

8. FUNZIONI VETTORIALI DIFFERENZIABILI

Studiamo le proprietà delle funzioni a valori vettoriali.

8.1. Matrice Jacobiana. Per estendere la definizione di differenziabilità al caso di funzioni vettoriali, usiamo la relazione (65) .

Def. 2.60. Una funzione $F : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$,

$$F(P) = (f_1(P), \dots, f_m(P)),$$

definita su un insieme aperto A , è detta differenziabile in $P_0 \in A$ se esiste una matrice M di taglia $m \times n$ tale che

$$F(P) = F(P_0) + M(P - P_0) + \|P - P_0\| \epsilon(P - P_0), \quad (79)$$

eq:6t

dove $\epsilon(P - P_0)$ è una funzione infinitesima a valori in \mathbb{R}^m . La matrice M è detta matrice jacobiana di F nel punto P_0 e si denota con $JF(P_0)$. La funzione F è differenziabile, se è differenziabile in tutti i punti di A .

La notazione $M(P - P_0)$ indica il vettore (colonna) in \mathbb{R}^m facendo il prodotto tra la matrice M di taglia $m \times n$ ed il vettore (colonna) $(P - P_0)$ in \mathbb{R}^n .

Ad esempio se $n = m = 2$, $P_0 = (x_0, y_0)$ e $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ la (79) diventa

$$\begin{bmatrix} f_1(x, y) \\ f_2(x, y) \end{bmatrix} = \begin{bmatrix} f_1(x_0, y_0) \\ f_2(x_0, y_0) \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \sqrt{(x - x_0)^2 + (y - y_0)^2} \begin{bmatrix} \epsilon_1(x - x_0, y - y_0) \\ \epsilon_2(x - x_0, y - y_0) \end{bmatrix},$$

dove $\epsilon_1(x - x_0, y - y_0)$ ed $\epsilon_2(x - x_0, y - y_0)$ sono due funzioni infinitesime.

La (79) che esprime il fatto che la funzione F è approssimata da una funzione lineare affine (vedi Esempio 2.13)

$$F(P_0) + JF(P_0)(P - P_0),$$

ed il resto

$$F(P) - \left(F(P_0) + JF(P_0)(P - P_0) \right) = \|P - P_0\| \epsilon(P - P_0)$$

va a zero più velocemente della distanza tra P e P_0 .

Il seguente risultato mostra come calcolare la matrice jacobiana.

2.40 **Lemma 2.61.** Una funzione $F : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$,

$$F(P) = (f_1(P), \dots, f_m(P))$$

definita su un insieme aperto A è differenziabile se e solo se tutte le sue componenti

$$f_1, \dots, f_m$$

sono differenziabili e, in tal caso,

$$JF(P) = \begin{bmatrix} \nabla f_1(P) \\ \dots \\ \nabla f_m(P) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(P) & \dots & \frac{\partial f_1}{\partial x_n}(P) \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1}(P) & \dots & \frac{\partial f_m}{\partial x_n}(P) \end{bmatrix}. \quad (80)$$

eq:68

Dimostrazione. Fissato $P^* \in A$ Posto $JF(P^*) = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$, allora la condizione

di differenziabilità diventa

$$\begin{bmatrix} f_1(P) \\ \dots \\ f_m(P) \end{bmatrix} = \begin{bmatrix} f_1(P^*) \\ \dots \\ f_m(P^*) \end{bmatrix} + \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 - x_1^* \\ \dots \\ x_n - x_n^* \end{bmatrix} + \begin{bmatrix} \epsilon_1(P - P^*) \\ \dots \\ \epsilon_m(P - P^*) \end{bmatrix}$$

dove $P = (x_1, \dots, x_n)$, $P^* = (x_1^*, \dots, x_n^*)$. Cioè, per ogni componente f_i

$$f_i(P) = f_i(P^*) + (a_{i1}, \dots, a_{in}) \cdot (P - P^*) + \epsilon_i(P - P^*).$$

dove, per l'unicità del gradiente,

$$\nabla f_i(P^*) = (a_{i1}, \dots, a_{in}).$$

La tesi segue osservando che ϵ è infinitesima se e solo se lo sono tutte le sue componenti ϵ_i . \square

Ad esempio se $n = m = 2$ la (80) diventa

$$JF(x, y) = \begin{bmatrix} \frac{\partial f_1}{\partial x}(x, y) & \frac{\partial f_1}{\partial y}(x, y) \\ \frac{\partial f_2}{\partial x}(x, y) & \frac{\partial f_2}{\partial y}(x, y) \end{bmatrix}.$$

8.2. Trasformazioni di coordinate. Una trasformazione (attiva) è una funzione

$$\Psi : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

che assegna ad ogni punto del dominio $P \in A$ il punto $\Psi(P)$. Esplicitamente, la trasformazione è data da

$$\Psi : \begin{cases} \bar{x} = \bar{x}(x, y) \\ \bar{y} = \bar{y}(x, y) \end{cases}.$$

dove le componenti $\bar{x}(x, y)$ ed $\bar{y}(x, y)$ sono le coordinate del punto trasformato.

Esempio 2.62. Data la matrice 2×2 , $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$, la trasformazione lineare associata è

$$\Psi : \begin{cases} \bar{x} = ax + cy \\ \bar{y} = bx + dy \end{cases},$$

in cui le componenti sono $\bar{x}(x, y) = ax + cy$ e $\bar{y}(x, y) = bx + dy$. La trasformazione Ψ manda il quadrato di vertici $(0, 0)$, $(1, 0)$, $(1, 1)$ e $(0, 1)$, nel parallelogramma di vertice l'origine ed i cui due lati adiacenti sono i vettori (a, b) e (c, d) . Se il determinante $ad - bc$ è zero, i vettori (a, b) e (c, d) sono paralleli ed il parallelogramma si riduce ad un segmento.

Un caso di particolare interesse sono le rotazioni. Fissato un angolo $\theta \in \mathbb{R}$, la corrispondente matrice di rotazione è

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

e la corrispondente trasformazione lineare è data da

$$\Psi : \begin{cases} \bar{x} = \cos \theta x - \sin \theta y \\ \bar{y} = \sin \theta x + \cos \theta y \end{cases}.$$

Se Ψ è invertibile, la trasformazione può essere pensata in modo passivo, interpretando i valori $\bar{x}(x, y)$ ed $\bar{y}(x, y)$ come coordinate del punto (x, y) nel sistema di riferimento trasformato. Solitamente si preferisce scrivere la funzione inversa che assegna ad (\bar{x}, \bar{y}) il corrispondente punto $\Phi(\bar{x}, \bar{y}) = (x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y}))$

$$\Phi : \begin{cases} x = x(\bar{x}, \bar{y}) \\ y = y(\bar{x}, \bar{y}) \end{cases},$$

dove (\bar{x}, \bar{y}) sono le coordinate del punto nel nuovo sistema di riferimento ed $x = x(\bar{x}, \bar{y})$ e $y = y(\bar{x}, \bar{y})$ le coordinate dello stesso punto nel vecchio sistema di riferimento. Questa costruzione motiva la seguente definizione.

Def. 2.63. Una funzione $\Phi : B \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita su un aperto B tale che

a) la funzione Φ è iniettiva e di classe C^1 ;

- b) l'immagine $A = \Phi(B)$ è un aperto;
 c) la funzione inversa $\Phi^{-1} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ è di classe C^1 ;

è detta una trasformazione di coordinate regolare tra gli aperti B ed A .

Ricordiamo che una trasformazione

$$\Phi : \begin{cases} x = x(\bar{x}, \bar{y}) \\ y = y(\bar{x}, \bar{y}) \end{cases},$$

è di classe C^1 se e solo se sono le funzioni $x(\bar{x}, \bar{y})$ e $y(\bar{x}, \bar{y})$ sono di classe C^1 . La matrice jacobiana è una matrice 2×2 data da

$$J\Phi(\bar{x}, \bar{y}) = \begin{bmatrix} \frac{\partial x}{\partial \bar{x}}(\bar{x}, \bar{y}) & \frac{\partial x}{\partial \bar{y}}(\bar{x}, \bar{y}) \\ \frac{\partial y}{\partial \bar{x}}(\bar{x}, \bar{y}) & \frac{\partial y}{\partial \bar{y}}(\bar{x}, \bar{y}) \end{bmatrix}.$$

Esempio 2.64. Sia $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la trasformazione lineare associata alla matrice $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\Phi : \begin{cases} x = a\bar{x} + b\bar{y} \\ y = c\bar{x} + d\bar{y} \end{cases}.$$

La trasformazione è regolare se e solo se $\det M = ad - bc \neq 0$. Infatti, $\Phi \in C^1(\mathbb{R}^2)$ poiché le componenti $x(\bar{x}, \bar{y}) = a\bar{x} + b\bar{y}$ ed $y(\bar{x}, \bar{y}) = c\bar{x} + d\bar{y}$ sono di classe $C^1(\mathbb{R}^2)$. Un teorema dell'algebra lineare assicura che Φ è iniettiva se e solo se $ad - bc \neq 0$ e, in tal caso, $\Phi(\mathbb{R}^2) = \mathbb{R}^2$, che è aperto. L'inversa di Φ è la trasformazione lineare associata alla matrice

$$M^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

cioè

$$\Phi^{-1} : \begin{cases} \bar{x} = \frac{d}{ad - bc}x + \frac{-b}{ad - bc}y \\ \bar{y} = \frac{-c}{ad - bc}x + \frac{a}{ad - bc}y \end{cases},$$

che è di classe C^1 .

