CAPITOLO 3

Calcolo differenziale per funzioni di più variabili reali

Scopo di questo capitolo è studiare le principali caratteristiche dei grafici di funzioni di più variabili, con particolare attenzione allo studio dei massimi e minimi.

Una funzione f dall'insieme A all'insieme B (brevemente: $f : A \to B$) è una legge che ad ogni elemento di A fa corrispondere al più un elemento di B. L'insieme A è detto dominio di f e B codominio di f. Il grafico della funzione è il sottoinsieme di $A \times B$ dato da

$$\{(a, b) \in A \times B \,:\, a \in \operatorname{dom} f \ \text{ e } \ f(a) = b\}$$

L'immagine di f è il sottoinsieme di B formato dagli elementi "raggiunti" da A tramite f. In formule

$$\operatorname{Im} f = \{b \in B \,:\, \exists\, a \in A \text{ tale che } f(a) = b\}\,.$$

Tratteremo funzioni definite su (sottoinsiemi di) $\mathbb{R}^n$ a valori in $\mathbb{R}^m$. Nel caso in cui $m = 1$, si parla di funzioni reali di n variabili reali. Nel caso in cui $m > 1$, si parla di funzioni a valori vettoriali. Il grafico di queste funzioni è dunque un sottoinsieme di $\mathbb{R}^n \times \mathbb{R}^m$; siccome possiamo visualizzare oggetti al più tridimensionali, oltre al caso di funzioni reali di una variabile reale (quelle di Analisi I), potremo disegnare grafici di funzioni reali di due variabili oppure di funzioni a valori in $\mathbb{R}^2$ di una variabile.

È consuetudine, nel caso di funzioni di una o più variabili reali, non specificare il sottoinsieme A su cui è definita la funzione, ma indicare solo il numero di variabili in gioco (diciamo n) e la legge (da cui si desume il numero di variabili per il codominio, diciamo m). In tal caso, si intende che il codominio è $\mathbb{R}^m$ e che il dominio di f sia il più grande sottoinsieme A di $\mathbb{R}^n$ su cui la legge ha senso, ovvero A è il sottoinsieme degli elementi a cui è assegnato un elemento di $\mathbb{R}^m$.

Solitamente disegneremo i grafici di funzioni reali di due variabili, usando i pacchetti di Maple o di `MatLab` .

Usiamo la seguente notazione: indichiamo con $\mathbf{x} = (x_1, \ldots, x_n)$ il punto di $\mathbb{R}^n$ di coordinate $(x_1, \ldots, x_n)$. È facile calcolare la distanza tra due punti $\mathbf{x} = (x_1, \ldots, x_n)$ e $\mathbf{y} = (y_1, \ldots, y_n)$ utilizzando più volte il teorema di Pitagora:

$$\operatorname{dist}(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}.$$

In particolare, la distanza di $\mathbf{x}$ dall'origine $\mathbf{0} = (0, \dots, 0)$ è $\sqrt{\sum_{k=1}^{n} x_k^2}$; talvolta si chiama norma di $\mathbf{x}$ e si indica con $\|\mathbf{x}\|$.

Indicheremo solitamente un punto di $\mathbb{R}^2$ con la coppia di coordinate (x, y), anziché (x_1, x_2) e un punto di $\mathbb{R}^3$ con la terna di coordinate (x, y, z), anziché (x_1, x_2, x_3).

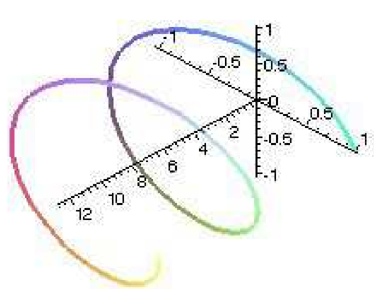
Se $\mathbf{x}$ e $\mathbf{y}$ sono due punti di $\mathbb{R}^n$, allora $\mathbf{x} + \mathbf{y}$ indica la somma dei due vettori $\mathbf{x}$ e $\mathbf{y}$, ossia $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n)$. Se α è un numero reale, $\alpha\mathbf{x}$ indica la moltiplicazione dello scalare α per il vettore $\mathbf{x}$, quindi $\alpha\mathbf{x} = (\alpha x_1, \dots, \alpha x_n)$.

Facilmente possiamo controllare le seguenti proprietà:

$$\|\alpha\mathbf{x}\| = |\alpha|\, \|\mathbf{x}\| \qquad \mathrm{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \qquad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

☞Esempio 3.1. Si disegni il grafico della funzione $f : [0, 4\pi] \to \mathbb{R}^2$ data dalla formula

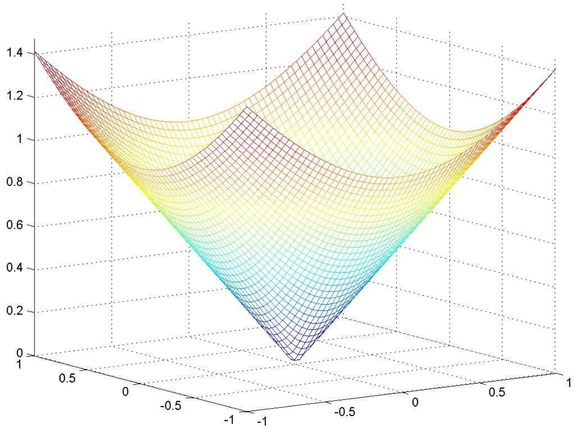
$$f(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \quad \forall t \in [0, 4\pi].$$



Si tratta di una funzione a valori vettoriali. Nell'immagine ci sono tutti i vettori di modulo 1, per le note proprietà di seno e coseno. Il grafico è un'elica che si avvolge due volte attorno all'asse delle t. ☺

☞Esempio 3.2. Si determinino grafico e immagine della funzione distanza in $\mathbb{R}^2$, ovvero della funzione

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



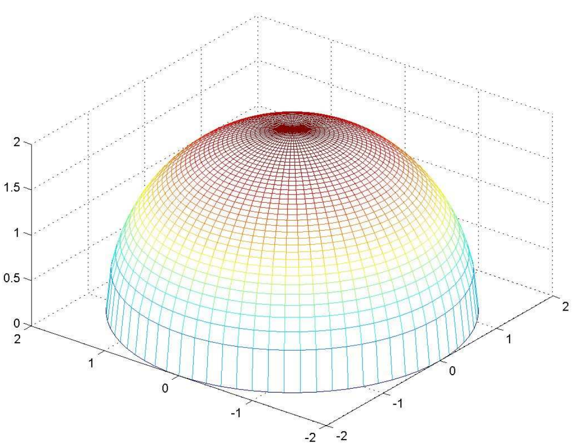
Essa ha dominio $\mathbb{R}^2$ e assume tutti i valori reali non negativi. Inoltre è costante sulle circonferenze centrate nell'origine e $f(x, 0) = |x|$. Pertanto il suo grafico è il cono disegnato in figura. ☺

Un metodo semplice per rendersi conto di come può essere fatto il grafico di una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ è quello di studiare le cosiddette linee di livello, ovvero i sottoinsiemi del piano del tipo

$$\{(x, y) \in \operatorname{dom} f \ : \ f(x, y) = k\}$$

al variare di k in $\mathbb{R}$.

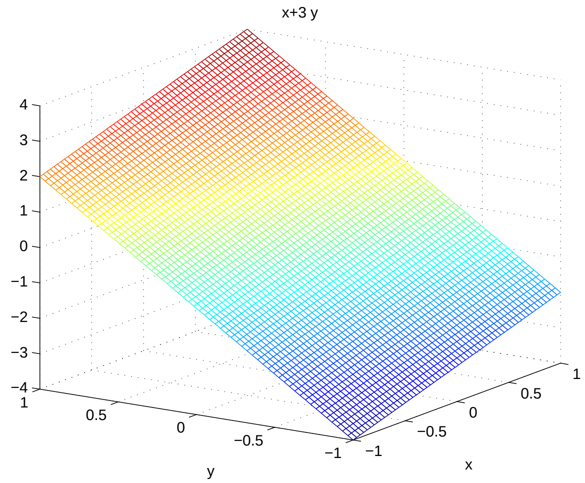
☞ESEMPIO 3.3. Determinare grafico e immagine della funzione $f(x, y) = \sqrt{4 - x^2 - y^2}$.



Essa ha dominio $\{(x, y) : x^2 + y^2 < 4\}$ e assume tutti i valori in $[0, 2]$. Infatti se $k \in \mathbb{R}$ è tale che $k = \sqrt{4 - x^2 - y^2}$, allora k deve essere maggiore o uguale a 0. Quadrando otteniamo $x^2 + y^2 = 4 - k^2$, quindi $k \leq 2$ e inoltre la curva di livello k è una circonferenza centrata nell'origine. Il grafico di f è la mezza sfera disegnata in figura.

☺

☞ESEMPIO 3.4. Determinare grafico e immagine della funzione $f(x, y) = 3x + y$.



Essa ha dominio $\mathbb{R}^2$ e assume tutti i valori in $\mathbb{R}$. Inoltre per $k \in \mathbb{R}$ la curva di livello k è una retta di coefficiente angolare -3. Il grafico di f è il piano inclinato disegnato in figura.

☺

In generale, una funzione che dipende solo dalla distanza dall'origine ha per linee di livello le circonferenze centrate nell'origine. Una funzione lineare ha per linee di livello rette parallele.

Supponiamo di voler ottenere il grafico dell'esempio 3.2 con Maple. Usiamo il comando

```
plot3d(sqrt(x ^2 + y^2),x=-1..1,y=-1..1);
```

Selezionando il grafico è possibile ruotarlo, cambiarne i colori, ecc. Possiamo anche memorizzare la funzione f e poi disegnarne il grafico scrivendo

```
f:=(x,y)->sqrt(x^2+y^2);
plot3d(f,-1..1,-1..1);
```

Per disegnare le linee di livello, occorre caricare il pacchetto `plots` (si fa una volta sola nel file) e poi usare il comando `contourplot`:

```
with(plots);
f:=(x,y)->sqrt(x^2+y^2);
contourplot(f,-1..1,-1..1);
```

I grafici di queste pagine sono stati ottenuti con `MatLab` . Ci sono diversi comandi, ma i principali sono `surf` (disegna la superficie), `mesh` (disegna la griglia) e `contour` (disegna le linee di livello) e le versioni "easy" `ezsurf`, `ezmesh` e `ezcontour`. Le versioni "easy" possono essere usate anche direttamente nella command window; il grafico del cono è stato ottenuto con il comando `ezmesh('(x^2 + y^2)^(1/2)',[-1,1,-1,1])`.

☞Esempio 3.5. Un esempio di funzione di $\mathbb{R}$ in $\mathbb{R}^n$ è legato alla forma parametrica della retta.

Tutti sappiamo che è unica la retta passante per due punti distinti P e Q; il vettore $\mathbf{v} = (Q - P)$ individua la direzione della retta; inizialmente supponiamo che P sia l'origine.

La retta in $\mathbb{R}^n$ passante per l'origine e con vettore direzionale $\mathbf{v}$ in $\mathbb{R}^n \setminus \{\mathbf{0}\}$ è costituita da tutti e soli i punti $\mathbf{x}$ di $\mathbb{R}^n$ allineati con $\mathbf{v}$. Pertanto $\mathbf{x}$ è sulla retta se e solo se $\mathbf{x}$ è un multiplo di $\mathbf{v}$, in formule: se e solo se esiste $t \in \mathbb{R}$ tale che $\mathbf{x} = t\,\mathbf{v}$.

La retta passante per il generico punto $\mathbf{x}_0$ e con vettore direzionale $\mathbf{v} \neq \mathbf{0}$ si ottiene traslando quella passante per l'origine, quindi $\mathbf{x}$ è su questa retta se e solo se esiste $t \in \mathbb{R}$ tale che $\mathbf{x} = \mathbf{x}_0 + t\,\mathbf{v}$. Questa è la forma parametrica della retta.

Il grafico della funzione $\mathbf{f} : \mathbb{R} \to \mathbb{R}^n$ definita da $\mathbf{f}(t) = \mathbf{x}_0 + t\,\mathbf{v}$ è quindi quello della retta in $\mathbb{R}^{n+1}$ passante per il punto $(0, \mathbf{x}_0)$ di $\mathbb{R}^{n+1}$ e con vettore direzionale $(1, \mathbf{v})$.

Viceversa, ogni retta di $\mathbb{R}^{n+1}$ con vettore direzionale $\mathbf{w}$ con $w_1 \neq 0$ si può pensare come grafico di una opportuna funzione.

Infatti, a meno di dividere ciascuna componente per w_1, possiamo supporre che $\mathbf{w} = (1, w_2, \ldots, w_{n+1}) = (1, \mathbf{w}')$ con $\mathbf{w}'$ in $\mathbb{R}^n$. Inoltre passerà per un punto $\mathbf{y}_0$ di $\mathbb{R}^{n+1}$ e quindi sarà costituita dai punti $\mathbf{x}$ di $\mathbb{R}^{n+1}$ per cui $\mathbf{x} = \mathbf{y}_0 + t\,\mathbf{w}$ con t reale.

In particolare la retta passa per un punto con prima coordinata nulla: $x_1 = y_{0_1} + t = 0$ ha soluzione $t = -y_{0_1}$. Quindi la retta passa per il punto $\mathbf{y}_0 - y_{0_1}\,\mathbf{w}$ che è del tipo $(0, \mathbf{z}_0)$ con $\mathbf{z}_0$ in $\mathbb{R}^n$ e ha vettore direzionale $(1, \mathbf{w}')$. In forma parametrica $x_1 = 0 + t \cdot 1 = t, x_j = z_{0_j} + tw'_j$, ovvero la retta è il grafico della funzione $\mathbf{f} : \mathbb{R} \to \mathbb{R}^n$ definita da $\mathbf{f}(t) = \mathbf{z}_0 + t\,\mathbf{w}'$. ☺

1. Limiti e continuità

Il concetto di limite per funzioni di più variabili è simile a quello per funzioni di una variabile.

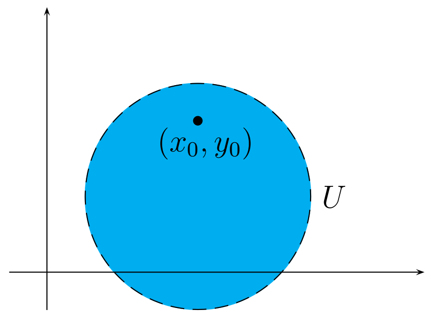
1.1. Intorni e punti di accumulazione. Iniziamo da alcune considerazioni sui sottoinsiemi di $\mathbb{R}^n$.

Definizione 3.1. Sia $\mathbf{x}_0 \in \mathbb{R}^n$. Un sottoinsieme U di $\mathbb{R}^n$ della forma

$$U = \{\mathbf{y} \in \mathbb{R}^n \,:\, \|\mathbf{y} - \mathbf{x}\| = \operatorname{dist}(\mathbf{x}, \mathbf{y}) < r\} \qquad \text{dove } r > 0 \quad \text{e} \quad \mathbf{x} \in \mathbb{R}^n \tag{3.1}$$

si dice intorno aperto del punto $\mathbf{x}_0$ se esso contiene $\mathbf{x}_0$.

Nel caso $n = 1$, gli intorni aperti di x_0 sono gli intervalli aperti che contengono x_0.



Nel caso $n = 2$, gli intorni aperti di (x_0, y_0) sono i dischi aperti (cioè senza il bordo) che contengono (x_0, y_0) all'interno.

In generale, chiamiamo disco aperto di centro $\mathbf{x}$ e raggio r un insieme della forma(3.1).

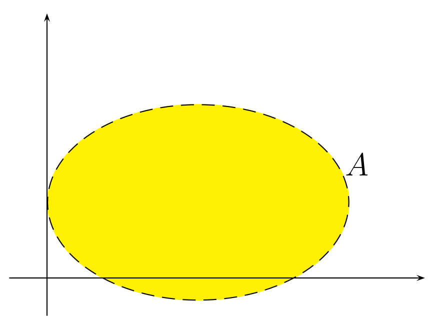
DEFINIZIONE 3.2. Un insieme V è un intorno di ∞ se è della forma

$$V = \{\mathbf{x} \in \mathbb{R}^n \; : \; \|\mathbf{x}\| > R\} \qquad \text{con } R > 0.$$

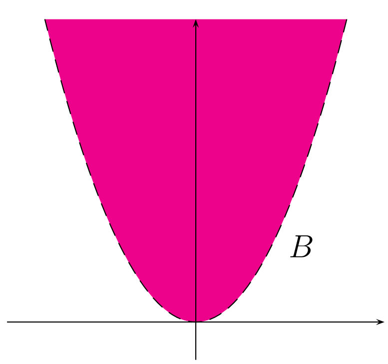
Nel caso $n = 2$, gli intorni di infinito sono i complementari dei dischi chiusi (ovvero considerati con il loro bordo) centrati nell'origine.

DEFINIZIONE 3.3. Sia A un sottoinsieme di $\mathbb{R}^n$. Un punto $\mathbf{x}$ di $\mathbb{R}^n \cup \{\infty\}$ si dice punto di accumulazione per A se ogni suo intorno interseca A in punti distinti da $\mathbf{x}$ stesso.

Come nel caso di funzioni di una variabile, i punti di accumulazione del dominio di una funzione sono quelli in cui possiamo calcolare i limiti.



Sia A l'interno dell'ellisse, come disegnato in figura. I punti di accumulazione di A sono i punti interni e dell'ellisse.



Sia $B = \{(x, y) \in \mathbb{R}^2 \; : \; x^2 < y\}$.

I punti di accumulazione di B sono ∞ e

$$\{(x, y) \in \mathbb{R}^2 \; : \; x^2 \leq y\}.$$

1.2. Definizione di limite. La definizione di limite è la stessa vista lo scorso anno, in termini dei "nuovi" intorni.

DEFINIZIONE 3.4. Sia $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ e sia $\mathbf{x}_0$ un punto di accumulazione per il dominio di $\mathbf{f}$. Dato $\mathbf{L} \in \mathbb{R}^m \cup \{\infty\}$ si dice che

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{L}$$

se per ogni intorno V di $\mathbf{L}$ esiste un intorno U di $\mathbf{x}_0$ tale che per ogni $\mathbf{x}$ in $U \cap \operatorname{dom} \mathbf{f}$ tranne al più x_0 si ha che $\mathbf{f}(\mathbf{x}) \in V$.

Non è difficile verificare che se $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ e $\mathbf{L} = (\ell_1, \dots, \ell_m) \in \mathbb{R}^m$, allora si ha

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{L} \quad \text{se e solo se} \quad \lim_{\mathbf{x}\to\mathbf{x}_0} f_k(\mathbf{x}) = \ell_k \qquad \forall k = 1, \dots, m.$$

In particolare, nel caso di $f : \mathbb{R}^n \to \mathbb{C} \simeq \mathbb{R}^2$, allora si ha $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{L}$ se e solo se $\lim_{\mathbf{x}\to\mathbf{x}_0} \operatorname{Re}(\mathbf{f}(\mathbf{x})) = \operatorname{Re}(\mathbf{L})$ e $\lim_{\mathbf{x}\to\mathbf{x}_0} \operatorname{Im}(\mathbf{f}(\mathbf{x})) = \operatorname{Im}(\mathbf{L})$.

