Calcolo differenziale in spazi normati

Appunti delle lezioni tenute dal Prof. Alessandro Fonda

Università di Trieste, a.a. 2014/2015

1 Il differenziale secondo Frechét

Siano X e Y due spazi vettoriali reali normati, Ω un sottoinsieme aperto di X e $f : \Omega \to Y$ una funzione.

Definizione 1 *La funzione f è* **differenziabile** *(secondo Frechét) in un punto $x_0 \in \Omega$ se esiste un'applicazione lineare limitata $L \in \mathcal{L}(X, Y)$ per cui si possa scrivere*

$$f(x) = f(x_0) + L(x - x_0) + r(x)\,,$$

con

$$\lim_{x \to x_0} \frac{r(x)}{\|x - x_0\|} = 0\,.$$

L'applicazione lineare L è il **differenziale** *di f in x_0, e si denoterà con $df(x_0)$, oppure con $f'(x_0)$.*

Il teorema seguente ci fornisce un modo per determinare il differenziale, facendo uso delle **derivate direzionali**, e ne garantisce allo stesso tempo l'unicità. Se $h \in X$, la derivata di f lungo h in x_0 è definita da

$$\partial_h f(x_0) = \lim_{\tau \to 0} \frac{f(x_0 + \tau h) - f(x_0)}{\tau}\,.$$

Teorema 1 *Se f è differenziabile in x_0, allora, per ogni $h \in X$, si ha*

$$df(x_0)(h) = \partial_h f(x_0)\,.$$

<u>Dimostrazione</u>. La formula è sicuramente vera se $h = 0$. Supponiamo ora $h \neq 0$. Usando la linearità di $L = df(x_0)$, si ha

$$\begin{aligned}
\lim_{\tau \to 0} \frac{f(x_0 + \tau h) - f(x_0)}{\tau} &= \lim_{\tau \to 0} \frac{L(\tau h) + r(x_0 + \tau h)}{\tau} \\
&= \lim_{\tau \to 0} \frac{\tau L(h) + r(x_0 + \tau h)}{\tau} \\
&= L(h) + \lim_{\tau \to 0} \frac{r(x_0 + \tau h)}{\tau}\,.
\end{aligned}$$

D'altra parte, si ha che

$$\lim_{\tau\to 0}\frac{r(x_0+\tau h)}{|\tau|}=\lim_{\tau\to 0}\frac{r(x_0+\tau h)}{\|\tau h\|}\|h\|=0\,,$$

da cui

$$\lim_{\tau\to 0}\frac{r(x_0+\tau h)}{\tau}=0\,,$$

e quindi

$$\lim_{\tau\to 0}\frac{f(x_0+\tau h)-f(x_0)}{\tau}=L(h)\,,$$

il che dimostra il risultato. ■

Segnaliamo che talvolta si usa scrivere $df(x_0)h$ invece di $df(x_0)(h)$. Vediamo ora che la differenziabilità di una funzione ne implica la continuità.

Teorema 2 *Se f è differenziabile in x_0, allora f è continua in x_0.*

Dimostrazione. Sappiamo che l'applicazione $L=df(x_0)$ è lineare e continua, e $L(0)=0$. Ne segue che

$$\begin{aligned}\lim_{x\to x_0}f(x)&=\lim_{x\to x_0}[f(x_0)+L(x-x_0)+r(x)]\\&=f(x_0)+L(0)+\lim_{x\to x_0}r(x)\\&=f(x_0)+\lim_{x\to x_0}\frac{r(x)}{\|x-x_0\|}\lim_{x\to x_0}\|x-x_0\|\\&=f(x_0)\,,\end{aligned}$$

cioè f è continua in x_0. ■

Diremo che $f:\Omega\to Y$ è **differenziabile** se lo è in ogni punto x_0 di Ω. In tal caso, resta definita la funzione $f':\Omega\to\mathcal{L}(X,Y)$, che a ogni $x\in\Omega$ associa $df(x)$: ricordiamo che, per ogni $x\in\Omega$, si ha

$$f'(x)(h)=\partial_h f(x)=\lim_{\tau\to 0}\frac{f(x+\tau h)-f(x)}{\tau}\,.$$

Se tale funzione f' è continua, si dice che f è di classe C^1.

Consideriamo ora il caso particolare in cui $Y=\mathbb{R}$. In tal caso, si usa dire che la funzione $f:\Omega\to\mathbb{R}$ è un **funzionale**. Risulta utile il seguente

Teorema 3 *Se $f:\Omega\to\mathbb{R}$ è differenziabile in un punto x_0 di minimo (o di massimo) locale, allora $df(x_0)=0$.*

Dimostrazione. Se $h \in X$, consideriamo la funzione $\phi_h :]-\delta, \delta[\to \mathbb{R}$, per un certo $\delta > 0$, definita da

$$\phi_h(t) = f(x_0 + th) .$$

Vediamo che ϕ_h è derivabile in 0:

$$\phi_h'(0) = \lim_{t \to 0} \frac{\phi_h(t) - \phi_h(0)}{t} = \lim_{t \to 0} \frac{f(x_0 + th) - f(x_0)}{t} = df(x_0)(h) ,$$

Siccome 0 è punto di minimo (o di massimo) locale per ϕ_h, deve essere $\phi_h'(0) = 0$. Quindi, $df(x_0)(h) = 0$, per ogni $h \in X$. ■

Osserviamo infine che se X è uno spazio di Hilbert reale, con prodotto scalare $\langle \cdot, \cdot \rangle$, e se $f : \Omega \to \mathbb{R}$ è differenziabile in x_0, il Teorema di Riesz permette di definire un vettore $\nabla f(x_0)$, detto **gradiente** di f in x_0, tale che

$$df(x_0)(h) = \langle \nabla f(x_0), h \rangle ,$$

per ogni $h \in X$.

2 Regole di calcolo

Cominciamo con alcune proposizioni di facile verifica.

1. Se $f : \Omega \to Y$ è costante, allora $df(x_0) = 0$, per ogni $x_0 \in \Omega$.
2. Se $\mathcal{A} : \Omega \to Y$ è lineare e limitata, allora $d\mathcal{A}(x_0) = \mathcal{A}$, per ogni $x_0 \in \Omega$.
3. Se $X = X_1 \times X_2$ [1] e $\mathcal{B} : \Omega \to Y$ è bilineare e limitata, scrivendo $x_0 = (x_1^0, x_2^0)$ e $h = (h_1, h_2)$, con $x_1^0, h_1 \in X_1$ e $x_2^0, h_2 \in X_2$, si ha:

$$d\mathcal{B}(x_0)(h) = \mathcal{B}(x_1^0, h_2) + \mathcal{B}(h_1, x_2^0) .$$

Tutto ciò si può chiaramente generalizzare per le funzioni n-lineari limitate.

Valgono le usuali regole di calcolo, che brevemente richiamiamo.

Teorema 4 *Se $f, g : \Omega \to Y$ sono differenziabili in x_0 e α, β sono due numeri reali, allora*

$$d(\alpha f + \beta g)(x_0) = \alpha df(x_0) + \beta dg(x_0) .$$

[1] Consideriamo su $X_1 \times X_2$ la norma

$$\|(x_1, x_2)\| = \sqrt{\|x_1\|^2 + \|x_2\|^2} .$$

Dimostrazione. Scrivendo

$$f(x) = f(x_0) + df(x_0)(x - x_0) + r_1(x)\,,\ g(x) = g(x_0) + dg(x_0)(x - x_0) + r_2(x)\,,$$

abbiamo che

$$(\alpha f + \beta g)(x) = (\alpha f + \beta g)(x_0) + (\alpha df(x_0) + \beta dg(x_0))(x - x_0) + r(x)\,,$$

con $r(x) = \alpha r_1(x) + \beta r_2(x)$, e

$$\lim_{x \to x_0} \frac{r(x)}{\|x - x_0\|} = \alpha \lim_{x \to x_0} \frac{r_1(x)}{\|x - x_0\|} + \beta \lim_{x \to x_0} \frac{r_2(x)}{\|x - x_0\|} = 0\,.$$

Ne segue che la funzione $\alpha f + \beta g$ è differenziabile in x_0 con differenziale $\alpha df(x_0) + \beta dg(x_0)$. ■