La mappa Φ definisce un cambio di coordinate

$$\text{vecchie: } (x, y) \longleftrightarrow \text{nuove: } (\bar{x}, \bar{y}) \quad \begin{cases} x = a\bar{x} + b\bar{y} \\ y = c\bar{x} + d\bar{y} \end{cases}$$

e, con questa notazione,

$$J\Phi(\bar{x}, \bar{y}) = \begin{bmatrix} \frac{\partial x}{\partial \bar{x}} & \frac{\partial x}{\partial \bar{y}} \\ \frac{\partial y}{\partial \bar{x}} & \frac{\partial y}{\partial \bar{y}} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

tra_lineari

Esempio 2.65. Sia $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la trasformazione di coordinate lineare definite nell'Esempio 2.65 con $ad - bc \neq 0$. Data $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ differenziabile su A aperto, la funzione composta

$$g(\bar{x}, \bar{y}) = f(a\bar{x} + b\bar{y}, c\bar{x} + d\bar{y}) = f(x, y) \quad (81)$$

f_cambia

ha dominio $B = \{(\bar{x}, \bar{y}) \in \mathbb{R}^2 \mid (a\bar{x} + b\bar{y}, c\bar{x} + d\bar{y}) \in A\}$, che risulta aperto. La regola di derivazione in catena mostra che g è differenziabile e nel punto $(\bar{x}, \bar{y}) \in B$ le derivate parziali valgono

$$\begin{aligned} \frac{\partial g}{\partial \bar{x}}(\bar{x}, \bar{y}) &= \frac{\partial f}{\partial x}(x, y) \frac{\partial x}{\partial \bar{x}}(\bar{x}, \bar{y}) + \frac{\partial f}{\partial y}(x, y) \frac{\partial y}{\partial \bar{x}}(\bar{x}, \bar{y}) = \frac{\partial f}{\partial x}(x, y) a + \frac{\partial f}{\partial y}(x, y) c \\ \frac{\partial g}{\partial \bar{y}}(\bar{x}, \bar{y}) &= \frac{\partial f}{\partial x}(x, y) \frac{\partial x}{\partial \bar{y}}(\bar{x}, \bar{y}) + \frac{\partial f}{\partial y}(x, y) \frac{\partial y}{\partial \bar{y}}(\bar{x}, \bar{y}) = \frac{\partial f}{\partial x}(x, y) b + \frac{\partial f}{\partial y}(x, y) d \end{aligned}$$

o, in forma compatta,

$$\nabla g(\bar{x}, \bar{y}) = \nabla f(x, y) \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Nelle precedenti espressioni $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ indicano le derivate parziali rispetto alla prima e seconda variabile di f , pensata come funzione delle variabili *vecchie* (x, y) , mentre $\frac{\partial g}{\partial x}$ e $\frac{\partial g}{\partial y}$ indicano le derivate parziali rispetto alla prima e seconda variabile di g , cioè f pensata come funzione delle variabili *nuove* (\bar{x}, \bar{y}) (come suggerisce l'equazione (81)). Tuttavia è utile usare due simboli diversi per f , funzione delle vecchie variabili, ed g , funzione delle nuove variabili, per evitare confusione notazionale, come mostra il seguente esempio.

◈. Sia

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},$$

allora

$$\begin{cases} x = \bar{x} \\ y = \bar{x} + \bar{y} \neq \bar{y} \end{cases} \quad \text{ma} \quad \begin{cases} \frac{\partial g}{\partial \bar{x}} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \neq \frac{\partial f}{\partial x} \\ \frac{\partial g}{\partial \bar{y}} = \frac{\partial f}{\partial y} \end{cases}.$$

Il seguente risultato fornisce una importante proprietà della funzione inversa.

Teo 2.66 (Teorema della mappa inversa). *Data una funzione $\Phi : B \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita su un aperto B e di classe C^1 . Se Φ è iniettiva e*

$$\det J\Phi(P) \neq 0 \quad \forall P \in B, \quad (82) \quad \boxed{\text{eq:21}}$$

allora $A = \Phi(B)$ è un insieme aperto, la funzione inversa $\Phi^{-1} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ è di classe C^1 e per ogni $Q \in A$

$$J\Phi^{-1}(Q) = (J\Phi(P))^{-1} \quad \text{con } P = \Phi^{-1}(Q),$$

dove $(J\Phi(P))^{-1}$ è la matrice inversa di $J\Phi(P)$.

Dimostrazione. Per il teorema della mappa inversa locale, per ogni $P_0 \in B$ esiste un aperto V_{P_0} contenente P_0 per cui la restrizione di Φ a V_{P_0} è una trasformazione regolare di coordinate con inversa di classe C^1 data da $\Phi_{P_0}^{-1} : W_{P_0} \rightarrow \mathbb{R}^n$ con $W_{P_0} = \Phi(V_{P_0})$ aperto (da notare che ogni restrizione ha un'inversa diversa). Tuttavia, osserviamo che se $Q \in W_{P_0} \cap W_{P_1}$, allora

$$\Phi(\Phi_{P_0}^{-1}(Q)) = Q \quad \Phi(\Phi_{P_1}^{-1}(Q)) = Q,$$

e l'ipotesi di iniettività di Φ implica che $\Phi_{P_0}^{-1}(Q) = \Phi_{P_1}^{-1}(Q)$. Definiamo $A = \cup_{P_0} W_{P_0}$, che è aperto, e $\Phi^{-1} : A \rightarrow \mathbb{R}^n$

$$\Phi^{-1}(Q) = \Phi_{P_0}^{-1}(Q) \quad Q \in W_{P_0}.$$

Poiché la restrizione di Φ^{-1} a ciascun aperto W_{P_0} coincide con $\Phi_{P_0}^{-1}$, allora Φ^{-1} è l'inversa di Φ , è di classe C^1 e vale

$$J\Phi^{-1}(Q) = (J\Phi(P))^{-1} \quad \text{con } P = \Phi^{-1}(Q).$$

□

◈. Nel caso $n = 1$, la condizione (82) si riduce all'assunzione che la derivata prima sia diversa da zero. Se il dominio è un intervallo, questo implica che la funzione è iniettiva. Tuttavia se $n > 1$, l'ipotesi che Φ sia iniettiva non si può togliere, anche se il dominio è connesso per archi, come mostra il seguente esempio.

Esempio 2.67 (Coordinate polari). Sia

$$\Phi : \begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases} \quad r \in (0, +\infty) \theta \in \mathbb{R}.$$

La funzione Φ non è iniettiva poiché $\Phi(r, \theta) = \Phi(r, \theta + 2\pi)$, ma

$$\det J\Phi(r, \theta) = \det \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} = r > 0 \quad r \in (0, +\infty) \quad \theta \in \mathbb{R}.$$

Tuttavia, Φ definisce una trasformazione di coordinate regolare sull'aperto $B = (0, +\infty) \times (-\pi, \pi)$ e la trasformazione inversa è

$$\Phi^{-1}(x, y) = \begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = 2 \arctan \frac{y}{\sqrt{x^2 + y^2} + x} \end{cases} \quad (x, y) \in A = \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 \mid y = 0, x \leq 0\},$$
(83) eq:111

dove la seconda relazione segue osservando che

$$\tan \frac{\theta}{2} = \frac{2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}}{2 \cos^2 \frac{\theta}{2}} = \frac{\sin \theta}{1 + \cos \theta} = \frac{\frac{y}{r}}{1 + \frac{x}{r}} \quad \frac{\theta}{2} \in (-\pi/2, \pi/2)$$

e l'arcotangente è l'inversa della tangente in $(-\pi/2, \pi/2)$.

La trasformazione Φ^{-1} non si può estendere in modo continuo al semiasse delle ascisse negative. Infatti, ad esempio,

$$\lim_{y \rightarrow 0^\pm} \theta(-1, y) = \lim_{y \rightarrow 0^\pm} 2 \arctan \frac{y}{\sqrt{1 + y^2} - 1} = 2 \lim_{y \rightarrow 0^\pm} \arctan \frac{\sqrt{1 + y^2} + 1}{y} = \pm\pi.$$

La matrice jacobiana della trasformazione inversa è data da

$$J\Phi^{-1}(x, y) = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}^{-1} = \frac{1}{r} \begin{bmatrix} r \cos \theta & r \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ -\frac{y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{bmatrix}.$$

La regola di derivazione in catena permette di calcolare come si trasformano le derivate parziali rispetto ad un cambiamento di coordinate.

Esempio 2.68. Sia $z = f(x, y)$ una funzione differenziabile definita nell'aperto A e Φ la trasformazione di coordinate polari definita nel precedente esempio. La funzione composta $g(r, \theta) = f(\Phi(r, \theta)) = f(r \cos \theta, r \sin \theta)$ è la funzione f in coordinate polari con dominio $B = \Phi^{-1}(A)$. La regola di derivazione in catena dà

$$\begin{aligned} \frac{\partial g}{\partial r}(r, \theta) &= \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) \frac{\partial x}{\partial r}(r, \theta) + \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) \frac{\partial y}{\partial r}(r \cos \theta, r \sin \theta) \\ &= \cos \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + \sin \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) \end{aligned}$$

$$\begin{aligned} \frac{\partial g}{\partial \theta}(r, \theta) &= \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) \frac{\partial x}{\partial \theta}(r, \theta) + \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta) \frac{\partial y}{\partial \theta}(r \cos \theta, r \sin \theta) \\ &= -r \sin \theta \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) + r \cos \theta \frac{\partial f}{\partial y}(r \cos \theta, r \sin \theta). \end{aligned}$$

Le formule precedenti hanno una interpretazione geometrica. Siano $n_r = (\cos \theta, \sin \theta)$ e $n_\theta = (-\sin \theta, \cos \theta)$ rispettivamente il versore radiale ed il versore angolare nel punto $P = (x, y) = (r \cos \theta, r \sin \theta)$, allora

$$\begin{aligned} \frac{\partial g}{\partial r}(r, \theta) &= \nabla f(P) \cdot n_r \\ \frac{\partial g}{\partial \theta}(r, \theta) &= r \nabla f(P) \cdot n_\theta \end{aligned}$$

Esempio 2.69 (Coordinate sferiche). Sia $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$

$$\Phi(r, \theta, \varphi) : \begin{cases} x = r \sin \theta \cos \varphi \\ y = r \sin \theta \sin \varphi \\ z = r \cos \theta \end{cases} \quad r \in \mathbb{R}, \theta \in \mathbb{R}, \varphi \in \mathbb{R}.$$

La funzione Φ è di classe C^1 con matrice jacobiana

$$J\Phi(r, \theta, \varphi) = \begin{bmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{bmatrix}$$

e $\det J\Phi(r, \theta, \varphi) = r^2 \sin \theta$. La funzione Φ definisce una trasformazione regolare di coordinate tra l'aperto

$$B = (0, +\infty) \times (0, \pi) \times (-\pi, \pi)$$

e l'aperto

$$A = \Phi(B) = \mathbb{R}^3 \setminus \{(x, y, z) \in \mathbb{R}^3 \mid y = 0, x \leq 0\}.$$

8.3. Curve. Ricordiamo che una curva è una funzione $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$. Nel caso di curve piane useremo spesso la notazione

$$\gamma(t) = x(t)\mathbf{i} + y(t)\mathbf{j} \quad \text{oppure} \quad \gamma : \begin{cases} x = x(t) \\ y = y(t) \end{cases}$$

⚠. Attenzione a non confondere γ con la sua traccia (o traiettoria)

$$\gamma(I) = \{\gamma(t) \in \mathbb{R}^n \mid t \in I\},$$

che è un sottoinsieme di \mathbb{R}^n . La funzione γ rappresenta la parametrizzazione dell'insieme $\gamma(I)$ e contiene le informazioni di tipo analitico.

