Capitolo Dodicesimo

CALCOLO DIFFERENZIALE PER FUNZIONI DI PIÙ VARIABILI

§ 1. CAMPI SCALARI

Sono dati: un insieme *aperto* $A \subset \mathbb{R}^n$, un punto $\underline{x}^0 = (x_1^0, x_2^0, \dots, x_n^0)^T \in A$ e una funzione $f: A \to \mathbb{R}$. Si pone allora il

PROBLEMA. Come si può estendere al nuovo contesto la nozione di derivata, in modo da ritrovare, nel caso $n = 1$, quella usuale e da far salve le importanti conseguenze che da essa abbiamo a suo tempo dedotto (quale la Formula di Taylor)?

Si vede subito che non possiamo riscrivere pari pari la vecchia definizione, dato che la scrittura $\dfrac{f(\underline{x}) - f(\underline{x}^0)}{\underline{x} - \underline{x}^0}$, avendo a denominatore un vettore, non ha alcun significato. Dobbiamo dunque seguire un'altra strada. Una possibilità è quella di considerare le restrizioni della f a sottoinsiemi di A formati da rette o segmenti per $\underline{x}^0$, ottenendo così funzioni di una variabile. Vediamo di essere un po' più precisi.

Derivate direzionali

Fissiamo un $\underline{v} \in \mathbb{R}^n$, con $\|\underline{v}\| = 1$, detto *versore* o *direzione orientata* e consideriamo un *segmento* del tipo $\{\underline{x} : \underline{x} = \underline{x}^0 + t\underline{v}$, con $t \in\ ]-\delta, \delta[, \delta > 0\}$ che sia contenuto in A. Un tale segmento esiste, dato che A è aperto.

DEFINIZIONE. Se la funzione $g :]-\delta, \delta[\ (\subset \mathbb{R}) \to \mathbb{R}$ definita da $g(t) = f(\underline{x}^0 + t\underline{v})$ è derivabile in $t = 0$, si dice che f *ammette derivata* (o che è *derivabile*) *in* $\underline{x}^0$ *secondo la direzione* $\underline{v}$ e si scrive

$$\frac{\partial f}{\partial \underline{v}}(\underline{x}^0) := g'(0) = \lim_{t\to 0} \frac{f(\underline{x}^0 + t\underline{v}) - f(\underline{x}^0)}{t}.$$

Si tenga presente che, per ipotesi, la derivata della g esiste *finita*.

Caso particolare: le derivate parziali

Sia $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ la base canonica di $\mathbb{R}^n$.

DEFINIZIONE. La derivata direzionale calcolata nella direzione di uno dei versori $\underline{e}_i$ prende il nome di *derivata parziale (prima) calcolata rispetto alla variabile* x_i; è dunque:

$$\frac{\partial f}{\partial \underline{e}_i}(\underline{x}^0) := \lim_{t\to 0} \frac{f(x_1^0, x_2^0, \dots, x_i^0 + t, \dots, x_n^0) - f(\underline{x}^0)}{t}.$$

La derivata $\dfrac{\partial f}{\partial \underline{e}_i}(\underline{x}^0)$ è spesso indicata con $\dfrac{\partial f}{\partial x_i}(\underline{x}^0)$ o con $f_{x_i}(\underline{x}^0)$.

Il calcolo delle derivate parziali è facile, in quanto basta considerare la f come funzione di una sola variabile, riguardando le altre come costanti, e utilizzare le ben note regole di derivazione. Quello delle derivate direzionali generiche è leggermente meno immediato.

ESEMPI. 1) Le derivate parziali della funzione $f : \mathbb{R}^2 \setminus \{\underline{0}\} \to \mathbb{R}$, con $f(x,y) = \dfrac{xy}{x^2 + y^2}$, sono:

$$\frac{\partial f}{\partial x}(\underline{x}) = \frac{y(x^2 + y^2) - 2x^2y}{(x^2 + y^2)^2} = \frac{y(y^2 - x^2)}{(x^2 + y^2)^2} \quad ; \qquad \frac{\partial f}{\partial y}(\underline{x}) = \frac{x(x^2 + y^2) - 2xy^2}{(x^2 + y^2)^2} = \frac{x(x^2 - y^2)}{(x^2 + y^2)^2}.$$

2) Le derivate parziali delle funzioni, di $\mathbb{R}^3$ in $\mathbb{R}$, $f(x,y,z) = x^2ye^z$ e $g(x,y,z) = x^2 + |y| + z$ sono:

$$\frac{\partial f}{\partial x}(\underline{x}) = 2xye^z; \quad \frac{\partial f}{\partial y}(\underline{x}) = x^2e^z; \quad \frac{\partial f}{\partial z}(\underline{x}) = x^2ye^z;$$

$$\frac{\partial g}{\partial x}(\underline{x}) = 2x \ ; \quad \frac{\partial g}{\partial y}(\underline{x}) = \frac{|y|}{y} \ ; \quad \frac{\partial g}{\partial z}(\underline{x}) = 1.$$

[La $\dfrac{\partial g}{\partial y}$ è, ovviamente, definita solo nei punti $(x,y,z)^T$ per cui è $y \neq 0$.]

3) Si vuole calcolare la derivata della funzione (di $\mathbb{R}^2$ in $\mathbb{R}$) $f(x,y) = x^2 + y^2$ nel punto $(x, y)^T$ secondo la direzione del versore $\underline{v} = (a, b)^T$; si vuole cioè la derivata della funzione $g(t) = (x + at)^2 + (y + bt)^2$ nel punto $t = 0$. Si ha: $g'(t) = 2a(x + at) + 2b(y + bt)$, da cui

$$\frac{\partial f}{\partial \underline{v}}(\underline{x}) = g'(0) = 2ax + 2by.$$

4) Si vuole calcolare la derivata della funzione (di $\mathbb{R}^2$ in $\mathbb{R}$) $f(x,y) = \sin(xy)$ nel punto $(x, y)^T$ secondo la direzione del versore $\underline{v} = (a, b)^T$; si ottiene la funzione

$$g(t) = \sin[(x + at)(y + bt)] = \sin[xy + (bx + ay)t + abt^2],$$

la cui derivata è: $g'(t) = (bx + ay + 2abt)\cos[xy + (bx + ay)t + abt^2]$; è dunque:

$$\frac{\partial f}{\partial \underline{v}}(\underline{x}) = g'(0) = (bx + ay)\cos(xy).$$

DEFINIZIONE. Se $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ è dotata di derivata parziale (finita) rispetto alla variabile x_i in un punto $\underline{x} \in A$, diremo che f è *derivabile* in quel punto rispetto a tale variabile. Se la f è derivabile rispetto a x_i in ogni punto di A, diremo che essa è *derivabile in A* rispetto a x_i.

Derivate seconde e derivate di ordine superiore

Se $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ è derivabile rispetto a x_i in ogni punto di A, si costruisce una funzione $\dfrac{\partial f}{\partial x_i} : A(\subset \mathbb{R}^n) \to \mathbb{R}$, [o $f_{x_i} : A(\subset \mathbb{R}^n) \to \mathbb{R}$].

DEFINIZIONE. Se la funzione $\dfrac{\partial f}{\partial x_i}$ è derivabile rispetto a x_j in $\underline{x}^0$, si pone

$$\frac{\partial^2 f}{\partial x_i\, \partial x_j}(\underline{x}^0) = f_{x_i x_j}(\underline{x}^0) := \left(\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)\right)(\underline{x}^0).$$

A questo numero si dà il nome di *derivata seconda della f in* $\underline{x}^0$ *rispetto a* x_i *e* x_j (nell'ordine).

Anche in questo caso, se questa derivata seconda è definita per ogni punto $\underline{x}$ di A, si ottiene una funzione $f_{x_i x_j} : A(\subset \mathbb{R}^n) \to \mathbb{R}$ detta *derivata seconda della f rispetto a* x_i *e* x_j .

Quest'ultima funzione può, a sua volta, essere derivabile rispetto a x_k e si parlerà di derivata terza e così via. Le derivate ottenute con m derivazioni successive sono dette derivate di *ordine m.*

Le derivate successive fatte sempre rispetto alla stessa variabile sono dette *pure,* mentre le altre sono dette *miste*.

ESEMPI. 5) Le derivate parziali seconde della funzione $f(x,y) = \dfrac{xy}{x^2 + y^2}$ sono:

$$f_{xx}(\underline{x}) = \frac{-2xy(3y^2 - x^2)}{(x^2 + y^2)^3} ; \quad f_{yy}(\underline{x}) = \frac{-2xy(3x^2 - y^2)}{(x^2 + y^2)^3} ; \quad f_{xy}(\underline{x}) = f_{yx}(\underline{x}) = \frac{6x^2y^2 - x^4 - y^4}{(x^2 + y^2)^3} .$$

6) Le derivate parziali seconde della funzione $f(x,y,z) = x^2ye^z$ sono:

$$f_{xx}(\underline{x}) = 2ye^z ;\ f_{yy}(\underline{x}) = 0; \quad f_{zz}(\underline{x}) = x^2ye^z ;$$
$$f_{xy}(\underline{x}) = f_{yx}(\underline{x}) = 2xe^z ; \quad f_{xz}(\underline{x}) = f_{zx}(\underline{x}) = 2xye^z ; \quad f_{yz}(\underline{x}) = f_{zy}(\underline{x}) = x^2e^z .$$

7) Le derivate seconde miste della funzione $f(x,y,z) = x \log x + y \log y + z \log z$ sono tutte nulle; quelle pure sono:

$$f_{xx}(\underline{x}) = \frac{1}{x} ;\ f_{yy}(\underline{x}) = \frac{1}{y} ;\ f_{zz}(\underline{x}) = \frac{1}{z} .$$

DEFINIZIONE. Sia data una funzione $f : A \to \mathbb{R}$, con A sottoinsieme aperto di $\mathbb{R}^n$. Se f ha in A tutte le derivate fino all'ordine k e queste sono continue, si dice che la f è di *classe* C^k in A [$f \in C^k(A)$]; se ciò vale per ogni k, f è detta di *classe* C^∞ in A [$f \in C^\infty(A)$].

Si dice inoltre che una funzione $f : cl\ A \to \mathbb{R}$, con A sottoinsieme aperto di $\mathbb{R}^n$, è di *classe* C^k in *cl A* se f è di *classe* C^k in A e tutte le sue derivate parziali, fino all'ordine k, sono prolungabili per continuità su *cl A*.

Si constata subito che, in tutti gli esempi sopra prodotti, le derivate seconde miste che differiscono solo per l'ordine con cui si effettuano le derivazioni sono fra loro uguali. È dunque naturale chiedersi se ciò accade sempre o, eventualmente, sotto quali condizioni. Ebbene, esistono funzioni con le derivate seconde miste diverse.

ESEMPIO. 8) Sia f: $\mathbb{R}^2 \to \mathbb{R}$ definita da:

$$f(x,y) = \begin{cases} \dfrac{xy^3}{x^2 + y^2} & \text{se è } \underline{x} \neq \underline{0} \\ 0 & \text{se è } \underline{x} = \underline{0} \end{cases} .$$

Si ha:

$$f_x(\underline{x}) = \begin{cases} \dfrac{y^3(y^2 - x^2)}{(x^2 + y^2)^2} & \text{se è } \underline{x} \neq \underline{0} \\ 0 & \text{se è } \underline{x} = \underline{0} \end{cases} ; \quad f_y(\underline{x}) = \begin{cases} \dfrac{xy^2(3x^2 + y^2)}{(x^2 + y^2)^2} & \text{se è } \underline{x} \neq \underline{0} \\ 0 & \text{se è } \underline{x} = \underline{0} \end{cases} ;$$

da cui si ottiene:

$$f_{xy}(0,0) = \lim_{y\to 0} \frac{f_x(0,y) - f_x(0,0)}{y} = \lim_{y\to 0} \frac{y^5}{y^5} = 1;$$
$$f_{yx}(0,0) = \lim_{x\to 0} \frac{f_y(x,0) - f_y(0,0)}{x} = \lim_{x\to 0} \frac{0}{x} = 0.$$

Al riguardo sussiste il seguente Teorema di cui omettiamo la dimostrazione.

TEOREMA 1 *(di Schwarz) - Se la funzione $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ è dotata in un intorno U di un punto $\underline{x}^0$ delle derivate seconde miste $f_{x_ix_j}$ e $f_{x_jx_i}$ e queste sono continue in $\underline{x}^0$, allora si ha $f_{x_ix_j}(\underline{x}^0) = f_{x_jx_i}(\underline{x}^0)$.* ∎

Il Teorema si estende anche alle derivate di ordine superiore; in particolare, si ha che:

Se la funzione $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ è di classe C^k in A, allora le derivate miste, di ordine minore o uguale a k, che differiscono solo per l'ordine di derivazione coincidono.

Torneremo più avanti su questo argomento (Cfr. § 3).

Sappiamo che, per le funzioni di una variabile, la derivabilità in un punto implica la continuità nel punto stesso. Sussiste un'analoga proprietà anche per le funzioni di più variabili? La risposta è negativa. Esistono cioè funzioni dotate di derivate parziali in $\underline{x}^0$ e che, tuttavia, non sono continue in tale punto.

ESEMPIO. 9) Sia f: $\mathbb{R}^2 \to \mathbb{R}$ definita da:

$$f(x,y) = \begin{cases} 1 & \text{se è } xy = 0 \\ 0 & \text{se è } xy \neq 0 \end{cases}.$$

Si vede subito che la f non è continua in $\underline{0}$, pur essendo $f_x(0,0) = f_y(0,0) = 0$.

