Calcolo differenziale per funzioni di piu' variabili reali

Davide Guidetti

1 Derivate

In questa sezione introduciamo le definizioni di base del calcolo differenziale.

Definizione 1.1. *Siano $A \subseteq \mathbb{R}^n$ ($n \in \mathbb{N}$), $v \in \mathbb{R}^n$, $x^0 \in A$. Poniamo*

$$A_{x^0,v} := \{t \in \mathbb{R} : x^0 + tv \in A\}. \tag{1.1}$$

Sia $f : A \to \mathbb{R}$. Per $t \in A_{x^0,v} \setminus \{0\}$, poniamo

$$r_{x^0,v}(t) := \frac{f(x^0 + tv) - f(x^0)}{t}. \tag{1.2}$$

e supponiamo che $0 \in D(A_{x^0,v})$. Se esiste

$$\lim_{t\to 0} r_{x^0,v}(t) = \lim_{t\to 0} \frac{f(x^0 + tv) - f(x^0)}{t}, \tag{1.3}$$

lo chiameremo **derivata di f rispetto al vettore v in x^0** *e lo indicheremo con la notazione $\frac{\partial f}{\partial v}(x^0)$.*

Esempio 1.1. Siano $A := \mathbb{R}^2$, $f : A \to \mathbb{R}$, $f(x_1, x_2) = x_1 x_2$, $x^0 = (1, 0)$, $v = (1, 1)$. Evidentemente, $A_{x^0,v} = \mathbb{R}$ e percio' $0 \in D(A_{x^0,v})$. Per $t \neq 0$, si ha

$$r_{x^0,v}(t) = \frac{f(1 + t, t)}{t} = 1 + t \to 1 \ (t \to 0).$$

Dunque, esiste $\frac{\partial f}{\partial v}(x^0)$ e vale 1.

Osservazione 1.1. Con riferimento alla definizione 1.1, si ha che il dominio di $r_{x^0,v}$ è l'insieme $A_{x^0,v} \setminus \{0\}$. Osserviamo che $0 \in D(A_{x^0,v})$ se e solo se $0 \in D(A_{x^0,v} \setminus \{0\})$. Dunque, la condizione $0 \in D(A_{x^0,v})$ è proprio cio' che serve per dare un senso a (1.3).

E' importante il fatto che, se $x^0 \in \mathring{A}$, la condizione $0 \in D(A_{x^0,v})$ è soddisfatta qualunque sia v in $\mathbb{R}^n$. Sia infatti $r \in \mathbb{R}^+$, tale che $B(x^0, r) \subseteq A$. Evidentemente, $A_{x^0,O} = \mathbb{R}$, e dunque $0 \in D(A_{x^0,O})$ (questo è vero supponendo semplicemente $x^0 \in A$). Se invece $v \neq 0$, $x^0 + tv \in A$ se

$$\|tv\| = d(x^0, x^0 + tv) < r.$$

Ora,

$$\|tv\| = |t|\|v\| < r$$

se $t \in]-r/\|v\|, r/\|v\|[$. Dunque,

$$]-r/\|v\|, r/\|v\|[\subseteq A_{x^0,v},$$

e questo implica $0 \in D(A_{x^0,v})$.

Consideriamo, invece, il caso $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}$, $x^0 = (0,0)$. Se $v^0 = (1,0)$, si ha

$$A_{x^0,v^0} = \mathbb{R},$$

per cui la condizione $0 \in D(A_{x^0,v^0})$ è soddisfatta. Invece, per $v^1 := (0,1)$, si ha

$$A_{x^0,v^1} = \{0\},$$

e quindi $0 \notin D(A_{x^0,v^1})$.

Nella prossima definizione, considereremo la cosi' detta **base canonica** di $\mathbb{R}^n$. Ricordiamo che con tale espressione si intende la famiglia di vettori $\{e^1, ..., e^j, ..., e^n\}$, con

$$\begin{array}{ccccccccc} e^j & = & (& 0, & ..., & 1, & ..., & 0 &), 1 \le j \le n. \\ & & & 1 & ... & j & ... & n & \end{array} \tag{1.4}$$

Definizione 1.2. *Siano $A \subseteq \mathbb{R}^n$ ($n \in \mathbb{N}$), $1 \le j \le n$, $x^0 \in A$, tale che $0 \in D(A_{x^0,e^j})$. Chiamiamo* **derivata parziale prima rispetto alla variabile x_j nel punto x^0** *la derivata $\frac{\partial f}{\partial e^j}(x^0)$ (se esiste).*

Tale derivata sarà indicata con uno dei simboli

$$\frac{\partial f}{\partial x_j}(x^0), \quad D_j f(x^0). \tag{1.5}$$

Osservazione 1.2. Siano $A \subseteq \mathbb{R}^n$, $1 \le j \le n$, $x^0 \in A$, tale che $0 \in D(A_{x^0,e^j})$, $f : A \to \mathbb{R}$. Se $x^0 = (x_1^0, ..., x_j^0, ..., x_n^0)$. Poniamo

$$B := \{y \in \mathbb{R} : (x_1^0, ..., y, ..., x_n^0) \in A\},$$

$$\begin{cases} g : B \to \mathbb{R}, \\ g(y) = f(x_1^0, ..., y, ..., x_n^0), \quad y \in B. \end{cases}$$

Allora, se $t \in A_{x^0,e^j} \setminus \{0\}$, $x_j^0 + te^j \in B$ e

$$\begin{aligned} \frac{f(x^0+te^j)-f(x^0)}{t} &= \frac{f(x_1^0,...,x_j^0+t,...,x_n^0)-f(x_1^0,...,x_j^0,...,x_n^0)}{t} \\ &= \frac{g(x_j^0+t)-g(x_j^0)}{t}. \end{aligned}$$

Percio' $D_j f(x^0)$ coincide (se esiste) con $g'(x_j^0)$. Nel processo di limite, la variabile x_k, con $k \ne j$, resta costantemente uguale a x_k^0. Percio', il calcolo di $D_j f(x^0)$ segue regole analoghe al calcolo della derivata ordinaria in $\mathbb{R}$, assumendo x_j come unica variabile. Le altre variabili vanno trattate come costanti. Per esempio, siano

$$A := \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 > 0\},$$

$$\begin{cases} f : A \to \mathbb{R}, \\ f(x) = f(x_1, x_2, x_3) = x_1^{x_2 x_3}, \quad x \in A. \end{cases}$$

Allora, $\forall x \in A$, si ha

$$D_1 f(x) = x_1^{x_2 x_3 - 1} x_2 x_3,$$

$$D_2 f(x) = x_1^{x_2 x_3} x_3 \ln(x_1),$$

$$D_3 f(x) = x_1^{x_2 x_3} x_2 \ln(x_1).$$

Osservazione 1.3. E' noto che, se $f : A \subseteq \mathbb{R} \to \mathbb{R}$ è derivabile in x^0, con $x^0 \in A \cap D(A)$, f è automaticamente continua in A.

Per funzioni di piu' variabili, è abbastanza chiaro che l'esistenza di tutte le derivate parziali prime finite in un punto non implica la continuita' della funzione in quel punto. Si consideri, in proposito, il seguente esempio: sia

$$\begin{cases} f : \mathbb{R}^2 \to \mathbb{R}, \\ f(x) = f(x_1, x_2) = \begin{cases} 1 & \text{se } x_1 x_2 = 0, \\ 0 & \text{altrimenti.} \end{cases} \end{cases} \tag{1.6}$$

Allora, è facile verificare che esistono $D_1 f(0,0)$ e $D_2 f(0,0)$ e valgono entrambe 0. Tuttavia, f non è continua in $(0,0)$. Infatti, sia $\epsilon = 1/2$. Non esiste alcun δ in $\mathbb{R}^+$ tale che, $\forall x \in B((0,0), \delta)$, valga $|f(x) - f(0,0)| < 1/2$. Infatti, se poniamo

$$x^\delta := (\delta/(2\sqrt{2}), \delta/(2\sqrt{2})),$$

si ha $\|x^\delta\| = \delta/2 < \delta$, ma

$$|f(x^\delta) - f(0,0)| = 1 > 1/2.$$

E' possibile presentare un esempio ancora piu' drastico di una funzione dotata in $(0,0)$ di tutte le derivate $\frac{\partial f}{\partial v}(0,0)$ finite qualunque sia v in $\mathbb{R}^2$, ma discontinua in $(0,0)$. Si consideri infatti la funzione definita nell'esempio 3.4 in [1]. Tale funzione f non è continua in $(0,0)$, in quanto non esiste $\lim_{x \to (0,0)} f(x)$. Tuttavia, qualunque sia v in $\mathbb{R}^2$ e qualunque sia $t \in \mathbb{R}$, tv appartiene a una certa retta A passante per l'origine, dipendente da v. Abbiamo visto che , qualunque sia la retta A, esiste $\delta > 0$, tale che, se $x \in A$ e $\|x\| < \delta$, vale $f(x) = 1$. Percio', qualunque sia $v \in \mathbb{R}^2$, esiste $\delta(v) \in \mathbb{R}^+$, tale che, comunque si prenda t in $]-\delta(v), \delta(v)[$, si ha $f(tv) = 1$. Ne segue che, se $t \in]-\delta(v), \delta(v)[\setminus \{0\}$,

$$\frac{f(tv) - f(0,0)}{t} = 0,$$

da cui $\frac{\partial f}{\partial v}(0,0) = 0$.

E' allora opportuno introdurre la seguente nozione, che implica (come vedremo) continuita' ed esistenza delle derivate rispetto a qualunque $v \in \mathbb{R}^n$.

Definizione 1.3. *Siano $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, $x^0 \in \mathring{A}$. Diremo che f e'* **differenziabile in** x^0 *se esiste $w \in \mathbb{R}^n$ tale che*

$$f(x^0 + h) - f(x^0) - w \cdot h = o(\|h\|) \quad (h \to 0),$$

cioe'

$$\lim_{h \to 0} \frac{f(x^0 + h) - f(x^0) - w \cdot h}{\|h\|} = 0. \tag{1.7}$$

Osservazione 1.4. Con riferimento alla definizione 1.3, poiche' $x^0 \in \mathring{A}$, esiste $\delta \in \mathbb{R}^+$ tale che, se $0 < \|h\| < \delta$, $x_0 + h \in A$. Quindi il limite in (1.7) ha senso.

Osservazione 1.5. La definizione 1.3 puo' essere riformulata dicendo che esiste $T : \mathbb{R}^n \to \mathbb{R}$ lineare tale che

$$f(x^0 + h) - f(x^0) - Th = o(\|h\|_m) \quad (h \to 0),$$

cioe'

$$\lim_{h \to O} \frac{f(x^0 + h) - f(x^0) - Th}{\|h\|_n} = 0.$$

L'interesse della nozione di funzione differenziale dipende dal seguente risultato:

Proposizione 1.1. *Siano $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, $x^0 \in \mathring{A}$. Supponiamo che f sia differenziabile in x^0. Allora:*
(I) f e' continua in x^0;
(II) $\forall v \in \mathbb{R}^n$ esiste $\frac{\partial f}{\partial v}(x^0)$ ed e' uguale a $w \cdot v$;
(III) w e' univocamente determinato e $w = (D_1 f(x^0), \ldots, D_n f(x^0))$.

Dimostrazione Poniamo, per $x \in A$,

$$r(x) := f(x) - f(x^0) - w \cdot (x - x_0).$$

Allora

$$|f(x) - f(x^0)| \leq |r(x)| + |w \cdot (x - x^0)| \leq |r(x)| + \|w\| \|x - x^0\|.$$

Siano, $\epsilon, \eta \in \mathbb{R}^+$. Allora, esiste $\delta(\eta) > 0$ tale che, se $\|x - x^0\| < \delta(\eta)$,

$$|r(x)| \leq \eta \|x - x^0\|,$$

per cui, se $x \in A$, e $\|x - x^0\| \leq \delta(\eta)$,

$$|f(x) - f(x^0)| \leq (\eta + \|w\|)\|x - x^0\|.$$

Quindi, se $x \in A$ e $\|x - x^0\| < \min\{\delta(\eta), \frac{\epsilon}{\eta + \|w\|}\}$,

$$|f(x) - f(x^0)| < \epsilon.$$

(II) Sia $v \in \mathbb{R}^n$. Sia $r \in \mathbb{R}^+$ tale che $B(x^0, r) \subseteq A$. Allora, se $t \in \mathbb{R}$, $|t|\|v\| < r$ e $t \neq 0$,

$$\frac{f(x^0 + tv) - f(x^0)}{t} = w \cdot v + \frac{r(x^0 + tv)}{t}.$$

Di conseguenza, se $v \neq O$,

$$|\frac{f(x^0 + tv) - f(x^0)}{t} - w \cdot v| = |\frac{r(x^0 + tv)}{t\|v\|}| \|v\| \leq \eta \|v\|,$$

se $\|tv\| < \delta(\eta)$, e $\eta\|v\| < \epsilon$ se $\eta < \frac{\epsilon}{\|v\|}$.
(III) Da (II) segue che, per $j = 1, \ldots, n$,

$$w_j = w \cdot e^j = D_j f(x^0).$$

□

Definizione 1.4. *Se sono soddisfatte le ipotesi della definizione 1.3 e f e' differenziabile in x^0 chiameremo* **gradiente di f in x^0** *e indicheremo col simbolo $\nabla f(x^0)$ l'elemento $w \in \mathbb{R}^n$.*

Osservazione 1.6. Siano soddisfatte le condizioni della definizione 1.3 e sia f differenziabile in x^0, con $\nabla f(x^0) \neq O$. Poniamo

$$v^0 = \frac{1}{\|\nabla f(x^0)\|} \nabla f(x^0).$$

Allora v^0 rappresenta la "direzione da x^0 di massima crescita di f". Infatti, se $v \in \mathbb{R}^n$ e $\|v\| = 1$ (in modo che v individua una direzione), si ha

$$f(x^0 + tv) = t\nabla f(x^0) \cdot v + o(t) \quad (t \to 0).$$

Si ha, per la disuguaglianza di Cauchy-Schwarz (vedi [1], teorema 1.1)

$$\nabla f(x^0) \cdot v \leq |\nabla f(x^0) \cdot v| \leq \|\nabla f(x^0)\| = \nabla f(x^0) \cdot v^0.$$

Si puo' invece verificare che, se $\|v\| = 1$ e $v \neq v^0$,

$$\nabla f(x^0) \cdot v < \|\nabla f(x^0)\|$$

(vedi l'esercizio 1.6).

Teorema 1.1. *Siano $A \subseteq \mathbb{R}^n$, $x^0 \in \mathring{A}$, $x^0 = (x_1^0, x_2^0, \dots, x_n^0)$, $r \in \mathbb{R}^+$ tale che $B(x^0, r) \subseteq A$. Supponiamo che $\forall x \in B(x^0, r)$, $\forall j \in \{1, \dots, n\}$ sia definita a valori reali $D_j f(x)$ e che le funzioni $D_1 f, \dots, D_n f$ siano continue in x^0. Allora f e' differenziabile in x^0.*

Di conseguenza, $\forall v \in \mathbb{R}^n$ esiste $\frac{\partial f}{\partial v}(x^0)$ ed e' uguale a $\nabla f(x^0) \cdot v$.

Dimostrazione L'ultima affermazione segue dalla prima e dalla proposizione 1.1(II).

Sia $x \in B(x^0, r)$. Allora

$$\begin{aligned} f(x) - f(x^0) &= f(x_1, x_2, \dots, x_n) - f(x_1^0, x_2^0, \dots, x_n^0) \\ &= f(x_1, x_2, \dots, x_n) - f(x_1^0, x_2, \dots, x_n) + f(x_1^0, x_2, \dots, x_n) - f(x_1^0, x_2^0, \dots, x_n) \\ &+ \dots + f(x_1^0, x_2^0, \dots, x_{i-1}^0, x_i, \dots, x_{n-1}, x_n) - f(x_1^0, x_2^0, \dots, x_{i-1}^0, x_i^0, \dots, x_{n-1}, x_n) \\ &+ \dots + f(x_1^0, x_2^0, \dots, x_{n-1}^0, x_n) - f(x_1^0, x_2^0, \dots, x_{n-1}^0, x_n^0). \end{aligned} \tag{1.8}$$

Osserviamo che l'espressione in (1.8) ha senso, valendo, per $i = 2, \dots, n$,

$$d((x_1^0, \dots, x_{i-1}^0, x_i, \dots, x_n), x^0) = (\sum_{k=i}^{n} (x_k - x_k^0)^2)^{1/2} \leq (\sum_{k=1}^{n} (x_k - x_k^0)^2)^{1/2} = d(x, x^0) < r.$$

Inoltre, poiche' $B(x^0, r)$ è convessa, $\forall t \in [x_i^0, x_i]$ (o $[x_i, x_i^0]$), $(x_1^0, \dots, x_{i-1}^0, t, x_{i+1}, \dots, x_n) \in B(x^0, r)$. Per le ipotesi fatte, ciascuna delle funzioni $g_i(t) = f(x_1^0, \dots, x_{i-1}^0, t, x_{i+1}, x_n)$ ($i \in \{1, \dots, n\}$) è derivabile in $[x_i^0, x_i]$ (o $[x_i, x_i^0]$) e

$$\begin{aligned} g_i'(t) &= \lim_{h \to 0} \tfrac{f(x_1^0, \dots, x_{i-1}^0, t+h, x_{i+1}, x_n) - f(x_1^0, \dots, x_{i-1}^0, t, x_{i+1}, x_n)}{h} \\ &= D_i f(x_1^0, \dots, x_{i-1}^0, t, x_{i+1}, x_n). \end{aligned}$$

Di conseguenza, per il teorema del valor medio, per ciascun $i \in \{1, \dots, n\}$, esiste $c_i \in [x_i^0, x_i]$ tale che

$$\begin{aligned} & f(x_1^0, \dots, x_{i-1}^0, x_i, \dots, x_n) - f(x_1^0, \dots, x_{i-1}^0, x_i^0, \dots, x_n) \\ &= g_i(x_i) - g_i(x_i^0) = g'(c_i)(x_i - x_i^0) \\ &= D_i f(x_1^0, \dots, x_{i-1}^0, c_i, , x_{i+1}, x_n)(x_i - x_i^0). \end{aligned}$$

Quindi

$$f(x_1, \dots, x_n) - f(x_1^0, \dots, x_n^0) = \sum_{i=1}^{n} D_i f(x_1^0, \dots, x_{i-1}^0, c_i, , x_{i+1}, x_n)(x_i - x_i^0).$$

e

$$f(x_1,\dots,x_n) - f(x_1^0,\dots,x_n^0) - \sum_{i=1}^n D_i f(x_1^0,\dots,x_{i-1}^0, x_i^0,,\dots,x_n^0)(x_i - x_i^0)$$

$$= \sum_{i=1}^n [D_i f(x_1^0,\dots,x_{i-1}^0, c_i,,x_{i+1},x_n) - D_i f(x_1^0,\dots,x_{i-1}^0,x_i^0,\dots,x_n^0)](x_i - x_i^0).$$

Poiche' le funzioni $D_i f$ $(1 \le i \le n)$ sono continue in x^0, fissato $\eta \in \mathbb{R}^+$, esiste $\delta(\eta) \in]0,r]$ tale che, se $y \in B(x^0,\delta(\eta))$, per $i = 1,\dots,n$,

$$|D_i f(y) - D_i f(x^0)| < \eta.$$

Se $x \in B(x^0,\delta(\eta))$, si ha, per ciascun $i \in \{1,\dots,n\}$,

$$d((x_1^0,\dots,x_{i-1}^0,c_i,x_{i+1},x_n),(x_1^0,\dots,x_{i-1}^0,x_i^0,\dots,x_{i+1}^0,\dots,x_n^0))$$

$$= [(c_i - x_i^0)^2 + \sum_{j=i+1}^n (x_i - x_i^0)^2]^{1/2} \le [\sum_{j=i}^n (x_i - x_i^0)^2]^{1/2}$$

$$\le d(x,x^0) < \delta(\eta).$$

Di conseguenza,

$$|f(x) - f(x^0) - \sum_{i=1}^n D_i f(x_1^0,\dots,x_{i-1}^0,x_i^0,,\dots,x_n^0)(x_i - x_i^0)|$$

$$\le \sum_{i=1}^n |D_i f(x_1^0,\dots,x_{i-1}^0,c_i,,x_{i+1},x_n) - D_i f(x_1^0,\dots,x_{i-1}^0,x_i^0,,\dots,x_n^0)||x_i - x_i^0| \tag{1.9}$$

$$\le \eta \sum_{i=1}^n |x_i - x_i^0| \le n\eta \|x - x^0\|.$$

Sia $\epsilon \in \mathbb{R}^+$. Preso $\eta < \frac{\epsilon}{n}$, da (1.9) si ricava

$$|f(x) - f(x^0) - \sum_{i=1}^n D_i f(x_1^0,\dots,x_0^{i-1},x_i^0,,\dots,x_n^0)(x_i - x_i^0)| \le \epsilon \|x - x^0\|.$$

Percio' f è differenziabile in x^0.

□

Introduciamo ora la classe di funzioni $C^1(A)$.

Definizione 1.5. *Siano $A \subseteq \mathbb{R}^n$ aperto, $f : A \to \mathbb{R}$. Diremo che f è di* **classe** C^1 *in A ($f \in C^1(A)$), se $\forall x \in A$, $\forall j \in \{1,...,n\}$ esiste a valori reali $D_j f(x)$. Inoltre le funzioni f, $D_1 f$,...,$D_n f$ sono continue su A.*

Dal teorema 1.1 si ottiene facilmente il seguente

Corollario 1.1. *Siano $A \subseteq \mathbb{R}^n$ aperto, $f \in C^1(A)$. Allora f e' differenziabile in ogni punto $x \in A$.*

Osservazione 1.7. Il risultato del teorema 1.1 non si puo' invertire: f puo' essere differenziabile in x^0 senza che le derivate siano continue in x^0. Per un esempio vedi l'esercizio 1.5.

Esempio 1.2. E' facile verificare che le funzioni polinomiali sono di classe C^1 in $\mathbb{R}^n$. Infatti, sono continue (vedi l'esempio 3.6 in [1]) e le loro derivate parziali prime sono ancora funzioni polinomiali, per cui sono anch'esse continue.

Esempio 1.3. Rivisitiamo l'esempio 1.1 alla luce del teorema 1.1. Si ha che $f \in C^1(\mathbb{R}^2)$ e, $\forall x = (x_1, x_2)$ in $\mathbb{R}^2$,

$$D_1 f(x) = x_2, \qquad D_2 f(x) = x_1,$$

da cui

$$\nabla f(1,0) = (0,1).$$

Di conseguenza, se $v = (1,1)$, si ha

$$\frac{\partial f}{\partial v}(1,0) = (0,1) \cdot (1,1) = 1.$$

Vediamo ora una generalizzazione del teorema del valor medio.