Data una curva $\gamma : [a, b] \rightarrow \mathbb{R}^n$

- a) i punti $\gamma(a)$ e $\gamma(b)$ si chiamano primo e secondo estremo della curva, rispettivamente;
- b) se $\gamma(a) = \gamma(b)$, si dice che γ è una curva chiusa.

Essendo una curva una funzione da \mathbb{R} ad \mathbb{R}^m , la nozione di differenziabilità è equivalente alla seguente definizione di derivabilità.

Def. 2.70. Una curva $\gamma : I \rightarrow \mathbb{R}^n$ definita in un intervallo I aperto è derivabile se per ogni $t_0 \in I$ esiste finito

$$\lim_{t \rightarrow t_0} \frac{\gamma(t) - \gamma(t_0)}{t - t_0} = \gamma'(t_0). \quad (84) \quad \boxed{\text{eq:62}}$$

Il vettore $\gamma'(t_0)$ è detto vettore tangente nel punto $\gamma(t_0)$.

La curva γ è detta di classe C^1 se il vettore tangente $\gamma'(t)$ è una funzione continua.

Se γ è una curva piana

$$\gamma : \begin{cases} x = x(t) \\ y = y(t) \end{cases} \quad t \in I$$

allora $\gamma(t)$ è derivabile se e solo se le componenti $x(t)$ e $y(t)$ sono derivabili (nel senso ordinario di funzioni di una variabile) ed il vettore tangente è

$$\gamma'(t) = (x'(t), y'(t)).$$

Inoltre la curva è di classe C^1 se e solo se le componenti $x(t)$ e $y(t)$ sono di classe C^1 .

Def. 2.71 (Curva regolare). Una curva $\gamma : I \rightarrow \mathbb{R}^n$, $I \subset \mathbb{R}$, intervallo si dice regolare se γ è di classe $C^1(I)$ ed il vettore tangente soddisfa la condizione

$$\gamma'(t) \neq 0 \quad \forall t \in I. \quad (85) \quad \boxed{\text{regolare}}$$

⚠. Se $I = [a, b]$, nella definizione si assume sempre che esistano (finiti) i limiti

$$\gamma'(a) = \lim_{t \rightarrow a^+} \frac{\gamma(t) - \gamma(a)}{t - a} \quad \gamma'(b) = \lim_{t \rightarrow b^-} \frac{\gamma(t) - \gamma(b)}{t - b}.$$

La condizione (85) equivale al fatto che la velocità scalare $v(t) = \|\gamma'(t)\|$ sia sempre positiva e garantisce l'esistenza della retta tangente alla traccia in tutti i punti $\gamma(t_0)$ al variare di $t_0 \in [a, b]$.

Dato $t_0 \in I$, l'equazione parametrica della retta tangente alla curva nel punto $P_0 = \gamma(t_0)$ è

$$P = \gamma(t_0) + h \gamma'(t_0) \quad h \in \mathbb{R}.$$

Nota. Nel caso di curve piane regolari, $\gamma(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, il vettore tangente è

$$\gamma'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j}$$

e la (85) diventa

$$v(t) = \sqrt{x'(t)^2 + y'(t)^2} > 0,$$

l'equazione parametrica della retta tangente alla curva nel punto $\gamma(t_0) = (x(t_0), y(t_0))$ è data da

$$\begin{cases} x = x(t_0) + hx'(t_0) \\ y = y(t_0) + hy'(t_0) \end{cases} \quad h \in \mathbb{R},$$

mentre la corrispondente equazione cartesiana è

$$y'(t_0)(x - x(t_0)) - x'(t_0)(y - y(t_0)) = 0.$$

Una classe importante di curve bidimensionali è descritta dal seguente esempio.

Esempio 2.72 (Grafico di una funzione). Data una funzione di classe C^1 , $\varphi : I \subset \mathbb{R} \rightarrow \mathbb{R}$ definita su un intervallo I , allora la curva

$$\gamma : \begin{cases} x = t \\ y = \varphi(t) \end{cases} \quad t \in I$$

è una curva regolare la cui traccia è il grafico di φ , $\{(x, y) \in \mathbb{R}^2 \mid x \in I, y = \varphi(x)\}$, e

$$\gamma'(t) = (1, \varphi'(t)) \quad v(t) = \sqrt{1 + \varphi'(t)^2}.$$

L'equazione della retta tangente in $\gamma(t_0) = (t_0, \varphi(t_0)) = (x_0, \varphi(x_0))$ risulta

$$\varphi'(x_0)(x - x_0) - (y - \varphi(x_0)) = 0,$$

in accordo con il significato geometrico di derivata.

Esempio 2.73 (Curve in coordinate polari). L'equazione parametrica di una curva piana può essere assegnata in coordinate polari

$$\begin{cases} r = r(t) \\ \theta = \theta(t) \end{cases} \quad t \in I$$

dove $r(t)$, $\theta(t)$ sono due funzioni di classe C^1 . Questo definisce una curva

$$\gamma : \begin{cases} x = r(t) \cos \theta(t) \\ y = r(t) \sin \theta(t) \end{cases} \quad t \in I$$

con vettore tangente

$$\gamma'(t) = (r'(t) \cos \theta(t) - r(t) \theta'(t) \sin \theta(t))\mathbf{i} + (r'(t) \sin \theta(t) + r(t) \theta'(t) \cos \theta(t))\mathbf{j}$$

e velocità

$$v(t) = \sqrt{r'(t)^2 + r(t)^2 \theta'(t)^2}.$$

Data una curva regolare γ si definiscono

a) il versore tangente

$$\boldsymbol{\tau}(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|},$$

b) se $\boldsymbol{\tau}(t)$ è derivabile e $\boldsymbol{\tau}'(t) \neq 0$, il versore normale

$$\mathbf{n}(t) = \frac{\boldsymbol{\tau}'(t)}{\|\boldsymbol{\tau}'(t)\|},$$

c) se $\boldsymbol{\tau}(t)$ è derivabile, la curvatura

$$\chi(t) = \frac{\|\boldsymbol{\tau}'(t)\|}{\|\gamma'(t)\|} = \frac{\|\boldsymbol{\tau}'(t)\|}{v(t)};$$

d) fissato t_0 , il piano passante per $\gamma(t_0)$ generato dai versori $\boldsymbol{\tau}(t_0)$ e $\mathbf{n}(t_0)$

$$\{\gamma(t_0) + h\boldsymbol{\tau}(t_0) + k\mathbf{n}(t_0) \mid h, k \in \mathbb{R}\}$$

è detto piano osculatore alla curva. Nel caso di curve nello spazio, il piano osculatore ha equazione cartesiana

$$(P - \gamma(t_0)) \cdot (\boldsymbol{\tau}(t_0) \wedge \mathbf{n}(t_0)) = 0.$$

Il significato geometrico è chiarito dalle seguenti relazioni

$$\gamma'(t) = v(t)\boldsymbol{\tau}(t) \quad (86a) \quad \boxed{\text{eq:86a}}$$

$$\boldsymbol{\tau}(t) \cdot \mathbf{n}(t) = 0 \quad \text{cioè} \quad \boldsymbol{\tau}(t) \perp \mathbf{n}(t) \quad (86b) \quad \boxed{\text{eq:86b}}$$

$$\boldsymbol{\tau}'(t) = \chi(t)v(t)\mathbf{n}(t) \quad (86c) \quad \boxed{\text{eq:86c}}$$

$$\gamma''(t) = v'(t)\boldsymbol{\tau}(t) + v(t)\boldsymbol{\tau}'(t) = v'(t)\boldsymbol{\tau}(t) + \chi(t)v(t)^2\mathbf{n}(t). \quad (86d) \quad \boxed{\text{eq:86d}}$$

La prima e la terza relazione sono ovvie, la seconda segue derivando $\boldsymbol{\tau}(t) \cdot \boldsymbol{\tau}(t) = 0$, la quarta si deduce derivando la prima ed usando la terza e mostra che l'accelerazione giace sempre nel piano osculatore. Nel caso di curve nel piano, se $\alpha(t)$ indica l'angolo tra la semiretta delle ascisse positive e la semiretta individuata dal vettore tangente $\gamma'(t)$ (positivo se misurato in senso antiorario), allora

$$\boldsymbol{\tau}(t) = (\cos \alpha(t), \sin \alpha(t))$$

$$\mathbf{n}(t) = \frac{\alpha'(t)}{|\alpha'(t)|} (-\sin \alpha(t), \cos \alpha(t))$$

$$\boldsymbol{\tau}'(t) = |\alpha'(t)|\mathbf{n}(t)$$

(il versore $(-\sin \alpha(t), \cos \alpha(t))$ è ottenuto ruotando $\boldsymbol{\tau}(t)$ di $\pi/2$ in senso antiorario).

La lunghezza di una curva può essere data in due modi. Il primo è comodo dal punto di vista operativo.