Inoltre $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \infty$ se e solo se $\lim_{\mathbf{x}\to\mathbf{x}_0} \|\mathbf{f}(\mathbf{x})\| = \infty$. Sembra una complicazione, ma notiamo che la funzione $\mathbf{x} \mapsto \|\mathbf{f}(\mathbf{x})\|$ è a valori scalari.

Per queste ragioni, nel seguito, tranne quando diversamente specificato, ci limitiamo a considerare il caso $m = 1$.

Inoltre, sempre nel caso $m = 1$, possiamo dare senso alla definizione di limite anche se $\ell = \pm\infty$.

1.3. Calcolo di limiti. Valgono le principali proprietà dei limiti viste per funzioni di una variabile: unicità del limite, permanenza del segno, regole di calcolo dei limiti. Vorrei ricordare una regola di limite per funzioni composte: se $g : \mathbb{R} \to \mathbb{R}$ è una funzione continua in ℓ e $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \ell$, allora $\lim_{\mathbf{x}\to\mathbf{x}_0} g(f(\mathbf{x})) = g(\ell)$. Vediamo ora alcuni esempi.

☞ESEMPIO 3.6. È banale verificare che $\lim_{(x,y)\to(x_0,y_0)} 1 = 1$. Proviamo che, nel caso (x_0, y_0) sia in $\mathbb{R}^2$, vale $\lim_{(x,y)\to(x_0,y_0)} x = x_0$.

Fissato un intorno di x_0 del tipo $V = (x_0 - \varepsilon, x_0 + \varepsilon)$ si ha $x \in V$ (cioè $|x - x_0| < \varepsilon$) per ogni (x, y) nell'intorno $U = \left\{(x, y) \,:\, \sqrt{(x - x_0)^2 + (y - y_0^2} < \varepsilon\right\}$. Infatti se (x, y) è in U, allora $-\varepsilon < |x - x_0| \le \sqrt{(x - x_0)^2 + (y - y_0^2} < \varepsilon$, cioè $x \in V$.

Analogamente $\lim_{(x,y)\to(x_0,y_0)} y = y_0$.

Dalle regole per il calcolo dei limiti, nel caso (x_0, y_0) sia in $\mathbb{R}^2$, segue che $\lim_{(x,y)\to(x_0,y_0)} P(x, y) = P(x_0, y_0)$ per ogni polinomio P. ☺

☞Esempio 3.7. Calcolare $\lim_{(x,y)\to(0,0)} \dfrac{\sin(x^2+y^2)}{x^2+y^2}$. Si ricordi che la funzione

$$g(t) = \begin{cases} \frac{\sin t}{t} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

è continua. Per le regole di limite del prodotto e della somma,

$$\lim_{(x,y)\to(0,0)} x^2 + y^2 = 0.$$

Per la regola di limite di funzioni composte appena ricordata, con $f(x,y) = x^2 + y^2$, si ha

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{(x^2+y^2)} = \lim_{(x,y)\to(0,0)} g(f(x,y)) = g(0) = 1.$$

☺

Nel caso di funzioni di una variabile abbiamo introdotto la nozione di limite sinistro e limite destro. Quando $n > 1$ ci si può avvicinare al punto $\mathbf{x}_0$ in un'infinità di modi e, se il limite esiste, in ognuno di questi modi dobbiamo ottenere sempre lo stesso valore limite.

Illustriamo più diffusamente la situazione nel caso $n = 2$. Supponiamo che $f : \mathbb{R}^2 \to \mathbb{R}$, (x_0, y_0) sia un punto di accumulazione per il dominio di f e $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = \ell$. Ci possiamo avvicinare al punto (x_0, y_0) lungo una qualsiasi curva che giaccia nel dominio di f. Se valutiamo il limite della funzione lungo una qualsiasi di queste curve, abbiamo da valutare il limite di una funzione di una sola variabile (che dovremmo saper fare). Ebbene: questo limite deve essere uguale a ℓ.

Più formalmente, utilizzando il limite della funzione composta nella forma di questa nota [1] si dimostra che

Proposizione 3.8 (Limite lungo una curva). *Sia $f : \mathbb{R}^n \to \mathbb{R}$ e sia $\mathbf{x}_0$ un punto di accumulazione per il dominio di f. Se (per ℓ in $\mathbb{R} \cup \{\pm\infty\}$)*

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \ell$$

allora per ogni $\mathbf{g} : (0, \varepsilon) \to \mathbb{R}^n$ tale che $\lim_{t\to 0^+} \mathbf{g}(t) = \mathbf{x}_0$, $\mathbf{g}(0,\varepsilon) \subset \operatorname{dom} f \setminus \{\mathbf{x}_0\}$ si ha

$$\lim_{t\to 0} f(\mathbf{g}(t)) = \ell.$$

Questa proprietà è soprattutto utile quando vogliamo dimostrare che un limite assegnato non esiste, come nell'esempio seguente. Inoltre ci permette di stabilire quanto possa valere un determinato limite.

[1] Siano $\mathbf{g} : (0, \varepsilon) \to \mathbb{R}^n$ e $f : \mathbb{R}^n \to \mathbb{R}$ due funzioni tali che l'immagine $\mathbf{g}(0,\varepsilon)$ sia contenuta nel dominio di f e $\lim_{t\to 0^+} \mathbf{g}(t) = \mathbf{x}_0$ con $\mathbf{g}(t) \neq \mathbf{x}_0$ per ogni t. Inoltre sia $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \ell$, allora $\lim_{t\to 0^+} f(\mathbf{g}(t)) = \ell$.

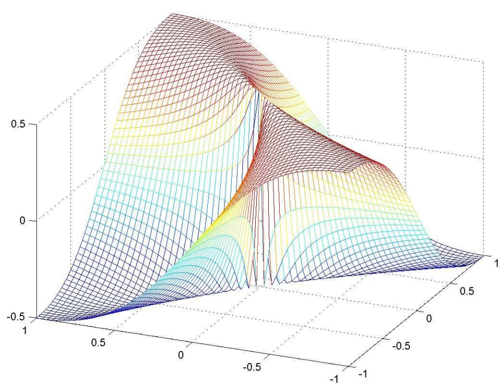
☞Esempio 3.9. Studiare $\lim_{(x,y)\to(0,0)} f(x,y)$ dove $f(x,y) = \frac{xy}{x^2+y^2}$.

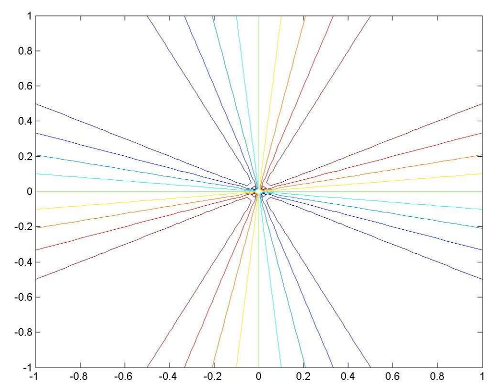
Innanzi tutto, il dominio di f è $\mathbb{R}^2$ privato di $(0,0)$, quindi l'origine è punto di accumulazione per il dominio di f.

Siccome la funzione sull'asse x (ossia in punti del dominio con $y = 0$) è nulla, possiamo iniziare pensando che se ci avviciniamo all'origine mantenendoci sull'asse x i valori della funzione sono prossimi a zero. In particolare, se il limite esiste, allora è 0.

Avviciniamoci all'origine lungo un'altra retta, ad esempio la retta $y = x$. Lungo questa retta abbiamo $f(x,y) = f(x,x) = \frac{1}{2}$, quindi se ci avviciniamo all'origine mantenendoci sulla retta $y = x$ ci sembra di dover star vicini al valore $\frac{1}{2}$.

Siccome non si può contemporaneamente star vicini a 0 e a $\frac{1}{2}$, ne dobbiamo concludere che il limite non esiste.





Nella figura di destra vediamo le linee di livello della funzione, che formano un fascio di rette per l'origine.

Per chiarire l'enunciato della proposizione, in questo esempio abbiamo scelto le funzioni $\mathbf{g}_1$ e $\mathbf{g}_2$ definite da

$$\mathbf{g}_1(t) = (t,0) \qquad \mathbf{g}_2(t) = (t,t) \qquad \forall t \in \mathbb{R}$$

e abbiamo notato che, anche se $\lim_{t\to 0} \mathbf{g}_j(t) = (0,0)$ per ogni $j = 1,2$, si ha

$$\lim_{t\to 0} f(\mathbf{g}_1(t)) = 0 \neq \frac{1}{2} = \lim_{t\to 0} f(\mathbf{g}_2(t)).$$

☺

Si potrebbe pensare che se esiste il limite su ogni retta e questi limiti sono uguali, allora esiste il limite della funzione, perché in questo modo esauriamo tutte le possibili direzioni con cui avvicinarsi al punto. Questo è assolutamente FALSO!!! Guardate cosa succede nel prossimo esempio. Ci sono infiniti modi di avvicinarsi a un punto dato, non esistono solo le rette.

☞Esempio 3.10. Studiare $\lim_{(x,y)\to(0,0)} f(x,y)$ dove $f(x,y) = x^2y/(x^4+y^2)$.

Innanzi tutto, il dominio di f è $\mathbb{R}^2$ privato di $(0,0)$, quindi l'origine è punto di accumulazione per il dominio di f.

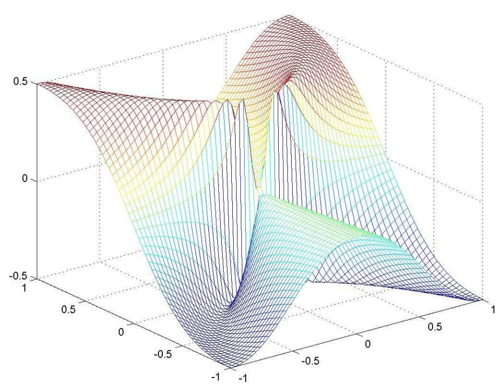
Siccome la funzione sull'asse x (ossia in punti del dominio con $y=0$) è nulla, possiamo iniziare pensando che se ci avviciniamo all'origine mantenendoci sull'asse x i valori della funzione sono prossimi a zero. In particolare, se il limite esiste, allora è 0.

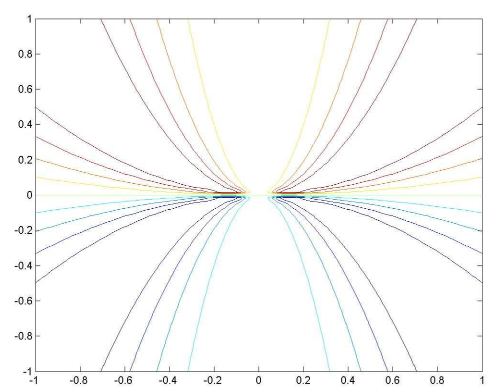
Avviciniamoci all'origine lungo un'altra curva, ad esempio la parabola $y = x^2$. Lungo questa curva abbiamo $f(x,y) = f(x,x^2) = \frac{1}{2}$, quindi se ci avviciniamo all'origine mantenendoci sulla parabola $y = x^2$ ci sembra di dover star vicini al valore $\frac{1}{2}$. Formalmente, se $\mathbf{g}_1(t) = (t,0)$ e $\mathbf{g}_2(t) = (t,t^2)$, allora $\lim_{t\to 0} \mathbf{g}_j(t) = (0,0)$ per $j = 1,2$, ma

$$\lim_{t\to 0} f(\mathbf{g}_1(t)) = 0 \neq \frac{1}{2} = \lim_{t\to 0} f(\mathbf{g}_2(t)).$$

Come prima, concludiamo che il limite non esiste.

Notare però che per ogni $\theta \in \mathbb{R}$ si ha $\lim_{t\to 0} f(t\cos\theta, t\sin\theta) = 0$, ovvero il limite valutato lungo qualsiasi retta per l'origine $t \mapsto (t\cos\theta, t\sin\theta)$ è 0.





Nella figura di destra vediamo le curve di livello della funzione, che sono parabole del tipo $y = a\,x^2$ con $a \in \mathbb{R}$. ☺

Abbiamo detto che la Proposizione 3.8 è anche utile perché ci permette di stabilire quanto possa valere un determinato limite: supponiamo ad esempio che $\mathbf{x}_0$ sia in $\mathbb{R}^n$ e che f sia definita in un intorno bucato di $\mathbf{x}_0$ (ovvero in un intorno di $\mathbf{x}_0$ tranne al più proprio $\mathbf{x}_0$). Allora possiamo restringere f a una retta per $\mathbf{x}_0$. Se lungo tale retta il limite risulta essere ℓ, allora il limite da calcolarsi, se esistente, deve valere ℓ. È il contenuto del seguente

Corollario 3.11. *Sia* $\mathbf{x}_0$ *in* $\mathbb{R}^n$ *e esista* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$*. Se, per un certo vettore* $\mathbf{u}$ *in* $\mathbb{R}^n$ *di lunghezza* 1*, i punti del tipo* $\mathbf{x}_0 + r\mathbf{u}$ *sono nel dominio di* f *per* $r \to 0^+$ *e*

$$\lim_{r\to 0^+} f(\mathbf{x}_0 + r\mathbf{u}) = \ell$$

allora

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \ell.$$

Nel caso di $\mathbb{R}^2$, un vettore $\mathbf{u}$ di lunghezza 1 è della forma

$$\mathbf{u} = (\cos\theta, \sin\theta) \qquad \text{per un certo } \theta.$$

Possiamo capire quale è il candidato valore del limite calcolando a θ fissato

$$\lim_{r\to 0^+} f(x_0 + r\cos\theta, y_0 + r\sin\theta).$$

Come più volte detto, se il risultato dipende da θ, possiamo subito dire che il limite non esiste. Se il risultato non dipende da θ, il limite potrebbe esistere. La domanda ora è: come stabilire se esiste?

Per poter stabilire l'esistenza di un limite, quando non funzionano le regole di calcolo già note, cerchiamo di valutare il limite "uniformemente rispetto alla direzione". Il trucco consiste nello stimare la funzione (di cui si desidera calcolare il limite) con un'altra funzione più semplice perché radiale. Di funzioni radiali si sa calcolare il limite utilizzando il teorema per le funzioni composte. Per concludere si utilizza il criterio del confronto.

Teorema 3.12. *Sia* $f : \mathbb{R}^n \to \mathbb{R}$ *e supponiamo che un disco bucato di centro* $\mathbf{x}_0 \in \mathbb{R}^n$ *e raggio positivo sia contenuto nel dominio di* f *(quindi* $\mathbf{x}_0$ *è un punto di accumulazione per il dominio di* f*). Se* $\ell \in \mathbb{R}$ *è tale che*

$$\lim_{r\to 0^+}\left(\sup_{\|\mathbf{u}\|=1}\left|f(\mathbf{x}_0 + r\mathbf{u}) - \ell\right|\right) = 0$$

allora

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \ell.$$

☞Esempio 3.13. Calcolare, se esiste, $\lim_{(x,y)\to(0,0)} f(x,y)$, dove $f(x,y) = \frac{x^2(y+1)+y^2}{x^2+y^2}$.

Per iniziare, valutiamo a θ fissato in $[0, 2\pi]$

$$\begin{aligned}\lim_{r\to 0^+} f(r\cos\theta, r\sin\theta) &= \lim_{r\to 0^+} \frac{r^2\cos^2\theta(r\sin\theta+1) + r^2\sin^2\theta}{r^2} \\ &= \lim_{r\to 0^+} r\,\cos^2\theta\sin\theta + \cos^2\theta + \sin^2\theta \\ &= \lim_{r\to 0^+} r\,\cos^2\theta\sin\theta + 1 = 1.\end{aligned}$$

IL valore limite così ottenuto non dipende dalla direzione θ. Valutiamo quindi la quantità $|f(x,y)-1|$, usando coordinate polari. Il nostro scopo è quello di maggiorare questa quantità con una funzione che dipenda solo dalla distanza dal punto $(0,0)$, dove stiamo calcolando il limite, e, al tendere a 0 di questa distanza risulti infinitesima. Per il teorema del confronto risulterà anche $|f(x,y)-1| \to 0$. Nel nostro caso abbiamo

$$\begin{aligned}
\sup_{\theta\in[0,2\pi]} |f(r\cos\theta, r\sin\theta)-1| &= \sup_{\theta\in[0,2\pi]} \left|\frac{r^2\cos^2\theta(r\sin\theta+1)+r^2\sin^2\theta}{r^2}-1\right| \\
&= \sup_{\theta\in[0,2\pi]} \left|\frac{r^2\cos^2\theta(r\sin\theta+1)+r^2\sin^2\theta-r^2}{r^2}\right| \\
&= \sup_{\theta\in[0,2\pi]} \left|\frac{r^3\cos^2\theta\sin\theta}{r^2}\right| \\
&= \sup_{\theta\in[0,2\pi]} r\,\cos^2\theta \\
&\le r \xrightarrow[r\to 0^+]{} 0.
\end{aligned}$$

Fra parentesi, notiamo che nell'ultima riga in questo caso vale un'uguaglianza (cioè vale $\sup_{\theta\in[0,2\pi]} r\,\cos^2\theta = r$), ma per i nostri scopi anche una semplice stima dell'estremo superiore sarebbe stata sufficiente.

Quindi

$$0 \le \lim_{(x,y)\to(0,0)} |f(x,y)-1| \le \lim_{r\to 0} r = 0$$

e allora $\displaystyle\lim_{(x,y)\to(0,0)} |f(x,y)-1| = 0$, cioè $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y) = 1$. ☺

Dimostrazione. È sufficiente verificare che

$$\lim_{\mathbf{x}\to\mathbf{x}_0} |f(\mathbf{x})-\ell| = 0.$$

Poniamo

$$g(r) = \sup_{\substack{\mathbf{u}\in\mathbb{R}^n \\ \|\mathbf{u}\|=1}} \big|f(\mathbf{x}_0 + r\mathbf{u}) - \ell\big|.$$

Allora per ipotesi $\lim_{r\to 0^+} g(r) = 0$ e per il teorema sul limite della funzione composta (vedi nota) si ha che

$$\lim_{\mathbf{x}\to\mathbf{x}_0 0} g(\|\mathbf{x}-\mathbf{x}_0\|) = 0.$$

Siccome per ogni $\mathbf{x}$ vale

$$0 \le |f(\mathbf{x})-\ell| \le g(\|\mathbf{x}-\mathbf{x}_0\|)$$

la tesi segue dal teorema del confronto dei limiti. □

Nel caso in cui il candidato valore limite sia $\pm\infty$, il criterio va modificato nella maniera seguente.

