

8. FUNZIONI DIFFERENZIABILI

Ricordiamo che una funzione scalare di una variabile $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, $y = f(x)$, definita in un intervallo aperto I , è derivabile in $x_0 \in A$ se esiste finito

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \stackrel{\text{def.}}{=} f'(x_0), \quad (59)$$

cioè il grafico della funzione è approssimato bene dalla retta tangente di equazione

$$z = f(x_0) + f'(x_0)(x - x_0)$$

che passa per il punto $(x_0, f(x_0))$ ed il cui coefficiente angolare è $f'(x_0)$. Più precisamente, per ogni $x \in I$

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + |x - x_0| \epsilon(x - x_0), \quad (60)$$

dove $\epsilon(x - x_0)$ è una funzione infinitesima.

Estendiamo la definizione al caso di funzioni di più variabili. Per semplicità trattiamo il caso di funzioni scalari di due variabili.

Def. 2.26. Una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ definita su un aperto A è derivabile nel punto $(x_0, y_0) \in A$ se

- a) la funzione di una variabile $f(x, y_0)$ è derivabile rispetto ad x nel punto x_0 ;
- b) la funzione di una variabile $f(x_0, y)$ è derivabile rispetto ad y nel punto y_0 .

In tal caso, le derivate parziali rispetto ad x ed y sono definite da

$$\frac{\partial f}{\partial x}(x_0, y_0) = \left. \frac{df(x, y_0)}{dx} \right|_{x=x_0} \quad (61a)$$

$$\frac{\partial f}{\partial y}(x_0, y_0) = \left. \frac{df(x_0, y)}{dy} \right|_{y=y_0}. \quad (61b)$$

ed il vettore

$$\nabla f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right)$$

è detto gradiente di f nel punto (x_0, y_0) . La funzione f è detta derivabile, se è derivabile in ogni punto di A .

Se f è una funzione è derivabile, le derivate parziali definiscono due funzioni

$$\frac{\partial f}{\partial x}(x, y) \quad \text{e} \quad \frac{\partial f}{\partial y}(x, y) \quad (x, y) \in A$$

che si calcolano usando le usuali regole di derivazione per le funzioni di una variabile in cui la variabile che non si derivata è considerata come una costante.

Esempio 2.27. La funzione

$$f(x, y) = \sin\left(\frac{x}{y}\right)$$

definita sull'insieme aperto

$$A = \{(x, y) \in \mathbb{R}^2 \mid y \neq 0\}$$

è derivabile con derivate

$$\frac{\partial f}{\partial x}(x, y) = \cos\left(\frac{x}{y}\right) \frac{1}{y} \quad \frac{\partial f}{\partial y}(x, y) = -\cos\left(\frac{x}{y}\right) \frac{x}{y^2}.$$

Le derivate parziali hanno un'interessante interpretazione geometrica. Poiché A è aperto, per ogni $P_0 = (x_0, y_0) \in A$ esiste $\delta > 0$ tale che $B(P_0, \delta) \subset A$ per cui la restrizione della funzione $f(x, y)$ alla retta $y = y_0$

$$f_{y_0}(x) = f(x, y_0)$$

è ben definita per ogni $x \in (x_0 - \delta, x_0 + \delta)$. Il grafico della restrizione

$$z = f_{y_0}(x) = f(x, y_0)$$

è l'intersezione il grafico di f con il piano $y = y_0$. La condizione a) della definizione equivale al fatto che $f_{y_0}(x)$ sia derivabile in x_0 e, dalla (59), segue che

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{x \rightarrow x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} = f'_{y_0}(x_0).$$

La derivata parziale $\frac{\partial f}{\partial x}(x_0, y_0)$ è quindi il coefficiente angolare della retta tangente alla curva ottenuta intersecando il grafico di f con il piano $y = y_0$. Nello spazio tale retta tangente giace nel piano $y = y_0$, passa per il punto $(x_0, y_0, f(x_0, y_0))$ ed ha vettore direzionale

$$v = (1, 0, \frac{\partial f}{\partial x}(x_0, y_0)).$$

Un analogo ragionamento vale per la derivata parziale rispetto ad y che rappresenta il coefficiente angolare della retta tangente alla curva ottenuta intersecando il grafico di f con il piano $x = x_0$. Nello spazio tale retta tangente giace nel piano $x = x_0$, passa per il punto $(x_0, y_0, f(x_0, y_0))$ ed ha vettore direzionale

$$w = (0, 1, \frac{\partial f}{\partial y}(x_0, y_0)).$$

Poiché i due vettori v e w non sono paralleli, definiscono un piano passante per il punto $Q_0 = (x_0, y_0, f(x_0, y_0))$ di equazione parametrica

$$Q = Q_0 + tv + sw \quad \Longleftrightarrow \quad \sigma : \begin{cases} x = x_0 + t \\ y = y_0 + s \\ z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)t + \frac{\partial f}{\partial y}(x_0, y_0)s \end{cases}.$$

Introdotta il vettore normale

$$N = v \wedge w = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \frac{\partial f}{\partial x}(x_0, y_0) \\ 0 & 1 & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix} = (-\frac{\partial f}{\partial x}(x_0, y_0), -\frac{\partial f}{\partial y}(x_0, y_0), 1) \neq 0,$$

che soddisfa $N \cdot v = N \cdot w = 0$, l'equazione cartesiana del piano diventa

$$\begin{aligned} N \cdot (Q - Q_0) &= 0 \\ \Longleftrightarrow \\ z &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0). \end{aligned} \quad (62)$$

Esempio 2.28. La funzione

$$f(x, y) = x^2 \cos(2\pi x) \quad (x, y) \in \mathbb{R}^2,$$

il cui grafico è rappresentato in Fig 6(a), è derivabile con

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

Consideriamo il punto $(x_0, y_0) = (1/2, 0)$. Le restrizioni alle rette $y = 0$ ed $x = 1/2$ sono rispettivamente una parabola ed un coseno

$$f_{x_0}(x) = x^2 \quad f_{y_0}(y) = \frac{1}{4} \cos(2\pi y),$$

si vedano i grafici Fig 6(b-c), In particolare

$$\frac{\partial f}{\partial x}(1/2, 0) = f'_{y_0}(1/2) = 1 \quad \frac{\partial f}{\partial y}(1/2, 0) = f'_{x_0}(0) = 0.$$

Le rette tangenti alle due curve hanno vettore direzionale dato da

$$v = (1, 0, 1) \quad w = (0, 1, 0) \implies N = v \wedge w = (-1, 0, 1).$$

Il piano (62) passante per il punto $(1/2, 0, 1/4) \in \mathbb{R}^3$ ha equazione

$$z = \frac{1}{2} + x$$

Il grafico di f e del piano sono rappresentati in Fig. 6(d-e).

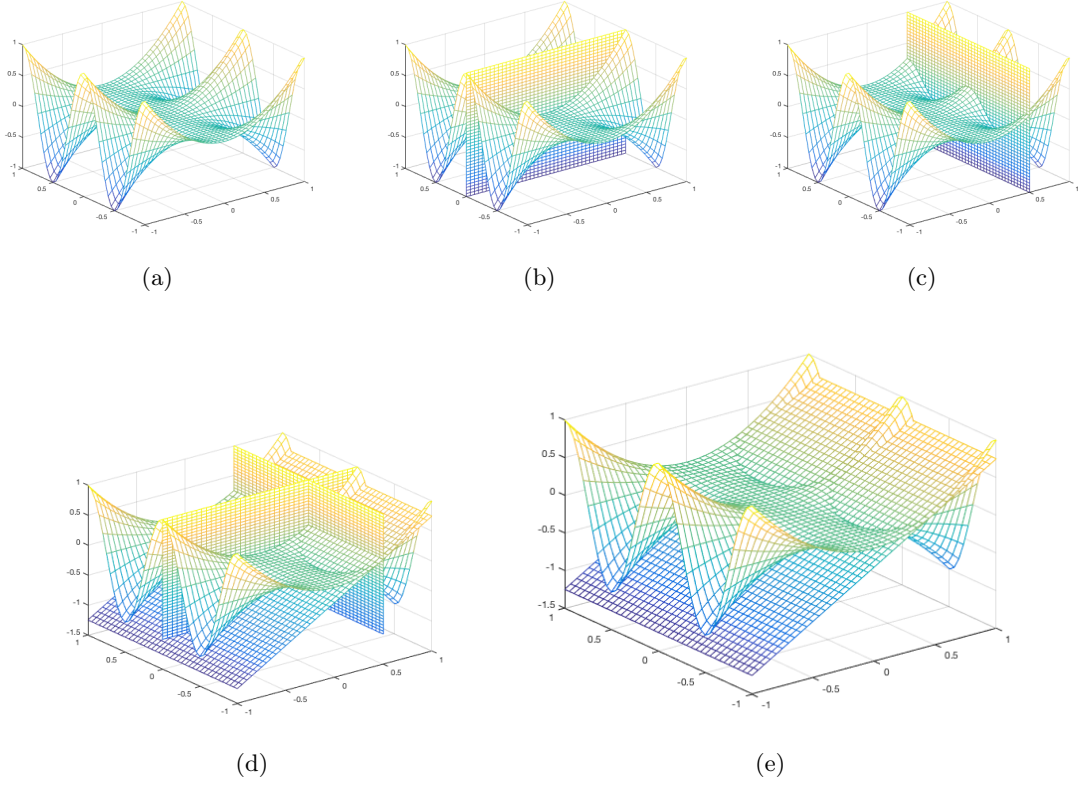


FIGURA 6. Grafico relativi a $f(x, y) = x^2 \cos(2\pi x)$.