Non solo, ma può accadere che una funzione sia dotata, in un punto $\underline{x}^0$, di derivate in tutte le direzioni, senza essere continua nel punto.

ESEMPIO. 10) Sia f: $\mathbb{R}^2 \to \mathbb{R}$ definita da:

$$f(x,y) = \begin{cases} \dfrac{x^2y}{x^4 + y^2} & \text{se è } \underline{x} \neq \underline{0} \\ 0 & \text{se è } \underline{x} = \underline{0} \end{cases}.$$

Si ha, intanto, $f_x(\underline{0}) = 0$. Dato poi il versore $\underline{v} = (a,b)^T$, con $b \neq 0$, si ha:

$$\frac{\partial f}{\partial \underline{v}}(\underline{0}) = \lim_{t\to 0} \frac{f(0 + at, 0 + bt) - f(\underline{0})}{t} = \lim_{t\to 0} \frac{a^2bt^3}{b^2t^3 + a^4t^5} = \frac{a^2}{b};$$

In $\underline{0}$ esistono dunque tutte le derivate direzionali. D'altra parte, se consideriamo la restrizione della f all'insieme $E = \{(x,x^2)^T \colon x \neq 0\}$, si vede subito che questa vale costantemente $\frac{1}{2} \neq 0 = f(\underline{0})$; pertanto la nostra funzione non è continua in $\underline{0}$.

Siamo perciò costretti a concludere che *la nozione di derivata parziale o direzionale non è la naturale estensione al caso delle funzioni di più variabili della nozione di derivata vista per le funzioni di una sola variabile.* Dobbiamo cercare un'altra strada.

Sappiamo che, per le funzioni di una variabile reale, la derivabilità in un punto x_0 equivale all'esistenza in x_0 dell'approssimante lineare. Quest'ultima nozione si estende in modo naturale al nuovo contesto.

DEFINIZIONE. Siano: A un sottoinsieme aperto di $\mathbb{R}^n$, $\underline{x}^0$ un prefissato punto di A e f una funzione di A in $\mathbb{R}$. Una funzione $\overline{f}(\underline{x}) = L(\underline{x} - \underline{x}^0) + q$, con $L \in \mathcal{L}(\mathbb{R}^n,\mathbb{R})$, $q \in \mathbb{R}$, è detta *approssimante lineare di f in $\underline{x}^0$* se

1) $\overline{f}(\underline{x}^0) = f(\underline{x}^0)$;
2) $f(\underline{x}) = \overline{f}(\underline{x}) + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|$, con $\lim\limits_{\underline{x}\to\underline{x}^0} \varepsilon(\underline{x}) = 0$.

Se una siffatta funzione $\overline{f}$ esiste, si ha:

$$(*) \qquad f(\underline{x}) = f(\underline{x}^0) + L(\underline{x} - \underline{x}^0) + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|, \text{ con } \varepsilon(\underline{x}) \to 0. \text{ se } \underline{x} \to \underline{x}^0.$$

DEFINIZIONE. La forma lineare L che compare nella (*) prende il nome di *differenziale di f in* $\underline{x}^0$. Per esprimere il fatto che la forma lineare L è il differenziale della funzione f *relativamente al punto* $\underline{x}^0$, si scrive $L = (df)(\underline{x}^0)$ o $L = df(\underline{x}^0)$.

DEFINIZIONE. Siano: A un sottoinsieme aperto di $\mathbb{R}^n$, $\underline{x}^0$ un prefissato punto di A e f una funzione di A in $\mathbb{R}$. Se la f ammette approssimante lineare in $\underline{x}^0$, si dice che la f è *differenziabile* in questo punto. Se la f è differenziabile in ogni punto di A, si dice che f è *differenziabile in A*.

OSSERVAZIONE. Ricordiamo che, come visto alla fine del § 3 del Capitolo 11, la matrice associata ad una forma lineare L di $\mathbb{R}^n$ è una matrice M a *una* riga e n colonne: è cioè $M = (a_1, a_2, \ldots, a_n)$. Dunque, per ogni $\underline{x} \in \mathbb{R}^n$, si ha:

$$(*) \qquad L(\underline{x}) = M\underline{x} = a_1x_1 + a_2x_2 + \ldots + a_nx_n = \langle \underline{a}, \underline{x} \rangle,$$

essendo $\underline{a} := (a_1, a_2, \ldots, a_n)^T$.
Notiamo che il differenziale della f, cioè la forma lineare L (o, equivalentemente, il vettore $\underline{a}$ che la individua) varia al variare dal punto $\underline{x}^0$. (Cfr. Teorema 3.)

TEOREMA 2. *Se* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *è differenziabile in un punto* $\underline{x}^0 \in A$, *allora f è continua in* $\underline{x}^0$.

DM. Se la f è differenziabile in $\underline{x}^0$, si ha $f(\underline{x}) = \overline{f}(\underline{x}) + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|$, che tende a $f(\underline{x}^0)$ al tendere di $\underline{x}$ a $\underline{x}^0$. ∎

TEOREMA 3. *Se* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *è differenziabile in un punto* $\underline{x}^0 \in A$, *allora f ha in* $\underline{x}^0$ *tutte le derivate direzionali e si ha* $\dfrac{\partial f}{\partial \underline{v}}(\underline{x}^0) = L(\underline{v})$.

DIM. Sia f differenziabile in $\underline{x}^0$. Qualunque sia il versore $\underline{v}$, si ha:

$$\frac{\partial f}{\partial \underline{v}}(\underline{x}^0) = \lim_{t\to 0} \frac{f(\underline{x}^0 + t\underline{v}) - f(\underline{x}^0)}{t} = \lim_{t\to 0} \frac{f(\underline{x}^0) + L(t\underline{v}) + \varepsilon(t)\cdot|t| - f(\underline{x}^0)}{t} =$$

$$= \lim_{t\to 0}\left(\frac{tL(\underline{v})}{t} + \frac{\varepsilon(t)|t|}{t}\right) = L(\underline{v}). \;∎$$

COROLLARIO 4. *Se* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *è differenziabile in un punto* $\underline{x}^0 \in A$, *allora f ha in* $\underline{x}^0$ *tutte le derivate parziali e si ha:*

$$\frac{\partial f}{\partial x_i}(\underline{x}^0) = L(\underline{e}_i) = \langle \underline{a}, \underline{e}_i \rangle = a_i. \;∎$$

COROLLARIO 5. *Se il differenziale di una funzione* $f : A\ (\subset \mathbb{R}^n) \to \mathbb{R}$ *esiste in un punto* $\underline{x}^0 \in A$, *allora esso è unico. Se è* $\underline{u} = (u_1, u_1, \dots, u_n)^T$, *si ha:*

$$L(\underline{u}) = (df(\underline{x}^0))(\underline{u}) = \frac{\partial f}{\partial x_1}(\underline{x}^0)u_1 + \frac{\partial f}{\partial x_2}(\underline{x}^0)u_2 + \dots + \frac{\partial f}{\partial x_n}(\underline{x}^0)u_n. \blacksquare$$

DEFINIZIONE. Se $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ è differenziabile in un punto $\underline{x}^0 \in A$, il vettore

$$\nabla f(\underline{x}^0) := \left(\frac{\partial f}{\partial x_1}(\underline{x}^0), \frac{\partial f}{\partial x_2}(\underline{x}^0), \dots, \frac{\partial f}{\partial x_n}(\underline{x}^0) \right)^T$$

è detto il *gradiente* di f in $\underline{x}^0$.

In base a tale definizione, si ha che, se $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ è una funzione differenziabile in un punto $\underline{x}^0 \in A$, allora:

$$(df(\underline{x}^0))(\underline{u}) = L(\underline{u}) = \langle \nabla f(\underline{x}^0), \underline{u} \rangle;$$

$$\frac{\partial f}{\partial \underline{v}}(\underline{x}^0) = L(\underline{v}) = \langle \nabla f(\underline{x}^0), \underline{v} \rangle, \quad \text{se è } \|\underline{v}\| = 1;$$

$$f(\underline{x}) = f(\underline{x}^0) + \langle \nabla f(\underline{x}^0), \underline{x} - \underline{x}^0 \rangle + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|, \text{ con } \lim_{\underline{x}\to\underline{x}^0} \varepsilon(\underline{x}) = 0.$$

ESEMPIO. 11) Consideriamo la funzione $f: \mathbb{R}^2 \to \mathbb{R}$ definita da $f(x,y) = e^x \cos y$. Vedremo tra poco che una funzione come questa è sicuramente differenziabile in ogni punto del suo dominio. Ammesso ciò, vediamo di calcolare il suo gradiente in un punto $\underline{x} = (x,y)^T$ e la derivata direzionale in tale punto secondo il versore $\underline{v} = (a,b)^T$. Si ha:

$$\nabla f(\underline{x}) = (e^x \cos y, -e^x \sin y)^T; \quad \frac{\partial f}{\partial \underline{v}}(\underline{x}) = \langle \nabla f(\underline{x}), \underline{v} \rangle = ae^x \cos y - be^x \sin y.$$

N.B. Non si confondano le notazioni

$$\left(\frac{\partial f}{\partial x_1}(\underline{x}^0), \frac{\partial f}{\partial x_2}(\underline{x}^0), \dots, \frac{\partial f}{\partial x_n}(\underline{x}^0) \right) \text{ e } \left(\frac{\partial f}{\partial x_1}(\underline{x}^0), \frac{\partial f}{\partial x_2}(\underline{x}^0), \dots, \frac{\partial f}{\partial x_n}(\underline{x}^0) \right)^T.$$

La prima è la *matrice* (a *una* riga e n colonne) associata alla forma lineare L del differenziale della f in $\underline{x}^0$; la seconda è il *vettore colonna* (matrice a n righe e *una* colonna) che è detto il *gradiente* della f in $\underline{x}^0$.

Può essere utile tener presente la seguente definizione che esprime l'interpretazione geometrica dell'approssimante lineare di una funzione di $\mathbb{R}^2$ in $\mathbb{R}$ in un punto $\underline{x}^0$ del suo dominio.

DEFINIZIONE. Siano $f : A(\subset \mathbb{R}^2) \to \mathbb{R}$ una funzione differenziabile in un punto $\underline{x}^0 \in A$ e $\overline{f}(\underline{x})$ il suo approssimante lineare in $\underline{x}^0$. La *superficie* di equazione

$$z = \overline{f}(\underline{x}) = f(\underline{x}^0) + \langle \nabla f(\underline{x}^0), \underline{x} - \underline{x}^0 \rangle$$

è detta *piano tangente* alla superficie di equazione $z = f(\underline{x})$ nel punto $P_0(\underline{x}^0, f(\underline{x}^0))$.

Si dimostra che questo piano contiene le rette per P_0 e tangenti alle curve di equazioni

$$\begin{cases} x = t \\ y = y_0 \\ z = f(t,y_0) \end{cases} \quad \text{e} \quad \begin{cases} x = x_0 \\ y = t \\ z = f(x_0,t) \end{cases} .$$

TEOREMA 6. *(del differenziale totale) - Se la funzione* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$, *con A insieme aperto, è dotata in un intorno U di un punto* $\underline{x}^0 \in A$ *di derivate parziali prime e queste sono continue in* $\underline{x}^0$, *allora f è differenziabile in* $\underline{x}^0$.

DIM. Limitiamoci al caso di una funzione di due variabili. Dato $\underline{x}^0 = (x_0, y_0)^T \in A$, esiste una sfera $S(\underline{x}^0, r)$ contenuta in U. Se $\underline{x} = (x,y)^T$ è un arbitrario punto di S, la differenza $f(\underline{x}) - f(\underline{x}^0)$ può essere scritta nella forma

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

Tutti i punti della poligonale di vertici $\underline{x}^0 = (x_0, y_0)^T$, $\underline{x}^* = (x,y_0)^T$ e $\underline{x} = (x,y)^T$ appartengono ancora a S. La restrizione della f a ciascuno dei due segmenti di questa poligonale può essere vista come una funzione di una sola variabile che, per le nostre ipotesi, risulta derivabile. Si può quindi applicare in entrambe i casi il Teorema di Lagrange; si ottiene:

$$f(x,y) - f(x,y_0) = f_y(x,\eta)(y - y_0)$$

e

$$f(x,y_0) - f(x_0, y_0) = f_x(\xi,y_0)(x - x_0).$$

Essendo le funzioni $f_x(x,y)$ e $f_y(x,y)$ continue in $\underline{x}^0$, si ha

$$f_y(x,\eta) = f_y(x_0,y_0) + \varepsilon_1(\underline{x}) \ \text{ e } \ f_x(\xi,y_0) = f_x(x_0,y_0) + \varepsilon_2(\underline{x}),$$

con $\varepsilon_1(\underline{x})$ ed $\varepsilon_2(\underline{x})$ tendenti a zero al tendere di $\underline{x}$ a $\underline{x}^0$. Si ottiene:

$$f(x,y) - f(x,y_0) = [f_y(x_0,y_0) + \varepsilon_1(\underline{x})](y - y_0) \ \text{ e } \ f(x,y_0) - f(x_0,y_0) = [f_x(x_0,y_0) + \varepsilon_2(\underline{x})](x - x_0).$$

In conclusione, è:

$$f(\underline{x}) - f(\underline{x}^0) = f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) + \varepsilon_2(\underline{x})(x - x_0) + \varepsilon_1(\underline{x})(y - y_0).$$

Posto $\overline{f}(\underline{x}) = f(\underline{x}^0) + f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0)$, si ha:

$$\frac{|f(\underline{x}) - \overline{f}(\underline{x})|}{||\underline{x} - \underline{x}^0||} = \frac{|\varepsilon_2(\underline{x})(x - x_0) + \varepsilon_1(\underline{x})(y - y_0)|}{||\underline{x} - \underline{x}^0||} \leq$$

$$\leq |\varepsilon_2(\underline{x})| \frac{|x - x_0|}{||\underline{x} - \underline{x}^0||} + |\varepsilon_1(\underline{x})| \frac{|y - y_0|}{||\underline{x} - \underline{x}^0||} \leq |\varepsilon_2(x)| + |\varepsilon_1(\underline{x})|,$$

che tende a zero al tendere di $\underline{x}$ a $\underline{x}^0$. ∎

Notiamo che non sussiste l'implicazione opposta di quest'ultimo Teorema; può anzi accadere che una funzione sia differenziabile in un punto $\underline{x}^0$ senza che *nessuna* delle sue derivate parziali sia continua in tale punto. Un controesempio è fornito dalla funzione di $\mathbb{R}$ in $\mathbb{R}$ che vale 0 in 0 mentre vale $x^2 \sin\frac{1}{x}$ negli altri punti.

