Analisi Matematica II

Appunti delle lezioni tenute dalla Prof.ssa R. Toader

Università di Trieste, CdL Ingegneria Industriale e CdL Ingegneria Navale

a.a. 2024/2025

1 Calcolo differenziale: funzioni da $\mathbb{R}^N$ a $\mathbb{R}$

In questa sezione, E sarà un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di E e $f : E \to \mathbb{R}$ una funzione. Vogliamo estendere il concetto di derivata già introdotto nel caso $N = 1$. Iniziamo con il fissare una "direzione", ossia un vettore $\boldsymbol{v} \in \mathbb{R}^N$ tale che $\|\boldsymbol{v}\| = 1$ (detto anche "versore"). Chiamiamo, se esiste, "derivata direzionale" di f in $\boldsymbol{x}_0$ nella direzione $\boldsymbol{v}$ il seguente limite

$$\lim_{t\to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t},$$

che verrà indicato con il simbolo

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0).$$

Se $\boldsymbol{v}$ coincide con un elemento $\boldsymbol{e}_k$ della base canonica $[\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N]$ di $\mathbb{R}^N$, la derivata direzionale si chiamerà "derivata parziale" k-esima di f in $\boldsymbol{x}_0$ e si indicherà con

$$\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0).$$

Se $\boldsymbol{x}_0 = (x_1^0, x_2^0, \ldots, x_N^0)$, si ha quindi:

$$\begin{aligned}\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) &= \lim_{t\to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{e}_k) - f(\boldsymbol{x}_0)}{t} \\ &= \lim_{t\to 0} \frac{f(x_1^0, x_2^0, \ldots, x_k^0 + t, \ldots, x_N^0) - f(x_1^0, x_2^0, \ldots, x_k^0, \ldots, x_N^0)}{t},\end{aligned}$$

per cui si usa parlare di "derivata rispetto alla k-esima variabile".

Esistono delle funzioni che, pur avendo derivate direzionali in tutte le possibili direzioni, non sono continue. Ad esempio, la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita da

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

ha tutte le derivate direzionali nulle in $\boldsymbol{x}_0 = (0,0)$, ma non è continua in tale punto, come si vede considerando la restrizione alla parabola $\{(x,y) \in \mathbb{R}^2 : y = x^2\}$. Questo fatto ci porta a cercare una generalizzazione più appropriata del concetto di derivata.

Definizione 1 *Diremo che la funzione f è "differenziabile" in $\boldsymbol{x}_0$ se esiste una applicazione lineare $\ell : \mathbb{R}^N \to \mathbb{R}$ per cui si possa scrivere*

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})\,,$$

dove r è una funzione tale che

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0\,.$$

Se f è differenziabile in $\boldsymbol{x}_0$, l'applicazione lineare ℓ si chiama "differenziale" di f in $\boldsymbol{x}_0$ e si indica con il simbolo

$$df(\boldsymbol{x}_0)\,.$$

Teorema 1 *Se f è differenziabile in $\boldsymbol{x}_0$, allora f è continua in $\boldsymbol{x}_0$.*

Dimostrazione. Sappiamo che l'applicazione $\ell = df(\boldsymbol{x}_0)$, essendo lineare, è continua e $\ell(\boldsymbol{0}) = 0$. Ne segue che

$$\begin{aligned} \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) &= \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} [f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})] \\ &= f(\boldsymbol{x}_0) + \ell(\boldsymbol{0}) + \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} r(\boldsymbol{x}) \\ &= f(\boldsymbol{x}_0) + \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \|\boldsymbol{x} - \boldsymbol{x}_0\| \\ &= f(\boldsymbol{x}_0)\,, \end{aligned}$$

il che dimostra che f è continua in $\boldsymbol{x}_0$. ■

Seguendo un'abitudine consolidata per le applicazioni lineari, si usa spesso scrivere $df(\boldsymbol{x}_0)\boldsymbol{h}$ invece di $df(\boldsymbol{x}_0)(\boldsymbol{h})$.

Teorema 2 *Se f è differenziabile in $\boldsymbol{x}_0$, allora esistono tutte le derivate direzionali di f in $\boldsymbol{x}_0$: per ogni direzione $\boldsymbol{v} \in \mathbb{R}^N$ si ha*

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = df(\boldsymbol{x}_0)\boldsymbol{v}\,.$$

Dimostrazione. Usando la definizione di differenziale, abbiamo

$$\begin{aligned}
\lim_{t\to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t} &= \lim_{t\to 0} \frac{df(\boldsymbol{x}_0)(t\boldsymbol{v}) + r(\boldsymbol{x}_0 + t\boldsymbol{v})}{t} \\
&= \lim_{t\to 0} \frac{t\, df(\boldsymbol{x}_0)\boldsymbol{v} + r(\boldsymbol{x}_0 + t\boldsymbol{v})}{t} \\
&= df(\boldsymbol{x}_0)\boldsymbol{v} + \lim_{t\to 0} \frac{r(\boldsymbol{x}_0 + t\boldsymbol{v})}{t} \,;
\end{aligned}$$

d'altra parte, essendo $\|\boldsymbol{v}\| = 1$, si ha

$$\lim_{t\to 0} \left| \frac{r(\boldsymbol{x}_0 + t\boldsymbol{v})}{t} \right| = \lim_{t\to 0} \frac{|r(\boldsymbol{x}_0 + t\boldsymbol{v})|}{\|(\boldsymbol{x}_0 + t\boldsymbol{v}) - \boldsymbol{x}_0\|} = \lim_{\boldsymbol{x}\to \boldsymbol{x}_0} \frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0\,,$$

da cui la tesi. ■

In particolare, se $\boldsymbol{v}$ coincide con un elemento $\boldsymbol{e}_k$ della base canonica $[\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_N]$, si ha:

$$\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) = df(\boldsymbol{x}_0)\boldsymbol{e}_k \,.$$

Scrivendo il vettore $\boldsymbol{h} \in \mathbb{R}^N$ come $\boldsymbol{h} = h_1\boldsymbol{e}_1 + h_2\boldsymbol{e}_2 + \cdots + h_N\boldsymbol{e}_N$, abbiamo

$$\begin{aligned}
df(\boldsymbol{x}_0)\boldsymbol{h} &= h_1 df(\boldsymbol{x}_0)\boldsymbol{e}_1 + h_2 df(\boldsymbol{x}_0)\boldsymbol{e}_2 + \cdots + h_N df(\boldsymbol{x}_0)\boldsymbol{e}_N \\
&= h_1 \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0) + h_2 \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0) + \cdots + h_N \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0)\,,
\end{aligned}$$

ossia

$$df(\boldsymbol{x}_0)\boldsymbol{h} = \sum_{k=1}^{N} \frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) h_k \,.$$

Introducendo il vettore "gradiente" di f in $\boldsymbol{x}_0$

$$\nabla f(\boldsymbol{x}_0) = \left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0), \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0), \dots, \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0) \right)\,,$$

si può scrivere

$$df(\boldsymbol{x}_0)\boldsymbol{h} = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{h} \,.$$

Possiamo quindi scrivere

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})\,,$$

con

$$\lim_{\boldsymbol{x}\to \boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0\,.$$

Analizziamo con maggiore attenzione il caso $N = 2$. Come di consueto, invece di usare la notazione (x_1, x_2), gli elementi di $\mathbb{R}^2$ verranno denotati con (x, y). Fissato quindi il punto $\boldsymbol{x}_0 = (x_0, y_0)$, possiamo scrivere

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

con

$$\lim_{(x,y)\to(x_0,y_0)} \frac{r(x,y)}{\sqrt{(x-x_0)^2+(y-y_0)^2}} = 0\,.$$

Ricordando che il grafico di f è l'insieme

$$G_f = \{(x,y,z)\in\mathbb{R}^3 : z = f(x,y)\}\,,$$

chiameremo "piano tangente" al grafico di f nel punto $(x_0, y_0, f(x_0,y_0))$ l'insieme

$$\left\{(x,y,z)\in\mathbb{R}^3 : z = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)\right\}.$$

2 Funzioni di classe $\mathcal{C}^1$

Il seguente risultato è noto come "teorema del differenziale totale".

