Capitolo 7

Derivate

Abbiamo già introdotto la nozione di derivata parziale e l'abbiamo applicata a problemi di massimo e minimo. Ricordiamo, prima di estendere la teoria della derivazione di funzioni di più variabili, la definizione di derivata parziale. Sia $f : \mathbb{R}^2 \to \mathbb{R}$ sia una funzione di due variabili x, e y e sia (x_0, y_0) un punto nel dominio.

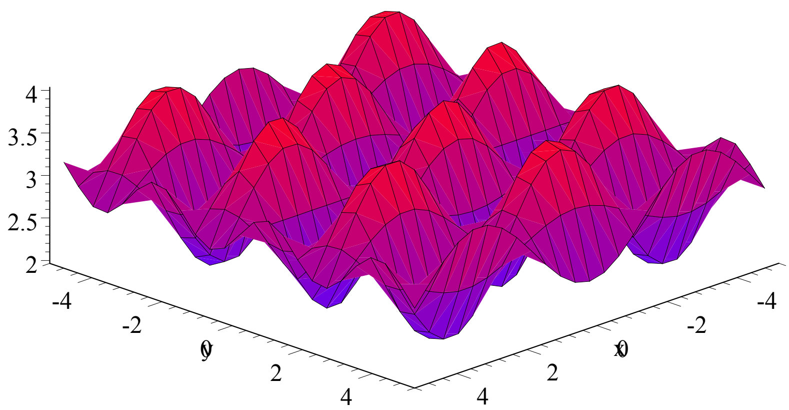
Definizione 7.1 *(Derivata parziale) La derivata parziale di f rispetto alla variabile x nel punto (x_0, y_0) è data da:*

$$\frac{\partial f}{\partial x}(x_0, y_0) = f_x(x_0, y_0) = \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}$$

se il limite esiste.

L'altra derivata parziale $\frac{\partial f}{\partial y}(x_0, y_0)$ $(f_x(x_0, y_0))$ è definita allo stesso modo. Nello stesso modo sono definite le derivate parziali delle funzioni di tre o più variabili.

Esempio 7.2 *Usiamo la funzione esempio $f(x, y) = 3 + \cos x \sin 2y$ per illustrare vari punti di vista ed usi delle derivate parziali.*
Qui di seguito è riportato una parte del suo grafico, una superficie ondosa in tre dimensioni



La superficie $z = 3 + \cos x \, \sin 2y$

Trovare le derivate parziali; qual'è il loro valore all'origine?

Soluzione La definizione ci dice semplicemente che per ottenere la derivata parziale rispetto ad una variabile bisogna differenziare la funzione rispetto ad una variabile considerando l'altra come costante. Si ha allora

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

Nell'origine si ha quindi

$$\frac{\partial f}{\partial x}(0,0) = 0\,, \quad \frac{\partial f}{\partial x}(0,0) = 2\ .$$

■

Fissare una Variabile; Sezionare. Fissare $y = y_{0,}$ come si fa per trovare $f_x(x_0, y_0)$, può essere pensato, dal punto di vista geometrico, come l'intersezione della superficie col piano $y = y_0$. L'intersezione della superficie col piano da luogo ad una curva; $f_x(x_0, y_0)$ rappresenta il coefficiente angolare alla curva nel punto $x = x_0$.

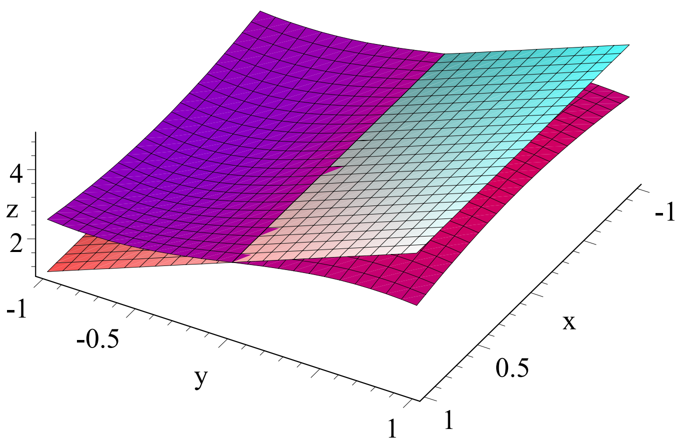
Per esempio, le derivate parziali $\dfrac{\partial f}{\partial x}(0,0) = 0\,, \ \dfrac{\partial f}{\partial x}(0,0) = 2$ della funzione di cui sopra, possono essere viste nel seguente modo: sezionare la superficie $z = 3 + \cos x \sin 2y$ col il piano $y = 0$ produce la superficie $z = 3$ che è un piano orizzontale e quindi la derivata in $x = 0$ è ovviamente zero. In modo simile, sezionando la superficie col piano $x = 0$ produce la curva $z = 3 + \sin 2y$ che ha derivata 2 in $y = 0$.

Approssimazione lineare. Le derivate parziali f_x, f_y ci dicono come varia la funzione relativamente alle direzioni determinare dalle direzioni positive degli assi x e y.
Usando le derivate parziali, come abbiamo già visto, possiamo scrivere l'equazione del piano tangente alla superficie grafico della funzione $z = f(x, y)$ nel punto (x_0, y_0, z_0) nella forma

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

Il grafico qui sotto rappresenta il piano tangente alla superficie $z = f(x, y) = 3 + \cos x \sin 2y$ nel punto $(0, 0, 3)$ (insieme con la superficie stessa).

$z = 3 + 2y$



Una superficie e l'approssimazione lineare

Le due superfici concordano molto bene nel punto $(0, 0, 3)$ ma non necessariamente negli altri.

7.0.1 Punti Stazionari, Massimi e Minimi

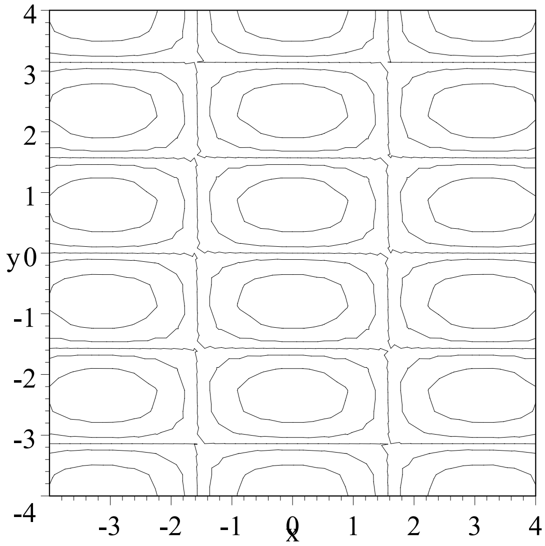
Un punto (x_0, y_0) appartenente al dominio di f è detto **punto stazionario** se le derivate di f sono entrambe nulle nel punto. Questo equivale a dire, dal punto di vista geometrico, che il piano tangente è orizzontale nel punto (x_0, y_0, z_0) od anche che l'approssimazione lineare è costante.

I punti stazionari si trovano quindi risolvendo il sistema

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

Tornando al caso dell'esempio $f(x,y) = 3 + \cos x \sin 2y$, significa risolvere il sistema

$$\begin{cases} f_x(x,y) = -\sin x \sin 2y = 0 \\ f_y(x,y) = 2\cos x \cos 2y = 0 \end{cases}$$



Curve di livello di $f(x,y) = 3 + \cos x \sin 2y$

L'interpretazione delle curve di livello permette di individuare, almeno qualitativamente, i massimi ed i minimi delle funzioni.

Derivate Parziali, Prodotto Vettoriale e Piano Tangente

Cerchiamo di affrontare in modo diverso il problema del piano tangente alla superficie $z = f(x,y)$ nel punto (x_0, y_0, z_0). Consideriamo ancora le curve che si ottengono intersecando la superficie con i piani della forma $x = x_0$ oppure $y = y_0$.

Parametrizziamo, per esempio, la curva intersezione la superficie $z = f(x, y)$ con il piano $y = y_0$ sapendo che $a \leq x \leq b$, $c \leq y \leq d$. I punti di tale curva hanno la forma $(x, y_0, f(x, y_0))$. Se vogliamo scriverla in forma parametrica si ha

$$x = t\,, \quad y = y_0\,, \quad z = f(t, y_0)\ ; \quad a \leq t \leq b.$$

Per $t = x_0$ la curva passa per il punto di coordinate (x_0, y_0, z_0). Il vettore velocità in questo punto ha valore uguale a

$$v(x_0) = (1, 0, f_x(x_0, y_0))$$

(L'ultima coordinata è stata trovata differenziando rispetto a t in $f(t, y_0)$, ma t gioca lo stesso ruolo di x in $f(x, y_0)$ Sappiamo che il vettore velocità è tangente alla curva nel punto e che la curva "appartiene" alla superficie così che il vettore $v(x_0)$ è tangente alla superficie nel punto (x_0, y_0, z_0).
Facciamo ora la stessa operazione fatta sopra, ma intersecando la superficie con il piano $x = x_0$. Ragionando nello stesso modo si ottiene il vettore $(0, 1, f_y(x_0, y_0))$ anch'esso tangente alla superficie nel punto (x_0, y_0, z_0). e non parallelo al primo.

Riepilogando: *Sia* $z = f(x, y)$ *l'equazione di una superficie, e sia* $z_0 = f(x_0, y_0)$. *Allora i due vettori*

$$(1, 0, f_x(x_0, y_0))\ , \quad e \quad (0, 1, f_y(x_0, y_0))$$

sono tangenti alla superficie nel punto (x_0, y_0, z_0).

Avendo ottenuto due vettori non paralleli, entrambi tangenti alla superficie nel punto (x_0, y_0, z_0) possiamo scrivere l'equazione del piano tangente alla superficie nel punto. In forma parametrica vettoriale si ottiene:

$$X(t, s) = (x_0, y_0, z_0) + s(1, 0, f_x(x_0, y_0)) + t(0, 1, f_y(x_0, y_0))\ ;$$

In forma parametrica si ha:

$$x = x_0 + s\,, \quad y = y_0 + t\,, \quad z = z_0 + s f_x(x_0, y_0) + t f_y(x_0, y_0)\ .$$

Si può scrivere l'equazione del piano nella solita forma di equazione scalare trovando prima il vettore normale al piano dato da

$$n = (1, 0, f_x(x_0, y_0)) \times (0, 1, f_y(x_0, y_0)) = (-f_x(x_0, y_0), -f_y(x_0, y_0), 1)$$

L'equazione del piano è data allora da:

$$((x, y, z) - (x_0, y_0, z_0)) \cdot (-f_x(x_0, y_0), -f_y(x_0, y_0), 1) = 0\,,$$

cioè

$$(x - x_0) f_x(x_0, y_0) + (y - y_0) f_y(x_0, y_0) = z - z_0$$

che dovrebbe essere nota.

Esempio 7.3 *Trovare il piano tangente alla superficie* $f(x, y) = x^2 - 2y^3$ *relativamente al punto nel piano di coordinate* $(1, 1)$.

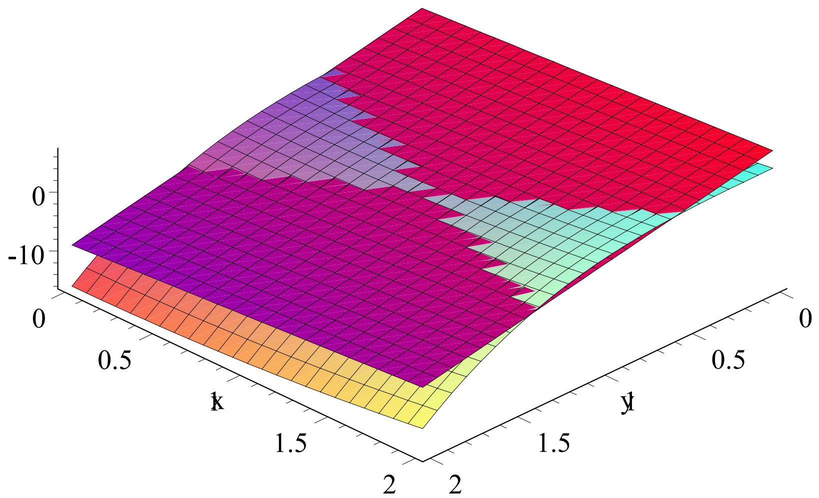
Soluzione Nel punto $(1, 1)$ la funzione vale $f(1, 1) = -1$. Il punto sulla superficie di cui si cerca il piano tangente è quindi $(1, 1, -1)$. Le derivate parziali di f sono $f_x(x, y) = 2x$ e $f_y(x, y) = -6y^2$ da cui $f_x(1, 1) = 2$, $f_y(1, 1) = -6$. Ne consegue che i vettori $(1, 0, 2)$ e $(0, 1, -6)$ sono tangenti alla superficie nel punto $(1, 1, -1)$.
Il vettore normale è allora dato da $n = (-2, 6, 1)$. Quindi l'equazione del piano tangente è data da

$$z = -2x + 4y - 3$$

In forma vettoriale da

$$(x, y, z) = (1, 1, -1) + s(1, 0, 2) + t(0, 1, -4)$$

■



La superficie $z = x^2 - 2y^3$ ed il suo piano tangente in $(1, 1, -1)$

7.0.2 Esercizi

1. In questo paragrafo è stato studiato il piano tangente alla superficie $z = 3 + \cos x \sin 2y$.

   (a) Trovare la funzione approssimazione lineare $L(x, y)$ di f nel punto $(0, 0)$ (il suo grafico è il piano tangente);

   (b) Scrivere l'equazione del piano in forma vettoriale (trovare prima il vettore normale).

2. Usando il grafico di contorno di $z = 3 + \cos x \sin 2y$ pagina 224:

   (a) Spiegare perché f ha un punto stazionario in ognuno dei punti di coordinate $(\pm\pi/2, k\pi/2)$ dove k è un intero;

   (b) Come appaiono questi punti nel grafico di contorno di f ?

   (c) Trovare di che natura sono i punti $(\pm\pi/2, k\pi/2)$ al variare di k .Cercare di capirlo valutando l'andamento del grafico della funzione dal grafico di contorno.

3. Per ognuna delle funzioni trovare il piano tangente alla superficie $z = f(x, y)$ nel punto dato, sia in forma scalare che in forma parametrica. Se possibile usare il computer per disegnare la superficie ed il piano tangente nello stesso punto.

   (a) $z = x^2 + y^2$, nel punto $(1, 2, 5)$;

   (b) $z = x^2 - y^2$, nel punto $(1, 2, -3)$;

   (c) $z = \sin(xy)$, nel punto $(1, \pi/2, 1)$;

   (d) $z = 1 + \cos xy$, nel punto $(0, 0, 2)$

4. Possono, i vettori $(-f_x(x_0, y_0), 0, 1)$, $(0, -f_y(x_0, y_0), 1)$ essere mai paralleli? Motivare la risposta.

5. Trovare l'equazione della retta ortogonale alle superfici date nei punti assegnati. Se possibile usare il computer per disegnare la superficie e la retta ortogonale.