Def. 2.74 (Lunghezza di una curva). Data una curva regolare $\gamma : [a, b] \rightarrow \mathbb{R}^n$

a) si chiama lunghezza della curva il valore positivo

$$\ell_\gamma = \int_a^b \|\gamma'(t)\| dt = \int_a^b v(t) dt,$$

b) si chiama ascissa curvilinea la funzione $s : [a, b] \rightarrow \mathbb{R}$

$$s(t) = \int_a^t \|\gamma'(\tau)\| d\tau = \int_a^t v(\tau) d\tau.$$

Poiché $v(t)$ è una funzione continua, per il teorema fondamentale del calcolo integrale, la funzione $s = s(t)$ è di classe $C^1([a, b])$ e $s'(t) = v(t) > 0$. Ne segue che, $s(t)$ è strettamente crescente e

$$s([a, b]) = [s(a), s(b)] = [0, \ell_\gamma]$$

Essendo $s(t)$ invertibile, se denotiamo con $t = t(s)$ l'inversa dell'ascissa curvilinea, la curva $\gamma_{\text{st}} : [0, \ell_\gamma] \rightarrow \mathbb{R}^n$

$$\gamma_{\text{st}}(s) = \gamma(t(s))$$

si chiama rappresentazione standard della curva γ ed è una curva equivalente a γ con lo stesso orientamento, quindi γ ed γ_{st} hanno la stessa traccia, gli stessi versori tangente e normale, e la stessa curvatura. Poiché $t'(s) = 1/v(t(s))$, dalle (86a) e (86d) segue che per ogni $s \in [0, \ell_\gamma]$

$$\gamma'_{\text{st}}(s) = \boldsymbol{\tau}(t(s)) \quad v_{\text{st}}(s) = 1 \quad \gamma''_{\text{st}}(s) = \chi(t(s))\mathbf{n}(t(s)).$$

Esempio 2.75 (Moto rettilineo). Siano $(x_0, y_0) \in \mathbb{R}^2$, $\theta \in [-\pi, \pi]$ e γ la curva

$$\gamma(t) = \begin{cases} x = x_0 + r(t) \cos \theta \\ y = y_0 + r(t) \sin \theta \end{cases} \quad t \in [0, 1]$$

$$\text{dove } r : [0, 1] \rightarrow \mathbb{R} \quad r \in C^1([0, 1]) \quad \begin{cases} r(0) = 0 \\ r(1) = \ell \end{cases} \quad r'(t) > 0 \quad t \in [0, 1].$$

Poiché $\gamma'(t) = r'(t)(\cos \theta, \sin \theta)$, γ è una curva regolare la cui traccia è il segmento di estremi $P_0 = (x_0, y_0)$ e $P_1 = (x_0 + \ell \cos \theta, y_0 + \ell \sin \theta)$, che giace sulla retta passante per (x_0, y_0) ed ha direzione il versore $(\cos \theta, \sin \theta)$. Inoltre si ha che

$$v(t) = r'(t) \quad \tau(t) = (\cos \theta, \sin \theta) \quad \tau'(t) = 0 \quad \chi(t) = 0.$$

L'ascissa curvilinea è data da $s(t) = \int_0^t r'(\tau) d\tau = r(t)$ e la lunghezza della curva è $s(1) = r(1) = \ell$. La rappresentazione standard della curva γ risulta

$$\gamma_{\text{st}}(s) = \gamma(t(s)) = \begin{cases} x(s) = x_0 + s \cos \theta \\ y(s) = y_0 + s \sin \theta \end{cases} \quad s \in [0, \ell],$$

che è l'equazione parametrica del segmento di estremi P_0 e P_1 .

Esempio 2.76 (Moto circolare). Sia $r > 0$ e γ la curva

$$\gamma(t) = \begin{cases} x = r \cos \theta(t) \\ y = r \sin \theta(t) \end{cases} \quad t \in [0, 1]$$

$$\text{dove } \theta : [0, 1] \rightarrow \mathbb{R} \quad \theta \in C^1([0, 1]) \quad \begin{cases} \theta(0) = \theta_0 \\ \theta(1) = \theta_1 \end{cases} \quad \theta'(t) > 0 \quad t \in [0, 1].$$

Poiché $\gamma'(t) = r\theta'(t)(-\sin \theta(t), \cos \theta(t))$, γ è una curva regolare la cui traccia è l'arco di circonferenza di raggio r , centro l'origine e di estremi $P_0 = (r \cos \theta_0, r \sin \theta_0)$ e $P_1 = (r \cos \theta_1, r \sin \theta_1)$. Inoltre si ha che

$$v(t) = r\theta'(t) \quad \tau(t) = (-\sin \theta(t), \cos \theta(t)) \quad \mathbf{n}(t) = -(\cos \theta(t), \sin \theta(t)) \quad \chi(t) = \frac{1}{r}.$$

L'ascissa curvilinea è data da $s(t) = \int_0^t r\theta'(\tau) d\tau = r(\theta(t) - \theta_0)$ e la lunghezza della curva è $s(1) = r(\theta_1 - \theta_0)$. La rappresentazione standard della curva γ risulta

$$\gamma_{\text{st}}(s) = \gamma(t(s)) = \begin{cases} x(s) = r \cos(\theta_0 + \frac{s}{r}) \\ y(s) = r \sin(\theta_0 + \frac{s}{r}) \end{cases} \quad s \in [0, r(\theta_1 - \theta_0)],$$

che è l'equazione parametrica dell'arco di circonferenza di estremi P_0 e P_1 .

Il seguente teorema mostra che la definizione di lunghezza ha un significato geometrico.

Teo 2.77. *Data una curva regolare $\gamma : [a, b] \rightarrow \mathbb{R}^n$, allora*

$$\ell_\gamma = \sup_{\Delta_{a,b}} \left(\sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| \right),$$

dove l'estremo superiore è fatto su tutte le partizioni $\Delta_{a,b} = \{t_0 = a < t_1 < \dots < t_n = b\}$ dell'intervallo $[a, b]$.

Dimostrazione. La dimostrazione è suddivisa in cinque passi. Per chiarezza supponiamo che $n = 2$.

Passo 1. Poiché la funzione γ' è una funzione continua è possibile integrarla componente per componente

$$\int_a^b \gamma'(t) dt = \left(\int_a^b x'(t) dt \right) \mathbf{i} + \left(\int_a^b y'(t) dt \right) \mathbf{j} \in \mathbb{R}^2$$

e per il teorema fondamentale del calcolo integrale

$$\int_a^b \gamma'(t) dt = \gamma(b) - \gamma(a). \quad (87) \quad \boxed{\text{eq:87}}$$

Passo 2. Vale la disuguaglianza

$$\left\| \int_a^b \gamma'(t) dt \right\| \leq \int_a^b \|\gamma'(t)\| dt. \quad (88) \quad \boxed{\text{eq:85}}$$

Infatti, posto $v = \int_a^b \gamma'(t) dt \in \mathbb{R}^n$, se $v = 0$ la disuguaglianza è evidente. Se $v \neq 0$, per la linearità dell'integrale e la disuguaglianza di Cauchy-Schwarz

$$\begin{aligned} \|v\|^2 &= v_1^2 + v_2^2 = \int_a^b v_1 x'(t) dt + \int_a^b v_2 y'(t) dt = \int_a^b v \cdot \gamma'(t) dt \\ &\leq \int_a^b \|v\| \|\gamma'(t)\| dt = \|v\| \int_a^b \|\gamma'(t)\| dt, \end{aligned}$$

da cui dividendo per $\|v\| \neq 0$ si ottiene la (88).

Passo 3. Mostriamo che per ogni partizione $\Delta_{a,b} = \{t_0 = a < t_1 < \dots < t_n = b\}$ vale

$$\sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| \leq \int_a^b \|\gamma'(t)\| dt.$$

Infatti, usando la (87) e la (88) su ogni intervallo $[t_{i-1}, t_i]$ si ha che

$$\begin{aligned} \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| &= \sum_{i=1}^n \left\| \int_{t_{i-1}}^{t_i} \gamma'(t) dt \right\| \\ &\leq \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt, \end{aligned}$$

dove si è usata l'additività degli integrali sui domini ed il fatto che $\Delta_{a,b}$ è una partizione di $[a, b]$. Ne segue che

$$\sup_{\Delta_{a,b}} \left(\sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| \right) \leq \int_a^b \|\gamma'(t)\| dt. \quad (89) \quad \boxed{\text{eq:86}}$$

Passo 4. Per ogni $t \in [a, b]$ definiamo

$$S(t) = \sup_{\Delta_{a,t}} \left(\sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| \right)$$

dove l'estremo superiore è fatto su tutte le partizioni $\Delta_{a,t} = \{t_0 = a < t_1 < \dots < t_n = t\}$ dell'intervallo $[a, t]$. Mostriamo che $S(t)$ è derivabile su $[a, b]$ ed $S'(t) = \|\gamma'(t)\|$. Per la (89),

$$S(t) \leq \int_a^t \|\gamma'(\tau)\| d\tau < +\infty.$$

Fissato $t \in [a, b]$ ed $h > 0$ tale che $t + h \leq b$, si dimostra che

$$S(t+h) - S(t) = \sup_{\Delta_{t,t+h}} \left(\sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| \right),$$

dove l'estremo superiore è fatto su tutte le partizioni $\Delta_{t,t+h} = \{t_0 = t < t_1 < \dots < t_n = t+h\}$ dell'intervallo $[t, t+h]$. Scegliendo la partizione banale $\Delta_{t,t+h} = \{t, t+h\}$, allora

$$\|\gamma(t+h) - \gamma(t)\| \leq S(t+h) - S(t),$$

dall'altra parte per la (89)

$$S(t+h) - S(t) \leq \int_t^{t+h} \|\gamma'(\tau)\| d\tau.$$

Dividendo per $h > 0$

$$\left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| \leq \frac{S(t+h) - S(t)}{h} \leq \frac{1}{h} \int_t^{t+h} \|\gamma'(\tau)\| d\tau$$

e passando al limite per h che tende a zero, poiché

$$\begin{aligned} \lim_{h \rightarrow 0} \left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| &= \|\gamma'(t)\| \\ \lim_{h \rightarrow 0} \frac{1}{h} \int_t^{t+h} \|\gamma'(\tau)\| d\tau &= \|\gamma'(t)\| \end{aligned}$$

per il teorema del confronto si ottiene che

$$\lim_{h \rightarrow 0} \frac{S(t+h) - S(t)}{h} = \|\gamma'(t)\|.$$

Quindi $S'(t) = \|\gamma'(t)\|$.