Teorema 1.2. *Siano $A \subseteq \mathbb{R}^n$ ($n \in \mathbb{N}$), x e y in A tali che $[x,y] \subseteq \mathring{A}$, f differenziabile in ogni punto di $[x,y]$. Allora esiste $z \in [x,y]$, tale che*

$$f(y) - f(x) = \nabla f(z) \cdot (y - x).$$

Dimostrazione Poniamo

$$\begin{cases} F : [0,1] \to \mathbb{R}, \\ F(t) = f(x + t(y-x)). \end{cases} \tag{1.10}$$

F è derivabile in $[0,1]$ e, $\forall t \in [0,1]$,

$$F'(t) = \nabla f(x + t(y-x)) \cdot (y - x), \tag{1.11}$$

Infatti, se $t \in [0,1]$, $h \in \mathbb{R} \setminus \{0\}$ e $t + h \in [0,1]$, si ha, per la proposizione 1.1(II),

$$\begin{aligned} \lim_{h\to 0} \tfrac{F(t+h)-F(t)}{h} &= \lim_{h\to 0} \tfrac{f(x+t(y-x)+h(y-x))-f(x+t(y-x))}{h} \\ = \tfrac{\partial f}{\partial (y-x)}(x+t(y-x)) &= \nabla f(x+t(y-x)) \cdot (y-x). \end{aligned}$$

Allora, applicando il teorema del valor medio unidimensionale, esiste $c \in]0,1[$ tale che

$$\begin{aligned} f(y) - f(x) &= F(1) - F(0) = F'(c) \\ &= \nabla f(x + c(y-x)) \cdot (y-x). \end{aligned}$$

Otteniamo allora la conclusione, ponendo $z = x + c(y - x)$.

□

Osservazione 1.8. Per uso futuro, osserviamo che la derivata della funzione F definita in (1.10) è espressa dalla formula (1.11).

Esercizio 1.1. Determinare il dominio naturale delle seguenti funzioni e calcolare, dove possibile, le derivate parziali prime:

(I) $f(x_1, x_2) = x_1^2 + 3x_1x_2$;
(II) $f(x_1, x_2) = \frac{x_1+x_2}{x_1-x_2}$;
(III) $f(x_1, x_2) = x_1^4 x_2^2 - 3x_1x_2 + 2x_2$;
(IV) $f(x_1, x_2) = \sin(x_1x_2)$;
(V) $f(x_1, x_2) = \ln(x_1 + x_2)$;
(VI) $f(x_1, x_2) = e^{-(x_1^2+x_2^2)}$;
(VII) $f(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$;

(VIII) $f(x_1, x_2) = x_1^2 x_2$;
(IX) $f(x_1, x_2) = \arctan(\frac{x_2}{x_1})$;
(X) $f(x_1, x_2) = x_2^{-x_1^2}$;
(XI) $f(x_1, x_2, x_3) = \frac{x_1^2+3x_1x_3+x_2^2}{x_3}$;
(XII) $f(x_1, x_2, x_3) = \frac{1}{\sqrt{x_1^2+x_2^2+x_3^2}}$;
(XIII) $f(x_1, x_2, x_3) = x_1x_2x_3$;
(XIV) $f(x) = \|x\|^2$ $(x \in \mathbb{R}^n)$;
(XV) $f(x) = \|x\|$ $(x \in \mathbb{R}^n)$;
(XVI) $f(x) = f(x_1, ..., x_n) = \prod_{i=1}^n x_i = x_1 \cdot ... \cdot x_n$.

Esercizio 1.2. Siano $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$, $x^0 \in \mathring{A}$. Provare che f è differenziabile in x^0 se e solo se f è derivabile in x^0.

Esercizio 1.3. Sia

$$\begin{cases} f : \mathbb{R}^2 \to \mathbb{R}, \\ f(x) = f(x_1, x_2) = \begin{cases} (\frac{x_1^2 x_2}{x_1^4+x_2^2})^2 & \text{se } (x_1, x_2) \neq (0,0), \\ 0 & \text{se } (x_1, x_2) = (0,0). \end{cases} \end{cases}$$

Verificare che, $\forall v \in \mathbb{R}^2$, esiste finita $\frac{\partial f}{\partial v}(0,0)$. Verificare poi che f non è continua in $(0,0)$. A tale proposito, considerare $f(x_1, x_1^2)$ $(x_1 \in \mathbb{R})$.

Esercizio 1.4. Siano $A \subseteq \mathbb{R}^n$, $\nu \in \mathbb{R}^n$, $x^0 \in A$, tali che $0 \in D(A_{x^0,\nu})$, $f, g : A \to \mathbb{R}$, tali che esistono in $\mathbb{R}$ $\frac{\partial f}{\partial \nu}(x_0)$ e $\frac{\partial g}{\partial \nu}(x_0)$. Verificare che:
(I) esiste $\frac{\partial (f+g)}{\partial \nu}(x_0)$ e coincide con $\frac{\partial f}{\partial \nu}(x_0) + \frac{\partial g}{\partial \nu}(x_0)$;
(II) esiste $\frac{\partial (fg)}{\partial \nu}(x_0)$ e coincide con $\frac{\partial f}{\partial \nu}(x_0)g(x^0) + f(x^0)\frac{\partial g}{\partial \nu}(x_0)$;
(III) se $g(x) \neq 0$ $\forall x \in A$, esiste $\frac{\partial (\frac{f}{g})}{\partial \nu}(x_0)$ e coincide con

$$\frac{\frac{\partial f}{\partial \nu}(x_0)g(x^0) - \frac{\partial g}{\partial \nu}(x_0)f(x^0)}{g(x^0)^2}.$$

Esercizio 1.5. Sia

$$\begin{cases} f : \mathbb{R} \to \mathbb{R}, \begin{cases} x^2 \sin(\frac{1}{x}) & \text{se} \quad x \neq 0, \\ 0 & \text{se} \quad x \neq 0. \end{cases} \end{cases}$$

Verificare che f è derivabile in ogni $x \in \mathbb{R}$, ma che f' non è continua in 0. Di conseguenza, per il risultato dell'esercizio 1.2, f è differenziabile in ogni punto di $\mathbb{R}$, ma non appartiene a $C^1(\mathbb{R})$.

Esercizio 1.6. Analizzando la dimostrazione della disuguaglianza di Cauchy-Schwarz, verificare che, se $v, w \in \mathbb{R}^n$, $v \neq O$, $\|w\| = 1$, $w \neq \|v\|^{-1}v$, allora

$$v \cdot w < \|v\|.$$

(Sugg.: da $v \cdot w = \|v\|$ segue

$$\|v + tw\|^2 = \|v\|^2 + 2tv \cdot w + t^2 = (\|v\| + t)^2 = 0$$

se $t = -\|v\|$. Quindi $v = \|v\|w$).

2 Derivate di ordine superiore

Le "derivate delle derivate" sono importanti, come vedremo, anche in dimensione superiore a 1. Cominciamo con la loro definizione.

Definizione 2.1. *Siano $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, $1 \leq j \leq n$. Poniamo*

$$A_j := \{x \in A : \exists \ \ D_j f(x) \in \mathbb{R}\}.$$

Su A_j è definita la funzione $D_j f$, che associa a ogni elemento x di A_j la derivata $D_j f(x)$. Se $x \in A_j$ e $1 \leq i \leq n$, può esistere la derivata $D_i(D_j f)(x)$ di $D_j f$ in x. Essa viene chiamata **derivata seconda** *di f in x rispetto alle variabili x_i, x_j (nell'ordine) e viene indicata con uno dei simboli $D_{ij} f(x)$, $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ $(D_j^2 f(x)$, $\frac{\partial^2 f}{\partial x_j^2}(x)$ se $i = j)$.*

Iterando il procedimento, si possono definire le derivate terze, quarte, ecc.. (o derivate di ordine 3, 4, ecc.). Ricorsivamente, dati $j_1, \dots, j_k, j_{k+1}$, con $k \in \mathbb{N}$, $j_1, j_2, ..., j_k, j_{k+1} \in \{1, ..., n\}$, non necessariamente a due a due distinti, e supposta nota la funzione $D_{j_k \dots j_1} f$ di dominio $A_{j_k \dots j_1}$, porremo

$$A_{j_{k+1} j_k \dots j_1} := \{x \in A_{j_k \dots j_1} : \exists D_{j_{k+1}}(D_{j_k \dots j_1} f)(x) \in \mathbb{R}\}$$

e

$$D_{j_{k+1} j_k \dots j_1} f(x) = D_{j_{k+1}}(D_{j_k \dots j_1} f)(x) = \frac{\partial^{k+1} f}{\partial x_{j_1} \partial x_{j_2} ... \partial x_{j_k} \partial x_{j_{k+1}}}(x).$$

Esempio 2.1. Sia $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = x_1^3 - 3x_1 x_2^2$. Allora si verifica facilmente che, $\forall x = (x_1, x_2) \in \mathbb{R}^2$, si ha

$$D_1 f(x) = 3x_1^2 - 3x_2^2, D_2 f(x) = -6x_1 x_2,$$

$$D_{11} f(x) = 6x_1, D_{12} f(x) = D_{21} f(x) = -6x_2, D_{22} f(x) = -6x_1,$$
$$D_{111} f(x) = 6, D_{112} f(x) = D_{121} f(x) = D_{211} f(x) = 0,$$
$$D_{122} f(x) = D_{212} f(x) = D_{221} f(x) = -6, D_{222} f(x) = 0.$$

Le derivate di ordine superiore a 3 sono poi identicamente nulle.

Questo esempio suggerisce che applicando lo stesso numero di volte le singole derivate del primo ordine, si ottiene lo stesso risultato (per esempio, $D_{122} f = D_{212} f$). Vedremo tra poco un esempio che mostra che cio' è falso in generale. Ma vedremo anche che diventa vero sotto opportune ulteriori condizioni.

Esempio 2.2. Sia

$$\begin{cases} f : \mathbb{R}^2 \to \mathbb{R}, \\ f(x_1, x_2) = \begin{cases} \frac{x_1 x_2 (x_1^2 - x_2^2)}{x_1^2 + x_2^2} & \text{se} \quad (x_1, x_2) \neq (0,0), \\ 0 & \text{se} \quad (x_1, x_2) = (0,0). \end{cases} \end{cases}$$

E' facile verificare che, $\forall x = (x_1, x_2)$ in $\mathbb{R}^2$, si ha

$$D_1 f(x) = \begin{cases} \frac{x_1^4 x_2 - x_2^5 + 4x_1^2 x_2^3}{(x_1^2 + x_2^2)^2} & \text{se} \quad (x_1, x_2) \neq (0,0), \\ 0 & \text{se} \quad (x_1, x_2) = (0,0), \end{cases}$$

$$D_2 f(x) = \begin{cases} \frac{x_1^5 - x_1 x_2^4 - 4x_1^3 x_2^2}{(x_1^2 + x_2^2)^2} & \text{se} \quad (x_1, x_2) \neq (0,0), \\ 0 & \text{se} \quad (x_1, x_2) = (0,0). \end{cases}$$

Segue che

$$\begin{aligned} D_{12}f(0,0) &= \lim_{t\to 0} \tfrac{D_2f(t,0)-D_2f(0,0)}{t} = \lim_{t\to 0} \tfrac{t}{t} \\ &= 1, \end{aligned}$$

mentre

$$\begin{aligned} D_{21}f(0,0) &= \lim_{t\to 0} \tfrac{D_1f(0,t)-D_1f(0,0)}{t} = \lim_{t\to 0} \tfrac{-t}{t} \\ &= -1. \end{aligned}$$

Dunque $D_{12}f(0,0) \neq D_{21}f(0,0)$.

Tuttavia, vale il seguente classico teorema di Schwarz:

Teorema 2.1. *(di Schwarz) Siano A un aperto in $\mathbb{R}^n$ ($n \in \mathbb{N}$), $f : A \to \mathbb{R}$, $1 \leq i, j \leq n$. Supponiamo che, $\forall x \in A$, siano definite e a valori reali $D_if(x)$, $D_jf(x)$, $D_{ij}f(x)$. Sia poi x^0 in A tale che $D_{ij}f$ e' continua in x^0. Allora esiste $D_{ji}f(x^0)$ e*

$$D_{ji}f(x^0) = D_{ij}f(x^0).$$

Per la dimostrazione, utilizzeremo il seguente

Lemma 2.1. *Siano A un aperto in $\mathbb{R}^2$, $f : A \to \mathbb{R}$. Supponiamo che, $\forall y = (y_1, y_2) \in A$, siano definite e a valori reali $D_1f(y)$, $D_2f(y)$, $D_{12}f(y)$. Siano poi $y^0 = (y_1^0, y_2^0)$ in A, $h, k \in \mathbb{R} \setminus \{0\}$ tali che*

$$Q(h,k) := [\min\{y_1^0, y_1^0 + h\}, \max\{y_1^0, y_1^0 + h\}] \times [\min\{y_2^0, y_2^0 + k\}, \max\{y_2^0, y_2^0 + k\}] \subseteq A.$$

Allora esiste $c \in Q(h,k)$ tale che

$$\Delta(h,k) := f(y_1^0 + h, y_2^0 + k) - f(y_1^0 + h, y_2^0) - f(y_1^0, y_2^0 + k) + f(y_1^0, y_2^0) = D_{12}f(c)hk.$$

Dimostrazione Poniamo

$$\begin{cases} g : [\min\{y_2^0, y_2^0 + k\}, \max\{y_2^0, y_2^0 + k\}] \to \mathbb{R}, \\ \\ g(z) = f(y_1^0 + h, z) - f(y_1^0, z). \end{cases}$$

g è derivabile in $[\min\{y_2^0, y_2^0 + k\}, \max\{y_2^0, y_2^0 + k\}]$ e

$$g'(z) = D_2f(y_1^0 + h, z) - D_2f(y_1^0, z) \quad \forall z \in [\min\{y_2^0, y_2^0 + k\}, \max\{y_2^0, y_2^0 + k\}]$$

Dunque, per il teorema del valor medio, esiste $c_2 \in [\min\{y_2^0, y_2^0 + k\}, \max\{y_2^0, y_2^0 + k\}]$ tale che

$$\Delta(h,k) = g(y_2^0 + k) - g(y_2^0) = g'(c_2)k = [D_2f(y_1^0 + h, c_2) - D_2f(y_1^0, c_2)]k.$$

Poniamo ora

$$\begin{cases} m : [\min\{y_1^0, y_1^0 + h\}, \max\{y_1^0, y_1^0 + h\}] \to \mathbb{R}, \\ \\ m(v) = D_2f(v, c_2)k. \end{cases}$$

m è derivabile in $[\min\{y_1^0, y_1^0 + h\}, \max\{y_1^0, y_1^0 + h\}]$ e

$$m'(v) = D_{12}f(v, c_2)k \quad \forall v \in [\min\{y_1^0, y_1^0 + h\}, \max\{y_1^0, y_1^0 + h\}].$$

Quindi, ancora per il teorema del valor medio, esiste $c_1 \in [\min\{y_1^0, y_1^0 + h\}, \max\{y_1^0, y_1^0 + h\}]$ tale che

$$\Delta(h,k) = m(y_1^0 + h) - m(y_1^0) = m'(c_1)h = D_{12}f(c_1, c_2)hk.$$

Segue dunque il risultato, con $c = (c_1, c_2)$.

$\square$

Dimostrazione del teorema 2.1 Poniamo

$$B = \{(y_1, y_2) \in \mathbb{R}^2 : (x_1^0, \ldots, x_{i-1}^0, y_1, x_{i+1}^0, \ldots, x_{j-1}^0, y_2, x_{j+1}^0, \ldots, x_n^0) \in A\}.$$

Lasciamo verificare al lettore che B è aperto in $\mathbb{R}^2$. Poniamo

$$\begin{cases} g : B \to \mathbb{R}, \\ \\ g(y_1, y_2) = f(x_1^0, \ldots, x_{i-1}^0, y_1, x_{i+1}^0, \ldots, x_{j-1}^0, y_2, x_{j+1}^0, \ldots, x_n^0). \end{cases}$$

E' chiaro che sono definite in B le derivate $D_1 g$, $D_2 g$, $D_{12} g$. Inoltre $\forall (y_1, y_2) \in B$,

$$D_1 g(y_1, y_2) = D_i f(x_1^0, \ldots, x_{i-1}^0, y_1, x_{i+1}^0, \ldots, x_{j-1}^0, y_2, x_{j+1}^0, \ldots, x_n^0),$$

$$D_2 g(y_1, y_2) = D_j f(x_1^0, \ldots, x_{i-1}^0, y_1, x_{i+1}^0, \ldots, x_{j-1}^0, y_2, x_{j+1}^0, \ldots, x_n^0),$$

$$D_{12} g(y_1, y_2) = D_{ij} f(x_1^0, \ldots, x_{i-1}^0, y_1, x_{i+1}^0, \ldots, x_{j-1}^0, y_2, x_{j+1}^0, \ldots, x_n^0). \tag{2.1}$$

Quindi, se esiste $D_{21} g(x_i^0, x_j^0)$, si ha

$$D_{21} g(x_i^0, x_j^0) = \lim_{t \to 0} \frac{D_1 g(x_i^0, x_j^0 + t) - D_1 g(x_i^0, x_j^0)}{t} =$$

$$\lim_{t \to 0} \frac{D_i f(x_1^0, \ldots, x_{i-1}^0, x_i^0, x_{i+1}^0, \ldots, x_{j-1}^0, x_j^0 + t, x_{j+1}^0, \ldots, x_n^0) - D_i f(x_1^0, \ldots, x_{i-1}^0, x_i^0, x_{i+1}^0, \ldots, x_{j-1}^0, x_j^0, x_{j+1}^0, \ldots, x_n^0)}{t}$$

e percio' esiste $D_{ji} f(x^0)$ e coincide con $D_{21} g(x_i^0, x_j^0)$. E' dunque sufficiente provare che esiste $D_{21} g(x_i^0, x_j^0)$ e coincide in $D_{12} g(x_i^0, x_j^0) = D_{ij} f(x^0)$.

Siano h_0, k_0 positivi, tali che $[x_i^0 - h_0, x_i^0 + h_0] \times [x_j^0 - k_0, x_j^0 + k_0] \subseteq B$. Sia $\epsilon \in \mathbb{R}^+$. Poiche' $D_{ij} f$ è continua in x^0, per (2.1) $D_{12} g$ è continua in (x_i^0, x_j^0). Quindi possiamo prendere h_0 e k_0 in modo che

$$|D_{12} g(c) - D_{12} g(x_i^0, x_j^0)| < \epsilon$$

$\forall c \in [x_i^0 - h_0, x_i^0 + h_0] \times [x_j^0 - k_0, x_j^0 + k_0]$. Siano $h, k \in \mathbb{R}$, tali che $0 < |h| \le h_0$, $0 < |k| \le k_0$. Allora, per il lemma 2.1 esiste c (dipendente da (h,k)) in

$$[\min\{x_i^0, x_i^0 + h\}, \max\{x_i^0, x_i^0 + h\}] \times [\min\{x_j^0, x_j^0 + k\}, \max\{x_j^0, x_j^0 + k\}]$$

$$\subseteq [x_i^0 - h_0, x_i^0 + h_0] \times [x_j^0 - k_0, x_j^0 + k_0]$$

tale che

$$\frac{g(x_i^0 + h, x_j^0 + k) - g(x_i^0 + h, x_j^0) - g(x_i^0, x_j^0 + k) + g(x_i^0, x_j^0)}{hk} = D_{12} g(c).$$

Segue

$$\left|\frac{g(x_i^0 + h, x_j^0 + k) - g(x_i^0 + h, x_j^0) - g(x_i^0, x_j^0 + k) + g(x_i^0, x_j^0)}{hk} - D_{12} g(x_i^0, x_j^0)\right| < \epsilon.$$

Passando al limite per h che tende a 0, otteniamo

$$|\frac{D_1 g(x_i^0, x_j^0 + k) - D_1 g(x_i^0, x_j^0)}{k} - D_{12} g(x_i^0, x_j^0)| \leq \epsilon. \tag{2.2}$$

Quindi $\forall \epsilon \in \mathbb{R}^+$ esiste $k_0 \in \mathbb{R}^+$ tale che, se $0 < |k| < k_0$, vale (2.2). Cio' vuol dire esattamente che esiste

$$D_{21} g(x_i^0, x_j^0) = \lim_{k \to 0} \frac{D_1 g(x_i^0, x_j^0 + k) - D_1 g(x_i^0, x_j^0)}{k} = D_{12} g(x_i^0, x_j^0).$$

□

Osservazione 2.1. Nella dimostrazione del teorema 2.1 abbiamo implicitamente supposto $i < j$. Al lettore dovrebbe essere chiaro come modicarla nel caso $j < i$.

Passiamo ora alla definizione delle classi $C^k(A)$.

Definizione 2.2. *Siano A un aperto di $\mathbb{R}^n$ ($n \in \mathbb{N}$), $f : A \to \mathbb{R}$. Diremo che f è di classe C^k ($k \in \mathbb{N}$) se possiede a valori reali tutte le derivate parziali di ordine non superiore a k in ogni punto di A e tali derivate sono continue in A. Scriveremo, per indicare tale eventualità, $f \in C^k(A)$. Se $f \in C^k(A)$ per ogni $k \in \mathbb{N}$ (e quindi possiede derivate parziali continue di ogni ordine), scriveremo $f \in C^\infty(A)$.*

Poniamo anche $C^0(A) := C(A)$.

Se $f \in C^k(A)$ per ogni $k \in \mathbb{N}$ (e quindi possiede derivate parziali continue di ogni ordine), scriveremo $f \in C^\infty(A)$.

Dal teorema di Schwarz segue subito la seguente

Proposizione 2.1. *Siano A un aperto di $\mathbb{R}^n$ ($n \in \mathbb{N}$), $k \in \mathbb{N}$, $k \geq 2$, $f \in C^k(A)$. Siano $j_1, j_2, \dots, j_k \in \{1, \dots, n\}$, non necessariamente a due a due distinti. Allora, $\forall x \in A$,*

$$D_{j_1 j_2 \dots j_k} f(x) = D_{j_2 j_1 \dots j_k} f(x).$$

Dimostrazione Se $k = 2$, la conclusione segue subito dal teorema di Schwarz.