Teorema 3 *Se f possiede le derivate parziali in un intorno di $\boldsymbol{x}_0$ ed esse sono continue in $\boldsymbol{x}_0$, allora f è differenziabile in $\boldsymbol{x}_0$.*

Dimostrazione. Supporremo per semplicità di notazioni $N = 2$. Definiamo l'applicazione lineare $\ell : \mathbb{R}^2 \to \mathbb{R}$ che ad ogni vettore $\boldsymbol{h} = (h_1, h_2)$ associa

$$\ell(\boldsymbol{h}) = \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)h_1 + \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)h_2\,.$$

Vedremo che ℓ è proprio il differenziale di f in $\boldsymbol{x}_0$. Intanto, è lineare, come si vede immediatamente. Inoltre, scrivendo $\boldsymbol{x}_0 = (x_1^0, x_2^0)$ e $\boldsymbol{x} = (x_1, x_2)$, per il Teorema di Lagrange si ha

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{x}_0) &= (f(x_1,x_2) - f(x_1^0,x_2)) + (f(x_1^0,x_2) - f(x_1^0,x_2^0)) \\ &= \frac{\partial f}{\partial x_1}(\xi_1,x_2)(x_1-x_1^0) + \frac{\partial f}{\partial x_2}(x_1^0,\xi_2)(x_2-x_2^0)\,, \end{aligned}$$

per un certo $\xi_1 \in [x_1^0, x_1]$ e un certo $\xi_2 \in [x_2^0, x_2]$. Quindi,

$$\begin{aligned} r(\boldsymbol{x}) &= f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - \ell(\boldsymbol{x}-\boldsymbol{x}_0) \\ &= \left[\frac{\partial f}{\partial x_1}(\xi_1,x_2) - \frac{\partial f}{\partial x_1}(x_1^0,x_2^0)\right](x_1-x_1^0) + \\ &\quad + \left[\frac{\partial f}{\partial x_2}(x_1^0,\xi_2) - \frac{\partial f}{\partial x_2}(x_1^0,x_2^0)\right](x_2-x_2^0)\,, \end{aligned}$$

ed essendo $|x_1 - x_1^0| \le \|\boldsymbol{x}-\boldsymbol{x}_0\|$ e $|x_2 - x_2^0| \le \|\boldsymbol{x}-\boldsymbol{x}_0\|$,

$$\frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \le \left|\frac{\partial f}{\partial x_1}(\xi_1,x_2) - \frac{\partial f}{\partial x_1}(x_1^0,x_2^0)\right| + \left|\frac{\partial f}{\partial x_2}(x_1^0,\xi_2) - \frac{\partial f}{\partial x_2}(x_1^0,x_2^0)\right|.$$

Facendo tendere $\boldsymbol{x}$ a $\boldsymbol{x}_0$, si ha che $(\xi_1,x_2)\to(x_1^0,x_2^0)$ e $(x_1^0,\xi_2)\to(x_1^0,x_2^0)$ per cui, essendo $\frac{\partial f}{\partial x_1}$ e $\frac{\partial f}{\partial x_2}$ continue in $\boldsymbol{x}_0 = (x_1^0, x_2^0)$, si ha

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} = 0\,,$$

da cui la tesi. ∎

Diremo che la funzione f è di classe $\mathcal{C}^1$ su E se f possiede le derivate parziali ed esse sono continue su tutto E. Dal teorema precedente segue che una funzione di classe $\mathcal{C}^1$ è "differenziabile su E", ossia in ogni punto di E.

3 Derivate parziali successive

Supponiamo, per semplicità, $N = 2$. Consideriamo E, un insieme aperto di $\mathbb{R}^2$ e una funzione $f : E \to \mathbb{R}$ che abbia le derivate parziali $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ in tutti i punti di E. Se esse posseggono a loro volta derivate parziali in un punto $\boldsymbol{x}_0$, queste si dicono "derivate parziali seconde" della f in $\boldsymbol{x}_0$ e si denotano con i simboli

$$\frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_1}\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\,, \quad \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_2}\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\,,$$
$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_1}\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\,, \quad \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_2}\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\,.$$

Teorema 4 (di Schwarz) *Se esistono le derivate parziali seconde* $\frac{\partial^2 f}{\partial x_2 \partial x_1}$, $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ *in un intorno di* $\boldsymbol{x}_0$ *ed esse sono continue in* $\boldsymbol{x}_0$, *allora*

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0)\,.$$

Dimostrazione. Sia $\rho > 0$ tale che $B(\boldsymbol{x}_0, \rho) \subseteq E$. Scriviamo $\boldsymbol{x}_0 = (x_1^0, x_2^0)$ e prendiamo un $\boldsymbol{x} = (x_1, x_2) \in B(\boldsymbol{x}_0, \rho)$ tale che $x_1 \neq x_1^0$ e $x_2 \neq x_2^0$. Possiamo allora definire

$$g(x_1, x_2) = \frac{f(x_1, x_2) - f(x_1, x_2^0)}{x_2 - x_2^0}\,, \quad h(x_1, x_2) = \frac{f(x_1, x_2) - f(x_1^0, x_2)}{x_1 - x_1^0}\,.$$

Si verifica che vale l'uguaglianza

$$\frac{g(x_1, x_2) - g(x_1^0, x_2)}{x_1 - x_1^0} = \frac{h(x_1, x_2) - h(x_1, x_2^0)}{x_2 - x_2^0}\,.$$

Per il Teorema di Lagrange, esiste un $\xi_1 \in]x_1^0, x_1[$ tale che

$$\frac{g(x_1, x_2) - g(x_1^0, x_2)}{x_1 - x_1^0} = \frac{\partial g}{\partial x_1}(\xi_1, x_2) = \frac{\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(\xi_1, x_2^0)}{x_2 - x_2^0}\,,$$

ed esiste un $\xi_2 \in]x_2^0, x_2[$ tale che

$$\frac{h(x_1, x_2) - h(x_1, x_2^0)}{x_2 - x_2^0} = \frac{\partial h}{\partial x_2}(x_1, \xi_2) = \frac{\frac{\partial f}{\partial x_2}(x_1, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, \xi_2)}{x_1 - x_1^0}\,.$$

Di nuovo per il Teorema di Lagrange, esiste un $\eta_2 \in]x_2^0, x_2[$ tale che

$$\frac{\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(\xi_1, x_2^0)}{x_2 - x_2^0} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \eta_2)\,,$$

ed esiste un $\eta_1 \in]x_1^0, x_1[$ tale che

$$\frac{\frac{\partial f}{\partial x_2}(x_1, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, \xi_2)}{x_1 - x_1^0} = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\eta_1, \xi_2)\,.$$

Quindi,

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \eta_2) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\eta_1, \xi_2)\,.$$

Facendo tendere $\boldsymbol{x} = (x_1, x_2)$ a $\boldsymbol{x}_0 = (x_1^0, x_2^0)$, si ha che sia (ξ_1, η_2) che (η_1, ξ_2) tendono a $\boldsymbol{x}_0$, e per la continuità delle derivate seconde miste si ha la tesi. ■

Diremo che la funzione f è di classe $\mathcal{C}^2$ su E se f possiede tutte le derivate parziali seconde ed esse sono continue su tutto E. Dal teorema precedente segue che se una funzione di classe $\mathcal{C}^2$, le derivate parziali "miste" sono uguali.

È utile definire la "matrice hessiana" di f nel punto $\boldsymbol{x}_0$:

$$Hf(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) \end{pmatrix};$$

se f è di classe $\mathcal{C}^2$, si tratta di una matrice simmetrica.

Quanto sopra si può estendere senza difficoltà alle funzioni di N variabili, con N qualunque. Se f è di classe $\mathcal{C}^2$, la matrice hessiana risulta allora una matrice simmetrica del tipo $N \times N$:

$$Hf(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) & \dots & \frac{\partial^2 f}{\partial x_N \partial x_1}(\boldsymbol{x}_0) \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) & \dots & \frac{\partial^2 f}{\partial x_N \partial x_2}(\boldsymbol{x}_0) \\ \\ \vdots & \vdots & \dots & \vdots \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_N}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_N}(\boldsymbol{x}_0) & \dots & \frac{\partial^2 f}{\partial x_N^2}(\boldsymbol{x}_0) \end{pmatrix}.$$

Procedendo per induzione, si possono definire le derivate parziali n-esime di una funzione. Si dice che la funzione f è di classe $\mathcal{C}^n$ su E se f possiede tutte le derivate parziali n-esime ed esse sono continue su tutto E.

4 La formula di Taylor

Supponiamo ora che $f : E \to \mathbb{R}$ sia una funzione di classe $\mathcal{C}^{n+1}$, per un certo $n \geq 1$.