Teorema 5 *Se $f : \Omega \to Y$ è differenziabile in x_0, U è un insieme aperto di Y contenente $f(x_0)$ e $g : U \to Z$ è differenziabile in $f(x_0)$, allora*

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0)\,.$$

Dimostrazione. Ponendo $y_0 = f(x_0)$, si ha

$$f(x) = f(x_0) + df(x_0)(x - x_0) + r_1(x)\,,\ g(y) = g(y_0) + dg(y_0)(y - y_0) + r_2(y)\,,$$

con

$$\lim_{x \to x_0} \frac{r_1(x)}{\|x - x_0\|} = 0\,, \quad \lim_{y \to y_0} \frac{r_2(y)}{\|y - y_0\|} = 0\,.$$

Introduciamo la funzione $R_2 : U \to Z$ così definita:

$$R_2(y) = \begin{cases} \dfrac{r_2(y)}{\|y - y_0\|} & \text{se } y \neq y_0\,, \\[2ex] 0 & \text{se } y = y_0\,. \end{cases}$$

Si noti che R_2 è continua in y_0. Allora

$$\begin{aligned} g(f(x)) &= g(f(x_0)) + dg(f(x_0))[f(x) - f(x_0)] + r_2(f(x)) \\ &= g(f(x_0)) + dg(f(x_0))[df(x_0)(x - x_0) + r_1(x)] + r_2(f(x)) \\ &= g(f(x_0)) + [dg(f(x_0)) \circ df(x_0)](x - x_0) + r_3(x)\,, \end{aligned}$$

dove

$$\begin{aligned} r_3(x) &= dg(f(x_0))(r_1(x)) + r_2(f(x)) \\ &= dg(f(x_0))(r_1(x)) + \|f(x) - f(x_0)\| R_2(f(x)) \\ &= dg(f(x_0))(r_1(x)) + \|df(x_0)(x - x_0) + r_1(x)\| R_2(f(x))\,. \end{aligned}$$

Quindi,

$$\frac{\|r_3(x)\|}{\|x-x_0\|} \leq \left\| dg(f(x_0)) \left(\frac{r_1(x)}{\|x-x_0\|} \right) \right\| + \\ + \left(\left\| df(x_0) \left(\frac{x-x_0}{\|x-x_0\|} \right) \right\| + \frac{\|r_1(x)\|}{\|x-x_0\|} \right) \|R_2(f(x))\| \,.$$

Se $x \to x_0$, il primo addendo tende a 0, poiché $dg(f(x_0))$ è continua; f è continua in x_0 e R_2 è continua in $y_0 = f(x_0)$ con $R_2(y_0) = 0$, per cui $\|R_2(f(x))\|$ tende a 0; $df(x_0)$, è limitata sulla palla $\overline{B}(0,1)$. Quindi, possiamo concludere che

$$\lim_{x \to x_0} \frac{\|r_3(x)\|}{\|x-x_0\|} = 0\,.$$

Ne segue che $g \circ f$ è differenziabile in x_0 con differenziale $dg(f(x_0)) \circ df(x_0)$. ■

Consideriamo ora il caso in cui sia $Y = Y_1 \times Y_2$. In tal caso, la funzione $f : \Omega \to Y_1 \times Y_2$ ha due componenti $f_1 : \Omega \to Y_1$ e $f_2 : \Omega \to Y_2$, per cui possiamo scrivere $f(x) = (f_1(x), f_2(x))$, per ogni $x \in \Omega$.

Teorema 6 *La funzione $f : \Omega \to Y_1 \times Y_2$ è differenziabile in x_0 se e solo se lo sono le sue componenti, e in tal caso si ha*

$$df(x_0)(h) = (df_1(x_0)(h), df_2(x_0)(h))\,,$$

per ogni $h \in X$.

Dimostrazione. Considerando le componenti nell'equazione

$$f(x) = f(x_0) + L(x-x_0) + r(x)\,,$$

possiamo scrivere

$$f_j(x) = f_j(x_0) + L_j(x-x_0) + r_j(x)\,,$$

con $j = 1, 2$, e sappiamo che

$$\lim_{x \to x_0} \frac{r(x)}{\|x-x_0\|} = 0 \iff \lim_{x \to x_0} \frac{r_j(x)}{\|x-x_0\|} = 0 \quad \text{per } j = 1, 2\,,$$

da cui la tesi. ■

Corollario 1 *Se $f_1 : \Omega \to Y_1$, $f_2 : \Omega \to Y_2$ sono differenziabili in x_0 e $\mathcal{B} : Y_1 \times Y_2 \to Y$ è bilineare e limitata, allora*

$$d[\mathcal{B}(f_1, f_2)](x_0)(h) = \mathcal{B}(df_1(x_0)(h), f_2(x_0)) + \mathcal{B}(f_1(x_0), df_2(x_0)(h)).$$

Dimostrazione. Abbiamo che $\mathcal{B}(f_1, f_2) = \mathcal{B} \circ f$, con $f : \Omega \to Y_1 \times Y_2$ definita da $f(x) = (f_1(x), f_2(x))$, per ogni $x \in \Omega$. La conclusione segue quindi dalla formula del differenziale di una funzione composta, tenendo conto della formula del differenziale di una funzione bilineare limitata. ■

Quanto visto sopra si può chiaramente generalizzare al prodotto di un numero finito qualunque di spazi.

3 Un sostituto al Teorema di Lagrange

Nel caso particolare in cui $X = \mathbb{R}$, si ha che

$$df(x_0)(h) = h\,df(x_0)(1) = h \lim_{\tau \to 0} \frac{f(x_0 + \tau) - f(x_0)}{\tau}\,.$$

In tal caso, si è soliti scrivere

$$f'(x_0) = \lim_{\tau \to 0} \frac{f(x_0 + \tau) - f(x_0)}{\tau}\,,$$

con leggero abuso di notazioni.

Avremo bisogno del seguente

Lemma 1 *Siano $a < b$ due numeri reali e $\varphi : [a, b] \to Y$ una funzione differenziabile: per ogni $t \in [a, b]$, esiste il limite*

$$\varphi'(t) = \lim_{\tau \to 0} \frac{\varphi(t + \tau) - \varphi(t)}{\tau} \in Y\,.$$

Se esiste una costante $C \geq 0$ tale che $\|\varphi'(t)\| \leq C$, per ogni $t \in [a, b]$, allora

$$\|\varphi(b) - \varphi(a)\| \leq C(b - a)\,.$$

Dimostrazione. Poniamo $I_0 = [a, b]$. Supponiamo per assurdo che sia

$$\|\varphi(b) - \varphi(a)\| - C(b - a) = M > 0\,.$$

Dividiamo l'intervallo $[a, b]$ in due parti uguali, prendendo il punto medio $m = \frac{a+b}{2}$. Allora deve valere una delle due disuguglianze

$$\|\varphi(m) - \varphi(a)\| - C(m - a) \geq \frac{M}{2}\,, \quad \|\varphi(b) - \varphi(m)\| - C(b - m) \geq \frac{M}{2}\,.$$

Se vale la prima, poniamo $I_1 = [a, m]$; se vale la seconda, $I_1 = [m, b]$. Procediamo ora allo stesso modo per definire I_2, poi I_3, ecc. Otteniamo così una successione di intervalli compatti $I_n = [a_n, b_n]$, con $I_0 \supseteq I_1 \supseteq I_2 \supseteq I_3 \supseteq \ldots$, di lunghezza

$$b_n - a_n = \frac{b - a}{2^n}\,,$$

per cui si ha

$$\|\varphi(b_n) - \varphi(a_n)\| - C(b_n - a_n) \geq \frac{M}{2^n}\,.$$

Pertanto, esiste un $c \in \mathbb{R}$ per cui $a_n \leq c \leq b_n$, per ogni n, e si ha che $\lim_n a_n = \lim_n b_n = c$. Essendo φ differenziabile in c, possiamo scrivere

$$\varphi(t) = \varphi(c) + \varphi'(c)(t - c) + r(t)\,,$$

con

$$\lim_{t\to c} \frac{r(t)}{t-c} = 0\,.$$

Sia $\varepsilon \in]0, \frac{M}{b-a}[$. Per n sufficientemente grande, si ha:

$$\begin{aligned}
M &\leq 2^n(\|\varphi(b_n)-\varphi(a_n)\| - C(b_n-a_n)) \\
&\leq 2^n(\|\varphi(b_n)-\varphi(c)\| + \|\varphi(c)-\varphi(a_n)\| - C(b_n-a_n)) \\
&= 2^n(\|\varphi'(c)(b_n-c)+r(b_n)\| + \|\varphi'(c)(a_n-c)+r(a_n)\| - C(b_n-a_n)) \\
&\leq 2^n(\|\varphi'(c)\|\,|b_n-c| + \|r(b_n)\| + \|\varphi'(c)\|\,|a_n-c| + \|r(a_n)\| - C(b_n-a_n)) \\
&\leq 2^n(C(|b_n-c|+|a_n-c|) + \|r(b_n)\| + \|r(a_n)\| - C(b_n-a_n)) \\
&= 2^n(\|r(b_n)\| + \|r(a_n)\|) \\
&\leq 2^n(\varepsilon|b_n-c| + \varepsilon|a_n-c|) \\
&= 2^n\varepsilon(b_n-a_n) \\
&= \varepsilon(b-a)\,,
\end{aligned}$$

una contraddizione. ■

Il lemma precedente si può generalizzare nel modo seguente, con dimostrazione analoga.