Supponiamo $k > 2$. Allora $g = D_{j_3 \dots j_k} f \in C^2(A)$. Dunque,

$$D_{j_1 j_2 \dots j_k} f = D_{j_1 j_2} g = D_{j_2 j_1} g = D_{j_2 j_1 \dots j_k} f.$$

□

Osservazione 2.2. Se $f \in C^k(A)$, con $k \geq 2$, due derivate qualunque di ordine non superiore a k, che si ottengono applicando lo stesso numero di volte le singole derivate parziali del primo ordine, coincidono. Lo verifichiamo nel caso $f \in C^3(A)$, con A aperto in $\mathbb{R}^3$, facendo vedere che

$$D_{123} f = D_{321} f.$$

Infatti, per la proposizione 2.1,

$$D_{123} f = D_{213} f.$$

Ma

$$D_{213} f = D_2(D_{13} f).$$

Poiche' $f \in C^3(A)$, $f \in C^2(A)$. Dunque $D_{13} f = D_{31} f$. Quindi,

$$D_{213} f = D_2(D_{13} f) = D_2(D_{31} f) = D_{231} f.$$

Infine, usando il fatto che $D_1 f \in C^2(A)$, concludiamo che

$$D_{231} f = D_{23}(D_1 f) = D_{32}(D_1 f) = D_{321} f.$$

Introduciamo ora la notazione dei cosi' detti multiindici.

Un multiindice con n elementi ($n \in \mathbb{N}$) è un elemento di $\mathbb{N}_0^n$. Se $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}_0^n$, $\beta = (\alpha_1, ..., \alpha_n) \in \mathbb{N}_0^n$ poniamo

$$|\alpha| := \alpha_1 + \cdots + \alpha_n. \tag{2.3}$$

Introduciamo ora la seguente notazione: siano $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}_0^n$, $A \subseteq \mathbb{R}^n$ aperto, $f \in C^k(A)$ ($k \in \mathbb{N}$). $|\alpha| \le k$. Poniamo

$$D^\alpha f = \frac{\partial^{\alpha_1 + ... + \alpha_n} f}{\partial x_1^{\alpha_1} ... \partial x_n^{\alpha_n}} := \underbrace{D_1 ... D_1}_{\alpha_1} ... \underbrace{D_n ... D_n}_{\alpha_n} f. \tag{2.4}$$

E' chiaro, che, se $j_1, \dots, j_k \in \{1, \dots, n\}$,

$$D_{j_1 \dots j_k} f = D^\alpha f,$$

con α_i cardinalita' di $\{l \in \{1, \dots, k\} : j_l = i\}$.

Il seguente risultato segue facilmente, per iterazione dal risultato dell'esercizio 1.4:

Teorema 2.2. *Siano $A \subseteq \mathbb{R}^n$ aperto, $k \in \mathbb{N}$, $f, g \in C^k(A)$, $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}_0^n$, con $|\alpha| \le k$, $c \in \mathbb{R}$. Allora:*

(I) $f + g, fg, cf \in C^k(A)$;

(II) $D^\alpha(f + g) = D^\alpha f + D^\alpha g$, $D^\alpha(cf) = cD^\alpha f$;

(III) quindi $C^k(A)$ (con le solite operazioni di somma di funzioni e moltiplicazione di uno funzione con un numero reale) è *uno spazio vettoriale su $\mathbb{R}$;*

(IV) inoltre l'applicazione $f \to D^\alpha f$ e' lineare da $C^k(A)$ a $C^{k-|\alpha|}(A)$.

Dimostrazione Vedi l'esercizio 2.1. □

Esercizio 2.1. (I) Dimostrare il teorema 2.2.

(II) Verificare che, se $f, g \in C^k(A)$, $D^\alpha(fg)$ è definita ed è combinazione lineare dei prodotti $D^\beta f D^\gamma g$, con $\beta, \gamma \in \mathbb{N}_0^n$, $|\beta| + |\gamma| = |\alpha|$. Ragionare per induzione su $|\alpha|$. Dedurre che, se $f, g \in C^k(A)$, $fg \in C^k(A)$.

(III) Sia $g \in C^k(A)$, con $g(x) \neq 0 \ \forall x \in A$. Sia $\alpha \in \mathbb{N}_0^n$, con $|\alpha| \le k$. Verificare che $D^\alpha(g^{-1})$ è definita ed è combinazione lineare di funzioni della forma $g^{-m} D^{\beta_1} g \dots D^{\beta_l} g$, con $m, l \in \mathbb{N}$, $\beta_1, \dots, \beta_l \in \mathbb{N}_0^n$, $\max_{1 \le j \le l} |\beta_j| \le m$.

Dedurre che $g^{-1} \in C^k(A)$ e, se $f \in C^k(A)$, $\frac{f}{g} \in C^k(A)$.

3 La formula di Taylor

Le motivazioni dell'introduzione della formula di Taylor per funzioni di piu' variabili sono simili a quelle viste per funzioni di una sola variabile: vogliamo approssimare localmente una funzione opportunamente regolare (dotata, cioè , di un congruo corredo di derivate) con una funzione dalla struttura semplice, che permetta quindi di studiarne il comportamento con relativa facilita'.

Veniamo allora subito al risultato principale, che andiamo a enunciare e dimostreremo nel seguito:

Teorema 3.1. *(Formula di Taylor) Siano A un aperto in $\mathbb{R}^n$, $k \in \mathbb{N}$, $f \in C^k(A)$, $x^0 = (x_1^0, ..., x_n^0) \in A$. Introduciamo il seguente polinomio di n variabili reali:*

$$\begin{aligned} & P(x_1, ..., x_n) \\ &= f(x^0) + \textstyle\sum_{j=1}^n D_j f(x^0)(x_j - x_j^0) + ... \\ &+ ... + \tfrac{1}{i!} \textstyle\sum_{j_1=1}^n ... \sum_{j_i=1}^n D_{j_1 ... j_i} f(x^0)(x_{j_1} - x_{j_1}^0)...(x_{j_i} - x_{j_i}^0) + \\ & ... + \tfrac{1}{k!} \textstyle\sum_{j_1=1}^n ... \sum_{j_k=1}^n D_{j_1 ... j_k} f(x^0)(x_{j_1} - x_{j_1}^0)...(x_{j_k} - x_{j_k}^0). \end{aligned} \tag{3.1}$$

Chiamiamo P **polinomio di Taylor di grado non superiore a k con punto iniziale x^0 per f.**

Allora $f - P = o(\|x - x^0\|^k)$ per $x \to x^0$, nel senso che vale

$$\lim_{x \to x^0} \frac{f(x) - P(x)}{\|x - x^0\|^k} = 0. \tag{3.2}$$

Osservazione 3.1. Nel caso particolare $n = 1$ $f - P = o((x - x^0)^k)$ per $x \to x^0$. In questo caso si verifica facilmente che le condizioni $f - P = o((x - x^0)^k)$ e $f - P = o(|x - x^0|^k)$ sono equivalenti.

Alla dimostrazione del teorema premettiamo il seguente ben noto risultato che, nel caso $n = 1$, fornisce una rappresentazione esplicita del resto nell'ambito della formula di Taylor (formula di Taylor con resto di Lagrange):

Lemma 3.1. *Siano I in intervallo con interno non vuoto in $\mathbb{R}$, $k \in \mathbb{N}$, $F \in C^k(I)$. Siano t_0 e t appartenenti a I, $t \neq t_0$. Allora esiste $c \in]\min\{t_0, t\}, \max\{t_0, t\}[$ tale che*

$$F(t) = \sum_{j=0}^{k-1} \frac{F^{(j)}(t_0)}{j!}(t - t_0)^j + \frac{F^{(k)}(c)}{k!}(t - t_0)^k.$$

Per una dimostrazione, vedi (ad esempio) [2], teorema 4.7.

Dimostrazione del teorema 3.1. Sia $r \in \mathbb{R}^+$, tale che $B(x^0, r) \subseteq A$. Per $x \in B(x^0, r)$, poniamo

$$\begin{cases} F : [0, 1] \to \mathbb{R}, \\ F(t) = f(x^0 + t(x - x^0)). \end{cases} \tag{3.3}$$

Applicando l'osservazione 1.8, si ha che $F \in C^1([0, 1])$ e

$$F'(t) = \sum_{j=1}^n D_j f(x^0 + t(x - x^0))(x_j - x_j^0), \quad t \in [0, 1].$$

Se $k > 1$, il metodo di derivazione si puo' iterare: se $f \in C^2(A)$, si ricava

$$F''(t) = \sum_{j_1=1}^n \sum_{j_2=1}^n D_{j_1, j_2} f(x^0 + t(x - x^0))(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0), \quad t \in [0, 1].$$

In generale, se $f \in C^k(A)$, si vede che $F \in C^k([0, 1])$ e, per $i = 1, \dots, k$, $t \in [0, 1]$,

$$F^{(i)}(t) = \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_i=1}^n D_{j_1 j_2 ... j_i} f(x^0 + t(x - x^0))(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0) \dots (x_{j_i} - x_{j_i}^0).$$

Di conseguenza, applicando il lemma 3.1, otteniamo, per un certo $c(x) \in]0,1[$,

$$f(x) = F(1) = \sum_{i=1}^{k-1} \frac{F^{(i)}(0)}{i!} + \frac{F^{(k)}(c(x))}{k!} = f(x^0) + \sum_{j=1}^{n} D_j f(x^0)(x_j - x_j^0)$$

$$+\frac{1}{2} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} D_{j_1 j_2} f(x^0)(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0)$$

$$+ \cdots + \frac{1}{i!} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \dots \sum_{j_i=1}^{n} D_{j_1 j_2 \dots j_i} f(x^0)(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0) \dots (x_{j_i} - x_{j_i}^0)$$

$$+ \cdots + \frac{1}{(k-1)!} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \dots \sum_{j_{k-1}=1}^{n} D_{j_1 j_2 \dots j_{k-1}} f(x^0)(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0) \dots (x_{j_{k-1}} - x_{j_{k-1}}^0)$$

$$+\frac{1}{k!} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \dots \sum_{j_k=1}^{n} D_{j_1 j_2 \dots j_k} f(x^0 + c(x)(x - x^0))(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0) \dots (x_{j_k} - x_{j_k}^0)$$

$$= P(x) + \frac{1}{k!} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \dots \sum_{j_k=1}^{n} [D_{j_1 j_2 \dots j_k} f(x^0 + c(x)(x - x^0)) - D_{j_1 j_2 \dots j_k} f(x^0)]$$

$$\times (x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0) \dots (x_{j_k} - x_{j_k}^0) \tag{3.4}$$

con P come in (3.1). Resta da provare che si ha

$$\lim_{x \to x^0} \frac{f(x) - P(x)}{\|x - x^0\|^k} = 0.$$

Siano ϵ e η in $\mathbb{R}^+$. Poiché la funzione f è di classe C^k, esiste $\delta(\eta) > 0$, tale che $B(x^0, \delta(\eta)) \subseteq A$ e, se $y \in B(x^0, \delta(\eta)) \subseteq A$, per ogni $j_1, \dots, j_k \in \{1, ..., n\}$, si ha

$$|D_{j_1 \dots j_k} f(y) - D_{j_1 \dots j_k} f(x^0)| < \eta. \tag{3.5}$$

Se $x \in B(x^0, \delta(\eta))$, si ha

$$\begin{aligned} \|(x^0 + c(x)(x - x^0)) - x^0\| &= c(x)\|x - x^0\| \\ &\leq \|x - x^0\| < \delta(\eta). \end{aligned}$$

Da (3.4) segue allora

$$\begin{aligned} |f(x) - P(x)| &\leq \frac{\eta}{k!} \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} |x_{j_1} - x_{j_1}^0| \dots |x_{j_k} - x_{j_k}^0| \\ &\leq \frac{n^k}{k!} \eta \|x - x^0\|^k. \end{aligned} \tag{3.6}$$

Da (3.6) segue subito, per $x \in B(x^0, \delta(\eta)) \setminus \{x^0\}$,

$$\frac{|f(x) - P(x)|}{\|x - x^0\|^k} \leq \frac{n^k}{k!} \eta,$$

da cui

$$\frac{|f(x) - P(x)|}{\|x - x^0\|^k} < \epsilon$$

se

$$\frac{n^k}{k!} \eta < \epsilon.$$

□

Osservazione 3.2. Il polinomio P che gode della proprieta' 3.2 è univocamente determinato. Cio' è ben noto nel caso $n = 1$. Nel caso generale, sia Q un polinomio di grado non superiore a k, tale che $f - Q = o(\|x - x^0\|^k)$ per $x \to x^0$. Allora, per differenza, $P - Q$ è un polinomio di grado non superiore a k e $P - Q = o(\|x - x^0\|^k)$ per $x \to x^0$. In particolare, per ogni $h \in \mathbb{R}^n$, posto $r(t) := P(x^0 + th) - Q(x^0 + th)$, r è un polinomio in $\mathbb{R}$ di grado non superiore a k tale che

$$\lim_{t \to 0} \frac{r(t)}{t^k} = 0.$$

Cio' implica (per il caso $n = 1$) che il polinomio di Taylor di grado non superiore a k per r in 0 è quello nullo. Ma r stesso è un polinomio di grado non superiore a k. Quindi $r(t) = 0 \ \forall t \in \mathbb{R}$. Di qui $P(x^0 + th) = Q(x^0 + th) \ \forall t \in \mathbb{R}$, $\forall h \in \mathbb{R}^n$. Cio' implica immediatamente $P(x) = Q(x)$ $\forall x \in \mathbb{R}^n$.

Osserviamo che ogni polinomio P di grado non superiore a k in $\mathbb{R}^n$ puo' essere rappresentato nella forma

$$P(x) = \sum_{\alpha \in N_0^n, |\alpha| \leq k} c_\alpha (x - x^0)^\alpha,$$

con la notazione, per $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$,

$$h^\alpha := h_1^{\alpha_1} \ldots h_n^{\alpha_n}, \tag{3.7}$$

e la convenzione $0^0 = 1$. Il coefficiente c_α è univocamente determinato da P, essendo

$$c_\alpha = \frac{D^\alpha P(x^0)}{\alpha!}, \tag{3.8}$$

con

$$\alpha! := \alpha_1! \ldots \alpha_n! \tag{3.9}$$

(vedi l'esercizio 3.1).

Osservazione 3.3. Con riferimento alla formula (3.1), utilizzando la notazione (3.7), si ha

$\frac{1}{i!} \sum_{j_1=1}^n ... \sum_{j_i=1}^n D_{j_1...j_i} f(x^0)(x_{j_1} - x_{j_1}^0)...(x_{j_i} - x_{j_i}^0) = \frac{1}{i!} \sum_{\alpha \in N_0^n, |\alpha| = i} n(\alpha) D^\alpha f(x^0)(x - x^0)^\alpha,$

con $n(\alpha)$ cardinalita' dell'insieme delle $i-$ple $(j_1, \ldots, j_i)$ a termini in $\{1, \ldots, n\}$ con α_1 termini uguali a 1, $\ldots, \alpha_n$ termini uguali a n.

Ricordiamo che, se $l, m \in \mathbb{N}_0$ e $l \leq m$, il numero di sottoinsiemi con l elementi di un insieme con m elementi è

$$\binom{m}{l} = \frac{m!}{l!(m-l)!};$$

stabilito il fatto che α_1 termini tra i j_i devono essere 1, α_2 termini tra i j_i devono essere 2, ecc. con $\alpha_1 + \cdots + \alpha_n = i$, la posizione degli elementi j_i con $j_i = 1$ puo' essere scelta in $\binom{i}{\alpha_1}$ modi diversi; per ciascuna di queste scelte, la posizione degli elementi j_i con $j_i = 2$ puo' essere scelta in $\binom{i-\alpha_1}{\alpha_2}$ modi diversi, eccetera. Dunque

$$n(\alpha) = \binom{i}{\alpha_1}\binom{i-\alpha_1}{\alpha_2} \ldots \binom{i-(\alpha_1+\ldots+\alpha_{n-1})}{\alpha_n}$$

$$= \frac{i!}{\alpha_1!(i-\alpha_1)!} \frac{(i-\alpha_1)!}{\alpha_2!(i-\alpha_1-\alpha_2)!} \cdots \frac{[i-(\alpha_1+\cdots+\alpha_{n-1}]!}{\alpha_n![i-(\alpha_1+\cdots+\alpha_{n-1}+\alpha_n)]!} = \frac{i!}{\alpha_1!\ldots\alpha_n!}.$$

Concludiamo che

$$P(x) = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq k} \frac{D^\alpha f(x^0)}{\alpha!}(x - x^0)^\alpha.$$

Esempio 3.1. Sia

$$\begin{cases} f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}, \\ f(x_1, x_2) = x_1^{x_2}, \quad (x_1, x_2) \in \mathbb{R}^+ \times \mathbb{R}. \end{cases} \tag{3.10}$$

Poniamo $x^0 = (x_1^0, x_2^0) = (1, 0)$ e scriviamo il polinomio di Taylor di grado non superiore a tre di f con punto iniziale x^0. Lasciamo al lettore la verifica che f è (almeno) di classe C^3 (esercizio 3). Si ha, applicando il teorema 3.1 e l'osservazione 2.2:

$$\begin{aligned} f(x_1, x_2) &= f(x^0) + \textstyle\sum_{j=1}^{2} D_j f(x^0)(x_j - x_j^0) + \frac{1}{2} \sum_{j_1=1}^{2} \sum_{j_2=1}^{2} D_{j_1 j_2} f(x^0) \\ &\times(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0) + \tfrac{1}{6} \textstyle\sum_{j_1=1}^{2} \sum_{j_2=1}^{2} \sum_{j_3=1}^{2} D_{j_1 j_2 j_3} f(x^0) \\ &(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0)(x_{j_3} - x_{j_3}^0) + r(x) \\ &= f(1, 0) + D_1 f(1, 0)(x_1 - 1) + D_2 f(1, 0)x_2 + \tfrac{1}{2}[D_{11} f(1, 0) \\ &(x_1 - 1)^2 + D_{12} f(1, 0)(x_1 - 1)x_2 + D_{21} f(1, 0)(x_1 - 1)x_2 \\ &+D_{22} f(1, 0)x_2^2] + \tfrac{1}{6}[D_{111} f(1, 0)(x_1 - 1)^3 + D_{112} f(1, 0) \\ &\times(x_1 - 1)^2 x_2 + D_{121} f(1, 0)(x_1 - 1)^2 x_2 + D_{122} f(1, 0) \\ &\times(x_1 - 1)x_2^2 + D_{211} f(1, 0)(x_1 - 1)^2 x_2 + D_{212} f(1, 0)(x_1 - 1)x_2^2 \\ &+D_{221} f(1, 0)(x_1 - 1)x_2^2 + D_{222} f(1, 0)x_2^3] + r(x) \\ &= f(1, 0) + \tfrac{\partial f}{\partial x_1}(1, 0)(x_1 - 1) + \tfrac{\partial f}{\partial x_2}(1, 0)x_2 + \tfrac{1}{2}\tfrac{\partial^2 f}{\partial x_1^2}(1, 0) \\ &\times(x_1 - 1)^2 + \tfrac{\partial^2 f}{\partial x_1 \partial x_2}(1, 0)(x_1 - 1)x_2 + \tfrac{1}{2}\tfrac{\partial^2 f}{\partial x_2^2}(1, 0)x_2^2 \\ &+\tfrac{1}{6}\tfrac{\partial^3 f}{\partial x_1^3}(1, 0)(x_1 - 1)^3 + \tfrac{1}{2}\tfrac{\partial^3 f}{\partial x_1^2 \partial x_2}(1, 0)(x_1 - 1)^2 x_2 \\ &+\tfrac{1}{2}\tfrac{\partial^3 f}{\partial x_1 \partial x_2^2}(1, 0)(x_1 - 1)x_2^2 + \tfrac{1}{6}\tfrac{\partial^3 f}{\partial x_2^3}(1, 0)x_2^3 + r(x), \end{aligned} \tag{3.11}$$

con $\lim\limits_{x \to x^0} \frac{r(x)}{\|x - x^0\|^3} = 0$.

Si ha, per $(x_1, x_2) \in \mathbb{R}^+ \times \mathbb{R}$,

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = x_2 x_1^{x_2 - 1}, \frac{\partial f}{\partial x_2}(x_1, x_2) = x_1^{x_2} \ln(x_1),$$

$$\frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) = x_2(x_2 - 1)x_1^{x_2 - 2}, \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) = x_1^{x_2 - 1}(1 + x_2 \ln(x_1)),$$

$$\frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) = x_1^{x_2} \ln^2(x_1), \frac{\partial^3 f}{\partial x_1^3}(x_1, x_2) = x_2(x_2 - 1)(x_2 - 2)x_1^{x_2 - 3},$$

$$\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(x_1, x_2) = (x_2 - 1)x_1^{x_2 - 2} + x_2[(x_2 - 1)x_1^{x_2 - 2} \ln(x_1) + \frac{x_1^{x_2 - 1}}{x_1}],$$

$$\frac{\partial^3 f}{\partial x_1 \partial x_2^2}(x_1,x_2) = x_1^{x_2-1}(x_2 \ln^2(x_1) + 2\ln(x_1)),$$

$$\frac{\partial^3 f}{\partial x_2^3}(x_1,x_2) = x_1^{x_2} \ln^3(x_1).$$

Segue

$$f(1,0) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(1,0) = 1,$$

$$\frac{\partial f}{\partial x_1}(1,0) = \frac{\partial f}{\partial x_2}(1,0) = \frac{\partial^2 f}{\partial x_1^2}(1,0) = \frac{\partial^2 f}{\partial x_2^2}(1,0) = \frac{\partial^3 f}{\partial x_1^3}(1,0)$$

$$= \frac{\partial^3 f}{\partial x_1 \partial x_2^2}(1,0) = \frac{\partial^3 f}{\partial x_2^3}(1,0) = 0,$$

$$\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(1,0) = -1,$$

da cui, per (3.11),

$$f(x_1,x_2) \quad = 1 + (x_1-1)x_2 - \tfrac{1}{2}(x_1-1)^2 x_2 + r(x).$$

Esercizio 3.1. Dimostrare nei dettagli le affermazioni contenute nell'osservazione 3.2.

Esercizio 3.2. Verificare che la funzione f definita in (3.10) è di classe C^3.

Cominciare col verificare che f è continua. A tale scopo, si puo' osservare che

$$f(x_1 x_2) = \exp(x_2 \ln(x_1))$$

e la continuita' puo' seguire dai teoremi 3.6 (III) e 3.7 (II) (in [SM]).

Esercizio 3.3. Sia $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x_1,x_2,x_3) = x_1 x_2 x_3^2$. Scrivere il polinomio di Taylor di f con punto iniziale $(1,-1,1)$ di grado non superiore a 3.

Stesso esercizio per $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x_1,x_2,x_3) = x_2^2 e^{x_1^2} + x_3 e^{x_2 x_3} + x_1^2 + 2x_2 - 3x_1$.

4 Estremanti relativi, punti critici, punti di sella

Definizione 4.1. *Siano $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, $x^0 \in A$. Diremo che x^0 è un* **punto di minimo (massimo) relativo** *per f se esiste $\delta \in \mathbb{R}^+$ tale che $f(x^0) \le f(x)$ (risp. $f(x^0) \ge f(x)$) $\forall x \in A \cap B(x^0,\delta)$.*

Chiameremo **estremanti relativi di** *f i suoi punti di massimo o minimo relativo.*

Introduciamo ora la nozione di punto critico.