Consideriamo come sopra, per semplicità, il caso $N = 2$. Introduciamo le seguenti notazioni:

$$D_{x_1} = \frac{\partial}{\partial x_1}\,, \quad D_{x_2} = \frac{\partial}{\partial x_2}\,,$$

$$D^2_{x_1} = \frac{\partial^2}{\partial x_1^2}\,, \quad D_{x_1}D_{x_2} = \frac{\partial^2}{\partial x_1 \partial x_2}\,, \quad D^2_{x_2} = \frac{\partial^2}{\partial x_2^2}\,,$$

e così via, per le derivate parziali successive. Si noti che, per un vettore $\boldsymbol{h} = (h_1, h_2) \in \mathbb{R}^2$, si ha

$$df(\boldsymbol{x}_0)\boldsymbol{h} = h_1 D_{x_1} f(\boldsymbol{x}_0) + h_2 D_{x_2} f(\boldsymbol{x}_0)\,,$$

che risulterà conveniente scrivere

$$df(\boldsymbol{x}_0)\boldsymbol{h} = [h_1 D_{x_1} + h_2 D_{x_2}] f(\boldsymbol{x}_0)\,.$$

In questo modo, possiamo pensare che f viene trasformata dall'operatore $[h_1 D_{x_1} + h_2 D_{x_2}]$ nella nuova funzione $[h_1 D_{x_1} + h_2 D_{x_2}]f = h_1 D_{x_1} f + h_2 D_{x_2} f$.

Dati due punti $\boldsymbol{x}_0$ e $\boldsymbol{x}$ in $\mathbb{R}^N$, si definisce il "segmento" che li congiunge:

$$[\boldsymbol{x}_0, \boldsymbol{x}] = \{\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0) : t \in [0,1]\}\,;$$

analogamente, scriveremo

$$]\boldsymbol{x}_0, \boldsymbol{x}[= \{\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0) : t \in]0,1[\,\}\,.$$

Supponiamo ora che $[\boldsymbol{x}_0, \boldsymbol{x}]$ sia un segmento contenuto in E e consideriamo la funzione $\phi : [0,1] \to \mathbb{R}$ definita da

$$\phi(t) = f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\,.$$

Dimostriamo che ϕ è derivabile $n+1$ volte su $[0,1]$. Per $t \in [0,1]$, essendo f differenziabile in $\boldsymbol{u}_0 = \boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0)$, si ha

$$f(\boldsymbol{u}) = f(\boldsymbol{u}_0) + df(\boldsymbol{u}_0)(\boldsymbol{u} - \boldsymbol{u}_0) + r(\boldsymbol{u})\,,$$

con

$$\lim_{\boldsymbol{u} \to \boldsymbol{u}_0} \frac{r(\boldsymbol{u})}{\|\boldsymbol{u} - \boldsymbol{u}_0\|} = 0\,.$$

Quindi,

$$\begin{aligned}
\lim_{s \to t} \frac{\phi(s) - \phi(t)}{s - t} &= \lim_{s \to t} \frac{f(\boldsymbol{x}_0 + s(\boldsymbol{x} - \boldsymbol{x}_0)) - f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))}{s - t} \\
&= \lim_{s \to t} \frac{df(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))((s-t)(\boldsymbol{x} - \boldsymbol{x}_0)) + r(\boldsymbol{x}_0 + s(\boldsymbol{x} - \boldsymbol{x}_0))}{s - t} \\
&= df(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))(\boldsymbol{x} - \boldsymbol{x}_0) + \lim_{s \to t} \frac{r(\boldsymbol{x}_0 + s(\boldsymbol{x} - \boldsymbol{x}_0))}{s - t}\,,
\end{aligned}$$

ed essendo

$$\lim_{s \to t} \left| \frac{r(\boldsymbol{x}_0 + s(\boldsymbol{x} - \boldsymbol{x}_0))}{s - t} \right| = \lim_{\boldsymbol{u} \to \boldsymbol{u}_0} \frac{|r(\boldsymbol{u})|}{\|\boldsymbol{u} - \boldsymbol{u}_0\|} \|\boldsymbol{x} - \boldsymbol{x}_0\| = 0\,,$$

si ha

$$\phi'(t) = \lim_{s \to t} \frac{\phi(s) - \phi(t)}{s - t} = df(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))(\boldsymbol{x} - \boldsymbol{x}_0)\,.$$

Con le nuove notazioni, ponendo $\boldsymbol{x} - \boldsymbol{x}_0 = \boldsymbol{h} = (h_1, h_2)$, abbiamo

$$\phi'(t) = [h_1 D_{x_1} + h_2 D_{x_2}] f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0)) = g(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\,,$$

dove g è la nuova funzione $[h_1 D_{x_1} + h_2 D_{x_2}] f$. Possiamo allora iterare il procedimento, e calcolare la derivata seconda di ϕ:

$$\begin{aligned}\phi''(t) &= [h_1 D_{x_1} + h_2 D_{x_2}] g(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0)) \\ &= [h_1 D_{x_1} + h_2 D_{x_2}][h_1 D_{x_1} + h_2 D_{x_2}] f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\,.\end{aligned}$$

Per brevità, scriveremo

$$\phi''(t) = [h_1 D_{x_1} + h_2 D_{x_2}]^2 f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\,.$$

Notiamo che, usando la linearità delle derivate parziali e l'uguaglianza delle derivate miste (Teorema di Schwarz), si ha

$$\begin{aligned}[h_1 D_{x_1} + h_2 D_{x_2}]^2 f &= h_1^2 D_{x_1}^2 f + 2h_1 h_2 D_{x_1} D_{x_2} f + h_2^2 D_{x_2}^2 f \\ &= [h_1^2 D_{x_1}^2 + 2h_1 h_2 D_{x_1} D_{x_2} + h_2^2 D_{x_2}^2] f\,.\end{aligned}$$

Osserviamo che l'espressione

$$[h_1 D_{x_1} + h_2 D_{x_2}]^2 = [h_1^2 D_{x_1}^2 + 2h_1 h_2 D_{x_1} D_{x_2} + h_2^2 D_{x_2}^2]$$

si ottiene formalmente come il quadrato di un binomio. Procedendo in questo modo, si può dimostrare per induzione che, per $k = 1, 2, \ldots, n+1$, la formula della derivata k-esima di ϕ è

$$\phi^{(k)}(t) = [h_1 D_{x_1} + h_2 D_{x_2}]^k f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\,,$$

e che, usando formalmente la formula del binomio di Newton

$$(a_1 + a_2)^k = \sum_{j=0}^{k} \binom{k}{j} a_1^{k-j} a_2^j,$$

si ha

$$[h_1 D_{x_1} + h_2 D_{x_2}]^k = \left[\sum_{j=0}^{k} \binom{k}{j} h_1^{k-j} h_2^j D_{x_1}^{k-j} D_{x_2}^j\right]$$

(in questa formula, i simboli $D_{x_1}^0$ e $D_{x_2}^0$ vanno interpretati come l'operatore identità).

Per poter scrivere agevolmente la formula di Taylor, introduciamo la notazione

$$d^k f(\boldsymbol{x}_0) \boldsymbol{h}^k = [h_1 D_{x_1} + h_2 D_{x_2}]^k f(\boldsymbol{x}_0)\,.$$

Teorema 5 *Sia $f : E \to \mathbb{R}$ di classe $\mathcal{C}^{n+1}$ e $[\boldsymbol{x}_0, \boldsymbol{x}]$ un segmento contenuto in E. Allora esiste un $\boldsymbol{\xi} \in]\boldsymbol{x}_0, \boldsymbol{x}[$ tale che*

$$f(\boldsymbol{x}) = p_n(\boldsymbol{x}) + r_n(\boldsymbol{x})\,,$$

dove

$$p_n(\boldsymbol{x}) = f(\boldsymbol{x}_0)+df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)+\frac{1}{2!}d^2f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^2+\cdots+\frac{1}{n!}d^nf(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^n$$

è il "polinomio di Taylor di grado n associato alla funzione f nel punto $\boldsymbol{x}_0$" e

$$r_n(\boldsymbol{x}) = \frac{1}{(n+1)!}d^{n+1}f(\boldsymbol{\xi})(\boldsymbol{x}-\boldsymbol{x}_0)^{n+1}$$

è il "resto di Lagrange".