Lemma 2 *Siano $a < b$ due numeri reali e $\varphi : [a,b] \to Y$, $g : [a,b] \to \mathbb{R}$ due funzioni differenziabili. Se $\|\varphi'(t)\| \leq g'(t)$, per ogni $t \in [a,b]$, allora*

$$\|\varphi(b)-\varphi(a)\| \leq g(b)-g(a)\,.$$

Supponiamo che Ω sia un aperto di X e che $f : \Omega \to Y$ sia una funzione differenziabile. Enunciamo il cosiddetto **Teorema del valor medio**.

Teorema 7 *Se $[x_0, x]$ è un segmento contenuto in Ω, allora*

$$\|f(x)-f(x_0)\| \leq \sup\Big\{\|df(v)\| : v \in [x_0,x]\Big\}\|x-x_0\|\,.$$

Dimostrazione. Se l'estremo superiore è uguale a $+\infty$, non c'è nulla da dimostrare. Supponiamo quindi che sia

$$\sup\Big\{\|df(v)\| : v \in [x_0,x]\Big\} = C \in \mathbb{R}\,.$$

Consideriamo la funzione $\varphi : [0,1] \to Y$, definita da $\varphi(t) = f((1-t)x_0 + tx)$. Si vede che

$$\begin{aligned}
\|\varphi'(t)\| &= \|df((1-t)x_0+tx)(x-x_0)\| \\
&\leq \|df((1-t)x_0+tx)\|\,\|x-x_0\| \\
&\leq C\|x-x_0\|\,,
\end{aligned}$$

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

$$\|f(x)-f(x_0)\| = \|\varphi(1)-\varphi(0)\| \leq C\|x-x_0\|\,,$$

che è quanto volevasi dimostrare. ■

4 Il differenziale secondo Gateaux

Può risultare utile introdurre la seguente nozione più debole di differenziablità.

Definizione 2 *La funzione f è* **$\mathcal{G}$-differenziabile** *(o differenziabile secondo Gateaux) in un punto $x_0 \in \Omega$ se, per ogni $h \in X$, esiste la derivata di f lungo h in x_0,*

$$\partial_h f(x_0) = \lim_{\tau \to 0} \frac{f(x_0 + \tau h) - f(x_0)}{\tau},$$

ed esiste un'applicazione lineare limitata $\tilde{L} \in \mathcal{L}(X, Y)$ per cui si possa scrivere

$$\tilde{L}(h) = \partial_h f(x_0).$$

L'applicazione lineare $\tilde{L}$ verrà in seguito indicata con il simbolo $d_{\mathcal{G}} f(x_0)$.

È evidente che, se f è differenziabile in x_0, allora $d_{\mathcal{G}} f(x_0) = df(x_0)$. Tuttavia, esistono delle funzioni che, pur essendo $\mathcal{G}$-differenziabili in un punto x_0, non sono ivi continue. Ad esempio, la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita da

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

ha tutte le derivate direzionali nulle in $x_0 = (0, 0)$, per cui $d_{\mathcal{G}} f(0, 0) = 0$, ma essa non è continua in tale punto, come si vede considerando la restrizione alla parabola $\{(x, y) \in \mathbb{R}^2 : y = x^2\}$.

Diremo che $f : \Omega \to Y$ è **$\mathcal{G}$-differenziabile** se lo è in ogni punto x_0 di Ω. Nel caso in cui $X = \mathbb{R}$, una funzione $\mathcal{G}$-differenziabile è automaticamente differenziabile.

Teorema 8 *Supponiamo che $f : \Omega \to Y$ sia $\mathcal{G}$-differenziabile. Se $d_{\mathcal{G}} f : \Omega \to \mathcal{L}(X, Y)$ è continua in x_0, allora f è differenziabile in x_0 e $df(x_0) = d_{\mathcal{G}} f(x_0)$.*

<u>Dimostrazione</u>. Fissato $\varepsilon > 0$, siccome Ω è aperto e $d_{\mathcal{G}} f$ è continua in x_0, esiste un $\delta > 0$ tale che

$$\|x - x_0\| < \delta \quad \Rightarrow \quad x \in \Omega \quad \text{e} \quad \|d_{\mathcal{G}} f(x) - d_{\mathcal{G}} f(x_0)\| < \varepsilon.$$

Scegliamo un x tale che $\|x - x_0\| < \delta$ e consideriamo la funzione $\phi : [0, 1] \to Y$ definita da

$$\phi(t) = f(x_0 + t(x - x_0)).$$

Vediamo che ϕ è differenziabile:

$$\begin{aligned}
\phi'(t) &= \lim_{\tau\to 0} \frac{\phi(t+\tau)-\phi(t)}{\tau} \\
&= \lim_{\tau\to 0} \frac{f(x_0+t(x-x_0)+\tau(x-x_0)) - f(x_0+t(x-x_0))}{\tau} \\
&= d_{\mathcal{G}} f(x_0+t(x-x_0))(x-x_0)\,.
\end{aligned}$$

Sia $\psi : [0,1] \to Y$ definita da

$$\psi(t) = d_{\mathcal{G}} f(x_0)(x_0+t(x-x_0))\,.$$

Si vede subito che ψ è differenziabile, con

$$\psi'(t) = d_{\mathcal{G}} f(x_0)(x-x_0)\,.$$

Infine, sia $\varphi : [0,1] \to Y$ definita da

$$\varphi(t) = \phi(t) - \psi(t)\,,$$

per cui

$$\begin{aligned}
\varphi'(t) &= \phi'(t) - \psi'(t) \\
&= d_{\mathcal{G}} f(x_0+t(x-x_0))(x-x_0) - d_{\mathcal{G}} f(x_0)(x-x_0) \\
&= [d_{\mathcal{G}} f(x_0+t(x-x_0)) - d_{\mathcal{G}} f(x_0)](x-x_0)\,.
\end{aligned}$$

Quindi, essendo $\|x-x_0\| < \delta$, abbiamo che $\|\varphi'(t)\| \leq \varepsilon\|x-x_0\|$ per ogni $t \in [0,1]$. Usando il Lemma 1, abbiamo che $\|\varphi(1)-\varphi(0)\| \leq \varepsilon\|x-x_0\|(1-0)$, per cui

$$\|f(x)-f(x_0)-d_{\mathcal{G}} f(x_0)(x-x_0)\| = \|\varphi(1)-\varphi(0)\| \leq \varepsilon\|x-x_0\|\,.$$

Abbiamo così dimostrato che per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che

$$0 < \|x-x_0\| < \delta \;\Rightarrow\; \left\| \frac{f(x)-f(x_0)-d_{\mathcal{G}} f(x_0)(x-x_0)}{\|x-x_0\|} \right\| \leq \varepsilon\,,$$

ossia che

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)-d_{\mathcal{G}} f(x_0)(x-x_0)}{\|x-x_0\|} = 0\,,$$

per cui f è differenziabile in x_0 e $df(x_0) = d_{\mathcal{G}} f(x_0)$. ■

Ne segue immediatamente il seguente

Corollario 2 *Se $f : \Omega \to Y$ è $\mathcal{G}$-differenziabile e $d_{\mathcal{G}} f : \Omega \to \mathcal{L}(X,Y)$ è continua, allora f è di classe C^1.*

5 Differenziali parziali

Siano X, Y, Z tre spazi vettoriali reali normati. Supponiamo che Ω sia un sottoinsieme aperto di $X \times Y$ e che (x_0, y_0) sia un punto di Ω. Consideriamo una funzione $f : \Omega \to Z$, e denotiamo i suoi valori con $f(x, y)$, dove $x \in X$ e $y \in Y$ sono tali che $(x, y) \in \Omega$.