Passo 5. Per quanto visto al passo precedente e per il teorema fondamentale del calcolo integrale

$$S(b) - S(a) = \int_a^b \|\gamma'(t)\| dt = \ell_\gamma.$$

Poiché $S(a) = 0$, la definizione di $S(b)$ implica che

$$\sup_{\Delta_{a,b}} \left(\sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| \right) = S(b) = \ell_\gamma.$$

□

APPENDICE A. COMPLEMENTI

A.1. Formula dell'accrescimento finito vettoriale. Per funzioni a valori vettoriali il teorema di Lagrange non è più valido. È tuttavia possibile dare il seguente risultato.

Teo 2.78 (formula dell'accrescimento finito vettoriale). *Data una funzione differenziabile $F : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ con A aperto e convesso, per ogni $P_1, P_2 \in A$ esiste $t \in (0, 1)$ tale che*

$$\|F(P_2) - F(P_1)\| \leq \|JF(P_1 + t(P_2 - P_1))\|_\infty \|P_2 - P_1\| \quad (90) \quad \boxed{\text{eq: 12b}}$$

Dimostrazione. Fissati $P_1, P_2 \in A$, se $F(P_2) = F(P_1)$ non c'è nulla da dimostrare. Altrimenti, posto $n = \frac{F(P_2) - F(P_1)}{\|F(P_2) - F(P_1)\|}$ con $\|n\| = 1$, definiamo $g : A \rightarrow \mathbb{R}$

$$g(P) = n \cdot F(P) = \sum_{i=1}^m n_i f_i(P),$$

che è differenziabile con gradiente

$$\nabla g(P) = \sum_{i=1}^m n_i \nabla f_i(P) = n JF(P)$$

dove $n \in \mathbb{R}^m$ e $\nabla g(F(P)) \in \mathbb{R}^n$ sono visti come vettori riga. Per la formula dell'accrescimento finito applicata a $g(F(P))$ nei punti P_1 e P_2 , esiste $t \in (0, 1)$ tale che

$$\begin{aligned} \|F(P_2) - F(P_1)\| &= g(P_2) - g(P_1) = \nabla g(P_1 + t(P_2 - P_1)) \cdot (P_2 - P_1) \\ &\stackrel{(*)}{=} n JF(P_1 + t(P_2 - P_1))(P_2 - P_1) \\ &\stackrel{(**)}{=} n \cdot \left(JF(P_1 + t(P_2 - P_1))(P_2 - P_1) \right) \\ &\stackrel{(***)}{\leq} \|JF(P_1 + t(P_2 - P_1))(P_2 - P_1)\| \|n\| \\ &\stackrel{****}{\leq} \|JF(P_1 + t(P_2 - P_1))\|_\infty \|P_2 - P_1\|, \end{aligned}$$

dove in $(*)$ $P_2 - P_1$ è pensato come vettore colonna \mathbb{R}^m ed in $(**)$ la matrice jacobiana è identificata con la corrispondente funzione lineare da \mathbb{R}^n ad \mathbb{R}^m , in $(***)$ si è usata la disuguaglianza di Cauchy-Schwarz e in $(****)$ la definizione di norma spettrale di una matrice ed il fatto che n è un versore. \square

A.2. Differenziabilità funzione composta. Il seguente teorema assicura che la composizione di due funzioni differenziabili è differenziabile.

Teo 2.79 (Derivazione in catena). *Date due funzioni differenziabili $F : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $G : B \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$, con A aperto di \mathbb{R}^n , B aperto di \mathbb{R}^m ed $F(A) \subset B$. Se F è differenziabile in P_0 e G è differenziabile in $Q_0 = F(P_0)$, allora la funzione composta $G \circ F : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ è differenziabile in P_0 e vale*

$$J(G \circ F)(P_0) = (JG)(F(P_0)) (JF)(P_0),$$

dove la moltiplicazione è quella righe per colonne tra la matrice $JG(Q_0)$ di taglia $p \times m$ e la matrice $JF(P_0)$ di taglia $m \times n$.

Dimostrazione. Denotiamo con $M = JF(P_0)$ ed $N = JG(Q_0) = JG(F(P_0))$ le matrici jacobiane di F e G . Per definizione di differenziabilità

$$\begin{aligned} F(P) &= F(P_0) + M(P - P_0) + \epsilon(P - P_0) & P \in A \\ G(Q) &= G(Q_0) + N(Q - Q_0) + \epsilon(Q - Q_0) & Q \in B, \end{aligned}$$

dove $\epsilon(P - P_0)$ e $\epsilon(Q - Q_0)$ sono funzioni infinitesime. Allora, per ogni $P \in A$, ponendo $Q = F(P) \in B$ e $Q_0 = F(P_0)$ nella seconda equazione, si ottiene

$$\begin{aligned} G(F(P)) &= G(F(P_0)) + N(F(P) - F(P_0)) + \epsilon(F(P) - F(P_0)) \\ &\stackrel{(*)}{=} G(F(P_0)) + N \left(M(P - P_0) + \epsilon(P - P_0) \right) + \epsilon(P - P_0) \\ &\stackrel{(**)}{=} G(F(P_0)) + NM(P - P_0) + N\epsilon(P - P_0) + \epsilon(P - P_0) \\ &\stackrel{(***)}{=} G(F(P_0)) + NM(P - P_0) + \epsilon(P - P_0) + \epsilon(P - P_0) \\ &= G(F(P_0)) + NM(P - P_0) + \epsilon(P - P_0) \end{aligned}$$

dove in $(*)$, essendo F continua in P_0 , si ha che

$$\epsilon(F(P) - F(P_0)) = \epsilon(\epsilon(P - P_0)) = \epsilon(P - P_0),$$

in $(**)$ si è usato il fatto che $M(v + w) = Mv + Mw$ ed in $(***)$

$$(N\epsilon(P - P_0)) = \epsilon(P - P_0),$$

poiché Nv è una funzione lineare e, quindi, infinitesima. \square

Esempio 2.80. Se una funzione differenziabile $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $\Phi : B \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$

$$\Phi : \begin{cases} x = x(\bar{x}, \bar{y}) \\ y = y(\bar{x}, \bar{y}), \end{cases}$$

è una trasformazione di coordinate tale che $\Phi(\bar{x}, \bar{y}) \in A$ per ogni $(\bar{x}, \bar{y}) \in B$, allora la funzione composta $g(\bar{x}, \bar{y}) = f(\Phi(\bar{x}, \bar{y})) = f(x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y}))$ con dominio B e codominio \mathbb{R} è differenziabile e vale per ogni $(x, y) \in B$

$$\begin{aligned}\frac{\partial g}{\partial \bar{x}}(\bar{x}, \bar{y}) &= \frac{\partial f}{\partial x}(x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y})) \frac{\partial x}{\partial \bar{x}}(\bar{x}, \bar{y}) + \frac{\partial f}{\partial y}(x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y})) \frac{\partial y}{\partial \bar{x}}(\bar{x}, \bar{y}) \\ \frac{\partial g}{\partial \bar{y}}(\bar{x}, \bar{y}) &= \frac{\partial f}{\partial x}(x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y})) \frac{\partial x}{\partial \bar{y}}(\bar{x}, \bar{y}) + \frac{\partial f}{\partial y}(x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y})) \frac{\partial y}{\partial \bar{y}}(\bar{x}, \bar{y})\end{aligned}$$

ovvero in forma compatta

$$\nabla g(\bar{x}, \bar{y}) = \nabla f(\Phi(\bar{x}, \bar{y})) J\Phi(\bar{x}, \bar{y}),$$

dove $\nabla f(\Phi(\bar{x}, \bar{y})) J\Phi(\bar{x}, \bar{y})$ è la moltiplicazione (riga per colonna) del vettore riga $\nabla f(\Phi(\bar{x}, \bar{y}))$ per la matrice $J\Phi(\bar{x}, \bar{y})$ di taglia 2×2 .

A.3. Formula di Taylor. Il seguente teorema estende al caso di funzioni di più variabili la formula di Taylor con il resto di Peano al second'ordine. Valgono risultati analoghi per sviluppi agli ordini successivi e con il resto di Lagrange. Ricordiamo che, data una funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^2 ed un punto $P_0 \in A$, per il teorema di Schwarz $Hf(P_0)$ è una matrice simmetrica, cui è associata la forma quadratica

$$v^t Hf(P_0) v = v \cdot Hf(P_0) v,$$

dove nel membro di sinistra v è pensato come vettore colonna e v^t come vettore riga, mentre in quello a destra $Hf(P_0)$ è identificata con la funzione lineare da \mathbb{R}^n in \mathbb{R}^n . Esplicitamente nel caso $n = 2$, se $P_0 = (x_0, y_0)$ ed $v = (h, k)$

$$\begin{aligned}v^t Hf(P_0) v &= \begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \\ \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &\stackrel{(*)}{=} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) h^2 + \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) h k + \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) k h + \frac{\partial^2 f}{\partial y^2}(x_0, y_0) k^2\end{aligned}$$

dove in $(*)$ il secondo ed il terzo addendo sono uguali per il teorema di Schwarz.

Teo 2.81 (Formula di Taylor). *Data una funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ definita su un aperto A e di classe C^2 e fissato $P_0 \in A$, allora per ogni $v \in \mathbb{R}^n$ tale che $P_0 + v \in A$*

$$f(P_0 + v) = f(P_0) + \nabla f(P_0) \cdot v + \frac{1}{2} v \cdot Hf(P_0) v + \|v\|^2 \epsilon(v). \quad (91) \quad \boxed{\text{eq:10}}$$

dove $\epsilon(v)$ è una funzione infinitesima.