Dimostrazione. Per la formula di Taylor applicata alla funzione ϕ, si ha

$$\phi(t) = \phi(0) + \phi'(0)t + \frac{1}{2!}\phi''(0)t^2 + \cdots + \frac{1}{n!}\phi^{(n)}(0)t^n + \frac{1}{(n+1)!}\phi^{(n+1)}(\xi)t^{n+1}\,,$$

per un certo $\xi \in]0,t[$. La formula cercata si ottiene prendendo $t = 1$ e sostituendo i valori delle derivate di ϕ trovati sopra. ■

Il polinomio di Taylor si può anche scrivere nella forma compatta

$$p_n(\boldsymbol{x}) = \sum_{k=0}^{n} \frac{1}{k!}d^kf(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^k\,,$$

con la convenzione che $d^0f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^0$, il primo addendo della somma, sia $f(\boldsymbol{x}_0)$. Si ha quindi

$$\begin{aligned} p_n(\boldsymbol{x}) &= \sum_{k=0}^{n} \frac{1}{k!}[(x_1-x_1^0)D_{x_1} + (x_2-x_2^0)D_{x_2}]^k f(\boldsymbol{x}_0) \\ &= \sum_{k=0}^{n} \frac{1}{k!}\left(\sum_{j=0}^{k}\binom{k}{j}\frac{\partial^k f}{\partial x_1^{k-j}\partial x_2^j}(\boldsymbol{x}_0)\;(x_1-x_1^0)^{k-j}(x_2-x_2^0)^j\right)\,. \end{aligned}$$

Può essere utile la seguente espressione per il polinomio di secondo grado:

$$p_2(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + \tfrac{1}{2}\Big(Hf(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\Big)\cdot(\boldsymbol{x}-\boldsymbol{x}_0)\,.$$

Il teorema sopra dimostrato resta valido per qualsiasi dimensione N, pur di interpretare correttamente le notazioni: ad esempio, per un vettore $\boldsymbol{h} = (h_1, h_2, \ldots, h_N)$, si dovrà leggere

$$d^kf(\boldsymbol{x}_0)\boldsymbol{h}^k = [h_1D_{x_1} + h_2D_{x_2} + \cdots + h_ND_{x_N}]^k f(\boldsymbol{x}_0)\,.$$

In questo caso, volendo esplicitare il polinomio di Taylor, sarà utile utilizzare la formula di Leibniz

$$(a_1+a_2+\cdots+a_N)^k = \sum_{m_1+m_2+\cdots+m_N=k} \frac{k!}{m_1!\,m_2!\cdots m_N!}\,a_1^{m_1}a_2^{m_2}\cdots a_N^{m_N}\,.$$

5 La ricerca di massimi e minimi

Come sopra, consideriamo un insieme aperto $E \subseteq \mathbb{R}^N$ e una funzione $f : E \to \mathbb{R}$. Diremo che $\boldsymbol{x}_0 \in E$ è un "punto di massimo locale" per la funzione f se esiste un intorno U di x_0 contenuto in E per cui x_0 è punto di massimo della restrizione di f a U. Equivalentemente, se

$$\exists \delta > 0 : \ \forall \boldsymbol{x} \in E \quad d(\boldsymbol{x}, \boldsymbol{x}_0) < \delta \ \Rightarrow \ f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0)\,.$$

Analogamente per "punto di minimo locale".

Teorema 6 (di Fermat) *Se $\boldsymbol{x}_0$ è un punto di massimo o di minimo locale e f è differenziabile in $\boldsymbol{x}_0$, allora $\nabla f(\boldsymbol{x}_0) = 0$.*

Dimostrazione. Se $\boldsymbol{x}_0$ è punto di massimo locale, per ogni direzione $\boldsymbol{v} \in \mathbb{R}^N$ avremo che

$$\frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t} \quad \begin{cases} \geq 0 & \text{se } t < 0\,, \\ \leq 0 & \text{se } t > 0\,. \end{cases}$$

Siccome f è differenziabile in $\boldsymbol{x}_0$, ne deduciamo che

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \lim_{t \to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t} = 0\,.$$

In particolare, sono nulle tutte le derivate parziali, per cui $\nabla f(\boldsymbol{x}_0) = 0$. Nel caso in cui $\boldsymbol{x}_0$ sia un punto di minimo locale, si procede in modo analogo. ■

Definizione 2 *Un punto il cui il gradiente si annulli è detto "punto stazionario".*

Naturalmente un tale punto potrebbe non essere nè di massimo nè di minimo.

Mostreremo ora come la formula di Taylor possa essere usata per stabilire un criterio affinchè un punto stazionario sia di massimo, o di minimo. Iniziamo con una definizione. Diremo che una matrice $\mathbb{A}$ simmetrica $N \times N$ è *definita positiva* se

$$[\mathbb{A}\boldsymbol{h}] \cdot \boldsymbol{h} > 0\,, \qquad \text{per ogni } \boldsymbol{h} \in \mathbb{R}^N \setminus \{\boldsymbol{0}\}\,.$$

Diremo che $\mathbb{A}$ è *definita negativa* se vale la disuguaglianza opposta, ossia se $-\mathbb{A}$ è definita positiva.

Teorema 7 *Se $\boldsymbol{x}_0$ è un punto stazionario e f è di classe $\mathcal{C}^2$, con matrice hessiana $Hf(\boldsymbol{x}_0)$ definita positiva, allora $\boldsymbol{x}_0$ è un punto di minimo locale. Se invece $Hf(\boldsymbol{x}_0)$ è definita negativa, allora $\boldsymbol{x}_0$ è un punto di massimo locale.*

Dimostrazione. Per la formula di Taylor, per $\boldsymbol{x} \neq \boldsymbol{x}_0$ in un intorno di $\boldsymbol{x}_0$ esiste un $\boldsymbol{\xi} \in]\boldsymbol{x}_0, \boldsymbol{x}[$ per cui

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \tfrac{1}{2}\Big(Hf(\boldsymbol{\xi})(\boldsymbol{x} - \boldsymbol{x}_0)\Big) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)\,.$$

Se $\mathbb{A} = Hf(\boldsymbol{x}_0)$ è definita positiva, esiste un $c > 0$ tale che, per ogni $\boldsymbol{v} \in \mathbb{R}^N$ con $\|\boldsymbol{v}\| = 1$,

$$[\mathbb{A}\boldsymbol{v}] \cdot \boldsymbol{v} \geq c\,.$$

(Abbiamo qui usato il Teorema di Weierstrass, e il fatto che la sfera $\{v \in \mathbb{R}^N : \|\boldsymbol{v}\| = 1\}$ è un insieme compatto.) Quindi

$$\Big(Hf(\boldsymbol{x}_0)\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}\Big) \cdot \frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \geq c\,.$$

Per la continuità delle derivate seconde, se $\boldsymbol{x}$ è sufficientemente vicino a $\boldsymbol{x}_0$,

$$\Big(Hf(\boldsymbol{\xi})\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}\Big) \cdot \frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \geq \tfrac{1}{2}c > 0\,.$$

(Lo si vede per assurdo, usando di nuovo la compattezza della sfera.) Essendo $\nabla f(\boldsymbol{x}_0) = \boldsymbol{0}$, per tali $\boldsymbol{x}$ abbiamo che

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + \tfrac{1}{2}\Big(Hf(\boldsymbol{\xi})(\boldsymbol{x}-\boldsymbol{x}_0)\Big) \cdot (\boldsymbol{x}-\boldsymbol{x}_0) \\ &\geq f(\boldsymbol{x}_0) + \tfrac{1}{2}c\|\boldsymbol{x}-\boldsymbol{x}_0\|^2 > f(\boldsymbol{x}_0)\,, \end{aligned}$$

per cui $\boldsymbol{x}_0$ è un punto di minimo locale.

La dimostrazione della seconda affermazione è analoga. ■

Enunciamo ora (senza dimostrazione) due criteri utili a stabilire quando una matrice $\mathbb{A}$ simmetrica $N \times N$ è definita positiva o negativa. Ricordiamo che gli autovalori di una matrice simmetrica sono tutti reali.

Primo criterio. *La matrice $\mathbb{A}$ è definita positiva se tutti i suoi autovalori sono positivi. Essa è definita negativa se tutti i suoi autovalori sono negativi.*

Secondo criterio. *La matrice $\mathbb{A} = (a_{ij})_{ij}$ è definita positiva se*

$$\begin{aligned} & a_{11} > 0\,, \\ & \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0\,, \\ & \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} > 0\,,\ \dots \\ & \det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{pmatrix} > 0\,. \end{aligned}$$

Essa è definita negativa se i determinanti scritti sopra hanno segno alternato: quelli delle sottomatrici con un numero dispari di righe e di colonne sono negativi, mentre quelli delle sottomatrici con un numero pari di righe e di colonne sono positivi.