Teorema 3.14. *Sia $f : \mathbb{R}^n \to \mathbb{R}$ e un disco bucato di centro $\mathbf{x}_0 \in \mathbb{R}^n$ e raggio positivo sia contenuto nel dominio di f (quindi $\mathbf{x}_0$ è un punto di accumulazione per il dominio di f).*

$$\lim_{r\to 0^+}\left(\inf_{\|\mathbf{u}\|=1} f(\mathbf{x}_0 + r\mathbf{u})\right) = +\infty \quad \Rightarrow \quad \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = +\infty.$$

$$\lim_{r\to 0^+}\left(\sup_{\|\mathbf{u}\|=1} f(\mathbf{x}_0 + r\mathbf{u})\right) = -\infty \quad \Rightarrow \quad \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = -\infty.$$

La verifica è lasciata per esercizio.

Come appena visto nell'esempio, il metodo per calcolare $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$, dove $(x_0, y_0) \in \mathbb{R}^2$, consiste nel:

1. passare in coordinate polari attorno al punto (x_0, y_0) dove si deve calcolare il limite — ovvero $x = x_0 + r\cos\theta$ e $y = y_0 + r\sin\theta$;
2. calcolare $\lim_{r\to 0^+} f(r\cos\theta, r\sin\theta)$; qualora tale limite dipenda da θ il procedimento si arresta, perché abbiamo stabilito che il lmite non esiste in base alla Proposizione 3.8; qualora questo limite non dipenda da θ — poniamo valga ℓ — si proceda con il punto 3;
3. nel caso in cui ℓ sia finito, calcolare o stimare uniformemente rispetto a θ
$$g(r) = \sup_{\theta\in[0,2\pi]} |f(x_0 + r\cos\theta, y_0 + r\sin\theta) - \ell|\,;$$
4. verificare che $\lim\limits_{r\to 0^+} g(r) = 0$. A questo punto concludiamo per il teorema del confronto
$$0 \le \lim_{(x,y)\to(x_0,y_0)} |f(x,y) - \ell| \le \lim_{r\to 0^+} g(r) = 0$$
che $\lim\limits_{(x,y)\to(x_0,y_0)} |f(x,y) - \ell| = 0$, ovvero che $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = \ell$.
Se invece $\lim\limits_{r\to 0^+} g(r) \neq 0$, può darsi che si sia fatta una stima troppo cruda; si provi a dimostrare la non esistenza del limite assegnato, adoperando restrizioni a curve opportune che non siano rette.

Il metodo va opportunamente modificato come segue nel caso in cui $\ell = \pm\infty$: se $\ell = +\infty$, allora occorre modificare i punti 3. e 4. in

3'. calcolare o stimare dal basso uniformemente rispetto a θ
$$g(r) = \inf_{\theta\in[0,2\pi]} f(x_0 + r\cos\theta, y_0 + r\sin\theta);$$

4'. verificare che $\lim\limits_{r\to 0^+} g(r) = +\infty$. Allora per il confronto
$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) \ge \lim_{r\to+\infty} g(r) = +\infty.$$

Nel caso in cui $\lim_{r\to 0^+} g(r) \neq +\infty$, si provi a dimostrare la non esistenza del limite assegnato, adoperando restrizioni a curve opportune che non siano rette.

(Analogamente se $\ell = -\infty$).

☞Esempio 3.15. Calcolare, se esiste, $\lim_{(x,y)\to(0,0)} f(x,y)$, dove $f(x,y) = \frac{\sin(|x|+2|y|)}{x^2+y^2}$.

Innanzi tutto semplifichiamo la situazione moltiplicando e dividendo per $|x|+2|y|$. Il vantaggio è che

$$\frac{\sin(|x|+2|y|)}{x^2+y^2} = \left|\frac{\sin(|x|+2|y|)}{|x|+2|y|}\right| \cdot \frac{|x|+2|y|}{x^2+y^2}$$

e il primo fattore tende a 1, quindi ci basta occuparci del secondo. Poniamo $h(x,y) = \frac{|x|+2|y|}{x^2+y^2}$ e passiamo in coordinate polari centrate in $(0,0)$. A θ fissato

$$\begin{aligned}\lim_{r\to 0^+} h(r\cos\theta, r\sin\theta) &= \lim_{r\to 0^+} \frac{|r\cos\theta| + 2r|\sin\theta|}{r^2}\\ &= \lim_{r\to 0^+} \frac{|\cos\theta| + 2|\sin\theta|}{r} = +\infty.\end{aligned}$$

Siccome la funzione $\theta \mapsto \varphi(\theta) = |\cos\theta| + 2|\sin\theta|$ è continua e $[0,2\pi]$ è un intervallo chiuso e limitato, per il Teorema di Weierstrass, φ ha minimo m; siccome φ è sempre strettamente positiva, il minimo è strettamente positivo e

$$\begin{aligned}\inf_{\theta\in[0,2\pi]} h(r\cos\theta, r\sin\theta) &= \inf_{\theta\in[0,2\pi]} \frac{|r\cos\theta| + 2r|\sin\theta|}{r^2}\\ &= \inf_{\theta\in[0,2\pi]} \frac{|\cos\theta| + 2|\sin\theta|}{r}\\ &\geq \frac{m}{r} \xrightarrow[r\to 0^+]{} +\infty.\end{aligned}$$

Quindi il limite è $+\infty$. In questo caso non abbiamo determinato precisamente l'estremo inferiore, ma lo abbiamo solamente stimato. ☺

Quando $\mathbf{x}_0 = \infty$ è sufficiente modificare il metodo visto facendo tendere r a infinito. Precisamente:

Corollario 3.16. *Supponiamo che esista* $\lim_{\mathbf{x}\to\infty} f(\mathbf{x})$. *Se, per un certo vettore* $\mathbf{u}$ *in* $\mathbb{R}^n$ *di lunghezza 1, i punti del tipo* $r\mathbf{u}$ *sono nel dominio di* f *per* $r \to +\infty$ *e*

$$\lim_{r\to+\infty} f(r\mathbf{u}) = \ell$$

allora

$$\lim_{\mathbf{x}\to\infty} f(\mathbf{x}) = \ell.$$

Teorema 3.17. *Sia $f : \mathbb{R}^n \to \mathbb{R}$ e supponiamo che il complementare di un disco centrato nell'origine sia contenuto nel dominio di f (quindi ∞ è un punto di accumulazione per il dominio di f). Se $\ell \in \mathbb{R}$ è tale che*

$$\lim_{r\to+\infty}\left(\sup_{\|\mathbf{u}\|=1}\left|f(r\mathbf{u})-\ell\right|\right)=0$$

allora

$$\lim_{\mathbf{x}\to\infty} f(\mathbf{x})=\ell.$$

Teorema 3.18. *Sia $f : \mathbb{R}^n \to \mathbb{R}$ e supponiamo che il complementare di un disco centrato nell'origine sia contenuto nel dominio di f (quindi ∞ è un punto di accumulazione per il dominio di f).*

$$\lim_{r\to+\infty}\left(\inf_{\|\mathbf{u}\|=1} f(r\mathbf{u})\right)=+\infty \quad\Rightarrow\quad \lim_{\mathbf{x}\to\infty} f(\mathbf{x})=+\infty.$$

$$\lim_{r\to+\infty}\left(\sup_{\|\mathbf{u}\|=1} f(r\mathbf{u})\right)=-\infty \quad\Rightarrow\quad \lim_{\mathbf{x}\to\infty} f(\mathbf{x})=-\infty.$$

In particolare, per $n = 2$ per calcolare un limite per $(x, y) \to \infty$, si scelgano coordinate polari centrate in $(0, 0)$ e si faccia tendere r a $+\infty$ nei punti 1, 2, 3 (o 3'), 4 (o 4').

☞Esempio 3.19. Verificare che $\lim_{(x,y)\to\infty} x$ non esiste. Infatti in coordinate polari

$$x = r\,\cos\theta \xrightarrow[r\to+\infty]{} \begin{cases} +\infty & \cos\theta>0\\ 0 & \cos\theta=0\\ -\infty & \cos\theta<0\end{cases}$$

Siccome il limite dipende dalla direzione, $\lim_{(x,y)\to\infty} x$ non esiste. ☺

☞Esempio 3.20. Calcolare, se esiste, $\lim_{(x,y)\to\infty} f(x, y)$, dove $f(x, y) = \frac{x}{x^2+y^2}$. In coordinate polari centrate in $(0, 0)$ valutiamo dapprima a θ fissato

$$\begin{aligned}\lim_{r\to+\infty} f(r\cos\theta, r\sin\theta) &= \lim_{r\to+\infty}\frac{r\cos\theta}{r^2}\\ &= \lim_{r\to+\infty}\frac{\cos\theta}{r}=0.\end{aligned}$$

Siccome il limite non dipende dalla direzione e è finito, calcoliamo

$$\sup_{\theta\in[0,2\pi]}|f(r\cos\theta, r\sin\theta)| = \sup_{\theta\in[0,2\pi]}\frac{|\cos\theta|}{r}=\frac{1}{r}\xrightarrow[r\to+\infty]{} 0.$$

Quindi il limite è 0. ☺

☞ESEMPIO 3.21. Calcolare, se esiste, $\lim_{(x,y)\to\infty} f(x,y)$, dove $f(x,y) = x^2 + y^4$. In coordinate polari centrate in $(0,0)$ valutiamo a θ fissato

$$\begin{aligned}\lim_{r\to+\infty} f(r\cos\theta, r\sin\theta) &= \lim_{r\to+\infty} r^2\cos^2\theta + r^4\sin^4\theta \\ &= \lim_{r\to+\infty} r^2(\cos^2\theta + r^2\sin^4\theta) = +\infty.\end{aligned}$$

Siccome il limite non dipende dalla direzione e è $+\infty$, calcoliamo

$$\begin{aligned}\inf_{\theta\in[0,2\pi]} f(r\cos\theta, r\sin\theta) &= \inf_{\theta\in[0,2\pi]} r^2(\cos^2\theta + r^2\sin^4\theta) \\ &= r^2 \inf_{\theta\in[0,2\pi]}(\cos^2\theta + r^2\sin^4\theta) \\ &\geq r^2 \inf_{\theta\in[0,2\pi]}(\cos^2\theta + r^2\sin^4\theta) \qquad \text{se } r\geq 1 \\ &\geq r^2 \inf_{\theta\in[0,2\pi]}(\cos^2\theta + \sin^4\theta) \\ &\geq m\,r^2\end{aligned}$$

dove il minimo $m > 0$ esiste per il Teorema di Weierstrass. Quindi

$$\lim_{r\to+\infty} \inf_{\theta\in[0,2\pi]} f(r\cos\theta, r\sin\theta) \geq \lim_{r\to+\infty} m\,r^2 = +\infty$$

e il limite assegnato vale $+\infty$. ☺

2. Continuità

DEFINIZIONE 3.5. Sia $f : \mathbb{R}^n \to \mathbb{R}$ e sia $\mathbf{x}_0$ un punto del dominio di f. Si dice che f è continua in $\mathbf{x}_0$ se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che se $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ e $\mathbf{x} \in \mathrm{dom} f$, allora $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \varepsilon$.

Questo equivale alla richiesta

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$$

nel caso in cui $\mathbf{x}_0$ sia un punto di accumulazione per il dominio di f. Per le regole di calcolo dei limiti, le funzioni polinomiali sono continue.

Per le funzioni continue valgono i seguenti teoremi (analoghi a quelli visti lo scorso anno) sulle generalizzazioni a più variabili degli intervalli e degli intervalli chiusi e limitati.

DEFINIZIONE 3.6. Un insieme $A \subseteq \mathbb{R}^n$ si dice connesso se comunque scegliamo due punti $\mathbf{a}, \mathbf{b}$ in A è possibile congiungere tali punti con una spezzata giacente in A.

Teorema 3.22 (degli zeri). *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione continua e sia A un insieme connesso. Inoltre esistano due punti $\mathbf{a}, \mathbf{b}$ in A tali che $f(\mathbf{a}) > 0$ e $f(\mathbf{b}) < 0$. Allora esiste uno zero di f, ovvero esiste $\mathbf{c}$ in A tale che $f(\mathbf{c}) = 0$.*

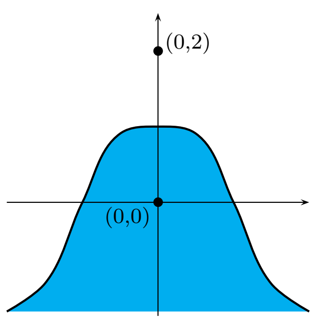
Corollario 3.23 (valori intermedi). *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione continua e sia A un insieme connesso. Dati $\mathbf{a}, \mathbf{b}$ in A, allora f assume in A tutti valori tra $f(\mathbf{a})$ e $f(\mathbf{b})$.*

Un importante uso del teorema degli zeri o del suo corollario è nello studio di disequazioni.

☞Esempio 3.24. Determinare e disegnare nel piano il dominio della funzione

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

Si ha $\operatorname{dom} f = \{(x, y) \in \mathbb{R}^2 : x^2 + y^3 - 1 \geq 0\}$. Dobbiamo quindi indicare i punti (x, y) di $\mathbb{R}^2$ per cui $x^2 + y^3 + 1 \geq 0$. Occupiamoci dapprima dei punti che soddisfano l'uguaglianza, ovvero $y = \sqrt[3]{1 - x^2}$.

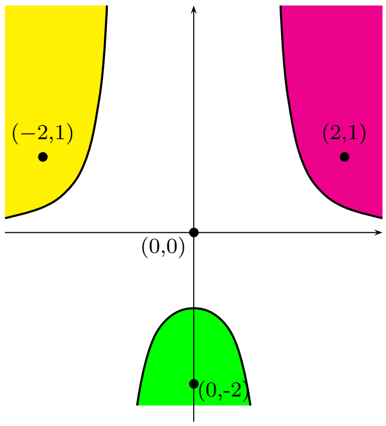


Il piano viene diviso dal grafico di $y = \sqrt[3]{1 - x^2}$ in due aperti connessi, quello bianco e quello azzurro. Scegliamo un punto qualsiasi in ognuno di questi aperti connessi. Ad esempio, scegliamo i punti $(0, 2)$ e $(0, 0)$. Siccome la funzione $g(x, y) = x^2 + y^3 - 1$ è continua e $g(0, 2) = 7 > 0$ allora si avrà anche $g(x, y) > 0$ in tutto l'aperto connesso che contiene $(0, 2)$, perché altrimenti in questo aperto connesso si dovrebbe trovare un punto (x, y) per cui $g(x, y) = 0$. Analogamente, siccome $g(0, 0) = -1 < 0$, allora $g(x, y) < 0$ su tutto l'aperto connesso che contiene $(0, 0)$. Quindi $\operatorname{dom} f$ risulta la regione evidenziata in bianco in figura, compreso il grafico di $y = \sqrt[3]{1 - x^2}$. ☺

☞Esempio 3.25. Determinare e disegnare nel piano il dominio della funzione

$$f(x, y) = \log(y(x^2 - 1) - 1).$$

Si ha $\operatorname{dom} f = \{(x, y) \in \mathbb{R}^2 : y(x^2 - 1) - 1 > 0\}$. Siccome la funzione $g(x, y) = y(x^2 - 1) - 1$ è continua, utilizziamo il metodo di prima. Iniziamo disegnando i punti (x, y) di $\mathbb{R}^2$ per cui $y(x^2 - 1) - 1 = 0$, ovvero $y(x^2 - 1) = 1$. Ma allora $x \neq \pm 1$ e $y = 1/(x^2 - 1)$.



Il piano viene diviso dal grafico di $y = 1/(x^2 - 1)$ in quattro aperti connessi. Scegliamo i punti $(0,0)$, $(\pm 2, 1)$ e $(0, -2)$ (uno in ogni regione) e calcoliamo

$$g(0,0) = -1 \quad g(\pm 2, 1) = 2 \quad g(0,-2) = 1.$$

Siccome g è continua, dovrà essere $g(x,y) < 0$ nella regione che contiene $(0,0)$ e $g(x,y) > 0$ in ciascuna delle regioni colorate. Quindi dom f risulta l'unione delle regioni colorate in figura, escluso il grafico di $y = \sqrt[3]{1 - x^2}$. ☺

DEFINIZIONE 3.7. Un insieme $A \subseteq \mathbb{R}^n$ si dice aperto se ogni suo punto risulta interno, ovvero se per ogni $\mathbf{a} \in A$ esiste un intorno aperto U di $\mathbf{a}$ tale che $U \subseteq A$. Un insieme $K \subseteq \mathbb{R}^n$ si dice chiuso se il suo complementare è aperto.

TEOREMA 3.26 (di Weierstrass). *Sia $K \subseteq \mathbb{R}^n$ un insieme chiuso e limitato e sia $f : K \to \mathbb{R}$ una funzione continua. Allora f è limitata e ammette massimo e minimo assoluti, ovvero esistono $\mathbf{y}_1$ e $\mathbf{y}_2$ in K tali che*

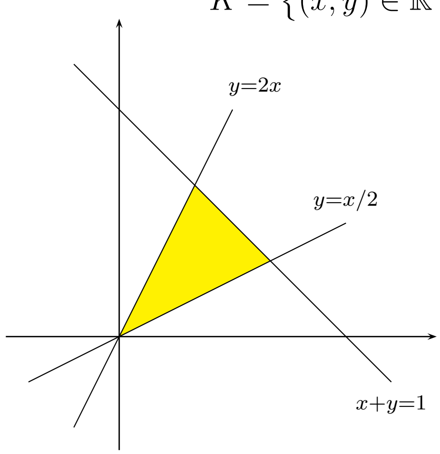
$$\min f = f(\mathbf{y}_1) \leq f(\mathbf{x}) \leq f(\mathbf{y}_2) = \max f \qquad \forall \mathbf{x} \in K.$$

☞ESEMPIO 3.27. Dire se la funzione

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

ha massimo e/o minimo assoluto sull'insieme

$$K = \left\{ (x,y) \in \mathbb{R}^2 \,:\, x + y \leq 1, \quad 2x - y \geq 0, \quad x - 2y \leq 0 \right\}$$



La funzione f è continua, come composta di funzioni continue.

Con il solito metodo possiamo disegnare l'insieme K e notare che è il triangolo chiuso colorato in giallo, quindi chiuso e limitato.

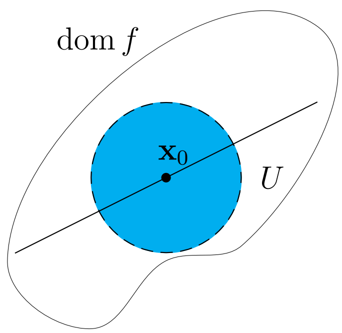
Per il Teorema di Weierstass, la funzione f ammette massimo e minimo assoluti in K. ☺

Nella prossima sezione svilupperemo metodi che ci permetteranno di calcolare i valori massimo o minimo assoluti di una funzione.

3. Derivabilità e differenziabilità

Iniziamo a affrontare il problema in maniera ingenua, cercando di trasferire le nostre conoscenze di funzioni di una variabile.

3.1. Derivate direzionali. Sia $f : \mathbb{R}^n \to \mathbb{R}$ e sia $\mathbf{x}_0$ un punto interno al dominio di f, ovvero esista un intorno aperto U contenuto in $\operatorname{dom} f$.