La funzione $f(\cdot, y_0)$ è definita in un intorno di x_0 e si può scrivere come funzione composta: $f(\cdot, y_0) = f \circ i_1$, con $i_1 : X \to X \times Y$ definita da $i_1(x) = (x, y_0)$. Se tale funzione $f(\cdot, y_0)$ è differenziabile in x_0, il suo differenziale $d[f(\cdot, y_0)](x_0)$, un'applicazione lineare limitata da X a Z, si chiama **differenziale parziale** rispetto a x di f in (x_0, y_0), e si denota con

$$\frac{\partial f}{\partial x}(x_0, y_0)\,.$$

Si noti che, se f è differenziabile in (x_0, y_0), allora, per il teorema sul differenziale di una funzione composta, si ha

$$d(f \circ i_1)(x_0) = df(i_1(x_0)) \circ di_1(x_0)\,,$$

quindi, per ogni $h \in X$, si ha

$$\frac{\partial f}{\partial x}(x_0, y_0)(h) = df(x_0, y_0)(h, 0)\,.$$

Un discorso del tutto analogo si può ora fare per definire il differenziale parziale rispetto a y di f in (x_0, y_0), applicazione lineare limitata da Y a Z, con notazione

$$\frac{\partial f}{\partial y}(x_0, y_0)\,.$$

Se f è differenziabile in (x_0, y_0), per ogni $k \in Y$ si ha

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

Ne segue la formula

$$df(x_0, y_0)(h, k) = \frac{\partial f}{\partial x}(x_0, y_0)(h) + \frac{\partial f}{\partial y}(x_0, y_0)(k)\,,$$

valida per ogni $(h, k) \in X \times Y$.

L'esistenza dei differenziali parziali non è sufficiente a garantire la differenziabilità di una funzione. Vale però il seguente

Teorema 9 *Se i differenziali parziali $\frac{\partial f}{\partial x}(x, y)$ e $\frac{\partial f}{\partial y}(x, y)$ esistono in tutti i punti (x, y) di un intorno U di (x_0, y_0) e le funzioni $\frac{\partial f}{\partial x} : U \to \mathcal{L}(X, Z)$, $\frac{\partial f}{\partial y} : U \to \mathcal{L}(Y, Z)$ sono continue in (x_0, y_0), allora f è differenziabile in (x_0, y_0).*

Dimostrazione. Sia $g : U \to Z$ la funzione definita da

$$g(x,y) = f(x,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)\,.$$

Vogliamo dimostrare che

$$\lim_{(x,y)\to(x_0,y_0)} \frac{g(x,y)}{\|(x,y)-(x_0,y_0)\|} = 0\,.$$

Fissiamo un $\varepsilon > 0$. Si vede subito che

$$\frac{\partial g}{\partial x}(x,y) = \frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial x}(x_0,y_0)\,, \quad \frac{\partial g}{\partial y}(x,y) = \frac{\partial f}{\partial y}(x,y) - \frac{\partial f}{\partial y}(x_0,y_0)\,.$$

Per la continuità dei differenziali parziali, esiste un $\delta > 0$ tale che, se $\|(x,y) - (x_0,y_0)\| \leq \delta$, allora

$$\left\|\frac{\partial g}{\partial x}(x,y)\right\| \leq \varepsilon\,, \quad \left\|\frac{\partial g}{\partial y}(x,y)\right\| \leq \varepsilon\,.$$

Allora, applicando il Teorema 7, essendo $g(x_0,y_0) = 0$, otteniamo

$$\begin{aligned}\|g(x,y)\| &\leq \|g(x,y) - g(x,y_0)\| + \|g(x,y_0) - g(x_0,y_0)\|\\ &\leq \varepsilon\|y-y_0\| + \varepsilon\|x-x_0\|\\ &\leq 2\varepsilon\|(x,y)-(x_0,y_0)\|\,,\end{aligned}$$

da cui la tesi. ■

Quanto visto sopra si può chiaramente generalizzare al prodotto di un numero finito qualunque di spazi.

Consideriamo ora una situazione particolare, in cui risulta utile saper calcolare i differenziali parziali. Come sopra, $\Omega \subseteq X \times Y$ sarà un insieme aperto. Siano $W_1, \ldots, W_n$ degli spazi vettoriali reali normati, e siano $g_j : \Omega \to W_j$ delle funzioni differenziabili nel punto (x_0,y_0), con $j = 1,\ldots,n$. Indichiamo con $(w_1,\ldots,w_n)$ gli elementi di $W_1 \times \ldots \times W_n$. Sia $V \subseteq W_1 \times \ldots \times W_n$ un intorno aperto del punto $v_0 = (g_1(x_0,y_0),\ldots,g_n(x_0,y_0))$, e sia $f : V \to Z$ differenziabile in v_0. Consideriamo la funzione

$$F(x,y) = f(g_1(x,y),\ldots,g_n(x,y))\,,$$

definita in un intorno di (x_0,y_0), a valori in Z. Essa è differenziabile in (x_0,y_0), in quanto composta di funzioni differenziabili.

Teorema 10 *Valgono le seguenti formule:*

$$\frac{\partial F}{\partial x}(x_0,y_0) = \frac{\partial f}{\partial w_1}(v_0) \circ \frac{\partial g_1}{\partial x}(x_0,y_0) + \ldots + \frac{\partial f}{\partial w_n}(v_0) \circ \frac{\partial g_n}{\partial x}(x_0,y_0)\,,$$

$$\frac{\partial F}{\partial y}(x_0,y_0) = \frac{\partial f}{\partial w_1}(v_0) \circ \frac{\partial g_1}{\partial y}(x_0,y_0) + \ldots + \frac{\partial f}{\partial w_n}(v_0) \circ \frac{\partial g_n}{\partial y}(x_0,y_0)\,.$$

Dimostrazione. Sia $g : \Omega \to W_1 \times \ldots \times W_n$ definita da

$$g(x,y) = (g_1(x,y), \ldots, g_n(x,y)) .$$

Abbiamo che $F = f \circ g$, per cui

$$dF(x_0,y_0)(h,k) = df(v_0)(dg(x_0,y_0)(h,k)) ,$$

per ogni $(h,k) \in X \times Y$. Quindi,

$$\begin{aligned}
&\frac{\partial F}{\partial x}(x_0,y_0)(h) + \frac{\partial F}{\partial y}(x_0,y_0)(k) = df(v_0)\left(\frac{\partial g}{\partial x}(x_0,y_0)(h) + \frac{\partial g}{\partial y}(x_0,y_0)(k)\right) \\
&= df(v_0)\left(\frac{\partial g_1}{\partial x}(x_0,y_0)(h) + \frac{\partial g_1}{\partial y}(x_0,y_0)(k), \ldots, \frac{\partial g_n}{\partial x}(x_0,y_0)(h) + \frac{\partial g_n}{\partial y}(x_0,y_0)(k)\right) \\
&= \frac{\partial f}{\partial w_1}(v_0)\left(\frac{\partial g_1}{\partial x}(x_0,y_0)(h) + \frac{\partial g_1}{\partial y}(x_0,y_0)(k)\right) + \ldots \\
&\quad \ldots + \frac{\partial f}{\partial w_n}(v_0)\left(\frac{\partial g_n}{\partial x}(x_0,y_0)(h) + \frac{\partial g_n}{\partial y}(x_0,y_0)(k)\right) .
\end{aligned}$$

La prima formula si ottiene prendendo $k = 0$, la seconda prendendo $h = 0$. ■

Similmente si può procedere se invece di due spazi X e Y si considera il prodotto di un numero finito qualsiasi di essi.

6 Il teorema della funzione implicita e del diffeomorfismo locale

Consideriamo qui tre spazi vettoriali reali normati X, Y, Z. Supporremo che Y e Z siano isomorfi, e siano entrambi completi, ossia spazi di Banach. Avremo bisogno del seguente risultato preliminare.