Dimostrazione. Per semplicità supponiamo che $n = 2$ e, senza perdita di generalità, che P_0 sia l'origine O e fissiamo $\delta > 0$ tale che $B(0, \delta) \subset A$. Definiamo il resto $g : B(O, \delta) \rightarrow \mathbb{R}$

$g(x, y) = f(x, y) -$

$$\left(f(0, 0) + \frac{\partial f}{\partial x}(0, 0)x + \frac{\partial f}{\partial y}(0, 0)y + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(0, 0)x^2 + \frac{\partial^2 f}{\partial x \partial y}(0, 0)xy + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(0, 0)y^2 \right).$$

La funzione g è di classe C^2 , $g(0, 0) = 0$

$$\begin{aligned}\frac{\partial g}{\partial x}(x, y) &= \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(0, 0) - \frac{\partial^2 f}{\partial x^2}(0, 0)x - \frac{\partial^2 f}{\partial x \partial y}(0, 0)y \\ \frac{\partial g}{\partial y}(x, y) &= \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(0, 0) - \frac{\partial^2 f}{\partial x \partial y}(0, 0)x - \frac{\partial^2 f}{\partial y^2}(0, 0)y\end{aligned}$$

e, quindi,

$$\nabla g(0, 0) = 0.$$

Infine

$$Hg(x, y) = Hf(x, y) - Hf(0, 0) \quad Hg(0, 0) = 0.$$

Essendo g di classe C^2 , ∇g è differenziabile in $(0, 0)$, cioè ricordando che $J\nabla g = Hg$, per ogni $(x, y) \in B(0, \delta)$

$$\nabla g(x, y) = \nabla g(0, 0) + Hg(0, 0)(x, y) + \|(x, y)\|\epsilon(x, y) = \|(x, y)\|\epsilon(x, y).$$

Essendo $B(O, \delta)$ convesso, dalla formula dell'accrescimento finito applicata a g per ogni $(x, y) \in B(O, \delta)$ esiste $t_{x,y} \in (0, 1)$ tale che

$$\begin{aligned} |g(x, y)| &= |g(x, y) - g(0, 0)| = |\nabla g(t_{x,y}x, t_{x,y}y) \cdot (x, y)| \leq \|\nabla g(t_{x,y}x, t_{x,y}y)\| \|(x, y)\| \\ &= t_{x,y} \|\epsilon(t_{x,y}x, t_{x,y}y)\| \|(x, y)\|^2 \leq \|(x, y)\|^2 \epsilon(x, y). \end{aligned}$$

dove si è sfruttato che $\epsilon(t_{x,y}x, t_{x,y}y)$ è infinitesima e $t_{x,y} \in (0, 1)$.

Abbiamo, quindi, dimostrato che il resto g soddisfa per ogni $(x, y) \in B(0, \delta)$

$$|g(x, y)| \leq \|(x, y)\|^2 \epsilon(x, y)$$

con ϵ infinitesima, che implica la tesi. \square

Se $n = 2$ e $P_0 = (x_0, y_0)$ la (91) diventa

$$\begin{aligned} f(x_0 + h, y_0 + k) &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x_0, y_0)h^2 + \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)hk + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(x_0, y_0)k^2 + \\ &\quad + (h^2 + k^2) \epsilon(h, k). \end{aligned}$$

ovvero, denotato con $P = P_0 + v = (x, y) = (x_0 + h, y_0 + k)$ il punto traslato

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x_0, y_0)(x - x_0)^2 + \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(x_0, y_0)(y - y_0)^2 + \\ &\quad + ((x - x_0)^2 + (y - y_0)^2) \epsilon(x - x_0, y - y_0), \end{aligned}$$

A.4. Caratterizzazione estremi relativi per funzioni scalari in \mathbb{R}^n . Un'analoga caratterizzazione vale per le funzioni di n variabili. A tal fine ricordiamo alcune proprietà delle forme quadratiche associate alle matrici simmetriche. Una matrice simmetrica M di taglia $n \times n$ è detta

- a) definita positiva se $v \cdot Mv > 0$ per ogni $v \in \mathbb{R}^n$, $v \neq 0$;
- b) definita negativa se $v \cdot Mv < 0$ per ogni $v \in \mathbb{R}^n$, $v \neq 0$;
- c) indefinita se esistono $v, w \in \mathbb{R}^n$ tali che $v \cdot Mv > 0$ e $w \cdot Mw < 0$.

Una matrice M simmetrica è diagonalizzabile su \mathbb{R} tramite una matrice ortogonale, cioè esiste una base ortonormale $\{e_1, \dots, e_n\}$ di \mathbb{R}^n ed un insieme di numeri reali $\lambda_1 \leq \dots \leq \lambda_n$ tali che

$$Me_1 = \lambda_1 e_1 \quad \dots \quad Me_n = \lambda_n e_n$$

ed inoltre per ogni $v \in \mathbb{R}^n$

$$\lambda_1 \|v\|^2 \leq v \cdot Mv \leq \lambda_n \|v\|^2. \quad (92) \quad \boxed{\text{eq:82}}$$

Dalla (92) segue immediatamente che

- a) una matrice è definita positiva se e solo se tutti i suoi autovalori sono strettamente positivi: $\lambda_1 > 0$;
- b) una matrice è definita negativa se e solo se tutti i suoi autovalori sono strettamente negativi: $\lambda_n < 0$;
- c) una matrice è indefinita se ha almeno un autovalore strettamente positivo ed uno strettamente negativo: $\lambda_1 < 0 < \lambda_n$.

Vale la seguente condizione sufficiente per l'esistenza di estremi relativi.

Teo 2.82. Data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe C^2 e definita su un insieme aperto A , sia $P_0 \in A$ un punto critico

- i) se $Hf(P_0)$ è definita positiva, allora P_0 è un punto di minimo relativo stretto;
- ii) se $Hf(P_0)$ è definita negativa, allora P_0 è un punto di massimo relativo stretto;
- iii) Se $Hf(P_0)$ è indefinita, P_0 è un punto di sella.

Dimostrazione. Sia $M = Hf(P_0)$, con la notazione introdotta sopra, osserviamo preliminarmente che le condizioni del teorema sono equivalenti a

$$\begin{aligned} i) & \iff \lambda_n \geq \lambda_1 > 0 \\ ii) & \iff \lambda_1 \leq \lambda_n < 0 \\ iii) & \iff \lambda_1 < 0 < \lambda_n \end{aligned}$$

Essendo P_0 interno esiste $\delta > 0$ tale che $B(P_0, \delta) \subset A$. Dalla formula di Taylor (91) per ogni $\|v\| < \delta$

$$\begin{aligned} f(P_0 + v) - f(P_0) &= \nabla f(P_0) \cdot v + \frac{1}{2} v \cdot Hf(P_0)v + \|v\|^2 \epsilon(v) \\ &= \frac{1}{2} v \cdot Hf(P_0)v + \|v\|^2 \epsilon(v), \end{aligned}$$

poiché $\nabla f(P_0) = 0$ essendo per ipotesi P_0 un punto critico.

Se vale l'assunzione i), cioè $\lambda_1 > 0$, allora per la (92)

$$f(P_0 + v) - f(P_0) \geq \left(\frac{\lambda_1}{2} + \epsilon(v)\right) \|v\|^2,$$

e, per definizione di funzione infinitesima, esiste $\delta' < \delta$ tale che $\epsilon(v) > -\lambda_1/4$ per ogni $v \in \mathbb{R}^n$ per cui $\|v\| < \delta'$. Allora, se $0 < \|v\| < \delta'$

$$f(P_0 + v) - f(P_0) \geq \left(\frac{\lambda_1}{2} - \frac{\lambda_1}{4}\right) \|v\|^2 = \frac{\lambda_1}{4} \|v\|^2 > 0$$

da cui $f(P_0 + v) > f(P_0)$, per cui P_0 è un punto di minimo stretto. Se vale l'assunzione ii) si ragiona in modo analogo.

Se vale iii) scegliendo $v = te_1$ con $0 < |t| < \delta$

$$f(P_0 + te_1) - f(P_0) = \left(\frac{\lambda_1}{2} + \epsilon(te_1)\right) t^2 \leq \frac{\lambda_1}{4} t^2 < 0.$$

Analogamente, scegliendo $v = te_n$ con $0 < |t| < \delta'$

$$f(P_0 + te_n) - f(P_0) = \left(\frac{\lambda_n}{2} + \epsilon(te_n)\right) t^2 \geq \frac{\lambda_n}{4} t^2 > 0,$$

per cui se $0 < |t| < \min\{\delta, \delta'\}$

$$f(P_0 + te_1) < f(P_0) < f(P_0 + te_n),$$

e, quindi P_0 è un punto di sella. □

Il seguente esempio mostra che nel caso $n = 2$ il Teo. 2.82 è equivalente a alla Prop. 2.49.

Esempio 2.83. Nel caso $n = 2$, l'equazione singolare della matrice $M = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ è

$$\det(M - \lambda I) = \det \begin{bmatrix} a - \lambda & b \\ b & d - \lambda \end{bmatrix} = \lambda^2 - (a + d)\lambda + (ad - b^2) = 0,$$

che ha sempre due soluzioni reali

$$\lambda_1 = \frac{(a + d) - \sqrt{(a - d)^2 + 4b^2}}{2} \quad \lambda_2 = \frac{(a + d) + \sqrt{(a - d)^2 + 4b^2}}{2}.$$

Da notare che

- a) $\lambda_1 \leq \lambda_2$;

- b) $\det M = \lambda_1 \lambda_2$ e $\text{tr}(M) = a + d = \lambda_1 + \lambda_2$;
c) se $\det M > 0$, cioè $ad > b^2 \geq 0$, allora a e d hanno lo stesso segno e sono entrambi non nulli;
d) la condizione che $\lambda_2 \geq \lambda_1 > 0$ equivale a $ad - b^2 > 0$ e $a > 0$ (oppure $d > 0$);
e) la condizione che $\lambda_1 \leq \lambda_2 < 0$ equivale a $ad - b^2 > 0$ e $a < 0$ (oppure $d < 0$).

