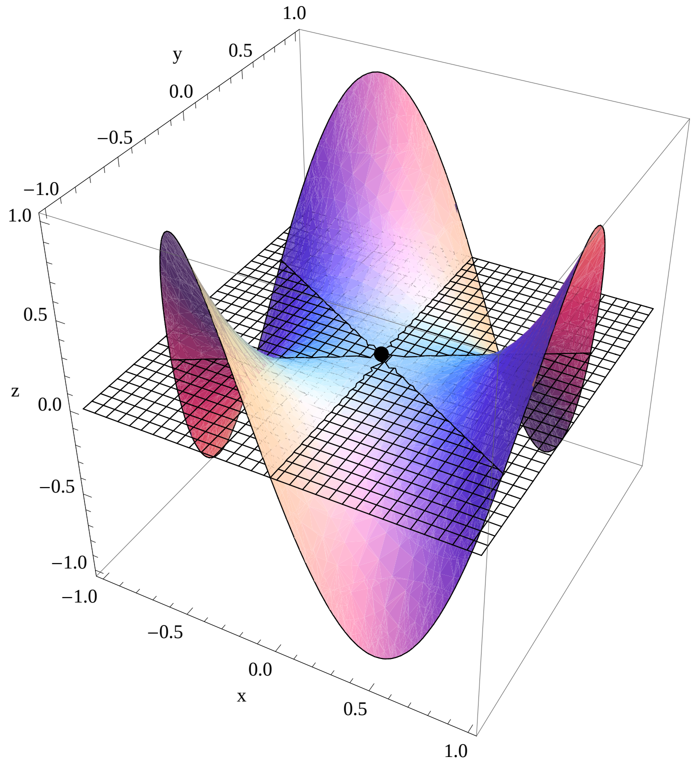


Figura 3.2. *Una "sella di scimmia" e il piano tangente*

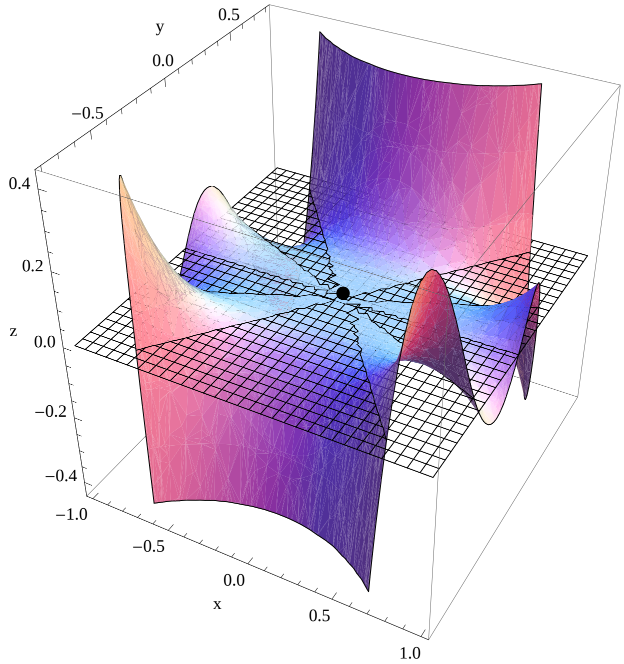


Figura 3.3. *Superficie con pronunciato "sfarfallio" rispetto al piano tangente orizzontale*

Quanto abbiamo detto si può riassumere nel seguente teorema.

Teorema 3.1 (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 3.2. *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*[1] *derivate seconde e si costruisce la seguente tabella (matrice), detta* matrice hessiana*,*

$$\begin{pmatrix} f''_{xx}(x_0,y_0) & f''_{xy}(x_0,y_0) \\ f''_{yx}(x_0,y_0) & f''_{yy}(x_0,y_0) \end{pmatrix}. \tag{3.1}$$

[1]In realtà ne bastano tre perché, nei casi che ci interessano, le due miste sono uguali.

Successivamente si calcola il seguente numero, detto determinante hessiano *o semplicemente* hessiano*, e indicato $H_f(x_0,y_0)$, o semplicemente con $H(x_0,y_0)$,*

$$f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)-f''_{xy}(x_0,y_0)f''_{yx}(x_0,y_0)=f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)-\Big(f''_{xy}(x_0,y_0)\Big)^2, \tag{3.2}$$

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*[2].

Si tenga ben presente che se $H(x_0,y_0)>0$, allora $f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)-\Big(f''_{xy}(x_0,y_0)\Big)^2>0$, da cui $f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)>\Big(f''_{xy}(x_0,y_0)\Big)^2\geq 0$, 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.

Esempio 3.1. 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[3].

3.2. Ottimizzazione vincolata in due variabili

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] E occorrerebbe un'indagine approfondita che di solito esula dagli scopi di questo corso.

[3] 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.

2. oppure sul bordo del dominio;
3. oppure sulla restrizione della funzione a una curva tracciata all'interno del dominio[4].

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 1.9 della pagina 6. 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 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 3.4 visualizza questa situazione. La figura di destra, 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.

Ebbene, è interessante 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.

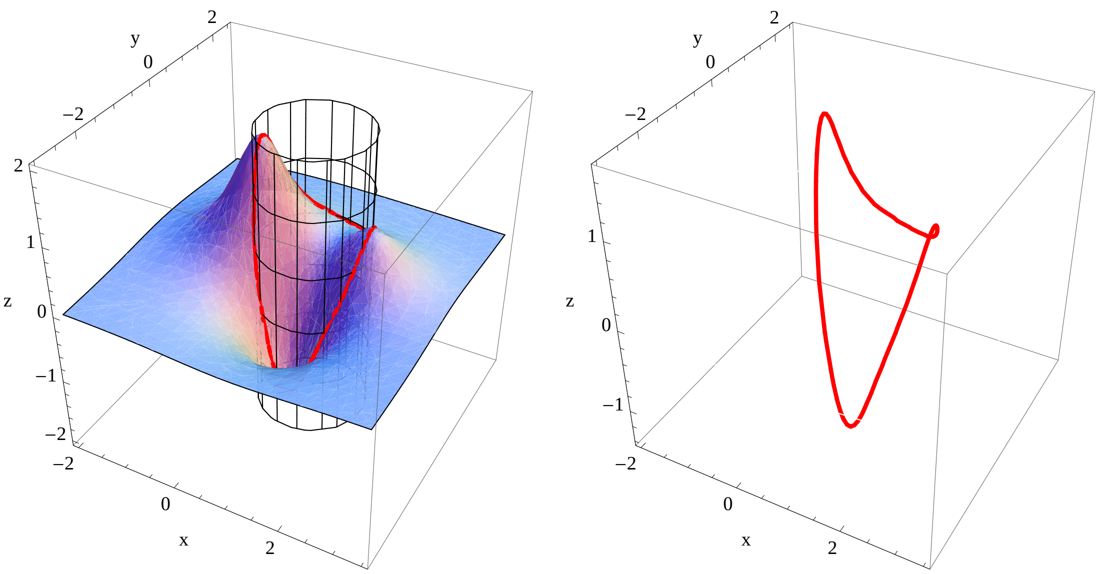


Figura 3.4. *Un problema di ottimizzazione vincolata*

[4] Ci sono anche alcune altre situazioni, ma ci limiteremo a trattare queste tre

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 $g(x, y) = 0$[(5)]. 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, e che, nella sostanza, si basa sul teorema di Dini, ovvero sulla possibilità, almeno teorica, di esplicitare una delle due variabili.

Esempio 3.2. 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 3.5.