6 Interpretazione duale del differenziale

Sia E un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di E e $f : E \to \mathbb{R}$ una funzione differenziabile in $\boldsymbol{x}_0$. Per definizione il differenziale di f in $\boldsymbol{x}_0$, $df(\boldsymbol{x}_0)$, è una applicazione lineare da $\mathbb{R}^N$ in $\mathbb{R}$; in quanto tale $df(\boldsymbol{x}_0)$ appartiene allo spazio duale $(\mathbb{R}^N)^*$ di $\mathbb{R}^N$.[1]

Lo spazio $(\mathbb{R}^N)^*$ è isomorfo a $\mathbb{R}^N$, dunque ha dimensione finita. Indichiamo con $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_N)$ la base canonica di $\mathbb{R}^N$. Per $i = 1, \dots, N$ sia $dx_i = \pi_i : \mathbb{R}^N \to \mathbb{R}$ la proiezione definita da $dx_i(\boldsymbol{x}) = \pi_i(\boldsymbol{x}) = x_i$, dove $\boldsymbol{x} = (x_1, \dots, x_N)$. La base canonica duale di $(\mathbb{R}^N)^*$ è data dalle proiezioni $(dx_1, \dots, dx_N)$. Scriviamo ora $df(\boldsymbol{x}_0)$ nella base canonica duale:

$$df(\boldsymbol{x}_0)\boldsymbol{h} = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{h} = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) h_i = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) \pi_i(\boldsymbol{h}) \,,$$

per cui

$$df(\boldsymbol{x}_0) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i} \pi_i \,,$$

quindi $df(\boldsymbol{x}_0)$ è rappresentato nella base canonica duale di $(\mathbb{R}^N)^*$ dal vettore

$$\Big(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_N}(x_0)\Big).$$

7 Interpretazione geometrica del gradiente

Data una funzione $f : E \subseteq \mathbb{R}^N \to \mathbb{R}$ definiamo l'insieme di livello $c \in \mathbb{R}$ di f come

$$\Sigma_c = \{\boldsymbol{x} \in E : f(\boldsymbol{x}) = c\} \,.$$

Consideriamo il caso $N = 2$. L'insieme di livello Σ_c si ottiene intersecando il grafico della funzione con il piano orizzontale di equazione $z = c$ e proiettando questa intersezione sul piano (x, y). Spesso gli insiemi di livello sono delle curve nel piano (x, y). Le curve di livello indicate sulle carte topografiche aiutano per esempio a capire la pendenza di una montagna.

Spostandosi nel piano (x, y) da un insieme di livello all'altro il valore della funzione aumenta o diminuisce. Vediamo ora che il gradiente in un punto (x_0, y_0) determina la direzione di massima pendenza sul grafico della funzione in un intorno del punto. Infatti, se consideriamo una direzione $\boldsymbol{v} = (v_1, v_2)$ e la restrizione della funzione alla retta passante per (x_0, y_0) e parallela a v: $\varphi(t) = f(x_0 + tv_1, y_0 + tv_2)$, la pendenza del suo grafico in $t = 0$ è data da

$$\varphi'(0) = \frac{\partial f}{\partial \boldsymbol{v}}(x_0, y_0) = \nabla f(x_0, y_0) \cdot \boldsymbol{v} \,.$$

[1] Per definizione lo spazio duale di $\mathbb{R}^N$ è lo spazio delle applicazioni lineari da $\mathbb{R}^N$ in $\mathbb{R}$.

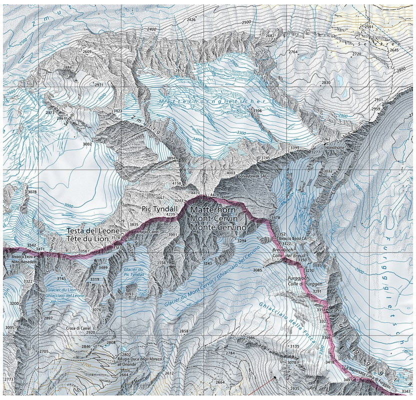

Figura 1: *Carta topografica del Monte Cervino* ©*Ufficio federale di topografia swisstopo*

Il gradiente di f individua la direzione $\boldsymbol{w} = \dfrac{\nabla f(x_0, y_0)}{\|\nabla f(x_0, y_0)\|}$. La derivata direzionale di f nella direzione $\boldsymbol{w}$ è data da

$$\begin{aligned}&\frac{\partial f}{\partial \boldsymbol{w}}(x_0, y_0) = \nabla f(x_0, y_0) \cdot \boldsymbol{w} = \nabla f(x_0, y_0) \cdot \frac{\nabla f(x_0, y_0)}{\|\nabla f(x_0, y_0)\|} = \frac{\|\nabla f(x_0, y_0)\|^2}{\|\nabla f(x_0, y_0)\|} \\ &= \|\nabla f(x_0, y_0)\| \, .\end{aligned}$$

Per la disuguaglianza di Schwarz si ha

$$\Big|\frac{\partial f}{\partial \boldsymbol{v}}(x_0, y_0)\Big| = |\nabla f(x_0, y_0) \cdot \boldsymbol{v}| \leq \|\nabla f(x_0, y_0)\| \, \|\boldsymbol{v}\| = \|\nabla f(x_0, y_0)\| = \frac{\partial f}{\partial \boldsymbol{w}}(x_0, y_0) \, ,$$

quindi per ogni direzione $\boldsymbol{v}$

$$-\frac{\partial f}{\partial \boldsymbol{w}}(x_0, y_0) \leq \frac{\partial f}{\partial \boldsymbol{v}}(x_0, y_0) \leq \frac{\partial f}{\partial \boldsymbol{w}}(x_0, y_0) \, ,$$

per cui la direzione del gradiente è quella in cui si ha il massimo accrescimento della funzione e la direzione opposta, individuata da $-\nabla f(x_0, y_0)$ è quella in cui si ha la massima decrescenza della funzione, in un intorno del punto (x_0, y_0). Da quanto detto segue che se in un punto il gradiente è non nullo allora spostandosi nelle direzione del gradiente i valori della funzione aumentano, spostandosi nella direzione opposta diminuiscono, quindi il punto in questione non può essere punto di massimo o punto di minimo della funzione. Di conseguenza i punti di massimo e di minimo locali interni al dominio vanno cercati tra i punti in cui il gradiente della funzione si annulla, come avevamo visto nel Teorema di Fermat.

8 Il differenziale di una funzione a valori vettoriali

Sia E un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di E e $f : E \to \mathbb{R}^M$ una funzione.

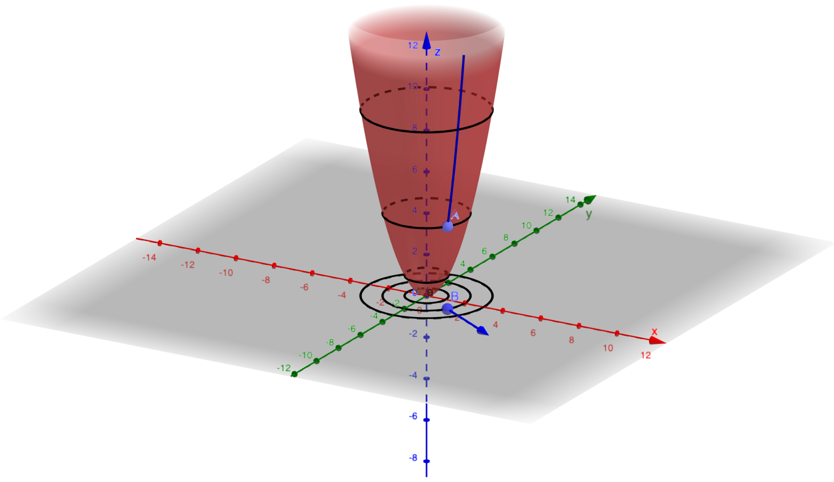

Figura 2: *Grafico di $f(x, y) = x^2 + y^2$, alcune curve di livello, il gradiente in un punto e la curva sul grafico corrispondente a un segmento nel piano (x, y) che parte dal punto fissato e va nella direzione del gradiente.*

Definizione 3 *Diremo che la funzione f è "differenziabile" in $\boldsymbol{x}_0$ se esiste una applicazione lineare $\ell : \mathbb{R}^N \to \mathbb{R}^M$ per cui si possa scrivere*

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})\,,$$

dove r è una funzione tale che

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \mathbf{0}\,.$$

Se f è differenziabile in $\boldsymbol{x}_0$, l'applicazione lineare ℓ si chiama "differenziale" di f in $\boldsymbol{x}_0$ e si indica con il simbolo

$$df(\boldsymbol{x}_0)\,.$$

Siano $f_1, f_2, \ldots, f_M$ le componenti di f rispetto alla base canonica di $\mathbb{R}^M$, per cui

$$f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_M(\boldsymbol{x}))\,.$$

Teorema 8 *La funzione f è differenziabile in $\boldsymbol{x}_0$ se e solo se lo sono tutte le sue componenti. In tal caso, per ogni vettore $\boldsymbol{h} \in \mathbb{R}^N$ si ha*

$$df(\boldsymbol{x}_0)\boldsymbol{h} = (df_1(\boldsymbol{x}_0)\boldsymbol{h}, df_2(\boldsymbol{x}_0)\boldsymbol{h}, \ldots, df_M(\boldsymbol{x}_0)\boldsymbol{h})\,.$$

Dimostrazione. Considerando le componenti nell'equazione

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})\,,$$

possiamo scrivere

$$f_k\boldsymbol{x}) = f_k(\boldsymbol{x}_0) + \ell_k(\boldsymbol{x} - \boldsymbol{x}_0) + r_k(\boldsymbol{x})\,,$$

con $k = 1, 2, \ldots, M$, e sappiamo che

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \mathbf{0} \quad \Leftrightarrow \quad \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{r_k(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0 \;\text{ per ogni } k = 1, 2, \ldots, M\,,$$

da cui la tesi. ∎

Il teorema precedente permette di ricondurre lo studio del differenziale di una funzione a valori vettoriali a quello delle sue componenti, che sono funzioni a valori scalari.

