Università di Pisa

Emanuele Pardini

Calcolo Differenziale in Spazi di Banach

Indice

Capitolo 1

Funzioni Differenziabili tra Spazi di Banach

1.1 Differenziabilità Secondo Frechèt

Qui E, F saranno spazi di Banach e $\mathcal{L}(E, F)$ lo spazio delle applicazioni lineari continue da E in F.

Definizione 1.1.1: Una $f : \Omega \subset E \to F$, Ω aperto, è detta *differenziabile secondo Frechèt* (e si dirà F-differenziabile o anche solo differenziabile) in $x \in \Omega$ se esiste $L \in \mathcal{L}(E, F)$ t.c.

$$f(x + h) = f(x) + Lh + o(\|h\|) \quad \text{per } \|h\| \to 0$$

In tal caso L è detto *differenziale (di Frechèt)* di f in x e si indica con $Df(x)$.

Mentre diremo che f è *differenziabile* se è differenziabile su ogni punto di ω.

Proposizione 1.1.2: *Se esiste L è unico.*

Dimostrazione. Se ho

$$f(x + h) = f(x) + L_1 h + o(\|h\|)$$

$$f(x + h) = f(x) + L_2 h + o(\|h\|)$$

allora $(L_1 - L_2)h = o(\|h\|)$ e dunque essendo $L_1 - L_2$ lineare si ha che $L_1 = L_2$. □

Vediamo adesso alcune regole di calcolo, la cui dimostrazione segue facilmente dalla definizione di differenziale di Frechèt:

Prodotti Cartesiani

Chiaramente se ho una funzione $f = (f_1, f_2) : \Omega \subset E \to F_1 \times F_2$ ho che f è differenziabile sse lo sono entrambe f_1, f_2.

Composizione

Siano E, F, G spazi di Banach e U aperto di E, V aperto di F. Siano $f : U \to F$ e $g : V \to G$ t.c. $F(U) \subset V$ t.c. f sia differenziabile in x con $Df(x) = L$ e g sia differenziabile in $y = f(x)$ con $Dg(y) = M$ allora $g \circ f : U \to G$ è differenziabile in x con

$$D(g \circ f)(x) = ML = Dg(f(x))Df(x)$$

Inversa

Sia $f : U \to V$ omeomorfismo tra gli aperti $U \subset E$ e $V \subset F$. Ora prendo $x \in U$ e sia $y = f(x)$, supponiamo che f sia differenziabile in x con $Df(x)$ invertibile allora f^{-1} è differenziabile in y con $Df^{-1}(y) = [Df(x)]^{-1}$.

1.2 Teoremi del valor medio

Teorema 1.2.1 (del valor medio di Lagrange): *Sia $f : \Omega \subset E \to \mathbb{R}$, con Ω aperto in E spazio di Banach e siano $x_0, x_1 \in \Omega$ t.c. tutto il segmento $[x_0, x_1] \subset \Omega$. Se f è differenziabile su tutto (x_0, x_1) (segmento aperto che congiunge x_0 e x_1) allora esiste un $y \in (x_0, x_1)$ t.c.*

$$f(x_1) - f(x_0) = Df(y)[x_1 - x_0]$$

Dimostrazione. Basta considerare la funzione

$$[0, 1] \ni t \longmapsto (tx_1 + (1 - t)x_0) \in [x_0, x_1]$$

per ricondurci immediatamente al caso reale. □

Teorema 1.2.2 (del valor medio): *Sia $\gamma : [a, b] \subset \mathbb{R} \to E$, con E spazio di Banach, continua su $[a, b]$ e derivabile in (a, b), allora*

$$\|\gamma(b) - \gamma(a)\| \leq \|\dot{\gamma}\|_{\infty,(a,b)}(b - a)$$

Dimostrazione. Se $\|\dot{\gamma}\|_{\infty,(a,b)} = \infty$ allora non c'è nulla da dimostrare. Supponiamo quindi $\|\dot{\gamma}\|_{\infty,(a,b)} < \infty$ e sia $M > \|\dot{\gamma}\|_{\infty,(a,b)}$. Considero l'applicazione

$$g(t) = \|\gamma(t) - \gamma(a)\| - Mt$$

per $t \in [a, b]$ e vediamo che tale funzione non ha minimo locale in $[a, b)$. Sia $t_0 \in [a, b)$ e considero lo sviluppo al primo ordine

$$\gamma(t) = \gamma(t_0) + \dot{\gamma}(t_0)|t - t_0| + o(|t - t_0|) \quad \text{per } t \to t_0$$

dunque se $t \to t_0$, $t > t_0$ si ha

$$g(t) - g(t_0) \leq \|\gamma(t) - \gamma(t_0)\| - M(t - t_0) \leq$$

$$\leq \|\dot\gamma(t_0) + o(1)\|(t - t_0) - M(t - t_0) \leq$$
$$\leq (\|\dot\gamma(t_0)\| - M + o(1))(t - t_0) < 0$$

in cui l'ultima è vera in quanto $M > \|\dot\gamma\|_{\infty,(a,b)}$, dunque t_0 non può essere un minimo locale per g. D'altra parte g è continua in $[a, b]$ compatto dunque ha minimo per Weierstrass che a questo punto è necessariamente b, e questo implica in particolare che $g(b) \leq g(a)$, cioè

$$\|\gamma(b) - \gamma(a)\| - Mb \leq -Ma$$

ossia

$$\|\gamma(b) - \gamma(a)\| \leq M(b - a)$$

e questo $\forall M > \|\dot\gamma\|_{\infty,(a,b)}$ e quindi vale anche per $M = \|\dot\gamma\|_{\infty,(a,b)}$. □

Il risultato precedente si può generalizzare con dimostrazione analoga al seguente

Teorema 1.2.3 (del valor medio generalizzato): *Sia $\gamma : [a, b] \subset \mathbb{R} \to E$, con E spazio di Banach, continua su $[a, b]$ e derivabile in (a, b) e sia $\psi : [a, b] \subset \mathbb{R} \to \mathbb{R}$ continua, crescente su $[a, b]$ e derivabile in (a, b) t.c. $\|\gamma'(t)\| \leq \psi'(t)$ allora*

$$\|\gamma(b) - \gamma(a)\| \leq \psi(b) - \psi(a)$$

Dimostrazione. La dimostrazione è del tutto analoga a quella precedente in cui si prende

$$g(t) = \|\gamma(t) - \gamma(a)\| - \psi(t)$$

con la quale si ha per $t_0 \in [a, b)$, $t \to t_0$ $t > t_0$

$$g(t) - g(t_0) \leq \|\gamma(t) - \gamma(t_0)\| - (\psi(t) - \psi(t_0)) \leq$$
$$\leq \|\dot\gamma(t_0) + o(1)\|(t - t_0) - (\psi'(t_0) + o(1))(t - t_0) \leq$$
$$\leq (\|\dot\gamma(t_0)\| - \psi'(t_0) + o(1))(t - t_0) < 0$$

e si conclude come nel teorema precedente. □

Ed ora un corollario molto importante

Proposizione 1.2.4 (Lipschitzianità in convessi): *Sia $\Omega \in E$ aperto e $f : \Omega \to F$ differenziabile in Ω con $\|Df\|_{\infty,\Omega} = K < +\infty$. Se Ω è convesso allora f è K-lipschitziana, cioè vale*

$$\|f(x) - f(y)\| \leq K\|x - y\|$$

Dimostrazione. Essendo Ω convesso si ha che $\forall x, y \in E$ $\forall t \in [0, 1]$ $tx + (1 - t)y \in \Omega$ e quindi è lecito unsare il teorema del valor medio e dire che

$$\|f(x) - f(y)\| \leq \sup_{0\leq t\leq 1} \|Df(x + (1 - t)y)\|\|x - y\| \leq$$
$$\leq K\|x - y\|$$

□

1.3 Differenziabilità Secondo Gateaux e Differenziali Parziali

Definizione 1.3.1: Sia $f : \Omega \subset E \to F$, $x \in \Omega$ e $v \in E$, dunque considero la curva $\gamma : t \mapsto x+tv$ e se esiste

$$\partial_v f(x) := \frac{d}{dt} f(x+tv)|_{t=0} = \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} \in F$$

è detta *derivata direzionale* di f in x lungo v.

Definizione 1.3.2: Una $f : \Omega \subset E \to F$ si dice *differenziabile secondo Gateaux* (G-differenziabile) in $x \in \Omega$ se $\forall v \in E$ esiste

$$\partial_v f(x) = \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t}$$

ed esiste un'applicazione lineare L_G t.c. $\forall v \in E$

$$L_G[v] = \partial_v f(x)$$

L_G è detta *differenziale di Gateaux* in x e si indica con $D_G f(x)$.

