CALCOLO DIFFERENZIALE

A. Brini

September 5, 2016

Contents

Part I

Spazi metrici, normati, di Banach

1 Spazi metrici

1.1 Definizione e primi esempi

Sia X un insieme. Si dice *distanza, o, metrica* una funzione

$$d : X \times X \to \mathbf{R}$$

tale che

1. $d(\underline{x}, \underline{x'}) \geq 0;\ d(\underline{x}, \underline{x'}) = 0$ *se e solo se* $\underline{x} = \underline{x'}$.
2. $d(\underline{x}, \underline{x'}) = d(\underline{x'}, \underline{x})$.
3. $d(\underline{x}, \underline{x'}) \leq d(\underline{x}, \underline{x''}) + d(\underline{x''}, \underline{x'})$.

La coppia (X, d) si dice *SPAZIO METRICO.*
ESEMPIO FONDAMENTALE 1 (La metrica euclidea in $\mathbf{R}^n$)
Consideriamo l'insieme delle $n-$ple ordinate di numeri reali:

$$\mathbf{R}^n = \{(x_1, x_2, \ldots, x_n); x_i \in \mathbf{R}, i = 1, \ldots, n\}.$$

L'insieme $\mathbf{R}^n$ possiede una struttura di spazio vettoriale (reale) di dimensione n, ove le operazioni di *somma* e *moltiplicazione esterna per scalari* sono definite nel modo ovvio.

$\mathbf{R}^n$ e' uno spazio metrico *detto spazio metrico euclideo*, quando munito della *metrica (o, distanza) euclidea*

$$d : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$$

definita come segue. Dati $\underline{x} = (x_1, x_2, \ldots, x_n)$, $\underline{x'} = (x'_1, x_2, \ldots, x_n) \in \mathbf{R}^n$, per definizione si ha:

$$d(\underline{x}, \underline{x'}) = \sqrt{(x_1 - x'_1)^2 + \cdots + (x_n - x'_n)^2}.$$

ESEMPIO FONDAMENTALE 2 (La metrica della TOPOLOGIA DISCRETA SU UN INSIEME X)

Sia X un insieme qualsiasi. Definiamo una funzione

$$d : X \times X \to \mathbf{R}$$

ponendo

$$d(x, y) = 0 \ se \ x = y, \quad d(x, y) = 1 \ se \ x \neq y,$$

per ogni $x, y \in X$.

E' immediato riconoscere che la funzione d e' una metrica, e, quindi, (X, d) e' spazio metrico.

1.2 Spazi metrici e topologia

Sia (X, d) spazio metrico.

Dati $x_0 \in X$ e $r \in \mathbf{R}^+$, si dice *intorno sferico (aperto)* di centro x_0 e raggio r il sottoinsieme

$$I(x_0, r) = \{x \in X; d(x, x_0) < r\} \subseteq X.$$

Un sottoinsieme $A \subseteq X$ si dice *limitato* se e solo se il suo *diametro*

$$d(A) = sup \ \{d(x, x'); x, x' \in A\}$$

e' FINITO.

Una ulteriore utile ed intuitiva caratterizzazione della limitatezza di un sottinsieme e' la seguente: A e' limitato se e solo se

$$\forall x \in X \ \exists \ r \in \mathbf{R}^+ \ tale \ che \ A \subseteq I(x, r).$$

Un sottoinsieme $A \subseteq X$ si dice *aperto* se e solo se

$$\forall \ x \in A \ \exists \ r \in \mathbf{R}^+ \ tale \ che \ I(x, r) \subseteq A.$$

Dati $B \subseteq X$ ed un punto $x_0 \in X$, si dice che x_0 e' *di accumulazione per* B se e solo se

$$\forall \ r \in \mathbf{R}^+ \ risulta \ (I(x_0, r) - \{x_0\}) \cap B \neq \emptyset.$$

Un sottoinsieme $D \subseteq X$ si dice *chiuso* se e solo se il suo complementare $D^c = X - D$ e' aperto in X.

Proposizione 1. *Un sottoinsieme $D \subseteq \mathbf{R}^n$ e' chiuso se e solo se contiene tutti i suoi punti di accumulazione.*

La dimostrazione e' lasciata come (utile) esercizio.

OSSERVAZIONI

- In generale, un sottoinsieme $C \subseteq X$ puo' essere non aperto ne' chiuso. Ad esempio, l'intervallo $[0,1[$ non e' chiuso ne' aperto in $\mathbf{R}$*euclideo*; infatti il punto 1 e' di accumulazione per $[0,1[$ ma non vi appartiene, mentre sul punto $0 \in [0,1[$ e' impossibile costruire un interno contenuto in $[0,1[$. (VERIFICARE)
- I sottoinsiemi $\emptyset$ e $\mathbf{R}^n$ sono contemporaneamente aperti e chiusi.

Nel caso degli spazi $\mathbf{R}^n$ *euclidei* vale un profondo Teorema, che puo' essere riguardato come un parziale "viceversa" delle affermazioni precedenti:

Teorema 1. *(Teorema di CONNESSIONE per $\mathbf{R}^n$ euclideo)*

In $\mathbf{R}^n$ euclideo, gli unici sottoinsiemi contemporaneamente aperti e chiusi sono quelli banali, vale a dire $\emptyset$ e $\mathbf{R}^n$.

Corollario 1. *Sia $C \subseteq \mathbf{R}^n$, $C \neq \emptyset$, $C \neq \mathbf{R}^n$.*

- *Se C e' aperto, allora C non e' chiuso.*
- *Se C e' chiuso, allora C non e' aperto.*

Proposizione 2. *(Proprieta' fondamentali delle famiglie di aperti e di chiusi) In uno spazio metrico qualsiasi (X,d) valgono le seguenti affermazioni:*

- *A_1) L'unione di una famiglia qualsiasi di insiemi aperti e' un insieme aperto.*
- *A_2) L'intersezione di una famiglia FINITA di insiemi aperti e' un insieme aperto.*
- *C_1) L'unione di una famiglia FINITA di insiemi chiusi e' un insieme chiuso.*
- *C_2) L'intersezione di una famiglia qualsiasi di insiemi chiusi e' un insieme chiuso.*

La dimostrazione segue direttamente dalle definizioni.

CONTROESEMPI IN R euclideo

- $\cap_{n \in \mathbf{Z}^+}]-\frac{1}{n}, \frac{1}{n}[= \{0\}$ e' intersezione di aperti, ma e' chiuso e, quindi *NON* aperto.
- $\cup_{n \in \mathbf{Z}^+} [-1+\frac{1}{n}, 1-\frac{1}{n}] =]-1,1[$ e' unione di chiusi, ma e' aperto e, quindi *NON* chiuso.

1.3 Funzioni continue $f : A \subseteq X \to Y$

Sia $f : A \subseteq X \to Y$ una funzione d allo spazio metrico (Y, d_Y), A aperto in X.

Dato $x_0 \in A$, si dice che la funzione f e' *continua nel punto* x_0 se e solo se

$$\forall\ \varepsilon \in \mathbf{R}^+,\ \exists\ \delta \in \mathbf{R}^+ \ \textit{tale che risulti}\ d_Y(f(x), f(x_0)) < \varepsilon,\ \textit{per ogni}\ x_0 \in I(x_0, \delta) \cap A.$$

Questa e' una definizione "locale"; globalmente, la funzione f si dice *continua sul dominio* A se e solo se risulta continua *ogni punto* di A.

Abbiamo una utilissima caratterizzazione della continuita' globale di una funzione

$$f : A \subseteq X \to Y.$$

In primo luogo, ricordiamo la definizione di *preimmagine (o, fibra)* di un sottoinsieme $C \subseteq Y$ rispetto alla funzione f. Dato $C \subseteq Y$, la sua preimmagine rispetto alla funzione f e' l'insieme:

$$f^{-1}[C] = \{x \in A; f(x) \in C\} \subseteq A \subseteq X.$$

Teorema 2. *Sia data una funzione* $f : A \subseteq X \to Y$, $A \neq \emptyset$.

Sono equivalenti le seguenti affermazioni:

- f *e' continua su* A.
- $\forall\ B \subseteq Y,\ B$ *aperto in* $Y,\ \exists\ B_1 \subseteq X,\ \ B_1$ *aperto in* $X,$ *tale che* $f^{-1}[B] = A \cap B_1$.
- $\forall\ D \subseteq Y,\ D$ *chiuso in* $Y,\ \exists\ D_1 \subseteq X,\ \ D_1$ *chiuso in* $X,$ *tale che* $f^{-1}[D] = A \cap D_1$.

Corollario 2. *Sia data una funzione* $f : X \to Y$.

Sono equivalenti le seguenti affermazioni:

- f *e' continua su* X.
- $\forall\ B \subseteq Y,\ B$ *aperto in* $Y,\ f^{-1}[B]$ *e' aperto in* X
- $\forall\ D \subseteq Y,\ D$ *chiuso in* $Y,\ f^{-1}[D]$ *e' chiuso in* X.

ESEMPIO FONDAMENTALE 3 (La metrica della CONVERGENZA UNIFORME SU $L(X, Y)$ E $C(X, Y)$ (X, Y spazi metrici)

1.4 Spazi metrici e successioni

Sia (X, d) spazio metrico, e sia $(x_n)_{n\in\mathbb{N}}$ una successione in X, cioe' $x_n \in X$, per ogni $n \in \mathbb{N}$.

La successione $(x_n)_{n\in\mathbb{N}}$ si dice *convergente* ad $x \in X$ se e solo se $\forall\varepsilon \in \mathbb{R}^+$ esiste $\upsilon \in \mathbb{N}$ tale che

$$d(x_n, x) < \varepsilon$$

per ogni $n > \upsilon$.

La successione $(x_n)_{n\in\mathbb{N}}$ si dice *di Cauchy* se e solo se $\forall\varepsilon \in \mathbb{R}^+$ esiste $\upsilon \in \mathbb{N}$ tale che

$$d(x_n, x_m) < \varepsilon$$

per ogni $n, m > \upsilon$.

Dalla disuguaglianza triangolare, segue che *ogni successione convergente e' di Cauchy.*

Il viceversa e', in generale, *FALSO*, e dipende dalle proprieta' intrinsrche dello spazio metrico in questione.

Part II

$\mathbf{R}^n$ euclideo

2 $\mathbb{R}^n$ euclideo come spazio metrico

2.1 La metrica euclidea in $\mathbf{R}^n$

Consideriamo l'insieme delle $n-$ple ordinate di numeri reali:

$$\mathbf{R}^n = \{(x_1, x_2, \dots, x_n); x_i \in \mathbf{R}, i = 1, \dots, n\}.$$

L'insieme $\mathbf{R}^n$ possiede una struttura di spazio vettoriale (reale) di dimensione n, ove le operazioni di *somma* e *moltiplicazione esterna per scalari* sono definite nel modo ovvio.

D' ora in avanti riguarderemo sempre $\mathbf{R}^n$ come *spazio metrico euclideo*, vale a dire muniremo $\mathbf{R}^n$ di una funzione, detta *metrica (o, distanza) euclidea*

$$d : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$$

definita come segue. Dati $\underline{x} = (x_1, x_2, \dots, x_n)$, $\underline{x'} = (x'_1, x_2, \dots, x_n) \in \mathbf{R}^n$, per definizione si ha:

$$d(\underline{x}, \underline{x'}) = \sqrt{(x_1 - x'_1)^2 + \cdots + (x_n - x'_n)^2}.$$

La metrica euclidea appena definita soddisfa appunto gli assiomi di metrica, vale a dire:

1. $d(\underline{x}, \underline{x'}) \geq 0$; $d(\underline{x}, \underline{x'}) = 0$ *se e solo se* $\underline{x} = \underline{x'}$.

2. $d(\underline{x}, \underline{x'}) = d(\underline{x'}, \underline{x})$.

3. $d(\underline{x}, \underline{x'}) \leq d(\underline{x}, \underline{x''}) + d(\underline{x''}, \underline{x'})$.

2.2 Primi concetti topologici in $\mathbf{R}^n$

Dati $x_0 \in \mathbf{R}^n$ e $r \in \mathbf{R}^+$, si dice *intorno sferico (aperto)* di centro x_0 e raggio r il sottoinsieme

$$I(x_0, r) = \{x \in \mathbf{R}^n; d(x, x_0) < r\} \subseteq \mathbf{R}^n.$$

Un sottoinsieme $A \subseteq \mathbf{R}^n$ si dice *limitato* se e solo se il suo *diametro*

$$d(A) = sup\ \{d(x, x'); x, x' \in A\}$$

e' FINITO. Una ulteriore utile ed intuitiva caratterizzazione della limitatezza di un sottinsieme e' la seguente: A e' limitato se e solo se

$$\forall x \in \mathbf{R}^n\ \exists\ r \in \mathbf{R}^+ \ \textit{tale che}\ A \subseteq I(x, r).$$

Un sottoinsieme $A \subseteq \mathbf{R}^n$ si dice *aperto* se e solo se

$$\forall\ x \in A\ \exists\ r \in \mathbf{R}^+ \ \textit{tale che}\ I(x, r) \subseteq A.$$

Dati $B \subseteq \mathbf{R}^n$ ed un punto $x_0 \in \mathbf{R}^n$, si dice che x_0 e' *di accumulazione per* B se e solo se

$$\forall\ r \in \mathbf{R}^+ \ \textit{risulta}\ (I(x_0, r) - \{x_0\}) \cap B \neq \emptyset.$$

Un sottoinsieme $D \subseteq \mathbf{R}^n$ si dice *chiuso* se e solo se il suo complementare $D^c = \mathbf{R}^n - D$ e' aperto in $\mathbf{R}^n$.

Proposizione 3. *Un sottoinsieme $D \subseteq \mathbf{R}^n$ e' chiuso se e solo se contiene tutti i suoi punti di accumulazione.*

La dimostrazione e' lasciata come (utile) esercizio.

OSSERVAZIONI

- In generale, un sottoinsieme $C \subseteq \mathbf{R}^n$ puo' essere non aperto ne' chiuso. Ad esempio, l'intervallo $[0, 1[$ non e' chiuso ne' aperto in $\mathbf{R}$; infatti il punto 1 e' di accumulazione per $[0, 1[$ ma non vi appartiene, mentre sul punto $0 \in [0, 1[$ e' impossibile costruire un interno contenuto in $[0, 1[$. (VERIFICARE)
- I sottoinsiemi $\emptyset$ e $\mathbf{R}^n$ sono contemporaneamente aperti e chiusi.

