

Note Integrative di Calcolo Differenziale ed Integrale Multidimensionale

per il corso di Analisi Matematica II — a.a. 2011-12

A. Visintin — Facoltà di Ingegneria di Trento

Premessa. Questo scritto raccoglie alcune osservazioni perlopiù frammentarie, volte ad integrare (e non certo a sostituire) il testo di Bramanti, Pagani e Salsa, che nel seguito sarà citato semplicemente come [BPS].

Queste pagine sono in corso d'opera, e verranno aggiornate durante lo svolgimento del corso. Sono scaricabili in pdf dal sito <http://www.science.unitn.it/~visintin/> ove sono disponibili anche altre informazioni sul corso.

Indice

1. Curve
2. Limiti
3. Funzioni di Più Variabili
4. Studio del Grafico di Funzioni di Più Variabili
5. Teorema della Funzione Implicita
6. Metodo dei Moltiplicatori di Lagrange
7. Integrali Multipli
8. Operatori Differenziali Vettoriali
9. Superfici
10. Altre Osservazioni

1 Curve

Le curve sono sottoinsiemi unidimensionali dello spazio euclideo \mathbf{R}^N ($N \geq 2$); quindi sono degli enti geometrici. In analisi matematica si dispone di tre diverse forme di rappresentazione delle curve: rappresentazione parametrica, implicita, ed esplicita.

Rappresentazione Parametrica. Una funzione continua $\vec{r}: [a, b] \rightarrow \mathbf{R}^N$ ($N \geq 1$) rappresenta parametricamente una curva, il cui sostegno è $\vec{r}([a, b]) = \{\vec{r}(t) \in \mathbf{R}^N : t \in [a, b]\}$: l'insieme immagine di \vec{r} .¹ La parametrizzazione \vec{r} individua non solo il sostegno, ma anche la legge oraria del moto lungo il sostegno.

Rappresentazione Implicita. Si dice che una curva $C \subset \mathbf{R}^N$ è rappresentata implicitamente se e solo se esistono $N - 1$ funzioni continue $\varphi_1, \dots, \varphi_{N-1} : \mathbf{R}^N \rightarrow \mathbf{R}$ tali che

$$C = \{x = (x_1, \dots, x_N) \in \mathbf{R}^N : \varphi_1(x) = \varphi_{N-1}(x) = 0\}.$$

Per $N = 2$ ovviamente $C = \{(x, y) \in \mathbf{R}^2 : \varphi(x, y) = 0\}$.

Rappresentazione Esplicita. Si dice che una curva $C \subset \mathbf{R}^N$ è rappresentata esplicitamente se e solo se esistono $N - 1$ funzioni continue $f_1, \dots, f_{N-1} : \mathbf{R}^N \rightarrow \mathbf{R}$ tali che

$$C = \{x = (f_1(x_N), \dots, f_{N-1}(x_N), x_N) : x_N \in [a, b]\}.$$

Per $N = 2$ ovviamente $C = \{(f(y), y) : y \in [a, b]\}$.

¹Ricordiamo che, data una funzione $f : A \rightarrow B$ (con A e B insiemi qualsiasi non vuoti), l'insieme $\{(x, f(x)) : x \in A\}$ è detto il *grafico* di f ; l'insieme $f(A) := \{f(x) : x \in A\}$ è l'*immagine* di f .

Si può anche scambiare il ruolo delle coordinate; ad esempio per $N = 2$ possiamo anche usare la rappresentazione $C = \{(x, g(x)) : x \in [c, d]\}$ (basta disporre di una delle due rappresentazioni).

Ogni curva esplicita (ovvero rappresentata esplicitamente) ammette un'ovvia rappresentazione implicita:

$$C = \{x = (x_1, \dots, x_N) \in \mathbf{R}^N : \varphi_1(x) = \varphi_{N-1}(x) = 0\}$$

per $\varphi_1(x) = x_1 - f_1(x_N), \quad \varphi_{N-1}(x) = x_{N-1} - f_{N-1}(x_N), \quad \forall x_N \in [a, b].$

Ogni curva esplicita ammette pure un'ovvia rappresentazione parametrica:

$$C = \alpha([a, b]), \quad \text{per } \alpha(t) = (f_1(t), \dots, f_{N-1}(t), t) \quad \forall t \in [a, b].$$

Curve Regolari. Una curva (parametrica) rappresentata mediante una funzione $\vec{r} : [a, b] \rightarrow \mathbf{R}^N$ è detta *regolare* se e solo se \vec{r} è di classe C^1 e $\vec{r}'(t) \neq \vec{0}$ per ogni $t \in [a, b]$.

La condizione di non annullamento della derivata serve ad individuare il versore tangente (quindi la retta tangente) ad ogni istante, escludendo la presenza di spigoli. Ad esempio la funzione $\vec{r} : [-1, 1] \rightarrow \mathbf{R}^2 : t \mapsto (t^3, |t^3|)$ è di classe C^1 , ma il suo grafico coincide con il grafico cartesiano della funzione $x : [-1, 1] \rightarrow \mathbf{R} : x \mapsto |x|$, che presenta uno spigolo nell'origine. D'altra parte $\vec{r}'(0) = (0, 0)$.

Come si spiega tale fatto? La curva parametrica può essere interpretata come la legge oraria di un punto materiale che si sposta lungo tale grafico, e che nell'origine ha velocità nulla; in tal modo il punto riesce a mutare istantaneamente la direzione del suo moto, senza introdurre alcuna discontinuità nella velocità.

2 Limiti

Funzioni Vettoriali. Si denoti con \hat{e}_i il versore ² dell'asse i -esimo, ovvero

$$\hat{e}_1 := (1, 0, \dots, 0), \quad \hat{e}_2 := (0, 1, \dots, 0), \quad \dots, \quad \hat{e}_n := (0, \dots, 0, 1).$$

Ogni vettore $\vec{f} \in \mathbf{R}^N$ ($N \geq 2$) può essere rappresentato nella forma $\vec{f} = \sum_{i=1}^N (\vec{f} \cdot \hat{e}_i) \hat{e}_i$, ove $f_i := \vec{f} \cdot \hat{e}_i$ è la componente di \vec{f} nella direzione i -esima. Quindi abbiamo le seguenti formule di trasformazione:

$$\vec{f} = \sum_{i=1}^N f_i \hat{e}_i, \quad f_i := \vec{f} \cdot \hat{e}_i \quad \text{per } i = 1, \dots, N, \quad \forall \vec{f} \in \mathbf{R}^N. \quad (2.1)$$

Ovviamente la stessa rappresentazione vale per le funzioni a valori vettoriali $\vec{f} : \mathbf{R}^M \rightarrow \mathbf{R}^N$. Questo permette di definire agevolmente le operazioni lineari di limite, derivazione, integrazione indefinita, integrazione definita (ed altre) per le funzioni vettoriali. Se \vec{f} è una funzione a valori vettoriali e Λ è un tale operatore lineare (ad esempio, il limite, la derivata, l'integrale, ecc.), allora si pone

$$\Lambda(\vec{f}) := \sum_{i=1}^N \Lambda(f_i) \hat{e}_i, \quad (2.2)$$

qualora $\Lambda(f_i)$ sia definito per ogni i . Ovvero si pone $\Lambda(\sum_{i=1}^N f_i \hat{e}_i) := \sum_{i=1}^N \Lambda(f_i) \hat{e}_i$, permutando l'operatore Λ con la somma, il che è sicuramente lecito poiché si tratta di una somma *finita*.

Limiti. Sia dato un campo $f : \mathbf{R}^N \rightarrow \mathbf{R}$ ($N \geq 1$). Per quanto appena esposto, il discorso si estenderà in modo ovvio ai campi a valori vettoriali.

Per ogni $x_0 \in A$, $\lim_{x \rightarrow x_0} f(x) = \ell$ con $\ell \in \mathbf{R} \cup \{\pm\infty\}$ se e solo se $\lim_{t \rightarrow t_0} f(r(t)) = \ell$, per ogni curva parametrica $r : [a, b] \rightarrow \mathbf{R}^N$ tale che $r(t_0) = x_0$ per un opportuno $t_0 \in [a, b]$.

²Indicherò con il *cappello* i versori, ovvero i vettori di modulo 1. In questo paragrafo mettiamo in evidenza le quantità vettoriali mediante la freccia. Nel seguito invece spesso ometteremo la freccia.

In questa definizione si fa riferimento al limite per $t \rightarrow t_0$. Il [BPS] invece parla di convergenza lungo una successione: $\lim_{x \rightarrow x_0} f(x) = \ell$ con $\ell \in \mathbf{R} \cup \{\pm\infty\}$ se e solo se $\lim_{n \rightarrow +\infty} f(z_n) = \ell$, per ogni successione $\{z_n\}$ in \mathbf{R}^N tale che $\lim_{n \rightarrow +\infty} z_n = x_0$. Si comprende facilmente che le due definizioni sono equivalenti.

Come già visto nel corso di Analisi I, il limite prescinde dal comportamento di f in x_0 . Si dice che f è continua in x_0 se e solo se $\lim_{x \rightarrow x_0} f(x) = f(x_0)$; si dice che f è continua tout court se e solo se f è continua in ogni punto.

Dalla precedente definizione di limite discende un ovvio criterio per la non esistenza del limite.

