
CAPITOLO 2

CALCOLO DIFFERENZIALE IN PIÙ VARIABILI

Un esercizio preliminare per familiarizzarsi con insiemi definiti da disuguaglianze. *Disegnare* i seguenti insiemi di $\mathbb{R}^2$:

$$\begin{aligned} &\{(x, y) : |x| + y^2 \geq 1\}; \quad \{(x, y) : |y| < \log(1 + |x|)\}; \quad \{(x, y) : y^2 - x^2 \geq 4\}; \\ &\{(x, y) : |y| + |x| \leq 2\}; \quad \{(x, y) : x^4 + y^4 \leq 2\}; \quad \{(x, y) : |x - 1| + |x| + |y + 1| \leq 1\}; \\ &\{(x, y) : |y - x^2| \geq 1\}; \quad \{(x, y) : |y + x| - |x^2 - y^2| \leq 2\}; \quad \{(x, y) : \sin x \geq \cos y\}; \\ &\{(x, y) : |y + \sin x| \leq 5\}. \end{aligned}$$

2-1 LIMITI E CONTINUITÀ

Richiamiamo alcuni concetti già noti.

$\mathbb{R}^n$ è l'insieme delle n -uple ordinate di numeri reali $x = (x_1, \dots, x_n)$; i suoi elementi si dicono punti o vettori. È uno spazio vettoriale (di dimensione n) con le operazioni naturali di somma componente per componente, e prodotto per uno scalare (cioè un numero reale) componente per componente. La *base canonica* di $\mathbb{R}^n$ è la base formata dai vettori $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, $\dots$, $e_n = (0, \dots, 0, 1)$. $\mathbb{R}^n$ è uno spazio di Banach con la norma euclidea $|\cdot|$ (corrispondente alla metrica euclidea $d(x, y) = |x - y|$). Un intorno (sferico) di un punto è una palla centrata in quel punto. Il prodotto scalare tra due vettori x, y è dato da $(x, y) = x_1 y_1 + \dots + x_n y_n$. Valgono le formule $(x, x) = |x|^2 \geq 0$; $|(x, y)| \leq |x| \cdot |y|$ (disuguaglianza di Cauchy–Schwartz).

Una successione in $\mathbb{R}^n$ è un'applicazione da $\mathbb{N}$ in $\mathbb{R}^n$; normalmente una successione si indica con x_k , ma le successioni di $\mathbb{R}^n$ presentano il problema dei doppi indici. Quando c'è rischio di confusione, useremo la notazione $\{x^{(k)}\}_{k \geq 1}$. Fare attenzione: $x^{(k)}$ è un elemento della successione, che scritto per esteso diventa $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})$. $x_j^{(k)}$, $j = 1, \dots, n$ si dicono le *componenti* della successione, e sono a loro volta n successioni di numeri reali.

Una successione $x^{(k)}$ converge a x (si scrive: $x^{(k)} \rightarrow x$) se $\forall \epsilon \exists k_\epsilon$ tale che $|x^{(k)} - x| < \epsilon$ per $k \geq k_\epsilon$. Una successione converge se e solo se converge ciascuna delle sue componenti: questa è una conseguenza immediata dell'equivalenza della norma euclidea $|x|$ con la norma

2 Calcolo differenziale

del $\max |x|_\infty$ vista in precedenza. Una (sotto)successione estratta da $x^{(k)}$ è una successione $x^{(\phi(k))}$, dove $\phi : \mathbb{N} \rightarrow \mathbb{N}$ è una funzione strettamente crescente.

Un aperto di $\mathbb{R}^n$ è un insieme tale che se $x \in A$ allora esiste una palla centrata in x contenuta in A ; un chiuso è il complementare di un aperto, ossia un insieme C tale che se una successione di punti di C converge, allora il limite appartiene a C . Un insieme è limitato se è contenuto in una palla. Un compatto è un insieme chiuso e limitato. Un punto x è aderente ad un insieme A se esiste una successione di punti di A convergente ad x ; è un punto di accumulazione di A se esiste una successione di punti di A , *distinti da x* , convergente ad x .

Sia $A \subseteq \mathbb{R}^n$, $x^{(0)} \in \mathbb{R}^n$ un punto di accumulazione per A . Data $f : A \rightarrow \mathbb{R}$, si dice che $f \rightarrow L$ (f tende a L , o ha limite L) per $x \rightarrow x^{(0)}$ (per x che tende a $x^{(0)}$) se $\forall \epsilon \exists \delta$ tale che $0 < |x - x^{(0)}| < \delta$ implica $|f(x) - L| < \epsilon$. Notare che l'eventuale valore di f in $x^{(0)}$ non interviene in questa definizione. Sia ora $x_0 \in A$ un punto di accumulazione per A . f è continua in x_0 se il limite di f per $x \rightarrow x_0$ esiste ed è uguale a $f(x_0)$. Una funzione continua su un compatto di $\mathbb{R}^n$ a valori in $\mathbb{R}$ ammette massimo e minimo.

Una funzione su $\mathbb{R}^n$ è spesso definita tramite una combinazione di funzioni elementari delle variabili $x_1, \dots, x_n$. Può accadere che tale espressione non sia definita per tutti gli $x = (x_1, \dots, x_n)$. In questi casi, l'*insieme di definizione* è l'insieme dei punti $x \in \mathbb{R}^n$ per i quali tutte le funzioni utilizzate nella definizione sono definite e si possono comporre come indicato. Invece l'*insieme di continuità* di una funzione $f : A \rightarrow \mathbb{R}$ è l'insieme dei punti di A nei quali f è continua.

Spesso le due variabili di $\mathbb{R}^2$ si indicano con (x, y) anziché con (x_1, x_2) , e le tre variabili di $\mathbb{R}^3$ con (x, y, z) . Diamo qui alcune funzioni su $\mathbb{R}^2$ che saranno utilizzate nel resto del capitolo come base per vari esercizi.

$$\begin{aligned} F_1(x, y) &= x^2 \log(x + y^2), & F_2 &= x \log(x/y), & F_3 &= \log(x/y) + \log(y/x), \\ F_4 &= \sin(x/y) + \cos(y/x), & F_5 &= \cos(xe^y) + x^5 y^4, & F_6 &= \arctan(x/y) \\ F_7 &= \log[\log x], & F_8 &= \exp[x - \log^2 y], & F_9 &= 1/\sin(x + y), \\ F_{10} &= \sqrt{xy^2}, & F_{11} &= \sqrt{x^2 + y^2 - 1}, & F_{12} &= \sqrt{x^2 - y^2}, \\ F_{13} &= \sqrt{x \sin y}, & F_{14} &= \sqrt{x\sqrt{y}}, & F_{15} &= \tan(x^2 - y), \\ F_{16} &= |x + y|, & F_{17} &= |x + y| + |x + 2y + 1|, & F_{18} &= \sqrt{|x - y^2| + 1}; \end{aligned}$$

$$f_0(x, y) = ax + by \quad (a, b \in \mathbb{R}); \quad f_1(x, y) = x^2 + xy + y^2; \quad f_2(x, y) = x^n y^m, \quad n, m \in \mathbb{N}$$

$$f_3(x, y) = P(x, y) \quad (\text{polinomio in due variabili})$$

$$f_4(x, y) = \begin{cases} 1 & \text{se } x = 0 \text{ oppure } y = 0 \\ 0 & \text{altrimenti} \end{cases}$$

$$f_5(x, y) = \begin{cases} \frac{x^2 - y^2}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases}$$

$$f_6(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases}$$

$$\begin{aligned}
f_7(x, y) &= \begin{cases} \frac{x^4}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_8(x, y) &= \begin{cases} \frac{x^7 - y^7 + x^6 y}{x^6 + y^6} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_9(x, y) &= \begin{cases} \frac{\sin(x^4 + 2y^4)}{x^2 + 3y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{10}(x, y) &= \begin{cases} (x^2 + y^2) \log(x^2 + y^2) & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{11}(x, y) &= \begin{cases} \frac{\sin(x - y)}{x - y} & \text{se } x \neq y \\ 1 & \text{altrimenti} \end{cases} \\
f_{12}(x, y) &= \begin{cases} \frac{e^{x^2 + y^2} - 1}{x^2 + 2y^2} & \text{se } (x, y) \neq (0, 0) \\ 1 & \text{in } (0, 0) \end{cases} \\
f_{13}(x, y) &= \begin{cases} (x^2 + y^2 - 1) \cos \frac{1}{x^2 + y^2 - 1} & \text{se } x^2 + y^2 \neq 1 \\ 0 & \text{altrimenti} \end{cases} \\
f_{14}(x, y) &= \begin{cases} \frac{\sin x^2 \cdot \sin y^2}{2x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{15}(x, y) &= \log 1 + x1 + y \\
f_{16}(x, y) &= \begin{cases} \frac{\sin(x + y)}{|x| + |y|} & \text{se } (x, y) \neq (0, 0) \\ 1 & \text{in } (0, 0) \end{cases} \\
f_{17}(x, y) &= \begin{cases} xy \log(x^2 + y^2) & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{18}(x, y) &= \begin{cases} \frac{x^2 - y^2}{2x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 1 & \text{in } (0, 0) \end{cases} \\
f_{19}(x, y) &= \begin{cases} \frac{x^2 + y^2}{\sin^2 x + \sin^2 y} & \text{se } (x, y) \neq (0, 0) \\ 1 & \text{in } (0, 0) \end{cases} \\
f_{20}(x, y) &= \begin{cases} xy \frac{\log(x^2 + 1) + \log(y^2 + 1)}{\log^2(x^2 + y^2)} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{21}(x, y) &= \begin{cases} \frac{\sin(x^2) \cdot \sin(y^2)}{2x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{22}(x, y) &= \begin{cases} \frac{x^3 \log |y| + y^3 \log |x|}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases}
\end{aligned}$$

4 Calcolo differenziale

$$\begin{aligned}
f_{23}(x, y) &= \begin{cases} \frac{(e^{(x^2)} - 1)(e^{(y^2)} - 1)}{|x| + |y|} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{24}(x, y) &= \begin{cases} (1 - (x + y)^2) \cos \frac{1}{(x + y)^2 - 1} & \text{se } x + y \neq \pm 1 \\ 0 & \text{sulle rette } x + y = \pm 1 \end{cases} \\
f_{25}(x, y) &= \begin{cases} (x^3 - 2y^3) \frac{e^{x^2 + y^2} - 1}{(x^2 + y^2)^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{26}(x, y) &= \begin{cases} \frac{\sin(x^2 + y^2)}{x^2 + 2y^2} + \frac{\cos(x^2 + y^2) - 1}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 1 & \text{in } (0, 0) \end{cases} \\
f_{27}(x, y) &= \begin{cases} [(x - 1)^2 + (y - 1)^2] \cos \frac{1}{(x - 1)^2 + (y - 1)^2} & \text{se } (x, y) \neq (1, 1) \\ 0 & \text{in } (1, 1) \end{cases} \\
f_{28}(x, y) &= \begin{cases} \frac{\cos(x^2 + y^2) - 1}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{29}(x, y) &= |x + y - 1| + |x - 1| + |xy| \\
f_{30}(x, y) &= \begin{cases} \frac{\sin x \cdot \sin y}{|\sin x| + |\sin y|} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{31}(x, y) &= \begin{cases} x & \text{se } \sqrt{x^2 + y^2} \in \mathbb{N} \\ y & \text{se } \sqrt{x^2 + y^2} \notin \mathbb{N} \end{cases} \\
f_{32}(x, y) &= \begin{cases} \frac{|x|^{|x|} - 1}{x^2 + y^2} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases} \\
f_{33}(x, y) &= \begin{cases} \frac{e^{x^3 + 2y^3} - \cos(x^3 + 2y^3)}{2x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{34}(x, y) &= \begin{cases} \frac{x^8 + y^4 - xy^3 + 4x^5y - 7x^3}{x^4 + 5y^4 + x^2 + 2y^6} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0) \end{cases} \\
f_{35}(x, y) &= \begin{cases} x \log \frac{x^2 + 2y^2}{2x^4 + y^4} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{in } (0, 0). \end{cases}
\end{aligned}$$

1. Osservazione Per lo studio di queste funzioni sarà utile ricordare le seguenti disuguaglianze: usando la notazione $r = |(x, y)| = \sqrt{x^2 + y^2}$,

$$|x^n y^m| \leq r^{n+m}; \quad |x|^m + |y|^m \geq cr^m$$

per una opportuna costante $c > 0$ che dipende solo da m .

ESERCIZI

- 1 Una successione converge se e solo se tutte le sue sottosuccessioni convergono allo stesso limite.
- 2 $x^{(k)} \rightarrow x$ se e solo se $|x^{(k)} - x| \rightarrow 0$, e se e solo se per ogni palla di centro x la successione è contenuta definitivamente (cioè a partire da un certo indice in poi) in essa.
- 3 Diciamo che la successione $x^{(k)}$ è l'unione delle successioni $x^{\phi(k)}$ e $x^{\psi(k)}$ se l'unione delle immagini di ϕ, ψ è tutto $\mathbb{N}$. Mostrare che $x^{(k)}$ converge ad $x^{(0)}$ se e solo se $x^{\phi(k)}$ e $x^{\psi(k)}$ convergono entrambe a $x^{(0)}$.
- 4 Determinare gli insiemi di definizione delle funzioni F_1 - F_{18} .
- 5 Determinare gli insiemi di definizione delle funzioni f_0 - f_{35} .
- 6 Determinare se le funzioni f_0 - f_{35} sono continue in $(0, 0)$.
- 7 Determinare gli insiemi di continuità delle funzioni f_0 - f_{35} .
- 8 Dimostrare le disuguaglianze dell'Osservazione 1 [La prima è banale. La seconda: chiaramente si ha $|x|^m + |y|^m \geq \max\{|x|, |y|\}^m$. Ma sappiamo che per vettori di $\mathbb{R}^n$ vale $|v| \leq n|v|_\infty$, e quindi su $\mathbb{R}^2$ abbiamo $r \leq 2 \max\{|x|, |y|\}$. Mettendo tutto insieme: $|x|^m + |y|^m \geq (r/2)^m$.]
- 9 Sia $A \subseteq \mathbb{R}^n$, $x \in A$. Allora f è continua in x se e solo se per ogni successione $x^{(k)}$ di punti di A convergente a x si ha $\lim_{k \rightarrow \infty} f(x^{(k)}) = f(x)$.
- 10 Una successione converge a x se e solo se da ogni sua sottosuccessione si può estrarre una sottosuccessione convergente a x .

2-2 DERIVAZIONE

1. Definizione Sia $A \subseteq \mathbb{R}^n$ aperto, $x^0 \in A$. f si dice *derivabile* rispetto a x_j ($j = 1, \dots, n$) in x^0 se esiste finito il limite

$$\lim_{t \rightarrow 0} \frac{f(x^0 + te_j) - f(x^0)}{t}.$$

Tale limite si indica con $D_j f(x^0)$ o $\frac{\partial f}{\partial x_j}(x^0)$ e si dice *derivata parziale* di f rispetto a x_j in x^0 . La funzione si dice *derivabile* in x^0 se è derivabile rispetto a $x_1, \dots, x_n$; allora il vettore

$$Df(x^0) = (D_1 f(x^0), \dots, D_n f(x^0)) = \left(\frac{\partial f}{\partial x_1}(x^0), \dots, \frac{\partial f}{\partial x_n}(x^0) \right)$$

si chiama *gradiente* di f in x^0 . Una notazione alternativa per il gradiente, spesso usata in fisica, è

$$\nabla f(x) = Df(x).$$

Una funzione derivabile in tutti i punti di un insieme A si dice *derivabile* in A ; l'insieme di *derivabilità* di una funzione è l'insieme di tutti i punti in cui essa è derivabile. Più in generale, se v è un vettore di modulo 1, f si dice *derivabile nella direzione v* in x^0 se esiste finito il limite

$$\lim_{t \rightarrow 0^+} \frac{f(x^0 + tv) - f(x^0)}{t}.$$

Tale limite si indica con $D_v f(x^0)$ o $\frac{\partial f}{\partial v}(x^0)$ e si dice *derivata direzionale* di f nella direzione v in x^0 . Notare che $D_v f(x^0)$ coincide con la derivata in $t = 0$ della restrizione di $f(x)$ alla retta $x = x^0 + vt$.

2. Osservazione È chiaro che le derivate parziali sono un caso particolare delle derivate direzionali. Il loro calcolo è molto semplice: infatti la definizione si può scrivere anche così

$$D_j f(x^0) = \lim_{t \rightarrow 0} \frac{1}{t} [f(x_1^0, \dots, x_{j-1}^0, x_j^0 + t, x_{j+1}^0, \dots, x_n^0) - f(x^0)];$$

da qui si vede che derivare f rispetto a x_j vuol dire tenere ferme tutte le variabili tranne x_j , cioè considerarle come costanti, e derivare f come se fosse funzione della sola variabile x_j . Calcolare le derivate parziali è abbastanza semplice: basta applicare le regole note per il caso di una sola variabile. Ad esempio:

$$\frac{\partial}{\partial x}(e^{(x+y)^3}) = e^{(x+y)^3} \frac{\partial}{\partial x}((x+y)^3) = e^{(x+y)^3} 3(x+y)^2.$$

Cosa abbiamo fatto? abbiamo trattato y come una costante, e derivato (rispetto a x) utilizzando le consuete regole di calcolo. È sufficiente un po' di pratica per imparare a calcolare le derivate parziali senza problemi.

3. Osservazione Diversamente da quanto accade per le funzioni di una sola variabile, una funzione di più variabili può essere derivabile in un punto ma non continua. Ad esempio la funzione f_4 è derivabile in $(0, 0)$, perché costante lungo gli assi; ma chiaramente non è continua perché ad esempio $f_4(x, x) = 0$ per $x \neq 0$. La situazione non migliora se si suppone che una funzione abbia tutte le derivate direzionali in un punto: potrebbe ancora essere discontinua in quel punto. Ad esempio

$$f_{36}(x, y) = \begin{cases} \left(\frac{x^2 y}{x^4 + y^2} \right)^2 & \text{se } (x, y) \neq 0 \\ 0 & \text{in } (0, 0) \end{cases}$$

non è continua in 0 perché la successione $(1/n, 1/n^2)$ tende a 0 mentre

$$\lim_{n \rightarrow +\infty} f_{36}\left(\frac{1}{n}, \frac{1}{n^2}\right) = \lim_{n \rightarrow +\infty} (n^{-4}n^{-4} + n^{-4})^2 = \frac{1}{4} \neq f(0, 0);$$

ma f_{36} è derivabile in tutte le direzioni in 0, infatti

$$D_v f_{36}(0, 0) = \lim_{t \rightarrow 0} \frac{t^2 v_1^4 v_2^2}{(t^2 v_1^4 + v_2^2)^2} = 0.$$

Un altro esempio, dovuto a Peano, è la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ che vale 1 se $x^2 < y < 2x^2$ e 0 altrimenti. Se si esamina il valore di f su una retta passante per l'origine, si vede che f è identicamente nulla per punti abbastanza vicini all'origine, quindi nell'origine tutte le derivate direzionali esistono e sono nulle. D'altra parte è chiaro che la funzione non è continua in $(0, 0)$.

4. Osservazione Se una funzione è assegnata come composizione di funzioni elementari la sua derivabilità nella maggior parte dei punti segue semplicemente dal fatto che è composizione di funzioni derivabili; non c'è quindi alcuna verifica da fare. Ad esempio $f = x^2 - yx^5$ è derivabile su tutto $\mathbb{R}^2$ perché composizione di funzioni derivabili (notare che qui si stanno usando teoremi per funzioni di *una sola* variabile, in quanto si suppone ad es. che x sia costante e si vuole derivare rispetto a y , o viceversa). Una funzione come $g = |x + y|$ è più delicata: la funzione $|\cdot|$ che compare nella definizione di g non è

derivabile in 0, quindi i punti $x + y = 0$ vanno studiati a parte. In tali casi per vedere se g è derivabile bisogna ritornare alla definizione, e fare il limite del rapporto incrementale. Ad esempio qui i punti da esaminare sono quelli della retta $x = -y$, cioè i punti del tipo $(\alpha, -\alpha)$:

$$\frac{1}{t}[f(\alpha + t, -\alpha) - f(\alpha, -\alpha)] = \frac{1}{t}[|t| - 0] = |t|;$$

da qui si vede che il limite del rapporto incrementale rispetto a x non esiste, quindi g non è derivabile rispetto a x in $(\alpha, -\alpha)$ (e quindi non è derivabile in $(\alpha, -\alpha)$).

ESERCIZI

- 1 Determinare gli insiemi di derivabilità di F_1 - F_{18} .
 - 2 Calcolare le derivate parziali D_1F , D_2F in un generico punto (x, y) dell'insieme di derivabilità per le funzioni F_1 - F_{18} .
 - 3 Determinare gli insiemi di derivabilità di f_1 - f_{36} .
 - 4 Calcolare le derivate parziali D_1f , D_2f in un generico punto (x, y) dell'insieme di derivabilità per le funzioni f_1 - f_{36} .
-
- 01 Siano $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$. Se f e g hanno derivate direzionali in 0 nella direzione v , allora anche $f \cdot g$, $f + g$ hanno derivata direzionale in 0 nella direzione v , e si ha $D_v(fg) = f(0) \cdot D_vg(0) + g(0) \cdot D_vf(0)$, $D_v(f + g) = D_vf(0) + D_vg(0)$.

2-3 DIFFERENZIALE

Come visto nell'Osservazione 2.3, le derivate parziali non danno una vera estensione a più variabili del concetto di derivata su $\mathbb{R}$ nel senso che una funzione derivabile in un punto può anch essere singolare in quel punto: le derivate parziali "non si accorgono" della singolarità. La vera estensione è il differenziale, per la cui definizione occorre richiamare un concetto di algebra lineare.

1. Definizione Sia V uno spazio vettoriale su $\mathbb{R}$. Una *forma lineare* o *funzionale* su V è un'applicazione lineare L da V in $\mathbb{R}$ (cioè tale che $L(x+y) = L(x) + L(y)$, $L(\lambda x) = \lambda L(x)$, $\forall x, y \in V$ e $\forall \lambda \in \mathbb{R}$). L'insieme di tutte le forme lineari su V si indica con V^* , ed è detto il *duale (algebrico)* di V . Se $L \in V^*$, il valore di L in v si può indicare sia (come di solito) con $L(v)$ che più semplicemente con Lv .

2. Osservazione V^* è uno spazio vettoriale su $\mathbb{R}$. Consideriamo il caso particolarmente importante $V = \mathbb{R}^n$. Per definire una forma $L \in (\mathbb{R}^n)^*$ basta assegnarne il valore sugli n elementi di una base di $\mathbb{R}^n$, $v_1, \dots, v_n$; infatti ogni elemento $v \in \mathbb{R}^n$ si scrive $v = \mu_1 v_1 + \dots + \mu_n v_n$ e quindi $L(v) = \mu_1 L(v_1) + \dots + \mu_n L(v_n)$. In particolare si indicano con e_j^* , $j = 1, \dots, n$, i funzionali che sulla base canonica $e_1, \dots, e_n$ hanno i seguenti valori: $e_j^*(e_i) = 1$ se $i = j$, 0 altrimenti. $\{e_1^*, \dots, e_n^*\}$ è una base (*canonica*) di $(\mathbb{R}^n)^*$; quindi vediamo che $(\mathbb{R}^n)^*$ ha dimensione n , e si può identificare con $\mathbb{R}^n$. Quindi per assegnare una forma lineare su $\mathbb{R}^n$ basta assegnare i suoi *coefficienti*, cioè i numeri $\lambda_j = L(e_j)$ dati dalla rappresentazione $L = \lambda_1 e_1^* + \dots + \lambda_n e_n^*$. Allora la forma si scrive semplicemente come

$$L(h) = L(h_1, \dots, h_n) = \lambda_1 h_1 + \dots + \lambda_n h_n \quad (1)$$

(perché $e_j^*(h) = h_j$). Conclusione *importante*: le forme lineari su $\mathbb{R}^n$ sono tutte e sole le espressioni del tipo (1). Nel seguito quando parleremo di forme lineari intenderemo sempre: su $\mathbb{R}^n$.

Sia A aperto di $\mathbb{R}^n$, $x^0 \in A$, $f : A \rightarrow \mathbb{R}$.

3. Definizione f si dice *differenziabile* in x^0 se esiste una forma lineare L tale che

$$\lim_{h \rightarrow 0} \frac{f(x^0 + h) - f(x^0) - L(h)}{|h|} = 0. \quad (2)$$

Allora L si dice *differenziale* di f in x^0 , e si indica con $df(x^0)$. Se f è differenziabile in ogni punto di A , f si dice *differenziabile* in A . L'*insieme di differenziabilità* di f è l'insieme di tutti i punti in cui f è differenziabile.