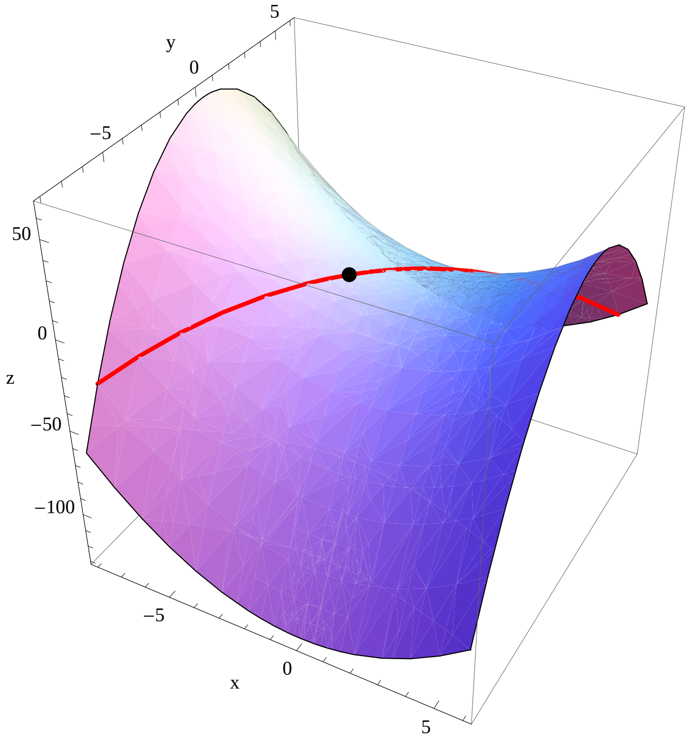


Figura 3.5. *Un massimo vincolato, con vincolo esplicitabile*

E veniamo ora al metodo dei *moltiplicatori di Lagrange*, da applicare quando non si può esplicitare alcuna variabile[(6)]. vale il seguente teorema.

Teorema 3.3. *Si supponga che le funzioni $f(x, y)$ e $g(x, y)$ siano continue con le loro derivate parziali in un aperto A, che $(x_0, y_0) \in A$ sia un punto soluzione dell'equazione $g(x, y) = 0$ e che in questo punto almeno una delle derivate parziali di g sia diversa da zero. Allora se (x_0, y_0) è di massimo o minimo vincolato per la f, con il vincolo $g(x, y) = 0$, esiste un numero reale λ tale che la funzione (detta "lagrangiana")*

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

[5]Attenzione: ridurre sempre il vincolo alla forma indicata, ovvero $g(x, y) = 0$.

[6]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.

abbia tutte le derivate nulle in (x_0, y_0, λ).

Si noti che questo teorema esprime solo una condizione necessaria per il massimo o minimo: se il punto è di massimo o minimo allora la funzione lagrangiana ha le derivate nulle, se però la funzione lagrangiana ha le derivate nulle, non è detto che il punto sia di massimo o minimo.

Non proponiamo la dimostrazione di questo teorema, ma facciamo notare come la condizione che almeno una delle derivate parziali di g sia diversa da zero, non è altro che la condizione richiesta dal teorema di Dini perché dall'equazione del vincolo $g(x, y) = 0$ si possa ricavare o la x o la y in funzione dell'altra variabile. In altre parole questo teorema sui massimi e minimi vincolati è una diretta conseguenza del teorema di Dini: anche se non si riesce a esplicitare direttamente o la x o la y, si può ugualmente concludere, in quanto il teorema di Dini permette comunque di avere il valore della derivata della funzione implicita almeno nel punto x_0 (o y_0). In sostanza il teorema di Dini consente, in un certo senso, di applicare il metodo dell'esplicitazione di una delle due variabili, anche quando questo non è tecnicamente possibile.

In pratica si procede nel seguente modo:

1. si costruisce la funzione lagrangiana (funzione di tre variabili)[7]: $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$;
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 3.3. 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:
$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) \quad \text{e} \quad \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)..$$

[7] 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.

Estraendo[8] da queste terne le coppie dei valori (x, y) si trovano i seguenti due punti:

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) \quad \text{e} \quad \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)..$$

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 3.6 illustra la situazione.

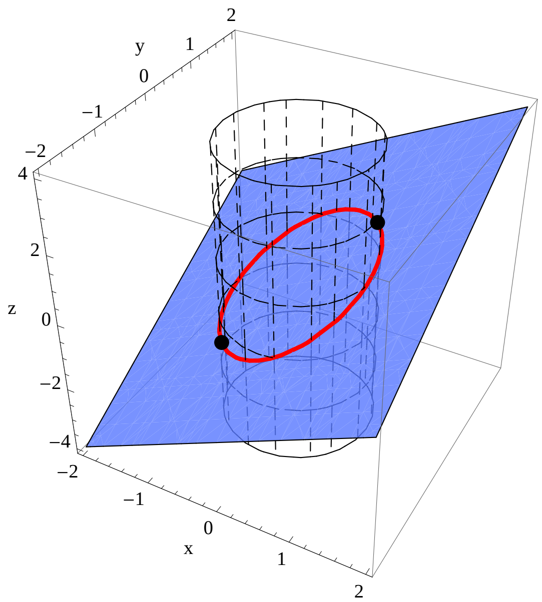


Figura 3.6. *Ottimizzazione vincolata, con il metodo dei moltiplicatori*

Esempio 3.4. 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 3.7 illustra la situazione.

[8] Si noti che, in realtà, il valore di λ non ha interesse in questa questione, bisogna però che un valore di λ esista.

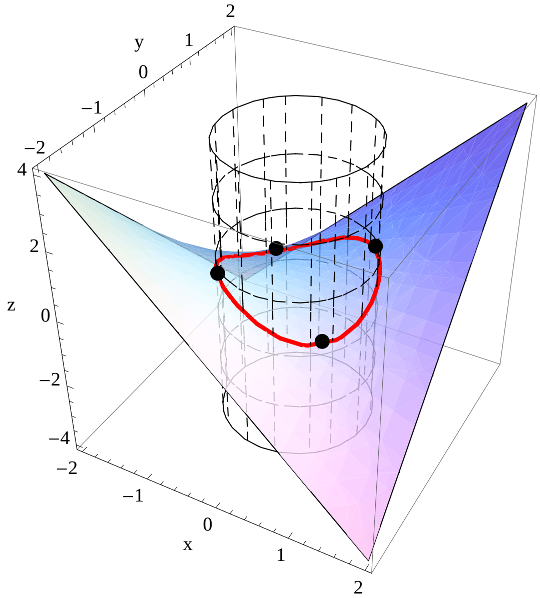


Figura 3.7. *Ancora una ottimizzazione vincolata, con il metodo dei moltiplicatori*

Esempio 3.5. 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[9] deve essere massimo. Quante macchine di ogni tipo deve acquistare?

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 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 3.8. Torneremo successivamente su questo punto.

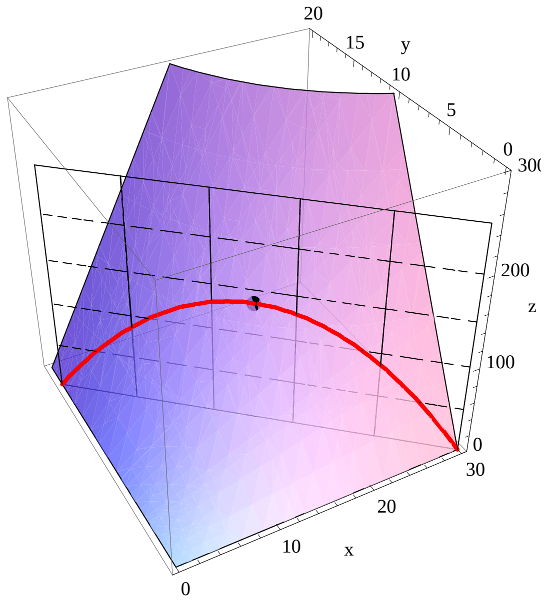


Figura 3.8. *Una ottimizzazione vincolata in un problema di economia*

[9] La funzione è detta proprio funzione di utilità.

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*:

$$\begin{pmatrix} 0 & g'_x & g'_y \\ g'_x & f''_{xx} + \lambda g''_{xx} & f''_{xy} + \lambda g''_{xy} \\ g'_y & f''_{yx} + \lambda g''_{yx} & f''_{yy} + \lambda g''_{yy} \end{pmatrix}. \tag{3.3}$$

Il determinante[10] 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.

3.3. Ottimizzazione globale su insiemi chiusi e limitati, in due variabili

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;
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.

[10] Il determinante di una matrice a 3 righe e 3 colonne come questa si fa con la regola di Sarrus, evidenziata nei passaggi seguenti.

$$\text{Dato} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}, \text{ costruisci } \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \begin{matrix} a & b \\ d & e \\ g & h \end{matrix}.$$

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

$$(aei + bfg + cdh) - (bdi + afh + crg).$$

Esempio 3.6. Trovare il massimo assoluto di

$$f(x,y) = x^2 + y^2 - xy + 3x - 7y$$

nell'insieme $T = \{\, (x, y \in \mathbb{R}^2 \mid 0 \le y \le 4 \wedge 0 \le x \le y \,\}$, rappresentato nella figura 3.9.