Dalla (60) segue che

$$f(x, y_0) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + |x - x_0| \epsilon(x - x_0)$$

$$f(x_0, y) = f(x_0, y_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + |y - y_0| \epsilon(y - y_0),$$

tuttavia non si può stabilire se il piano definito dalla (62) approssimi o meno il grafico di f quando (x, y) è vicino a (x_0, y_0) , come mostrato nell'Esempio 2.31. Questo motiva la seguente definizione.

Def. 2.29. Data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ definita su un aperto A e derivabile, è detta differenziabile nel punto $(x_0, y_0) \in A$ se

$$f(x, y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + \sqrt{(x - x_0)^2 + (y - y_0)^2} \epsilon(x - x_0, y - y_0) \quad (63)$$

dove $\epsilon(x - x_0, y - y_0)$ è una funzione infinitesima. In tale caso il piano di equazione cartesiana

$$z = f(x_0, y_0) + \cancel{f(x_0, y_0)} + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \quad (64)$$

è detto piano tangente al grafico di f nel punto $(x_0, y_0, f(x_0, y_0))$.

La funzione f è detta differenziabile se è differenziabile in tutti i punti di A .

In forma compatta la (63) diventa

$$f(P) = f(P_0) + \nabla f(P_0) \cdot (P - P_0) + \|P - P_0\| \epsilon(P - P_0) \quad P \in A. \quad (65)$$

Dal punto geometrico la condizione che f sia differenziabile in (x_0, y_0) significa che la distanza in \mathbb{R}^3 tra il punto $(x, y, f(x, y))$ sul grafico di f ed il corrispondente punto $(x, y, f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0))$ sul piano (64) tende a zero più velocemente della distanza in \mathbb{R}^2 di (x, y) da (x_0, y_0) . Dalla (64) segue che il vettore normale al piano tangente è

$$N = \left(-\frac{\partial f}{\partial x}(x_0, y_0), -\frac{\partial f}{\partial y}(x_0, y_0), 1\right),$$

le cui due prime componenti sono $-\nabla f(x_0, y_0)$ e la terza componente è 1.

Il seguente risultato mostra che le funzioni differenziabili sono sempre continue.

Prop. 2.30. *Se $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ è una funzione differenziabile definita su un insieme aperto, allora f è continua.*

Dimostrazione. Fissato $P_0 \in A$, la definizione di differenziabilità implica che

$$f(P) = f(P_0) + \nabla f(P_0) \cdot (P - P_0) + \|P - P_0\| \epsilon(P - P_0) \quad P \in A.$$

Osserviamo che la funzione

$$g(P) = \nabla f(P_0) \cdot (P - P_0) + \|P - P_0\| \epsilon(P - P_0) \quad P \in A$$

è continua in P_0 e $g(P_0) = 0$, cioè $g(P) = \epsilon(P - P_0)$ è una funzione infinitesima e, quindi,

$$f(P) = f(P_0) + \epsilon(P - P_0)$$

è continua in P_0 . □

Come mostra il seguente esempio, ci sono funzioni patologiche che sono derivabili, ma non sono differenziabili.

Esempio 2.31. La funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

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

è derivabile in $(0, 0)$, ma non è differenziabile in $(0, 0)$. Infatti,

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x} &= \lim_{x \rightarrow 0} \frac{x \cdot 0}{(x^2 + 0)x} = 0 \\ \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{y} &= \lim_{y \rightarrow 0} \frac{0 \cdot y}{(0 + y^2)y} = 0, \end{aligned}$$

per cui f è derivabile in $(0, 0)$ e

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

Il piano (62) è $z = 0$, cioè il piano xy . Tuttavia la funzione f non è continua nell'origine. Infatti

$$f(x, x) = \begin{cases} \frac{1}{2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

quindi, per la Proposizione 2.30 f non è differenziabile, vedi Fig. 7.

Il seguente teorema dà una condizione sufficiente per la differenziabilità, espressa in termini di continuità delle derivate parziali.

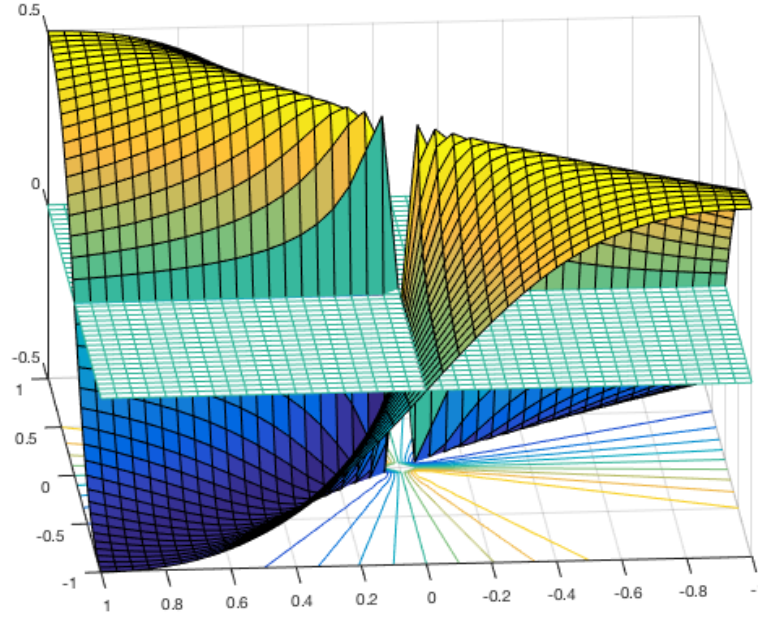


FIGURA 7. Grafico della funzione $\frac{xy}{x^2+y^2}$.

Teo 2.32. Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ con A aperto. Se sono soddisfatte le seguenti condizioni

- a) f è derivabile,
- b) le derivate parziali $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ sono continue,

allora f è differenziabile.

Dimostrazione. A meno di una traslazione, è sufficiente mostrare che f è differenziabile in $(0,0) \in A$. Poiché $(0,0)$ è interno ad A , esiste $r > 0$ tale che $[-r, r] \times [-r, r] \subset A$. Per ogni $(x, y) \in [-r, r] \times [-r, r]$

$$\begin{aligned} f(x, y) - f(0, 0) &= \underbrace{f(x, y) - f(0, y)}_{\text{I}} + \underbrace{f(0, y) - f(0, 0)}_{\text{II}} \\ &= \frac{\partial f}{\partial x}(t_{x,y}, y) x + \frac{\partial f}{\partial y}(0, \tau_y) y \end{aligned}$$

dove $0 \leq |t_{x,y}| \leq |x|$ e $0 \leq |\tau_y| \leq |y|$. Infatti, per l'addendo I, fissato y , si è applicato il teorema di Lagrange alla funzione di una variabile

$$\varphi_y : [-r, r] \rightarrow \mathbb{R} \quad \varphi_y(x) = f(x, y)$$

che per l'ipotesi a) risulta derivabile con derivata $\varphi'_y(x) = \frac{\partial f}{\partial x}(x, y)$. Analogamente per l'addendo II si è applicato il teorema di Lagrange alla funzione di una variabile

$$\psi : [-r, r] \rightarrow \mathbb{R} \quad \psi(y) = f(0, y) \quad \psi'(y) = \frac{\partial f}{\partial y}(0, y).$$

Osserviamo che per costruzione $t_{x,y}$ e τ_y sono funzioni infinitesime di (x, y) . L'ipotesi b) e c) di continuità assicurano che