Talvolta è più comodo scrivere il limite della definizione di differenziale nella forma seguente:

$$\lim_{x \rightarrow x^0} \frac{f(x) - f(x^0) - L(x - x^0)}{|x - x^0|} = 0.$$

4. Osservazione Notare che la forma $L = df(x^0)$ non è sempre la stessa, ma dipende dal punto x^0 . Fare attenzione alla scrittura: $df(x^0)(h)$ è il differenziale di f in x^0 applicato al vettore h . Quando si scrive df , si intende una funzione che dipende dal punto $x \in A$, e per ogni x è una forma lineare in h :

$$df : A \rightarrow (\mathbb{R}^n)^*.$$

5. Osservazione Se f è differenziabile in x^0 , allora è anche derivabile in x^0 in tutte le direzioni, infatti preso un vettore di modulo unitario v

$$\lim_{t \rightarrow 0} \frac{f(x^0 + tv) - f(x^0)}{t} = \lim_{t \rightarrow 0} \frac{f(x^0 + tv) - f(x^0) - L(tv)}{t} + L(v) = L(v).$$

Otteniamo così che $D_v f(x^0) = L(v) = df(x^0)(v)$. In particolare $D_j f(x^0) = L(e_j) = \lambda_j$, ossia i coefficienti del differenziale sono proprio le derivate parziali di f in x^0 :

$$L(h) = df(x^0)(h) = D_1 f(x^0)h_1 + \cdots + D_n f(x^0)h_n. \quad (3)$$

Ne segue che: se vogliamo verificare che f è differenziabile in un punto, calcoliamo anzitutto le derivate parziali in quel punto, quindi prendiamo la forma L come in (3), e proviamo a fare il limite (2); se otteniamo 0, la funzione è differenziabile, se il limite non esiste (o non fa 0), non lo è. Naturalmente, se qualcuna delle derivate parziali in un punto non esiste, la funzione non può essere differenziabile in quel punto.

Usando la notazione del prodotto scalare, si può scrivere anche $df(x^0)(h) = (Df(x^0), h) = Df(x^0) \cdot h$ dove Df è il gradiente di f .

6. Osservazione Ricordiamo che la notazione $o(\phi)$ per $x \rightarrow 0$, letta *o piccolo di ϕ* , indica una qualunque funzione $\psi(x)$ definita vicino a $0 \in \mathbb{R}$ che tende a zero per $x \rightarrow 0$ *più velocemente della funzione ϕ* , ossia tale che

$$\lim_{x \rightarrow 0} \frac{\psi(x)}{\phi(x)} = 0.$$

Quindi ad esempio le funzioni x^3 e $|x|^{5/2}$ sono $o(x^2)$ per $x \rightarrow 0$; le funzioni x e x^2 *non sono* $o(x^2)$ per $x \rightarrow 0$.

Ricordiamo anche che una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ è derivabile in un punto x^0 se e solo se si ha

$$f(x^0 + h) = f(x^0) + f'(x^0)h + o(h).$$

Possiamo scrivere la definizione di differenziale esattamente nello stesso modo: f definita su un aperto di $\mathbb{R}^n$ contenente il punto x^0 è differenziabile in x^0 se e solo se esiste una forma lineare L tale che

$$f(x^0 + h) = f(x^0) + L(h) + o(|h|).$$

Quanto vale esattamente $o(|h|)$ in questa formula? Basta ricavarlo dalla definizione:

$$o(|h|) = \frac{f(x^0 + h) - f(x^0) - L(h)}{|h|}.$$

Questo modo di scrivere la definizione mostra come il differenziale sia la naturale generalizzazione del concetto di derivata a funzioni di più variabili. Naturalmente la forma lineare $L(h)$ nel caso $n = 1$ si riduce al prodotto $f'(x^0)h$, $h \in \mathbb{R}$.

Quindi possiamo dire che una funzione è differenziabile quando nel suo incremento si può isolare una parte lineare in h e il resto tende a 0 più velocemente di h .

7. Esempio Se f non è differenziabile in un punto, allora non è detto che le derivate direzionali si calcolino secondo la formula $D_v f(x^0) = D_1 f(x^0)v_1 + \dots + D_n f(x^0)v_n$. Ad esempio la funzione $f = f_6$

$$f_6(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & \text{se } (x, y) \neq 0 \\ 0 & \text{in } (0, 0) \end{cases}$$

è derivabile in tutte le direzioni in 0, e si ha $D_v f(0, 0) = v_1^2 v_2$ per $v = (v_1, v_2)$ di modulo 1. In particolare $D_1 f(0, 0) = D_2 f(0, 0) = 0$. Quindi non vale la formula suesposta (e quindi f non è differenziabile, altrimenti la formula sarebbe valida). La funzione è anche continua in 0 perché $|f(x, y)| \leq |y|$.

8. Proposizione Se f è differenziabile in x^0 , allora è continua in x^0 .

DIM. Sappiamo che $f(x^0 + h) = f(x^0) + L(h) + o(|h|)$. Se calcoliamo il limite per $h \rightarrow 0$ di ambo i membri, e osserviamo che $L(h) \rightarrow 0$ e $o(|h|) \rightarrow 0$, otteniamo subito che $\lim_{h \rightarrow 0} f(x^0 + h) = f(x^0)$.

9. Teorema (del differenziale totale) Se f è derivabile in un intorno di x^0 e le derivate sono continue in x^0 , allora f è differenziabile in x^0 .

DIM. Ci limitiamo a trattare il caso di una funzione di due variabili $f(x_1, x_2)$; il caso generale è completamente analogo. Se le derivate parziali esistono nell'intorno $B = B(x^0, r)$ di $x^0 = (x_1^0, x_2^0)$, possiamo scrivere per $x \in B$

$$f(x) - f(x^0) = f(x_1, x_2) - f(x_1^0, x_2) + f(x_1^0, x_2) - f(x_1^0, x_2^0).$$

Per il Teorema del valor medio abbiamo

$$f(x_1, x_2) - f(x_1^0, x_2) = (x_1 - x_1^0)D_1 f(\xi, x_2)$$

$$f(x_1^0, x_2) - f(x_1^0, x_2^0) = (x_2 - x_2^0)D_2f(x_1^0, \eta)$$

con ξ tra x_1 e x_1^0 , η tra x_2 e x_2^0 . Dobbiamo mostrare che f è differenziabile, quindi consideriamo la forma $L(h) = D_1f(x^0)h_1 + D_2f(x^0)h_2$ (unico candidato possibile come differenziale) e calcoliamo il limite della definizione. Possiamo scrivere, usando le identità precedenti,

$$\begin{aligned} f(x) - f(x^0) - L(x - x^0) &= (x_1 - x_1^0)(D_1f(\xi, x_2) - D_1f(x_1^0, x_2^0)) \\ &\quad + (x_2 - x_2^0)(D_2f(x_1^0, \eta) - D_2f(x_1^0, x_2^0)) \end{aligned}$$

e quindi

$$\left| \frac{f(x) - f(x^0) - L(x - x^0)}{|x - x^0|} \right| \leq |D_1f(\xi, x_2) - D_1f(x_1^0, x_2^0)| + |D_2f(x_1^0, \eta) - D_2f(x_1^0, x_2^0)|$$

da cui la tesi passando al limite per $x \rightarrow x^0$ e usando la continuità delle derivate parziali.

10. Osservazione In particolare, se f è derivabile con derivate continue in un aperto A , allora f è differenziabile in A .

11. Osservazione Sia f differenziabile in x^0 . Allora la funzione $\phi(x) = f(x^0) + df(x^0)(x - x^0)$ ha per grafico un piano passante per il punto $(x^0, f(x^0)) \in \mathbb{R}^{n+1}$, detto il *piano tangente* al grafico di f nel punto x^0 . Notare infatti che

$$\lim_{x \rightarrow x^0} \frac{f(x) - \phi(x)}{|x - x^0|} = 0.$$

12. Esempio Una funzione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}$ ovviamente è derivabile in ogni punto, e il differenziale dL in ogni punto è esattamente uguale ad L , cioè non dipende dal punto [basta applicare la definizione, e si vede che il rapporto incrementale nella definizione di df è identicamente nullo per $f = L$]. Un caso speciale molto interessante è dato dalle funzioni $\pi_j : \mathbb{R}^n \rightarrow \mathbb{R}$ che ad ogni punto x associano la sua j -esima coordinata, ossia $\pi_j(x) = x_j$. Naturalmente π_j è lineare e quindi $d\pi_j$ in ogni punto x^0 è ancora la funzione che ad ogni vettore associa la sua coordinata j -esima:

$$d\pi_j(x^0)(h) = h_j$$

cioè, per ogni x^0 , $d\pi_j(x^0) = e_j^*$ (si potrebbe anche scrivere $d\pi_j = \pi_j$, ma la prima π_j opera sulle x , la seconda sul vettore h , allora per evitare confusioni è meglio tenere le notazioni distinte).

Spesso la funzione π_j si indica semplicemente con x_j , e quindi si scrive $d\pi_j = dx_j = e_j^*$. Dato che gli e_j^* sono una base di $(\mathbb{R}^n)^*$, vediamo che $\{dx_1, \dots, dx_n\}$ è (in ogni punto) una base di $(\mathbb{R}^n)^*$, e quindi ogni differenziale $df(x^0)$ si scrive come combinazione lineare:

$$df(x^0) = D_1f(x^0)dx_1 + \dots + D_nf(x^0)dx_n.$$

Ad esempio

$$d(x^2ye^z) = 2xye^zdx + x^2e^zdy + x^2ye^zdz.$$

13. Esempio Non vale il viceversa del Teorema del differenziale totale: anzi può accadere che una funzione sia differenziabile in un punto senza essere derivabile in nessun altro punto. Per vederlo, osserviamo che una qualunque funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che

$$|f(x, y)| \leq x^2 + y^2$$

per tutti gli x, y è sicuramente differenziabile nell'origine con $df(0, 0) = 0$ (verificare!). Tuttavia il comportamento di f fuori dall'origine può essere pessimo. Ad esempio la funzione

$$f(x, y) = \begin{cases} x^2 + y^2 & \text{se } x, y \in \mathbf{Q} \\ -x^2 - y^2 & \text{altrimenti} \end{cases}$$

è differenziabile in $(0, 0)$ per quanto appena detto, ma gli insiemi di continuità, derivabilità e differenziabilità di f coincidono tutti e tre con $\{(0, 0)\}$.

14. Esempio Abbiamo definito in precedenza le *curve continue (archi)* in $\mathbb{R}^n$, ossia le funzioni continue $\phi : [a, b] \rightarrow \mathbb{R}^n$ definite su un intervallo chiuso $[a, b]$ a valori in $\mathbb{R}^n$. Naturalmente una curva $\phi(t)$ è descritta dalle sue n componenti

$$\phi(t) = (\phi_1(t), \dots, \phi_n(t))$$

ed è continua se e solo se tutte le ϕ_j sono funzioni continue (da $[a, b]$ a valori reali).

Questo suggerisce di definire una *curva* C^1 detta anche *arco* C^1 , come una funzione $\phi : [a, b] \rightarrow \mathbb{R}^n$ tale che tutte le sue componenti siano funzioni C^1 (ossia continue con derivata prima continua) da $[a, b]$ a valori reali. La *derivata* di $\phi(t)$ si definisce come il vettore delle derivate

$$\phi'(t) := (\phi'_1(t), \dots, \phi'_n(t))$$

ed è una curva continua. Se immaginiamo la curva $\phi(t)$ come la traiettoria di un punto in funzione del tempo, il vettore $\phi'(t)$ esprime la velocità del punto al tempo t .

Se uniamo due archi C^1 non otteniamo in generale un arco C^1 (potrebbe esserci uno spigolo nel punto di giunzione) ma questo non è un problema perché possiamo comunque calcolare la derivata tranne che nel punto di giunzione. In generale possiamo definire una *curva* C^1 *a tratti* detta anche *arco* C^1 *a tratti* come l'unione di un numero finito di archi C^1 .

Notiamo un fatto importante: *un aperto di $\mathbb{R}^n$ è connesso se e solo se è connesso per archi C^1 a tratti*. Abbiamo già dimostrato nel Capitolo 1 (Proposizione 7.9) che questo è vero per archi continui, ma la dimostrazione funziona perfettamente e senza modifiche anche per archi C^1 a tratti!

15. Esempio Consideriamo ora la situazione seguente: abbiamo una funzione $f : A \rightarrow \mathbb{R}$ definita su un aperto A di $\mathbb{R}^n$, e una curva $\phi : [a, b] \rightarrow A$ a valori in A ; quindi possiamo considerare la funzione composta $g(t) = f(\phi(t))$, ossia possiamo studiare i valori di f lungo la curva ϕ . Se f è differenziabile e ϕ è una curva C^1 , la funzione $g : [a, b] \rightarrow \mathbb{R}$ dovrebbe essere derivabile. Proviamo a calcolare $g'(t)$. Fissiamo un punto t , allora per $h \in \mathbb{R}$ piccolo possiamo scrivere, usando la definizione di differenziale di f nel punto $x^0 = \phi(t)$,

$$f(\phi(t+h)) = f(\phi(t)) + Df(\phi(t))k + o(|k|), \quad k := \phi(t+h) - \phi(t)$$

e quindi abbiamo

$$g'(t) = \lim_{h \rightarrow 0} \frac{f(\phi(t+h)) - f(\phi(t))}{h} = Df(\phi(t)) \lim_{h \rightarrow 0} \frac{\phi(t+h) - \phi(t)}{h} + \lim_{h \rightarrow 0} \frac{o(|k|)}{h}.$$

Esaminiamo i termini a sinistra. Il rapporto incrementale di ϕ ovviamente tende al vettore $\phi'(t)$. Il secondo limite tende a 0. Infatti, quando $k = 0$ abbiamo ovviamente

12 Calcolo differenziale

$o(|k|) = 0$; se ci limitiamo ai $k \neq 0$, possiamo scrivere

$$\frac{o(|k|)}{h} = \frac{o(|k|)}{|k|} \cdot \frac{|k|}{h}.$$

La frazione $o(|k|)/|k|$ tende a zero perché $k \rightarrow 0$, mentre il rapporto $|k|/h$ è limitato grazie al Teorema di Lagrange (di Calcolo I) applicato a ciascuna componente. In conclusione otteniamo la formula

$$g'(t) = Df(\phi(t))\phi'(t) \quad (4)$$

e vediamo che la derivata di $f \circ \phi$ è il "prodotto dei differenziali" di f e ϕ . Possiamo scrivere la formula precedente nella forma più suggestiva

$$g'(t) = df(x^0)(v), \quad x^0 = \phi(t), \quad v = \phi'(t),$$

dalla quale vediamo che il differenziale di f nel punto x^0 , applicato alla direzione v , esprime la variazione della funzione f calcolata lungo una curva ϕ passante per x^0 che esce nella direzione $v = \phi'(t)$ da x^0 . Questo punto di vista è molto importante per le applicazioni del calcolo differenziale.

Ricordiamo che se $f : [a, b] \rightarrow \mathbb{R}$ è derivabile con derivata nulla in ogni punto, allora f è costante. Se vogliamo estendere questa proprietà al caso di più variabili, ci scontriamo con un problema geometrico: è più complicato descrivere la connessione e la disconnessione degli insiemi in dimensione maggiore di uno. La versione multidimensionale del risultato è la seguente:

16. Proposizione *Siano A un aperto di $\mathbb{R}^n$, $f : A \rightarrow \mathbb{R}$ differenziabile. Se $df \equiv 0$, allora f è costante su ciascuna componente connessa di A .*

DIM. Sia A_1 una componente connessa di A , quindi A_1 è anche connessa per archi C^1 a tratti in quanto è aperta. Prendiamo due punti qualunque $x, y \in A_1$ e uniamoli con una curva $\phi(t)$ C^1 a tratti tutta contenuta in A_1 . Usando la formula dell'esempio precedente possiamo scrivere su ognuno dei tratti C^1 di ϕ

$$\frac{d}{dt}[f(\phi(t))] = Df(\phi(t)) \cdot \phi'(t) = 0$$

e quindi $f(\phi(t))$ è costante. Ne segue che $f(x) = f(y)$ ossia f è costante su A_1 .

ESERCIZI

- 1 Verificare le affermazioni dell'Osservazione 2.
- 2 Siano f, g differenziabili in x^0 . Dimostrare che $f+g, f \cdot g$ sono differenziabili in x^0 , e si ha: $d(f+g)(x^0) = df(x^0) + dg(x^0)$, $d(fg)(x^0) = f(x^0) \cdot dg(x^0) + g(x^0) \cdot df(x^0)$.
- 3 Una funzione costante è differenziabile con differenziale nullo.
- 4 Determinare gli insiemi di differenziabilità di F_1 - F_{18} , f_0 - f_{36} .
- 5 Verificare l'affermazione dell'Osservazione 10.
- 6 $f : \mathbb{R} \rightarrow \mathbb{R}$ è derivabile se e solo se è differenziabile. Scrivere il differenziale di f usando la derivata di f .
- 7 Sia A aperto limitato non vuoto di $\mathbb{R}^n$, sia $f : \bar{A} \rightarrow \mathbb{R}$ continua su $\bar{A}$, differenziabile su A e costante su ∂A . Allora esiste $x^0 \in A$ in cui il differenziale di f si annulla (Teorema di Rolle in $\mathbb{R}^n$).
- 8 Sia $L : \mathbb{R}^n \rightarrow \mathbb{R}$ una forma lineare. Calcolare il differenziale dL di L .
- 9 Calcolare il differenziale e il gradiente della forma quadratica $f(x) = \sum_{i,j=1}^n a_{ij}x_i x_j$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

-
- 01** Siano $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$, f differenziabile in 0, $f(0) = 0$, g continua in 0. Allora fg è differenziabile in 0, e si ha $d(fg)(0) = g(0) \cdot df(0)$.
- 02** Siano $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$ derivabili, e sia $g : \mathbb{R}^n \rightarrow \mathbb{R}$ definita come $g(x) = f_1(x_1) + \dots + f_n(x_n)$. Allora g è differenziabile e $dg(x)(h) = \sum f'_j(x_j)h_j$.
- 03** Siano $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabili in 0 e $g(0) \neq 0$. Dimostrare che f/g è ben definita in un intorno di 0 ed è differenziabile in 0, e calcolarne il differenziale.

2-4 DERIVATE SUCCESSIVE, FORMULA DI TAYLOR

Se una funzione f è derivabile su un insieme, le sue derivate $D_j f(x)$ sono n funzioni, che possono essere a loro volta derivabili. Quindi si possono calcolare le derivate seconde di f : $D_i D_j f = \frac{\partial^2 f}{\partial x_i \partial x_j}$, le derivate terze e così via. Si dice ordine di una derivata il numero totale di derivate fatte rispetto alle singole variabili. Una notazione comoda è la seguente: se $v = (v_1, \dots, v_n)$ è un vettore e $\alpha = (\alpha_1, \dots, \alpha_n)$ è un *multiindice*, ossia una n -upla di interi ≥ 0 , si scrive

$$v^\alpha = v_1^{\alpha_1} \dots v_n^{\alpha_n}$$

e con D^α indicheremo la derivata

$$D^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

ossia α_1 volte rispetto a x_1 , $\dots$, α_n volte rispetto a x_n ; in totale D^α è una derivata di ordine $|\alpha| = \alpha_1 + \dots + \alpha_n$, e l'intero $|\alpha|$ si chiama *lunghezza* del multiindice α . Ad esempio, se $f(x, y, z)$ è una funzione su $\mathbb{R}^3$,

$$D^{(1,1,1)} f = \frac{\partial^3 f}{\partial x \partial y \partial z}, \quad D^{(2,0,3)} f = \frac{\partial^5 f}{\partial^2 x \partial^3 z}$$

e così via. Inoltre dato un multiindice $\alpha = (\alpha_1, \dots, \alpha_n)$, il suo *fattoriale* è definito come

$$\alpha! = \alpha_1! \dots \alpha_n!.$$

1. Osservazione In generale, anche se le derivate seconde di f esistono, non è detto che $D_i D_j f = D_j D_i f$, ossia le derivate miste possono dipendere dall'ordine di derivazione. Ad esempio, se $f = f_{37}$ è la funzione

$$f_{37}(x, y) = \begin{cases} y^2 \arctan \frac{x}{y} & \text{se } y \neq 0 \\ 0 & \text{se } y = 0 \end{cases}$$

si ha $D_1 D_2 f(0, 0) = 0$, $D_2 D_1 f(0, 0) = 1$.

2. Teorema (di Schwartz) Sia $A \subseteq \mathbb{R}^2$ aperto, $(x^0, y^0) \in A$, $f : A \rightarrow \mathbb{R}$. Supponiamo che le derivate prime e le derivate parziali miste $D_1 D_2 f$, $D_2 D_1 f$ esistano in un intorno di (x^0, y^0) e siano continue in (x^0, y^0) . Allora

$$D_1 D_2 f(x^0, y^0) = D_2 D_1 f(x^0, y^0).$$

DIM. Poniamo

$$w(h, k) = \frac{1}{hk} [f(x^0 + h, y^0 + k) - f(x^0, y^0 + k) - f(x^0 + h, y^0) + f(x^0, y^0)];$$

è chiaro che la funzione w è ben definita se $|h|, |k| < r$ con r sufficientemente piccolo. Osserviamo ora che ponendo

$$\phi(h, k) = f(x^0 + h, y^0 + k) - f(x^0, y^0 + k)$$

e applicando due volte il Teorema del valor medio, si ha

$$\begin{aligned} w(h, k) &= \frac{1}{hk} [\phi(h, k) - \phi(h, 0)] = \frac{1}{hk} k D_2 \phi(h, \eta) \\ &= \frac{1}{h} [D_2 f(x^0 + h, y^0 + \eta) - D_2 f(x^0, y^0 + \eta)] = D_1 D_2 f(x^0 + \xi, y^0 + \eta) \end{aligned}$$

per ξ compreso tra 0 e h , e η compreso tra 0 e k . Invece, ponendo

$$\psi(h, k) = f(x^0 + h, y^0 + k) - f(x^0 + h, y^0)$$

e procedendo in modo simile si ha

$$\begin{aligned} w(h, k) &= \frac{1}{hk} [\psi(h, k) - \psi(0, k)] = \frac{1}{hk} h D_1 \psi(\xi', k) \\ &= \frac{1}{k} [D_1 f(x^0 + \xi', y^0 + k) - D_1 f(x^0 + \xi', y^0)] = D_2 D_1 f(x^0 + \xi', y^0 + \eta') \end{aligned}$$

per ξ' compreso tra 0 e h , e η' compreso tra 0 e k . Quindi abbiamo

$$D_1 D_2 f(x^0 + \xi, y^0 + \eta) = D_2 D_1 f(x^0 + \xi', y^0 + \eta');$$

se ora facciamo tendere (h, k) a $(0, 0)$, anche ξ, ξ', η, η' tendono a 0 e per la continuità delle derivate miste otteniamo la tesi.

Applicando il teorema precedente si dimostra subito che le derivate parziali miste $D_i D_j f$ e $D_j D_i f$ di una funzione di n variabili sono uguali in un punto se esistono in un intorno e sono continue nel punto. In particolare, se f è di classe C^2 , la matrice delle sue derivate seconde in x

$$Hf(x) = [D_i D_j f(x)]_{i,j=1}^n$$

detta *l'Hessiana* di f , è una matrice simmetrica.

Allo stesso modo si ottiene il seguente risultato generale:

3. Proposizione Sia A un aperto di $\mathbb{R}^n$ e $f \in C^k(A)$. Allora le derivate $D^\alpha f$ con $|\alpha| \leq k$ non dipendono dall'ordine di derivazione.

4. Osservazione Per familiarizzarci con la notazione dei multiindici studiamo la *potenza di un multinomio*:

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha. \quad (1)$$

La somma è estesa a tutti i multiindici di lunghezza k . Il caso $k = 2$ è semplicemente la potenza di un binomio, che conosciamo già:

$$(x_1 + x_2)^k = \sum_{j=0}^k \frac{k!}{j!(k-j)!} x_1^j x_2^{k-j} = \sum_{\alpha_1 + \alpha_2 = k} \frac{k!}{\alpha_1! \alpha_2!} x_1^{\alpha_1} x_2^{\alpha_2}.$$

La formula generale si dimostra per induzione su n , tenendo k fissato. Infatti, se la formula è vera per $n - 1$ addendi, possiamo scrivere, indicando con $\hat{x} = (x_1, \dots, x_{n-1})$ e $\hat{\alpha} = (\alpha_1, \dots, \alpha_{n-1})$ i vettori di dimensione $n - 1$,

$$\begin{aligned} (x_1 + \dots + x_n)^k &= ((x_1 + \dots + x_{n-1}) + x_n)^k = \sum_{i+j=k} \frac{k!}{i!j!} (x_1 + \dots + x_{n-1})^i x_n^j \\ &= \sum_{i+j=k} \frac{k!}{i!j!} \sum_{|\hat{\alpha}|=i} \frac{i!}{\hat{\alpha}!} \hat{x}^{\hat{\alpha}} x_n^j = \sum_{|\hat{\alpha}|+j=k} \frac{k!}{\hat{\alpha}!j!} \hat{x}^{\hat{\alpha}} x_n^j \end{aligned}$$

e questa è proprio la formula (1) (basta porre $\alpha_n = j$).