Nel caso degli spazi $\mathbf{R}^n$ *euclidei* vale un profondo Teorema, che puo' essere riguardato come un parziale "viceversa" delle affermazioni precedenti:

Teorema 3. *(Teorema di CONNESSIONE per $\mathbf{R}^n$ euclideo)*

In $\mathbf{R}^n$ euclideo, gli unici sottoinsiemi contemporaneamente aperti e chiusi sono quelli banali, vale a dire $\emptyset$ e $\mathbf{R}^n$.

Corollario 3. *Sia $C \subseteq \mathbf{R}^n$, $C \neq \emptyset$, $C \neq \mathbf{R}^n$.*

- *Se C e' aperto, allora C non e' chiuso.*

- *Se C e' chiuso, allora C non e' aperto.*

Proposizione 4. *(Proprieta' fondamentali delle famiglie di aperti e di chiusi)*

- A_1) *L'unione di una famiglia qualsiasi di insiemi aperti e' un insieme aperto.*
- A_2) *L'intersezione di una famiglia FINITA di insiemi aperti e' un insieme aperto.*
- C_1) *L'unione di una famiglia FINITA di insiemi chiusi e' un insieme chiuso.*
- C_2) *L'intersezione di una famiglia qualsiasi di insiemi chiusi e' un insieme chiuso.*

CONTROESEMPI IN R euclideo

- $\cap_{n\in\mathbf{Z}^+}\ ]-\frac{1}{n},\frac{1}{n}[=\{0\}$ e' intersezione di aperti, ma e' chiuso e, quindi *NON* aperto.
- $\cup_{n\in\mathbf{Z}^+}\ [-1+\frac{1}{n},1-\frac{1}{n}]=]-1,1[$ e' unione di chiusi, ma e' aperto e, quindi *NON* chiuso.

2.3 Funzioni continue $f: A\subseteq\mathbf{R}^n\to\mathbf{R}$

Sia $f: A\subseteq\mathbf{R}^n\to\mathbf{R}$ una funzione a valori reali. Nel seguito, supporremo sempre che il *dominio* $\mathbf{R}^n$ ed il codominio $\mathbf{R}$ siano, come spazi metrici, muniti delle rispettive *metriche euclidee.*

Dato $x_0\in A$, si dice che la funzione f e' *continua nel punto* x_0 se e solo se

$$\forall\ \varepsilon\in\mathbf{R}^+,\ \exists\ \delta\in\mathbf{R}^+\ \textit{tale che risulti}\ |f(x)-f(x_0)|<\varepsilon,\ \textit{per ogni}\ x_0\in I(x_0,\delta)\cap A.$$

Questa e' una definizione "locale"; globalmente, la funzione f si dice *continua sul dominio* A se e solo se risulta continua *ogni punto* di A.

Abbiamo una utilissima caratterizzazione della continuita' globale di una funzione

$$f: A\subseteq\mathbf{R}^n\to\mathbf{R}.$$

In primo luogo, ricordiamo la definizione di *preimmagine (o, fibra)* di un sottoinsieme $C\subseteq\mathbf{R}$ rispetto alla funzione f. Dato $C\subseteq\mathbf{R}$, la sua preimmagine rispetto alla funzione f e' l'insieme:

$$f^{-1}[C]=\{x\in A; f(x)\in C\}\subseteq A\subseteq\mathbf{R}^n.$$

Teorema 4. *Sia data una funzione $f: A\subseteq\mathbf{R}^n\to\mathbf{R}$, $A\neq\emptyset$.*

Sono equivalenti le seguenti affermazioni:

- *f e' continua su A.*
- $\forall\ B\subseteq\mathbf{R},\ B$ *aperto in* $\mathbf{R},\ \exists\ B_1\subseteq\mathbf{R}^n,\ \ B_1$ *aperto in* $\mathbf{R}^n,$ *tale che* $f^{-1}[B]=A\cap B_1.$

- $\forall\ D \subseteq \mathbf{R},\ D$ *chiuso in* $\mathbf{R},\ \exists\ D_1 \subseteq \mathbf{R}^n,\ \ D_1$ *chiuso in* $\mathbf{R}^n$, *tale che* $f^{-1}[D] = A \cap D_1$.

Corollario 4. *Sia data una funzione* $f : \mathbf{R}^n \to \mathbf{R}$.

Sono equivalenti le seguenti affermazioni:

- f *e' continua su* $\mathbf{R}^n$.
- $\forall\ B \subseteq \mathbf{R},\ B$ *aperto in* $\mathbf{R},\ f^{-1}[B]$ *e' aperto in* $\mathbf{R}^n$.
- $\forall\ D \subseteq \mathbf{R},\ D$ *chiuso in* $\mathbf{R},\ f^{-1}[D]$ *e' chiuso in* $\mathbf{R}^n$.

2.4 Successioni in $\mathbf{R}^n$ euclideo; successioni convergenti e successioni Cauchy

Consideriamo una *successione* $(x_n)_{n\in\mathbf{N}}$ ad elementi in $\mathbf{R}^n$, vale a dire $x_n \in \mathbf{R}^n$, per ogni $n \in \mathbf{N}$.

Si dice che la successione $(x_n)_{n\in\mathbf{N}}$ e' *convergente* al punto $x_0 \in \mathbf{R}^n$ se e solo se

$$\forall\ \varepsilon \in \mathbf{R}^+,\ \exists\ \nu \in \mathbf{N} \text{ tale che } d(x_n, x_0) < \varepsilon,\ \text{per ogni } n > \nu.$$

In questo caso si scrive anche, per brevita',

$$\lim_{n\to\infty}\ x_n = x_0.$$

Una successione $(x_n)_{n\in\mathbf{N}}$ ad elementi in $\mathbf{R}^n$ si dice *di Cauchy* se e solo se

$$\forall\ \varepsilon \in \mathbf{R}^+,\ \exists\ \nu \in \mathbf{N} \text{ tale che } d(x_n, x_m) < \varepsilon,\ \text{per ogni } n, m > \nu.$$

Dalla disuguaglianza triangolare segue subito (utile e semplice esercizio) che una successione convergente e' sempre di Cauchy. Il "viceversa" e' tuttavia FALSO, per spazi metrici generali.

Tuttavia, per $\mathbf{R}^n$ euclideo, vale il seguente fondamentale risultato:

Teorema 5. *(Teorema di COMPLETEZZA PER* $\mathbf{R}^n$ *EUCLIDEO)*

Una successione $(x_n)_{n\in\mathbf{N}}$ *ad elementi in* $\mathbf{R}^n$ *e' convergente se e solo se e' di Cauchy.*

In generale, si dice che uno spazio metrico e' *completo* se e solo se ogni successione di Cauchy e' convergente (questo non avviene sempre; ad esempio, nello spazio $\mathbf{Q}$ dei numeri razionali, munito (per restrizione) della metrica euclidea, questa affermazone risulta FALSA.

La nozione di successione convergente permette di fornire una utile caratterizzazione alternativa della continuita' di una funzione in un punto.

Proposizione 5. *Sia data una funzione* $f : A \subseteq \mathbf{R}^n \to \mathbf{R}$, $A \neq \emptyset$, *e sia* $x_0 \in A$.

La funzione f *e' continua in* $x_0 \in A$ *se e solo se per ogni successione* $(x_n)_{n\in\mathbf{N}}$ *in* A *tale che*

$$\lim_{x_n\to x_0}\ x_n = x_0$$

risulta

$$\lim_{x_n\to x_0}\ f(x_n) = f(x_0).$$

3 $\mathbf{R}^n$ euclideo come spazio normato

D' ora in avanti riguarderemo sempre $\mathbf{R}^n$ anche come *spazio normato euclideo*, vale a dire muniremo $\mathbf{R}^n$ di una funzione, detta *norma euclidea*

$$|| \; || : \mathbf{R}^n \to \mathbf{R}$$

definita come segue. Dato $\underline{x} = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n$, per definizione si ha:

$$||\underline{x}|| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

La norma euclidea appena definita soddisfa appunto gli assiomi di norma, vale a dire:

1. $||\underline{x}|| \geq 0$; $||\underline{x}|| = 0$ *se e solo se* $\underline{x} = \underline{0}$.
2. $||\lambda \cdot \underline{x}|| = |\lambda| \cdot ||\underline{x}||, \;\lambda \in \mathbf{R}$.
3. $||\underline{x} + \underline{x'}|| \leq ||\underline{x}|| + ||\underline{x'}||$.

3.1 Norme e metriche

In generale, assegnata una funzione norma $|| \; ||$, la funzione di due variabili definita ponendo $d(x, x') = ||x - y||$ risulta essere una metrica.

Nel caso da noi considerato di $\mathbf{R}^n$ euclideo, la situazione e' particolarmente semplice ed intuitiva.

Dati $\underline{x} = (x_1, x_2, \dots, x_n)$, $\underline{x'} = (x'_1, x_2, \dots, x_n) \in \mathbf{R}^n$, $|| \; ||$ la norma euclidea, si ha

$$||\underline{x} - \underline{x'}|| = \sqrt{(x_1 - x'_1)^2 + \cdots + (x_n - x'_n)^2} = d(\underline{x}, \underline{x'}),$$

che e' appunto la metrica euclidea. In breve, si dice che la metrica canonicamemte associata alla norma euclidea e' la metrica euclidea.

4 $\mathbf{R}^n$ euclideo come spazio con prodotto interno

D' ora in avanti riguarderemo sempre $\mathbf{R}^n$ anche come *spazio con prodotto interno euclideo*, vale a dire muniremo $\mathbf{R}^n$ di una funzione, detta *prodotto interno euclideo*

$$< \; >: \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$$

definita come segue. Dati $\underline{x} = (x_1, x_2, \dots, x_n)$, $\underline{x'} = (x'_1, x_2, \dots, x_n) \in \mathbf{R}^n$, per definizione si ha:

$$< \underline{x}, \underline{x'} >= x_1 x'_1 + \cdots + x_n x'_n = \sum_{i=1}^{n} \; x_i x'_i.$$

Il prodotto interno euclideo appena definito soddisfa appunto gli assiomi di prodotto interno, vale a dire:

1. $$< \underline{x} + \underline{x}', \underline{y} > = < \underline{x}, \underline{y} > + < \underline{x}', \underline{y} >,$$
   e
   $$< \underline{x}, \underline{y} + \underline{y}' > = < \underline{x}, \underline{y} > + < \underline{x}, \underline{y}' >,$$
2. $$< \lambda \cdot x, y >= < x, \lambda \cdot y > = \lambda \cdot < \underline{x}, \underline{y} >, \quad \lambda \in \mathbf{R}$$
3. $$< \underline{x}, \underline{y} > = < \underline{y}, \underline{x} > .$$
4. $$< \underline{x}, \underline{x} > \geq 0; \; < \underline{x}, \underline{x} > \neq 0 \; \textit{per ogni} \; x \neq \underline{0}.$$

Le proprieta' 1) e 2) si esprimono sinteticamente dicendo che il prodotto interno e' *bilineare*, la proprieta' 3) dicendo che e' *simmetrico*, la proprieta' 4) dicendo che e' *definito positivo*.

4.1 Interpretazione geometrica: prodotti interni ed angoli

Il prodotto interno euclideo in $\mathbf{R}^n$ permette di dare una definizione trasparente della nozione di *angolo* tra due vettori non nulli (elementi di $\mathbf{R}^n$ diversi dal vettore nullo $\underline{0} \in \mathbf{R}^n$).

Il punto di partenza e' il seguente risultato.

Teorema 6. *(Disuguaglianza di Cauchy-Schwarz)*
Siano x, y vettori in $\mathbf{R}^n$. Allora:

$$| < x, y > | \; \leq \sqrt{< x, x >} \, \sqrt{< y, y >} \; = \; ||x|| \cdot ||y||.$$

Geometricamente, possiamo interpretare la disuguaglianza di Cauchy-Schwarz come segue.

Siano x, y vettori in $\mathbf{R}^n$ diversi dal vettore nullo: allora la precedente disuguaglianza puo' essere riscritta nella forma

$$-1 \; \leq \; \frac{< x, y >}{||x|| \cdot ||y||} \; \leq 1.$$

Percio' esiste *unico* un numero reale $\theta \in [0, \pi]$ tale che

$$\cos(\theta) \; = \; \frac{< x, y >}{||x|| \cdot ||y||}.$$

Il valore θ si dice *angolo* tra i vettori non nulli $x, y \in \mathbf{R}^n$.

ESEMPIO in $\mathbf{R}^2$ Siano $x = (x_1, 0), y = (0, y_2)$ vettori in $\mathbf{R}^2$, $x_1, y_2 \neq 0$. Allora

$$< x, y > = \; 0 \; = \; \cos(\frac{\pi}{2}),$$

cioe' i vettori x e y sono *ortogonali*.

4.2 Prodotti interni, norme e metriche

In generale, assegnata una funzione prodotto interno $< \, , \, >$, la funzione di una variabile definita ponendo $||x|| = \sqrt{< x, x >}$ risulta essere una norma.

Nel caso da noi considerato di $\mathbf{R}^n$ euclideo, la situazione e' particolarmente semplice ed intuitiva.

Dato $\underline{x} = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$, $< \, , \, >$ il prodotto interno euclideo, si ha

$$\sqrt{< \underline{x}, \underline{x} >} = \sqrt{x_1^2 + \cdots + x_n^2} = ||\underline{x}||,$$

che e' appunto la norma euclidea. In breve, si dice che la norma canonicamemte associata al prodotto interno euclideo e' la norma euclidea. D'altra parte, ricordiamo che la metrica canonicamemte associata alla norma euclidea e' la metrica euclidea; in conclusione, $\mathbf{R}^n$ munito del prodotto interno euclideo da' luogo, canonicamente, ad $\mathbf{R}^n$ euclideo come spazio normato e ad $\mathbf{R}^n$ euclideo come spazio metrico.