   (a) $z = x^2 - y^3$ nel punto $(1, 1, 0)$;

   (b) $z = \sin 2x \cos y$ nel punto $(\pi/4, 0, 1)$;

   (c) $z = \log(x + 2y)$ nel punto $(1/2, 1/4, 0)$;

   (d) $z = x^2 + y^2$ nel punto $(-1, 1, 2)$.

7.1 Il Gradiente

Abbiamo visto che una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ può ammettere , in un punto (x_0, y_0) del dominio, derivate parziali $f_x(x_0, y_0)$ e $f_y(x_0, y_0)$. Diventa naturale considerare il vettore di $\mathbb{R}^2$ che ha come componente queste due derivate

$$\nabla f(x_0, y_0) = (f_x(x_0, y_0), f_y(x_0, y_0))$$

chiamato il **gradiente** di f in (x_0, y_0) .

Nel paragrafo precedente, abbiamo visto ed interpretato il significato delle derivate parziali e quindi delle componenti del vettore gradiente. Vogliamo ora capire come interpretare il *vettore gradiente,* capire cosa il suo modulo e la sua direzione ci possono dire; in che modo collegarlo agli oggetti matematici che abbiamo già studiato.

Definizione 7.4 *(Gradiente di una funzione in un punto). Sia $f(x, y)$ una funzione di due variabili e (x_0, y_0) un punto del dominio. Assumiamo che entrambe le derivate parziali esistano in (x_0, y_0) . Il gradiente di f nel punto (x_0, y_0) è il vettore del piano*

$$\nabla f(x_0, y_0) = (f_x(x_0, y_0), f_y(x_0, y_0)) \ .$$

Per una funzione di tre variabili $g(x, y, z)$ il gradiente è il vettore tridimensionale

$$\nabla g(x_0, y_0, z_0) = (g_x(x_0, y_0), g_y(x_0, y_0), g_z(x_0, y_0)) \ .$$

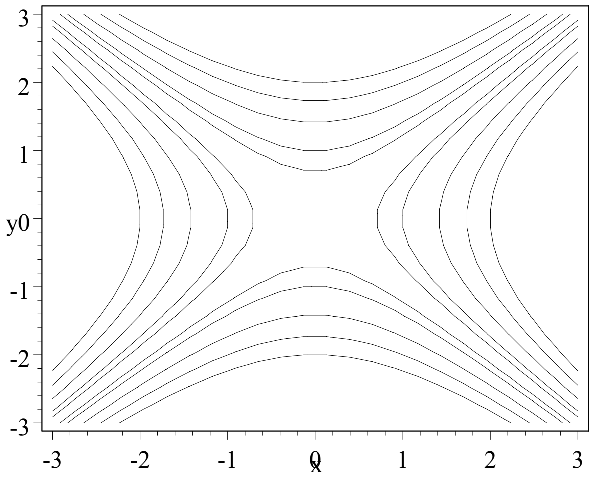
Nota. *Ricordarsi che il grafico di una funzione $f : R^2 \to R$ è un oggetto bidimensionale (superficie) in R^3 e che $\nabla f(X_0)$ è un vettore della stessa dimensione della dimensione del dominio di f. Di fatto il vettore gradiente giace "naturalmente" nel dominio di f . Spesso tracceremo $\nabla f(X_0)$ applicato nel punto X_0.*

Il gradiente come funzione. La derivata $f'(x)$, di una funzione di una variabile $f(x)$, è anch'essa una funzione di una variabile. Nel caso di una funzione di due variabili $f(x, y)$ si ha invece che il vettore gradiente, visto come funzione della coppia (x, y), $(x, y) \to \nabla f(x, y)$ è una funzione da $\mathbb{R}^2 \to \mathbb{R}^2$. Si ha così che nonostante f sia una funzione scalare, la funzione ∇f è una funzione vettoriale che a volte chiameremo **campo vettoriale.**

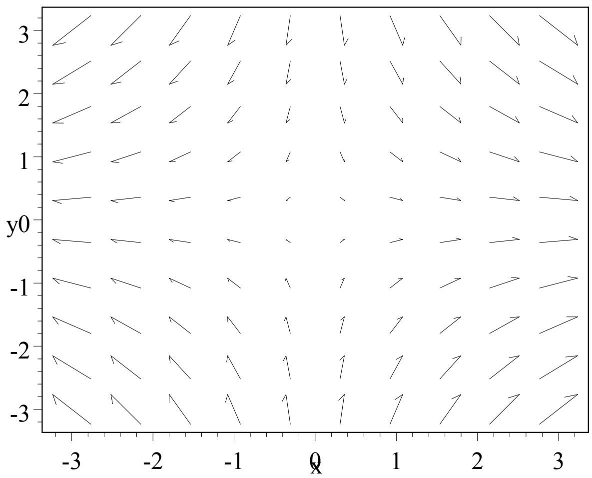
Il calcolo del gradiente di $f : \mathbb{R}^2 \to \mathbb{R}$, o anche $f : \mathbb{R}^n \to \mathbb{R}$, $n > 2$ è semplice una volta che si sappiano fare le derivate delle funzioni di una variabile.

Esempio 7.5 *Sia $f(x,y) = x^2 - y^2$. Calcolare il gradiente di f. Cosa ci dicono modulo e direzione del vettore gradiente?*

Soluzione. Si ha che $\nabla f(x,y) = (2x, -2y)$. Per esempio, $\nabla f(0,0) = (0,0)$; $\nabla f(1,1) = (2,-2)$; $\nabla f(-1,-1) = (-2,2)$; $\nabla f(2,4) = (4,-8)$



Mappa di contorno di $f(x,y) = x^2 - y^2$



Campo vettoriale gradiente di $f(x,y) = x^2 - y^2$

Guardando le due figure si osserva che la funzione gradiente assegna un vettore ad ogni punto del dominio. Per ovvie ragioni la figura ne mostra

solo alcune. Da notare che, per esempio, lungo l'asse x, $f(x,y) = x^2$, così che $f(x,y)$ cresce dapprima lentamente, poi sempre più velocemente, allontanandosi dall'origine sia verso destra che verso sinistra. Queste informazioni appaiono in figura. Lungo l'asse x il vettore gradiente punta in direzione opposta all'origine. Lungo l'asse y accade esattamente l'opposto.
In un punto stazionario di f il gradiente è nullo (come accade nell'origine per l'esempio dato). Le figure ci fanno anche capire che l'origine è un *punto di sella.*

■

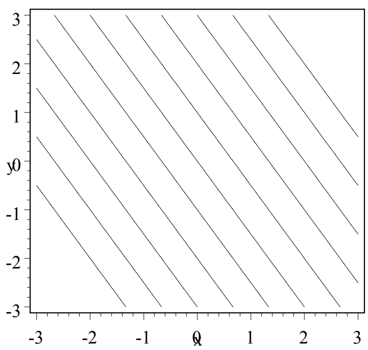
L'osservazione più importante da fare è comunque:

In ogni punto (x_0, y_0) *del dominio, il vettore gradiente è perpendicolare alla curva di livello passante per* (x_0, y_0).

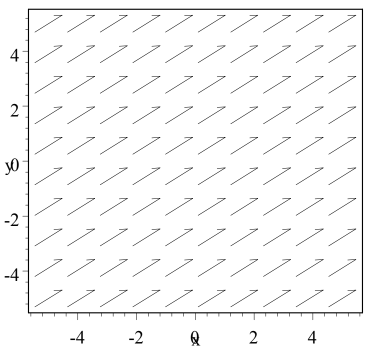
Vedremo più avanti una dimostrazione rigorosa di questo fatto. Intuitivamente esso ci dice che le curve di livello sono perpendicolari alle direzioni di massima pendenza.

Il Gradiente di una Funzione Lineare

Una funzione lineare ha la forma $L(x,y) = ax + by + c$ e quindi ha derivate parziali costanti. Il vettore gradiente è dato perciò da $\nabla L(x,y) = (a,b)$ per tutte le coppie (x,y).
Vediamo il grafico, per esempio, della funzione $L(x,y) = 3x + 2y$



Mappa di contorno di $L(x,y)$



Campo vettoriale $(3,2)$

Anche in questo caso, come si vede bene osservando i due grafici, il vettore gradiente $(3,2)$ appare essere perpendicolare alle curve di livello $3x + 2y = k$ Come è ben noto, il coefficiente angolare di questa retta è $-3/2$ e quindi il vettore $(2,-3)$ è un vettore tangente che è ortogonale al vettore $(3,2)$, come affermato.

Funzioni lineari in tre variabili: gradienti e superfici di livello

Una funzione lineare in tre variabili è data da $L(x,y,z) = ax+by+cz+d$. Il vettore gradiente è dato da $\nabla L = (a,b,c)$, vettore costante tridimensionale. Consideriamo l'insieme $\{(x,y,z) \in \mathbb{R}^3 : L(x,y,z) = w_0\}$, cioè la superficie di livello $L(x,y,z) = w_0$. Si ha, $ax+by+cz+d = w_0$ che rappresenta il piano di equazione $ax+by+cz = w_0 - d$. Come è noto dalla geometria elementare il vettore (a,b,c) è perpendicolare al piano stesso. Questo mostra, come nel caso di due variabili, che il vettore gradiente nel punto di coordinate (x_0,y_0,z_0) è ortogonale alla linea di livello per lo stesso punto.

7.1.1 Gradiente ed Approssimazione Lineare

Sia $f : \mathcal{D} \subset \mathbb{R}^2 \to R$ una funzione differenziabile, (x_0,y_0) un punto nel dominio $\mathcal{D}$. Abbiamo precedentemente definito l'approssimazione lineare di f nell'intorno di (x_0,y_0) come la funzione definita da

$$L(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0).$$

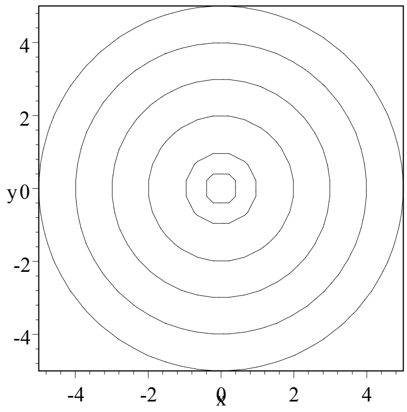
Possiamo scrivere questa espressione nella forma vettoriale,

$$L(X) = f(X_0) + \nabla f(X_0) \cdot (X - X_0)$$

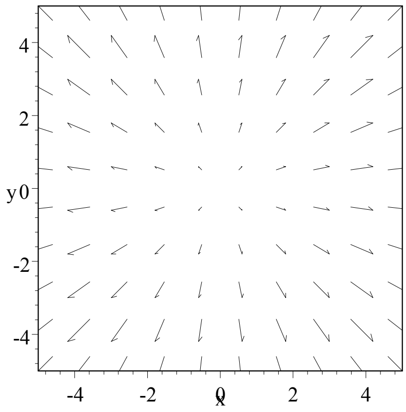
essendo $X_0 = (x_0,y_0)$, $X = (x,y)$.

(Da notare che, in forma vettoriale, la formula in tre variabili è scritta nella identica forma).

Come abbiamo visto e detto ripetutamente, ogni funzione di più variabili che sia differenziabile può essere approssimata, in ogni punto del dominio X_0 con una funzione lineare. Prima di proseguire osserviamo ancora, per esempio, il caso di $f(x,y) = x^2 + y^2$



Mappa di contorno di $x^2 + y^2$



Il campo gradiente $(2x, 2y)$.

Come negli altri casi, il vettore gradiente è perpendicolare alle curve di livello.

La proprietà di perpendicolarità del gradiente è molto utile quando si voglia trovare il piano tangente ad un punto di una superficie in $\mathbb{R}^3$.

Esempio 7.6 *Trovare l'equazione del piano tangente alla sfera* $x^2+y^2+z^2 = 14$ *nel punto di coordinate* $(1, 2, 3)$

Soluzione. Possiamo pensare alla sfera come la superficie di livello 14 della funzione $f(x, y, z) = x^2 + y^2 + z^2$. Il gradiente di f è $\nabla f(x, y, z) = (2x, 2y, 2z)$, in particolare si ha $\nabla f(1, 2, 3) = (1, 4, 6)$. Questo è anche il vettore normale al piano cercato che ha perciò equazione $(x - 1)+2(y - 2)+ 6(y - 3) = 0$, o anche $x + 2y + 3z = 23$.

■

E' lecito domandarsi quale errore si commette, sostituendo, localmente, $f(X)$ con la sua approssimazione lineare $L(X)$. Quello che è possibile affermare al nostro livello di conoscenza è la seguente affermazione

Teorema 7.7 *Indichiamo con* $\varepsilon(X - X_0)$ *la differenza tra* $f(X)$ *ed* $L(X)$ *in un intorno di* X_0. *Si ha*

$$\lim_{X \to X_0} \frac{|\varepsilon(X - X_0)|}{\|X - X_0\|} = 0$$

Derivate Direzionali

Le derivate direzionali ci dicono come varia una funzione quando la variabile indipendente varia lungo le direzioni degli assi coordinati. Ma gli assi coordinati, oltre ad essere usati come elemento di orientamento del piano non sono direzioni privilegiate rispetto alle altre ed è quindi ovvio chiedersi come si individua la variazione della funzioni lungo direzioni che non siano quelle degli assi coordinati. La definizione di **Derivata direzionale** risponde alla domanda che ci siamo appena fatti. Scriviamo la definizione in forma vettoriale che è una scrittura unificante rispetto alla dimensione dello spazio

Definizione 7.8 ***(Derivata direzionale)****. Sia f una funzione, X_0 un punto nel dominio, u un vettore unitario. La derivata di f in X_0 nella direzione determinata da u è data da*

$$D_u(X_0) = \lim_{h\to 0} \frac{f(X_0 + hu) - f(X_0)}{h}$$

se tale limite esiste.

Notare che se $u = i$ (i versore dell'asse x) la definizione sopra è quella di $f_x(X_0)$. Di più

$$D_i(X_0) = f_x(X_0)\ ,\ D_j(X_0) = f_y(X_0)\ ,\ D_k(X_0) = f_z(X_0)$$

avendo indicato con j il versore dell'asse y e con k quello dell'asse z.

Sebbene la definizione soddisfi il nostro bisogno di descrivere la variazione di una funzione lungo una direzione qualsiasi, non è semplice da usarsi.
Si pone allora il problema di come calcolare le derivate direzionali. Non volendo entrare in dettagli tecnici assumiamo che la funzione f abbia derivate continue in X_0.