Diamo qualche cenno sulla *formula di Taylor* per funzioni di più variabili. Sia $B = B(x^0, r)$ una palla aperta di $\mathbb{R}^n$ e $f : B \rightarrow \mathbb{R}$ una funzione di classe C^N . Fissiamo un punto $x \in B$ distinto da x^0 , e chiamando $w = x - x^0$, consideriamo la funzione di una sola variabile

$$F(t) = f(x^0 + t(x - x^0)) = f(x^0 + tw),$$

definita per tutti i $|t| < r/|x - x^0|$. Si tratta naturalmente della restrizione di f alla retta che passa per x^0 e x . Come visto in precedenza, sappiamo calcolare le derivate di F : per la derivata prima abbiamo subito

$$F'(t) = w \cdot Df(x^0 + tw)$$

e ripetendo il procedimento otteniamo

$$F^{(k)}(t) = (w \cdot D)^k f(x^0 + tw)$$

dove $(w \cdot D)f = w_1 D_1 f + \dots + w_n D_n f$ mentre $(w \cdot D)^k f$ denota la stessa operazione ripetuta k volte. Vediamo così che anche $F(t)$ è di classe C^N . Applichiamo adesso a F la formula di Taylor in $t = 0$ calcolata nel punto $t = 1$:

$$F(1) = \sum_{k=0}^N \frac{F^{(k)}(0)}{k!} + R_N \quad \Rightarrow \quad f(x) = \sum_{k=0}^N \frac{(w \cdot D)^k f(x^0)}{k!} + R_N(x^0, x)$$

dove naturalmente R_N è il resto di Taylor. A questo punto basta sviluppare la potenza del multinomio $(w \cdot D)^k$ tramite la formula (1), e ricordare la forma di Lagrange del resto di Taylor, per ottenere il risultato seguente:

5. Teorema Sia $B = B(x^0, r)$ una palla aperta e $f : B \rightarrow \mathbb{R}$ una funzione di classe C^{N+1} . Allora per tutti gli $x \in B$ vale la formula di Taylor

$$f(x) = \sum_{|\alpha| \leq N} \frac{D^\alpha f(x^0)}{\alpha!} (x - x^0)^\alpha + R_N(x^0, x)$$

dove il resto R_N è dato (in forma di Lagrange) da

$$R_N(x^0, x) = \sum_{|\alpha|=N+1} \frac{D^\alpha f(x^0 + \xi(x - x^0))}{\alpha!} (x^0 - x)^\alpha$$

per qualche $\xi \in (0, 1)$ che dipende da x^0 e x .

6. Osservazione Un'altra formula utile per il resto è quella integrale di Peano:

$$R_N(x^0, x) = (N+1) \sum_{|\alpha|=N+1} \frac{(x-x^0)^\alpha}{\alpha!} \int_0^1 (1-t)^N D^\alpha f(x^0 + t(x-x^0)) dt.$$

Anche questa formula è conseguenza immediata di quella per funzioni di una variabile.

7. Osservazione Dalla forma di Lagrange, vediamo subito che se B_1 è la palla chiusa $\overline{B}(x^0, r_1)$, con $r_1 < r$, il resto di Taylor verifica la stima

$$|R_N(x^0, x)| \leq C_N \left(\max_{B_1} \sum_{|\alpha|=N+1} |D^\alpha f| \right) |x - x^0|^{N+1} \quad \text{per ogni } x \in B_1,$$

dove C_N è una costante che dipende solo da N . In particolare, si ha $R_N = o(|x - x^0|^N)$ esattamente come per le funzioni di una variabile.

8. Osservazione I casi più utili e importanti della formula di Taylor sono gli sviluppi al primo e al secondo ordine. Al primo ordine si ha

$$f(x) = f(x^0) + \sum_{j=1}^n D_j f(x^0)(x_j - x_j^0) + R_1(x^0, x) = f(x^0) + Df(x^0) \cdot (x - x^0) + R_1(x^0, x);$$

in questo caso il resto in forma di Lagrange si può scrivere semplicemente

$$R_1(x^0, x) = \frac{1}{2} \sum_{j,k=1}^n D_j D_k f(x^0 + \xi(x - x^0))(x_j - x_j^0)(x_k - x_k^0)$$

o in forma più compatta

$$R_1(x^0, x) = \frac{1}{2} (Hf(x^0 + \xi(x - x^0))(x - x^0), (x - x^0))$$

usando l'Hessiana di f . Lo sviluppo al secondo ordine analogamente si può scrivere

$$f(x) = f(x^0) + Df(x^0) \cdot (x - x^0) + \frac{1}{2} (Hf(x^0)(x - x^0), (x - x^0)) + R_2(x^0, x).$$

ESERCIZI

- 1 Verificare l'Osservazione 1.
- 2 Verificare la Proposizione 3.
- 3 Sviluppare in formula di Taylor nell'origine $(0,0)$ le seguenti funzioni: e^{x+y} fino all'ordine 3, $\log(1 + x \sin y^2)$ fino all'ordine 4, $x^3 \sin y + y \log(1 + x)$ fino all'ordine 4.
- 4 Data la funzione $f(x) = e^{-|x|^2}$ da $\mathbb{R}^n$ a valori reali, calcolare i polinomi di Taylor nel punto $x^0 = 0$ fino all'ordine 3. Qual è la formula di Taylor di ordine N qualunque?
- 5 Stessa domanda dell'esercizio precedente per la funzione $f(x) = \frac{1}{1+|x|^2}$.

- 01 Vale la seguente versione del Teorema di Schwartz: se D_1f , D_2f esistono in un intorno di (x^0, y^0) e sono differenziabili in (x^0, y^0) , allora $D_1D_2f(x^0, y^0) = D_2D_1f(x^0, y^0)$.
- 02 Calcolare tutte le derivate quarte di $f(x, y) = x^4 + y^4 - 4x^2y^2$.

2-5 FUNZIONI A VALORI VETTORIALI

Abbiamo già ricordato che una funzione $f : A \rightarrow \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, è equivalente ad una m -upla di funzioni $f = (f_1, \dots, f_m)$. f è continua se e solo se lo sono $f_1, \dots, f_m$, ed è derivabile se e solo se lo sono $f_1, \dots, f_m$.

Diciamo che f è *differenziabile* in x^0 se esiste un'applicazione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tale che

$$\lim_{h \rightarrow 0} \frac{f(x^0 + h) - f(x^0) - L(h)}{|h|} = 0;$$

L si chiama come di consueto il *differenziale* di f in x^0 e si indica con $df(x^0)$; esso è rappresentato da una matrice $m \times n$, detta la *matrice Jacobiana* di f e indicata con $Df(x^0)$. È facile vedere che f è differenziabile in x^0 se e solo se sono differenziabili in x^0 le sue componenti $f_1, \dots, f_m$, e il differenziale df è la m -upla dei differenziali $df_1, \dots, df_m$. Analogamente, Df è una matrice le cui righe sono $Df_1, \dots, Df_m$:

$$Df(x^0) = \begin{pmatrix} D_1f_1(x^0) & \dots & D_nf_1(x^0) \\ \dots & \dots & \dots \\ D_1f_m(x^0) & \dots & D_nf_m(x^0) \end{pmatrix} = [D_jf_i(x^0)].$$

1. Esempio Se $\phi : [a, b] \rightarrow \mathbb{R}^n$ è una curva C^1 , abbiamo già definito la derivata $\phi'(t)$ come il vettore delle derivate $(\phi'_1(t), \dots, \phi'_n(t))$ delle singole componenti. In questo caso è facile verificare che ϕ è differenziabile, e il suo differenziale è rappresentato dalla matrice $n \times 1$ (vettore colonna)

$$D\phi(t) = \begin{pmatrix} \phi'_1(t) \\ \vdots \\ \phi'_n(t) \end{pmatrix}.$$

Vale il seguente teorema sul differenziale delle funzioni composte. Abbiamo già visto questa proprietà in un caso molto particolare, nell'Esempio 3.15.

2. Teorema Siano A aperto di $\mathbb{R}^n$, B aperto di $\mathbb{R}^m$, C aperto di $\mathbb{R}^l$. Siano $f : A \rightarrow B$, $g : B \rightarrow C$, e sia f differenziabile in $x^0 \in A$, g differenziabile in $f(x^0)$. Allora $g \circ f : A \rightarrow C$ è differenziabile in x^0 , e il differenziale di $g \circ f$ è la composizione dei differenziali:

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

ossia, la matrice Jacobiana di $g \circ f$ è il prodotto delle matrici Jacobiane

$$D(g \circ f)(x^0) = Dg(f(x^0)) \cdot Df(x^0)$$

ossia

$$D_h g_i(f(x^0)) = \sum_{j=1}^m D_j g_i(f(x^0)) D_h f_j(x^0), \quad i = 1, \dots, k, \quad h = 1, \dots, n.$$

DIM. Chiamiamo $y^0 = f(x^0)$, $L = df(x^0)$, $M = dg(y^0)$. Poniamo

$$k = f(x^0 + h) - f(x^0)$$

e notiamo che quando $h \rightarrow 0$ allora $k \rightarrow 0$ dato che f è continua, e inoltre

$$k = f(x^0 + h) - f(x^0) = L(h) + o(|h|) \quad \text{per } h \rightarrow 0 \quad (1)$$

dato che f è differenziabile in x^0 . Ne segue che

$$o(|k|) = o(|h|) \quad \text{per } h \rightarrow 0.$$

Per dimostrarlo, basta scrivere

$$\frac{o(|k|)}{|h|} = \begin{cases} 0 & \text{se } k = 0, \\ \frac{o(|k|)}{|k|} \frac{|k|}{|h|} & \text{se } k \neq 0 \end{cases}$$

e notare che il rapporto $|k|/|h|$ resta limitato quando $h \rightarrow 0$, dato che

$$\frac{|k|}{|h|} \leq \frac{|L(h)|}{|h|} + \frac{o(|h|)}{|h|}$$

e il primo termine è limitato perché L è Lipschitziana, il secondo tende addirittura a 0.

Ora, dato che g è differenziabile in $y^0 = f(x^0)$, abbiamo

$$g(y^0 + k) = g(y^0) + M(k) + o(|k|) \quad \text{per } k \rightarrow 0.$$

Scegliendo k come in (1) otteniamo

$$g(f(x^0 + h)) = g(f(x^0)) + ML(h) + M(o(|h|)) + o(|k|)$$

ossia la funzione $F = g \circ f$ soddisfa

$$F(x^0 + h) = F(x^0) + ML(h) + M(o(|h|)) + o(|k|) \quad \text{per } h \rightarrow 0.$$

Ne segue la tesi, perché $M(o(|h|))$ è un $o(|h|)$ (essendo M Lipschitziana) e anche $o(|k|)$ è un $o(|h|)$ (vedi sopra).

3. Esempio Se $\phi : \mathbb{R} \rightarrow \mathbb{R}^n$, $F : \mathbb{R}^n \rightarrow \mathbb{R}$, si ha applicando il Teorema 2

$$\frac{d}{dt} F(\phi(t)) = DF(\phi(t)) \cdot \phi'(t) = \sum_{j=1}^n D_j F(\phi(t)) \phi'_j(t).$$

Con una notazione più antica, meno precisa ma più facile da ricordare e da usare,

$$\frac{d}{dt} F(\phi_1(t), \dots, \phi_n(t)) = \frac{\partial F}{\partial x_1} \frac{d\phi_1}{dt} + \dots + \frac{\partial F}{\partial x_n} \frac{d\phi_n}{dt}.$$

Ad esempio ($x, y \in \mathbb{R}^n$)

$$\frac{d}{dt} F(x + ty) = DF(x + ty) \cdot y = \sum_{j=1}^n D_j F(x + ty) y_j.$$

Analogamente, se $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^n$, $F : \mathbb{R}^n \rightarrow \mathbb{R}$, quindi $F \circ \phi : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$D(F \circ \phi)(x) = DF(\phi(x)) \cdot D\phi(x)$$

ossia, per $i = 1, 2$

$$D_i(F \circ \phi)(x) = \sum_{j=1}^n D_j F(\phi(x)) D_i \phi_j(x).$$

Con la notazione tradizionale, $F = F(y_1, \dots, y_n)$, $F \circ \phi = F(\phi_1(x_1, x_2), \dots, \phi_n(x_1, x_2))$ e

$$\frac{\partial(F \circ \phi)}{\partial x_1} = \frac{\partial F}{\partial y_1} \frac{\partial \phi_1}{\partial x_1} + \frac{\partial F}{\partial y_2} \frac{\partial \phi_2}{\partial x_1} + \dots + \frac{\partial F}{\partial y_n} \frac{\partial \phi_n}{\partial x_1},$$

$$\frac{\partial(F \circ \phi)}{\partial x_2} = \frac{\partial F}{\partial y_1} \frac{\partial \phi_1}{\partial x_2} + \frac{\partial F}{\partial y_2} \frac{\partial \phi_2}{\partial x_2} + \dots + \frac{\partial F}{\partial y_n} \frac{\partial \phi_n}{\partial x_2}.$$

4. Esempio (Coordinate polari in $\mathbb{R}^2$). Consideriamo l'applicazione $\Phi : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}^2$ definita da

$$\begin{cases} x = \Phi_1(\rho, \theta) = \rho \cos \theta \\ y = \Phi_2(\rho, \theta) = \rho \sin \theta. \end{cases}$$

Data una funzione $f(x, y)$ su $\mathbb{R}^2$, *scrivere f in coordinate polari* vuol dire considerare la funzione

$$\tilde{f}(\rho, \theta) = f \circ \Phi = f(\rho \cos \theta, \rho \sin \theta).$$

La matrice Jacobiana di Φ è

$$D\Phi = \begin{pmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{pmatrix}.$$

Quindi si hanno le formule

$$\frac{\partial \tilde{f}}{\partial \rho} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \rho} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \rho} = \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta,$$

$$\frac{\partial \tilde{f}}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = -\frac{\partial f}{\partial x} \rho \sin \theta + \frac{\partial f}{\partial y} \rho \cos \theta$$

(in forma sintetica, $D_{\rho, \theta} \tilde{f} = D_{x, y} f \cdot D\Phi$). Notare che la funzione Φ non è iniettiva. Se chiamiamo $I = \{(x, y) : x \leq 0, y = 0\}$ il semiasse delle x negative, e restringiamo Φ nel modo seguente:

$$\Phi :]0, +\infty[\times]-\pi, \pi[\rightarrow \mathbb{R}^2 \setminus I$$

allora Φ è suriettiva e invertibile e si può esprimere Φ^{-1} come

$$\rho = \sqrt{x^2 + y^2},$$

$$\theta = \begin{cases} \arctan(y/x) & \text{per } x > 0, \\ \pi/2 - \arctan(x/y) & \text{per } y > 0, \\ -\pi/2 - \arctan(x/y) & \text{per } y < 0. \end{cases}$$

La funzione θ è continua su $\mathbb{R}^2 \setminus I$, ed è esattamente l'angolo del vettore (x, y) con il semiasse delle x positive, misurato in radianti in senso antiorario.

ESERCIZI

- 1** Verificare il Teorema 2 nel caso seguente: $f :]a, b[\rightarrow A$ differenziabile in t_0 , $A \subseteq \mathbb{R}^n$, $g : A \rightarrow \mathbb{R}$ differenziabile in $f(t_0)$. Si chiede cioè di dimostrare che $g \circ f$ è differenziabile in t_0 e si ha

$$(g \circ f)'(t_0) = \sum_{j=1}^n D_j g(f(t_0)) f'(t_0)_j.$$

- 2** Calcolare $\frac{\partial(F \circ \phi)}{\partial x}$, $\frac{\partial(F \circ \phi)}{\partial y}$, dove $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ e $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ sono date da

$$F(x, y) = x^2 - y^2, \quad f_1(x, y) = \sin(x + y), \quad f_2(x, y) = \cos(x - y);$$

$$F(x, y) = xe^y, \quad f_1(x, y) = x^2 + y, \quad f_2(x, y) = y^2 + x;$$

$$F(x, y) = x - \log y, \quad f_1(x, y) = y^n, \quad f_2(x, y) = x^m.$$

- 3** Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f = (f_1, \dots, f_m)$; dimostrare che f è differenziabile in x^0 se e solo se lo sono $f_1, \dots, f_m$. In tal caso dimostrare che si ha $df(x^0) = (df_1(x^0), \dots, df_m(x^0))$ e la matrice che rappresenta $df(x^0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ha come j -esima riga $Df_j(x^0)$, il gradiente di f_j in x^0 .
- 4** Siano $f, g : \mathbb{R} \rightarrow \mathbb{R}$ di classe C^1 , e poniamo $F_1(x, y) = f(f(x) - g(y))$, $F_2(x, y) = g(g(x) - f(y))$. Scrivere la matrice Jacobiana di $F = (F_1, F_2)$.
- 5** Siano $g : \mathbb{R} \rightarrow \mathbb{R}$ e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ due funzioni di classe C^2 . Scrivere gradiente ed Hessiana di $g \circ f$.
- 6** Data la funzione $u(x) = |x|^{2-n}$ da $\mathbb{R}^n$ in $\mathbb{R}$, $n \geq 3$, verificare che $\frac{\partial^2 u}{\partial x_1^2} + \dots + \frac{\partial^2 u}{\partial x_n^2} = 0$ per $x \neq 0$.

- 01** Trovare tutte le funzioni C^1 su $\mathbb{R}^2$ tali che

$$x \frac{\partial u}{\partial y} = y \frac{\partial u}{\partial x}.$$

[Scriviamo u in coordinate polari, l'equazione diventa

$$\rho \cos \theta \frac{\partial u}{\partial y} = \rho \sin \theta \frac{\partial u}{\partial x};$$

ricordando le formule dell'Esempio 3 otteniamo

$$\frac{\partial \tilde{u}}{\partial \theta} = -\rho \sin \theta \frac{\partial u}{\partial x} + \rho \cos \theta \frac{\partial u}{\partial y} \equiv 0$$

cioè $\tilde{u} = \tilde{u}(\rho)$ non dipende da θ .]

- 02** Trovare tutte le funzioni $u(x, y)$ C^1 su $\mathbb{R}^2$ tali che

$$\frac{\partial u}{\partial x} = 0.$$

Inoltre, dato $v \in \mathbb{R}^2$ di modulo 1, trovare tutte le funzioni $u(x, y)$ C^1 su $\mathbb{R}^2$ tali che $D_v f = 0$ in tutti i punti. [Per la prima domanda: fissato y , la funzione $u(x, y)$ è costante in x ; indicando con $\alpha(y)$ tale valore costante, abbiamo che $u(x, y)$ è una qualunque funzione $\alpha(y)$ della sola y .]

- 03** Trovare tutte le soluzioni C^1 $u(x, y)$ dell'equazione

$$\frac{\partial^2 u}{\partial x \partial y} = 0.$$

[Sia $v(x, y) = D_y u$; l'equazione data ci dice che $D_x v = 0$, quindi v è una funzione della sola y : abbiamo cioè $D_y u = \alpha(y)$. Se adesso $A(y)$ è una qualunque primitiva di $\alpha(y)$, abbiamo $D_y u = D_y A$, ossia $D_y(u - A) = 0$. Quindi $u - A$ è una funzione della sola x , $B(x)$; otteniamo allora che u deve avere necessariamente la forma $u(x, y) = A(y) + B(x)$, e d'altra parte è chiaro che una tale u risolve l'equazione proposta.]

04 Trovare tutte le soluzioni dell'equazione delle onde

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0.$$

[Eseguiamo il cambiamento di variabile $(x, y) \rightarrow (\xi, \eta)$ dato da $\xi = ct - x$, $\eta = ct + x$. Abbiamo allora

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial t} + \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial t} = c \frac{\partial u}{\partial \xi} + c \frac{\partial u}{\partial \eta},$$

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial x} = -\frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \eta},$$

da cui

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &= \frac{\partial}{\partial t} \left(c \frac{\partial u}{\partial \xi} + c \frac{\partial u}{\partial \eta} \right) = \frac{\partial}{\partial \xi} \left(c \frac{\partial u}{\partial \xi} + c \frac{\partial u}{\partial \eta} \right) \cdot \frac{\partial \xi}{\partial t} + \frac{\partial}{\partial \eta} \left(c \frac{\partial u}{\partial \xi} + c \frac{\partial u}{\partial \eta} \right) \cdot \frac{\partial \eta}{\partial t} \\ &= c^2 \left[\frac{\partial^2 u}{\partial \xi^2} + \frac{\partial^2 u}{\partial \eta^2} + 2 \frac{\partial^2 u}{\partial \xi \partial \eta} \right], \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x} \left(-\frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \eta} \right) = \frac{\partial}{\partial \xi} \left(-\frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \eta} \right) \cdot \frac{\partial \xi}{\partial x} + \frac{\partial}{\partial \eta} \left(-\frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \eta} \right) \cdot \frac{\partial \eta}{\partial x} \\ &= \frac{\partial^2 u}{\partial \xi^2} + \frac{\partial^2 u}{\partial \eta^2} - 2 \frac{\partial^2 u}{\partial \xi \partial \eta}, \end{aligned}$$

e quindi

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} \cdot c^2 \equiv 4c^2 \frac{\partial^2 u}{\partial \xi \partial \eta}.$$

Nelle nuove coordinate l'equazione è diventata

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = 0$$

che abbiamo risolto nell'esercizio precedente. Abbiamo così $u(\xi, \eta) = A(\xi) + B(\eta)$ da cui

$$u(x, y) = A(ct + x) + B(ct - y).]$$

05 Sia $u(x, y)$ soluzione dell'equazione

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

e sia $v(\rho, \theta) = u(\rho \cos \theta, \rho \sin \theta)$. Mostrare che v soddisfa l'equazione

$$v_{\rho\rho} + \frac{1}{\rho} v_{\rho} + \frac{1}{\rho^2} v_{\theta\theta} = 0.$$

Trovare tutte le soluzioni che dipendono solo da θ e quelle che dipendono solo da ρ . [$v(\theta) = a\theta + b$; invece $v(\rho)$ deve verificare $v_{\rho\rho} + v_{\rho}/\rho = 0$ cioè $(\rho v_{\rho})_{\rho} = 0$ quindi $v_{\rho} = a/\rho$ quindi $v(\rho) = a \log \rho + b$. Le uniche soluzioni definite su tutto $\mathbb{R}^2$ sono le costanti.]

06 Sia $u(x, y)$ soluzione dell'equazione

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0,$$

e sia $v(\rho, \theta) = u(\rho \cos \theta, \rho \sin \theta)$. Mostrare che v soddisfa l'equazione

$$\frac{4}{\rho^2} \sin \theta \cos \theta v_{\theta} + (\cos^2 \theta - \sin^2 \theta) \left[v_{\rho\rho} + \frac{1}{\rho} v_{\rho} - \frac{1}{\rho^2} v_{\theta\theta} \right] = 0.$$

Trovare tutte le soluzioni che dipendono solo da θ e quelle che dipendono solo da ρ . [$v(\rho)$: si deve avere $v_{\rho\rho} = v_{\rho}/\rho$ quindi $v(\rho) = a\rho^2/2 + b$, che sono definite su tutto $\mathbb{R}^2$; $v(\theta)$: si ha $4 \sin \theta \cos \theta v_{\theta} = (\cos^2 \theta - \sin^2 \theta) v_{\theta\theta}$ da cui $v_{\theta\theta}/v_{\theta} = 4 \sin \theta \cos \theta / (\cos^2 \theta - \sin^2 \theta)$ che integrata dà $v_{\theta} = a/\cos(2\theta)$ da cui

$$v(\theta) = \frac{a}{2} \log \left| \frac{\cos \theta + \sin \theta}{\cos \theta - \sin \theta} \right| + b$$

che però non è definita dappertutto.]

2-6 MASSIMI E MINIMI

Sia A un aperto di $\mathbb{R}^n$, $f : A \rightarrow \mathbb{R}$.

1. Definizione Un punto $x^0 \in A$ si dice di *massimo (minimo) locale* per f se esiste un intorno di x^0 tale che per ogni punto x di tale intorno si ha $f(x) \leq f(x^0)$ ($f(x) \geq f(x^0)$).

2. Proposizione Sia f derivabile in x^0 punto di minimo o massimo locale per f . Allora $Df(x^0) = 0$, ossia tutte le derivate di f in x^0 si annullano.

DIM. Dimostriamo che $D_1 f(x^0) = 0$: la funzione $\phi(t) = f(t, x_2^0, \dots, x_n^0)$ di una sola variabile è derivabile in $t = x_1^0$ ed ha un minimo (massimo) locale, quindi $\phi'(x_1^0) = 0$ cioè $D_1 f(x^0) = 0$. Analogamente per le altre derivate.

Un punto in cui tutte le derivate si annullano si dice *punto stazionario*. La proposizione precedente afferma che un punto di massimo o minimo locale è necessariamente un punto stazionario.

Naturalmente, non è vero in generale il viceversa: tutte le derivate possono annullarsi in un punto che non è né di massimo né di minimo. Consideriamo ad esempio le due funzioni $f_1(x, y) = x^2 + y^2$ e $f_2(x, y) = x^2 - y^2$; per entrambe l'origine è un punto stazionario, ma mentre la prima ha un minimo in $(0, 0)$ in quanto $f_1 \geq 0$ su tutto $\mathbb{R}^2$, la seconda non ha né minimo né massimo in $(0, 0)$ come si verifica subito calcolando $f(0, y) = -y^2$ e $f(x, 0) = x^2$.