Teorema 2.1 (Criterio per la non esistenza del limite in \mathbf{R}^N) *Se esistono due curve parametriche $r_1, r_2 : [a, b] \rightarrow \mathbf{R}^N$ tali che $r_1(t_0) = r_2(t_0) = x_0$ per un opportuno $t_0 \in [a, b]$, e tali che $\lim_{t \rightarrow t_0} f(r_1(t)) \neq \lim_{t \rightarrow t_0} f(r_2(t))$, allora il $\lim_{x \rightarrow x_0} f(x)$ non esiste.*

Disponiamo anche di un criterio per l'esistenza del limite, per $N = 2$. Mediante una traslazione, ci si può sempre ridurre al caso in cui $(x_0, y_0) = (0, 0)$; si possono quindi trasformare le coordinate cartesiane (x, y) in coordinate polari (ρ, θ) . Pertanto ³ $f(x, y) = f(\rho \cos \theta, \rho \sin \theta) =: \tilde{f}(\rho, \theta)$. Si noti che

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y) = \lim_{\rho \rightarrow 0} \tilde{f}(\rho, \theta); \quad (2.3)$$

più precisamente, se un limite esiste allora esiste anche l'altro ed i due limiti sono uguali; se un limite non esiste allora pure l'altro non esiste.

Teorema 2.2 (Criterio per l'esistenza del limite in \mathbf{R}^2) *Sia $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ e $\ell \in \mathbf{R}$. Se esiste una funzione ⁴ $g : \mathbf{R}^+ \rightarrow \mathbf{R}^+$ tale che, definendo \tilde{f} come in (2.3),*

$$\lim_{s \rightarrow 0} g(s) = 0, \quad |\tilde{f}(\rho, \theta) - \ell| \leq g(\rho) \quad \forall (\rho, \theta), \quad (2.4)$$

allora $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = \lim_{\rho \rightarrow 0} \tilde{f}(\rho, \theta) = \ell$.

Si noti che l'esistenza di $\lim_{\rho \rightarrow 0} \tilde{f}(\rho, \theta) = \ell$ per ogni θ non basta a garantire l'esistenza di una g come richiesto in (2.5), e quindi non garantisce l'esistenza del limite.

Non esiste una strategia ottimale per comprendere se un limite esiste o meno. Può comunque essere utile sondare la possibilità di applicare il criterio per la non esistenza del limite utilizzando una classe abbastanza ampia di curve passanti per il punto x_0 . Ad esempio, per $N = 2$, nel piano di coordinate (x, y) , si può provare con le rette, o le parabole passanti per (x_0, y_0) ; o più in generale con curve di equazione esplicita della forma $y - y_0 = \lambda(x - x_0)^\alpha$ (oppure $x - x_0 = \lambda(y - y_0)^\alpha$, a secondo del caso), con λ ed α costanti reali non nulle da determinare.

Un Esempio. Si studi l'esistenza del limite

$$L := \lim_{(x,y) \rightarrow (0,0)} \frac{x^\alpha y}{x^2 + y^b} \quad \text{per } a, b \in \mathbf{N} \text{ dati, e } b \text{ pari.}^5 \quad (2.5)$$

Ponendo $y = \lambda x^\alpha$, con λ e α parametri reali non nulli a disposizione, ci riduciamo al calcolo del limite

$$\tilde{L} := \lim_{x \rightarrow 0^+} \frac{\lambda x^{a+\alpha}}{x^2 + \lambda^b x^{\alpha b}}$$

³I simboli “:=” e “=” indicano delle definizioni. Qui si sta definendo la \tilde{f} . (Si noti la “tilde” \sim : non è la stessa f !)

⁴Si indica con \mathbf{R}^+ l'insieme dei numeri reali non negativi.

⁵Se b non fosse pari, il denominatore si annullerebbe anche in punti vicini all'origine.

Ci limitiamo ad $x > 0$, poichè la funzione $x^{a+\alpha}$ potrebbe avere esponente frazionario; in tal caso le potenze con base negativa non sarebbero numeri reali. Questo significa che consideriamo solo *semicurve*, che hanno un estremo nell'origine. Ponendo $\alpha = 2/b$ ($\in \mathbf{N}$) otteniamo

$$\tilde{L} = \lim_{x \rightarrow 0^+} \frac{\lambda x^{a+2/b}}{x^2(1+\lambda^b)} = \lim_{x \rightarrow 0^+} x^{a+2/b-2} \frac{\lambda}{1+\lambda^b}.$$

Se $a + 2/b - 2 < 0$ allora $\tilde{L} = +\infty$; quindi L non esiste.

Se $a + 2/b - 2 = 0$ allora $\tilde{L} = \lambda/(1 + \lambda^b)$. Questo limite dipende da λ ; per il Teorema 2.1 quindi L non esiste.

Se $a + 2/b - 2 > 0$ allora $\tilde{L} = 0$. Questo limite è finito e non dipende da λ ; non possiamo quindi applicare il Teorema 2.1.

In quest'ultimo caso è naturale cercare di applicare il Teorema 2.2. Per usare le coordinate polari, conviene ridursi ad un denominatore potenza di ρ . A questo scopo, introduciamo prima la trasformazione di coordinate

$$X = x, \quad Y = |y|^{b/2} \text{sign}(y) \quad \text{per } y \neq 0, \quad Y = 0 \quad \text{per } y = 0,$$

la cui trasformazione inversa è

$$x = X, \quad y = |Y|^{2/b} \text{sign}(Y) \quad \text{per } Y \neq 0, \quad y = 0 \quad \text{per } Y = 0.$$

Pertanto

$$f(x, y) := \frac{x^a y}{x^2 + y^b} = \frac{X^a |Y|^{2/b} \text{sign}(Y)}{X^2 + Y^2} =: F(X, Y).$$

Introduciamo ora le coordinate polari nel piano X, Y , ottenendo

$$\begin{aligned} F(X, Y) &= \rho^{a+2/b} (\cos \theta)^a (\sin \theta)^{2/b} \text{sign}(\sin \theta) / \rho^2 \\ &= \rho^{a+2/b-2} (\cos \theta)^a (\sin \theta)^{2/b} \text{sign}(\sin \theta) =: \tilde{F}(\rho, \theta); \end{aligned}$$

pertanto $|\tilde{F}(\rho, \theta)| \leq \rho^{a+2/b-2}$. Ricordando che abbiamo supposto $a + 2/b - 2 > 0$, possiamo ora applicare il Teorema 2.2 ponendo $g(\rho) = \rho^{a+2/b-2}$ e $\ell = 0$. Quindi $L = 0$.

Riassumendo:

$$\text{se } a + 2/b \leq 2 \text{ allora } L \text{ non esiste;} \quad \text{se } a + 2/b > 2 \text{ allora } L = 0.$$

3 Funzioni di Più Variabili

*** Elementi di Topologia.** La geometria euclidea può essere considerata come la geometria dei corpi rigidi, poiché studia proprietà che sono invarianti per rotazioni e traslazioni. Per contro la topologia è quella branca della geometria che studia proprietà invarianti per deformazioni continue, quindi può essere pensata come la geometria dei corpi deformabili. Le proprietà che resistono a questa classe più vasta di trasformazioni sono una classe più ristretta di quelle euclidee, e quindi possono essere considerate come più fondamentali.

Elementi di Topologia di \mathbf{R}^N . Per ogni $x_0 \in \mathbf{R}^N$ ($N \geq 1$) ed ogni $a > 0$, si dice *palla* (aperta) di centro x_0 e raggio a l'insieme $B(x_0, a) := \{x \in \mathbf{R}^N : \|x - x_0\| < a\}$.

Sia $A \subset \mathbf{R}^N$ ed $x \in \mathbf{R}^N$. Si dice che x è *interno* ad A se e solo se esiste una palla di centro x interamente inclusa in A . Si dice che x è *esterno* ad A se e solo se x è interno ad $\mathbf{R}^N \setminus A$ (il complementare di A). La *frontiera* (o bordo) di A è l'insieme dei punti che non sono né interni né esterni ad A , ed è indicata con ∂A .

Si dice *aperto* un insieme in cui tutti i punti sono interni, mentre si dice *chiuso* un insieme il cui complementare è aperto. Si verifica facilmente che un insieme è aperto se e solo se non contiene alcun punto della sua frontiera, e che è chiuso se e solo se contiene tutta la sua frontiera. Si dice interno di A l'insieme dei suoi punti interni; questo è un aperto. Si dice chiusura di A l'unione di A con la sua frontiera; questo è un chiuso. Si dice *intorno* di un punto un qualsiasi aperto che contiene il punto.

In \mathbf{R}^N , per convenzione \emptyset ed \mathbf{R}^N stesso sono sia aperti che chiusi; essi sono i soli insiemi a godere di tale proprietà. Si osservi che un insieme può non essere né aperto né chiuso.

Continuità e Differenziabilità. Sia Ω un sottoinsieme aperto di \mathbf{R}^N ($N \geq 1$), $f : \Omega \rightarrow \mathbf{R}$, $x_0 \in \Omega$.

Ricordiamo la definizione di differenziabilità. Si dice che f è differenziabile in x_0 se e solo se esiste una funzione lineare L (detta *differenziale*) tale che

$$f(x_0 + h) = f(x_0) + L(h) + o(h) \quad \text{per } h \rightarrow 0,$$

ovvero, per definizione di “o piccolo”,⁶

$$E(x_0; h) := \frac{f(x_0 + h) - f(x_0) - L(h)}{|h|} \rightarrow 0 \quad \text{per } h \rightarrow 0.$$

Se questo vale, allora la funzione L è detta il *differenziale di f in x_0* , e si scrive $L = df(x_0)$.⁷

Si considerino le seguenti proprietà:

(C) f è globalmente continua in x_0 (ovvero $\lim_{x \rightarrow x_0} f(x) = f(x_0)$),

(CD) f è continua in x_0 in tutte le direzioni,

(DD) f è derivabile in x_0 in tutte le direzioni,

(D) f è differenziabile in x_0 .

