Funzioni di più variabili 1

Funzioni, grafici, insiemi di livello

$$f : \mathbb{R}^n \to \mathbb{R}\,; \qquad \mathbf{x} = (x_1, x_2, ..., x_n) \mapsto f(\mathbf{x}) = f(x_1, x_2, ..., x_n)$$

Esempi ($n = 2, 3$).

$$f : \mathbb{R}^2 \to \mathbb{R}\,, \quad f(x, y) = e^x \sin y\,.$$

$$g : D \to \mathbb{R}, \quad g(x, y) = \frac{1}{x - y}\,; \quad D = \{(x, y) \,|\, x \neq y\}\,.$$

$$h : B \to \mathbb{R}\,, \quad h(x, y) = \sqrt{1 - x^2 - y^2}\,; \quad B = \{(x, y) \,|\, x^2 + y^2 \leq 1\}\,.$$

$$U : \mathbb{R}^3 \backslash \{\mathbf{0}\} \to \mathbb{R}\,; \quad U(x, y, z) = \frac{k}{\sqrt{x^2 + y^2 + z^2}}$$

Diversi *domini di definizione* (insiemi di esistenza).

Grafico: Data $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, si dice *grafico* di f l'insieme

$$\big\{ \big(\mathbf{x}, f(\mathbf{x})\big) \,\big|\, \mathbf{x} \in D \big\} \subset \mathbb{R}^{n+1} \,.$$

Se $n = 2$, il grafico si chiama "superficie cartesiana" di equazione $z = f(x, y)$.

Esempio. Il grafico di $h(x, y) = \sqrt{1 - x^2 - y^2}$ è la superficie della *semisfera* di equazione $z = \sqrt{1 - x^2 - y^2}$ di raggio unitario e centro nell'origine $(0, 0, 0)$.

Insiemi di livello: Si definisce *insieme di livello c di f* l'insieme

$$\big\{ \mathbf{x} \in D \,\big|\, f(\mathbf{x}) = c \big\} \subset \mathbb{R}^n \,.$$

Se $n = 2$, curve di livello; se $n = 3$, superfici di livello.

Esercizio. Disegnare, al variare di $c \in \mathbb{R}$, le curve di livello c delle funzioni

$$f(x, y) = 2x^2 + y^2 \quad \text{e} \quad g(x, y) = x^2 - y^2 \,.$$

Topologia dei sottoinsiemi di $\mathbb{R}^n$

Si dice intorno sferico di $\mathbf{x}_0 \in \mathbb{R}^n$ un insieme

$$B_r(\mathbf{x}_0) := \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; |\mathbf{x} - \mathbf{x}_0| < r \right\}$$

per qualche $r > 0$.

Sia E è un sottoinsieme di $\mathbb{R}^n$, e sia $E^c := \mathbb{R}^n \backslash E$.

Un punto $\mathbf{x}_0 \in \mathbb{R}^n$ si dice:

i) *interno* ad E se esiste un intorno di $\mathbf{x}_0$ contenuto in E ;

ii) *esterno* ad E se esiste un intorno di $\mathbf{x}_0$ contenuto in E^c ;

iii) *di frontiera* per E se non è interno né esterno .

Equivalente a iii): un punto $\mathbf{x}_0$ è di frontiera per E se *ogni* intorno di $\mathbf{x}_0$ contiene sia punti di E che punti di E^c.

Un punto interno ad E appartiene ad E; un punto esterno ad E appartiene ad E^c.
Un punto di frontiera può appartenere ad E o a E^c.

- $\mathring{E} :=$ insieme dei punti interni ad E ;
- $\partial E :=$ insieme dei punti di frontiera per E ;
- $\bar{E} := E \cup \partial E$, chiusura di E .

Ovviamente: $\mathring{E} \subseteq E \subseteq \bar{E}$.

<u>Definizione</u> Un punto $\mathbf{x}_0$ si dice **punto di accumulazione** per E se *ogni* intorno di $\mathbf{x}_0$ contiene punti di E *diversi da* $\mathbf{x}_0$.

oppure...se in ogni intorno di $\mathbf{x}_0$ esistono *infiniti punti di* E.

Osservazioni.

Un punto di accumulazione può appartenere o non appartenere a E;

ogni punto interno di E è di accumulazione;

un punto di E che *non* è di accumulazione si dice *punto isolato*; un punto isolato è necessariamente di frontiera.

Esercizio.

Disegnare nel piano cartesiano l'insieme

$$E = \left\{(x, y) \,\middle|\, 1 < x^2 + y^2 \leq 4\right\} \cup \left\{(0, 0)\right\}.$$

e identificare i punti interni, esterni, di frontiera, di accumulazione, isolati.

Definizioni.

Un insieme $E \subseteq \mathbb{R}^n$ si dice:

- **aperto** se $E = \mathring{E}$;
- **chiuso** se E^c è aperto;
- **limitato** se $E \subseteq B_R(\mathbf{0})$ per qualche $R > 0$;
- **connesso** (per archi) se per ogni coppia di punti $\mathbf{x}$, $\mathbf{y}$ di E esiste una curva continua *contenuta in E* che ha per estremi i due punti.

Caratterizzazioni utili:

E è aperto $\Leftrightarrow$ E non contiene punti della sua frontiera;

E è chiuso $\Leftrightarrow \partial E \subseteq E$, cioè se e solo se $E = \bar{E}$.

E è chiuso $\Leftrightarrow$ E contiene tutti i suoi punti di accumulazione.

E è limitato $\Leftrightarrow$ esiste $R > 0$ tale che $\mathbf{x} \in E \Rightarrow |\mathbf{x}| \leq R$.

($\emptyset$ e $\mathbb{R}^n$ sono aperti *e* chiusi.)

Esempi.

Consideriamo gli insiemi

$$B = \{(x,y) \,|\, x^2 + y^2 \leq 1\}\,, \qquad D = \{(x,y) \,|\, x \neq y\} \quad \text{e} \quad \mathbb{R}^3 \backslash \{\mathbf{0}\}\,.$$

Il cerchio B include la sua frontiera, per cui è chiuso e limitato; inoltre B, essendo un insieme *convesso*, è anche connesso.

L'insieme D è unione dei due semipiani aperti $x - y > 0$ e $x - y < 0$; poiché l'unione di due (o più) aperti è ancora un insieme aperto, D è aperto.

Ovviamente, D non è limitato.

Infine, D *non* è connesso perché qualunque curva continua che unisce punti dei due semipiani *deve attraversare* la retta $x = y$ (dimostrarlo formalmente usando il teorema degli zeri.)

Verificare che l'insieme $\mathbb{R}^3 \backslash \{\mathbf{0}\}$ è aperto, non limitato e connesso.

Limiti e continuità

Sia $f : D \to \mathbb{R}$, con $D \subseteq \mathbb{R}^n$ e sia $\mathbf{x}_0$ punto di accumulazione per D.

Definizione. Si dice che

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

se vale

$$\forall\, \epsilon > 0 \quad \exists\, r \,:\, \mathbf{x} \in D \text{ e } 0 < |\mathbf{x} - \mathbf{x}_0| < r \;\Rightarrow\; |f(\mathbf{x}) - L| < \epsilon\,.$$

$$\big(\forall\, \epsilon > 0 \quad \exists\, B_r(\mathbf{x}_0) \,:\, \mathbf{x} \in D \cap B_r(\mathbf{x}_0) \setminus \{\mathbf{x}_0\} \;\Rightarrow\; |f(\mathbf{x}) - L| < \epsilon\,.\big)$$

Si scrive in breve:

$$f(\mathbf{x}) \to L \quad \text{per} \quad \mathbf{x} \to \mathbf{x}_0$$

Si dice che

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = +\infty \quad (-\infty)$$

se

$$\forall\, M > 0 \quad \exists\, r \;:\; \mathbf{x} \in D \text{ e } 0 < |\mathbf{x} - \mathbf{x}_0| < r \;\Rightarrow\; f(\mathbf{x}) > M \quad (< -M)\,.$$

Dal punto di vista formale, sono definizioni simili a quelle dei limiti di funzioni di una variabile reale.

$\Rightarrow$ *Valgono i noti teoremi sui limiti*.

Maggiori difficoltà nelle *verifiche e nei calcoli*.

<u>Definizione</u>. Si dice che f è **continua** in $\mathbf{x}_0 \in D \subseteq \mathbb{R}^n$ se

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

Se f è continua in ogni punto di D si dice che è continua in D.

Verifiche di esistenza e non esistenza dei limiti (per $n = 2$).

$$\mathbf{x} = (x, y); \qquad |\mathbf{x} - \mathbf{x}_0| = \sqrt{(x - x_0)^2 + (y - y_0)^2} \equiv \rho\,.$$

Per verificare che

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L\,,$$

è spesso comodo utilizzare le coordinate polari centrate in (x_0, y_0):

$$x = x_0 + \rho\cos\theta\,, \qquad y = y_0 + \rho\sin\theta\,, \qquad \rho \geq 0\,,\ \ 0 \leq \theta < 2\pi\,.$$

Il limite è verificato se (c.s.) esiste $g : \mathbb{R}_+ \to \mathbb{R}$ tale che

$$|f(x_0 + \rho\cos\theta, y_0 + \rho\sin\theta) - L| \leq g(\rho) \to 0$$

per $\rho \to 0$.

Infatti:

$$(x, y) \to (x_0, y_0) \;\Leftrightarrow\; \rho \to 0$$

e, per la precedente maggiorazione,

$$g(\rho) \to 0 \;\Rightarrow\; f(x, y) \to L\,.$$

<u>Esempi</u>.

Verificare che

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

Ponendo $x = \rho\cos\theta$, $y = \rho\sin\theta$,

$$\left|\frac{x^2 y}{x^2 + y^2} - 0\right| = \left|\frac{\rho^3\cos^2\theta\,\sin\theta}{\rho^2}\right| = \rho|\cos^2\theta\,\sin\theta| \le \rho\,.$$

Poiché $\rho \to 0$, il limite è verificato.

Verificare che

$$\lim_{(x,y)\to(0,1)} x\cos\left(\frac{1}{y-1}\right)+y \;=\; 1\,.$$

Qui conviene usare le disuguaglianze:

$$\left|x\cos\left(\frac{1}{y-1}\right)+y-1\right| \le \left|x\cos\left(\frac{1}{y-1}\right)\right| + \left|y-1\right| \le |x|+|y-1|\,,$$

e ricavare poi il risultato ponendo $x=\rho\cos\theta$, $y=1+\rho\sin\theta$.

Esempio di non esistenza del limite.

Sia

$$f(x,y)=\frac{xy}{x^2+y^2} \qquad (x,y)\neq(0,0)$$

Allora $\lim_{(x,y)\to(0,0)} f(x,y)$ non esiste.

Per verificarlo, avviciniamoci all'origine lungo le due rette $x = t$, $y = t$ e $x = t$, $y = -t$, dove $t \to 0$.

Le *restrizioni* della funzione f alle due rette valgono rispettivamente

$$f(t,t) = 1/2 \quad \text{e} \quad f(t,-t) = -1/2\,,$$

per cui hanno *limiti diversi* ($\pm 1/2$) per $t \to 0$. Ma il limite di f, se esiste, è *unico*. □

Esercizio

Verificare che

$$\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^4 + y^2} = 0\,,$$

mentre

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^4 + y^2}$$

non esiste.

Dai teoremi sui limiti si deducono le usuali *proprietà delle funzioni continue* (continuità della somma, del prodotto...).

In particolare:

Se $f : \mathbb{R}^n \to \mathbb{R}$ è continua in un punto $\mathbf{x}_0$ e $f(\mathbf{x}_0) > 0$, allora esiste un intorno $B_r(\mathbf{x}_0)$ tale che $f(\mathbf{x}) > 0$ per ogni $\mathbf{x} \in B_r(\mathbf{x}_0)$ (permanenza del segno).

<u>Esercizio</u>

Dimostrare che se f è continua in $\mathbb{R}^n$, i sottoinsiemi

$$\{\mathbf{x} \,|\, f(\mathbf{x}) > 0\}\,, \qquad \{\mathbf{x} \,|\, f(\mathbf{x}) < 0\}\,, \qquad \{\mathbf{x} \,|\, f(\mathbf{x}) \neq 0\}\,,$$

sono *aperti*.

Dimostrare che ogni insieme di livello di una funzione continua in $\mathbb{R}^n$ è un insieme *chiuso*.

Composizione di funzioni:

Se $f : \mathbb{R}^n \to \mathbb{R}$ è continua in $\mathbf{x}_0$ e $g : \mathbb{R} \to \mathbb{R}$ è continua in $f(\mathbf{x}_0)$, la funzione $g \circ f : \mathbb{R}^n \to \mathbb{R}$ è continua in $\mathbf{x}_0$;

Se $\mathbf{r} : \mathbb{R} \to \mathbb{R}^m$ è continua in t_0 e $f : \mathbb{R}^m \to \mathbb{R}$ è continua in $\mathbf{r}(t_0)$, la funzione $f \circ \mathbf{r} : \mathbb{R} \to \mathbb{R}$ è continua in t_0.