Supponiamo ora che $f \in C^2(A)$, e sia $x^0 \in A$ un punto stazionario. Scriviamo lo sviluppo di Taylor al primo ordine con resto di Lagrange; dato che le derivate prime si annullano, otteniamo

$$f(x) = f(x^0) + R_1(x^0, x), \quad R_1(x^0, x) = \frac{1}{2}(Hf(x^0 + \xi(x - x^0))(x - x^0), (x - x^0)) \quad (1)$$

dove $\xi \in (0, 1)$ dipende da x^0 e x , e Hf è la matrice Hessiana, formata dalle derivate seconde di f . Notare che l'Hessiana di una funzione di classe C^2 è simmetrica per il Teorema di Schwartz.

Ricordiamo che data una matrice $H = [h_{ij}]$ simmetrica (ossia $h_{ij} = h_{ji}$), la *forma quadratica* associata ad H è la funzione su $\mathbb{R}^n$

$$(Hv, v) = \sum_{i,j=1}^n h_{ij}v_i v_j.$$

La matrice H , o la forma (Hv, v) , si dice *definita positiva* se $(Hv, v) > 0 \forall v \neq 0$, *semidefinita positiva* se $(Hv, v) \geq 0 \forall v$ (e analogamente (semi)definita negativa). Notare il fatto importante che basta verificare queste proprietà per i vettori v di modulo uno, dato che $(H(tv), (tv)) = t^2(Hv, v)$ per qualunque numero reale t .

Positività e negatività di una matrice simmetrica sono collegate ai suoi autovalori. Detti $\lambda_1, \dots, \lambda_n$ gli autovalori di H , reali perché H è simmetrica, si ha che H è definita positiva se e solo se $\lambda_i > 0 \forall i$, e semidefinita positiva se e solo se $\lambda_i \geq 0 \forall i$ (analogamente per le forme negative). Vediamo quindi che se H è definita positiva si ha $(Hv, v) \geq \lambda|v|^2$ per ogni v , con $\lambda > 0$ uguale al più piccolo autovalore (per vederlo, basta scrivere v in una base ortogonale in cui H è diagonale, che esiste perché H è simmetrica. Vedi anche es.01).

3. Proposizione Sia $f \in C^2(A)$, $A \subseteq \mathbb{R}^n$ aperto, $x^0 \in A$, e supponiamo che x^0 sia un punto stazionario ossia $Df(x^0) = 0$.

(a) Se x^0 è un minimo locale, allora $Hf(x^0) \geq 0$.

(b) Se $Hf(x) \geq 0$ per x in un intorno di x^0 , allora x^0 è un minimo locale.

(c) Se $Hf(x^0) > 0$, allora x^0 è un minimo locale.

Analoghi criteri valgono quando l'Hessiana è (semi)definita negativa per i punti di massimo locale.

DIM. Criterio (a): sia v un vettore di modulo 1 e sia $x = x^0 + tv$. Dalla formula di Taylor (1) segue che $R_1(x^0, x) = t^2(Hf(x + t\xi v)v, v) \geq 0$ per x in un intorno di x^0 ossia per t vicino a 0. Dividendo per t^2 e mandando $t \rightarrow 0$ otteniamo la tesi (dato che Hf è continua).

Criterio (b): segue immediatamente dalla formula di Taylor (1), infatti per ipotesi $R_1 \geq 0$ in un intorno di x^0 e quindi $f(x) \geq f(x^0)$.

Criterio (c): basta dimostrare che $Hf(x) > 0$ per x in un intorno di x^0 , e la conclusione allora segue dal punto precedente. Se ciò non fosse vero, esisterebbe una successione $x_k \rightarrow x^0$ di punti di A e una successione v_k di vettori di modulo 1 tali che $(Hf(x_k)v_k, v_k) \leq 0$; ma dato che la palla chiusa unitaria è compatta, possiamo estrarre da v_k una sottosuccessione convergente, che chiamiamo ancora v_k ; sia v^0 il suo limite; allora mandando $k \rightarrow \infty$ in $(Hf(x_k)v_k, v_k) \leq 0$ otteniamo $(Hf(x^0)v^0, v^0) \leq 0$ e questo va contro l'ipotesi.

4. Osservazione Se la matrice Hessiana ha sia un autovalore > 0 , sia un autovalore < 0 , allora sicuramente il punto non è né di massimo né di minimo, come segue dal punto (a). In questo caso si dice che la funzione ha un *punto sella*.

5. Esempio Consideriamo in dettaglio il caso di funzioni di due variabili. La matrice Hessiana H allora è una matrice 2×2 ; per studiarne il segno non c'è bisogno di calcolare esplicitamente gli autovalori. Come è noto dalla teoria delle matrici, se si cambia base la traccia (=somma degli elementi sulla diagonale) e il determinante delle matrici non cambiano; essendo l'Hessiana simmetrica (Teorema di Schwartz) esiste una base in cui essa è diagonale, e sulla diagonale troviamo esattamente i suoi autovalori λ_1, λ_2 ; vediamo così che $\lambda_1 + \lambda_2 = \text{tr } H$, $\lambda_1 \cdot \lambda_2 = \det H$. Seguono subito le regole:

- 1) $\det H > 0$, $\text{tr } H > 0 \Rightarrow \lambda_1 > 0, \lambda_2 > 0 \Rightarrow$ minimo;
- 2) $\det H > 0$, $\text{tr } H < 0 \Rightarrow \lambda_1 < 0, \lambda_2 < 0 \Rightarrow$ massimo;
- 3) $\det H < 0 \Rightarrow \lambda_1$ e λ_2 hanno segno opposto $\Rightarrow$ punto sella;
- 4) $\det H = 0 \Rightarrow$ un autovalore è nullo $\Rightarrow$ non si può stabilire il carattere del punto.

Naturalmente, nel caso 4) può darsi che il punto sia di massimo, minimo, o un punto sella; semplicemente i metodi dati qui non consentono di stabilirne il carattere. Infatti qui esaminiamo solo i termini della funzione fino al secondo ordine; quando questi si annullano, bisogna passare allo studio dei termini di ordine superiore, che però necessita di metodi più complicati. Vedi a questo proposito l'Esempio 6. Per completare l'elenco delle conseguenze della Proposizione 3, abbiamo anche che se un punto è un

- 1) minimo $\Rightarrow \lambda_1, \lambda_2 \geq 0 \Rightarrow \det H \geq 0$, $\text{tr } H \geq 0$;
- 2) massimo $\Rightarrow \lambda_1, \lambda_2 \leq 0 \Rightarrow \det H \geq 0$, $\text{tr } H \leq 0$.

Possiamo infine riassumere quanto visto in una regola per trovare i massimi e i minimi di una funzione f definita su un insieme $A \subseteq \mathbb{R}^2$, interni all'insieme A :

Cercare anzitutto i punti in cui $Df(x^0) = 0$; quindi passare allo studio di $Hf(x^0)$.

Se $\det H = 0$, non si può dire nulla. Quando $\det H < 0$, si ha un punto sella. Quando $\det H > 0$, si ha un massimo se $\text{tr } H < 0$, un minimo se $\text{tr } H > 0$ (notare che in questo caso il segno della traccia di H è lo stesso dell'elemento in alto a sinistra $D_1 D_1 f(x^0)$, dato che $D_1 D_1 f(x^0), D_2 D_2 f(x^0)$ devono avere lo stesso segno).

Per i punti non interni all'insieme A , la regola non vale; si devono usare altri metodi, tra cui il metodo dei *moltiplicatori di Lagrange* (Paragrafo 10).

6. Esempio Studiamo le tre funzioni $f(x, y) = x^4 + y^4$, $g(x, y) = -x^4 - y^4$, $h(x, y) = x^4 - y^4$ nel punto $(0, 0)$. Il differenziale si annulla in $(0, 0)$ per tutte e tre; ma la matrice Hessiana è identicamente nulla per tutte e tre, quindi i nostri metodi non ci permettono di stabilire il carattere del punto. Anche se è evidente che f ha un minimo, g un massimo e h un punto sella.

7. Esempio (Minima distanza tra due rette). Date due rette in $\mathbb{R}^n$, descritte in forma parametrica

$$X(t) = At + B, \quad Y(s) = Cs + D$$

(A, B, C, D vettori fissati di $\mathbb{R}^n$, $A \neq 0$, $C \neq 0$), vogliamo calcolare la loro distanza. Quindi vogliamo calcolare il minimo della funzione di due variabili

$$f(s, t) = |Y(s) - X(t)|^2 = |Cs - At + D - B|^2$$

(naturalmente f è il *quadrato* della distanza; è più comodo lavorare con f perché è una funzione regolare). Abbiamo

$$\frac{\partial f}{\partial s} = 2C \cdot (Cs - At + D - B),$$

$$\frac{\partial f}{\partial t} = -2A \cdot (Cs - At + D - B),$$

quindi imponendo la condizione $Df = 0$ otteniamo il sistema

$$\begin{cases} |C|^2 s - A \cdot Ct = B \cdot C - D \cdot C \\ -A \cdot Cs + |A|^2 t = A \cdot D - A \cdot B; \end{cases}$$

notiamo che il determinante del sistema è

$$\Delta = |A|^2 |C|^2 - (A \cdot C)^2.$$

L'Hessiana di f è costante (non dipende da s, t) e vale

$$Hf = 2 \begin{pmatrix} |C|^2 & -A \cdot C \\ -A \cdot C & |A|^2 \end{pmatrix}.$$

Si hanno allora due casi:

1) Rette parallele, ossia A, C paralleli, il che equivale a $A \cdot C = |A| \cdot |C|$: il determinante del sistema si annulla, quindi esso ha infinite soluzioni. Per ogni $s \in \mathbb{R}$ si può prendere $t = [A \cdot D - A \cdot B + A \cdot Cs] / |A|^2$ e si ottengono così delle coppie (s, t) a cui corrispondono punti di minima distanza

$$\min_{\mathbb{R}} f = \frac{1}{|A|^2} [|A|^2 \cdot |B - D|^2 - |A \cdot (B - D)|^2].$$

2) Rette non parallele, quindi $\Delta \neq 0$: allora il sistema ha un'unica soluzione

$$s_0 = \frac{(D-B)[A(A \cdot C) - C|A|^2]}{|A|^2|C|^2 - (A \cdot C)^2}, \quad t_0 = \frac{(D-B)[A|C|^2 - C(A \cdot C)]}{|A|^2|C|^2 - (A \cdot C)^2},$$

e a (s_0, t_0) corrispondono i due punti delle rette di minima distanza. Per trovare il valore della distanza cioè $\min f$ basta sostituire s_0, t_0 nella definizione di $f(s, t)$.

ESERCIZI

- 1 Per le seguenti funzioni, trovare i punti stazionari (ossia nei quali $Df = 0$), e dire se essi sono punti di massimo, minimo, sella, o se non è possibile stabilirne il carattere dall'esame dell'Hessiana:

$$f(x, y) = x(x+y)e^{x-y}$$

$$f(x, y) = y^2 - x^3$$

$$f(x, y) = y^2 + x^2y + x^4$$

$$f(x, y) = x^2 + xy^3 + 1$$

$$f(x, y) = (x+y)^2 - (x+5y+xy)$$

$$f(x, y) = xy^2(a-x-y)^2$$

$$f(x, y) = (5x+7y-25) \exp\{-(x^2+xy+y^2)\}$$

$$f(x, y) = \cos x + \cos y + \cos(x+y)$$

$$f(x, y) = \cos(x+y) + \cos(x-y)$$

$$f(x, y) = (x-2)^4 + (x-y)^4$$

$$f(x, y) = 4xy + \frac{1}{x} + \frac{1}{y}$$

$$f(x, y) = 1 - \sqrt{x^2 + y^2}$$

$$f(x, y) = (6-x-y)x^2y^3$$

$$f(x, y) = x^3 + y^3 - 3axy.$$

- 2 Stesso esercizio per le funzioni

$$f(x, y, z) = x^2 + y^2 + 2z^2 + yz$$

$$f(x, y, z) = (x+z)^2 - x(y+3z) + y^3$$

$$f(x, y, z) = x^2 + y^3 + z^2 - xy - xz$$

- 3 Dividere un numero $\alpha \geq 0$ in tre parti positive il cui prodotto sia massimo [parti uguali].
 4 Trovare la massima area di un triangolo di semiperimetro fissato s [triangolo equilatero].
 5 Trovare le coordinate dei punti della parabola $2y - x^2 = 0$ la cui distanza dal punto $(4, 1)$ è massima o minima $[(2, 2)]$.

- 01 Se H è simmetrica e $(Hv, v) > 0$ per tutti i v , allora esiste una costante $c > 0$ tale che $(Hv, v) \geq c|v|^2$ per ogni v .
 [L'insieme $S = \{|v| = 1\}$ è compatto; la forma ha dunque un minimo su S che deve essere positivo.]
 02 Sia $f: \mathbb{R}^n \rightarrow \mathbb{R}$ continua tale che $\lim_{|x| \rightarrow \infty} f(x) = +\infty$ (questo vuol dire che per ogni M esiste R tale che se $|x| > R$ allora $f(x) > M$). Dimostrare che f ammette un minimo.
 03 Tra tutti i parallelepipedi di superficie totale fissata, qual è quello di volume massimo?
 04 Sia A una matrice $n \times n$ simmetrica definita positiva e $b \in \mathbb{R}^n$, e sia $f(x) = \frac{1}{2}Ax \cdot x - b \cdot x$. Dimostrare che f ha un unico punto stazionario, dato dalla soluzione del sistema $Ax = b$. Verificare che tale punto è un minimo locale per f .
 Stessa domanda per la funzione $g(x) = |Ax - b|^2$.

2-7 FUNZIONI CONVESSE

Richiamiamo per cominciare la nozione di convessità per funzioni di una variabile reale e alcune sue caratterizzazioni mediante proprietà delle derivate.

1. Definizione Una funzione $f : I =]a, b[\rightarrow \mathbb{R}$ è *convessa* se e solo se $\forall x', x'' \in I$ e $\forall t \in [0, 1]$ posto $x_t = tx'' + (1 - t)x'$ si ha

$$f(x_t) \leq tf(x'') + (1 - t)f(x'). \quad (1)$$

Una funzione f si dice *concava* se la funzione $-f$ è convessa.

2. Esempio Dalla definizione è immediato verificare che $f(x) = |x|$ e $f(x) = x^2$ sono convesse.

3. Osservazione Dati tre punti qualsiasi $x_0 < x_1 < x_2$ in I e posto $x' = x_0$, $x'' = x_2$ e $t = (x_1 - x_0)/(x_2 - x_0)$ la condizione (1) è equivalente per terne ordinate qualsiasi a

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}. \quad (2)$$

Dati quattro punti qualsiasi $x_0 < x_1 < x_2 < x_3$ in I anche (2) si può riformulare come

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}. \quad (3)$$

La condizione (2) (rispettivamente (3)) si può riformulare nel caso di funzioni derivabili, passando al limite per $x_1 \downarrow x_0$ e $x_1 \uparrow x_2$ (rispettivamente, per $x_1 \downarrow x_0$ e $x_2 \uparrow x_3$) ottenendo la caratterizzazione seguente.

4. Proposizione Data una funzione $f : I =]a, b[\rightarrow \mathbb{R}$ di classe C^1 sono equivalenti

- (i) f è convessa.
- (ii) il grafico di f è al di sopra delle sue tangenti ovvero $\forall x, x_1 \in I$ si ha

$$f(x) \geq f(x_1) + f'(x_1)(x - x_1).$$

- (iii) f' è crescente ossia $f'(x_0) \leq f'(x_2)$ per ogni $x_0 < x_2$ in I .

Nel caso di funzioni due volte derivabili l'ultima condizione porta alla caratterizzazione seguente.

5. Proposizione Data una funzione $f : I =]a, b[\rightarrow \mathbb{R}$ di classe C^2 sono equivalenti

- (i) f è convessa.
- (ii) $f''(x) \geq 0$ per ogni $x \in I$.

6. Esempio Dalle precedenti caratterizzazioni segue immediatamente che le funzioni $f(x) = e^x$, $f(x) = |x|^p$, $p > 1$ sono convesse, la funzione $f(x) = \log x$ è concava.

Consideriamo ora il caso n -dimensionale. Dapprima evidenziamo la proprietà del dominio delle funzioni usata implicitamente in una dimensione.

7. Definizione Un sottoinsieme $\Omega \subset \mathbb{R}^n$ è *convesso* se e solo se $\forall x', x'' \in \Omega$ il punto $x_t = tx'' + (1-t)x' \in \Omega$ per ogni $t \in [0, 1]$. Il punto x_t è detto combinazione convessa di x' e x'' di parametro t .

8. Osservazione La definizione precedente risulta a ben vedere un rafforzamento della connessione per archi. Siccome il punto x_t al variare di $t \in [0, 1]$ descrive il segmento congiungente x' ed x'' , l'insieme Ω è convesso se e solo se è "connesso per segmenti".

9. Esempio Dalle proprietà della norma segue che ogni palla $B(x^0, r)$ è un insieme convesso. Più in generale, per ogni A matrice simmetrica, $A \geq 0$, l'insieme $\{(x, Ax) < 1\}$ è un insieme convesso.

La definizione di convessità per funzioni è ora analoga a quella in una dimensione.

10. Definizione Dato $\Omega \subset \mathbb{R}^n$ un insieme convesso, una funzione $f : \Omega \rightarrow \mathbb{R}$ è *convessa* se e solo se $\forall x', x'' \in \Omega$ e $\forall t \in [0, 1]$ posto $x_t = tx'' + (1-t)x'$ si ha

$$f(x_t) \leq tf(x'') + (1-t)f(x').$$

Una funzione f si dice *concava* se la funzione $-f$ è convessa.

11. Esempio Dalla definizione è immediato verificare ancora che $f(x) = |x|$ e $f(x) = |x|^2$ sono convesse.

Per dare caratterizzazioni della convessità di verifica più agevole ci restringiamo a funzioni definite su aperti $A \subset \mathbb{R}^n$ e ivi differenziabili una o due volte con continuità. L'osservazione chiave è che dati $x^0 \in A$ e $v \in \mathbb{R}^n$ qualsiasi, la funzione convessa f ristretta alla retta $x = x^0 + tv$, $g(t) = f(x^0 + tv)$, è definita in un intorno di $t = 0$ ed ivi convessa. Inoltre la convessità di f in A è equivalente alla convessità di g per ogni scelta di x^0 e di v , nell'intervallo (dipendente da x^0 e v) in cui ogni g è definita. Per funzioni f di classe $C^1(A)$ o $C^2(A)$ rispettivamente dalle regole di derivazione abbiamo

$$g'(t) = v \cdot Df(x^0 + tv), \quad g''(t) = (v, Hf(x^0 + tv)v). \quad (4)$$

Applicando la Proposizione 4 e le (4) con $t = 1$ e $v = x - x^0$, $x \in A$ arbitrario, si ha

12. Proposizione Dato $A \subset \mathbb{R}^n$ convesso e aperto ed $f \in C^1(A)$ sono equivalenti

- (i) f è convessa;
- (ii) il grafico di f è al di sopra dei suoi piani tangenti ovvero $\forall x, x^0 \in A$ vale la disuguaglianza

$$f(x) \geq f(x^0) + Df(x^0)(x - x^0).$$

Invece, dalla Proposizione 5 e le (4) con $t = 1$ e $v = x - x^0$, $x \in A$ arbitrario, si ha

13. Proposizione Dato $A \subset \mathbb{R}^n$ convesso e aperto ed $f \in C^2(A)$ sono equivalenti

- (i) f è convessa;
- (ii) $Hf(x^0) \geq 0$ per ogni $x^0 \in A$.

14. Esempio Le funzioni $f(x) = e^{|x|^2}$, $f(x) = |x|^p$, $p > 1$, sono convesse in $\mathbb{R}^n$. Le funzioni $f(x, y) = x^2 - y^2$ e $f(x, y) = \log(x^2 + y^2)$, $(x, y) \neq (0, 0)$, non sono né concave né convesse.

15. Osservazione In base alle Proposizioni 12 e 13 per una funzione C^2 su una palla $B(x^0, r)$, quando $Df(x^0) = 0$, la convessità (concavità) di f nella palla è una condizione sufficiente perchè x^0 sia punto di minimo (massimo).

ESERCIZI

- 1 Provare che l'intersezione di convessi è sempre un convesso mentre l'unione non sempre lo è.
- 2 Provare che la chiusura di un convesso è un convesso.
- 3 Provare che se $C \subset \mathbb{R}^n$ è un convesso chiuso ed $x' \notin C$ allora esiste un unico *punto di distanza minima* $x'' \in C$ tale che $|x' - x''| = \text{dist}(x', C)$
- 4 Dati $a \in \mathbb{R}^n \setminus \{0\}$ e $b \in \mathbb{R}$, provare che un semispazio $S = \{x \in \mathbb{R}^n \mid a \cdot x \leq b\}$ è un insieme convesso.
- 5 Provare che, se $C \subset \mathbb{R}^n$ è un convesso chiuso, $x' \notin C$ ed x'' è la sua proiezione allora $C \subset S$ dove $S = \{(x - x'') \cdot (x' - x'') \leq 0\}$ è un semispazio chiuso. Dedurre che un convesso chiuso è l'intersezione dei semispazi chiusi che lo contengono.
- 6 Provare che $f : \mathbb{R}^n \rightarrow \mathbb{R}$ è convessa se e solo se il suo *epigrafico* $\text{epi}(f) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid y \geq f(x)\} \subset \mathbb{R}^{n+1}$ è convesso.
- 7 Date $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, convesse su Ω convesso, provare che la somma $f + g$ è una funzione convessa ed il massimo di f e g lo è. Provare che in generale né la differenza né in minimo in generale lo sono.
- 8 Data $g : \mathbb{R} \rightarrow \mathbb{R}$ convessa e crescente e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convessa, provare che la composizione $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}$ è convessa. Dare una dimostrazione alternativa nel caso che g ed f siano funzioni C^2 .
- 9 Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^2 che è sia concava che convessa. Provare che esistono $a \in \mathbb{R}^n$ e $b \in \mathbb{R}$ tali che $f(x) = a \cdot x + b$ per ogni $x \in \mathbb{R}^n$.

-
- 01 (Trasformata di Legendre) Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continua e tale che $\frac{f(x)}{|x|} \rightarrow +\infty$ per $|x| \rightarrow \infty$ (superlinearità all'infinito). Per ogni $p \in \mathbb{R}^n$ poniamo $f^*(p) = \max_{x \in \mathbb{R}^n} (p \cdot x - f(x))$. Provare che $f^* : \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione convessa.
 - 02 Sia $f(x) = \frac{|x|^p}{p}$, $p > 1$, oppure $f(x) = \frac{1}{2}Ax \cdot x$, A matrice simmetrica definita positiva. Calcolare $f^*(p)$ e verificare che la sua trasformata soddisfa $f^{**}(x) = f(x)$ per ogni $x \in \mathbb{R}^n$.

2-8 TEOREMI DELLA FUNZIONE INVERSA E DEL DINI

Il Teorema della funzione inversa è uno dei teoremi centrali dell'Analisi. Prima di enunciarlo dimostriamo due lemmi di preparazione:

1. Lemma (*Teorema del valor medio su $\mathbb{R}^n$*) Sia A un aperto di $\mathbb{R}^n$, $g : A \rightarrow \mathbb{R}$ di classe C^1 , e siano $x, y \in A$ tali che il segmento $[x, y]$ che unisce i punti x, y sia contenuto in A . Allora esiste $\xi \in [x, y]$ tale che

$$g(x) - g(y) = Dg(\xi)(x - y).$$

DIM. Restringiamo g alla retta che unisce x e y e consideriamo la funzione $\phi(t) = g(x + t(y - x))$. $\phi(t)$ è definita e C^1 per $t \in [0, 1]$, quindi per il Teorema del valor medio esiste $s \in (0, 1)$ tale che

$$\phi(1) - \phi(0) = \phi'(s).$$

Ponendo $\xi = x + s(y - x)$ e osservando che $\phi'(t) = Dg(x + t(y - x))(y - x)$, otteniamo la tesi.

2. Lemma Sia A un aperto di $\mathbb{R}^n$, $G : A \rightarrow \mathbb{R}^n$ una funzione di classe C^1 e $x^0 \in A$. Supponiamo che $DG(x^0) = 0$. Allora $\forall \epsilon > 0$ possiamo trovare una palla chiusa $B = \overline{B(x^0, r)} \subset A$ tale che per tutti i punti $x, x' \in B$ si ha

$$|G(x) - G(x')| \leq \epsilon |x - x'|.$$

DIM. Chiamiamo $B(r) = \overline{B(x^0, r)}$ la palla chiusa centrata in x^0 , che per r piccolo è contenuta in A . Se indichiamo con $m(r)$ il massimo di tutte le derivate prime di G su $B(r)$, dato che DG è continua e si annulla in x^0 , vediamo che $m(r) \rightarrow 0$ per $r \rightarrow 0$ (perché?)