Definizione 4.2. *Siano $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $x^0 \in \mathring{A}$, f differenziablie in x^0 . Diremo che A è un* **punto critico** *di f se $\nabla f(x^0) = O$.*

Teorema 4.1. *Siano $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $x^0 \in \mathring{A}$, f differenziablie in x^0 . Supponiamo che x^0 sia un estremante relativo di f. Allora x^0 è un punto critico di f.*

Dimostrazione Supponiamo che (ad esempio) x^0 sia un punto di minimo relativo per f. Allora, poiche' $x^0 \in \mathring{A}$, esiste $\delta \in \mathbb{R}^+$ tale che $B(x^0, r) \subseteq A$ e $f(x^0) \leq f(x)$ $\forall B(x^0, r)$. Quindi, se $h \in \mathbb{R}^n$, $t \in \mathbb{R}$ e $t\|h\| < r$,

$$f(x^0) \leq f(x^0 + th).$$

Poniamo $g(t) = f(x^0 + th)$ (definita in $\{t \in \mathbb{R} : t\|h\| < r\}$). Per noti risultati validi per funzioni di una variabile reale, se g è derivabile in 0, $g'(0) = 0$. Ma

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

Per la proposizione 1.1, tale limite esiste ed è uguale a $\nabla f(x^0) \cdot h$. Quindi $\nabla f(x^0) \cdot h = 0$ $\forall h \in \mathbb{R}^n$. Cio' implica $\nabla f(x^0) = O$.

$\square$

Osservazione 4.1. Il risultato del teorema 4.1 non è invertibile. In altre parole, un punto critico non è necessariamente un estremante relativo. Vediamone un paio di esempi.

Si consideri $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$. Allora 0 è un punto critico di f ($f'(0) = 0$), ma non è un estremante relativo, perché, per ogni $r \in \mathbb{R}^+$, $\pm r/2 \in I(0, r)$ e

$$f(-r/2) = -r^3/8 < 0 = f(0) < r^3/8 = f(r/2).$$

Un altro esempio si ha ponendo

$$\begin{cases} f : \mathbb{R}^2 \to \mathbb{R}, \\ f(x_1, x_2) = x_1^2 - x_2^2. \end{cases}$$

$(0, 0)$ è un punto critico di f, ma per ogni $r \in \mathbb{R}^+$ i punti $(r/2, 0)$ e $(0, -r/2)$ appartengono a $B((0, 0), r)$ e si ha

$$f(0, -r/2) = -r^2/4 < 0 = f(0, 0) < r^2/4 = f(r/2, 0).$$

Dunque, $(0, 0)$ non è un estremante relativo di f.

Data $f \in C^1(A)$, con A aperto in $\mathbb{R}^n$, chiameremo **punti di sella** di f i suoi punti critici che non sono estremanti relativi.

Nella parte successiva di questa sezione studieremo, utilizzando come strumento principale la formula di Taylor, i punti critici di una funzione, cercando ulteriori condizioni necessarie (oltre al risultato del teorema 4.1) e condizioni sufficienti affinché siano estremanti relativi.

Siano $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$ aperto, $f \in C^2(A)$. Sia poi $x^0 \in A$ un punto critico di f. Allora, dalla formula di Taylor, otteniamo, per $x \in A$,

$$f(x) = f(x^0) + Q(x - x^0) + r(x), \tag{4.1}$$

con

$$\begin{cases} Q : \mathbb{R}^n \to \mathbb{R}, \\ Q(h_1, ..., h_n) = (1/2) \sum_{i=1}^n \sum_{j=1}^n D_{ij} f(x^0) h_i h_j, \end{cases} \tag{4.2}$$

e

$$\lim_{x \to x^0} \frac{r(x)}{\|x - x^0\|^2} = 0. \tag{4.3}$$

La funzione Q definita in (4.2) è una forma quadratica, nel senso della prossima definizione.

Definizione 4.3. *Una* **forma quadratica** *in* $\mathbb{R}^n$ *e' una funzione polinomiale* $Q : \mathbb{R}^n \to \mathbb{R}$, *della forma*

$$Q(h_1, ..., h_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j,$$

con $a_{ij} \in \mathbb{R}$ *(*$1 \leq i, j \leq n$*).*

Osservazione 4.2. Presentiamo qui alcune semplici osservazioni, riguardo alle forme quadratiche.

Prima di tutto, se Q è una forma quadratica in $\mathbb{R}^n$,

$$Q(O) = Q(0, ..., 0) = 0. \tag{4.4}$$

Inoltre, se $h = (h_1, ..., h_n) \in \mathbb{R}^n$ e $t \in \mathbb{R}$,

$$Q(th) = t^2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j = t^2 Q(h). \tag{4.5}$$

Infine, nella definizione 4.3 non è restrittivo supporre che, per $1 \leq i, j \leq n$, valga

$$a_{ij} = a_{ji}. \tag{4.6}$$

Qualora tale condizione non fosse soddisfatta, bastera' sostituire a_{ij} e a_{ji} con $(a_{ij} + a_{ji})/2$.

Passiamo ora a definire alcune classi di forme quadratiche.

Definizione 4.4. *Sia Q una forma quadratica in $\mathbb{R}^n$. Diremo che:*
(I) Q è **semidefinita positiva** *se* $Q(h) \geq 0 \ \forall h \in \mathbb{R}^n$*;*
(II) Q è **semidefinita negativa** *se* $Q(h) \leq 0 \ \forall h \in \mathbb{R}^n$*;*
(III) Q è **definita positiva** *se* $Q(h) > 0 \ \forall h \in \mathbb{R}^n \setminus \{O\}$*;*
(IV) Q è **definita negativa** *se* $Q(h) < 0 \ \forall h \in \mathbb{R}^n \setminus \{O\}$*.*

Esempio 4.1. Presentiamo alcuni esempi di forme quadratiche soddisfacenti le condizioni della definizione 4.4.

Sia

$$\begin{cases} Q_0 : \mathbb{R}^2 \to \mathbb{R}, \\ Q_0(h_1, h_2) = h_1^2. \end{cases} \tag{4.7}$$

Ovviamente, Q_0 è semidefinita positiva. Q_0 non è definita positiva, in quanto, se $h_2 \in \mathbb{R} \setminus \{0\}$, $Q_0(0, h_2) = 0$.

Sia

$$\begin{cases} Q_1 : \mathbb{R}^2 \to \mathbb{R}, \\ Q_1(h_1, h_2) = h_1^2 + h_2^2. \end{cases} \tag{4.8}$$

E' facile verificare che Q_1 è definita positiva.

Sia

$$\begin{cases} Q_2 : \mathbb{R}^2 \to \mathbb{R}, \\ Q_2(h_1, h_2) = -h_1^2. \end{cases} \tag{4.9}$$

Q_2 è semidefinita negativa, ma non definita negativa.

Sia

$$\begin{cases} Q_3 : \mathbb{R}^2 \to \mathbb{R}, \\ Q_3(h_1, h_2) = -h_1^2 - h_2^2. \end{cases} \tag{4.10}$$

Q_3 è definita negativa.

Sia
$$\begin{cases} Q_4 : \mathbb{R}^2 \to \mathbb{R}, \\ Q_4(h_1, h_2) = h_1 h_2. \end{cases} \tag{4.11}$$
Q_4 non è semidefinita, in quanto $Q_4(1,1) = 1 > 0$, $Q_4(1,-1) = -1 < 0$.

Vediamo ora il primo risultato che lega il segno di una forma quadratica al problema principale di questa sezione.

Teorema 4.2. *Siano $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$ aperto, $f \in C^2(A)$, $x^0 \in A$. Allora:*
(I) se x^0 è un punto di minimo relativo per f, la forma quadratica Q definita in (4.2) è semidefinita positiva;
(II) se x^0 è un punto di massimo relativo per f, la forma quadratica Q definita in (4.2) è semidefinita negativa.

Dimostrazione Supponiamo, per fissare le idee, che x^0 sia un punto di minimo relativo per f. Verifichiamo che
$$Q(h) \geq 0, \qquad \forall h \in \mathbb{R}^n. \tag{4.12}$$
In base a (4.4), ci si puo' limitare al caso $h \in \mathbb{R}^n \setminus \{O\}$.

Sia $r > 0$, tale che $B(x^0, r) \subseteq A$ e $f(x^0) \leq f(x)$ $\forall x \in B(x^0, r)$. Poiché x^0 è un punto critico di f (teorema 4.1), f ammette lo sviluppo di Taylor (4.1). Se $t \in \mathbb{R}$ e $|t|\|h\| < r$, si ha allora
$$f(x^0 + th) = f(x^0) + Q(th) + r(x^0 + th) = f(x^0) + t^2 Q(h) + r(x^0 + th), \tag{4.13}$$
da cui, se $t \in]-r/\|h\|, r/\|h\|[\setminus\{0\}$,
$$Q(h) = (f(x^0 + th) - f(x^0))/t^2 - r(x^0 + th)/t^2. \tag{4.14}$$
Poiché $f(x^0 + th) - f(x^0) \geq 0$, da (4.14) segue, $\forall t \in]-r/\|h\|, r/\|h\|[\setminus\{0\}$,
$$Q(h) \geq -|r(x^0 + th)|/t^2 \tag{4.15}$$
D'altra parte,
$$\lim_{t \to 0} \frac{r(x^0 + th)}{t^2} = \lim_{t \to 0} \frac{r(x^0 + th)}{\|(x^0 + th) - x^0\|^2} \|h\|^2 = 0,$$
come conseguenza di (4.3). Sia allora $\epsilon \in \mathbb{R}^+$. Esiste $\delta(\epsilon) \in \mathbb{R}^+$, che possiamo supporre non superiore a $r/\|h\|$, tale che, se $t \in \mathbb{R} \setminus \{0\}$ e $|t| < \delta(\epsilon)$,
$$|r(x^0 + th)|/t^2 < \epsilon. \tag{4.16}$$
Con questa scelta di t, da (4.15) otteniamo
$$Q(h) > -\epsilon \tag{4.17}$$
e, poiché ϵ è un arbitrario elemento di $\mathbb{R}^+$, possiamo concludere che $Q(h) \geq 0$.
□

Esempio 4.2. Siano $A \subseteq \mathbb{R}^n$ aperto, $f \in C^2(A)$, $x^0 \in A$ tale che $\nabla f(x^0) = O$ e la forma quadratica Q definita in (4.2) è semidefinita. Allora, x^0 non è necessariamente un estremante relativo per f. Si consideri, in proposito, il seguente esempio: sia
$$\begin{cases} f : \mathbb{R}^2 \to \mathbb{R}, \\ f(x_1, x_2) = x_1^2 + x_2^3. \end{cases}$$

Si verifica facilmente che $x^0 = (0,0)$ è un punto critico e che la forma quadratica Q soddisfa

$$Q(h_1, h_2) = h_1^2, \quad \forall (h_1, h_2) \in \mathbb{R}^2.$$

Dunque, Q è semidefinita positiva. Pero' $(0,0)$ non è un estremante relativo, perché, $\forall r \in \mathbb{R}^+$, $(0, \pm r/2) \in B(x^0, r)$ e si ha

$$f(0, -r/2) = -r^3/8 < 0 = f(0,0) < r^3/8 = f(0, r/2).$$

Dunque, le condizioni $\nabla f(x^0) = O$ e Q semidefinita sono necessarie, ma non sufficienti, affinché x^0 sia un estremante relativo.

Presentiamo ora una condizione sufficiente per avere un estremante relativo.

Teorema 4.3. *Siano $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$ aperto, $f \in C^2(A)$, $x^0 \in A$. Allora:*
(I) se x^0 è un punto critico di f e la forma quadratica Q definita in (4.2) è definita positiva, x^0 è un punto di minimo relativo per f;
(II) se x^0 è un punto critico di f e la forma quadratica Q definita in (4.2) è definita negativa, x^0 è un punto di massimo relativo per f.

Dimostrazione Verifichiamo solo (I). La dimostrazione di (II) è analoga. Utilizziamo, ancora una volta, lo sviluppo di Taylor (4.1). Cominciamo allora con l'osservare che esiste $\nu \in \mathbb{R}^+$, tale che

$$Q(h) \geq \nu \|h\|^2, \quad \forall h \in \mathbb{R}^n. \tag{4.18}$$

Infatti, poiché $S := \{h \in \mathbb{R}^n : \|h\| = 1\}$ è chiuso e limitato e Q è continua, esiste

$$\nu := \min_S Q$$

e si ha $\nu > 0$, perché Q è definita positiva. Allora, se $h \in \mathbb{R}^n \setminus \{O\}$, da (4.5) segue

$$Q(h) = \|h\|^2 Q(h/\|h\|) \geq \nu \|h\|^2.$$

(4.18) è poi banalmente soddisfatta se $h = O$.

Da (4.3) ricaviamo invece che esiste $\delta \in \mathbb{R}^+$, che possiamo supporre tale che $B(x^0, \delta) \subseteq A$, tale che, se $x \in B(x^0, \delta) \setminus \{x^0\}$, vale

$$\frac{|r(x)|}{\|x - x^0\|^2} < \nu/2. \tag{4.19}$$

Allora, se $x \in B(x^0, \delta) \setminus \{x^0\}$, otteniamo da (4.18) e (4.19)

$$\begin{aligned} f(x) &= f(x^0) + Q(x - x^0) + r(x) \\ &\geq f(x^0) + \nu\|x - x^0\|^2 - |r(x)| \\ &\geq f(x^0) + \nu\|x - x^0\|^2 - (\nu/2)\|x - x^0\|^2 \\ &= f(x^0) + (\nu/2)\|x - x^0\|^2 \\ &> f(x^0). \end{aligned}$$

Quindi x^0 è un punto di minimo relativo.

□

Osservazione 4.3. Dalla dimostrazione del teorema 4.3, si vede che vale, di fatto, un risultato piu' forte di quello dichiarato nell'enunciato: vale a dire, che esiste $\delta \in \mathbb{R}^+$, tale che, se $x \in A \cap B(x^0, \delta)$ e $x \neq x^0$, si ha $f(x^0) < f(x)$ (risp. $f(x^0) > f(x)$). In tal caso, si dice che x^0 è un **estremante relativo forte** di f.

Osservazione 4.4. Sia

$$\begin{cases} f : \mathbb{R}^2 \to \mathbb{R}, \\ f(x_1, x_2) = x_1^2 + x_2^4. \end{cases}$$

Si verifica immediatamente che f ammette minimo e l'unico punto di minimo è $(0,0)$. Come deve essere in base ai teoremi 4.1 e 4.2, $(0,0)$ è un punto critico e la forma quadratica Q definita in (4.2) ponendo $x^0 = (0,0)$, che è

$$Q(h_1, h_2) = h_1^2,$$

è semidefinita positiva. Tuttavia, non è definita positiva.

Questo esempio dimostra che le condizioni del teorema 4.3 sono sufficienti, ma non necessarie, per avere un estremante relativo.

Resta il problema di stabilire se una certa forma quadratica sia semidefinita o definita. Tale forma puo' essere rappresentata nella forma

$$Q(h) = \frac{1}{2} h \cdot Ah,$$

dove $A : \mathbb{R}^n \to \mathbb{R}^n$

$$A(h_1, \dots, n_n) = (\sum_{j=1}^{n} D_{1j} f(x^0) h_j, \dots, \sum_{j=1}^{n} D_{ij} f(x^0) h_j, \dots, \sum_{j=1}^{n} D_{nj} f(x^0) h_j).$$

A è l'applicazione lineare associata alla matrice $n \times n$ (detta hessiana) $(D_{ij} f(x^0))_{1 \leq i,j \leq n}$, che è simmetrica per il teorema di Schwarz. Da cio' segue che A è simmetrica rispetto al solito prodotto scalare in $\mathbb{R}^n$, nel senso che

$$Ah \cdot k = h \cdot Ak \quad \forall h, k \in \mathbb{R}^n.$$

Per tali applicazioni vale il cosi' detto teorema spettrale (vedi [3], cap. 11):

Teorema 4.4. *Siano V uno spazio vettoriale di dimensione $n \geq 1$ su $\mathbb{R}$, munito di un prodotto scalare $< \cdot, \cdot >$ e $A : V \to V$ un'applicazione lineare simmetrica. Allora esiste una base ortonormale di V costituita da autovalori di A.*

Quindi, esistono $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ reali e $v_1, \dots, v_n$ in V tali che, se $1 \leq i, j \leq n$,

$$Av_i = \lambda_i v_i$$

e $< v_i, v_j >= \delta_{ij}$ (simbolo di Kronecker).

Sia allora $h \in \mathbb{R}^n$. h si puo' rappresentare nella forma

$$h = \sum_{i=1}^{n} y_i v_i,$$

con $y_1, \dots, y_n \in \mathbb{R}$. Segue

$$Q(h) = (\sum_{i=1}^{n} y_i v_i) \cdot (\sum_{j=1}^{n} \lambda_j y_j v_j) = \sum_{i=1}^{n} \lambda_i y_i^2.$$

E' allora chiara la seguente

Proposizione 4.1. *La forma quadratica Q è :*
(I) semidefinita positiva se e solo se $min_{1 \leq i \leq n} \lambda_i \geq 0$;
(II) semidefinita negativa se e solo se $max_{1 \leq i \leq n} \lambda_i \leq 0$;
(III) definita positiva se e solo se $min_{1 \leq i \leq n} \lambda_i > 0$;
(IV) definita negativa se e solo se $max_{1 \leq i \leq n} \lambda_i < 0$.

Ora gli autovalori $\lambda_1, \ldots, \lambda_n$ sono le soluzioni dell'equazione algebrica di grado n nell'incognita λ

$$det((\delta_{ij}\lambda - D_{ij}f(x^0))_{1 \leq i,j \leq n}) = 0, \tag{4.20}$$

Tale equazione non e', in generale, risolubile esplicitamente. Tuttavia, per i nostri scopi ci basta sapere se le soluzioni siano tutte non negative (o non positive). Possiamo usare inoltre l'informazione che le soluzioni complesse di (4.20) sono tutte reali, perchè cosi' sono gli autovalori di A per il teorema spettrale. Quindi è possibile applicare il seguente elementare risultato:

Proposizione 4.2. *Sia $p(\lambda) = \lambda^n + \sum_{j=0}^{n-1} a_j \lambda^j$ un polinomio a coefficienti reali, i cui zeri complessi sono tutti reali. Allora:*
(I) gli zeri di p sono tutti positivi se e solo se $(-1)^{n-j} a_j > 0 \; \forall j = 0, \ldots, n$ (ponendo $a_n = 1$);
(II) gli zeri di p sono tutti negativi se e solo se $a_j > 0$ per ciascun $j = 0, \ldots, n-1$.

Dimostrazione (I) Siano $0 < \lambda_1 \leq \ldots \leq \lambda_n$ gli zeri di p (ciascuno di essi compare tante volte quanto è la rispettiva molteplicita'); allora

$$p(\lambda) = (\lambda - \lambda_1) \ldots (\lambda - \lambda_n).$$

Per il principio di identita' dei polinomi, ciascun a_j è somma di addendi della forma $(-1)^{n-j} \lambda_{i_1} \ldots \ldots \lambda_{i_{n-j}}$, con $1 \leq i_1 < \cdots < i_j \leq n$. Segue che $(-1)^{n-j} a_j > 0 \; \forall j = 0 \ldots, n$. Viceversa, se $(-1)^{n-j} a_j > 0 \; \forall j = 0, \ldots, n$, si ha che, se $\lambda < 0$, gli addendi $a_j \lambda^j$ hanno tutti lo stesso segno (uguale a quello di $(-1)^n$). Pertanto $p(\lambda) \neq 0$. D'altra parte, $p(0) = a_0 \neq 0$. Poiche' tutti gli zeri sono reali, devono essere tutti positivi.

(II) Siano $\lambda_1 \leq \cdots \leq \lambda_n < 0$ gli zeri di p (ciascuno di essi compare tante volte quanto è la rispettiva molteplicita'); allora

$$p(\lambda) = (\lambda + |\lambda_1|) \ldots (\lambda + |\lambda_n|).$$

Per il principio di identita' dei polinomi, ciascun a_j è somma di addendi della forma $|\lambda_{i_1}| \ldots |\lambda_{i_{n-j}}|$ e percio' $a_j > 0$ per $j = 0, \ldots, n-1$.

Viceversa, è chiaro che, se $a_j > 0$ per ciascun $j = 0, \ldots, n-1$, $p(\lambda) > 0$ se $\lambda \geq 0$. Usando ancora il fatto che tutti gli zeri di p sono reali, concludiamo che sono tutti negativi.

$\square$

Esempio 4.3. Sia A la trasformazione lineare in $\mathbb{R}^2$ associata alla matrice a coefficienti reali

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}. \tag{4.21}$$

Quindi

$$Ah = A(h_1, h_2) = (a_{11}h_1 + a_{12}h_2, a_{12}h_1 + a_{22}h_2).$$

Da A otteniamo la forma quadratica

$$Q(h_1, h_2) = h_1(a_{11}h_1 + a_{12}h_2) + h_2(a_{12}h_1 + a_{22}h_2) = a_{11}h_1^2 + 2a_{12}h_1h_2 + a_{22}h_2^2.$$

Si ha

$$p(\lambda) = det \begin{pmatrix} \lambda - a_{11} & -a_{12} \\ -a_{12} & \lambda - a_{22} \end{pmatrix}$$

$$= (\lambda - a_{11})(\lambda - a_{22}) - a_{12}^2$$

$$= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}^2.$$

Dalla proposizione 4.1 ricaviamo allora che

(a) Q è definita positiva se e solo se $a_{11} + a_{22} > 0$ e $a_{11}a_{22} - a_{12}^2 = det(A) > 0$;
(b) Q è definita negativa se e solo se $a_{11} + a_{22} < 0$ e $det(A) > 0$;
(c) Q è semidefinita positiva se e solo è definita positiva, oppure $det(A) = 0$ e $a_{11} + a_{22} \geq 0$;
(d) Q è semidefinita negativa se e solo è definita negativa, oppure $det(A) = 0$ e $a_{11} + a_{22} \leq 0$.

Esercizio 4.1. Determinare e studiare i punti critici delle seguenti funzioni $f : \mathbb{R}^2 \to \mathbb{R}$:

(I) $f(x_1, x_2) = 2x_1^3 + x_1x_2^2 + 5x_1^2 + x_2^2$;
(II) $f(x_1, x_2) = e^{2x_1}(x_1 + x_2^2 + 2x_2)$;
(III) $f(x_1, x_2) = \sin(x_1x_2)$;
(IV) $f(x_1, x_2) = \cos(x_1)\cosh(x_2)$;
(V) $f(x_1, x_2) = x_1 - x_1^2 - x_2^2$;
(VI) $f(x_1, x_2) = (x_1 + 1)(x_2 - 2)$;
(VII) $f(x_1, x_2) = x_1x_2(x_1 - 1)$;
(VIII) $f(x_1, x_2) = x_1^3 + x_1 - 4x_1x_2 - 2x_2^2$.