Dalla regola di composizione si vede che nel caso in cui f è F-differenziabile vale che $\exists \partial_v f(x) = Df(x)[v]$. Dunque se f è F-differenziabile è anche G-differenziabile con $D_G f = Df$, ma una funzione potrebbe avere tutte le derivate direzionali in un punto ma non essere differenziabile su esso come è mostrato nel seguente esempio, dunque una funzione potrebbe essere G-differenziabile ma non F-differenziabile.

Esempio 1.3.3: Prendendo $f : \mathbb{R}^2 \to \mathbb{R}$ t.c.

$$f(x,y) = \begin{cases} \frac{x^4 y^2}{(x^4+y^2)^2} & \text{per } (x,y) \neq (0,0) \\ 0 & \text{per } (x,y) = (0,0) \end{cases}$$

si ha che f ammette tutte le derivate direzionali nulle in $(0,0)$ e quindi $\exists D_G f(0,0) = 0$ ma non è nemmeno continua in tale punto. Infatti se $\mathbf{v} = (l,h)$ si ha che

$$\partial_{\mathbf{v}} f(0,0) = \lim_{t\to 0} \frac{t^4 l^4 t^2 l^2}{t(t^4 l^4 + t^2 h^2)^2} = \lim_{t\to 0} \frac{l^4 h^2}{t^3(l^4 + h^2/t^2)} = \lim_{t\to 0} \frac{t l^4 h^2}{t^2 l^4 + h^2} = 0$$

e non è continua in $(0,0)$ in quanto considerando la restrizione alla parabola $\{(x,y) \in \mathbb{R}^2 | y = x^2\}$

$$\lim_{x\to 0} \frac{x^8}{(2x^4)^2} = \frac{1}{4}$$

Definizione 1.3.4: Sia $\Omega \subset E_1 \times E_2$ aperto, $\forall x \in E_1$ posso considerare $\Omega_2^x = \{y \in E_2 \mid (x,y) \in \Omega\} = (J_2^x)^{-1}(\Omega)$, dove J_2^x è l'immersione di E_2 in $E_1 \times E_2$ tramite $x \in E_1$. E dunque $f : \Omega \to F$ ha *differenziale parziale nella seconda variabile* in (x,y) sse la composizione $f \circ J_2^x : \Omega_2^x \to F$ è differenziabile in $y \in \Omega_2^x$. Ed in tal caso si scrive:

$$D_2 f(x,y) = D(f \circ J_2^x)(y)$$

Analogamente si dice che $f : \Omega \to F$ ha *differenziale parziale nella prima variabile* in (x, y) sse la composizione $f \circ J_1^y : \Omega_1^y \to F$ è differenziabile in $x \in \Omega_1^y$. Ed in tal caso si scrive:

$$D_1 f(x, y) = D(f \circ J_1^y)(x)$$

Osservazione 1.3.5: Dalla definizione è immediato notare che se f è differenziabile in $(x, y) \in \Omega$ allora f ha differenziali parziali nella prima e nella seconda variabile.

Ed inoltre guardando agli sviluppi dati dalla definizione, tenendo conto della regola di composizione e del fatto che $DJ_2^x = J_2^0$ e $DJ_1^y = J_1^0$, si ha che se f è differenziabile in (x, y) allora $D_1 f(x, y) = Df(x, y)[\cdot, 0]$ e $D_2 f(x, y) = Df(x, y)[0, \cdot]$ e quindi in tal caso vale che:

$$Df(x, y)[u, v] = Df(x, y)[u, 0] + Df(x, y)[0, v] = D_1 f(x, y)[u] + D_2 f(x, y)[v]$$

Osservazione 1.3.6: Inoltre, quindi, se f è differenziabile in (x, y) si ha che ad esempio

$$D_1 f(x, y)[u] = Df(x, y)[u, 0] = \partial_{(u,0)} f(x, y)$$

ed analogamente per l'altro (o gli altri se gli spazi sono più di due) differenziale parziale.

Adesso vediamo un teorema importante.

Teorema 1.3.7 (del differenziale totale): *Siano E_1, E_2, F spazi di Banach e $\Omega \subset E_1 \times E_2$ un aperto. Sia $f : \Omega \to F$ e sia $(x_0, y_0) \in \Omega$, se*

- *esiste $D_1 f(x_0, y_0) \in \mathcal{L}(E_1, F)$*
- *$\forall (x, y) \in \Omega$ esiste $D_2 f(x, y)$ continuo in (x_0, y_0)*

allora f è differenziabile in (x_0, y_0) e $\forall (u, v) \in E_1 \times E_2$

$$Df(x_0, y_0)[u, v] = D_1 f(x_0, y_0)[u] + D_2 f(x_0, y_0)[v]$$

Dimostrazione. Vediamo che vale lo sviluppo della definizione di differenziale

$$f(x_0 + u, y_0 + v)f(x_0, y_0) - D_1 f(x_0, y_0)[u] - D_2 f(x_0, y_0)[v] =$$

$$= [f(x_0 + u, y_0 + v) - f(x_0 + u, y_0) - D_2 f(x_0, y_0)[v]] + [f(x_0 + u, y_0) - f(x_0, y_0) - D_1 f(x_0, y_0)[u]]$$

qui il secondo termine è per definizione di differenziale parziale un $o(\|u\|_{E_1})$.

Per il primo termine si può scrivere che è uguale a

$$[f(x_0 + u, y_0 + tv) - tD_2 f(x_0, y_0)[v]]_{t=0}^{t=1} \leq$$

$$\leq \sup_{0 \leq t \leq 1} \|D_2 f(x_0 + u, y_0 + tv)[v] - tD_2 f(x_0, y_0)[v])\|_F$$

in cui si è usato il teorema del valor medio.

Ora

$$\|D_2f(x_0+u,y_0+tv)[v]-tD_2f(x_0,y_0)[v])\|_F \leq$$

$$\leq \|D_2f(x_0+u,y_0+tv)-tD_2f(x_0,y_0))\|_{L(E_2,F)}\|v\|_{E_2}$$

dunque essendo $t\in[0,1]$

$$\|f(x_0+u,y_0+v)-f(x_0+u,y_0)-D_2f(x_0,y_0)[v]\| \leq$$

$$\leq \sup_{\|(x,y)-(x_0,y_0)\|\leq\|(u,v)\|} \|D_2f(x_0+u,y_0+tv)-tD_2f(x_0,y_0))\|_{L(E_2,F)}\|v\|_{E_2} = o(\|v\|_{E_2})$$

per $(u,v)\to 0$, da cui la tesi. □

Definizione 1.3.8: Se $f:\Omega\to F$, $\Omega\subset E$ aperto è differenziabile in ogni $x\in\Omega$ di diciamo che f è differenziabile in Ω ed in tal caso è definito il *differenziale di f in Ω*

$$Df:\Omega\to\mathcal{L}(E,F)$$

Inoltre diciamo che *f è di classe C^1 in Ω* sse esiste il differenziale di f in Ω continuo su Ω.

Dunque dal teorema del differenziale totale si ha il seguente risultato.

Corollario 1.3.9: *Sia $f:\Omega\subset E_1\times E_2\to F$, Ω aperto. f è C^1 sse*

- $\exists D_1f\in C^0(\Omega,\mathcal{L}(E_1,F))$
- $\exists D_2f\in C^0(\Omega,\mathcal{L}(E_2,F))$

Dimostrazione. Se esistono i differenziali parziali continui allora il differenziale di f è la somma dei due parziali per il teorema del differenziale totale e quindi è continuo. Se il differenziale di f esiste continuo allora chiaramente anche i due parziali sono continui. □

E grazie al teorema del valor medio possiamo dare la seguente

Proposizione 1.3.10: *Sia $\Omega\in E$ aperto e $f:\Omega\to F$ differenziabile in Ω con $\|Df\|_{\infty,\Omega}=K<+\infty$. Se Ω è convesso allora f è K-lipschitziana.*

Dimostrazione. Essendo Ω convesso si ha che $\forall x,y\in E$ $\forall t\in[0,1]$ $tx+(1-t)y\in\Omega$ e quindi è lecito unsare il teorema del valor medio e dire che

$$\|f(x)-f(y)\| \leq \sup_{0\leq t\leq 1}\|Df(x+(1-t)y)\|\|x-y\| \leq$$

$$\leq K\|x-y\|$$

□

1.4 Teorema di limite sotto il segno di differenziale

PRIMA O POI ARRIVERÀ.