Se $x, x' \in B(r)$, applicando il Teorema del valor medio alla componente G_1 di G possiamo trovare $\xi \in B$ tale che $G_1(x) - G_1(x') = DG_1(\xi)(x - x')$, e quindi otteniamo

$$|G_1(x) - G_1(x')| \leq |DG_1(\xi)| \cdot |x - x'| \leq \sqrt{n} \cdot m(r) \cdot |x - x'|$$

(dato che $|v|^2 = v_1^2 + \dots + v_n^2 \leq n \max_j v_j^2$). La stessa disuguaglianza vale per le altre componenti di G . Elevando tutte le disuguaglianze al quadrato e sommando otteniamo

$$|G(x) - G(x')|^2 \leq n^2 m(r)^2 |x - x'|^2 \quad \Rightarrow \quad |G(x) - G(x')| \leq nm(r) |x - x'|.$$

Infine, preso r abbastanza piccolo abbiamo $nm(r) \leq \epsilon$ e otteniamo la tesi.

Sappiamo che un'applicazione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è invertibile se e solo se il determinante della sua matrice è diverso da 0. Dato che $DL = L$, questo fatto si può enunciare dicendo che L è invertibile se e solo se il determinante della sua matrice Jacobiana è diverso da 0. Il teorema seguente è la versione "curva" di questa proprietà: se il determinante della Jacobiana di una funzione C^1 è diverso da 0 in un punto, allora la funzione è invertibile in un intorno di quel punto. Diversamente dal caso dell'applicazione lineare L non possiamo aspettarci che la funzione sia invertibile dappertutto, però possiamo trovare un intorno del punto su cui essa lo è: si dice anche che la funzione è *localmente invertibile* in quel punto.

3. Teorema (Teorema della funzione inversa, o di invertibilità locale) Sia A un aperto di $\mathbb{R}^n$ con $x^0 \in A$, sia $F : A \rightarrow \mathbb{R}^n$ una funzione di classe C^1 , e supponiamo che

$$\det DF(x^0) \neq 0.$$

Allora esistono un aperto U con $x^0 \in U \subseteq A$ e un aperto V in $\mathbb{R}^n$ tali che $F : U \rightarrow V$ è biunivoca, e F^{-1} è di classe C^1 .

DIM. Per ipotesi, la matrice $L = DF(x^0)$ è invertibile. Se definiamo la nuova funzione $\tilde{F} := L^{-1} \circ F$, vediamo subito che dimostrare il teorema per F o per $\tilde{F}$ è equivalente (infatti $F^{-1}(y) = \tilde{F}^{-1}(L^{-1}y)$). Dato che $D\tilde{F}(x^0) = L^{-1}L = I$, questo ragionamento ci dice che è sufficiente dimostrare il teorema nel caso speciale in cui $DF(x^0) = I$ è la matrice identica. Notiamo anche che $\det DF(x)$ è una funzione continua che non si annulla in $x = x^0$, quindi possiamo trovare un intorno $A' \subseteq A$ di x^0 sul quale $\det DF(x) \neq 0$.

Calcolare la funzione inversa di F vuol dire, fissato y , risolvere l'equazione $F(x) = y$, che si può riscrivere nella forma $y + x - F(x) = x$. Quindi basta trovare un punto fisso $G(x) = x$ per la nuova funzione $G(x) := y + x - F(x)$ ossia $G = y + I - F$ (che naturalmente dipende dal punto y fissato). Il vantaggio è che possiamo utilizzare il Teorema delle contrazioni.

Dato che $DG(x^0) = I + 0 - DF(x^0) = 0$, applicando il Lemma 2 con $\epsilon = 1/2$ possiamo trovare una palla chiusa $\overline{B(x^0, r)} \subset A'$ (notare che r è indipendente da y), tale che

$$|G(x) - G(x')| \leq \frac{1}{2}|x - x'| \quad \forall x, x' \in \overline{B(x^0, r)}. \quad (1)$$

Inoltre dato che $G(x) - x^0 = G(x) - G(x^0) + G(x^0) - x^0 = G(x) - G(x^0) + y - F(x^0)$ abbiamo

$$|G(x) - x^0| \leq \frac{1}{2}|x - x^0| + |y - F(x^0)|$$

e se $y \in B(F(x^0), r/2)$ otteniamo

$$|G(x) - x^0| < \frac{r}{2} + \frac{r}{2} = r \quad \forall x \in \overline{B(x^0, r)}.$$

Abbiamo così dimostrato che se $y \in B(F(x^0), r/2)$ la funzione G è una contrazione dalla palla chiusa $\overline{B(x^0, r)}$ in sé, la quale è uno spazio metrico completo. Ne segue che G ha un unico punto fisso x che quindi soddisfa $F(x) = y$, e $x \in B(x^0, r)$. In altri termini, la funzione F è una biiezione da U a V , dove $V = B(F(x^0), r/2)$ e $U = F^{-1}(V) \cap B(x^0, r)$.

Dimostriamo adesso che F^{-1} è continua. Da (1) abbiamo

$$\frac{1}{2}|x - x'| \geq |G(x) - G(x')| = |x - x' - F(x) + F(x')| \geq |x - x'| - |F(x) - F(x')|$$

per tutti gli $x, x' \in U$, e quindi

$$|x - x'| \leq 2|F(x) - F(x')| \quad \Rightarrow \quad |F^{-1}(y) - F^{-1}(y')| \leq 2|y - y'| \quad \forall y, y' \in V$$

ossia $F^{-1} : V \rightarrow U$ è Lipschitziana (di costante 2) e quindi continua.

Infine dimostriamo che F^{-1} è differenziabile, e che se $y = F(x)$ vale la formula $D(F^{-1})(y) = (DF(x))^{-1}$; notare che allora F^{-1} è automaticamente anche C^1 (perché se una matrice è una funzione continua anche la sua inversa è una funzione continua). Dunque, fissiamo un punto $y = F(x) \in V$, chiamiamo $M = DF(x)$, che sappiamo essere invertibile, e cerchiamo di dimostrare che tende a 0 il seguente rapporto quando $y' = F(x') \rightarrow y$:

$$\frac{F^{-1}(y') - F^{-1}(y) - M^{-1}(y' - y)}{|y' - y|} = M^{-1} \left(\frac{M(x' - x) - (F(x') - F(x))}{|x' - x|} \right) \cdot \frac{|F^{-1}(y') - F^{-1}(y)|}{|y' - y|}.$$

Notiamo che quando $y' \rightarrow y$ anche $x' \rightarrow x$ perché F^{-1} è continua; allora la frazione fra parentesi tende a 0 (perché F è differenziabile in x) e se applichiamo M^{-1} il limite fa ancora 0 (perché le applicazioni lineari sono continue): invece l'ultimo rapporto a destra è ≤ 2 perché la funzione F^{-1} è Lipschitziana di costante 2, e di conseguenza tutta l'espressione a secondo membro tende a 0.

Si noti che la formula $D(F^{-1})(F(x^0)) = (DF(x^0))^{-1}$ si può dimostrare facilmente anche derivando l'identità $F \circ F^{-1} = I$ e usando la regola di derivazione per funzioni composte (Teorema 2). Attenzione però: se F è C^1 con $\det DF(x^0) = 0$, la funzione F può anche essere invertibile in un intorno di x_0 , ma l'inversa F^{-1} sicuramente non è C^1 [altrimenti $DF \cdot D(F^{-1}) = I$, quindi DF sarebbe una matrice invertibile]. Per esempio $f(x) = x^3$ è biunivoca da $\mathbb{R}$ in $\mathbb{R}$ e C^1 , $Df(0)$ ossia $f'(0)$ è uguale a 0, e l'inversa $f^{-1}(x) = \sqrt[3]{x}$ non è differenziabile in $x = 0$.

Il Teorema dell'applicazione inversa consente di risolvere il seguente problema: risolvere il sistema

$$\begin{cases} F_1(x_1, \dots, x_n) = y_1 \\ F_2(x_1, \dots, x_n) = y_2 \\ \dots \\ F_n(x_1, \dots, x_n) = y_n \end{cases} \quad \text{ossia} \quad F(x) = y$$

e cioè calcolare $x_1, \dots, x_n$ in funzione di $y_1, \dots, y_n$. Naturalmente, il teorema non ci dice come calcolare esplicitamente i valori di x dati quelli di y ; il teorema ci garantisce solo che, se $\det DF \neq 0$ in un punto, allora per ogni valore di $y = (y_1, \dots, y_n)$ esiste uno e un solo valore di $x = (x_1, \dots, x_n)$ che risolve il sistema (in un intorno del punto considerato).

Nelle applicazioni si incontra spesso un problema apparentemente più omplicato: si ha un sistema di equazioni

$$\begin{cases} f_1(x_1, \dots, x_n; y_1, \dots, y_m) = z_1 \\ f_2(x_1, \dots, x_n; y_1, \dots, y_m) = z_2 \\ \dots \\ f_m(x_1, \dots, x_n; y_1, \dots, y_m) = z_m \end{cases} \quad \text{ossia} \quad f(x, y) = z$$

e si vuole sapere se è possibile ricavare le $y_1, \dots, y_m$ in funzione delle altre variabili; ad esempio, fissato z , se è possibile ricavare y come funzione di x . Dato che le equazioni sono m e anche le incognite sono m , questo dovrebbe essere possibile. Nel caso speciale in cui le f_j sono funzioni lineari, la teoria dei sistemi lineari ci dice che il sistema si può risolvere a patto che il minore $m \times m$ formato dalle ultime m colonne del sistema abbia determinante non nullo. Il teorema seguente estende questo risultato al caso di un sistema in cui le funzioni f_j sono funzioni regolari qualsiasi; la matrice da esaminare per sapere se il problema si può risolvere è, come al solito, la matrice Jacobiana di $f = (f_1, \dots, f_m)$.

Useremo qui la seguente notazione: abbiamo una funzione $f(x, y)$ che dipende dalle variabili $x = (x_1, \dots, x_n)$ e dalle $y = (y_1, \dots, y_m)$; la matrice Jacobiana di f è come al solito Df ; con $D_x f$, $D_y f$ indichiamo invece le prime n colonne e le ultime m colonne di Df , cioè le matrici delle derivate di f rispetto alle x e rispetto alle y .

4. Teorema (Teorema del Dini, o della funzione implicita) Sia $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ una funzione, di classe C^1 su un intorno del punto (x^0, y^0) . Supponiamo che $f(x^0, y^0) = z^0$ e che

$$\det D_y f(x^0, y^0) \neq 0.$$

Allora esistono $U \subseteq \mathbb{R}^n$ intorno aperto di x^0 , $A \subseteq \mathbb{R}^n \times \mathbb{R}^m$ intorno aperto (x^0, y^0) , e una funzione $g : U \rightarrow \mathbb{R}^m$ di classe C^1 , tali che per ogni $(x, y) \in A$ si ha

$$f(x, y) = z^0 \quad \Leftrightarrow \quad y = g(x).$$

DIM. Per ricondurre al Teorema dell'applicazione inversa, definiamo una nuova funzione $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ come $F(x, y) = (x, f(x, y))$. La matrice Jacobiana di F è

$$DF(x, y) = \begin{pmatrix} I & 0 \\ D_x f(x, y) & D_y f(x, y) \end{pmatrix}$$

e quindi $\det DF(x^0, y^0) = \det D_y f(x^0, y^0) \neq 0$. Possiamo applicare il Teorema della funzione inversa e otteniamo un intorno A di (x^0, y^0) e un intorno B di $F(x^0, y^0) = (x^0, f(x^0, y^0))$ tali che $F : A \rightarrow B$ è invertibile, con inversa C^1 . Ovviamente, restringendo B (e quindi restringendo anche A), possiamo prenderlo della forma speciale $U \times V$, con U intorno aperto di x^0 e W intorno aperto di $f(x^0, y^0) = z^0$; quindi $F^{-1} : U \times W \rightarrow A$. Per capire la struttura di F^{-1} separiamone le componenti: per $(x, z) \in U \times W$ scriviamo

$$F^{-1}(x, z) = (h(x, z), k(x, z))$$

dove h è a valori in $\mathbb{R}^n$ e k a valori in $\mathbb{R}^m$. Notiamo che

$$(x, z) = F(F^{-1}(x, z)) = F(h(x, z), k(x, z)) = (h(x, z), f(h(x, z), k(x, z)))$$

quindi le prime n componenti di F^{-1} sono $h(x, z) = x$ e possiamo scrivere più semplicemente

$$(x, z) = F(x, k(x, z)) = (x, f(x, k(x, z))) \quad \Rightarrow \quad z = f(x, k(x, z)) \quad \forall x \in U, z \in W.$$

Vediamo quindi che per concludere basta prendere $g(x) = k(x, z^0)$.

5. Osservazione Nella situazione del Teorema, si dice che la funzione $y = g(x)$ è *definita implicitamente* dalla condizione $f(x, y) = z^0$. Notare che abbiamo dimostrato un risultato più forte: abbiamo costruito una funzione $y = k(x, z)$ di classe C^1 , che risolve $f(x, y) = z$ al variare di z in un intorno di z^0 (e non solo per $z = z^0$ fissato).

6. Esempio Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ data da $f(x, y) = x^2 + y^2$. Vogliamo risolvere l'equazione $f(x, y) = 1$ cioè $x^2 + y^2 = 1$ rispetto a y . Il Teorema del Dini ci dice che questo è possibile in un intorno di un punto (x_0, y_0) con $f(x_0, y_0) = 1$, se $\frac{\partial f}{\partial y}(x_0, y_0) \equiv 2y_0 \neq 0$. Verifichiamo che cosa vuol dire questo geometricamente. Sappiamo bene che $f(x, y) = 0$ individua i punti della circonferenza di centro l'origine e raggio 1. Vogliamo descrivere questi punti con una relazione del tipo $y = \phi(x)$ (ossia "ricavare y rispetto a x "): ora è chiaro che questo non si può fare vicino ai due punti della circonferenza che hanno ordinata $y_0 = 0$, in quanto vicino a tali punti per ogni valore di x ci sono *due* valori di y che risolvono l'equazione: non si può "ricavare" y in nessun intorno del punto $(1, 0)$. Negli altri punti della circonferenza, invece, se ci restringiamo ad un intorno abbastanza piccolo, questo si può fare (anche esplicitamente: $y = \sqrt{1 - x^2}$ per i punti della semicirconferenza superiore e $y = -\sqrt{1 - x^2}$ per la semicirconferenza inferiore). Si noti però che $Df(x, y) = (2x, 2y)$ non si annulla mai; anche nei due punti suddetti, in cui $\partial f / \partial y = 0$, si ha pur sempre $\partial f / \partial x \neq 0$; quindi si può applicare Dini all'altra variabile e ricavare x in funzione di y (come è evidente geometricamente).

7. Esempio Più in generale: sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe C^1 che verifica le ipotesi del Teorema del Dini in (x_0, y_0) :

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

Allora esistono un intorno U di x_0 e uno V di y_0 tali che per $(x, y) \in U \times V$ i punti in cui $f(x, y) = 0$ si possono scrivere nella forma $y = \phi(x)$, con $\phi : U \rightarrow V$ funzione C^1 :

$$f(x, y) = 0 \quad \Leftrightarrow \quad y = \phi(x).$$

Si usa dire che la funzione ϕ è *definita implicitamente dalla funzione f* , e in effetti non possiamo calcolare ϕ esplicitamente, ma sappiamo solo che essa soddisfa l'identità

$f(x, \phi(x)) = 0$. Quanto fa la derivata $\phi'(x)$? Derivando $f(x, \phi(x)) = 0$ rispetto a x si ha

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

e ricavando $\phi'(x)$ (dato che per ipotesi $\partial f/\partial y$ non si annulla in (x_0, y_0) , non si annulla neppure in un intorno di tale punto) si ottiene

$$\phi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \phi(x))}{\frac{\partial f}{\partial y}(x, \phi(x))};$$

notare che questa relazione ci consente di calcolare effettivamente $\phi'(x)$ in un punto (x, y) in cui f si annulla (cioè un punto della forma $(x, \phi(x))$).

Analogamente, se $f(x, y, z)$ è C^1 , $f(x_0, y_0, z_0) = 0$ e $\frac{\partial f}{\partial z}(x_0, y_0, z_0) \neq 0$, allora per (x, y) vicino a (x_0, y_0) è definita una funzione $z = \phi(x, y)$ con la proprietà che, per i punti (x, y, z) in un intorno di (x_0, y_0, z_0) ,

$$f(x, y, z) = 0 \quad \Leftrightarrow \quad z = \phi(x, y).$$

Vale dunque l'identità $f(x, y, \phi(x, y)) = 0$; derivandola rispetto a x e y si ottiene

$$\frac{\partial \phi}{\partial x}(x, y) = -\frac{\frac{\partial f}{\partial x}(x, y, \phi(x, y))}{\frac{\partial f}{\partial z}(x, y, \phi(x, y))}, \quad \frac{\partial \phi}{\partial y}(x, y) = -\frac{\frac{\partial f}{\partial y}(x, y, \phi(x, y))}{\frac{\partial f}{\partial z}(x, y, \phi(x, y))};$$

In generale, se la funzione $y = g(x)$ è definita implicitamente da $f(x, y) = z^0$ come nell'enunciato del Teorema del Dini, derivando l'identità $f(x, g(x)) = z^0$ rispetto a x otteniamo

$$D_x f(x, g(x)) + D_y f(x, g(x))D_x g(x) = 0$$

da cui la formula generale

$$D_x g(x) = -(D_y f(x, g(x)))^{-1} D_x f(x, g(x)).$$

ESERCIZI

- 1 Un'applicazione lineare da $\mathbb{R}^n$ in $\mathbb{R}^m$ è continua.
- 2 $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ lineare invertibile manda aperti in aperti [L^{-1} è continua].
- 3 Se $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è lineare, $DL(x_0) = L$ per ogni x_0 [dalla definizione].
- 4 Sia $y = \phi(x)$ definita implicitamente da $f(x, y) = 0$, dove

$$f(x, y) = x^3 + y^3 - 2xy.$$

Calcolare $\phi'(1)$, osservando che $f(1, 1) = 0$, e dopo aver verificato che le ipotesi del Teorema del Dini sono soddisfatte in tale punto. Analogamente, data $y = \phi(x)$ definita da $f(x, y) = 0$ con

$$f(x, y) = e^{x+y} - x^{10} - e,$$

calcolare $\phi'(0)$ (osservando che $f(0, 1) = 0$).

34 Calcolo differenziale

- 5 Siano $F(x, y, z)$, $G(x, y, z)$ di classe C^1 e supponiamo che si annullino in (x_0, y_0, z_0) . Supponiamo che la Jacobiana di F, G rispetto a y, z

$$\frac{\partial(F, G)}{\partial(y, z)} = \begin{pmatrix} D_2F & D_3F \\ D_2G & D_3G \end{pmatrix}$$

abbia determinante non nullo nel punto. Allora per il Teorema del Dini sono definite implicitamente due funzioni $y = \phi(x)$, $z = \psi(x)$ equivalenti al sistema $F = 0, G = 0$. Calcolare ϕ', ψ' .

- 01 Vale ancora il Teorema del valor medio per $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$? [No. Basta prendere una curva $\phi : [0, 1] \rightarrow \mathbb{R}^2$ tale che $\phi'(t) \neq 0$ ma $\phi(0) = \phi(1)$, ad esempio $\phi(t) = (\cos 2\pi t, \sin 2\pi t)$].
- 02 Dimostrare la seguente forma del Teorema del valor medio. Sia $g : A \rightarrow \mathbb{R}^m$, A aperto di $\mathbb{R}^n$, $[x, y]$ segmento contenuto in A . Fissiamo $v \in \mathbb{R}^m$. Allora esiste $\xi \in [x, y]$ (dipendente da v) tale che $v \cdot (f(x) - f(y)) = v \cdot df(\xi)(x - y)$.
- 03 Sia $B = B(0, 1) \subseteq \mathbb{R}^n$, $f : B \rightarrow \mathbb{R}$ tale che $D_v f(x) = 0$ per ogni x in B e per ogni direzione v . Allora f è costante.
- 04 Siano f, g come nel Teorema del Dini. Allora vale la relazione

$$Dg(x) = -[D_y f(x, g(x))]^{-1} D_x f(x, g(x)).$$

- 05 Determinare in quali punti di $\mathbb{R}^2$ sono localmente invertibili le funzioni $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ date da

$$\begin{aligned} F(x, y) &= (x + y, x^2 + y^2), & F(x, y) &= (xy, x^2 + y^2) \\ F(x, y) &= (xe^y, ye^x), & F(x, y) &= (x + \cos y, y + \cos x). \end{aligned}$$

2-9 CURVE E SUPERFICI

Ritorniamo sul tema delle curve per uno studio più approfondito. Cos'è una curva? per esempio, cos'è la circonferenza unitaria C in $\mathbb{R}^2$? Possiamo descriverla in tre modi. In *forma parametrica*: C è l'immagine dell'arco

$$\phi : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \phi(t) = (\cos t, \sin t).$$

In *forma implicita*: C è il luogo degli zeri di $f(x, y) = x^2 + y^2 - 1$

$$C = \{(x, y) : x^2 + y^2 - 1 = 0\}.$$

In *forma cartesiana*: C è il grafico di

$$f : [-1, 1] \rightarrow \mathbb{R}, \quad y = f(x) = \sqrt{1 - x^2}$$

(naturalmente questo descrive solo il semicerchio superiore, per quello inferiore dobbiamo usare $f(x) = -\sqrt{1 - x^2}$). Vedremo che tutte le curve (e più in generale tutte le superfici) ammettono queste tre descrizioni, e che sotto opportune ipotesi le tre descrizioni sono *localmente equivalenti*. Cominciamo ripassando le basi.

1. Definizione Una *curva* (in *forma parametrica*) in $\mathbb{R}^n$ è un'applicazione continua $\phi \equiv (\phi_1, \dots, \phi_n) : I \rightarrow \mathbb{R}^n$ da un intervallo $I \subseteq \mathbb{R}$ a valori in $\mathbb{R}^n$. $\phi_1, \dots, \phi_n$ si dicono *componenti* della curva, mentre l'immagine di ϕ si dice *sostegno* della curva; spesso il sostegno di ϕ si indica ancora con ϕ . La curva si dice *regolare* se ϕ è C^1 in I e $\phi' \neq 0$ in $\overset{\circ}{I}$. Si dice *semplice* se per $t \neq s$, $t, s \in \overset{\circ}{I}$, si ha $\phi(t) \neq \phi(s)$. Se $I = [a, b]$, $\phi(a)$ e $\phi(b)$ si dicono *punto iniziale* e *punto finale* della curva (*estremi*), e se coincidono la curva si dice *chiusa*.

2. Esempio a) La funzione $\phi(t) \equiv (0, 0)$ è continua da $\mathbb{R}$ in $\mathbb{R}^2$, quindi è una curva; ma la sua immagine è il solo punto $(0, 0)$. Se invece imponiamo che ϕ sia regolare, allora otteniamo una curva nel senso intuitivo del termine.

b) $\phi : \mathbb{R} \rightarrow \mathbb{R}^2$ data da $\phi(t) = (\cos t, \sin t)$ descrive la circonferenza unitaria con centro nell'origine. Curva regolare, non semplice; diventa regolare, semplice, chiusa se la restringiamo a $\phi : [0, 2\pi] \rightarrow \mathbb{R}^2$.

c) $\phi : \mathbb{R} \rightarrow \mathbb{R}^n$, $\phi(t) = At + B$ con A, B vettori fissati di $\mathbb{R}^n$, rappresenta una retta in forma parametrica in $\mathbb{R}^n$, passante per B e parallela al vettore A . Curva regolare e semplice.

d) $\phi : [0, 2\pi] \rightarrow \mathbb{R}^2$, $\phi(t) = (\cos(t) \cos(3t), \sin(t) \cos(3t))$ è chiusa, regolare, non semplice (trifoglio).

e) $\phi : [-1, 1] \rightarrow \mathbb{R}^2$, $\phi(t) = (t^2, t^3)$ è C^1 , semplice, ma non è regolare (perché?). Per $t = 0$ si ha una *cuspid*e (parabola semicubica).

f) $\phi : [-1, 1] \rightarrow \mathbb{R}^2$, $\phi(t) = (t^2, t^2)$. Si tratta del segmento che unisce $(0, 0)$ con $(1, 1)$, percorso due volte. Non regolare (perché?), non semplice, chiusa.

g) $\phi : \mathbb{R} \rightarrow \mathbb{R}^3$, $\phi(t) = (r \cos t, r \sin t, kt)$. Regolare, semplice (elica cilindrica di raggio r e passo k).

h) $\phi : [-1, 1] \rightarrow \mathbb{R}^2$, $\phi(t) = (|t|, t)$. Non regolare (perché non è C^1), semplice.

i) $\phi : [-2, 2] \setminus \{-1\} \rightarrow \mathbb{R}^2$, $\phi(t) = \left(\frac{3t}{1+t^3}, \frac{3t^2}{1+t^3}\right)$, ($t \neq -1$). C^1 , regolare, semplice (attenzione!), detta folium di Cartesio.

j) Si può costruire una curva continua $\phi : [0, 1] \rightarrow \mathbb{R}^2$ la cui immagine è tutto il quadrato $[0, 1] \times [0, 1]$ (Peano). Vediamo quindi che è necessaria qualche ipotesi in più per essere sicuri che una curva abbia un aspetto “normale”; l'ipotesi $\phi \in C^1$ serve proprio a questo.