Lemma 3 *Sia $A \in \mathcal{L}(Y,Z)$ invertibile, con $A^{-1} \in \mathcal{L}(Z,Y)$. Se $B \in \mathcal{L}(Y,Z)$ è tale che*

$$\|B - A\| < \frac{1}{\|A^{-1}\|} ,$$

allora anche B è invertibile, con $B^{-1} \in \mathcal{L}(Z,Y)$. Inoltre, l'applicazione

$$A \mapsto A^{-1}$$

è continua.

Dimostrazione. Sia $C = (A - B)A^{-1}$. Allora $C \in \mathcal{L}(Z)$ è tale che

$$\|C\| \le \|A - B\| \, \|A^{-1}\| < 1 .$$

Vediamo che $I - C$ è invertibile, con inversa data dalla serie di Neumann

$$(I - C)^{-1} = \sum_{k=0}^{\infty} C^k \, .$$

Tale serie converge, in quanto converge totalmente, e $\mathcal{L}(Z)$ è uno spazio metrico completo. Inoltre, si ha

$$\begin{aligned}
(I - C)\bigg(\sum_{k=0}^{\infty} C^k \bigg) &= (I - C)\bigg(\lim_n \sum_{k=0}^{n} C^k \bigg) \\
&= \lim_n \bigg[(I - C)\bigg(\sum_{k=0}^{n} C^k \bigg) \bigg] \\
&= \lim_n (I - C^{n+1}) = I \, ,
\end{aligned}$$

e allo stesso modo

$$\begin{aligned}
\bigg(\sum_{k=0}^{\infty} C^k \bigg)(I - C) &= \bigg(\lim_n \sum_{k=0}^{n} C^k \bigg)(I - C) \\
&= \lim_n \bigg[\bigg(\sum_{k=0}^{n} C^k \bigg)(I - C) \bigg] \\
&= \lim_n (I - C^{n+1}) = I \, .
\end{aligned}$$

Abbiamo così dimostrato che $I - C$ è invertibile. Essendo $I - C = BA^{-1}$, si conclude facilmente.

Vediamo ora che l'applicazione che ad ogni A associa la sua inversa A^{-1} è continua. Infatti, fissata A invertibile, se $\|B - A\| < 1/\|A^{-1}\|$, abbiamo che anche B è invertibile, e

$$\begin{aligned}
\|B^{-1} - A^{-1}\| &= \|(I - A^{-1}(A - B))^{-1} A^{-1} - A^{-1}\| \\
&= \bigg\| \bigg(\sum_{k=0}^{\infty} (A^{-1}(A - B))^k \bigg) A^{-1} - A^{-1} \bigg\| \\
&= \bigg\| \bigg(\sum_{k=1}^{\infty} (A^{-1}(A - B))^k \bigg) A^{-1} \bigg\| \\
&\leq \sum_{k=1}^{\infty} \Big(\|A^{-1}\|^{k+1} \|A - B\|^k \Big) \\
&= \sum_{k=0}^{\infty} \Big(\|A^{-1}\| \, \|A - B\| \Big)^k \|A^{-1}\|^2 \|A - B\| \, .
\end{aligned}$$

Si vede dunque che l'applicazione considerata è Lipschitziana in ogni palla centrata in A di raggio minore di $1/\|A^{-1}\|$. ∎

Enunciamo ora il **Teorema della funzione implicita**.

Teorema 11 *Siano Ω un sottoinsieme di $X \times Y$ e $f : \Omega \to Z$ una funzione continua, tale che $\frac{\partial f}{\partial y} : \Omega \to \mathcal{L}(Y, Z)$ esista e sia continua. Supponiamo che Y e Z siano completi. Inoltre, sia $(x_0, y_0) \in \Omega$ un punto in cui $f(x_0, y_0) = 0$ e $\frac{\partial f}{\partial y}(x_0, y_0) : Y \to Z$ sia un isomorfismo. Allora esistono un intorno aperto U di x_0, un intorno aperto V di y_0, con $U \times V \subseteq \Omega$, e una funzione continua $\eta : U \to V$ il cui grafico coincide con l'insieme*

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

Se $x \in U$ e $y \in V$, si ha quindi

$$f(x, y) = 0 \iff y = \eta(x) .$$

Se inoltre f è di classe C^1, allora il dominio U può essere scelto in modo che anche η sia di classe C^1. In tal caso, si ha

$$d\eta(x) = -\left(\frac{\partial f}{\partial y}(x, \eta(x))\right)^{-1} \circ \frac{\partial f}{\partial x}(x, \eta(x)) ,$$

per ogni $x \in U$.

<u>Dimostrazione</u>. Sarà conveniente porre $\mathcal{A} = \dfrac{\partial f}{\partial y}(x_0, y_0)$ e

$$R(x, y) = f(x, y) - \mathcal{A}(y - y_0) , \quad \mathcal{F}(x, y) = y_0 - \mathcal{A}^{-1} R(x, y) .$$

Allora

$$f(x, y) = 0 \iff y = \mathcal{F}(x, y) .$$

Vogliamo dimostrare che esistono un intorno aperto U di x_0, un intorno aperto V di y_0, con $U \times \overline{V} \subseteq \Omega$ tali che, per ogni $x \in U$, si ha che $\mathcal{F}(x, \cdot)$ è una contrazione che manda $\overline{V}$ in sè stesso.

Poniamo

$$\epsilon = \frac{1}{2\|\mathcal{A}^{-1}\|} .$$

Per la continuità di $\frac{\partial f}{\partial y}$ in (x_0, y_0), esistono un intorno aperto W di x_0 e una palla aperta $V = B(y_0, \gamma)$, di raggio $\gamma > 0$, tali che $W \times \overline{V} \subseteq \Omega$ e, per ogni $(x, y) \in W \times \overline{V}$, si ha che

$$\left\| \frac{\partial f}{\partial y}(x, y) - \mathcal{A} \right\| \leq \epsilon .$$

Inoltre, essendo $f(x_0, y_0) = 0$, per la continuità di f in (x_0, y_0), esiste un intorno aperto U di x_0, che possiamo supporre essere contenuto in W, tale che, per ogni $x \in U$,

$$\|f(x, y_0)\| < \gamma\epsilon .$$

Dimostriamo ora che, per ogni $x \in U$ e ogni $y_1, y_2 \in \overline{V}$, si ha

$$\|R(x, y_2) - R(x, y_1)\| \leq \epsilon \|y_2 - y_1\| .$$

A tal scopo, fissati $x \in U$ e $y_1, y_2 \in \overline{V}$, consideriamo la funzione $g : [0, 1] \to Z$ definita da

$$g(t) = R(x, (1 - t)y_1 + ty_2) .$$

Allora

$$\begin{aligned}
\|g'(t)\| &= \left\| \frac{\partial R}{\partial y}(x, (1 - t)y_1 + ty_2)(y_2 - y_1) \right\| \\
&\leq \left\| \frac{\partial R}{\partial y}(x, (1 - t)y_1 + ty_2) \right\| \|y_2 - y_1\| \\
&= \left\| \frac{\partial f}{\partial y}(x, (1 - t)y_1 + ty_2) - \mathcal{A} \right\| \|y_2 - y_1\| \\
&\leq \epsilon \|y_2 - y_1\| .
\end{aligned}$$

Per il Lemma 1, abbiamo

$$\|R(x, y_2) - R(x, y_1)\| = \|g(1) - g(0)\| \leq \epsilon \|y_2 - y_1\| ,$$

che è quanto volevamo dimostrare. Ne segue che, per ogni $x \in U$ e ogni $y_1, y_2 \in \overline{V}$, si ha

$$\|\mathcal{F}(x, y_2) - \mathcal{F}(x, y_1)\| \leq \|\mathcal{A}^{-1}\| \, \|R(x, y_2) - R(x, y_1)\| \leq \frac{1}{2} \|y_2 - y_1\| .$$