Consideriamo una retta per il punto $\mathbf{x}_0$. Allora possiamo restringere la funzione f a questa retta e ottenere una funzione di una variabile, che sappiamo derivare. In formule, se $\mathbf{v} = (v_1, \ldots, v_n)$ è un vettore in $\mathbb{R}^n \setminus \{\mathbf{0}\}$, allora la retta per $\mathbf{x}_0$ e parallela a $\mathbf{v}$ è descritta in forma parametrica dall'equazione (vettoriale) $\mathbf{x}(t) = \mathbf{x}_0 + t\mathbf{v}$ al variare di t in $\mathbb{R}$. Per t sufficientemente piccolo i punti $\mathbf{x}(t)$ sono nell'intorno aperto U, quindi sono nel $\operatorname{dom} f$. Possiamo chiederci se la funzione di una variabile $f(\mathbf{x}(t))$ risulta derivabile in $t = 0$.

Definizione 3.8. Si dice che $f : \mathbb{R}^n \to \mathbb{R}$ è derivabile (o derivabile Gateaux) in $\mathbf{x}_0$ interno al $\operatorname{dom} f$ secondo il vettore $\mathbf{v} \in \mathbb{R}^n \setminus \{0\}$ se esiste finito

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

Qualora questo limite esista finito, lo si chiama derivata di f in $\mathbf{x}_0$ secondo il vettore $\mathbf{v}$ e si indica con $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$.

Tra le derivate direzionali sono importanti quelle secondo i vettori della base canonica di $\mathbb{R}^n$ (quelli che hanno tutte le componenti nulle, tranne una che è uguale a 1). Una derivata direzionale secondo un vettore della base canonica si chiama derivata parziale. Solitamente se $\mathbf{e}_j = (0, \ldots, \overset{j}{1}, \ldots, 0)$, cioè $\mathbf{e}_j$ è il vettore della base canonica che ha un 1 al posto j, allora la derivata parziale j-esima viene indicata con uno dei simboli $\partial_j f$, $\frac{\partial f}{\partial x_j}$, f_{x_j}, f_j. Calcolare la derivate parziale j-esima è molto semplice: basta trattare le altre coordinate come costanti.

Se $n = 2$ si usano anche i simboli $\frac{\partial f}{\partial x}$, $\partial_x f$ e $\frac{\partial f}{\partial y}$, $\partial_y f$ per indicare le derivate parziali di f.

☞Esempio 3.28. Calcolare le derivate parziali di $\sin(x^2 y)$. Trattando y come costante e derivando rispetto a x, otteniamo $\frac{\partial}{\partial x}\sin(x^2 y) = 2xy\cos(x^2 y)$. Trattando x come costante e derivando rispetto a y otteniamo $\frac{\partial}{\partial y}\sin(x^2 y) = x^2 y \cos(x^2 y)$. ☺

☞Esempio 3.29. Calcolare le derivate direzionali nell'origine della funzione

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Fissiamo un vettore (v_1, v_2) con $v_1^2 + v_2^2 \neq 0$. Allora se $t \neq 0$ abbiamo

$$\begin{aligned} \frac{f(0+tv_1, 0+tv_2) - f(0,0)}{t} &= \frac{t^2 v_1^2 \, t v_2}{t^4 v_1^4 + t^2 v_2^2} \frac{1}{t} \\ &= \frac{v_1^2 \, v_2}{t^2 v_1^4 + v_2^2} \xrightarrow[t \to 0]{} \begin{cases} 0 & v_2 = 0 \\ v_1^2 / v_2 & v_2 \neq 0. \end{cases} \end{aligned}$$

☺

Osserviamo che la funzione dell'esempio precedente è derivabile secondo tutte le direzioni nell'origine, tuttavia non è continua in $(0,0)$, perché $f(x, x^2) = \frac{1}{2}$ per ogni $x \neq 0$. Questo esempio ci mostra che una nozione di derivabilità sensata per funzioni di più variabili deve essere più forte (vorremmo che le funzioni derivabili risultino più regolari delle funzioni continue).

L'anno scorso abbiamo visto che la nozione di derivabilità in x_0 è equivalente al fatto di poter approssimare la funzione con un polinomio di grado 1 vicino a x_0 a meno di infinitesimi di ordine superiore. Geometricamente, questo fatto si esprime dicendo che il grafico della funzione vicino a x_0 è "molto vicino" a quello della retta tangente al grafico in x_0.

L'idea dell'approssimazione tramite la retta tangente, che nel caso di due variabili diventa un piano tangente, è quella che si dimostra vincente nel caso di funzioni di più variabili.

Iniziamo a richiamare alcune nozioni sui grafici di funzioni lineari in $\mathbb{R}^n$.

3.2. Piani in $\mathbb{R}^3$ – iperpiani in $\mathbb{R}^{n+1}$. Un piano in $\mathbb{R}^3$ può essere pensato individuato mediante un punto P_0 e un vettore V dello spazio vettoriale $\mathbb{R}^3$. Infatti possiamo dire che un punto P sta nel piano Π passante per P_0 se il vettore $P - P_0$ è ortogonale al vettore V. In formule questo si esprime dicendo che il prodotto scalare tra $P - P_0$ e V deve essere nullo:

$$\langle V, P - P_0 \rangle = 0 \qquad \forall P \in \Pi.$$

Se P_0 ha coordinate (x_0, y_0, z_0) e $V = (v_1, v_2, v_3)$ nella base canonica, allora troviamo l'equazione dei punti $P = (x, y, z)$ del piano Π

$$v_1(x - x_0) + v_2(y - y_0) + v_3(z - z_0) = 0.$$

Il vettore V si dice vettore normale al piano Π.

Viceversa, un'equazione lineare in x, y, z della forma

$$ax + by + cz = d$$

dove a, b, c non sono tutti nulli, rappresenta l'equazione di un piano con vettore normale (a, b, c) (passante per $(0, 0, d/c)$ se $c \neq 0$, passante per $(d/a, 0, 0)$ se $a \neq 0$, per $(0, d/b, 0)$ se $b \neq 0$).

Si noti che se $c \neq 0$, allora il piano non è verticale e è il grafico della funzione lineare

$$\ell(x,y) = \frac{d}{c} - \frac{a}{c}\,x - \frac{b}{c}\,y\,.$$

Un iperpiano in $\mathbb{R}^{n+1}$ è un sottoinsieme di $\mathbb{R}^{n+1}$ formato da tutti quei punti che hanno prodotto scalare costante con un vettore fissato dato, detto anche vettore normale. Pertanto un'equazione lineare della forma

$$\sum_{j=1}^{n+1} a_j x_j = b \qquad \text{dove } a_j, b \in \mathbb{R}$$

rappresenta un iperpiano in $\mathbb{R}^{n+1}$ e precisamente quell'iperpiano che ha prodotto costante ($= b$) con il vettore $(a_1, \dots, a_{n+1})$. Quando $a_{n+1} \neq 0$ il piano non è verticale e possiamo scriverlo nella forma

$$x_{n+1} = b' - \sum_{j=1}^{n} a'_j x_j \qquad \text{dove } a'_j, b' \in \mathbb{R}$$

ovvero come grafico di una funzione lineare.

3.3. Differenziabilità. Vogliamo dire che una funzione risulta differenziabile in un punto $\mathbf{x}_0$ se in quel punto può essere approssimata tramite una funzione lineare a meno di infinitesimi di ordine superiore a uno, in termini geometrici se ha un iperpiano tangente non verticale.

Definizione 3.9. Sia $f : \mathbb{R}^n \to \mathbb{R}$ e sia $\mathbf{x}_0$ un punto interno al dominio di f. Si dice che f è è differenziabile in $\mathbf{x}_0$ se esistono un vettore $\mathbf{d}$ in $\mathbb{R}^n$ e una funzione ω definita in un opportuno intorno U_0 di $\mathbf{0}$ in $\mathbb{R}^n$ tali che

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \langle \mathbf{d}, \mathbf{x} - \mathbf{x}_0 \rangle + \|\mathbf{x} - \mathbf{x}_0\| \omega(\mathbf{x} - \mathbf{x}_0) \qquad \forall \mathbf{x} \text{ tale che } \mathbf{x} - \mathbf{x}_0 \in U_0$$

e $\lim_{\mathbf{t} \to 0} \omega(\mathbf{t}) = \omega(\mathbf{0}) = 0$.

Notare che il vettore $\mathbf{d}$ dipende sia da f sia da $\mathbf{x}_0$ e solitamente si chiama gradiente di f in $\mathbf{x}_0$; in formule $\nabla f(\mathbf{x}_0)$. Nel prossimo teorema vediamo che la nozione di differenziabilità ora introdotta implica la continuità e è più forte della nozione di derivabilità secondo un vettore.

Teorema 3.30 (del differenziale). *Sia f differenziabile in un punto $\mathbf{x}_0$ interno al dominio di f. Allora*

a) *f è continua in $\mathbf{x}_0$;*
b) *il vettore $\nabla f(\mathbf{x}_0)$ è $\left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0)\right)$;*
c) *f è derivabile in $\mathbf{x}_0$ secondo ogni vettore $\mathbf{v}$ non nullo e $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \langle \nabla f(\mathbf{x}_0), \mathbf{v} \rangle$.*

Dimostrazione. Siccome i polinomi sono funzioni continue, il prodotto scalare e il quadrato della norma sono funzioni continue. Inoltre per composizione con la funzione radice, anche la norma è continua. Quindi $\lim_{\mathbf{x}\to\mathbf{x}_0}\langle\mathbf{d},\mathbf{x}-\mathbf{x}_0\rangle = \langle\mathbf{d},\mathbf{0}\rangle = 0$ e $\lim_{\mathbf{x}\to\mathbf{x}_0}\|\mathbf{x}-\mathbf{x}_0\| = \|\mathbf{0}\| = 0$. Da cui

$$\begin{aligned}\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) &= \lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}_0) + \langle\mathbf{d},\mathbf{x}-\mathbf{x}_0\rangle + \|\mathbf{x}-\mathbf{x}_0\|\omega(\mathbf{x}-\mathbf{x}_0)\\ &= f(\mathbf{x}_0),\end{aligned}$$

ovvero f è continua in $\mathbf{x}_0$. Occupiamoci ora di c). Si ha

$$\begin{aligned}\lim_{t\to 0}\frac{f(\mathbf{x}_0+t\mathbf{v})-f(\mathbf{x}_0)}{t} &= \lim_{t\to 0}\frac{f(\mathbf{x}_0)+\langle\mathbf{d},t\mathbf{v}\rangle+\|t\mathbf{v}\|\omega(t\mathbf{v})-f(\mathbf{x}_0)}{t}\\ &= \lim_{t\to 0}\langle\mathbf{d},\mathbf{v}\rangle+\frac{|t|}{t}\,\|\mathbf{v}\|\,\omega(t\mathbf{v})\\ &= \langle\mathbf{d},\mathbf{v}\rangle,\end{aligned}$$

perché ω è continua e infinitesima in $\mathbf{0}$ e $\frac{|t|}{t}$ è limitato (vale ± 1). Ma allora f è derivabile in $\mathbf{x}_0$ secondo ogni vettore $\mathbf{v}$ non nullo e $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \langle\mathbf{d},\mathbf{v}\rangle$. Da questo ricaviamo anche b), prendendo come vettori $\mathbf{v}$ i vettori della base canonica: si ha $\frac{\partial f}{\partial x_j}(\mathbf{x}_0) = \frac{\partial f}{\partial \mathbf{e}_j}(\mathbf{x}_0) = \langle\mathbf{d},\mathbf{e}_j\rangle = d_j$, ovvero $\mathbf{d} = (\frac{\partial f}{\partial x_1}(\mathbf{x}_0),\ldots,\frac{\partial f}{\partial x_n}(\mathbf{x}_0))$. □

Si noti che l'iperpiano di $\mathbb{R}^{n+1}$ di equazione

$$x_{n+1} = f(\mathbf{x}_0) + \langle\nabla f(\mathbf{x}_0),\mathbf{x}-\mathbf{x}_0\rangle$$

è tangente al grafico di f nel punto $(\mathbf{x}_0, f(\mathbf{x}_0))$ e ha vettore normale $(\nabla f(\mathbf{x}_0),-1)$.

Da un punto di vista geometrico, l'interpretazione di quanto abbiamo visto è la seguente: se una funzione in un determinato punto possiede un iperpiano tangente, allora la restrizione di questa funzione a ogni retta parallela agli assi $\mathbf{x}(t) = \mathbf{x}_0 + t\mathbf{e}_j$ dovrà risultare derivabile, come funzione di una variabile in $t = 0$. Infatti la retta intersezione tra l'iperpiano tangente e il "cilindro" sopra $\mathbf{x}(t)$ giace sull'iperpiano tangente, quindi è tangente al grafico di f. Pertanto la funzione f dovrà risultare parzialmente derivabile e

$$\partial_j f(\mathbf{x}_0) = \lim_{t\to 0}\frac{f(\mathbf{x}_0+t\mathbf{e}_j)-f(\mathbf{x}_0)}{t}.$$

La retta tangente, nel piano individuato da $\mathbf{e}_j$ e $\mathbf{e}_{n+1}$, sarà del tipo $x_{n+1} = f(\mathbf{x}_0)+\partial_j f(\mathbf{x}_0)\,t$; vista nello spazio $\mathbb{R}^{n+1}$ avrà equazione parametrica

$$\begin{cases} x_k = x_{0k} + t\delta_{kj} \qquad k = 1,\ldots,n\\ x_{n+1} = f(\mathbf{x}_0)+\partial_j f(\mathbf{x}_0)\,t\end{cases}$$

quindi ha vettore direzionale $\mathbf{v}_j = (0,\ldots,\overset{j}{1},\ldots,\partial_j f(\mathbf{x}_0))$. L'iperpiano tangente dovrà contenere tutte queste rette tangenti al variare di $j = 1,\ldots,n$, quindi il suo vettore normale $\mathbf{N}$ dovrà essere ortogonale a tutti i $\mathbf{v}_j$. Questo vuol dire che

$$\langle\mathbf{N},\mathbf{v}_j\rangle = 0 \qquad \forall j = 1,\ldots,n.$$

Allora le componenti di $\mathbf{N} = (N_1, \dots, N_{n+1})$ soddisfano le equazioni

$$0 = \langle \mathbf{N}, \mathbf{v}_j \rangle = N_j + N_{n+1}\partial_j f(\mathbf{x}_0) \qquad \forall j = 1, \dots, n,$$

da cui $N_j = -N_{n+1}\partial_j f(\mathbf{x}_0)$. Se scegliamo, ad esempio, $N_{n+1} = -1$, otteniamo che $N_j = \partial_j f(\mathbf{x}_0)$ ovvero l'iperpiano tangente deve avere equazione

$$x_{n+1} = f(\mathbf{x}_0) + \sum_{j=1}^{n} \partial_j f(\mathbf{x}_0)(x_j - x_{0j}) = f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle.$$

Per verificare la differenziabilità di f in un punto fissato, possiamo procedere in questo modo.

Proposizione 3.31. *Sia $f : \mathbb{R}^n \to \mathbb{R}$ derivabile parzialmente nel punto $\mathbf{x}_0$ interno al dominio di f. Se*

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle}{\|\mathbf{x} - \mathbf{x}_0\|} = 0,$$

allora f è differenziabile in $\mathbf{x}_0$. Se il limite invece non esiste o è diverso da 0, allora f non è differenziabile in $\mathbf{x}_0$.

Dimostrazione. Si ponga

$$\mathbf{d} = \nabla f(\mathbf{x}_0) \qquad \omega(\mathbf{t}) = \begin{cases} \frac{f(\mathbf{t}+\mathbf{x}_0) - f(\mathbf{x}_0) - \langle \nabla f(\mathbf{x}_0), \mathbf{t} \rangle}{\|\mathbf{t}\|} & \mathbf{t} \neq \mathbf{0} \quad \mathbf{t} + \mathbf{x}_0 \in \operatorname{dom} f \\ 0 & \mathbf{t} = \mathbf{0} \end{cases}$$

e si controllino le proprietà della definizione. □

☞Esempio 3.32. Si studi la differenziabilità in $(0,0)$ della funzione

$$f(x,y) = \begin{cases} \frac{x^3 y}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

Siccome la funzione è nulla sugli assi, allora le derivate parziali in $(0,0)$ sono nulle: $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$. La funzione f è differenziabile in $(0,0)$ se risulta essere 0 il

$$\begin{aligned} \lim_{(x,y)\to(0,0)} \omega(x,y) &= \lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - \langle \nabla f(0,0), (x,y) \rangle}{\sqrt{x^2+y^2}} \\ &= \lim_{(x,y)\to(0,0)} \frac{x^3 y}{(x^2+y^2)^{3/2}}. \end{aligned}$$

Valutando il modulo di questa espressione in coordinate polari otteniamo:

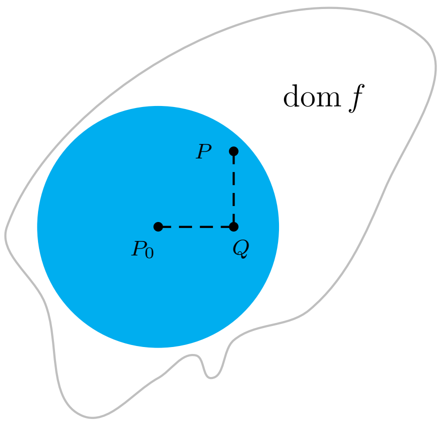
$$\begin{aligned} \left| \frac{x^3 y}{(x^2+y^2)^{3/2}} \right| &= \left| \frac{r^4 \cos^3\theta \sin\theta}{r^3} \right| \\ &\leq r \xrightarrow[r\to 0]{} 0. \end{aligned}$$

Pertanto $\lim\limits_{(x,y)\to(0,0)} \omega(x,y) = 0$ e f risulta quindi differenziabile in $(0,0)$. Il piano tangente in $(0,0)$ al grafico della funzione è $z = 0$. ☺

In alcuni casi è molto semplice verificare che una funzione è differenziabile.

TEOREMA 3.33. *Sia $\mathbf{x}_0$ un punto interno al dominio di f e in un intorno di tale punto esistano e siano continue tutte le derivate parziali di f. Allora f è differenziabile in $\mathbf{x}_0$.*

DIMOSTRAZIONE. Diamo un'idea della dimostrazione nel caso $n = 2$. Siano, come in figura, $P_0 = (x_0, y_0)$, $P = (x, y)$, $Q = (x, y_0)$ punti nell'intorno di (x_0, y_0) dove le derivate parziali sono continue. Dobbiamo verificare che $f(P) - f(P_0) - \langle \nabla f(P_0), P - P_0 \rangle$ tende a zero con ordine maggiore di $\|P - P_0\|$, per $P \to P_0$.