Percio' tutti i concetti definiti per $\mathbf{R}^n$ come spazio metrico euclideo (ad es. *insiemi limitati, insiemi aperti, punti di accumulazione, insiemi chiusi, funzioni continue, successioni convergenti e di Cauchy, completezza*) mantengono invariato il loro significato quando riguardiamo $\mathbf{R}^n$ come spazio con prodotto interno euclideo.

Part III

Funzioni differenziabili a valori reali, Formula di Taylor ed ottimizzazione libera

5 Applicazioni $f : A \ (A \subset \mathbf{R}^n) \to \mathbf{R}$ differenziabili

5.1 Alcune osservazioni preliminari

In questa sezione, estenderemo le nostre conoscenze sul calcolo differenziale per funzioni di *una variabile reale DERIVABILI* al caso di funzioni in n variabili, $n \geq 1$.

Le principali variazioni possono essere sintetizzate come segue:

- Nel caso $n \geq 1$, NON HA PIU' SENSO parlare di di funzioni derivabili in termini "assoluti": come vedremo, anche con esempi elementari e geometrici, nel caso di piu' variabili, il concetto di *derivabilita'* DEVE essere SEMPRE associato alla scelta di una *direzione*, o, *versore*.

  Nel caso di funzioni di una variabile reale, questo fenomeno non appare poiche' in $\mathbf{R}$ abbiamo solo DUE direzioni, espresse dai vettori 1 e -1 e, di conseguenza (essendo 1 e -1 semplicemente l'uno l'opposto dell'altro) le due possibili definizioni di derivabilita' (e, di *derivata*) di fatto coincidono, essendo la stessa a meno di un segno.

- Il caso di funzioni in n variabili che ammettano derivate secondo *tutte le direzioni* NON fornisce il "vero analogo" della nozione di funzione derivabile in UNA variabile: ad esempio (FATTO FONDAMENTALE), la esistenza anche di tutte le derivate secondo ogni direzione in un punto *NON implica* la *continuita'* in tale punto.

- Diviene percio' necessario introdurre un nuovo concetto che risulti il "vero" analogo del concetto di derivabilita' nel caso di funzioni di piu' variabili: questo concetto e' appunto il concetto di di funzione *DIFFERENZIABILE*, che costituira' il tema principale del presente corso.

- Nel caso $n = 1$, la tradizionale nozione di derivabilita' e la nozione di differenziabilta' *coincidono*: questo spiega la ragione per cui, al livello di funzioni di una variabile, NON SIA NECESSARIO introdurre il concetto di differenziabilita'.

- Per comprendere la teoria delle funzioni differenziabili in piu' variabili, NON e' necessario conoscere la teoria delle funzioni derivabili in una variabile: come vedremo, tutti i risultati della teoria delle funzioni derivabili in una variabile possono essere riottenuti da quelli della teoria delle funzioni differenziabili semplicemente ponendo il numero n delle variabili uguale ad 1.

5.2 Derivate direzionali e derivate parziali

Un vettore $v \in \mathbf{R}^n$ si dice *direzione, (o, versore)* se e solo se e' di norma euclidea uguale ad 1: in simboli, $||v|| = 1$.

Se $\underline{x} \in \mathbf{R}^n$ e v e' una direzione, l'insieme

$$r_{\underline{x},v} = \{x \in \mathbf{R}^n; x = \underline{x} + tv,\ t \in \mathbf{R}\}$$

e' una RETTA, in particolare la (*unica*) retta passante per il punto $\underline{x}$ ed avente direzione v.

Sia A un aperto di $\mathbf{R}^n$, $\underline{x}$ A, e si condideri una funzione

$$f : A \to \mathbf{R}.$$

Diremo che f e' *derivabile in* $\underline{x} \in A$ *secondo la direzione* v se e solo se

$$\textit{ESISTE FINITO il limite} \lim_{t \to 0 \in \mathbf{R}} \frac{f(\underline{x} + tv) - f(\underline{x})}{t}.$$

Questo limite, se esiste ed e' finito, si indica con il simbolo

$$\frac{\partial\ f}{\partial\ v}\ (\underline{x}),$$

e si dice *derivata direzionale di* f *nel punto* $\underline{x}$ *secondo la direzione* v.

Se v e' un vettore della "base canonica" $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$

(ove $\mathbf{e}_i$ indica il vettore di $\mathbf{R}^n$ che ha tutte componenti 0 salvo la $i-$esima che vale 1, per $i = 1, \dots, n$),

la derivata direzionale rispetto ad $\mathbf{e}_i$ si dice $i-$esima *DERIVATA PARZIALE* e si denota con uno dei due simboli

$$\frac{\partial\ f}{\partial\ x_i}\ (\underline{x}), \quad D_i(f)(\underline{x}).$$

OSSERVAZIONE FONDAMENTALE. In generale, se $n > 1$, le derivate direzionali (se esistono) sono *INFINITE*. Invece le derivate parziali (se esistono) sono in numero *FINITO* (al massimo n), cioe' la dimensione dello spazio dominio della funzione, o, equivalentemente, il numero delle variabili.

Sia ora

$$\underline{x} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) \in \mathbf{R}^n,$$

per cui

$$\underline{x} + t\mathbf{e}_i = (\underline{x}_1, \dots, \underline{x}_i + t, \dots, \underline{x}_n).$$

Fissato $i = 1, \dots, n$, si ha quindi

$$\frac{\partial\ f}{\partial\ x_i}\ (\underline{x})\ = \lim_{t \to 0 \in \mathbf{R}}\ \frac{f(\underline{x} + t\mathbf{e}_i) - f(\underline{x})}{t} =$$

$$\lim_{t \to 0 \in \mathbf{R}}\ \frac{f(\underline{x}_1, \dots, \underline{x}_i + t, \dots, \underline{x}_n) - f(\underline{x}_1, \dots, \underline{x}_i, \dots, \underline{x}_n)}{t}.$$

Percio' la derivata parziale $i-$esima $\frac{\partial\ f}{\partial\ x_i}\ (\underline{x})$ (nel punto $\underline{x}$) di f rispetto ad x_i si ottiene considerando le variabili $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ come *COSTANTI*, e derivando "in modo tradizionale" rispetto alla variabile x_i.

5.3 Funzioni differenziabili

Una funzione $f : A \to \mathbf{R}$, A aperto di $\mathbf{R}^n$ si dice *DIFFERENZIABILE* nel punto $\underline{x} \in A$ se e solo se esiste un funzionale lineare

$$L_{\underline{x}} : \mathbf{R}^n \to \mathbf{R}$$

(dipendente dal punto $\underline{x}$) tale che

$$\lim_{h \to \underline{0} \in \mathbf{R}^n}\ \frac{f(\underline{x} + h) - f(\underline{x}) - L_{\underline{x}}(h)}{||h||}\ =\ 0.$$

Il funzionale lineare $L_{\underline{x}}$ si dice *differenziale* di f nel punto $\underline{x}$, e si denota anche con il simbolo $df(\underline{x})$.

Se f e' differenziabile in ogni punto di A, diremo che f e' differenziabile in A.

5.4 Interpretazione "geometrica" della differenziabilita'

Sia $f : A \to \mathbf{R}$ differenziabile in $\underline{x} \in A$, e sia $L_{\underline{x}}$ il suo differenziale in $\underline{x}$.

Consideriamo il numeratore della frazione che compare nel limite precedente, cioe'

$$E_{\underline{x}}(h) \;=\; f(\underline{x}+h) - f(\underline{x}) - L_{\underline{x}}(h),$$

riguardato come funzione dell' incremento vettoriale $h \in \mathbf{R}^n$.

Il valore $E_{\underline{x}}(h)$ puo' essere interpretato come l' *errore* (anche esso funzione di h) che si commette, nel punto $\underline{x}+h$, "approssimando" la funzione f con la funzione

$$f(\underline{x}) + L_{\underline{x}},$$

che e' somma di una funzione costante $f(\underline{x})$ e di una funzione lineare $L_{\underline{x}}$.

D'altra parte, il denominatore $||h|| = ||(\underline{x}+h) - \underline{x}||$ fornisce la distanza del punto $\underline{x}+h$ dal punto $\underline{x}$.

Percio', la condizione di differenziabilita' di f nel punto $\underline{x}$ puo' essere riscritta come segue:

$$\textit{esiste} \quad L_{\underline{x}} : \mathbf{R}^n \to \mathbf{R} \quad \textit{lineare}$$

ed

$$E_{\underline{x}} \quad \textit{funzione di } h \textit{ in un intorno di } \underline{x}$$

tali che

$$f(\underline{x}+h) = (f(\underline{x}) + L_{\underline{x}}(h)) + E_{\underline{x}}(h)$$

con

$$\lim_{h \to \underline{0} \in \mathbf{R}^n} \frac{E_{\underline{x}}(h)}{||h||} = 0.$$

A PAROLE: in prossimita' di $\underline{x}$ e' possibile approssimare f con il polinomio $f(\underline{x}) + L_{\underline{x}}$ commettendo un errore $E_{\underline{x}}$ (funzione di $||h|| = d(\underline{x}+h, \underline{x})$) che, per $h \to \underline{0}$, "tende a zero piu' velocemente" della norma di h.

5.5 Differenziabilita' e derivabilita' direzionale

Proposizione 6. *Sia A un aperto di $\mathbf{R}^n$, e sia $f : A \to \mathbf{R}$ differenziabile in $\underline{x} \in A$. Allora f ha, in $\underline{x}$, tutte le derivate direzionali $\frac{\partial f}{\partial v}(\underline{x})$, per ogni versore v. Inoltre*

$$\frac{\partial f}{\partial v}(\underline{x}) \;=\; L_{\underline{x}}(v),$$

per ogni versore v.

DIMOSTRAZIONE. Sia v un qualsiasi versore in $\mathbf{R}^n$. *Specializzando* la condizione di differenziabilita' al caso $h = tv$, otteniamo

$$\lim_{t \to 0 \in \mathbf{R}} \frac{f(\underline{x}+tv) - f(\underline{x}) - L_{\underline{x}}(tv)}{|t|} \;=\; 0.$$

Questa condizione e' EQUIVALENTE alla condizione

$$\lim_{t\to 0\in\mathbf{R}} \frac{f(\underline{x}+tv)-f(\underline{x})-L_{\underline{x}}(tv)}{t} \;=\; 0.$$

(PERCHE'? Verificare usando la definizione di limite)

Per linearita' di $L_{\underline{x}}$, l'ultima condizione e' equivalente alla condizione

$$\lim_{t\to 0\in\mathbf{R}} \frac{f(\underline{x}+tv)-f(\underline{x})-t\cdot L_{\underline{x}}(v)}{t} \;=\; 0.$$

Per proprieta' elementari dei limiti, ne segue

$$\lim_{t\to 0\in\mathbf{R}} \frac{f(\underline{x}+tv)-f(\underline{x})}{t} \;=\; \lim_{t\to 0\in\mathbf{R}} \frac{t\cdot L_{\underline{x}(v)}}{t} \;=\; L_{\underline{x}(v)};$$

percio'

$$\lim_{t\to 0\in\mathbf{R}} \frac{f(\underline{x}+tv)-f(\underline{x})}{t}$$

esiste finito ed e' uguale alla valutazione $L_{\underline{x}}(v)$ del differenziale $L_{\underline{x}}$ sul versore v. In conclusione, non solo esiste la derivata di f nel punto $\underline{x}$ secondo la direzione v, ma, di piu', abbiamo

$$\frac{\partial\, f}{\partial\, v}\;(\underline{x}) \;=\; L_{\underline{x}}(v).$$

5.6 Valutazioni del differenziale, vettore gradiente e prodotti interni

Sia A un aperto di $\mathbf{R}^n$, e sia $f : A \to \mathbf{R}$ differenziabile in $\underline{x} \in A$, $L_{\underline{x}}$ il differenziale di f nel punto $\underline{x} \in A$.

Da quanto appena dimostrato, sappiamo che f ha tutte le derivate direzionali in $\underline{x} \in A$ e quindi, a fortiori, tutte le n derivate parziali.

Si dice *gradiente* di f in $\underline{x}$ il vettore

$$grad\; f(\underline{x}) \;=\; (\;\frac{\partial\, f}{\partial\, x_1}\;(\underline{x}), \frac{\partial\, f}{\partial\, x_2}\;(\underline{x}), \ldots, \frac{\partial\, f}{\partial\, x_n}\;(\underline{x})\;) \in \mathbf{R}^n.$$

Sia $v \;=\; (v_1, v_2, \ldots, v_n) \;=\; \sum_{i=1}^{n}\; v_i \cdot \mathbf{e}_i$ un vettore in $\mathbf{R}^n$.

(ricordiamo: $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ denota la base canonica di $\mathbf{R}^n$)

Per linearita', il differenziale di f nel punto $\underline{x} \in A$ valutato su $v \in \mathbf{R}^n$ risulta

$$L_{\underline{x}}(v) = \sum_{i=1}^{n}\; v_i \cdot L_{\underline{x}}(\mathbf{e}_i) = \sum_{i=1}^{n}\; v_i \cdot\; \frac{\partial\, f}{\partial\, x_i}\;(\underline{x}).$$

In sintesi, la valutazione del differenziale $L_{\underline{x}}$ su ogni vettore $v \in \mathbf{R}^n$ puo' essere espressa, in modo algebrico assai semplice, come prodotto interno

$$< gradf(\underline{x}), v > = \sum_{i=1}^{n}\; v_i \cdot\; \frac{\partial\, f}{\partial\, x_i}\;(\underline{x}) \;\;=\; L_{\underline{x}}(v).$$

Come immediata conseguenza, abbiamo

Corollario 5. *Sia A un aperto di $\mathbf{R}^n$, e sia $f : A \to \mathbf{R}$ differenziabile in $\underline{x} \in A$. Allora*

$$\frac{\partial\ f}{\partial\ v}\ (\underline{x})\ \ = < gradf(\underline{x}), v >,$$

per ogni versore $v \in \mathbf{R}^n$.

5.7 Derivabilita' direzionale e differenziabilita'

Se $n > 1$, il "viceversa" del precedente risultato e' FALSO.

Esplicitamente, una funzione f puo' avere tutte le derivate direzionali in un punto $\underline{x}$ MA NON ESSERE differenziabile in tale punto!