Esercizio 4.2. Determinare e studiare i punti critici delle seguenti funzioni $f : \mathbb{R}^3 \to \mathbb{R}$:

(I) $f(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2$;
(II) $f(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + 2x_3 - x_1x_2 - x_1x_3$.

Esercizio 4.3. Siano I un intervallo aperto in $\mathbb{R}$, $f \in C^2(I)$, $x_0 \in I$. Provare quanto segue :

(I) se x_0 è un punto di minimo relativo per f, $f'(x_0) = 0$ e $f''(x_0) \geq 0$;
(II) se x_0 è un punto di massimo relativo per f, $f'(x_0) = 0$ e $f''(x_0) \leq 0$;
(III) se $f'(x_0) = 0$ e $f''(x_0) > 0$, x_0 è un punto di minimo relativo per f;
(IV) se $f'(x_0) = 0$ e $f''(x_0) < 0$, x_0 è un punto di massimo relativo per f.

5 Derivate di funzioni a valori vettoriali

Fino ad ora abbiamo esclusivamente considerato derivate di funzioni a valori in $\mathbb{R}$. Qui estenderemo alcune delle definizioni e dei risultati relativi al calcolo differenziale a funzioni a valori in $\mathbb{R}^m$.

Cominciamo con l'estendere la definizione 1.1 :

Definizione 5.1. *Siano $m, n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $v \in \mathbb{R}^n$, $x^0 \in A$. Sia $f : A \to \mathbb{R}^m$. Per $t \in A_{x^0,v} \setminus \{0\}$, poniamo*

$$r_{x^0,v}(t) := \frac{f(x^0 + tv) - f(x^0)}{t}. \tag{5.1}$$

e supponiamo che $0 \in D(A_{x^0,v})$. Se esiste

$$\lim_{t \to 0} r_{x^0,v}(t) = \lim_{t \to 0} \frac{f(x^0 + tv) - f(x^0)}{t}, \tag{5.2}$$

lo chiameremo **derivata di f rispetto al vettore v in x^0** *e lo indicheremo con la notazione* $\frac{\partial f}{\partial v}(x^0)$.

Nel caso $v = e^j$ ($j-$esimo elemento della base canonica di $\mathbb{R}^n$), parleremo di **derivata parziale prima rispetto alla variabile x_j** *e useremo le notazioni in (1.5).*

Osservazione 5.1. Sotto le ipotesi della definizione 5.1, se $t \in A_{x^0,v} \setminus \{0\}$, si ha

$$\frac{f(x^0+tv)-f(x^0)}{t} = (\frac{f_1(x^0+tv)-f_1(x^0)}{t}, ..., \frac{f_m(x^0+tv)-f_m(x^0)}{t}).$$

Dunque, in base al teorema 3.2 in [1], $\frac{\partial f}{\partial v}(x^0)$ esiste se e solo se, per ciascun $i = 1, ..., m$, esistono a valori reali le derivate $\frac{\partial f_i}{\partial v}(x^0)$. Si ha inoltre

$$\frac{\partial f}{\partial v}(x^0) = (\frac{\partial f_1}{\partial v}(x^0), ..., \frac{\partial f_m}{\partial v}(x^0)). \tag{5.3}$$

In particolare, per $j = 1, ..., n$,

$$D_j f(x^0) = (D_j f_1(x^0), ..., D_j f_m(x^0)). \tag{5.4}$$

Ad esempio, se

$$\begin{cases} f : \mathbb{R}^3 \to \mathbb{R}^2, \\ f(x_1, x_2, x_3) = (\sin(x_1)x_2x_3, (1+x_3^2)^{x_2}), \end{cases}$$

si ha, $\forall x = (x_1, x_2, x_3) \in \mathbb{R}^3$,

$$D_2 f(x) = (\sin(x_1)x_3, (1+x_3^2)^{x_2}\ln(1+x_3^2)).$$

Veniamo ora alla nozione di funzione differenziabile a valori vettoriale:

Definizione 5.2. *Siano $n, m \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}^m$, $x^0 \in \mathring{A}$. Diremo che f e'* **differenziabile in x^0** *se esiste $T : \mathbb{R}^n \to \mathbb{R}^m$ lineare tale che*

$$f(x^0+h) - f(x^0) - Th = o(\|h\|_n) \quad (h \to 0),$$

cioe'

$$\lim_{h\to 0} \frac{f(x^0+h)-f(x^0)-Th}{\|h\|_n} = 0. \tag{5.5}$$

(equivalente a

$$\lim_{h\to 0} \frac{\|f(x^0+h)-f(x^0)-Th\|_m}{\|h\|_n} = 0.)$$

Osservazione 5.2. Nel caso $m = 1$ le applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}$ sono le funzioni della forma $T : \mathbb{R}^n \to \mathbb{R}$, $Th = w \cdot h$, con $w \in \mathbb{R}^n$. Quindi la definizione (5.2), nel caso $m = 1$, è equivalente alla definizione 1.3.

Proposizione 5.1. *Siano $n, m \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}^m$, $x^0 \in \mathring{A}$, $f(x) = (f_1(x), \dots, f_m(x))$. Allora f e' differenziabile in x^0 se e solo se ciascuna delle funzioni $f_1, \dots, f_m$ è differenziabile in x^0.*

Dimostrazione Le applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}^m$ sono le applicazioni della forma

$$Th = (w^1 \cdot h, \dots, w^m \cdot h), \quad h \in \mathbb{R}^n,$$

con $w^1, \dots, w^m \in \mathbb{R}^n$. Quindi, se f è differenziabile in x^0 e

$$r(h) = f(x^0 + h) - f(x^0) - Th = r(h) = (r_1(h), \dots, r_m(h)),$$

per $i = 1, \dots, m$ si ha

$$f_i(x^0 + h) - f_i(x^0) - w^i \cdot h = r_i(h).$$

Inoltre,

$$\frac{|r_i(h)|}{\|h\|_n} \leq \frac{\|r(h)\|_m}{\|h\|_n} \to 0 \quad (h \to 0).$$

Quindi, se f è differenziabile in x^0, ciascuna delle $f_1, \dots, f_m$ è differenziabile in x^0.

Viceversa, supponiamo che le $f_1, \dots, f_m$ siano differenziabili in x^0 e, per ciascun i, $\nabla f_i(x^0) = w^i \in \mathbb{R}^n$ e

$$r_i(h) = f_i(x^0 + h) - f_i(x^0) - w^i \cdot h \quad \forall i \in \{1, \dots, m\}.$$

Poniamo

$$\begin{cases} T : \mathbb{R}^n \to \mathbb{R}^m, \\ \\ Th := (w^1 \cdot h, \dots, w^m \cdot h). \end{cases}$$

Allora $T \in L(\mathbb{R}^n, \mathbb{R}^m)$. Inoltre

$$f(x^0 + h) - f(x^0) - Th = r(h),$$

con

$$r(h) = (r_1(h), \dots, r_m(h)).$$

Siano ϵ, η positivi. Esiste $\delta(\eta)$ tale che, se $x^0 + h \in A$ e $\|h\|_n < \delta(\eta)$,

$$|r_i(h)| \leq \eta \|h\|_n \quad \forall i \in \{1, \dots, m\}.$$

Con tale scelta di h,

$$\|r(h)\|_m = (\sum_{i=1}^m r_i(h)^2)^{1/2} \leq m^{1/2} \eta \|h\|_n \leq \epsilon \|h\|_n$$

se $m^{1/2}\eta \leq \epsilon$. Quindi f è differenziabile in x^0.

□

Dal teorema 3.2 in [1] e dalla proposizione 1.1 segue immediatamente la seguente

Proposizione 5.2. *Siano $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}^m$, $x^0 \in \mathring{A}$. Supponiamo che f sia differenziabile in x^0. Allora:*

(I) f e' continua in x^0;

(II) $\forall v \in \mathbb{R}^n$ esiste $\frac{\partial f}{\partial v}(x^0)$ ed e' uguale a Tv, con $T \in L(\mathbb{R}^n; \mathbb{R}^m)$ come nella definizione 5.2;

(III) T e' univocamente determinata e

$$Th = (\nabla f_1(x^0) \cdot h, \dots, \nabla f_m(x^0) \cdot h).$$

Definizione 5.3. *Se f e' differenziabile in x^0, chiameremo l'elemento T di $L(\mathbb{R}^n;\mathbb{R}^m)$ di cui alla definizione 5.2* **differenziale di f in x^0** *e lo indicheremo con la notazione $df(x^0)$.*

Osservazione 5.3. Se f è differenziabile in x^0, il differenziale $df(x^0)$ è l'elemento di $L(\mathbb{R}^n, \mathbb{R}^m)$ associato alla matrice $m \times n$

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^0) & ... & \frac{\partial f_1}{\partial x_n}(x^0) \\ ... & \frac{\partial f_i}{\partial x_j}(x^0) & ... \\ \frac{\partial f_m}{\partial x_1}(x^0) & ... & \frac{\partial f_m}{\partial x_n}(x^0) \end{pmatrix} \tag{5.6}$$

Naturalmente, cio' significa che, se $h = (h_1, \dots, h_n)$, Th puo' essere ottenuto (scritto come vettore colonna $m \times 1$) come prodotto (righe per colonne)

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^0) & ... & \frac{\partial f_1}{\partial x_n}(x^0) \\ ... & \frac{\partial f_i}{\partial x_j}(x^0) & ... \\ \frac{\partial f_m}{\partial x_1}(x^0) & ... & \frac{\partial f_m}{\partial x_n}(x^0) \end{pmatrix} \begin{pmatrix} h_1 \\ ... \\ h_n \end{pmatrix}$$

Chiameremo la matrice (5.6) **matrice jacobiana** di f in x^0, e la indicheremo con la notazione $J_f(x^0)$.

Si possono definire senza difficolta' le derivate di ordine superiore di funzioni a valori vettoriali, ricalcando la definizione 2.1 :

Definizione 5.4. *Siano $m, n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}^m$, $1 \leq j \leq n$. Poniamo*

$$A_j := \{x \in A : \exists \;\; D_j f(x) \in \mathbb{R}^m\}.$$

Su A_j è definita la funzione $D_j f$, che associa a ogni elemento x di A_j la derivata $D_j f(x)$. Se $x \in A_j$ e $1 \leq i \leq n$, può esistere la derivata $D_i(D_j f)(x)$ di $D_j f$ in x. Essa viene chiamata **derivata seconda** *di f in x rispetto alle variabili x_i, x_j (nell'ordine) e viene indicata con uno dei simboli $D_{ij} f(x)$, $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ $(D_j^2 f(x)$, $\frac{\partial^2 f}{\partial x_j^2}(x)$ se $i = j)$.*

Iterando il procedimento, si possono definire le derivate terze, quarte, ecc.. (o derivate di ordine 3, 4, ecc.). Dato $k \in \mathbb{N}$,$j_1, j_2, ..., j_k \in \{1, ..., n\}$, non necessariamente a due a due distinti, porremo

$$D_{j_1 j_2 ... j_k} f(x) = D_{j_1}(D_{j_2}(...(D_{j_k} f))...)(x) = \frac{\partial^k f}{\partial x_{j_1} \partial x_{j_2} ... \partial x_{j_k}}(x).$$

Osservazione 5.4. Iterando le considerazioni contenute nell'osservazione 5.1, è facile verificare che la derivata $\frac{\partial^k f}{\partial x_{j_1} \partial x_{j_2} ... \partial x_{j_k}}(x)$ esiste se e solo se esiste ciascuna delle derivate $\frac{\partial^k f_1}{\partial x_{j_1} \partial x_{j_2} ... \partial x_{j_k}}(x), ...,$ $\frac{\partial^k f_m}{\partial x_{j_1} \partial x_{j_2} ... \partial x_{j_k}}(x)$. Vale, inoltre,

$$\frac{\partial^k f}{\partial x_{j_1} \partial x_{j_2} ... \partial x_{j_k}}(x) = (\frac{\partial^k f_1}{\partial x_{j_1} \partial x_{j_2} ... \partial x_{j_k}}(x), ..., \frac{\partial^k f_m}{\partial x_{j_1} \partial x_{j_2} ... \partial x_{j_k}}(x)). \tag{5.7}$$

Osservazione 5.5. Il teorema di Schwarz (teorema 2.1) è facilmente estendibile a funzioni a valori in $\mathbb{R}^m$. La dimostrazione puo' essere ottenuta ragionando componente per componente.

La definizione di funzione di classe C^k a valori vettoriali ricalca quella data nel caso di funzioni a valori scalari (definizione 2.2):

Definizione 5.5. *Siano $m, n \in \mathbb{N}$, A un aperto di $\mathbb{R}^n$, $f : A \to \mathbb{R}^m$. Diremo che f è di classe C^k ($k \in \mathbb{N}$) se possiede a valori reali tutte le derivate parziali di ordine non superiore a k in ogni punto di A e tali derivate sono continue in A. Scriveremo, per indicare tale eventualità, $f \in C^k(A; \mathbb{R}^m)$. Se $f \in C^k(A; \mathbb{R}^m)$ per ogni $k \in \mathbb{N}$ (e quindi possiede derivate parziali continue di ogni ordine), scriveremo $f \in C^\infty(A; \mathbb{R}^m)$.*

Osservazione 5.6. Ragionando componente per componente, si vede facilmente che se A un aperto di $\mathbb{R}^n$ e $f \in C^1(A; \mathbb{R}^m)$, f è differenziabile in ogni punto di A. Nello stesso modo si vede che, se $f \in C^k(A; \mathbb{R}^m)$, due derivate qualunque di ordine non superiore a k, che si ottengono applicando lo stesso numero di volte le singole derivate parziali del primo ordine coincidono. Agli elementi di $C^k(A; \mathbb{R}^m)$ è estendibile la notazione (2.4).

Dall'osservazione 5.4 e dai teoremi 3.6 in [1] e 2.2 segue subito il risultato:

Teorema 5.1. *Siano $k, m, n \in \mathbb{N}$, A un aperto di $\mathbb{R}^n$, $f : A \to \mathbb{R}^m$, $f(x) = (f_1(x), ..., f_m(x))$ $\forall x \in A$. Allora*

(I) $f \in C^k(A; \mathbb{R}^m)$ se e solo se ciascuna delle funzioni f_1,...,f_m appartiene a $C^k(A)$;

(II) se $f, g \in C^k(A; \mathbb{R}^m)$ e $c \in \mathbb{R}$, $f + g, cf \in C^k(A; \mathbb{R}^m)$;

(III) di conseguenza, $C^k(A; \mathbb{R}^m)$ è uno spazio vettoriale su $\mathbb{R}$ con le solite operazioni di somma di funzioni e di moltiplicazione di una funzione per uno scalare;

(IV) inoltre, se $\alpha \in \mathbb{N}_0^n$ e $|\alpha| \leq k$, l'applicazione $f \to D^\alpha f$ e' lineare da $C^k(A; \mathbb{R}^m)$ a $C^{k-|\alpha|}(A; \mathbb{R}^m)$.

Concludiamo questa sezione esaminando il caso $n = 1$, che è fisicamente assai importante. Solo in tale caso si puo' parlare di "derivata" (usando il singolare):

Definizione 5.6. *Siano $m \in \mathbb{N}$, $A \subseteq \mathbb{R}$, $t^0 \in A \cap D(A)$. Sia poi $f : A \to \mathbb{R}^m$. Per $t \in A \setminus \{t^0\}$, poniamo*

$$r_{t^0}(t) := \frac{f(t) - f(t^0)}{t - t^0}. \tag{5.8}$$

Se esiste

$$\lim_{t \to t^0} r_{t^0}(t) = \lim_{t \to t^0} \frac{f(t) - f(t^0)}{t - t^0}, \tag{5.9}$$

lo chiameremo **derivata di f in t^0** *e lo indicheremo con una delle notazioni $\frac{df}{dt}(x^0)$,$f'(t^0)$, $Df(t^0)$. In tal caso, diremo che f è* **derivabile** *in t^0.*

Osservazione 5.7. Con riferimento alla definizione 5.6, è appena il caso di osservare che, se $f(t) = (f_1(t), ..., f_m(t))$, f è derivabile in t^0 se e solo se ciascuna delle funzioni f_1,...,f_m è derivabile in t^0. In tal caso, si ha

$$f'(t^0) = (f_1'(t^0), ..., f_m'(t^0)). \tag{5.10}$$

E' poi possibile definire anche le derivate di ordine superiore di f. Lasciamo il compito al lettore, come semplice esercizio.

6 Derivate di funzioni composte ("chain rule")

Il risultato fondamentale per la derivazione di funzioni composte è il seguente

Teorema 6.1. *Siano n, m, l numeri naturali, $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$, $f : A \to \mathbb{R}^m$, $g : B \to \mathbb{R}^l$, con $f(A) \subseteq B$. Sia $x^0 \in \mathring{A}$, con $f(x^0) \in \mathring{B}$. Supponiamo che f sia differenziabile in x^0 e g sia differenziabile in $f(x^0)$. Allora $g \circ f$ è differenziabile in x^0 e*

$$d(g \circ f)(x^0) = dg(f(x^0)) \circ df(x^0).$$

Dimostrazione Sia $h \in \mathbb{R}^n$, tale che $x^0 + h \in A$. Poniamo

$$r(h) := f(x^0 + h) - f(x^0) - df(x^0)h.$$

Allora

$$\lim_{h \to 0} \frac{\|r(h))\|_m}{\|h\|_n} = 0. \tag{6.1}$$

Sia $k \in \mathbb{R}^m$, tale che $f(x^0) + k \in A$. Poniamo

$$s(k) := g(f(x^0) + k) - g(f(x^0)) - dg(f(x^0))k.$$

Allora

$$\lim_{k \to 0} \frac{\|s(k)\|_l}{\|k\|_m} = 0. \tag{6.2}$$

Sia $h \in \mathbb{R}^n$, tale che $x^0 + h \in A$. Allora

$$(g \circ f)(x^0 + h) = g(f(x^0)) + dg(f(x^0))[f(x^0 + h) - f(x^0)] + s[f(x^0 + h) - f(x^0)]$$

$$= g(f(x^0)) + dg(f(x^0))[df(x^0)h + r(h)] + s[f(x^0 + h) - f(x^0)]$$

$$= g(f(x^0)) + [dg(f(x^0)) \circ df(x^0)]h + dg(f(x^0))r(h) + s[f(x^0 + h) - f(x^0)].$$

Resta da provare che

$$\lim_{h \to 0} \frac{\|dg(f(x^0))r(h) + s[f(x^0 + h) - f(x^0)]\|_l}{\|h\|_n} = 0 \quad (\|h\|_n \to 0). \tag{6.3}$$

Siano $\epsilon, \eta \in \mathbb{R}^+$. Per la disuguaglianza (3.8) in [1], esiste $L \in \mathbb{R}^+$ tale che

$$\|dg(f(x^0))r(h)\|_l \le L\|r(h)\|_m.$$

Per (6.1), esiste $\delta_1(\eta) \in \mathbb{R}^+$ tale che, se $\|h\|_n \le \delta_1(\eta)$,

$$\|r(h)\|_m \le \eta\|h\|_n.$$

Con tale scelta di h, si ha

$$\|dg(f(x^0))r(h)\|_l \le L\eta\|h\|_n.$$

Inoltre, per (6.2), esiste $\delta_2(\eta) \in \mathbb{R}^+$ tale che, se $\|k\|_m \le \delta_2(\eta) \in \mathbb{R}^+$, $f(x^0) + k \in B$ e

$$\|s(k)\|_l \le \eta\|k\|_m.$$

Se $\|h\|_n \le \delta_1(\eta)$, si ha

$$\|f(x^0 + h) - f(x^0)\|_m \le \|df(x^0)h\|_m + \|r(h)\|_m \le (M + \eta)\|h\|_n,$$

per un certo $M \ge 0$ (stiamo di nuovo applicando la disuguaglianza (3.8) in [1]).

Sia $\delta_3(\eta)$ positivo, con

$$\delta_3(\eta) \le \min\{\delta_1(\eta), \frac{\delta_2(\eta)}{M + \eta}\}.$$

Si ha allora, per $\|h\|_n \le \delta_3(\eta)$,

$$\|dg(f(x^0))r(h) + s[f(x^0 + h) - f(x^0)]\|_l \le (L + M + \eta)\eta\|h\|_n.$$

Scelto η in modo che
$$(L + M + \eta)\eta \leq \epsilon,$$
si conclude che
$$\|dg(f(x^0))r(h) + s[f(x^0 + h) - f(x^0)]\|_l \leq \epsilon \|h\|_n.$$
Quindi vale (6.3).

$\square$

Dal teorema 6.1 si ricava facilmente la "chain rule" per il calcolo delle derivate di funzioni composte:

Corollario 6.1. *Siano m e n naturali, $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$, $f : B \to \mathbb{R}^n$, $g : A \to \mathbb{R}$. Supponiamo inoltre che $f(B) \subseteq A$, $x^0 \in \mathring{B}$, $f(x^0) \in \mathring{A}$, f sia differenziabile in x^0, g sia differenziabile in $f(x^0)$. Allora $\forall i \in \{1, ..., m\}$ e' definita la derivata $D_i(g \circ f)(x^0)$. Inoltre*
$$D_i(g \circ f)(x^0) = \nabla g(f(x^0)) \cdot D_i f(x^0) = \sum_{j=1}^{n} D_j g(f(x^0)) \frac{\partial f_j}{\partial x_i}(x^0).$$
(chain rule), con $\cdot$ prodotto scalare in $\mathbb{R}^n$.

Dimostrazione $g \circ f$ è differenziabile in x^0 per il teorema 6.1. Quindi, per la proposizione 5.2, è definita la derivata $D_i(g \circ f)(x^0)$. Inoltre
$$D_i(g \circ f)(x^0) = d(g \circ f)(x^0)e^i = (dg(f(x^0)) \circ df(x^0))e^i$$
$$= dg(f(x^0))[df(x^0)e^i] = dg(f(x^0))[D_i f(x^0)] = \nabla g(f(x^0)) \cdot D_i f(x^0).$$
$\square$

Proposizione 6.1. *Siano m, n, p, k naturali, A aperto in $\mathbb{R}^n$, B aperto in $\mathbb{R}^m$, $g \in C^k(A; \mathbb{R}^p)$, $f \in C^k(B; \mathbb{R}^n)$, tale che $f(B) \subseteq A$. Allora $g \circ f \in C^k(B; \mathbb{R}^p)$.*

Dimostrazione Osserviamo, innanzi tutto che, per il teorema 5.1(I), ci si puo' limitare al caso $p = 1$. Proviamo il risultato per induzione su k.