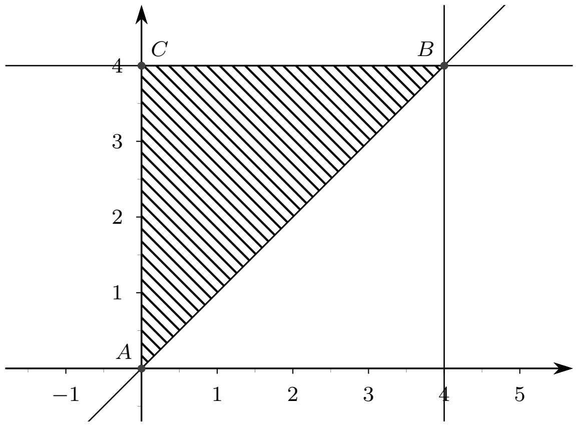


Figura 3.9. *Il triangolo* $T = \{\, (x, y \in \mathbb{R}^2 \mid 0 \le y \le 4 \wedge 0 \le x \le y \,\}$

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,

$$P\left(\frac{11}{3}, \frac{11}{3}\right).$$

Vediamo allora cosa succede sul bordo di T, distinguendo 3 casi.

1. Sul segmento AB si ha $y = x$, con $0 \le x \le 4$: 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 $0 \le x \le 4$: 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 $0 \le y \le 4$: 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.

3.4. Ottimizzazione libera in tre o più variabili

La condizione necessaria per la ricerca di massimi e minimi liberi in più di due variabili è una immediata generalizzazione di quella già esposta per il caso di due variabili.

Teorema 3.4 (Condizione necessaria per i massimi e minimi in n variabili)**.** *Se una funzione* $f(x_1, x_2, \dots, x_n)$ *dotata di derivate parziali ha, in corrispondenza a un punto* $(x_{1_0}, x_{2_0}, \dots, x_{n_0})$ interno *al dominio, un massimo o un minimo, allora necessariamente le derivate sono* contemporaneamente *nulle in* $(x_{1_0}, x_{2_0}, \dots, x_{n_0})$.

Tenendo conto della definizione di gradiente, la condizione può anche essere espressa dicendo che il gradiente della funzione in un punto di massimo o minimo deve essere il vettore nullo.

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. Il teorema precedente si può allora riformulare dicendo che: condizione necessaria perché un punto P_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, come già succedeva anche per le funzioni di una variabile.

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, come già osservato, per le funzioni di più variabili, dove i concetti di funzione crescente e decrescente *non hanno alcun senso.* Per risolvere il problema si può seguire una strategia simile (ma non identica) a quella del caso di due variabili, basata sul calcolo delle derivate seconde.

Data una funzione $f(x_1, x_2, \dots, x_n)$ dotata almeno di derivate seconde. Se $P_0 = (x_{1_0}, x_{2_0}, \dots, x_{n_0})$ è un punto stazionario per f, interno al dominio, calcoliamo, in quel punto, tutte le derivate seconde e costruiamo con esse la matrice già considerata a pagina 43, detta *matrice hessiana*.

$$H(x_{1_0}, x_{2_0}, \dots, x_{n_0}) = \begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} & f''_{x_1x_3} & \cdots & f''_{x_1x_n} \\ f''_{x_2x_1} & f''_{x_2x_2} & f''_{x_2x_3} & \cdots & f''_{x_2x_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f''_{x_nx_1} & f''_{x_nx_2} & f''_{x_nx_3} & \cdots & f''_{x_nx_n} \end{pmatrix} \tag{3.4}$$

e le seguenti sottomatrici

$$\begin{aligned} H_1(x_{1_0}, x_{2_0}, \dots, x_{n_0}) &= f''_{x_1x_1} \\ H_2(x_{1_0}, x_{2_0}, \dots, x_{n_0}) &= \begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} \\ f''_{x_2x_1} & f''_{x_2x_2} \end{pmatrix} \\ H_3(x_{1_0}, x_{2_0}, \dots, x_{n_0}) &= \begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} & f''_{x_1x_3} \\ f''_{x_2x_1} & f''_{x_2x_2} & f''_{x_2x_3} \\ f''_{x_3x_1} & f''_{x_3x_2} & f''_{x_3x_3} \end{pmatrix} \\ H_4(x_{1_0}, x_{2_0}, \dots, x_{n_0}) &= \cdots \end{aligned}$$

Indichiamo con M_1, M_2, M_3, ecc. i determinanti (vedi Appendice A) di queste matrici, detti *minori principali*. Vale allora il seguente teorema, che esprime una condizione sufficiente per i massimi o minimi liberi.

Teorema 3.5. *Se tutti i minori principali sono positivi, il punto P_0 è di minimo, se tutti i minori principali di ordine dispari sono negativi e quelli di ordine pari sono positivi, il punto P_0 è di massimo.*

Si danno le seguenti definizioni relativamente alla matrice hessiana.

Definizione 3.6. *La matrice Hessiana di una funzione di $\mathbb{R}^n$ in $\mathbb{R}$ si dice* definita positiva *se tutti i suoi minori principali sono positivi,* definita negativa *se tutti i minori principali di ordine dispari sono negativi e quelli di ordine pari sono positivi.*

Definizione 3.7. *La matrice Hessiana di una funzione di $\mathbb{R}^n$ in $\mathbb{R}$ si dice* semidefinita positiva *se tutti i suoi minori principali sono maggiori o uguali a* 0, semidefinita negativa *se tutti i minori principali di ordine dispari sono minori o uguali a* 0 *e quelli di ordine pari sono maggiori o uguali a* 0.

Si presti attenzione alla differenza delle due definizioni: nella prima si richiede che i minori siano strettamente maggiori o minori di 0, nella seconda che siano maggiori o uguali oppure minori o uguali a 0.

Con queste definizioni si può riformulare il teorema precedente dicendo che se la matrice hessiana è definita positiva, il punto critico P_0 è di minimo, se è definita negativa è di massimo.

Si noti che l'analogo teorema per le funzioni di *solo due* variabili è "più efficiente", in quanto contiene anche la condizione che se $M_2 < 0$, il punto è di sella. Per più di due variabili non vale una condizione simile.

Esempio 3.7. Sia

$$f(x,y,z) = \frac{1}{x} + y^2 + \frac{1}{z} + xz.$$

Si chiede di trovare i punti stazionari, precisando se si tratta di massimi o minimi.

Il dominio della funzione è costituito da tutti i punti di $\mathbb{R}^3$ tali che $x \neq 0 \wedge z \neq 0$. Si ha

$$\nabla f(x,y,z) = \left(-\frac{1}{x^2} + z \,,\, 2y \,,\, -\frac{1}{z^2} + x \right).$$

Uguagliando a zero le tre derivate parziali si ottiene come unico punto stazionario $P_0 = (1,0,1)$. Calcolando le derivate seconde si trova la matrice Hessiana

$$H(x,y,z) = \begin{pmatrix} \frac{2}{x^3} & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & \frac{2}{z^3} \end{pmatrix}, \quad \text{da cui} \quad H(1,0,1) = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Si ha allora $M_1 = 2$, $M_2 = 4$, $M_3 = 6$: il punto è di minimo relativo.

3.5. Funzioni convesse e concave

Per le funzioni a valori reali il concetto di funzione convessa o concava è particolarmente importante e consente di capire meglio il perché una funzione ha un massimo invece che un minimo, o viceversa, in un punto. Per maggiore chiarezza ripetiamo le definizioni e le proprietà già note per le funzioni convesse di una sola variabile, per poi passare all'estensione a funzioni di più variabili. Occorre tenere presente che il concetto di funzione convessa richiede l'uso di disuguaglianze sul codominio, e quindi si può porre solo per funzioni a valori reali.

Definizione 3.8. *Sia data una funzione $f\colon A \subseteq \mathbb{R} \to \mathbb{R}$. Chiameremo* epigrafico *o* sopragrafico *della funzione il sottoinsieme dei punti (x,y) del piano tali che $x \in A$ e $y \geq f(x)$, cioè l'insieme dei punti del piano che stanno sopra al grafico di una funzione o sul grafico stesso. Analoga la definizione per* sottografico.

Definizione 3.9. *Sia data una funzione $f\colon I \subseteq \mathbb{R} \to \mathbb{R}$, dove I è un intervallo. La funzione si dice* convessa *in I se il suo epigrafico è un insieme convesso, si dice* concava *se il suo sottografico è un insieme convesso.*

La richiesta che una funzione sia convessa (o concava) è molto stringente: si può infatti provare che una funzione convessa su un intervallo è sempre almeno continua.