Percio', la condizione di differenziabilta' e' *strettamente* piu' forte di quella di avere tutte le derivate direzionali in tale punto. (Un modo semplice di capire e ricordare questo fatto fondamentale e' il seguente: la differenziabilita' in $\underline{x}$ IMPLICA la continuita' in tale punto - come vedremo tra poco -, mentre una funzione puo' avere tutte le derivate direzionali in $\underline{x}$ ma NON essere continua in tale punto.)

5.8 Differenziabilta' e continuita'

Proposizione 7. *Sia A un aperto di $\mathbf{R}^n$, e sia $f : A \to \mathbf{R}$ differenziabile in $\underline{x} \in A$. Allora f e' continua in $\underline{x}$.*

DIMOSTRAZIONE. Poiche' f e' differenziabile in $\underline{x}$, ricordiamo che si ha

$$f(\underline{x} + h) = f(\underline{x}) + L_{\underline{x}}(h) + E_{\underline{x}}(h)$$

con

$$\lim_{h \to \underline{0} \in \mathbf{R}^n} \frac{E_{\underline{x}}(h)}{||h||} = 0.$$

Ora

$$f(\underline{x} + h) - f(\underline{x}) = L_{\underline{x}}(h) + E_{\underline{x}}(h) =< grad\ f(\underline{x}), h > + E_{\underline{x}}(h),$$

da cui

$$|f(\underline{x} + h) - f(\underline{x})| \leq |< grad\ f(\underline{x}), h > | + |E_{\underline{x}}(h)|.$$

Dalla disuguaglianza di Cauchy-Schwarz, abbiamo

$$|< grad\ f(\underline{x}), h > | \leq ||grad\ f(\underline{x})|| \cdot ||h||;$$

ne segue

$$|f(\underline{x} + h) - f(\underline{x})| \leq ||grad\ f(\underline{x})|| \cdot ||h|| + |E_{\underline{x}}(h)|.$$

Ovviamente $0 \leq |f(\underline{x} + h) - f(\underline{x})|$ ed *e' maggiorata dalla somma di due funzioni che tendono entrambe a* 0 per $h \to \underline{0} \in \mathbf{R}^n$. Quindi $|f(\underline{x}+h) - f(\underline{x})| \to 0$ per $h \to \underline{0} \in \mathbf{R}^n$, cioe' f e' continua in $\underline{x}$.

Il caso di $||grad\ f(\underline{x})|| \cdot ||h||$ e' ovvio, per definizione. In quanto a $|E_{\underline{x}}(h)|$, si ragiona come segue:

$$\lim_{h \to \underline{0} \in \mathbf{R}^n} \frac{E_{\underline{x}}(h)}{||h||} = 0\ \Rightarrow\ \lim_{h \to \underline{0} \in \mathbf{R}^n} E_{\underline{x}}(h) = 0\ \Leftrightarrow\ \lim_{h \to \underline{0} \in \mathbf{R}^n} |E_{\underline{x}}(h)| = 0.$$

5.9 Il Teorema del Differenziale Totale

Ci interessa fornire una condizione sufficente alla differenziabilita' di una funzione in un punto.

Teorema 7. *Sia A un aperto di $\mathbf{R}^n$, $\underline{x} \in A$ e sia $f : A \to \mathbf{R}$.*

Supponiamo che f abbia tutte le derivate parziali in un intorno del punto $\underline{x}$ e che queste siano continue in $\underline{x}$.

Allora f e' differenziabile in $\underline{x}$.

5.10 Operazioni "legittime" tra funzioni differenziabili in un punto

Proposizione 8. *Sia A un aperto di $\mathbf{R}^n$, e siano $f, g : A \to \mathbf{R}$ differenziabili in $\underline{x} \in A$. Allora :*

- *$f + g$ e' differenziabile in $\underline{x}$. Inoltre*

$$d(f+g)(\underline{x}) \ = \ d(f)(\underline{x}) + d(g)(\underline{x}).$$

- *$f \cdot g$ e' differenziabile in $\underline{x}$. Inoltre*

$$d(f \cdot g)(\underline{x}) \ = \ d(f)(\underline{x}) \cdot g(\underline{x}) + f(\underline{x}) \cdot d(g)(\underline{x}).$$

- *Per ogni $\lambda \in \mathbf{R}$, λf e' differenziabile in $\underline{x}$. Inoltre*

$$d(\lambda f)(\underline{x}) \ = \lambda \cdot d(f)(\underline{x}).$$

5.11 Come si "scrivono in modo intrinseco" i differenziali?

Dato lo spazio vettoriale $\mathbf{R}^n$, il suo *spazio duale* e', per definizione, l'insieme

$$(\mathbf{R}^n)^* = \{\varphi : \mathbf{R}^n \to \mathbf{R}; \quad \varphi \quad funzionale \quad lineare\},$$

munito delle operazioni di somma:

$$(\varphi_1 + \varphi_2)(v) = \varphi_1(v) + \varphi_2(v), \quad \forall v \in \mathbf{R}^n,$$

e di moltiplicazione esterna per scalare:

$$(\lambda\varphi_1)(v) = \lambda\varphi(v), \quad \forall v \in \mathbf{R}^n, \quad \forall \lambda \in \mathbf{R}.$$

La struttura che risulta e' chiaramente uno spazio vettoriale.

I differenziali sono funzionali lineari, e, quindi, elementi dello spazio duale $(\mathbf{R}^n)^*$; di conseguenza, se determiniamo una "base canonica" per lo spazio duale $(\mathbf{R}^n)^*$ avremo una "scrittura canonica" per i differenziali.

5.12 "Base canonica" dello spazio duale $(\mathbf{R}^n)^*$

Indicata con $\{\underline{e}_1, \ldots, \underline{e}_n\}$ la base canonica di $\mathbf{R}^n$, consideriamo l'insieme di funzionali lineari

$$\{dx_1, \ldots, dx_n\}$$

cosi' definiti:

$$dx_i : \mathbf{R}^n \to \mathbf{R},$$

ove

$$dx_i(\underline{e}_i) = 1 \quad e \quad dx_i(\underline{e}_j) = 0 \ se \ i \neq j.$$

OSSERVAZIONE FONDAMENTALE.

Sia $v = (v_1, \ldots, v_n) = \sum_{i=1}^n \ v_i \underline{e}_i \in \mathbf{R}^n$.

Allora

$$dx_i(v) = v_i;$$

per questa ragione i funzionali dx_i vengono spesso detti *funzionali coordinata*.

Teorema 8. *L'insieme*

$$\{dx_1, \ldots, dx_n\}$$

e' base per lo spazio duale $(\mathbf{R}^n)^*$.

In particolare, lo spazio duale $(\mathbf{R}^n)^*$ *e' uno spazio vettoriale di dimensione finita* n.

DIMOSTRAZIONE. Dobbiamo provare che $\{dx_1, \ldots, dx_n\}$ e' un sistema di generatori e che e' un insieme linearmente indipendente nello spazio vettoriale $(\mathbf{R}^n)^*$.

sistema di generatori)

Sia $\varphi \in (\mathbf{R}^n)^*$, e sia $v = (v_1, \ldots, v_n) = \sum_{i=1}^n \ v_i \underline{e}_i \in \mathbf{R}^n$. Allora

$$\varphi(v) = \varphi(\sum_{i=1}^n \ v_i \underline{e}_i) = \sum_{i=1}^n \ v_i \varphi(\underline{e}_i) = \sum_{i=1}^n \ \varphi(\underline{e}_i) dx_i(v), \quad \forall v \in \mathbf{R}^n.$$

Questa famiglia (infinita) di identita' tra valutazioni implica la seguente identita' nello spazio duale $(\mathbf{R}^n)^*$:

$$\varphi = \sum_{i=1}^n \ \varphi(\underline{e}_i) \ dx_i.$$

lineare indipendenza)

Dobbiamo provare che la condizione

$$\sum_{i=1}^n \ c_i \ dx_i \equiv 0$$

implica

$$c_1 = c_2 = \ldots = c_n = 0.$$

Ora

$$\sum_{i=1}^{n} c_i \ dx_i(\underline{e}_1) = c_1 = 0,$$

$$\sum_{i=1}^{n} c_i \ dx_i(\underline{e}_2) = c_2 = 0,$$

....

$$\sum_{i=1}^{n} c_i \ dx_i(\underline{e}_n) = c_n = 0,$$

da cui la tesi.

Corollario 6. *Sia $f : A \to \mathbf{R}$, A aperto di $\mathbf{R}^n$, $\underline{x} \in A$, e sia f differenziabile in $\underline{x}$. Denotiamo con $L_{\underline{x}}$ il differenziale di f in $\underline{x} \in A$.*

Allora il differenziale si scrive (in modo unico nello spazio duale $(\mathbf{R}^n)^*$) come segue:

$$L_{\underline{x}} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underline{x}) dx_i \in (\mathbf{R}^n)^*.$$

DIMOSTRAZIONE. Segue immediatamente dalla equivalenza tra le identita' (puntuali, sulle valutazioni):

$$L_{\underline{x}}(v) = < grad\ f(\underline{x}), v > = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underline{x}) v_i \ = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underline{x}) dx_i(v), \quad \forall v \in \mathbf{R}^n$$

e la identita' (vettoriale, tra funzioni):

$$L_{\underline{x}} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underline{x}) dx_i \in (\mathbf{R}^n)^*.$$

A PAROLE, il differenziale $L_{\underline{x}}$ si scrive, in modo unico, *come combinazione lineare delle funzioni coordinata dx_i con coefficienti le derivate parziali* $\frac{\partial f}{\partial x_i}(\underline{x})$.

5.13 Derivate successive (o, miste)

A titolo di esempio, consideriamo la funzione

$$f : \mathbf{R}^2 \to \mathbf{R}, \quad f(x, y) = xy + |y|.$$

Si noti che la derivata parziale $\frac{\partial f}{\partial x} = y$ e' definita su tutti i punti del dominio $\mathbf{R}^2$, e' ovunque continua (come funzione delle variabili x e y) ed e' *derivabile* rispetto alla variabile y.

Piu' precisamente, si ha:

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = 1$$

in ogni punto del dominio $\mathbf{R}^2$.

TUTTAVIA, la derivata parziale $\frac{\partial}{\partial y}$ rispetto alla variabile y NON ESISTE nei punti per cui $y = 0$ e, qundi, non puo' esistere la derivata successiva $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$.

In conclusione e', in generale, *FALSO* che sussista l'identita'

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

A PAROLE, non e' vero (in generale) che, "scambiando l'ordine delle derivate parziali" (successivamente applicate), si ottenga il medesimo risultato.

Addirittura, come abbiamo appena visto, puo' accadere che una derivata successiva esista (nel caso specifico, $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$), mentre, in certi punti, l'altra non esista (nel caso specifico, $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$, per $y = 0$).

Tuttavia, sotto opportune ipotesi (strutturalmennte analoghe a quelle del Teorema del Differenziale Totale), sussiste il seguente fondamentale risultato:

Teorema 9. *(di Schwartz, o, sulla scambiabilita' dell'ordine di derivazione parziale)*
Siano A un aperto di $\mathbf{R}^n$, $f : A \to \mathbf{R}$, $\underline{x} = (\underline{x}_1, \ldots, \underline{x}_n) \in A$.
Se, per $i, j = 1, 2; \ldots, , n$, le derivate successive

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right), \quad \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$$

esistono in un INTORNO di $\underline{x}$ e sono continue in $\underline{x}$, allora le loro valutazioni nel punto $\underline{x}$ coincidono. In simboli:

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(\underline{x}) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\underline{x}).$$

5.14 Derivazione di funzioni composte

Siano A un aperto di $\mathbf{R}^n$, $f : A \to \mathbf{R}$. Sia $r : I \to \mathbf{R}^n$, $I \subset \mathbf{R}$ un intervallo aperto,

$$r(t) = (r_1(t), r_2(t), \ldots, r_n(t))$$

tale che l'immagine di r

$$r[I] = \{(r_1(t), r_2(t), \ldots, r_n(t)) \in \mathbf{R}^n; t \in I\}$$

sia contenuta in A.

Sia $t_0 \in I$ tale che:

- $r_i(t)$ e' derivabile in t_0, per ogni $i = 1, 2, \ldots, n$.
- f e' differenziabile in $r(t_0) \in A$.

Allora, posto

$$g(t) = f(r_1(t), r_2(t), \dots, r_n(t))$$

per ogni $t \in I$, la *funzione composta*

$$g = f \circ r :\to \mathbf{R}, \quad g : t \mapsto g(t)$$

e' DERIVABILE in t_0. Di piu', risulta:

$$g'(t_0) =< grad\ f(r(t_0)), (r_1'(t_0), r_2'(t_0), \dots, r_n'(t_0)) >= \sum_{i=1}^{n} \frac{\partial f(r(t_0))}{\partial x_i} r_i'(t_0).$$

5.15 Polinomi di Taylor di grado k

Siano A un aperto di $\mathbf{R}^n$, $f : A \to \mathbf{R}$. Sia f di classe $C^{(k+1)}$.

Fissato il punto base $\underline{x}$, ci chiediamo come costruire il $k-$esimo polinomio approssimante di Taylor $T_k(\underline{x} + h)$ (di grado k e relativo alla scelta del punto base $\underline{x}$).

Tale polinomio e' univocamente determintato, ed e' il polinomio:

$$T_k(\underline{x} + h) = \sum_{p=0}^{k} \quad \sum_{(i_1, i_2, \dots, i_n),\ i_1+i_2+\dots+i_n = p} \frac{1}{(i_1)!(i_2)! \cdots (i_n)!} \frac{\partial^{(p)}\ f(\underline{x})}{\partial x_1^{i_1} \partial x_2^{i_2} \cdots \partial x_n^{i_n}}\ h_1^{i_1} h_2^{i_2} \cdots h_n^{i_n}.$$

Proposizione 9. *Per ogni $p \leq k$, il polinomio di Taylor T_k ha in $h = 0$ (o, equivalentemente, riguardato come funzione di $\underline{x} + h$, cioe' in $\underline{x}$) le stesse derivate seccessive di ordine p di f in $\underline{x}$, fino ad ordine $p \leq k$.*

Proposizione 10. *Sia f di classe $C^{(k)}$.*

Il resto $k-$esimo $R_k(\underline{x}; h)$ soddisfa la seguente proprieta' di andamento a ZERO:

$$\frac{R_k(\underline{x}; h)}{||h||^k} \to 0 \quad per \quad h \to \underline{0} \in \mathbf{R}^n$$

6 Ottimizzazione libera: massimi e minimi relativi per una funzione $f : A \to \mathbf{R}$ sul proprio dominio (A un aperto di $\mathbf{R}^n$)

6.1 Massimi e minimi relativi per funzioni in piu' variabili

Sia funzione $f : A \to \mathbf{R}$, $A \subset \mathbf{R}^n$.