1.5 Caso di $\mathbb{R}^n$

1.5.1 $\mathbb{R} \to F$: Derivata

Nel caso $\gamma : I \subset \mathbb{R} \to F$, con I intervallo aperto si ha che il differenziale di una tale funzione sta in $\mathcal{L}(\mathbb{R}, F)$ il quale è è in corrispondenza con F tramite la mappa

$$\mathcal{L}(\mathbb{R}, F) \ni (t \mapsto tv) \longmapsto v \in F$$

Denotiamo il vettore corrispondente alla curva γ con la notazione $\dot{\gamma}(x) := D\gamma(x)[1]$ e lo chiamiamo *derivata* di γ in x. In particolare se $F = \mathbb{R}^n$ dalla definizione di differenziale di Frechèt e di derivata (di funzioni $\mathbb{R} \to \mathbb{R}$) si ha che

$$\dot{\gamma}(x) = (\dot{\gamma}_i(x))_{i=1,\dots,n}$$

con $\gamma = (\gamma_i)_{i=1,\dots,n}$.

1.5.2 $E \to \mathbb{R}$: Gradiente

Qui si prende $E = \mathbb{R}^n$ o più in generale spazio di Hilbert (spazio di Banach la cui norma è indotta da un prodotto scalare).

Si ha una corrispondenza tra $\mathcal{L}(E, \mathbb{R}) = E^*$, duale topologico di E, spazio in cui sta il differenziale, ed E tramite la mappa

$$E^* \ni \phi_v \longmapsto v \in E$$

in cui $\phi_v(u) = (u \cdot v)$, prodotto scalare di cui è dotato E. Denotiamo il vettore relativo al differenziale della funzione f in x con la notazione $\nabla f(x)$ e lo chiamiamo *gradiente* di f in x.

Ora nel caso di $E = \mathbb{R}^n$, chiamando $\{\mathbf{e_i}\}_{i=1,\dots,m}$ la base canonica, se ho una funzione differenziabile in un qualche $\mathbf{x} \in \mathbb{R}^n$ si ha che

$$\partial_{\mathbf{e_i}} f(\mathbf{x}) = Df(\mathbf{x})[\mathbf{e_i}] = (\nabla f(x) \cdot \mathbf{e_i})$$

con

$$\partial_{\mathbf{e_i}} f(\mathbf{x}) = Df(\mathbf{x})[\mathbf{e_i}]$$

che è l'i-esimo differenziale parziale della funzione f per quanto detto nell'osservazioni 1 e 2. Dunque chiamando

$$\partial_i f(\mathbf{x}) := \partial_{e_i} f(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{e_i}) - f(\mathbf{x})}{t}$$

la *i-esima derivata parziale* di f si ha che

$$\nabla f(\mathbf{x}) = \sum_{i=1}^{n}(\nabla f(x)\cdot \mathbf{e_i})\mathbf{e_i} = \sum_{i=1}^{n} \partial_i f(\mathbf{x})\mathbf{e_i} = (\partial_i f(\mathbf{x}))_{i=1,\dots,n}$$

Talvolta, per comodità, se x_i è l'i-esima variabile da cui dipende la nostra funzione f, indicherò con $\partial_{x_i} f(\mathbf{x})$ la i-esima derivata parziale di f in $\mathbf{x}$.

1.5.3 $\mathbb{R}^n \to \mathbb{R}^m$: Matrice jacobiana

Qui il differenziale di una funzione $\mathbf{f}: \mathbb{R}^n \to \mathbb{R}^m$ sta in $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ che è in corrispondenza con $M(m, n, \mathbb{R})$ spazio delle matrici $m \times n$ ad entrate reali tramite la mappa

$$\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \ni D \longmapsto (d_{ij} := (\mathbf{e_i} \cdot D\mathbf{e_j}))_{i=1,\dots,m}^{j=1,\dots,n} \in M(m, n, \mathbb{R}).$$

Denotiamo la matrice che corrisponde al differenziale di $\mathbf{f}$ in $\mathbf{x}$ con la notazione $J\mathbf{f}(\mathbf{x})$ e la chiamiamo *matrice jacobiana* di $\mathbf{f}$ in $\mathbf{x}$. Anche quì vale lo stesso discorso fatto nel caso precedente, dunque

$$\partial_{\mathbf{e_i}}\mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{x})[\mathbf{e_i}] = J\mathbf{f}(\mathbf{x})\mathbf{e_i}$$

ma $\mathbf{f} = (f_j)_{j=1,\dots,m}$, con $f_j : \mathbb{R}^n \to \mathbb{R} \quad \forall j = 1, \dots, m$, quindi per il caso precedente componente per componente

$$\partial_{\mathbf{e_i}}\mathbf{f}(\mathbf{x}) = \lim_{t\to 0} \frac{\mathbf{f}(\mathbf{x} + t\mathbf{e_i}) - \mathbf{f}(\mathbf{x})}{t} = (\partial_i f_j(\mathbf{x}))_{j=1,\dots,m}$$

e quindi

$$(J_{\mathbf{f}}(\mathbf{x}))_{ij} = (\mathbf{e_i} \cdot J_{\mathbf{f}}(\mathbf{x})\mathbf{e_j}) = \partial_i f_j(\mathbf{x})$$

i-esima derivata parziale della j-esima componente di $\mathbf{f}$.

Talvolta, per comodità, se x_i è l'i-esima variabile da cui dipende la nostra funzione $\mathbf{f}$ e f_j è la j-esima componente di $\mathbf{f}$, indicherò con $\partial_{x_i} f_j(\mathbf{x})$ la i-esima derivata parziale di f_j in $\mathbf{x}$

Osservazione 1.5.1: La condizione di olomorfia è un caso particolare (quello complesso) della condizione di differenziabilità già data.

1.6 Teorema d'inversione locale

Nel seguito chiamerò $GL(E)$ lo spazio delle applicazioni lineari continue invertibili da $E \to E$.

Lemma 1.6.1: *L'applicazione $Inv : GL(E) \to GL(E)$ t.c. $Inv(A) = A^{-1}$ è differenziabile C^1.*

Dimostrazione. Iniziamo col ricordare che $GL(E)$ è un aperto in $\mathcal{L}(E)$ con la topologia della norma operatoriale, dunque quanto detto ha senso.

Vale che $\forall A \in GL(E)$ $B(A, \|A^{-1}\|^{-1}) \subset GL(E)$ quindi $\forall H \in GL(E)$ con $\|H\| < \|A^{-1}\|^{-1}$ si ha che

$$\|A + H\| \leq \|A(I + A^{-1}H)\| \leq \|A\|\|I + A^{-1}H\| \leq \|A\|(1 + \|A^{-1}\|\|H\|) < \|A\|$$

e quindi $A + H \in GL(E)$. Dunque

$$Inv(A+H) = (A+H)^{-1} = (I + A^{-1}H)^{-1}A^{-1} = (I - (-A^{-1}H))^{-1}A^{-1} =$$

per la serie di Newmann

$$= \left(\sum_{k \geq 0} (-1)^k (A^{-1}H)^k \right) A^{-1} = A^{-1} - A^{-1}HA^{-1} + A^{-1}HA^{-1}HA^{-1} - ... =$$

$$= A^{-1} - A^{-1}HA^{-1} + (A^{-1}H)^2(A+H)^{-1}$$

e quindi

$$Inv(A+H) = Inv(A) - A^{-1}HA^{-1} + o(\|H\|) \quad \text{per } \|H\| \to 0$$

cioè Inv è differenziabile in $A \in GL(E)$ con $DInv(A) = -A^{-1}[\cdot]A^{-1}$.

Infine $DInv : GL(E) \to \mathcal{L}(\mathcal{L}(E), \mathcal{L}(E))$ è chiaramente continua in ogni $A \in GL(E)$ in quanto solamente composizione di composizioni. □

Definizione 1.6.2: Siano $U \subset E$, $V \subset F$ due aperti. Diciamo che $f : U \to V$ è un **diffeomorfismo di classe** C^k sse è un omeomorfismo con $f, f^{-1} \in C^k$.

Ricordiamo un importante Lemma

Proposizione 1.6.3 (Perturbazioni lipschitziane dell'identità): *Sia E uno spazio di Banach, $\Omega \subset E$ un aperto e $g : \Omega \to E$ un'applicazione $K-$lipschitziana di costante $K < 1$. Sia $f = I + g$ allora:*

(1) $f(\Omega)$ è aperto in E.

Precisamente: $\forall a \in \Omega$ $\forall r > 0$ con $\overline{B}(a,r) \subset \Omega$ si ha che $f(\Omega) \supset f(\overline{B}(a,r)) \supset \overline{B}(f(a),(1-k)r)$

(2) $f : \Omega \to f(\Omega)$ è un omeomorfismo bi-lipschitziano.

Precisamente: f è $(1+K)-$lipschitz. e f^{-1} è $\frac{1}{1-K}-$lipschitz. .