Esempio

La funzione $\mathbf{x} \mapsto |\mathbf{x}|$ è continua in $\mathbb{R}^n$ (verificarlo per esercizio).

Se g è continua in $\mathbb{R}_+$, la funzione *radiale* $g(|\mathbf{x}|)$ è continua in $\mathbb{R}^n$.

La funzione composta $f \circ \mathbf{r}\,(t) = f\big(\mathbf{r}(t)\big)$, si dice *restrizione di f alla curva* $\mathbf{r}$.

Funzioni continue: proprietà topologiche

<u>Teorema</u> (Weierstrass).

Sia $E \subset \mathbb{R}^n$ *chiuso e limitato* (compatto) e $f: E \to \mathbb{R}$ *continua*. Allora

$$\text{Esistono} \quad \mathbf{x}_m, \mathbf{x}_M \in E \quad t.c. \quad f(\mathbf{x}_m) \le f(\mathbf{x}) \le f(\mathbf{x}_M) \quad \forall\, \mathbf{x} \in E\,.$$

Cioè, f *assume massimo e minimo in* E.

<u>Osservazioni</u>

i) I punti $\mathbf{x}_m$, $\mathbf{x}_M$ si dicono *punti di minimo e di massimo* per f, mentre $f(\mathbf{x}_m)$ e $f(\mathbf{x}_M)$ sono i *valori* minimo e massimo (valori estremi) di f.

ii) Le ipotesi del teorema sono condizioni *sufficienti* per l'esistenza dei massimi e dei minimi

iii) Nessuna informazione dal teorema sulla molteplicità dei punti di minimo o di massimo.

Teorema (degli zeri).

Sia $E \subset \mathbb{R}^n$ *connesso*, e $f:\ E \to \mathbb{R}$ *continua*.
Supponiamo che esistano due punti $\mathbf{x}$, $\mathbf{y}$ in E tali che $f(\mathbf{x}) > 0$ e $f(\mathbf{y}) < 0$.

Allora esiste un punto $\mathbf{z} \in E$ tale che $f(\mathbf{z}) = 0$.

Dimostrazione:

Poiché E è connesso, esiste una curva $\mathbf{r} : [a, b] \to E$ tale che $\mathbf{r}(a) = \mathbf{x}$ e $\mathbf{r}(b) = \mathbf{y}$. La funzione composta

$$g(t) = f\big(\mathbf{r}(t)\big)$$

è continua in $[a, b]$ e tale che $g(a) > 0$, $g(b) < 0$.

Per il teorema degli zeri unidimensionale, esiste $\bar{t} \in (a, b)$ tale che $g(\bar{t}) = 0$.

Ma:

$$g(\bar{t}) = f\big(\mathbf{r}(\bar{t})\big)\,,$$

per cui f si annulla nel punto $\mathbf{z} = \mathbf{r}(\bar{t})$. $\square$

Osservazione. Se $f:\ \mathbb{R}^n \to \mathbb{R}$ è continua, l'insieme (chiuso) $\{\mathbf{x} \mid f(\mathbf{x}) = 0\}$ divide lo spazio in *componenti connesse* dove f ha segno costante.

Funzioni di più variabili 2

March 9, 2021

Derivate parziali e direzionali

$$f : \mathbb{R}^n \to \mathbb{R}\,; \qquad \mathbf{x} = (x_1, x_2, ..., x_n) \mapsto f(\mathbf{x}) = f(x_1, x_2, ..., x_n)$$

Se $n = 2$:

$$f : D \subseteq \mathbb{R}^2 \to \mathbb{R}\,, \qquad (x, y) \mapsto f(x, y)\,.$$

Assumiamo D aperto e fissiamo $(x_0, y_0) \in D$; possiamo valutare f in (x_0, y_0) e in un intorno $B_r(x_0, y_0)$.

Qual'è il *tasso di variazione di f* se ci spostiamo dal punto fissato lungo una qualunque direzione ?

Consideriamo inizialmente le *direzioni degli assi*; fissiamo $y = y_0$ e muoviamo la x a partire da x_0, oppure fissiamo $x = x_0$ e muoviamo la y a partire da y_0.

Consideriamo quindi i limiti:

$$\lim_{h\to 0} \frac{f(x_0+h,y_0)-f(x_0,y_0)}{h}\,; \qquad \lim_{k\to 0} \frac{f(x_0,y_0+k)-f(x_0,y_0)}{k}\,.$$

Se tali limiti esistono finiti, si chiamano **derivate parziali** di f rispetto a x e rispetto a y, calcolate nel punto (x_0, y_0):

$$\frac{\partial f}{\partial x}(x_0,y_0) := \lim_{h\to 0} \frac{f(x_0+h,y_0)-f(x_0,y_0)}{h}$$

$$\frac{\partial f}{\partial y}(x_0,y_0) := \lim_{k\to 0} \frac{f(x_0,y_0+k)-f(x_0,y_0)}{k}$$

Altre notazioni di uso comune:

$$\partial_x f(x_0,y_0)\,,\;\partial_y f(x_0,y_0)\,; \quad f_x(x_0,y_0)\,,\;f_y(x_0,y_0)\,; \quad D_1 f(x_0,y_0)\,,\;D_2 f(x_0,y_0)\,.$$

Le derivate parziali sono semplicemente le derivate (ordinarie) delle funzioni $x \mapsto f(x, y_0)$, $y \mapsto f(x_0, y)$, cioè delle *restrizioni* di f rispettivamente alle rette $y = y_0$ e $x = x_0$.

Dunque valgono le stesse regole di calcolo delle derivate ordinarie, con l'avvertenza: quando si deriva rispetto ad una variabile, l'altra viene considerata come una *costante*.

Esempio.

Derivate parziali di

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

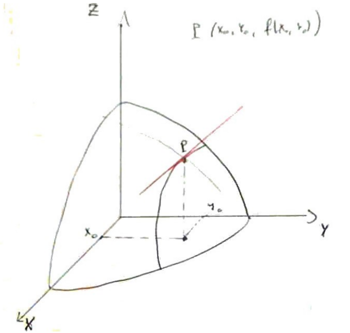
nel generico punto (x, y):

$$\partial_x f(x, y) = (x^2 + 2x + y)e^x + y\cos(xy)\,;$$

$$\partial_y f(x, y) = e^x + x\cos(xy)\,.$$

Significato geometrico di $\partial_x f(x_0, y_0)$ $\left(\partial_y f(x_0, y_0)\right)$:

Pendenza (coefficiente angolare) della retta tangente alla curva $z = f(x, y_0)$ $\left(z = f(x_0, y)\right)$, intersezione della superficie $z = f(x, y)$ con il piano (verticale) $y = y_0$ $(x = x_0)$.



Esercizio: verificare che i *vettori tangenti* alle due sezioni della superficie, considerate come curve regolari, sono rispettivamente:

$$\mathbf{i} + f_x(x_0, y_0)\mathbf{k}\,, \quad \mathbf{j} + f_y(x_0, y_0)\mathbf{k}\,.$$

Derivate parziali per $n \geq 2$.

$$f : D \subseteq \mathbb{R}^n \to \mathbb{R}\,; \qquad \mathbf{x} = (x_1, x_2, ..., x_n) \mapsto f(\mathbf{x}) = f(x_1, x_2, ..., x_n)$$

Sia D aperto, $\mathbf{x}_0 \in D$; $\mathbf{e}_i = (0, 0, ..., 1, ...0)$ denota il *versore dell'asse* x_i.

La derivata parziale di f rispetto a x_i in $\mathbf{x}_0$ si definisce

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) := \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\,\mathbf{e}_i) - f(\mathbf{x}_0)}{h}$$

Si muove solo l'i-esima variabile indipendente x_i.

Altre notazioni:

$$\partial_{x_i} f(\mathbf{x}_0)\,; \quad f_{x_i}(\mathbf{x}_0)\,; \quad D_i f(\mathbf{x}_0)\,.$$

Se esistono tutte le derivate parziali $\partial_{x_i} f(\mathbf{x}_0)$, $i = 1, 2, ..., n$, la funzione si dice **derivabile** in $\mathbf{x}_0$. Se f è derivabile in ogni punto di D, si dice derivabile in D.

Se f è derivabile in un punto $\mathbf{x}$, si definisce **gradiente** di f il vettore

$$\nabla f(\mathbf{x}) := \big(\partial_{x_1} f(\mathbf{x}), \partial_{x_2} f(\mathbf{x}), ..., \partial_{x_n} f(\mathbf{x})\big) = \sum_{i=1}^{n} \partial_{x_i} f(\mathbf{x})\, \mathbf{e}_i \,.$$

Esempi

Il gradiente della funzione $f(x, y) = x\, e^y$ nel generico punto (x, y) è

$$\nabla f(x, y) = e^y\, \mathbf{i} + x\, e^y\, \mathbf{j}\,.$$

Perciò:

$$\nabla f(0, 0) = \mathbf{i}\,, \quad \nabla f(1, 0) = \mathbf{i} + \mathbf{j}\,, \quad \nabla f(-1, 2) = e^2\, \mathbf{i} - e^2\, \mathbf{j}\,, ...$$

Se $F(x, y, z) = \sqrt{x^2 + y^2 + z^2} \equiv r$, il gradiente

$$\nabla F(x, y, z) = \frac{x}{r}\, \mathbf{i} + \frac{y}{r}\, \mathbf{j} + \frac{z}{r}\, \mathbf{k} \equiv \frac{\mathbf{r}}{r}\,,$$

definito in $R^3 \backslash \{(0, 0, 0)\}$, è in ogni punto fuori dall'origine il *versore radiale*.

E se volessimo valutare la variazione di f quando ci spostiamo da $\mathbf{x}_0$ lungo una direzione *qualsiasi* ?

Se $\mathbf{x}_0 \in D$ (D aperto) e $\mathbf{v}$ è un *versore*, consideriamo la retta $\mathbf{x}_0 + t\mathbf{v}$.

Definizione

Se esiste finito il limite

$$D_{\mathbf{v}} f(\mathbf{x}_0) := \lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\,\mathbf{v}) - f(\mathbf{x}_0)}{t} \,,$$

si dice *derivata direzionale* di f nella direzione di $\mathbf{v}$ calcolata nel punto $\mathbf{x}_0$.

Osservazioni

i) Il limite è la derivata in $t = 0$ della restrizione di f alla retta;

ii) Le derivate parziali sono particolari derivate direzionali (scegliere $\mathbf{v} = \mathbf{e}_i$).

Se $n = 2$, $\mathbf{x} = (x_0, y_0)$, $\mathbf{v} = (\cos\theta, \sin\theta)$ e si può scrivere:

$$D_{\mathbf{v}} f(x_0, y_0) = \lim_{t\to 0} \frac{f(x_0 + t\cos\theta, y_0 + t\sin\theta) - f(x_0, y_0)}{t}\,.$$

<u>Esempio</u>

Sia $f(x, y) = (x^2 y)^{1/3}$ e $\mathbf{v} = (\cos\theta, \sin\theta)$.

$$D_{\mathbf{v}} f(0, 0) = \lim_{t\to 0} \frac{\big((t\cos\theta)^2\, t\sin\theta\,\big)^{1/3} - 0}{t} =$$

$$= \lim_{t\to 0} \frac{t\big(\cos\theta)^{2/3}\,(\sin\theta)^{1/3}}{t} = (\cos\theta)^{2/3}\,(\sin\theta)^{1/3}\,.$$

Osservare che le derivate parziali di f nell'origine ($\theta = 0$ e $\theta = \pi/2$) sono nulle, in accordo con le restrizioni di f agli assi: $f(x, 0) = 0$, $f(0, y) = 0$.

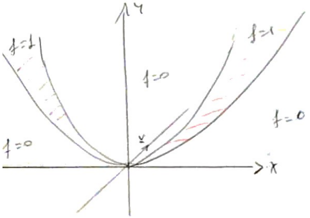
Attenzione:

Se $n = 1$, è noto che: f derivabile in $x_0 \Rightarrow f$ continua in x_0.
Se $n > 1$ l'implicazione *non* vale. Nemmeno se esistono *tutte* le derivate direzionali nel punto $\mathbf{x}_0$.

Esempio

$$f(x,y) = \begin{cases} 1, & \text{if } x^2 < y < 2x^2; \\ 0, & \text{otherwise.} \end{cases}$$

Dalla definizione segue che $f(0,0) = 0$; inoltre, per ogni θ, si vede che per t abbastanza piccolo $f(t\cos\theta, t\sin\theta) = 0$.



Dunque $D_{\mathbf{v}} f(0,0) = 0$ per ogni $\mathbf{v}$, ma f *non* è continua nell'origine.

Differenziabilità

Se $n > 1$, la derivabilità di f è un'informazione 'debole'...

Ripassiamo una definizione già introdotta nel caso $n = 1$:

Sia $f : \mathbb{R} \to \mathbb{R}$ derivabile in $x_0 \in \mathbb{R}$ con derivata $f'(x_0)$.