Se $e_1, e_2 \in \mathbb{R}^2$ è la base di autovettori di M

$$\begin{aligned} M e_1 &= \lambda_1 e_1 \\ M e_2 &= \lambda_2 e_2 \\ e_i \cdot e_j &= \delta_{ij}, \end{aligned}$$

allora, a meno di cambiare e_1 in $-e_1$, esiste una matrice di rotazione

$$V = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

tale che

$$\begin{aligned} e_1 &= V \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \\ e_2 &= V \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}, \end{aligned} \tag{93} \quad \boxed{\text{eq:35}}$$

da cui segue che M si può scrivere come

$$M = V \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} V^t.$$

Inoltre, poiché $v = h e_1 + k e_2$ con $h^2 + k^2 = \|v\|^2$, allora

$$M v = \lambda_1 h e_1 + \lambda_2 k e_2$$

e la forma quadratica risulta

$$v^t M(P_0) v = \lambda_1 h^2 + \lambda_2 k^2$$

e la (92) diventa

$$\lambda_1 \|v\|^2 \leq v^t M v \leq \lambda_2 \|v\|^2. \tag{94} \quad \boxed{\text{eq:15}}$$

Dalla (94) segue che, se $M \neq 0$, vale una delle seguenti condizioni

$$\begin{array}{llll} M \text{ definita positiva} & \iff & \lambda_2 \geq \lambda_1 > 0 & \iff & ad - bc > 0 \text{ e } a > 0 \\ M \text{ definita negativa} & \iff & \lambda_1 \leq \lambda_2 < 0 & \iff & ad - bc > 0 \text{ e } a < 0 \\ M \text{ indefinita} & \iff & \lambda_1 < 0 < \lambda_2 & \iff & ad - bc < 0 \end{array} \tag{95} \quad \boxed{\text{eq:16}}$$

Se $\det M = 0$, cioè se $\lambda_1 \lambda_2 = 0$, esiste un vettore $v \in \mathbb{R}^2$, $v \neq 0$ tale che $M v = 0$.

Osserviamo che, data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe C^2 e definita su un insieme aperto A , se $P_0 \in A$ un punto critico di f , allora con la notazione introdotta,

$$\begin{aligned} f(P) - f(P_0) &= \frac{\lambda_1}{2} ((P - P_0) \cdot e_1)^2 + \frac{\lambda_2}{2} ((P - P_0) \cdot e_2)^2 + \|P - P_0\|^2 \epsilon(P - P_0) \\ &= g(P) + \|P - P_0\|^2 \epsilon(P - P_0). \end{aligned}$$

Le quantità

$$h = (P - P_0) \cdot e_1 \quad k = (P - P_0) \cdot e_2.$$

descrivono le componenti dell'incremento $P - P_0$ lungo la base e_1, e_2 degli autovettori della matrice Hessiana nel punto critico P_0 . La funzione

$$g(P) = \frac{\lambda_1}{2} ((P - P_0) \cdot e_1)^2 + \frac{\lambda_2}{2} ((P - P_0) \cdot e_2)^2$$

descrive, a meno del resto $\|P - P_0\|^2 \epsilon(P - P_0)$, il corrispondente incremento della funzione f . Se $\lambda_1 \lambda_2 > 0$, per cui P_0 è un punto estremo relativo, l'insieme di livello di g di quota $c/2$

$$C = \{P \in \mathbb{R}^2 \mid g(P) = c/2\};$$

- i) è un'ellisse con centro P_0 , asse maggiore lungo e_1 ed asse minore lungo e_2 , se c ha segno concorde con λ_1 e λ_2 ($\sqrt{c/\lambda_1}$ e $\sqrt{c/\lambda_2}$ sono le lunghezze dei semi-assi);
- ii) si riduce al solo punto P_0 , se $c = 0$;
- iii) è l'insieme vuoto, se $c < 0$.

Nel caso $\lambda_1 < 0 < \lambda_2$, per cui P_0 non è un estremo relativo,

- i) C è un'iperbole con assi le rette con vettori direzionali e_1 ed e_2 , e vertici nei punti $P_0 \pm \sqrt{c/\lambda_2} e_2$, se $c > 0$;
- ii) C si riduce ad una coppia di rette, se $c = 0$;
- iii) C è un'iperbole con assi le rette con vettori direzionali e_1 ed e_2 , e vertici nei punti $P_0 \pm \sqrt{c/\lambda_1} e_1$, se $c < 0$.

Se $\lambda_1 \lambda_2 = 0$, l'approssimazione g non è sufficiente per descrivere il comportamento di f vicino a P_0 .

ese:1

Esempio 2.84. Consideriamo la funzione $f(x, y) = x^2 + xy + y^2$, definita su \mathbb{R}^2 e di classe C^2 . Il gradiente e la matrice hessiana sono

$$\nabla f(x, y) = (2x + y, x + 2y) \quad Hf(x, y) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

L'unico punto critico è l'origine $(0, 0)$ e gli autovalori della matrice hessiana (che non dipende dal punto) sono le soluzioni dell'equazione

$$(2 - \lambda)^2 - 1 = 0 \quad \text{cioè} \quad \lambda_1 = 1 \quad \lambda_2 = 3,$$

ed i rispettivi autovettori sono

$$e_1 = \frac{1}{\sqrt{2}}(1, -1) \quad e_2 = \frac{1}{\sqrt{2}}(1, 1),$$

che si ottengono ruotando la base canonica $(1, 0), (0, 1)$ di un angolo $\theta = -\pi/4$, vedi (93). Poiché entrambi gli autovalori sono strettamente positivi, l'origine $(0, 0)$ è un punto di minimo. Dato $c > 0$, l'insieme di livello di f di quota c è un'ellisse con asse maggiore lungo e_1 e asse minore lungo e_2 .

In effetti, completando a quadrato

$$f(x, y) = (x + \frac{1}{2}y)^2 + \frac{3}{4}y^2 = g\left(\begin{bmatrix} 1 & 0 \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}\right),$$

dove $g(\bar{x}, \bar{y}) = \bar{x}^2 + \bar{y}^2$. Gli insiemi di livello di g ed f si ottengono l'uno dall'altro con la trasformazione di coordinate lineare

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{2}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}$$

Ad esempio, l'insieme di livello di f di quota 1 è ottenuto dilatando la circonferenza unitaria (di centro l'origine) lungo l'asse delle y di un fattore $2/\sqrt{3}$ e quindi effettuando uno "shearing" $\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}$ parallelo all'asse delle x , vedi Figura 9.

A.5. Teorema della mappa inversa locale. Il seguente teorema caratterizza le trasformazioni regolari di coordinate localmente.

Teo 2.85 (Teorema della mappa inversa locale). *Data una funzione $\Phi : B \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita su un aperto B e di classe C^1 ed un punto $P_0 \in B$ tale che*

$$\det J\Phi(P_0) \neq 0, \tag{96}$$

eq:21b

allora esistono due aperti $V \subset B$ e $W \subset \mathbb{R}^n$ che soddisfano le seguenti condizioni:

- a) $P_0 \in V$ e $\Phi(P_0) \in W$;
- b) la funzione Φ ristretta a V è iniettiva e $\Phi(V) = W$;

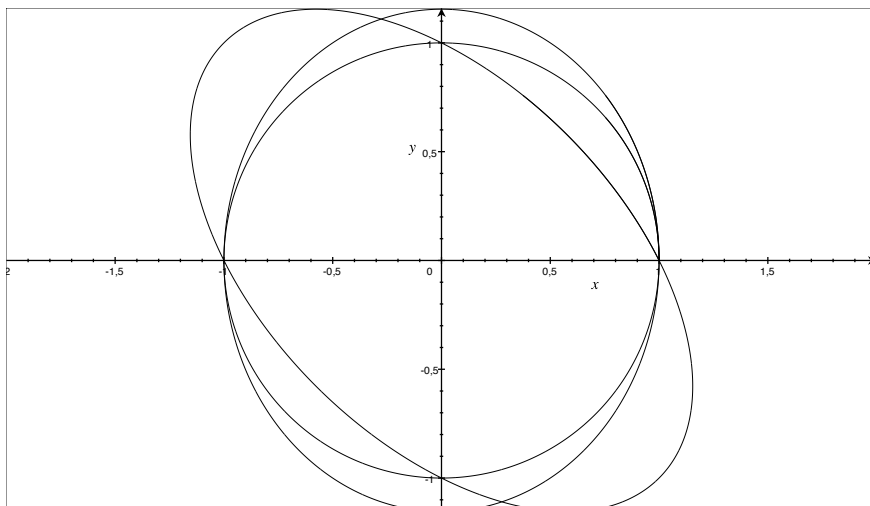


fig:11

FIGURA 9. Esempio 2.84

c) la funzione inversa $\Phi^{-1} : W \rightarrow \mathbb{R}^n$ è di classe C^1 e per ogni $Q \in W$

$$J\Phi^{-1}(Q) = (J\Phi(P))^{-1} \quad \text{con } P = \Phi^{-1}(Q),$$

dove $(J\Phi(P))^{-1}$ è la matrice inversa di $J\Phi(P)$.

Dimostrazione. La dimostrazione è divisa in quattro passi.

1 Passo. Proviamo la tesi nell'ipotesi aggiuntiva che

$$P_0 = 0 \quad \Phi(P_0) = 0 \quad J\Phi(0) = I, \quad (97)$$

eq:65

dove I è la matrice identica. Per dimostrare il teorema, bisogna trovare per ogni $w \in \mathbb{R}^n$ sufficientemente piccolo un unico $v \in B$ tale che $\Phi(v) = w$, cioè che soddisfa l'equazione di punto fisso

$$w - \Phi(v) + v = v.$$

È, quindi, sufficiente dimostrare che $T(v) = w + v - \Phi(v)$ soddisfa le ipotesi del teorema di punto fisso. Definiamo

$$\Omega : B \rightarrow \mathbb{R}^n \quad \Omega(v) = v - \Phi(v),$$

che è di classe C^1 e soddisfa

$$\Omega(0) = 0 \quad J\Omega(0) = I - J\Phi(0) = 0,$$

poiché per la (97) $J\Phi(0)$ è la matrice identica.