§ 2. CAMPI VETTORIALI

Il concetto di differenziale si estende in modo naturale anche ai campi vettoriali.

DEFINIZIONE. Data una funzione g: $A \to \mathbb{R}^m$, con A sottoinsieme aperto di $\mathbb{R}^n$, e fissato un punto $\underline{x}^0 \in A$, una funzione $\overline{g}(\underline{x}) = L(\underline{x} - \underline{x}^0) + \underline{q}$, con $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ e $\underline{q} \in \mathbb{R}^m$, è detta *approssimante lineare della g in* $\underline{x}^0$ se:

1) $\overline{g}(\underline{x}^0) = g(\underline{x}^0)$;

2) $g(\underline{x}) = \overline{g}(\underline{x}) + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|$, con $\lim\limits_{\underline{x}\to\underline{x}^0} \varepsilon(\underline{x}) = \underline{0} \in \mathbb{R}^m$.

Se una siffatta funzione $\overline{g}$ esiste, si ha:

(*) $$g(\underline{x}) = g(\underline{x}^0) + L(\underline{x} - \underline{x}^0) + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|, \text{ con } \varepsilon(\underline{x}) \to \underline{0}. \text{ se } \underline{x} \to \underline{x}^0.$$

DEFINIZIONE. Se g: $A \to \mathbb{R}^m$ è dotata di approssimante lineare in un punto $\underline{x}^0 \in A$, si dice che g è *differenziabile in* $\underline{x}^0$ e l'applicazione lineare L che compare nella (*) è detta il *differenziale della g in* $\underline{x}^0$.

Dunque, se la funzione g: $A(\subset \mathbb{R}^n) \to \mathbb{R}^m$ è differenziabile in un punto $\underline{x}^0 \in A$, indicata con $M \in \mathcal{M}(m, n)$ la matrice (a m righe e n colonne) associata all'applicazione lineare L, si ha:

$$g(\underline{x}) = g(\underline{x}^0) + M(\underline{x} - \underline{x}^0) + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|, \text{ con } \lim_{\underline{x}\to\underline{x}^0} \varepsilon(\underline{x}) = \underline{0} \in \mathbb{R}^m.$$

Se è: $$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

si ottiene:

$$g(\underline{x}) = g(x_1, x_2, \dots, x_n) = (g_1(\underline{x}), g_2(\underline{x}), \dots, g_m(\underline{x}))^T = \begin{pmatrix} g_1(x_1, x_2, \dots, x_n) \\ g_2(x_1, x_2, \dots, x_n) \\ \dots, \\ g_m(x_1, x_2, \dots, x_n) \end{pmatrix} =$$

$$= \begin{pmatrix} g_1(\underline{x}^0) \\ g_2(\underline{x}^0) \\ \dots, \\ g_m(\underline{x}^0) \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 - x_1^0 \\ x_2 - x_2^0 \\ \dots \\ x_n - x_n^0 \end{pmatrix} + \begin{pmatrix} \varepsilon_1(\underline{x})\cdot\|\underline{x} - \underline{x}^0\| \\ \varepsilon_2(\underline{x})\cdot\|\underline{x} - \underline{x}^0\| \\ \dots \\ \varepsilon_m(\underline{x})\cdot\|\underline{x} - \underline{x}^0\| \end{pmatrix}.$$

La componente i - ima della g è quindi espressa da:

$$g_i(x_1, x_2, \dots, x_n) = g_i(\underline{x}^0) + a_{i1}(x_1 - x_1^0) + a_{i2}(x_2 - x_2^0) + \dots + a_{in}(x_n - x_n^0) + \varepsilon_i(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|,$$

con $\lim\limits_{\underline{x}\to\underline{x}^0} \varepsilon_i(\underline{x}) = 0$, per $i = 1, 2, \dots, m$.

Da ciò segue immediatamente il

TEOREMA 7. *Una funzione* $g = (g_1, g_2, \dots, g_m)^T : A(\subset \mathbb{R}^n) \to \mathbb{R}^m$ *è differenziabile in un punto* $\underline{x}^0 \in A$ *se e solo se lo è ciascuna delle sue componenti* g_i. ∎

In virtù del Corollario 4 possiamo concludere col

COROLLARIO 8. *Se il differenziale di una funzione* $g = (g_1, g_2, \ldots, g_m)^T : A(\subset \mathbb{R}^n) \to \mathbb{R}^m$ *esiste in un punto* $\underline{x}^0 \in A$, *allora esso è unico e nella corrispondente matrice* $M = (a_{ij})$ *si ha:*

$$a_{ij} = \frac{\partial g_i}{\partial x_j}(\underline{x}^0). \blacksquare$$

DEFINIZIONE. Se la funzione g: $A(\subset \mathbb{R}^n) \to \mathbb{R}^m$ è differenziabile in un punto $\underline{x}^0 \in A$, la matrice M che definisce il differenziale prende il nome di *matrice jacobiana* della g in $\underline{x}^0$ e si indica con $(Jg)(\underline{x}^0)$. È dunque, per definizione,

$$(Jg)(\underline{x}^0) := \left(\frac{\partial g_i}{\partial x_j}(\underline{x}^0)\right)_{i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n}.$$

Casi particolari

$\boxed{n = m = 2}$ Sia $g = (g_1, g_2)^T$: $A(\subset \mathbb{R}^2) \to \mathbb{R}^2$. Si ha $g(x, y) = \begin{pmatrix} g_1(x,y) \\ g_2(x,y) \end{pmatrix}$ Se g è differenziabile in $\underline{x}^0 \in A$, si ha:

$$(Jg)(\underline{x}^0) = \begin{pmatrix} \frac{\partial g_1}{\partial x}(x_0,y_0) & \frac{\partial g_1}{\partial y}(x_0,y_0) \\ \frac{\partial g_2}{\partial x}(x_0,y_0) & \frac{\partial g_2}{\partial y}(x_0,y_0) \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{pmatrix}(\underline{x}^0).$$

Sia, per esempio, g: $\mathbb{R}^2 \to \mathbb{R}^2$ definita da $\begin{cases} g_1(\rho, \vartheta) = \rho \cos \vartheta \\ g_2(\rho, \vartheta) = \rho \sin \vartheta \end{cases}$.

Si ha: $$(Jg)(\rho, \vartheta) = \begin{pmatrix} \cos \vartheta & -\rho \sin \vartheta \\ \sin \vartheta & \rho \cos \vartheta \end{pmatrix}.$$

$\boxed{n = 1}$ Una funzione $g = (g_1, g_2, \ldots, g_m)^T$: $A(\subset \mathbb{R}) \to \mathbb{R}^m$ è differenziabile in un punto $x_0 \in A$ se e solo se ogni g_i è derivabile in x_0 ed è

$$(Jg)(x_0) = \begin{pmatrix} g'_1(x_0) \\ g'_2(x_0) \\ \ldots, \\ g'_m(x_0) \end{pmatrix} =: g'(x_0).$$

Differenziabilità della funzione composta

Il noto Teorema di derivazione delle funzioni composte è generalizzato dal seguente risultato di cui omettiamo la dimostrazione:

TEOREMA 9. *Siano date le funzioni* $g : B(\subset \mathbb{R}^p) \to A \subset \mathbb{R}^n$, *differenziabile in* $\underline{u}^0 \in B$, *e* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}^m$, *differenziabile in* $\underline{x}^0 = g(\underline{u}^0) \in A$, *allora la funzione composta* $h = f \circ g$: $B(\subset \mathbb{R}^p) \to \mathbb{R}^m$ *è differenziabile nel punto* $\underline{u}^0 \in B$, *e la sua matrice Jacobiana è data da*

$$(Jh)(\underline{u}^0) = (Jf)(\underline{x}^0)\,(Jg)(\underline{u}^0)\,,$$

dove il secondo membro è dato dal prodotto (righe per colonne) delle matrici Jacobiane della f *in* $\underline{x}^0$ *e della* g *in* $\underline{u}^0$. ▮

Si ha cioè:

$$\begin{pmatrix} \dfrac{\partial h_1}{\partial u_1}(\underline{u}^0) & \dots & \dfrac{\partial h_1}{\partial u_p}(\underline{u}^0) \\ \dots & \dots & \dots \\ \dfrac{\partial h_m}{\partial u_1}(\underline{u}^0) & \dots & \dfrac{\partial h_m}{\partial u_p}(\underline{u}^0) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\underline{x}^0) & \dots & \dfrac{\partial f_1}{\partial x_n}(\underline{x}^0) \\ \dots & \dots & \dots \\ \dfrac{\partial f_m}{\partial x_1}(\underline{x}^0) & \dots & \dfrac{\partial f_m}{\partial x_n}(\underline{x}^0) \end{pmatrix} \begin{pmatrix} \dfrac{\partial g_1}{\partial u_1}(\underline{u}^0) & \dots & \dfrac{\partial g_1}{\partial u_p}(\underline{u}^0) \\ \dots & \dots & \dots \\ \dfrac{\partial g_n}{\partial u_1}(\underline{u}^0) & \dots & \dfrac{\partial g_n}{\partial u_p}(\underline{u}^0) \end{pmatrix}$$

Caso particolare: $\boxed{\boldsymbol{m = p = 1}}$

Siano: g: $I(\subset \mathbb{R}) \to A \subset \mathbb{R}^n$, differenziabile in $u_0 \in I$, e $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$, differenziabile in $\underline{x}^0 = g(u_0) \in A$, allora la funzione composta $h = f \circ g$: $I(\subset \mathbb{R}) \to \mathbb{R}$ è derivabile nel punto $u_0 \in I$ e si ha:

$$h'(u_0) = (Jf)(\underline{x}^0) \cdot (Jg)(u_0) =$$

$$= \left(\frac{\partial f}{\partial x_1}(\underline{x}^0), \quad \dots, \quad \frac{\partial f}{\partial x_n}(\underline{x}^0) \right) \begin{pmatrix} g'_1(u_0) \\ \dots \\ g'_n(u_0) \end{pmatrix} = \langle \nabla f(\underline{x}^0), g'(u_0) \rangle .$$

ESEMPIO. 1) Siano g: $\mathbb{R} \to \mathbb{R}^3$ la funzione definita da $g(u) = \begin{pmatrix} u \\ u^2 \\ u^3 \end{pmatrix}$ e $f : \mathbb{R}^3 \to \mathbb{R}$, la funzione definita da $f(x,y,z) = e^x y \cos z$; per la funzione composta $h(u) = f \circ g(u)$ si ha:

$$h'(u) = \langle \nabla f(\underline{x}), g'(u) \rangle = (e^x y \cos z,\ e^x \cos z, - e^x y \sin z) \begin{pmatrix} 1 \\ 2u \\ 3u^2 \end{pmatrix} =$$

$$= 1 e^x y \cos z + 2u e^x \cos z - 3u^2 e^x y \sin z = e^u u^2 \cos u^3 + 2u e^u \cos u^3 - 3u^4 e^u \sin u^3 .$$

Come esercizio, si verifichi che, derivando $h(u) = e^u u^2 \cos u^3$, si ottiene lo stesso risultato.

Applicazione: La formula del valor medio

Siano $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ differenziabile, $\underline{x}^0 \in A$, e $S(\underline{x}^0, r)$ una sfera di centro $\underline{x}^0$ contenuta in A. Dato $\underline{x}^1 \in S$, sia $J = [\underline{x}^0, \underline{x}^1] = \{\underline{x}: \underline{x} = \underline{x}(t) = \underline{x}^0 + t(\underline{x}^1 - \underline{x}^0),\ t \in I = [0,1]\}$ il segmento di estremi $\underline{x}^0$ a $\underline{x}^1$. La funzione $F(t) = f(\underline{x}(t))$ è derivabile e si ha $F'(t) = \langle \nabla f(\underline{x}), \underline{x}'(t) \rangle$. Alla funzione $F(t)$ è applicabile su I il Teorema di Lagrange e si ha

$$F(1) - F(0) = F'(\tau)(1 - 0), \text{ con } 0 < \tau < 1.$$

Posto $\xi = \underline{x}(\tau)$, si ottiene il

TEOREMA 10 *(Formula del valor medio) - Siano* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *una funzione differenziabile su un aperto A,* $\underline{x}^0 \in A$, *e* $S(\underline{x}^0,r)$ *una sfera di centro* $\underline{x}^0$ *contenuta in A. Per ogni* $x \in S$, *esiste un punto* ξ *interno al segmento di estremi* $\underline{x}$ *e* $\underline{x}^0$, *per cui si ha*

$$f(\underline{x}) - f(\underline{x}^0) = \langle \nabla f(\xi), \underline{x} - \underline{x}^0 \rangle. \blacksquare$$

Sappiamo che, se una funzione di una variabile reale ha in un intervallo *I* la derivata identicamente nulla, allora essa è costante su *I*. Vediamo di studiare l'analogo problema per le funzioni di più variabili.