Vale allora

$$f(x_0 + h) - f(x_0) = f'(x_0)\, h + o(h)\,,$$

dove l'applicazione lineare

$$h \mapsto f'(x_0)\, h = df(x_0)\,,$$

è il *differenziale* di f in x_0 e il termine $o(h)$ verifica

$$\lim_{h \to 0} o(h)/h = 0\,.$$

Definizione

Sia $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ aperto e $\mathbf{x}_0 \in D$. Sia inoltre $\mathbf{h} \in \mathbb{R}^n$ tale che $\mathbf{x}_0 + \mathbf{h} \in D$.

La funzione f si dice **differenziabile** in $\mathbf{x}_0$ se esiste $\mathbf{a} \in \mathbb{R}^n$ tale che

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \mathbf{a} \cdot \mathbf{h} + o(|\mathbf{h}|)\,,$$

dove $o(|\mathbf{h}|)/|\mathbf{h}| \to 0$ per $\mathbf{h} \to \mathbf{0}$. $\square$

Con $\mathbf{a} \cdot \mathbf{h}$ si indica il prodotto scalare dei due vettori:

$$\text{se } \mathbf{a} = (a_1, a_2, ..., a_n)\,, \qquad \mathbf{h} = (h_1, h_2, ..., h_n)\,,$$

allora

$$\mathbf{a} \cdot \mathbf{h} = a_1 h_1 + a_2 h_2 + ... + a_n h_n\,.$$

L'applicazione lineare $\mathbf{h} \mapsto \mathbf{a} \cdot \mathbf{h}$, si dice **differenziale** di f nel punto $\mathbf{x}_0$ e si denota con il simbolo $df(\mathbf{x}_0)$.

Teorema.

Sia f differenziabile in $\mathbf{x}_0 \in D$ (aperto in $\mathbb{R}^n$). Allora:

i) f è continua in $\mathbf{x}_0$;

ii) f è derivabile in $\mathbf{x}_0$ e, nella definizione di differenziabilità, $\mathbf{a} = \nabla f(\mathbf{x}_0)$;

iii) esistono tutte le derivate direzionali di f in $\mathbf{x}_0$ e vale la formula

$$D_{\mathbf{v}} f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v}$$

Dimostrazione:

i) Ponendo $\mathbf{x} = \mathbf{x}_0 + \mathbf{h}$, riscriviamo la definizione di differenziabilità nella forma

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot \left(\mathbf{x} - \mathbf{x}_0\right) + o(|\mathbf{x} - \mathbf{x}_0|)\,,$$

da cui si ottiene facilmente $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$.

ii) Scegliendo $\mathbf{h} = h\,\mathbf{e}_i$ nella definizione di differenziabilità si ottiene

$$f(\mathbf{x}_0 + h\,\mathbf{e}_i) - f(\mathbf{x}_0) = \mathbf{a} \cdot h\,\mathbf{e}_i + o(|h\,\mathbf{e}_i|) = h\,a_i + o(|h|)\,.$$

Dividendo per h e facendo il limite per $h \to 0$ si trova

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = a_i + \lim_{h\to 0} \frac{o(|h|)}{h} = a_i\,, \quad \forall\, i = 1, 2, ..., n\,.$$

iii) Per il punto ii), possiamo ora scrivere nella definizione

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + o(|\mathbf{h}|)\,,$$

Scegliamo ora $\mathbf{h} = t\,\mathbf{v}$, con $\mathbf{v}$ *versore* qualsiasi:

$$f(\mathbf{x}_0 + t\,\mathbf{v}) - f(\mathbf{x}_0) = t\,\nabla f(\mathbf{x}_0) \cdot \mathbf{v} + o(|t|)\,,$$

Dividendo per t e facendo il limite per $t \to 0$ si ottiene ora

$$D_{\mathbf{v}} f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v}$$

Osservazioni importanti.

Se anche una sola delle tre proprietà *i*), *ii*), *iii*) del teorema non è verificata, la f *non* è differenziabile.

Dal punto *ii*) segue $df(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{h}$; denotando con $d\mathbf{x}$ lo spostamento arbitrario $\mathbf{h}$ si scrive anche

$$df(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot d\mathbf{x} = \sum_{i=1}^{n} \partial_{x_i} f(\mathbf{x})\, dx_i \,.$$

Da *iii*) e dalle proprietà prodotto scalare segue:

$$D_{\mathbf{v}} f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v} = \left|\nabla f(\mathbf{x}_0)\right| \cos\alpha \,,$$

dove α è l'angolo tra il vettore $\nabla f(\mathbf{x}_0)$ (se $\neq \mathbf{0}$) ed il versore $\mathbf{v}$.

Dunque la derivata è massima ($= |\nabla f(\mathbf{x}_0)|$) per $\alpha = 0$.
La *direzione di massima crescita* di f in $\mathbf{x}_0$ è quella del versore $\dfrac{\nabla f(\mathbf{x}_0)}{\left|\nabla f(\mathbf{x}_0)\right|}$,

cioè quella del *gradiente*.

Per $n = 2$, ponendo $\mathbf{x}_0 = (x_0, y_0)$, $\mathbf{h} = (h, k)$, la definizione di differenziabilità si scrive

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \partial_x f(x_0, y_0)\, h + \partial_y f(x_0, y_0)\, k + o\big(\sqrt{h^2 + k^2}\big)$$

Introducendo le variabili $x = x_0 + h$, $y = y_0 + k$, abbiamo la relazione equivalente

$$\begin{aligned} f(x, y) \quad &= \quad f(x_0, y_0) + \partial_x f(x_0, y_0)\,(x - x_0) + \partial_y f(x_0, y_0)\,(y - y_0) \\ &+ \quad o\big([(x - x_0)^2 + (y - y_0)^2]^{1/2}\big) \end{aligned}$$

che ha un'importante *intepretazione geometrica*:
il piano di equazione

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

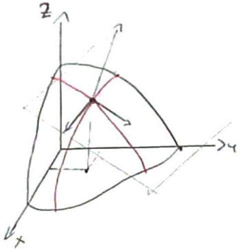
si dice **piano tangente** alla superficie di equazione $z = f(x, y)$ nel punto (x_0, y_0) e rappresenta la migliore *approssimazione lineare* di f in un intorno del punto.

Dalla geometria, il vettore

$$-\partial_x f(x_0, y_0)\,\mathbf{i} - \partial_y f(x_0, y_0)\,\mathbf{j} + \mathbf{k}$$

è perpendicolare al piano tangente. La sua direzione è la direzione *normale* alla superficie $z = f(x, y)$ nel punto (x_0, y_0).
Osservare che

$$-\partial_x f(x_0, y_0)\,\mathbf{i} - \partial_y f(x_0, y_0)\,\mathbf{j} + \mathbf{k} = \left[\mathbf{i} + \partial_x f(x_0, y_0)\,\mathbf{k}\right] \wedge \left[\mathbf{j} + \partial_y f(x_0, y_0)\,\mathbf{k}\right]$$



Verifiche di Differenziabilità.

La verifica diretta non è generalmente agevole.

Ricordando la definizione e il punto *ii*) del teorema, occorre verificare che nel punto considerato la funzione sia derivabile, calcolare il gradiente e infine verificare il limite:

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{f(\mathbf{x}_0+\mathbf{h})-f(\mathbf{x}_0)-\nabla f(\mathbf{x}_0)\cdot\mathbf{h}}{|\mathbf{h}|}=0\,.$$

Nel caso $n=2$:

$$\lim_{(h,k)\to(0,0)} \frac{f(x_0+h,y_0+k)-f(x_0,y_0)-\partial_x f(x_0,y_0)\,h-\partial_y f(x_0,y_0)\,k}{\sqrt{h^2+k^2}}=0\,.$$

Esempio

Verifichiamo che

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

è differenziabile in $(1,0)$.

Calcoliamo:

$$f(1,0) = 1\,, \qquad \partial_x f(1,0) = 1\,, \qquad \partial_y f(1,0) = 2\,.$$

Il limite da verificare è allora:

$$\lim_{(h,k)\to(0,0)} \frac{(1+h)(2k+1) - 1 - \ h - 2\,k}{\sqrt{h^2+k^2}} = \lim_{(h,k)\to(0,0)} \frac{2h\,k}{\sqrt{h^2+k^2}} = 0\,.$$

Teorema (c.s. di differenziabilità)

Se f è derivabile *in un intorno* di $\mathbf{x}_0 \in \mathbb{R}^n$ e le derivate parziali di f sono continue in $\mathbf{x}_0$, allora f è differenziabile in $\mathbf{x}_0$.

In particolare, se f è derivabile in un aperto D con derivate parziali continue $\big(f$ di classe $\mathcal{C}^1(D)\big)$, allora f è differenziabile in (tutti i punti di) D.

Attenzione. Se non valgono le ipotesi del teorema occorre tornare alla verifica diretta:
la funzione $f(x, y) = |xy|$ è differenziabile nell'origine (fare la verifica con la definizione) ma *non* valgono le ipotesi del teorema in un intorno dell'origine.

Esercizio.

Verificare che la funzione $f(x, y) = xe^y$ è differenziabile in tutto $\mathbb{R}^2$. Scrivere l'equazione del piano tangente alla superficie $z = xe^y$ in $(2, 0)$. Calcolare nello stesso punto tutte le derivate direzionali e determinare la direzione di massima crescita di f.

Funzioni di più variabili 3

March 18, 2021

Derivazione delle funzioni composte

Se $f : \mathbb{R}^n \to \mathbb{R}$ è *differenziabile* in $\mathbf{x}_0$ e $g : \mathbb{R} \to \mathbb{R}$ è derivabile in $f(\mathbf{x}_0)$, la funzione $g \circ f : \mathbb{R}^n \to \mathbb{R}$ è differenziabile in $\mathbf{x}_0$ e vale

$$\nabla(g \circ f)(\mathbf{x}_0) = g'\big(f(\mathbf{x}_0)\big)\,\nabla f(\mathbf{x}_0)\,.$$

Se $\mathbf{r} : \mathbb{R} \to \mathbb{R}^m$ è derivabile in t_0 e $f : \mathbb{R}^m \to \mathbb{R}$ è *differenziabile* in $\mathbf{r}(t_0)$, la funzione $f \circ \mathbf{r} : \mathbb{R} \to \mathbb{R}$ è derivabile in t_0 e vale

$$(f \circ \mathbf{r})'(t_0) = \nabla f(\mathbf{r}(t_0)) \cdot \mathbf{r}'(t_0)\,.$$

Gradiente di funzioni radiali.

La funzione $\mathbf{x} \mapsto |\mathbf{x}|$ è differenziabile in $\mathbb{R}^n \backslash \{\mathbf{0}\}$ (le derivate parziali sono continue al di fuori dell'origine).

Se g è derivabile in $\mathbb{R}_+$, la funzione radiale $g(|\mathbf{x}|)$ è differenziabile in $\mathbb{R}^n \backslash \{\mathbf{0}\}$ e vale

$$\nabla g(|\mathbf{x}|) = g'(|\mathbf{x}|) \, \nabla |\mathbf{x}| = g'(|\mathbf{x}|) \, \frac{\mathbf{x}}{|\mathbf{x}|} \, .$$

In particolare, se $n = 3$, $r = \sqrt{x^2 + y^2 + z^2}$, $g(r) = \dfrac{k}{r}$:

$$\nabla g(r) = -\frac{k}{r^2}\left(\frac{x}{r}\,\mathbf{i} + \frac{y}{r}\,\mathbf{j} + \frac{z}{r}\,\mathbf{k}\right) = -k\,\frac{\mathbf{r}}{r^3} \qquad (\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}).$$

Gradiente e curve di livello.

Sia $f(x, y)$ differenziabile (in un aperto di $\mathbb{R}^2$) e supponiamo che l'insieme di livello

$$\{(x, y) \,|\, f(x, y) = c\}$$

sia (sostegno di) una curva regolare $\mathbf{r} = x(t)\,\mathbf{i} + y(t)\,\mathbf{j}$, $t \in I$.

Allora:

$$g(t) := f\big(x(t), y(t)\big) = c, \qquad \forall\, t \in I\,.$$

Derivando l'equazione rispetto a t:

$$g'(t) = \nabla f\big(x(t), y(t)\big) \cdot \mathbf{r}'(t) = 0\,.$$

Dunque, in ogni punto regolare di una curva di livello il gradiente di f (se $\neq \mathbf{0}$) è un vettore *perpendicolare* alla tangente alla curva.

(Si dice in breve che il gradiente è ortogonale alle curve di livello).

Derivate (e differenziali) di ordine superiore.

Sia f una funzione definita in un aperto di $\mathbb{R}^n$ e supponiamo che una derivata parziale $\dfrac{\partial f}{\partial x_i}$ esista in un *intorno* di un punto $\mathbf{x} \in \mathbb{R}^n$.

Se la funzione $\dfrac{\partial f}{\partial x_i}$ è a sua volta derivabile rispetto a x_j in $\mathbf{x}$, è definita la **derivata parziale seconda** rispetto a x_i e a x_j

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\mathbf{x}) \, .$$

Notazioni alternative:

$$\partial_{x_j x_i} f \, , \quad f_{x_j x_i} \, , \quad D^2_{ji} f \, , \ldots$$

Se f è derivabile in un aperto di R^n e le derivate parziali sono a loro volta derivabili, sono definite n^2 derivate parziali seconde.