Si noti che il concetto di funzione convessa viene dato solo per funzioni definite su un intervallo, e si ricordi che gli intervalli sono gli unici insiemi convessi di $\mathbb{R}$.

Nel caso di funzioni differenziabili (che in $\mathbb{R}$ significa semplicemente derivabili), il concetto di funzione convessa o concava si può esprimere in una maniera particolarmente significativa per le applicazioni.

Teorema 3.10. *Una funzione derivabile $f\colon I \subseteq \mathbb{R} \to \mathbb{R}$, dove I è un intervallo, è convessa se per ogni x_0 di I il grafico della funzione sta al di sopra della retta tangente al grafico nel punto x_0; la funzione è invece concava se il suo grafico sta al di sotto della retta tangente al grafico nel punto x_0. In formule: una funzione è convessa su un intervallo I se*

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0), \quad \forall x \in I \land \forall x_0 \in I.$$

Una funzione è concava su un intervallo I se

$$f(x) \leq f(x_0) + f'(x_0)(x - x_0), \quad \forall x \in I \land \forall x_0 \in I.$$

I concetti di funzione convessa e concava si possono estendere alle funzioni di più variabili, quasi con le stesse parole. Si deve solo tenere presente che il concetto di epigrafico o sottografico hanno ancora un significato geometrico solo per funzioni di due variabili (in cui il grafico, l'epigrafico e il sottografico sono sottoinsiemi dello spazio), mentre non hanno più alcuna rappresentazione geometrica intuitiva nel caso di funzioni di tre o più variabili.

Definizione 3.11. *Sia data una funzione $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$, dove A è un insieme convesso. La funzione si dice* convessa *in A se il suo epigrafico è un insieme convesso, si dice* concava *se il suo sottografico è un insieme convesso.*

Anche qui per le funzioni differenziabili si può esprimere la convessità con riguardo al piano (o "iperpiano") tangente al grafico della funzione.

Teorema 3.12. *Una funzione differenziabile $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$, dove A è un insieme convesso, è convessa se per ogni $\vec{x}_0$ di A il grafico della funzione sta al di sopra dell'iperpiano tangente al grafico nel punto $\vec{x}_0$; la funzione è invece concava se il suo grafico sta al di sotto dell'iperpiano tangente al grafico nel punto $\vec{x}_0$. In formule: una funzione è convessa su un insieme convesso A se*

$$f(\vec{x}) \geq f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0), \quad \forall \vec{x} \in A \land \forall \vec{x}_0 \in A.$$

Una funzione è concava su un intervallo A se

$$f(\vec{x}) \leq f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0), \quad \forall \vec{x} \in A \land \forall \vec{x}_0 \in A.$$

Tra i vari motivi per cui le funzioni convesse o concave sono importanti segnaliamo il fatto che se una funzione convessa è differenziabile in un insieme convesso e $\vec{x}_0$ è un punto critico (cioè dove si annullano le derivate parziali), allora $\vec{x_0}$ è di minimo assoluto, Se invece è concava, nelle stesso condizioni $\vec{x}_0$ è di massimo assoluto.

Valgono i seguenti teoremi di caratterizzazione per le funzioni due volte differenziabili.

1. Per le funzioni di una variabile: una funzione f è convessa su un intervallo I se e solo se $f''(x) \geq 0$, $\forall x \in I$ (concava se e solo se $f''(x) \leq 0$, $\forall x \in I$).
2. Per le funzioni di più variabili: una funzione f è convessa su un insieme convesso A se e solo se l'Hessiano è semidefinito positivo (concava se e solo se l'Hessiano è semidefinito negativo). Inoltre se l'Hessiano è definito positivo la funzione è (strettamente) convessa, se è definito negativo è (strettamente) concava.

3.6. Ottimizzazione vincolata in tre variabili

I problemi di ottimizzazione vincolata in più di due variabili sono decisamente più complessi che non nel caso di due variabili. Ci limiteremo a considerare solo il caso di tre variabili, nelle due situazioni possibili.

3.6.1. Funzione di tre variabili con un vincolo

Sia data una funzione $f(x, y, z)$ definita in un dominio D e si consideri un vincolo dato da[11] $g(x, y, z) = 0$. Si vogliono determinare i massimi e minimi di f con il vincolo dato. Vale il seguente teorema.

Teorema 3.13. *Si supponga che le funzioni f e g siano continue con le loro derivate parziali in un aperto A, che $(x_0, y_0, z_0) \in A$ sia un punto soluzione dell'equazione $g(x, y, z) = 0$ e che in questo punto almeno una delle derivate parziali di g sia diversa da zero. Allora se (x_0, y_0, z_0) è di massimo o minimo vincolato per la f, con il vincolo $g(x, y, z) = 0$, esiste un numero reale λ tale che la funzione (detta "lagrangiana")*

$$L(x, y, z, \lambda) = f(x, y, z) + \lambda g(x, y, z)$$

abbia tutte le derivate nulle in (x_0, y_0, z_0, λ).

Si noti che questo teorema esprime solo una condizione necessaria per il massimo o minimo: se il punto è di massimo o minimo allora la funzione lagrangiana ha le derivate nulle, se però la funzione lagrangiana ha le derivate nulle, non è detto che il punto sia di massimo o minimo.

In pratica si procede come segue:

1. se da $g(x, y, z) = 0$ si può ricavare la x, o la y, oppure la z, in funzione delle altre due variabili, la si sostituisce nella funzione f che diventa funzione di due variabili, e si procede appunto come nella funzioni di due variabili;
2. se questo non è possibile, o è troppo complesso, si usa il teorema appena enunciato:

1. si costruisce la funzione lagrangiana (funzione di quattro variabili): $L(x, y, z, \lambda) = f(x, y, z) + \lambda g(x, y, z)$;
2. si calcolano le derivate parziali rispetto alle quattro variabili: L'_x, L'_y, L'_z, L'_λ;
3. si risolve il sistema
$$\begin{cases} L'_x = 0 \\ L'_y = 0 \\ L'_y = 0 \\ L'_\lambda = 0 \end{cases};$$
4. se ci sono punti di massimo o minimo vincolato questi si trovano tra le terne (x, y, z) estratte dalle quaterne (x, y, z, λ) che risolvono il sistema. Si noti come la situazione sia simile a quella dei massimi e minimi liberi: le terne (x, y, z) estratte dalle quaterne (x, y, z, 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 tre 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.

[11] Anche se non possiamo entrare nei dettagli, segnaliamo che un'equazione del tipo $g(x, y, z) = 0$, sotto opportune condizioni di regolarità, rappresenta una superficie dello spazio. Poiché il dominio D è un insieme, generalmente tridimensionale, di punti dello spazio, la considerazione del vincolo restringe il calcolo dei massimi e minimi solamente ai punti di D appartenenti alla superficie rappresentata dal vincolo: è la generalizzazione del caso di due variabili in cui cercavamo i massimi e minimi non su tutti i punti di un insieme bidimensionale, ma solamente su quelli di una curva tracciata in quell'insieme.

Esempio 3.8. Si vogliano trovare il massimo e il minimo della funzione $f(x, y, z) = x^2 + y^2 + z^2$ con il vincolo $x^4 + y^4 + z^4 - 1 = 0$.