3. Esempio Data una curva $\phi : I \rightarrow \mathbb{R}^n$ di classe C^1 , I intervallo, e $t_0 \in I$, l'equazione parametrica $x(t) = \phi(t_0) + \phi'(t_0)(t - t_0)$ descrive la *retta tangente* alla curva ϕ nel punto $\phi(t_0)$.

Se $\phi : I \rightarrow \mathbb{R}^2$, e una delle componenti è l'identità, allora la curva è il grafico cartesiano di una funzione. Ad esempio se $\phi(t) = (t, \phi_2(t))$, allora (il sostegno di) ϕ è il grafico di $\phi_2 : I \rightarrow \mathbb{R}$. In questo caso diremo che la curva è *rappresentata in forma cartesiana*. Ad esempio la funzione $f(x) = x^2$ ha come grafico una parabola; diremo che essa è rappresentata in forma cartesiana come $y = x^2$. In forma parametrica essa è data dalla curva $\phi(t) = (t, t^2)$. Notiamo che una curva C^1 in forma cartesiana è sempre semplice e regolare.

Si osservi che due curve possono avere lo stesso sostegno ma essere formalmente diverse: ad es. $\phi : [0, 1] \rightarrow \mathbb{R}^2$ ha lo stesso sostegno di $\psi : [0, 2] \rightarrow \mathbb{R}^2$ data da $\psi(t) = \phi(t/2)$.

4. Definizione Dati I, J intervalli di $\mathbb{R}$, un *cambiamento di parametro* è un'applicazione $\tau : I \rightarrow J$ di classe C^1 , biunivoca, con inversa C^1 . Ne segue in particolare che τ' non si annulla mai, quindi o $\tau' > 0$, e allora si dice che τ *conserva il verso*, o $\tau' < 0$ e si dice che *inverte il verso*. Due curve $\phi : I \rightarrow \mathbb{R}^n$, $\psi : J \rightarrow \mathbb{R}^n$ tali che $\phi(t) = \psi(\tau(t))$ si dicono *equivalenti* se $\tau' > 0$.

5. Esempio Se $\phi : [a, b] \rightarrow \mathbb{R}^2$ è una curva regolare semplice e $t_0 \in]a, b[$, allora per t

vicino a t_0 è sempre possibile esprimerla in forma cartesiana. Infatti essendo $\phi'(t_0) \neq 0$ dev'essere ad esempio $\phi'_1(t_0) \neq 0$, quindi $\phi'_1(t) \neq 0$ per t in un intorno $I_\epsilon =]t_0 - \epsilon, t_0 + \epsilon[$ di t_0 . Ne segue che $\phi_1 : I_\epsilon \rightarrow \mathbb{R}$ è invertibile (strettamente crescente o decrescente). Allora $\tau(t) = \phi_1^{-1}(t)$ è un cambiamento di parametro, e la curva ϕ è equivalente alla curva

$$\phi(\tau(t)) = (t, \phi_2(\phi_1^{-1}(t))).$$

Abbiamo cioè rappresentato ϕ in forma cartesiana come $y = \phi_2(\phi_1^{-1}(x))$, in un intorno di t_0 . (Se invece $\phi'_2(t_0) \neq 0$, allora potremo rappresentare ϕ nella forma cartesiana $x = g(y) = \phi_1(\phi_2^{-1}(y))$).

6. Esempio Spesso per definire una curva si utilizza la *forma polare*, ossia la curva è descritta in coordinate polari, assegnando per ogni valore dell'angolo θ il corrispondente valore del raggio vettore $\rho(\theta)$. Si tratta sempre di un caso particolare della forma parametrica, in cui si prende come parametro l'angolo θ e si ha $\phi(\theta) = (\rho(\theta) \cos \theta, \rho(\theta) \sin \theta)$.

Un modo molto comune di definire una curva, oltre alla forma parametrica, che ha come casi particolari le forme cartesiana e polare, è l'espressione in *forma implicita*: ossia come luogo di zeri di una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Ad esempio se $f(x, y) = x^2 + y^2$, l'equazione $f(x, y) = 1$ descrive chiaramente la circonferenza unitaria di centro l'origine, ossia l'insieme $C = \{(x, y) : x^2 + y^2 = 1\}$.

Ma in generale il luogo di zeri $C = \{f(x, y) = 0\}$ di una funzione non è una curva; ad esempio se f è la funzione nulla allora C è tutto $\mathbb{R}^2$, se $f = x^2 + y^2 + 1$ allora C è l'insieme vuoto, se $f = x^2 + y^2$ allora C è il solo punto $(0, 0)$. Il Teorema del Dini dà una condizione sufficiente perché C sia localmente una curva. Infatti, se f è C^1 , $(x_0, y_0) \in C$ ossia $f(x_0, y_0) = 0$, e inoltre $Df(x_0, y_0) \neq 0$, ad esempio $\partial f / \partial y \neq 0$, allora in un intorno di (x_0, y_0) l'insieme $C = \{f = 0\}$ si può rappresentare come $y = \phi(x)$, ossia è una curva in forma cartesiana di classe C^1 , quindi in particolare è semplice e regolare (argomento simile se $D_1 f \neq 0$).

In conclusione, data $f(x, y)$ di classe C^1 in un intorno di (x_0, y_0) in cui $f(x_0, y_0) = 0$ ma $Df(x_0, y_0) \neq 0$, diremo che $f = 0$ definisce una *curva in forma implicita* in un intorno di tale punto.

7. Esempio (Minicorso di geometria differenziale delle curve in tre dimensioni). Sia $\phi :]-1, 1[\rightarrow \mathbb{R}^3$ una curva regolare di classe C^3 . Un parametro privilegiato è la *lunghezza d'arco* $s(t) = \int_0^t |\phi'(\sigma)| d\sigma$; vedremo successivamente che $s(t)$ esprime la lunghezza dell'arco di curva tra $\phi(0)$ e $\phi(t)$. Si ha $s'(t) > 0$, quindi $s(t)$ conserva il verso; ponendo $\tilde{\phi}(\sigma) = \phi(s^{-1}(\sigma))$ otteniamo una curva equivalente a ϕ con lo stesso verso. Notiamo che da $\tilde{\phi}(s(t)) = \phi(t)$ segue $\tilde{\phi}'(s(t)) \cdot s'(t) = \phi'(t)$, e dato che $s'(t) = |\phi'(t)|$ otteniamo $|\tilde{\phi}'(s(t))| = 1$ cioè $|\tilde{\phi}'(s)| = 1$. In altri termini se ϕ è parametrizzata secondo la lunghezza d'arco, il modulo di ϕ' è costantemente uguale ad uno. Nel seguito supporremo che ϕ sia una curva C^3 (regolare, semplice) parametrizzata con la lunghezza d'arco.

Derivando $|\phi'(s)|^2 = 1$ si ha $\phi''(s) \cdot \phi'(s) = 0$, ossia i vettori $\phi''(s)$ e $\phi'(s)$ sono ortogonali. Possiamo ora ruotare e traslare gli assi di $\mathbb{R}^3$ (il che non cambia la forma della curva) portando l'origine nel punto $\phi(0)$, l'asse x parallelo a $\phi'(0)$ e l'asse y parallelo

a $\phi''(0)$ (supponendo per semplicità $\phi''(0) \neq 0$). Abbiamo così $\phi(0) = 0$, $\phi'_2(0) = \phi'_3(0) = \phi''_1(0) = \phi''_3(0) = 0$. Se sviluppiamo le componenti di ϕ in polinomio di Taylor, le prime due al secondo ordine e la terza al terzo ordine, otteniamo

$$\begin{aligned}\phi_1(s) &= \phi'_1(0)s + R_2(s), \\ \phi_2(s) &= \frac{1}{2}\phi''_2(0)s^2 + R_2(s), \\ \phi_3(s) &= \frac{1}{6}\phi'''_3(0)s^3 + R_3(s).\end{aligned}$$

Infine notiamo che $\phi'_1(0) = |\phi'(0)| = 1$, poniamo $\kappa = \phi''_2(0)$ e otteniamo il seguente risultato: in un opportuno sistema di riferimento, e con la scelta del parametro d'arco, una curva regolare semplice C^3 si scrive localmente come

$$\begin{aligned}\phi_1(s) &= s + R_2(s), \\ \phi_2(s) &= \frac{1}{2}\kappa s^2 + R_2(s), \\ \phi_3(s) &= \frac{1}{6}\phi'''_3(0)s^3 + R_3(s).\end{aligned}$$

La quantità $\kappa = \phi''_2(0)$ si dice *curvatura* di ϕ in 0. Se la curvatura è diversa da 0, la quantità $\tau = \phi'''_3(0)/\kappa$ si dice *torsione* di ϕ in 0.

Quanto detto per le curve si generalizza facilmente alle superfici: una superficie in forma parametrica in $\mathbb{R}^n$ è un'applicazione $\phi : A \rightarrow \mathbb{R}^n$, A aperto di $\mathbb{R}^\ell$, $\ell < n$. Come nel caso delle curve regolari, dobbiamo imporre delle condizioni su ϕ perché l'immagine di ϕ sia una vera superficie nel senso intuitivo del termine (ad es. se le funzioni $\phi_1, \dots, \phi_n$ sono identicamente nulle, l'immagine di ϕ è $\{(0, 0)\}$).

8. Definizione Una *superficie regolare* di dimensione ℓ in $\mathbb{R}^n$ ($\ell < n$), in *forma parametrica*, è un'applicazione $\phi : A \rightarrow \mathbb{R}^n$, A aperto di $\mathbb{R}^\ell$, di classe C^1 , tale che la matrice jacobiana $D\phi$ ha rango ℓ in ogni punto di A . La superficie si dice *semplice* se $\phi : A \rightarrow \mathbb{R}^n$ è iniettiva.

Se $t_0 \in A$, il *piano tangente* alla superficie nel punto $\phi(t_0)$ si scrive in forma parametrica come $x = \phi(t_0) + D\phi(t_0)(t - t_0)$.

Ricordiamo che una matrice ha rango ℓ se esiste un minore quadrato di lato ℓ con determinante non nullo, mentre tutti i minori con dimensione maggiore di ℓ hanno determinante nullo. Nel caso precedente $D\phi$ è una matrice $n \times \ell$, quindi il suo rango è il massimo possibile. Notare che il piano tangente risulta un piano di dimensione ℓ .

9. Osservazione 1) Una curva regolare $\phi : (a, b) \rightarrow \mathbb{R}^n$ è una superficie regolare di dimensione 1.

2) Una superficie in *forma cartesiana* in $\mathbb{R}^n$, è il grafico di una funzione C^1 $f : A \rightarrow \mathbb{R}^m$, A aperto di $\mathbb{R}^\ell$, dove $n = \ell + m$. Tale superficie si può vedere come una superficie

in forma parametrica $\phi : A \rightarrow \mathbb{R}^n$ dove

$$\phi(t_1, \dots, t_\ell) = (t_1, \dots, t_\ell, f_1(t), \dots, f_m(t)) \quad (1)$$

ossia $\phi(t) = (t, f(t))$, quindi la forma cartesiana è un caso particolare di quella parametrica. Notiamo che il rango di ϕ è sempre ℓ , perché non può essere maggiore di ℓ e inoltre

$$D\phi = \begin{pmatrix} I \\ Df \end{pmatrix}$$

e il minore $\ell \times \ell$ delle prime ℓ righe ha determinante 1, quindi la superficie è automaticamente una superficie regolare; chiaramente, essa è anche semplice.

10. Osservazione Sia S una superficie regolare semplice di dimensione ℓ descritta in forma parametrica come $\phi : A \rightarrow \mathbb{R}^n$, $A \subseteq \mathbb{R}^\ell$ aperto, $x^0 = \phi(t^0)$ un suo punto con $t^0 \in A$. Anche in questo caso è possibile esprimere S in un intorno di x^0 in forma cartesiana, ossia si possono spezzare le variabili x in due gruppi y, z di ℓ e $m = n - \ell$ variabili, e costruire una funzione $h : \mathbb{R}^\ell \rightarrow \mathbb{R}^m$ di classe C^1 tale che in un intorno di x^0 S è descritta dall'equazione $z = h(y)$. [La semplice dimostrazione è identica al caso delle curve: dato che $D\phi(t^0)$ ha rango massimo ℓ , possiamo trovare ℓ righe il cui determinante è diverso da 0, ad esempio le prime ℓ . Scriviamo allora $\psi = (\phi_1, \dots, \phi_\ell)$ e $\chi = (\phi_{\ell+1}, \dots, \phi_n)$ ossia $\phi = (\psi, \chi)$. Allora l'applicazione ψ da $\mathbb{R}^\ell$ in $\mathbb{R}^\ell$ è localmente invertibile per il Teorema dell'applicazione inversa, e se poniamo $h = \chi \circ \psi^{-1}$ vediamo subito che S si può descrivere come $z = h(y)$.]

11. Definizione Anche una superficie si può descrivere in forma implicita. Se A è un aperto di $\mathbb{R}^n$, una *superficie regolare* di dimensione ℓ in $\mathbb{R}^n$ ($\ell < n$), *in forma implicita*, è il luogo di zeri $S = \{x \in A : F(x) = 0\}$ di una funzione $F : A \rightarrow \mathbb{R}^{n-\ell}$ di classe C^1 , con $DF(x)$ di rango $n - \ell$ (cioè massimo) in tutti i punti di S .

Notiamo subito che grazie al Teorema del Dini si può sempre passare dalla forma implicita a quella cartesiana, almeno localmente: infatti fissato $x^0 \in S$, dall'ipotesi segue subito che possiamo spezzare le variabili in due gruppi (y, z) di ℓ e $m = n - \ell$ variabili tali che $D_z F(x^0)$ ha determinante diverso da 0; e quindi esiste una funzione $z = h(y)$ di classe C^1 tale che S si può descrivere come $z = h(y)$ in un intorno di x^0 . In particolare, una superficie descritta in forma implicita è sempre regolare semplice, e la dimensione di S è ℓ , ossia $n - \text{rango di } DF$.

12. Osservazione Riassumendo, abbiamo visto che di una curva o una superficie si possono dare tre definizioni diverse:

- (1) forma parametrica;
- (2) forma cartesiana;
- (3) forma implicita.

Abbiamo però visto che, in sostanza ossia localmente, le tre definizioni sono equivalenti.

Nella definizione di superficie in forma parametrica, notiamo che non è essenziale prendere ϕ a valori in $\mathbb{R}^n$; avremmo potuto considerare una qualunque $\phi : A \rightarrow S$ definita su un aperto $A \subseteq \mathbb{R}^\ell$ a valori in uno spazio topologico S , che sia un omeomorfismo. In altri termini si può studiare la “superficie” S dimenticando l'esistenza dei “punti circostanti” di $\mathbb{R}^n$; la “dimensione” di S è ancora ℓ , per studiare una funzione $f : S \rightarrow \mathbb{R}$ ci si riconduce

allo studio della funzione composta $f \circ \phi : A \rightarrow \mathbb{R}$, e così via. Quest'idea, opportunamente sviluppata, è alla base del concetto di varietà astratta studiato in Geometria Differenziale. Le superfici da noi considerate sono un caso particolare, quello delle varietà *immerse* in $\mathbb{R}^n$ (ma neanche troppo particolare, infatti un teorema di Whitney afferma che ogni varietà differenziabile di dimensione ℓ si può vedere come una varietà immersa in $\mathbb{R}^{2\ell+1}$).

ESERCIZI

- 1 Sia C la curva in $\mathbb{R}^3$ intersezione dei cilindri $\{y^2 + z^2 = r^2\}$, $\{x^2 + z^2 = R^2\}$ ($r, R > 0$). Per quali r, R essa è regolare? [C è definita in forma implicita da $\{F = 0\}$, dove $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ è $F(x, y, z) = (y^2 + z^2 - r^2, x^2 + z^2 - R^2)$. Si ha

$$DF = \begin{pmatrix} 0 & 2y & 2z \\ 2x & 0 & 2z \end{pmatrix}$$

che ha rango 2 tranne che in $(0, 0, z)$, $(0, y, 0)$, $(x, 0, 0)$. I punti del secondo e terzo tipo non appartengono a C . I primi appartengono a C solo se $r = R$. Quindi se $r \neq R$, C è una curva regolare. Se $r = R$ i punti $(0, 0, \pm R)$ sono singolari: le curve $(t, \pm t, \pm \sqrt{R^2 - t^2})$ sono contenute in C e passano per tutti e due i punti, quindi sicuramente in nessun intorno di tali punti è possibile descrivere C come una curva regolare.]

- 2 Data $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, una *curva di livello* di f è un insieme del tipo $\{f(x, y) = k\}$, per qualche costante k . Dire per quali valori di k sono regolari le curve di livello di $f(x, y) = (x^2 - 1)^2 + (y^2 + 1)^2$.
- 3 Un *toro* in $\mathbb{R}^3$ è la superficie definita in forma implicita da

$$(\sqrt{x^2 + y^2} - R)^2 + z^2 = r^2.$$

Per quali valori di r, R si ha una superficie regolare di dimensione 2?

- 4 Sia $C \subseteq \mathbb{R}^4$ la superficie in forma implicita

$$C = \{(x, y, u, v) : (x - u)^2 + (y - v)^2 = 1\}.$$

Dire se C è regolare e di che dimensione. Stessa domanda per

$$C = \{(x, y, u, v) : (x - u)^2 + (y - v)^2 = 1, y = 0\}$$

$$C = \{(x, y, u, v) : (x - u)^2 + (y - v)^2 = 1, x^2 + y^2 = k^2\}$$

$$C = \{(x, y, u, v) : (x - u)^2 + (y - v)^2 = 1, x^2 + y^2 = k^2, v = 0\}.$$

- 5 Sia S superficie regolare di dimensione ℓ in $\mathbb{R}^n$, descritta in un intorno di un suo punto x_0 in forma implicita o in forma parametrica. Dimostrare che esiste un'applicazione G da un intorno di x_0 a valori in $\mathbb{R}^n$, localmente invertibile, tale che $G(S)$ è il piano coordinato $y_{\ell+1} = \dots = y_n = 0$. Ossia, cambiando coordinate è possibile *spianare* localmente una superficie regolare.

[Usiamo la notazione $x = (x', x'')$ per indicare che abbiamo spezzato le variabili in due gruppi di ℓ e $m = n - \ell$ variabili. Dopo aver espresso localmente la S in forma cartesiana come $x'' = h(x')$, è sufficiente porre $G(x) = (x', x' - h(x''))$.]

2-10 MASSIMI E MINIMI VINCOLATI

I massimi e minimi di una funzione continua e derivabile $f : [a, b] \rightarrow \mathbb{R}$ possono essere o interni all'intervallo $[a, b]$, nel qual caso devono essere punti stazionari (in cui si annulla la derivata), oppure possono essere al bordo dell'intervallo; il secondo caso si studia molto facilmente perché basta calcolare i valori $f(a)$ e $f(b)$ e confrontarli con i valori di f nel resto dell'intervallo.

La situazione per una $f : K \rightarrow \mathbb{R}$ quando K è un chiuso di $\mathbb{R}^n$ naturalmente è più complicata. Eventuali massimi e minimi interni a K vanno cercati fra i punti stazionari, utilizzando le tecniche studiate nel Paragrafo 6. Ma per completare lo studio ci serve un metodo per calcolare massimi e minimi di f ristretta al bordo di K , che può essere una superficie. Introduciamo una terminologia appropriata per questa situazione:

1. Definizione Sia S una superficie regolare in $\mathbb{R}^n$ di dimensione $\ell < n$, U un intorno di $x^0 \in S$, e $f : U \rightarrow \mathbb{R}$ una funzione di classe C^1 . Si dice che $x_0 \in S$ è un punto di *massimo per f su S* , o *vincolato ad S* , se per ogni $x \in S$ appartenente ad un opportuno intorno di x^0 vale $f(x^0) \geq f(x)$. La superficie S è detta anche *vincolo* del problema di massimo. Analoga la definizione di minimo vincolato.

Se il vincolo S è descritto in forma parametrica come immagine di una funzione $\phi(t_1, \dots, t_\ell)$ da un aperto $A \subset \mathbb{R}^\ell$ a valori in $\mathbb{R}^n$, trovare i massimi e minimi vincolati è piuttosto semplice: basta studiare la funzione composta $f \circ \phi : A \rightarrow \mathbb{R}$ con i soliti metodi, ossia basta cercare i punti stazionari di $f \circ \phi$ e studiarne il carattere. Vediamo un esempio elementare.

2. Esempio Trovare massimi e minimi di $f : K \rightarrow \mathbb{R}$ dove K è il disco unitario di $\mathbb{R}^2$ e $f(x, y) = xy$. Cerchiamo anzitutto i punti stazionari interni al disco:

$$D_1 f(x, y) = D_2 f(x, y) = 0 \quad \Leftrightarrow \quad x = y = 0.$$

L'unico punto stazionario è l'origine che però è un punto di sella. Quindi massimo e minimo (che devono esistere per Weierstraß) si trovano sul bordo ∂K ossia sulla circonferenza unitaria. Come facciamo a trovarli?

Fortunatamente conosciamo una semplice parametrizzazione della circonferenza unitaria come curva $\phi(t) = (\cos t, \sin t)$ per $t \in [0, 2\pi]$. Se restringiamo f alla curva ϕ vediamo che basta studiare la funzione $g(t) = f(\phi(t)) = \cos t \cdot \sin t$ per $t \in \mathbb{R}$. La derivata $g'(t) = \cos^2 t - \sin^2 t = \cos(2t)$ si annulla per $t = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}$ (gli ulteriori valori di t ridanno gli stessi quattro punti). Dato che $g(\frac{\pi}{4}) = g(\frac{5\pi}{4}) = 1$ mentre $g(\frac{3\pi}{4}) = g(\frac{7\pi}{4}) = -1$, concludiamo che i primi sono punti di massimo per f su K mentre gli altri due sono minimi.

Se invece il vincolo S è dato in forma implicita come luogo di zeri di una funzione F , e non abbiamo una parametrizzazione semplice di S , si può ricorrere al *metodo dei moltiplicatori di Lagrange*. Per descrivere il metodo sono utili le nozioni di *vettore tangente* e *vettore normale* a una superficie.

3. Definizione Sia $S \subset \mathbb{R}^n$ una superficie regolare di dimensione $\ell < n$ e sia $x^0 \in S$. Diciamo che un vettore $v \in \mathbb{R}^n$ è *tangente* a S in x^0 se esiste una curva $\gamma : (-\epsilon, \epsilon) \rightarrow S$ di classe C^1 tale che $\gamma(0) = x^0$ e $\gamma'(0) = v$ (ossia una curva tutta contenuta in S passante per x^0 nella direzione v). Diciamo che un vettore $w \in \mathbb{R}^n$ è *normale* a S in x^0 se è ortogonale a tutti i vettori tangenti a S in x^0 .

4. Osservazione È facile dimostrare che l'insieme di tutti i vettori tangenti a una superficie regolare S di dimensione ℓ in un punto forma uno spazio vettoriale di dimensione ℓ , detto lo *spazio tangente* a S in x^0 [Infatti basta scrivere la superficie in forma parametrica come immagine di una $\psi : A \rightarrow \mathbb{R}^n$, A aperto di $\mathbb{R}^\ell$, e osservare che tutte le curve

contenute in S passanti per x^0 si possono scrivere nella forma $\psi \circ \gamma$ dove $\gamma : (-\epsilon, \epsilon) \rightarrow A$ è una curva contenuta in A ; quindi l'insieme dei vettori tangenti coincide con l'immagine dell'applicazione lineare $D\psi(x^0)$ che per ipotesi ha rango ℓ . Di conseguenza, anche i vettori normali a S in x^0 formano uno spazio vettoriale, di dimensione $n - \ell$.

Tuttavia qui siamo più interessati a una caratterizzazione concreta dei vettori tangenti e normali per superfici definite in forma implicita.

5. Proposizione Sia $S \subset \mathbb{R}^n$ una superficie regolare di dimensione $\ell < n$ definita in forma implicita come $\{x : F(x) = 0\}$ (ossia, F è una funzione C^1 da un aperto di $\mathbb{R}^n$ a valori in $\mathbb{R}^{n-\ell}$ tale che $DF(x)$ ha rango $n - \ell$ per ogni $x \in S$), e sia $x^0 \in S$.

Allora i vettori tangenti a S in x^0 sono tutti i $v \in \mathbb{R}^n$ tali che

$$DF(x^0)v = 0 \quad (\text{ossia } v \in \ker DF(x^0)) \quad (1)$$

mentre i vettori normali a S in x^0 sono tutti i $w \in \mathbb{R}^n$ combinazioni lineari di $DF_1(x^0), \dots, DF_{n-\ell}(x^0)$: $\exists \lambda_1, \dots, \lambda_{n-\ell}$ tali che

$$w = \lambda_1 DF_1(x^0) + \dots + \lambda_{n-\ell} DF_{n-\ell}(x^0) \quad (\text{ossia } w \in \text{img } DF(x^0)^T). \quad (2)$$

DIM. Se $\gamma(t)$ è una curva C^1 a valori in S abbiamo $F(\gamma(t)) = 0$, quindi derivando otteniamo

$$DF(\gamma(t))\gamma'(t) = 0.$$

Se prendiamo in particolare $\gamma : (-\epsilon, \epsilon) \rightarrow S$ con $\gamma(0) = x^0$ e $\gamma'(0) = v$ otteniamo $DF(x^0)v = 0$, cioè tutti i vettori tangenti soddisfano (1). Dato che lo spazio tangente ha dimensione ℓ proprio come $\ker DF(x^0)$, possiamo concludere che i vettori tangenti sono tutti e soli i vettori che verificano (1).