Sia $k = 1$. Allora, per il corollario 6.1, se $i \in \{1, ..., m\}$, $\forall x \in B$ esiste $D_i(g \circ f)(X)$ e
$$D_i(g \circ f)(x) = \sum_{j=1}^{n} D_j g(f(x)) \frac{\partial f_j}{\partial x_i}(x).$$
Poiche' tutte le derivate che compaiono sono continue, $D_i(g \circ f)$ è continua in virtu' dei teoremi 3.6 e 3.7 in [1].

Supponiamo che l'enunciato sia valido per un certo $k \in \mathbb{N}$. Supponiano inoltre che, se $\alpha \in \mathbb{N}_0^m$, con $|\alpha| \leq k$, $D^\alpha(g \circ f)(x)$ sia esprimibile come combinazione lineare di funzioni del tipo
$$D^\beta g(f(x)) D^{\gamma^1} g(x) \cdots D^{\gamma^l} g(x), \tag{6.4}$$
con $\beta \in \mathbb{N}_0^n$, $l \in \mathbb{N}$, $\gamma^1, \ldots, \gamma^l \in \mathbb{N}_0^m$,
$$\max\{|\beta|, |\gamma^1|, \ldots, |\gamma^l|\} \leq k.$$
Cio' accade se $k = 1$. Ne segue che, se $f \in C^{k+1}(B; \mathbb{R}^n)$, $g \in C^{k+1}A)$ e $j \in \{1, \ldots, m\}$, $D^{\alpha + e^j}(g \circ f)(x)$ è definita e, per il corollario 6.1, si puo' scrivere come combinazione lineare di termini del tipo (6.4), con
$$\max\{|\beta|, |\gamma^1|, \ldots, |\gamma^l|\} \leq k + 1.$$

Applicando quindi ancora i teoremi 3.6 e 3.7 in [1], si conclude che, se l'enunciato è valido per k, è valido anche per $k+1$.

Percio' la conclusione generale segue per induzione.

$\square$

Esercizio 6.1. Siano

$$\begin{cases} f : \mathbb{R}^3 \to \mathbb{R}^3 \\ f(\rho, \theta, \phi) = (\rho \sin(\theta) \cos(\phi), \rho \sin(\theta) \sin(\phi), \rho \cos(\theta)), \end{cases}$$

$g \in C^1(\mathbb{R}^3)$. Calcolare le derivate parziali del primo ordine di $g \circ f$.

Ripetere l'esercizio, prendendo

$$f(\rho, \theta, \zeta) := (\rho \cos(\theta), \rho \sin(\theta), \zeta).$$

7 Il teorema di invertibilita' locale

Sia A un aperto in $\mathbb{R}^n$ e sia $f \in C^1(A; \mathbb{R}^n)$. Sia poi $x^0 \in A$. Consideriamo l'equazione

$$f(x) = y, \tag{7.1}$$

con $y \in \mathbb{R}^n$. Poniamo

$$r(h) = f(x^0 + h) - f(x^0) - df(x^0)h$$

e scriviamo (7.1) nella forma

$$f(x^0) + df(x^0)h + r(h) = y, \tag{7.2}$$

con $h = x - x^0$. $df(x^0)$ è una trasformazione lineare da $\mathbb{R}^n$ a $\mathbb{R}^n$. Supponiamo che sia invertibile. Allora (7.2) è equivalente a

$$h = df(x^0)^{-1}[y - f(x^0) - r(h)]. \tag{7.3}$$

Si ha che $r(h) = o(\|h\|)$ per $h \to 0$; in generale, quindi, questo termine è "trascurabile" se $\|h\|$ è "piccolo", vale a dire, se $x = x^0 + h$ è prossimo a x^0. Ma se x è prossimo a x^0, per continuita', y sara' prossimo a $f(x^0)$.

In conclusione, consideriamo l'equazione (7.1), con $\|y - f(x^0\|$ piccolo; allora possiamo sperare di trovare una soluzione $x = x^0 + h$ con $\|h\|$ piccolo. In tal caso ci aspettiamo che il termine $r(h)$ sia trascurabile. Se lo ignoriamo, troviamo la "pseudosoluzione"

$$h = df(x^0)^{-1}[y - f(x^0)],$$

da cui

$$x = x^0 + df(x^0)^{-1}[y - f(x^0)].$$

Ci aspettiamo quindi un risultato di questo genere: se $df(x^0)$ è invertibile in $L(\mathbb{R}^n, \mathbb{R}^n)$ e y è "prossimo" a $f(x^0)$, (7.1) ha un'unica soluzione x "prossima" a x^0. Inoltre, x è vicina a $x^0 + df(x^0)^{-1}[y - f(x^0)]$.

Un risultato di questo genere è il seguente "teorema di invertibilita' locale":

Teorema 7.1. *(di invertibilità locale) Siano A un aperto di $\mathbb{R}^n$, $f \in C^1(A, \mathbb{R}^n)$, $x^0 \in A$. Supponiamo che la matrice jacobiana $J_f(x^0)$ abbia determinante diverso da 0. Allora esistono A_1 e A_2 aperti in $\mathbb{R}^n$, tali che:*

(I) $A_1 \subseteq A$, $x^0 \in A_1$, $f(x^0) \in A_2$;

(II) se $g := f_{|A_1}$, g e' una biiezione tra A_1 e A_2;

(III) $g^{-1} : A_2 \to A_1$ è di classe C^1 e, se $f \in C^k(A; \mathbb{R}^n)$ $(k \geq 1)$, anche g^{-1} e' di classe C^k.

Osservazione 7.1. L'ipotesi che la matrice jacobiana $J_f(x^0)$ abbia determinante diverso da 0 equivale a richiedere che la trasformazione lineare $df(x^0)$ sia un isomorfismo lineare di $\mathbb{R}^n$ in se'.

Alla dimostrazione del teorema 7.1 premettiamo alcuni risultati preliminari.

Lemma 7.1. *Siano $n, m \in \mathbb{N}$, A un aperto in $\mathbb{R}^n$, $f \in C^1(A; \mathbb{R}^m)$. Siano poi $x^0 \in A$ e*

$$r(x) = f(x) - f(x^0) - df(x^0)(x - x^0), \quad x \in A.$$

Allora $\forall \epsilon \in \mathbb{R}^+$ esiste $\delta(\epsilon) \in \mathbb{R}^+$ tale che $B(x^0, \delta(\epsilon)) \subseteq A$ e, $\forall x^1, x^2 \in B(x^0, \delta(\epsilon))$

$$\|r(x_1) - r(x_2)\|_m \leq \epsilon \|x_1 - x_2\|_n.$$

Dimostrazione Ragionando sulle componenti di f, si vede che è sufficiente considerare il caso $m = 1$. Quindi

$$df(x^0)h = \nabla f(x^0) \cdot h, \quad h \in \mathbb{R}^n.$$

Sia $r \in \mathbb{R}^+$ tale che $B(x^0, r) \subseteq A$. Dati $x^1, x^2 \in B(x^0, r)$, si ha

$$r(x^1) - r(x^2) = f(x_1) - f(x_2) - \nabla f(x^0) \cdot (x^1 - x^2).$$

(Qui e nel seguito indicheremo con $\cdot$ il prodotto scalare in $\mathbb{R}^n$).

Posto $F(t) := f(x^2 + t(x^1 - x^2))$, si ha

$$f(x^1) - f(x^2) = F(1) - F(0) = \int_0^1 F'(t)dt = \int_0^1 \nabla f(x^2 + t(x^1 - x^2)) \cdot (x^1 - x^2)dt.$$

Ne segue

$$r(x^1) - r(x^2) = \int_0^1 [\nabla f(x^2 + t(x^1 - x^2)) - \nabla f(x^0)] \cdot (x^1 - x^2)dt.$$

Poiche' $f \in C^1(A)$, per ogni $\epsilon \in \mathbb{R}^+$ esiste $\delta(\epsilon) \in (0, R]$, tale che, se $y \in B(x^0, \delta(\epsilon))$,

$$\|\nabla f(y) - \nabla f(x^0)\|_n \leq \epsilon.$$

Allora, se $x^1, x^2 \in B(x^0, \delta(\epsilon))$, $[x^2, x^1] \subseteq B(x^0, \delta(\epsilon))$, per cui

$$|r(x^1) - r(x^2)| \leq \int_0^1 |[\nabla f(x^2 + t(x^1 - x^2)) - \nabla f(x^0)] \cdot (x^1 - x^2)|dt$$

$$\leq \int_0^1 \|\nabla f(x^2 + t(x^1 - x^2)) - \nabla f(x^0)]\|dy \cdot \|x^1 - x^2\|dt$$

$$\leq \epsilon \|x^1 - x^2\|.$$

□

Lemma 7.2. *Siano A un aperto in $\mathbb{R}^n$, $f \in C^1(A; \mathbb{R}^n)$, $x^0 \in A$ tale che $det(J_f(x^0)) \neq 0$. Allora:*

(I) esistono $\rho_0, \rho_1 \in \mathbb{R}^+$ tali che $\overline{B(x^0, \rho_1)} \subseteq A$ e $\forall y \in \overline{B(f(x^0), \rho_0)}$ esiste un'unica $x \in \overline{B(x^0, \rho_1)}$) tale che $f(x) = y$;

(II) inoltre esiste $\lambda > 0$ tale che, se $y \in \overline{B(f(x^0), \rho_0)}$, $x \in \overline{B(x^0, \rho_1)}$) e $f(x) = y$,

$$\|x - x^0\| \leq \lambda \|y - f(x^0)\|.$$

Dimostrazione L'equazione

$$f(x) = y$$

è equivalente a

$$x = x^0 + df(x^0)^{-1}[y - f(x^0) - r(x)], \tag{7.4}$$

con

$$r(x) = f(x) - f(x^0) - df(x^0)(x - x^0).$$

Per quanto visto nell'esempio 3.6 in [1], esiste $L \in \mathbb{R}^+$ tale che

$$\|df(x^0)^{-1}k\| \leq L\|k\| \quad \forall k \in \mathbb{R}^n.$$

Dunque, per il lemma 7.1, esiste $r_1 \in \mathbb{R}^+$ tale che $\overline{B(x^0, r_1)} \subseteq A$ e, se x^1, x^2 stanno in questa palla,

$$\|df(x^0)^{-1}r(x^1)\| \leq \frac{1}{2}\|x^1 - x^0\|, \quad \|df(x^0)^{-1}[r(x^1) - r(x^2)]\| \leq \frac{1}{2}\|x^1 - x^2\|.$$

Siano $0 \leq \rho \leq r_1$ e $y \in \mathbb{R}^n$. Consideriamo l'applicazione

$$\begin{cases} F : \overline{B(x^0, \rho)} \to \mathbb{R}^n, \\ \\ F(x) = x^0 + df(x^0)^{-1}[y - f(x^0) - r(x)]. \end{cases}$$

Allora, se $x \in \overline{B(x^0, \rho)}$,

$$\|F(x) - x^0\| = \|df(x^0)^{-1}[y - f(x^0) - r(x)]\| \leq \|df(x^0)^{-1}[y - f(x^0)]\| + \|df(x^0)^{-1}r(x)\|$$

$$\leq L\|y - f(x^0\| + \tfrac{1}{2}\|x - x^0\| \leq L\|y - f(x^0\| + \tfrac{\rho}{2} \leq \rho$$

se

$$\|y - f(x^0\| \leq \frac{\rho}{2L}.$$

Inoltre, se $x^1, x^2 \in \overline{B(x^0, \rho)}$,

$$\|F(x^1) - F(x^2)\| = \|df(x^0)[r(x^2) - r(x^1)] \leq \frac{1}{2}\|x^1 - x^2\|.$$

Quindi, se $\|y - f(x^0)\| = \frac{\rho}{2L} \leq \frac{r_1}{2L}$, per le teorema delle contrazioni esiste un unico $x \in \overline{B(x^0, \rho)}$ tale che

$$F(x) = x.$$

Vale percio' (7.4), per cui $f(x) = y$. Allora la conclusione (I) è provata, con $\rho_0 = \frac{r_1}{2L}$, $\rho_1 = r_1$, $\lambda = 2L$.

□

Lemma 7.3. *Siano A un aperto in $\mathbb{R}^n$, $f \in C^1(A; \mathbb{R}^n)$, $x^0 \in A$ tale che $det(J_f(x^0)) \neq 0$. Allora esiste $r \in \mathbb{R}^+$ tale che $B(x^0, r) \subseteq A$ e $f_{|B(x^0,r)}$ e' iniettiva.*

Dimostrazione Siano ρ_0 e ρ_1 come nel lemma 7.2. Poiche' f è continua, esiste $r \in (0, \rho_1]$ tale che $f(B(x^0, r)) \subseteq B(f(x^0), \rho_0)$. La conclusione segue allora dal lemma 7.2 (I) (unicita' di x).

□

Dimostrazione del teorema di invertibilita' locale Osserviamo, innanzi tutto, che $\{x \in A : det(J_f(x)) \neq 0\}$ è un insieme aperto in $\mathbb{R}^n$. Infatti, è ben noto che

$$det \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & a_{ij} & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix} = P(a_{11}, \ldots, a_{nn}),$$

con P polinomio (vedi [3]). Percio',

$$det(J_f(x)) = P(D_1 f_1(x), \ldots, D_n f_n(x))$$

è una funzione continua della variabile x, essendo $f \in C^1(A; \mathbb{R}^n)$. Di conseguenza, per il risultato dell'esercizio 3.3 (I) in [1], $\{x \in A : det(J_f(x)) \neq 0\}$ è aperto. Poniamo allora

$$A_1 = B(x^0, r),$$

tale che $B(x^0, r) \subseteq A$, $det(J_f(x)) \neq 0$ $\forall x \in B(x^0, r)$ e $f_{|B(x^0,r)}$ è iniettiva. L'esistenza di r cosi' fatto segue dal Lemma 7.3.

Poniamo

$$A_2 := f(A_1).$$

Allora f è una biiezione tra A_1 e A_2. Poniamo

$$g := f_{|A_1}.$$

Per il lemma 7.2 (I), A_2 è un insieme aperto.

Proviamo che g^{-1} è continua. Sia $y^0 \in A_2$, $y^0 = g(x^0)$, $\epsilon \in \mathbb{R}^+$. Per il lemma 7.2, esistono ρ_0, ρ_1 positivi tali che $\overline{B(x^0, \rho_1)} \subseteq A_1$ e

$$\overline{B(y^0, \rho_0)} \subseteq g(\overline{B(x^0, \rho_1)}) \subseteq A_2$$

ed esiste $\lambda \in \mathbb{R}^+$ tale che, se $y \in \overline{B(y^0, \rho_0)}$,

$$\|g^{-1}(y) - g^{-1}(y^0)\| \leq \lambda \|y - y^0\|.$$

Dunque g^{-1} è continua in y^0, generico elemento di A_2.

Verifichiamo che g^{-1} è differenziabile in ogni punto di A_2 e che, se $y^0 \in A_2$,

$$dg^{-1}(y^0) = df(g^{-1}(y^0))^{-1}.$$

Sia $y^0 \in A_2$, $y^0 = g(x^0)$. Da

$$g(x) - f(x^0) = df(x^0)(x - x^0) + r(x), \quad x \in A_1$$

segue, posto $y := g(x)$,

$$y - y^0 = df(x^0)[g^{-1}(y) - g^{-1}(y^0)] + r(g^{-1}(y))$$

e

$$g^{-1}(y) - g^{-1}(y^0) = df(x^0)^{-1}(y - y^0) - df(x^0)^{-1} r(g^{-1}(y)).$$

Siano ϵ, η positivi. Dalla continuita' di g^{-1} e dal lemma 7.2(II) segue che esiste $\delta(\eta) > 0$ tale che, se $\|y - y^0\| \leq \delta(\eta)$,

$$\|r(g^{-1}(y))\| \leq \eta \|g^{-1}(y) - g^{-1}(y^0)\| \leq \lambda \eta \|y - y^0\|.$$

Scelto η in modo che $\lambda\eta \leq \epsilon$, segue la conclusione.

Verifichiamo che $g^{-1} \in C^1(A_2, \mathbb{R}^n)$. Abbiamo visto che, se $y \in A_2$,

$$dg(y) = df(g^{-1}(y))^{-1}$$

Dalle ben note formule che danno l'inversa di una matrice quadrata (vedi [3]), cio' implica che, $\forall i, j \in \{1, \ldots, n\}$, $\forall y \in A_2$,

$$D_i g_j(y) = \frac{P_{ij}(((D_k f_l)(g(y))_{1 \leq k,l \leq n})}{Q_{ij}((D_k f_l)(g(y))_{1 \leq k,l \leq n}}),$$

over P_{ij} e Q_{ij} sono opportuni polinomi. Poiche' $f \in C^1(A; \mathbb{R}^n))$ e g è continua, i teoremi 3.6 e 3.7 in [1] implicano la continuita' delle funzioni $D_i g_j$.

Verifichiamo, infine, che, se $f \in C^k(A; \mathbb{R}^n)$ $(k \in \mathbb{N})$, $g \in C^k(A_2; \mathbb{R}^n)$. Possiamo vederlo per induzione su k: se $k = 1$, l'abbiamo gia' dimostrato. Supponiamo che, per un certo k, se $f \in C^k(A; \mathbb{R}^n)$, $g \in C^k(A_2; \mathbb{R}^n)$. Sia $f \in C^{k+1}(A; \mathbb{R}^n)$. Allora le funzioni $D_k f_l$ $(1 \leq k, l \leq n)$ appartengono a $C^k(A)$ e $g^{-1} \in C^k(A_2; \mathbb{R}^n)$. Quindi, per la proposizione 6.1 e il risultato dell'esercizio 2.1, abbiamo che ogni funzione $D_i g_j$ appartiene a $C^k(A_2)$. Ne segue che $g \in C^{k+1}(A_2; \mathbb{R}^n)$ e la dimostrazione è completa.

$\square$

Osservazione 7.2. Poniamo

$$\begin{cases} f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^2, \\ \\ f(x_1, x_2) = (x_1 \cos(x_2), x_1 \sin(x_2)). \end{cases}$$

Allora è facile verificare che le ipotesi del teorema di incertibilita' locale sono soddisfatte in corrispondenza di ogni $(x_1, x_2) \in \mathbb{R}^+ \times \mathbb{R}$. Tuttavia f non è iniettiva.

Questo esempio mostra che l'invertibilita' di f vale, in generale, solo localmente.

8 Il teorema delle funzioni implicite

.

Siano A un aperto in $\mathbb{R}^n$ e $f \in C^1(A; \mathbb{R}^m)$, con $1 \leq m < n$. Sia

$$x^0 = (x_1^0, \ldots, x_{n-m}^0, x_{n-m+1}^0, \ldots, x_n^0) \in A,$$

tale che

$$f(x^0) = O = (0, \ldots, 0)$$

(lo zero di $\mathbb{R}^m$). Vogliamo descrivere, almeno localmente,

$$\{x \in A : f(x) = O\}.$$

Essendo $f(x^0) = O$, si ha

$$f(x) = df(x^0)(x - x^0) + r(x),$$

con

$$\lim_{x \to 0} \|x - x^0\|_n^{-1} r(x) = O$$

in $\mathbb{R}^m$. Quindi l'equazione (sistema)

$$f(x) = 0 \tag{8.1}$$

si puo scrivere per esteso nella forma (posto $x = (x_1, \dots, x_m, x_{m+1}, \dots x_n)$)

$$\begin{cases} D_1 f_1(x^0)(x_1 - x_1^0) + \dots + D_{n-m} f_1(x^0)(x_{n-m} - x_{n-m}^0) \\ +D_{n-m+1} f_1(x^0)(x_{n-m+1} - x_{n-m+1}^0) + \dots + D_n f_1(x^0)(x_n - x_n^0) = -r_1(x_1, \dots, x_n), \\ \dots, \\ D_1 f_m(x^0)(x_1 - x_1^0) + \dots + D_{n-m} f_m(x^0)(x_{n-m} - x_{n-m}^0) \\ +D_{n-m+1} f_m(x^0)(x_{n-m+1} - x_{n-m+1}^0) + \dots + D_n f_m(x^0)(x_n - x_n^0) = -r_m(x_1, \dots, x_n). \end{cases} \tag{8.2}$$

Introduciamo ora l'ipotesi che la matrice $J_f(x^0)$ abbia rango massimo m e, piu' precisamente, che il suo minore costituito dalle ultime m righe e colonne abbia determinante diverso da 0. Questo e'

$$M := \begin{pmatrix} D_{n-m+1} f_1(x^0) & ... & D_n f_1(x^0) \\ ... & ... & ... \\ D_{n-m+1} f_m(x^0) & ... & D_n f_m(x^0) \end{pmatrix}.$$

Ci interessano, nello spirito del teorema di invertibilita' locale, le soluzioni di (8.1) vicine a x^0. Trascuriamo allora il resto $r(x)$ e consideriamo, invece di (8.2), la sua approssimazione

$$\begin{aligned} &\begin{pmatrix} D_{n-m+1} f_1(x^0) & ... & D_n f_1(x^0) \\ ... & ... & ... \\ D_{n-m+1} f_m(x^0) & ... & D_n f_m(x^0) \end{pmatrix} \begin{pmatrix} x_{n-m+1} - x_{n-m+1}^0 \\ ... \\ x_n - x_n^0 \end{pmatrix} \\ &= - \begin{pmatrix} D_1 f_1(x^0) & ... & D_{n-m} f_1(x^0) \\ ... & ... & ... \\ D_1 f_m(x^0) & ... & D_{n-m} f_m(x^0) \end{pmatrix} \begin{pmatrix} x_1 - x_1^0 \\ ... \\ x_{n-m} - x_{n-m}^0 \end{pmatrix}. \end{aligned} \tag{8.3}$$

Indichiamo con N la matrice

$$\begin{pmatrix} D_1 f_1(x^0) & ... & D_{n-m} f_1(x^0) \\ ... & ... & ... \\ D_1 f_m(x^0) & ... & D_{n-m} f_m(x^0) \end{pmatrix}.$$

Allora, poiche' M è invertibile, l'equazione (8.3) ammette, $\forall (x_1, \dots, x_{n-m}) \in \mathbb{R}^{n-m}$ l'unica soluzione

$$\begin{pmatrix} x_{n-m+1} \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} x_{n-m+1}^0 \\ ... \\ x_n^0 \end{pmatrix} - M^{-1} N \begin{pmatrix} x_1 - x_1^0 \\ ... \\ x_{n-m} - x_{n-m}^0 \end{pmatrix}.$$

Quindi $\{x \in \mathbb{R}^n : f(x) = O\}$ sembrerebbe un oggetto "$n-m-$dimensionale", nel senso che i suoi punti x sono univocamente determinati dagli $n-m$ parametri $x_1, \dots, x_{n-m}$.