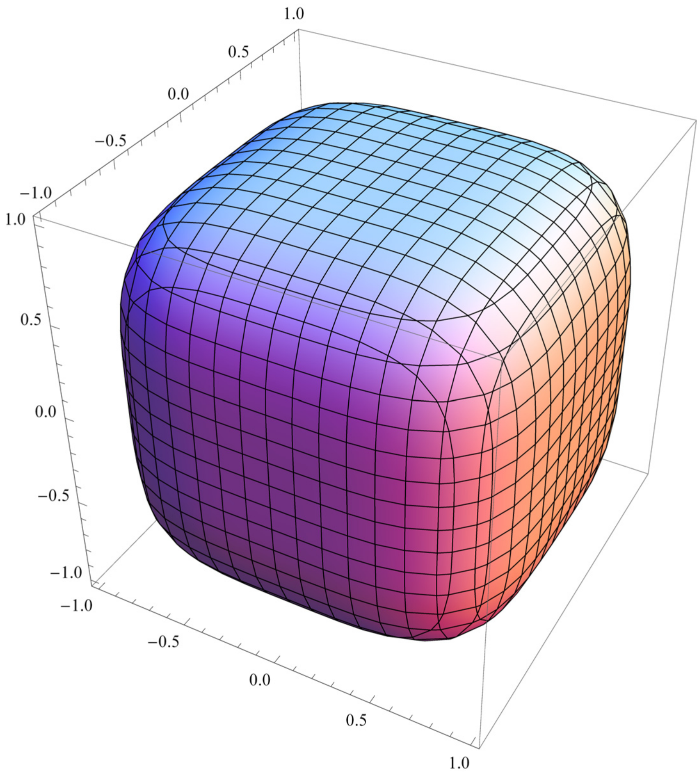


Figura 3.10. *Superficie vincolo per l'esempio* 3.8

Considerata la funzione Lagrangiana

$$L(x, y, z, \lambda) = x^2 + y^2 + z^2 + \lambda(x^4 + y^4 + z^4 - 1),$$

si perviene a considerare il sistema

$$\begin{cases} 2x + 4\lambda x^3 = 0 \\ 2y + 4\lambda y^3 = 0 \\ 2z + 4\lambda z^3 = 0 \\ x^4 + y^4 + z^4 - 1 = 0 \end{cases}$$

Poiché c'è simmetria rispetto al cambiamento di segno di x, y e z, ci limitiamo a trovare solo le soluzioni con $x \ge 0$, $y \ge 0$ e $z \ge 0$ (punti del primo ottante). Con un po' di pazienza si trovano sette punti che riportiamo di seguito, con accanto il valore della funzione f in ciascuno.

$$P_1 = (1, 0, 0) \qquad f(P_1) = 1;$$

$$P_2 = (0, 1, 0) \qquad f(P_2) = 1;$$

$$P_3 = (0, 0, 1) \qquad f(P_1) = 1;$$

$$P_4 = \left(\frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}}, 0\right) \qquad f(P_4) = \sqrt{2};$$

$$P_5 = \left(\frac{1}{\sqrt[4]{2}}, 0, \frac{1}{\sqrt[4]{2}}\right) \qquad f(P_5) = \sqrt{2};$$

$$P_6 = \left(0, \frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}}\right) \qquad f(P_4) = \sqrt{2};$$

$$P_7 = \left(\frac{1}{\sqrt[4]{3}}, \frac{1}{\sqrt[4]{3}}, \frac{1}{\sqrt[4]{3}} \right) \qquad f(P_7) = \sqrt{3}.$$

Considerato che l'insieme rappresentato dal vincolo è chiuso e limitato, si conclude che il minimo vale 1, mentre il massimo vale $\sqrt{3}$.

Solo per curiosità abbiamo riportato nella figura 3.8 l'illustrazione grafica della superficie-vincolo su cui abbiamo calcolato il massimo e minimo vincolato in questo esempio.

3.6.2. Funzione di tre variabili con due vincoli

Sia data una funzione $f(x, y, z)$ definita in un dominio D e si consideri il sistema costituito da due vincoli[12]

$$\begin{cases} g(x, y, z) = 0 \\ h(x, y, z) = 0 \end{cases} . \tag{3.5}$$

Si vogliono determinare i massimi e minimi di f con i vincoli dati. Vale il seguente teorema.

Teorema 3.14. *Si supponga che le funzioni f, g e h siano continue con le loro derivate parziali in un aperto A, che $(x_0, y_0, z_0) \in A$ sia un punto soluzione del sistema (3.5) e che in questo punto la matrice*

$$\begin{pmatrix} g'_x(x_0, y_0, z_0) & g'_y(x_0, y_0, z_0) & g'_z(x_0, y_0, z_0) \\ h'_x(x_0, y_0, z_0) & h'_y(x_0, y_0, z_0) & h'_z(x_0, y_0, z_0) \end{pmatrix}$$

abbia rango 2. Allora se (x_0, y_0, z_0) è di massimo o minimo vincolato per la f, con i vincoli espressi dal sistema (3.5), esistono due numeri reali λ e μ tali che la funzione (detta "lagrangiana")

$$L(x, y, z, \lambda, \mu) = f(x, y, z) + \lambda g(x, y, z) + \mu h(x, y, z)$$

abbia tutte le derivate nulle in $(x_0, y_0, z_0, \lambda, \mu)$.

Si noti che questo teorema esprime solo una condizione necessaria per il massimo o minimo: se il punto è di massimo o minimo allora la funzione lagrangiana ha le derivate nulle, se però la funzione lagrangiana ha le derivate nulle, non è detto che il punto sia di massimo o minimo.

In pratica si procede come segue:

1. se dal sistema dei due vincoli si possono ricavare due delle tre variabili in funzione della terza, le si sostituiscono nella funzione f che diventa funzione di una sola variabile, e si procede appunto come nella funzioni di una variabile;
2. se questo non è possibile, o è troppo complesso, si usa il seguente metodo dei moltiplicatori di Lagrange.

Si procede nel seguente modo:

1. si costruisce la funzione lagrangiana (funzione di cinque variabili): $L(x, y, z, \lambda, \mu) = f(x, y, z) + \lambda g(x, y, z) + \mu h(x, y, z)$;
2. si calcolano le derivate parziali rispetto alle cinque variabili: L'_x, L'_y, L'_z, L'_λ, L'_μ;

[12] Anche se non possiamo entrare nei dettagli, segnaliamo che un sistema di due vincoli come quelli considerati, sotto opportune condizioni di regolarità, rappresenta una curva dello spazio, ottenuta come intersezione di due superfici. Poiché il dominio D è un insieme, generalmente tridimensionale, di punti dello spazio, la considerazione del vincolo restringe il calcolo dei massimi e minimi solamente ai punti di D appartenenti alla curva rappresentata dal vincolo: è la generalizzazione del caso di due variabili in cui cercavamo i massimi e minimi non su tutti i punti di un insieme bidimensionale, ma solamente su quelli di una curva tracciata in quell'insieme.

3. si risolve il sistema
$$\begin{cases} L'_x = 0 \\ L'_y = 0 \\ L'_y = 0 \\ L'_\lambda = 0 \\ L'_\mu = 0 \end{cases} ;$$

4. se ci sono punti di massimo o minimo vincolato questi si trovano tra le terne (x, y, z) estratte dalle cinquine (x, y, z, λ, μ) che risolvono il sistema. Si noti come la situazione sia simile a quella dei massimi e minimi liberi: le terne (x, y, z) estratte dalle cinquine (x, y, z, λ, μ) 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 tre 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 3.9. Calcolare gli estremi della funzione $f(x, y, z) = x + y^2 z$ con i vincoli $x^2 + y^2 - 2 = 0$ e $z - x = 0$.

Scritta la funzione lagrangiana e calcolate le derivate parziali si ottiene il sistema
$$\begin{cases} 1 + 2\lambda x - \mu = 0 \\ 2yz + 2\lambda y = 0 \\ y^2 + \mu = 0 \\ x^2 + y^2 - 2 = 0 \\ z - x = 0 \end{cases} .$$

Con un po' di pazienza si trovano sei punti che riportiamo di seguito, con accanto il valore della funzione f in ciascuno.

$$\begin{array}{ll} P_1 = (\sqrt{2}, 0, \sqrt{2}) & f(P_1) = \sqrt{2}; \\ P_2 = (-\sqrt{2}, 0, -\sqrt{2}) & f(P_2) = \sqrt{2}; \\ P_3 = (1, 1, 1) & f(P_3) = 2; \\ P_4 = (1, -1, 1) & f(P_4) = 2; \\ P_5 = (-1, 1, -1) & f(P_5) = -2; \\ P_6 = (-1, -1, -1) & f(P_6) = -2. \end{array}$$

Il minimo vale dunque -2 e il massimo 2.

Solo per curiosità riportiamo l'illustrazione grafica della curva-vincolo su cui abbiamo calcolato il massimo e minimo vincolato in questo esempio. Questa curva è ottenuta come intersezione di due superfici.