Sia Q il punto che ha stessa ordinata di P e stessa ascissa di P_0, come in figura. Scriviamo

$$f(P) - f(P_0) = f(P) - f(Q) + f(Q) - f(P_0).$$

In questo modo, arriviamo al punto P, partendo da P_0, muovendoci dapprima a y costante e dopo a x costante. Siccome la funzione f ha derivate parziali continue in un intorno di P_0, possiamo applicare il teorema di Lagrange e dire che esistono punti ξ, compreso tra x e x_0, e η compreso tra y e y_0, tali che

$$\begin{aligned} f(P) - f(Q) &= \partial_y f(x, \eta)(y - y_0) \\ f(Q) - f(P_0) &= \partial_x f(\xi, y_0)(x - x_0). \end{aligned}$$

Possiamo quindi scrivere $f(P) - f(P_0) - \langle \nabla f(P_0), P - P_0 \rangle$ come

$$(\partial_x f(\xi, y_0) - \partial_x f(x_0, y_0))(x - x_0) + (\partial_y f(x, \eta) - \partial_y f(x_0, y_0))(y - y_0).$$

Dobbiamo controllare che questo tenda a zero più velocemente di $\|P - P_0\|$ per $P \to P_0$. D'altra parte

$$\left| \frac{(\partial_x f(\xi, y_0) - \partial_x f(x_0, y_0))(x - x_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \right| \leq |\partial_x f(\xi, y_0) - \partial_x f(x_0, y_0)|$$

e, siccome le derivate parziali sono continue nell'intorno di P_0, questa quantità tende a zero per $(x, y) \to (x_0, y_0)$. Analogamente per il secondo addendo. $\square$

☞ESEMPIO 3.34. La funzione $f(x, y) = e^{x^2+y^2} \cos(x + y)$ è differenziabile in $\mathbb{R}^2$, perché le sue derivate parziali sono

$$\begin{aligned} \partial_x f(x, y) &= e^{x^2+y^2} \cos(x + y)\, 2x + e^{x^2+y^2} \sin(x + y) \\ \partial_y f(x, y) &= e^{x^2+y^2} \cos(x + y)\, 2y + e^{x^2+y^2} \sin(x + y) \end{aligned}$$

e sono continue come composte di funzioni continue. ☺

☞Esempio 3.35. La funzione definita da

$$f(x,y) = \begin{cases} y\,e^{-x^2/y^2} & y \neq 0 \\ 0 & y = 0 \end{cases}$$

è differenziabile in $\mathbb{R}^2 \setminus \{(x,0) : x \in \mathbb{R}\}$, perché le sue derivate parziali per $y \neq 0$ sono

$$\begin{aligned} \partial_x f(x,y) &= e^{-x^2/y^2}(-2x/y) \\ \partial_y f(x,y) &= e^{-x^2/y^2} + e^{-x^2/y^2}(2x^2/y^2) \end{aligned}$$

e sono continue in $\mathbb{R}^2 \setminus \{(x,0) : x \in \mathbb{R}\}$. Ci chiediamo se f è differenziabile in $(x_0,0)$ per qualche x_0. Si ha

$$\begin{aligned} \partial_x f(x_0,0) &= \lim_{t\to 0} \frac{f(x_0+t,0)-f(x_0,0)}{t} = \lim_{t\to 0} 0 = 0 \\ \partial_y f(x_0,0) &= \lim_{t\to 0} \frac{f(x_0,t)-f(x_0,0)}{t} = \lim_{t\to 0} e^{-x_0^2/t^2} = \begin{cases} 0 & x_0 \neq 0 \\ 1 & x_0 = 0 \end{cases} \end{aligned}$$

Per $x_0 = 0$ dovrebbe essere nullo il

$$\lim_{(x,y)\to(0,0)} \frac{y\,e^{-x^2/y^2} - y}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{y\,(e^{-x^2/y^2}-1)}{\sqrt{x^2+y^2}}.$$

Notiamo che restringendosi a $x = y$ il limite precedente diventa

$$\lim_{x\to 0} \frac{x\,(e^{-1}-1)}{|x|\sqrt{2}} \qquad \text{che non esiste.}$$

Quindi f non è differenziabile in $(0,0)$.

Per $x_0 \neq 0$ dobbiamo invece considerare

$$\lim_{(x,y)\to(x_0,0)} \frac{y\,e^{-x^2/y^2}}{\sqrt{(x-x_0)^2+y^2}} = \lim_{(x,y)\to(x_0,0)} h(x,y).$$

Possiamo scrivere h come prodotto di e^{-x^2/y^2} che è infinitesima per $(x,y) \to (x_0,0)$ e $\frac{y}{\sqrt{(x-x_0)^2+y^2}}$ che è limitata, quindi h è infinitesima per $(x,y) \to (x_0,0)$. Oppure possiamo seguire il metodo solito delle coordinate polari centrate in $(x_0,0)$:

$$\begin{aligned} \sup_{\substack{\theta\in(0,2\pi) \\ \theta\neq\pi}} |h(x_0+r\cos\theta, r\sin\theta)| &= \sup_{\substack{\theta\in(0,2\pi) \\ \theta\neq\pi}} \frac{r|\sin\theta|\,e^{-(x_0+r\cos\theta)^2/(r\sin\theta)^2}}{r} \\ &= \sup_{\substack{\theta\in(0,2\pi) \\ \theta\neq\pi}} |\sin\theta|\,e^{-(x_0+r\cos\theta)^2/(r\sin\theta)^2} \\ &\leq \sup_{\substack{\theta\in(0,2\pi) \\ \theta\neq\pi}} e^{-(x_0+r\cos\theta)^2/r^2} \end{aligned}$$

Per stimare questa funzione, notiamo che ci possiamo limitare a considerare $r < |x_0|/2$ (perché ci interessa r piccolo), da cui $(x_0 + r\cos\theta)^2 \geq |x_0|^2/4$; inoltre $0 < (r\sin\theta)^2 \leq r^2$ per cui

$$(x_0 + r\cos\theta)^2/(r\sin\theta)^2 \geq |x_0|^2/(4r^2).$$

Siccome $t \mapsto e^{-t}$ è decrescente, ricaviamo che $e^{-(x_0+r\cos\theta)^2/(r\sin\theta)^2} \leq e^{-|x_0|^2/(4r^2)}$. Allora

$$\sup_{\substack{\theta\in(0,2\pi)\\ \theta\neq\pi}} |h(x_0 + r\cos\theta, r\sin\theta)| \leq e^{-|x_0|^2/(4r^2)} \xrightarrow[r\to 0^+]{} 0.$$

Quindi f è differenziabile anche in $(x_0, 0)$ con $x_0 \neq 0$. ☺

3.4. Derivazione della funzione composta. Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita sull'insieme A aperto e sia $\mathbf{g} = (g_1, \ldots, g_n)$ una funzione a valori vettoriali definita sull'intervallo (a, b) e tale che per ogni t in (a, b) si abbia $\mathbf{g}(t) \in A$. Allora possiamo considerare la funzione composta $f \circ \mathbf{g}$, cioè $t \mapsto f(\mathbf{g}(t))$, che è una funzione di una variabile reale.

Teorema 3.36 (di derivazione della funzione composta—prima versione). *Sia t_0 in (a, b) e supponiamo che tutte le funzioni g_j siano derivabili in t_0, $j = 1, \ldots, n$. Sia inoltre f differenziabile in $\mathbf{g}(t_0)$. Allora la funzione composta $f \circ \mathbf{g}$ è derivabile in t_0 e*

$$\frac{d}{dt}(f \circ \mathbf{g})(t_0) = \sum_{j=1}^{n} \partial_j f(\mathbf{g}(t_0))\, g_j'(t_0) = \langle \nabla f(\mathbf{g}(t_0)), \mathbf{g}'(t_0)\rangle,$$

dove $\mathbf{g}'(t_0)$ denota il vettore $(g_1'(t_0), \ldots, g_n'(t_0))$.

☞Esempio 3.37. Siano $f(x, y) = e^{-(x^2+y^2)}$ e $\mathbf{g}(t) = (\cos t, \sin t)$. Calcolare la derivata della funzione $f \circ \mathbf{g}$.

Si ha $\nabla f(x, y) = (-2x\, e^{-(x^2+y^2)}, -2y\, e^{-(x^2+y^2)})$, quindi

$$\nabla f(\cos t, \sin t) = (-2\cos t\, e^{-1}, -2\sin t\, e^{-1})$$

e $\mathbf{g}'(t) = (-\sin t, \cos t)$. Allora

$$\frac{d}{dt}(f \circ \mathbf{g})(t) = -2\cos t\, e^{-1}(-\sin t) - 2\sin t\, e^{-1}(\cos t) = 0.$$

(Del resto $f \circ \mathbf{g}$ è la funzione che vale costantemente e^{-1}) ☺

Le funzioni dei prossimi esempi sono tutte derivabili o differenziabili in tutti i punti del loro dominio. Vogliamo solo illustrare il significato delle formule.

☞Esempio 3.38. Siano $f(x, y) = \cos(x + 2y)$ e $\mathbf{g}(t) = (t^2, t - 2)$. Calcolare la derivata della funzione $f \circ \mathbf{g}$.

Si ha $\nabla f(x, y) = (-\sin(x + y), -2\sin(x + y))$, quindi

$$\nabla f(t^2, t - 2) = (-\sin(t^2 + t - 2), -2\sin(t^2 + t - 2))$$

e $\mathbf{g}'(t) = (2t, 1)$. Allora

$$\frac{d}{dt}(f \circ \mathbf{g})(t) = (-\sin(t^2+t-2))(2t) + (-2\sin(t^2+t-2)) = -2(t+1)\sin(t^2+t-2).$$

☺

4. Derivate di ordine successivo

Sia data una funzione $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e supponiamo che esistano in A, insieme aperto, le derivate parziali di f. Le derivate parziali si possono considerare a loro volta come funzioni definite in A e ci potremmo chiedere se tali funzioni sono a loro volta derivabili parzialmente.

Definizione 3.10. Se le funzioni $\partial_j f$ sono derivabili parzialmente, si dice che f ammette derivate parziali del secondo ordine (o derivate parziali seconde).

Si indica con $\partial_{kj} f$ la funzione $\partial_k(\partial_j f)$. Nel caso in cui $k = j$ si parla di derivate parziali seconde pure (ovvero si deriva rispetto alla stessa variabile) e talvolta si usa la notazione ∂_j^2 per indicare ∂_{jj}; nel caso in cui $k \neq j$ si parla di derivate parziali seconde miste.

In generale, se $j \neq k$, non è detto che $\partial_{kj} f = \partial_{jk} f$. Tuttavia, se la funzione f è sufficientemente regolare, l'ordine di derivazione non conta; questo è il contenuto del prossimo teorema.

Teorema 3.39 (di Schwarz sulle derivate miste). *Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione e sia $\mathbf{x}_0$ un punto interno al dominio di f. Supponiamo che in un intorno di $\mathbf{x}_0$ esistano le derivate parziali $\partial_{kj} f$ e $\partial_{jk} f$ e siano continue in $\mathbf{x}_0$. Allora $\partial_{kj} f(\mathbf{x}_0) = \partial_{jk} f(\mathbf{x}_0)$.*

☞Esempio 3.40. Sia

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

Verificare che $\partial_{xy} f(0,0) \neq \partial_{yx} f(0,0)$. ☺

Possiamo scrivere le derivate parziali seconde di una funzione in un punto $\mathbf{x}_0$ raggruppandole in una matrice in questo modo:

$$\begin{pmatrix} \partial_{11} f(\mathbf{x}_0) & \partial_{12} f(\mathbf{x}_0) & \dots & \partial_{1n} f(\mathbf{x}_0) \\ \partial_{21} f(\mathbf{x}_0) & \partial_{22} f(\mathbf{x}_0) & \dots & \partial_{2n} f(\mathbf{x}_0) \\ \vdots & \vdots & & \vdots \\ \partial_{n1} f(\mathbf{x}_0) & \partial_{n2} f(\mathbf{x}_0) & \dots & \partial_{nn} f(\mathbf{x}_0) \end{pmatrix}$$

Questa matrice si chiama matrice Hessiana di f in $\mathbf{x}_0$ e si indica con $\mathrm{Hess}_f(\mathbf{x}_0)$ o anche $H_{f,\mathbf{x}_0}$. Se la funzione f ha derivate parziali seconde continue, allora questa matrice è simmetrica. A ogni matrice simmetrica H si associa la forma quadratica $Q_H(\mathbf{x}) = {}^t\mathbf{x}\, H\, \mathbf{x}$.

Abbiamo visto che se f è differenziabile in $\mathbf{x}_0$, allora f si può approssimare in un intorno U_0 di questo punto tramite il suo piano tangente, ovvero un polinomio in n variabili di grado 1:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \|\mathbf{x} - \mathbf{x}_0\| \omega(\mathbf{x} - \mathbf{x}_0) \qquad \forall \mathbf{x} \in U_0$$

e $\lim_{\mathbf{t}\to 0} \omega(\mathbf{t}) = \omega(\mathbf{0}) = 0$.

Cosa succede se la funzione è più regolare, ad esempio con derivate parziali seconde continue? Analogamente al caso di funzioni di una variabile, possiamo approssimare la funzione nell'intorno di un punto fissato mediante un polinomio (in n variabili) di grado 2.

Teorema 3.41 (formula di Taylor di ordine 2 con resto di Peano). *Sia f una funzione con derivate parziali continue sino all'ordine 2 in un intorno U_0 del punto $\mathbf{x}_0$. Allora*

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \tfrac{1}{2}\, Q_{H_{f,\mathbf{x}_0}}(\mathbf{x} - \mathbf{x}_0) + \|\mathbf{x} - \mathbf{x}_0\|^2 \omega(\mathbf{x} - \mathbf{x}_0) \qquad \forall \mathbf{x} \in U_0$$

e $\lim_{\mathbf{t}\to 0} \omega(\mathbf{t}) = \omega(\mathbf{0}) = 0$.

Dimostrazione. Se $\mathbf{x} = \mathbf{x}_0 + \mathbf{h}$ con $\mathbf{h}$ in $\mathbb{R}^n$. Restringiamo l'attenzione alla funzione ristretta alla retta per $\mathbf{x}_0$ e $\mathbf{x}$ (o alla porzione di retta che cade nel dominio di f): consideriamo la funzione $\mathbf{g}(t) = \mathbf{x}_0 + t\mathbf{h}$ per $t \in (-\varepsilon, 1+\varepsilon)$. Allora $\mathbf{g}(0) = \mathbf{x}_0$ e $\mathbf{g}(1) = \mathbf{x}$ e $\mathbf{g}(t)$ è un punto della retta passante per $\mathbf{x}_0$ e $\mathbf{x}$. Allora la funzione $\phi = f \circ \mathbf{g}$ è una funzione di una variabile reale, cui applichiamo la formula di Taylor del primo ordine, centrata in $t_0 = 0$ con resto di Lagrange: esiste $c_{\mathbf{x}} \in (0, t)$ tale che

$$\phi(t) = \phi(0) + \phi'(0)t + \tfrac{1}{2}\phi''(c_{\mathbf{x}})\, t^2. \tag{3.2}$$

Ma dal Teorema 3.36 di derivazione della funzione composta abbiamo

$$\begin{aligned}
\phi'(t) &= \sum_{j=1}^{n} h_j\, \partial_j f(\mathbf{x}_0 + t\mathbf{h}) \\
\phi''(t) &= \sum_{j=1}^{n} h_j\, \frac{d}{dt} \partial_j f(\mathbf{x}_0 + t\mathbf{h}) \\
&= \sum_{j=1}^{n} h_j \sum_{k=1}^{n} h_k\, \partial_k \partial_j f(\mathbf{x}_0 + t\mathbf{h}).
\end{aligned}$$

Quindi

$$\phi'(0) = \sum_{j=1}^{n} h_j\,\partial_j f(\mathbf{x}_0) = \langle \nabla f(\mathbf{x}_0), \mathbf{h}\rangle$$

$$\phi''(0) = \sum_{j,k=1}^{n} h_j\,h_k\,\partial_{kj} f(\mathbf{x}_0) = Q_{H_{f,\mathbf{x}_0}}(\mathbf{h}).$$

Notiamo che $\phi(1) = f(\mathbf{x})$ e quindi sostituendo $t = 1$ nella (3.2), aggiungendo e togliendo il termine del secondo ordine, otteniamo

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{h}\rangle + \tfrac{1}{2}\phi''(c_{\mathbf{x}}) \\ &= f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{h}\rangle + \tfrac{1}{2} Q_{H_{f,\mathbf{x}_0}}(\mathbf{h}) + \tfrac{1}{2}\left\{\phi''(c_{\mathbf{x}}) - Q_{H_{f,\mathbf{x}_0}}(\mathbf{h})\right\}. \end{aligned}$$

Dobbiamo ancora controllare che il termine tra parentesi graffe tende a 0 abbastanza rapidamente quando $\mathbf{x} \to \mathbf{x}_0$, cioè $\mathbf{h} \to \mathbf{0}$. Si ha

$$\frac{\left\{\phi''(c_{\mathbf{x}}) - Q_{H_{f,\mathbf{x}_0}}(\mathbf{h})\right\}}{\|\mathbf{h}\|^2} = \sum_{j,k=1}^{n} \frac{h_j\,h_k}{\|\mathbf{h}\|^2}\,\left[\partial_{kj} f(\mathbf{x}_0 + c_{\mathbf{x}}\mathbf{h}) - \partial_{kj} f(\mathbf{x}_0)\right]$$

Ogni addendo è il prodotto di $\frac{h_j\,h_k}{\|\mathbf{h}\|^2}$, che è una quantità limitata, per $[\cdots]$ che tende a 0 per la continuità delle derivate seconde. $\square$

5. Ricerca di massimi e minimi, relativi e assoluti, punti sella

Sia $f : \mathbb{R}^n \to \mathbb{R}$ e sia $\mathbf{x}_0$ un punto nel dominio di f.

Definizione 3.11. Si dice che $\mathbf{x}_0$ è un punto di massimo relativo per f se esiste un intorno U di $\mathbf{x}_0$ tale che

$$f(\mathbf{x}_0) \geq f(\mathbf{x}) \qquad \forall \mathbf{x} \in U \cap \operatorname{dom} f.$$

Se vale la diseguaglianza stretta $f(\mathbf{x}_0) > f(\mathbf{x})$ per ogni $\mathbf{x} \in U \cap \operatorname{dom} f$ escluso $\mathbf{x}_0$, si dice che $\mathbf{x}_0$ è un punto di massimo relativo forte.

Definizione 3.12. Si dice che $\mathbf{x}_0$ è un punto di minimo relativo per f se esiste un intorno U di $\mathbf{x}_0$ tale che

$$f(\mathbf{x}_0) \leq f(\mathbf{x}) \qquad \forall \mathbf{x} \in U \cap \operatorname{dom} f.$$

Se vale la diseguaglianza stretta $f(\mathbf{x}_0) < f(\mathbf{x})$ per ogni $\mathbf{x} \in U \cap \operatorname{dom} f$ escluso $\mathbf{x}_0$, si dice che $\mathbf{x}_0$ è un punto di minimo relativo forte.