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{\partial f}{\partial x}(0, 0) + \epsilon(x, y) \\ \frac{\partial f}{\partial y}(x, y) &= \frac{\partial f}{\partial y}(0, 0) + \epsilon(x, y) \end{aligned}$$

dove le funzioni ϵ (eventualmente diverse tra di loro) sono infinitesime. Ne segue che

$$\begin{aligned}
 f(x, y) &= f(0, 0) + \frac{\partial f}{\partial x}(t_{x,y}, y) x + \frac{\partial f}{\partial y}(0, \tau_y) y \\
 &= f(0, 0) + \frac{\partial f}{\partial x}(0, 0) x + \frac{\partial f}{\partial y}(0, 0) y + \epsilon(t_{x,y}, y) x + \epsilon(x, t_y) y \\
 &\stackrel{(*)}{=} f(0, 0) + \frac{\partial f}{\partial x}(0, 0) x + \frac{\partial f}{\partial y}(0, 0) y + \\
 &\quad + \sqrt{x^2 + y^2} \left(\epsilon(x, y) \frac{x}{\sqrt{x^2 + y^2}} + \epsilon(x, y) \frac{y}{\sqrt{x^2 + y^2}} \right), \\
 &\stackrel{(**)}{=} f(0, 0) + \frac{\partial f}{\partial x}(0, 0) x + \frac{\partial f}{\partial y}(0, 0) y + \sqrt{x^2 + y^2} \epsilon(x, y),
 \end{aligned}$$

dove in $(*)$ si è sfruttato il fatto che $t_{x,y}$ e τ_y sono infinitesime ed in $(**)$ ϵ è ancora infinitesima poiché

$$\left| \frac{x}{\sqrt{x^2 + y^2}} \right| \leq 1 \quad \text{e} \quad \left| \frac{y}{\sqrt{x^2 + y^2}} \right| \leq 1.$$

□

Una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ è detta di classe C^1 se f è derivabile e le derivate parziali $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ sono continue. Il precedente teorema assicura che se f è di classe C^1 , allora f è differenziabile.

Dalle proprietà delle funzioni continue e derivabili segue che somma, differenza, prodotto e rapporto di funzioni di classe C^1 è di classe C^1 e, quindi, differenziabile.

Esempio 2.33. La funzione

$$f(x, y) = -(x + 2)y^2 \quad (x, y) \in \mathbb{R}^2$$

è derivabile e le sue derivate

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

sono funzioni continue, quindi è differenziabile. Consideriamo il punto $(x_0, y_0) = (0, -1/2)$. Poiché

$$f(0, -1/2) = -\frac{1}{2} \quad \frac{\partial f}{\partial x}(0, -1/2) = -\frac{1}{4} \quad \frac{\partial f}{\partial y}(0, -1/2) = 2,$$

l'equazione del piano tangente al grafico di f nel punto $(0, -1/2, -1/2)$ ha equazione

$$z = -\frac{1}{2} - \frac{1}{4}x + 2(y + \frac{1}{2}).$$

Il grafico di f e del piano tangente sono rappresentati in Fig. 8.

Il seguente risultato mostra come calcolare la variazione di f lungo una generica direzione.

Cor. 2.34. Data una funzione differenziabile $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ definita su un insieme aperto A , fissato $P_0 = (x_0, y_0) \in A$, per ogni vettore $v = (v_1, v_2) \in \mathbb{R}^2$ esiste finito

$$\lim_{t \rightarrow 0} \frac{f(P_0 + tv) - f(P_0)}{t} = \nabla f(P_0) \cdot v = \frac{\partial f}{\partial x}(x_0, y_0)v_1 + \frac{\partial f}{\partial y}(x_0, y_0)v_2. \quad (66)$$

Dimostrazione. Dalla (65) scegliendo $P = P_0 + tv$

$$\frac{f(P_0 + tv) - f(P_0)}{t} = \nabla f(P_0) \cdot v + \frac{|t|}{t} \|v\| \epsilon(tv).$$

Poiché $\lim_{t \rightarrow 0} \frac{|t|}{t} \epsilon(tv) = 0$ segue la tesi.

□

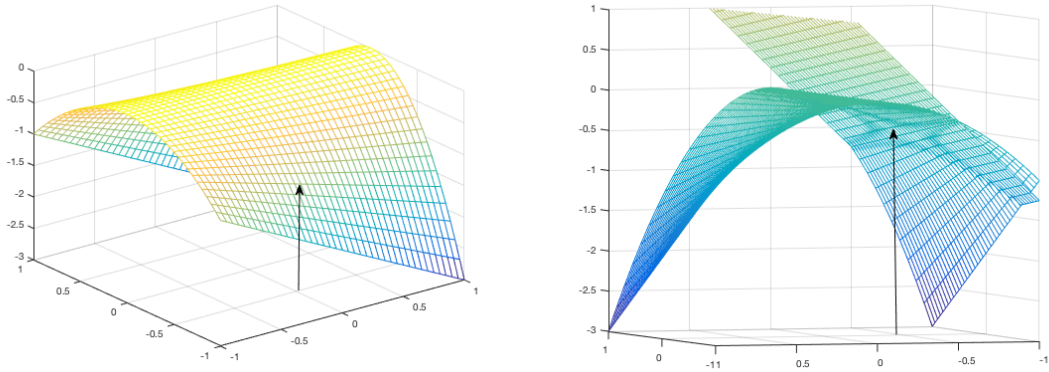


FIGURA 8. Grafico di $f(x) = f(x, y) = -(x+2)y^2$ e del piano tangente nel punto $(0, -1/2, -1/2)$.

Se esiste finito il limite (66) si pone

$$\frac{\partial f}{\partial v}(P_0) = \lim_{t \rightarrow 0} \frac{f(P_0 + tv) - f(P_0)}{t} = \lim_{t \rightarrow 0} \frac{f(x_0 + tv_1, y_0 + tv_2) - f(x_0, y_0)}{t},$$

e si chiama derivata direzionale lungo il vettore $v = (v_1, v_2) \in \mathbb{R}^2$ nel punto $P_0 = (x_0, y_0)$. In particolare,

$$\begin{aligned} \frac{\partial f}{\partial v}(x_0, y_0) &= \frac{\partial f}{\partial x}(x_0, y_0) & \text{se } v &= (1, 0) \\ \frac{\partial f}{\partial v}(x_0, y_0) &= \frac{\partial f}{\partial y}(x_0, y_0) & \text{se } v &= (0, 1). \end{aligned}$$

Il seguente esempio mostra che l'esistenza di tutte le derivate direzionali sia solo una condizione necessaria.

Esempio 2.35. La funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

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

ammette derivate direzionali lungo ogni vettore $v = (v_1, v_2) \in \mathbb{R}^2$ in $(0, 0)$

$$\frac{\partial f}{\partial v}(0, 0) = \begin{cases} \frac{v_1^2 v_2}{v_1^2 + v_2^2} & (v_1, v_2) \neq (0, 0) \\ 0 & (v_1, v_2) = (0, 0) \end{cases},$$

ma non è differenziabile in $(0, 0)$.

Infatti, dato $v = (v_1, v_2) \neq (0, 0)$

$$\lim_{t \rightarrow 0} \frac{f(tv_1, tv_2) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{(tv_1)^2 tv_2}{((tv_1)^2 + (tv_2)^2) t} = \frac{v_1^2 v_2}{v_1^2 + v_2^2}.$$

In particolare si ha che $\frac{\partial f}{\partial x}(0, 0) = 0$ (se $v = (1, 0)$) e $\frac{\partial f}{\partial y}(0, 0) = 0$ (se $v = (0, 1)$).

Tuttavia f non è differenziabile in $(0, 0)$. Infatti, se lo fosse, l'equazione (66) implicherebbe $\frac{\partial f}{\partial v}(0, 0) = 0$ per ogni v .

Esempio 2.36. Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la funzione $f(x, y) = x^2 + y^2 = \|P - O\|^2$ (distanza al quadrato di P dall'origine O). Poiché f è derivabile e le derivate parziali

$$\frac{\partial f}{\partial x}(x, y) = 2x \quad \text{e} \quad \frac{\partial f}{\partial y}(x, y) = 2y \quad \forall (x, y) \in \mathbb{R}^2$$

sono funzioni continue, f è di classe C^1 e, quindi, è differenziabile. In particolare, dato un punto $P_0 = (x_0, y_0) \in \mathbb{R}^2$ il gradiente vale

$$\nabla f(x_0, y_0) = 2(x_0, y_0) \quad \text{in forma compatta } \nabla f(P_0) = 2(P_0 - O)$$

e la differenziabilità assicura l'esistenza del piano tangente al grafico di f nel punto $(x_0, y_0, x_0^2 + y_0^2)$ di equazione

$$z = x_0^2 + y_0^2 + 2x_0(x - x_0) + 2y_0(y - y_0) = 2x_0 x + 2y_0 y - (x_0^2 + y_0^2)$$

e la derivata direzionale lungo ogni vettore $v = (v_1, v_2) \in \mathbb{R}^2$ in P_0

$$\frac{\partial f}{\partial v}(x_0, y_0) = 2x_0 v_1 + 2y_0 v_2 \quad \text{in forma vettoriale} \quad \frac{\partial f}{\partial v}(P_0) = 2(P_0 - O) \cdot v.$$

Nel caso di funzioni di due variabili, il teorema di derivazione di funzioni composte dà luogo a due regole di derivazione, note come derivazione in catena.