La seconda proprietà segue subito dalla prima: infatti i vettori normali sono i vettori ortogonali a tutti i vettori tangenti, mentre i vettori che soddisfano (2) sono i vettori ortogonali a tutti quelli che soddisfano (1).

6. Osservazione Sia F una funzione C^1 definita su un aperto di $\mathbb{R}^n$ a valori reali. Un insieme di livello di F è un insieme del tipo $S = \{x : F(x) = c\}$ per $c \in \mathbb{R}$ fissata. Se $DF \neq 0$ in tutti i punti di S , per quanto visto sopra S è una superficie regolare di dimensione $n - 1$. In questo caso la Proposizione 5 mostra che il vettore $DF(x)$ è ortogonale ad S per tutti gli $x \in S$.

Siamo pronti per dimostrare la proprietà che è alla base del metodo dei moltiplicatori di Lagrange:

7. Teorema (Teorema dei moltiplicatori di Lagrange) Sia $S \subset \mathbb{R}^n$ una superficie regolare di dimensione $\ell < n$ definita in forma implicita come $S = \{x : F(x) = 0\}$, sia $x^0 \in S$, e sia f una funzione C^1 definita su un intorno di x^0 a valori reali.

Se f ha massimo o minimo vincolato a S in x_0 , allora $Df(x_0)$ è normale ad S in x_0 .

DIM. Sia v un vettore tangente a S in x_0 , quindi esiste $\gamma : (-\epsilon, +\epsilon) \rightarrow S$ tale che $\gamma(0) = x_0$ e $\gamma'(0) = v$. Chiaramente la funzione $f(\gamma(t))$ ha massimo o minimo in $t = 0$, quindi la sua derivata in 0 si annulla:

$$0 = \frac{d}{dt} f(\gamma(t))|_{t=0} = Df(\gamma(0))\gamma'(0) = Df(x_0)v.$$

Vediamo così che $Df(x_0)$ è ortogonale ad un qualunque vettore tangente v ossia è un vettore normale.

Spieghiamo il significato del termine "moltiplicatori": abbiamo dimostrato che in un punto di massimo vincolato a S il gradiente Df deve essere normale alla superficie. Ricordando la Proposizione 5, questo vuol dire che esistono $\lambda_1, \dots, \lambda_{n-\ell} \in \mathbb{R}$ tali che

$$Df(x^0) = \lambda_1 DF_1(x^0) + \dots + \lambda_{n-\ell} DF_{n-\ell}(x^0)$$

vale a dire, più esplicitamente,

$$\frac{\partial f}{\partial x_j}(x^0) = \sum_{i=1}^{n-\ell} \lambda_i \frac{\partial F_i}{\partial x_j}(x^0), \quad j = 1, \dots, n. \quad (3)$$

I numeri λ_j si chiamano per l'appunto i *moltiplicatori di Lagrange*. Notare che il numero dei moltiplicatori è esattamente il "numero dei vincoli", in quanto il vincolo $F = 0$ è vettoriale e riassume $n - \ell$ condizioni.

Per rendere più scorrevole il calcolo dei massimi o minimi di una funzione $f(x)$ vincolati alle condizioni $\{F(x) = 0\}$, si procede nel modo seguente:

(1) Si forma la *Lagrangiana* del problema, definita come

$$G(x, \lambda) := f(x) - \lambda \cdot F(x);$$

(2) Si imposta e si risolve il sistema di equazioni

$$DG(x, \lambda) = 0$$

dove DG indica tutte le derivate di G sia rispetto a $x_1, \dots, x_n$ sia rispetto a $\lambda_1, \dots, \lambda_{n-\ell}$. Si noti infatti che le derivate rispetto a x danno le n condizioni (3) mentre le derivate rispetto a λ danno il vincolo $F = 0$.

Notiamo che la tecnica dei moltiplicatori di Lagrange ci permette di determinare dei punti sul vincolo nei quali la funzione da massimizzare è stazionaria, ma non ci aiuta a capire il carattere di questi punti. Per determinare se si tratta effettivamente di punti di massimo o di minimo, abbiamo bisogno di ulteriori argomenti.

8. Esempio Cerchiamo (di nuovo!) massimo e minimo di $f(x, y) = xy$ sulla circonferenza $C = \{x^2 + y^2 = 1\}$. Poniamo $F = x^2 + y^2 - 1$; si ha $n = 2$, $\ell = n - \ell = 1$ ("un solo vincolo"). La Lagrangiana è

$$G(x, y, \lambda) = xy - \lambda(x^2 + y^2 - 1).$$

Impostiamo il sistema $DG(x, y, \lambda) = 0$:

$$\begin{cases} y = 2\lambda x \\ x = 2\lambda y \\ x^2 + y^2 = 1 \end{cases}$$

(tre equazioni in tre incognite x, y, λ). Risolvendo otteniamo le quattro soluzioni

$$\lambda = -\frac{1}{2}, \quad \left(\mp \frac{\sqrt{2}}{2}, \pm \frac{\sqrt{2}}{2} \right) \quad \text{e} \quad \lambda = \frac{1}{2}, \quad \left(\pm \frac{\sqrt{2}}{2}, \pm \frac{\sqrt{2}}{2} \right).$$

Dato che l'insieme è compatto, massimo e minimo assoluto devono essere raggiunti, quindi ci basta calcolare il valore di f nei quattro punti trovati per concludere che i primi due sono punti di minimo, gli altri due di massimo. Ovviamente la soluzione coincide con quella ottenuta nel primo Esempio.

9. Esempio Determinare il parallelepipedo di massimo volume con superficie totale c assegnata. Se $x, y, z \geq 0$ sono i lati, dobbiamo studiare la funzione $V = xyz$ con la condizione $S = 2(xy + yz + xz) = c$; notiamo anche che possiamo supporre $x, y, z \neq 0$ in quanto se un lato si annulla certamente non abbiamo un massimo. Quindi abbiamo $n = 3$, $\lambda = 1$ ("un solo vincolo"), $f = xyz$, $F = S - c$, la Lagrangiana è data da

$$G(x, y, \lambda) = xyz - \lambda(2(xy + yz + xz) - c).$$

Impostando le equazioni $DG(x, y, z, \lambda) = 0$ otteniamo il sistema

$$\begin{cases} yz = 2\lambda(y + z) \\ xz = 2\lambda(x + z) \\ xy = 2\lambda(x + y) \\ xy + yz + xz = c/2 \end{cases}$$

da cui $(x, y, z \neq 0)$

$$\frac{y+z}{yz} = \frac{x+z}{xz} = \frac{x+y}{xy}$$

e

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

che dà $x = y = z = \sqrt{c}/6$ (cubo).

ESERCIZI

- 1 Trovare massimi e minimi di $f(x) = \prod_{j=1}^n x_j^{x_j}$ sul sottoinsieme di $\mathbb{R}^n$

$$A = \{x : x_j \geq 0, j = 1, \dots, n, x_1 + \dots + x_n = 1\}.$$

[A è chiuso e limitato, quindi f ha massimo e minimo. Inoltre $0 \leq f \leq 1$ su A essendo $0 \leq x_j \leq 1$, e si ha $f < 1$ tranne che nei punti $e_j = (0, \dots, 0, 1, 0, \dots, 0)$. Quindi i punti di massimo sono $e_1, \dots, e_n$. Troviamo i minimi: basta studiare f sull'insieme

$$A_1 = \{x : x_j > 0, j = 1, \dots, n, x_1 + \dots + x_n = 1\}$$

il quale è una superficie regolare di dimensione $n - 1$. Lagrange $\Rightarrow$

$$\begin{cases} f(x)(1 + \log x_j) = \lambda, & j = 1, \dots, n \\ x_1 + \dots + x_n = 1 \end{cases}$$

da cui $x_1 = \dots = x_n = 1/n$; il minimo si ha in $(1/n, \dots, 1/n)$.

- 2 Dati $a \in \mathbb{R}^n$ e $f(x) = a \cdot x$, determinare il massimo di f sull'insieme $C = \{|x|^2 = 1\} \subseteq \mathbb{R}^n$. Se poi A è una matrice $n \times n$ simmetrica e definita positiva, determinare il massimo di f sull'insieme $C_1 = \{Ax \cdot x = 1\} \subseteq \mathbb{R}^n$.
- 3 Minimo di $f(x) = x_1^2 + \dots + x_n^2$ su $\{x_1 x_2 \dots x_n = k\}$, $k \geq 0$ [limitarsi agli $x_j \geq 0$, dove si ha un solo minimo; gli altri minimi si ottengono da questo in modo semplice; quanti sono?] Che si può dire del massimo? Usare il risultato sul minimo per dimostrare la nota disuguaglianza: media aritmetica di n numeri positivi $\geq$ media geometrica.
- 4 Massimi di $f(x) = (x_1 + \dots + x_n)^2$ su $\{x_1^2 + \dots + x_n^2 = 1\}$ [ce ne sono 2]. Quanti sono i minimi?
- 5 Calcolare la distanza di $a \in \mathbb{R}^3$ dal piano di $\mathbb{R}^3$ dato da $b \cdot x + b_0 = 0$, con $b, b_0 \in \mathbb{R}^3$, $b \neq 0$.
- 6 Determinare il punto dell'insieme $\{x : a \cdot x = \alpha, b \cdot x = \beta\}$ di $\mathbb{R}^3$ più vicino all'origine, per $a, b \in \mathbb{R}^3 \setminus 0$ e $\alpha, \beta \in \mathbb{R}$ fissati.
- 7 Determinare il punto più vicino e quello più lontano dall'origine di $\mathbb{R}^2$ sulla curva data implicitamente da $2x^4 + y^4 - xy = 0$.

8 Trovare il minimo delle seguenti funzioni sugli insiemi indicati a fianco:

$$\begin{aligned} x^2 + y^2 + z^2 & \quad \text{su} \quad \{x + 2y - 3z = 5\}, \\ x^2 + 5y^2 + z^2 + t^2 & \quad \text{su} \quad \{x - 3y - z = 1\} \cap \{x + 3y - t = 2\}, \\ x^2 + y^2 + 2z^2 & \quad \text{su} \quad \{x + y = 1\} \cap \{2x + y - z = 2\}, \\ 2x^2 + y^2 + 2z^2 + t^2 & \quad \text{su} \quad \{x - y - t = 4\}. \end{aligned}$$

2-11 LUNGHEZZA DI UNA CURVA

Definire la lunghezza di una curva non richiede tecniche sofisticate: è sufficiente approssimare la curva con linee spezzate inscritte nella curva, e prendere il sup delle lunghezze di tali spezzate. Naturalmente non per tutte le curve il procedimento dà un risultato finito, ad esempio la curva di Peano che riempie tutto il quadrato $[0, 1]^2$ non può avere lunghezza finita. Le curve di lunghezza finita si chiamano *rettificabili*, e in particolare tutte le curve C^1 lo sono. In questo caso si può dare anche una semplice formula per il calcolo della lunghezza, che sarà di importanza fondamentale nel seguito.

1. Definizione Data una *partizione* di $[a, b]$, cioè un insieme finito di punti $\pi = \{t_0, t_1, \dots, t_m\}$ con $a = t_0 < t_1 < \dots < t_{m-1} < t_m = b$, la lunghezza della spezzata inscritta nella curva ϕ , avente vertici $\phi(t_0), \dots, \phi(t_m)$ è data da

$$L_\pi(\phi) = \sum_{k=1}^m |\phi(t_k) - \phi(t_{k-1})|.$$

La *lunghezza* della curva ϕ è il numero

$$L(\phi) = \sup\{L_\pi(\phi) : \pi \text{ partizione di } [a, b]\}.$$

Se $L(\phi) < +\infty$, la curva si dice *rettificabile*.

2. Osservazione Nella dimostrazione che segue calcoleremo integrali di funzioni continue a valori vettoriali. Se $\phi : [a, b] \rightarrow \mathbb{R}^n$ è una curva continua, il suo integrale da a a b (nel senso di Riemann) si definisce semplicemente come il vettore degli integrali componente per componente:

$$\int_a^b \phi(t) dt = \left(\int_a^b \phi_1(t) dt, \dots, \int_a^b \phi_n(t) dt \right).$$

Notiamo che vale la disuguaglianza

$$\left| \int_a^b \phi(t) dt \right| \leq \int_a^b |\phi(t)| dt.$$

Per dimostrarla basta osservare che per ogni $v \in \mathbb{R}^n$ il vettore $w = \int_a^b \phi(t) dt$ soddisfa

$$|v \cdot w| \leq \int_a^b |v \cdot \phi(t)| dt \leq |v| \int_a^b |\phi(t)| dt$$

e prendendo $v = w$ e semplificando otteniamo la tesi.

3. Teorema Una curva $\phi : [a, b] \rightarrow \mathbb{R}^n$ di classe C^1 è rettificabile, e si ha

$$L(\phi) = \int_a^b |\phi'(t)| dt.$$

DIM. Chiamiamo I il numero $\int_a^b |\phi'(t)| dt$: vogliamo dimostrare che $L(\phi) \leq I$ e che $I \leq L(\phi)$.

Sia $\pi = \{t_0, \dots, t_m\}$ una qualunque partizione di $[a, b]$. Si ha

$$|\phi(t_k) - \phi(t_{k-1})| = \left| \int_{t_{k-1}}^{t_k} \phi'(t) dt \right| \leq \int_{t_{k-1}}^{t_k} |\phi'(t)| dt$$

da cui sommando per $k = 1, \dots, m$,

$$L_\pi(\phi) \leq I.$$

Prendendo il sup al variare di π otteniamo la prima disuguaglianza

$$L(\phi) \leq I.$$

In particolare, ϕ è rettificabile.

Dimostriamo ora la disuguaglianza inversa. La funzione ϕ' è uniformemente continua, quindi per ogni $\epsilon > 0$ esiste $\delta > 0$ tale che se $|s - t| < \delta$ si ha $|\phi'(s) - \phi'(t)| < \epsilon$. Se prendiamo una partizione di $[a, b]$ tale che $|t_k - t_{k-1}| < \delta$ per tutti i k , si ha, per ogni $s \in [t_{k-1}, t_k]$,

$$\begin{aligned} \phi(t_k) - \phi(t_{k-1}) &= \int_{t_{k-1}}^{t_k} \phi'(t) dt = \int_{t_{k-1}}^{t_k} [\phi'(s) + (\phi'(t) - \phi'(s))] dt \\ &= \phi'(s)(t_k - t_{k-1}) + \int_{t_{k-1}}^{t_k} (\phi'(t) - \phi'(s)) dt. \end{aligned}$$

Ne segue

$$|\phi(t_k) - \phi(t_{k-1})| \geq |\phi'(s)|(t_k - t_{k-1}) - \int_{t_{k-1}}^{t_k} |\phi'(t) - \phi'(s)| dt \geq |\phi'(s)|(t_k - t_{k-1}) - \epsilon(t_k - t_{k-1}).$$

Dividiamo per $t_k - t_{k-1}$ e integriamo in s da t_{k-1} a t_k : otteniamo

$$|\phi(t_k) - \phi(t_{k-1})| \geq \int_{t_{k-1}}^{t_k} |\phi'(s)| ds - \epsilon(t_k - t_{k-1})$$

e sommando su k otteniamo

$$L_\pi(\phi) \geq I - \epsilon(b - a) \quad \Rightarrow \quad L(\phi) \geq I - \epsilon(b - a)$$

e l'arbitrarietà di ϵ ,

$$L(\phi) \geq I.$$

4. Osservazione Notiamo che il teorema vale ovviamente anche per le curve C^1 a tratti, cioè continue e tali che $[a, b]$ si può spezzare in un numero finito di intervalli su ognuno dei quali ϕ è di classe C^1 .

5. Esempio Calcoliamo la lunghezza della curva $\phi : [0, 1] \rightarrow \mathbb{R}^3$ data da $\phi_1(t) = x(t) = t^2/\sqrt{2}$, $\phi_2(t) = y(t) = t^3/3$, $\phi_3(t) = z(t) = 1 - t$. Si ha

$$L = \int_0^1 \sqrt{2t^2 + t^4 + 1} dt = \int_0^1 (1 + t^2) dt = \frac{4}{3}.$$

6. Esempio Se una curva $\phi : [a, b] \rightarrow \mathbb{R}^2$ è grafico di una funzione $y = f(x)$, di classe C^1 , quindi $\phi(t) = (t, f(t))$, si ha

$$L = \int_a^b |\phi'(t)| dt = \int_a^b \sqrt{1 + |f'(x)|^2} dx.$$

Se una curva è rappresentata in coordinate polari come $\rho = f(\theta)$, $\theta \in [\alpha, \beta]$, allora scrivendola in forma parametrica $\phi(\theta)$ si ha $\phi_1(\theta) = f(\theta) \cos \theta$, $\phi_2(\theta) = f(\theta) \sin \theta$, quindi

$$L = \int_\alpha^\beta |\phi'(\theta)| d\theta = \int_\alpha^\beta \sqrt{|f(\theta)|^2 + |f'(\theta)|^2} d\theta.$$

ESERCIZI

1 Calcolare la lunghezza dei seguenti archi di curva.

- 1) $9y^2 = 4x^3$, $0 \leq x \leq 3$ $[14/3]$
- 2) $y = 2\sqrt{3x}$, $0 \leq x \leq 1$ $[1 + (3 \log 3)/2]$
- 3) $y^2 = 2x - x^2$, $0 \leq x \leq 1$ $[\pi/2]$
- 4) $y = \log \sin x$, $\pi/3 \leq x \leq \pi/2$ $[(\log 3)/2]$
- 5) $y = 1 - \log \cos x$, $0 \leq x \leq \pi/4$ $[\log(\sqrt{2} + 1)]$
- 6) arco di catenaria: $\phi(x) = (x, \cosh x)$, $x \in [0, a]$
- 7) arco di parabola: $\phi(x) = (x, x^2)$, $x \in [0, a]$
- 8) arco di elica cilindrica ($r > 0, k > 0$): $\phi(t) = (r \cos t, r \sin t, kt)$, $t \in [0, a]$
- 9) arco di cicloide ($k > 0$): $\phi(t) = (k(t - \sin t), k(1 - \cos t))$, $t \in [0, a]$
- 10) spirale di Archimede ($k > 0$): $\rho = k\theta$, $\theta \in [0, a]$
- 11) spirale logaritmica: $\rho = e^{-\theta}$, $\theta \in [\alpha, \beta]$ oppure $\theta \in [0, +\infty[$
- 12) cardiode ($k > 0$): $\rho = k(1 + \cos \theta)$, $\theta \in [-\pi, \pi]$.

01 Una curva in $\mathbb{R}^2$ che sia grafico di una funzione continua e monotona è rettificabile.

[Sia $f : [a, b] \rightarrow \mathbb{R}$ monotona, ad esempio crescente, la curva è $\phi(t) = (t, f(t))$; sia $\pi = \{t_0, \dots, t_m\}$ una partizione di $[a, b]$, si ha

$$|\phi(t_k) - \phi(t_{k-1})| = \sqrt{(t_k - t_{k-1})^2 + (f(t_k) - f(t_{k-1}))^2} \leq t_k - t_{k-1} + f(t_k) - f(t_{k-1})$$

e pertanto

$$L_\pi(\phi) \leq b - a + f(b) - f(a) < +\infty.]$$

2-12 1-FORME

In questo Paragrafo, Ω è un aperto di $\mathbb{R}^n$. Se $f : \Omega \rightarrow \mathbb{R}$ è di classe C^1 , abbiamo visto che il differenziale è un'applicazione $df : \Omega \rightarrow (\mathbb{R}^n)^*$, che scritta nella base $dx_1, \dots, dx_n$ (che in ogni punto di Ω è la base duale $e_1^*, \dots, e_n^*$), diventa

$$df = D_1 f dx_1 + \dots + D_n f dx_n.$$

1. Definizione In generale, un'applicazione $\omega : \Omega \rightarrow (\mathbb{R}^n)^*$ si dice *1-forma* o *forma differenziale di grado 1*; si può sempre scrivere nella base

$$\omega(x) = \omega_1(x)dx_1 + \cdots + \omega_n(x)dx_n$$

per opportune funzioni $\omega_j(x) : \Omega \rightarrow \mathbb{R}$ dette i *coefficienti* della forma. La forma si dice *continua*, $C^k(\Omega)$, $C^\infty(\Omega)$ se lo sono i suoi coefficienti. Nel caso particolare che ω sia proprio il differenziale di una certa funzione $f : \Omega \rightarrow \mathbb{R}$ di classe C^1 , la forma si dice *esatta* e la funzione f si chiama *primitiva* della forma su Ω .

2. Osservazione Il problema di trovare una primitiva di una forma ω assomiglia molto al problema di trovare una primitiva di una funzione di una sola variabile, ma ci sono alcune differenze importanti che mettono in gioco la geometria degli insiemi di $\mathbb{R}^n$. Vediamone due:

(i) È sempre possibile trovare una primitiva? No. Infatti se $\omega \in C^1(\Omega)$ e $f \in C^2(\Omega)$ è una sua primitiva ossia $\omega_j(x) = D_j f(x)$, dal Teorema di Schwartz abbiamo $D_i \omega_j = D_i D_j f = D_j D_i f = D_j \omega_i$. Quindi la condizione $D_i \omega_j = D_j \omega_i$ è *necessaria* affinché $\omega \in C^1(\Omega)$ sia esatta.

(ii) Quante primitive ci possono essere? Se f e g sono due primitive della stessa forma ω su Ω , ne segue che $d(f-g) = 0$. Questo vuol dire che su ogni componente connessa di Ω le funzioni f e g differiscono al più per una costante (vedi Proposizione 3.16); naturalmente la costante può essere diversa in ciascuna componente. Se Ω è connesso allora possiamo dire che $g = f + C$.

3. Definizione Una forma $\omega \in C^1(\Omega)$ si dice *chiusa* se per ogni coppia di indici i, j si ha $D_i \omega_j = D_j \omega_i$ su Ω .

Riassumendo, una forma C^1 esatta deve essere anche chiusa. Vale il viceversa? ossia, questa condizione è anche sufficiente perché esista una primitiva? e come si calcola? La risposta dipende non solo dalle proprietà di ω ma anche dalle proprietà topologiche dell'aperto Ω , cioè su alcuni aperti ogni forma chiusa è esatta, su altri no. Sapere di quanto le forme chiuse sono più delle forme esatte, dà informazioni sulla forma dell'aperto (in sintesi, la dimensione dello spazio quoziente chiuse/esatte dipende dal “numero dei buchi” di Ω). Questo è un primo esempio di un fenomeno generale: le forme differenziali sono uno strumento analitico per ottenere informazioni geometriche. In questo senso, esse rappresentano un ponte di collegamento fra l'analisi e la geometria.

Inoltre, noi qui ci limiteremo a lavorare con forme differenziali definite su aperti di $\mathbb{R}^n$. Ma la teoria si può estendere in modo naturale a superfici qualunque (e anche a strutture più generali). La teoria delle forme differenziali permette di estendere l'analisi classica, definita su oggetti “piatti” come gli aperti di $\mathbb{R}^n$, a oggetti “curvi” ossia varietà differenziabili.

4. Esempio Diamo l'esempio più semplice, e in un certo senso il prototipo, di una forma chiusa che non è esatta. Questo esempio suggerisce che il problema non è tanto trovare una primitiva localmente, il che si può sempre fare se la forma è chiusa (vedi più avanti); il problema è “incollare” le varie primitive locali e definirne una globale su tutto l'aperto dove la forma è definita.

Consideriamo la forma

$$\omega(x, y) = -\frac{y}{x^2 + y^2}dx + \frac{x}{x^2 + y^2}dy, \quad \omega \in C^2(\Omega), \quad \Omega = \mathbb{R}^2 \setminus 0.$$

ω è chiusa perché $D_2\omega_1 = D_1\omega_2 = \frac{y^2 - x^2}{(x^2 + y^2)^2}$. Proviamo a calcolare una primitiva. Introduciamo la funzione

$$\theta(x, y) = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{se } y \in \mathbb{R} \text{ e } x > 0, \\ \frac{\pi}{2} - \arctan\left(\frac{x}{y}\right) & \text{se } y > 0 \text{ e } x \in \mathbb{R}, \\ -\frac{\pi}{2} - \arctan\left(\frac{x}{y}\right) & \text{se } y < 0 \text{ e } x \in \mathbb{R}. \end{cases}$$

La funzione $\theta(x, y)$, detta anche $\arg(x, y)$ o *argomento* del vettore (x, y) esprime l'angolo formato dal vettore con la direzione positiva dell'asse delle x (misurato in radianti in senso antiorario); si tratta esattamente della stessa funzione θ utilizzata per descrivere la coordinata angolare in coordinate polari. Notare che le tre espressioni sono definite su insiemi che si sovrappongono parzialmente, ma le definizioni coincidono sulle zone di sovrapposizione e la funzione θ è continua e anzi C^1 su tutto l'aperto A dato da

$$A = \mathbb{R}^2 \setminus (-\infty, 0] \times \{0\}$$

ossia per tutti i punti del piano tranne il semiasse x negativo. Tuttavia non è possibile estendere con continuità θ a tutto $\Omega = \mathbb{R}^2 \setminus 0$, perché se ci avviciniamo ad un punto del semiasse x negativo dal di sopra, θ tende a $+\pi$, mentre se ci avviciniamo dal di sotto, θ tende a $-\pi$: c'è un salto di 2π (come è ovvio).