Inoltre, per ogni $y \in \overline{V}$,

$$\begin{aligned}
\|\mathcal{F}(x, y) - y_0\| &\leq \|\mathcal{F}(x, y) - \mathcal{F}(x, y_0)\| + \|\mathcal{F}(x, y_0) - y_0\| \\
&\leq \frac{1}{2} \|y - y_0\| + \|\mathcal{A}^{-1} R(x, y_0)\| \\
&= \frac{1}{2} \|y - y_0\| + \|\mathcal{A}^{-1} f(x, y_0)\| \\
&\leq \frac{\gamma}{2} + \|\mathcal{A}^{-1}\| \, \|f(x, y_0)\| \\
&< \frac{\gamma}{2} + \frac{\gamma}{2} = \gamma .
\end{aligned}$$

Abbiamo così dimostrato che, per ogni $x \in U$, si ha che $\mathcal{F}(x, \cdot)$ è una contrazione che manda $\overline{V}$ in V. Per il teorema delle contrazioni, per ogni $x \in U$ esiste un unico $y \in \overline{V}$ tale che $\mathcal{F}(x, y) = y$. Tale y appartiene a V, e verrà indicato con $\eta(x)$. Resta così definita la funzione $\eta : U \to V$. Essa è tale che, per $(x, y) \in U \times V$,

$$y = \eta(x) \iff \mathcal{F}(x, y) = y \iff f(x, y) = 0 .$$

Vediamo ora che η è continua. Fissiamo un certo $\hat{x} \in U$ e verifichiamo che η è continua in $\hat{x}$. Per ogni $x \in U$,

$$\begin{aligned}\|\eta(x) - \eta(\hat{x})\| &= \|\mathcal{F}(x, \eta(x)) - \mathcal{F}(\hat{x}, \eta(\hat{x}))\| \\ &\leq \|\mathcal{F}(x, \eta(x)) - \mathcal{F}(x, \eta(\hat{x}))\| + \|\mathcal{F}(x, \eta(\hat{x})) - \mathcal{F}(\hat{x}, \eta(\hat{x}))\| \\ &\leq \frac{1}{2}\|\eta(x) - \eta(\hat{x})\| + \|\mathcal{F}(x, \eta(\hat{x})) - \mathcal{F}(\hat{x}, \eta(\hat{x}))\|\,.\end{aligned}$$

Quindi,

$$\begin{aligned}\frac{1}{2}\|\eta(x) - \eta(\hat{x})\| &\leq \|\mathcal{F}(x, \eta(\hat{x})) - \mathcal{F}(\hat{x}, \eta(\hat{x}))\| \\ &\leq \|\mathcal{A}^{-1}\|\,\|R(x, \eta(\hat{x})) - R(\hat{x}, \eta(\hat{x}))\| \\ &= \|\mathcal{A}^{-1}\|\,\|f(x, \eta(\hat{x})) - f(\hat{x}, \eta(\hat{x}))\|\,.\end{aligned}$$

Essendo f continua, ne segue che

$$\lim_{x \to \hat{x}} \|\eta(x) - \eta(\hat{x})\| = 0\,,$$

ossia che η è continua in $\hat{x}$.

Supponiamo ora che f sia di classe C^1. Notiamo che, essendo η continua con $\eta(x_0) = y_0$, restringendo eventualmente l'insieme U, per il precedente lemma possiamo supporre che, per ogni $x \in U$, l'applicazione lineare $\frac{\partial f}{\partial y}(x, \eta(x))$ sia invertibile e che

$$\left\| \left(\frac{\partial f}{\partial y}(x, \eta(x)) \right)^{-1} \right\| \leq 2 \left\| \left(\frac{\partial f}{\partial y}(x_0, y_0) \right)^{-1} \right\| := C\,.$$

Fissiamo un punto $\hat{x} \in U$, e scriviamo $\hat{y} = \eta(\hat{x})$. Poniamo inoltre

$$\Gamma = -\left(\frac{\partial f}{\partial y}(\hat{x}, \eta(\hat{x})) \right)^{-1} \circ \frac{\partial f}{\partial x}(\hat{x}, \eta(\hat{x}))\,.$$

Vogliamo dimostrare che η è differenziabile in $\hat{x}$: verificheremo che

$$\lim_{h \to 0} \frac{\eta(\hat{x} + h) - \eta(\hat{x}) - \Gamma(h)}{\|h\|} = 0\,.$$

Fissiamo $\varepsilon > 0$. Sia

$$k(h) = \eta(\hat{x} + h) - \eta(\hat{x})\,,$$

definita in un intorno di 0. Notiamo che, essendo η continua, se $h \to 0$, allora anche $k(h) \to 0$. Essendo f differenziabile in $(\hat{x}, \hat{y})$, esiste un $\delta > 0$ tale che, se $\|h\| \leq \delta$, allora

$$\begin{aligned}\left\| f(\hat{x} + h, \hat{y} + k(h)) - f(\hat{x}, \hat{y}) - \frac{\partial f}{\partial x}(\hat{x}, \hat{y})(h) - \frac{\partial f}{\partial y}(\hat{x}, \hat{y})(k(h)) \right\| &\leq \\ &\leq \varepsilon(\|h\| + \|k(h)\|)\,.\end{aligned}$$

Supponiamo $\|h\| \leq \delta$. Siccome

$$f(\hat{x}+h, \hat{y}+k(h)) = f(\hat{x}+h, \eta(\hat{x}+h)) = 0\,, \quad f(\hat{x},\hat{y}) = f(\hat{x},\eta(\hat{x})) = 0\,,$$

abbiamo quindi che

$$\left\| \frac{\partial f}{\partial x}(\hat{x},\hat{y})(h) + \frac{\partial f}{\partial y}(\hat{x},\hat{y})(k(h)) \right\| \leq \varepsilon(\|h\| + \|k(h)\|)\,.$$

Scrivendo

$$\begin{aligned} \eta(\hat{x}+h) - \eta(\hat{x}) - \Gamma(h) &= k(h) + \left(\frac{\partial f}{\partial y}(\hat{x},\hat{y})\right)^{-1} \left(\frac{\partial f}{\partial x}(\hat{x},\hat{y})(h)\right) \\ &= \left(\frac{\partial f}{\partial y}(\hat{x},\hat{y})\right)^{-1} \left(\frac{\partial f}{\partial x}(\hat{x},\hat{y})(h) + \frac{\partial f}{\partial y}(\hat{x},\hat{y})(k(h))\right)\,, \end{aligned}$$

abbiamo che

$$\begin{aligned} \|\eta(\hat{x}+h) - \eta(\hat{x}) - \Gamma(h)\| &\leq \left\| \left(\frac{\partial f}{\partial y}(\hat{x},\hat{y})\right)^{-1} \right\| \varepsilon(\|h\| + \|k(h)\|) \\ &\leq C\varepsilon(\|h\| + \|k(h)\|) \\ &= C\varepsilon(\|h\| + \|\eta(\hat{x}+h) - \eta(\hat{x})\|) \\ &\leq C\varepsilon(\|h\| + \|\eta(\hat{x}+h) - \eta(\hat{x}) - \Gamma(h)\| + \|\Gamma(h)\|)\,, \end{aligned}$$

da cui

$$(1 - C\varepsilon)\|\eta(\hat{x}+h) - \eta(\hat{x}) - \Gamma(h)\| \leq C\varepsilon(1 + \|\Gamma\|)\|h\|\,.$$

Potendo supporre che ε sia piccolo in modo tale che $C\varepsilon < \frac{1}{2}$, otteniamo quindi che

$$\frac{\|\eta(\hat{x}+h) - \eta(\hat{x}) - \Gamma(h)\|}{\|h\|} \leq 2C\varepsilon(1 + \|\Gamma\|)\,.$$

Questo ci permette di concludere che η è differenziabile in $\hat{x}$, con differenziale

$$d\eta(\hat{x}) = \Gamma = -\left(\frac{\partial f}{\partial y}(\hat{x},\eta(\hat{x}))\right)^{-1} \circ \frac{\partial f}{\partial x}(\hat{x},\eta(\hat{x}))\,.$$

Siccome $\eta : U \to Y$ è continua e l'applicazione $A \mapsto A^{-1}$ è anch'essa continua, dalla formula dimostrata segue che anche $d\eta : U \to \mathcal{L}(X,Y)$ è continua, e la dimostrazione è così completata. ■

Chiaramente, si può enunciare un teorema analogo, in cui il ruolo delle variabili x e y risulta scambiato.