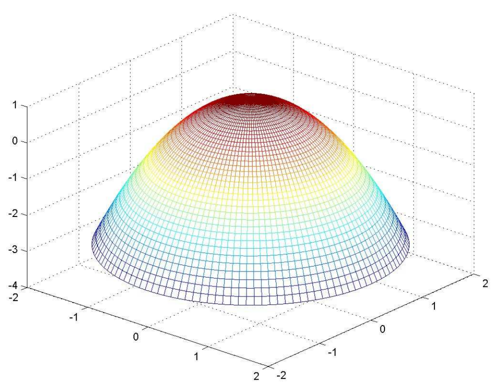
Supponiamo per un momento che $n = 2$, ossia $f : \mathbb{R}^2 \to \mathbb{R}$, e che (x_0, y_0) sia un punto di massimo relativo per f. La funzione $f(x, y_0)$, restrizione di f alla retta passante per (x_0, y_0) della forma $y = y_0$, deve avere un massimo relativo per $x = x_0$. Quindi la derivata (rispetto alla variabile x) della funzione $f(x, y_0)$ deve essere nulla in x_0. Analogamente, la restrizione della funzione f alla retta passante per (x_0, y_0) della forma $x = x_0$ deve avere un massimo relativo per $y = y_0$ e quindi la derivata (rispetto a y) della funzione $f(x_0, y)$ deve essere nulla in y_0. Nel caso generale, l'enunciato dice:

TEOREMA 3.42. *Se f è differenziabile in $\mathbf{x}_0$ e $\mathbf{x}_0$ è un punto di massimo o di minimo relativo per f, allora $\nabla f(\mathbf{x}_0) = \mathbf{0}$.*

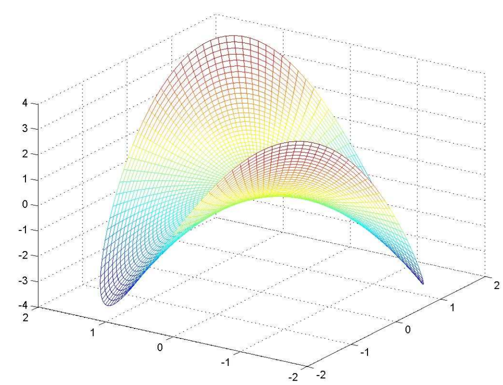
I punti $\mathbf{x}_0$ tali che $\nabla f(\mathbf{x}_0) = 0$ si dicono punti critici.

DEFINIZIONE 3.13. Si dice che un punto critico $\mathbf{x}_0$ è un punto sella se per ogni intorno U di $\mathbf{x}_0$ esistono punti $\mathbf{x}$ e $\mathbf{y}$ in U tali che

$$f(\mathbf{x}) < f(\mathbf{x}_0) < f(\mathbf{y}).$$



$(0, 0)$ è punto di max rel per $1 - x^2 - y^2$



$(0, 0)$ è punto sella per xy

Se la funzione f è di classe C^2, tramite un test sulla matrice Hessiana, possiamo stabilire se un punto critico è di massimo relativo, minimo relativo, o sella, almeno nella maggior parte dei casi. Infatti, la formula di Taylor di ordine 2 centrata in un punto critico $\mathbf{x}_0$ diventa più semplice e, attraverso tale formula riusciamo a stabilire il segno di $f(\mathbf{x}) - f(\mathbf{x}_0)$ in un intorno di $\mathbf{x}_0$ tramite lo studio della forma quadratica associata alla matrice Hessiana.

Iniziamo con il caso $n = 2$ e, per semplificare le notazioni, supponiamo che $\mathbf{x}_0 = (0, 0)$ e che l'approssimazione data dall'approssimazione di Taylor del secondo ordine sia "buona" in un intorno dell'origine, ovvero che f sia di classe C^2. Allora

$$f(x, y) - f(0, 0) \simeq \tfrac{1}{2}(ax^2 + 2bxy + cy^2)$$

dove $\mathrm{Hess}_f(0, 0) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Ci siamo ricondotti al problema di studiare il segno di un polinomio omogeneo di grado 2. Si avrà $f(x, y) - f(0, 0) > 0$ in ogni punto (x, y) in un intorno di $(0,0)$ tranne al più $(0,0)$ se $ax^2 + 2bxy + cy^2 > 0$ in ogni punto $(x,y) \neq (0,0)$. Sotto quali condizioni su a, b, c questo accade?

Se $y = 0$ questo succede quando $a > 0$. Supponiamo allora $a > 0$ e $y \neq 0$. Si ha $ax^2 + 2bxy + cy^2 > 0$ in ogni punto (x,y) con $y \neq 0$ se e solo se $at^2 + 2bt + c > 0$ per ogni $t \in \mathbb{R}$, quindi deve essere $ac - b^2 > 0$. In sintesi abbiamo controllato che se valgono entrambe le condizioni

1. $\nabla f(0,0) = (0,0)$,
2. $a = \partial_x^2 f(0,0) > 0$ e $\det(\mathrm{Hess}_f(0,0)) = ac - b^2 > 0$,

allora $(0,0)$ è un punto di minimo relativo per f.

Analogamente, se valgono entrambe le condizioni

1'. $\nabla f(0,0) = (0,0)$,
2'. $a = \partial_x^2 f(0,0) < 0$ e $\det(\mathrm{Hess}_f(0,0)) = ac - b^2 > 0$,

allora $(0,0)$ è un punto di massimo relativo per f.

Infine, se

1". $\nabla f(0,0) = (0,0)$,
2". $\det(\mathrm{Hess}_f(0,0)) = ac - b^2 < 0$,

allora $(0,0)$ è un punto sella.

Torniamo al caso generale e, per brevità, indichiamo con H la matrice Hessiana di f in $\mathbf{x}_0$. Per $\mathbf{x} \neq \mathbf{x}_0$ possiamo scrivere

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}_0) &= \tfrac{1}{2} Q_H(\mathbf{x} - \mathbf{x}_0) + \|\mathbf{x} - \mathbf{x}_0\|^2 \omega(\mathbf{x} - \mathbf{x}_0) \\ &= \tfrac{1}{2}\, {}^t(\mathbf{x} - \mathbf{x}_0) H (\mathbf{x} - \mathbf{x}_0) + \|\mathbf{x} - \mathbf{x}_0\|^2 \omega(\mathbf{x} - \mathbf{x}_0) \\ &= \|\mathbf{x} - \mathbf{x}_0\|^2 \left(\tfrac{1}{2}\, {}^t\!\left(\tfrac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|}\right) H \left(\tfrac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|}\right) + \omega(\mathbf{x} - \mathbf{x}_0) \right), \end{aligned}$$

dove $\lim_{\mathbf{t} \to 0} \omega(\mathbf{t}) = \omega(\mathbf{0}) = 0$. Vedremo che il segno di $f(\mathbf{x}) - f(\mathbf{x}_0)$ in un intorno piccolo di $\mathbf{x}_0$ è determinato dal segno di $\langle H\mathbf{t}, \mathbf{t} \rangle$, dove $\mathbf{t} = \frac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|}$ è un vettore di lunghezza 1 in $\mathbb{R}^n$. È il contenuto del seguente teorema.

Teorema 3.43 (Condizioni sufficienti per massimo/minimo relativo sella). *Sia f una funzione con derivate parziali continue fino al secondo ordine in un aperto di $\mathbb{R}^n$. Sia $\mathbf{x}_0$ un punto in questo aperto tale che $\nabla f(\mathbf{x}_0) = 0$ e sia H la matrice Hessiana di f in $\mathbf{x}_0$.*

a) *Se ${}^t\mathbf{t}H\mathbf{t} > 0$ per ogni $\mathbf{t} \in \mathbb{R}^n$ con $\|\mathbf{t}\| = 1$, allora $\mathbf{x}_0$ è un punto di minimo relativo per f.*
b) *Se ${}^t\mathbf{t}H\mathbf{t} < 0$ per ogni $\mathbf{t} \in \mathbb{R}^n$ con $\|\mathbf{t}\| = 1$, allora $\mathbf{x}_0$ è un punto di massimo relativo per f.*
c) *Se esistono $\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{R}^n$ con $\|\mathbf{t}_1\| = \|\mathbf{t}_2\| = 1$, e ${}^t\mathbf{t}_1H\mathbf{t}_1 < 0$, ${}^t\mathbf{t}_2H\mathbf{t}_2 > 0$, allora $\mathbf{x}_0$ è un punto sella.*

Il caso a) si presenta quando la forma quadratica associata alla matrice simmetrica H è definita positiva, ovvero quando H ha tutti autovalori positivi.

Il caso b) si presenta quando la forma quadratica associata alla matrice simmetrica H è definita negativa, ovvero quando H ha tutti autovalori negativi.

Il caso c) si presenta quando la forma quadratica associata alla matrice simmetrica H è indefinita perché H ha sia autovalori positivi sia autovalori negativi.

Dimostrazione. Esiste un intorno V di x_0 in cui possiamo scrivere

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \|\mathbf{x} - \mathbf{x}_0\|^2 \left(\tfrac{1}{2}\, {}^t\!\left(\tfrac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|}\right) H \left(\tfrac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|}\right) + \omega(\mathbf{x} - \mathbf{x}_0)\right) \qquad \forall \mathbf{x} \in V,\ \mathbf{x} \neq \mathbf{x}_0$$

dove $\lim\limits_{\mathbf{t}\to 0} \omega(\mathbf{t}) = \omega(\mathbf{0}) = 0$.

Nel caso a), quando H ha tutti autovalori positivi, possiamo dire che H ha un autovalore più piccolo λ e allora ${}^t\mathbf{t}H\mathbf{t} \geq \lambda > 0$ per ogni $\mathbf{t} \in \mathbb{R}^n$ con $\|\mathbf{t}\| = 1$. [2] Siccome $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} \lambda/2 + \omega(\mathbf{x} - \mathbf{x}_0) = \lambda/2 > 0$, esiste un intorno $U \subseteq V$ di $\mathbf{x}_0$ tale che $\lambda/2 + \omega(\mathbf{x} - \mathbf{x}_0) > 0$ per ogni $\mathbf{x}$ in U.

[2] Se non avete visto questo fatto in ALGA, potete anche ragionare in questo modo: la forma quadratica $\mathbf{v} \mapsto Q(\mathbf{v})$ è una funzione continua e sempre positiva sulla sfera $S = \{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\| = 1\}$. La sfera è un insieme chiuso e limitato, quindi per il Teorema di Weierstrass la funzione Q ha minimo assoluto λ, ovvero

$$\begin{aligned} &\lambda = Q(\mathbf{v}_0) && \text{per un opportuno } \mathbf{v}_0 \in S \\ &\lambda \leq Q(\mathbf{v}) && \text{per ogni } \mathbf{v} \in S. \end{aligned}$$

Siccome $Q(\mathbf{v}) > 0$ per ogni $\mathbf{v}$ in S, si ha $\lambda > 0$.

A questo punto per $\mathbf{x}$ in U, $\mathbf{x} \neq \mathbf{x}_0$

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \underbrace{\|\mathbf{x}-\mathbf{x}_0\|^2}_{>0} \left(\underbrace{\tfrac{1}{2}\, {}^t\!\left(\tfrac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|}\right) H \left(\tfrac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|}\right)}_{\geq \lambda/2} + \omega(\mathbf{x}-\mathbf{x}_0) \right)$$
$$\geq \|\mathbf{x}-\mathbf{x}_0\|^2 \left(\frac{\lambda}{2} + \omega(\mathbf{x}-\mathbf{x}_0) \right) > 0.$$

Il caso b) è del tutto analogo.

Per il caso c), supponiamo che $\lambda_1 = {}^t\mathbf{t}_1 H \mathbf{t}_1 < 0$ e $\lambda_2 = {}^t\mathbf{t}_2 H \mathbf{t}_2 > 0$.

Osserviamo che fissato un qualsiasi intorno U di $\mathbf{x}_0$, esiste $\alpha > 0$ abbastanza piccolo tale che un punto del tipo $\mathbf{x} = \mathbf{x}_0 + \alpha \mathbf{t}_1$ stia in U. Allora per un tale $\mathbf{x} = \mathbf{x}_0 + \alpha \mathbf{t}_1$

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \underbrace{|\alpha|^2}_{>0} \left(\underbrace{\tfrac{1}{2}\, {}^t\!\left(\mathbf{t}_1\right) H \left(\mathbf{t}_1\right)}_{=\lambda_1/2} + \omega(\alpha \mathbf{t}_1) \right)$$
$$= |\alpha|^2 \left(\frac{\lambda_1}{2} + \omega(\alpha \mathbf{t}_1) \right).$$

Osserviamo che la quantità fra parentesi tonde tende a $\frac{\lambda_1}{2} < 0$ quando $\mathbf{x} \to \mathbf{x}_0$ (cioè quando $\alpha \to 0$) quindi si mantiene negativa per α piccolo abbastanza. Questo prova che esiste un punto $\mathbf{x} = \mathbf{x}_0 + \alpha \mathbf{t}_1$ in U tale che

$$f(\mathbf{x}) - f(\mathbf{x}_0) = |\alpha|^2 \left(\frac{\lambda_1}{2} + \omega(\alpha \mathbf{t}_1) \right) < 0.$$

Analogamente si trova $\mathbf{y}$ del tipo $\mathbf{y} = \mathbf{x}_0 + \beta \mathbf{t}_2$ con β piccolo in modo che $\mathbf{y} = \mathbf{x}_0 + \beta \mathbf{t}_2$ stia in U e

$$f(\mathbf{y}) - f(\mathbf{x}_0) = \underbrace{|\beta|^2}_{>0} \left(\tfrac{1}{2}\, {}^t\!\left(\mathbf{t}_2\right) H \left(\mathbf{t}_2\right) + \omega(\beta \mathbf{t}_2) \right) = |\beta|^2 \underbrace{\left(\frac{\lambda_2}{2} + \omega(\alpha \mathbf{t}_1) \right)}_{\to \frac{\lambda_2}{2} > 0} > 0$$

Quindi in ogni intorno U abbiamo determinato punti $\mathbf{x}$ e $\mathbf{y}$ in U tali che

$$f(\mathbf{x}) < f(\mathbf{x}_0) < f(\mathbf{y})$$

ovvero $\mathbf{x}_0$ è un punto sella. □

☞ESEMPIO 3.44. Determinare massimi e minimi relativi della funzione $f(x, y) = xy\, e^{-(x^2+y^2)/2}$.

La funzione è super-regolare, perché è composta di polinomi con la funzione esponenziale. Calcoliamo il gradiente e la matrice Hessiana.

$$
\begin{aligned}
\partial_x f(x,y) &= y\, e^{-(x^2+y^2)/2} - x^2 y\, e^{-(x^2+y^2)/2} = y(1-x^2)\, e^{-(x^2+y^2)/2} \\
\partial_y f(x,y) &= x\, e^{-(x^2+y^2)/2} - x y^2\, e^{-(x^2+y^2)/2} = x(1-y^2)\, e^{-(x^2+y^2)/2} \\
\partial_x^2 f(x,y) &= -2xy\, e^{-(x^2+y^2)/2} - xy(1-x^2)\, e^{-(x^2+y^2)/2} = xy(x^2-3)\, e^{-(x^2+y^2)/2} \\
\partial_{xy} f(x,y) &= \partial_{yx} f(x,y) = (1-x^2)\, e^{-(x^2+y^2)/2} - y^2(1-x^2)\, e^{-(x^2+y^2)/2} \\
&\qquad\qquad\qquad\; = (1-y^2)(1-x^2)\, e^{-(x^2+y^2)/2} \\
\partial_y^2 f(x,y) &= -2xy\, e^{-(x^2+y^2)/2} - xy(1-y^2)\, e^{-(x^2+y^2)/2} = xy(y^2-3)\, e^{-(x^2+y^2)/2}.
\end{aligned}
$$

Iniziamo a selezionare i punti per cui $\nabla f(x,y) = 0$. Questo accade quando

$$
\begin{cases} y\,(1-x^2) = 0 \\ x(1-y^2) = 0 \end{cases} \qquad \text{ovvero} \qquad (x,y) \in \{(0,0),(1,1),(-1,1),(1,-1),(-1,-1)\}\,.
$$

Analizziamo la matrice Hessiana in ciascuno di questi casi:

$$
H(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad H(1,1) = \begin{pmatrix} -2e^{-1} & 0 \\ 0 & -2e^{-1} \end{pmatrix} = H(-1,-1)
$$
$$
H(-1,1) = \begin{pmatrix} 2e^{-1} & 0 \\ 0 & 2e^{-1} \end{pmatrix} = H(1,-1).
$$

Dall'esame degli autovalori o dai criteri visti nel caso $n = 2$ abbiamo che $(0,0)$ è un punto sella, $(1,1)$ e $(-1,-1)$ sono punti di massimo relativo (con massimo relativo $1/e$), $(-1,1)$ e $(1,-1)$ sono punti di minimo relativo (con minimo relativo $-1/e$).

Si può anche notare che $\pm 1/e$ sono massimo e minimo assoluto di f su $\mathbb{R}^2$.

Infatti $\lim_{(x,y)\to\infty} f(x,y) = 0$, quindi esiste $R > 0$ tale che $|f(x,y)| < 1/2e$ se $x^2+y^2 \geq R$. Nel disco chiuso $\{x^2+y^2 \leq R\}$ la funzione f assume massimo e minimo assoluti per il Teorema di Weierstrass. Se il massimo o il minimo assoluti sono assunti in punti interni, allora ritroviamo i valori precedenti. D'altra parte $|f(x,y)| < 1/2e$ se $x^2+y^2 = R$, quindi $\pm 1/e$ sono massimo e minimo assoluti. ☺

E se fosse $\det H = 0$? Il seguente teorema permette di dare una risposta almeno parziale.

TEOREMA 3.45 (Condizioni necessarie per massimo/minimo relativo). *Sia f una funzione con derivate parziali continue fino al secondo ordine in un aperto A di $\mathbb{R}^n$. Sia $\mathbf{x}_0$ un punto in A tale che $\nabla f(\mathbf{x}_0) = 0$ e sia H la matrice Hessiana di f in $\mathbf{x}_0$.*

a) *Se $\mathbf{x}_0$ è un punto di minimo relativo per f allora ${}^t\mathbf{t}H\mathbf{t} \geq 0$ per ogni $\mathbf{t} \in \mathbb{R}^n$ con $\|\mathbf{t}\| = 1$;*
b) *Se $\mathbf{x}_0$ è un punto di massimo relativo per f allora ${}^t\mathbf{t}H\mathbf{t} \leq 0$ per ogni $\mathbf{t} \in \mathbb{R}^n$ con $\|\mathbf{t}\| = 1$.*

Il caso a) si presenta quando la forma quadratica è semidefinita positiva, ossia quando gli autovalori della matrice Hessiana sono tutti maggiori o uguali a zero.

Il caso b) si presenta quando la forma quadratica è semidefinita negativa, ossia quando gli autovalori della matrice Hessiana sono tutti minori o uguali a zero.