Proposizione 7.9 *Siano f, X_0, u, come sopra. Supponiamo che f ammetta derivate parziali continue in X_0, allora:*

$$D_u(X_0) = \nabla f(X_0) \cdot u$$

Dimostrazione. Per provare che quanto abbiamo detto è vero vediamo dapprima cosa accade se f è una funzione lineare, cioè se $f(x, y, z) = ax + by + cz + d = \nabla f \cdot (x, y, z) + d$. In questo caso si ha

$$\begin{aligned} \frac{f(X_0 + hu) - f(X_0)}{h} &= \frac{\nabla f \cdot (X_0 + hu) - \nabla f \cdot (X_0)}{h} \\ &= \frac{\nabla f \cdot hu}{h} = \frac{h\nabla f \cdot u}{h} \\ &= \nabla f \cdot u \end{aligned}$$

Questo dimostra che ciò che abbiamo affermato vale nel caso di funzioni lineari.
Nel caso di funzioni non lineari, ricordando la proprietà della differenza tra valore della funzione e approssimazione lineare, la dimostrazione la si ottiene scrivendo

$$\begin{aligned}
\frac{f(X_0+hu)-f(X_0)}{h}-\nabla f\cdot u &= \frac{f(X_0+hu)-f(X_0)-h\nabla f\cdot u}{h}\\
&= \frac{f(X_0+hu)-f(X_0)-\nabla f\cdot hu}{h}\\
&= \frac{\varepsilon(hu)}{h}\\
e,\ \textit{come noto}\quad 0 &= \lim_{h\to 0}\frac{\varepsilon(hu)}{h}
\end{aligned}$$

■

Interpretazione del vettore gradiente. Sia u un vettore unitario. Ricordando le proprietà del prodotto scalare si ha

$$D_u(X_0)=\nabla f(X_0)\cdot u=|\nabla f(X_0)|\cos\theta$$

dove θ è l'angolo tra $\nabla f(X_0)$ e u. In particolare

$$D_u(X_0)\leq|\nabla f(X_0)|\ ;$$

con l'uguaglianza che vale solo se il vettore u è parallelo a $\nabla f(X_0)$. Ne seguono due importanti proprietà che vale la pena evidenziare:

(a) Il vettore $\nabla f(X_0)$ punta nella direzione di massima crescita di f rispetto al valore $f(X_0)$;

(b) Il modulo $|\nabla f(X_0)|$ è la massima velocità di cambiamento di f.

Esempio 7.10 *Trovare la derivata direzionale di $f(x,y)=x^2+y^2$ in varie direzioni nel punto $(2,1)$. In quale direzione f cresce maggiormente? In quale decresce maggiormente?*

Soluzione Il gradiente di f in $(2,1)$ è dato da $\nabla f(2,1)=(4,2)$. Lungo questa direzione, quindi, f cresce ad una velocità di $|(4,2)|=\sqrt{20}$ unità di uscita per unità di ingresso. Nella direzione opposta (la direzione indicata cioè dal vettore $(-4,-2)$, $\cos\theta=-1$) la derivata direzionale vale $-\sqrt{20}$. Nella direzione data da $\left(1/\sqrt{2},1/\sqrt{2}\right)$ la derivata direzionale vale:

$$D_u(2,1)=(4,2)\cdot\left(1/\sqrt{2},1/\sqrt{2}\right)=6/\sqrt{2}.$$

■

7.1.2 Esercizi

Nota. *Maple* o altri pacchetti software possono essere usati per gli esercizi. Riportiamo alcuni comandi utili di *Maple.*

```
>with(plots);with(linalg);
>gradplot(x^2+y^2,x=-5..5,y=-5..5,grid=[10,10],scaling=constrained);
>grad(x^2+y^2, [x,y]);
>fieldplot([2*x,2*y],x=-5..5,y=-5..5,grid=[10,10],scaling=constrained);
```

(porre la griglia (grid) come negli esempi determina il numero di frecce tracciate. Imporre che la scala sia la stessa sui due assi (`scaling=constrained`) mantiene la perpendicolarità dei vettori).

1. Tracciare (a mano) la mappa gradiente nel quadrato $[0, 2] \times [0, 2]$. In ogni punto a coordinate intere calcolare e tracciare il vettore gradiente. Tracciate inoltre le curve di livello passanti per tali punti.

   (a) $f(x, y) = (x + y)$;

   (b) $f(x, y) = (x^2 - y)/2$;

   (c) $f(x, y) = (y - x^2)/2$;

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

2. Trovare il gradiente delle funzioni nei punti indicati. Tracciare anche le curve di livello per i punti assegnati e mostrare che il vettore gradiente è perpendicolare alle curve nei punti assegnati.

   (a) $f(x, y) = x + y$; $(x_0, y_0) = (2, 2)$;

   (b) $f(x, y) = x^2 + y$; $(x_0, y_0) = (1, 2)$;

   (c) $f(x, y) = x - y^2$; $(x_0, y_0) = (2, 1)$;

   (d) $f(x, y) = x^2 + y^2$; $(x_0, y_0) = (1, 1)$.

3. Sia $f(x, y) = ax + by + c$ una funzione lineare e (x_0, y_0) un punto nel dominio.

   (a) Scrivere l'equazione della linea di livello per (x_0, y_0) .

   (b) Mostrare che la linea di livello è perpendicolare al gradiente.

4. Trovare il piano tangente alle superfici date nei punti assegnati. Provare ad usare il software per controllare il risultato, disegnando superficie e piano tangente.

(a) $x^2 + y^2 + 2z^2 = 4$ nel punto $(1, 1, 2)$;

(b) $x^2 + y^2 + 2z^2 = 4$ nel punto $(0, 2, 0)$;

(c) $x^2 - y^2 + z$ nel punto $(1, 1, 2)$;

(d) $z = x^2 + y^2$ nel punto $(2, 1, 5)$.

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

   (a) Trovare la derivata direzionale di f nel punto $(2, 1)$ in ognuna delle direzioni $\theta = k\frac{\pi}{4}, k = 0, \cdots, 7$. Esprimere le soluzioni anche in forma decimale.

   (b) Disegnare i risultati trovati come funzione dell'angolo θ . Qual'è la forma del grafico?

   (c) Dato il punto $(2, 0)$, trovare la direzione (o le direzioni) nelle quali f cresce con un tasso di 3 unità di uscita per unità di ingresso.

6. Data la funzione $f(x, y) = x + y + \sin y$ trovare e disegnare gradiente e linee di livello nel quadrato $[-2, 3] \times [-2, 3]$. Confrontare con il grafico che si ottiene con il software.

7.2 Linearità Locale: Teoria della Derivazione

Ci siamo limitati, fino ad ora, ad operare con le derivate parziali tenendo la teoria al minor livello compatibile con le necessità del calcolo. In questa sezione svilupperemo la teoria delle funzioni di più variabili, cercando di precisare la definizione di differenziabilità e quindi la nozione di approssimazione lineare locale. Non intendiamo ovviamente sviluppare la teoria in tutta la sua completezza, ma focalizzare meglio alcune questioni teoriche (e loro ricadute). Nel farlo ci limiteremo (per ragioni di semplicità) alle sole funzioni di due variabili.

7.2.1 Approssimazione Lineare e Funzioni Differenziabili

Sia $(x, y) \to f(x, y)$ una funzione e (x_0, y_0) un punto del dominio. Abbiamo definito come approssimazione lineare la funzione lineare

$$\begin{aligned} L(x, y) &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ &= f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (X - X_0) \end{aligned}$$

Per scrivere L tutto ciò che abbiamo bisogno di conoscere sono il valore della funzione e le sue derivate parziali nel punto (x_0, y_0). In queste condizioni L ed f hanno lo stesso valore e le stesse derivate parziali in (x_0, y_0). Per questo ci aspettiamo che L approssimi bene f non solo in (x_0, y_0) anche in un intorno del punto. Gli esempi visti fino ad ora tutto sembrava filare liscio. Esaminiamo però quest'altro esempio

Esempio 7.11 *Sia $f(x, y)$ la funzione definita nel seguente modo*

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{se } (x, y) = (0, 0) \end{cases}$$

Trovare l'approssimazione lineare in $(0, 0)$ e verificare se essa approssima "bene" f in un intorno dell'origine.

Soluzione. Si vede immediatamente che $f(x, 0) = f(0, y) = 0$ che implica $f_x(0, 0) = f_y(0, 0) = 0$ Poiché $f(0, 0) = 0$ ne segue che l'approssimazione lineare in $(0, 0)$ è data da $L(x, y) \equiv 0$.
D'altra parte se (x, y) giace sulla retta $x = y$ si ha che $f(x, x) = 1/2$, oppure se è $x = -y$ è $f(x, -x) = -1/2$.
Si vede allora immediatamente che, mentre in qualsiasi intorno dell'origine

L vale zero, la funzione si comporta molto irregolarmente in ogni intorno dell'origine, assumendo valori costanti, diversi tra loro, quando ci si muove verso l'origine per segmenti.
Questo avviene perché la funzione **non è continua** in $(0,0)$. Non vi allora alcuna possibilità di approssimare f localmente intorno all'origine con una funzione lineare.

■

L'esempio precedente ci mostra come, per una funzione di più variabili, la continuità non sia necessaria per l'esistenza delle derivate parziali. Questo contrasta con quanto studiato per le funzioni di una variabile dove l'esistenza della derivata in un punto implicava la continuità della funzione nel punto stesso. Infatti se $f(x)$ è una funzione di una variabile, la definizione di derivata è

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

se tale limite esiste. E' ovvio che per funzioni di più variabili non ha senso considerare il rapporto

$$\frac{f(X) - f(X_0)}{X - X_0}$$

essendo (nel caso di due variabili) $X = (x, y)$, $X_0 = (x_0, y_0)$.

Possiamo però, partendo dalla definizione di derivata, scrivere un limite equivalente:

$$\begin{aligned} f'(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \Longleftrightarrow \\ &\Longleftrightarrow \lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0 \end{aligned}$$

Quest'ultima condizione è equivalente ad affermare l'esistenza di un numero $f'(x_0)$ per il quale vale

$$\lim_{x \to x_0} \frac{f(x) - (f(x_0) - f'(x_0)(x - x_0))}{|x - x_0|} = 0 \tag{7.1}$$

(Il valore assoluto al denominatore non modifica nulla rispetto alla definizione precedente, ma è essenziale nella definizione che daremo per le funzioni di più variabili).

Notare invece che il numeratore che abbiamo costruito è del tipo $f(x) - L(x)$ dove $L(x) = f(x_0) - f'(x_0)(x - x_0)$ è una funzione *lineare*.

Possiamo leggere il quoziente 7.1 nel seguente modo: quando $x \to x_0$ la differenza $f(x) - L(x)$ tende a zero *più rapidamente* del denominatore $|x - x_0|$ (in altre parole è come dire che $L(x)$ approssima $f(x)$ meglio di quanto x non faccia con x_0). Questa è la condizione chiave per la definizione di differenziabilità per funzioni di più variabili.

Definizione 7.12 *Sia $f(x,y)$ una funzione e $X_0 = (x_0, y_0)$ un punto del suo dominio. Sia*

$$\begin{aligned} L(x,y) &= f(x_0,y_0) + f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) \\ &= f(X_0) + \nabla f(X_0) \cdot (X - X_0) = L(X) \end{aligned}$$

l'approssimazione lineare di f in (x_0, y_0). Se

$$\lim_{X \to X_0} \frac{f(X) - L(X)}{|X - X_0|} = 0 ,$$

*diremo che f è **differenziabile in** x_0 ed il vettore $\nabla f(X_0)$ è **il gradiente** di f in X_0.*

Osservazione: Qui, come nel caso di una variabile, il limite garantisce che l'approssimazione lineare $L(X)$ approssima bene la funzione $f(X)$.
Non solo questo, ma si richiede anche che la quantità dentro l'operazione di limite tenda a zero *qualunque* sia il modo con cui X tende ad X_0.

La domanda che si pone, visto l'esempio precedente è la seguente: *sotto quali condizioni l'esistenza delle derivate parziali implica la differenziabilità di una funzione nel punto X_0?*
Diamo qui di seguito, senza dimostrazione ,una condizione sufficiente facilmente verificabile.

Teorema 7.13 *Se le derivate parziali f_x e f_y sono continua in X_0 allora la funzione è differenziabile in X_0 e $L(X) = f(X_0) + \nabla f(X_0) \cdot (X - X_0)$.*

Nota 7.14 ***Una nota sulla terminologia usata.*** *Abbiamo chiamato funzione lineare una funzione del tipo $L(x,y,z) = Ax + By + Cz + D$ Il nome è ragionevole perché grafici e curve di livello di funzioni lineari sono rette, piani o altri oggetti "piatti".*
In realtà va ricordato che la parola lineare in geometria viene usata per indicare trasformazioni con la proprietà : $\mathcal{L}(ax + by) = a\mathcal{L}(x) + b\mathcal{L}(y)$. Le due nozioni coincidono perciò solo nel caso che $D = 0$. Nei testi di geometria, nel caso $D \neq 0$ una funzione come L viene chiamata **affine**.

7.2.2 Esercizi

Come prima, *Maple* o altro software può essere utile in alcuni esercizi.

1. Consideriamo di nuovo la funzione

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

   (a) Verificare che f è costante lungo le linee della forma $y = mx$ $(x \neq 0)$;

   (b) Disegnare le linee di livello $f(x,y) = A$, $A = \pm 1/2,\ \pm 2/5,\ \pm 3/10$. Provare poi a vedere come si comportano i pacchetti software in questo caso. Quali sono i problemi che si presentano?

   (c) Sia $u = \left(1/\sqrt{2}, 1/\sqrt{2}\right)$. Esiste la derivata direzionale $D_u f(0,0)$? Perché o perché no?

   (d) Esiste una direzione u per la quale esiste $D_u f(0,0)$?

   (e) Provare a disegnare col software la superficie $z = \dfrac{xy}{x^2+y^2}$ nell'intorno dell'origine. Analizzare il risultato.

2. Considerare la funzione

$$g(x,y) = \begin{cases} \dfrac{x^2}{x^2+y^2} & se\ (x,y) \neq (0,0) \\ 0 & se\ (x,y) = (0,0) \end{cases}$$

   (a) Esistono le derivate parziali $g_x(0,0)$, $g_y(0,0)$? Perché o perché no?

   (b) Lungo quali curve $g(x,y)$ è costante? (esclusa l'origine). Qual'è il valore di $g(x,y)$ lungo le curve $y = mx$?

   (c) Tracciare le linee di livello $g(x,y) = A$, $A = \pm 1/2,\ \pm 2/5,\ \pm 3/10$. Provare a vedere come si comportano i pacchetti software in questo caso. Quali sono i problemi che si presentano?

   (d) Sia $u = \left(1/\sqrt{2}, 1/\sqrt{2}\right)$. Esiste la derivata direzionale $D_u g(0,0)$? Perché o perché no?

   (e) Esiste una direzione u per la quale esiste $D_u g(0,0)$?

   (f) Provare a far disegnare dal software la superficie $z = \dfrac{x^2}{x^2+y^2}$ nell'intorno dell'origine.

3. Per ognuna delle funzioni sotto trovare l'approssimazione lineare $L(x, y)$ nel punto $(0, 0)$. Quindi valutare il

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - L(x,y)}{\sqrt{x^2+y^2}}$$

Inoltre far disegnare dal software la quantità $\dfrac{f(x,y) - L(x,y)}{\sqrt{x^2+y^2}}$ nell'intorno dell'origine.

   (a) $f(x, y) = \sin(x + y)$;
   (b) $f(x, y) = \sin(xy)$;
   (c) $f(x, y) = x^2 + y$;
   (d) $f(x, y) = x^2 + y^2$.