Teorema 12 *Siano Ω un sottoinsieme di $X \times Y$ e $f : \Omega \to Z$ una funzione continua, tale che $\frac{\partial f}{\partial x} : \Omega \to \mathcal{L}(X,Z)$ esista e sia continua. Supponiamo che Y e Z siano completi. Inoltre, sia $(x_0,y_0) \in \Omega$ un punto in cui $f(x_0,y_0) = 0$*

e $\frac{\partial f}{\partial x}(x_0, y_0) : X \to Z$ sia un isomorfismo. Allora esistono un intorno aperto U di x_0, un intorno aperto V di y_0, con $U \times V \subseteq \Omega$, e una funzione continua $\eta : V \to U$ il cui grafico coincide con l'insieme

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

Se $x \in U$ e $y \in V$, si ha quindi

$$f(x, y) = 0 \iff x = \eta(y) .$$

Se inoltre f è di classe C^1, allora il dominio U può essere scelto in modo che anche η sia di classe C^1. In tal caso, si ha

$$d\eta(y) = -\left(\frac{\partial f}{\partial x}(y, \eta(y))\right)^{-1} \circ \frac{\partial f}{\partial y}(y, \eta(y)) ,$$

per ogni $y \in V$.

Dati due insiemi aperti A e B in due spazi normati X e Y, una funzione $g : A \to B$ si dice essere un **diffeomorfismo** se, oltre a essere di classe C^1, è biiettiva e la sua inversa $g^{-1} : B \to A$ è anch'essa di classe C^1.

Useremo ora il teorema precedente per dimostrare il seguente **Teorema del diffeomorfismo locale**. Supponiamo qui che X e Y siano entrambi spazi di Banach.

Teorema 13 *Sia Ω un sottoinsieme aperto di X e $g : \Omega \to Y$ una funzione di classe C^1. Sia x_0 un punto di Ω in cui il differenziale $dg(x_0)$ risulta essere un isomorfismo tra X e Y. Allora esistono un intorno aperto U di x_0 e un intorno aperto V di $g(x_0)$ tali che la restrizione di g ad U è un diffeomorfismo tra U e V.*

Dimostrazione. Consideriamo la funzione $f : \Omega \times Y \to Y$ definita da

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

Posto $y_0 = g(x_0)$, si ha che $f(x_0, y_0) = 0$, e $\frac{\partial f}{\partial x}(x_0, y_0) = dg(x_0)$ è un isomorfismo tra X e Y. Per il teorema della funzione implicita, esistono un intorno aperto U di x_0, con $U \subseteq \Omega$, un intorno aperto V di y_0, e una funzione $\eta : V \to U$, di classe C^1, con la seguente proprietà: se $x \in U$ e $y \in V$, si ha

$$f(x, y) = 0 \iff x = \eta(y) .$$

Pertanto, η è l'inversa della della funzione ottenuta restringendo il dominio di g ad U e il suo codominio a V. ∎

7 Differenziali di ordine superiore

Sia $f : \Omega \to Y$ una funzione differenziabile, e sia $f' : \Omega \to \mathcal{L}(X, Y)$ la funzione che a ogni $x \in \Omega$ associa $df(x)$. Se f' è a sua volta differenziabile in un punto x_0, diremo che f è due volte differenziabile in x_0. In tal caso, il differenziale $df'(x_0)$ si chiama **differenziale secondo** di f in x_0, e si denota con $d^2f(x_0)$, o con $f''(x_0)$.

Si ha quindi che $d^2f(x_0)$ è un elemento di $\mathcal{L}(X, \mathcal{L}(X, Y))$. Spesso, però, si preferisce considerare $d^2f(x_0)$ come un'applicazione bilineare limitata da $X \times X$ in Y, e si scrive $d^2f(x_0)(h, k)$ invece di $d^2f(x_0)(h)(k)$, con h, k in X.

Se f è due volte differenziabile in x_0, si ha

$$\begin{aligned} d^2f(x_0)(h,k) &= \lim_{\tau\to 0} \frac{df(x_0+\tau h) - df(x_0)}{\tau}(k) \\ &= \lim_{\tau\to 0} \frac{df(x_0+\tau h)(k) - df(x_0)(k)}{\tau} \\ &= \lim_{\tau\to 0} \frac{\partial_k f(x_0+\tau h) - \partial_k f(x_0)}{\tau} \\ &= \partial_h(\partial_k f)(x_0)\,, \end{aligned}$$

che si scrive anche $\partial_h \partial_k f(x_0)$. Vediamo ora la simmetria di $d^2f(x_0)$.

Teorema 14 *Se f è due volte differenziabile in x_0, l'applicazione bilineare $d^2f(x_0)$ è simmetrica: per ogni $h, k \in X$, si ha*

$$d^2f(x_0)(h,k) = d^2f(x_0)(k,h)\,.$$

<u>Dimostrazione</u>. Dimostreremo dapprima che

$$\lim_{(u,v)\to(0,0)} \frac{f(x_0+u+v) - f(x_0+u) - f(x_0+v) + f(x_0) - d^2f(x_0)(v,u)}{(\|u\|+\|v\|)^2} = 0\,.$$

Fissiamo un $\varepsilon > 0$. Siccome df è differenziabile in x_0, esiste un $\delta > 0$ tale che

$$\|h\| \le \delta \;\Rightarrow\; \|df(x_0+h) - df(x_0) - d^2f(x_0)(h)\| \le \varepsilon\|h\|\,.$$

Consideriamo la funzione g_v, definita in una palla $B(0, r)$, per $\|v\|$ sufficientemente piccola, in questo modo:

$$g_v(u) = f(x_0+u+v) - f(x_0+u) - f(x_0+v) + f(x_0) - d^2f(x_0)(v,u)\,.$$

Notiamo che $g_v(0) = 0$. Osseviamo che la funzione $\phi_v = d^2f(x_0)(v, \cdot)$ è lineare, per cui $d\phi_v(u) = \phi_v = d^2f(x_0)(v)$. Quindi,

$$\|dg_v(u)\| = \|df(x_0+u+v) - df(x_0+u) - d^2f(x_0)(v)\|$$

$$
\begin{aligned}
&= \|df(x_0+u+v) - df(x_0) - d^2f(x_0)(u+v)] - \\
&\quad -[df(x_0+u) - df(x_0) - d^2f(x_0)(u)]\| \\
&\le \|df(x_0+u+v) - df(x_0) - d^2f(x_0)(u+v)\| + \\
&\quad +\|df(x_0+u) - df(x_0) - d^2f(x_0)(u)]\| \\
&\le \varepsilon\|u+v\| + \varepsilon\|u\| \\
&\le 2\varepsilon(\|u\| + \|v\|)\,.
\end{aligned}
$$

Applicando il Teorema 7,

$$\|g_v(u)\| = \|g_v(u) - g_v(0)\| \le 2\varepsilon(\|u\| + \|v\|)^2\,.$$

Abbiamo così dimostrato quanto volevamo, cioè che il limite di partenza è uguale a zero; equivalentemente, scambiando u e v,

$$\lim_{(u,v)\to(0,0)} \frac{f(x_0+v+u) - f(x_0+v) - f(x_0+u) + f(x_0) - d^2f(x_0)(u,v)}{(\|v\| + \|u\|)^2} = 0\,.$$

Ma allora anche

$$\lim_{(u,v)\to(0,0)} \frac{d^2f(x_0)(v,u) - d^2f(x_0)(u,v)}{(\|u\| + \|v\|)^2} = 0\,.$$

Prendiamo arbitrariamente due vettori h, k con $\|h\| = \|k\| = 1$, e fissiamo $\varepsilon > 0$. Per quanto sopra, esiste un $\delta > 0$ tale che, se $\|u\| + \|v\| \le 2\delta$, allora

$$\|d^2f(x_0)(v,u) - d^2f(x_0)(u,v)\| \le \varepsilon(\|v\| + \|u\|)^2.$$

Allora, per ogni $\tau \in]0,\delta[$ è $\|\tau h\| + \|\tau k\| = 2\tau \le 2\delta$, per cui

$$
\begin{aligned}
\|d^2f(x_0)(k,h) - d^2f(x_0)(h,k)\| &= \frac{1}{\tau^2}\|d^2f(x_0)(\tau k,\tau h) - d^2f(x_0)(\tau h,\tau k)\| \\
&\le \frac{1}{\tau^2}\varepsilon(\|\tau k\| + \|\tau h\|)^2 = \varepsilon(\|k\| + \|h\|)^2.
\end{aligned}
$$

Ne segue che $d^2f(x_0)(k,h) - d^2f(x_0)(h,k) = 0$, e questo è vero comunque si prendano h, k con $\|h\| = \|k\| = 1$. La dimostrazione è così completata. ∎

Il teorema precedente afferma quindi che, se f è differenziabile due volte in x_0, allora, per ogni h e k in X, si ha

$$\partial_h\partial_k f(x_0) = \partial_k\partial_h f(x_0)\,.$$

Enunciamo ora un secondo teorema in cui vale tale uguaglianza, il noto **Teorema di Schwarz**.