Teo 2.37. Data una funzione differenziabile $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$

a) se

$$\varphi : I \subset \mathbb{R} \rightarrow \mathbb{R}$$

è una funzione derivabile definita su un intervallo I e $f(x, y) \in I$ per ogni $(x, y) \in A$, allora la funzione composta

$$g(x, y) = \varphi(f(x, y)) \quad (x, y) \in A$$

è differenziabile e

$$\begin{cases} \frac{\partial \varphi(f(x, y))}{\partial x} = \varphi'(f(x, y)) \frac{\partial f}{\partial x}(x, y) \\ \frac{\partial \varphi(f(x, y))}{\partial y} = \varphi'(f(x, y)) \frac{\partial f}{\partial y}(x, y) \end{cases} \quad \text{cioè} \quad \nabla g(P) = \varphi'(f(P)) \nabla f(P);$$

per ogni $P = (x, y) \in A$;

b) se $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ è una curva

$$\gamma : \begin{cases} x = x(t) \\ y = y(t) \end{cases} \quad t \in I$$

definita su intervallo I , le cui componenti $x(t)$ ed $y(t)$ sono derivabili e $(x(t), y(t)) \in A$ per ogni $t \in I$, allora la funzione composta

$$\psi(t) = f(\gamma(t)) = f(x(t), y(t)) \quad t \in I$$

è derivabile e vale

$$\begin{aligned} f(x(t), y(t))' &= \frac{\partial f}{\partial x}(x(t), y(t)) x'(t) + \frac{\partial f}{\partial y}(x(t), y(t)) y'(t) \\ &= \nabla f(\gamma(t)) \cdot \gamma'(t) \quad \text{dove } \gamma'(t) = (x'(t), y'(t)) \end{aligned} \quad (67)$$

per ogni $t \in I$.

Esempio 2.38. Sia $g(x, y) = \sqrt{x^2 + y^2} = \|P - O\|$ con dominio \mathbb{R}^2 . Poiché la funzione $f(x, y) = x^2 + y^2$ è differenziabile in \mathbb{R}^2 con $\nabla f(x, y) = (2x, 2y) = 2(P - O)$ e la funzione $\varphi(t) = \sqrt{t}$ è derivabile in $(0, +\infty)$ con $\varphi'(t) = \frac{1}{2\sqrt{t}}$, per il teorema di derivazione in catena g è differenziabile in $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \neq 0\}$ e vale

$$\begin{cases} \frac{\partial g}{\partial x}(x, y) = \frac{x}{\sqrt{x^2 + y^2}} \\ \frac{\partial g}{\partial y}(x, y) = \frac{y}{\sqrt{x^2 + y^2}} \end{cases} \quad \text{cioè} \quad (\nabla g)(P) = \frac{P - O}{\|P - O\|}.$$

Nell'origine $g(x, y)$ non è derivabile poiché non esistono le derivate parziali in $(0, 0)$: infatti le funzioni $g(x, 0) = |x|$ e $g(0, y) = |y|$ non sono derivabili in 0.

Analogamente, la funzione $h(x, y) = \frac{1}{\sqrt{x^2+y^2}} = \frac{1}{\|P-O\|}$ con dominio $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \neq 0\}$ è differenziabile e, poiché $(t^{-\frac{1}{2}})' = -\frac{1}{2(\sqrt{t})^3}$, vale

$$(\nabla h)(P) = -\frac{P-O}{\|P-O\|^3} = \frac{-1}{\|P-O\|^2} \frac{P-O}{\|P-O\|}.$$

Esempio 2.39. Sia $\gamma(t) = (x_0 + v_1 t, y_0 + v_2 t) = P_0 + vt$, $t \in \mathbb{R}$, la retta passante per $P_0 = (x_0, y_0)$ con direzione $v = (v_1, v_2)$. La curva è differenziabile e $\gamma'(t) = v$ per ogni $t \in \mathbb{R}$.

Se $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ è differenziabile in P_0 punto interno di A , allora la funzione composta $\psi(t) = f(x_0 + v_1 t, y_0 + v_2 t)$ è derivabile in 0 e l'equazione (67) con $t = 0$ dà

$$\psi'(0) = f(x_0 + v_1 t, y_0 + v_2 t)' \Big|_{t=0} = \nabla f(P_0) \cdot v$$

che, ovviamente, coincide con (66).

Il concetto di derivate di ordine successivo si definisce in modo ricorsivo. Trattiamo esplicitamente solo il caso delle derivate seconde per funzioni scalari di due variabili. Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione derivabile definita su un insieme aperto A . La funzione f è detta derivabile due volte se le derivate parziali $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ sono a loro volta derivabili e si definisce la matrice hessiana di f come

$$\begin{aligned} Hf(x, y) &= \begin{bmatrix} \nabla(\frac{\partial f}{\partial x}(x, y)) \\ \nabla(\frac{\partial f}{\partial y}(x, y)) \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x}(\frac{\partial f}{\partial x}(x, y)) & \frac{\partial}{\partial y}(\frac{\partial f}{\partial x}(x, y)) \\ \frac{\partial}{\partial x}(\frac{\partial f}{\partial y}(x, y)) & \frac{\partial}{\partial y}(\frac{\partial f}{\partial y}(x, y)) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial y \partial x}(x, y) \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{bmatrix}. \end{aligned}$$

Se, inoltre, tutte e quattro le derivate seconde sono continue, si dice che f è una funzione di classe C^2 . In tal caso la matrice hessiana è simmetrica, come mostra il seguente risultato.

Teo 2.40 (Teorema di Schwarz). *Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione di classe C^2 definita su un insieme aperto A , allora*

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y) \quad (x, y) \in A.$$

Dimostrazione. Senza perdita di generalità, assumiamo che l'origine $(0, 0) \in A$ e dimostriamo che $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \frac{\partial^2 f}{\partial x \partial y}(0, 0)$.

Dato $\delta > 0$ tale che $(-\delta, \delta) \times (-\delta, \delta) \subset A$, definiamo $h : (-\delta, \delta) \times (-\delta, \delta) \rightarrow \mathbb{R}$

$$\begin{aligned} h(x, y) &= f(x, y) - f(x, 0) - f(0, y) + f(0, 0) \\ &= g(x, y) - g(0, y) \\ &= \hat{g}(x, y) - \hat{g}(x, 0), \end{aligned}$$

dove

$$\begin{aligned} g : (-\delta, \delta) \times (-\delta, \delta) &\rightarrow \mathbb{R} & g(x, y) &= f(x, y) - f(x, 0) \\ \hat{g} : (-\delta, \delta) \times (-\delta, \delta) &\rightarrow \mathbb{R} & \hat{g}(x, y) &= f(x, y) - f(0, y) \end{aligned}$$

sono funzioni derivabili che soddisfano

$$\begin{aligned} \frac{\partial g}{\partial x}(x, y) &= \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(x, 0) \\ \frac{\partial \hat{g}}{\partial y}(x, y) &= \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(0, y). \end{aligned}$$

Fissati $x, y \in (0, \delta)$, per il teorema di Lagrange esiste $0 \leq t_{x,y} \leq x$ tale che

$$\begin{aligned} h(x, y) &= g(x, y) - g(0, y) = \frac{\partial g}{\partial x}(t_{x,y}, y) x \\ &= \left(\frac{\partial f}{\partial x}(t_{x,y}, y) - \frac{\partial f}{\partial x}(t_{x,y}, 0) \right) x \\ &= \frac{\partial^2 f}{\partial y \partial x}(t_{x,y}, \tau_{x,y}) xy \\ &= \left(\frac{\partial^2 f}{\partial y \partial x}(0, 0) + \epsilon(t_{x,y}, \tau_{x,y}) \right) xy \end{aligned}$$

dove $0 \leq \tau_{x,y} \leq y$ per il teorema di Lagrange applicato a $\frac{\partial f}{\partial x}(x, y)$ ed ϵ è infinitesima per l'ipotesi di continuità di $\frac{\partial^2 f}{\partial y \partial x}$. Poiché anche $t_{x,y}$ e $\tau_{x,y}$ sono infinitesime, per ogni $x, y \in (0, \delta)$

$$\frac{h(x, y)}{xy} = \frac{\partial^2 f}{\partial y \partial x}(0, 0) + \epsilon(x, y). \quad (68a)$$

Ripetendo lo stesso ragionamento con \hat{g} al posto di g si deduce che

$$\frac{h(x, y)}{xy} = \frac{\partial^2 f}{\partial x \partial y}(0, 0) + \epsilon(x, y). \quad (68b)$$

Confrontando la (68a) e la (68b), segue che

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

□

⚡. Senza l'ipotesi che le derivate seconde siano continue, può succedere che le derivate miste non siano uguali tra di loro, come mostra il seguente esempio

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

8.1. Proprietà delle funzioni differenziabili.

Il seguente risultato è l'analogo del teorema di Lagrange. Ricordiamo che un insieme $A \subset \mathbb{R}^2$ è convesso se dati $P_1, P_2 \in A$, il segmento di estremi P_1 e P_2 è tutto contenuto in A , cioè se per ogni $t \in [0, 1]$ allora $P_1 + t(P_2 - P_1) \in A$.