7.2.3 Derivazione di Funzioni Composte

La regola di derivazione di funzioni composte, così come la regola di derivazione del prodotto sono tutti risultati di tipo "combinatorio". Ci dicono come trovare le derivate di funzioni che si ottengono componendo tra loro funzioni di cui si conoscono le derivate. Le derivate che otteniamo sono anche loro "combinazioni" delle derivate delle funzioni componenti.
La combinazione di funzioni e la regola di derivazione, nel caso di funzioni di una variabile è relativamente semplice. Componendo due funzioni di una sola variabile f e g si ottiene ancora una funzione di una variabile $f \circ g$ definita da $f \circ g(x) = f(g(x))$, se $f(x) = x^2$ e $g(x) = e^x$, allora $f \circ g(x) = (e^x)^2 = e^{2x}$.
Per funzioni di più variabili si opera nello stesso modo, l'uscita di una funzione viene usata come ingresso per l'altra. Nel calcolo di più variabili, tuttavia sia l'ingresso che l'uscita possono essere sia scalari che vettori, così che risulta importante tener conto di qual'è la dimensione del dominio e dell'immagine. La notazione $\to$ che ci dice tra quali spazi operano le funzioni può essere d'aiuto.

Esempio 7.15 *Consideriamo le funzioni* $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R}^2 \to \mathbb{R}$, $h : \mathbb{R} \to \mathbb{R}^2$ *definite da*

$$f(t) = t^2, \quad g(x,y) = x^2 + y^2, \quad h(t) = (\cos t, \sin t).$$

Quali composizioni hanno senso?

Soluzione La notazione $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R}^2 \to \mathbb{R}$ ci dice immediatamente che la composizione $g \circ f$ non ha senso. L'uscita di f è uno scalare mentre il dominio di g è un vettore del piano. Ha invece senso la composizione $f \circ g$, infatti l'uscita di g è uno scalare a cui si può applicare f. Simbolicamente abbiamo $f \circ g : \mathbb{R}^2 \to \mathbb{R} \to \mathbb{R}$ o più semplicemente $f \circ g : \mathbb{R}^2 \to \mathbb{R}$ definita da

$$f \circ g(x,y) = f\left(x^2 + y^2\right) = \left(x^2 + y^2\right)^3 .$$

Analogamente, se consideriamo le due funzioni $g : \mathbb{R}^2 \to \mathbb{R}$ e $h : \mathbb{R} \to \mathbb{R}^2$ vediamo che sono possibili due differenti tipi di composizioni, le quali sono strutturalmente profondamente diverse tra di loro. Si ha $h \circ g : \mathbb{R}^2 \to \mathbb{R}^2$ e $g \circ h = \mathbb{R} \to \mathbb{R}$ definite da

$$h \circ g(x,y) = h\left(x^2 + y^2\right) = \left(\cos\left(x^2 + y^2\right), \sin\left(x^2 + y^2\right)\right)$$

e

$$g \circ h(t) = g(\cos t, \sin t) = \cos^2 t + \sin^2 t = 1.$$

Si può anche definire la composizione $h \circ f : \mathbb{R} \to \mathbb{R}^2$

$$h \circ f\,(t) = h\left(t^3\right) = \left(\cos t^3, \sin t^3\right).$$

■

Prima di arrivare a definire e calcolare la derivazione della composizione per funzioni di più variabili ricordiamo quello che accade per funzioni di una sola variabile.

Derivazione della Composizione per Funzioni di una Variabile

Proposizione 7.16 *Siano f e g funzioni differenziabili con a elemento del dominio di g. Allora*

$$(f \circ g)'\,(a) = f'\,(g\,(a)) \cdot g'\,(a)\ .$$

Ci sono altre notazioni per dire le stesse cose. Se scriviamo $y = f\,(u)$ e $u = g\,(x)$, allora la regola di derivazione della composizione può essere scritta come

$$\frac{dy}{dx}\,(a) = \frac{dy}{du}\,(g\,(a)) \cdot \frac{du}{dx}\,(a)\ .$$

Qualunque sia la forma simbolica con cui la scriviamo l'idea chiave è che: *la derivazione della composizione* $f \circ g$ è il *prodotto* delle derivate f' e g'.

La regola di derivazione è semplice. Da notare tuttavia che le due derivate del prodotto $f'\,(g\,(a))$ e $g'\,(a)$ sono valutate in punti diversi; g in $x = a$ ed f in $g\,(a)$.
Il seguente diagramma

$$a \xrightarrow{g} g\,(a) \xrightarrow{f} f\,(g\,(a))$$

mostra perché queste scelte hanno senso. Le due derivate sono valutate nei corrispondenti punti del dominio.

Perché la regola di derivazione funziona? La risposta è semplice e cercheremo di darla in modo che possa essere estesa con immediatezza al caso di più variabili.

- la regola funziona per le funzioni lineari; sia cioè $f\,(x) = A+Bx,\ g\,(x) = C + Dx\,,\ A, B, C, D$ costanti. In questo caso è $f'\,(x) = B$ e $g\prime\,(x) = D$. La composizione $f \circ g\,(x)$ da luogo a

  $$f \circ g\,(x) = f\,(C + Dx) = A + B\,(C + Dx) = A + BC + BDx\ .$$

  Allora $(f \circ g)\,(x) = BD$ come il prodotto di f' con g'.

- Come abbiamo visto le funzioni differenziabili sono localmente lineari, nel senso che in ogni punto del dominio possono essere localmente approssimate con funzioni lineari. Sia f una funzione differenziabile e indichiamo con L_f l'approssimazione lineare.
  Consideriamo la composizione $f \circ g$; sia a un punto del dominio di g e indichiamo con $b = g(a)$, è
  $$a \overset{g}{\to} b \overset{f}{\to} f(b)$$
  L'approssimazione lineare di g in a è
  $$L_g(x) = g(a) + g'(a)(x-a) = b + g'(a)(x-a) \ ;$$
  notiamo che
  $$L_g(a) = g(a) \text{ e } L_g'(a) = g'(a) \ .$$
  In modo analogo l'approssimazione di f in b è
  $$L_f(x) = f(b) + f'(b)(x-b)$$
  con
  $$L_f(b) = f(b) \text{ e } L_f'(b) = f'(b) \ .$$
  Se componiamo $L_f \circ L_g$ abbiamo
  $$a \overset{L_g}{\to} b \overset{L_f}{\to} f(b) \ .$$
  come per le funzioni di cui sono approssimazione . Esplicitando si ha
  $$L_f \circ L_g(x) = L_f(b + g'(a)(x-a)) = f(b) + f'(b)\, g'(a)(x-a)$$
  Poiché L_f e L_g sono lineari sappiamo che la composizione $L_f \circ L_g(x)$ ha derivata $f'(b)\, g'(a)$.
  Si ottiene cioè che anche sostituendo alle funzioni la loro approssimazione lineare locale, il risultato della derivazione della composizione ripropone la formula che avevamo indicato nella proposizione iniziale.

Il risultato sopra oltre ad essere vero è scritto in una forma che ci permette la sua generalizzazione al caso di funzioni di più variabili. La sua dimostrazione rigorosa va al di là degli scopi di questo corso.

Derivazione di Funzioni Composte: Moltiplicazione fra Matrici

La derivazione della composizione di funzioni, come abbiamo cercato di illustrare, porta sempre allo stesso risultato: *la derivata della composizione di* $f \circ g$ si trova moltiplicando, nel senso appropriato, le derivate di f e g.
Nel caso di funzioni di più variabili le derivate sono vettori e/o matrici, quindi in questo caso moltiplicazione significa moltiplicazione tra matrici o prodotto scalare di vettori.

Derivate come matrici. Funzioni di più variabili e funzioni a valori vettoriali generano una intera collezione di derivate e derivate parziali.
Consideriamo, per esempio, il caso di una funzione $K : \mathbb{R}^2 \to \mathbb{R}^2$ definita da

$$k(x,y) = (u(x,y), v(x,y)) = \left(x^2 + y, 3x - y^2\right) .$$

Ognuna delle due funzioni $u(x,y)$ e $v(x,y)$ ammette come gradiente il vettore (u_x, u_y) e (v_x, v_y). Con questi due elementi costruiamo la matrice 2×2

$$\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} 2x & 1 \\ 3 & -2y \end{pmatrix}$$

L'idea di derivata come matrice ha senso indipendentemente dalle dimensioni di dominio e codominio, compreso il caso di funzioni reali di variabile reale (riflettere sul perché).
Diamo di seguito la definizione generale

Definizione 7.17 *Sia $f : \mathbb{R}^n \to \mathbb{R}^m$ una funzione a valori vettoriali, di n variabili indipendenti $x_1, x_2, \ldots, x_n$ data da*

$$f\ (X) = (f_1(x_1, x_2, \ldots, x_n),\ f_2(x_1, x_2, \ldots, x_n), \ldots,\ f_m(x_1, x_2, \ldots, x_n)) .$$

*. Sia $X_0 = (x_1^0, x_2^0, \ldots, x_n^0)$ un punto nel dominio di f . La derivata (che chiameremo **derivata totale)** di f ed indicheremo con f' è data da*

$$f'(X_0) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(X_0) & \dfrac{\partial f_1}{\partial x_2}(X_0) & \cdots & \dfrac{\partial f_1}{\partial x_n}(X_0) \\ \dfrac{\partial f_2}{\partial x_1}(X_0) & \dfrac{\partial f_2}{\partial x_2}(X_0) & \cdots & \dfrac{\partial f_2}{\partial x_n}(X_0) \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(X_0) & \dfrac{\partial f_m}{\partial x_2}(X_0) & \cdots & \dfrac{\partial f_m}{\partial x_n}(X_0) \end{pmatrix}$$

Notare che ogni riga della matrice è data dal gradiente di una delle componenti di f . Infatti a volte la derivata totale viene scritta nella forma

$$f'(X_0) = \begin{pmatrix} \nabla f_1(X_0) \\ \nabla f_2(X_0) \\ \vdots \\ \nabla f_m(X_0) \end{pmatrix}$$

La matrice $f'(X_0)$ di f viene chiamata ***matrice Jacobiana,*** *ed è a volte indicata nella forma:*

$$f' = \frac{\partial(f_1, \cdots, f_m)}{\partial(x_1, \cdots, x_n)}$$

Esempio 7.18 *Consideriamo le seguenti funzioni*

$$\begin{aligned} f(x) &= \log(1+x)\ ,\ \ g(x,y) = x^2 - y^2\ ,\ \ h(t) = (\sin t, \cos t) \\ p(s,t) &= \left(1 + 2s^2 + 3t, s - t^3\right)\ ,\ \ q(x,y,z) = (yz,\ xz,\ xy)\,. \end{aligned}$$

Trovare le loro matrici Jacobiane.

Soluzione. Si ha

$$\begin{aligned} f'(x) &= \left(\frac{1}{1+x}\right)\ ,\ \ g'(x,y) = (2x,\ -2y)\ ,\ \ h'(t) = \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix} \\ p'(s,t) &= \begin{pmatrix} 4s & 3 \\ 1 & -3t^2 \end{pmatrix}\ ,\ \ q'(x,y,z) = \begin{pmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix} \end{aligned}$$

Notate forma e disposizione delle matrici. In particolare $f'(x)$ è una matrice 1×1 (cioè uno scalare), $g'(x,y)$ è il gradiente di g o anche una matrice 1×2. Notare infine che $h'(t)$ è una matrice 2×1.

■

Ancora su funzioni lineari, Matrici e Derivate

C'è una connessione stretta tra funzioni lineari e matrici. Per esempio, l'espressione

$$L(x,y) = (1 + 2x + 3y, 4 + 5x + 6y)$$

dice esattamente la stessa cosa dell'equazione matriciale

$$L(x,y) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} + \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

dove il punto sta ad indicare la moltiplicazione di matrici.

Più in generale, ogni funzione lineare può essere scritta nella forma

$$L(X) = C + M \cdot X$$

dove M è una matrice, spesso chiamata **matrici dei coefficienti**, X è il vettore degli ingressi e C è un vettore costante.

Scrivere le funzioni lineari in questa forma ha due vantaggi:

(A) **composizione e prodotto tra matrici.** Consideriamo due funzioni lineari espresse in forma matriciale

$$L_1(X) = C_1 + M_1 \cdot X \quad \text{e} \quad L_2(X) = C_2 + M_2 \cdot X$$

La composizione $L_1 \circ L_2$ ha allora la forma

$$\begin{aligned} L_1 \circ L_2(X) &= L_1(C_2 + M_2 \cdot X) \\ &= C_1 + M_1 \cdot (C_2 + M_2 \cdot X) \\ &= \underbrace{C_1 + M_1 \cdot C_2} + \underbrace{(M_1 \cdot M_2)} \cdot X \\ &= C + M \cdot X \end{aligned}$$

dove $C = C_1 + M_1 \cdot C_2$ ed $M = M_1 \cdot M_2$ I passaggi seguono dalle proprietà algebriche della moltiplicazione tra matrici.

(B) **Derivate di funzioni lineari.** Riprendendo il punto (A) si vede ancora una volta un fatto semplice ma importante delle funzioni lineari, e cioè che se esprimiamo L nella forma matriciale

$$L(X) = C + M \cdot X$$

e notiamo che la sua derivata è data da

$$L' = M,$$

si ha che la derivata della composizione $L_1 \circ L_2$ è la matrice prodotto $M_1 \cdot M_2$.

Approssimazione Lineare di Funzioni e Derivazione della Composizione

Per funzioni differenziabili a valori reali abbiamo visto che l'approssimazione lineare, nell'intorno di un punto X_0 è data da

$$L(X) = f(X_0) + \nabla f(X_0) \cdot (X - X_0)$$

dove il punto rappresenta il prodotto scalare tra vettori. Se la funzione è a valori vettoriali, $f = (f_1, f_2)$ (o $f = (f_1, f_2, f_3)$) , allora l'approssimazione lineare di f in X_0 è l'analogo matriciale dell'equazione precedente

$$L(X) = f(X_0) + f'(X_0) \cdot (X - X_0)$$

dove ora f' rappresenta la derivata totale (matriciale) di f ed il punto il prodotto tra matrici. Da notare che adesso $L(X)$ (così come f) è una funzione a valori vettoriali tale che

$$L(X_0) = f(X_0) \ , \quad L'(X_0) = f'(X_0) \ .$$

Ciò considerato possiamo formulare il seguente teorema

Teorema 7.19 ***(Teorema di derivazione della composizione)*** *Siano f e g funzioni differenziabili tali che X_0 appartiene al dominio di g e $g(X_0)$ al dominio di f. Si ha che*

$$(f \circ g)'(X_0) = f'(g(X_0)) \cdot g'(X_0)$$

dove il punto rappresenta il prodotto tra matrici.