Si ha

$$(D) \Rightarrow (DD) \Rightarrow (CD), \quad (D) \Rightarrow (C) \Rightarrow (CD),$$

pertanto $(D) \Rightarrow (CD)$. Nessuna di queste implicazioni è invertibile; inoltre

$$(C) \not\Rightarrow (DD), \quad (DD) \not\Rightarrow (C).$$

Infatti una proprietà di continuità (quale (C) o (CD)) non può implicarne una di derivabilità (quale (D) o (DD)); ed una proprietà di tipo direzionale (quale (CD) o (DD)) non può implicarne una globale (quale (C) o (D)).

Queste affermazioni valgono anche per i campi vettoriali (ovvero per $\vec{f} : \Omega \rightarrow \mathbf{R}^M$ per $M > 1$). Infatti, indicando con (P) una qualsiasi delle quattro proprietà di sopra, si ha:

(P) vale per \vec{f} se e solo se essa vale per f_i , per ogni $i \in \{1, \dots, N\}$.

Condizioni Sufficienti per la Differenziabilità. Come si è visto, l'esistenza di tutte le derivate direzionali in un punto non implica la differenziabilità in quel punto. Comunque vale il seguente risultato.

Teorema 3.1 *Sia A un aperto di \mathbf{R}^N e $f : A \rightarrow \mathbf{R}$. Se esistono tutte le derivate parziali prime di f in una palla $B(x_0, r) \subset A$ e queste sono continue in x_0 , allora f è differenziabile in x_0 .*

Corollario 3.2 *Sia A un aperto di \mathbf{R}^N e $f : A \rightarrow \mathbf{R}$. Se tutte le derivate parziali prime di f esistono e sono continue in A (ovvero, se $f \in C^1(A)$),⁸ allora f è differenziabile in tutti i punti di A .*

⁶Ricordiamo che, date due funzioni definite in un intervallo che contiene x_0 , Si dice che “ $f(x)$ è o piccolo di $g(x)$ ” se e solo se $f(x)/g(x) \rightarrow 0$ per $x \rightarrow x_0$.

⁷Questa definizione è leggermente diversa da quella, comunque corretta, data da [BPS]: ivi si assume che $L(h) = \nabla f(x_0) \cdot h$. In effetti, quest'ultima formula può essere verificata a posteriori.

⁸Il concetto di funzione di classe C^k si estende alle funzioni di più variabili nel modo seguente.

Sia $k \in \mathbf{N}$ ed $A \subset \mathbf{R}^N$ (aperto). Si dice che una funzione f è di classe C^k (e si scrive $f \in C^k(A)$) se e solo se f ammette tutte le derivate parziali fino all'ordine k , e tutte queste funzioni sono continue in A . Se f ammette derivate parziali di ogni ordine, f è detta di classe C^∞ .

Per la dimostrazione del corollario basta osservare che per ogni punto di un aperto esiste una palla centrata in quel punto ed interamente contenuta nell'aperto.

Questo teorema è particolarmente utile poiché di solito l'ipotesi “ $f \in C^1(A)$ ” è di ben più facile verifica che la differenziabilità, ed è soddisfatta in un gran numero di casi.

*** Una Funzione Differenziabile con Derivata Discontinua.** Il teorema precedente non può essere invertito, come mostra il seguente controesempio.

Sia $f : \mathbf{R} \rightarrow \mathbf{R}$ definita da $f(x) := x^2 \sin(1/x)$ per ogni $x \neq 0$, $f(0) := 0$.

f è differenziabile in 0, infatti

$$f'(0) = \lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h} = \lim_{h \rightarrow 0} \frac{h^2 \sin(1/h)}{h} = \lim_{h \rightarrow 0} h \sin(1/h) = 0.$$

D'altra parte f' non è continua in 0; infatti per ogni $x \neq 0$, $f'(x) = 2x \sin(1/x) - \cos(1/x)$, e questa funzione non ammette limite per $x \rightarrow 0$.

*** La Formula del Differenziale.** Per $i = 1, \dots, N$ si ponga⁹ $P_i : \mathbf{R}^N \rightarrow \mathbf{R} : \vec{x} \mapsto x_i$, la proiezione sull'asse i -esimo. Tale funzione è lineare, quindi coincide con il suo differenziale: $dP_i(\vec{x}) = P_i(\vec{x}) = x_i$.

Applicando la linearità del differenziale e l'identità $df(\hat{e}_i) = D_i f(\vec{x})$, si ha

$$df(\vec{x}) = df\left(\sum_{i=1}^N x_i \hat{e}_i\right) = \sum_{i=1}^N x_i df(\hat{e}_i) = \sum_{i=1}^N x_i D_i f(\vec{x}) = \sum_{i=1}^N D_i f(\vec{x}) dP_i(\vec{x}).$$

Scrivendo (un po' impropriamente) $dP_i(\vec{x}) = dx_i$, si perviene alla classica formula

$$df(\vec{x}) = \sum_{i=1}^N D_i f(\vec{x}) dx_i.$$

Il Teorema di Derivazione Sotto Integrale. Questo risultato riguarda la derivazione di un integrale dipendente da un parametro (qui indicato con x).

Teorema 3.3 Sia Ω un sottoinsieme aperto di \mathbf{R}^N ($N \geq 1$), $f : \Omega \times [a, b] \rightarrow \mathbf{R}$ tale che f e ∇f siano di classe C^1 . Allora la funzione $\Omega \rightarrow \mathbf{R} : x \mapsto \int_a^b f(x, t) dt$ è di classe C^1 e

$$\frac{\partial}{\partial x_i} \int_a^b f(x, t) dt = \int_a^b \frac{\partial}{\partial x_i} f(x, t) dt \quad \text{per } i = 1, \dots, N.$$

Se $N = 1$ occorre scrivere $\frac{d}{dx} \int_a^b f(x, t) dt = \int_a^b \frac{\partial}{\partial x} f(x, t) dt$, poiché in tal caso $\int_a^b f(x, t) dt$ è una funzione della sola x .

Una Formula di Derivazione. Siano $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ e $\alpha, \beta : \mathbf{R} \rightarrow \mathbf{R}$ di classe C^1 . Allora

$$\frac{d}{dx} \int_{\alpha(x)}^{\beta(x)} f(x, y) dy = \int_{\alpha(x)}^{\beta(x)} \frac{\partial f(x, y)}{\partial x} dy - \alpha'(x) f(x, \alpha(x)) + \beta'(x) f(x, \beta(x)). \quad (3.1)$$

Per verificarlo si ponga $\Phi(x, w, z) := \int_w^z f(x, y) dy$ per ogni $(x, w, z) \in \mathbf{R}^3$, si osservi che Φ è di classe C^1 , e che

$$\begin{aligned} \frac{d}{dx} \int_{\alpha(x)}^{\beta(x)} f(x, y) dy &= \frac{d}{dx} \Phi(x, \alpha(x), \beta(x)) \\ &= D_1 \Phi(x, \alpha(x), \beta(x)) + D_2 \Phi(x, \alpha(x), \beta(x)) \alpha'(x) + D_3 \Phi(x, \alpha(x), \beta(x)) \beta'(x), \\ D_1 \Phi(x, w, z) &= \int_w^z \frac{\partial f(x, y)}{\partial x} dy, \quad D_2 \Phi(x, w, z) = -f(x, w), \quad D_3 \Phi(x, w, z) = f(x, z) \\ &\quad \forall (x, w, z) \in \mathbf{R}^3. \end{aligned}$$

⁹Qui scriviamo \vec{x} in luogo di x , mettendo in evidenza che si tratta di un vettore.

4 Studio del Grafico di Funzioni di più Variabili

Studio dei Punti Stazionari di un Campo Scalare. Sia Ω un aperto di \mathbf{R}^N , $f : \Omega \rightarrow \mathbf{R}$ di classe C^2 . La matrice Hessiana di f (qui denotata con H_f) è simmetrica, pertanto i suoi autovalori λ_i sono reali. Se x_0 è un punto stazionario di f , valgono le seguenti proprietà:

$$\forall i, \lambda_i > 0 \implies x_0 \text{ è un punto di minimo relativo,} \quad (4.1)$$

$$\forall i, \lambda_i < 0 \implies x_0 \text{ è un punto di massimo relativo,} \quad (4.2)$$

$$\exists i, \exists j : \lambda_i < 0 < \lambda_j \implies x_0 \text{ è un punto di sella.} \quad (4.3)$$

Inoltre

$$x_0 \text{ è un punto di minimo relativo} \implies \forall i, \lambda_i \geq 0, \quad (4.4)$$

$$x_0 \text{ è un punto di massimo relativo} \implies \forall i, \lambda_i \leq 0, \quad (4.5)$$

$$x_0 \text{ è un punto di sella} \implies \exists i, \exists j : \lambda_i \leq 0 \leq \lambda_j \quad (4.6)$$

(nell'ultima affermazione non si esclude $i = j$).¹⁰ Nessuna di queste implicazioni può essere invertita: le affermazioni sul comportamento della funzione non implicano le disuguaglianze con i segni stretti, d'altra parte le disuguaglianze con i segni non stretti non bastano a determinare il comportamento della funzione. Inoltre la sola informazione “ $\forall i, \lambda_i = 0$ ” (e pertanto anche la più debole “ $\exists i : \lambda_i = 0$ ”) non permette di escludere nessuno dei tre comportamenti.¹¹ Basta considerare i seguenti controesempi:

$$f_1(x, y) := x^4 + y^4, \quad f_2(x, y) := -x^4 - y^4, \quad f_3(x, y) := x^4 - y^4, \quad (4.7)$$

per i quali $(0, 0)$ è rispettivamente punto di minimo relativo, di massimo relativo, di sella. In ciascuno di questi casi $(0, 0)$ è un punto stazionario, ivi la matrice Hessiana coincide con la matrice nulla, e pertanto tutti i suoi autovalori sono nulli.¹²

Si ricorda che una matrice A è detta

definita positiva se $v \cdot A \cdot v \left(:= \sum_{i,j=1}^M v_i A_{ij} v_j \right) > 0$ per ogni $v \in \mathbf{R}^N \setminus \{0\}$,

semidefinita positiva se $v \cdot A \cdot v \geq 0$ per ogni $v \in \mathbf{R}^N$,

indefinita se esistono $v_1, v_2 \in \mathbf{R}^N$ tali che $v_1 \cdot A \cdot v_1 > 0$, $v_2 \cdot A \cdot v_2 < 0$.