Teorema 15 *Dati due vettori h e k di X, supponiamo che esistano $\partial_h\partial_k f(x)$ e $\partial_k\partial_h f(x)$, per tutti gli x in un intorno U di x_0, e che le due funzioni $\partial_h\partial_k f : U \to Y$ e $\partial_k\partial_h f : U \to Y$ siano continue in x_0. Allora*

$$\partial_h\partial_k f(x_0) = \partial_k\partial_h f(x_0)\,.$$

Dimostrazione. Possiamo supporre che $\|h\| = \|k\| = 1$. Consideriamo la funzione

$$g_1(\sigma,\tau) = f(x_0 + \sigma h + \tau k) - \sigma\tau\, \partial_k \partial_h f(x_0)\,,$$

definita in una palla $B((0,0),r) \subseteq \mathbb{R}^2$, a valori in Y. Dimostriamo che

$$\lim_{(\sigma,\tau)\to(0,0)} \frac{g_1(\sigma,\tau) - g_1(\sigma,0) - g_1(0,\tau) + g_1(0,0)}{\sigma\tau} = 0\,.$$

Fissiamo un $\varepsilon > 0$. Per la continuità di $\partial_h \partial_k f$ in x_0, esiste un $\delta > 0$ tale che

$$\sigma^2 + \tau^2 < \delta^2 \quad \Rightarrow \quad \left\| \frac{\partial}{\partial\tau} \frac{\partial g_1}{\partial\sigma}(\sigma,\tau) \right\| < \varepsilon\,.$$

Allora, se $\sigma^2 + \tau^2 < \delta^2$, usando il Teorema 7, si ha

$$\begin{aligned}
\|(g_1(\sigma,\tau) & - g_1(\sigma,0)) - (g_1(0,\tau) - g_1(0,0))\| \leq \\
& \leq \sup\left\{ \left\| \frac{\partial g_1}{\partial\sigma}(s,\tau) - \frac{\partial g_1}{\partial\sigma}(s,0) \right\| : s \in [0,\sigma] \right\} |\sigma| \\
& \leq \sup\left\{ \sup\left\{ \left\| \frac{\partial}{\partial\tau} \frac{\partial g_1}{\partial\sigma}(s,t) \right\| : t \in [0,\tau] \right\} |\tau| : s \in [0,\sigma] \right\} |\sigma| \\
& < \varepsilon |\tau|\, |\sigma|\,.
\end{aligned}$$

In modo del tutto analogo si dimostra anche che, definendo la funzione

$$g_2(\sigma,\tau) = f(x_0 + \sigma h + \tau k) - \sigma\tau\, \partial_h \partial_k f(x_0)\,,$$

si ha che

$$\lim_{(\sigma,\tau)\to(0,0)} \frac{g_2(\sigma,\tau) - g_2(\sigma,0) - g_2(0,\tau) + g_2(0,0)}{\sigma\tau} = 0\,.$$

Mettendo assieme i due limiti e sottraendo, ne segue direttamente la tesi. ■

Il differenziale n-esimo di f in x_0 si denota con $d^n f(x_0)$ o con $f^{(n)}(x_0)$, e si definisce per induzione:

$$d^n f(x_0) = d(d^{n-1} f)(x_0)\,.$$

Si considera $d^n f(x_0)$ come un'applicazione n-lineare limitata da $X \times \ldots \times X$ in Y. Si ha

$$df(x_0)(h_1, \ldots, h_n) = \partial_{h_1} \ldots \partial_{h_n} f(x_0)$$

con $h_1, \ldots, h_n \in X$. Se f è n volte differenziabile in x_0, si dimostra la seguente proprietà di simmetria per il differenziale n-esimo: se $(h_1, \ldots, h_n)$ è una qualsiasi n-upla di vettori di X e $(h_{s(1)}, \ldots, h_{s(n)})$ ne è una permutazione, allora

$$d^n f(x_0)(h_1, \ldots, h_n) = d^n f(x_0)(h_{s(1)}, \ldots, h_{s(n)})\,.$$

8 La formula di Taylor

Supponiamo che Ω sia un aperto di X e che $f : \Omega \to Y$ sia una funzione $(n+1)$-volte differenziabile, per un certo $n \geq 1$. Sarà comodo usare la seguente notazione:

$$d^j f(x_0)(h)^j = d^j f(x_0)(\underbrace{h, \dots, h}_{j \text{ volte}}).$$

Teorema 16 *Se $[x_0, x]$ è un segmento contenuto in Ω, allora*

$$\|f(x) - p_n(x)\| \leq \frac{1}{(n+1)!} \sup\Big\{\|d^{n+1} f(v)\| : v \in [x_0, x]\Big\} \|x - x_0\|^{n+1},$$

dove

$$p_n(x) = f(x_0) + df(x_0)(x - x_0) + \frac{1}{2!} d^2 f(x_0)(x - x_0)^2 + \dots + \frac{1}{n!} d^n f(x_0)(x - x_0)^n$$

è il "polinomio di Taylor" di grado n associato alla funzione f nel punto x_0.

<u>Dimostrazione</u>. Se l'estremo superiore è uguale a $+\infty$, non c'è nulla da dimostrare. Supponiamo quindi che sia

$$\sup\Big\{\|d^{n+1} f(v)\| : v \in [x_0, x]\Big\} = C \in \mathbb{R}.$$

Sia $u : [0,1] \to Y$ definita da $u(t) = f((1-t)x_0 + tx)$. Essa è una funzione $(n+1)$-volte differenziabile. Vediamo che

$$\begin{aligned}
&u'(t) = df((1-t)x_0 + tx)(x - x_0), \\
&u''(t) = d^2 f((1-t)x_0 + tx)(x - x_0)^2 \\
&\dots \\
&u^{(j)}(t) = d^j f((1-t)x_0 + tx)(x - x_0)^j \\
&\dots \\
&u^{(n+1)}(t) = d^{n+1} f((1-t)x_0 + tx)(x - x_0)^{n+1},
\end{aligned}$$

per cui

$$\|u^{(n+1)}(t)\| \leq C\|x - x_0\|^{n+1}, \quad \text{per ogni } t \in [0,1].$$

Siano $\varphi : [0,1] \to Y$ e $g : [0,1] \to \mathbb{R}$ definite da

$$\varphi(t) = u(t) + (1-t)u'(t) + \frac{1}{2!}(1-t)^2 u''(t) + \dots + \frac{1}{n!}(1-t)^n u^{(n)}(t),$$

$$g(t) = -\frac{C\|x - x_0\|^{n+1}}{(n+1)!}(1-t)^{n+1}.$$

Si vede che

$$\|\varphi'(t)\| = \left\| \frac{1}{n!}(1-t)^n u^{(n+1)}(t) \right\| \leq \frac{1}{n!} C \|x - x_0\|^{n+1}(1-t)^n = g'(t),$$

per ogni $t \in [0,1]$. Applicando il Lemma 2, abbiamo che $\|\varphi(1) - \varphi(0)\| \leq g(1) - g(0)$, ossia

$$\left\| u(1) - u(0) - u'(0) - \frac{1}{2!}u''(0) - \ldots - \frac{1}{n!}u^{(n)}(0) \right\| \leq \frac{C\|x - x_0\|^{n+1}}{(n+1)!} \,,$$

ossia ancora

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

che è quanto si voleva dimostrare. ∎

Concludiamo con il seguente

Corollario 3 *Se $f : \Omega \to Y$ è di classe C^{N+1} e $[x_0, x]$ è un segmento contenuto in Ω, allora si può scrivere*

$$f(x) = p_n(x) + r_n(x) \,,$$

con

$$\lim_{x \to x_0} \frac{r_n(x)}{\|x - x_0\|^n} = 0 \,.$$