Naturalmente, poiche' abbiamo completamente trascurato il resto r, potremo (al piu') aspettarci che il risultato valga localmente, considerando solo elementi x prossimi a x^0. Vale, in effetti il seguente

Teorema 8.1. *(delle funzioni implicite di Dini) Siano m, $n \in \mathbb{N}$, con $m < n$, A aperto in $\mathbb{R}^n$, $f \in C^1(A;\mathbb{R}^m)$. Sia poi $x^0 = (x_1^0, ..., x_{n-m}^0, x_{n-m+1}^0, ..., x_n^0) \in A$, tale che $f(x^0) = O$. Supponiamo che la matrice jacobiana $J_f(x^0)$ abbia rango massimo ($= m$) e che il minore costituito dalle sue ultime m righe e colonne abbia determinante diverso da 0. Allora esistono A_1 e A_2 aperti in $\mathbb{R}^{n-m}$ e $\mathbb{R}^m$ rispettivamente, tali che:*
(I) $(x_1^0, ..., x_{n-m}^0) \in A_1$, $(x_{n-m+1}^0, ..., x_n^0) \in A_2$;
(II) $A_1 \times A_2 \subseteq A$;
(III) $\forall (x_1, ..., x_{n-m}) \in A_1$ esiste unico

$$\phi(x_1, ..., x_{n-m}) = (\phi_1(x_1, ..., x_{n-m}), ..., \phi_m(x_1, ..., x_{n-m})) \in A_2$$

tale che

$$f(x_1, ..., x_{n-m}, \phi_1(x_1, ..., x_{n-m}), ..., \phi_m(x_1, ..., x_{n-m})) = O;$$

(IV) la funzione ϕ è di classe C^1 e, se $f \in C^k(A;\mathbb{R}^m)$ ($k \in \mathbb{N}$), $\phi \in C^k(A_1;\mathbb{R}^m)$.

Dimostrazione Poniamo, dato $x = (x_1, \ldots, x_{n-m}, x_{n-m+1}, \ldots, x_n) \in \mathbb{R}^n$, $x = (y, z)$, con

$$y = (x_1, \ldots, x_{n-m}), \quad z = (x_{n-m+1}, \ldots, x_n).$$

In quest'ordine di idee, sia

$$x^0 = (y^0, z^0),$$

con

$$y^0 = (x_1^0, \ldots, x_{n-m}^0), \quad z^0 = (x_{n-m+1}^0, \ldots, x_n^0).$$

Poniamo anche

$$\begin{cases} F : A \to \mathbb{R}^n, \\ \\ F(y, z) = (y, f(y, z)). \end{cases}$$

Allora $F \in C^1(A;\mathbb{R}^n)$. Inoltre,

$$J_F(y^0, z^0) = \begin{pmatrix} 1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & \ldots & 0 \\ ... & & & & & \\ 0 & 0 & \ldots & 1 & \ldots & 0 \\ D_1 f_1(x^0) & \ldots & \ldots & D_{n-m} f_1(x^0) & \ldots & D_n f_1(x^0) \\ ... & ... & ... & ... & ... & ... \\ D_1 f_m(x^0) & \ldots & \ldots & D_{n-m} f_m(x^0) & \ldots & D_n f_m(x^0) \end{pmatrix}.$$

Si ha quindi che

$$det(J_F(y^0, z^0)) = \det \begin{pmatrix} D_{n-m+1} f_1(x^0) & \ldots & \ldots & D_{n-m} f_1(x^0) & \ldots & D_n f_1(x^0) \\ ... & ... & ... & ... & ... & ... \\ D_{n-m+1} f_m(x^0) & \ldots & \ldots & D_{n-m} f_m(x^0) & \ldots & D_n f_m(x^0) \end{pmatrix} \neq 0.$$

Per il teorema di invertibilita' locale, esistono Ω_1 e Ω_2 aperti in $\mathbb{R}^n$ tali che:
(I) $\Omega_1 \subseteq A$, $x^0 \in \Omega_1$;
(II) $F(x^0) = (y^0, f(y^0, z^0)) = (y^0, O) \in \Omega_2$;
(III) se $G = F_{|\Omega_1}$, G è una biiezione tra Ω_1 e Ω_2;
(IV) $G^{-1} \in C^1(\Omega_2;\mathbb{R}^n)$ e, se $f \in C^k(A;\mathbb{R}^m)$ ($k \in \mathbb{N}$), $G^{-1} \in C^k(\Omega_2;\mathbb{R}^n)$.

Poniamo

$$H(y,v) := (G^{-1}_{m-n+1}(y,v), \dots, G^{-1}_n(y,v)), \quad (y,v) \in \Omega_2.$$

Osserviamo che, per come è definita F,

$$G^{-1}(y,v) = (y, H(y,v)),$$

con $H : \Omega_2 \to \mathbb{R}^m$. Quindi, se $(y,O) \in \Omega_2$,

$$F(G^{-1}(y,O)) = G(G^{-1}(y,O)) = (y,O),$$

da cui

$$f(G^{-1}(y,O)) = O. \tag{8.4}$$

Fissiamo allora A_1 e A_2 aperti rispettivamente in $\mathbb{R}^{n-m}$ e $\mathbb{R}^m$ tali che

(V) $y^0 \in A_1$, $z^0 \in A_2$;
(VI) $A_1 \times A_2 \subseteq \Omega_1$, $A_1 \times \{O\} \subseteq \Omega_2$ e $H(y,O) \in A_2$ $\forall y \in A_1$.

Cio' è possibile perche' $(y^0,O) \in \Omega_2$ e H è continua. Inoltre $F(y^0,z^0) = (y^0,O)$, da cui $H(y^0,O) = z^0$.

Poniamo

$$\begin{cases} \phi : A_1 \to \mathbb{R}^m, \\ \phi(y) = H(y,O). \end{cases} \tag{8.5}$$

Allora, se $y \in A_1$, da (8.4) segue

$$f(y,\phi(y)) = f(G^{-1}(y,O)) = O.$$

Per definizione, $\phi(A_1) \subseteq A_2$. Sia, inoltre, $(y,z) \in A_1 \times A_2$ tale che

$$f(y,z) = O.$$

Segue allora, poiche' $(y,z) \in A_1 \times A_2 \subseteq \Omega_1$,

$$G(y,z) = F(y,z) = (y,O),$$

da cui

$$(y,z) = G^{-1}(y,O) = (y, H(y,O)).$$

Quindi

$$z = H(y,O) = \phi(y).$$

Resta solo da provare la regolarita' di ϕ. Questa segue facilmente dal fatto che, per il teorema di invertibilita' locale, $G^{-1} \in C^1(\Omega_2;\mathbb{R}^n)$ e, se $f \in C^k(A;\mathbb{R}^m)$, $F \in C^k(A;\mathbb{R}^m)$, e percio' $G^{-1} \in C^k(\Omega_2;\mathbb{R}^n)$. Ne segue che $H \in C^k(\Omega_2;\mathbb{R}^m)$ e quindi, per (8.5), $\phi \in C^k(A_1;\mathbb{R}^m)$.

□

Osservazione 8.1. Con riferimento all'enunciato del teorema 8.1, chiaramente il minore costituito dalle ultime m colonne non svolge alcun ruolo privilegiato: qualora uno qualunque dei minori costituiti da m colonne di $J_f(x^0)$ avesse determinante diverso da 0, si potrebbero "esplicitare" localmente le variabili corrispondenti in funzione delle altre.

Esempio 8.1. Sia $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1$. Se

$$x^0 := (0, 0, 1),$$

si ha $f(x^0) = 0$. Inoltre,

$$J_f(x^0) = \begin{pmatrix} 0 & 0 & 2 \end{pmatrix},$$

che ha rango massimo ($= 1$). La sottomatrice (2) è , ovviamente, invertibile.

Dunque,il teorema di Dini è applicabile: esistono A_1 e A_2 aperti in $\mathbb{R}^2$ e $\mathbb{R}$ rispettivamente, tali che:

(I) $(0, 0) \in A_1$, $1 \in A_2$;

(II) $\forall (x_1, x_2) \in A_1$ esiste unico $\phi(x_1, x_2) \in A_2$ tale che

$$x_1^2 + x_2^2 + \phi(x_1, x_2)^2 - 1 = 0.$$

In effetti, se poniamo

$$A_1 := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 < 1\},$$

$$A_2 = \mathbb{R}^+,$$

$$\begin{cases} \phi : A_1 \to \mathbb{R}, \\ \phi(x_1, x_2) = \sqrt{1 - x_1^2 - x_2^2}, \end{cases}$$

è facile verificare che $\forall (x_1, x_2) \in A_1$ esiste unico $x_3 \in A_2$ tale che

$$f(x_1 x_2, x_3) = O$$

e $x_3 = \phi(x_1, x_2)$.

Osserviamo che, se $(x_1, x_2) \in A_1$, $(x_1, x_2, \phi(x_1, x_2))$ non è l'unico punto della forma (x_1, x_2, x_3) in $\mathbb{R}^3$ che annulla f: c'è anche $(x_1, x_2, -\phi(x_1, x_2))$. Quindi il teorema 8.1 è di carattere locale.

Esempio 8.2. Sia

$$\begin{cases} f : \mathbb{R}^4 \to \mathbb{R}^2, \\ \\ f(x_1, x_2, x_3, x_4) = (3x_1^2 + x_2 - x_3 + x_4^3, x_1 - x_2^2 + 2x_3 + x_4). \end{cases}$$

Si ha $f(0, 0, 0, 0) = (0, 0)$. Inoltre

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

Questa matrice ha rango massimo (2) e il minore costituito dalle ultime due colonne ha determinante non nullo. Possiamo quindi applicare il teorema 8.1 per descrivere localmente (vicino a $(0, 0, 0, 0)$)

$$V = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : f(x_1, x_2, x_3, x_4) = (0, 0).$$

Possiamo dire che esistono A_1 e A_2 aperti in $\mathbb{R}^2$ contenenti entrambi $(0, 0)$ e una funzione $\phi : A_1 \to \mathbb{R}^2$ di classe C^∞ tali che $\forall (x_1, x_2) \in A_1$

$$f(x_1, x_2, \phi(x_1, x_2)) = (0, 0) \quad \forall (x_1, x_2) \in A_1, \phi(0, 0) = (0, 0),$$

e, se $(x_1, x_2, x_3, x_4) \in A_1 \times A_2$ e $f(x_1, x_2, x_3, x_4) = (0, 0)$, necessariamente $(x_3, x_4) = \phi(x_1, x_2)$.

Poniamo

$$\phi(x_1, x_2) = (\phi_1(x_1, x_2), \phi_2(x_1, x_2)).$$

Allora, $\forall (x_1, x_2) \in A_1$

$$\begin{cases} 3x_1^2 + x_2 - \phi_1(x_1, x_2) + \phi_2(x_1, x_2)^3 = 0, \\ \\ x_1 - x_2^2 + 2\phi_1(x_1, x_2) + \phi_2(x_1, x_2) = 0, \end{cases}$$

da cui, derivando rispetto a x_1 e x_2 e prendendo $(x_1, x_2) = (0, 0)$, si ottiene

$$\begin{cases} -D_1\phi_1(0,0) = 0, \\ 1 - D_2\phi_1(0,0) = 0, \\ 1 + 2D_1\phi_1(0,0) + D_1\phi_2(0,0) = 0, \\ 2D_2\phi_1(0,0) + D_2\phi_2(0,0) = 0, \end{cases}$$

da cui

$$D_1\phi_1(0,0) = 0, \quad D_2\phi_1(0,0) = 1, \quad D_1\phi_2(0,0) = -1, D_2\phi_2(0,0) = -2.$$

Dalla formula di Taylor segue allora, per $(x_1, x_2) \to (0, 0)$,

$$\phi_1(x_1, x_2) = -x_2 + o((x_1^2 + x_2^2)^{1/2}),$$
$$\phi_2(x_1, x_2) = -x_1 - 2x_2 + o((x_1^2 + x_2^2)^{1/2}).$$

9 Il teorema dei moltiplicatori di Lagrange

Col teorema 4.1 abbiamo visto che, se $f : A \to \mathbb{R}$ è differenziabile in $x^0 \in \mathring{A}$ e x^0 è un estremante relativo per f, allora x^0 è un punto critico di f, cioe' $\nabla f(x^0) = O$. Il teorema dei moltiplicatori di Lagrange fornisce una condizione necessaria affinche' x^0 sia un estremante relativo per f nel caso la condizione $x^0 \in \mathring{A}$ non sia piu' soddisfatta, ma

$$A = \{x \in \Omega : g(x) = 0\},$$

con

$$A = \{x \in \Omega : g(x) = O\},$$

con Ω aperto in $\mathbb{R}^n$ e $g : \Omega \to \mathbb{R}^m$, con $m < n$, opportunamente regolare.

Cominciamo allora con alcuni risultati e definizioni preliminari, interessanti di per se'.

Definizione 9.1. *Siano $A \subseteq \mathbb{R}^n$, $x^0 \in A$, $v \in \mathbb{R}^n$. Diremo che v e' tangente ad A in x^0 se esistono $\delta \in \mathbb{R}^+$ e $\psi :]-\delta, \delta[\to \mathbb{R}^n$ tali che:*

(I) $\psi(t) \in A$ $\forall t \in]-\delta, \delta[$ e $\psi(0) = x^0$;

(II) ψ è derivabile in 0 e $\psi'(0) = v$.

Indicheremo con $T_A(x^0)$ lo **spazio tangente ad A in x^0**, *vale a dire, l'insieme dei v in $\mathbb{R}^n$ che sono tangenti ad A in x^0.*

Caratterizziamo $T_A(x^0)$ in un caso piuttosto generale, particolarmente importante.

Proposizione 9.1. *Siano Ω un aperto in $\mathbb{R}^n$, $g \in C^1(\Omega; \mathbb{R}^m)$, con $1 \le m < n$. Poniamo*

$$A = \{x \in \Omega : g(x) = O\}$$

Sia $x^0 \in A$ tale che la matrice jacobiana $J_g(x^0)$ abbia rango massimo ($= m$). Allora

$$T_A(x^0) = Ker(dg(x^0)) = \{v \in \mathbb{R}^n : dg(x^0)v = O\}.$$

Dimostrazione Sia $v \in T_A(x^0)$. Allora esistono $\delta \in \mathbb{R}^+$ e $\psi :]-\delta, \delta[\to \mathbb{R}^n$ tali che $\psi(t) \in A$ $\forall t \in]-\delta, \delta[$ e $\psi(0) = x^0$, ψ è derivabile in 0 e $\psi'(0) = v$. Si ha allora $g(\psi(t)) = O$ $\forall t \in]-\delta, \delta[$. Poiche' ψ è derivabile in 0, $g \circ \psi$ è derivabile in 0 e, essendo $g(\psi(t)) = 0$ $\forall t \in]-\delta, \delta[$,

$$(g \circ w)'(0) = O.$$

Ma per il corollario 6.1 (applicato componente per componente di g) si ha

$$O = (g \circ w)'(0) = dg(w(0))w'(0) = dg(x^0)v.$$

Quindi $v \in Ker(dg(x^0))$.

Sia, viceversa, $v \in \mathbb{R}^n$ tale che $dg(x^0)v = 0$. Proviamo che $v \in T_A(x^0)$. Per ipotesi, $J_g(x^0)$ ha rango m. Supponiamo allora (per esempio) che il minore costituito dalle sue ultime righe e colonne sia invertibile. Allora, per il teorema delle funzioni implicite, se $x^0 = (y^0, z^0)$, con $y^0 \in \mathbb{R}^{n-m}$ e $z^0 \in \mathbb{R}^m$, esistono A_1, A_2 aperti in $\mathbb{R}^{n-m}$ e $\mathbb{R}^m$ rispettivamente, contenenti y^0 e z^0, e una funzione $\phi : A_1 \to A_2$ di classe C^1 tali che

$$\{x = (y, z) \in A_1 \times A_2 : g(y, z) = O\} = \{(y, \phi(y)) : y \in A_1\}.$$

Sia $v = (v_1, \ldots, v_n)$. Poniamo allora

$$\begin{cases} \psi :]-\delta, \delta[\to \mathbb{R}^n, \\ \psi(t) = (y_1^0 + tv_1, \ldots, y_{n-m}^0 + tv_{n-m}, \phi(y_1^0 + tv_1, \ldots, y_{n-m}^0 + tv_{n-m})) \\ = (y_1^0 + tv_1, \ldots, y_{n-m}^0 + tv_{n-m}, \psi_{m-n+1}(t), \ldots, \psi_n(t)). \end{cases}$$

Prendiamo δ abbastanza piccolo, in modo che $(y_1^0 + tv_1, \ldots, y_{n-m}^0 + tv_{n-m}) \in A_1$ $\forall t \in]-\delta, \delta[$. Allora, per definizione, $\psi'(0) \in T_A(x^0)$ e percio', per la prima parte della dimostrazione, $dg(x^0)\psi'(0) = O$, cioe',

$$\begin{cases} D_{n-m+1}g_1(x^0)\psi'_{n-m+1}(0) + \cdots + D_n g_1(x^0)\psi'_n(0) \\ = -D_1 g_1(x^0)v_1 - \cdots - D_{n-m}g_1(x^0)v_{n-m}, \\ \ldots \\ D_{n-m+1}g_{n-m}(x^0)\psi'_{n-m+1}(0) + \cdots + D_n g_{n-m}(x^0)\psi'_n(0) \\ = -D_1 g_{n-m}(x^0)v_1 - \cdots - D_{n-m}g_{n-m}(x^0)v_{n-m}, \end{cases}$$

Ma $dg(x^0)v = 0$. Quindi

$$\begin{cases} D_{n-m+1}g_1(x^0)v_{n-m+1} + \cdots + D_n g_1(x^0)w_n \\ = -D_1 g_1(x^0)v_1 - \cdots - D_{n-m}g_1(x^0)v_{n-m}, \\ \ldots \\ D_{n-m+1}g_{n-m}(x^0)w'_{n-m+1}(0) + \cdots + D_n g_{n-m}(x^0)w'_n \\ = -D_1 g_{n-m}(x^0)v_1 - \cdots - D_{n-m}g_{n-m}(x^0)v_{n-m}, \end{cases}$$

Poiche' la matrice

$$\begin{pmatrix} D_{n-m+1}g_1(x^0) & \ldots & D_n g_1(x^0) \\ \ldots & \ldots & \ldots \\ D_{n-m+1}g_m(x^0) & \ldots & D_n g_m(x^0) \end{pmatrix},$$

ha rango massimo m, segue che

$$\psi'_{n-m+1}(0) = v_{n-m+1} \quad \ldots \quad \psi'_n(0) = v_n,$$

da cui $\psi'(0) = v$ e percio' $v \in T_A(x^0)$.

□

Osservazione 9.1. Supponiamo che A soddisfi le ipotesi della proposizione 9.1. Allora $T_A(x^0) = Ker(dg(x^0))$. Quindi $T_A(x^0)$ è un sottospazio vettoriale di $\mathbb{R}^n$. Poiche' $J_g(x^0)$ ha rango m, per la formula di Grassmann (vedi [3]),

$$dim(Ker(dg(x^0))) = n - m.$$

Osservazione 9.2. Siano $m, n \in \mathbb{N}$, con $m < n$, $V \subseteq \mathbb{R}^n$, $p \in \mathbb{N} \cup \{\infty\}$. Diremo che V è una **sottovarieta' differenziale di $\mathbb{R}^n$ di dimensione m e di classe C^p** se $\forall x \in V$ esistono Ω_x aperto in $\mathbb{R}^n$ e $\phi_x \in C^p(\Omega_x; \mathbb{R}^{n-m})$ tali che:
(I) $x \in \Omega_x$;
(II) $V \cap \Omega_x = \{y \in \Omega_x : \phi_x(y) = O\}$;
(III) $J_{\phi_x}(x)$ ha rango $n - m$.
Per quanto visto nell'esempio 8.2, $T_x(V) = ker(d\phi_x)$ è un sottospazio vettoriale di $\mathbb{R}^n$ di dimensione m.

Esempio 9.1. Sia

$$V = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 = 1, x_3 = 0\} = \{x \in \mathbb{R}^3 : g(x) = 0\}$$

con

$$\begin{cases} g : \mathbb{R}^3 \to \mathbb{R}^2, \\ g(x_1, x_2, x_3) = (x_1^2 + x_2^2 - 1, x_3). \end{cases}$$

Si ha, $\forall (x_1, x_2, x_3) \in \mathbb{R}^3$,

$$J_g(x_1, x_2, x_3) = \begin{pmatrix} 2x_1 & 2x_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{9.1}$$

che ha rango massimo ($= 2$) se $x \in V$. Infatti, in tal caso, almeno uno dei due valori x_1 e x_2 è diverso da 0. Se (ad esempio) $x_1 \neq 0$, il minore

$$\begin{pmatrix} 2x_1 & 0 \\ 0 & 1 \end{pmatrix}$$

ha determinante $2x_1 \neq 0$. Quindi la matrice (9.1) ha rango 2 $\forall x \in V$. Possiamo allora prendere $\Omega_x = \mathbb{R}^3$ e $\phi_x = g$ $\forall x \in V$. Dunque V e' una sottovarieta' differenziale di $\mathbb{R}^3$ di dimensione $3 - 2 = 1$ e classe C^∞.

Se $x = (x_1, x_2, x_3) \in V$ e $h = (h_1, h_2, h_3)$, si ha

$$dg(x)h = (2x_1h_1 + 2x_2h_2, h_3).$$

Quindi

$$T_V(x) = ker(dg(x)) = \{h = (h_1, h_2, h_3) \in \mathbb{R}^3 : 2x_1h_1 + 2x_2h_2 = 0, h_3 = 0\}.$$

Se (ad esempio) $x_1 \neq 0$, $T_V(x)$ coincide con

$$\{t(-x_2, x_1, 0) : t \in \mathbb{R}\}.$$

Definizione 9.2. *Siano $A \subseteq \mathbb{R}^n$, $x^0 \in A$. Si definisce* **spazio normale ad A in x^0** *e si indica con $N_A(x^0)$.*

$$\{n \in \mathbb{R}^n : v \cdot n = 0 \quad \forall v \in T_A(x^0)\}.$$

Proposizione 9.2. *Sia A soddisfacente le ipotesi della proposizione 9.1. Sia $x^0 \in A$. Allora $N_A(x^0)$ è un sottospazio vettoriale di $\mathbb{R}^n$ di dimensione m e coincide precisamente con il sottospazio generato da*

$$\{\nabla g_1(x^0), \ldots, \nabla g_m(x^0)\}.$$

Dimostrazione Sappiamo dalla proposizione 9.1 che $T_A(x^0)$ coincide esattamente con $Ker(dg(x^0))$. Quindi $N_A(x^0)$ coincide col suo complemento ortogonale. Ne segue che, poiche' la dimensione di $T_A(x^0)$ è $n - m$, la dimensione di $N_A(x^0)$ è

$$n - (n - m) = m.$$

Inoltre

$$Ker(dg(x^0)) = \{v \in \mathbb{R}^n : \nabla g_1(x^0) \cdot v = \cdots = \nabla g_m(x^0) \cdot v = 0\}$$

Ne segue che lo spazio vettoriale $span(\{\nabla g_1(x^0), \ldots, \nabla g_m(x^0)\})$ generato da $\{\nabla g_1(x^0), \ldots, \nabla g_m(x^0)\}$ è contenuto in $N_A(x^0)$. Ma, poiche' $J_g(x^0)$ ha rango m, la dimensione di $span(\{\nabla g_1(x^0), \ldots, \nabla g_m(x^0)\})$ è proprio m. Ne segue che

$$N_A(x^0) = span(\{\nabla g_1(x^0), \ldots, \nabla g_m(x^0)\}).$$

□

Siamo ora in grado di enunciare e dimostrare il risultato principale di questa sezione.