I concetti di matrice *definita negativa* e quello di matrice *semidefinita negativa* sono definiti analogamente.

Per le matrici simmetriche sussistono ovvi legami tra queste proprietà della matrice ed il segno dei suoi autovalori. Ad esempio, A è definita positiva se e solo se tutti gli autovalori sono positivi; A è semidefinita positiva se e solo se tutti gli autovalori sono non negativi; A è indefinita se e solo se esistono due autovalori di segno opposto.¹³

¹⁰Si può notare che l'ultima affermazione è una conseguenza logica delle (4.1) e (4.2).

¹¹Ovviamente con questa affermazione non si esclude che dell'informazione aggiuntiva possa permettere di trarre conclusioni!

¹²Al solito, per padroneggiare i possibili scenari occorre conoscere le affermazioni *in positivo* (ovvero le implicazioni valide, qui le (4.1), ..., (4.6)) e disporre di un armamentario di casi critici (qui le f_i della (4.7)), che possono fornire controesempi per le affermazioni *in negativo*. In questo modo si ha un'idea sia delle possibilità offerte dagli strumenti a disposizione che dei loro limiti.

¹³Infatti ogni matrice simmetrica A è equivalente ad una matrice diagonale; più precisamente, esiste una matrice ortogonale Q tale che la matrice $D = Q^{-1}AQ$ sia diagonale.

Una matrice Q è detta ortogonale (qualcuno dice *ortonormale*) se e solo se è invertibile e la sua inversa coincide con la sua trasposta Q^* (Q è così denominata poiché la trasformazione associata $v \mapsto Qv$ conserva l'ortogonalità, ovvero se v_1 e v_2 sono tra loro ortogonali allora lo stesso vale per Qv_1 e Qv_2). In tal caso la trasformazione $A \mapsto Q^{-1}AQ$ rappresenta il passaggio da un sistema di coordinate cartesiane ortogonali ad un altro sistema dello stesso tipo, ottenuto ad esempio mediante una rotazione. Quindi la sostituzione di A con $Q^{-1}AQ$ corrisponde a descrivere gli stessi oggetti in un diverso sistema di coordinate cartesiane ortogonali.

Le precedenti affermazioni si possono allora riscrivere come segue in termini di proprietà di segno della matrice Hessiana:

$$H_f(x_0) \text{ è definita positiva} \implies x_0 \text{ è un punto di minimo relativo,} \quad (4.8)$$

$$H_f(x_0) \text{ è definita negativa} \implies x_0 \text{ è un punto di massimo relativo,} \quad (4.9)$$

$$H_f(x_0) \text{ è indefinita} \implies x_0 \text{ è un punto di sella.} \quad (4.10)$$

Inoltre

$$x_0 \text{ è un punto di minimo relativo} \implies H_f(x_0) \text{ è semidefinita positiva,} \quad (4.11)$$

$$x_0 \text{ è un punto di massimo relativo} \implies H_f(x_0) \text{ è semidefinita negativa.} \quad (4.12)$$

Ovviamente tutte queste affermazioni si possono scrivere in modo equivalente applicando la regola logica secondo cui “P \implies Q” è equivalente a “(non Q) \implies (non P)”.

*** Regola di Sylvester.** Sia A una matrice simmetrica di \mathbf{R}^{N^2} ; per $k = 1, \dots, N$ sia A_k la matrice di \mathbf{R}^{k^2} ottenuta sopprimendo righe e colonne di A di indice $> k$.

Teorema 4.1 (*di Sylvester*) *Nelle ipotesi di sopra, A è definita positiva se e solo se $\text{Det}(A_k) > 0$ per $k = 1, \dots, N$.*

Un Controesempio di Peano. Per stabilire la proprietà di minimo globale non basta verificarla in ogni direzione. Ad esempio sia data la funzione $f : \mathbf{R}^2 \rightarrow \mathbf{R} : (x, y) \mapsto (y - x^2)(y - 2x^2)$. Il punto $(0, 0)$ è di minimo (stretto) rispetto ad ogni direzione, ma non è punto di minimo globale.

Il Modello della Scatola Nera. Un’importante applicazione dello studio dei minimi di una funzione di più variabili è costituito dal *metodo dei minimi quadrati*. Questo metodo è di notevole utilità pratica; ad esempio è spesso impiegato per derivare relazioni funzionali approssimate da dati sperimentali. Esso può essere illustrato facendo riferimento al modello della *scatola nera*.

Quest’ultima rappresenta un *sistema*¹⁴ del quale non si conosce (né si cerca di conoscere) la struttura interna, e col quale si può interagire solo mediante due canali, uno di ingresso (*input*) ed uno di uscita (*output*). Si può investigare il comportamento di tale sistema mediante *esperimenti* in cui si forniscono dei valori di input (indicati con x), e si registrano i corrispondenti valori di output (indicati con y). Nell’ambito di una descrizione quantitativa, si assume che x ed y varino in prescritte classi numeriche (ad esempio, possono essere scalari interi o reali o complessi, o vettori di \mathbf{R}^N , ecc.).

Il sistema si può comportare in modo *deterministico* oppure *stocastico* (o *aleatorio*). Nel primo caso il valore di input determina univocamente quello di output. Nel secondo caso diversi valori di output possono corrispondere ad uno stesso valore di input, ad esempio a causa di un *rumore* che altera i valori di output; si noti che allora ha senso ripetere più esperimenti con uno stesso valore di input.¹⁵

Si pone il problema di realizzare degli esperimenti ed elaborarne i risultati, in modo da prevedere al meglio il comportamento della scatola nera per esperimenti successivi. In particolare, dato un insieme di coppie input-output (x_i, y_i) , occorre congetturare una relazione funzionale tra input ed output, ovvero rappresentare y come funzione di x . È qui che interviene il metodo dei minimi quadrati.

Il Metodo dei Minimi Quadrati. Sia dato un insieme finito di punti $\{(x_i, y_i)\}_{i=1, \dots, M} \subset \mathbf{R}^2$. In una classe di funzioni \mathcal{F} da specificare, si cerchi una funzione $f : \mathbf{R} \rightarrow \mathbf{R}$ il cui grafico passi per tali punti, ovvero tale che $y_i = f(x_i)$ per ogni i . Se vi fossero delle coppie con la stessa x_i e diverse y_i

¹⁴In questa breve digressione non definiremo il concetto di *sistema*, che per altro è ormai divenuto di uso corrente.

¹⁵Si assume che il rumore sia abbastanza debole e per così dire *imparziale*, e che quindi *si compensi* tra i diversi esperimenti. Naturalmente tutto questo andrebbe precisato!

(ad esempio (1, 2) e (1, 3)), questo problema non avrebbe soluzione. Comunque, anche se ciò fosse escluso, per M grande la classe \mathcal{F} dovrebbe essere alquanto ampia per contenere una soluzione; ad esempio, se si cercasse f tra i polinomi, in generale questi dovrebbero avere grado $m = M - 1$, ovvero f dipenderebbe da M parametri.

È più conveniente formulare un problema diverso, accontentandosi di una soluzione approssimata; d'altra parte di solito pure i dati (x_i, y_i) sono solo approssimati. Pertanto si cerca in una classe ristretta la funzione che meglio approssima i punti dati, ovvero che minimizza una certa *distanza* tra le y_i e le $f(x_i)$. Come misura di tale distanza si assume l'*errore quadratico totale* $S := \sum_i \varepsilon_i^2$ (nelle somme si sottintenderà che l'indice varia tra 1 e M), ove $\varepsilon_i := y_i - f(x_i)$ rappresenta l'errore dell'esperimento i -esimo.

Ad esempio si potrebbe cercare f lineare (più precisamente *affine*): $f(x; a, b) = ax + b$. Il problema di determinare i parametri $a, b \in \mathbf{R}$ in modo tale che essi minimizzino $S(a, b) := \sum_i (y_i - ax_i - b)^2$ è detto *regressione lineare*, e la retta di equazione $y = ax + b$ è detta *retta di regressione*. Si può mostrare che basta minimizzare S in un opportuno insieme chiuso e limitato.¹⁶ Poiché S è continua, esiste almeno un punto (a, b) di minimo assoluto. Ivi

$$\frac{\partial S}{\partial a}(a, b) = 0, \quad \frac{\partial S}{\partial b}(a, b) = 0,$$

ovvero

$$\begin{cases} a \sum_i x_i^2 + b \sum_i x_i = \sum_i x_i y_i \\ a \sum_i x_i + bM = \sum_i y_i \end{cases}$$

(si è utilizzato il fatto che $\sum_i b = Mb$). Questo sistema ha soluzione

$$a = \frac{M \sum_i x_i y_i - \sum_i x_i \sum_j y_j}{M \sum_i x_i^2 - (\sum_i x_i)^2}, \quad b = \frac{\sum_i x_i^2 \sum_j y_j - \sum_i x_i \sum_j x_j y_j}{M \sum_i x_i^2 - (\sum_i x_i)^2}, \quad (4.13)$$

se il comune denominatore $M \sum_i x_i^2 - (\sum_i x_i)^2$ non si annulla.