LEMMA 11. *Se la funzione* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *ha su un insieme aperto A le derivate parziali identicamente nulle, allora, per ogni* $\underline{x}^0 \in A$, *esiste una sfera di centro* $\underline{x}^0$ *in cui la f è costante.*

DIM. Poiché le derivate parziali della *f* sono continue, *f* è differenziabile in *A*, con $\nabla f \equiv \underline{0}$. Fissiamo un punto $\underline{x}^0 \in A$. Essendo *A* aperto, esiste una sfera $S(\underline{x}^0,r)$ contenuta in *A*. Per ogni $\underline{x}^1 \in S$, il segmento di equazione $\underline{x} = \underline{x}(t)$ che lo unisce a $\underline{x}^0$ è contenuto in *A*. La restrizione di *f* a questo segmento è una funzione di una variabile con derivata $\langle \nabla f(\underline{x}), x'(t) \rangle$ identicamente nulla ed è quindi costante, con valore $f(\underline{x}^0)$. $\blacksquare$

TEOREMA 12. *Se la funzione* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *ha su un insieme aperto e connesso A le derivate parziali identicamente nulle (che implica* $\nabla f \equiv \underline{0}$ *in A), allora la f è costante in A.*

DIM. Fissiamo ancora un punto $\underline{x}^0 \in A$. Per il Lemma 11, esiste una sfera $S(\underline{x}^0,r)$ contenuta in *A*. in cui la *f* è costante. Sia *A'* il sottoinsieme di *A* formato dai punti $\underline{x}$ per cui è $f(\underline{x}) = f(\underline{x}^0)$. Si ha, intanto, $S \subset A'$. Se è $\underline{x} \in A'$ $(\subset A)$, esiste, ancora per il Lemma 11, una sfera di centro $\underline{x}$ contenuta in *A'*; dunque *A'* è un sottoinsieme aperto di *A*. Sia ora $\underline{x}^1$ un generico punto di *A*. Essendo *A* connesso, esiste un'applicazione continua $\gamma : I = [0, 1] \to A$, con $\gamma(0) = \underline{x}^0$, $\gamma(1) = \underline{x}^1$ e $\gamma(I) \subset A$. Siano $t^* = \sup\{t \in I : \gamma(t) \in A'\}$ e $\underline{x}^* = \gamma(t^*) \in A$. Per il Lemma 11, esiste una sfera *S* di centro $\underline{x}^*$ contenuta in *A* in cui la *f* è costante. In *S* devono cadere punti di *A'*; si ottiene $f(\underline{x}) = f(\underline{x}^0)$, da cui $\underline{x} \in A'$. Se fosse $t^* < 1$; esisterebbero dei $t > t^*$ con $\gamma(t) \in A'$. Ma ciò andrebbe contro la definizione di t^*; si conclude che è $t^* = 1$ e che $\underline{x}^1 \in A'$, ossia $f(\underline{x}^1) = f(\underline{x}^0)$. $\blacksquare$

N.B. Può accadere che una funzione $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ differenziabile in un insieme aperto e connesso *A* abbia la derivata parziale f_{x_i} identicamente nulla in *A*, senza che la *f* sia costante rispetto a x_i.

ESEMPIO. Siano $A = \{(x, y)^T : y < 0\} \cup \{(x, y)^T : y \geq 0, |x| > 1\}$ e $f: A(\subset \mathbb{R}^2) \to \mathbb{R}$ definita da

$$f(x,y) = \begin{cases} 0 & \text{se è } y < 0 \\ y^2 & \text{se è } y \geq 0 \text{ e } x > 1 \\ -y^2 & \text{se è } y \geq 0 \text{ e } x < -1 \end{cases} .$$

Si ha $f_x(\underline{x}) \equiv 0$, pur essendo, per esempio, $f(2,1) = 1$ e $f(-2,1) = -1$.

§ 3. IL DIFFERENZIALE SECONDO PER I CAMPI SCALARI

Sia $f : I(\subset \mathbb{R}) \to \mathbb{R}$ una funzione derivabile; dunque, ad ogni $x \in I$ resta associato il numero reale $f'(x)$. Si ha così una nuova funzione f', sempre di I in $\mathbb{R}$. Se anche la f' è derivabile, la sua derivata è la *derivata seconda della f* su I. Come vanno le cose per le funzioni di più variabili?

Sia $f : A \to \mathbb{R}$ una funzione differenziabile sul sottoinsieme aperto A di $\mathbb{R}^n$. Ad ogni $\underline{x} \in A$ associamo il vettore $\nabla f(\underline{x}) \in \mathbb{R}^n$. Viene così definita una nuova funzione

$$g = \nabla f: A \to \mathbb{R}^n.$$

DEFINIZIONE. Se la funzione $g = \nabla f: A \to \mathbb{R}^n$ è differenziabile in un punto $\underline{x}^0 \in A$, si dice che f è *due volte differenziabile in* $\underline{x}^0$ e la matrice Jacobiana di g in $\underline{x}^0$ si chiama *matrice hessiana di f in* $\underline{x}^0$ e si indica con $(Hf)(\underline{x}^0)$.

È dunque, per definizione,

$$(Hf)(\underline{x}^0) := (Jg)(\underline{x}^0) = \begin{pmatrix} \dfrac{\partial g_1}{\partial x_1}(\underline{x}^0) & \dots & \dfrac{\partial g_1}{\partial x_n}(\underline{x}^0) \\ \dots & \dots & \dots \\ \dfrac{\partial g_n}{\partial x_1}(\underline{x}^0) & \dots & \dfrac{\partial g_n}{\partial x_n}(\underline{x}^0) \end{pmatrix}.$$

Essendo $g_1(\underline{x}) := \dfrac{\partial f}{\partial x_1}(\underline{x})$, $g_2(\underline{x}) := \dfrac{\partial f}{\partial x_2}(\underline{x})$, …, $g_n(\underline{x}) := \dfrac{\partial f}{\partial x_n}(\underline{x})$, si ha:

$$(Hf)(\underline{x}^0) := \begin{pmatrix} \dfrac{\partial}{\partial x_1}\dfrac{\partial f}{\partial x_1}(\underline{x}^0) & \dots & \dfrac{\partial}{\partial x_n}\dfrac{\partial f}{\partial x_1}(\underline{x}^0) \\ \dots & \dots & \dots \\ \dfrac{\partial}{\partial x_1}\dfrac{\partial f}{\partial x_n}(\underline{x}^0) & \dots & \dfrac{\partial}{\partial x_n}\dfrac{\partial f}{\partial x_n}(\underline{x}^0) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(\underline{x}^0) & \dots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\underline{x}^0) \\ \dots & \dots & \dots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\underline{x}^0) & \dots & \dfrac{\partial^2 f}{\partial x_n^2}(\underline{x}^0) \end{pmatrix}.$$

DEFINIZIONE. L'applicazione che ad ogni $\underline{u} \in \mathbb{R}^n$ associa il numero $\langle (Hf)(\underline{x}^0)\underline{u}, \underline{u} \rangle \in \mathbb{R}$ prende il nome di *differenziale secondo della f in* $\underline{x}^0$ e si indica con $(d^2f)(\underline{x}^0)$. (Il perché verrà chiarito tra poco, Teorema 14.)

OSSERVAZIONE. Risulta:

$$(d^2f)(\underline{x}^0)\,(\underline{u}) = \langle (Hf)(\underline{x}^0)\underline{u}, \underline{u} \rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i\,\partial x_j}(\underline{x}^0)u_i u_j\,;$$

dunque, se $(d^2f)(\underline{x}^0)$ non è il polinomio nullo, è un polinomio omogeneo di secondo grado.

ESEMPIO. 1) Sia $f: \mathbb{R}^2 \to \mathbb{R}$ l'applicazione definita da $f(x,y) = x^2y + xy^3$. Si ha:

$$\nabla f(\underline{x}) = (2xy + y^3,\ x^2 + 3xy^2)^{\mathrm{T}}; \quad (Hf)(\underline{x}) = \begin{pmatrix} 2y & 2x + 3y^2 \\ 2x + 3y^2 & 6xy \end{pmatrix}.$$

Dato il vettore $\underline{u} = (u, v)^T$, si ha:

$$(Hf)(\underline{x})\underline{u} = (2yu + (2x + 3y^2)v,\ (2x + 3y^2)u + 6xyv)^T;$$

$$\langle (Hf)(\underline{x}^0)\underline{u}\ ,\ \underline{u}\rangle = 2yu^2 + 2(2x + 3y^2)uv + 6xyv^2.$$

I seguenti risultati descrivono le proprietà del differenziale secondo .

TEOREMA 13. *Se* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *è due volte differenziabile in un punto* $\underline{x}^0 \in A$, *allora la matrice Hessiana* $(Hf)(\underline{x}^0)$ *è simmetrica; è cioè* $f_{x_ix_j}(\underline{x}^0) = f_{x_jx_i}(\underline{x}^0)$ (Cfr. § 4). ∎

Si ritrova così il risultato del Teorema di Schwarz, ma sotto ipotesi diverse.

TEOREMA 14. *Se* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *è due volte differenziabile in un punto* $\underline{x}^0 \in A$, *allora sussiste la seguente Formula di Taylor:*

$$f(\underline{x}) = f(\underline{x}^0) + \langle \nabla f(\underline{x}^0), \underline{x} - \underline{x}^0\rangle + \frac{1}{2}\langle (Hf)(\underline{x}^0)(\underline{x} - \underline{x}^0), \underline{x} - \underline{x}^0\rangle + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|^2,$$

con $\lim_{\underline{x}\to\underline{x}^0} \varepsilon(\underline{x}) = 0$.

DIM[1]**.** Proveremo il Teorema sotto l'ulteriore ipotesi che la funzione ∇f sia continua in A. Siano S una sfera di centro $\underline{x}^0$ contenuta in A e $\underline{x}$ un punto di S. Consideriamo la restrizione della f al segmento $[\underline{x}^0, \underline{x}] = \{\underline{x}^0 + t(\underline{x} - \underline{x}^0): 0 \le t \le 1\}$ di estremi $\underline{x}$ e $\underline{x}^0$ e poniamo $F(t) = f(\underline{x}^0 + t(\underline{x} - \underline{x}^0))$. La $F(t)$ è funzione, di classe C^1, di una sola variabile; ad essa si può dunque applicare la formula di Torricelli (cfr. il Teor. 12 del Cap. 13). Si ha:

$$(*)\qquad f(\underline{x}) - f(\underline{x}^0) = F(1) - F(0) = \int_0^1 F'(t)\,dt = \int_0^1 \langle \nabla f(\underline{x}^0 + t(\underline{x} - \underline{x}^0)), \underline{x} - \underline{x}^0\rangle\, dt.$$

Essendo, per ipotesi, $\nabla f(\underline{x})$ differenziabile in $\underline{x}^0$, si ha:

$$\nabla f(\underline{x}^0 + t(\underline{x} - \underline{x}^0)) = \nabla f(\underline{x}^0) + (Hf)(\underline{x}^0)(t(\underline{x} - \underline{x}^0)) + \varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0))\, t\,\|\underline{x} - \underline{x}^0\|,$$

con $\varepsilon(\underline{x})$ che tende a zero al tendere di $\underline{x}$ a $\underline{x}^0$. Sostituendo nella (*), si ottiene facilmente:

$$f(\underline{x}) - f(\underline{x}^0) = \int_0^1 \langle \nabla f(\underline{x}^0), \underline{x} - \underline{x}^0\rangle dt +$$

$$+ \int_0^1 \langle (Hf)(\underline{x}^0)(\underline{x} - \underline{x}^0),\ \underline{x} - \underline{x}^0\rangle t\,dt + \int_0^1 \langle \varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0)), \underline{x} - \underline{x}^0\rangle t\,\|\underline{x} - \underline{x}^0\|\,dt =$$

$$= \langle \nabla f(\underline{x}^0), \underline{x} - \underline{x}^0\rangle \int_0^1 dt + \langle (Hf)(\underline{x}^0)(\underline{x} - \underline{x}^0), \underline{x} - \underline{x}^0\rangle \int_0^1 t\,dt +$$

$$+ \|\underline{x} - \underline{x}^0\| \int_0^1 \langle \varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0)), \underline{x} - \underline{x}^0\rangle t\,dt =$$

---

[1] Questa dimostrazione presuppone la conoscenza di alcuni degli argomenti che verranno esposti nel prossimo Capitolo.

$$=\langle \nabla f(\underline{x}^0), \underline{x} - \underline{x}^0 \rangle + \frac{1}{2} \langle (Hf)(\underline{x}^0)(\underline{x} - \underline{x}^0), \underline{x} - \underline{x}^0 \rangle + \|\underline{x} - \underline{x}^0\| \int_0^1 \langle \varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0)), \underline{x} - \underline{x}^0 \rangle t\, dt.$$

Avendosi

$$\|\underline{x} - \underline{x}^0\| \left| \int_0^1 \langle \varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0)), \underline{x} - \underline{x}^0 \rangle t\, dt \right| \leq$$

$$\leq \|\underline{x} - \underline{x}^0\| \int_0^1 \left| \langle \varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0)), \underline{x} - \underline{x}^0 \rangle \right| t\, dt \leq \|\underline{x} - \underline{x}^0\|^2 \int_0^1 \|\varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0))\| t\, dt,$$

basta provare che $\int_0^1 \|\varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0))\| t\, dt$ tende a 0 al tendere di $\underline{x}$ a $\underline{x}^0$. Fissato un $\eta > 0$, esiste un $\delta > 0$ tale che da $0 < \|\underline{x} - \underline{x}^0\| < \delta$ segue $\|\varepsilon(\underline{x})\| < \eta$. Per tali $\underline{x}$ si ha

$$\int_0^1 \|\varepsilon(\underline{x}^0 + t(\underline{x} - \underline{x}^0))\| t\, dt < \int_0^1 \eta t\, dt = \frac{\eta}{2}. \blacksquare$$

TEOREMA 15. *Se* $f : A \to \mathbb{R}$, *con A sottoinsieme aperto di* $\mathbb{R}^n$, *è di classe* C^2 *in A, allora f è due volte differenziabile in A.*

DIM. Basta applicare il Teorema del differenziale totale alle derivate parziali $\frac{\partial f}{\partial x_i}$, con $i = 1, 2, \ldots, n$. $\blacksquare$

§ 4. FORME QUADRATICHE

DEFINIZIONE. Una matrice quadrata

$$M = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

è detta *simmetrica* se è $a_{ij} = a_{ji}$, con $i, j = 1, 2, \ldots, n$.