Teorema 9.1. *Siano $n, m \in \mathbb{N}$, con $m < n$, Ω un aperto di $\mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $g \in C^1(\Omega; \mathbb{R}^m)$. Sia poi*

$$A = \{x \in \Omega : g(x) = O\}.$$

Supponiamo che $J_g(x)$ abbia rango massimo m $\forall x \in A$. Sia $x^0 \in A$ tale che f è differenziabile in x^0 e x^0 è un estremante relativo per $f_{|A}$.

Allora $\nabla f(x^0) \in N_A(x^0)$. Quindi esistono $c_1, \ldots, c_m$ reali (i cosi' detti "moltiplicatori di Lagrange") tali che

$$\nabla f(x^0) = \sum_{i=1}^{m} c_i \nabla g_i(x^0).$$

Dimostrazione Supponiamo (ad esempio) che x^0 sia un punto di minimo relativo per $f_{|A}$. Quindi esiste $r > 0$ tale che $\forall x \in A \cap B(x^0, r)$

$$f(x^0) \leq f(x).$$

Sia $v \in T_A(x^0)$. Allora esiste $\psi :]-\delta, \delta[\to \mathbb{R}^n$ tale che $\psi(t) \in A$ $\forall t \in]-\delta, \delta[$, $\psi(0) = x^0$, ψ è derivabile in 0 e $\psi'(0) = v$. Esiste $\delta_0 \in (0, \delta]$ tale che $\psi(t) \in A \cap B(x^0, r)$ $\forall t \in]-\delta_0, \delta_0[$ e quindi

$$f(\psi(0)) \leq f(\psi(t)) \quad \forall t \in]-\delta_0, \delta_0[.$$

Percio' 0 è un punto di minimo relativo per $f \circ \psi$. Inoltre, per il teorema 6.1, $f \circ \psi$ è derivabile in 0. Segue

$$0 = (f \circ \psi)'(0) = \nabla f(\psi(0)) \cdot \psi'(0) = \nabla f(x^0) \cdot v.$$

Cio' vale $\forall v \in T_A(x^0)$. Quindi $\nabla f(x^0) \in N_A(x^0)$.

La seconda affermazione segue dalla proposizione 9.2.

□

Osservazione 9.3. Con riferimento alle ipotesi del teorema 9.1, se x^0 appartiene a A e $\nabla f(x^0)$ è combinazione lineare di $\nabla g_1(x^0)$,..., $\nabla g_m(x^0)$, x^0 non è necessariamente un estremante relativo per $f_{|A}$. Consideriamo, in proposito, il seguente esempio.

Siano $\Omega = \mathbb{R}^2$, $f : \Omega \to \mathbb{R}$, $f(x_1, x_2) = x_1^3$, $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x_1, x_2) = x_2$. Allora

$$A = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}$$

e $\nabla g(x_1, x_2) = (0, 1)$ $\forall (x_1, x_2) \in \mathbb{R}^2$. Si ha $\nabla f(0, 0) = O$, da cui

$$\nabla f(0, 0) = \lambda \nabla g(0, 0)$$

se $\lambda = 0$. Tuttavia, $(0, 0)$ non è un estremante relativo per $f_{|V}$. Infatti, $\forall \delta \in \mathbb{R}^+$, i punti $(\pm\delta/2, 0)$ appartengono a $B_\delta(0, 0) \cap V$ e

$$-\delta^3/8 = f(-\delta/2, 0) < 0 = f(0, 0) < \delta^3/8 = f(\delta/2, 0).$$

Esempio 9.2. Siano

$$A := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}, \tag{9.2}$$

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x_1, x_2) = x_1 x_2. \tag{9.3}$$

f è continua e A è chiuso e limitato (il fatto che A sia chiuso puo' essere verificato applicando il risultato dell'esercizio 3.3 in [1]). Dunque, per il teorema di Weierstrass (teorema 4.2 in [1]) f ammette minimo e massimo in B. Per determinarli, poniamo

$$\begin{cases} g : \mathbb{R}^2 \to \mathbb{R}, \\ g(x_1, x_2) = x_1^2 + x_2^2 - 1. \end{cases} \tag{9.4}$$

$\mathbb{R}^2$ è aperto, $A = \{x \in \mathbb{R}^2 : g(x) = 0\}$. Inoltre, $\nabla g(x) = 2x$ $\forall x \in \mathbb{R}^2$. Poiché $O \notin A$, le ipotesi del teorema dei moltiplicatori di Lagrange sono tutte soddisfatte. Di conseguenza, se $x = (x_1, x_2)$ è un estremante relativo per $f_{|V}$, esiste $\lambda \in \mathbb{R}$, tale che

$$\nabla f(x) = \lambda \nabla g(x), \tag{9.5}$$

vale a dire,

$$(x_2, x_1) = 2\lambda(x_1, x_2).$$

Dunque (x_1, x_2, λ) è una soluzione del sistema

$$\begin{cases} x_2 = 2\lambda x_1, \\ x_1 = 2\lambda x_2, \\ x_1^2 + x_2^2 - 1 = 0. \end{cases} \tag{9.6}$$

(9.6) ammette le soluzioni (x_1, x_2, λ) seguenti:

$$(1/\sqrt{2}, 1/\sqrt{2}, 1/2), (-1/\sqrt{2}, -1/\sqrt{2}, 1/2),$$

$$(1/\sqrt{2}, -1/\sqrt{2}, -1/2), (-1/\sqrt{2}, 1/\sqrt{2}, -1/2).$$

Si ha

$$f(1/\sqrt{2}, 1/\sqrt{2}) = f(-1/\sqrt{2}, -1/\sqrt{2}) = 1/2,$$

$$f(1/\sqrt{2}, -1/\sqrt{2}) = f(-1/\sqrt{2}, 1/\sqrt{2}) = -1/2.$$

Poiché i punti di massimo e di minimo per $f_{|A}$ sono necessariamente tra quelli trovati, ne deduciamo che $\min_A f = -1/2$ e i punti di minimo sono $(1/\sqrt{2}, -1/\sqrt{2})$ e $(-1/\sqrt{2}, 1/\sqrt{2})$, mentre $\max_A f = 1/2$ e i punti di massimo sono $(1/\sqrt{2}, 1/\sqrt{2})$ e $(-1/\sqrt{2}, -1/\sqrt{2})$.

Osserviamo, infine, che A è connesso per archi (vedi l'esercizio 9.1). Dal teorema di Bolzano (teorema 5.3 in [1]), sappiamo che $f(A)$ è un intervallo. Concludiamo allora che

$$f(A) = [-1/2, 1/2].$$

Sia

$$A := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}. \tag{9.7}$$

Determiniamo $f(B)$. Anche in questo caso B è chiuso, limitato e connesso per archi. Ne segue che

$$f(B) = [\min(f(B)), max(f(B))].$$

I punti di massimo e di minimo possono essere nell'interno di B o in A. Quelli in A sono, evidentemente, estremanti relativi per $f_{|A}$ e sono percio' quelli della forma $(\pm 1/\sqrt{2}, \pm 1/\sqrt{2})$ (prendendo $+$ o $-$ in tutti i modi possibili). Quelli interni sono necessariamente punti critici di f. L'unico punto critico di f è $(0,0)$ e

$$-\frac{1}{2} < 0 = f(0,0) < \frac{1}{2}.$$

Quindi $f(B) = f(A) = [-1/2, 1/2]$.

Esempio 9.3. Ricaviamo la formula classica della distanza di un punto $x^0 \in \mathbb{R}^3$ da un piano A.

Sia

$$A := \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : a_1x_1 + a_2x_2 + a_3x_3 + a_4 = 0\},$$

con $a_j \in \mathbb{R}$ $(1 \leq j \leq 4)$, $(a_1, a_2, a_3) \neq (0, 0, 0)$. Sia poi

$$x^0 = (x_1^0, x_2^0, x_3^0).$$

Cerchiamo $x \in A$, tale che

$$d(x, x^0) = \min_{y \in A} d(y, x^0). \tag{9.8}$$

Il fatto che il problema (9.8) ammetta soluzioni non è ovvio. Infatti, la funzione

$$\begin{cases} d : \mathbb{R}^3 \to \mathbb{R}, \\ d(y) = d(y, x^0) \end{cases}$$

$y \to d(y, x^0)$ è continua da $\mathbb{R}^3$ a $\mathbb{R}$, A è chiuso, ma non limitato. Dunque, il teorema 4.2 in [1] non è direttamente applicabile. Per verificare che d ammette minimo su A, possiamo ragionare come segue.

Fissiamo arbitrariamente $x^1 \in A$. Sia $\rho := d(x^1, x^0)$. Sia $x \in A$, tale che $\|x\| > \rho + \|x^0\|$. Allora si ha

$$d(x) \quad = \|x - x^0\| \geq \|x\| - \|x^0\| > \rho. \tag{9.9}$$

Osserviamo che (9.9) implica

$$\|x^1\| \leq \rho + \|x^0\|.$$

Poniamo

$$A' := \{x \in A : \|x\| \leq \rho + \|x^0\|\}.$$

Allora $A' \neq \emptyset$, perché $x^1 \in A'$, è chiuso e limitato (sul fatto che A' sia effettivamente chiuso, si veda l'esercizio 9.2). Percio', d ammette minimo in A'. Poiché $x^1 \in A'$,

$$\min_{A'} d \leq d(x^1, x^0) = \rho.$$

Sia x^2 un punto di minimo per d in A'. Sia poi $x \in A$. Se $x \in A'$, ovviamente, $d(x^2) \leq d(x)$. Se $x \notin A'$, allora $\|x\| > \|x^0\| + \rho$, per cui

$$d(x^2) \leq \rho < d(x).$$

Concludiamo che d ammette minimo in A e vale

$$\min_A d = \min_{A'} d.$$

Poniamo ora

$$\begin{cases} f : \mathbb{R}^3 \to \mathbb{R}, \\ f(x) = d(x, x^0)^2 = \|x - x^0\|^2. \end{cases}$$

E' chiaro che, poiché d ammette minimo su A, anche f ammette minimo su A. Inoltre, i punti di minimo sono gli stessi e

$$\min_A f = (\min_A d)^2.$$

Se x è un punto di minimo per f su A, per il teorema dei moltiplicatori di Lagrange, esiste $\lambda \in \mathbb{R}$, tale che

$$\nabla f(x) = \lambda \nabla g(x), \tag{9.10}$$

con $g(x) = a_1x_1 + a_2x_2 + a_3x_3 + a_4$, cioè ,

$$2(x - x^0) = \lambda(a_1, a_2, a_3). \tag{9.11}$$

Se $x = (x_1, x_2, x_3)$ imponendo la condizione $x \in A$, si ottiene da (9.11)

$$a_1x_1^0 + a_2x_2^0 + a_3x_3^0 + \frac{\lambda(a_1^2 + a_2^2 + a_3^2)}{2} + a_4 = 0,$$

da cui

$$\lambda = -\frac{2(a_1x_1^0 + a_2x_2^0 + a_3x_3^0 + a_4)}{a_1^2 + a_2^2 + a_3^2}.$$

Essendo

$$x = x^0 + \frac{\lambda}{2}(a_1, a_2, a_3),$$

si ha allora

$$\min_A(d) = \|x - x^0\| = \frac{|\lambda|}{2}\sqrt{a_1^2 + a_2^2 + a_3^2} = \frac{|a_1x_1^0 + a_2x_2^0 + a_3x_3^0 + a_4|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}.$$

Osservazione 9.4. Concludiamo questa sezione con un'applicazione del teorema dei moltiplicatori di Lagrange alla dimostrazione del teorema spettrale, (teorema 4.4). Per semplicita', ci limiteremo a considerare solo il caso particolare $V = \mathbb{R}^n$ con il prodotto scalare euclideo $< \cdot, \cdot >$. Supporremo, per evitare il caso banale $n = 1$, $n \geq 2$. Quindi $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ e

$$< Av, w >=< v, Aw > \forall v, w \in \mathbb{R}^n. \tag{9.12}$$

Osserviamo che cio' equivale a chiedere che la matrice associata ad A coincida con la sua trasposta. Infatti, se

$$v = (v_1, \dots, v_n), \quad w = (w_1, \dots, w_n),$$

e se tale matrice è

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & & \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

si ha

$$< Av, w >= \sum_{i=1}^{n} (Av)_i w_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_j w_i,$$

mentre

$$< v, Aw >= \sum_{j=1}^{n} v_j (Aw)_j = \sum_{j=1}^{n} (v_j \sum_{i=1}^{n} a_{ji} w_i) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji} v_j w_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_i w_j.$$

Deve quindi essere

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_j w_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_i w_j$$

$\forall v, w \in \mathbb{R}^n$. Siano $i_0, j_0 \in \{1, \dots, n\}$. Prendendo $v = e^{j_0}$, $w = e^{i_0}$, si ha allora

$$a_{i_0 j_0} =< Ae^{j_0}, e^{i_0} >=< e^{j_0}, Ae^{i_0} >= a_{j_0 i_0},$$

per cui $a_{ij} = a_{ji}$ $\forall i, j \in \{1, \dots, n\}$.

Il viceversa, cioe' se $a_{ij} = a_{ji}$ $\forall i, j \in \{1, \dots, n\}$, vale (9.12), è chiaro.

Consideriamo la forma quadratica

$$Qv :=< v, Av > \quad v \in \mathbb{R}^n.$$

Si ha, per $j = 1, \dots, n$, $v \in \mathbb{R}^n$, $t \in \mathbb{R} \setminus \{0\}$,

$$\tfrac{Q(v+te^j)-Qv}{t} = \tfrac{<v+te^j, A(v+te^j)> - <v, Av>}{t}$$

$$=< e^j, Av > + < v, Ae^j > +t < e^j, Ae^j > .$$

Quindi

$$(D_j Q)(v) = \lim_{t \to 0} \frac{Q(v + te^j) - Qv}{t} =< e^j, Av > + < v, Ae^j >= 2 < Av, e^j >,$$

essendo

$$< e^j, Av >=< Ae^j, v >=< v, Ae^j > .$$

Quindi

$$(\nabla Q)(v) = 2Av \quad \forall v \in \mathbb{R}^n.$$

Q è continua e ammette quindi minimo in

$$A_1 = \{v \in \mathbb{R}^n : \|v\| = 1\} = \{v \in \mathbb{R}^n : g^1(v) = 0\},$$

con $g^1(v) = \|v\|^2 - 1$. Osserviamo che, se $v = (v_1, \ldots, v_n)$,

$$J_{g^1}(v) = \begin{pmatrix} 2v_1 & ... & 2v_n \end{pmatrix}$$

che ha rango 1 se $v \neq O$. Allora, se v^1 è un punto di minimo per $Q_{|A_1}$, per il teorema dei moltiplicatori di Lagrange, esiste $c_1 \in \mathbb{R}$ tale che

$$2Av^1 = 2c_1 v^1,$$

cioè v^1 è un autovettore di A corrispondente all'autovalore c_1. Osserviamo anche che

$$\min_{A_1}(Q) = Q(v^1) =< v^1, Av^1 >= c_1\|v^1\|^2 = c_1.$$

In generale, supponiamo che $1 \leq k < n$ e di aver determinato k autovettori di A $v^1, \ldots, v^k$ di norma 1, a due a due ortogonali, con corrispondenti autovalori reali $c_1, \ldots, c_k \in \mathbb{R}$, $c_1 \leq \cdots \leq c_k$. Poniamo

$$A_{k+1} := \{v \in \mathbb{R}^n : \|v\| = 1, < v, v^1 >= \cdots =< v, v^k >= 0\}$$

$A_{k+1} = A_n$ ha due elementi opposti tra loro. Siano questi $\pm v^n$. Ciascuno di essi è un autovettore di A. Infatti, per $i = 1, \ldots, n-1$ si ha

$$< Av^n, v^i >=< v^n, Av^i >= c_i < v^n, v^i >= 0.$$

Quindi $Av^n \in \{v^1, \ldots, v^{n-1}\}^{\perp}$ che ha dimensione uno, e percio' esiste $c_n \in \mathbb{R}$ tale che

$$Av^n = c_n v^n.$$

Supponiamo $k < n-1$. A^{k+1} è chiuso, limitato, non vuoto. Quindi $Q_{|A_{k+1}}$ ammette minimo. Si ha

$$A_{k+1} = \{v \in \mathbb{R}^n : g^{k+1}(v) = O\},$$

con

$$g^{k+1}(v) = (\|v\|^2 - 1, < v, v^1 >, \ldots, < v, v^k >).$$

Se $v^i = (v^i_1, \ldots, v^i_n)$ $(i = 1, \ldots, k)$, $v = (v_1, \ldots, v_n)$, si ha

$$J_{g^{k+1}}(v) = \begin{pmatrix} 2v_1 & ... & 2v_n \\ v^1_1 & ... & v^1_n \\ \ldots & & \\ v^i_1 & ... & v^i_n \\ \ldots & & \\ v^k_1 & ... & v^k_n \end{pmatrix}$$

La matrice $J_{g^{k+1}}(v)$ ha rango $k+1$. Per verificarlo, utilizzando il fatto ben noto che il rango per righe coincide col rango per colonne (vedi [3], cap. 7), basta far vedere che, $\forall v \in A_{k+1}$. $\{v, v^1, \ldots, v^k\}$ sono linearmente indipendenti. Siano infatti $c, c_1, \ldots, c_k \in \mathbb{R}$ tali che

$$cv + \sum_{i=1}^{k} c_i v^i = O.$$

Per $j \in \{1, \dots, k\}$ si ha

$$0 =< cv + \sum_{i=1}^{k} c_i v^i, v^j >= c_j.$$

Da cio' segue $cv = 0$, da cui $c = 0$, essendo $\|v\| = 1$.

Allora, per il teorema dei moltiplicatori di Lagrange, esistono $c_{k+1,1}, \dots, c_{k+1,k+1}$ in $\mathbb{R}$ tali che, se v^{k+1} è un punto di minimo per $Q_{|A_{k+1}}$

$$2Av^{k+1} = \sum_{i=1}^{k} c_{k+1,i} v^i + 2c_{k+1,k+1} v^{k+1}.$$

Per $j = 1, \dots, k$, si ha

$$c_{k+1,j} =< \sum_{i=1}^{k} c_{k+1,i} v^i + 2c_{k+1,k+1} v^{k+1}, v^j >$$

$$= 2 < Av^{k+1}, v^j >= 2 < v^{k+1}, Av^j >= 2c_j < v^{k+1}, v^j >= 0.$$

Quindi

$$Av^{k+1} = c_{k+1,k+1} v^{k+1}.$$

Concludiamo che v^{k+1} è un autovettore di A ortogonale a $v^1, \dots, v^k$, con autovalore corrispondente $c_{k+1} = c_{k+1,k+1}$.

Nel modo descritto si costruisce una base ortonormale $(v^i)_{i=1,\dots,n}$ costituita di auvettori di A.

Osserviamo che, per $i = 1, \dots, n$,

$$c_i = \min_{A_i} Q$$

e $A_n \subseteq A_{n-1} \subseteq ... \subseteq A_1$. Percio'

$$c_1 \le \dots \le c_n.$$

Esercizio 9.1. Provare che $\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$ è connesso per archi.

(Sugg: osservare che è l'immagine della funzione $h : [0, 2\pi] \to \mathbb{R}$, $h(t) = (\cos(t), \sin(t)))$.

Esercizio 9.2. Provare che l'insieme A' nell'esempio 9.3 è chiuso e limitato in $\mathbb{R}^3$.

Esercizio 9.3. Siano $A \subseteq \mathbb{R}^n$ chiuso, $f : A \to \mathbb{R}$ continua.

(I) Supponiamo che esistano $x^0 \in A$ e $R \ge 0$ tale che $f(x^0) \le f(x)$ $\forall x \in A \cap \{y \in \mathbb{R}^n : \|y\| \ge R\}$. Provare che f ammette minimo.

(II) Supponiamo che esistano $x^0 \in A$ e $R \ge 0$ tale che $f(x^0) \ge f(x)$ $\forall x \in A \cap \{y \in \mathbb{R}^n : \|y\| \ge R\}$. Provare che f ammette massimo.

Esercizio 9.4. Sia $f : A \to \mathbb{R}$. Determinare l'immagine $f(A)$ nei casi seguenti:

(I) $A = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 \le 1, x_3^2 \ge 1/4\}$, $f(x_1, x_2, x_3) = x_3^2 - x_1^2 - x_2^2$.
(Attenzione ! A non è connesso per archi.)

(II) $A = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 \le (x_3 - 4)^2, 0 \le x_3 \le 8\}$, $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$.

(III) $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 \le \max\{1, x_2^2\}, 0 \le x_2 \le 2\}$, $f(x_1, x_2) = x_1^2 + x_2^2$.

(IV) $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \le g^2\}$ $(g \in \mathbb{R}^+)$, $f(x_1, x_2) = x_1^4 + x_2^4 - 8(x_1^2 + x_2^2)$.

(V) $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + 4x_2^2 = 4\}$, $f(x_1, x_2) = x_1^2 - x_2$.

(VI) $A = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 - 1 = 0, x_2 \le 1/2\}$, $f(x_1, x_2, x_3) = x_1 - 2x_2 + 2x_3$.

(VII) $A = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$, $f(x_1, x_2, x_3) = x_1 x_2 x_3$.

(VIII) $A = \{(x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2| \le 1\}$, $f(x_1, x_2) = x_1 + 2x_2$.

(IX) $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + 4x_2^2 \le 1\}$, $f(x_1, x_2) = x_1^2 - x_2^2 + x_1$.

(X) $A = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + 3x_2^2 \le 1, R \ge x_3 \ge 1\}$ $(R \ge 1)$, $f(x_1, x_2, x_3) = \arctan(\frac{1}{x_1^2 + x_2^2 + x_3^2})$.

References

[1] D. Guidetti, "$\mathbb{R}^n$ e spazi metrici".

[2] E. Lanconelli, "Lezioni di analisi matematica 1", Pitagora Editrice (1994).

[3] S. Lang, "Algebra lineare".