* Quest'ultima eventualità si verifica se e solo se tutti gli x_i sono uguali, poiché

$$M \sum_i \left(x_i - \frac{1}{M} \sum_j x_j \right)^2 = M \sum_i x_i^2 - 2 \sum_i x_i \sum_j x_j + \frac{1}{M} \sum_i \left(\sum_j x_j \right)^2 = M \sum_i x_i^2 - \left(\sum_i x_i \right)^2,$$

che ovviamente si annulla se e solo se $x_i = \tilde{x} := \sum_j x_j / M$ per ogni i . In tal caso il sistema ha infinite soluzioni. Infatti l'informazione rappresentata dagli $f(x_i)$ ($i = 1, \dots, M$) determina un valore ottimale \tilde{y} di $f(x)$ solo per $x = \tilde{x}$; quindi ogni retta che passa per il punto (\tilde{x}, \tilde{y}) rappresenta una funzione che minimizza l'errore quadratico medio.

Si può poi verificare facilmente che la matrice Hessiana di S è definita positiva nel punto (a, b) definito da (4.13), a meno che (ancora una volta) tutti gli x_i siano uguali. In tal modo si ritrova il fatto che (a, b) è un punto di minimo relativo per S .

* **Generalizzazioni.** Il metodo di sopra può essere generalizzato in vari modi.

(i) Siano $X_i \in \mathbf{R}^P$ e $Y_i \in \mathbf{R}^Q$. In tal caso il problema della regressione lineare consiste nella determinazione di una matrice A di Q righe e P colonne e di un vettore $B \in \mathbf{R}^Q$ che minimizzino

$$S(A, B) := \sum_i \|Y_i - AX_i - B\|^2 = \sum_i \sum_{\ell=1}^Q \left(Y_{i\ell} - \sum_{m=1}^P A_{\ell m} X_{im} - B_\ell \right)^2.$$

¹⁶ * Infatti, a meno che i punti (x_i, y_i) non siano tutti allineati, $S(a, b) \rightarrow +\infty$ se $a \rightarrow +\infty$ e $b \rightarrow +\infty$ congiuntamente. Quindi per minimizzare S non occorre andare all'infinito (in \mathbf{R}^2) ...

Derivando rispetto a ciascuno dei parametri A_{ij} , B_j si ottiene un sistema di $QP + Q$ equazioni nelle incognite $A_{\ell m}$, B_ℓ .

(ii) Si possono considerare diverse forme di regressione non lineare. La funzione affine $f(x; a, b) = ax + b$ può essere sostituita da un polinomio di un prefissato grado m :

$$f(x; a_0, \dots, a_m) := a_m x^m + a_{m-1} x^{m-1} + \dots + a_0,$$

nel qual caso si minimizza la funzione

$$S(a_0, \dots, a_m) := \sum_i \left(y_i - a_m x_i^m - a_{m-1} x_i^{m-1} - \dots - a_0 \right)^2.$$

Derivando rispetto a ciascuno dei parametri si ottiene un sistema di $m + 1$ equazioni lineari nelle incognite a_0, \dots, a_m .

(iii) Più in generale si potrebbero considerare funzioni non lineari della forma

$$f(x; a_0, \dots, a_m) := a_1 f_1(x) + \dots + a_m f_m(x),$$

con f_1, \dots, f_m funzioni date. La linearità rispetto ai coefficienti a_1, \dots, a_m riduce il problema della loro identificazione alla risoluzione di un sistema lineare (con una matrice che dipende dai dati $y_{i1} = f_1(x_i), \dots, y_{im} = f_m(x_i)$, per $i = 1, \dots, M$).

5 Teorema della Funzione Implicita

Incominciamo enunciando il teorema nella sua forma più semplice.

Teorema 5.1 (*Teorema della Funzione Implicita in \mathbf{R}^2 , di Dini*) Sia Ω un aperto di \mathbf{R}^2 ed $f : \Omega \rightarrow \mathbf{R}$ una funzione continua dotata di derivata continua nella direzione y . Sia $(x^0, y^0) \in \Omega$ tale che

$$f(x^0, y^0) = 0, \quad \frac{\partial f}{\partial y}(x^0, y^0) \neq 0. \quad (5.1)$$

Allora esistono

- (i) un intorno U di x^0 in \mathbf{R} ,
- (ii) un intorno V di y^0 in \mathbf{R} ,
- (iii) una funzione $g : U \rightarrow V$ continua,

tali che

$$\{(x, y) \in U \times V : f(x, y) = 0\} = \{(x, g(x)) : x \in U\}. \quad (5.2)$$

Inoltre, se f è di classe C^1 in Ω , allora g è di classe C^1 in U e

$$\frac{dg}{dx}(x) = - \frac{\frac{\partial f}{\partial x}(x, g(x))}{\frac{\partial f}{\partial y}(x, g(x))} \quad \forall x \in U. \quad (5.3)$$

Questo risultato ha solo carattere *locale*, ovvero riguarda solo il comportamento in un opportuno intorno del punto (x^0, y^0) .

La formula finale discende immediatamente dalla identità $f(x, g(x)) = 0$ per ogni $x \in U$, che definisce la g *implicitamente*. Infatti, posto $\varphi(x) := f(x, g(x))$, si ha $\varphi(x) = 0$ per ogni $x \in U$, da cui derivando

$$0 = \varphi'(x) = \frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x)) \frac{dg}{dx}(x), \quad (5.4)$$

che ovviamente fornisce la (5.3).

Questo risultato può essere facilmente esteso al caso di funzioni scalari di più variabili.

Teorema 5.2 (Teorema della Funzione Implicita in \mathbf{R}^N) Sia Ω un aperto di \mathbf{R}^N ed $f : \Omega \rightarrow \mathbf{R}$ una funzione continua dotata di derivata continua nella direzione x_N .

Sia $(x_1^0, \dots, x_N^0) \in \Omega$ tale che

$$f(x_1^0, \dots, x_N^0) = 0, \quad \frac{\partial f}{\partial x_N}(x_1^0, \dots, x_N^0) \neq 0. \quad (5.5)$$

Allora esistono

- (i) un intorno U di $(x_1^0, \dots, x_{N-1}^0)$ in \mathbf{R}^{N-1} ,
- (ii) un intorno V di x_N^0 in \mathbf{R} ,
- (iii) una funzione $g : U \rightarrow V$ continua,

tali che

$$\begin{aligned} & \{(x_1, \dots, x_N) \in U \times V : f(x_1, \dots, x_N) = 0\} \\ & = \{(x_1, \dots, x_{N-1}, g(x_1, \dots, x_{N-1})) : (x_1, \dots, x_{N-1}) \in U\}. \end{aligned} \quad (5.6)$$

Inoltre, se f è di classe C^1 in Ω , allora g è di classe C^1 in U e

$$\frac{\partial g}{\partial x_i}(x_1, \dots, x_{N-1}) = - \frac{\frac{\partial f}{\partial x_i}(x_1, \dots, x_{N-1}, g(x_1, \dots, x_{N-1}))}{\frac{\partial f}{\partial x_N}(x_1, \dots, x_{N-1}, g(x_1, \dots, x_{N-1}))} \quad \forall (x_1, \dots, x_{N-1}) \in U, \quad (5.7)$$

per $i = 1, \dots, N-1$.

Anche qui la formula finale discende immediatamente dalla identità

$$f(x_1, \dots, x_{N-1}, g(x_1, \dots, x_{N-1})) = 0 \quad \forall (x_1, \dots, x_{N-1}) \in U,$$

che definisce la g implicitamente.

Una ulteriore importante generalizzazione (qui omessa) riguarda funzioni vettoriali di più variabili.

Osservazioni. (i) Il Teorema 5.1 si generalizza in modo ovvio al caso in cui la condizione $f(x^0, y^0) = 0$ è sostituita da $f(x^0, y^0) = C$, con $C \in \mathbf{R}$ fissato. In tal caso, per ottenere una rappresentazione cartesiana locale dell'insieme $\{(x, y) : f(x, y) = C\}$, basta applicare il Teorema 5.1 alla funzione $\tilde{f}(x, y) := f(x, y) - C$. Un'analogia estensione vale per il Teorema 5.2.

(ii) Un risultato analogo al Teorema 5.1 vale se la condizione $\frac{\partial f}{\partial y}(x^0, y^0) \neq 0$ è sostituita da $\frac{\partial f}{\partial x}(x^0, y^0) \neq 0$. In tal caso in un intorno del punto (x^0, y^0) si rappresenterà l'insieme $\{(x, y) : f(x, y) = 0\}$ come grafico di una funzione $y \mapsto x$. Questo è un'ovvia conseguenza della simmetria del ruolo giocato dalle variabili x e y nell'enunciato del teorema. Anche in questo caso un'analogia estensione vale per il Teorema 5.2.