Esempio (n=2).

Sia $f(x,y) = x^2 \cos y\,,\ \ (x,y) \in \mathbb{R}^2\,;$ le derivate parziali prime sono:

$$\partial_x f(x,y) = 2x\cos y, \qquad \partial_y f(x,y) = -x^2 \sin y\,, \qquad \text{da cui:}$$

$$\partial_{xx} f(x,y) = 2\cos y\,, \quad \partial_{xy} f(x,y) = -2x\sin y\,,$$

$$\partial_{yx} f(x,y) = -2x\sin y\,, \quad \partial_{yy} f(x,y) = -x^2\cos y\,.$$

in ogni punto $(x,y) \in \mathbb{R}^2$.

L'uguaglianza delle derivate miste non è casuale:

Teorema (Schwarz).

Sia f definita in un aperto di $\mathbb{R}^n$ e assumiamo che le derivate parziali seconde $\partial_{x_i x_j} f,\ \ \partial_{x_j x_i} f\ \ (i \neq j)$ esistano in un *intorno* di un punto $\mathbf{x}$ e siano *continue in* $\mathbf{x}$.
Allora $\partial_{x_i x_j} f(\mathbf{x}) = \partial_{x_j x_i} f(\mathbf{x})$.

Se una funzione ha derivate seconde continue in un aperto $D \subseteq \mathbb{R}^n$, si dice che è *di classe* $\mathcal{C}^2(D)$.

Le definizioni di derivate parziali successive si generalizzano alle derivate di ordine k.

Se f ha derivate parziali continue fino all'ordine k in un aperto $D \subseteq \mathbb{R}^n$, si dice che $f \in \mathcal{C}^k(D)$. Anche per queste funzioni, le derivate rispetto a variabili differenti si possono commutare.

Esempio

Le derivate parziali terze di $f(x, y) = x^2 \cos y$, sono:

$$\partial_{xxx} f(x, y) = 0\,, \quad \partial_{yxx} f(x, y) = \partial_{xyx} f(x, y) = \partial_{xxy} f(x, y) = -2 \sin y\,,$$
$$\partial_{xyy} f(x, y) = \partial_{yxy} f(x, y) = \partial_{yyx} f(x, y) = -2x \cos y\,, \quad \partial_{yyy} f(x, y) = x^2 \sin y\,.$$

Esercizio

Calcolare tutte le derivate parziali della funzione $g(x, y, z) = xyz$.

Differenziale secondo

Se f è differenziabile in un aperto $D \subseteq \mathbb{R}^n$, esistono nello stesso aperto le derivate parziali $\partial_{x_i} f$, $i = 1, 2, ..., n$. Se queste derivate sono a loro volta differenziabili in $\mathbf{x}_0 \in D$, si dice che f è *due volte differenziabile* in $\mathbf{x}_0$.

Si definisce *differenziale secondo* di f in $\mathbf{x}_0$ la funzione (forma quadratica)

$$d^2 f(\mathbf{x}_0) \,:\, \mathbf{h} \mapsto \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} f(\mathbf{x}_0) h_i\, h_j\,, \qquad \mathbf{h} = (h_1, h_2, ..., h_n) \in \mathbb{R}^n\,.$$

Più in breve scriveremo

$$d^2 f(\mathbf{x}_0) \,=\, \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0) h_i\, h_j$$

o anche

$$d^2 f(\mathbf{x}_0) \,=\, \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0) dx_i\, dx_j\,.$$

Osservazioni:

Se $f \in \mathcal{C}^2(D)$, le derivate prime $\partial_{x_i} f$ sono differenziabili (per la condizione sufficiente applicata alle derivate paziali).

Dunque: $f \in \mathcal{C}^2(D) \;\Rightarrow\; f$ due volte differenziabile.

Possiamo scrivere il differenziale secondo (come qualsiasi forma quadratica) nel seguente modo:

$$d^2 f(\mathbf{x}_0) \,=\, H_f(\mathbf{x}_0)\,\mathbf{h}\,\cdot\mathbf{h} \quad \big(= \mathbf{h}^T\, H_f(\mathbf{x}_0)\,\mathbf{h}\,\big)\,,$$

dove ora $\mathbf{h}$ è pensato come *vettore colonna* e $H_f(\mathbf{x}_0)$ è la *matrice dei coefficienti*

$$H_f(\mathbf{x}_0) = \begin{pmatrix} f_{x_1 x_1}(\mathbf{x}_0) & f_{x_1 x_2}(\mathbf{x}_0) \;\dots & f_{x_1 x_n}(\mathbf{x}_0) \\ f_{x_2 x_1}(\mathbf{x}_0) & f_{x_2 x_2}(\mathbf{x}_0) \;\dots & f_{x_2 x_n}(\mathbf{x}_0) \\ \dots & \dots & \dots \\ f_{x_n x_1}(\mathbf{x}_0) & f_{x_n x_2}(\mathbf{x}_0) \;\dots & f_{x_n x_n}(\mathbf{x}_0) \end{pmatrix}$$

detta **matrice Hessiana** di f in $\mathbf{x}_0$.
Per quanto visto sopra, questa matrice è *simmetrica*.

Caso $n = 2$: $f : D \to \mathbb{R}$, $\mathbf{x}_0 = (x_0, y_0) \in D$, $\mathbf{h} = (h, k)$:

$$df(x_0, y_0) = \partial_x f(x_0, y_0)h + \partial_y f(x_0, y_0)k = \nabla f(x_0, y_0) \cdot \mathbf{h}\,;$$

$$d^2 f(x_0, y_0) \,=\, \partial_{xx} f(x_0, y_0)h^2 + \partial_{xy} f(x_0, y_0)hk + \partial_{yx} f(x_0, y_0)kh + \partial_{yy} f(x_0, y_0)k^2$$

$$= \,\partial_{xx} f(x_0, y_0)h^2 + 2\partial_{xy} f(x_0, y_0)hk + \partial_{yy} f(x_0, y_0)k^2 \,=\, H_f(\mathbf{x}_0)\,\mathbf{h}\,\cdot \mathbf{h}\,,$$

dove

$$H_f(\mathbf{x}_0) = \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix}.$$

Esempio

Calcoliamo il differenziale secondo di $f(x,y) = x^2 e^{3y}$ nel punto $(1,0)$.

Derivate prime: $f_x(x,y) = 2xe^{3y}\,, \quad f_y(x,y) = 3x^2 e^{3y}\,;$

Derivate seconde:

$$f_{xx}(x,y) = 2e^{3y}\,, \quad f_{xy}(x,y) = f_{yx}(x,y) = 6xe^{3y}\,, \quad f_{yy}(x,y) = 9x^2 e^{3y}\,;$$

Dunque:

$$f_{xx}(1,0) = 2\,, \quad f_{xy}(1,0) = f_{yx}(1,0) = 6\,, \quad f_{yy}(1,0) = 9\,;$$

$$d^2 f(1,0) \,=\, 2h^2 + 12hk + 9k^2\,.$$

Matrice Hessiana nel punto:

$$H_f(1,0) = \begin{pmatrix} 2 & 6 \\ 6 & 9 \end{pmatrix}$$

Formula di Taylor (2° ordine)

Per una funzione due volte differenziabile, possiamo ottenere approssimazioni locali più accurate. Per semplicità assumeremo $f \in \mathcal{C}^2(D)$, $D \subseteq \mathbb{R}^n$ aperto.

Sia $\mathbf{x}_0 \in D$ e $\mathbf{h} \in \mathbb{R}^n$ tale che $\mathbf{x}_0 + \mathbf{h} \in D$. Valgono allora le formule:

Taylor, resto secondo Lagrange

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \partial_{x_i} f(\mathbf{x}_0) h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0 + \delta\, \mathbf{h}) h_i\, h_j$$

per qualche $\delta \in (0, 1)$.

Taylor, resto secondo Peano

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \partial_{x_i} f(\mathbf{x}_0) h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0) h_i\, h_j + o(|\mathbf{h}|^2)$$

Le due formule si possono scrivere in forma più compatta.

Lagrange:

$$\begin{aligned} f(\mathbf{x}_0+\mathbf{h}) &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\cdot\mathbf{h} + \frac{1}{2}H_f(\mathbf{x}_0+\delta\,\mathbf{h})\mathbf{h}\cdot\,\mathbf{h} \\ &= f(\mathbf{x}_0) + df(\mathbf{x}_0) + \frac{1}{2}d^2 f(\mathbf{x}_0+\delta\,\mathbf{h})\,, \end{aligned}$$

Peano:

$$\begin{aligned} f(\mathbf{x}_0+\mathbf{h}) &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\cdot\mathbf{h} + \frac{1}{2}H_f(\mathbf{x}_0)\mathbf{h}\cdot\,\mathbf{h} + o(|\mathbf{h}|^2) \\ &= f(\mathbf{x}_0) + df(\mathbf{x}_0) + \frac{1}{2}d^2 f(\mathbf{x}_0) + o(|\mathbf{h}|^2)\,. \end{aligned}$$

Dimostrazione.

Nel caso di una variabile, se $g(t) \in \mathcal{C}^2(I)$ e $t_0 \in I$, vale la formula

$$g(t) = g(t_0) + g'(t_0)(t - t_0) + \frac{1}{2}g''(\delta)(t - t_0)^2 \,,$$

per qualche $\delta \in (t_0, t) \subset I$.

In particolare, se $t_0 = 0$ e $t = 1$,

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(\delta) \,,$$

con $\delta \in (0, 1)$.

Definiamo ora

$$g(t) = f(\mathbf{x}_0 + t\mathbf{h}) \,;$$

osserviamo che la funzione composta è definita in $[0, 1]$ e che vale:

$$g(0) = f(\mathbf{x}_0)\,; \qquad g(1) = f(\mathbf{x}_0 + \mathbf{h})\,.$$

Inoltre g è due volte derivabile con continuità in $[0, 1]$ e applicando (due volte) la formula di derivazione delle funzioni composte si ottiene

$$g'(t) = \sum_{i=1}^{n} \partial_{x_i} f(\mathbf{x}_0 + t\,\mathbf{h}) h_i\,,$$

$$g''(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0 + t\,\mathbf{h}) h_i\, h_j$$

Ponendo $t = 0$ nella prima, $t = \delta$ nella seconda e inserendo tutte le espressioni nella formula

$$g(1) = g(0) + g'(0) + \frac{1}{2} g''(\delta)\,,$$

si ottiene Taylor con il resto di Lagrange. □

Traccia della dimostrazione della formula con il resto secondo Peano.
Si parte dalla formula con il resto di Lagrange

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} H_f(\mathbf{x}_0 + \delta\, \mathbf{h})\mathbf{h} \cdot \mathbf{h}$$

e la si scrive nella forma:

$$= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} + \frac{1}{2}\Big[H_f(\mathbf{x}_0 + \delta\, \mathbf{h}) - H_f(\mathbf{x}_0)\Big]\mathbf{h} \cdot \mathbf{h}\,.$$

Per la continuità delle derivate seconde (coefficienti della matrice Hessiana) si dimostra che

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\Big[H_f(\mathbf{x}_0 + \delta\, \mathbf{h}) - H_f(\mathbf{x}_0)\Big]\mathbf{h} \cdot \mathbf{h}}{|\mathbf{h}|^2} = 0\,,$$

cioè che l'ultimo termine è $o(|\mathbf{h}|^2)$, come richiesto.

Esempio.

Sviluppo di Taylor di $f(x,y) = x^2 e^{3y}$ nell'intorno di $(1,0)$.

Valore di f nel punto: $f(1,0) = 1$; derivate nel punto (calcolate a p.11):

$$f_x(1,0) = 2\,, \quad f_y(1,0) = 3\,, \quad f_{xx}(1,0) = 2\,, \quad f_{xy}(1,0) = 6\,, \quad f_{yy}(1,0) = 9\,.$$

$$\Downarrow$$

$$(1+h)^2 e^{3k} = 1 + 2h + 3k + h^2 + 6hk + \frac{9}{2}k^2 + o(h^2 + k^2)\,.$$

Ponendo $x = 1 + h$, $y = k$, si può anche scrivere:

$$x^2 e^{3y} = 1 + 2(x-1) + 3y + (x-1)^2 + 6(x-1)y + \frac{9}{2}y^2 + o\big((x-1)^2 + y^2\big)\,.$$

Estremi liberi

Problema.

Costruire la scatola di superficie totale minima che racchiude un volume assegnato V.

Detti x, y e z gli spigoli del parallelepipedo, si cerca il minimo della funzione $2(xy + xz + yz)$ con la condizione $xyz = V$.
Ricavando $z = V/xy$, si pone il problema nella forma:

Trovare il minimo di

$$f(x, y) = xy + \frac{V}{x} + \frac{V}{y} \quad \text{in} \quad D = \{(x, y) \,|\, x > 0\,,\ y > 0\,\}\,.$$

Problemi di estremi liberi: trovare massimi e minimi di una funzione in un insieme *aperto* (o nei punti *interni* di un dominio).