Sia $\underline{x} \in A$.

Diremo che $\underline{x}$ e' *punto di massimo relativo* per f se e solo se *esiste un intorno* $U_{\underline{x}}$ di $\underline{x}$ in $\mathbf{R}^n$ tale che risulti

$$f(x) \leq f(\underline{x}),$$

per ogni $x \in A \cap U_{\underline{x}}$.

Analogamente, diremo che $\underline{x}$ e' *punto di minimo relativo* per f se e solo se *esiste un intorno* $U_{\underline{x}}$ di $\underline{x}$ in $\mathbf{R}^n$ tale che risulti

$$f(x) \geq f(\underline{x}),$$

per ogni $x \in A \cap U_{\underline{x}}$.

6.2 Matrici Hessiane

Sia funzione $f : A \to \mathbf{R}$, A aperto di $\mathbf{R}^n$, e sia $\underline{x} \in A$. Sia f di classe $C^{(2)}$.

La matrice

$$\mathbf{H}_{f(\underline{x})} = \begin{pmatrix} \frac{\partial^{(2)} f(\underline{x})}{\partial x_1^2} & \cdots\cdots & \frac{\partial^{(2)} f(\underline{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^{(2)} f(\underline{x})}{\partial x_2 \partial x_1} & \cdots\cdots & \frac{\partial^{(2)} f(\underline{x})}{\partial x_2 \partial x_n} \\ \vdots & & \\ \frac{\partial^{(2)} f(\underline{x})}{\partial x_n \partial x_1} & \cdots\cdots & \frac{\partial^{(2)} f(\underline{x})}{\partial x_n^2} \end{pmatrix} = \left(\frac{\partial^{(2)} f(\underline{x})}{\partial x_i \partial x_j} \right)_{i,j=1,\ldots,n}$$

e' ovviamente una matrice *simmetrica* (per il teorema di Schwartz) e si dice *MATRICE HESSIANA della funzione* f *in* $\underline{x}$.

6.3 Alcune osservazioni preliminari

Siano A un aperto di $\mathbf{R}^n$, $f : A \to \mathbf{R}$. Sia $\underline{x} \in A$.

Sia $\delta > 0$ tale che $I(\underline{x}, \delta) \subseteq A$.

Fissato $\underline{v} \in \mathbf{R}^n$, $\underline{v} \neq \underline{0}$, consideriamo la funzione (della unica variabile reale t, con $|t| < \frac{\delta}{||\underline{v}||}$):

$$F(t) = f(\underline{x} + t\underline{v}) = (f \circ r)(t),$$

ove

$$r(t) :]-\frac{\delta}{||\underline{v}||}, \frac{\delta}{||\underline{v}||}[\to \mathbf{R}, \quad r(t) = (r_1(t), \ldots, r_n(t))$$

con

$$r_i(t) = \underline{x}_i + t\underline{v}_i$$

$(\underline{x}_i = (\underline{x}_1, \ldots, \underline{x}_n), \quad \underline{v} = (\underline{v}_1, \ldots, \underline{v}_n))$.

- Sia f di classe $C^{(1)}$. Allora F e' derivabile e, di piu', risulta:

$$F'(t) =< grad\ f(r(t)), (r_1'(t), r_2'(t), \ldots, r_n'(t)) >= \sum_{i=1}^{n} \frac{\partial f(\underline{x} + t\underline{v})}{\partial x_i} \underline{v}_i. \qquad (*)$$

- Sia f di classe $C^{(2)}$. Allora F ammette derivata seconda e, di piu', risulta:

$$F''(t) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x} + t\underline{v}))v_i v_j =< \underline{v} \times \mathbf{H}_{f(\underline{x}+t\underline{v})}, \underline{v} > .$$

  Infatti, essendo f di classe $C^{(2)}$, le derivate parziali $\frac{\partial f}{\partial x_j}$ sono di classe $C^{(1)}$ e, quindi, differenziabili; l'asserto si ottiene applicando il teorema sulle funzioni composte a ciascun addendo della formula $(*)$.

6.4 Condizioni necessarie

Proposizione 11. *Sia funzione $f : A \to \mathbf{R}$, A aperto di $\mathbf{R}^n$.*

Sia $\underline{x} \in A$ punto di massimo (minimo) relativo per la funzione f.

Se f e' differenziabile in $\underline{x}$, allora il differenziale di f in $\underline{x}$ e' il funzionale identicamente nullo: in simboli

$$df(\underline{x}) \equiv 0.$$

DIMOSTRAZIONE

Se il punto $\underline{x}$ e' punto di massimo (minimo) relativo per la funzione f (e la funzione e' differenziabile in $\underline{x}$), allora ammette tutte le derivate direzionali $\frac{\partial f}{\partial v}(\underline{x})$ ed inoltre deve aversi

$$\frac{\partial f}{\partial v}(\underline{x}) = 0,$$

per *ogni versore* $v \in \mathbf{R}^n$.

Poiche' $\frac{\partial f}{\partial v}(\underline{x}) = df(\underline{x})(v)$, ne segue che $df(\underline{x}) \equiv 0$.

NB I punti $\underline{x} \in A$ del dominino di una funzione $f : A \to \mathbf{R}$ per cui il differenziale risulta identicamente nullo (in simboli, $df(\underline{x}) \equiv 0$) si dicono *PUNTI CRITICI* per la funzione f.

Proposizione 12. *Sia data una funzione $f : A \to \mathbf{R}$, A aperto di $\mathbf{R}^n$. Sia f di classe $C^{(2)}$.*

Sia $\underline{x}$ punto critico per f.

Se $\underline{x}$ e' punto di minimo (massimo) relativo per f, allora:

per ogni vettore $v \in \mathbf{R}^n$ risulta

$$< v \times \mathbf{H}_{f(\underline{x})}, v > \geq 0, \quad \left(< v \times \mathbf{H}_{f(\underline{x})}, v > \leq 0\right),$$

cioe' la matrice Hessiana e' semidefinita positiva (cioe' la matrice Hessiana e' semidefinita negativa).

DIMOSTRAZIONE

Sia $\delta > 0$ tale che $I(\underline{x}, \delta) \subseteq A$, e tale che

$$f(x) \geq f(\underline{x}), \quad \forall x \in I(\underline{x}, \delta).$$

Fissato $\underline{v} \in \mathbf{R}^n$, $\underline{v} \neq \underline{0}$, la funzione (della unica variabile reale t, con $|t| < \frac{\delta}{||\underline{v}||}$):

$$F(t) = f(\underline{x} + t\underline{v})$$

e' di classe $C^{(2)}$ e, ovviamente,

$$F(t) = f(\underline{x} + t\underline{v}) \geq F(0) = f(\underline{x}), \quad \forall t \in]-\frac{\delta}{||\underline{v}||}, \frac{\delta}{||\underline{v}||}[.$$

Allora 0 e' punto di minimo per F, interno al suo dominio. Deve essere pertanto $F'(0) = 0$ e $F''(0) \geq 0$.

Per ipotesi abbiamo gia' che $F'(0) = \sum_{i=1}^{n} \frac{\partial f(\underline{x})}{\partial x_i} \underline{v}_i = df(\underline{x})(\underline{v}) = 0$.

Resta quindi la condizione

$$F''(0) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}) v_i v_j = < \underline{v} \times \mathbf{H}_{f(\underline{x})}, \underline{v} > \geq 0;$$

per la genericita' di $\underline{v}$, la tesi e' dimostrata.

6.5 Condizioni sufficienti

Teorema 10. *Sia data una funzione $f : A \to \mathbf{R}$, A aperto di $\mathbf{R}^n$. Sia f di classe $C^{(2)}$.*

Sia $\underline{x}$ punto critico per f.

Se, per ogni vettore $v \in \mathbf{R}^n, v \neq \underline{0} \in \mathbf{R}^n$ risulta

$$< v \times \mathbf{H}_{f(\underline{x})}, v > \; > \; 0, \quad \left(< v \times \mathbf{H}_{f(\underline{x})}, v > \; < \; 0\right),$$

cioe' la matrice Hessiana e' definita positiva (cioe' la matrice Hessiana e' definita negativa), allora $\underline{x}$ e' punto di minimo (massimo) relativo per f.

La dimostrazione e' omessa.

7 Esercizi su ottimizzazione libera

Sia data una funzione f di due variabili x e y; nel seguito, per brevita' di notazione denoteremo con f_x, f_y, f_{xx}, f_{xy},f_{yx}, f_{yy} le derivate $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial^{(2)} f}{\partial x^2}$, $\frac{\partial^{(2)} f}{\partial x \partial y}$, $\frac{\partial^{(2)} f}{\partial y \partial x}$, $\frac{\partial^{(2)} f}{\partial y^2}$ nel "generico" punto (x, y) del dominio (se esistono).

Dato un punto $\underline{x}$ del dominio in cui tali derivate esistano, scriveremo

$f_x(\underline{x})$, $f_y(\underline{x})$, $f_{xx}(\underline{x})$, $f_{xy}(\underline{x})$, $f_{yx}(\underline{x})$, $f_{yy}(\underline{x})$

per le *valutazioni* di dette derivate nel punto $\underline{x}$.

1)

Si consideri la funzione $f : R^2 \to R$, definita da $f(x, y) = x^3 - 3x^2 - 9x + \frac{1}{y^2+1}$.

- La funzione f e' differenziabile su R^2?
- Si scriva il gradiente di f nel punto $\alpha = (0,0)$.
- Si calcoli la derivata direzionale di f nel punto $\alpha = (0,0)$, secondo la direzione $v = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$.
- Si determinino gli eventuali punti di minimo relativo e di massimo relativo per f.

SOLUZIONE.

Le derivate parziali nel punto generico (x,y) sono

$$f_x = 3x^2 - 6x - 9, \quad f_y = \frac{-2y}{(y^2+1)^2}.$$

La funzione f e' ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale).

Si ha *grad* $f(0,0) = (-9,0)$, da cui

$$\frac{\partial f}{\partial v}(0,0) =< (-9,0), (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}) > .$$

Per determinare gli eventuali punti di estremo relativo, dobbiamo prima trovare i punti critici.

In concreto, porre $f_x = 3x^2 - 6x - 9 = 0$, da cui (due soluzioni) $x = 3, -1$ e $f_y = \frac{-2y}{(y^2+1)^2} = 0$, da cui $y = 0$.

Abbiamo percio' due punti critici:

$$\underline{x} = (3,0), \quad \underline{x'} = (-1,0).$$

Gli elementi della matrice Hessiana (nel punto generico (x,y)) sono:

$$f_{xx} = 6x - 6, \quad f_{yx} = f_{xy} = 0, \quad f_{yy} = \frac{6y^2 - 2}{(y^2+1)^3},$$

da cui, valutando nei punti critici, abbiamo

$$f_{xx}(\underline{x}) = f_{xx}((3,0)) = 12 \quad e \quad f_{xx}(\underline{x'}) = f_{xx}((-1,0)) = -12$$

e

$$f_{yy}(\underline{x}) = f_{yy}((3,0)) = f_{yy}(\underline{x'}) = f_{yy}((-1,0)) = -2.$$

Le matrici Hessiane nei punti critici sono percio':

$$\mathbf{H}_{f(\underline{x})} = \begin{pmatrix} 12 & 0 \\ 0 & -2 \end{pmatrix}$$

e

$$\mathbf{H}_{f(\underline{x}')} = \left(\begin{array}{cc} -12 & 0 \\ 0 & -2 \end{array} \right).$$

La prima matrice e' NON SEMIDEFINITA, per cui $\underline{x}$ e' punto di SELLA, mentre la seconda matrice e' DEFINITA NEGATIVA, per cui $\underline{x}'$ e' punto di MASSIMO.

2)

- Si consideri la funzione $f : R^2 \to R$, $f(x,y) = (x^2+1)(y^3-3y)$.
- (*) f e' differenziabile in ogni punto di R^2?
- (*) Si determini il gradiente di f in $\alpha = (0,0)$ e la derivata di f lungo la direzione $v = (\frac{3}{5}, -\frac{4}{5})$ nel punto α.
- (*) Quale fra i punti α, $\beta = (0,1)$ e $\gamma = (0,-1)$ e' un estremo relativo per f?

SOLUZIONE.

Le derivate parziali nel punto generico (x,y) sono

$$f_x = 2x(y^3-3y), \quad f_y = (x^2+1)(3y^2-3).$$

La funzione f e' ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale).

Si ha $grad\ f(0,0) = (0,-3)$, da cui

$$\frac{\partial f}{\partial v}(0,0) =< (0,-3), (\frac{3}{5}, -\frac{4}{5}) > .$$

Il punto $\alpha = (0,0)$ non e' critico, quindi non e' punto di estremo relativo.

I punti $\beta = (0,1)$ e $\gamma = (0,-1)$ sono critici e dobbiamo quindi esaminare le matrici Hessiane.

Gli elementi della matrice Hessiana (nel punto generico (x,y)) sono:

$$f_{xx} = 2(y^3-3y), \quad f_{yx} = f_{xy} = 2x(3y^2-3), \quad f_{yy} = 6y(x^2+1).$$

Le matrici Hessiane nei punti critici β e γ sono percio':

$$\mathbf{H}_{f(\beta)} = \left(\begin{array}{cc} -4 & 0 \\ 0 & 6 \end{array} \right)$$

e

$$\mathbf{H}_{f(\gamma)} = \left(\begin{array}{cc} 4 & 0 \\ 0 & -6 \end{array} \right).$$

Entrambe le matrici sono NON SEMIDEFINITE, per cui β e γ sono punti di SELLA.

3)

- Si consideri la funzione
$$f : R^2 \to R$$
, definita da
$$f(x, y) = (x^2 + 1)(y^3 + 3y^2 + 1).$$

- La funzione f e' differenziabile su R^2?