Dimostrazione. (Diamo solo un'idea della dimostrazione).
L'idea è sostanzialmente la stessa che per le funzioni reali di variabile reale. Prima approssimiamo f e g con appropriate funzioni lineari L_f e L_g per le quali il teorema vale (come abbiamo visto sopra). Dopo concludiamo che il teorema vale in generale. Per g in X_0 ed f in $g(X_0)$ abbiamo le seguenti approssimazioni

$$\begin{aligned} L_g(X) &= g(X_0) + g'(X_0)(X - X_0) \\ L_f(X) &= f(g(X_0)) + f'(g(X_0))(X - X_0) \end{aligned}$$

La natura dell'approssimazione $f \approx L_f$ e $g \approx L_g$ implicano che $f \circ g \approx L_f \circ L_g$ ed anche che $(f \circ g)'(X_0) = L_f \circ L_g(X_0)$. D'altra parte, come abbiamo visto, quest'ultima derivata corrisponde al prodotto delle matrici derivate. Perciò

$$(f \circ g)'(X_0) = L_f \circ L_g(X_0) = f'(g(X_0)) \cdot g'(X_0) \ .$$

■

Esempio 7.20 *Consideriamo le funzioni*

$$f(u,v) = (uv, u-v) \quad e \;\; g(x,y) = \left(x+y, x^2+y^2\right) .$$

Trovare $(f \circ g)'(x,y)$ *e* $(f \circ g)'(3,4)$.

Soluzione. Le derivate sotto forma di matrice sono:

$$f'(u,v) = \begin{pmatrix} v & u \\ 1 & -1 \end{pmatrix} \;\; \text{e} \;\; g'(x,y) = \begin{pmatrix} 1 & 1 \\ 2x & 2y \end{pmatrix} .$$

La regola di derivazione delle funzioni composte ci dice allora che

$$(f \circ g)'(x,y) = \begin{pmatrix} v & u \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2x & 2y \end{pmatrix} = \begin{pmatrix} v+2ux & v+2y \\ 1-2x & 1-2y \end{pmatrix} ;$$

Sostituendo adesso $u = x+y$ e $v = x^2+y^2$ si ottiene

$$(f \circ g)'(x,y) = \begin{pmatrix} x^2+y^2+2(x+y)x & x^2+y^2+2y \\ 1-2x & 1-2y \end{pmatrix} .$$

Per trovare $(f \circ g)'(3,4)$ basta sostituire i valori di $x = 3$ e $y = 4$ sopra. Alternativamente, osserviamo che $g(3,4) = (7,25)$ quindi

$$f'(7,25) = \begin{pmatrix} 25 & 7 \\ 1 & -1 \end{pmatrix} , \;\; \text{e} \;\; g'(3,4) = \begin{pmatrix} 1 & 1 \\ 6 & 8 \end{pmatrix}$$

da cui

$$(f \circ g)'(3,4) = \begin{pmatrix} 25 & 7 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 67 & 81 \\ -5 & -7 \end{pmatrix}$$

■

Esempio 7.21 *A volte si opera una composizione di funzione senza esplicitare nominalmente le funzioni. Per esempio, supponiamo che u sia funzione di x ed y mentre x ed y sono funzioni di s e t . Trovare le derivate parziali $\partial u/\partial s$ e $\partial u/\partial t$.*

Soluzione. Vediamo come applicare la regola di derivazione delle funzioni composte. Scriviamo dapprima

$$u = u(x,y) \;, \quad X(s,t) = (x(s,t), y(s,t)) \;.$$

Allora è

$$u'(x,y)=\left(\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}\right)\ ,\quad \text{e}\ \ X'(s,t)=\begin{pmatrix}\dfrac{\partial x}{\partial s} & \dfrac{\partial x}{\partial t}\\ \dfrac{\partial y}{\partial s} & \dfrac{\partial y}{\partial t}\end{pmatrix}\ ,$$

la regola di derivazione del prodotto ci da

$$u'(s,t)=\left(\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}\right)\cdot\begin{pmatrix}\dfrac{\partial x}{\partial s} & \dfrac{\partial x}{\partial t}\\ \dfrac{\partial y}{\partial s} & \dfrac{\partial y}{\partial t}\end{pmatrix}$$

da cui segue

$$\frac{\partial u}{\partial s}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial s}\ ,\quad \frac{\partial u}{\partial t}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial t}$$

(notate la ”cancellazione simbolica” delle derivate parziali).

■

7.2.4 Esercizi

1. Sia $f(x) = a + bx$, $g(x) = c + dx$, $h(x) = x^2$; a, b, c, d costanti.

   (a) Determinare valori di a, b, c, e d in modo tale che $f \circ g(x) \neq g \circ f(x)$. (le possibilità sono molte).

   (b) Determinare valori di a, b, c,e d in modo tale che f e g siano funzioni diverse ma tali che $f \circ g(x) = g \circ f(x)$.

   (c) Quali condizioni su a, b, c, e d garantiscono che $f \circ g(x) = g \circ f(x)$?

   (d) Sotto quali condizioni per a e b si ha $f \circ h(x) = h \circ f(x)$?

2. Sia $f = ax^2$ e $g = bx^3$, dove a e b sono costanti non nulle. Sotto quali condizioni si ha che $f \circ g(x) = g \circ f(x)$?

3. Scrivere le derivate di ognuna delle funzioni qui di seguito e valutarle nei punti assegnati.

   (a) $f(x, y) = (x + 2y + 3, 4x + 5y + 6)$; $X_0 = (0, 0)$;

   (b) $f(x, y) = (x + 2y + 3, 4x + 5y + 6)$; $X_0 = (1, 2)$;

   (c) $g(x, y, z) = (y + z, x + z, y + z)$; $X_0 = (1, 2, 3)$;

   (d) $h(t) = (\cos t, \sin t, t)$; $t_0 = \pi/2$;

   (e) $k(s, t) = (1, 2, 3) + s(4, 5, 6) + t(7, 8, 9)$; $(s_0, t_0) = (1, 1)$.

4. Siano f, g, h, k come sopra. In ognuna delle parti sotto valutare se la composizione è sensata o meno. In caso affermativo trovare la funzione composta ed usare il teorema di derivazione delle funzioni composte per calcolare la derivata nel punto assegnato.

   (a) $k \circ f$; $(x_0, y_0) = (0, 0)$;

   (b) $f \circ g$; $(x_0, y_0, z_0) = (1, 2, 3)$;

   (c) $g \circ k$; $(s_0, t_0) = (1, 1)$.

5. In questo esercizio è $f(x) = x + x^2$, $g(x) = \sin x$ e $x_0 = 0$

   (a) Trovare L_g , approssimazione lineare di g in x_0;

   (b) Trovare L_f , approssimazione lineare di f in $g(x_0)$;

   (c) Trovare le formule per $f \circ g$ e $L_f \circ L_g$;

   (d) Mostrare che $(f \circ g)'(x_0) = (L_f \circ L_g)'(x_0)$;

(e) Usare il software per disegnare $f \circ g$ e $L_f \circ L_g$ nell'intorno di x_0 Valutare come stanno tra loro i grafici.

6. Ripetere l'esercizio precedente essendo $f(x) = x + x^2$, $g(x) = e^x$, e $x_0 = 0$.

7. Ripetere l'esercizio precedente essendo $f(t) = t^2 - 9t + 20$, $g(x, y) = x^2 + y^2$ e $X_0 = (2, 1)$. In (e) disegnare le funzioni come superfici dello spazio xyz.

8. Siano f, g, h e L le funzioni $f(t) = t^3$, $g(x, y) = x^2 + y^2$, $h(t) = (\cos t, \sin t)$, $L(x, y) = (1 + 2x + 3y, 4 + 5x + 6y)$.

   (a) Usando la regola di derivazione composta calcolare $(f \circ g)'(x_0, y_0)$;

   (b) Usando la regola di derivazione composta calcolare $(h \circ g)'(x_0, y_0)$;

   (c) Usando la regola di derivazione composta calcolare $(g \circ L)'(x_0, y_0)$; [**Sugg.**: per evitare di imbrogliarsi nei nomi riscrivete g come $g(u, v) = u^2 + v^2$.]

9. Siano g ed h come nell'esercizio precedente.

   (a) Usare la regola di derivazione composta per valutare $(g \circ h)'(t)$;

   (b) Calcolare la composizione $(g \circ h)(t)$ e valutare poi la derivata delle funzione ottenuta. Confrontare il risultato con quello ottenuto in (a).

10. Sia $f(t) = t^3$ e $h(t) = (\cos t, \sin t)$.

   (a) Calcolare $(h \circ f)'(t)$ usando la regola di derivazione della composizione;

   (b) Calcolare la composizione $(h \circ f)(t)$ e valutare poi la derivata delle funzione ottenuta. Confrontare il risultato con quello ottenuto in (a).

7.3 Derivate di Ordine Superiore e Approssimazione Quadratica.

Per funzioni di una variabile, sufficientemente regolari, non è difficile calcolare le derivate di ordine superiore al primo. Le derivate di ordine superiore, d'altra parte, rivestono un interesse non solo di tipo teorico. La derivata seconda f'' , per esempio, ha un importante significato geometrico; ci dice quanto rapidamente ed in quale direzione varia la pendenza del grafico data dal valore di f' e ci da quindi la concavità del grafico di f. Questo dato ci permette anche di distinguere tra i vari tipi di punti stazionari di f. Supponiamo, per esempio, che sia $f'(x_0) = 0$ e $f''(x_0) < 0$. Allora in x_0 il grafico è concavo e quindi f ammette un *massimo locale* in x_0.
Un altro uso che se ne fa è l'estensione del polinomio approssimante di Mac Laurin o Taylor.
Cerchiamo di vedere cosa accade nel caso di funzioni di più variabili.

7.3.1 Derivate Seconde e Superiori

Le funzioni di più variabili possono avere derivate parziali ripetute di vari ordini, ma come abbiamo già visto si passa da scalari a vettori e matrici. Vediamo con un esempio.

Esempio 7.22 *Sia $f = x^2 + xy^2$. Trovare tutte le possibili derivate seconde.*

Soluzione: Le derivate parziali prime sono

$$f_x = \frac{\partial f}{\partial x} = 2x + y^2\,, \quad f_y = \frac{\partial f}{\partial y} = 2xy$$

Derivando ancora otteniamo i seguenti quattro risultati.
Da f_x si ha

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} = 2\,, \quad f_{xy} = \frac{\partial^2 f}{\partial y \partial x} = 2y$$

mentre da f_y si ottiene

$$f_{yx} = \frac{\partial^2 f}{\partial x \partial y} = 2y\,, \quad f_{yy} = \frac{\partial^2 f}{\partial y^2} = 2x$$

(Da notare che i simboli f_{yx} e $\dfrac{\partial^2 f}{\partial x \partial y}$ hanno lo stesso significato anche se l'ordine dei simboli può sembrare rovesciato).

Si deve osservare che se f è una funzione da $\mathbb{R}^2 \to \mathbb{R}$, il suo gradiente $\nabla f = (f_x, f_y)$ è una funzione $\nabla f : \mathbb{R}^2 \to \mathbb{R}^2$ e quindi la sua derivata (derivata seconda per f) è data dalla matrice

$$f'' = (\nabla f)' = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 2y \\ 2y & 2x \end{pmatrix}.$$

Questa matrice è anche chiamata **matrice Hessiana** di f (La matrice Hessiana di f calcolata ne punto X_0 viene indicata come $H_f(X_0)$.
(Il nome di matrice Hessiana deriva da quello del matematico tedesco Ludwig Otto Hesse (1811-1874)).

■

Osserviamo alcuni fatti relativi alla matrice delle derivate seconde:

- **Dimensione.** Per una funzione $f(x_1, x_2, \ldots, x_n)$ di n variabili l'Hessiana è una matrice $n \times n$. L'elemento di posto j nella riga $i - esima$ è $f_{x_i x_j}$ che si ottiene derivando f prima rispetto ad x_i e poi rispetto ad x_j.
  Per esempio se $f(x, y, z) = xz + yz^2$ allora l'Hessiana di f è una matrice 3×3 della forma

$$f''(x, y, z) = \begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2z \\ 0 & 0 & 1 \\ 2z & 1 & 2x \end{pmatrix}.$$

- **Le righe sono gradienti.** Le righe delle matrice Hessiana sono i vettori gradienti delle derivate parziali $f_x,\ f_y,\ f_z$ che sono funzioni da $\mathbb{R}^3 \to \mathbb{R}$. La seconda riga, per esempio è ∇f_y.

- **L'ordine di derivazione (di solito) non conta.** In entrambi gli esempi che abbiamo proposto la matrice Hessiana è simmetrica rispetto alla diagonale principale, in altre parole è

$$f_{xy} = f_{yx}\ , \quad f_{xz} = f_{zx}\ , \quad f_{yz} = f_{zy}\ ;$$

  cioè l'ordine di derivazione nelle derivate parziali miste sembra non essere importante (almeno ad ora). E' un dato interessante che questo fatto vale per tutte le funzioni di più variabili con un comportamento "sufficientemente regolare".
  Chiariremo i termini del problema per una funzione di due variabili (anche se il risultato vale qualunque sia il numero delle variabili indipendenti).

Teorema 7.23 *(Eguaglianza delle derivate miste). Sia* $f : \mathbb{R}^2 \to \mathbb{R}$ *una funzione con la proprietà che le derivate seconde miste* $f_{xy}, f_{yx} : \mathbb{R}^2 \to \mathbb{R}$ *sono definite e continue nel dominio di* f *. Allora per ogni* (x, y) *si ha*

$$f_{xy}(x, y) = f_{yx}(x, y) .$$

Dimostrazione. (Idea della dimostrazione). Ci sono molti modi di dimostrare il teorema. Noi ne useremo uno che coinvolge l'uso di un integrale doppio. Diamo solo l'idea della dimostrazione, lasciando i dettagli per esercizio.
Dimostreremo che $f_{xy}(0,0) = f_{yx}(0,0)$. Questo è sufficiente visto che il punto $(0,0)$ non ha niente di particolare.
Consideriamo il quadrato $R = [0,h] \times [0,h]$, con h generico, vogliamo mostrare che è

$$\iint_R f_{xy}(x, y)\, dA = \iint_R f_{yx}(x, y)\, dA . \tag{7.2}$$

Fermiamoci prima a capire come questa uguaglianza ci può aiutare a risolvere il problema. Supponiamo, per esempio, che sia $f_{xy}(0,0) > f_{yx}(0,0)$. Allora per la continuità delle due funzioni si ha che $f_{xy}(x,y) > f_{yx}(x,y)$ in tutto un intorno del punto $(0,0)$. In particolare si può trovare un rettangolo $R = [0,h] \times [0,h]$ nel quale è $f_{xy}(x,y) > f_{yx}(x,y)\ \forall (x,y) \in R$. In questo caso, ovviamente l'uguaglianza integrale non può valere. Infatti, consideriamo il lato sinistro dell'uguaglianza 7.2 e calcoliamo l'integrale in modo iterato, ricordando che, per definizione, f_{xy} è la derivata fatta rispetto ad y di f_x.Si ha

$$\begin{aligned} \iint_R f_{xy}(x, y)\, dA &= \int_0^h \left(\int_0^h f_{xy}(x,y)\, dy \right) dx \\ &= \int_0^h f_x(x,y)\big|_0^h\, dx \\ &= f(x,h) - f(x,0)\big|_0^h \\ &= f(h,h) - f(0,h) - f(h,0) + f(0,0) . \end{aligned}$$

Un calcolo simile mostra che il lato destro dell'eguaglianza integrale 7.2 ha lo stesso valore. ■

Polinomi di Taylor ed Approssimazione Quadratica

Abbiamo già visto quale sia l'approssimazione lineare di una funzione di più variabili. Data $f(x, y)$ l'approssimazione lineare (nel caso di due variabili) in $X_0 = (x_0, y_0)$ è stata definita come

$$\begin{aligned} L(x,y) &= f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \\ &= f(X_0) + \nabla f(X_0)\cdot(X-X_0) \end{aligned}$$

Il passo successivo, che implica l'uso delle derivate seconde(in analogia a quanto fatto e all'analogia con il caso di una variabile), è dato da

$$\begin{aligned} Q(x,y) &= f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \\ &\quad + \frac{f_{xx}(x_0,y_0)}{2}(x-x_0)^2 + f_{xy}(x_0,y_0)(x-x_0)(y-y_0) \\ &\quad + \frac{f_{yy}(x_0,y_0)}{2}(y-y_0)^2 \end{aligned}$$

Osserviamo che:

La definizione di Q garantisce che nel punto (x_0, y_0) la funzione f e q hanno le stesse derivate prime e seconde. Infatti, cerchiamo per esempio Q_{xy} si ha

$$Q_x(x,y) = f_x(x_0,y_0) + f_{xx}(x_0,y_0)(x-x_0) + f_{xy}(x_0,y_0)(y-y_0)$$

da cui

$$Q_{xy} = f_{xy}(x_0,y_0)$$

In forma vettoriale l'approssimazione quadratica può essere scritta nel seguente modo

$$Q(x,y) = f(X_0) + \nabla f(X_0)\cdot(X-X_0) + \frac{1}{2}(X-X_0)^T \cdot f''(X_0)\cdot(X-X_0)$$

dove il termine

$$(X-X_0)\cdot f''(X_0)\cdot(X-X_0)^T$$

indica il prodotto del vettore riga $(X - X_0)$ per la matrice $f''(X_0)$ per il vettore colonna (vettore trasposto) $(X - X_0)^T$.