DIMOSTRAZIONE. Supponiamo che $\mathbf{x}_0$ sia un punto di minimo relativo per f. Allora per ogni vettore $\mathbf{v}$ di lunghezza 1, la funzione definita da

$$g(t) = f(\mathbf{x}_0 + t\mathbf{v}) \qquad \forall t \, : \, \mathbf{x}_0 + t\mathbf{v} \in A$$

ha un minimo relativo per $t = 0$. Quindi $g'(0) = 0$ e $g''(0) \geq 0$. Calcolando $g''(0)$ come nella dimostrazione del Teorema 3.41 di Taylor, otteniamo

$$0 \leq g''(0) = \sum_{j,k=1}^{n} v_j v_k \partial^2_{kj} f(\mathbf{x}_0) = {}^t\mathbf{v}H\mathbf{v}$$

□

☞ESEMPIO 3.46. Si consideri il punto critico $(0, 0)$ nei seguenti casi:

$$f(x, y) = x^2 + y^4 \qquad g(x, y) = x^2 - y^4 \qquad h(x, y) = (x - y^2)(x - 2y^2).$$

È facile controllare che f, g, h hanno la stessa matrice hessiana in $(0, 0)$, data da

$$H = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

che è semidefinita positiva. Allora possiamo escludere che $(0, 0)$ sia un punto di massimo relativo, ma non possiamo decidere se sia di minimo o sella.

Verifichiamo che $(0, 0)$ è punto di minimo per f e di sella per g.

Infatti, $f(x, y) > 0 = f(0, 0)$ se $(x, y) \neq (0, 0)$, quindi $(0, 0)$ è punto di minimo per f. Invece se $x, y \neq 0$, $g(0, y) = -y^4 < 0$ e $g(x, 0) = x^2 > 0$, quindi $(0, 0)$ è punto di sella per g.

Per quanto riguarda h, notiamo che che ogni restrizione della funzione h a una retta per $(0, 0)$ del tipo $t \mapsto \phi(tv_1, tv_2)$ ha un minimo per $t = 0$. Questo non basta per concludere che $(0,0)$ è punto di minimo per f. In effetti h ha una sella in $(0,0)$, perché, se $x, y \neq 0$, $h(x,0) = x^2 > 0$ e $h(3/2y^2, y) = -1/4y^4 < 0$. ☺

In generale, se vogliamo determinare il massimo e il minimo **assoluti** di una funzione continua su un insieme dato, possiamo ragionare in questo modo. Innanzi tutto, analizziamo l'insieme: se l'insieme è chiuso e limitato, allora possiamo usare il Teorema di Weierstrass e dire che senz'altro esistono massimo e minimo assoluti. Tali valori saranno assunti in punti del tipo:

- punti di non differenziabilità
- punti in cui il gradiente è nullo (detti "punti critici")
- punti della frontiera (del bordo) dell'insieme.

Nella prossima sezione svilupperemo un metodo che ci permetterà di trovare punti di massimo o di minimo sulla frontiera di un insieme.

Se poi l'insieme non è chiuso oppure non è limitato, dobbiamo anche analizzare i limiti (all'infinito se l'insieme non è limitato) della funzione agli estremi dell'insieme.

6. Estremi vincolati

Un esempio di problema vincolato è il seguente: dobbiamo costruire un recipiente per conservare 100mc. di acqua. Per comodità lo costruiremo a forma di parallelepipedo. Il materiale che adoperiamo però è costoso, se desideriamo che l'acqua mantenga inalterate le sue proprietà, quindi decidiamo di costruirlo senza coperchio e di superficie più piccola possibile. Quanto devono misurare gli spigoli, affinché la superficie sia minima?

Possiamo formalizzare il problema precedente in questo modo: se indichiamo con x, y, z le lunghezze dei lati del parallelepipedo allora il volume $V = xyz = 100$mc. e la superficie è $S = xy+2xz+2yz$. Ovviamente le lunghezze sono positive. Desideriamo quindi trovare il minimo assoluto della funzione $xy+2xz+2yz$ definita sull'insieme $\{(x,y,z) \in \mathbb{R}^3 \,:\, x > 0, y > 0, z > 0\}$ soggetta al vincolo $xyz = 100$.

Un primo metodo, consigliabile quando il vincolo è particolarmente semplice, consiste nell'utilizzare il vincolo per ridurre il numero di variabili. Nel nostro caso, possiamo ricondurci al problema di determinare il minimo di una funzione f di due variabili ponendo $z = V/(xy)$. Allora $f(x,y) = S(x,y,V/(xy)) = xy + 2V/y + 2V/x$ è la funzione di cui occorre trovare il minimo sull'insieme $A = \{(x,y) \in \mathbb{R}^2 \,:\, x > 0 \; y > 0\}$.

È facile rendersi conto che ci deve essere il minimo assoluto di f su A, perché se $\bar{x}, \bar{y} \geq 0$ e i limiti si intendono con $(x,y) \in A$, allora

$$\lim_{(x,y)\to\infty} f(x,y) = +\infty \qquad \lim_{(x,y)\to(\bar{x},0)} f(x,y) = +\infty \qquad \lim_{(x,y)\to(0,\bar{y})} f(x,y) = +\infty.$$

Cerchiamo allora tale minimo. Siccome f è differenziabile in A, se (x, y) è punto di minimo allora deve essere $\nabla f(x, y) = 0$, ovvero

$$\partial_x f(x, y) = y - 2V/x^2 = 0 \qquad \text{e} \qquad \partial_y f(x, y) = x - 2V/y^2 = 0$$

che ha come (unica) soluzione in A il punto di coordinate $x = \sqrt[3]{2V}$, $y = \sqrt[3]{2V}$. Siccome tale soluzione è unica, è senz'altro il punto di minimo cercato.

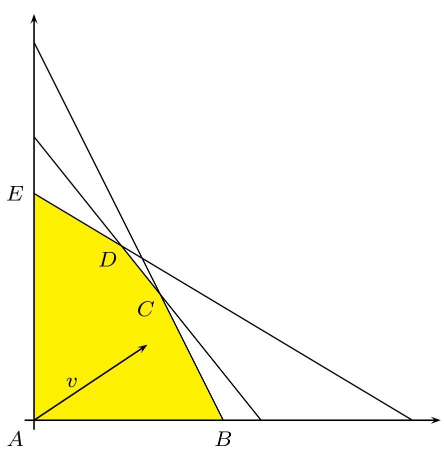
La nostra scatola deve quindi avere base quadrata di lato $\sqrt[3]{2V}$ e altezza $\sqrt[3]{V/4}$. La superficie minima trovata è $3\sqrt[3]{(2V)^2}$.

Un altro esempio di problema vincolato è il seguente. Un produttore di tessuti prepara due qualità di cotone misto lana: una, più pregiata con il 20% in peso di lana, il 50% di cotone e il rimanente 30% di poliestere; l'altra, meno pregiata, con il 10% di lana, il 40% di cotone e il rimanente 50% di poliestere. La qualità più pregiata è venduta a 3 EURO al chilo e quella meno pregiata a 2 EURO al chilo. Il magazzino contiene 2 tonnellate di lana, 6 tonnellate di cotone e 6 tonnellate di poliestere. Quanti chili di ciascun tessuto dovrà produrre per massimizzare il suo guadagno?

Possiamo pensare che x siano le tonnellate di tessuto pregiato da produrre e y le tonnellate di quello meno pregiato. Si desidera massimizzare $f(x, y) = 3000x + 2000y$. I materiali a disposizione in magazzino forniscono i vincoli

$$20\%x + 10\%y \le 2 \qquad 50\%x + 40\%y \le 6 \qquad 30\%x + 50\%y \le 6,$$

insieme alle ovvie richieste $x \ge 0$, $y \ge 0$. In questo caso i vincoli sono espressi in forma di disuguaglianze e quindi otterremo una regione interna, in cui cercare punti critici, e la sua frontiera.



Dobbiamo analizzare i punti interni al poligono in figura per cui $\nabla f = 0$ (non ce ne sono). Infine dobbiamo cercare il massimo assoluto sulla frontiera, ovvero restringere la funzione ai lati AB, BC, CD, DE, EA e su ciascun lato cercare il massimo assoluto della funzione.

In questo caso, f è una funzione lineare, quindi gli eventuali massimo e minimo assoluti si potranno trovare solo nei vertici. Calcoliamo quindi i valori di f in ciascuno dei vertici e otteniamo

$$f(A) = f(0, 0) = 0 \qquad f(B) = f(10, 0) = 30\,000 \qquad f(C) = f(\tfrac{20}{3}, \tfrac{20}{3}) = \tfrac{100\,000}{3} \simeq 33\,333.33$$
$$f(D) = f(\tfrac{60}{13}, \tfrac{120}{13}) = \tfrac{420\,000}{13} \simeq 32\,307.69 \qquad f(E) = f(0, 12) = 24\,000.$$

Il massimo profitto è quindi di 33 333 EURO e si ottiene confezionando 6.6666 tonnellate di tessuto pregiato e 6.6666 tonnellate di tessuto meno pregiato.

Allo stesso risultato saremmo potuti arrivare con un poco meno calcoli utilizzando il significato geometrico del gradiente. Abbiamo visto, nel Teorema 3.30 (del differenziale), che se f è differenziabile in un punto $\mathbf{x}$, allora la derivata rispetto al vettore non nullo $\mathbf{v}$ è data dal prodotto scalare del gradiente di f in $\mathbf{x}$ e il vettore $\mathbf{v}$. In formule, se θ è l'angolo formato da $\nabla f(\mathbf{x})$ e dal vettore $\mathbf{v}$ di norma 1,

$$\partial_{\mathbf{v}} f(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \mathbf{v} \rangle = \|\nabla f(\mathbf{x})\| \, \|\mathbf{v}\| \, \cos\theta = \|\nabla f(\mathbf{x})\| \, \cos\theta.$$

Pertanto f cresce più rapidamente nel verso del gradiente (corrispondente a $\cos\theta = 1$) e decresce più rapidamente nel verso opposto al gradiente (corrispondente a $\cos\theta = -1$). Infine la variazione di f è nulla nella direzione tangente all'ipersuperficie di livello per il punto $\mathbf{x}$ (casi corrispondenti a $\cos\theta = 0$), ovvero il gradiente in $\mathbf{x}$ è ortogonale all'ipersuperficie di livello passante per $\mathbf{x}$.

Tornando all'esempio del produttore di tessuti, il gradiente di f in ogni punto è $(3000, 2000)$. In modulo è molto grande, ma ha la stessa direzione del vettore v disegnato in figura. La freccia del vettore ci indica che il punto di massimo assoluto tra i vertici deve essere C. Questo metodo fornisce una buona indicazione quando sia la funzione sia i vincoli sono lineari.

In sintesi, un problema vincolato consiste nel massimizzare o minimizzare una certa funzione $f(\mathbf{x})$ soggetta a un vincolo definito da una o più equazioni della forma $g(\mathbf{x}) = 0$.

Definizione 3.14. Siano $f, g_1, \ldots, g_k : \mathbb{R}^n \to \mathbb{R}$ funzioni definite in un aperto A di $\mathbb{R}^n$. Si ponga $S = \{\mathbf{x} \in A \, : \, g_1(\mathbf{x}) = \cdots = g_k(\mathbf{x}) = 0\}$. Un punto $\mathbf{x}_0$ in S è un punto di massimo relativo per f soggetta al vincolo S se esiste un intorno U di $\mathbf{x}_0$ tale che $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ per ogni $\mathbf{x} \in S \cap U$.

Analoga è la definizione di punto di minimo relativo vincolato. In generale si dice punto di estremo (relativo) vincolato un punto di massimo o di minimo (relativo) vincolato.

Teorema 3.47 (dei moltiplicatori di Lagrange). *Siano $f, g_1, \ldots, g_k : \mathbb{R}^n \to \mathbb{R}$ funzioni con derivate parziali continue in un aperto A e sia $k < n$. Indichiamo con S il vincolo, ossia $S = \{\mathbf{x} \in \mathbb{R}^n \, : \, g_1(\mathbf{x}) = \cdots = g_k(\mathbf{x}) = 0\}$. Supponiamo che $\mathbf{x}_0 \in S$ sia un punto di estremo relativo vincolato per f e che la matrice*

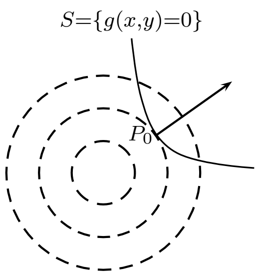
$$\begin{pmatrix} \partial_1 g_1(\mathbf{x}_0) & \partial_1 g_2(\mathbf{x}_0) & \ldots & \partial_1 g_k(\mathbf{x}_0) \\ \partial_2 g_1(\mathbf{x}_0) & \partial_2 g_2(\mathbf{x}_0) & \ldots & \partial_2 g_k(\mathbf{x}_0) \\ \vdots & \vdots & & \vdots \\ \partial_n g_1(\mathbf{x}_0) & \partial_n g_2(\mathbf{x}_0) & \ldots & \partial_n g_k(\mathbf{x}_0) \end{pmatrix}$$

abbia rango k. Allora esistono $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ tali che $\nabla f(\mathbf{x}_0) + \sum_{j=1}^{k} \lambda_j \, \nabla g_j(\mathbf{x}_0) = 0$.

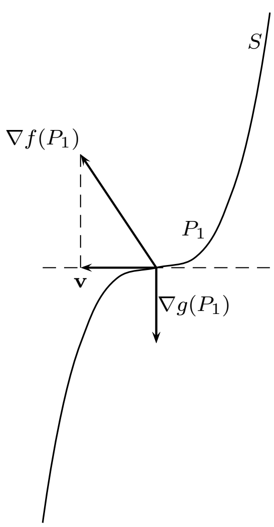
I numeri $\lambda_1, \ldots, \lambda_k$ vengono detti moltiplicatori di Lagrange.

Useremo solo il caso in cui il vincolo è definito da una sola equazione $g(\mathbf{x}) = 0$. In questo caso il teorema afferma che se $\mathbf{x}_0$ è un estremo vincolato e $\nabla g(\mathbf{x}_0) \neq \mathbf{0}$, allora esiste $\lambda_0 \in \mathbb{R}$ tale che $\nabla f(\mathbf{x}_0) + \lambda_0 \nabla g(\mathbf{x}_0) = 0$, ovvero $(\lambda_0, \mathbf{x}_0)$ è un punto critico per la funzione (detta lagrangiana)

$$L(\lambda, \mathbf{x}) = f(x, y) + \lambda g(x, y).$$



Nel disegno, la situazione nel caso $n = 2$. Tratteggiate sono alcune linee di livello della funzione. La condizione trovata vuol semplicemente dire che se un punto è un estremo relativo vincolato, allora i gradienti di f e g in quel punto devono essere paralleli, come in P_0.



Altrimenti, se ∇g e ∇f nel punto di estremo relativo non fossero paralleli, come in P_1, allora la proiezione di $\nabla f(P_1)$ lungo la retta tangente al vincolo in P_1 sarebbe non nulla. Chiamiamo $\mathbf{v}$ questa proiezione.

Ma allora f avrebbe una derivata direzionale positiva nella direzione di $\mathbf{v}$ e una negativa nella direzione opposta.

Quindi muovendosi lungo il vincolo S, la funzione f crescerebbe o decrescerebbe allontanandosi da P_1 nella direzione di $\mathbf{v}$ o di $-\mathbf{v}$. Quindi P_1 non può essere né punto di minimo né punto di massimo relativo per f soggetta al vincolo a S.

Si noti che il teorema dei moltiplicatori di Lagrange non garantisce che una soluzione esista: esso fornisce un metodo per trovare una soluzione che, per altre considerazioni fatte, deve esistere. Inoltre non si occupa dei punti in cui $\nabla g = 0$. Questo è lo scopo dei due esercizi seguenti: far vedere che non è detto che il metodo fornisca una soluzione (occorre stabilire per altra via che questa debba esistere) e far vedere il ruolo della condizione $\nabla g = 0$.

☞Esempio 3.48. Trovare, se esistono, massimi e minimi assoluti di $f(x, y) = x^2 y$ soggetta al vincolo $S = \{g(x, y) = y - x = 0\}$.

È facile rendersi conto che il problema dato non ha soluzione, perché la funzione f ristretta al vincolo coincide con $f(x, x) = x^3$ che non è limitata né inferiormente né superiormente.

Tuttavia se scriviamo $L(x, y, \lambda) = x^2y + \lambda(y - x)$ e calcoliamo

$$\frac{\partial L}{\partial \lambda} = y - x = 0$$
$$\frac{\partial L}{\partial x} = 2xy - \lambda = 0$$
$$\frac{\partial L}{\partial y} = x^2 + \lambda = 0$$

troviamo come soluzione $x = 0, y = 0, \lambda = 0$.

☞Esempio 3.49. Verificare che $(0, 0)$ è punto di minimo assoluto per $f(x, y) = x$ soggetta al vincolo $S = \{(x, y) \in \mathbb{R}^2 \, : \, g(x, y) = y^2 - x^3 = 0\}$.

Il fatto che $f(0, 0) = 0$ sia minimo assoluto vincolato a S è ovvio, dal momento che se $(x, y) \in S$, allora $x \geq 0$, quindi $f(x, y) = x \geq 0$.

Col metodo della lagrangiana $L(x, y, \lambda) = x + \lambda(y^2 - x^3)$, siamo ricondotti al trovare le soluzioni di

$$\frac{\partial L}{\partial \lambda} = y^2 - x^3 = 0$$
$$\frac{\partial L}{\partial x} = 1 - 3\lambda x^2 = 0$$
$$\frac{\partial L}{\partial y} = 2\lambda y = 0.$$

La lagrangiana non ha punti critici (l'ultima equazione ha soluzioni $y = 0$ oppure $\lambda = 0$; ma: se $y = 0$, si ricava dalla prima che $x = 0$, ma allora la seconda non è soddisfatta; se invece è $\lambda = 0$, la seconda di nuovo non vale).

La ragione per cui il metodo dei moltiplicatori di Lagrange non funziona è che il gradiente di g nel punto di minimo è nullo. Questo corrisponde al fatto che la funzione g è regolare, ma la sua curva di livello 0 presenta una cuspide in $(0, 0)$. ☺

Quando usiamo il metodo dei moltiplicatori di Lagrange dobbiamo quindi tener conto del fatto che un punto di estremo vincolato può essere:

- un punto in cui non sono continue le derivate parziali di f o di g;
- un punto in cui $\nabla g = 0$;
- un punto "estremo" per il vincolo;
- un punto critico della lagrangiana.

☞Esempio 3.50. Determinare, se esiste, il punto della curva $y = x^2 - 4x + \frac{7}{2}$ più vicino all'origine.

Possiamo considerare $g(x, y) = x^2 - 4x + \frac{7}{2} - y$ come la funzione che definisce il vincolo $S = \{(x, y) \in \mathbb{R}^2 \, : \, g(x, y) = 0\}$ e come funzione da minimizzare la $f(x, y) = x^2 + y^2$, che è il quadrato della distanza dall'origine (e ha gli stessi punti di minimo della distanza dall'origine).

Osserviamo innanzi tutto che il minimo assoluto esiste, perché f tende a $+\infty$ se $(x,y) \to \infty$. Inoltre sia f sia g sono funzioni con derivate parziali continue e $\frac{\partial g}{\partial y}(x,y) = -1$, quindi ∇g non è mai nullo.