Dimostrazione. (1) Provare l'inclusione $f(\overline{B}(a,r)) \supset \overline{B}(f(a),(1-K)r)$, significa provare che: $\forall y \in \overline{B}(f(a),(1-K)r)$ l'equazione $f(x) = y$ ha soluzione $x \in \overline{B}(a,r)$.

Ora fissato $y \in \overline{B}(f(a),(1-K)r)$, si può scrivere che $y = f(x) = x + g(x)$ e quindi che $x = y - g(x)$.

Quindi il problema si può ricondurre ad un problema di punto fisso per l'applicazione

$$x \overset{T_y}{\longmapsto} y - g(x)$$

la quale ha costante di Lipschitz:

$$lip(T_y) = lip(-g) = lip(g) = K < 1$$

Inoltre si ha che $T_y(\overline{B}(a,r)) \subset \overline{B}(a,r)$.

Infatti se $x \in \overline{B}(a,r)$ allora:

$$\|T_y(x) - a\| = \|y - g(x) - a\| = \|y - a - g(a) + g(a) - g(x)\| \leq$$

$$\leq \|y - (a + g(a))\| + \|g(a) - g(x)\| \leq (1-K)r + Kr = r$$

cioè $T_y(x) \in \overline{B}(a,r)$.

Dunque è una contrazione e ammette punto fisso per il Teo. delle Contrazioni.

(2) $f = I + g$ è $(1+K)-$lipschitziana ed inoltre è iniettiva, infatti $\forall x, x' \in \Omega$:

$$\|f(x) - f(x')\| = \|x + g(x) - x' - g(x')\| \geq \|x - x'\| - \|g(x) - g(x')\| \geq$$

$$\geq \|x - x'\| - k\|x - x'\| = (1-k)\|x - x'\| \geq 0$$

ed è $= 0$ sse $x = x'$.

Dunque f è iniettiva e subordina di conseguenza una bigezione $\Omega \xrightarrow{\sim} f(\Omega)$ ovvero $\forall y, y' \in f(\Omega)$ $\exists x, x' \in \Omega$ t.c. $y = f(x)$ e $y' = f(x')$ ossia t.c. $x = f^{-1}(y)$ e $x' = f^{-1}(y')$.

Dunque, riprendendo anche quanto fatto prima:

$$\|y - y'\| \geq (1-K)\|f^{-1}(y) - f^{-1}(y')\|$$

e quindi:

$$\|f^{-1}(y) - f^{-1}(y')\| \leq \frac{1}{(1-K)}\|y - y'\|$$

ovvero f^{-1} è $\frac{1}{(1-K)}-$lipschitziana. □

Teorema 1.6.4 (d'inversione locale): *Siano E, F spazi di Banach e sia $\Omega \subset E$ aperto, $x_0 \in \Omega$. Se $f : \Omega \to F$ è di classe C^1 con $Df(x_0)$ invertibile allora f è un diffeomorfismo tra un intorno aperto U di x_0 e un intorno aperto V di $f(x_0)$.*

Dimostrazione. Sia $A = Df(x_0)$, che è invertibile per ipotesi, dunque considero $g = A^{-1} \circ f$, basta mostrare che g è un diffeomorfismo tra un intorno aperto U di x_0 e un aperto $g(U)$, in quanto poi basta prendere $V = A(g(U)) = f(U)$. Intanto si può dire che g è differenziabile con, in particolare:

$$Dg(x_0) = D(A^{-1} \circ f)(x_0) = A^{-1} \circ Df(x_0) = A^{-1}A = I.$$

Ora $g = I + (g - I)$ e scegliamo $U = B(x_0, r) \subset \Omega$ t.c. $\|Dg - I\|_{\infty, B(x_0,r)} = K < 1$ e questo si può fare perchè g è C^1 e quindi Dg è continuo, in particolare in x_0. Ora $g - I : B(x_0, r) = u \to E$ è K-lipscitz., in quanto $B(x_0, r)$ è convesso, dunque $g = I + (g - I)$ è un perturbazione lipschitziana dell'identità, dunque $g(U) = g(B(x_0, r))$ è aperto e g è omeomorfismo tra U e $g(U)$. Vediamo che è anche diffeomorfismo. $\forall x \in B(x_0, r)$, essendo $\|Dg(x) - I\| \leq K < 1$, si ha per Newmann $Dg(x) = I + (Dg(x) - I) \in GL(E)$ e quindi per la regola di differenziazione dell'inversa $g^{-1} : g(B(x_0, r)) \to B(x_0, r)$ è differenziabile in $g(B(x_0, r))$ con $Dg^{-1}(y) = [Dg(f^{-1}(y))]^{-1}$ che è continua in quanto composizione di applicazioni continue: g_{-1}, Dg, Inv. □

1.7 Teorema della funzione implicita

Teorema 1.7.1 (della Funzione Implicita): *Siano E, F, G spazi di Banach, $\Omega \subset E \times F$ aperto e $f \in C^1(\Omega, G)$. Sia $\Gamma = \{(x,y) \in \Omega | f(x,y) = 0\}$. Sia $(x_0, y_0) \in \Gamma$ e supponiamo che $D_2 f(x_0, y_0) \in \mathcal{L}(E, F)$ sia invertibile. Allora $\exists U$ intorno aperto di x_0, $\exists V$ intorno aperto di y_0, $\exists u : U \to F$ C^1 t.c.*

$$\Gamma \cap (U \times V) = graph(u)$$

e $Du(x_0) = -D_2^{-1} f(x_0, y_0) \circ D_1 f(x_0, y_0)$.

Dimostrazione. Se tale u esiste allora $\forall x \in U \quad f(x, u(x)) = 0$ in quanto $(x, u(x)) \in \Gamma$ da cui differenziando:

$$Df(x, u(x)) = D_1 f(x, u(x)) + D_2 f(x, u(x)) Du(x) = 0$$

in particolare in (x_0, y_0)

$$Du(x_0) = -D_2^{-1} f(x_0, y_0) \circ D_1 f(x_0, y_0)$$

Ora per la prima parte si applica il teorema d'inversione locale alla funzione

$$\Phi : \Omega \subset E \times F \to E \times G$$

t.c.

$$\Phi(x, y) = (x, f(x, y))$$

nel punto (x_0, y_0). Ora $\Phi \in C^1(\Omega, E \times G)$ in quanto sono C^1 le sue componenti e

$$D\Phi(x, y)[u, v] = (u, D_1 f(x, y)[u] + D_2 f(x, y)[v]) \in E \times G$$

vediamo che $D\Phi(x_0, y_0)$ è invertibile. Chiamando $L = D_1 f(x_0, y_0)$ e $M = D_2 f(x_0, y_0)$ si ha che

$$D\Phi(x_0, y_0)[u, v] = (u, L[u] + M[v]) = (a, b)$$

ha unica soluzione

$$(u, v) = (a, -M^{-1} L[a] + M^{-1}[b])$$

dunque $D\Phi(x_0, y_0)$ è bigettiva dunque è invertibile. Dunque per il teorema di inversione locale $\exists W$ intorno aperto di x_0, $\exists V$ intorno di y_0, in E e F rispettivamente, t.c. $W \times V \subset \Omega$ e $\Phi|_{W \times V}$ è diffeomorfismo fra $W \times V$ e un aperto $\Phi(W \times V)$.

Ora sia

$$U = \{x \in W | (x, 0) \in \Phi(W \times V)\} = \{x \in W | \exists y \in V \text{ t.c. } \Phi(x, y) = (x, 0)\} =$$

$$= \{x \in W | \Phi^{-1}(\{x\}) \cap \Gamma \neq \emptyset\} \subset W$$

U è un intorno aperto di x_0 e $\Phi|_{U \times V}; U \times V \to \Phi(U \times V)$ è ancora un diffeomorfismo in quanto ho solo ristretto. Dunque $\forall (x, y) \in U \times V$ vale:

$$(x, y) \in \Gamma \iff f(x, y) = 0 \iff \Phi(x, y) = (x, 0) \iff$$

$$\iff (x,y)=\Phi^{-1}(x,0) \iff y=p_F\circ\Phi^{-1}(x,0)=p_F\circ\Phi^{-1}\circ J_E^0(x)$$

dunque il teorema è dimostrato con $u=p_F\circ\Phi^{-1}\circ J_E^0$ in cui p_F è la proiezione su F e $J_E^0(x)=(x,0)$. □

Osservazione 1.7.2: Lo stesso risultato vale per qualsiasi curva d livello $\Gamma_a=\{(x,y)\in\Omega|f(x,y)=a\}$, la generalizzazione è immediata con $u=p_F\circ\Phi\circ J_E^a$. Inoltre, ovviamente, vale il risultato analogo nel caso di $D_1f(x_0,y_0)$ invertibile, trovando U e V come nell'enunciato ed una funzione $v:V\to U$ C^1.