La scrittura vettoriale ci aiuta, in parte perché più simile all'analoga scrittura per le funzioni di una variabile, soprattutto perché questa scrittura è svincolata dalla dimensione dello spazio di arrivo che può essere di dimensione 2 come 3 o altro essendo il significato dei simboli lo stesso.

Approssimazioni di Ordine Superiore Ovviamente non è necessario fermarsi ad una approssimazione del secondo ordine. Se la regolarità della funzione lo permette possiamo scrivere l'approssimazione cubica ed oltre. Il vero problema è la complessità della scrittura. Infatti, se consideriamo i termini di ordine 3 già nel caso di una funzione di due variabili si ha

$$\begin{aligned}&\frac{1}{3!}\left(f_{xxx}(x_0,y_0)(x-x_0)^3+3f_{xxy}(x_0,y_0)(x-x_0)^2(y-y_0)\right.\\&\left.+3f_{xyy}(x_0,y_0)(x-x_0)(y-y_0)^2+f_{yyy}(x_0,y_0)(y-y_0)^3\right).\end{aligned}$$

Ricordiamo comunque che *Maple* o altri software non hanno problemi anche nel calcolare i termini di ordine superiore.

Esempio 7.24 *Trovare l'approssimazione del secondo ordine per la funzione* $f(x,y)=ye^x$ *nel punto* $(0,0)$.

Soluzione. Calcoliamo f e le derivate prime e seconde in $(0,0)$. Si ha

$$\begin{aligned}f(0,0)&=0,\ \nabla f(x,y)=(ye^x,e^x)\ \text{da cui}\ \nabla f(0,0)=(0,1),\\ f''(x,y)&=\begin{pmatrix}ye^x & e^x\\ e^x & 0\end{pmatrix}\ \text{da cui}\ f''(0,0)=\begin{pmatrix}0&1\\1&0\end{pmatrix}\end{aligned}$$

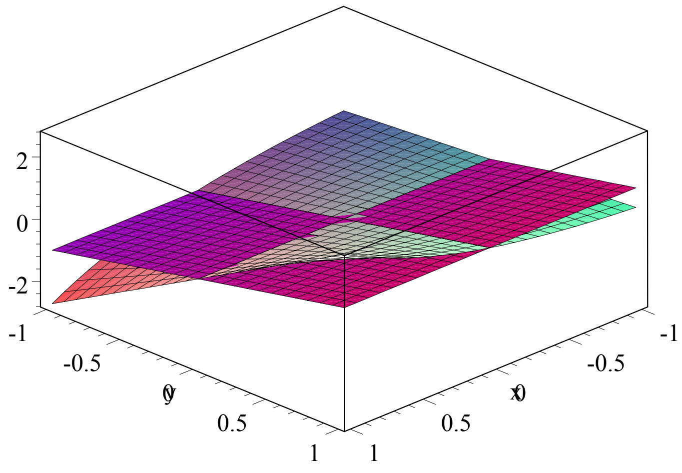
perciò è

$$\begin{aligned}Q(x,y)&=(0,1)\cdot(x,y)+\frac{1}{2}(x,y)\cdot\begin{pmatrix}0&1\\1&0\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}\\&=y+xy\end{aligned}$$

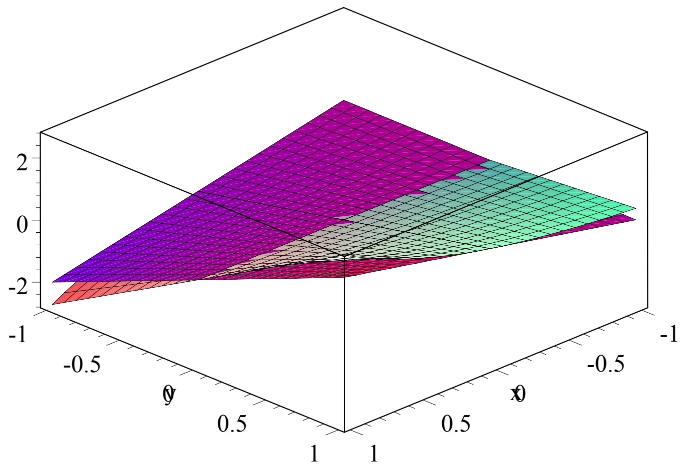
Come l'approssimazione lineare, l'approssimazione quadratica approssima bene f nell'intorno del punto $(0,0)$

■

Qui di seguito diamo il grafico di f , L, Q per un confronto. Il grafico di Q fornisce un'approssimazione migliore di quello di L.



Grafici di f e di L



Grafici di f e Q

7.3.2 Esercizi

L'uso di software può essere utile. Riportiamo per comodità alcuni comandi di *Maple*. I primi due caricano comandi. Sperimentare i comandi per vedere come operano, usate inoltre, se necessario, l'help in linea.

```
>with(linalg);
>readlib(mtaylor);
>grad(x^2+y^2, [x,y]);
>grad(x^2+y^2+z^2, [x,y,z]);
>hessian(x^2*y, [x,y]);
>hessian(x^2*y*z, [x,y,z]);
>mtaylor(sin(x), x=0, 5);
>mtaylor(sin(x)+cos(y), [x,y]);
>mtaylor(x^2+y^2, [x=2,y=1]);
```

1. Trovare (a mano) i polinomi di Taylor di primo, secondo e terzo grado $p_1,\ p_2,\ p_3$ per ognuna delle seguenti funzioni. Se possibile usare la tecnologia per disegnare $f,\, p_1,\, p_2,\, p_3$.

   (a) $f\,(x) = \cos x\,;\ x_0 = 0;$

   (b) $f\,(x) = \log x\,;\ x_0 = 1;$

   (c) $f\,(x) = \sin \sqrt{x}\,;\ x_0 = 0.$

2. Per ognuna delle funzioni calcolare la matrice hessiana $f''\,(X_0)$ nel punto X_0 assegnato. Usare *Maple* (o altro software) per controllare il risultato.

   (a) $f\,(x,y) = \sin(x\,y)\,;\ X_0 = (0,0)\,;$

   (b) $f\,(x,y) = x\,y\,;\ X_0 = (0,0)\,;$

   (c) $f\,(x,y) = \sin\,(x\,) + \cos\,(2y)\ ;\ X_0 = (0,0)\,;$

   (d) $f\,(x,y) = x^{\,2} + y^2\,;\ X_0 = (0,0)\,;$

   (e) $f\,(x,y) = x^{\,2} - y^2\,;\ X_0 = (0,0)\,;$

   (f) $f\,(x,y) = Ax^2 + By^2 + Cxy + Dx + Ey + F\,;\ X_0 = (x_0, y_0)\,;$

   (g) $f\,(x,y,z) = \sin\,(x + y + z^2)\ ;\ X_0 = (0,0,0)\,.$

3. Per ognuna delle funzioni, calcolare l'approssimazione quadratica $Q\,(X_0)$ nel punto X_0 assegnato. Usare *Maple* (o altro software) per controllare il risultato. Se possibile, usare la tecnologia per disegnare f e Q.

   (a) $f\,(x,y) = \sin(x\,y)\,;\ X_0 = (0,0)\,;$

(b) $f(x,y) = x\,y\,;\ X_0 = (0,0)\,;$

(c) $f(x,y) = \sin(x) + \cos(2y)\ ;\ X_0 = (0,0)\,;$

(d) $f(x,y) = x^{2} + y^2\,;\ X_0 = (0,0)\,;$

(e) $f(x,y) = x^{2} - y^2\,;\ X_0 = (0,0)\,;$

(f) $f(x,y,z) = \sin(x + y + z^2)\ ;\ X_0 = (0,0,0)\,.$

4. Abbiamo, nel capitolo, affermato che $Q(X)$ può essere scritto, in forma vettoriale, nella forma

   $Q(X) = f(X_0) + \nabla f(X_0) \cdot (X - X_0) + \frac{1}{2}(X - X_0) \cdot f''(X_0) \cdot (X - X_0)^T$

   Esplicitare tutti i dettagli di calcolo per verificare che la formula è vera.

7.4 Massimi, Minimi ed Approssimazione Quadratica

Una funzione, come noto, ha un **massimo locale** in X_0 se $f(X_0) \geq f(X)$ per tutti gli ingressi X in un intorno di X_0. In questo caso il valore $f(X_0)$ è chiamato **valore di massimo locale** di f. Le definizioni di **minimo locale** e **valore minimo locale** sono definite in modo simile. Il problema che ci poniamo è quello di dare le condizioni necessarie e quelle sufficienti per determinare i punti di massimo e di minimo di una funzione di più variabili. Per semplicità ci limiteremo essenzialmente alle funzioni di due variabili.

Sia f una funzione di due variabili ; per evitare questioni tecniche, che esulano lo scopo di questo corso, assumeremo che tutte le derivate parziali esistano e siano continue. Consideriamo inoltre il caso che il punto X_0 sia l'origine $(0,0)$. L'approssimazione quadratica di $f(x,y)$ ha la forma

$$\begin{aligned} & Q(x,y) \\ = \; & f(0,0) + f_x(0,0)\,x + f_y(0,0)\,y + \frac{f_{xx}(0,0)}{2}x^2 + f_{xy}(0,0)\,xy + \frac{f_{yy}(0,0)}{2}y^2 \end{aligned}$$

Se $(0,0)$ è un punto stazionario, allora i termini del primo ordine scompaiono (sono zero) e si ha

$$Q(x,y) = f(0,0) + \frac{f_{xx}(0,0)}{2}x^2 + f_{xy}(0,0)\,xy + \frac{f_{yy}(0,0)}{2}y^2$$

La domanda è capire come i valori di f_{xx} , f_{xy} , f_{yy} determinano il tipo di punto stazionario.
In $Q(x,y)$ il primo termine è costante, quindi ciò che conta sono gli altri termini che hanno la forma del tipo

$$Ax^2 + Bxy + Cy^2 \quad \left(A = \frac{f_{xx}(0,0)}{2}\ ,\ B = f_{xy}(0,0)\ ,\ C = \frac{f_{yy}(0,0)}{2} \right)$$

Analisi di $Ax^2 + Bxy + Cy^2$

Sia $f(x,y) = Ax^2 + Bxy + Cy^2$. Il punto $(0,0)$ è un punto stazionario di f qualunque siano i valori di A, B, e C. Per vedere come il tipo di punto stazionario dipende da questi valori, studieremo alcuni esempi semplici ma importanti. In ognuno degli esempi valuteremo la matrice hessiana in $(0,0)$, utile anche nel seguito.

Esempio 7.25 *Sia* $f(x,y) = x^2 + y^2$. *Come si comporta* f *nell'intorno del punto stazionario* $(0,0)$*? Descrivere la superficie* $z = f(x,y)$. *Cosa cambia se consideriamo* $f(x,y) = -(x^2 + y^2)$.

Soluzione. La matrice hessiana della funzione è semplice

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Chiaramente, f ha un minimo locale in $(0,0)$ poiché per tutte le coppie (x,y) si ha

$$f(x,y) = x^2 + y^2 \geq 0 = f(0,0) \ .$$

■

La superficie $z = f(x,y)$, come già detto, è chiamata **paraboloide di rotazione** (o semplicemente paraboloide), ha vertice in $(0,0)$ e le curve di livello di f sono circonferenze centrate nell'origine. Il cambio di $f(x,y)$ in $-f(x,y)$ cambia il minimo in massimo, il paraboloide è rivolto verso il basso e la matrice hessiana cambia di segno.

Esempio 7.26 *Sia* $g(x,y) = 3x^2 + 2y^2$. *Come si comporta* g *nell'intorno di* $(0,0)$? *Descrivere la superficie* $z = g(x,y)$

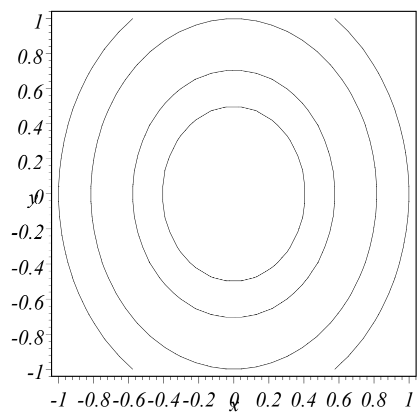


Grafico di contorno di $3x^2 + 2y^2$

Soluzione La differenza con l'esempio precedente è solo nella presenza dei coefficienti 2 e 3. Così , visto che è

$$g(x,y) = 3x^2 + 2y^2 \geq 0 = g(0,0) \ .$$

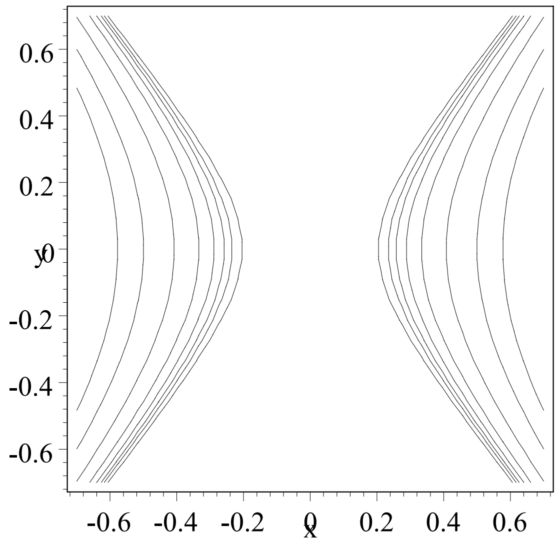
Questa volta, tuttavia, i differenti coefficienti di x^2 e y^2 significa che le curve di livello $3x^2 + 2y^2 = c$ sono ellissi. La superficie $z = 3x^2 + 2y^2$ è chiamata **paraboloide ellittico**. La matrice hessiana è data da

$$g''(0,0) = \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix}$$

■

Esempio 7.27 *Sia* $h(x,y) = 3x^2 - 2y^2$*. Qual'è il comportamento di* h *nell'intorno del punto stazionario* $(0,0)$*? Descrivere la superficie* $z = h(x,y)$.