DEFINIZIONE. Data una matrice simmetrica M si dice *forma quadratica associata a M* la funzione $\varphi: \mathbb{R}^n \to \mathbb{R}$ definita da

$$\varphi(\underline{u}) = \langle M\underline{u}, \underline{u} \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} u_i u_j .$$

Dunque, se $\varphi(u)$ non è il polinomio nullo, è un polinomio omogeneo di secondo grado.

ESEMPIO. 1) $n = 1$; $\varphi(u) = au^2$;
$n = 2$; $\varphi(u_1, u_2) = au_1^2 + 2bu_1u_2 + cu_2^2$;
$n = 3$; $\varphi(u_1, u_2, u_3) = a_{11}u_1^2 + a_{22}u_2^2 + a_{33}u_3^2 + 2a_{12}u_1u_2 + 2a_{13}u_1u_3 + 2a_{23}u_2u_3$.

DEFINIZIONE. Una forma quadratica φ è detta:

definita positiva se è $\varphi(\underline{u}) > 0$ per ogni $\underline{u} \neq \underline{0}$;
definita negativa se è $\varphi(\underline{u}) < 0$ per ogni $\underline{u} \neq \underline{0}$;
semidefinita positiva se è $\varphi(\underline{u}) \geq 0$ per ogni $\underline{u}$;
semidefinita negativa se è $\varphi(\underline{u}) \leq 0$ per ogni $\underline{u}$;
indefinita (di segno) se $\exists\ \underline{u}, \underline{v}$ tali che $\varphi(\underline{u}) > 0$ e $\varphi(\underline{v}) < 0$.

ESEMPIO. 2) Si constata subito che:

$\varphi(u_1, u_2) = 3u_1^2 + 2u_2^2$ è definita positiva; $\varphi(u_1, u_2) = -u_1^2 - u_2^2$ è definita negativa;
$\varphi(u_1, u_2) = u_1^2 - u_2^2$ è indefinita; $\varphi(u_1, u_2) = u_1^2$ è semidefinita positiva.

Sussiste al riguardo il seguente risultato

TEOREMA 16. *(di Jacobi) - Data la matrice simmetrica*

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

si ponga: $M_1 := a_{11}$, $M_2 := \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, …, $M_n := |M|$ *(dunque M_i è il minore principale di ordine i, con i* = 1, 2, …, *n). Allora, per la forma quadratica φ associata a M si ha che:*

φ *è definita positiva se e solo se è* $M_1 > 0, M_2 > 0, M_3 > 0, \dots, M_n > 0$;
φ *è definita negativa se e solo se è* $M_1 < 0, M_2 > 0, M_3 < 0, \dots, (-1)^n M_n > 0$. ∎

Caso particolare, *n* = 2

TEOREMA 17. *Data la matrice simmetrica non nulla*

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

per la forma quadratica $\varphi(u_1, u_2) = au_1^2 + 2bu_1u_2 + cu_2^2$ *associata a M si ha che:*

φ *è definita positiva se e solo se è* $a > 0$ *e* $ac - b^2 > 0$;
φ *è definita negativa se e solo se è* $a < 0$ *e* $ac - b^2 > 0$;
φ *è indefinita se e solo se è* $ac - b^2 < 0$;
φ *è semidefinita (ma non definita) positiva se e solo se è* $ac - b^2 = 0$, *con* $a > 0$ *o* $c > 0$;
φ *è semidefinita (ma non definita) negativa se e solo se è* $ac - b^2 = 0$, *con* $a < 0$ *o* $c < 0$.

DIM. Sia $\underline{u} = (u_1,u_2)^T \neq \underline{0}$ e con $u_2 \neq 0$. Si ha:

$$\varphi(u_1, u_2) = u_2^2 \left(a \left(\frac{u_1}{u_2}\right)^2 + 2b\,\frac{u_1}{u_2} + c \right).$$

Posto $t = \dfrac{u_1}{u_2}$, si ottiene che il segno di $\varphi(u_1, u_2)$ è dato dal segno della funzione polinomiale $\psi(t) = at^2 + 2bt + c$. Ora la funzione ψ cambia segno se e solo se è $ac - b^2 < 0$, mentre è di segno costante se e solo se è $ac - b^2 > 0$, che implica $ac > 0$; in questo caso il segno di $\psi(t)$ è dato dal segno di a (e quindi di c). Ne viene che la φ è semidefinita (ma non definita) se e solo se è $ac - b^2 = 0$, con $(a \neq 0) \vee (c \neq 0)$. Essendo $\varphi(u_1, 0) = au_1^2$, si perviene alla conclusione anche nel caso che sia $u_2 = 0$. ▮

TEOREMA 18. *Una forma quadratica* φ *è definita positiva* [*negativa*] *se e solo se esiste un numero positivo m tale che* $\varphi(\underline{u}) \geq m\|\underline{u}\|^2$ [*se e solo se esiste un numero negativo M tale che* $\varphi(\underline{u}) \leq M\|\underline{u}\|^2$] *per ogni* $\underline{u}$.

DIM. Sia $\underline{u} \neq \underline{0}$; posto $\underline{v} = \frac{\underline{u}}{\|\underline{u}\|}$, si ha $\varphi(\underline{u}) = \|\underline{u}\|^2\varphi(\underline{v})$, con $\|\underline{v}\| = 1$. La funzione $\varphi(\underline{v})$ è definita e continua sulla sfera unitaria di $\mathbb{R}^n$ che è un insieme compatto; per il Teorema di Weierstrass, essa assume dunque un valore minimo m e uno massimo M. È dunque

$$m\|\underline{u}\|^2 \leq \|\underline{u}\|^2\varphi(\underline{v}) = \varphi(\underline{u}) \leq M\|\underline{u}\|^2.$$

Si ottiene così la tesi, dato che la φ è definita positiva [negativa] se e solo se è $m > 0$ [se e solo se è $M < 0$]. ▮

§ 5. ESTREMI LIBERI PER FUNZIONI SCALARI

Si pone in modo molto naturale il seguente:

PROBLEMA. Data la funzione $f : E(\subset \mathbb{R}^n) \to \mathbb{R}$, ricercare l'estremo superiore e l'estremo inferiore dei valori assunti dalla funzione, ossia $\sup f(E)$ e $\inf f(E)$. Si vuole, in particolare, decidere se la f è limitata o no su E.

Sappiamo che se la f è continua e l'insieme E è compatto (cioè chiuso e limitato), allora, per il Teorema di Weierstrass, l'insieme $f(E)$ ammette massimo e minimo. E se E non è compatto?

Chiaramente, se si trova un sottoinsieme di E in cui la restrizione della f è superiormente [inferiormente] illimitata, è tale anche la f su tutto E. Provare che la f è limitata è, di regola, più delicato, in quanto richiede un lavoro di maggiorazioni e minorazioni da escogitare di caso in caso.

ESEMPI. 1) Si consideri la funzione $f(x,y) = \frac{xy}{x^2 + y^2}$, definita in $E = \mathbb{R}^2 \setminus \{\underline{0}\}$. Dalla ben nota disuguaglianza $|xy| \leq \frac{x^2 + y^2}{2}$, si ha $\frac{|xy|}{x^2 + y^2} \leq \frac{1}{2}$. Dunque la f è limitata e si vede subito che $-\frac{1}{2}$ e $\frac{1}{2}$ sono, rispettivamente, il minimo e il massimo della f.

2) Si consideri la funzione di $\mathbb{R}^2$ in $\mathbb{R}$ definita da $f(x,y) = x^4 + y^4 - 4xy$. La sua restrizione all'asse delle ascisse dà luogo alla funzione x^4 che è superiormente illimitata; è dunque $\sup f(\mathbb{R}^2) = +\infty$. Si ha, inoltre, $f(x,y) = x^4 + y^4 - 4xy \geq x^4 + y^4 - 2(x^2 + y^2) = (x^4 - 2x^2) + (y^4 - 2y^2) \geq -2$ (come si vede studiando brevemente la funzione $x^4 - 2x^2$). Abbiamo così provato che la f è inferiormente limitata; si vede anzi che essa ha addirittura un valore minimo, dato che è $f(1,1) = -2$.

In analogia con quanto fatto per le funzioni di una variabile, si dà la seguente

DEFINIZIONE. Siano dati: una funzione $f : E(\subset \mathbb{R}^n) \to \mathbb{R}$ e un punto $\underline{x}^0 \in E$. Si dice che il punto $\underline{x}^0$ è di *massimo* [*minimo*] *relativo* per la f se esiste un intorno U di $\underline{x}^0$ tale che

$$\underline{x} \in U \cap E \setminus \{\underline{x}^0\} \Rightarrow f(\underline{x}) < f(\underline{x}^0) \qquad [\Rightarrow f(\underline{x}) > f(\underline{x}^0)].$$

Un punto $\underline{x}^0 \in E$ che sia di massimo o di minimo relativo per la f è detto un *punto di estremo per la f.*

DEFINIZIONE. Siano dati: una funzione $f : E(\subset \mathbb{R}^n) \to \mathbb{R}$ e un punto $\underline{x}^0$ interno ad E. Si dice che il punto $\underline{x}^0$ è di *sella* per la f se esistono due rette r e s passanti per $\underline{x}^0$ tali che questo punto sia di massimo relativo per la restrizione della f a $r \cap E$ e di minimo relativo per la restrizione della f a $s \cap E$.

Il Teorema di Fermat per le funzioni di una variabile può essere così generalizzato:

TEOREMA 19 *(Test delle derivate prime) - Siano dati: un sottoinsieme aperto A di* $\mathbb{R}^n$, *un punto* $\underline{x}^0 \in A$ *e una funzione* $f : A \to \mathbb{R}$ *differenziabile in* $\underline{x}^0$. *Se* $\underline{x}^0$ *è punto di estremo per la f, si ha necessariamente* $\nabla f(\underline{x}^0) = \underline{0}$.

DIM. Se il punto $\underline{x}^0$ è di estremo per la f lo è anche per le sue restrizioni alle rette per $\underline{x}^0$ e parallele agli assi; a tali restrizioni si può applicare il Teorema di Fermat. Dunque la f ha nulle in $\underline{x}^0$ tutte le sue derivate parziali prime. ∎

DEFINIZIONE. Data $f : E(\subset \mathbb{R}^n) \to \mathbb{R}$, un pu nto $\underline{x}^0 \in E$ in cui la f è differenziabile ed è $\nabla f(\underline{x}^0) = \underline{0}$ è detto un punto *critico* per la f.

Il Teorema precedente ci dice dunque che un punto di estremo per una funzione f a valori reali, definita e differenziabile su un aperto di $\mathbb{R}^n$, è un punto critico per la f.

N.B. Non sussiste l'implicazione opposta. Basta pensare ad una funzione del tipo $f(\underline{x}) = x^3$.