- Si determini il gradiente di di f in $(1, 1)$.

- Si determinino gli eventuali punti di massimo e minimo relativo per la funzione f.

SOLUZIONE.

Le derivate parziali nel punto generico (x, y) sono
$$f_x = 2x(y^3 + 3y^2 + 1), \quad f_y = (x^2 + 1)(3y^2 + 6y).$$

La funzione f e' ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale).

Si ha $grad\ f(1, 1) = (10, 18)$.

La derivata parziale f_y si annulla per $y = 0, -2$. Poche', per $y = 0, -2$, si ha $y^3+3y+1 > 0$, la condizione $grad\ f = (0, 0)$ implica $x = 0$.

I punti critici sono percio' i punti $\underline{x} = (0, 0)$ e $\underline{x}' = (0, -2)$

Gli elementi della matrice Hessiana (nel punto generico (x, y)) sono:
$$f_{xx} = 2(y^3 - 3y + 1) > 0$$
nei punti critici,
$$f_{yx} = f_{xy} = 0,$$
e
$$f_{yy} = (x^2 + 1)(6y + 6),$$
che vale 6 in $\underline{x} = (0, 0)$ e vale -6 in $\underline{x}' = (0, -2)$.

Percio' $\underline{x}$ e' di minimo, $\underline{x}'$ e' di sella.

4)

Sia
$$f : R^2 \to R, \qquad f(x, y) = x^2 + xy - y^2.$$

- f e' differenziabile in R^2?

- Si calcoli il gradiente di f nel punto $x_0 = (1, 1)$.

- Si calcoli la derivata direzionale di f in $x_0 = (1, 1)$ secondo la direzione $v = (\sqrt{2}/2, \sqrt{2}/2)$.

- La funzione f ammette punti di estremo relativo?

SOLUZIONE.

Le derivate parziali nel punto generico (x, y) sono

$$f_x = 2x + y, \quad f_y = x - 2y.$$

La funzione f e' ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale).

Si ha $grad\ f(1, 1) = (3, -1)$.

L'unico punto critico e' percio' il punto $\underline{x} = (0, 0)$.

Gli elementi della matrice Hessiana (nel punto generico (x, y)) sono:

$$f_{xx} = 2 \quad f_{yx} = f_{xy} = 1, \quad f_{yy} = -2.$$

Si noti che il determinante della matrice Hessiana vale -5; poiche' tale determinante fornisce il prodotto degli autovalori, gli autovalori hanno segno discorde e quindi la matrice e' NON SEMIDEFINITA. Percio' $\underline{x} = (0, 0)$ e' di sella.

5)

Sia

$$f : R^2 \to R, \qquad f(x, y) = 3x^3 + 3y^2 - x.$$

Si determinino i punti critici di f. Quali tra questi sono punti di estremo relativo?

SOLUZIONE.

Le derivate parziali nel punto generico (x, y) sono

$$f_x = 9x^2 - 1, \quad f_y = 6y.$$

La funzione f e' ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale).

I punti critici sono percio' i punti $\underline{x} = (\frac{1}{3}, 0)$ e $\underline{x}' = (-\frac{1}{3}, 0)$.

Gli elementi della matrice Hessiana (nel punto generico (x, y)) sono:

$$f_{xx} = 18x, \quad f_{yx} = f_{xy} = 0, \quad f_{yy} = 6.$$

Percio' $\underline{x}$ e' di minimo, $\underline{x}'$ e' di sella.

6)

Si consideri la funzione

$$f = R^2 \to R,$$

definita da

$$f(x, y) = x \sin(y).$$

- La funzione f e' differenziabile su R^2?
- Si calcoli il gradiente di f nel punto $\alpha = (\pi, \pi)$.

- Si calcoli la derivata direzionale di f nel punto α, secondo la direzione $v = (-1/2, \sqrt{3}/2)$.
- Si deteminino i punti critici di f.
- Si determinio gli eventuali punti di estremo relativo per f.

SOLUZIONE.

Le derivate parziali nel punto generico (x, y) sono

$$f_x = sin(y), \quad f_y = xcos(y).$$

La funzione f e' ovunque differenziabile, in quanto le sue derivate parziali esistono e sono continue in ogni punto del dominio (cfr., Teorema del Differenziale Totale).

I punti critici sono i punti della forma

$$(0, k\pi), \quad k \in \mathbb{Z}.$$

Gli elementi della matrice Hessiana (nel punto generico (x, y)) sono:

$$f_{xx} = 0, \quad f_{yx} = f_{xy} = cos(y), \quad f_{yy} = -xsin(y).$$

In ogni punto critico la matrice Hessiana e':

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

la quale, avendo determinante negativo, e' NON SEMIDEFINITA: percio' tutti i punti critici sono punti di sella.

Part IV

Ottimizzazione vincolata

8 Applicazioni $f : A(A \subset \mathbf{R}^r) \to \mathbf{R}^n$ differenziabili

8.1 Funzioni a valori vettoriali

Sia $A \subset \mathbf{R}^r$, $f : A \to \mathbf{R}^n$. Nel seguito, per semplicita', supporremo sempre che A sia un *aperto* di $\mathbf{R}^r$.

Dato un vettore $\underline{x} = (x_1, \ldots, x_r) \in A$, la sua *immagine mediante* f (ovvero, la "valutazione di f in $\underline{x}$") e' un vettore $f(\underline{x})$ in $\mathbf{R}^n$; percio', rispetto alla base canonica di $\mathbf{R}^n$, $f(\underline{x})$ si scrive come $n-$pla

$$f(\underline{x}) = (f_1(\underline{x}), f_2(\underline{x}), \ldots, f_n(\underline{x}));$$

gli scalari $f_i(\underline{x})$ si diranno *coordinate* di $f(\underline{x})$.

Di conseguenza, la funzione $f : A \to \mathbf{R}^n$ *individua ed e' individuata* da n funzioni

$$f_1, f_2, \ldots, f_n$$

a valori in $\mathbf{R}$; queste funzioni $f_1, f_2, \ldots, f_n$ si dicono *componenti scalari della funzione f a valori vettoriali.*

Usualmente, identificheremo una funzione $f : A \to \mathbf{R}^n$, $A \subset \mathbf{R}^r$, con la $n-$pla delle sue componenti scalari, e scriveremo

$$f \equiv (f_1, f_2, \ldots, , f_n).$$

8.2 Funzioni differenziabili a valori vettoriali

Sia $f : A \to \mathbf{R}^n$, A aperto di $\mathbf{R}^r$.

Si dice che f e' *differenziabile nel punto* x_0 se e solo se esiste un operatore lineare

$$L_{x_0} : \mathbf{R}^r \to \mathbf{R}^n$$

(*dipendente da* x_0) tale che risulti

$$\lim_{h \to \underline{0} \in \mathbf{R}^r} \frac{f(x_0 + h) - f(x_0) - L_{x_0}(h)}{||h||} = \underline{0} \in \mathbf{R}^n,$$

ove, come sempre, $||h||$ indica la norma euclidea del vettore $h \in \mathbf{R}^r$.

L'operatore $h \mapsto L_{x_0}(h)$ si dice ancora *differenziale* di f in x_0 e si denota brevemente anche con il simbolo $df(x_0)$, mentre le sue "valutazioni" si indicano con $df(x_0)(h), h \in \mathbf{R}^r$.

Dalla definizione, segue pressoche' immediatamente che *la funzione a valori vettoriali $f \equiv (f_1, f_2, \ldots, , f_n)$ e' differenziabile in x_0 se e solo se le sue componenti scalari $f_1, f_2, \ldots, , f_n$ sono tutte differenziabili in x_0.*

Inoltre, per ogni $j = 1, 2, \ldots, n$, la $j-$esima componente scalare $L_{x_0,j}$ dell'operatore L_{x_0} coincide con il differenziale della $j-$esima componente scalare f_j della funzione f.

Piu' esplicitamente, per ogni $h = (h_1, h_2, \ldots, h_r) \in \mathbf{R}^r$, si ha

$$L_{x_0,j}(h) = \sum_{k=1}^{r} \frac{\partial f_j(x_0)}{\partial x_k} \cdot h_k =< grad\ f_j(x_0), h > .$$

Ne segue che la matrice dell'operatore lineare L_{x_0} (rispetto alle basi canoniche di $\mathbf{R}^r$ e $\mathbf{R}^n$, rispettivamente) e' la matrice le cui righe sono i *gradienti* delle componenti scalari $f_1, f_2, \ldots, , f_n$; esplicitamente, abbiamo quindi la matrice

$$\mathbf{J}_{f(x_0)} = \begin{pmatrix} \frac{\partial f_1(x_0)}{\partial x_1} & \ldots\ldots & \frac{\partial f_1(x_0)}{\partial x_r} \\ \frac{\partial f_2(x_0)}{\partial x_1} & \ldots\ldots & \frac{\partial f_2(x_0)}{\partial x_r} \\ \vdots & & \\ \frac{\partial f_n(x_0)}{\partial x_1} & \ldots\ldots & \frac{\partial f_n(x_0)}{\partial x_r} \end{pmatrix}.$$

Questa matrice si chiama *MATRICE JACOBIANA della funzione f in x_0.*

La condizione di differenziabilita' si potra' allora scrivere anche come segue:

$$\lim_{h \to \underline{0} \in \mathbf{R}^r} \frac{f(x_0 + h) - f(x_0) - \mathbf{J}_{f(x_0)} \times h^t}{||h||} = \underline{0} \in \mathbf{R}^n.$$

Utilizzando gli analoghi risultati per funzioni a valori scalari, si deduce immediatamente:

Proposizione 13. *Se f e' differenziabile in $x_0 \in A$, allora f e' continua in x_0; se f e' differenziabile in A, allora f e' continua in A. Se $f \in C_A^{(1)}$, allora f e' differenziabile in A.*

N.B. L'ultima affermazione e' la versione "multidimensionale" del *Teorema del Differenziale Totale.*

Teorema 11. *(sulla composizione di funzioni differenziabili)*

Sia $A \subset \mathbf{R}^r$, $B \subset \mathbf{R}^n$, a punto interno ad A, b punto interno a B.

Siano $g : A \to B$, $g(a) = b$, $f : B \to \mathbf{R}^p$, g differenziabile in a, f differenziabile in b.

Allora la funzione composta

$$f \circ g : A \to \mathbf{R}^p$$

e' differenziabile in a.

Inoltre

$$d(f \circ g)(a) = df(b) \circ dg(b) \qquad (*).$$

DIMOSTRAZIONE.

Siano date $g \equiv (g_1, \ldots, g_n) : A \to B \subseteq \mathbb{R}^n$ ed $f \equiv (f_1, \ldots, f_p) : B \to \mathbb{R}^p$. Sia $a \in A$ e $b = g(a) \in B$, con g differenziabile in a ed f differenziabile in $b = g(a)$.

Allora la funzione composta $f \circ g$ e' differenziabile in $a \in A$.

La matrice Jacobiana $J_{(f \circ g)(a)}$ e' una matrice $p \times r$: Dati $i = 1, \ldots, p$ e $k = 1, \ldots, r$, vogliamo provare che l'elemento di posto (i, k) in $J_{(f \circ g)(a)}$:

$$\frac{\partial (f \circ g)_i(a)}{\partial x_k}$$

esiste e vogliamo calcolarlo in termini delle matrici Jacobiane

$$J_{f(g(a))}, \quad J_{g(a)}.$$

Notiamo, in primo luogo, che la componente scalare $i-$esima $(f \circ g)_i$ coincide con la funzione composta $(f_i \circ g)$.

Inoltre la derivata parziale

$$\frac{\partial (f \circ g)_i(a)}{\partial x_k} = \frac{\partial (f_i \circ g)(a)}{\partial x_k}$$

puo' essere pensata come la derivata ordinaria della funzione $f_i \circ g$ *riguardata come funzione della singola variabile* x_k, la quale, a sua volta, e' la funzione composta della funzione f_i con la funzione $g \equiv (g_1, \ldots, g_n)$ *riguardata come funzione della singola variabile* x_k.

Ora, essendo g differenziabile in a, le derivate (parziali) di $g_1, \ldots, g_n$ rispetto alla variabile x_k esistono in a e, di conseguenza - essendo f_i differenziabile in $g(a)$ - la derivata (parziale)

$$\frac{\partial (f_i \circ g)(a)}{\partial x_k}$$

esiste in a; di piu', si ha:

$$\frac{\partial (f_i \circ g)(a)}{\partial x_k} = \sum_{j=1}^{n} \frac{\partial f_i(g(a))}{\partial x_j} \frac{\partial g_j(a)}{\partial x_k}.$$

Pertanto l'elemento

$$\frac{\partial (f \circ g)_i(a)}{\partial x_k} = \frac{\partial (f_i \circ g)(a)}{\partial x_k} \in J_{(f \circ g)(a)}$$

coincide con il prodotto tra la $i-$esima riga di $J_{f(g(a))}$ e la $k-$esima colonna di $J_{g(a)}$.

N.B. L'identita' $(*)$ deve leggersi: *l'operatore lineare* $d(f \circ g)(a) : \mathbf{R}^r \to \mathbf{R}^p$ *si ottiene COMPONENDO gli operatori lineari* $dg(a) : \mathbf{R}^r \to \mathbf{R}^n$ *e* $df(b) : \mathbf{R}^r n \to \mathbf{R}^p$.