Soluzione Poiché i coefficienti di x^2 e y^2 hanno segno differente, le curve di livello, che corrispondono ad equazioni della forma $h(x,y) = 3x^2 - 2y^2 = c$, sono delle iperboli e la superficie è chiamato **paraboloide iperbolico.** Qui di seguito un esempio di mappa di contorno



Mappa di contorno di $3x^2 - 2y^2$

La mappa di contorno ci fa vedere che il punto $(0,0)$ è una sella, cioè un punto che rappresenta sia un massimo ed un minimo per h dipende dalla direzione in cui si osserva il fenomeno. Se fissiamo $x = 0$ si ha $h(0,y) = -2y^2$, quindi questa sezione è una parabola rivolta verso il basso e l'origine è un *massimo*; d'altra parte se consideriamo la sezione $y = 0$ si ha $h(x,0) = 3x^2$

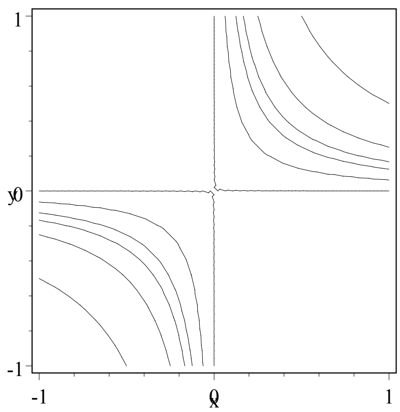
cioè una parabola rivolta verso l'alto e l'origine è un *minimo*. La matrice hessiana è adesso

$$h''(0,0) = \begin{pmatrix} 6 & 03x^2 - 2y^2 = 1/4 \\ 0 & -4 \end{pmatrix}$$

■

Esempio 7.28 *Consideriamo ora* $j(x,y) = xy$ *Qual'è il comportamento di* j *nell'intorno del punto stazionario* $(0,0)$. *Descrivere la superficie* $z = j(x,y)$.

Soluzione La funzione j si comporta come la funzione h dell'esempio precedente. Le curve di livello sono anche in questo caso delle iperboli del tipo $xy = c$.



Mappa di contorno di xy

Prendendo una sezione della superficie col piano $x = y$ si ottiene $j(x,y) = x^2$, mentre rispetto al piano $x = -y$ si ha $(x,y) = -x^2$. Queste opposte tendenze ci dicono che anche in questo caso il punto $(0,0)$ è una *sella* e la superficie è un altro paraboloide iperbolico.
La matrice hessiana questa volta è

$$j''(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

■

Con questo ultimo esempio vogliamo illustrare una tecnica importante che useremo nel seguito.

Esercizio 7.29 *Sia* $k(x,y) = x^2 + xy + y^2$. *Discutere il punto stazionario* $(0,0)$.

Soluzione. Completiamo il quadrato in x ed y.

$$k(x,y) = x^2 + xy + y^2 = \left(x + \frac{y}{2}\right)^2 + \frac{3}{4}y^2$$

Questa nuova scrittura mostra che $(0,0)$ è un punto di minimo, poiché per tutti gli (x,y)

$$k(x,y) = \left(x + \frac{y}{2}\right)^2 + \frac{3}{4}y^2 \geq 0 = k(0,0)$$

La matrice hessiana è data da

$$k''(0,0) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Il caso generale. Vogliamo considerare il caso generale $f(x,y) = Ax^2 + Bxy + Cy^2$, e vedere, come nel caso precedente, come completare il quadrato. Supponiamo, per convenienza che sia $A \neq 0$. Si ha

$$\begin{aligned} f(x,y) &= Ax^2 + Bxy + Cy^2 = A\left(x^2 + \frac{B}{A}xy + \frac{C}{A}y^2\right) \\ &= A\left(\left(x + \frac{B}{2A}y\right)^2 + \left(\frac{C}{A} - \frac{B^2}{4A^2}\right)y^2\right) \end{aligned}$$

Questo mostra che il tipo di punto stazionario dipende dal segno dei coefficienti di y^2.
Si ha che

$$\frac{C}{A} - \frac{B^2}{4A^2} \geq 0 \Longleftrightarrow \frac{C}{A} \geq \frac{B^2}{4A^2} \Longleftrightarrow 4AC - B^2 \geq 0$$

- Se $4AC - B^2 > 0$ allora f ha un minimo locale in $(0,0)$ se $A > 0$, un massimo locale se $A < 0$.
- Se $4AC - B^2 < 0$, allora f ha un punto di sella in $(0,0)$.

In termini di derivate. Se riscriviamo le conclusioni di cui sopra in termini di derivate, ricordando che $2A = f_{xx}(0,0)$, $B = f_{xy}(0,0)$, $2C = f_{yy}(0,0)$, si ha che

$$4AC - B^2 = f_{xx}f_{yy} - f_{xy}^2\,;$$

in altre parole, $4AC - B^2$ è il **determinante** della matrice hessiana

$$f''(0,0) = \begin{pmatrix} f_{xx}(0,0) & f_{xy}(0,0) \\ f_{xy}(0,0) & f_{yy}(0,0) \end{pmatrix}.$$

■

Riscriviamo adesso il risultato ottenuto come teorema generale. Assumiamo, come sopra, che la funzione f ha derivate seconde continue.

Teorema 7.30 ***(Punti stazionari e matrice Hessiana)*** *Sia* (x_0, y_0) *un punto stazionario di una funzione* $f(x,y)$. *Sia* $f''(x_0,y_0)$ *la matrice hessiana di* f , *e sia*

$$D = f_{xx}(x_0,y_0)\, f_{yy}(x_0,y_0) - f_{xy}^2(x_0,y_0)$$

il determinante di $f''(x_0,y_0)$. *Allora*

(▶) Se $D > 0$ *e* $f_{xx}(x_0,y_0) > 0$, *allora* f *ha un minimo locale in* (x_0,y_0);

(▶) Se $D > 0$ *e* $f_{xx}(x_0,y_0) < 0$, *allora* f *ha un massimo locale in* (x_0,y_0);

(▶) Se $D < 0$ *allora* f *ha un punto di sella nel punto* (x_0,y_0);

(▶) Se $D = 0$ *si ha bisogno di ulteriori informazioni.*

Questo teorema rende, in molti casi, routine il calcolo di massimi e minimi.

Esempio 7.31 *La funzione* $f(x,y) = xy - y - 2x + 2$ *ha un punto stazionario. Trovarlo e dire di che tipo di punto stazionario si tratta.*

Soluzione. Per trovare il punto stazionario risolviamo il sistema

$$\nabla f(x,y) = (y-2, x-1) = (0,0) \ ;$$

chiaramente l'unica soluzione è il punto $(1,2)$. In questo punto la matrice hessiana ha la forma

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

Ne segue che $D = -1$ e quindi il punto $(1,2)$ è un punto di sella.

■

Nota 7.32 *La stessa idea, basata sulla matrice hessiana, può essere applicata in dimensioni superiori a 2. Ma le conclusioni sono più complicate e macchinose. Non le presenteremo qui. In molti casi un'analisi diretta permette comunque di risolvere il problema.*

Esempio 7.33 *Sia $f(x, y, z)$ una funzione di tre variabili e supponiamo che sia $\nabla f(0, 0, 0) = (0, 0, 0)$ così che f ha un punto stazionario nell'origine. Supponiamo che sia $f_{xx}(0, 0) > 0$ e $f_{yy}(0, 0) < 0$. Mostrare che l'origine non è né massimo né minimo.*

Soluzione. Consideriamo la funzione $g(t) = f(t, 0, 0)$. Si ha che $g(0) = 0$, $g'(0) = 0$ e $g''(0) = f_{xx}(0, 0) > 0$. Ne segue che g ha un minimo locale per $t = 0$. In modo simile consideriamo la funzione $h(t) = f(0, t, 0)$; questa ha un massimo locale per $t = 0$. Ne segue che la funzione f non può avere né massimo né minimo in $(0, 0, 0,)$.

■

7.4.1 Esercizi

1. Supponiamo che $f(x,y)$ abbia un punto stazionario in (x_0,y_0) e che sia $f_{xx}(x_0,y_0)\, f_{yy}(x_0,y_0) < 0$. Mostrare che (x_0,y_0) è un punto di sella.

2. Per ogni costante $A \neq 0$, il grafico di $z = Ax^2 + Ay^2$ è un paraboloide circolare.

   (a) Spiegare cosa si ottiene tagliando la superficie con un piano $z = c$;

   (b) Qual'è la forma della sezione che si ottiene sezionando la superficie con un piano $x = c$.

   (c) Qual'è la forma della sezione che si ottiene sezionando la superficie con un piano $y = c$.

3. Ripetere l'esercizio precedente per il paraboloide $z = x^2 - y^2$.

4. Sia $f(x,y) = 3x^2 - 6xy + 2y^3$. Trovare e classificare i punti stazionari di f.

5. Trovare e classificare i punti stazionari di $f(x,y) = x^3 + y^3 + 3x^2 - 3y^2 - 8$.

6. Trovare e classificare i punti stazionari di $f(x,y) = x^2 - xy - y^2$.

7. Trovare e classificare i punti stazionari di $f(x,y) = x^4 + y^4$. E' possibile usare il teorema di classificazione dei punti stazionari?

8. Sia $f(x,y) = x^2 + axy + by^2$, con $a,\ b$ costanti reali.

   (a) Spiegare perché $(0,0)$ è stazionario indipendentemente dal valore delle costanti;

   (b) Per quali valori di $a,\ b$ l'origine è un massimo locale? Un minimo locale? Una sella? Dare una risposta esauriente fornendo esempi dei tre casi;

   (c) Supponiamo adesso che sia $b = a^2/4$. Che tipo di punto critico è l'origine in questo caso?

9. Sia $f(x,y) = \sin x + \cos 2y$. Mostrare che $(\pi/2, 0)$ è un punto stazionario. Di che tipo è?

10. Sia $f(x,y) = x^2$; il grafico di questa funzione è detto cilindro.

    (a) Descrivere il grafico di f vicino al punto stazionario $(0,0)$;

(b) Mostrare che f ha un minimo locale in $(0,0)$ (Notare che il minimo locale non è "stretto" nel senso che $f(x,y) \geq f(0,0)$ ma non è detto che sia $f(x,y) > f(0,0)$);

(c) Cosa dice il test delle derivate seconde?

11. Considerare la funzione $f(x,y) = x^2 + bxy + y^2$ dove b è una qualsiasi costante.

    (a) Per quali valori di b la funzione ha in zero un massimo? Un minimo? Una sella?

    (b) Se è possibile usare il software per disegnare curve di livello e superfici nei vari casi;

    (c) Per quali valori di b l'hessiano è zero? Qual'è il comportamento della funzione nell'intorno dell'origine, in questo caso? Provare a disegnare (a mano) varie curve di livello nell'intorno dell'origine per capire cosa accade.

7.5 Moltiplicatori di Lagrange e Ottimizzazione Vincolata

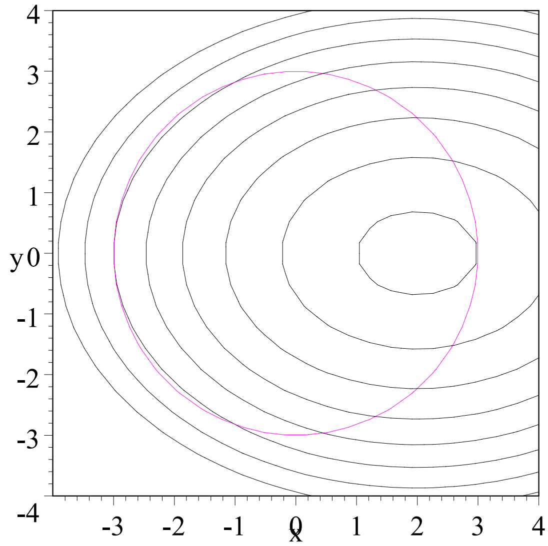
Sia $f(x,y)$ una funzione definita su un dominio di $\mathbb{R}^2$; abbiamo visto come trovare i massimi e minimi relativi della funzione. I possibili candidati sono i punti stazionari ed il test delle derivate seconde ci aiuta (in molti casi) a scegliere tra i vari casi. Se, per esempio, $f(x,y) = x^2 - 4x + 2y^2$, allora $\nabla f(x,y) = (2x-4, 4y)$ e $(2,0)$ è l'unico punto stazionario, ed è facile vedere che tale punto è un minimo locale.

In alcune situazioni è interessante trovare il valore massimo e minimo di una funzione sottoposta a qualche "vincolo" rispetto al dominio. Per esempio potremmo voler trovare il massimo ed il minimo di $f(x,y)$ quando (x,y) è vincolato a stare sulla circonferenza $x^2 + y^2 = 9$. Risolvere l'equazione per il gradiente, come sopra, non ci da nulla di buono; infatti il punto $(2,0)$ non appartiene alla circonferenza e non risolve quindi il problema richiesto. Vedremo che l'uso della tecnica di soluzione dell'equazione del gradiente è ancora lo strumento da usare, ma in modo appropriato al nuovo tipo di problema in esame.

Il problema che vogliamo studiare è chiamato di **ottimizzazione vincolata**, la funzione di cui si cerca massimo o minimo è detta **funzione obiettivo**, la restrizione sugli ingressi è descritta da un'equazione detta **equazione del vincolo** (a volte i vincoli sono dati da disuguaglianza o più equazioni).

Esempio 7.34 *Cerchiamo massimi e minimi di* $f(x,y) = x^2 - 4x + 2y^2$ *soggetta al vincolo* $x^2 + y^2 = 9$.

Soluzione Il vincolo descrive una circonferenza nel piano xy centrata nell'origine e di raggio 3. Vediamo in uno stesso grafico il vincolo e le curve di livello di $f(x,y)$



Mappa di contorno di $f(x, y)$ e vincolo

Osserviamo il disegno più da vicino.

Funzione Obiettivo Le curve di livello di f sono ellissi con centro in $(2, 0)$ dove f ha il suo valore minimo. Più grandi sono le ellissi maggiori sono i valori di f.

Il Vincolo Gli ingressi che soddisfano il vincolo sono quelli che giacciono sulla circonferenza C, $x^2 + y^2 = 9$. Se immaginiamo di camminare sulla superficie $z = f(x, y)$ sopra la curva C il problema è decidere in quale punto l'altezza della superficie, descritta dalle curve di livello, è massima o minima.