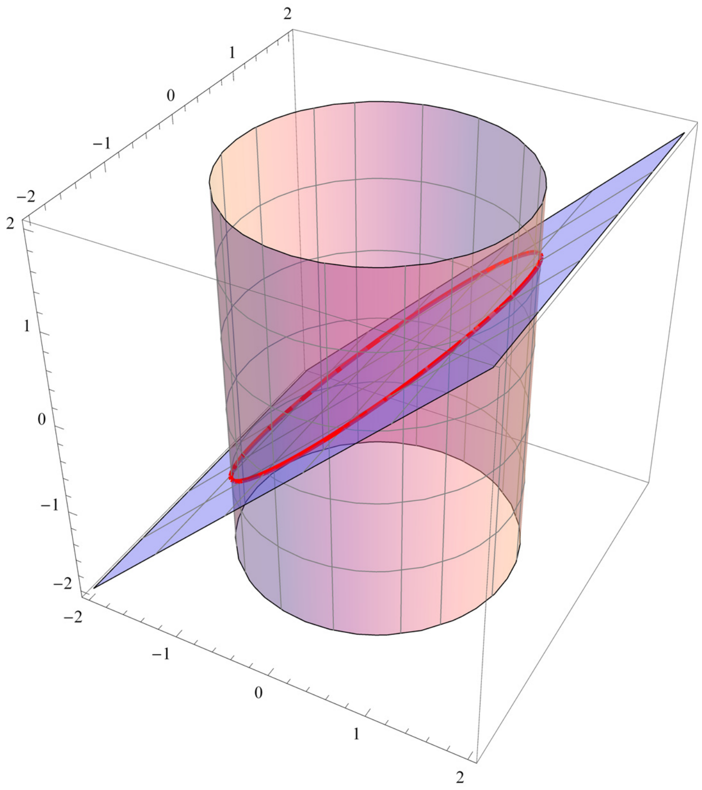


Figura 3.11. *Curva vincolo per l'esempio* 3.9

4. Cenno ai campi vettoriali

4.1. Definizioni

Parlando di funzioni di più variabili normalmente si sottintende il fatto che si considerano funzioni di $\mathbb{R}^n$ in $\mathbb{R}$: si dovrebbe più propriamente parlare di funzioni "reali di più variabili reali". Nelle applicazioni hanno però interesse anche funzioni a valori in $\mathbb{R}^m$: non ci sono grosse difficoltà nella loro trattazione: è sufficiente considerare anziché una sola funzione di $\mathbb{R}^n$ in $\mathbb{R}$, una m-pla funzioni di $\mathbb{R}^n$ in $\mathbb{R}$. Se per esempio consideriamo due funzioni di due variabili:

$$f_1(x,y) = x^2 + y^2 \quad \text{e} \quad f_2(x,y) = x - 2y,$$

potremo considerare la coppia[1] di funzioni (f_1, f_2), come una funzione, diciamola F, il cui dominio è $\mathbb{R}^2$ e il cui codominio é ancora $\mathbb{R}^2$:

$$F(x,y) = \big(f_1(x,y), f_2(x,y)\big).$$

Questa funzione alla coppia, per esempio, $(1,2)$ fa corrispondere la coppia $(5,-3)$. Potremo scrivere $F(1,2) = (5,-3)$. In una situazione come questa le due funzioni f_1 e f_2 si chiamano *componenti* della funzione F.

Per queste funzioni potremo usare una parte della nomenclatura già utilizzata per le funzioni a valori reali. Per esempio, data la funzione

$$F\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad F(x,y) = \left(\sqrt{x^2+y^2-1}\,, \sqrt{2x+1} + y^2\right),$$

potremo cercare il suo "dominio naturale": si dovranno considerare le condizioni da imporre a *ciascuna* delle funzioni componenti e poi farne l'intersezione. Si otterrebbe, nel caso in esame,

$$\begin{cases} x^2+y^2-1 \geq 0 \\ 2x+1 \geq 0 \end{cases}.$$

Il dominio nel piano x, y è rappresentato nella figura 4.1.

Trattando funzioni a valori in $\mathbb{R}^m$ dovremo però prestare attenzione al fatto che, non essendo il codominio l'insieme dei numeri reali, *non* sarà possibile in generale parlare di somma, prodotto, quoziente di due funzioni. Tratteremo, in un caso di interesse, il problema della composizione di due funzioni. Nessun problema ci sarà invece per introdurre i concetti di limite, continuità, derivate parziali: basterà riferirsi a ciascuna delle funzioni componenti. Per esempio, data la funzione

$$F\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad F(x,y) = \big(f_1(x,y), f_(x,y)\big) = (x^2+xy, x-2x^2y),$$

potremo considerare le sue quattro derivate parziali:

$$\frac{\partial f_1}{\partial x} = 2x + y, \quad \frac{\partial f_1}{\partial y} = x, \quad \frac{\partial f_2}{\partial x} = 1 - 4xy, \quad \frac{\partial f_2}{\partial y} = -2x^2.$$

[1] Attenzione. Abbiamo volutamente parlato di "coppia" di funzioni, in quanto interessa l'ordine in cui le funzioni sono considerate.

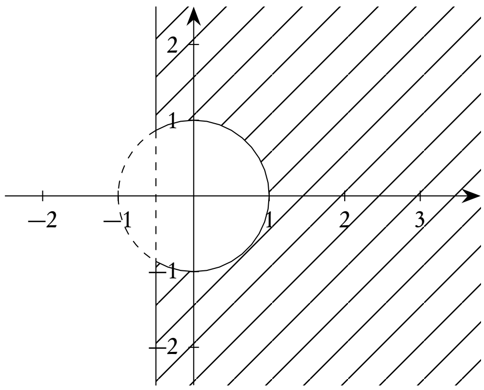


Figura 4.1. *Dominio della funzione* $F(x,y)=\left(\sqrt{x^2+y^2-1}\,,\,\sqrt{2x+1}+y^2\right)$

Abitualmente si usa scrivere queste quattro derivate utilizzando una matrice in cui sulle righe ci sono le derivate delle diverse componenti della funzione:

$$J_F=\begin{pmatrix}\dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y}\\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y}\end{pmatrix}=\begin{pmatrix}2x+y & x\\ 1-4xy & -2x^2\end{pmatrix} \tag{4.1}$$

e questa matrice si chiama *matrice jacobiana* o semplicemente *Jacobiano*. Nel caso delle funzioni a valori in $\mathbb{R}$, cioè funzioni che hanno "una sola" componente, la matrice jacobiana si riduce semplicemente al gradiente. Nel caso ancora più particolare di funzioni reali di una sola variabile reale, la matrice jacobiana si riduce alla semplice derivata della funzione.

Una osservazione particolarmente importante riguarda il fatto che, essendo il codominio di queste funzioni un insieme non ordinato, *non* si potranno introdurre i concetti di massimo e minimo, come invece succedeva per le funzioni a valori reali.

Le funzioni a valori in $\mathbb{R}^m$, con $m>1$, sono anche chiamate *campi vettoriali*, per contro le funzioni a valori in $\mathbb{R}$ sono chiamate *campi scalari*.

4.2. Funzioni composte

Avendo a disposizione anche funzioni di $\mathbb{R}^n\to\mathbb{R}^m$, si potranno considerare funzioni composte in situazioni più complesse rispetto al casi di funzioni solo a valori in $\mathbb{R}$.

Più precisamente: date due funzioni

$$f\colon\mathbb{R}^n\to\mathbb{R}^p \quad \text{e} \quad g\colon\mathbb{R}^p\to\mathbb{R}^m,$$

potremo considerare la funzione

$$F=g\circ f\colon\mathbb{R}^n\to\mathbb{R}^m.$$

Per esempio se

$$f\colon\mathbb{R}^2\to\mathbb{R}^2,\ (x,y)\mapsto\big(f_1(x,y),f_2(x,y)\big)=(x^2+y,x-2y) \text{ e } g\colon\mathbb{R}^2\to\mathbb{R},\ (x,y)\mapsto 2x+3y,$$

la funzione composta

$$g\circ f\colon\mathbb{R}^2\to\mathbb{R}$$

è data da

$$(x,y) \mapsto g\big(f(x,y)\big) = g\big(f_1(x,y), f_2(x,y)\big) = 2(x^2+y) + 3(x-2y) = 2x^2 + 3x - 4y.$$

Non sarà invece possibile considerare la funzione composta $f \circ g$.

Il teorema di derivazione delle funzioni composte, utilizzando la matrice jacobiana, si esprime in una forma particolarmente significativa e che generalizza, anche nel suo aspetto formale, la nota formula sulla derivata delle funzioni composte di una sola variabile. Precisamente vale il seguente teorema.