Capitolo 2

Differenziale secondo

2.1 Differenziale secondo e sue proprietà

Definizione 2.1.1: Siano E, F spazi di Banach, $\Omega \subset E$ aperto e $f : \Omega \to F$ differenziabile su Ω e supponiamo che $Df : \Omega \to \mathcal{L}(E, F)$ sia a sua volta differenziabile in x_0, dunque definiamo il *differenziale secondo* di f in x_0 come

$$D^2 f(x_0) := D(Df)(x_0) \in \mathcal{L}(E, \mathcal{L}(E, F))$$

dove $\mathcal{L}(E, \mathcal{L}(E, F))$ è uno spazio di Banach (con la norma operatoriale) naturalmente isometrico allo spazio delle applicazioni bilineari continue $\mathcal{L}^2(E^2, F)$ con la norma $\|b\|_{\mathcal{L}^2(E^2,F)} := \|b\|_{\infty, B_E \times B_E}$ in cui $B_E := B(\mathbf{0}, 1)$, che non è altro che l'analogo della norma operatoriale per il caso bilineare.

Vediamo una proposizione molto utile.

Proposizione 2.1.2: *Siano E, F spazi di Banach, $\Omega \subset E$ aperto e $f : \Omega \to F$ differenziabile su Ω. Supponiamo che $Df : \Omega \to \mathcal{L}(E, F)$ sia a sua volta differenziabile in x_0, allora vale lo sviluppo*

$$f(x_0 + u + v) = f(x_0) + f(x_0 + u) + f(x_0 + v) + D^2 f(x_0)[u, v] + o([\|u\| + \|v\|]^2)$$

per $\|(u, v)\| \to 0$. Inoltre vale

$$D^2 f(x)[u, v] = \frac{\partial^2}{\partial t \partial s} f(x + tu + sv)|_{s,t=0} = \partial_u \partial_v f(x)$$

Dimostrazione.

$$\|f(x_0 + u + v) - f(x_0) - f(x_0 + u) - f(x_0 + v) - D^2 f(x_0)[u, v]\|_F =$$

$$= \|[f(x_0 + u + tv) + f(x_0 + tv) - tD^2 f(x_0)[u, v]_{t=0}^{t=1}\|_F \leq$$

per il teorema del valor medio

$$\leq \sup_{t \in [0,1]} \|Df(x_0 + u + tv)[v] - Df(x_0 + tv)[v] - D^2 f(x_0)[u, v]\|_F =$$

$$= \sup_{t\in[0,1]} \|Df(x_0+u+tv)[v] - Df(x_0+tv)[v] - D(Df)(x_0)[u,v]\|_F \leq$$

$$\leq \sup_{t\in[0,1]} \|Df(x_0+u+tv) - Df(x_0+tv) - D(Df)(x_0)[u]\|_{\mathcal{L}(E,F)}\|v\|_E =$$

sostituendo con lo sviluppo del differenziale per Df:

$$= \sup_{t\in[0,1]} \|Df(x_0) + D(Df)(x_0)[u+tv] + o(\|u\|+\|v\|) -$$

$$-Df(x_0) - D(Df)(x_0)[tv] - o(\|v\|) - D(Df)(x_0)[u]\|_{\mathcal{L}(E,F)}\|v\|_E =$$

cancellando il cancellabile

$$= \sup_{t\in[0,1]} \|o(\|u\|+\|v\|)\|\|v\|_E = o(\|u\|+\|v\|)\cdot\|v\|_E \leq$$

$$\leq o(\|u\|+\|v\|)\cdot(\|u\|+\|v\|) = o([\|u\|+\|v\|]^2) \text{ per } \|(u,v)\| \to 0$$

Per l'altra parte si ha che

$$\frac{\partial}{\partial s} f(x+tu+sv)|_{s=0} = Df(x+tu+sv)[v]|_{s=0} = Df(x+tu)[v]$$

dunque

$$\frac{\partial^2}{\partial t\partial s} f(x+tu+sv)|_{t,s=0} = \frac{\partial}{\partial t}\left(\frac{\partial}{\partial s} f(x+tu+sv)|_{s=0}\right)|_{t=0} =$$

$$= \frac{\partial}{\partial t}(Df(x+tu)[v])|_{t=0} =$$

qui considerando la valutazione in v, EV_v, si ha che questa è chiaramente un'applicazione lineare, dunque se chiamo $\Gamma(t) = Df(x+tu)$ si ha che

$$\frac{\partial}{\partial t}(EV_v \circ \Gamma)(t) = DEV_v\left[\frac{\partial}{\partial t}\Gamma(t)\right] = EV_v\left[\frac{\partial}{\partial t}\Gamma(t)\right]$$

dunque è lecito scrivere

$$= \left(\frac{\partial}{\partial t}(Df(x+tu))\right)[v]|_{t=0} =$$

$$= D(Df)(x)[u][v] = D^2 f(x)[u,v].$$

□

Adesso l'importante proprietà di simmetria.

Lemma 2.1.3: $b \in \mathcal{L}^2(E^2, F)$.

$$b(u,v) = o([\|u\|+\|v\|]^2) \textit{ per } \|(u,v)\| \to 0 \iff b = 0$$

Dimostrazione. ($\Leftarrow$) ovvia.

($\Rightarrow$) $\forall u, v \in E$ prendo $t \in \mathbb{R}$, $t \to 0$ allora:

$$b(tu, tv) = o(t^2[\|u\| + \|v\|]^2) = o(t^2)\text{per } t \to 0$$

ma $b(tu, tv) = t^2 b(u, v)$ dunque

$$b(u, v) = o(1)$$

e quindi $b = 0$. $\square$

Proposizione 2.1.4 (Simmetria del differenziale Secondo): *Siano E, F spazi di Banach, $\Omega \subset E$ aperto e $f : \Omega \to F$ differenziabile su Ω, se $\exists D^2 f(x_0)$, con $x_0 \in \Omega$, allora è un'applicazione bilineare simmetrica.*

Dimostrazione. Per lo sviluppo dimostrato precedentemente si ha che

$$D^2 f(x_0)[u, v] = D^2 f(x_0)[v, u] = o([\|u\| + \|v\|]^2)$$

per $\|(u, v)\| \to 0$, dunque si conclude col lemma precedente.

$\square$

2.2 Teorema di Schwarz

In questa sezione intenderemo, per una funzione $f : I \times J \to E$ con $I, J \subset \mathbb{R}$ ed E spazio di Banach, con

$$(I \ni t \mapsto \partial_1 f(x, y)t \in E) = \partial_1 f(x, y)[\cdot] := D_1 f(x, y)$$

il differenziale parziale nella prima variabile di f in (x, y) che per definizione, ad esempio nella prima variabile, è

$$\partial_1 f(x, y)[\cdot] = D(f \circ J_1^y)(x)$$

che è il differenziale di una funzione da $\mathbb{R} \to E$.

$\partial_1 f(x, y)$ è quindi una derivata nella nomenclatura introdotta in 1.5.1.

Discorso analogo vale anche nella seconda variabile chiaramente.

Ora se f ammette derivata parziale nella prima variabile (ad esempio) su tutto $I \times J$, allora esiste la sua derivata nella prima variabile su tutto $I \times J$

$$\partial_1 f : I \times J \to E$$

se questa è a sua volta differenziabile nella seconda variabile (ad esempio) in (x, y), posso definire

$$\partial_{21} f(x, y) := \partial_2(\partial_1 f)(x, y) \in E$$

derivata seconda nella seconda variabile della derivata nella prima variabile di f.

Detto questo siamo pronti per il teorema

Teorema 2.2.1 (di Schwarz): *Siano $I, J \subset \mathbb{R}$ intervalli, E spazio di Banach, $x_0 \in I$ e $f : I \times J \to E$ t.c.*

- $\exists \partial_1 f(x,y) \quad \forall (x,y) \in I \times J$
- $\exists \partial_{21} f(x,y) \quad \forall (x,y) \in I \times J$ *continua in* (x_0, y_0)
- $\exists \partial_2 f(x, y_0) \quad \forall x \in I$

allora esiste anche $\partial_{12} f(x_0, y_0)$ e coincide con $\partial_{21} f(x_0, y_0)$.