(iii) Questi due teoremi non sono invertibili. Più precisamente, riferendoci ad esempio al Teorema 5.1, la possibilità di rappresentare esplicitamente la funzione definita implicitamente (ovvero l'esistenza di una funzione g tale che $f(x, g(x)) = 0$ per ogni $x \in U$) non implica la condizione sulla $\frac{\partial f}{\partial y}$. Ad esempio sia $f(x, y) := y^3$ e $(x^0, y^0) := (0, 0)$. Allora l'insieme $\{(x, y) : f(x, y) = 0\}$ coincide con l'asse delle x , ovvero è il grafico della funzione $x \mapsto y = 0$; nondimeno $\frac{\partial f}{\partial y}(0, 0) = 0$.

(iv) Nel Teorema 5.1 (ed analogamente nel Teorema 5.2) la condizione $\frac{\partial f}{\partial y}(x^0, y^0) \neq 0$ è irrinunciabile, come testimoniato dai seguenti controesempi.

Sia $f(x, y) = 0$ per ogni $(x, y) \in \mathbf{R}^2$. In questo caso in un intorno di qualsiasi punto (x^0, y^0) l'insieme $\{(x, y) : f(x, y) = 0\}$ non è rappresentabile con una curva, poiché coincide con un insieme aperto (addirittura con tutto \mathbf{R}^2). Ed infatti $\frac{\partial f}{\partial y}(x^0, y^0) = 0$.

Un controesempio meno banale è costituito dalla funzione $f(x, y) := x^2 - y^2$ per ogni $(x, y) \in \mathbf{R}^2$, che rappresenta l'unione di due rette, la bisettrice del primo e del terzo quadrante e quella del secondo e

del quarto quadrante. Nell'origine l'insieme in cui $f = 0$ non è rappresentabile come grafico cartesiano in alcun intorno dell'origine, ed infatti $\nabla f(0, 0) = (0, 0)$.

Curve Regolari in Forma Implicita. Per quanto visto, se in un punto $(x^0, y^0) \in \Omega$ si ha $f(x^0, y^0) = 0$ e $\nabla f(x^0, y^0) \neq \vec{0}$, allora in un intorno di tale punto l'insieme $\{(x, y) : f(x, y) = 0\}$ è costituito da una curva rappresentabile esplicitamente e regolare in quel punto, nella forma esplicita $y = g(x)$ o $x = g(y)$ a seconda di quale componente del gradiente non si annulla (in entrambe le forme se entrambe le componenti non si annullano).

Si dice allora che una curva $C = \{(x, y) : f(x, y) = 0\}$ è *regolare* in un punto $(x^0, y^0) \in \Omega$ se e solo se

$$f \in C^1, \quad \nabla f(x^0, y^0) \neq \vec{0}, \quad (5.8)$$

e che C è regolare se e solo se è regolare in ogni punto.

6 Metodo dei Moltiplicatori di Lagrange

Vincoli Scalari. Si consideri una curva C del piano rappresentata in forma implicita da una funzione $g : \mathbf{R}^2 \rightarrow \mathbf{R}$ di classe C^1 ; ovvero $C := \{(x, y) \in \mathbf{R}^2 : g(x, y) = 0\}$. Sia $(\tilde{x}, \tilde{y}) \in S$ un punto regolare, ovvero tale che $\nabla g(\tilde{x}, \tilde{y}) \neq \vec{0}$. Per il teorema della funzione implicita C può essere rappresentata in forma esplicita in un intorno di (\tilde{x}, \tilde{y}) ; ad esempio, se $D_2 g(\tilde{x}, \tilde{y}) \neq 0$ allora ivi C è il grafico cartesiano di una funzione $y = h(x)$, con h di classe C^1 (ovviamente $\tilde{y} = h(\tilde{x})$).

Sia $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ di classe C^1 . Si ponga $\varphi(x) := f(x, h(x))$ in un intorno U di \tilde{x} , e si noti che anche φ è di classe C^1 .

$(\tilde{x}, \tilde{y}) \in C$ è detto *punto stazionario (vincolato) per la funzione f sulla curva C* se e solo se $\varphi'(\tilde{x}) = 0$. Si verifica facilmente che se (\tilde{x}, \tilde{y}) è un punto di massimo (o minimo) locale per f su C , allora \tilde{x} è un punto di massimo (o minimo) locale per φ , quindi $\varphi'(\tilde{x}) = 0$, ovvero (\tilde{x}, \tilde{y}) è un punto stazionario per f su C . Per contro è ovvio che possono esistere dei punti stazionari per f su C che non sono né di massimo né di minimo locale per f .

Si noti che $\varphi'(x) := \nabla f(x, h(x)) \cdot (1, h'(x))$, e che $(1, h'(x))$ è tangente a C in (x, y) . Quindi $\varphi'(\tilde{x}) = 0$ se e solo se $\nabla f(\tilde{x}, \tilde{y})$ è ortogonale a C in (\tilde{x}, \tilde{y}) . Pure $\nabla g(\tilde{x}, \tilde{y})$ è ortogonale a C in (\tilde{x}, \tilde{y}) , pertanto si ha il seguente risultato.

Teorema 6.1 *Si assumano le precedenti ipotesi (in particolare, sia $\nabla g(\tilde{x}, \tilde{y}) \neq (0, 0)$). Allora $(\tilde{x}, \tilde{y}) \in C$ è un punto stazionario (vincolato) per f su C se e solo se $\nabla f(\tilde{x}, \tilde{y})$ è parallelo a $\nabla g(\tilde{x}, \tilde{y})$, ovvero se e solo se*

$$\exists \tilde{\lambda} \in \mathbf{R} : \nabla f(\tilde{x}, \tilde{y}) = \tilde{\lambda} \nabla g(\tilde{x}, \tilde{y}). \quad (6.1)$$

(Un tale λ esiste necessariamente, poiché $\nabla g(\tilde{x}, \tilde{y}) \neq (0, 0)$.)

Come già detto, l'appartenenza di (\tilde{x}, \tilde{y}) a C corrisponde a $g(\tilde{x}, \tilde{y}) = 0$. Quest'ultima condizione e (6.1) si possono esprimere elegantemente dicendo che in $(\tilde{x}, \tilde{y}, \tilde{\lambda})$ la funzione

$$\mathcal{L}(x, y, \lambda) := f(x, y) - \lambda g(x, y) \quad (6.2)$$

ha un punto stazionario (non vincolato). \mathcal{L} e $\tilde{\lambda}$ sono rispettivamente detti *funzione Lagrangiana* e *moltiplicatore di Lagrange*.

L'esistenza dei moltiplicatori di Lagrange dipende non solo dai vincoli ma anche dalla loro rappresentazione analitica, ovvero dalla scelta delle funzioni g_i .

Infatti, se g è tale che $\nabla g(\tilde{x}, \tilde{y}) \neq \vec{0}$ allora sappiamo che esiste il moltiplicatore di Lagrange. Si ponga ora $\hat{g} := g^2$; allora $\hat{g}(x, y) = 0$ se e solo se $g(x, y) = 0$, ma in tali punti $\nabla \hat{g} = 2g \nabla g = \vec{0}$. Cioè usando la rappresentazione in termini di \hat{g} si ottiene lo stesso vincolo, ma viene meno la condizione di non annullamento del gradiente, e pertanto il metodo di Lagrange non è applicabile.

7 Integrali Multipli

A Proposito del Cambiamento di Variabile negli Integrali Multipli. Si consideri la trasformazione di coordinate definita dalle seguenti relazioni tra le “vecchie” coordinate (x, y) e le “nuove” coordinate (u, v) :

$$x = X(u, v), \quad y = Y(u, v), \quad (7.1)$$

con X, Y funzioni note di classe C^1 . Pertanto

$$dx = X_u du + X_v dv, \quad dy = Y_u du + Y_v dv. \quad (7.2)$$

Sia $J = J(u, v)$ la corrispondente matrice jacobiana. Supporremo che il determinante jacobiano, denotato $\det(J(u, v))$, si annulli al più in un numero finito di punti.

Le coordinate (x, y) saranno supposte cartesiane ed ortogonali, mentre le coordinate (u, v) saranno supposte solo ortogonali (il che significa che in ogni punto le linee coordinate sono ortogonali). Questo implica che $\hat{x} \wedge \hat{y}$ e $\hat{u} \wedge \hat{v}$ sono dei versori. Ad esempio le coordinate polari (ρ, θ) sono ortogonali, ed ovviamente non cartesiane.

Denotiamo $\tilde{\Omega}$ il dominio nel sistema di coordinate (u, v) che è rappresentato da Ω nel sistema di coordinate (x, y) , e poniamo

$$\tilde{f}(u, v) := f(X(u, v), Y(u, v)) \quad \forall (u, v) \in \tilde{\Omega}.$$

Questa è la formula di trasformazione degli integrali doppi:

$$\iint_{\Omega} f(x, y) dx dy = \iint_{\tilde{\Omega}} \tilde{f}(u, v) |\det(J(u, v))| du dv. \quad (7.3)$$

Questo risultato si estende in modo ovvio agli integrali tripli, quadrupli, ecc..