TEOREMA 20 *(Test delle derivate seconde) - Sia* $f : A \to \mathbb{R}$, *con A sottoinsieme aperto di* $\mathbb{R}^n$, *una funzione due volte differenziabile in un punto* $\underline{x}^0 \in A$ *e sia* $\nabla f(\underline{x}^0) = \underline{0}$. *Allora, detta* $\varphi(\underline{u})$ *la forma quadratica* $\langle (Hf)(\underline{x}^0)\underline{u}, \underline{u}\rangle$, *si ha che:*

i) se φ *è definita positiva,* $\underline{x}^0$ *è punto di minimo relativo per la f;*

ii) se φ *è definita negativa,* $\underline{x}^0$ *è punto di massimo relativo per la f;*

iii) se φ *è indefinita,* $\underline{x}^0$ *è punto di sella per la f;*

DIM. Posto $\underline{v} = \dfrac{\underline{x} - \underline{x}^0}{\|\underline{x} - \underline{x}^0\|}$ e utilizzando la formula di Taylor, si ha:

$$f(\underline{x}) - f(\underline{x}^0) = \langle \nabla f(\underline{x}^0), \underline{x} - \underline{x}^0\rangle + \frac{1}{2}\langle (Hf)(\underline{x}^0)(\underline{x} - \underline{x}^0), \ \underline{x} - \underline{x}^0\rangle + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|^2 =$$

$$= \frac{1}{2}\langle (Hf)(\underline{x}^0)\underline{v}, \underline{v}\rangle\cdot\|\underline{x} - \underline{x}^0\|^2 + \varepsilon(\underline{x})\cdot\|\underline{x} - \underline{x}^0\|^2 =$$

$$= \left(\frac{1}{2}\langle (Hf)(\underline{x}^0)\underline{v}, \underline{v}\rangle + \varepsilon(\underline{x})\right)\|\underline{x} - \underline{x}^0\|^2.$$

i) Per il Teorema 18, esiste un $m > 0$ tale che $\varphi(\underline{v}) = \langle (Hf)(\underline{x}^0)\underline{v}, \underline{v}\rangle \geq m$. È dunque:

$$f(\underline{x}) - f(\underline{x}^0) \geq \left(\frac{1}{2}m + \varepsilon(\underline{x})\right)\|\underline{x} - \underline{x}^0\|^2.$$

Dato che $\varepsilon(\underline{x})$) tende a zero al tendere di $\underline{x}$ a $\underline{x}^0$, si ha $\lim_{\underline{x}\to\underline{x}^0}\left(\frac{1}{2}m + \varepsilon(\underline{x})\right) = \frac{m}{2} > 0$. Per il Teorema della permanenza del segno, esiste dunque un intorno U di $\underline{x}^0$ in cui è $\frac{m}{2} + \varepsilon(\underline{x}) > 0$. Nello stesso intorno, per $\underline{x} \neq \underline{x}^0$, si ha $f(\underline{x}) - f(\underline{x}^0) > 0$.

ii) Si prova in modo perfettamente analogo, sfruttando il fatto che, sempre per il Teorema 18, esiste un $M < 0$ tale che $\varphi(\underline{v}) = \langle(Hf)(\underline{x}^0)\underline{v}, \underline{v}\rangle \leq M$.

iii) Se la forma quadratica $\langle(Hf)(\underline{x}^0)\underline{u}, \underline{u}\rangle$ è indefinita, si ha $m < 0 < M$. Esistono perciò due versori $\underline{v}_1$ e $\underline{v}_2$ tali che $\varphi(\underline{v}_1) > 0$ e $\varphi(\underline{v}_2) < 0$. Il punto $\underline{x}^0$ è di minimo per la restrizione di f ad $A \cap \{\underline{x}: \underline{x} = \underline{x}^0 + \underline{v}_1 t\}$ e di massimo per la restrizione di f ad $A \cap \{\underline{x}: \underline{x} = \underline{x}^0 + \underline{v}_2 t\}$. ∎

N.B. Sia $\underline{x}^0$ un punto in cui è $\nabla f(\underline{x}^0) = \underline{0}$. Se la forma quadratica $\langle(Hf)(\underline{x}^0)\underline{u}, \underline{u}\rangle$ è semidefinita, o se è il polinomio nullo, non si può dire, senza ulteriori informazioni, se il punto $\underline{x}^0$ è di estremo o meno.

ESEMPI. 3) Si cercano gli estremi della funzione di $\mathbb{R}^2$ in $\mathbb{R}$ definita da $f(x,y) = x^4 + y^4 - 4xy$. Si ha $\nabla f(\underline{x}) = (4x^3 - 4y, 4y^3 - 4x)^T$; esso si annulla in $\underline{x}^0 = \underline{0}$, in $\underline{x}^1 = (1,1)^T$ e in $\underline{x}^2 = (-1,-1)^T$. Avendosi

$$(Hf)(\underline{x}) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

si ottiene

$$(Hf)(\underline{0}) = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix}, \qquad (Hf)(\underline{x}^1) = (Hf)(\underline{x}^2) = \begin{pmatrix} 12 & -4 \\ -4 & 12 \end{pmatrix}.$$

Applicando il Teor. 17, si ottiene che le forma quadratiche $\langle(Hf)(\underline{x}^1)\underline{u}, \underline{u}\rangle$ e $\langle(Hf)(\underline{x}^2)\underline{u}, \underline{u}\rangle$ sono definite positive ed i punti $\underline{x}^1$ e $\underline{x}^2$ sono di minimo. Anzi, si ha $f(\underline{x}^1) = f(\underline{x}^2) = -2$; d'altra parte abbiamo visto più su che è $f(\underline{x}) \geq -2$; si riottiene così che -2 è il minimo della funzione. Si vede poi subito che la forma $\langle(Hf)(\underline{0})\underline{u}, \underline{u}\rangle = -8uv$ è indefinita e quindi il punto $\underline{0}$ è di sella.

4) Si cercano gli estremi della funzione di $\mathbb{R}^2$ in $\mathbb{R}$ definita da $f(x,y) = x^4 - 2x^2y + y^2$. Si ha $\nabla f(\underline{x}) = (4x^3 - 4xy, -2x^2 + 2y)^T$; i punti di annullamento si trovano risolvendo il sistema

$$\begin{cases} x^3 - xy = 0 \\ y - x^2 = 0 \end{cases} \Leftrightarrow \begin{cases} x(x^2 - y) = 0 \\ y - x^2 = 0 \end{cases}.$$

Il gradiente si annulla dunque in tutti e soli i punti del tipo $(x, x^2)^T$. Si ha:

$$(Hf)(\underline{x}) = \begin{pmatrix} 12x^2 - 4y & -4x \\ -4x & 2 \end{pmatrix} \qquad (Hf)(x,x^2) = \begin{pmatrix} 8x^2 & -4x \\ -4x & 2 \end{pmatrix},$$

da cui: $\det (Hf)(x, x^2) \equiv 0$. In tutti questi punti, la forma quadratica $\langle(Hf)(x, x^2)\underline{u}, \underline{u}\rangle$ è semidefinita. Per questa via, non possiamo perciò concludere nulla. Basta però osservare che è $f(x,y) = (x^2 - y)^2$ per stabilire che i punti $(x, x^2)^T$ sono tutti di minimo relativo in senso debole, che non ci sono punti di massimo relativo, che è $\min f = 0$ e $\sup f = +\infty$.

§ 6. ESTREMI VINCOLATI PER FUNZIONI SCALARI

Si cercano gli estremi della funzione $f : E(\subset \mathbb{R}^2) \to \mathbb{R}$ definita da $f(x,y) = x + y$, se:
1) $E = E_1 := \{(x,y)^T: x^2 + y^2 \leq 1\}$;
2) $E = E_2 := \{(x,y)^T: x^4 + y^4 - 2xy \leq 1\}$.

Nel primo caso, si arriva facilmente al risultato. Si vede che gli estremi vanno ricercati fra i punti per cui è $x^2 + y^2 = 1$. Si può allora esplicitare una delle due variabili su due semicirconferenze, oppure si può scrivere l'equazione parametrica della circonferenza [cioè $x = \cos \vartheta$, $y = \sin \vartheta$]; in ogni caso ci si riduce a studiare funzioni di una sola variabile. Ma nel secondo caso la faccenda è molto più complicata. Come possiamo procedere?

DEFINIZIONE. Data le funzione $f : E(\subset \mathbb{R}^n) \to \mathbb{R}$, diremo che un sottoinsieme proprio e non vuoto Γ di E è un *vincolo* per la *f*.

ESEMPIO. 1) I vincoli tipici (ma non gli unici possibili) sono:
a) $n = 2$; $\Gamma := \{(x,y)^T: \varphi(x,y) = 0\}$, curva piana.
b) $n = 3$; $\Gamma := \{(x,y,z)^T: \varphi(x,y,z) = 0\}$, superficie nello spazio.
c) $n = 3$; $\Gamma := \{(x,y,z)^T: \varphi(x,y,z) = 0,\ \psi(x,y,z) = 0\}$, curva nello spazio data come intersezione di due superfici.
[Si veda quanto detto nel $ 2 del Cap. 11 a proposito dei termini "curva" e "superficie".]

DEFINIZIONE. Sono dati una: funzione $f : E(\subset \mathbb{R}^n) \to \mathbb{R}$, un vincolo Γ e un punto $\underline{x}^0 \in \Gamma$. Si dice che $\underline{x}^0$ è *di estremo vincolato* o *condizionato* per *f* su Γ se $\underline{x}^0$ è punto di estremo per la restrizione di *f* a Γ.

ESEMPI. 2) Trovare gli estremi condizionati della funzione $f(x,y,z) = x + y^2 + z^3$, con i vincoli: $z = x^2 + y^2$, $x + y + z = 1$.

3) Trovare i punti della curva di equazione $3x^2 + 2xy + 3y^2 = 1$ che hanno distanza massima o minima dal punto origine $\underline{0}$.

Stabiliamo, intanto, il seguente risultato:

TEOREMA 21. *Siano:* $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ *una funzione di classe* C^1 *sull'insieme aperto* A, $\Gamma = \{(x,y)^T \in A: \varphi(x,y) = 0\} \subset A$ *un vincolo per f e* $\gamma: I = [a,b] \to \mathbb{R}^n$ *una curva regolare* (cioè *differenziabile* e con $\gamma'(t) \neq \underline{0}$, $\forall t \in I$), *con sostegno contenuto in* Γ. *Sia poi* $\underline{x}^0 \in \Gamma$, *con* $\underline{x}^0 = \gamma(t_0)$, $t_0 \in\ ]a,b[$. *Se* $\underline{x}^0$ *è punto di estremo condizionato per f su* Γ, *allora si ha*

$$\langle \nabla f(\underline{x}^0), \gamma'(\underline{x}^0) \rangle = 0.$$

DIM. La funzione $\psi(t) = f(\gamma(t)) : [a,b] \to \mathbb{R}$ è di classe C^1 ed ha in t_0 un punto di estremo interno. Dunque, per il Teorema di Fermat, si ha

$$0 = \psi'(t_0) = \langle \nabla f(\underline{x}^0), \gamma'(\underline{x}^0) \rangle. \blacksquare$$

Ciò ci mostra che i punti di estremo vincolato vanno ricercati fra quelli in cui il ∇f o non è definito o è ortogonale alla tangente di ogni curva regolare passante per il punto stesso e avente il sostegno contenuto nel vincolo Γ. Ci si esprime dicendo che ∇f è *ortogonale al vincolo* Γ.

Il metodo dei moltiplicatori di Lagrange

Una condizione necessaria affinché un punto sia di estremo è fornita dal seguente risultato, attribuito a Lagrange. Non produrremo la dimostrazione di questo Teorema; inoltre, anziché darne un unico enunciato generale, preferiamo spezzarlo in tre diverse proposizioni, allo scopo di renderne più chiaro l'utilizzo pratico.

$\boxed{n = 2.}$ **TEOREMA 22.** *Siano:* $f, \varphi : A(\subset \mathbb{R}^2) \to \mathbb{R}$ *due funzioni di classe* C^1 *sull'insieme aperto* A, $\Gamma := \{(x,y)^T \in A: \varphi(x,y) = 0\}$, $\underline{x}^0 = (x_0, y_0)^T \in \Gamma$, *con* $\nabla\varphi(\underline{x}^0) \neq \underline{0}$. *Allora, se* $\underline{x}^0$ *è di estremo condizionato per la* f *su* Γ, *esiste un numero reale* λ_0 *tale che*

$$\nabla f(\underline{x}^0) + \lambda_0 \nabla\varphi(\underline{x}^0) = \underline{0},$$

cioè (x_0, y_0, λ_0) *è soluzione del sistema:*

$$(*) \qquad \begin{cases} f_x(x,y) + \lambda\varphi_x(x,y) = 0 \\ f_y(x,y) + \lambda\varphi_y(x,y) = 0 \\ \varphi(x,y) = 0 \end{cases} . \blacksquare$$

OSSERVAZIONE. Notiamo che la condizione necessaria espressa dal Teorema precedente dice che ∇f è *parallelo* a $\nabla\varphi$. Tenuto presente che $\nabla\varphi$ è *ortogonale* alla tangente alla curva di sostegno Γ (cfr. l'Esercizio 6), si vede che il risultato concorda con quanto visto nel Teorema 21.

ESEMPIO. 4) Trovare i punti della curva di equazione $3x^2 + 2xy + 3y^2 = 1$ che hanno distanza massima o minima dal punto origine $\underline{0}$. Poiché la radice quadrata è una funzione crescente, il problema è equivalente a quello di trovare il valore massimo e il valore minimo della funzione $f(x,y) = x^2 + y^2$, con il vincolo $\varphi(x,y) = 3x^2 + 2xy + 3y^2 - 1 = 0$. Il sistema (*) diventa:

$$\begin{cases} 2x + \lambda(6x + 2y) = 0 \\ 2y + \lambda(6y + 2x) = 0 \\ 3x^2 + 2xy + 3y^2 - 1 = 0 \end{cases} \Leftrightarrow \begin{cases} (1 + 3\lambda)x + \lambda y = 0 \\ (1 + 3\lambda)y + \lambda x = 0 \\ 3x^2 + 2xy + 3y^2 - 1 = 0 \end{cases} \Leftrightarrow \begin{cases} (1 + 2\lambda)(x - y) = 0 \\ (1 + 3\lambda)y + \lambda x = 0 \\ 3x^2 + 2xy + 3y^2 - 1 = 0 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} x = y \\ (1 + 4\lambda)x = 0 \\ 8x^2 = 1 \end{cases} (\Rightarrow\ x = y = \pm\frac{\sqrt{2}}{4}) \quad \vee \quad \begin{cases} \lambda = \frac{-1}{2} \\ y = -x \\ 4x^2 = 1 \end{cases} \quad (\Rightarrow\ x = -y = \pm\frac{1}{2}).$$

Si ha poi $\nabla\varphi(\underline{x}) = \underline{0}$ se e solo se è $\underline{x} = \underline{0}$, ma $\underline{0}$ non appartiene a Γ. La funzione f è continua ed è ristretta ad un insieme chiuso e limitato; esiste perciò un valore massimo ed uno minimo. Gli unici punti dove la funzione può assumere questo massimo e questo minimo sono:

$$\underline{x}^1 = \left(\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}\right)^T, \ \underline{x}^2 = \left(-\frac{\sqrt{2}}{4}, -\frac{\sqrt{2}}{4}\right)^T \ \underline{x}^3 = \left(\frac{1}{2}, \frac{-1}{2}\right)^T \ \underline{x}^4 = \left(\frac{-1}{2}, \frac{1}{2}\right)^T.$$

Avendosi $f(\underline{x}^1) = f(\underline{x}^2) = 1/4$ e $f(\underline{x}^3) = f(\underline{x}^4) = 1/2$, si conclude che questi due valori sono, rispettivamente, il minimo e il massimo di quelli assunti dalla f nella restrizione studiata. In conclusione, i punti della curva che hanno distanza massima da $\underline{0}$ sono $\underline{x}^3$ e $\underline{x}^4$, mentre quelli che hanno distanza minima sono $\underline{x}^1$ e $\underline{x}^2$.