Essendo la funzione $\det J\Phi(v)$ continua e, per ipotesi, $\det J\Phi(0) = 1$, esiste $\delta' > 0$ tale che per ogni $v \in B(0, \delta') \subset B$ vale $\det J\Phi(v) \neq 0$ e, quindi, $J\Phi(v)$ è invertibile. Inoltre, poiché la funzione Ω è di classe C^1 , esiste¹ $0 < \delta < \delta'$ tale che

$$\|J\Omega(v)\|_\infty \leq 1/2 \quad \text{per ogni } v \in \overline{B(0, \delta)} \subset B.$$

¹Ad esempio, se $n = 2$, ponendo $v = (x, y)$ e $\Omega(x, y) = (f(x, y), g(x, y))$, la condizione che $J\Omega(0, 0) = 0$ equivale al fatto che

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = \frac{\partial g}{\partial x}(0, 0) = \frac{\partial g}{\partial y}(0, 0).$$

Poiché tutte le derivate parziali sono continue, esiste $\delta > 0$ tale che per ogni $(x, y) \in B(0, \delta)$

$$\|J\Omega(x, y)\|_\infty \leq \|J\Omega(x, y)\|_2 = \sqrt{\frac{\partial f}{\partial x}(x, y)^2 + \frac{\partial f}{\partial y}(x, y)^2 + \frac{\partial g}{\partial x}(x, y)^2 + \frac{\partial g}{\partial y}(x, y)^2} \leq 1/2.$$

Essendo $B(0, \delta')$ convesso, la (90) implica che per ogni $v_1, v_2 \in \overline{B(0, \delta)}$ esiste $w \in \overline{B(0, \delta)}$ (poiché w è sul segmento di estremi v_1 e v_2)

$$\|\Omega(v_2) - \Omega(v_1)\| \leq \|J\Omega(w)\|_\infty \|v_2 - v_1\| \leq \frac{1}{2} \|v_2 - v_1\|. \quad (98) \quad \boxed{\text{eq:17}}$$

In particolare, se $v_2 = v$ e $v_1 = 0$ poiché $\Omega(0) = 0$,

$$\|\Omega(v)\| = \|\Omega(v) - \Omega(0)\| \leq \frac{1}{2} \|v\|.$$

Fissato $w \in \overline{B(0, \delta/2)}$, mostriamo che esiste unico $v \in \overline{B(0, \delta)}$ tale che

$$\Phi(v) = w \iff w + \Omega(v) = v.$$

Definiamo $T : \overline{B(0, \delta)} \rightarrow \mathbb{R}^n$

$$T(v) = w + \Omega(v).$$

Poiché $\|w + \Omega(v)\| \leq \delta/2 + \delta/2 = \delta$, allora $T(\overline{B(0, \delta)}) \subset \overline{B(0, \delta)}$. Inoltre, per ogni $v_2, v_1 \in \overline{B(0, \delta)}$ applicando la (98),

$$\|T(v_2) - T(v_1)\| = \|\Omega(v_2) - \Omega(v_1)\| \leq \frac{1}{2} \|v_2 - v_1\|. \quad (99) \quad \boxed{\text{eq:18}}$$

Ne segue che T soddisfa le ipotesi del teorema di punto fisso con $L = 1/2$, per cui esiste unico $v \in \overline{B(0, \delta)}$ tale che $T(v) = v$, cioè $w = \Phi(v)$.

Riassumendo, abbiamo dimostrato che, per ogni $w \in \mathbb{R}^n$, $\|w\| \leq \delta/2$, esiste unico $v \in \mathbb{R}^n$, $\|v\| \leq \delta$, dipendente da w , tale che

$$\Phi(v) = w \iff w + \Omega(v) = v. \quad (100) \quad \boxed{\text{eq:19}}$$

Definiamo

$$\begin{aligned} W &= B(0, \delta/2) \\ V &= B(0, \delta) \cap \Phi^{-1}(W) \subset B \end{aligned}$$

Evidentemente, W è aperto e $0 \in W$. Inoltre, poiché $\Phi(0) = 0$, $0 \in V$ e V è aperto essendo $\Phi^{-1}(W)$ aperto per la Proposizione 2.20. Dalla (100) segue che Φ ristretta a V è iniettiva e $\Phi(V) = W$, infatti per ogni $w \in W$ esiste un unica $v \in \overline{B(0, \delta)}$ tale che $\Phi(v) = w$ ed inoltre

$$\|v\| = \|v + \Omega(v)\| \leq \|w\| + \|\Omega(v)\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta,$$

che garantisce che $v \in V$. La restrizione di Φ a V è, quindi, invertibile e la sua inversa è data da

$$\Phi^{-1} : W \rightarrow \mathbb{R}^n \quad \Phi^{-1}(w) = v,$$

dove v soddisfa la (100).

Passo 2. Mostriamo che Φ^{-1} è lipschitziana e, quindi, continua. Dati $w_1, w_2 \in W$

$$\begin{aligned} \|\Phi^{-1}(w_2) - \Phi^{-1}(w_1)\| &= \|v_2 - v_1\| = \|(w_2 + \Omega(v_2)) - (w_1 + \Omega(v_1))\| \\ &\leq \|w_2 - w_1\| + \|\Omega(v_2) - \Omega(v_1)\| \\ &\leq \|w_2 - w_1\| + \frac{1}{2} \|v_2 - v_1\| = \|w_2 - w_1\| + \frac{1}{2} \|\Phi^{-1}(w_2) - \Phi^{-1}(w_1)\|, \end{aligned}$$

dove nell'ultima disuguaglianza si è usata la (99). Quindi

$$\|v_2 - v_1\| = \|\Phi^{-1}(w_2) - \Phi^{-1}(w_1)\| \leq 2\|w_2 - w_1\|. \quad (101) \quad \boxed{\text{eq:20}}$$

Passo 3. Ricordiamo che per definizione di δ' e del fatto che $\delta < \delta'$, per ogni $v \in B(0, \delta)$ la matrice $J\Phi(v)$ è invertibile. Mostriamo che Φ^{-1} è differenziabile e $J\Phi^{-1}(w) = J\Phi(v)^{-1}$ dove $v = \Phi^{-1}(w)$. Fissiamo $w_0 \in W$, allora per ogni $w \in W$, posto $v = \Phi^{-1}(w)$ e $v_0 = \Phi^{-1}(w_0)$

$$\begin{aligned} \Phi^{-1}(w) - \left(\Phi^{-1}(w_0) + J\Phi(v_0)^{-1}(w - w_0) \right) &= v - v_0 - J\Phi(v_0)^{-1}(\Phi(v) - \Phi(v_0)) \\ &\stackrel{(*)}{=} v - v_0 - J\Phi(v_0)^{-1} \left(J\Phi(v_0)(v - v_0) + \|v - v_0\|\epsilon(v - v_0) \right) \\ &= -\|v - v_0\| J\Phi(v_0)^{-1} \epsilon(\Phi^{-1}(w) - \Phi^{-1}(w_0)) \\ &\stackrel{(**)}{=} \frac{\|v - v_0\|}{\|w - w_0\|} \|w - w_0\| \epsilon(w - w_0) \stackrel{(***)}{=} \|w - w_0\| \epsilon(w - w_0) \end{aligned}$$

dove in (*), per la differenziabilità di Φ in v_0 , $\epsilon(v - v_0)$ è una funzione infinitesima, in (**), poiché $\Phi^{-1}(w)$ è continua in w_0 e $(J\Phi(v_0))^{-1}$ è una matrice, allora

$$-(J\Phi(v_0))^{-1} \epsilon(\Phi^{-1}(w) - \Phi^{-1}(w_0)) = \epsilon(w - w_0).$$

In (***), per la (101), $\|v - v_0\| \leq 2\|w - w_0\|$. Ne segue che

$$\Phi^{-1}(w) = \Phi^{-1}(w_0) + (J\Phi(v_0))^{-1}(w - w_0) + \|w - w_0\| \epsilon(w - w_0),$$

per cui Φ^{-1} è differenziabile in w_0 e $J\Phi^{-1}(w_0) = J\Phi(v_0)^{-1}$. Infine, essendo Φ di classe C^1 , tutte le derivate parziali di Φ sono continue, quindi sono continui tutti gli elementi della matrice $(J\Phi(\Phi^{-1}(w)))^{-1}$, per cui anche Φ^{-1} è di classe C^1 .

Passo 4. Dimostriamo la tesi nel caso generale. Denotiamo con $M = Jf(P_0)$ la matrice jacobiana di Φ in P_0 che è invertibile per l'ipotesi (96), e consideriamo la funzione

$$\Phi_0 : \{v \in \mathbb{R}^n \mid P_0 + v \in B\} \rightarrow \mathbb{R}^n \quad \Phi_0(v) = M^{-1}(\Phi(P_0 + v) - \Phi(P_0)),$$

che è definita in un insieme aperto contenente l'origine, è di classe C^1 e soddisfa la (97). Applicando a Φ_0 quanto dimostrato nei passi precedenti esistono V_0, W_0 aperti di \mathbb{R}^n tali che $0 \in V_0$, $0 \in \Phi_0(V_0) = W_0$ e Φ_0 ristretta a V_0 è invertibile con inversa Φ_0^{-1} di classe C^1 . Definiamo

$$V = P_0 + V_0 \quad W = MW_0 + \Phi(P_0),$$

che sono insiemi aperti che contengono rispettivamente P_0 e $\Phi(P_0)$. Si verifica immediatamente che la funzione di classe C^1

$$\Psi : W \rightarrow \mathbb{R}^n \quad \Psi(Q) = \Phi_0^{-1}(M^{-1}(Q - \Phi(P_0))) + P_0$$

è l'inversa di Φ . □