*** Una Giustificazione Sintetica della Formula (7.3).** Il vettore che nel sistema di coordinate (x, y) è rappresentato da $dx \hat{x}$ nel nuovo sistema di coordinate (u, v) è rappresentato da $X_u du \hat{u} + X_v dv \hat{v}$. Analogamente, il vettore che nel sistema di coordinate (x, y) è rappresentato da $dy \hat{y}$ nel nuovo sistema di coordinate (u, v) è rappresentato da $Y_u du \hat{u} + Y_v dv \hat{v}$. Ovvero: ¹⁷

$$dx \hat{x} = X_u du \hat{u} + X_v dv \hat{v}, \quad dy \hat{y} = Y_u du \hat{u} + Y_v dv \hat{v}. \quad (7.4)$$

Essendo $\hat{u} \wedge \hat{u} = \hat{v} \wedge \hat{v} = \vec{0}$ e $\hat{v} \wedge \hat{u} = -\hat{u} \wedge \hat{v}$, abbiamo

$$\begin{aligned} (dx \hat{x}) \wedge (dy \hat{y}) &= (X_u du \hat{u} + X_v dv \hat{v}) \wedge (Y_u du \hat{u} + Y_v dv \hat{v}) \\ &= (X_u \hat{u} + X_v \hat{v}) \wedge (Y_u \hat{u} + Y_v \hat{v}) du dv \\ &= (X_u Y_v - X_v Y_u) du dv \hat{u} \wedge \hat{v} = \det(J) du dv \hat{u} \wedge \hat{v}. \end{aligned} \quad (7.5)$$

Infine, essendo $\hat{x} \wedge \hat{y}$ e $\hat{u} \wedge \hat{v}$ dei versori,

$$dx dy = \|dx \hat{x} \wedge dy \hat{y}\| = \|\det(J) \hat{u} \wedge \hat{v}\| du dv = |\det(J)| du dv. \quad (7.6)$$

Questa formula fornisce la (7.3).

*** Confronto col Cambiamento di Variabile negli Integrali Semplici.** Si ricordi la formula per il cambiamento di variabile negli integrali semplici:

$$\int_{\varphi(a)}^{\varphi(b)} f(y) dy = \int_a^b f(\varphi(x)) \varphi'(x) dx \quad (7.7)$$

¹⁷Questo è il punto fondamentale. La formula di trasformazione cercata discende da queste eguaglianze mediante un semplice calcolo.

ove $f : \mathbf{R} \rightarrow \mathbf{R}$ è continua, $\varphi : [a, b] \rightarrow \mathbf{R}$ è di classe C^1 . Si noti che non si richiede $\varphi' \neq 0$.

Si confronti tale formula con quella per il cambiamento di variabile negli integrali multipli:

$$\int_{\varphi(A)} f(y)dy = \int_A f(\varphi(x))|\det(J_\varphi(x))|dx \quad (7.8)$$

ove $A \subset \mathbf{R}^2$ è un insieme limitato con frontiera di misura bidimensionale nulla, $f : A \rightarrow \mathbf{R}$ è continua, $\varphi : A \rightarrow \mathbf{R}^2$ è di classe C^1 ed iniettiva (ipotesi queste che possono essere un po' indebolite...), tale che l'insieme immagine $\varphi(A)$ sia limitato. J_φ indica la matrice jacobiana di φ .

Si noti che il determinante jacobiano è una naturale estensione della derivata al caso di più variabili, ed in effetti la seconda formula estende la prima. Sorge allora spontanea una domanda:

*perché nella (7.8) compare il valore assoluto di J_φ ,
mentre nella (7.7) φ' compare senza il segno di valore assoluto?*

Per risolvere questo apparente paradosso, si osservi che l'integrale (definito) semplice è effettuato su un intervallo *orientato*, infatti per definizione $\int_d^c g(x)dx = -\int_c^d g(x)dx$ per ogni $c, d \in \mathbf{R}$ con $c < d$ (e per ogni funzione g). Invece gli integrali multipli sono effettuati su insiemi pluridimensionali *non orientati*. Ed allora per poter confrontare i due tipi di integrali riduciamo anche gli integrali semplici ad integrali su insiemi non orientati, ponendo

$$\int_{[c,d]} g(x)dx := \int_c^d g(x)dx \quad \forall c, d \in \mathbf{R}, c < d.$$

Se $\varphi' \geq 0$ allora $|\varphi'| = \varphi'$ e non si pone nessun problema. più interessante è il caso in cui $\varphi' < 0$. Allora $\varphi(a) > \varphi(b)$ (poiché $a < b$), quindi per la (7.7) si ha

$$\int_{[\varphi(b), \varphi(a)]} f(y)dy = -\int_{\varphi(a)}^{\varphi(b)} f(y)dy = -\int_a^b f(\varphi(x))\varphi'(x)dx = \int_{[a,b]} f(\varphi(x))|\varphi'(x)|dx,$$

ovvero

$$\int_{[\varphi(b), \varphi(a)]} f(y)dy = \int_{[a,b]} f(\varphi(x))|\varphi'(x)|dx.$$

Quest'ultima formula è un caso particolare della (7.8); si noti che essa contiene integrali su insiemi non orientati, e che compare il valore assoluto di φ' .

8 Operatori Differenziali Vettoriali

Useremo la notazione tensoriale. Fissato un sistema di coordinate cartesiane in \mathbf{R}^N ($N > 1$), i tensori¹⁸ si presentano come tabelle ad un numero prescritto di indici, ciascuno dei quali varia tra 1 e N . Il numero di indici è detto *ordine* del tensore; ad esempio, gli scalari i vettori e le matrici sono rispettivamente tensori di ordine 0, 1, 2. Applicheremo la convenzione della somma sugli indici ripetuti.

¹⁹ Ad esempio, nell'espressione $a_i b_i$ l'indice i è ripetuto, quindi tale scrittura sta per $\sum_{i=1}^N a_i b_i$.

In particolare abbiamo il tensore di Kronecker

$$\delta_{ij} := \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j, \end{cases} \quad (8.1)$$

e quello di Levi Civita, che qui introduciamo solo nel caso di $N = 3$,

$$\epsilon_{ijk} := \begin{cases} 1 & \text{se } (i, j, k) = (1, 2, 3), \text{ o } (i, j, k) = (2, 3, 1), \text{ o } (i, j, k) = (3, 1, 2), \\ 0 & \text{se } i = j, \text{ o } j = k, \text{ o } i = k \quad (i = j = k \text{ non escluso}), \\ -1 & \text{se } (i, j, k) = (1, 3, 2), \text{ o } (i, j, k) = (3, 2, 1), \text{ o } (i, j, k) = (2, 1, 3). \end{cases} \quad (8.2)$$

¹⁸Questa denominazione è dovuta al fatto che questa teoria ha avuto origine in meccanica del continuo.

¹⁹Questa è spesso detta la *convenzione di Einstein*, in quanto da questi usata nella presentazione della teoria della relatività.

Quindi ogni volta che useremo il tensore di Levi Civita dovremo assumere $N = 3$.

Indichiamo con \hat{e}_k il versore dell'asse k -esimo ($k = 1, \dots, N$), ovvero $\hat{e}_1 := (1, 0, \dots, 0)$, $\hat{e}_2 := (0, 1, \dots, 0)$, ecc..²⁰ Vediamo ora come i due fondamentali tensori di Kronecker e di Levi Civita permettono di esprimere le operazioni di moltiplicazione tra vettori:

(i) essendo $\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$ per ogni $i, j \in \{1, \dots, N\}$, abbiamo

$$\vec{A} \cdot \vec{B} = (A_i \hat{e}_i) \cdot (B_j \hat{e}_j) = A_i B_j \hat{e}_i \cdot \hat{e}_j = A_i B_j \delta_{ij} = A_i B_i \quad \forall \vec{A}, \vec{B} \in \mathbf{R}^N; \quad (8.3)$$

(ii) analogamente, essendo $\hat{e}_i \times \hat{e}_j = \epsilon_{ijk} \hat{e}_k$ per ogni $i, j, k \in \{1, \dots, 3\}$, abbiamo

$$\vec{A} \times \vec{B} = (A_i \hat{e}_i) \times (B_j \hat{e}_j) = A_i B_j \hat{e}_i \times \hat{e}_j = A_i B_j \epsilon_{ijk} \hat{e}_k \quad \forall \vec{A}, \vec{B} \in \mathbf{R}^3. \quad (8.4)$$

Poniamo ora $A = \nabla$ (ovvero $A_k = \partial_k := \frac{\partial}{\partial x_k}$ per $k = 1, \dots, N$), e sostituiamo \vec{B} con un campo vettoriale (regolare) $\vec{F} : \Omega(\subset \mathbf{R}^3) \rightarrow \mathbf{R}^3$. Otteniamo

$$\nabla \varphi := \partial_i \varphi \hat{e}_i (= (\partial_1 \varphi, \dots, \partial_N \varphi)), \quad (8.5)$$

$$\nabla \cdot \vec{F} := \partial_i F_i (= \partial_1 F_1 + \dots + \partial_N F_N), \quad (8.6)$$

$$\Delta \varphi (= \nabla \cdot \nabla \varphi = \nabla^2 \varphi) := \partial_i \partial_i \varphi (= \partial_1^2 \varphi + \dots + \partial_N^2 \varphi), \quad (8.7)$$

$$\nabla \times \vec{F} := \epsilon_{ijk} \partial_i F_j \hat{e}_k \quad (\text{solo per } N = 3). \quad (8.8)$$

Alcune Identità. Applicando la formula di derivazione del prodotto, si possono verificare le seguenti identità, valide per ogni coppia di campi φ e \vec{F} di classe C^1 ,

$$\nabla \cdot (\varphi \vec{F}) = (\nabla \varphi) \cdot \vec{F} + \varphi \nabla \cdot \vec{F}, \quad \nabla \times (\varphi \vec{F}) = (\nabla \varphi) \times \vec{F} + \varphi \nabla \times \vec{F}.$$

Inoltre se φ e \vec{F} sono di classe C^2 , si ha

$$\nabla \times \nabla \varphi = \vec{0}, \quad (8.9)$$

$$\nabla \cdot \nabla \times \vec{F} = 0, \quad (8.10)$$

$$\nabla \times \nabla \times \vec{F} = \nabla(\nabla \cdot \vec{F}) - \Delta \vec{F}. \quad (8.11)$$

Dimostriamo ad esempio la (8.9). Scambiando prima l'ordine di derivazione e poi ridefinendo gli indici (ovvero cambiando la loro denominazione – ad esempio $A_i B_j C_{ij} = A_j B_i C_{ji}$ per ogni A, B, C) si ha

$$\nabla \times \nabla \varphi = \epsilon_{ijk} \partial_i \partial_j \varphi \hat{e}_k = \epsilon_{ijk} \partial_j \partial_i \varphi \hat{e}_k = \epsilon_{jik} \partial_i \partial_j \varphi \hat{e}_k;$$

quindi

$$\nabla \times \nabla \varphi = \frac{1}{2} (\epsilon_{ijk} + \epsilon_{jik}) \partial_i \partial_j \varphi \hat{e}_k,$$

e per la definizione del tensore ϵ questo coincide con il vettore $\vec{0}$.