Si verifica facilmente che la funzione θ è una primitiva di ω su A :

$$d\theta(x, y) = \omega(x, y) \quad \forall (x, y) \in A.$$

Se esistesse una primitiva f di ω su tutto Ω , in particolare sarebbe $d(f - \theta) = 0$ su A ed essendo A connesso ne seguirebbe che f è della forma $f = \theta + C$ per una costante C ; ma allora anche f avrebbe un salto di 2π in corrispondenza dei punti del semiasse x negativo, e non potrebbe essere $C^1(\Omega)$. Dunque ω non possiede primitive definite su tutto Ω .

Andiamo più a fondo nel problema. A questo scopo abbiamo bisogno del concetto di integrale di una forma lungo una curva.

5. Definizione Breve sommario delle definizioni utilizzate nel seguito. Una *curva* in Ω è un'applicazione continua $\gamma : [a, b] \rightarrow \Omega$. γ si dice C^1 a tratti se $[a, b]$ si spezza in un numero finito di intervalli su ciascuno dei quali γ è C^1 . $\gamma(a)$ e $\gamma(b)$ si dicono *estremo iniziale* e *finale* della curva, e se coincidono la curva si dice *chiusa*. Il *sostegno* di γ è la sua immagine $\gamma([a, b])$, spesso indicata ancora con γ . Un *cambiamento di parametro* è un'applicazione $\tau : [a, b] \rightarrow [c, d]$ di classe C^1 con $\tau' \neq 0$ (quindi invertibile con inversa C^1). Due curve $\phi : [a, b] \rightarrow \Omega$, $\psi : [c, d] \rightarrow \Omega$ si dicono *equivalenti* se esiste un cambiamento di parametro $\tau : [a, b] \rightarrow [c, d]$ tale che $\phi(t) = \psi(\tau(t))$ con $\tau' > 0$; è chiaro che si tratta di una relazione di equivalenza [verificare!]. Notiamo che curve equivalenti hanno lo stesso sostegno.

Con $-\gamma$ indichiamo la curva $\gamma(-t)$ definita su $[-b, -a]$; date due curve $\phi : [a, b] \rightarrow \Omega$, $\psi : [c, d] \rightarrow \Omega$ tali che $\phi(b) = \psi(c)$, con $\phi + \psi$ indichiamo la curva definita su $[a, b + d - c]$

come

$$(\phi + \psi)(t) = \begin{cases} \phi(t) & \text{per } t \in [a, b], \\ \psi(t + c - b) & \text{per } t \in [b, b + d - c]. \end{cases}$$

Con $\phi - \psi$ si indica la curva $\phi + (-\psi)$.

6. Definizione Sia ω una forma continua su Ω , sia $\gamma : [a, b] \rightarrow \Omega$ una curva C^1 . L'integrale di ω lungo γ è il numero

$$\int_{\gamma} \omega = \sum_{j=1}^n \int_a^b \omega_j(\gamma(t)) \gamma'_j(t) dt.$$

Se γ è C^1 a tratti, si definisce l'integrale spezzando nei tratti in cui γ è C^1 , integrando su ciascun tratto e sommando. Nel seguito supporremo *tutte* le curve C^1 a tratti.

7. Osservazione In fisica, una funzione $F = (F_1, F_2, F_3) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ si chiama anche un *campo di forze*. Consideriamo la forma differenziale $\omega = F_1 dx_1 + F_2 dx_2 + F_3 dx_3$ associata ad F . L'integrale di ω lungo una curva γ si dice allora il *lavoro* di un punto materiale che si muove lungo γ , sottoposto al campo di forze F .

8. Proposizione Sia ω una forma continua su Ω e ϕ, ψ curve C^1 a tratti in Ω . Allora

(a) Se ϕ e ψ sono equivalenti, $\int_{\phi} \omega = \int_{\psi} \omega$.

(b) $\int_{-\phi} \omega = -\int_{\phi} \omega$.

(c) Se $\omega = df$ è esatta e $\phi : [a, b] \rightarrow \Omega$, allora $\int_{\phi} \omega = f(\phi(b)) - f(\phi(a))$.

DIM. (a): sia $\phi : [a, b] \rightarrow \Omega$, $\psi : [c, d] \rightarrow \Omega$, $\tau : [a, b] \rightarrow [c, d]$ il cambiamento di variabili, e $\phi = \psi \circ \tau$. Allora

$$\int_{\phi} \omega = \int_a^b \sum \omega_j(\phi(t)) \phi'_j(t) dt = \int_a^b \sum \omega_j(\psi(\tau(t))) \psi'_j(\tau(t)) \tau'(t) dt = \int_c^d \sum \omega_j(\psi(s)) \psi'_j(s) ds = \int_{\psi} \omega$$

effettuando il cambiamento di variabile $s = \tau(t)$. (b) è ovvia. (c) segue dal fatto che

$$\sum_{j=1}^n \omega_j(\phi(t)) \phi'_j(t) = Df(\phi(t)) \cdot \phi'(t) = \frac{d}{dt} [f(\phi(t))],$$

quindi $\int_{\phi} \omega = \int_a^b \frac{d}{dt} [f(\phi(t))] dt = f(\phi(b)) - f(\phi(a))$.

Nel seguito individueremo una classe di aperti sui quali ogni forma chiusa è anche esatta. A questo scopo abbiamo bisogno di qualche risultato di topologia.

9. Osservazione Conosciamo già le nozioni di insieme connesso e insieme connesso per archi, e abbiamo dimostrato in precedenza (Proposizione 9 del Paragrafo 7 del Capitolo 1) che le due nozioni coincidono per sottoinsiemi aperti di $\mathbb{R}^n$. La definizione di insieme *connesso per archi* C^1 a tratti è completamente analoga un sottoinsieme $A \subseteq \mathbb{R}^n$ si dice connesso per archi C^1 a tratti se dati due punti $x, y \in A$ è sempre possibile congiungerli con un arco C^1 a tratti. Anche in questo caso si ha:

10. Lemma Un aperto di $\mathbb{R}^n$ è connesso se e solo se è connesso per archi C^1 a tratti.

DIM. La dimostrazione è, parola per parola, identica a quella della Proposizione 7.9 del Capitolo 1. L'unica differenza è che l'insieme A_p va definito come l'insieme dei punti che possono essere connessi a p da un arco C^1 a tratti contenuto in A .

Vediamo anzitutto un criterio molto semplice per riconoscere le forme esatte.

11. Proposizione *Una forma continua in un aperto è esatta se e solo se il suo integrale su ogni curva chiusa C^1 a tratti è nullo.*

DIM. Se $\omega = df$ è esatta e ϕ è chiusa, allora $\int_{\phi} \omega = f(\phi(a)) - f(\phi(b)) = 0$.

Viceversa, supponiamo di sapere che $\int_{\phi} \omega = 0$ per ogni curva chiusa. Supponiamo anzitutto che Ω sia connesso. Fissato $x_0 \in \Omega$, per ogni altro $x \in \Omega$ esiste una curva ϕ C^1 a tratti da x_0 a x per il Lemma precedente, quindi possiamo porre $f(x) = \int_{\phi} \omega$. Perché questa definizione abbia senso, dobbiamo verificare che se ψ è un'altra curva da x_0 a x , si ha $\int_{\psi} \omega = \int_{\phi} \omega$; ma basta osservare che $\psi - \phi$ è una curva chiusa e applicare la Prop. 6 per avere $0 = \int_{\psi - \phi} \omega = \int_{\psi} \omega - \int_{\phi} \omega$.

Resta da verificare che $f(x)$ è una primitiva di ω su Ω , cioè che $D_j f = \omega_j$ per ogni j . Calcoliamo il rapporto incrementale di f in x nella direzione e_j : scegliamo una curva ϕ da x_0 a x , quindi $f(x) = \int_{\phi} \omega$; sia poi ψ il segmento che unisce x a $x + he_j$, cioè $\psi(t) = x + the_j$, $t \in [0, 1]$, quindi $f(x + he_j) = \int_{\phi + \psi} \omega$; allora il rapporto incrementale si scrive [notare che $\psi'_i(t) = 0$ se $i \neq j$, $\psi'_j(t) = h$]

$$\begin{aligned} \frac{1}{h} [f(x + he_j) - f(x)] &= \frac{1}{h} \left[\int_{\phi + \psi} \omega - \int_{\phi} \omega \right] = \frac{1}{h} \int_{\psi} \omega = \frac{1}{h} \int_0^1 \omega_j(x + the_j) h dt \\ &= \int_0^1 \omega_j(x + the_j) dt = \frac{1}{h} \int_0^h \omega_j(x + se_j) ds \end{aligned}$$

che tende a $\omega_j(x)$ se $h \rightarrow 0$. Dunque f è primitiva di ω .

Infine, se Ω non è connesso, abbiamo dimostrato che ω ammette una primitiva su ogni componente connessa di Ω ; queste funzioni definiscono una funzione C^1 su Ω che è una primitiva di ω su tutto Ω .

12. Esempio Utilizzando il criterio precedente è molto più facile verificare che la forma ω dell'esempio 3 non è esatta. Ad esempio il suo integrale sulla circonferenza ϕ di centro 0 e raggio 1 non fa 0: parametrizziamola come $\phi(t) = (\cos t, \sin t)$ con $0 \leq t \leq 2\pi$, ed abbiamo

$$\int_{\phi} \omega = \int_0^{2\pi} (\sin^2 t + \cos^2 t) dt = 2\pi.$$

13. Definizione Una curva chiusa $\phi : [a, b] \rightarrow \Omega$ si dice *omotopa in Ω ad un punto $x_0 \in \Omega$* se esiste un'applicazione continua $h : [a, b] \times [0, 1] \rightarrow \Omega$ tale che

$$h(t, 0) = \phi(t), \quad h(t, 1) = x_0 \quad \forall t \in [a, b]$$

$$h(a, s) = h(b, s) \quad \forall s \in [0, 1].$$

Se h è di classe C^k per $s \in]0, 1[$, ϕ si dice *omotopa C^k al punto x_0* . Non è difficile dimostrare [ma noi non lo faremo] che se Ω è un aperto di $\mathbb{R}^n$ e ϕ è una curva chiusa omotopa in Ω ad un punto, allora l'omotopia $h(t, s)$ si può prendere di classe $C^k \forall k$ per tutti i $t \in [0, 1]$ e gli $s \in (0, 1]$; inoltre, se la curva è C^1 a tratti, l'omotopia si può prendere in modo che su ogni tratto $\partial_t h(t, s)$ converga uniformemente a $\phi'(t)$ per $s \rightarrow 0$.

L'aperto Ω si dice *semplicemente connesso* se ogni curva chiusa in Ω è omotopa ad un punto. Se $\Omega \subseteq \mathbb{R}^2$, è molto facile visualizzare gli aperti semplicemente connessi: sono quelli “senza buchi”.

Concludiamo con il teorema fondamentale della teoria delle 1-forme.

14. Teorema *In un aperto semplicemente connesso ogni forma C^1 chiusa è esatta.*

DIM. Dimostriamo che l'integrale di ω su una qualunque curva C^1 a tratti chiusa $\phi : [a, b] \rightarrow \Omega$ è nullo, da cui la tesi per la proposizione precedente. Sappiamo che ϕ è omotopa C^2 ad un punto; sia $h(t, s)$ l'omotopia. Chiaramente per ogni $s \in]0, 1[$ fissato, $t \mapsto h(t, s)$ è una curva C^2 in Ω , chiusa; poniamo

$$\alpha(s) = \int_a^b \sum_{j=1}^n \omega_j(h(t, s)) \partial_t h_j(t, s) dt \quad (1)$$

ossia l'integrale di ω lungo tale curva. Notiamo poi che $\alpha(1) = 0$ [per $s = 1$ la curva si riduce ad un punto], mentre per $s \rightarrow 0$ la funzione $\alpha(s)$ tende a $\int_\phi \omega$, dato che $\partial_t h(t, s)$ converge uniformemente a $\phi'(t)$ su ognuno dei tratti in cui ϕ è di classe C^1 . Quindi per concludere la dimostrazione ci basta dimostrare che $\alpha'(s) = 0$ (ossia α è costante) per $s > 0$. Derivando sotto il segno di integrale [lecito perché l'integrando è C^1] si ha

$$\alpha'(s) = \int_a^b \sum_j \omega_j(h) \partial_s \partial_t h_j dt + \int_a^b \sum_{i,j} D_i \omega_j(h) \partial_s h_i \partial_t h_j dt.$$

Dato che ω è chiusa, possiamo sostituire l'identità $D_i \omega_j = D_j \omega_i$ nel secondo integrale e integrando per parti abbiamo

$$\int_a^b \sum_{i,j} D_j \omega_i(h) \partial_s h_i \partial_t h_j dt = \int_a^b \sum_i \partial_t [\omega_i(h)] \partial_s h_i dt = [\sum_i \omega_i(h) \partial_s h_i]_{t=a}^{t=b} - \int_a^b \sum_i \omega_i(h) \partial_t \partial_s h_i dt$$

da cui

$$\alpha'(s) = \sum_i \omega_i(h(b, s)) \partial_s h_i(b, s) - \sum_i \omega_i(h(a, s)) \partial_s h_i(a, s).$$

Ora basta ricordare che $h_i(a, s) = h_i(b, s)$ per ogni s e quindi anche $\partial_s h_i(a, s) = \partial_s h_i(b, s)$, da cui $\alpha'(s) = 0$.

15. Osservazione Chiaramente una palla è semplicemente connessa [una curva chiusa ϕ in $B(x_0, r)$ è omotopa al centro x_0 tramite l'omotopia $h(t, s) = x_0 + (1-s)(\phi(t) - x_0)$]. Quindi un corollario del teorema precedente è: ogni forma chiusa è *localmente* esatta, vale a dire ammette una primitiva nell'intorno di ogni punto. Il teorema invece afferma che se l'aperto di definizione è semplicemente connesso, esiste un'unica primitiva globale su tutto l'aperto. Il problema evidentemente è “incollare” le diverse primitive nei diversi interni; notare che se due interni si hanno intersezione non vuota (e connessa), le due primitive differiscono al più per una costante. Nel punto 3 abbiamo studiato un esempio concreto di questo fenomeno.

16. Osservazione Un insieme Ω si dice *stellato* rispetto ad un suo punto x_0 se per ogni $x \in \Omega$ il segmento che unisce x_0 a x è tutto contenuto in Ω . Gli insiemi stellati sono semplicemente connessi, e dato che è molto facile verificare che un insieme è stellato, questa proprietà li rende particolarmente utili nello studio delle forme differenziali. La dimostrazione è immediata: se Ω è stellato rispetto a $x_0 \in \Omega$ e γ è una curva chiusa in Ω , l'omotopia cercata si definisce semplicemente come $h(t, s) = (1-s)\gamma(t) + sx_0$.

17. Esempio Dimostrare che $\omega(x, y) = 2xydx + x^2dy$ su $\mathbb{R}^2$ è esatta e calcolarne una primitiva. Si ha $D_2\omega_1 = 2x = D_1\omega_2$, quindi ω è chiusa; $\mathbb{R}^2$ è stellato, quindi ω è esatta. Una primitiva si può calcolare scegliendo arbitrariamente una curva ϕ da 0 a (x, y) e

ponendo $f(x, y) = \int_{\phi} \omega$; ma c'è un modo più semplice. Infatti dev'essere $D_1 f = \omega_1 = 2xy$, quindi $f(x, y) = x^2 y + g(y)$ con g da determinare; ma $D_2 f = \omega_2$, quindi $x^2 + g'(y) = x^2 \Rightarrow g' = 0 \Rightarrow g$ è costante. Possiamo così prendere $f(x, y) = x^2 y$ o più in generale $f(x, y) = x^2 y + C$ (la primitiva è sempre determinata a meno di una costante).

ESERCIZI

- 1 Calcolare $\int_{\phi} \omega$, dove

$$\begin{array}{ll} \omega(x, y) = x^2 y \, dx + (x - y) \, dy, & \phi(t) = (\cos t, \sin t), \quad t \in [0, 2\pi] \\ \omega(x, y) = xy \, dx + ye^x \, dy, & \phi(t) = (t, t^2), \quad t \in [0, 1] \\ \omega(x, y) = y \, dx + \log x \, dy, & \phi(t) = (t, \sqrt{1+t}), \quad t \in [1, 2] \\ \omega(x, y) = d(xy^2), & \phi(t) = (t^3, t^4), \quad t \in [0, 1] \\ \omega(x, y) = \frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy, & \phi(t) = (\cos t, \sin t), \quad t \in [0, 2\pi]. \end{array}$$

- 2 Determinare quali delle seguenti forme sono chiuse e quali esatte sugli aperti indicati a fianco, e calcolare (se possibile) una primitiva.

$$\begin{array}{ll} \omega = (x+1)dx + (y+2)dy, & \Omega = \mathbb{R}^2 \\ \omega = (y+1)dx + (x+2)dy, & \Omega = \mathbb{R}^2 \\ \omega = x \, dx - y \, dy, & \Omega = \mathbb{R}^2 \\ \omega = y \, dx - x \, dy, & \Omega = \mathbb{R}^2 \\ \omega = \frac{2x}{2x^2 + y^2} dx + \frac{y}{2x^2 + y^2} dy, & \Omega = \mathbb{R}^2 \setminus 0 \\ \omega = \frac{2y}{x^2 + 2y^2} dx - \frac{x}{x^2 + 2y^2} dy, & \Omega = \mathbb{R}^2 \setminus 0 \\ \omega = (y+z)dx + (x+z)dy + (x+y)dz, & \Omega = \mathbb{R}^3 \\ \omega = (3y-1)dx + (3x+1)dy + 3z \, dz, & \Omega = \mathbb{R}^3 \\ \omega = \frac{2xy}{(1+x^2)^2} dx - \frac{1}{1+x^2} dy, & \Omega = \mathbb{R}^2. \end{array}$$

- 3 Trovare per quali valori di $a \in \mathbb{R}$ la forma su $\mathbb{R}^2$

$$\omega(x, y) = -\frac{2xy}{1+y^2+2x^2+x^4} dx + \frac{a+x^2}{1+y^2+2x^2+x^4} dy$$

è chiusa ed esatta, e calcolarne una primitiva [$a = 1$, $f = \arctan(y/(1+x^2))$]; per calcolare f procedere come nell'Esempio 17].

- 01 Sia Ω aperto connesso, ω una forma esatta su Ω , e siano f, g due primitiva di ω . Allora $f - g$ è costante. Che succede se Ω non è connesso?
- 02 Sia ω una forma su $\mathbb{R}^2$ del tipo $\omega(x, y) = \alpha(y)dx + \beta(x)dy$, con α, β di classe C^1 . Determinare per quali funzioni α, β la forma è chiusa, per quali è esatta, e calcolarne una primitiva.
- 03 Stessa domanda per $\omega(x, y) = y\alpha(x)dx + x\beta(y)dy$.
- 04 Stessa domanda per la forma su $\mathbb{R}^3$ $\omega(x, y, z) = ydx + xdy + \alpha(x, y, z)dz$.
- 05 Stessa domanda per la forma $\omega(x, y, z) = xzdx + 2yzdy + \alpha(x, y)dz$.
- 06 Sia $\Omega = \{(x, y) : x + y \neq 0\}$, e sia

$$\omega = \frac{y}{(x+y)^2} dx - \frac{x}{(x+y)^2} dy.$$

Determinare se essa è chiusa, esatta, e in caso affermativo calcolarne *tutte* le primitive [si; si; una primitiva è $f = -y/(x+y)$; le altre si ottengono...]

- 07 (*Microcorso di variabile complessa*).

(i) Sia Lz un'applicazione lineare (su $\mathbb{C}$) da $\mathbb{C}$ in $\mathbb{C}$, quindi $Lz = \alpha \cdot z$ per un fissato $\alpha \in \mathbb{C}$. Allora essa può essere

vista come un'applicazione lineare da $\mathbb{R}^2$ in $\mathbb{R}^2$ rappresentata da una matrice 2×2 . Che forma ha tale matrice, e come si scrive nei termini di α ?

[Se $\alpha = a + ib$, $z = x + iy$, si ha

$$\alpha \cdot z = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.]$$

(ii) Sia A un aperto di $\mathbb{C}$ e $f : A \rightarrow \mathbb{C}$. f si dice *derivabile in senso complesso* in $z_0 \in A$ se esiste il limite

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0},$$

detto *derivata* (in senso complesso) di f in z_0 . Se f è derivabile in senso complesso in tutti i punti di A , si dice che essa è *olomorfa* in A . Possiamo considerare f come una funzione da A aperto di $\mathbb{R}^2$ in $\mathbb{R}^2$; dimostrare che si ha $f'(z_0) = D_1 f(z_0) = \frac{1}{i} D_2 f(z_0)$. Ponendo $f(x + iy) = u(x, y) + iv(x, y)$, questa relazione equivale alle due relazioni $u_x = v_y$, $u_y = -v_x$, dette *equazioni di Cauchy-Riemann*.

[Basta fare il limite del rapporto incrementale complesso scegliendo $z = x + iy_0 \rightarrow z_0$ oppure $z = x_0 + iy \rightarrow z_0$; i due limiti ovviamente devono coincidere]

iii) Se f è derivabile in senso complesso in z_0 allora f è differenziabile in z_0 , e $df(z_0) \cdot h = f'(z_0) \cdot h$, ossia l'applicazione lineare $df(z_0)$ coincide con la moltiplicazione per il numero complesso $f'(z_0)$ nel senso di (i). Viceversa, se f è differenziabile in z_0 e si ha $D_2 f(z_0) = i D_1 f(z_0)$, allora f è derivabile in senso complesso in z_0 .

[Dalla definizione

$$\frac{f(z) - f(z_0) - f'(z_0) \cdot (z - z_0)}{|z - z_0|} = \left(\frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right) \frac{z - z_0}{|z - z_0|} \rightarrow 0.]$$

(iv) Somma, prodotto e quoziente di funzioni olomorfe sono funzioni olomorfe. $f(z) = z$ è olomorfa, $f(z) = \bar{z}$ non lo è.

(v) Si possono definire le 1-forme a valori complessi su $\Omega \subseteq \mathbb{R}^2 = \mathbb{C}$ semplicemente considerando le $\omega = \omega_1(z)dx + \omega_2(z)dy$ con ω_j a valori complessi. Le nozioni e i risultati relativi alle 1-forme si possono estendere senza difficoltà alle forme complesse; in particolare l'integrale è definito come l'integrale della parte reale di ω più i volte l'integrale della parte immaginaria. Usando le notazioni $dz = dx + idy$, $d\bar{z} = dx - idy$, mostrare che ogni forma si scrive come $\omega = \alpha(z)dz + \beta(z)d\bar{z}$. Dimostrare che la forma $\omega = f(z)dz$ di classe C^1 è chiusa se e solo se $f(z)$ è olomorfa.

(vi) Data $f : \Omega \rightarrow \mathbb{C}$ continua, e γ curva C^1 a tratti in Ω , l'integrale di f su γ è definito come l'integrale su γ della forma $\omega = f(z)dz$; si usa la notazione $\int_\gamma f(z)dz$. Dimostrare il Teorema Integrale di Cauchy: se l'integrale di $f(z)$ su ogni curva chiusa in Ω è nullo, ed $f \in C^1$, allora f è olomorfa in Ω . Inoltre se Ω è semplicemente connesso, vale anche il viceversa.

(vii) Dimostrare che $\int_{\gamma_R} \frac{1}{z} dz = 2\pi i$ se $\gamma_R(t) = Re^{2\pi it}$, $0 \leq t \leq 2\pi$ (circonferenza di raggio R).

(viii) Dimostrare che se $f(z)$ è C^1

$$\lim_{R \rightarrow 0} \int_{\gamma_R} \frac{f(z)}{z} dz = 2\pi i f(0).$$

[Usare Taylor: $f(z) = f(0) + R_1(z)$ con $|R_1(z)| \leq |z| \cdot \sup_{|z| \leq \epsilon} |Df(z) - Df(0)|$ per $|z| \leq \epsilon$.]

(ix) Sia $f \in C^1$. Dimostrare che per f olomorfa l'integrale $\int_{\gamma_R} \frac{f(z)}{z} dz$ non dipende dal raggio R e dedurne la *formula di Cauchy*

$$f(0) = \frac{1}{2\pi i} \int_{\gamma_R} f(z) dz.$$

PIERO D'ANCONA: DIPARTIMENTO DI MATEMATICA, SAPIENZA UNIVERSITÀ DI ROMA, PIAZZALE A. MORO 2, 00185 ROMA, ITALY

E-mail address: dancona@mat.uniroma1.it