N.B. Si tenga presente che, quando risolviamo il sistema (*), non siamo interessati a determinare i valori di λ, lo facciamo solo se questo ci è utile per determinare i valori di x e di y (e, quando è il caso, di z) che sono quelli che stiamo cercando.

$\boxed{\textbf{n = 3, 1 vincolo}}$. **TEOREMA 23.** *Siano: $f, \varphi : A(\subset \mathbb{R}^3) \to \mathbb{R}$ due funzioni di classe C^1 sull'insieme aperto A, $\Gamma := \{(x,y,z)^T \in A: \varphi(x,y,z) = 0\}$, $\underline{x}^0 = (x_0, y_0, z_0)^T \in \Gamma$, con $\nabla\varphi(\underline{x}^0) \neq \underline{0}$. Allora, se $\underline{x}^0$ è di estremo condizionato per la f su Γ, esiste un numero reale λ_0 tale che*

$$\nabla f(\underline{x}^0) + \lambda_0 \nabla\varphi(\underline{x}^0) = \underline{0},$$

cioè $(x_0, y_0, z_0, \lambda_0)$ è soluzione del sistema:

$$(*) \qquad \begin{cases} f_x(x,y,z) + \lambda\varphi_x(x,y,z) = 0 \\ f_y(x,y,z) + \lambda\varphi_y(x,y,z) = 0 \\ f_z(x,y,z) + \lambda\varphi_z(x,y,z) = 0 \\ \varphi(x,y,z) = 0 \end{cases} . \blacksquare$$

ESEMPIO. 5) Trovare gli estremi della funzione $f(x,y,z) = x + y + z$, su

$$E := \left\{(x,y,z)^T\!: \ \frac{x^2}{4} + y^2 + \frac{z^2}{9} = 1\right\}.$$

Il sistema (*) diventa:

$$\begin{cases} 1 + \frac{2}{4}\lambda x = 0 \\ 1 + 2\lambda y = 0 \\ 1 + \frac{2}{9}\lambda z = 0 \\ \frac{x^2}{4} + y^2 + \frac{z^2}{9} = 1 \end{cases} \Leftrightarrow \begin{cases} \lambda = -\frac{2}{x} \\ y = \frac{x}{4} \\ z = \frac{9}{4}x \\ \frac{14}{16}x^2 = 1 \end{cases} \Leftrightarrow (1) \vee (2)$$

$$(1) \begin{cases} x = -2\sqrt{\frac{2}{7}} \\ y = -\frac{1}{2}\sqrt{\frac{2}{7}} \\ z = -\frac{9}{2}\sqrt{\frac{2}{7}} \end{cases} ; \quad (2) \begin{cases} x = 2\sqrt{\frac{2}{7}} \\ y = \frac{1}{2}\sqrt{\frac{2}{7}} \\ z = \frac{9}{2}\sqrt{\frac{2}{7}} \end{cases} .$$

Si ha poi $\nabla\varphi(\underline{x}) = \underline{0}$ se e solo se è $\underline{x} = \underline{0}$, ma $\underline{0}$ non appartiene a Γ. La funzione f è continua ed è ristretta ad un insieme chiuso e limitato; esiste perciò un valore massimo ed uno minimo. Gli unici punti dove la funzione può assumere questo massimo e questo minimo sono:

$$\underline{x}^1 = \left(-2\sqrt{\frac{2}{7}}, -\frac{1}{2}\sqrt{\frac{2}{7}}, -\frac{9}{2}\sqrt{\frac{2}{7}}\right)^T, \quad \underline{x}^2 = \left(2\sqrt{\frac{2}{7}}, \frac{1}{2}\sqrt{\frac{2}{7}}, \frac{9}{2}\sqrt{\frac{2}{7}}\right)^T,$$

Avendosi $f(\underline{x}^2) = 7\sqrt{\frac{2}{7}}$ e $f(\underline{x}^1) = -7\sqrt{\frac{2}{7}}$ si conclude che questi due valori sono, rispettivamente, il massimo e il minimo di quelli assunti dalla f nella restrizione studiata.

n = 3, 2 vincoli. **TEOREMA 24.** *Siano: f, φ, ψ: $A(\subset \mathbb{R}^3) \to \mathbb{R}$ tre funzioni di classe C^1 nell'insieme aperto A, $\Gamma := \{(x,y,z)^T \in A: \varphi(x,y,z) = 0, \psi(x,y,z) = 0\}$, $\underline{x}^0 = (x_0, y_0, z_0)^T \in \Gamma$, con la matrice Jacobiana $\begin{pmatrix} \varphi_x(\underline{x}^0) & \varphi_y(\underline{x}^0) & \varphi_z(\underline{x}^0) \\ \psi_x(\underline{x}^0) & \psi_y(\underline{x}^0) & \psi_z(\underline{x}^0) \end{pmatrix}$ di rango (o caratteristica) 2. Allora, se $\underline{x}^0$ è di estremo condizionato per la f su Γ, esistono due numeri reali λ_0 e μ_0 tali che*

$$\nabla f(\underline{x}^0) + \lambda_0 \nabla\varphi(\underline{x}^0) + \mu_0 \nabla\psi(\underline{x}^0) = \underline{0},$$

cioè $(x_0, y_0, z_0, \lambda_0, \mu_0)$ è soluzione del sistema:

$$(*) \qquad \begin{cases} f_x(x,y,z) + \lambda\varphi_x(x,y,z) + \mu\psi_x(x,y,z) = 0 \\ f_y(x,y,z) + \lambda\varphi_y(x,y,z) + {}_0\psi_y(x,y,z) = 0 \\ f_z(x,y,z) + \lambda\varphi_z(x,y,z) + \mu\psi_z(x,y,z) = 0 \\ \varphi(x,y,z) = 0 \\ \psi(x,y,z) = 0 \end{cases} \quad . \blacksquare$$

ESEMPIO. 6) Si vogliono trovare gli estremi condizionati della funzione $f(x,y,z) = x + y^2z$, con i vincoli: $x^2 + y^2 - 2 = 0$ e $z - x = 0$. Il sistema (*) diventa

$$\begin{cases} 1 + 2\lambda x - \mu = 0 \\ 2yz + 2\lambda y = 0 \\ y^2 + \mu = 0 \\ x^2 + y^2 - 2 = 0 \\ z - x = 0 \end{cases} \Leftrightarrow \begin{cases} \mu = -y^2 \\ 2y(z + \lambda) = 0 \\ 1 + 2\lambda x + y^2 = 0 \\ x^2 + y^2 - 2 = 0 \\ z - x = 0 \end{cases} \Leftrightarrow \ (1) \vee (2);$$

$$(1) \begin{cases} \mu = 0 = y \\ 1 + 2\lambda x = 0 \\ x^2 = 2 \\ z - x = 0 \end{cases} \Rightarrow \begin{cases} y = 0 \\ x = z = \pm\sqrt{2} \end{cases};$$

$$(2) \begin{cases} \mu = -y^2 \\ \lambda = -z \\ 1 - 2xz + y^2 = 0 \\ x^2 + y^2 - 2 = 0 \\ z - x = 0 \end{cases} \Rightarrow \begin{cases} 1 - 2x^2 + y^2 = 0 \\ x^2 + y^2 - 2 = 0 \\ z - x = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 = 1 \\ x^2 + y^2 - 2 = 0 \\ z - x = 0 \end{cases} \Leftrightarrow \begin{cases} x = z = \pm 1 \\ y = \pm 1 \end{cases} .$$

Gli unici punti in cui il rango della matrice jacobiana del vincolo è minore di 2 sono, come si constata facilmente, quelli del tipo $(0, 0, z)^T$ che però non appartengono a Γ. I punti che possono essere di estremo condizionato per la nostra funzione sono dunque i seguenti:

$$\underline{x}^1 = (\sqrt{2}, 0, \sqrt{2})^T; \quad \underline{x}^2 = (-\sqrt{2}, 0, -\sqrt{2})^T; \quad \underline{x}^3 = (1, 1, 1)^T;$$

$$\underline{x}^4 = (1, -1, 1)^T; \quad \underline{x}^5 = (-1, 1, -1)^T; \quad \underline{x}^6 = (-1, -1, -1)^T.$$

Si ha: $f(\underline{x}^1) = \sqrt{2}$, $f(\underline{x}^2) = -\sqrt{2}$; $f(\underline{x}^3) = f(\underline{x}^4) = 2$; $f(\underline{x}^5) = f(\underline{x}^6) = -2$. I valori minimo e massimo sono dunque -2 e 2.

OSSERVAZIONE. (Ricetta per la ricerca dei punti di estremo). I punti di estremo per una funzione f: E $(\subset \mathbb{R}^n) \to \mathbb{R}$, con $\mathcal{F}E = \{\underline{x}: \Phi(\underline{x}) = 0\}$, vanno ricercati tra:

1) i punti interni in cui è $\nabla f = \underline{0}$;
2) i punti interni in cui f non è differenziabile;
3) i punti di frontiera in cui è applicabile il metodo dei moltiplicatori di Lagrange;
4) i punti di frontiera in cui non è applicabile il metodo dei moltiplicatori di Lagrange e, in particolare, i punti in cui il rango della matrice jacobiana $J\Phi$ non è massimo.

ESEMPIO. 7) Cercare gli estremi della funzione $f(x,y,z) = x + y + z$, ristretta al cubo $E = \{(x,y,z)^T: 0 \le x \le 1, 0 \le y \le 1, 0 \le z \le 1\}$.

Occupiamoci dapprima dei punti interni. Si vede subito che il gradiente della f è definito in ogni punto e non è mai nullo. Non ci sono punti di estremo interni.

Veniamo ai punti di frontiera. Cominciamo con i punti interni alle facce del cubo.

Faccia $F_1 = \{(x,y,z)^T: 0 \le x \le 1, 0 \le y \le 1, z = 0\}$. Si ottiene la funzione $f(x,y) = x + y$. Questa è una funzione di due variabili definita su un quadrato; il gradiente di questa funzione è sempre definito e mai nullo; non ci sono punti di estremo interni. In modo analogo si procede per le altre facce del cubo.

Si passa allora agli spigoli. Sia, per esempio, $S_1 = \{(x,y,z)^T: 0 \le x \le 1, y = 0, z = 0\}$. Si ottiene la funzione $f(x) = x$, che non ha punti di estremo per $0 < x < 1$. Non ci resta che da calcolare i valori della f negli 8 vertici del cubo. Minimo: $f(0, 0, 0) = 0$; massimo: $f(1, 1, 1) = 3$.

§ 7. E S E R C I Z I

1) *a*) Siano: $g : \mathbb{R}^2 \to \mathbb{R}^2$ la funzione $\begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}$, $f : \mathbb{R}^2 \to \mathbb{R}$, una funzione differenziabile e $h(\underline{u}) = f \circ g(\underline{u})$. Si calcoli $(Jh)(\underline{u}^0)$, con $\underline{u}^0 = (\rho_0, \vartheta_0)^T$.

b) Stesso problema con g: $\mathbb{R} \to \mathbb{R}^3$ data da $\begin{pmatrix} \cos u \\ \sin u \\ u \end{pmatrix}$ e f: $\mathbb{R}^3 \to \mathbb{R}$ differenziabile.

[$\Re$. *a*)
$$\left(\frac{\partial h}{\partial \rho}(\underline{u}^0), \frac{\partial h}{\partial \vartheta}(\underline{u}^0)\right) = \left(\frac{\partial f}{\partial x_1}(\underline{x}^0), \frac{\partial f}{\partial x_2}(\underline{x}^0)\right) \begin{pmatrix} \cos \vartheta_0 & -\rho_0 \sin \vartheta_0 \\ \sin \vartheta_0 & \rho_0 \cos \vartheta_0 \end{pmatrix} =$$
$$= \left(\frac{\partial f}{\partial x_1}(\underline{x}^0)\cos \vartheta_0 + \frac{\partial f}{\partial x_2}(\underline{x}^0)\sin \vartheta_0, -\frac{\partial f}{\partial x_1}(\underline{x}^0)\rho_0 \sin \theta_0 + \frac{\partial f}{\partial x_2}(\underline{x}^0)\rho_0 \cos\vartheta_0\right).$$
b)
$$\frac{dh}{dx}(u_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \frac{\partial f}{\partial x_3}(x_0)\right)\begin{pmatrix} -\sin u_0 \\ \cos u_0 \\ 1 \end{pmatrix} = -\frac{\partial f}{\partial x_1}(x_0)\sin u_0 + \frac{\partial f}{\partial x_2}(x_0)\cos u_0 + \frac{\partial f}{\partial x_3}(x_0).]$$

2) Calcolare il gradiente delle seguenti funzioni di $\mathbb{R}^2$ o $\mathbb{R}^3$ in $\mathbb{R}$:

a) $f(x,y) = x^2 + y^2\sin(xy)$; *b*) $g(x,y,z) = \log(x^2 + y^2 - 2z^2)$; *c*) $h(x,y,z) = x^{(y - z)}$.