In questo caso la funzione lagrangiana è

$$L(x,y,\lambda) = x^2 + y^2 + \lambda \left(x^2 - 4x + \tfrac{7}{2} - y\right),$$

e ha come punti critici i punti tali che

$$\begin{aligned} \frac{\partial L}{\partial \lambda} &= x^2 - 4x + \tfrac{7}{2} - y = 0 \\ \frac{\partial L}{\partial x} &= 2x + \lambda(2x-4) = 0 \\ \frac{\partial L}{\partial y} &= 2y - \lambda = 0. \end{aligned}$$

C'è un unico punto critico della lagrangiana, in corrispondenza di $x = 1, y = \frac{1}{2}, \lambda = 1$. Pertanto il punto $\left(1, \frac{1}{2}\right)$ è il punto di minimo assoluto per f. Si ha $f\left(1, \frac{1}{2}\right) = \frac{5}{4}$, quindi la distanza minima della parabola $y = x^2 - 4x + \frac{7}{2}$ dall'origine è $\frac{\sqrt{5}}{2}$. ☺

☞ESEMPIO 3.51. Determinare estremi assoluti di $f(x,y) = 2x^2 + y^2$ nell'insieme

$$K = \left\{(x,y) \in \mathbb{R}^2 \,:\, x^4 - x^2 + y^2 - 5 \leq 0\right\}.$$

Sia $g(x,y) = x^4 - x^2 + y^2 - 5$. Notiamo innanzi tutto che f è continua e K è un insieme chiuso e limitato, perché

$$g(x,y) \leq 0 \iff (x^2 - \tfrac{1}{2})^2 + y^2 \leq \tfrac{21}{4},$$

quindi $K \subset \left\{(x,y) \,:\, |x| \leq \sqrt{\frac{1+\sqrt{21}}{2}}, |y| \leq \frac{\sqrt{21}}{2}\right\}$.

Allora per il Teoreme di Weierstrass f ha estremi assoluti in K. Occupiamoci dapprima dei punti interni, valutando i punti critici di f. Si ha

$$\nabla f(x,y) = (4x, 2y) = (0,0) \iff (x,y) = (0,0) \quad \text{e } f(0,0) = 0.$$

Occupiamoci ora dei punti sul bordo, ovvero dei punti su

$$S = \left\{(x,y) \in \mathbb{R}^2 \,:\, g(x,y) = x^4 - x^2 + y^2 - 5 = 0\right\}.$$

Notiamo che

$$\nabla g(x,y) = (4x^3 - 2x, 2y) = (0,0) \iff (x,y) \in \left\{(0,0), (\pm\tfrac{1}{\sqrt{2}}, 0)\right\}$$

e che $(0,0), (\pm\frac{1}{\sqrt{2}}, 0)$ non sono su S, quindi

$$\nabla g(x,y) = (4x^3 - 2x, 2y) \neq (0,0) \qquad \forall (x,y) \in S.$$

Utilizziamo il Teorema dei moltiplicatori di Lagrange e risolviamo il sistema

$$\begin{cases} x^4 - x^2 + y^2 - 5 = 0 \\ 4x + \lambda(4x^3 - 2x) = 0 \\ 2y + \lambda 2y = 0 \end{cases} \qquad \begin{cases} x^4 - x^2 + y^2 - 5 = 0 \\ x(2 + \lambda(2x^2 - 1)) = 0 \\ 2y(1 + \lambda) = 0 \end{cases}$$

L'ultima equazione ha soluzioni $y = 0$ oppure $\lambda = -1$. Se $y = 0$ si ricava dalla prima che $x = \pm\sqrt{\frac{1+\sqrt{21}}{2}}$ e un certo valore di λ dalla seconda.

Se $\lambda = -1$ si ricava dalla seconda equazione che $x = 0$ oppure $x = \pm\sqrt{\frac{3}{2}}$. Per $x = 0$ nella prima si ottiene $y = \pm\sqrt{5}$; per $x = \pm\sqrt{\frac{3}{2}}$ nella prima si ricava $y = \pm\sqrt{17}/2$.

Valutiamo ora f in tutti questi punti (e ricordiamo che $f(0,0) = 0$)

$$f(0, \pm\sqrt{5}) = 5 \qquad f(\pm\sqrt{\tfrac{3}{2}}, \pm\tfrac{\sqrt{17}}{2}) = \tfrac{29}{4} \qquad f(\pm\sqrt{\tfrac{1+\sqrt{21}}{2}}, 0) = 1 + \sqrt{21}.$$

Ne ricaviamo che 0 è il minimo assoluto e $\frac{29}{4}$ è il massimo assoluto. ☺

7. Esercizi

1) Determinare il dominio e i punti di accumulazione per il dominio delle seguenti funzioni $f : \mathbb{R}^2 \to \mathbb{R}$ e, eventualmente con l'aiuto di un software, disegnare le linee di livello:

$$f(x,y) = \frac{x-y}{x+y}; \qquad f(x,y) = \sqrt{xy}; \qquad f(x,y) = \log(1+xy); \qquad f(x,y) = \frac{xy}{x^2 - y^2}.$$

2) Calcolare, se esistono, i seguenti limiti:

$$\lim_{(x,y)\to(0,0)} \frac{x^2+y^2}{y}; \qquad \lim_{(x,y)\to(0,0)} \frac{x^2y^2}{2x^4+y^4}; \qquad \lim_{(x,y)\to(0,0)} \frac{\sin(x-y)}{\cos(x+y)}; \qquad \lim_{(x,y)\to(1,2)} \frac{2x^2-xy}{4x^2-y^2};$$

$$\lim_{(x,y)\to\infty} x - y; \qquad \lim_{(x,y)\to\infty} x^2 + y; \qquad \lim_{(x,y)\to\infty} x^2 + \sqrt{|y|} + \cos y; \qquad \lim_{(x,y)\to\infty} \frac{x^2 y}{x^2 + (y-1)^2}.$$

3) Dire se i seguenti sottoinsiemi di $\mathbb{R}^2$ sono connessi e/o compatti

$$\left\{(x,y) \in \mathbb{R}^2 \,:\, x^2 + y \geq 0\right\}, \qquad \left\{(x,y) \in \mathbb{R}^2 \,:\, x^2 + 9y^2 \leq 4\right\}, \qquad \left\{(x,y) \in \mathbb{R}^2 \,:\, xy > 2\right\}.$$

4) Dire dove le seguenti funzioni sono differenziabili:

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

$$f(x,y) = \begin{cases} xy\frac{y^2-x^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0); \end{cases} \qquad f(x,y) = \begin{cases} xy & x \geq 0 \\ y & x < 0; \end{cases}$$

$$f(x,y) = \begin{cases} \frac{x^3y}{x^4+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0); \end{cases} \qquad f(x,y) = \begin{cases} \frac{x^4y}{x^4+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

5) Determinare e classificare i punti critici delle funzioni:

$$f(x,y) = 2x^3 - 6xy + 3y^2 \qquad f(x,y) = \frac{x}{y} + \frac{8}{x} - y \qquad f(x,y) = \cos x + \cos y.$$

6) In quale direzione aumentano maggiormente le seguenti funzioni in ciascuno dei punti assegnati

$$\begin{aligned} &f(x,y) = 3x - 4y \qquad \text{in } (0,2) \\ &f(x,y) = \frac{x}{1+y} \qquad \text{in } (0,0) \\ &f(x,y) = \log(x^2+y^2) \qquad \text{in } (1,1)? \end{aligned}$$

7) Determinare massimo e minimo assoluti di $2xy$ sull'insieme definito dalla disequazione $|x| + |y| \leq 4$.

8) (difficilino) Determinare, se esistono, massimi e minimi assoluti su $\mathbb{R}^2$ per le funzioni dell'esercizio 1, seconda riga.

9) (difficile) La temperatura nel punto (x,y) è data dalla funzione

$$f(x,y) = \frac{140 + 30x^2 - 60x + 120y^2}{8 + x^2 - 2x + 4y^2}.$$

Determinare, se esiste, la temperatura minima.

10) Dobbiamo preparare un pacco postale. L'ufficio richiede che i pacchi siano a forma di parallelepipedo rettangolo in modo che la somma dell'altezza più il perimetro della base non ecceda 6 metri. Qual è il volume massimo che possiamo spedire?

11) Massimizzare, se possibile, $x^3\, y^5$ soggetta al vincolo $x + y = 8$.

12) Determinare la distanza minima dall'origine della superficie $xyz^2 = 2$.

13) Determinare i valori massimo e minimo di $xy + z^2$ sulla palla $x^2 + y^2 + z^2 \leq 1$.

14) Si deve costruire una scatola di base rettangolare senza coperchio e di volume fissato V (in metri cubi), impiegando due materiali diversi. Il materiale usato per la base della scatola e per la parte frontale costa cinque volte tanto (al metro quadro) quello da usare per le rimanenti facce. Quali devono essere le dimensioni della scatola, per renderne minimo il costo?

15) Determinare i punti di massimo e di minimo assoluto della funzione $f(x, y) = (x - 2y)^2$ soggetta al vincolo $x^2 + 4y^2 - 1 \leq 0$.

16) Sia S l'insieme $S = \{(x, y) \in \mathbb{R}^2 : x(x^2 - 1) = y^2\}$. Determinare i punti di S aventi ordinata minima e quelli aventi ordinata massima.

17) Determinare i punti di massimo e minimo assoluto di $f(x, y) = e^{y^2 - x^2}$ nell'insieme $S = \{(x, y) \in \mathbb{R}^2 : y \leq 0,\ x^2 - y^2 \leq 0,\ x^2 - y - 4 \leq 0\}$.

8. Complementi su gradiente e curve di livello, $n = 2$

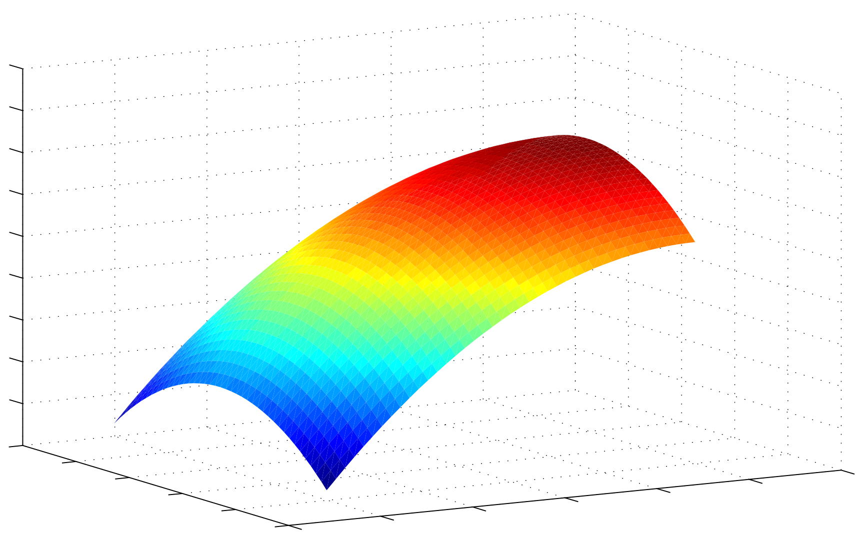
Si ricordi che il grafico di f è un sottoinsieme di $\mathbb{R}^{n+1}$. Abbiamo detto che se $f : A \to \mathbb{R}$ è una funzione definita su un aperto $A \subseteq \mathbb{R}^n$ e differenziabile nel punto $\mathbf{x}_0$ di A, allora l'iperpiano di equazione

$$x_{n+1} = f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle$$

è tangente al grafico di f nel punto $(\mathbf{x}_0, f(\mathbf{x}_0))$.

Il vettore $(\nabla f(\mathbf{x}_0), -1)$ è uno dei due vettori normali al piano tangente al grafico di f e quindi al grafico stesso di f nel punto $(\mathbf{x}_0, f(\mathbf{x}_0))$.

In particolare per $n = 2$ consideriamo il grafico di f in $\mathbb{R}^3$:

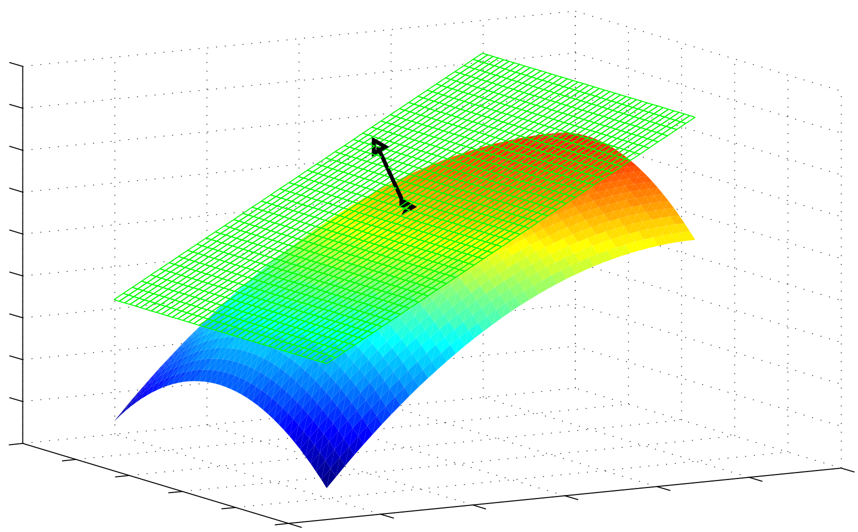


il piano tangente al grafico di f nel punto $(x_0, y_0, f(x_0, y_0))$ è la funzione lineare $\ell(x, y) = f(x_0, y_0) + \partial_x f(x_0, y_0)\,(x - x_0) + \partial_y f(x_0, y_0)\,(y - y_0)$ e ha come grafico l'insieme dei punti (x, y, z) tali che

$$z = f(x_0, y_0) + \partial_x f(x_0, y_0)\,(x - x_0) + \partial_y f(x_0, y_0)\,(y - y_0).$$

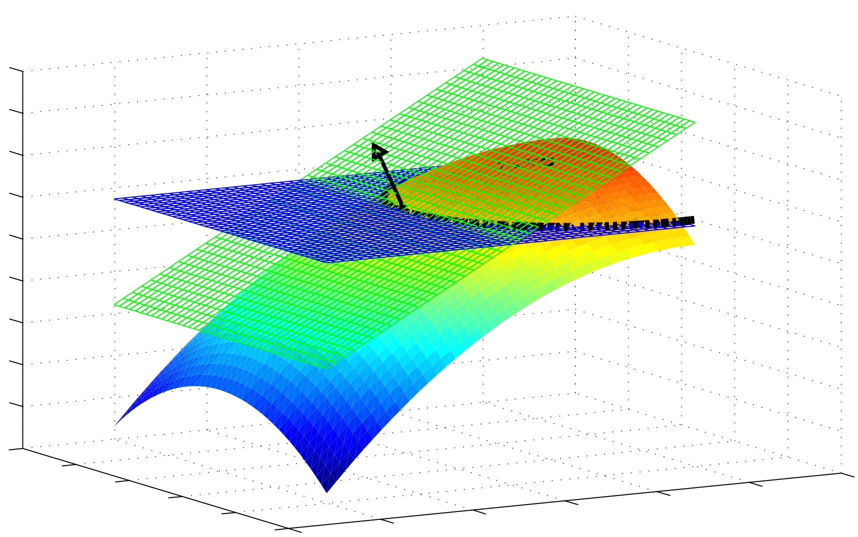
Per una migliore visualizzazione, il vettore normale al piano che abbiamo disegnato è

$$(-\nabla f(x_0, y_0), 1)$$



Sia $\nabla f(x_0, y_0) \neq 0$. Sezioniamo ora il grafico di f con il piano parallelo al piano x, y a quota $f(x_0, y_0)$ e otteniamo una curva C, in formule C è descritta da:

$$\begin{cases} z = f(x,y) \\ z = f(x_0, y_0) \end{cases}$$



Proiettiamo la curva C sul piano x, y: otteniamo così la curva di livello di f passante per il punto (x_0, y_0), ossia la curva

$$C_0 = \big\{(x, y) \in \mathbb{R}^2 \,:\, f(x, y) = f(x_0, y_0)\big\}\,.$$

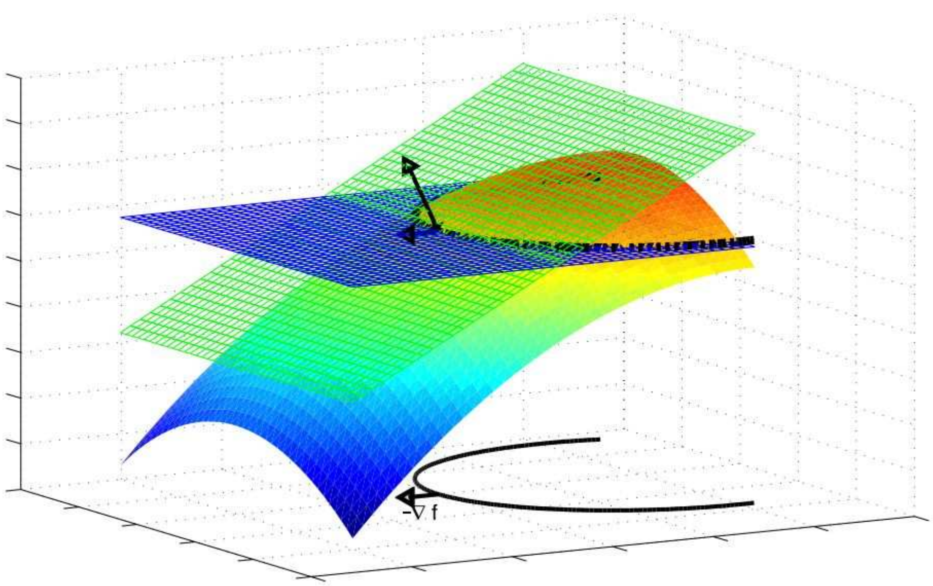
Ripetiamo la stessa operazione con il piano tangente: sezioniamo col piano $z = f(x_0, y_0)$ e otteniamo l'intersezione di due piani

$$\begin{cases} z = f(x_0, y_0) + \partial_x f(x_0, y_0)\,(x - x_0) + \partial_y f(x_0, y_0)\,(y - y_0) \\ z = f(x_0, y_0) \end{cases}$$

Siccome $\nabla f(x_0, y_0) \neq 0$, i due piani non sono coincidenti e la loro intersezione è una retta. Tale retta sta nel piano tangente al grafico, quindi è tangente alla curva C. Proiettiamo ancora sul piano x, y e otteniamo la retta r_0 di equazione

$$\partial_x f(x_0, y_0)\,(x - x_0) + \partial_y f(x_0, y_0)\,(y - y_0) = 0$$

che risulta quindi essere tangente alla curva di livello C_0. Il vettore $\nabla f(x_0, y_0)$ è ortogonale a r_0 e quindi a C_0 nel punto (x_0, y_0).



Abbiamo quindi verificato che

> PROPOSIZIONE 3.52. *Sia f differenziabile in (x_0, y_0) e $\nabla f(x_0, y_0) \neq 0$. Allora la curva di livello $\{(x, y) \,:\, f(x, y) = f(x_0, y_0)\}$ ha un vettore tangente* $\mathbf{t}$ *nel punto (x_0, y_0) e il gradiente $\nabla f(x_0, y_0)$ è perpendicolare a* $\mathbf{t}$ *nel punto (x_0, y_0).*