È utile considerare la matrice associata all'applicazione lineare $\ell = df(\boldsymbol{x}_0)$, data da

$$\begin{pmatrix} \ell_1(\boldsymbol{e}_1) & \ell_1(\boldsymbol{e}_2) & \ldots & \ell_1(\boldsymbol{e}_N) \\ \ell_2(\boldsymbol{e}_1) & \ell_2(\boldsymbol{e}_2) & \ldots & \ell_2(\boldsymbol{e}_N) \\ \vdots & \vdots & & \vdots \\ \ell_M(\boldsymbol{e}_1) & \ell_M(\boldsymbol{e}_2) & \ldots & \ell_M(\boldsymbol{e}_N) \end{pmatrix},$$

dove $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots \boldsymbol{e}_N$ sono i vettori della base canonica di $\mathbb{R}^N$. Tale matrice si chiama "matrice jacobiana" associata alla funzione f nel punto $\boldsymbol{x}_0$ e si denota con $Jf(\boldsymbol{x}_0)$. Ricordando che

$$\frac{\partial f_k}{\partial x_j}(\boldsymbol{x}_0) = df_k(\boldsymbol{x}_0)\boldsymbol{e}_j\,,$$

con $k = 1, 2, \ldots, M$ e $j = 1, 2, \ldots, N$, si ottiene la matrice

$$Jf(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_1}{\partial x_N}(\boldsymbol{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_2}{\partial x_N}(\boldsymbol{x}_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_M}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_M}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_M}{\partial x_N}(\boldsymbol{x}_0) \end{pmatrix}.$$

Studiamo ora la differenziabilità di una funzione composta.

Teorema 9 *Se $f : E \to \mathbb{R}^M$ è differenziabile in $\boldsymbol{x}_0$, E' è un aperto di $\mathbb{R}^M$ contenente $f(E)$ e $g : E' \to \mathbb{R}^L$ è differenziabile in $f(\boldsymbol{x}_0)$, allora $g \circ f$ è differenziabile in $\boldsymbol{x}_0$, e si ha*

$$d(g \circ f)(\boldsymbol{x}_0) = dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0)\,.$$

Dimostrazione. Ponendo $\boldsymbol{y}_0 = f(\boldsymbol{x}_0)$, si ha

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0)+df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)+r_1(\boldsymbol{x})\,, \quad g(\boldsymbol{y}) = g(\boldsymbol{y}_0)+dg(\boldsymbol{y}_0)(\boldsymbol{y}-\boldsymbol{y}_0)+r_2(\boldsymbol{y})\,,$$

con

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{r_1(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} = \boldsymbol{0}\,, \quad \lim_{\boldsymbol{y}\to\boldsymbol{y}_0} \frac{r_2(\boldsymbol{y})}{\|\boldsymbol{y}-\boldsymbol{y}_0\|} = \boldsymbol{0}\,.$$

Introduciamo la funzione $R_2 : E' \to \mathbb{R}^L$ così definita:

$$R_2(\boldsymbol{y}) = \begin{cases} \dfrac{r_2(\boldsymbol{y})}{\|\boldsymbol{y}-\boldsymbol{y}_0\|} & \text{se } \boldsymbol{y} \neq \boldsymbol{y}_0\,, \\ \\ \boldsymbol{0} & \text{se } \boldsymbol{y} = \boldsymbol{y}_0\,. \end{cases}$$

Si noti che R_2 è continua in $\boldsymbol{y}_0$. Allora

$$\begin{aligned} g(f(\boldsymbol{x})) &= g(f(\boldsymbol{x}_0)) + dg(f(\boldsymbol{x}_0))[f(\boldsymbol{x}) - f(\boldsymbol{x}_0)] + r_2(f(\boldsymbol{x})) \\ &= g(f(\boldsymbol{x}_0)) + dg(f(\boldsymbol{x}_0))[df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0) + r_1(\boldsymbol{x})] + r_2(f(\boldsymbol{x})) \\ &= g(f(\boldsymbol{x}_0)) + [dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0)](\boldsymbol{x}-\boldsymbol{x}_0) + r_3(\boldsymbol{x})\,, \end{aligned}$$

dove

$$\begin{aligned} r_3(\boldsymbol{x}) &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + r_2(f(\boldsymbol{x})) \\ &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| R_2(f(\boldsymbol{x})) \\ &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + \|df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0) + r_1(\boldsymbol{x})\| R_2(f(\boldsymbol{x}))\,. \end{aligned}$$

Quindi,

$$\begin{aligned} \frac{\|r_3(\boldsymbol{x})\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \leq & \left\| dg(f(\boldsymbol{x}_0)) \left(\frac{r_1(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \right) \right\| + \\ & + \left(\left\| df(\boldsymbol{x}_0) \left(\frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \right) \right\| + \frac{\|r_1(\boldsymbol{x})\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \right) \|R_2(f(\boldsymbol{x}))\|\,. \end{aligned}$$

Se $\boldsymbol{x} \to \boldsymbol{x}_0$, il primo addendo tende a 0, poiché $dg(f(\boldsymbol{x}_0))$ è continua; f è continua in $\boldsymbol{x}_0$ e R_2 è continua in $\boldsymbol{y}_0 = f(\boldsymbol{x}_0)$ con $R_2(\boldsymbol{y}_0)) = \boldsymbol{0}$, per cui $\|R_2(f(\boldsymbol{x}))\|$ tende a 0; $df(\boldsymbol{x}_0)$, essendo continua, è limitata sull'insieme compatto $\bar{B}(\boldsymbol{0}, 1)$. Quindi, si ha che

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{\|r_3(\boldsymbol{x})\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} = 0\,.$$

Ne segue che $g \circ f$ è differenziabile in x_0 con differenziale $dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0)$. ■

Come noto, la matrice associata alla composizione di due applicazioni lineari è il prodotto delle due matrici corrispondenti. Dal teorema precedente abbiamo quindi la seguente formula per le matrici jacobiane:

$$J(g \circ f)(\boldsymbol{x}_0) = Jg(f(\boldsymbol{x}_0)) \cdot Jf(\boldsymbol{x}_0)\,,$$

ossia

$$\begin{pmatrix} \frac{\partial(g\circ f)_1}{\partial x_1}(\boldsymbol{x}_0) \cdots \frac{\partial(g\circ f)_1}{\partial x_N}(\boldsymbol{x}_0) \\ \vdots \quad \cdots \quad \vdots \\ \frac{\partial(g\circ f)_L}{\partial x_1}(\boldsymbol{x}_0) \cdots \frac{\partial(g\circ f)_L}{\partial x_N}(\boldsymbol{x}_0) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(\boldsymbol{x}_0)) \cdots \frac{\partial g_1}{\partial y_M}(f(\boldsymbol{x}_0)) \\ \vdots \quad \cdots \quad \vdots \\ \frac{\partial g_L}{\partial y_1}(f(\boldsymbol{x}_0)) \cdots \frac{\partial g_L}{\partial y_M}(f(\boldsymbol{x}_0)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0) \cdots \frac{\partial f_1}{\partial x_N}(\boldsymbol{x}_0) \\ \vdots \quad \cdots \quad \vdots \\ \frac{\partial f_M}{\partial x_1}(\boldsymbol{x}_0) \cdots \frac{\partial f_M}{\partial x_N}(\boldsymbol{x}_0) \end{pmatrix}.$$

Ne segue la formula per le derivate parziali:

$$\begin{aligned} &\frac{\partial(g\circ f)_i}{\partial x_j}(\boldsymbol{x}_0) = \\ &= \frac{\partial g_i}{\partial y_1}(f(\boldsymbol{x}_0))\frac{\partial f_1}{\partial x_j}(\boldsymbol{x}_0) + \frac{\partial g_i}{\partial y_2}(f(\boldsymbol{x}_0))\frac{\partial f_2}{\partial x_j}(\boldsymbol{x}_0) + \cdots + \frac{\partial g_i}{\partial y_M}(f(\boldsymbol{x}_0))\frac{\partial f_M}{\partial x_j}(\boldsymbol{x}_0) \\ &= \sum_{k=1}^{M} \frac{\partial g_i}{\partial y_k}(f(\boldsymbol{x}_0))\frac{\partial f_k}{\partial x_j}(\boldsymbol{x}_0)\,, \end{aligned}$$

dove $i = 1, 2, \ldots, L$ e $j = 1, 2, \ldots, N$.