Una Risposta Grafica Una osservazione attenta alla curva C suggerisce di considerare in particolare i quattro punti della circonferenza $(-3, 0)$, $(3, 0)$, e $(-2, \sqrt{5})$, $(-2, -\sqrt{5})$. In questi punti si ottiene il massimo ed il minimo dell'altezza, mentre tra questi punti sulla superficie si sale e si scende. I valori di f in questi punti sono

$$f(-3, 0) = 21, \; f(3, 0) = -3, \; e \; f\left(-2, \sqrt{5}\right) = f\left(-2, -\sqrt{5}\right) = 22 .$$

Seguendo questi punti, il massimo vincolato vale 22 nei punti $(-2, \pm\sqrt{5})$, mentre il minimo vale -3 in $(3, 0)$.

■

Osservazione La scelta di quei quattro punti **P** è dovuta al fatto che **le curve di livello di f passanti per P sono tangenti al vincolo C.**

L'osservazione appena fatta, che cercheremo di reinterpretare in termini di gradiente, è l'idea principale di questo paragrafo. Vedremo poi come usare questa proprietà che chiameremo **condizione di Lagrange**

E' possibile usare un approccio diretto come mostra il seguente esempio.

Esempio 7.35 *Cerchiamo massimi e minimi di* $f(x,y) = x^2 - 4x + 2y^2$ *soggetta al vincolo* $x^2 + y^2 = 9$ *parametrizzando il vincolo.*

Soluzione La circonferenza $x^2 + y^2 = 9$ può essere parametrizzata, per esempio, come

$$X(t) = (x(t), y(t)) = (3\cos t, 3\sin t) \ , \ 0 \leq t \leq 2\pi$$

Vincolare (x,y) a giacere sulla circonferenza significa vincolare la funzione f a stare sulla circonferenza, cioè

$h(t) = f(X(t)) = 9\cos^2 t - 12\cos t + 18\sin^2 t = 9\sin^2 t - 12\cos t\,,\ 0 \leq t \leq 2\pi\,.$

Questo riduce il problema al caso di una sola variabile, che si risolve semplicemente

$$h'(t) = 18\sin t\cos t + 12\sin t = \sin t\,(18\cos t + 12)\,.$$

Si ha $h'(t) = 0$ se $\sin t = 0$ (cioè $y = 0$) oppure $\cos t = -2/3$ (cioè $x = -2$) che sono gli stessi punti trovati prima.
Osserviamo anche che dalla regola di derivazione della composizione si ha

$$h'(t) = \nabla f(X(t)) \cdot X'(t) \ .$$

Dal punto di vista geometrico, nei punti in cui si ha $h'(t) = 0$, si ha che il gradiente ∇f è perpendicolare al vettore $X'(t)$ che è il vettore tangente alla curva che rappresenta il vincolo. Questo è un altro modo di enunciare la condizione di Lagrange, perché in ogni punto $P(x,y)$ il gradiente di f è perpendicolare alla curva di livello di f passante per P.

$\blacksquare$

7.5.1 Gradienti e Condizioni di Lagrange

Il punto fondamentale, illustrato nei due esempi precedenti, è il seguente.

In un problema di ottimizzazione vincolata, un punto di massimo o minimo è caratterizzato dall'avere il gradiente della funzione obbiettivo deve essere perpendicolare all'insieme dei vincoli.

La condizione è complicata ad esprimersi verbalmente ma semplice ad usarsi con l'aiuto del gradiente. Il fatto fondamentale che unisce tutte le idee principali, è la connessione fra gradienti ed insiemi di livello.

Teorema 7.36 ***Gradienti ed insiemi di livello.*** *Sia $g(x,y)$ una funzione differenziabile, (x_0, y_0) un punto nel dominio di g. Sia C la curva di livello di g passante per (x_0, y_0). Se $\nabla g(x_0, y_0) \neq (0,0)$ allora $\nabla g(x_0, y_0)$ è perpendicolare a C in (x_0, y_0).*

Dimostrazione. La dimostrazione è una semplice e elegante applicazione della regola di derivazione della composizione di funzioni. Supponiamo che la curva C sia parametrizzata da una funzione a valori vettoriali $X(t)$ tale che $X(t_0) = (x_0, y_0)$. Il vettore $X'(t)$ è tangente a tangente a C in (x_0, y_0). (Stiamo assumendo il fatto tecnico che tale parametrizzazione esista poiché abbiamo assunto che $\nabla g(x_0, y_0) \neq (0,0)$). Poiché g é costante lungo C la funzione composta $g(X(t))$ è costante in t. Allora si ha

$$0 = \frac{d}{dt} g(X(t))\bigg|_{(t_0)} = \nabla g(x_0, y_0) \cdot X'(t_0)$$

Ne segue che $\nabla g(x_0, y_0)$ è perpendicolare a $X'(t_0)$ e quindi a C. ■

Cosa accade in dimensioni maggiori Nonostante la maggior parte della teoria sia sviluppata in dimensione due, il teorema precedente vale anche in dimensioni maggiori di due, eccetto che in questi casi l'insieme di livello è una superficie, non una curva. Per esempio, in dimensione tre l'insieme di livello $g(x,y,z) = x^2 + y^2 + z^2 = 1$ è una sfera in $\mathbb{R}^3$. Il teorema ci dice che in ogni punto (x,y,z) della sfera il vettore gradiente $(2x, 2y, 2z)$ è perpendicolare alla sfera.

Il gradiente della funzione vincolo. Il vincolo è di solito descritto da un' equazione. Se scriviamo l'equazione del vincolo nella forma $g(x,y) = 0$ dove $g(x,y)$ è una funzione, allora la curva che descrive il vincolo è la curva di livello zero di $g(x,y)$. Il teorema aggiunge inoltre che in ogni punto (x,y) della curva di livello, il vettore gradiente $\nabla g(x,y)$ della funzione vincolo o è il vettore zero oppure è perpendicolare alla curva di livello.

Il gradiente della funzione obiettivo. Sia (x_0, y_0) un punto della curva $g(x,y) = 0$ e supponiamo che la funzione obiettivo $f(x,y)$ assuma un minimo o un massimo locale in (x_0, y_0) (in confronto ai punti vicino sulla curva vincolo). Si ha allora che

Il gradiente $\nabla f(x_0, y_0)$ è perpendicolare alla curva di vincolo in (x_0, y_0).

Il primo esempio fatto illustra la situazione. In tutti e quattro i punti "candidati" ad essere massimi o minimi le curve di livello di f sono parallele alla curva di vincolo. I calcoli alla fine del secondo esempio spiega il perché di questo fatto.

7.5.2 Moltiplicatori di Lagrange

Ciò che abbiamo cercato di indicare sopra ci dice una cosa importane. Se (x_0, y_0) è un punto di massimo o minimo vincolato allora entrambi i vettori $\nabla f(x_0, y_0)$ e $\nabla g(x_0, y_0)$ sono perpendicolari alla curva di livello $g(x, y) = 0$. Ne segue che questi due vettori sono paralleli tra di loro, cioè multipli scalari uno dell'altro. Scriviamo formalmente il risultato nel caso bidimensionale, ricordando però che esso vale in qualunque dimensione.

Teorema 7.37 *(**Moltiplicatori di Lagrange**). Siano $f(x, y)$ e $g(x, y)$ funzioni da $\mathbb{R}^2 \to \mathbb{R}$. Consideriamo il problema di ottimizzare $f(x, y)$ soggetta al vincolo $g(x, y) = 0$. Se f assume un massimo od un minimo vincolato in (x_0, y_0) allora esiste $\lambda \in \mathbb{R}$ tale che*

$$\nabla f(x_0, y_0) = \lambda g(x_0, y_0) \ .$$

*Lo scalare λ è chiamato **moltiplicatore di Lagrange.***

Vediamo come funziona il Teorema in alcuni casi semplici.

Esempio 7.38 *Ottimizzare $f(x, y) = x + y$ soggetta al vincolo $x^2 + y^2 = 9$.*

Soluzione Se scriviamo $x^2 + y^2 - 9$, allora il vincolo diventa $g(x, y) = 0$ come nel teorema. (Questo trucco funziona sempre; notare che la costante 9 è "assorbita" in g. Si ha allora $\nabla f(x, y) = (1, 1)$ e $\nabla g(x, y) = (2x, 2y)$.
Il teorema afferma che il massimo o minimo vincolato si ha, se esiste, nei punti (x, y) nei quali è

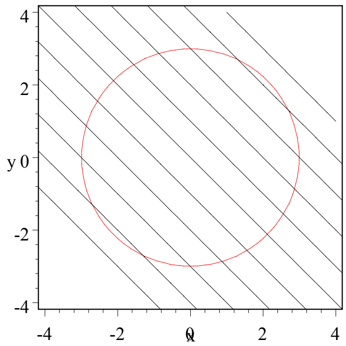
$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

per qualche valore dello scalare λ. Inoltre deve essere soddisfatta l'equazione del vincolo. Si ha allora

$$(1, 1) = \lambda (2x, 2y) \ , \text{ e } x^2 + y^2 = 9.$$

Si ottengono così tre equazioni nelle tre incognite x, y, λ.
La soluzione di questo sistema si ottiene, per esempio, notando che $(1, 1) = \lambda (2x, 2y)$ implica $x = y$. Sostituendo nell'equazione del vincolo si ottiene $2x^2 = 9$, o $x = \pm 3/\sqrt{2}$.(non importa trovare λ, l'importante è trovare x, y).
I punti candidati ad essere massimo o minimo sono allora $\left(3/\sqrt{2}, 3/\sqrt{2}\right)$ e $\left(-3/\sqrt{2}, -3/\sqrt{2}\right)$.
I valori di f sono $f\left(3/\sqrt{2}, 3/\sqrt{2}\right) = 3/\sqrt{2}$ e $f\left(-3/\sqrt{2}, -3/\sqrt{2}\right) = -3/\sqrt{2}$.
Il primo è quindi il massimo vincolato, il secondo il minimo vincolato. ■

Il disegno che segue suggerisce la stessa conclusione



Mappa di contorno di $x + y$ e vincolo.

ATTENZIONE!

Il teorema è spesso utile, ma va usato con molta attenzione e con alcuni distinguo. E' specialmente importante capire cosa il teorema *non dice.*

La condizione è necessaria ma non sufficiente. Il teorema di Lagrange afferma che la condizione $\nabla f = \lambda \nabla g$ è necessaria perché un punto (x_0, y_0) sia di massimo o minimo vincolato, ma NON sufficiente (Ricordate che anche nel caso di una funzione di una sola variabile la condizione $f'(x) = 0$ era necessaria ma non sufficiente).

Potrebbe non esserci soluzione. Non tutti i problemi di ottimo vincolato hanno soluzione. Anche in questo caso, comunque, la condizione di Lagrange può essere utile.

Esempio 7.39 *Ottimizzare $f(x, y) = x + y$ soggetta al vincolo $g(x, y) = y = 0$.*

Soluzione. E' chiaro che $f(x, 0) = x$ può assumere tutti i valori, quindi f non ha né massimo ne minimo vincolato. D'altra parte se scriviamo la condizione di Lagrange si ha $\nabla f(x, y) = (1, 1) = \lambda (1, 0) = \nabla g(x, y)$ che è chiaramente impossibile, quindi il teorema ci dice che non esiste né massimo né minimo. ■

Quando esiste una soluzione? Il problema di ottimo dell'esempio precedente non aveva soluzione. Il fatto è che il vincolo illimitato, lascia libera la funzione di crescere senza limiti.
La teoria generale (che supera i nostri scopi) garantisce tuttavia che se f e g sono funzioni differenziabili, ed il vincolo $g(x, y) = 0$ è limitato, allora f assume (finito) un massimo e minimo vincolato. In questo caso, il teorema garantisce che questi valori devono occorrere dove la condizione di Lagrange è soddisfatta

Se il vincolo è illimitato, come nell'esempio precedente, allora la funzione obiettivo *può o meno* assumere un massimo o minimo vincolato; dipende dalla funzione obiettivo e non c'è una regola semplice per decidere.

Risolvere può essere difficile. Data la funzione $f(x, y)$ e il vincolo $g(x, y) = 0$, la condizione di Lagrange e l'equazione del vincolo formano un sistema di tre equazioni, non necessariamente lineari, in tre incognite. Nel caso di funzioni di tre variabili sono coinvolte quattro variabili. Risolvere tali sistemi può essere complicato o anche impossibile. Fortunatamente, molti problemi interessanti portano a sistemi di equazioni semplici. Il fatto che il particolare valore di λ usualmente non importi, può a volte aiutare.

Esempio 7.40 *Ottimizzare* $f(x, y, z) = x + y + z$ *vincolata da* $g(x, y, z) = x^2 + y^2 + z^2 = 3$.

Soluzione. Il vincolo è la sfera di raggio $\sqrt{3}$, quindi limitata e poiché le funzioni sono differenziabili il massimo e minimo vincolato esiste. Le condizioni di Lagrange sono

$$\nabla f = (1, 1, 1) = \lambda \nabla g = \lambda (2x, 2y, 2z)$$

da cui segue immediatamente $x = y = z$. Mettendo questo risultato nell'equazione del vincolo si ha

$$x^2 + y^2 + z^2 - 3 = 0 \Longrightarrow 3x^2 = 3 \Longrightarrow x = \pm 1 .$$

Allora i possibili candidati sono i punti $(1, 1, 1)$ e $(-1, -1, -1)$ che sono in realtà massimo e minimo rispettivamente. ■

7.5.3 Esercizi

1. Usa il metodo dei moltiplicatori di Lagrange per risolvere l'esercizio, quando possibile. Quindi rifare l'esercizio usando metodi elementari, usando il vincolo per riscrivere la funzione obiettivo come funzione di una variabile.

   (a) $f(x, y) = xy$, soggetta a $g(x, y) = x + y - 1 = 0$.

   (b) $f(x, y) = x + y$, soggetta a $g(x, y) = xy - 1 = 0$.

2. Usa il metodo dei moltiplicatori di Lagrange per risolvere l'esercizio, quando possibile. Se non esistono massimi e minimi, spiegarne il motivo. Calcolare ∇f e ∇g in ogni punto di minimo e massimo vincolato.

   (a) $f(x, y) = x - y$, soggetta a $g(x, y) = x^2 + y^2 - 1 = 0$.

   (b) $f(x, y) = xy$, soggetta a $g(x, y) = x^2 + y^2 - 1 = 0$.

   (c) $f(x, y) = x^2 + y^2$ soggetta a $g(x, y) = x + y - 2 = 0$

   (d) $f(x, y) = x^2 + xy + y^2$ soggetta a $g(x, y) = x + y - 2 = 0$

   (e) $f(x, y) = x + 2y$ soggetta a $g(x, y) = x^2 + y^2 - 9 = 0$.

   (f) $f(x, y, z) = 2x + y + z$ soggetta a $g(x, y) = x^2 + y^2 + z^2 - 6 = 0$.

3. Il contadino Rossi ha una rete di 100 m e vuole usarla per racchiudere una porcilaia rettangolare. Aiutare il contadino Rossi a risolvere il problema (usare i moltiplicatori di Lagrange per risolvere il problema).

4. Il contadino Bianchi ha una rete di 100 m e vuole usarla per racchiudere una porcilaia a forma di triangolo rettangolo. Aiutare il contadino Bianchi a risolvere il problema [**Sugg.:**Usare i moltiplicatori di Lagrange per risolvere il problema].