Osservazione: nella formulazione originale in *tre* variabili, la funzione era definita su un sottoinsieme chiuso di $\mathbb{R}^3$. Quando l'insieme in cui si cercano gli estremi di una funzione non è aperto, si parla di *estremi vincolati*.

Definizione (Estremi locali)

Sia $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ aperto. Un punto $\mathbf{x}_0 \in D$ si dice *punto di massimo (minimo) relativo o locale* di f se esiste $B_r(\mathbf{x}_0)$ tale che

$$\mathbf{x} \in B_r(\mathbf{x}_0) \quad \Rightarrow \quad f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \big(f(\mathbf{x}) \geq f(\mathbf{x}_0)\big)$$

Se le disuguaglianze sono strette, si parla di massimi o minimi (locali) *forti*.

Come trovare i punti di massimo e minimo locale ?

Teorema (Fermat)

Se $\mathbf{x}_0$ è un punto di massimo o minimo locale di f *e se* f è derivabile in $\mathbf{x}_0$, allora $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

Dimostrazione:

Nelle ipotesi del teorema, la funzione $g(t) := f(\mathbf{x}_0 + t\,\mathbf{e}_i)$ è definita in un intorno di $t = 0$ ed ha un estremo locale in 0 per ogni $i = 1, 2, ..., n$.

Inoltre, g è derivabile e vale: $g'(0) = \partial_{x_i} f(\mathbf{x}_0)$.

Il risultato segue allora dal teorema di Fermat in una variabile.

Osservazioni

i) La condizione $\nabla f(\mathbf{x}_0) = \mathbf{0}$ è solo *necessaria* perché il punto sia un estremo locale.

ii) Può succedere che in un punto di estremo la funzione *non* sia derivabile. Per esempio, $f(x, y) = \sqrt{x^2 + y^2}$ ha un minimo (globale) nell'origine, dove non è derivabile.

iii) Se f è differenziabile in un punto di estremo $\mathbf{x}_0$, il piano tangente è orizzontale.

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

Se un punto critico non è un punto di estremo, si dice che è un *punto di sella o colle*.

<u>Esempi</u>. Le funzioni:

$$f(x, y) = x^2 + y^2\,, \quad g(x, y) = -x^2 - y^2\,, \quad h(x, y) = x^2 - y^2\,,$$

hanno tutte $(0, 0)$ come punto critico. L'origine è: un punto di minimo per f, di massimo per g e un punto di sella per h.

Classificazione dei punti stazionari.

Se ora $f \in \mathcal{C}^2(D)$ e $\mathbf{x}_0 \in D$ è un punto critico di f, applicando la formula di Taylor abbiamo

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \frac{1}{2} H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} + o(|\mathbf{h}|^2)\,,$$

dove

$$H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} = \sum_{i,j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}_0) h_i\, h_j\,.$$

Dunque, la variazione $\Delta f := f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0)$ in un intorno di un punto critico $\mathbf{x}_0$ *dipende dalla forma quadratica* $H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h}$ (differenziale secondo).

Dobbiamo allora considerare il problema (algebrico) dello *studio del segno di una forma quadratica*

$$q(\mathbf{h}) = A\mathbf{h} \cdot \mathbf{h} = \sum_{i,j=1}^{n} a_{i\,j} h_i\, h_j$$

associata ad una matrice simmetrica A.

Una forma quadratica

$$q(\mathbf{h}) = A\mathbf{h} \cdot \mathbf{h} \ \left(= \ \mathbf{h}^T A\mathbf{h} \right) \qquad \mathbf{h} \in \mathbb{R}^n,$$

si dice

i) definita positiva (negativa) se $\forall\, \mathbf{h} \neq \mathbf{0} \ \ q(\mathbf{h}) > 0$ (< 0);

ii) semidefinita positiva (negativa) se $\forall\, \mathbf{h} \neq \mathbf{0} \ \ q(\mathbf{h}) \geq 0$ (≤ 0) ed esiste $\mathbf{h} \neq \mathbf{0}$ t.c. $q(\mathbf{h}) = 0$;

iii) indefinita se $\exists \ \mathbf{h}_1, \mathbf{h}_2$ t.c. $q(\mathbf{h}_1) > 0$ e $q(\mathbf{h}_2) < 0$.

Esempi ($n = 3$)

Se $\mathbf{h} = (h_1, h_2, h_3)$, le forme quadratiche

$$q(\mathbf{h}) = \begin{cases} h_1^2 + 3h_2^2 + 2h_3^2, \\ -2h_1^2 - h_2^2 - 4h_3^2, \\ h_1^2 + h_3^2, \\ -h_2^2 - h_3^2, \\ h_1^2 - h_2^2 + h_3^2, \end{cases}$$

sono rispettivamente: definita positiva, definita negativa, semidefinita positiva, semidefinita negativa e indefinita.

Test degli autovalori

Remind: un numero λ si dice *autovalore* di una matrice A ($n \,\mathrm{x}\, n$) se esiste un vettore non nullo $\mathbf{v}$, detto *autovettore*, tale che $A\mathbf{v} = \lambda\,\mathbf{v}$.
Questo succede se e solo se λ è soluzione di

$$\det(A - \lambda\, I) = 0 \qquad \text{(equazione caratteristica)}$$

Se A è *simmetrica*:

i) Gli autovalori λ_1, λ_2,...,λ_n, sono reali;
ii) Esistono n autovettori ortogonali che formano una base in $\mathbb{R}^n$;
iii) Esiste una matrice S che trasforma la base canonica di $\mathbb{R}^n$ nella base di autovettori e diagonalizza A:

$$S^T A S = \begin{pmatrix} \lambda_1 & 0 \ldots & 0 \\ 0 & \lambda_2 \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & 0 \ldots & \lambda_n \end{pmatrix}$$

Usando queste proprietà si dimostra che la forma $q(\mathbf{h}) = A\mathbf{h} \cdot \mathbf{h}$ è:

i) definita positiva (negativa) se e solo se tutti gli autovalori di A sono > 0 (< 0);
ii) semidefinita positiva (negativa) se e solo se tutti gli autovalori di A sono ≥ 0 (≤ 0) ed esiste almeno un autovalore nullo;
iii) indefinita se esistono autovalori di segno opposto.

Inoltre:

Se q è definita positiva, allora

$$q(\mathbf{h}) \geq \lambda_{min}|\mathbf{h}|^2 ,$$

dove $\lambda_{min} > 0$ è il minimo autovalore di A.

Se q è definita negativa, allora

$$q(\mathbf{h}) \leq \lambda_{max}|\mathbf{h}|^2 ,$$

dove $\lambda_{max} < 0$ è il massimo autovalore di A.

Esempio

Classificare la forma quadratica in $\mathbb{R}^2$: $q(h,k) = h^2 - hk + k^2$.

La matrice dei coefficienti è

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

Gli autovalori sono $\lambda_1 = 1/2$, $\lambda_2 = 3/2$. Quindi q è definita positiva.

Osserviamo che

$$q(h,k) = \frac{1}{2}(h-k)^2 + \frac{1}{2}(h^2+k^2) \geq \frac{1}{2}(h^2+k^2)\,.$$

Esercizio

Verificare che la forma quadratica in $\mathbb{R}^3$:

$$q(\mathbf{h}) = 6h_1^2 + 3h_2^2 - 8h_1h_3 + 4h_3^2\,,$$

è definita positiva.

Classificazione senza il calcolo degli autovalori (<u>solo</u> per $n = 2$).

Se

$$q(h,k) = ah^2 + 2bhk + ck^2,$$

la matrice dei coefficienti è

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

e l'equazione agli autovalori ha la forma

$$\det \begin{pmatrix} a-\lambda & b \\ b & c-\lambda \end{pmatrix} = \lambda^2 - (a+c)\lambda + ac - b^2 = 0\,.$$

Detti λ_1, λ_2 gli autovalori, abbiamo:

$$\lambda_1 + \lambda_2 = a + c \; = \mathrm{Tr}(A); \quad \lambda_1\lambda_2 = ac - b^2 \; = \det(A)\,.$$

Dunque:

$\det(A) > 0$ e $\mathrm{Tr}(A) > 0$ (< 0) $\Leftrightarrow$ q è definita positiva (negativa);

$\det(A) = 0$ e $\mathrm{Tr}(A) > 0$ (< 0) $\Leftrightarrow$ q è semidefinita positiva (negativa);

$\det(A) < 0 \Leftrightarrow$ q è indefinita.

Teorema (test delle derivate seconde).

Sia $f \in \mathcal{C}^2(D)$, $D \subseteq \mathbb{R}^n$ e sia $\mathbf{x}_0 \in D$ un punto critico di f.
Se la forma quadratica $q(\mathbf{h}) = H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h}$ è

- definita positiva (negativa) $\Rightarrow$ $\mathbf{x}_0$ punto di minimo (massimo) locale forte;
- indefinita $\Rightarrow$ $\mathbf{x}_0$ punto di sella.

Dimostrazione:

Nelle ipotesi del teorema, se $q(\mathbf{h})$ è definita positiva:

$$\Delta f := f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \frac{1}{2} H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h} + o(|\mathbf{h}|^2)$$

$$\geq \frac{\lambda_{min}}{2}|\mathbf{h}|^2 + o(|\mathbf{h}|^2) = \frac{\lambda_{min}}{2}|\mathbf{h}|^2\big(1 + o(1)\big) .$$

L'ultimo termine è > 0 per ogni $\mathbf{h}$ di norma *sufficientemente piccola.*

Dunque, in un intorno di $\mathbf{x}_0$ vale $\Delta f > 0$, cioè $\mathbf{x}_0$ è punto di minimo locale forte.

Se $q(\mathbf{h})$ è definita negativa, si dimostra in modo simile che $\Delta f < 0$ in un intorno di $\mathbf{x}_0$ e quindi che abbiamo un massimo locale forte.

Se $q(\mathbf{h})$ è indefinita, esistono due vettori $\mathbf{h}_1$, $\mathbf{h}_2$ tali che $q(\mathbf{h}_1) > 0$ e $q(\mathbf{h}_2) < 0$.
Possiamo supporre che siano di lunghezza unitaria (versori).
Allora, per t sufficientemente piccolo:

$$f(\mathbf{x}_0 + t\mathbf{h}_1) - f(\mathbf{x}_0) = \frac{1}{2}t^2 H_f(\mathbf{x}_0)\mathbf{h}_1 \cdot \mathbf{h}_1 + o(t^2)$$

$$= t^2\big(q(\mathbf{h}_1)/2 + o(1)\big) > 0\,,$$

$$f(\mathbf{x}_0 + t\mathbf{h}_2) - f(\mathbf{x}_0) = \frac{1}{2}t^2 H_f(\mathbf{x}_0)\mathbf{h}_2 \cdot \mathbf{h}_2 + o(t^2)$$

$$= t^2\big(q(\mathbf{h}_2)/2 + o(1)\big) < 0\,.$$

Dunque, lungo due direzioni uscenti da $\mathbf{x}_0$ la variazione Δf ha segno *opposto*.
Si conclude che $\mathbf{x}_0$ è un punto di sella.

Osservazione.

Se la forma quadratica $q(\mathbf{h}) = H_f(\mathbf{x}_0)\mathbf{h} \cdot \mathbf{h}$ è *semidefinita* (positiva o negativa) il test non decide e occorre un'ulteriore analisi.

Per esempio, le funzioni $f(x, y) = x^2 + y^4$ e $g(x, y) = x^2 - y^4$, hanno entrambe un unico punto critico nell'origine e per entrambe la matrice Hessiana nell'origine è

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

Quindi, *in entrambi casi* la forma quadratica associata è $q(h, k) = 2h^2$, semidefinita positiva.
Ma l'origine è punto di minimo per f e un colle per g.

La situazione cambia se la forma quadratica è semidefinita positiva (negativa) in un *intorno* di un punto critico $\mathbf{x}_0$

In questo caso, usando la formula di Taylor con il resto di Lagrange:

$$\Delta f := f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = \frac{1}{2} H_f(\mathbf{x}_0 + \delta\mathbf{h})\mathbf{h} \cdot \mathbf{h} \ \geq 0 \ \ (\leq 0)$$

per ogni $\mathbf{h}$ sufficientemente piccolo.

Questo significa che $\mathbf{x}_0$ è un punto di minimo (massimo) locale debole.

Funzioni di più variabili 4

March 22, 2021

Ricerca di massimi e minimi liberi

Come esempio di applicazione dei metodi descritti, risolviamo il problema della scatola di superficie minima.