Teo 2.41 (formula dell'accrescimento finito). *Data una funzione differenziabile $f : A \rightarrow \mathbb{R}$ definita su insieme A aperto e convesso, allora per ogni $P_1, P_2 \in A$, esiste $Q \in A$ tale che*

$$f(P_2) - f(P_1) = \nabla f(Q) \cdot (P_2 - P_1), \quad (69)$$

dove $Q = P_1 + t(P_2 - P_1)$ con $t \in (0, 1)$.

Dimostrazione. Fissati $P_1, P_2 \in A$, definiamo la curva $\gamma(t) = P_1 + t(P_2 - P_1)$ con $t \in [0, 1]$ che è derivabile, $\gamma'(t) = (P_2 - P_1)$ e la cui traccia è contenuta in A per l'ipotesi di convessità. Inoltre, la funzione composta $f(\gamma(t))$, definita da $[0, 1]$ a valori in \mathbb{R} , è continua in $[0, 1]$ e derivabile in $(0, 1)$. Per il teorema di Lagrange esiste $t \in (0, 1)$ tale che

$$f(P_2) - f(P_1) = f(\gamma(1)) - f(\gamma(0)) = f(\gamma(t))'(1 - 0) = \nabla f(P_1 + t(P_2 - P_1)) \cdot (P_2 - P_1)$$

dove l'ultima uguaglianza segue dalla (67). La tesi segue ponendo $Q = P_1 + t(P_2 - P_1)$. □

Se $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$, la formula (69) diventa

$$f(x_2, y_2) - f(x_1, y_1) = \frac{\partial f}{\partial x}(x^*, y^*)(x_2 - x_1) + \frac{\partial f}{\partial y}(x^*, y^*)(y_2 - y_1).$$

dove

$$x^* = x_1 + t(x_2 - x_1) \quad y^* = y_1 + t(y_2 - y_1) \quad \text{con } t \in (0, 1).$$

La seguente proposizione fornisce un'interpretazione geometrica del gradiente.

Prop. 2.42. *Data una funzione differenziabile $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ definita in un aperto A , fissato $P_0 \in A$ tale che $\nabla f(P_0) \neq 0$, allora*

a) *fissato un versore $v \in \mathbb{R}^2$, sia θ l'angolo compreso tra le semirette orientate individuate dai vettori v e $\nabla f(P_0)$, allora*

$$\frac{\partial f}{\partial v}(P_0) = \|\nabla f(P_0)\| \cos \theta,$$

b) *fissata una curva $\gamma : I \rightarrow \mathbb{R}^2$*

$$\gamma : \begin{cases} x = x(t) \\ y = y(t) \end{cases} \quad t \in I$$

definita su intervallo I , le cui componenti $x(t)$ ed $y(t)$ sono derivabili e $\gamma(t) \in A$ per ogni $t \in I$, se

$$\begin{cases} \gamma(t_0) = P_0 & \text{per qualche } t_0 \in I \\ f(\gamma(t)) = f(P_0) & \text{per ogni } t \in I \end{cases},$$

allora posto $\gamma'(t) = (x'(t), y'(t))$

$$\nabla f(P_0) \cdot \gamma'(t_0) = 0.$$

Dimostrazione. L'affermazione in a) segue dalla (66) e dal fatto che il prodotto scalare tra due vettori è il prodotto delle corrispondenti norme per il coseno dell'angolo compreso. Il punto b) segue dalla (67) ponendo $t = t_0$. \square

\diamond . Se $\nabla f(P_0) = 0$, la (66) implica che $\frac{\partial f}{\partial v}(P_0) = 0$ per ogni vettore v .

La proposizione fornisce un'importante interpretazione geometrica del gradiente. Infatti, dal punto a) segue che la derivata direzionale $\frac{\partial f}{\partial v}$ è massima se v ha la stessa direzione e verso di $\nabla f(P_0)$, è nulla se v è perpendicolare a $\nabla f(P_0)$ ed è minima se v ha la stessa direzione e verso opposto di $\nabla f(P_0)$. Quindi il vettore $\nabla f(P_0)$, diverso da zero per ipotesi, individua la direzione di massima crescita per f . Inoltre il punto b) mostra che se γ è una curva passante per P_0 e la cui traccia è contenuta nell'insieme di livello di quota $c = f(P_0)$

$$A_c = \{P \in A \mid f(P) = f(P_0)\},$$

allora il gradiente $\nabla f(P_0)$ è perpendicolare al vettore tangente alla curva in P_0 e quindi la retta di equazione

$$\nabla f(P_0) \cdot (P - P_0) = 0$$

è la retta tangente all'insieme di livello A_c nel punto $P_0 = (x_0, y_0)$. Se $P_0 = (x_0, y_0)$ la retta tangente all'insieme di livello A_c ha equazione

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0.$$

Esempio 2.43. Sia $f(x, y) = x^2 + y^2 = \|P - O\|^2$ e $P_0 = (x_0, y_0) \neq (0, 0)$. L'insieme di livello $A_c = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = x_0^2 + y_0^2\}$ di quota $c = x_0^2 + y_0^2$ è una circonferenza di raggio \sqrt{c} . Ricordando che $\nabla f(P_0) = 2(x_0, y_0) = 2(P_0 - O) \neq (0, 0)$, la funzione ha la massima crescita nella direzione e verso del vettore $P_0 - O$, applicato in P_0 , e la retta tangente in P_0 ad A_c ha equazione

$$2x_0(x - x_0) + 2y_0(y - y_0) = 0 \quad \text{cioè} \quad (P - P_0) \cdot (P_0 - O) = 0.$$

In $(0, 0)$ il gradiente è nullo e l'insieme di livello di quota 0 si riduce ad un punto.

Il seguente teorema caratterizza le funzioni che hanno gradiente nullo in tutti i punti del dominio, purché sia connesso per archi.

Prop. 2.44. Data una funzione differenziabile $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ definita su un insieme A aperto e connesso per archi, se

$$\nabla f(P) = 0 \quad \forall P \in A \quad \Longleftrightarrow \quad \begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases} \quad \forall (x, y) \in A, \quad (70)$$

allora f è una funzione costante, cioè esiste $c \in \mathbb{R}$ tale che $f(P) = c$ per ogni $P \in A$.

Dimostrazione. La dimostrazione è divisa in due passi.

Passo 1. Mostriamo che f è localmente costante. Poiché A è aperto, fissato $P^* \in A$ esiste $\delta > 0$ tale che $B(P^*, \delta) \subset A$ e, essendo $B(P^*, \delta)$ convesso, la formula dell'accrescimento finito (69) implica che per ogni $P \in B(P^*, \delta)$ esiste $Q \in B(P^*, \delta)$ tale che

$$f(P) - f(P^*) = \nabla f(Q) \cdot (P - P^*) = 0$$

poiché per ipotesi il gradiente di f è nullo su A , da cui segue che $f(P) = f(P^*)$ per ogni $P \in B(P^*, \delta)$.

Passo 2. Proviamo che dati due punti $P_0, P_1 \in A$, allora $f(P_1) = f(P_0)$. Per ipotesi di connessione per archi esiste una curva continua γ di estremi $\gamma(0) = P_0$ e $\gamma(1) = P_1$, la cui traccia è contenuta tutta in A . Essendo l'insieme

$$I = \{t \in [0, 1] \mid f(\gamma(t)) = f(P_0)\}$$

non vuoto e contenuto in $[0, 1]$, l'estremo superiore t^* di I è finito e $t^* \in [0, 1]$. Denotato con $P^* = \gamma(t^*)$, per quanto visto esiste $\delta > 0$ tale che per ogni $P \in B(P^*, \delta) \subset A$ allora $f(P) = f(P^*)$. Per la continuità di γ , esiste δ' tale che se $t \in [0, 1] \cap (t^* - \delta', t^* + \delta')$, allora $\gamma(t) \in B(P^*, \delta)$ e, quindi,

$$f(\gamma(t)) = f(P^*) \quad t \in [0, 1] \cap (t^* - \delta', t^* + \delta'). \quad (71)$$

La definizione di estremo superiore implica che esiste $t \in [0, 1]$ e $t - \delta' < t \leq t^*$ tale che $f(\gamma(t)) = f(P_0)$. Dalla (71) segue che $f(P^*) = f(P_0)$ e, quindi, che $f(\gamma(t)) = f(P_0)$ per ogni $t \geq t^*$ e $t \in [0, 1]$. Poiché t^* è un maggiorante, allora necessariamente $t^* = 1$, da cui segue che $f(P_1) = f(P_0)$. □

La condizione che A sia connesso per archi non si può eliminare dalle ipotesi del proposizione.