Dimostrazione. Vale lo sviluppo

$$\|f(x_0+u, y_0+v) - f(x_0, y_0+v) + f(x_0+u, y_0) + f(x_0, y_0) - \partial_{21} f(x_0, y_0)uv\| =$$

$$= \|[f(x_0+tu, y_0+v) + f(x_0+tu, y_0) - t\partial_{21} f(x_0, y_0)uv]_{t=0}^{t=1}\| \leq$$

per il teorema del valor medio

$$\leq \sup_{t\in[0,1]} \|\partial_1 f(x_0+tu, y_0+v)u - \partial_1 f(x_0+tu, y_0)u - \partial_{21} f(x_0, y_0)uv\| =$$

$$= \sup_{t\in[0,1]} \|[\partial_1 f(x_0+tu, y_0+sv)u - s\partial_{21} f(x_0, y_0)uv]_{s=0}^{s=1}\| \leq$$

$$\leq \sup_{t,x\in[0,1]} \|\partial_{21} f(x_0+tu, y_0+sv)uv - \partial_{21} f(x_0, y_0)uv\| =$$

$$\sup_{t,x\in[0,1]} \|\partial_{21} f(x_0+tu, y_0+sv) - \partial_{21} f(x_0, y_0)\||u||v| = o(|u||v|)$$

per $(u, v) \to (0, 0)$.

Dunque dividendo per v

$$\left\|\frac{f(x_0+u, y_0+v) - f(x_0+u, y_0)}{v} + \frac{f(x_0, y_0+v) - f(x_0, y_0)}{v} - \partial_{21} f(x_0, y_0)u\right\| =$$

$$= \|\partial_2 f(x_0+u, y_0) - \partial_2 f(x_0, y_0) + \partial_{21} f(x_0, y_0)u\| = o(|u|)$$

e passando al limite per $(u, v) \to (0, 0)$ e poi dividendo per u

$$\left\|\frac{\partial_2 f(x_0+u, y_0) - \partial_2 f(x_0, y_0)}{u} + \partial_{21} f(x_0, y_0)\right\| =$$

$$\|\partial_{12} f(x_0, y_0) - \partial_{21} f(x_0, y_0) = o(1)$$

per $(u, v) \to (0, 0)$, dunque $\exists \partial_{12} f(x_0, y_0) = \partial_{21} f(x_0, y_0)$.

□

2.2.1 Caso $\mathbb{R}^n \to \mathbb{R}$: matrice hessiana

Se $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, Ω aperto, ha differenziale secondo in $\mathbf{x} \in \Omega$, $D^2 f(\mathbf{x})$,allora sappiamo che è simmetrico come elemento di $\mathcal{L}^2(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$.

Ora sia $b \in \mathcal{L}^2(R^n \times \mathbb{R}^n, \mathbb{R})$ e $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, se $\{\mathbf{e_i}\}_{i=1,\dots,n}$ è la base canonica di $\mathbb{R}^n$

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e_i} \qquad \mathbf{y} = \sum_{j=1}^{n} y_j \mathbf{e_j}$$

dunque

$$b[\mathbf{x}, \mathbf{y}] = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j b[\mathbf{e_i}, \mathbf{e_j}]$$

Dunque si ha una corrispondenza

$$\mathcal{L}^2(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}) \ni b \longleftrightarrow B = (b_{ij})_{i,j=1,\dots,n} \in M(n, n, \mathbb{R})$$

in cui

$$b_{ij} = (\mathbf{e_i} \cdot B\mathbf{e_j}) = b[\mathbf{e_i}, \mathbf{e_j}]$$

La matrice che rappresenta $D^2 f(\mathbf{x})$ è detta *matrice hessiana* di f in $\mathbf{x}$ e viene indicata con $H_f(\mathbf{x})$.

Ora, dalla Proposizione 4

$$(H_f(\mathbf{x}))_{ij} = D^2 f(\mathbf{x})[\mathbf{e_i}, \mathbf{e_j}] = \frac{\partial^2}{\partial t \partial s} f(x + t\mathbf{e_i} + s\mathbf{e_j})|_{s,t=0} = \partial_{ij} f(\mathbf{x})$$

Capitolo 3

Formula di Taylor

3.1 Differenziali di ordine superiore

Analogamente a come si è definito il differenziale secondo, si può naturalmente definire differenziali di ogni ordine per una funzione $f : E \to F$ tra spazi di Banach.

Chiaramente per ogni intero m vale che $\mathcal{L}(E, \mathcal{L}(E, ..., \mathcal{L}(E, F))...)$ m volte è identificabile con $\mathcal{L}^m(E^m, F)$, spazio delle applicazioni $m-$lineari continue da $E^m \to F$.

$\mathcal{L}^m(E^m, F)$ è sempre uno spazio di Banach con la norma

$$\|L\|_{\mathcal{L}^m(E^m,F)} := \sup_{\|h_i\|_E=1} \|L[h_1, ..., h_m]\|_F$$

Definizione 3.1.1: Siano E, F spazi di Banach, $\Omega \subset E$ aperto e $f : \Omega \to F$ una funzione che ammette differenziale $(m-1)$-esimo su Ω e che tale differenziale

$$D^{m-1}f : \Omega \to \mathcal{L}(E^{m-1}, F)$$

sia a sua volta differenziabile in un $x \in \Omega$, allora è definito il *differenziale m-esimo* di f in x

$$D^m f(x) \in \mathcal{L}^m(E^m, F)$$

.

Inoltre dico che f *è di classe* C^m *in* Ω se esiste il differenziale m-esimo di f in ogni punto di Ω e $D^m f$ è continuo su Ω.

E vale il seguente risultato di simmetria

Proposizione 3.1.2: *Se esiste* $D^m f(x)$ *è simmetrico.*

Dimostrazione. Per induzione su m.

$(m = 1)$

ovvio.

$(m-1) \Rightarrow (m)$

$D^{m-1}f$ esiste in un intorno di x per ipotesi induttiva simmetrico.

Ora

$$D^m f(x) = D^2(D^{m-2}f)(x)$$

dunque è simmetrico nelle prime due variabili per la simmetria del differenziale secondo, ma allora è simmetrico nelle prime due variabili e dalla seconda all'm-esima, dunque è totalmente simmetrico.

□

3.2 Formula di Taylor

Se $D^m f(x)$ è il differenziale m-esimo di f in x useremo la notazione

$$D^m f(x)[h]^m := D^m f(x) \underbrace{[h...h]}_{\text{m volte}}$$

Definizione 3.2.1: Preso $\Omega \subset E$ aperto e $f : \Omega \to F$ t.c. sia differenziabile n volte in x_0, definisco il *polinomio di Taylor* di f di ordine n centrato in x_0, la funzione

$$T_n f(x_0, x) := \sum_{i=0}^{n} \frac{1}{i!} D^i f(x_0)[x - x_0]^i$$

Dunque si ha il seguente

Teorema 3.2.2 (Formula di Taylor con resto di Peano): *Siano E, F spazi di Banach, $\Omega \subset E$ aperto e $f : \Omega \to F$ di classe C^n.*

Se $x_0 \in \Omega$ allora vale

$$f(x_0 + h) = T_n f(x_0)(x_0 + h) + o(\|h\|^n) = \sum_{i=0}^{n} \frac{1}{i!} D^i f(x_0)[h]^i + o(\|h\|^n)$$

Dimostrazione. Sia $r > 0$ t.c. $B(x_0.r) \subset \Omega$ e fisso $\varphi \in F^*$ t.c. $\|\varphi\|_{F^*} \leq 1$, dunque $\forall h \in B(0, r)$ considero la funzione reale $g_h : [-1, 1] \to \mathbb{R}$ t.c

$$g_h(t) = \langle \varphi, f(x_0 + th) \rangle$$

ora per la regola della differenziazione di funzioni composte questa è C^n con

$$g_h'(t) = \langle \varphi, Df(x_0 + th)[h] \rangle$$

$$g_h''(t) = \langle \varphi, D^2 f(x_0 + th)[h, h] \rangle$$

e proseguendo si ottiene che $\forall k_1^n$

$$g_h^{(k)}(t) = \langle \varphi, D^k f(x_0 + th)[h]^k \rangle$$

ora per la formula di Taylor per g_h all'ordine n-esimo del lemma precedente

$$g_h(1) - g_h(0) = \sum_{i=1}^{n-1} \frac{1}{i!} g_h^{(i)}(0) + \eta(1)$$

con $\eta(t) = o(1)$ per $t \to 0$.