9 Il teorema della funzione implicita - primo enunciato

Il seguente risultato porta il nome di Ulisse Dini.

Teorema 10 *Siano $\Omega \subseteq \mathbb{R} \times \mathbb{R}$ un aperto, $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$ e (x_0, y_0) un punto di Ω per cui si abbia:*

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

Allora esistono un intorno aperto U di x_0, un intorno aperto V di y_0 e una funzione $\eta : U \to V$ di classe $\mathcal{C}^1$ tali che $U \times V \subseteq \Omega$ e, presi $x \in U$ e $y \in V$, si ha:

$$g(x, y) = 0 \quad \Leftrightarrow \quad y = \eta(x)\,.$$

Inoltre, la funzione η è di classe $\mathcal{C}^1$ e vale la formula

$$\eta'(x) = -\frac{\frac{\partial g}{\partial x}(x, \eta(x))}{\frac{\partial g}{\partial y}(x, \eta(x))}\,.$$

La funzione η risulta definita "implicitamente" dall'equazione $g(x, y) = 0$; il suo grafico è l'insieme

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

Dimostrazione. Supponiamo ad esempio $\frac{\partial g}{\partial y}(x_0, y_0) > 0$. Per la proprietà di permanenza del segno, esiste un $\delta > 0$ tale che, se $|x - x_0| \leq \delta$ e $|y - y_0| \leq \delta$, allora $\frac{\partial g}{\partial y}(x, y) > 0$. Quindi, per ogni $x \in [x_0 - \delta, x_0 + \delta]$, la funzione $g(x, \cdot)$ è strettamente crescente su $[y_0 - \delta, y_0 + \delta]$. Essendo $g(x_0, y_0) = 0$, avremo che

$$g(x_0, y_0 - \delta) < 0 < g(x_0, y_0 + \delta)\,.$$

Per la permanenza del segno, esiste un $\delta' > 0$ tale che, se $x \in [x_0 - \delta', x_0 + \delta']$, allora

$$g(x, y_0 - \delta) < 0 < g(x, y_0 + \delta)\,.$$

Definiamo $U =]x_0 - \delta', x_0 + \delta'[$ e $V =]y_0 - \delta, y_0 + \delta[$. Quindi, per ogni $x \in U$, siccome $g(x, \cdot)$ è strettamente crescente, esiste uno ed un solo $y \in]y_0 - \delta, y_0 + \delta[$ per cui $g(x, y) = 0$; chiamo $\eta(x)$ tale y. Resta così definita una funzione $\eta : U \to V$ tale che, presi $x \in U$ e $y \in V$, si ha:

$$g(x, y) = 0 \quad \Leftrightarrow \quad y = \eta(x)\,.$$

Per vedere che η è continua, fissiamo ora un $\bar{x} \in U$ e dimostriamo la continuità in $\bar{x}$. Preso un $x \in U$ e considerata la funzione $\gamma : [0, 1] \to \mathbb{R} \times \mathbb{R}$ definita da

$$\gamma(t) = (\bar{x} + t(x - \bar{x}), \eta(\bar{x}) + t(\eta(x) - \eta(\bar{x}))),$$

applicando il Teorema di Lagrange alla funzione $g \circ \gamma$ si ha che esiste un $\xi \in]0, 1[$ per cui

$$g(x, \eta(x)) - g(\bar{x}, \eta(\bar{x})) = \frac{\partial g}{\partial x}(\gamma(\xi))(x - \bar{x}) + \frac{\partial g}{\partial y}(\gamma(\xi))(\eta(x) - \eta(\bar{x}))\,.$$

Essendo $g(x, \eta(x)) = g(\bar{x}, \eta(\bar{x})) = 0$, si ha che

$$|\eta(x) - \eta(\bar{x})| = \left| \frac{\frac{\partial g}{\partial x}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))} \right| |x - \bar{x}|\,.$$

Siccome le derivate parziali di g sono continue e $\frac{\partial g}{\partial y}$ è non nulla sul compatto $\overline{U} \times \overline{V}$, si ha che $|\frac{\partial g}{\partial x}(\gamma(\xi))(\frac{\partial g}{\partial y}(\gamma(\xi)))^{-1}|$ è limitato superiormente e ne segue la continuità di η in $\bar{x}$. Resta da vedere la derivabilità: procedendo come sopra si ha che

$$\frac{\eta(x) - \eta(\bar{x})}{x - \bar{x}} = -\frac{\frac{\partial g}{\partial x}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))}\,,$$

con $\gamma(\xi)$ appartenente al segmento che congiunge $(\bar{x}, \eta(\bar{x}))$ con $(x, \eta(x))$. Se x tende a $\bar{x}$, si ha che $\gamma(\xi)$ tende a $(\bar{x}, \eta(\bar{x}))$ e quindi

$$\eta'(\bar{x}) = \lim_{x \to \bar{x}} \frac{\eta(x) - \eta(\bar{x})}{x - \bar{x}} = -\frac{\frac{\partial g}{\partial x}(\bar{x}, \eta(\bar{x}))}{\frac{\partial g}{\partial y}(\bar{x}, \eta(\bar{x}))}\,.$$

Ne segue che η è di classe $\mathcal{C}^1$. ∎

Vale naturalmente anche il seguente enunciato simmetrico rispetto al precedente.

Teorema 11 *Siano $\Omega \subseteq \mathbb{R} \times \mathbb{R}$ un aperto, $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$ e (x_0, y_0) un punto di Ω per cui si abbia:*

$$g(x_0, y_0) = 0 \qquad \frac{\partial g}{\partial x}(x_0, y_0) \neq 0\,.$$

Allora esistono un intorno aperto U di x_0, un intorno aperto V di y_0 e una funzione $\eta : V \to U$ di classe $\mathcal{C}^1$ tali che $U \times V \subseteq \Omega$ e, presi $x \in U$ e $y \in V$, si ha:

$$g(x, y) = 0 \quad \Leftrightarrow \quad x = \eta(y)\,.$$

Inoltre, la funzione η è di classe $\mathcal{C}^1$ e vale la formula

$$\eta'(y) = -\frac{\frac{\partial g}{\partial y}(\eta(y), y)}{\frac{\partial g}{\partial x}(\eta(y), y)}\,.$$

10 Il teorema della funzione implicita - caso generale

Vediamo come si generalizza il teorema della funzione implicita. Consideremo un insieme aperto Ω di $\mathbb{R}^M \times \mathbb{R}^N$ e una funzione $g : \Omega \to \mathbb{R}^N$, di classe $\mathcal{C}^1$. Quindi, g ha N componenti

$$g(\boldsymbol{x}, \boldsymbol{y}) = (g_1(\boldsymbol{x}, \boldsymbol{y}), \ldots, g_N(\boldsymbol{x}, \boldsymbol{y}))\,.$$

Qui $\boldsymbol{x} = (x_1, \ldots, x_M) \in \mathbb{R}^M$ e $\boldsymbol{y} = (y_1, \ldots, y_N) \in \mathbb{R}^N$. Useremo la seguente notazione per le matrici jacobiane:

$$\frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\boldsymbol{x}, \boldsymbol{y}) & \cdots & \frac{\partial g_1}{\partial x_M}(\boldsymbol{x}, \boldsymbol{y}) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_N}{\partial x_1}(\boldsymbol{x}, \boldsymbol{y}) & \cdots & \frac{\partial g_N}{\partial x_M}(\boldsymbol{x}, \boldsymbol{y}) \end{pmatrix}, \quad \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\boldsymbol{x}, \boldsymbol{y}) & \cdots & \frac{\partial g_1}{\partial y_N}(\boldsymbol{x}, \boldsymbol{y}) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_N}{\partial y_1}(\boldsymbol{x}, \boldsymbol{y}) & \cdots & \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}, \boldsymbol{y}) \end{pmatrix}.$$

Possiamo ora enunciare il Teorema di Dini in questo caso più generale.