Esempio 2.45. Sia $A = \{(x, y) \in \mathbb{R}^2 \mid xy \neq 0\}$ ed $f : A \rightarrow \mathbb{R}$

$$f(x, y) = \arctan\left(\frac{x}{y}\right) + \arctan\left(\frac{y}{x}\right) = \varphi\left(\frac{x}{y}\right),$$

dove $\varphi(t) = \arctan t + \arctan(1/t)$ definita su $\mathbb{R} \setminus \{0\}$. Il gradiente di f è nullo poiché

$$\nabla f = \varphi'\left(\frac{x}{y}\right) \left(\frac{1}{y}, -\frac{x}{y^2}\right) = 0$$

essendo $\varphi'(t) = 0$. Tuttavia f non è costante poiché $f(1, 1) = \pi/2$ ed $f(1, -1) = -\pi/2$.

La seguente proposizione dà una condizione necessaria per l'esistenza di estremi relativi, di cui richiamiamo la definizione.

Def. 2.46. Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $P_0 \in A$. Se esiste $\delta > 0$ tale che per ogni punto $P \in A \cap B(P_0, \delta)$, $P \neq P_0$

$f(P) \geq f(P_0)$	P_0 è detto punto di minimo relativo
	$f(P_0)$ è detto minimo relativo
$f(P) \leq f(P_0)$	P_0 è detto punto di massimo relativo
	$f(P_0)$ è detto massimo relativo
$f(P) > f(P_0)$	P_0 è detto punto di minimo relativo stretto
	$f(P_0)$ è detto minimo relativo stretto
$f(P) < f(P_0)$	P_0 è detto punto di massimo relativo stretto
	$f(P_0)$ è detto massimo relativo stretto

Infine, P_0 è detto punto di sella se esistono due versori $\hat{v}, \hat{w} \in \mathbb{R}^2$ e $\delta > 0$ tali che per ogni $0 < |t| < \delta$

$$\begin{aligned} f(P_0 + t\hat{v}) &> f(P_0) \\ f(P_0 + t\hat{w}) &< f(P_0). \end{aligned}$$

In analogia con il caso delle funzioni di una variabile, vale la seguente condizione necessaria.

Prop. 2.47 (Condizione necessaria del prim'ordine per estremi relativi). *Data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ differenziabile in P_0 punto interno di A . Se $P_0 = (x_0, y_0)$ è un punto di massimo o minimo relativo, allora*

$$\nabla f(P_0) = 0 \quad \Longleftrightarrow \quad \begin{cases} \frac{\partial f}{\partial x}(x_0, y_0) = 0 \\ \frac{\partial f}{\partial y}(x_0, y_0) = 0 \end{cases}.$$

Dimostrazione. Consideriamo il caso in cui P_0 sia un punto di massimo relativo. Per definizione, esiste $\delta > 0$ tale che $f(P) - f(P_0) \leq 0$ per ogni $P \in B(P_0, \delta) \subset A$. Posto $v = \nabla f(P_0)$, se $0 < t < \frac{\delta}{1+\|v\|}$, i punti $P_0 + tv$ appartengono a $B(P_0, \delta) \subset A$ e, dalla definizione di differenziabilità,

$$0 \geq \frac{f(P_0 + tv) - f(P_0)}{t} = \frac{\nabla f(P_0) \cdot tv + \|tv\|\epsilon(tv)}{t} = v \cdot v + \|v\|\epsilon(tv) = \|v\|^2 + \|v\|\epsilon(tv)$$

Passando al limite per t che tende a zero (da destra) si ottiene che $\|v\|^2 \leq 0$, da cui $v = 0$. \square

Se $\nabla f(P_0) = 0$, P_0 è detto un punto critico per f in A . La proposizione implica che gli estremi relativi interni al dominio di f sono da cercare tra i punti critici. Tuttavia tale condizione è solo necessaria, come mostra il seguente esempio.

Esempio 2.48. La funzione $f(x, y) = y^2 - x^2$ ha come punti critici solo l'origine $(0, 0)$. Tuttavia, $f(0, 0) = 0$ e

$$\begin{aligned} f(x, y) &> 0 = f(0, 0) && \text{se } |y| > |x| \\ f(x, y) &< 0 = f(0, 0) && \text{se } |y| < |x|, \end{aligned}$$

per cui $f(x, y)$ cambia segno in ogni palla di centro $(0, 0)$ e, quindi, $(0, 0)$ non è un punto di estremo relativo.

La seguente proposizione dà una condizione sufficiente affinché un punto critico sia di estremo relativo.

Prop. 2.49 (Condizione sufficiente del second'ordine per estremi relativi).

Data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe C^2 e definita su un insieme aperto A , sia $(x_0, y_0) \in A$ un punto critico di f .

a) Se $\det Hf(x_0, y_0) > 0$ e $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$, allora P_0 è un punto di minimo relativo stretto.

- b) Se $\det Hf(x_0, y_0) > 0$ e $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$, allora P_0 è un punto di massimo relativo stretto.
 c) Se $\det Hf(x_0, y_0) < 0$, P_0 è un punto di sella.

⚡. Se P_0 non è un punto critico, non può essere un punto di massimo o minimo relativo.

⚡. Se $\det Hf(x_0, y_0) = 0$, la condizione del second'ordine non permette di stabilire se un punto critico sia un punto di estremo relativo o meno.

Esempio 2.50. Le funzioni $f(x, y) = x^4 + y^4$ e $g(x, y) = x^4 - y^4$ sono di classe C^2 , hanno come unico punto critico l'origine $(0, 0)$ e la matrice hessiana nell'origine è

$$Hf(0, 0) = Hg(0, 0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Tuttavia, $(0, 0)$ è un punto di minimo assoluto per f poiché $x^4 + y^4 \geq 0$ e $(0, 0)$ non è un punto di estremo relativo per g poiché $g(x, 0) = x^4$ ha un minimo assoluto in $x = 0$ e $g(0, y) = -y^4$ ha un massimo assoluto in $y = 0$.

8.2. Funzioni implicite. Il seguente teorema fornisce una condizione sufficiente affinché, data un'equazione della forma $f(x, y) = 0$, sia possibile determinare y come funzione della variabile x .

Teo 2.51 (Teorema della funzione implicita o di Dini). *Data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe C^1 definita su un insieme aperto A aperto, se un punto $(x_0, y_0) \in A$ soddisfa*

$$f(x_0, y_0) = 0 \quad e \quad \frac{\partial f}{\partial y}(x_0, y_0) \neq 0,$$

allora esiste una funzione

$$\varphi : I \subset \mathbb{R} \rightarrow \mathbb{R} \quad y = \varphi(x)$$

di classe C^1 e definita su insieme aperto I cui appartiene x_0 , tale che tale che

$$f(x, \varphi(x)) = 0 \tag{72a}$$

$$\varphi(x_0) = y_0 \tag{72b}$$

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \varphi(x))}{\frac{\partial f}{\partial y}(x, \varphi(x))} \tag{72c}$$

per ogni $x \in I$. Inoltre, esiste un insieme aperto $V \subset \mathbb{R}^2$ cui appartiene (x_0, y_0) , tale che

a) *per ogni $x \in I$ il punto $(x, \varphi(x)) \in V$,*

b) *se $(x, y) \in V$ è soluzione dell'equazione $f(x, y) = 0$, allora $x \in I$ e $y = \varphi(x)$.*

Se, inoltre f è di classe C^2 , allora φ è di classe C^2 .

⚡. La (72a) e la (72c) sottintendono che $(x, \varphi(x)) \in A$ e $\frac{\partial f}{\partial y}(x, \varphi(x)) \neq 0$ per ogni $x \in I$.