[$\Re$. *a*) $\nabla f(\underline{x}) = (2x + y^3\cos(xy), 2y \sin(xy) + xy^2\cos(xy))^T$;

b) $\nabla g(\underline{x}) = \left(\frac{2x}{x^2 + y^2 - 2z^2}, \frac{2y}{x^2 + y^2 - 2z^2}, \frac{-4z}{x^2 + y^2 - 2z^2}\right)^T$;

c) $\nabla h(\underline{x}) = (x^{(y - z - 1)}, x^{(y - z)}\log x, -x^{(y - z)}\log x)^T$.]

3) *a*) Calcolare la derivata direzionale della funzione $f(x,y)$ dell'Esercizio (2*a*) nel punto $(1,1)^T$ secondo la direzione del vettore $(1,2)^T$.
b) Analoga domanda per le funzioni degli esercizi (2*b*) e (2*c*) nel punto $(1,1,0)^T$ secondo la direzione del vettore $(1,0,1)^T$.

[$\Re$. *a*) Il versore della direzione assegnata è $\underline{v} = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)^T$; dall'Esercizio (2*a*) si ha: $\nabla f(1,1) = (2 + \cos 1, 2\sin 1 + \cos 1)^T$. Si ottiene: $\frac{\partial f}{\partial \underline{v}}(1,1) = \frac{2 + \cos 1}{\sqrt{5}} + 2\,\frac{2\sin 1 + \cos 1}{\sqrt{5}}$.

b) In questi casi è $\underline{v} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^T$; poi si procede esattamente come sopra.]

4) Si calcoli la matrice Jacobiana della funzione composta $f\circ g$, con $f(\underline{x})$ di $\mathbb{R}^3$ in $\mathbb{R}$ differenziabile e $g : \mathbb{R}^3 \to \mathbb{R}^3$ data da $g(\rho,\vartheta,\varphi) = \begin{pmatrix} \rho \sin\varphi \cos\vartheta \\ \rho \sin\varphi \sin\vartheta \\ \rho \cos\varphi \end{pmatrix}$.

[$\Re$. Basta calcolare il prodotto (righe per colonne) delle matrici Jacobiane della f e della g:

$$\left(\frac{\partial f}{\partial x_1}(\underline{x}),\ \frac{\partial f}{\partial x_2}(\underline{x}),\ \frac{\partial f}{\partial x_3}(\underline{x})\right)\begin{pmatrix} \sin\varphi\cos\vartheta & \rho\cos\varphi\cos\vartheta & -\rho\sin\varphi\sin\vartheta \\ \sin\varphi\sin\vartheta & \rho\cos\varphi\sin\vartheta & \rho\sin\varphi\cos\vartheta \\ \cos\varphi & -\rho\sin\varphi & 0 \end{pmatrix}.]$$

5) Sia $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$ differenziabile in $\underline{x}^0 \in A$, con $\nabla f(\underline{x}^0) \neq \underline{0}$. In quale direzione è massima (minima) la derivata direzionale della f in $\underline{x}^0$?

[$\Re$. Sia dato un versore $\underline{v}$. Sappiamo che è:

$$(*) \qquad \left|\frac{\partial f}{\partial \underline{v}}\right| = |\langle \nabla f, \underline{v}\rangle| \leq \|\nabla f\|\cdot\|\underline{v}\| = \|\nabla f\| \quad (\neq \underline{0} \text{ per ipotesi}).$$

Il valore massimo (e quello minimo) si hanno quando nella (*) vale il segno di uguaglianza e sappiamo che ciò accade se e solo se i vettori $\underline{v}$ e ∇f sono paralleli (Cfr. Teorema 19 del Capitolo 11).]

6) *Curve di livello.* Data $f : A(\subset \mathbb{R}^n) \to \mathbb{R}$, f differenziabile, fissiamo un punto $\underline{x}^0 \in A$. È detto *insieme di livello* l'insieme

$$\Gamma := \{\underline{x} \in A: f(\underline{x}) = f(\underline{x}^0)\}.$$

Sia $n = 2$. Si può dimostrare che, se è $\nabla f(\underline{x}^0) \neq \underline{0}$, Γ è, almeno localmente, il sostegno di una *curva regolare* γ (detta *curva di livello*) esprimibile nella forma cartesiana $y = g(x)$ [oppure $x = h(y)$], definita in un intervallo I. Si trovi, sotto queste ipotesi, l'espressione di $g'(x)$ [di $h'(y)$].

[$\Re$. Sia, per esempio, γ esprimibile nella forma $y = g(x)$; si ha dunque $\gamma(t) = (t,g(t))^T$. Posto $F(t) = f(t, g(t))$, si ottiene un'applicazione di I in $\mathbb{R}$ derivabile, con $F'(t) \equiv 0$. Ne viene:

$$F'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle = \frac{\partial f}{\partial x}(\gamma(t)) + \frac{\partial f}{\partial y}(\gamma(t))g'(t) \equiv 0,$$

da cui; $$g'(t) = -\frac{f_x(x,g(x))}{f_y(x,g(x))}.]$$

7) Si trovino gli estremi della funzione $f(x,y) = x + y$ ristretta all'insieme dei punti del piano per cui è $x^4 + y^4 - 2xy - 1 \leq 0$.

[$\Re$. Si vede subito che non ci sono punti di estremo interni al dominio. Passando ai punti di frontiera. Si constata che è $\nabla\varphi(\underline{x}) = \underline{0}$ solo nei punti $(k,k)^T$ con $k \in \{0, \pm\sqrt{1/2}\}$ che però non appartengono al vincolo. Applicando il metodo dei moltiplicatori di Lagrange, si ottiene il sistema formato dalle 3 equazioni:

(*) $\quad 1 + 2\lambda[2x^3 - y] = 0; \quad 1 + 2\lambda[2y^3 - x] = 0; \quad x^4 + y^4 - 2xy - 1 = 0.$

Sottraendo la seconda dalla prima, si ottiene:

$$2\lambda[2(x^3 - y^3) + (x - y)] = 2\lambda(x - y)[2x^2 + 2xy + 2y^2 + 1] = 0.$$

Se è $\lambda = 0$, la prima delle (*) diventa $1 = 0$. Deve perciò essere $\lambda \neq 0$. Si vede che l'ultima equazione è soddisfatta solo dai punti per cui è $x = y$ (infatti è $2x^2 + 2xy + 2y^2 + 1 = x^2 + y^2 + (x + y)^2 + 1 > 0$). Sostituendo nella terza delle (*), si ottiene l'equazione: $2x^4 - 2x^2 - 1 = 0$. Si trovano così i punti $\underline{x}^1 = (t,t)^T$ e $\underline{x}^2 = (-t,-t)^T$ con $t = \sqrt{\frac{1+\sqrt{3}}{2}}$. Conclusione: il minimo condizionato della f è $f(\underline{x}^2) = -2\sqrt{\frac{1+\sqrt{3}}{2}}$ e il massimo è $f(\underline{x}^1) = 2\sqrt{\frac{1+\sqrt{3}}{2}}$.]

8) Si trovino gli estremi della funzione $f(x,y) = x^2 + y^2$ ristretta all'insieme dei punti del piano per cui è $x^4 + y^4 - 2xy - 1 = 0$.

[$\Re$. … Si ottiene il sistema formato dalle 3 equazioni

(*) $\quad x + \lambda[2x^3 - y] = 0; \quad y + \lambda[2y^3 - x] = 0; \quad x^4 + y^4 - 2xy - 1 = 0.$

Non può essere $\lambda = 0$, perché si otterrebbe l'uguaglianza $1 = 0$. Non può essere $x = 0$, perché dalla prima si otterrebbe anche $y = 0$, ma $(0,0)$ non è soluzione della terza. Dunque x, y e λ sono tutti diversi da zero. Dalle prime due (*) si ottiene:

$$-\frac{1}{\lambda} = \frac{2x^3 - y}{x} = \frac{2y^3 - x}{y},$$

da cui $$(2xy + 1)(x^2 - y^2) = 0.$$

Se fosse $2xy + 1 = 0$, la terza delle (*) diventerebbe $x^4 + y^4 = 0$, che non dà soluzioni, essendo $x \neq 0 \neq y$.

Da $x = y$ si ottengono i punti $\underline{x}^1$ e $\underline{x}^2$ trovati nell'Esercizio precedente.

Da $x = -y$, si trovano i punti $\underline{x}^3 = (t',t')^T$ e $\underline{x}^4 = (-t',-t')^T$, con $t' = \sqrt{\frac{-1+\sqrt{3}}{2}}$.

Minimo: $f(\underline{x}^3) = f(\underline{x}^4) = -1 + \sqrt{3}$; massimo: $f(\underline{x}^1) = f(\underline{x}^2) = 1 + \sqrt{3}$.]

9) Si ricerchino il massimo e il minimo della funzione $f(x,y,z) = xyz$ ristretta all'insieme

$$E_k := \{(x,y,z)^T: x + y + z = k, x \geq 0, y \geq 0, z \geq 0\}.$$

Si sfrutti il risultato ottenuto per dimostrare che la media geometrica di 3 numeri positivi è sempre minore o uguale alla loro media aritmetica.

[$\Re$. La funzione f è continua e definita su un insieme compatto, quindi, per il Teorema di Weierstrass, ammette un massimo e un minimo. Si vede subito che il minimo è 0 ed è assunto nei punti in cui è nulla almeno una delle coordinate. Cerchiamo il massimo fra i punti del dominio per cui è $x > 0$, $y > 0$, $z > 0$. Applicando il metodo dei moltiplicatori di Lagrange, si ottiene il sistema formato dalle 4 equazioni:

$$(*) \qquad yz = -\lambda; \qquad xz = -\lambda; \qquad xy = -\lambda; \qquad x + y + z - k = 0.$$

Dalle prime 3 delle (*), dividendo membro a membro, si ottiene il sistema

$$x = y = z.$$

Sostituendo nella quarta delle (*), si ottiene $x = y = z = \frac{k}{3}$. Il valore massimo della f è dunque assunto nel punto $\underline{x}^0 = \left(\frac{k}{3}, \frac{k}{3}, \frac{k}{3}\right)^T$ e si ha $f(\underline{x}^0) = \frac{k^3}{3^3}$.

Dunque, da $x > 0$, $y > 0$, $z > 0$, $x + y + z = 3M$ $(= k)$, segue $xyz \leq M^3$ e quindi $\sqrt[3]{xyz} \leq M$.]

10) Si studi il comportamento nell'origine delle funzioni di $\mathbb{R}^2$ in $\mathbb{R}$:

$$f_1(x,y) = x^2 + y^4; \quad f_2(x,y) = x^2 - y^4; \quad f_3(x,y) = x^2; \quad f_4(x,y) = x^2 + y^3;$$
$$g_1(x,y) = x^4 + y^4; \quad g_2(x,y) = x^4 - y^4; \quad g_3(x,y) = x^4; \quad g_4(x,y) = x^3.$$

[$\Re$. Per tutte le funzioni è $\nabla f(\underline{0}) = \underline{0}$. Per le prime 4 funzioni è $\langle (Hf_i)(\underline{0})\underline{u},\underline{u}\rangle = 2u^2$ che è una forma quadratica semidefinita positiva, ma $\underline{0}$ è punto di minimo per f_1, di sella per f_2, di minimo relativo in senso debole per f_3 e nulla di tutto ciò per f_4. Per le ultime 4 funzioni $\langle (Hg_i)(\underline{0})\underline{u},\underline{u}\rangle$ è il polinomio nullo (forma quadratica *nulla*); ora $\underline{0}$ è punto di minimo per g_1, di sella per g_2, di minimo relativo in senso debole per g_3 e nulla di tutto ciò per g_4.]

11) **3.-** Si determinino gli estremi assoluti e relativi della funzione $f(x,y) = \log(x-+ 2y)$,

sull'insieme $\qquad E = \{(x,y)^T: x^2 + y^2 \leq 1\} \cap D$, con D = dominio di f.

12) È data la funzione $\qquad f(x,y) = x^2 - x - y^2$.

a) Trovare gli estremi assoluti e relativi di f.
b) Trovare gli estremi assoluti di f ristretta all'insieme $E = \{(x, y)^T: x^2 + y^2 \leq 1\}$.

13) Si determinino gli estremi assoluti e relativi della funzione $\quad f(x,y,z) = x^2 + y^2 + z^2$,

sull'insieme $\qquad E = \{(x,y,z)^T: \frac{x^2}{4} + y^2 + \frac{z^2}{9} \leq 1\}$.

14) Si determinino gli estremi assoluti e relativi della funzione $f(x,y) = x^2 - 2xy^2 - y^2$,

sull'insieme $\qquad E = \{(x,y)^T: x^2 + y^2 \leq 1\}$.