La (8.10) si dimostra analogamente, mentre la (8.11) discende dalla seguente identità tensoriale:

$$\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}. \quad (8.12)$$

²⁰Si noti che $(e_k)_i = \delta_{ki}$ per ogni k, i (ovviamente $(e_k)_i$ è la componente i -esima del vettore \hat{e}_k). Inoltre, per ogni vettore $\vec{F} := (F_1, \dots, F_N)$, le formule che legano un vettore alle sue componenti si possono scrivere come segue:

$$\vec{F} = F_k \hat{e}_k, \quad F_k = \vec{F} \cdot \hat{e}_k.$$

9 Superfici

Ogni superficie di \mathbf{R}^3 può essere rappresentata o parametricamente, o esplicitamente, o implicitamente.

Rappresentazione Parametrica. Sia T un sottoinsieme aperto di \mathbf{R}^2 ; la parametrizzazione di una superficie \mathcal{S} di \mathbf{R}^3 è una funzione continua $\vec{R} : T \rightarrow \mathbf{R}^3$ tale che $\vec{R}(T) = \mathcal{S}$. Se \vec{R} è di classe \mathcal{C}^1 si definisce il *prodotto vettoriale fondamentale*:

$$\vec{\Pi}(u, v) := \vec{R}_u(u, v) \times \vec{R}_v(u, v) \quad \forall (u, v) \in T \quad (9.1)$$

(con \vec{R}_u e \vec{R}_v denotiamo le derivate parziali). La parametrizzazione \vec{R} è detta *regolare* se e solo se \vec{R} è di classe \mathcal{C}^1 e $\vec{\Pi}(u, v) \neq \vec{0}$ per ogni $(u, v) \in T$; ogni punto in cui $\vec{\Pi}(u, v) = \vec{0}$ è detto *singolare*.

Una superficie elementare di area $du dv$ è trasformata da \vec{R} in una superficie di area $\|(\vec{R}_u du) \times (\vec{R}_v dv)\| = \|\vec{\Pi}\| du dv$, a meno di infinitesimi di ordine superiore. Il modulo del prodotto vettoriale fondamentale, ovvero $\|\vec{\Pi}\|$, può quindi essere interpretato come il *fattore di ingrandimento o rimpicciolimento delle aree*.

Il piano tangente ad \mathcal{S} in un suo punto $\vec{R}(u_0, v_0)$ è generato da due rette tangenti alle linee coordinate, le quali hanno rappresentazione parametrica

$$\mathbf{R} \rightarrow \mathbf{R}^3 : u \mapsto (x, y, z) = \vec{R}(u_0, v_0) + (u - u_0)\vec{R}_u(u_0, v_0), \quad (9.2)$$

$$\mathbf{R} \rightarrow \mathbf{R}^3 : v \mapsto (x, y, z) = \vec{R}(u_0, v_0) + (v - v_0)\vec{R}_v(u_0, v_0).$$

Il piano pertanto ha rappresentazione parametrica

$$\mathbf{R}^2 \rightarrow \mathbf{R}^3 : (u, v) \mapsto (x, y, z) = \vec{R}(u_0, v_0) + (u - u_0)\vec{R}_u(u_0, v_0) + (v - v_0)\vec{R}_v(u_0, v_0). \quad (9.3)$$

Si noti che l'insieme dei parametri del piano coincide con tutto \mathbf{R}^2 .

La retta normale ad \mathcal{S} in un suo punto $\vec{R}(u_0, v_0)$ è ortogonale alle due rette tangenti alle linee coordinate in quel punto, ed quindi ha la direzione di $\vec{\Pi}(u_0, v_0) = \vec{R}_u(u_0, v_0) \times \vec{R}_v(u_0, v_0)$. Quindi essa ha rappresentazione parametrica

$$\mathbf{R} \rightarrow \mathbf{R}^3 : t \mapsto (x, y, z) = \vec{R}(u_0, v_0) + t\vec{\Pi}(u_0, v_0), \quad (9.4)$$

e rappresentazione implicita

$$[(x, y, z) - \vec{R}(u_0, v_0)] \times \vec{\Pi}(u_0, v_0) = \vec{0}. \quad (9.5)$$

Il piano tangente ad \mathcal{S} in un suo punto $\vec{R}(u_0, v_0)$ è ortogonale alla retta normale ad \mathcal{S} nello stesso punto $\vec{R}(u_0, v_0)$, ovvero è ortogonale al vettore $\vec{\Pi}(u_0, v_0)$. Esso pertanto ha rappresentazione implicita

$$[(x, y, z) - \vec{R}(u_0, v_0)] \cdot \vec{\Pi}(u_0, v_0) = 0, \quad (9.6)$$

ovvero

$$[(x, y, z) - \vec{R}(u_0, v_0)] \cdot \vec{R}_u(u_0, v_0) \times \vec{R}_v(u_0, v_0) = 0; \quad (9.7)$$

o anche, posto $\vec{R}(u_0, v_0) = (x_0, y_0, z_0)$ e $\vec{\Pi}(u_0, v_0) = (\Pi_1, \Pi_2, \Pi_3)$,

$$\frac{x - x_0}{\Pi_1} = \frac{y - y_0}{\Pi_2} = \frac{z - z_0}{\Pi_3}. \quad (9.8)$$

Con questa scrittura si intende che, se un denominatore si annulla, allora il corrispondente numeratore è pure nullo.

Rappresentazione Esplicita. Se la superficie \mathcal{S} ammette una rappresentazione esplicita

$$z = f(x, y) \quad \forall (x, y) \in \Omega, \text{ con } f \text{ di classe } \mathcal{C}^1, \quad (9.9)$$

allora si ricava la corrispondente rappresentazione parametrica

$$x = u, \quad y = v, \quad z = f(u, v), \quad (9.10)$$

ovvero $\vec{R}(u, v) = (u, v, f(u, v))$ per ogni $(u, v) \in T = \Omega$. Il *prodotto vettoriale fondamentale* è pertanto

$$\vec{\Pi} = \vec{R}_u \times \vec{R}_v = (1, 0, f_u) \times (0, 1, f_v) = (-f_u, -f_v, 1) \quad \text{in } T; \quad (9.11)$$

il suo modulo è quindi pari a $\sqrt{1 + (f_u)^2 + (f_v)^2} = \sqrt{1 + |\nabla f_u|^2}$.

Rappresentazione Implicita. Diverse superfici possono essere espresse nella forma

$$\mathcal{S} = \{(x, y, z) : g(x, y, z) = 0\} \quad \text{con } g \in C^0(\mathbf{R}^3). \quad (9.12)$$

Un insieme così rappresentato può consistere in una superficie, che in tal caso si dice rappresentata implicitamente (dalla funzione g), o anche degenerare in un insieme di dimensione inferiore a due. Se la funzione g è di classe C^1 ed il suo gradiente non si annulla in alcun punto della superficie, \mathcal{S} è detta una superficie regolare. (Questa definizione è coerente con quella fornita per le superfici parametriche, poiché mediante il teorema della funzione implicita si può verificare che ogni superficie che può essere rappresentata in entrambe le forme è regolare secondo una definizione se e solo se lo è seconda l'altra.)

Sia (x_0, y_0, z_0) un punto di una superficie regolare \mathcal{S} rappresentata implicitamente da una funzione g . Poiché \mathcal{S} è un insieme di livello per la funzione g , come noto $\nabla g(x, y, z)$ è ortogonale a \mathcal{S} . Un punto (x, y, z) appartiene quindi alla retta normale ad \mathcal{S} in (x_0, y_0, z_0) se e solo se

$$(x, y, z) = (x_0, y_0, z_0) + t\nabla g(x_0, y_0, z_0) \quad \text{per } t \in \mathbf{R}, \quad (9.13)$$

ovvero

$$\frac{x - x_0}{D_1 g(x_0, y_0, z_0)} = \frac{y - y_0}{D_2 g(x_0, y_0, z_0)} = \frac{z - z_0}{D_3 g(x_0, y_0, z_0)}. \quad (9.14)$$

Anche qui si intende che, se un denominatore si annulla, allora il corrispondente numeratore è pure nullo. Questa è la rappresentazione parametrica della retta normale.

Un punto (x, y, z) appartiene al piano tangente ad \mathcal{S} in (x_0, y_0, z_0) se e solo se è ortogonale alla retta normale, ovvero

$$[(x, y, z) - (x_0, y_0, z_0)] \cdot \nabla g(x_0, y_0, z_0) = 0. \quad (9.15)$$

Questa è la rappresentazione implicita del piano tangente.