Da cui

$$\langle \varphi, f(x_0 + h) - f(x_0) \rangle = \sum_{i=1}^{n-1} \frac{1}{i!} \langle \varphi, D^i f(x_0)[h]^i \rangle + \frac{1}{n!} \langle \varphi, D^n f(x_0 + \xi_h h)[h]^n \rangle =$$

$$= \sum_{i=1}^{n} \frac{1}{i!} \langle \varphi, D^i f(x_0)[h]^i \rangle + \frac{1}{n!} \langle \varphi, D^n f(x_0 + \xi_h h)[h]^n \rangle - \langle \varphi, D^n f(x_0)[h]^n \rangle$$

dunque sfruttando la linearità di φ e prendendo i valori assoluti ho

$$\left| \left\langle \varphi, f(x_0 + h) - \sum_{i=0}^{n} \frac{1}{i!} D^i f(x_0)[h]^i \right\rangle \right| = |\langle \varphi, \eta(h) \rangle|$$

con $\eta(h) = \frac{1}{n!}(D^n f(x_0 + \xi_h h)[h]^n - D^n f(x_0)[h]^n)$, dunque

$$\left| \left\langle \varphi, f(x_0 + h) - \sum_{i=0}^{n} \frac{1}{i!} D^i f(x_0)[h]^i \right\rangle \right| \leq \|\varphi\|_{F^*} \|\eta(h)\|_F$$

e per l'arbitrarietà di $\varphi \in F^*$ (è un corollario del Teorema di Hahn-Banach, vedi Brezis, Corollary 1.4) si ottiene

$$\left\| f(x_0 + h) - \sum_{i=0}^{n} \frac{1}{i!} D^i f(x_0)[h]^i \right\|_F \leq \|\eta(h)\|_F$$

Dunque per la continuità di $D^n f$ si ha che

$$\|\eta(h)\|_F = \left\| \frac{1}{n!} (D^n f(x_0 + \xi_h h)[h]^n - D^n f(x_0)[h]^n) \right\|_F =$$

$$= \left\| \frac{1}{n!} (D^n f(x_0 + \xi_h h) - D^n f(x_0))[h]^n \right\|_F \leq$$

$$\leq \frac{1}{n!} \|D^n f(x_0 + \xi_h h) - D^n f(x_0)\|_{\mathcal{L}^n(E^n, F)} \|h\|_E^n = o(\|h\|_E^n)$$

e quindi la tesi. □

Inoltre vale il seguente, che è la cosa più vicina ad un resto di Lagrange che si riesce ad ottenere in questo contesto

Teorema 3.2.3 (Maggiorazione del resto di Lagrange): *Sia $\Omega \subset E$ aperto, $x_0, x \in \Omega$ t.c. il segmento $[x_0, x] \subset \Omega$ e $f : \Omega \to F$ differenziabile $n+1$ volte sui punti di $[x_0, x]$ allora vale che*

$$\|f(x) - T_n f(x_0, x)\| \leq \frac{1}{(n+1)!} \|D^{n+1} f\|_{\infty, [x_0, x]} \|x - x_0\|^{n+1}$$

in cui $\|D^{n+1} f\|_{\infty, [x_0, x]} = \sup_{v \in [x_0, x]} \|D^{n+1} f(v)\|_{\mathcal{L}^{n+1}(E^{n+1}, F)}$.

Dimostrazione. Se $\|D^{n+1}f\|_{\infty,[x_0,x]} = +\infty$ allora non c'è niente da dimostrare.

Supponiamo quindi $M = \|D^{n+1}f\|_{\infty,[x_0,x]} < +\infty$ e consideriamo la funzione

$$g(t) = f(tx + (1-t)x_0)$$

questa per ipotesi è una funzione $n+1$ volte differenziabile $\forall t \in [0,1]$ e si vede bene che

$$g^{(k)}(t) = D^k f(tx + (1-t)x_0)[x - x_0]^k$$

per ogni $k = 1, ..., n$, infatti:

$$\frac{d}{dt} D^{k-1} f(tx + (1-t)x_0)[x - x_0]^{k-1} = \frac{d}{dt}\{(EV_{x-x_0} \circ D^{k-1} f \circ \gamma)(t)\} =$$

in cui $EV^i_{x-x_0}$ è l'operatore di valutazione in $[x-x_0]^i$ che opera in $\mathcal{L}^i(E^i, F)$ e $\gamma(t) = tx-(1-t)x_0$, che per la regola di composizione è

$$= EV^{k-1}_{x-x_0}[D^k f(tx - (1-t)x_0)[x - x_0, \cdot, ..., \cdot]] = D^k f(tx + (1-t)x_0)[x - x_0]^k$$

dunque si ha in particolare

$$\|g^{(n+1)}(t)\| \leq M\|x - x_0\|^{n+1}$$

per ogni $t \in [0,1]$.

Ora siano $\varphi : [0,1] \to F$, $\psi : [0,1] \to \mathbb{R}$ t.c.

$$\varphi(t) = \sum_{i=0}^{n} \frac{1}{i!} g^{(i)}(t)(1-t)^i$$

$$\psi(t) = -\frac{M\|x - x_0\|^{n+1}}{(n+1)!}(1-t)^{n+1}$$

ora ψ è crescente in t e $\forall t \in [0,1]$

$$\|\varphi'(t)\| = \|\frac{1}{n!} g^{(n+1)}(t)(1-t)^n\| \leq$$

per quanto detto prima

$$\leq \frac{1}{n!} M\|x - x_0\|^{n+1}(1-t)^n = \psi'(t)$$

dunque si può applicare il teorema del valor medio generalizzato e scrivere

$$\|\varphi(1) - \varphi(0) \leq \psi(1) - \psi(0)$$

ossia, essendo $\varphi(1) = g(1)$ e $\psi(1) = 0$

$$\|g(1) - \sum_{i=0}^{n} \frac{1}{i!} g^{(i)}(0)(1-t)^i\| \leq \frac{M\|x - x_0\|^{n+1}}{(n+1)!}$$

che è

$$\|f(x) - T_n f(x_0, x)\| \leq \frac{M\|x - x_0\|^{n+1}}{(n+1)!}$$

la tesi. □

Capitolo 4

Massimi e Minimi

In questa sezione tratteremo di condizioni necessarie o sufficienti per avere massimi o minimi per funzioni $f : E \to \mathbb{R}$ con E spazio di Banach.

4.1 Massimi e minimi liberi

Definizione 4.1.1: Sia E spazio di Banach, diciamo che $b \in \mathcal{L}_s^2(E, \mathbb{R})$ è *positiva* se $\forall x \in E$

$$b(x, x) \geq 0$$

Teorema 4.1.2 (Condizioni necessarie per estremi liberi): *Sia E spazio di Banach, $\Omega \subset E$ aperto e $f : \Omega \to \mathbb{R}$. Se $x_0 \in \Omega$ è un punto di minimo (massimo) locale allora*

- *se esiste $Df(x_0)$ è nullo*
- *se esiste $D^2 f(x_0)$ è positivo (negativo).*

Dimostrazione. Per ogni $u \in E$ posso considerare la funzione

$$\mathbb{R} \ni t \longmapsto f(x_0 + tu) \in \mathbb{R}$$

se f ha un minimo (massimo) in x_0 allora anche questa curva deve averne uno, dunque per quanto noto sulle funzioni reali si ha che

$$\frac{d}{dt} f(x_0 + tu)|_{t=0} = \partial_u f(x_0) = 0$$

dunque se f ha $Df(x_0)$, essendo $\partial_u f(x_0) = Df(x_0)[u]$, si ha necessariamente $Df(x_0)$ nullo.

Inoltre sempre per quanto noto sulle funzioni reali si ha che

$$\frac{d^2}{dt^2} f(x_0 + tu)|_{t=0} \geq 0 \quad (\leq 0)$$

dunque se esiste $D^2 f(x_0)$, essendo $\frac{d^2}{dt^2} f(x_0 + tu)|_{t=0} = D^2 f(x_0)[u, u]$, si ha necessariamente $D^2 f(x_0) \geq 0 \quad (\leq 0)$. □

Definizione 4.1.3: Sia E spazio di Banach, diciamo che $b \in \mathcal{L}_s^2(E, \mathbb{R})$ è *definita positiva* se $\exists \alpha > 0$ t.c. $\forall x \in E$
$$b(x,x) \geq \alpha \|x\|^2$$
Mentre diciamo che b è *definita negativa* se $\exists \alpha < 0$ t.c. $\forall x \in E$
$$b(x,x) \leq \alpha \|x\|^2$$

Osservazione 4.1.4: Nel caso in cui E abbia dimensione finita la definizione di applicazione definita positiva data è equivalente a dire che $\forall x \in E - \{0\}$
$$b(x,x) > 0$$
Infatti E è uno spazio di Banach finito dimensionale, dunque la palla unitaria è compatta (vedi Clarke pag. 14) e per ogni $x \neq 0$ si ha
$$b(x,x) = \|x\|^2 b\left(\frac{x}{\|x\|}, \frac{x}{\|x\|}\right) \geq \alpha \|x\|^2$$
con $\alpha = \min_{y \in B(0,1)} b(y,y)$, che esiste per compattezza di $B(0,1)$. Il viceversa è ovvio.