Trovare il minimo di

$$f(x,y) = xy + \frac{V}{x} + \frac{V}{y} \quad \text{in} \quad D = \{(x,y)\,|\,x > 0\,,\ y > 0\,\}\,.$$

La funzione è regolare in tutto l'aperto, per cui possiamo cercare i punti critici e applicare il test delle derivate seconde. Calcoliamo il gradiente di f in un punto generico:

$$\nabla f(x,y) = (y - V/x^2)\,\mathbf{i} + (x - V/y^2)\,\mathbf{j}\,;$$

Cerchiamo quindi le soluzioni (in D) del *sistema*

$$\begin{cases} y - V/x^2 = 0 \\ x - V/y^2 = 0 \end{cases}$$

Dopo semplici passaggi algebrici si ottiene l'unica soluzione

$$x = y = \sqrt[3]{V}$$

Per classificare il punto critico calcoliamo:

$$\partial_{xx} f(x,y) = 2V/x^3\,; \quad \partial_{xy} f(x,y) = 1\,; \quad \partial_{yy} f(x,y) = 2V/y^3\,.$$

Quindi:

$$H_f(\sqrt[3]{V}, \sqrt[3]{V}) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\,,$$

Poiché $\det H_f(\sqrt[3]{V}, \sqrt[3]{V}) = 3 > 0$ e $\partial_{xx} f(\sqrt[3]{V}, \sqrt[3]{V}) = 2 > 0$, si tratta di un *punto di minimo*.

Si può dimostrare che il punto trovato è di minimo globale. Il *valore* minimo della funzione è $f(V^{1/3}, V^{1/3}) = 3V^{2/3}$.

Dunque, la soluzione del problema è il cubo ($x = y = z$) di lato $V^{1/3}$ e di superficie totale $S = 6V^{2/3}$.

Esercizi

Trovare i punti critici ed eventuali estremi (locali e globali) in $\mathbb{R}^2$ delle funzioni

$$f(x,y) = x^4 + y^4 - 4xy + 1\,;$$

$$g(x,y) = x^2 + y^2(x+1)^3\,.$$

Lo stesso per la funzione

$$F(x,y,z) = \frac{1}{3}x^3 + yz - x - z \qquad \text{in}\,\mathbb{R}^3$$

e per la funzione

$$h(x,y) = \frac{xy}{x^2+y^2} \qquad \text{in}\,\mathbb{R}^2\backslash\{(0,0)\}\,.$$

Massimi e minimi con vincoli

Consideriamo il problema:

trovare i massimi e i minimi della funzione $\mathrm{h}(x, y) = xy$ nell'insieme $B = \{(x, y) \,|\, x^2 + y^2 \leq 1\}$.

Poiché h è continua e B è chiuso e limitato, i punti di estremo esistono per il teorema di Weierstrass.

Dividiamo il problema in due parti:

1) ricerca degli estremi liberi nei punti interni $\mathring{B}$ (il disco aperto);
2) studio della restrizione di h alla frontiera ∂B (la circonferenza).

1) $\nabla \mathrm{h}(x, y) = y\,\mathbf{i} + x\,\mathbf{j} = \mathbf{0} \;\Rightarrow\; (0, 0)$ è l'unico punto critico in $\mathring{B}$.

Calcolando le derivate seconde e la matrice Hessiana si trova:

$h_{xx}(x,y) = h_{yy}(x,y) = 0$, $h_{xy}(x,y) = 1$, $\Rightarrow$ $\det H_h(0,0) = -1 < 0$.

Dunque l'origine è un colle e non ci sono estremi locali nel disco aperto.

2) Parametrizzando ∂B con

$$x = \cos t, \quad y = \sin t, \quad t \in [0, 2\pi],$$

abbiamo

$$\mathrm{h}(\cos t, \sin t) = \cos t \, \sin t = \frac{1}{2}\sin(2t) \qquad t \in [0, 2\pi]\,.$$

Il massimo è raggiunto per $t = \pi/4$ e $t = 5\pi/4$, cioè nei punti

$$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) \quad \text{e} \quad (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})\,.$$

Il minimo per $t = 3\pi/4$ e $t = 7\pi/4$, cioè in

$$(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) \quad \text{e} \quad (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})\,.$$

In generale, si cercano gli estremi di una funzione di due variabili su un sottoinsieme del piano *definito da un'equazione* (*vincolo di uguaglianza*)

$$\mathrm{f}(x, y) = 0$$

per qualche funzione f assegnata. Nel caso precedente, la frontiera è definita da $\mathrm{f}(x, y) := x^2 + y^2 - 1 = 0$.

Esiste un metodo per studiare il problema di ottimizzazione vincolata anche se non si 'risolve' esplicitamente l'equazione del vincolo.

Definizione (funzione implicita)

Una funzione $g : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$, si dice definita implicitamente dall'equazione $\mathrm{f}(x, y) = 0$ se

$$\mathrm{f}\big(x, g(x)\big) = 0\,, \quad \text{per ogni } x \in I\,.$$

Il risultato che segue fornisce condizioni *sufficienti* per esistenza e unicità (locali) di una funzione implicita

Teorema della funzione implicita (Dini).

Sia $\mathrm{f} \in \mathcal{C}^1(D)$, $D \subseteq \mathbb{R}^2$ aperto. Supponiamo che esista $(x_0, y_0) \in D$ tale che

$$\mathrm{f}(x_0, y_0) = 0 \quad e \quad \partial_y \mathrm{f}(x_0, y_0) \neq 0\,.$$

Allora esiste un intorno I di x_0 e un'unica funzione $g : I \to \mathbb{R}$ tale che $y_0 = g(x_0)$ e

$$\mathrm{f}\big(x, g(x)\big) = 0 \quad \forall\, x \in I\,.$$

Inoltre $g \in \mathcal{C}^1(I)$ e

$$g'(x) = -\frac{\partial_x \mathrm{f}(x, g(x))}{\partial_y \mathrm{f}(x, g(x))} \quad \forall\, x \in I\,.$$

Osservazioni:

i) La formula per g' si può ricavare derivando l'identità $\mathrm{f}(x, g(x)) = 0$ e applicando il teorema di derivazione delle funzioni composte;

ii) dalla formula si ottiene l'equazione della retta tangente al grafico di g in (x_0, y_0):

$$y = y_0 - \frac{\partial_x \mathrm{f}(x_0, y_0)}{\partial_y \mathrm{f}(x_0, y_0)}(x - x_0)\,,$$

iii) se l'ultima ipotesi del teorema si cambia con $\partial_x \mathrm{f}(x_0, y_0) \neq 0$, la tesi diventa:

esiste un intorno J di y_0 e un'unica funzione $h : J \to \mathbb{R}$ tale che $x_0 = h(y_0)$ e

$$\mathrm{f}\big(h(y), y\big) = 0 \quad \forall\, y \in J\,.$$

Inoltre $h \in \mathcal{C}^1(J)$ e

$$h'(y) = -\frac{\partial_y \mathrm{f}(h(y), y)}{\partial_x \mathrm{f}(h(y), y)} \quad \forall\, y \in J\,.$$

Esempio

L'equazione

$$\mathrm{f}(x, y) := e^{xy} + x - y + 1 = 0$$

è verificata nel punto $(0, 2)$. Inoltre

$$\partial_y \mathrm{f}\,(0, 2) = (xe^{xy} - 1)\big|_{(0,2)} = -1 \neq 0\,.$$

Esiste allora $g(x)$, definita (e regolare) in un intorno I di $x = 0$, tale che

$$g(0) = 2 \qquad \text{e} \qquad \mathrm{f}(x, g(x)) = 0\,,\;\; \forall x \in I\,.$$

Nell'origine abbiamo: $g'(0) = -\frac{\partial_x \mathrm{f}\,(0,2)}{\partial_y \mathrm{f}\,(0,2)} = -\frac{3}{-1} = 3$,

per cui in I vale

$$g(x) = 2 + 3x + o(x)$$

Le ipotesi del teorema di Dini sono verificate anche nel punto $(-2,0)$:

$$\mathrm{f}(-2,0)=0\,;\quad \partial_y\mathrm{f}(-2,0)=-3\neq 0\,.$$

In questo caso :

$$g(-2)=0,\ \ g'(-2)\ =\ -\frac{\partial_x\mathrm{f}(-2,0)}{\partial_y\mathrm{f}(-2,0)}\ =\ -\frac{1}{-3}\ =\ 1/3$$

e quindi

$$g(x)=\frac{1}{3}(x+2)+o(x+2)$$

in un intorno di $x=-2$.

Se in (x_0, y_0) vale $\mathrm{f}(x_0, y_0) = 0$ e $\nabla\mathrm{f}(x_0, y_0) \neq \mathbf{0}$, il teorema della funzione implicita garantisce che il vincolo $\mathrm{f}(x, y) = 0$ è *localmente* il grafico di una funzione di una variabile. Chiameremo *regolari* questi punti.

I punti (x_0, y_0) tali che

$$\mathrm{f}(x_0, y_0) = 0 \quad \text{e} \quad \nabla\mathrm{f}(x_0, y_0) = \mathbf{0},$$

si dicono *punti singolari* del vincolo (o dell'insieme di livello $\mathrm{f} = 0$).

Esempi

L' origine è un punto singolare dei vincoli

$$x^2 + y^2 = 0\,, \qquad x^2 - y^2 = 0\,, \qquad x^2 - y^3 = 0\,,$$

poiché in tutti e tre i casi $\mathrm{f}(0, 0) = 0$ e $\nabla\mathrm{f}(0, 0) = \mathbf{0}$.

Dagli esempi si vede che nell'intorno di un punto singolare *non sono garantite* esistenza, unicità e regolarità di una curva che rappresenti il vincolo.

Osservazione. Se (x_0, y_0) è un *punto critico* di una funzione di $f(x, y)$, esso è anche *punto singolare dell'insieme di livello* $\{(x, y) \mid f(x, y) = c\}$, dove $c = f(x_0, y_0)$.

Funzione implicita in più variabili

Consideriamo il caso di funzioni definite implicitamente dall'equazione $\mathrm{f}(x, y, z) = 0$.

Teorema (Dini).

Sia $\mathrm{f} \in \mathcal{C}^1(D)$, $D \subseteq \mathbb{R}^3$ aperto. Supponiamo che esista $(x_0, y_0, z_0) \in D$ tale che

$$\mathrm{f}(x_0, y_0, z_0) = 0 \quad e \quad \partial_z \mathrm{f}(x_0, y_0, z_0) \neq 0\,.$$

Allora esiste un'unica funzione $g(x, y)$ definita in un intorno U di (x_0, y_0) e tale che $z_0 = g(x_0, y_0)$ e

$$\mathrm{f}\big(x, y, g(x, y)\big) = 0 \qquad \forall\,(x, y) \in U\,.$$

Inoltre $g \in \mathcal{C}^1(U)$ e

$$\partial_x g(x, y) = -\frac{\partial_x \mathrm{f}(x, y, g(x, y))}{\partial_z \mathrm{f}(x, y, g(x, y))}\,, \qquad \partial_y g(x, y) = -\frac{\partial_y \mathrm{f}(x, y, g(x, y))}{\partial_z \mathrm{f}(x, y, g(x, y))}\,,$$

$\forall\,(x, y) \in U$.

Osservazioni:

i) Dalle formule per le derivate $\partial_x g$, $\partial_y g$, si ricava l'equazione del *piano tangente* alla superficie $z = g(x, y)$ in (x_0, y_0, z_0):

$$z = z_0 - \frac{\partial_x \mathrm{f}(x_0, y_0, z_0)}{\partial_z \mathrm{f}(x_0, y_0, z_0)}(x - x_0) - \frac{\partial_y \mathrm{f}(x_0, y_0, z_0)}{\partial_z \mathrm{f}(x_0, y_0, z_0)}(y - y_0)\,,$$

ii) se l'ultima ipotesi del teorema si cambia con $\partial_x \mathrm{f}(x_0, y_0, z_0) \neq 0$ (o con $\partial_y \mathrm{f}(x_0, y_0, z_0) \neq 0$), si dimostra l'esistenza di una funzione $h(y, z)$ $(\,k(x, z)\,)$ definita e regolare in un intorno di (y_0, z_0) (di (x_0, z_0)) e che soddisfa

$$\mathrm{f}\big(h(y, z), y, z\big) = 0 \qquad \big(\,\mathrm{f}\big(x, k(x, z), z\big) = 0\,\big)$$

<u>Esercizio</u>.

Usando l'equazione del piano tangente alla funzione implicita dimostrare che il gradiente ∇f di una funzione di tre variabili (di classe $\mathcal{C}^1$) è *ortogonale* alla *superficie di livello* $\{(x, y, z) \,|\, f(x, y, z) = c\}$ in ogni punto regolare.

Vincoli di uguaglianza: moltiplicatori di Lagrange.

Torniamo ora al problema di determinare gli estremi di una funzione $f(x, y)$ sotto una condizione di vincolo del tipo $g(x, y) = c$.

Teorema (moltiplicatori di Lagrange)

Siano f, $g \in \mathcal{C}^1(\mathbb{R}^2)$ e sia $(x_0.y_0)$ un punto di estremo vincolato per f con il vincolo $g(x, y) = c$ ($c \in \mathbb{R}$).