Dimostrazione. Definiamo la funzione $\Phi : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$

$$\Phi(x, y) = (x, f(x, y)),$$

che è di classe C^1 con matrice Jacobiana

$$J\Phi(x, y) = \begin{bmatrix} 1 & 0 \\ \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{bmatrix}$$

il cui determinante è $\frac{\partial f}{\partial y}(x, y)$. L'ipotesi $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ implica che $\det J\Phi(x_0, y_0) \neq 0$ ed il teorema della funzione inversa locale assicura che esistono due aperti V e W tali che $(x_0, y_0) \in V \subset A$, $(x_0, f(x_0, y_0)) = (x_0, 0) \in W$, $\Phi(V) = W$, Φ ristretta a V è invertibile e l'inversa $\Phi^{-1} : W \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$

$$\Phi^{-1}(u, v) = (g(u, v), h(u, v)) \tag{73}$$

è di classe C^1 . Poiché $\Phi(\Phi^{-1}(u, v)) = (u, v)$, segue che

$$\begin{aligned} g(u, v) &= u \\ f(g(u, v), h(u, v)) &= v = f(u, h(u, v)). \end{aligned} \quad (74)$$

Definiamo $I = \{x \in \mathbb{R} \mid (x, 0) \in W\}$ e $\varphi : I \rightarrow \mathbb{R}$

$$\varphi(x) = h(x, 0).$$

Poiché W è un intorno aperto di $(x_0, 0)$, allora I è un intorno aperto di x_0 e φ è di classe C^1 per costruzione. La (72a) segue dalla (74) con $u = x$ e $v = 0$.

Essendo $f(x, \varphi(x)) = 0$ e, quindi, $f(x, \varphi(x))' = 0$ per ogni $x \in I$, dalla regola di derivazione in catena segue che

$$f(x, \varphi(x))' = \frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x))\varphi'(x) = 0, \quad (75)$$

sa cui segue la (72c).

Evidentemente, se $x \in I$, $(x, 0) \in W$ e $\Phi^{-1}(x, 0) = (x, \varphi(x)) \in V$. Viceversa se $(x, y) \in V$ e $f(x, y) = 0$

$$(x, y) = \Phi^{-1}(\Phi(x, y)) = \Phi^{-1}(x, f(x, y)) = \Phi^{-1}(x, 0),$$

sa cui segue che $x \in I$ e $y = \varphi(x)$.

Se f è di classe C^2 , dall'equazione (72c), segue che φ è di classe C^2 . Per derivazione successive si dimostra il caso $f \in C^k$.

□

L'applicazione $y = \varphi(x)$ è detta *funzione definita implicitamente* dall'equazione $f(x, y) = 0$ in un intorno di (x_0, y_0) . Se la derivata parziale rispetto $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, il precedente teorema assicura che, per ogni $x \in I$ l'equazione

$$f(x, y) = 0$$

ammette un'unica soluzione y , data da $y = \varphi(x)$, tale che $(x, y) \in V$. Osserviamo che, essendo I e V aperti contenenti x_0 ed (x_0, y_0) , rispettivamente, allora esistono $\sigma > 0$ e $\delta > 0$ tali che

$$(x, \varphi(x)) \in B((x_0, y_0), \sigma) \subset V \quad \text{per ogni } (x_0 - \delta, x_0 + \delta) \subset I,$$

allora per ogni $x \in \mathbb{R}$ tale che $|x - x_0| < \delta$ esiste unica $y \in \mathbb{R}$ soluzione dell'equazione $f(x, y) = 0$ tale che $|y - y_0| < \sigma$ ed è data da $y = \varphi(x)$. In particolare, poiché $f(x_0, y_0) = 0$ se $x = x_0$ la corrispondente soluzione è $y_0 = \varphi(x_0)$. Inoltre, essendo $\varphi(x)$ di classe C^1 , la soluzione y dipende con regolarità da x .

Dal punto di vista geometrico la condizione $f(x, \varphi(x)) = 0$ equivale al fatto che l'insieme di livello di f di quota 0 in un intorno di (x_0, y_0)

$$\{(x, y) \in A \cap V \mid f(x, y) = 0\}$$

è il grafico $y = \varphi(x)$ di una funzione di una variabile.

Analogamente, se $f(x_0, y_0) = 0$ e $\frac{\partial f}{\partial x}(x_0, y_0) \neq 0$, è possibile esplicitare la variabile x come funzione della variabile y . Più precisamente, esiste $\psi : J \rightarrow \mathbb{R}$, J intorno aperto di y_0 , di classe C^1 tale che

$$\psi(y_0) = x_0 \quad f(\psi(y), y) = 0 \quad \psi'(y) = -\frac{\frac{\partial f}{\partial y}(\psi(y), y)}{\frac{\partial f}{\partial x}(\psi(y), y)}$$

per ogni $y \in J$.

Esempio 2.52. Sia f è una funzione di classe C^2 e $(x_0, y_0) \in A$ tale che

$$f(x_0, y_0) = 0 \quad \text{e} \quad \begin{cases} \frac{\partial f}{\partial x}(x_0, y_0) = 0 \\ \frac{\partial f}{\partial y}(x_0, y_0) \neq 0 \end{cases}.$$

Il teorema della funzione implicita assicura che esiste $y = \varphi(x)$ definita in un intorno aperto I di x_0 , derivabile due volte in I , tale che

$$\varphi(x_0) = y_0 \quad , \quad \varphi'(x_0) = 0 \quad \text{e} \quad \varphi''(x_0) = -\frac{\frac{\partial^2 f}{\partial x^2}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)},$$

dove la seconda uguaglianza segue da (72c) e $\frac{\partial f}{\partial x}(x_0, y_0) = 0$, mentre la terza si ottiene derivando ulteriormente la (75) e ponendo $x = x_0$. In particolare, se $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) \neq 0$, x_0 è un punto di minimo o massimo relativo per φ .

Il seguente esempio mostra i problemi che sorgono se $\frac{\partial f}{\partial y}(x_0, y_0) = 0$.

Esempio 2.53. L'equazione $y^2 - x = 0$ ha come soluzione $(x_0, y_0) = (0, 0)$. Tuttavia, per ogni $x > 0$ esistono due soluzioni $y = \sqrt{x}$ e $y = -\sqrt{x}$, mentre per $x < 0$ non esiste alcuna soluzione. Il teorema della funzione implicita non si può applicare: infatti, posto $f(x, y) = y^2 - x$, $\nabla f(x, y) = (-1, 2y)$ per cui $\frac{\partial f}{\partial y}(0, 0) = 0$. Tuttavia, poiché $\frac{\partial f}{\partial x}(0, 0) = -1$, è possibile esplicitare x come funzione della y , infatti $x = y^2 = \psi(y)$ per ogni $y \in \mathbb{R}$.

Il seguente esempio mostra i problemi che sorgono se $\nabla f(x_0, y_0) = 0$.

Esempio 2.54. L'equazione $y^2 - x^2 = 0$ ha come soluzione $(x_0, y_0) = (0, 0)$. Tuttavia in un intorno di $(0, 0)$ non è possibile esplicitare né la y come funzione della x , né la y come funzione della x poiché $\{(x, y) \in \mathbb{R}^2 \mid y^2 - x^2 = 0\}$ è la coppia di rette $y = x$ e $y = -x$ che si intersecano nell'origine. Il teorema della funzione implicita non si può applicare: infatti, posto $f(x, y) = y^2 - x^2$, $\nabla f(x, y) = (-2x, 2y)$ per cui $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$.

Il teorema della funzione implicita può essere utilizzato per mostrare come le soluzioni di un'equazione dipendano con regolarità dai coefficienti, come mostra il seguente esempio.

Esempio 2.55. Dato $a \in \mathbb{R}$, si consideri l'equazione $x^3 + ax^2 + x - 1 = 0$. Se $a = -1$, $x^3 - x^2 + x - 1 = (x^2 + 1)(x - 1)$ per cui l'unica soluzione reale dell'equazione è $x = 1$. Posto $f(x, a) = x^3 + ax^2 + x - 1$, $\nabla f(x, a) = (3x^2 + 2ax + 1, x^2)$, poiché $f(1, -1) = 0$ e $\frac{\partial f}{\partial x}(1, -1) = 2 \neq 0$, il teorema della funzione implicita assicura che, per ogni a sufficientemente vicino a -1 , esiste unico $x_a = \psi(a)$, sufficientemente vicino a 1, soluzione di $x^3 + ax^2 + x - 1 = 0$. Inoltre, poiché $\psi'(-1) = -\frac{1}{2} < 0$, per $a > -1$ si ha $x_a < 1$ e per $a < -1$ si ha $x_a > 1$.

8.3. Massimi e minimi vincolati. Il seguente teorema dà una condizione necessaria per l'esistenza di estremi relativi di una funzione f su un insieme C espresso come insieme di livello di una funzione g .