In termini di matrici Jacobiane, l'identita' $(*)$ assume una forma assai semplice:

Corollario 7.

$$\mathbf{J}_{(f \circ g)(a)} = \mathbf{J}_{f(b)} \times \mathbf{J}_{g(a)}.$$

Esempio 1. *Consideriamo le funzioni (differenziabili, in virtu' delle osservazioni precedenti - sono funzioni polinomiali e, quindi di classe* $C^{(\infty)}$*)*

$$g : \mathbf{R}^2 \to \mathbf{R}^3, \quad g(x_1, x_2) = (x_1^2 - x_2, x_1 + 1, x_1 - x_2 + 2),$$

e

$$f : \mathbf{R}^3 \to \mathbf{R}^3, \quad f(x_1, x_2, x_3) = (x_2, x_3 - 1, x_1 x_2).$$

Sia

$$F = f \circ g : \mathbf{R}^2 \to \mathbf{R}^3;$$

risulta

$$F(x_1, x_2) = (g_2(x_1, x_2), g_3(x_1, x_2) - 1, g_1(x_1, x_2) g_2(x_1, x_2)) = (x_1 + 1, x_1 - x_2 + 1, x_1^3 + x_1^2 - x_1 x_2 - x_2).$$

Posto $a = (1, 2) \in \mathbf{R}^2$*, si ha* $b = g(a) = (-1, 2, 1)$*.*

Per un "generico" punto $(x_1, x_2) \in \mathbf{R}^2$*, la matrice Jacobiana di* g *risulta essere*

$$\mathbf{J}_{g(x_1, x_2))} = \begin{pmatrix} 2x_1 & -1 \\ 1 & 0 \\ 1 & -1 \end{pmatrix},$$

per cui, valutando la matrice Jacobiana nel punto $a = (1, 2)$, *risulta*

$$\mathbf{J}_{g(1,2))} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ 1 & -1 \end{pmatrix}.$$

Per $(x_1, x_2, x_3) \in \mathbf{R}^3$, *risulta*

$$\mathbf{J}_{f(x_1,x_2,x_3))} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ x_2 & x_1 & 0 \end{pmatrix},$$

per cui, valutando la matrice Jacobiana nel punto $b = (-1, 2, 1)$, *risulta*

$$\mathbf{J}_{f(-1,2,1))} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}.$$

Per $(x_1, x_2) \in \mathbf{R}^2$, *risulta*

$$\mathbf{J}_{F(x_1,x_2)} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 3x_1^2 + 2x_1 - x_2 & -x_1 - 1 \end{pmatrix},$$

per cui, valutando la matrice Jacobiana nel punto $a = (1, 2)$, *risulta*

$$\mathbf{J}_{F(1,2)} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 3 & -2 \end{pmatrix}.$$

Come asserito dal Teorema, si ha:

$$\mathbf{J}_{f(b)} \times \mathbf{J}_{g(a)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix} \times \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 3 & -2 \end{pmatrix} = \mathbf{J}_{F(1,2)}.$$

9 Curve in $\mathbf{R}^n$

Si dice *curva* in $\mathbf{R}^n$ una applicazione continua φ da un intervallo chiuso $J = [a, b] \subset \mathbf{R}$ ad $\mathbf{R}^n$.

Posto $\varphi \equiv (\varphi_1, \ldots, \varphi_n)$, $\varphi_i : [a, b] \to \mathbf{R}$, la curva φ si dice di classe $C^{(k)}$ se e solo se le sue componenti scalari φ_i sono di classe $C^{(k)}$, per ogni $i = 1, \ldots, n$.

Data una curva $\varphi : [a, b] \to \mathbf{R}$, l'insieme $\varphi([a, b]) = \{\varphi(t); t \in [a, b]\} \subset \mathbf{R}^n$ si dice *sostegno* della curva φ.

Una curva $\varphi : [a, b] \to \mathbf{R}$ si dice *regolare* se e solo se:

- L'applicazione φ e' di classe $C^{(1)}$.

- $\varphi'(t) \equiv (\varphi_1'(t), \ldots, \varphi_n'(t))$ e' diverso dal vettore nullo di $\mathbf{R}^n$, per ogni $t \in]a,b[$ ed inoltre, se $\varphi(a) = \varphi(b)$ allora $(\varphi_1'(a), \ldots, \varphi_n'(a)) = (\varphi_1'(b), \ldots, \varphi_n'(b)) \neq \underline{0}$.

Una curva $\varphi : [a,b] \to \mathbf{R}^n$ si dice *semplice aperta* se e solo se definisce un *omeomorfismo* da $[a,b]$ a $\varphi([a,b]) = \{\varphi(t); t \in [a,b]\} \subset \mathbf{R}^n$.

OSSERVAZIONE. Si ha il seguente risultato generale per funzioni continue tra spazi metrici: *sia $f : X \to Y$ continua e biiettiva, X spazio metrico compatto. Allora f e' un omeomorfismo.* Poiche' $[a,b]$ e' un compatto, la curva continua $\varphi : [a,b] \to \mathbf{R}^n$ e' semplice aperta se e solo se e' INIETTIVA.

9.1 Vettori tangenti

Nel seguito, per semplicita', ci limiteremo a considerare curve regolari semplici aperte $\varphi : [a,b] \to \mathbf{R}^n$. Data tale curva e dato $t_0 \in]a,b[$ si dice *VERSORE TANGENTE* alla curva nel punto $\varphi(t_0)$ il versore

$$T(t_0) = (\varphi_1'(t_0)^2 + \cdots + \varphi_n{}'(t_0)^2)^{-\frac{1}{2}} (\varphi_1'(t_0), \ldots, \varphi_n'(t_0)).$$

Consistentemente, la retta

$$\{\varphi(t_0) + \lambda T(t_0); \lambda \in \mathbf{R}\}$$

si dice "retta tangente" al sostegno della curva φ nel punto $\varphi(t_0)$, ed ogni vettore (non nullo) proporzionale a $T(t_0)$ si dice "vettore tangente" alla curva nel punto $\varphi(t_0)$.

10 Varieta' in $\mathbf{R}^n$

10.1 Matrici Jacobiane

Sia F una funzione

$$F \equiv (F_1, \ldots, F_{n-r}) : I(\underline{\alpha}, \delta) \to \mathbf{R}^{n-r},$$

ove $I(\underline{\alpha}, \delta) \subset \mathbf{R}^n$ e' un intorno sferico aperto del punto $\underline{\alpha} \in \mathbf{R}^n$.

Supponiamo che F sia di classe $C^{(k)}$, $k \geq 1$ (cioe' F_i di classe $C^{(k)}$, $k \geq 1$, per ogni $i = 1, \ldots, n-r$).

Ricordiamo che la *MATRICE JACOBIANA* di F nel punto $\underline{\alpha}$ e' la matrice $(n-r) \times n$ cosi' definita:

$$\mathbf{J}_{F(\underline{\alpha})} = \begin{pmatrix} \frac{\partial F_1(\underline{\alpha})}{\partial x_1} & \cdots\cdots & \frac{\partial F_1(\underline{\alpha})}{\partial x_n} \\ \frac{\partial F_2(\underline{\alpha})}{\partial x_1} & \cdots\cdots & \frac{\partial F_2(\underline{\alpha})}{\partial x_n} \\ \vdots & & \\ \frac{\partial F_{n-r}(\underline{\alpha})}{\partial x_1} & \cdots & \frac{\partial F_{n-r}(\underline{\alpha})}{\partial x_n} \end{pmatrix}.$$

10.2 Definizione di varieta'. Punti regolari

Siano $1 \leq r \leq n$, $k \geq 1$, e sia V un sottoinsieme di $\mathbf{R}^n$.

Si dice che V e' una *VARIETA' di* $\mathbf{R}^n$ *di DIMENSIONE* r *e CLASSE* $C^{(k)}$ se e solo se, per ogni $\underline{\alpha} \in V$, esistono un intorno sferico aperto $I(\underline{\alpha}, \delta) \subset \mathbf{R}^n$ di $\underline{\alpha}$ ed una funzione

$$F \equiv (F_1, \ldots, F_{n-r}) : I(\underline{\alpha}, \delta) \to \mathbf{R}^{n-r},$$

F di classe $C^{(k)}$, tali che:

1. La matrice Jacobiana $\mathbf{J}_{F(\underline{\alpha})}$ ha rango $n - r$.
2. $V \cap I(\underline{\alpha}, \delta) = \{x \in I(\underline{\alpha}, \delta); F(x) = \underline{0}\}$.

Brevemente, spesso si dice che $F \equiv (F_1, \ldots, F_{n-r})$ e' la funzione che fornisce ("localmente", vicino ad $\underline{\alpha}$) le "equazioni" della varieta' V.

Piu' in generale, dato un sottoinsieme $V \subset \mathbf{R}^n$, i punti $\underline{\alpha}$ soddisfacenti le precedenti condizioni sono detti *PUNTI REGOLARI*. Di conseguenza, in questo linguaggio, una varieta' e' un sottoinsieme $V \subset \mathbf{R}^n$ i cui punti risultino tutti regolari.

10.3 Curve, varieta', spazi tangenti e spazi normali

Sia V una varieta' di $\mathbf{R}^n$ di dimensione r ($r \leq n$) e classe $C^{(k)}$ ($k \geq 1$).

Un vettore $h \in \mathbf{R}^n$ si dice *TANGENTE a* V *in* $\underline{\alpha}$ se e solo se esistono un reale positivo $\delta \in \mathbf{R}^+$ ed un *curva regolare semplice aperta*

$$\varphi : [-\delta, \delta] \to \mathbf{R}^n$$

tale che

- i) $\varphi([-\delta, \delta]) \subset V$.
- *ii)* $\varphi(0) = \underline{\alpha}$ *e* $\varphi'(0) = h$.

Geometricamente, la condizione i) significa che la curva φ "descrive una traiettoria" sulla varieta' V "passante per $\underline{\alpha}$" (in corrispondenza del valore $t = 0$); la condizione ii) significa che il vettore h e' proporzionale al "versore tangente" al sostegno di φ nel punto $\varphi(0) = \underline{\alpha}$.

Si dice *SPAZIO TANGENTE* alla varieta' V nel punto $\underline{\alpha}$ l'insieme

$$\mathbf{T}(\underline{\alpha}) = \{h \in \mathbf{R}^n; h \quad tangente\ a\ V\ in\ \underline{\alpha}\} \cup \{\underline{0}\}.$$

Teorema 12. *Sia* V *e' una varieta" di* $\mathbf{R}^n$ *di dimensione* r *e classe* $C^{(k)}$, $k \geq 1$.

Siano $\underline{\alpha} \in V$ *ed* F *una funzione*

$$F \equiv (F_1, \ldots, F_{n-r}) : I(\underline{\alpha}, \delta) \to \mathbf{R}^{n-r},$$

F di classe $C^{(k)}$, tali che:

1. *La matrice Jacobiana* $\mathbf{J}_{F(\underline{\alpha})}$ *ha rango* $n-r$.

2. $V \cap I(\underline{\alpha}, \delta) = \{x \in I(\underline{\alpha}, \delta); F(x) = \underline{0}\}$.

Allora $\mathbf{T}(\underline{\alpha}) = Ker\ dF(\underline{\alpha})$, *ove* $dF(\underline{\alpha})$ *denota, al solito, il differenziale di* F *nel punto* $\underline{\alpha}$.

Corollario 8. *Lo spazio tangente* $\mathbf{T}(\underline{\alpha})$ *e' il sottospazio vettoriale di dimensione* r

$$\{h \in \mathbf{R}^n; \mathbf{J}_{F(\underline{\alpha})} \cdot h = \underline{0}\} \subset \mathbf{R}^n.$$

Poiche' la $i-$esima riga della matrice jacobiana $\mathbf{J}_{F(\underline{\alpha})}$ e' il gradiente *grad* $F_i(\underline{\alpha})$ della $i-$esima componente scalare F_i della funzione F, $i = 1, \ldots, n-r$, il precedente risultato puo' essere riformulato come segue.

Corollario 9. *Lo spazio tangente* $\mathbf{T}(\underline{\alpha})$ *e' lo spazio di tutti i vettori* $h \in \mathbf{R}^n$ *tali che*

$$< grad\ F_i(\underline{\alpha}), h >= 0, \quad per\ ogni\ i = 1, \ldots, n-r,$$

cioe' lo spazio dei vettori ortogonali ai gradienti delle $n-r$ *componenti scalari di* F, *valutati nel punto* $\underline{\alpha}$.

Sia V e' una varieta" di $\mathbf{R}^n$ di dimensione r e classe $C^{(k)}$, $k \geq 1$.

Si dice *SPAZIO NORMALE* alla varieta' V nel punto $\underline{\alpha}$ il *complemento ortogonale* (in $\mathbf{R}^n$) dello spazio tangente $\mathbf{T}(\underline{\alpha})$, cioe' il sottospazio vettoriale

$$\mathbf{N}(\underline{\alpha}) = \{u \in \mathbf{R}^n; < u, h >= 0 \quad per\ ogni\ h \in \mathbf{T}(\underline{\alpha})\} \subset \mathbf{R}^n.$$

Osservazione 1. *Se* V *ha dimensione* r, *anche lo spazio tangente* $\mathbf{T}(\underline{\alpha})$ *ha dimensione* r *(come sottospazio vettoriale di* $\mathbf{R}^n$*): di conseguenza, lo spazio normale* $\mathbf{N}(\underline{\alpha})$ *e' un sottospazio vettoriale di dimensione* $n-r$.

Ora , dal Corollario 2, sappiamo che gli $n-r$ *vettori*

$$grad\ F_i(\underline{\alpha}), \quad i = 1, \ldots, n-r$$

appartengono allo spazio normale $\mathbf{N}(\underline{\alpha})$.

Inoltre, poiche' la matrice Jacobiana $\mathbf{J}_{F(\underline{\alpha})}$ *ha rango* $n-r$, *questi vettori sono linearmente indipendenti.*

Ne segue allora immediatamente:

Corollario 10. *L' insieme*

$$\{grad\ F_i(\underline{\alpha}); \quad i = 1, \ldots, n-r\}$$

e' base dello spazio normale $\mathbf{N}(\underline{\alpha})$.

11 Punti critici vincolati. Moltiplicatori di Lagrange

Sia V e' una varieta" di $\mathbf{R}^n$ di dimensione r e classe $C^{(k)}$, $k \geq 1$.

Sia A un aperto di $\mathbf{R}^n$, $V \subset A$ e si consideri una funzione $\Phi : A \to \mathbf{R}$ di classe $C^{(1)}$ su A.

Un massimo o minimo relativo della restrizione Φ/V di Φ a V si dice *ESTREMO RELATIVO VINCOLATO* di Φ rispetto al "vincolo" V.