Se (x_0, y_0) è un punto regolare del vincolo ($\nabla g(x_0, y_0) \neq \mathbf{0}$) esiste $\lambda_0 \in \mathbb{R}$ tale che

$$\nabla f(x_0, y_0) = \lambda_0 \nabla g(x_0, y_0) \,.$$

Dimostrazione

Poichè (x_0, y_0) è un punto regolare, per il teorema di Dini l'equazione $g(x, y) - c = 0$ definisce localmente un arco di curva regolare

$$x = x(t), \quad y = y(t), \qquad t \in I,$$

dove si può assumere I un intorno dell'origine e $(x(0), y(0)) = (x_0, y_0)$.

Per ipotesi, $t = 0$ è punto di estremo locale per la funzione composta $f(x(t), y(t))$, per cui

$$0 = \frac{d}{dt} f(x(t), y(t))\Big|_{t=0} = \partial_x f(x_0, y_0) x'(0) + \partial_y f(x_0, y_0) y'(0) \,.$$

Quindi, nel punto considerato, il gradiente $\nabla f(x_0, y_0)$ è *ortogonale* al vettore tangente $x'(0)\,\mathbf{i} + y'(0)\,\mathbf{j}$ alla curva di livello $g(x, y) = c$.

Ma, come è noto, anche il vettore ∇g è ortogonale (in ogni punto regolare) alla curva di livello.

Si conclude che i vettori $\nabla f(x_0, y_0)$ e $\nabla g(x_0, y_0)$ *devono essere paralleli*, ovvero che

$$\nabla f(x_0, y_0) = \lambda_0 \nabla g(x_0, y_0)$$

per qualche $\lambda_0 \in \mathbb{R}$. $\square$

Un punto regolare del vincolo in cui vale la condizione $\nabla f(x, y) = \lambda \nabla g(x, y)$ si dice **punto critico vincolato**.

Per il teorema precedente, se un punto (regolare) è di estremo vincolato, allora è un punto critico vincolato.

Osservazioni.

Introducendo la funzione detta *Lagrangiana*

$$\mathcal{L}(x, y, \lambda) := f(x, y) - \lambda \big(g(x, y) - c\big)$$

possiamo dire che i punti critici vincolati si trovano risolvendo il sistema

$$\begin{cases} \partial_x \mathcal{L}(x, y, \lambda) = \partial_x f(x, y) - \lambda\, \partial_x g(x, y) = 0, \\ \partial_y \mathcal{L}(x, y, \lambda) = \partial_y f(x, y) - \lambda\, \partial_y g(x, y) = 0, \\ \partial_\lambda \mathcal{L}(x, y, \lambda) = c - g(x, y) = 0. \end{cases}$$

Il sistema equivale alla ricerca dei punti critici (liberi) della Lagrangiana come funzione delle *tre* variabili (x, y, λ).

La ricerca degli estremi vincolati con il *metodo dei moltiplicatori di Lagrange* consiste nei seguenti passi:

i) Si identificano eventuali punti singolari del vincolo e si valuta la funzione in tali punti.

ii) Si cercano i punti critici vincolati risolvendo il sistema per i punti critici della Lagrangiana.

iii) Si valuta la funzione nei punti critici trovati.

Esempio

Ritorniamo al problema di determinare gli estremi di

$$f(x, y) = xy$$

con il vincolo

$$x^2 + y^2 = 1\,,$$

che non presenta punti singolari.

In questo caso la *Lagrangiana* si scrive

$$\mathcal{L}(x, y, \lambda) := xy - \lambda\left(x^2 + y^2 - 1\right).$$

I punti critici vincolati si trovano risolvendo il sistema

$$\begin{cases} \partial_x \mathcal{L}(x, y, \lambda) = y - 2\lambda\, x = 0, \\ \partial_y \mathcal{L}(x, y, \lambda) = x - 2\lambda\, y = 0, \\ \partial_\lambda \mathcal{L}(x, y, \lambda) = -(x^2 + y^2 - 1) = 0. \end{cases}$$

Eliminando λ nelle prime due equazioni si ricava $x = y$ o $x = -y$; sostituendo nella terza si trovano le stesse soluzioni di p. 6.

Osservare che le curve di livello

$$xy = 1/2 \qquad \text{e} \qquad xy = -1/2\,,$$

sono *tangenti al vincolo* $x^2 + y^2 = 1$ nei punti di estremo vincolato, in accordo con il teorema dei moltiplicatori.

Problemi.

Trovare il rettangolo di perimetro massimo inscritto nell'ellisse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\,.$$

Trovare il minimo di

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

con il vincolo

$$x^3 - y^2 = 0\,.$$

Funzioni a valori vettoriali.

Una funzione di più variabili a valori vettoriali

$$\mathbf{f} : D \subseteq \mathbb{R}^n \to \mathbb{R}^m, \qquad (m > 1) \qquad \mathbf{x} \mapsto \mathbf{f}(\mathbf{x}).$$

è definita da m funzioni a valori reali $f_i(x_1, x_2, ..., x_n)$ $(i = 1, 2, ..., m)$ sul comune dominio D.

Limiti e continuità

Sia $\mathbf{x}_0 \in \mathbb{R}^n$ un punto di accumulazione per D, e sia $\mathbf{L} = (L_1, L_2, ..., L_m) \in \mathbb{R}^m$. Si dice che

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{L} \qquad \text{se} \qquad \lim_{\mathbf{x} \to \mathbf{x}_0} \left| \mathbf{f}(\mathbf{x}) - \mathbf{L} \right| = 0.$$

Segue dalla definizione che **f** ha limite **L** per $\mathbf{x} \to \mathbf{x}_0$ se e solo se $\lim_{\mathbf{x} \to \mathbf{x}_0} f_i(\mathbf{x}) = L_i$, per ogni $i = 1, 2, ..., m$.

Si dice che **f** è **continua** in $\mathbf{x}_0 \in D$, se

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

Differenziabilità

Si dice che $\mathbf{f}$ è differenziabile in un punto $\mathbf{x}_0 \in D$ se tutte le componenti f_i, $i = 1, 2, ..., m$ sono differenziabili in quel punto.

Avremo allora, per ogni $\mathbf{h} \in \mathbb{R}^n$ t.c. $\mathbf{x}_0 + \mathbf{h} \in D$:

$$f_i(\mathbf{x}_0 + \mathbf{h}) - f_i(\mathbf{x}_0) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) h_j + o(|\mathbf{h}|)\,, \qquad i = 1, 2, ..., m\,.$$

Conviene scrivere queste relazioni in forma vettoriale introducendo i *vettori colonna*

$$\mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \\ . \\ . \\ f_m \end{pmatrix} \qquad \mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \\ . \\ . \\ h_n \end{pmatrix}$$

e la matrice *mxn*, detta **matrice Jacobiana**, che ha per elementi di posto *ij* le derivate parziali $\dfrac{\partial f_i}{\partial x_j}$:

$$\mathbf{D}\,\mathbf{f}(\mathbf{x}) := \begin{pmatrix} \partial_{x_1} f_1(\mathbf{x}) & ... & \partial_{x_n} f_1(\mathbf{x}) \\ \partial_{x_1} f_2(\mathbf{x}) & ... & \partial_{x_n} f_2(\mathbf{x}) \\ . & ... & . \\ . & ... & . \\ \partial_{x_1} f_m(\mathbf{x}) & \dots & \partial_{x_n} f_m(\mathbf{x}) \end{pmatrix}$$

Per la matrice Jacobiana si usa spesso la notazione $J_{\mathbf{f}}(\mathbf{x})$.

La relazione di differenziabilità in un punto si scrive ora in forma vettoriale:

$$\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) = \mathbf{D}\,\mathbf{f}(\mathbf{x}_0)\,\mathbf{h} + \mathbf{o}(|\mathbf{h}|)\,,$$

dove $\mathbf{o}(|\mathbf{h}|)$ è un *vettore* le cui m componenti sono infinitesime di ordine superiore a $|\mathbf{h}|$ per $\mathbf{h} \to \mathbf{0}$.

Il *differenziale* di $\mathbf{f}$ in $\mathbf{x}_0$ è la funzione lineare:

$$d\mathbf{f}(\mathbf{x}_0)\ :\ \mathbb{R}^n \to \mathbb{R}^m\,, \qquad \mathbf{h} \mapsto \mathbf{D}\,\mathbf{f}(\mathbf{x}_0)\,\mathbf{h}\,.$$

Osservazioni.

Nel caso particolare di $f : \mathbb{R}^n \to \mathbb{R}$ (cioè per $m = 1$) la matrice Jacobiana in un punto $\mathbf{x} = (x_1, x_2, ..., x_n)$ diviene

$$\mathbf{Df}(\mathbf{x}) = \left(\partial_{x_1} f(\mathbf{x}), \partial_{x_2} f(\mathbf{x})..., \partial_{x_n} f(\mathbf{x})\right)$$

ed è quindi il gradiente $\nabla f(\mathbf{x})$ scritto come vettore riga.

Nel caso particolare di $\mathbf{f} : \mathbb{R} \to \mathbb{R}^m$ (cioè per $n = 1$) la matrice Jacobiana in un punto $t \in \mathbb{R}$ diviene

$$\mathbf{Df}(t) = \begin{pmatrix} f_1'(t) \\ f_2'(t) \\ . \\ . \\ f_m'(t) \end{pmatrix}$$

ed è quindi il vettore tangente $\mathbf{f}'(t)$ scritto come vettore colonna.

Osservazione sulle notazioni

In casi di particolari applicazioni (superfici parametriche, trasformazioni di coordinate, campi vettoriali) si utilizzano notazioni diverse per rappresentare le funzioni a valori vettoriali.

Per esempio, una trasformazione di coordinate $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^n$ è usualmente rappresentata con le equazioni

$$\begin{cases} x_1 = & x_1(u_1, u_2, ..., u_n) \\ x_2 = & x_2(u_1, u_2, ..., u_n) \\ ... & ... \\ ... & .... \\ x_n = & x_n(u_1, u_2, ..., u_n) \end{cases}$$

Qui il ruolo delle variabili indipendenti è giocato dalle $(u_1, u_2, ..., u_n)$, mentre le $(x_1, x_2, ..., x_n)$ sono le componenti del vettore $\mathbf{T}(u_1, u_2, ..., u_n)$.

Il determinante della matrice Jacobiana della trasformazione si scrive anche

$$det\, J_{\mathbf{T}}(u_1, u_2, ..., u_n) \equiv \frac{\partial(x_1, x_2, ..., x_n)}{\partial(u_1, u_2, ..., u_n)}$$

Esempi

La matrice Jacobiana di una funzione *lineare* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \mapsto \mathbf{f}(\mathbf{x}) := M\mathbf{x}$, è (in ogni punto) la stessa matrice M :

$$J_{\mathbf{f}}(\mathbf{x}) = M\,.$$

Calcoliamo la matrice Jacobiana della *trasformazione in coordinate polari*:

$$\mathbf{T} \ : \quad \begin{cases} x = \rho\cos\theta, \\ y = \rho\sin\theta, \end{cases} \qquad\qquad (\rho,\theta) \in [0,+\infty) \times [0,2\pi)\,.$$

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

Osserviamo che

$$\det J_{\mathbf{T}}(\rho,\theta) = \rho\,,$$

per cui la matrice Jacobiana della trasformazione diventa singolare nell'origine.

Condizione sufficiente di differenziabilità

Se tutte le componenti f_i di una funzione a valori vettoriali $\mathbf{f}$ hanno derivate continue in un punto $\mathbf{x}_0$, allora $\mathbf{f}$ è differenziabile in $\mathbf{x}_0$.

In altri termini, se tutti gli elementi della matrice Jacobiana sono funzioni continue in $\mathbf{x}_0$, allora $\mathbf{f}$ è differenziabile in $\mathbf{x}_0$.

Anche il *teorema di derivazione delle funzioni composte* si può convenientemente enunciare con le matrici jacobiane:

Siano $\mathbf{f} =: \mathbb{R}^n \to \mathbb{R}^m$ e $\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^k$ tali che:

$\mathbf{f}$ è differenziabile in $\mathbf{x}_0 \in \mathbb{R}^n$ e $\mathbf{g}$ è differenziabile in $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0) \in \mathbb{R}^m$.

Allora la funzione composta $\mathbf{g} \circ \mathbf{f}$ è differenziabile in $\mathbf{x}_0$ e vale

$$\mathbf{D}(\mathbf{g} \circ \mathbf{f})(\mathbf{x}_0) = \mathbf{Dg}(\mathbf{y}_0)\, \mathbf{Df}(\mathbf{x}_0)\,.$$

Osservare che (in ogni punto) la matrice $J_{\mathbf{g}}$ è di tipo $k \times m$, mentre $J_{\mathbf{f}}$ è di tipo $m \times n$ per cui il prodotto (righe per colonne) è ben definito ed è una matrice $k \times n$.