Analogamente, ma prendendo il massimo, si ottiene che in dimensione finita definita negativa è equivalente a dire che $\forall x \in E - \{0\}$
$$b(x,x) < 0.$$

Teorema 4.1.5 (Condizioni sufficienti per estremi liberi): *Sia E spazio di Banach, $\Omega \subset E$ aperto e $f : \Omega \to \mathbb{R}$. Se f ammette $D^2 f(x_0)$ definito positivo (negativo) e $Df(x_0) = 0$ allora x_0 è un punto di minimo (massimo) stretto locale (stretto nel senso che è l'unico dentro ad un'opportuna palla).*

Dimostrazione. Per la formula di Taylor con resto di Peano si ha per $\|u\| \to 0$
$$f(x_0 + u) = f(x_0) + Df(x_0)[u] + \frac{1}{2} D^2 f(x_0)[u,u] + o(\|u\|^2) =$$
$$= f(x_0) + \frac{1}{2} D^2 f(x_0)[u,u] + o(\|u\|^2) \geq$$
$$\geq f(x_0) + \|u\|^2 \left(\frac{\alpha}{2} + o(1)\right)$$
con $\alpha > 0 \ \ (< 0)$ da cui la tesi. $\square$

4.2 Moltiplicatori di Lagrange

Ricordiamo (senza dimostrarlo) il seguente teorema riguardante la compattezza delle palle chiuse in uno spazio di Banach.

Teorema 4.2.1: *Sia E uno spazio di Banach. Le palle chiuse di E sono compatte se e solo se E ha dimensione finita.*

Adesso il risultato della sezione.

Teorema 4.2.2 (Moltiplicatori di Lagrange): *Siano E spazio di Banach di dimensione finita, $\Omega \subset E$ un aperto ed $f_0, f_1, ..., f_h \in C^1(\Omega, \mathbb{R})$. Consideriamo*

$$\Sigma = \{x \in \Omega \mid f_1(x) = ... = f_h(x) = 0\}.$$

Se $\overline{x} \in \Sigma$ minimizza localmente $f_{0|\Sigma}$, cioè $\exists \overline{B} = \overline{B}(\overline{x}, \rho) \subset \Omega$ t.c. $\forall x \in \overline{B} \cap \Sigma$ vale $f_0(x) \geq f_0(\overline{x})$. Allora i gradienti $\nabla f_0(\overline{x}), \nabla f_1(\overline{x}), ..., \nabla f_h(\overline{x})$ sono linearmente dipendenti, ossia $\exists \lambda_0, \lambda_1, ..., \lambda_h \in \mathbb{R}$ non tutti nulli t.c.

$$\sum_{i=0}^{h} \nabla f_i(\overline{x}) = 0.$$

Dimostrazione. Per $k \in \mathbb{N}$ consideriamo la funzione

$$g_k(x) = \|x - \overline{x}\|^2 + f_0(x) + k \sum_{i=1}^{h} f_i(x)^2$$

e notiamo che essendo E di dimensione finita la palla $\overline{B}$ è compatta, dunque per continuità g_k vi ammette un minimo x_k e questo per ogni $k \in \mathbb{N}$.

Passo 1: Dico che $(g_k(x_k))_{n\in\mathbb{N}}$ è una successione limitata in $\mathbb{R}$.

Da una parte si ha

$$g_k(x_k) \leq g_k(\overline{x}) = f_0(\overline{x})$$

dall'altra invece

$$g_k(x_k) \geq \|x_k - \overline{x}\|^2 + f_0(x_k) \geq f_0(x_k) \geq \min_{x\in\overline{B}} f_0(x)$$

da cui quanto voluto.

NON CONCLUSA, PRIMA O POI ARRIVERÀ.

$\square$

Appendice A

Spazi Vettoriali e Normati Complessi

A.1 $\mathbb{C}$-Spazi Vettoriali

Se E è un $\mathbb{C}$-spazio vettoriale, questo è anche un $\mathbb{R}$-spazio vettoriale, per restrizione del campo degli scalari e

$$J : E \to E \quad \text{t.c.} \quad x \mapsto ix$$

è un endomorfismo di E t.c. $J^2 = -I$.

Viceversa un $\mathbb{R}$-spazio vettoriale E con un endomorfismo $J \in \mathcal{L}(E, E)$ t.c. $J^2 = -I$ ha una struttura di $\mathbb{C}$-spazio vettoriale in cui, se $x \in E$

$$(a + ib)x := ax + bJx$$

dunque possiamo dare la seguente caratterizzazione

Proposizione A.1.1: *Uno spazio vettoriale E è un $\mathbb{C}$-spazio vettoriale se e solo se è un $\mathbb{R}$-spazio vettoriale munito di un endomorfismo $J \in \mathcal{L}(E, E)$ t.c. $J^2 = -I$.*

A.2 Complessificazione

Se E è un $\mathbb{R}$-spazio vettoriale, si può considerare

$$E_{\mathbb{C}} := (E \times E, J)$$

con $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ il quale è un endomorfismo t.c. $J^2 = -I_{E\times E}$ che rende $E \times E$ un $\mathbb{C}$-spazio vettoriale in cui per $a + ib \in \mathbb{C}$ e $(x, y) \in E_{\mathbb{C}}$

$$(a + ib)(x, y) := a(x, y) + bJ(x, y) = (ax - by, ay + bx)$$

$E_{\mathbb{C}}$ è detto **complessificato di** E.

Osservazione A.2.1: $\mathbb{C}$, come $\mathbb{C}$- spazio vettoriale, è identificabile con $\mathbb{R}_{\mathbb{C}}$ infatti qui $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ e se $(a,b),(x,y) \in \mathbb{R}^2$

$$(a+ib)(x+iy) = (ax-by) + i(ay+bx)$$

e, come definito sopra

$$(a+ib)(x,y) = a(x,y) + bJ(x,y) = (ax-by, ay+bx)$$

A.3 Applicazioni $\mathbb{C}$-Lineari

Siano E, F $\mathbb{C}$-spazi vettoriali e $L : E \to F$ un'applicazione $\mathbb{R}$-lineare, chiaramente si ha

Proposizione A.3.1: *L è $\mathbb{C}$-lineare se e solo se $iL = Li$, cioè, considerando E ed F come $\mathbb{R}$-spazi vettoriali con i rispettivi endomorfismi J_E, J_F t.c.per $x \in E$ e $y \in F$*

$$J_E x = ix \quad J_F y = iy$$

si ha che $J_F L = L J_E$.

Ora, dunque, considerando il caso di $E_{\mathbb{C}}$ con E $\mathbb{R}$-spazio vettoriale, un'applicazione $\mathbb{R}$-lineare

$$L = \begin{pmatrix} A & C \\ B & D \end{pmatrix} : E \times E \to E \times E$$

con A, B, C, D endomorfismi di E $\mathbb{R}$-lineari (in cui, come in precedenza per J, si è usata la notazione matriciale per comodità, sperando di non confondere il lettore) è $\mathbb{C}$-lineare se e solo se commuta con $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, condizione che equivale a

$$\begin{cases} A = D \\ B = -C \end{cases}$$

Osservazione A.3.2: Prendiamo $\mathbb{R}_{\mathbb{C}}$ e sia $L = \begin{pmatrix} a & c \\ b & d \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2$, per quanto appena detto L è $\mathbb{C}$-lineare se e solo se $\begin{cases} a = d \\ b = -c \end{cases}$ ossia se e solo se è **conforme speciale** nel qual caso è la moltiplicazione per $a + ib$, infatti si è definito il prodotto per $a + ib$ come

$$(a+ib)(x,y) = a(x,y) + bJ(x,y) = (ax-by, ay+bx) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Dunque riassumendo in $\mathbb{C}(= \mathbb{R}_{\mathbb{C}})$ un'applicazione $\mathbb{R}-$lineare è $\mathbb{C}-$lineare se e solo se è la moltiplicazione per un numero complesso.

A.4 Spazi Normati Complessi

Una **norma complessa** (o solamente **norma**) su un $\mathbb{C}$-spazio vettoriale E è una norma reale (su E come $\mathbb{R}-$spazio vettoriale)

$$\|\cdot\| : E \to [0, +\infty)$$

t.c. $\forall \lambda \in \mathbb{C}\ \forall x \in E\ \|\lambda x\| = |\lambda|\|x\|$

Osservazione A.4.1: Sostanzialmente una norma complessa è una norma che deve fa quello che deve fare una norma anche con i numeri complessi.

Chiaramente una norma complessa induce una metrica sul $\mathbb{C}$-spazio vettoriale esattamente come nel caso reale.

Uno spazio di Banach su $\mathbb{C}$ è proprio un $\mathbb{C}-$spazio vettoriale normato con una norma complessa completo.

Chiamo $\mathcal{L}_{\mathbb{C}}(E, F)$ lo spazio delle applicazioni $\mathbb{C}$-lineari continue tra due spazi di Banach su $\mathbb{C}$ (anche solo normati in realtà) E ed F.