Teorema 13. *(dei moltiplicatori di Lagrange)*

Se Φ ha estremo relativo vincolato in $\underline{\alpha} \in V$ (rispetto al vincolo V), allora esistono degli scalari

$$\lambda_1, \lambda_2, \ldots, \lambda_{n-r}$$

tali che $\underline{\alpha}$ sia punto critico per la funzione

$$\Phi - \sum_{j=1}^{n-r} \lambda_j f_j,$$

ove

$$f \equiv (f_1, \ldots, f_{n-r}) : I(\underline{\alpha}, \delta) \to \mathbf{R}^{n-r}$$

e' la funzione che fornisce ("localmente", vicino ad $\underline{\alpha}$) le "equazioni" della varieta' V.

DIMOSTRAZIONE. Sia h un vettore tangente a V in $\underline{\alpha}$ e sia $\varphi : [-\delta, \delta] \to \mathbf{R}^n$ una curva regolare semplice aperta tale che

- i) $\varphi([-\delta, \delta]) \subset V$.
- *ii)* $\varphi(0) = \underline{\alpha}$ *e* $\varphi'(0) = h$.

Posto $\omega(t) = \Phi(\varphi)(t)$, poiche' $\varphi(t) \in V$ e Φ/V ha estremo relativo in $\underline{\alpha}$, ne segue che ω ha estremo relativo in $t = 0$, e quindi deve aversi $\omega'(0) = 0$.

Risulta, d'altronde,

$$\omega'(0) = \sum_{j=1}^{n} \frac{\partial \Phi(\varphi(0))}{\partial x_j} \cdot \varphi_j'(0) = < grad\ \Phi(\varphi(0)), \varphi'(0) > = < grad\ \Phi(\underline{\alpha}), h > = 0.$$

(Si noti che quest' ultima relazione vale *per ogni* vettore tangente h a V in $\underline{\alpha}$.)

Percio' *grad* $\Phi(\underline{\alpha})$ *appartiene allo spazio normale a* V *in* $\underline{\alpha}$ e, poiche' questo e' generato dai vettori

$$grad\ f_1(\underline{\alpha}), grad\ f_2(\underline{\alpha}), \ldots, grad\ f_{n-r}(\underline{\alpha}),$$

esistono degli scalari

$$\lambda_1, \lambda_2, \ldots, \lambda_{n-r}$$

tali che

$$grad\ \Phi(\underline{\alpha}) = \sum_{j=1}^{n-r} \lambda_j\ grad\ f_j(\underline{\alpha}).$$

Percio' *grad* $(\Phi - \sum_{j=1}^{n-r} \lambda_j\ f_j)(\underline{\alpha}) = 0$, ossia $\underline{\alpha}$ e' punto critico per $\Phi - \sum_{j=1}^{n-r} \lambda_j\ f_j$.

12 Esempi/Applicazioni

1) Sia $\Phi : \mathbf{R}^2 \to \mathbf{R}$, $\Phi(x_1, x_2) = x_1 + x_2$ e sia

$$V = \{(x_1, x_2) \in \mathbf{R}^2; f(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0\}.$$

L'insieme V e' una varieta'di $\mathbf{R}^2$ di dimensione 1 (tutti i suoi punti sono *regolari, VERIFICARE*).

I punti critici di V sono tutti e soli i punti per i quali *grad* $\Phi(x_1, x_2) = (1, 1)$ e' proporzionale a *grad* $f = (2x_1, 2x_2)$; percio' deve aversi $x_1 = x_2$ e quindi, dovendo tali punti appartenere a V si hanno due casi:

$$(x_1, x_2) = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}),\ (-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}).$$

Essendo V un compatto, in virtu' del teorema di Weierstrass, questi due punti saranno uno punto di minimo (assoluto rispetto al vincolo V) e l'altro punto di massimo (assoluto rispetto al vincolo V).

2) Sia $\Phi : \mathbf{R}^3 \to \mathbf{R}$, $\Phi(x_1, x_2, x_3) = x_1 + x_2 + x_3^2$ e sia

$$V = \{(x_1, x_2, x_3) \in \mathbf{R}^3; f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1 = 0\}.$$

L'insieme V e' una varieta'di $\mathbf{R}^2$ di dimensione 1 (tutti i suoi punti sono *regolari, VERIFICARE*).

I punti critici di V sono tutti e soli i punti per i quali *grad* $\Phi(x_1, x_2, x_3) = (1, 1, 2x_3)$ e' proporzionale a *grad* $f(2x_1, 2x_2, x_3) = (2x_1, 2x_2, 2x_3)$; percio' deve risultare

$$1 = \lambda 2x_1, \quad 1 = \lambda 2x_2, \quad 2x_3 = \lambda 2x_3;$$

Se $x_3 \neq 0$ ne consegue che $\lambda = 1$ e, quindi, $x_1 = \frac{1}{2}$, $x_2 = \frac{1}{2}$; dovendo i punti appartenere a V, deve essere $x_3 = \pm\frac{1}{\sqrt{2}}$.

Si hanno due punti critici

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

con moltiplicatore $\lambda = 1$.

Se $x_3 = 0$, deve aversi $x_1^2 + x_2^2 = 1$, ma anche $x_1 = x_2$; percio' $x_1 = x_2 = \pm\frac{1}{\sqrt{2}}$.

Si hanno altri due punti critici

$$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), \quad (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)$$

con moltiplicatori $\lambda = \frac{\sqrt{2}}{2}$ e $\lambda = -\frac{\sqrt{2}}{2}$, rispettivamente.

Essendo V un compatto, in virtu' del teorema di Weierstrass, la funzione Φ ammette punti di minimo (assoluto rispetto al vincolo V) e punti di massimo (assoluto rispetto al vincolo V).

Andando a valutare la funzione nei quattro punti critici trovati, si verifica immediatamente che il minimo assoluto si ottiene nel punto

$$(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0),$$

mentre il massimo assoluto si ottiene nei punti

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

3) Sia $\Phi : \mathbf{R}^3 \to \mathbf{R}$, $\Phi(x_1, x_2, x_3) = x_1 + x_2 + x_3$ e sia

$$f \equiv (f_1, f_2) : \mathbf{R}^3 \to \mathbf{R}^2,\ f_1(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 4,\ f_2(x_1, x_2, x_3) = x_1 - 1.$$

Sia

$$V = \{(x_1, x_2, x_3) \in \mathbf{R}^3; f(x_1, x_2, x_3) = 0\}.$$

L'insieme V e' una varieta'di $\mathbf{R}^3$ di dimensione 1 (tutti i suoi punti sono *regolari, VERIFICARE*).

Si ha *grad* $\Phi(x_1, x_2, x_3) = (1, 1, 1)$, *grad* $f_1 = (2x_1, 2x_2, 2x_3)$, *grad* $f_2 = (1, 0, 0)$ e dobbiamo determinare i punti di V tali che esistano $\lambda_1, \lambda_2 \in \mathbf{R}$ tali che

$$(1, 1, 1) = \lambda_1(2x_1, 2x_2, 2x_3) + \lambda_2(1, 0, 0).$$

Segue immediatamente che $x_2 = x_3$; poiche' il punto deve appartenere a V, si ha inoltre $x_1 = 1$ e quindi $x_2 = x_3 = \pm\frac{\sqrt{3}}{2}$. Inoltre, valutando le prime coordinate, si ha $\lambda_2 = 1 - 2\lambda_1$.

In conclusione, i punti critici sono $(1, \frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2})$ e $(1, -\frac{\sqrt{3}}{2}, -\frac{\sqrt{3}}{2})$ con moltiplicatori

$$\lambda_1 = \frac{1}{\sqrt{3}},\ \lambda_2 = 1 - \frac{2}{\sqrt{3}}$$

e

$$\lambda_1 = -\frac{1}{\sqrt{3}},\ \lambda_2 = 1 + \frac{2}{\sqrt{3}},$$

rispettivamente.

Essendo V un compatto, in virtu' del teorema di Weierstrass, questi due punti saranno uno punto di minimo (assoluto rispetto al vincolo V) e l'altro punto di massimo (assoluto rispetto al vincolo V).

4) Sia $\Phi : \mathbf{R}^2 \to \mathbf{R}$, $\Phi(x_1, x_2) = x_1 x_2$ e sia

$$V = \{(x_1, x_2) \in \mathbf{R}^2; f(x_1, x_2) = x_1 + 2x_2 - 3 = 0\}.$$

Risulta *grad* $\Phi(x_1, x_2) = (x_2, x_1)$ e *grad* $f(x_1, x_2) = (1, 2)$; ne segue

$$x_2 = \lambda, \quad x_1 = 2\lambda.$$

I punti critici sono percio' della forma $(x_1, x_2) = (2\lambda, \lambda)$; dovendo appartenere alla varieta' V, si ha che $2\lambda + 2\lambda - 3 = 0$ e, quindi, $\lambda = \frac{3}{4}$.

In conclusione, l'unico punto critico e' il punto $(\frac{3}{2}, \frac{3}{4})$, con moltiplicatore $\lambda = \frac{3}{4}$.

5) Sia $\Phi : \mathbf{R}^3 \to \mathbf{R}$, $\Phi(x_1, x_2, x_3) = x_1 x_2 + x_3^2$ e sia

$$f \equiv (f_1, f_2) : \mathbf{R}^3 \to \mathbf{R}^2, \ f_1(x_1, x_2, x_3) = x_1 + x_2 - 1, \ f_2(x_1, x_2, x_3) = x_2^2 + x_3^2 - 1.$$

Sia

$$V = \{(x_1, x_2, x_3) \in \mathbf{R}^3 ; f(x_1, x_2, x_3) = 0\}.$$

L'insieme V e' una varieta'di $\mathbf{R}^3$ di dimensione 1 (tutti i suoi punti sono *regolari, VERIFICARE*).

Si ha *grad* $\Phi(x_1, x_2, x_3) = (x_2, x_1, 2x_3)$, *grad* $f_1 = (1, 1, 0)$, *grad* $f_2 = (0, 2x_2, 2x_3)$ e dobbiamo determinare i punti di V tali che esistano $\lambda_1, \lambda_2 \in \mathbf{R}$ tali che

$$(x_2, x_1, 2x_3) = \lambda_1(1, 1, 0) + \lambda_2(0, 2x_2, 2x_3).$$

Percio' $\lambda_1 = x_2$ e, se $x_3 \neq 0$, $\lambda_2 = 1$, da cui $x_1 = \lambda_1 + 2x_2 = 3x_2$; dalla condizione $f_1(x_1, x_2, x_3) = 0$ evinciamo che $x_2 = \frac{1}{4}$ e $x_1 = \frac{3}{4}$.

Dalla condizione $f_1(x_1, x_2, x_3) = 0$ segue allora $x_3 = \pm\frac{\sqrt{15}}{4}$.

In conclusione, vi sono due punti critici con $x_3 \neq 0$, precisamente

$$(\frac{3}{4}, \frac{1}{4}, \frac{\sqrt{15}}{4}) \quad (\frac{3}{4}, \frac{1}{4}, -\frac{\sqrt{15}}{4}).$$

Se $x_3 = 0$, allora $x_2 = \pm 1$. Otteniamo allora altri due punti critici, precisamente

$$(2, -1, 0)$$

con moltiplicatori $\lambda_1 = -1$ e $\lambda_2 = -\frac{3}{2}$ e

$$(0, 1, 0)$$

con moltiplicatori $\lambda_1 = 1$ e $\lambda_2 = -\frac{1}{2}$.

Essendo V un compatto, in virtu' del teorema di Weierstrass, la funzione Φ ammette punti di minimo (assoluto rispetto al vincolo V) e punti di massimo (assoluto rispetto al vincolo V).

Andando a valutare la funzione nei quattro punti critici trovati, si verifica immediatamente che il minimo assoluto si ottiene nel punto

$$(2, -1, 0),$$

mentre il massimo assoluto si ottiene nei punti

$$(\frac{3}{4}, \frac{1}{1}, \frac{\sqrt{15}}{4}), \quad (\frac{3}{4}, \frac{1}{1}, -\frac{\sqrt{15}}{4}).$$

6) Sia $\Phi : \mathbf{R}^3 \to \mathbf{R}$, $\Phi(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ e sia

$$f \equiv (f_1, f_2) : \mathbf{R}^3 \to \mathbf{R}^2, \ f_1(x_1, x_2, x_3) = x_1^2 + x_2 - 1, \ f_2(x_1, x_2, x_3) = x_2 + x_3 - 1.$$

Sia

$$V = \{(x_1, x_2, x_3) \in \mathbf{R}^3; f(x_1, x_2, x_3) = 0\}.$$

L'insieme V e' una varieta'di $\mathbf{R}^3$ di dimensione 1 (tutti i suoi punti sono *regolari, VERIFICARE*).

Si ha *grad* $\Phi(x_1, x_2, x_3) = (2x_1, 2x_2, 2x_3)$, *grad* $f_1 = (2x_1, 1, 0)$, *grad* $f_2 = (0, 1, 1)$ e dobbiamo determinare i punti di V tali che esistano $\lambda_1, \lambda_2 \in \mathbf{R}$ tali che

$$(2x_1, 2x_2, 2x_3) = \lambda_1(2x_1, 1, 0) + \lambda_2(0, 1, 1).$$

Se $x_1 \neq 0$, si ha chiaramente $\lambda_1 = 1$. Inoltre

$$2x_2 = \lambda_1 + \lambda_2, \quad 2x_3 = \lambda_2;$$

percio'

$$2x_2 - 2x_3 = 1, \quad x_2 + x_3 = 1,$$

da cui $x_3 = \frac{1}{4}$ e $x_2 = \frac{3}{4}$.

Poiche' il punto deve appartenere a V, ne consegue che $x_1 = \pm\frac{1}{2}$.

In conclusione, abbiamo due punti critici

$$(\frac{1}{2}, \frac{3}{4}, \frac{1}{4}), \quad (-\frac{1}{2}, \frac{3}{4}, \frac{1}{4}),$$

con moltiplicatori

$$\lambda_1 = 1, \quad \lambda_2 = \frac{1}{2}.$$

Se $x_1 = 0$, si ha $x_2 = 1$ e $x_3 = 0$. Percio' abbiamo un terzo punto critico

$$(0, 1, 0)$$

con moltiplicatori

$$\lambda_1 = 1, \quad \lambda_2 = 0.$$

Andrea Brini

Dipartimento di Matematica, Universita di Bologna

40126 Bologna, Italy

E-mail: brini@dm.unibo.it