Ponendo $\mathbf{y} = \mathbf{f}(\mathbf{x})$, $\mathbf{z} = \mathbf{g}(\mathbf{y}) = \mathbf{g} \circ \mathbf{f}\,(\mathbf{x})$, il risultato equivale alle relazioni

$$\frac{\partial z_i}{\partial x_j} = \sum_{l=1}^{m} \frac{\partial z_i}{\partial y_l} \frac{\partial y_l}{\partial x_j}, \qquad \forall \;\; i = 1, 2, .., k\,, \; j = 1, 2, .., n\,.$$

Esempio

Data $g(x, y)$ differenziabile e la trasformazione in coordinate polari

$$x = \rho \cos\theta\,, \qquad y = \rho \sin\theta\,,$$

calcoliamo le derivate della funzione composta

$$\tilde{g}(\rho, \theta) := g(\rho\cos\theta, \rho\sin\theta)\,.$$

Osserviamo che in questo caso : $\quad n = m = 2,\;\; k = 1,$

$$(x_1, x_2) \;\Rightarrow\; (\rho, \theta)\,; \qquad (y_1, y_2) \;\Rightarrow\; (x, y)\,,$$

la funzione $\mathbf{f}$ è la trasformazione di coordinate $\mathbf{T}$ di p. 26 e $g \circ \mathbf{T} = \tilde{g}$.

Abbiamo allora :

$$\partial_\rho \tilde{g}(\rho,\theta) = \partial_\rho g(\rho\cos\theta, \rho\sin\theta) = \partial_x g(x,y)\cos\theta + \partial_y g(x,y)\sin\theta\,,$$

$$\partial_\theta \tilde{g}(\rho,\theta) = \partial_\theta g(\rho\cos\theta, \rho\sin\theta) = \partial_x g(x,y)(-\rho\sin\theta) + \partial_y g(x,y)\,\rho\cos\theta\,.$$

Osserviamo che le due relazioni si scrivono in forma vettoriale:

$$\left(\partial_\rho \tilde{g},\, \partial_\theta \tilde{g}\right) = \left(\partial_x g,\, \partial_y g\right) \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$$

ovvero

$$J_{\tilde{g}}(\rho,\theta) = J_g(x,y)\, J_{\mathbf{T}}(\rho,\theta)\,,$$

in accordo con il teorema a pag. 27.

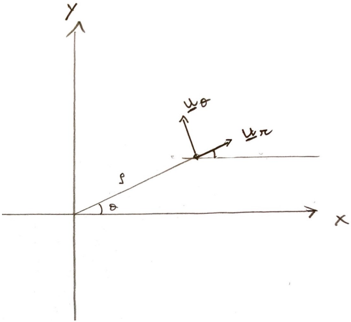
Osservazioni

Detto $\mathbf{u}_r = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}$ il *versore della direzione radiale*, la prima relazione della pagina precedente si può scrivere

$$\partial_\rho \tilde{g}(\rho,\theta) = D_{\mathbf{u}_r} g(x,y)\,,$$

dove (ρ,θ) e (x,y) sono sempre i punti che si corrispondono nella trasformazione.

In modo analogo, la seconda relazione ci dice che $\dfrac{1}{\rho}\partial_\theta \tilde{g}$ è la derivata di g nella direzione del versore $\mathbf{u}_\theta = -\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j}$.



Superfici parametriche.

Sia $D \subseteq \mathbb{R}^2$ con interno non vuoto ($\mathring{D} \neq \emptyset$). Si dice *superficie in forma parametrica* una funzione

$$\mathbf{r}: \ D \to \mathbb{R}^3 \,, \qquad (u,v) \mapsto \mathbf{r}(u,v) = x(u,v)\,\mathbf{i} + y(u,v)\,\mathbf{j} + z(u,v)\,\mathbf{k} \,.$$

Se $\mathbf{r} \in \mathcal{C}^1(\mathring{D})$, la superficie si dice di classe $\mathcal{C}^1$.

Per queste superfici sono definiti (per ogni $(u,v) \in \mathring{D}$) i vettori

$$\mathbf{r}_u(u,v) := \partial_u \mathbf{r}(u,v) \,, \qquad \mathbf{r}_v(u,v) := \partial_v \mathbf{r}(u,v) \,,$$

'tangenti alle linee coordinate' sulla superficie.

Una superficie di classe $\mathcal{C}^1$ si dice **regolare** se

$$\mathbf{r}_u(u,v) \wedge \mathbf{r}_v(u,v) \neq \mathbf{0}$$

per ogni $(u,v) \in \mathring{D}$.

Si può anche dire che una superficie è regolare se la matrice Jacobiana

$$J_{\mathbf{r}}(u,v) = \begin{pmatrix} x_u(u,v) & x_v(u,v) \\ y_u(u,v) & y_v(u,v) \\ z_u(u,v) & z_v(u,v) \end{pmatrix}$$

ha rango 2 (massimo) per ogni $(u,v) \in \mathring{D}$.

In ogni punto $\mathbf{r}(u,v)$ $\big((u,v) \in \mathring{D}\big)$ di una superficie regolare è definito un *versore normale*

$$\mathbf{n}(u,v) = \frac{\mathbf{r}_u(u,v) \wedge \mathbf{r}_v(u,v)}{|\mathbf{r}_u(u,v) \wedge \mathbf{r}_v(u,v)|}\,,$$

perpendicolare al piano tangente che contiene i vettori $\mathbf{r}_u(u,v)$ e $\mathbf{r}_v(u,v)$.

Esempi

La funzione

$$\mathbf{r}(u,v) = R\cos u\,\mathbf{i} + R\sin u\,\mathbf{j} + v\,\mathbf{k}\,, \qquad u \in [0,2\pi)\,,\ v \in \mathbb{R}\,,$$

rappresenta la *superficie cilindrica* di equazione $x^2 + y^2 = R^2$ nello spazio.

La superficie è regolare perché

$$\mathbf{r}_u(u,v) = -R\sin u\,\mathbf{i} + R\cos u\,\mathbf{j}\,, \quad \mathbf{r}_v(u,v) = \mathbf{k}\,,$$

$$\mathbf{r}_u(u,v) \wedge \mathbf{r}_v(u,v) = R\cos u\,\mathbf{i} + R\sin u\,\mathbf{j} \neq \mathbf{0}\,.$$

Il *versore*

$$\mathbf{n}(u,v) = \cos u\,\mathbf{i} + \sin u\,\mathbf{j} \neq \mathbf{0}\,,$$

è in ogni punto perpendicolare alla superficie e ha verso 'uscente' dal cilindro.

La funzione

$$\mathbf{r}(\phi,\theta) = R\sin\phi\cos\theta\,\mathbf{i} + R\sin\phi\sin\theta\,\mathbf{j} + R\cos\phi\,\mathbf{k}\,, \quad \phi\in[0,\pi]\,,\ \theta\in[0,2\pi)\,,$$

parametrizza la *superficie sferica* S_R di equazione $x^2+y^2+z^2=R^2$.

Abbiamo ora:

$$\mathbf{r}_\phi(\phi,\theta) = R\cos\phi\cos\theta\,\mathbf{i} + R\cos\phi\sin\theta\,\mathbf{j} - R\sin\phi\,\mathbf{k}\,,$$

$$\mathbf{r}_\theta(\phi,\theta) = -R\sin\phi\sin\theta\,\mathbf{i} + R\sin\phi\cos\theta\,\mathbf{j}\,.$$

Verificare che $\mathbf{r}_\phi\wedge\mathbf{r}_\theta$ è perpendicolare alla superficie, punta verso l'esterno della sfera e ha lunghezza

$$|\mathbf{r}_\phi(\phi,\theta)\wedge\mathbf{r}_\theta(\phi,\theta)| = R^2\sin\phi\,.$$

Superfici cartesiane

Il grafico di una funzione $f(x, y)$ si può rappresentare in forma parametrica:

$$\mathbf{r}(u, v) = u\,\mathbf{i} + v\,\mathbf{j} + f(u, v)\,\mathbf{k}\,, \qquad (u, v) \in D\,.$$

Se $f \in \mathcal{C}^1(D)$, la superficie è regolare:

$$\mathbf{r}_u(u, v) \wedge \mathbf{r}_v(u, v) = \big(\mathbf{i} + f_x(u, v)\,\mathbf{k}\big) \wedge \big(\mathbf{j} + f_y(u, v)\,\mathbf{k}\big) = -f_x\,\mathbf{i} - f_y\,\mathbf{j} + \mathbf{k}\,.$$

$$|\mathbf{r}_u(u, v) \wedge \mathbf{r}_v(u, v)| = \sqrt{f_x(u, v)^2 + f_y(u, v)^2 + 1} = \sqrt{1 + |\nabla f|^2}$$

Il versore normale (che punta verso l'alto) si scrive

$$\mathbf{n} = \frac{-f_x}{\sqrt{1 + |\nabla f|^2}}\,\mathbf{i} + \frac{-f_y}{\sqrt{1 + |\nabla f|^2}}\,\mathbf{j} + \frac{1}{\sqrt{1 + |\nabla f|^2}}\,\mathbf{k}\,.$$

Invertibilità delle funzioni da $\mathbb{R}^n$ in $\mathbb{R}^n$

Richiesta necessaria per le trasformazioni di coordinate, ma che in generale non è soddisfatta 'globalmente'.

Vale per le trasformazioni *lineari*

$$\mathbf{x} \mapsto A\,\mathbf{x}$$

se e solo se A è una matrice $n \times n$ non singolare, cioè se

$$\det A \neq 0\,.$$

Nel caso generale abbiamo:

Teorema (Inversione locale)

Sia $\mathbf{f} : D \to \mathbb{R}^n$, con $D \subseteq \mathbb{R}^n$ aperto e $f \in \mathcal{C}^1(D)$. Supponiamo che in $\mathbf{x}_0 \in D$ valga

$$\det \mathbf{Df}(\mathbf{x}_0) \neq 0\,.$$

Allora esistono un intorno U di $\mathbf{x}_0$ e un intorno V di $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ tali che $\mathbf{f}$ è una corrispondenza *biunivoca* tra U e V.

Inoltre, detta $\mathbf{g} : V \to U$ la funzione inversa di $\mathbf{f}|_U$, risulta $\mathbf{g} \in \mathcal{C}^1(V)$ e vale la formula

$$\mathbf{Dg}(\mathbf{y}) = \mathbf{Df}(\mathbf{x})^{-1}\,,$$

per ogni $\mathbf{y} \in V$ e $\mathbf{x} = \mathbf{g}(\mathbf{y}) \in U$.

Esempi

Per la trasformazione $\mathbf{T}$ in coordinate polari si ha $\det D_{\mathbf{T}}(\rho, \theta) = \rho$ (v. pag.26);

segue dal teorema che la trasformazione è invertibile nell'intorno di ogni punto con $\rho > 0$, cioè al di fuori dell'origine nel piano (x, y).

In effetti, dalle formule della trasformazione segue che $\mathbf{T}(0, \theta) = (0, 0)$ per ogni θ, per cui $\mathbf{T}$ non è localmente biunivoca intorno ai punti dell'asse $\rho = 0$.

Problema

Verificare che $\mathbf{T}$ *ristretta all'aperto*

$$A = (0, +\infty) \times (0, 2\pi)\,,$$

è biunivoca tra A ed il piano $\mathbb{R}^2$ privato della semiretta $\{(x, 0)\,,\ x \geq 0\}$ e che $\mathbf{T}^{-1}$ è di classe $\mathcal{C}^1$ (diffeomorfismo globale).

Anche per le *coordinate cilindriche* in $\mathbb{R}^3$:

$$\begin{cases} x = \rho\cos\theta\,, \\ y = \rho\sin\theta\,, \\ z = t\,, \end{cases} \qquad \rho \in [0,+\infty),\ \theta \in [0,2\pi),\ t \in \mathbb{R}\,,$$

risulta $\det J_{\mathbf{T}}(\rho,\theta,t) = \rho$ (fare la verifica); in questo caso, i punti singolari della trasformazione sono tutti quelli dell'asse z.

Infine, per le *coordinate sferiche* in $\mathbb{R}^3$

$$\begin{cases} x = \rho\sin\phi\cos\theta\,, \\ y = \rho\sin\phi\sin\theta\,, \\ z = \rho\cos\phi\,, \end{cases} \qquad \rho \in [0,+\infty),\ \ \phi \in [0,\pi],\ \ \theta \in [0,2\pi),$$

risulta

$$\det J_{\mathbf{T}}(\rho,\theta,t) = \rho^2\sin\phi\,.$$

Anche in questo caso i punti dell'asse z sono singolari.

Attenzione: Anche se una trasformazione soddisfa in ogni punto le ipotesi del teorema di inversione locale, *non è garantito* che sia *globalmente* invertibile.

Per esempio

$$\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2 \qquad \mathbf{f}(x, y) = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}$$

ha matrice Jacobiana

$$J_{\mathbf{f}}(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}$$

Abbiamo $\det J_{\mathbf{f}}(x, y) = e^{2x} \neq 0$ in ogni punto (x, y); quindi **f** è localmente invertibile in un intorno di ogni punto del piano.

Tuttavia, **f** non è globalmente invertibile poiché

$$\mathbf{f}(x, y + 2n\pi) = \mathbf{f}(x, y)$$

per ogni $n \in \mathbb{Z}$.

Esercizio. Studiare l'invertibilità locale della trasformazione $\mathbf{f}(x, y) = \begin{pmatrix} x \\ xy \end{pmatrix}$.