Teorema 4.1 (Derivazione di funzioni composte)**.** *Date due funzioni*[2]

$$f\colon \mathbb{R}^n \to \mathbb{R}^p \quad e \quad g\colon \mathbb{R}^p \to \mathbb{R}^m,$$

entrambe differenziabili, la funzione composta $F = g \circ f$ è anch'essa differenziabile e la sua matrice jacobiana si ottiene facendo il prodotto, righe per colonne, delle matrici jacobiane di g e f, nell'ordine:

$$J_F = J_{g\circ f} = J_g \cdot J_f. \tag{4.2}$$

Naturalmente bisognerà tenere in debito conto il fatto che le derivate di f vanno calcolate in

$$(x_1, x_2, \dots, x_n),$$

mentre le derivate di g andranno calcolate nel punto

$$\big(f_1(x_1,x_2,\dots,x_n), \dots, f_p(x_1,x_2,\dots,x_n)\big).$$

Consideriamo per esempio le funzioni f e g

$$f\colon \mathbb{R}^2 \to \mathbb{R}^2,\ (x,y) \mapsto \big(f_1(x,y), f_2(x,y)\big) = (x^2+y^2, -2y^2) \text{ e } g\colon \mathbb{R}^2 \to \mathbb{R},\ (x,y) \mapsto 2x^2 - 3y^2.$$

Avremo, intanto

$$F = g\circ f\colon \mathbb{R}^2 \to \mathbb{R}, \quad F(x,y) = 2(x^2+y^2)^2 - 3(-2y^2)^2 = 2x^4 - 10y^4 + 4x^2y^2,$$

e dunque

$$J_F = \nabla F = (8x^3 + 8xy^2, -40y^3 + 8x^2y).$$

Abbiamo inoltre

$$J_f = \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ 0 & -4y \end{pmatrix}$$

e

$$J_g = \nabla g = \big(4x, -6y\big).$$

Lo Jacobiano di g va calcolato nel punto $f(x,y)$. Si ottiene

$$J_g(f(x,y)) = \big(4(x^2+y^2), -6(-2y^2)\big) = (4x^2+4y^2, 12y^2).$$

Quindi

$$J_g \cdot J_f = (4x^2+4y^2, 12y^2) \cdot \begin{pmatrix} 2x & 2y \\ 0 & -4y \end{pmatrix} = (8x^3 + 8xy^2, -40y^3 + 8x^2y),$$

esattamente come prima.

[2] Si noti che la funzione f è composta da p funzioni di n variabili, mentre la funzione g è composta da m funzioni di p variabili.

4.3. Il caso delle funzioni di una variabile

Tra tutti i campi vettoriali ci interessa in modo particolare il caso delle funzioni

$$f\colon \mathbb{R} \to \mathbb{R}^m,$$

cioè il caso il cui tutte le funzioni componenti sono funzioni di una sola variabile. In questo caso è tradizione indicare la variabile[3] con t. Inoltre le funzioni componenti sono di solito indicate con $x_1(t), x_2(t), \dots, x_m(t)$. La situazione dal punto di vista delle derivate è ora particolarmente semplice, in quanto ciascuna delle funzioni componenti è funzione di una sola variabile e dunque avrà una sola derivata "ordinaria", anziché derivate parziali. A motivo di questo fatto si può parlare di "derivata" anche per la funzione f, anziché di Jacobiano, anche se si deve tenere presente che la derivata in questo caso non è un numero, ma un vettore a m componenti (matrice jacobiana con una sola colonna e m righe).

Per esempio si consideri la funzione

$$f\colon \mathbb{R} \to \mathbb{R}^3,\ f(t) = (x_1(t), x_2(t), x_3(t)) = (x(t), y(t), z(t)) = (t^2, 3t+1, 2t^2-t).$$

Si ha

$$f'(t) = J_f(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix} = \begin{pmatrix} 2t \\ 3 \\ 4t-1 \end{pmatrix}.$$

Il problema di tracciare il grafico per questo tipo di funzioni si potrebbe porre solo nel caso $f\colon \mathbb{R} \to \mathbb{R}^2$ (tre variabili in totale), ma la cosa non ha alcun interesse pratico. Nel caso di funzioni $f\colon \mathbb{R} \to \mathbb{R}^2$ oppure di funzioni $f\colon \mathbb{R} \to \mathbb{R}^3$, ha invece un grande interesse applicativo la considerazione dell'insieme immagine, che è costituito da un sottoinsieme di $\mathbb{R}^2$ (cioè del piano) nel primo caso e di $\mathbb{R}^3$ (cioè dello spazio) nel secondo caso. Nei casi di nostro interesse questi sottoinsiemi sono sempre costituiti da curve del piano o dello spazio rispettivamente.

Proponiamo due esempi, nel piano.

Esempio 4.1. Sia $f\colon \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t, t^2)$, funzione che si usa scrivere come

$$\begin{cases} x = t \\ y = t^2 \end{cases} . \tag{4.3}$$

Queste formule si chiamano anche "equazioni parametriche" della curva immagine della funzione. Ricavando t dalla prima equazione e sostituendo nella seconda si trova l'equazione della parabola $y = x^2$, che è proprio l'insieme immagine della funzione f in esame. Il lettore può verificare, mediante sostituzione diretta, che prendendo un valore di t qualunque, dalle formule 4.3 si ottiene sempre un punto della parabola e che, viceversa, ogni punto della parabola si ottiene dalle 4.3 con un opportuno valore di t.

Esempio 4.2. Sia $f\colon \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t^2-1, t^3-t)$, funzione che si usa scrivere come

$$\begin{cases} x = t^2 - 1 \\ y = t^3 - t \end{cases} .$$

In questo caso è più complesso eliminare il "parametro" t. La curva immagine di questa funzione è rappresentata nella figura 4.2.

[3] Il motivo di questo fatto è da ricercarsi nel fatto che, in fisica, una funzione di $\mathbb{R}$ in $\mathbb{R}^3$ può essere interpretata come il moto di un punto nello spazio al variare del tempo t.

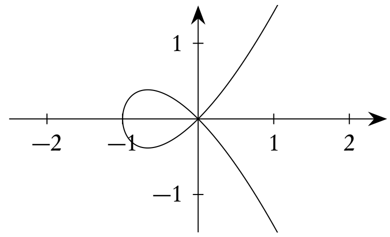


Figura 4.2. *Curva immagine della funzione* $f(t) = \left(t^2 - 1, t^3 - t\right)$

Utilizzando le funzioni $f\colon \mathbb{R} \to \mathbb{R}^n$ e le funzioni $g\colon \mathbb{R}^n \to \mathbb{R}$, si può costruire sia la funzione composta

$$f \circ g\colon \mathbb{R}^n \to \mathbb{R}^n$$

che la funzione composta

$$g \circ f\colon \mathbb{R} \to \mathbb{R}.$$

Siamo particolarmente interessati al secondo caso, in cui il risultato finale è una "normale" funzione reale di una sola variabile reale.

Esempio 4.3. Siano date le due funzioni

$$f\colon \mathbb{R} \to \mathbb{R}^2,\ f(t) = (t^3 - t^2, 2t + t^2) \quad \text{e} \quad g\colon \mathbb{R}^2 \to \mathbb{R},\ g(x, y) = \sin x + xy,$$

di cui la prima si usa scrivere come

$$\begin{cases} x(t) = t^3 - t^2 \\ y(t) = 2t + t^2 \end{cases}.$$

Si ha allora

$$F = g \circ f\colon \mathbb{R} \to \mathbb{R},\ F(t) = \sin(t^3 - t^2) + (t^3 - t^2)(2t + t^2) = \sin(t^3 - t^2) + t^5 + t^4 - 2t^3.$$

Si ha poi

$$F'(t) = (3t^2 - 2t)\cos(t^3 - t^2) + 5t^4 + 4t^3 - 6t^2.$$

Questa derivata si può anche calcolare con la regola di derivazione delle funzioni composte che abbiamo visto nel teorema 4.1. Si ha

$$f'(t) = J_f = \begin{pmatrix} 3t^2 - 2t \\ 2 + 2t \end{pmatrix}, \quad J_g = \nabla g = \left(\cos x + y, x\right).$$

Calcoliamo lo Jacobiano di g nel punto $(t^3 - t^2, 2t + t^2)$:

$$J_g(t^3 - t^2, 2t + t^2) = \left(\cos(t^3 - t^2) + 2t + t^2, t^3 - t^2\right).$$

Allora

$$\begin{aligned} F'(t) = J_g \cdot J_f &= \left(\cos(t^3 - t^2) + 2t + t^2, t^3 - t^2\right) \cdot \begin{pmatrix} 3t^2 - 2t \\ 2 + 2t \end{pmatrix} = \\ &= \left(\cos(t^3 - t^2) + 2t + t^2\right)\left(3t^2 - 2t\right) + \left(t^3 - t^2\right)\left(2 + 2t\right) = \\ &= (3t^2 - 2t)\cos(t^3 - t^2) + 5t^4 + 4t^3 - 6t^2, \end{aligned}$$

esattamente come prima.