Teo 2.56 (teorema dei moltiplicatori di Lagrange). *Data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe C^1 definita su un insieme aperto A aperto ed un insieme*

$$C = \{(x, y) \in A \mid g(x, y) = 0\},$$

dove $g : A \rightarrow \mathbb{R}$ è una funzione di classe C^1 , se un punto $(x_0, y_0) \in C$ soddisfa le seguenti condizioni:

a) *esiste $\delta > 0$ tale che per ogni $(x, y) \in C \cap B((x_0, y_0), \delta)$*

$$f(x, y) \geq f(x_0, y_0) \quad [\text{oppure } f(x, y) \leq f(x_0, y_0)];$$

b) *il gradiente $\nabla g(x_0, y_0) \neq 0$,*

allora

$$\det \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \\ \frac{\partial g}{\partial x}(x_0, y_0) & \frac{\partial g}{\partial y}(x_0, y_0) \end{bmatrix} = 0. \quad (76)$$

Dimostrazione. Supponiamo, ad esempio, che $\frac{\partial g}{\partial y}(x_0, y_0) \neq 0$, per il teorema della funzione implicita esiste $\varphi : I \rightarrow \mathbb{R}$, $y = \varphi(x)$, di classe C^1 tale che $\varphi(x_0) = y_0$ e

$$C \cap V = \{(x, \varphi(x)) \in \mathbb{R}^2 \mid x \in I\},$$

dove I è un intorno aperto di x_0 e V un intorno aperto di (x_0, y_0) . Allora x_0 è un punto (interno) di estremo relativo per la funzione $f(x, \varphi(x))$ definita su I e di classe C^1 e, quindi la derivata prima $f(x, \varphi(x))$ in $x = x_0$ è zero. Dalla regola di derivazione in catena, si ha che

$$\begin{aligned} f(x_0, \varphi(x_0))' &= \frac{\partial f}{\partial x}(x_0, \varphi(x_0)) + \frac{\partial f}{\partial y}(x_0, \varphi(x_0))\varphi'(x_0) \\ &= \frac{\partial f}{\partial x}(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0) \frac{\frac{\partial g}{\partial x}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)}, \end{aligned}$$

dove per la (72c)

$$\varphi'(x_0) = -\frac{\frac{\partial g}{\partial x}(x_0, \varphi(x_0))}{\frac{\partial g}{\partial y}(x_0, \varphi(x_0))}.$$

La condizione $f(x_0, \varphi(x_0))' = 0$ implica che

$$\frac{\partial f}{\partial x}(x_0, y_0) \frac{\partial g}{\partial y}(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0) \frac{\partial g}{\partial x}(x_0, y_0) = \det \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \\ \frac{\partial g}{\partial x}(x_0, y_0) & \frac{\partial g}{\partial y}(x_0, y_0) \end{bmatrix} = 0.$$

□

La condizione a) esprime il fatto che P_0 è un punto di minimo (risp. massimo) relativo per la funzione f vincolata su C .

◈. Il teorema assicura che i punti di estremo relativo vincolato sono da cercare tra i punti $(x, y) \in A$ che sono soluzione del sistema

$$\begin{cases} \det \begin{bmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \\ \frac{\partial g}{\partial x}(x, y) & \frac{\partial g}{\partial y}(x, y) \end{bmatrix} = 0 \\ g(x, y) = 0 \end{cases}.$$

La condizione (76) è solo necessaria, quindi in generale ci sono soluzione del sistema che non sono estremi relativi vincolati.

Poiché $\nabla g(x_0, y_0) \neq 0$, la condizione (76)

$$\det \begin{bmatrix} \nabla f(x_0, f_0) \\ \nabla g(x_0, f_0) \end{bmatrix} = 0$$

equivale al fatto che esiste $\lambda \in \mathbb{R}$ tale che

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0), \quad (77)$$

in cui λ prende il nome di moltiplicatore di Lagrange. La (77) esprime il fatto che $\nabla f(x_0, y_0)$ è parallelo a $\nabla g(x_0, y_0)$, cioè perpendicolare alla retta tangente a C in (x_0, y_0) .

Il seguente esempio mostra come calcolare gli autovalori di una matrice simmetrica.

Esempio 2.57. Data una matrice simmetrica $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$, la corrispondente forma quadratica $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x, y) = ax^2 + 2bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

ammette minimo $f(x_1, y_1)$ e massimo $f(x_2, y_2)$ assoluti sull'insieme

$$C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\},$$

poiché f è continua e C è chiuso e limitato. Il vincolo C è l'insieme di livello di quota 0 della funzione $g(x, y) = x^2 + y^2 - 1$. Le funzioni f e g sono di classe $C^1(\mathbb{R}^2)$ e

$$\nabla f(x, y) = 2(ax + by, bx + cy) = 2 \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\nabla g(x, y) = 2(x, y).$$

Poiché $\nabla g(x, y) \neq 0$ per ogni $(x, y) \in C$, il teorema dei moltiplicatori di Lagrange implica che i punti di minimo e massimo (x_1, y_1) e (x_2, y_2) sono soluzioni rispettivamente di

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \lambda_1 \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \lambda_2 \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

con la condizione $x_1^2 + y_1^2 = x_2^2 + y_2^2 = 1$. In altre parole, (x_1, y_1) e (x_2, y_2) sono gli autovettori di norma 1 della matrice simmetrica $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$. I corrispondenti moltiplicatori di Lagrange sono i rispettivi autovalori λ_1 e λ_2 . Inoltre è immediato verificare che

$$f(x_1, y_1) = \lambda_1 \quad \text{e} \quad f(x_2, y_2) = \lambda_2,$$

da cui segue che $\lambda_1 \leq f(x, y) \leq \lambda_2$ per ogni $(x, y) \in C$. Infine, osservando che $f(\alpha x, \alpha y) = \alpha^2 f(x, y)$ per ogni $\alpha \in \mathbb{R}$, segue che

$$\lambda_1(x^2 + y^2) \leq \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \lambda_2(x^2 + y^2) \quad \forall (x, y) \in \mathbb{R}^2.$$

Il seguente esempio mostra come calcolare i vertici di un'ellisse.

Esempio 2.58. La funzione $f(x, y) = x^2 + y^2$ ammette massimo e minimo assoluti sull'insieme

$$C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 + xy = 1\},$$

poiché f è continua e C è chiuso e limitato, dove la limitatezza segue dal fatto che

$$\frac{x^2 + y^2}{2} \leq x^2 + y^2 + xy$$

per cui $C \subset B(O, 2)$. Posto $g(x, y) = x^2 + y^2 + xy - 1$, f e g sono di classe $C^1(\mathbb{R}^2)$ e

$$\nabla f(x, y) = (2x, 2y)$$

$$\nabla g(x, y) = (2x + y, 2y + x) \neq 0 \quad \forall (x, y) \in C.$$

Il teorema dei moltiplicatori di Lagrange implica che i punti di minimo e massimo sono necessariamente soluzioni del sistema

$$\begin{cases} \det \begin{bmatrix} 2x & 2y \\ 2x + y & 2y + x \end{bmatrix} = 0 \\ x^2 + y^2 + xy = 1 \end{cases} \iff \begin{cases} x^2 - y^2 = 0 \\ x^2 + y^2 + xy = 1 \end{cases},$$

che ha come soluzioni

$$\begin{cases} P_1 = \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right) \\ P_2 = \left(-\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}\right) \end{cases} \quad f(P_1) = f(P_2) = \frac{2}{3}$$

$$\begin{cases} P_3 = (1, -1) \\ P_4 = (-1, 1) \end{cases} \quad f(P_3) = f(P_4) = 2.$$

Dal punto di vista geometrico P_1 e P_2 sono i punti di C che hanno distanza minore dall'origine, per cui sono i vertici dell'ellisse che giacciono sull'asse minore, mentre P_3 e P_4 sono i punti di C che hanno distanza maggiore dall'origine, per cui sono i vertici che giacciono sull'asse maggiore. Ne segue che C è un'ellisse con assi posti lungo le rette $y = -x$ e $y = x$.

Il teorema dei moltiplicatori di Lagrange può essere esteso al caso di m vincoli in \mathbb{R}^n , purché $m < n$.

Teo 2.59. *Data una funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^1 definita su un aperto A aperto ed un insieme*

$$C = \{P \in A \mid g_1(P) = \dots = g_m(P) = 0\},$$

dove $g_1, \dots, g_m : A \rightarrow \mathbb{R}$ sono di classe C^1 e $m < n$, se $P_0 \in C$ punto di estremo relativo di f vincolata su C ed i vettori $\nabla g_1(P_0), \dots, \nabla g_m(P_0)$ sono linearmente indipendenti, allora esistono unici $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ che soddisfano

$$\nabla f(P_0) = \sum_{i=1}^m \lambda_i \nabla g_i(P_0). \quad (78)$$