Si usano le seguenti scritture semplificate per riassumere quanto detto.

$$F(t) = g(x(t), y(t)),$$

e

$$F'(t) = \frac{\partial g}{\partial x}((x(t), y(t))x'(t) + \frac{\partial g}{\partial y}((x(t), y(t))y'(t),$$

o anche

$$F'(t) = \frac{\partial g}{\partial x}((x(t), y(t))\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial g}{\partial y}((x(t), y(t))\frac{\mathrm{d}y}{\mathrm{d}t},$$

o ancora

$$F'(t) = \frac{\partial g}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial g}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t},$$

sottintendendo il fatto che le derivate parziali della g vanno calcolate nel punto $(x(t), y(t))$.

5. Esercizi

Proponiamo qui di seguito alcuni semplici esercizi introduttivi. Altri esercizi saranno proposti durante il corso.

Esercizio 5.1. *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. $f(x,y) = \mathrm{e}^x y$.
4. $f(x,y) = \mathrm{e}^{xy}xy$.
5. $f(x,y) = y \ln x$.
6. $f(x,y) = \ln(xy$.
7. $f(x,y) = \dfrac{\ln x}{y}$.
8. $f(x,y) = \mathrm{e}^{x+xy^2}$.

Esercizio 5.2. *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. $f(x,y) = x^2\mathrm{e}^y$, $P(0,0)$.
5. $f(x,y) = xy\mathrm{e}^{x+y}$, $P(1,0)$, $Q(-1,1)$, $R(1,-1)$.
6. $f(x,y) = x^2\mathrm{e}^{3y-x}$, $P(1,0)$, $Q(0,0)$.
7. $f(x,y) = \mathrm{e}^{xy}$, $P(0,0)$.
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. $f(x,y) = \mathrm{e}^{xy-x}$, $P(0,1)$.
11. $f(x,y) = x^2 - \mathrm{e}^{y^2}$, $P(0,0)$.
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. $f(x,y) = \mathrm{e}^{xy} - y$, $P(1,0)$, $Q(0,1)$.

16. $f(x,y) = \mathrm{e}^{xy} - y^2$, $P(0,0)$.
17. $f(x,y) = \mathrm{e}^{xy} - xy$, $P(0,1)$, $Q(1,0)$, $R(1,1)$.
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 5.3. *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. $f(x,y) = \mathrm{e}^{x-y+2}$, $\sqrt{x-y+xy} = 1$.
5. $f(x,y) = \ln(x-y)$, $x^2 - y^2 = 3$.

Esercizio 5.4. *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* $(x-1)^2 + (y-1)^2 = 1$; $P(1,1)$, $Q\left({(\sqrt{2}+1)}/{\sqrt{2}}, {(\sqrt{2}+1)}/{\sqrt{2}}\right)$.
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 5.5. *Come per l'esercizio 5.4; 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* $x^3 - y^3 = 0$; $P(0,0)$, $Q(1,1)$, $R(-1,-1)$.

A. Richiami su determinante e rango di una matrice

A.1. Determinante di una matrice quadrata

Daremo solo una definizione ricorsiva di determinante, basata su un teorema dovuto a Laplace: data la definizione per la matrici di ordine 1, la definizione per matrici di ordine superiore si riconduce, mediante passaggi successivi, al caso $n = 1$.

Per una matrice quadrata di ordine 2 (come la matrice hessiana di una funzione di due variabili) abbiamo già dato la definizione di determinante (vedi la definizione di determinante hessiano nella pagina 48). Per una matrice quadrata di ordine 3 abbiamo già usato la regola di Sarrus (vedi la nota nella pagina 54).

Per il determinante di una matrice A useremo le seguenti notazioni:

$$|A| = \det \boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \cdot \end{vmatrix} . \tag{A.1}$$

Definizione A.1 (Minori di una matrice)**.** *Data una matrice A si chiama sottomatrice di A ogni matrice ottenuta da A sopprimendo un certo numero di righe e un certo numero di colonne (anche non consecutive). Si chiama* minore *di A una sottomatrice quadrata di A. Si può usare il termine minore anche per indicare il determinante del minore stesso.*

Definizione A.2 (Complemento algebrico o cofattore)**.** *Dato un elemento a_{ij} di una matrice quadrata A si chiama suo* complemento algebrico *o* cofattore*, e si indica con A_{ij}, il determinante, moltiplicato per $(-1)^{i+j}$, del minore che si ottiene sopprimendo la riga e la colonna di A che si intersecano in a_{ij}; questo minore è anche detto minore complementare di a_{ij}.*

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & m_{m,j} & a_{m,j+1} & \cdots & a_{m,n} \end{pmatrix} , \tag{A.2}$$

$$\boldsymbol{A}_{ij} = (-1)^{i+j} \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{m,n} \end{vmatrix}$$

Definizione A.3 (Determinante)**.** *Premesso che per una matrice di ordine* 1 *il determinante coincide con l'unico elemento, data una matrice quadrata A di ordine n e considerata una sua riga o colonna qualunque, il* determinante *di A è il numero ottenuto moltiplicando gli elementi della riga o colonna scelta per i rispettivi cofattori e sommando i risultati ottenuti. A partire dalla riga r-esima di A si ha:*

$$|A| = a_{r1}A_{r1} + a_{r2}A_{r2} + \cdots + a_{rn}A_{rn}\,; \tag{A.3}$$

a partire dalla colonna p-esima di A si ha invece:

$$|A| = a_{1p}A_{1p} + a_{2p}A_{2p} + \cdots + a_{np}A_{np}\,. \tag{A.4}$$

In sostanza, per calcolare il determinante una matrice di ordine n si devono calcolare n determinanti di matrici di ordine $n-1$, per calcolare i quali si devono calcolare $n-1$ determinanti di matrici di ordine $n-2$, e così via. Naturalmente si prova che nella definizione data la scelta della riga o colonna è ininfluente ai fini del risultato. È immediato che per matrici di ordine 2 la definizione porge il già noto risultato:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}\,.$$

Esempio A.1. Per calcolare il determinante di

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

scegliendo la prima colonna, si ha

$$\begin{aligned} |A| &= 1\cdot(-1)^{1+1}\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} + 3\cdot(-1)^{2+1}\begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} + (-2)\cdot(-1)^{3+1}\begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = \\ &= 1\cdot(1-2) - 3\cdot(2+1) - 2\cdot(4+1) = -17\,. \end{aligned}$$

A.2. Rango di una matrice 2x3

Si consideri una matrice a due righe e tre colonne

$$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \tag{A.5}$$

e si suoi minori di ordine 2

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad , \quad \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \quad , \quad \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix}.$$

Se almeno uno dei determinanti di questi minori è diverso da zero, allora si dice che la matrice ha rango 2. Se tutti e tre i determinanti di questi minori sono nulli, ma la matrice ha almeno un elemento diverso da zero, si dice che la matrice ha rango 1.

Il concetto di rango si estende a matrici rettangolari di un qualunque numero di righe e colonne, ma il concetto ci interessa solo nel caso particolare di due righe e tre colonne, in relazione al teorema dei massimi e minimi vincolati in tre variabili con due vincoli.

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.
$[a,b]$	$\{\, x \in \mathbb{R} \mid a \leq x \leq b \,\}$.
$]a,b[$	$\{\, x \in \mathbb{R} \mid a < x < b \,\}$.
$]a,b]$	$\{\, x \in \mathbb{R} \mid a < x \leq b \,\}$.
$[a,b[$	$\{\, x \in \mathbb{R} \mid a \leq x < b \,\}$.
$[a,+\infty[$	$\{\, x \in \mathbb{R} \mid x \geq a \,\}$.
$]a,+\infty[$	$\{\, x \in \mathbb{R} \mid x > a \,\}$.
$]-\infty,a]$	$\{\, x \in \mathbb{R} \mid x \leq a \,\}$.

Continua nella pagina successiva

Segue dalla pagina precedente

$]-\infty, a[$	$\{\, x \in \mathbb{R} \mid x < a \,\}$.
$f\colon D \to C\,,\ x \mapsto f(x)$	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: $\mathbf{A}$, $\mathbf{B}$, $\dots$ 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 $\operatorname{Log}(x)$ 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	κ	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