Teorema 12 *Siano $\Omega \subseteq \mathbb{R}^M \times \mathbb{R}^N$ un aperto, $g : \Omega \to \mathbb{R}^N$ una funzione di classe $\mathcal{C}^1$ e $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ un punto di Ω per cui si abbia:*

$$g(\boldsymbol{x}_0, \boldsymbol{y}_0) = \boldsymbol{0}\,, \qquad \det \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0\,.$$

Allora esistono un intorno aperto U di $\boldsymbol{x}_0$, un intorno aperto V di $\boldsymbol{y}_0$ e una funzione $\eta : U \to V$ di classe $\mathcal{C}^1$ tali che $U \times V \subseteq \Omega$ e, presi $\boldsymbol{x} \in U$ e $\boldsymbol{y} \in V$, si ha:

$$g(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0} \quad \Leftrightarrow \quad \boldsymbol{y} = \eta(\boldsymbol{x})\,.$$

Inoltre, la funzione η è di classe $\mathcal{C}^1$ e vale la formula

$$J\eta(\boldsymbol{x}) = -\left(\frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}, \eta(\boldsymbol{x}))\right)^{-1} \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x}))\,.$$

Dimostrazione. [2] Faremo la dimostrazione per induzione su N.

Nel caso $N = 1$ e $M \geq 2$, si procede in modo del tutto analogo a quanto già fatto nel caso $M = 1$. Basterà prendere, al posto dell'intervallo $[x_0-\delta, x_0+\delta]$, la palla chiusa $\overline{B}(\boldsymbol{x}_0, \delta)$, e similmente per gli intorni aperti di $\boldsymbol{x}_0$, per dimostrare l'esistenza e la continuità della funzione η. Resta da vedere la derivabilità: considerato $\bar{\boldsymbol{x}} = (\bar{x}_1, \dots, \bar{x}_M)$, prendiamo ora $\boldsymbol{x} = (\bar{x}_1+h, \dots, \bar{x}_M)$; procedendo come in precedenza, si ha che

$$\frac{\eta(\bar{x}_1 + h, \dots, \bar{x}_M) - \eta(\bar{x}_1, \dots, \bar{x}_M)}{h} = -\frac{\frac{\partial g}{\partial x_1}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))},$$

con $\gamma(\xi)$ appartenente al segmento che congiunge $(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))$ con $(\boldsymbol{x}, \eta(\boldsymbol{x}))$. Se h tende a 0, si ha che $\gamma(\xi)$ tende a $(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))$ e quindi

$$\frac{\partial \eta}{\partial x_1}(\bar{\boldsymbol{x}}) = \lim_{h \to 0} \frac{\eta(\bar{x}_1 + h, \dots, \bar{x}_M) - \eta(\bar{x}_1, \dots, \bar{x}_M)}{h} = -\frac{\frac{\partial g}{\partial x_1}(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))}{\frac{\partial g}{\partial y}(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))}.$$

Analogamente si calcolano le derivate parziali rispetto a $x_2, \dots, x_M$, per cui si vede che η è di classe $\mathcal{C}^1$ e

$$J\eta(\boldsymbol{x}) = -\frac{1}{\frac{\partial g}{\partial y}(\boldsymbol{x}, \eta(\boldsymbol{x}))} \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x})).$$

Supponiamo ora l'enunciato valido fino a $N - 1$, per un certo $N \geq 2$ (e $M \geq 1$ qualsiasi) e dimostriamo che vale anche per N. Useremo la notazione

$$\tilde{\boldsymbol{y}}_1 = (y_1, \dots, y_{N-1}),$$

per cui scriveremo $y = (\tilde{\boldsymbol{y}}_1, y_N)$. Siccome

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) & \cdots & \frac{\partial g_1}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_N}{\partial y_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) & \cdots & \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \end{pmatrix} \neq 0,$$

almeno uno degli elementi dell'ultima colonna è non nullo. Possiamo supporre senza perdita di generalità, eventualmente permutando le righe, che sia $\frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0$. Scrivendo $\boldsymbol{y}_0 = (\tilde{\boldsymbol{y}}_1^0, y_N^0)$, con $\tilde{\boldsymbol{y}}_1^0 = (y_1^0, \dots, y_{N-1}^0)$, sarà

$$g_N(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0, y_N^0) = 0, \qquad \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0, y_N^0) \neq 0.$$

Allora (caso unidimensionale) esistono un intorno aperto U_1 di $(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0)$, un intorno aperto V_N di y_N^0 e una funzione $\eta_1 : U_1 \to V_N$ di classe $\mathcal{C}^1$ tali che $U_1 \times V_N \subseteq \Omega$, per cui si abbia: se $(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \in U_1$ e $y_N \in V_N$,

$$g_N(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = 0 \quad \Leftrightarrow \quad y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1),$$

[2] non fatta a lezione

e

$$J\eta_1(\boldsymbol{x},\tilde{\boldsymbol{y}}_1) = -\frac{1}{\frac{\partial g_N}{\partial y_N}(\boldsymbol{x},\tilde{\boldsymbol{y}}_1,\eta_1(\boldsymbol{x},\tilde{\boldsymbol{y}}_1)))}\,\frac{\partial g_N}{\partial(\boldsymbol{x},\tilde{\boldsymbol{y}}_1)}(\boldsymbol{x},\tilde{\boldsymbol{y}}_1,\eta_1(\boldsymbol{x},\tilde{\boldsymbol{y}}_1))\,.$$

Possiamo supporre U_1 della forma $\tilde{U}\times\tilde{V}_1$, con $\tilde{U}$ intorno aperto di $\boldsymbol{x}_0$ e $\tilde{V}_1$ intorno aperto di $\tilde{\boldsymbol{y}}_1^0$. Definiamo la funzione $\phi : \tilde{U}\times\tilde{V}_1 \to \mathbb{R}^{N-1}$, ponendo

$$\phi(\boldsymbol{x},\tilde{\boldsymbol{y}}_1) = (g_1(\boldsymbol{x},\tilde{\boldsymbol{y}}_1,\eta_1(\boldsymbol{x},\tilde{\boldsymbol{y}}_1)),\dots,g_{N-1}(\boldsymbol{x},\tilde{\boldsymbol{y}}_1,\eta_1(\boldsymbol{x},\tilde{\boldsymbol{y}}_1)))\,.$$

Per brevità, scriveremo

$$g_{(1,\dots,N-1)}(\boldsymbol{x},\boldsymbol{y}) = (g_1(\boldsymbol{x},\boldsymbol{y}),\dots,g_{N-1}(\boldsymbol{x},\boldsymbol{y}))\,.$$

Notiamo che $\phi(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0)=0$ e che, essendo $\eta_1(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0)=y_N^0$,

$$\frac{\partial\phi}{\partial\tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0) = \frac{\partial g_{(1,\dots,N-1)}}{\partial\tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0,\boldsymbol{y}_0) + \frac{\partial g_{(1,\dots,N-1)}}{\partial y_N}(\boldsymbol{x}_0,\boldsymbol{y}_0)\frac{\partial\eta_1}{\partial\tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0)\,. \qquad (*)$$

Inoltre, siccome $g_N(\boldsymbol{x},\tilde{\boldsymbol{y}}_1,\eta_1(\boldsymbol{x},\tilde{\boldsymbol{y}}_1))=0$, per ogni $(\boldsymbol{x},\tilde{\boldsymbol{y}}_1)\in U_1$, differenziando si ha:

$$0 = \frac{\partial g_N}{\partial\tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0,\boldsymbol{y}_0) + \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0,\boldsymbol{y}_0)\frac{\partial\eta_1}{\partial\tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0)\,. \qquad (**)$$

Scriviamo

$$\det\frac{\partial\phi}{\partial\tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0) = \frac{1}{\frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0,\boldsymbol{y}_0)}\det\left(\begin{array}{c|c}\frac{\partial\phi}{\partial\tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0) & \frac{\partial g_{(1,\dots,N-1)}}{\partial y_N}(\boldsymbol{x}_0,\boldsymbol{y}_0)\\ \hline 0 & \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0,\boldsymbol{y}_0)\end{array}\right),$$

avendo usato la notazione di matrice suddivisa a blocchi. Sostituendo le due uguaglianze $(*)$, $(**)$ e usando le proprietà dei determinanti, si ha:

$$\begin{aligned}
&\det\left(\begin{array}{c|c}\frac{\partial\phi}{\partial\tilde{y}_1}(x_0,\tilde{y}_1^0) & \frac{\partial g_{(1,\dots,N-1)}}{\partial y_N}(x_0,y_0)\\ \hline 0 & \frac{\partial g_N}{\partial y_N}(x_0,y_0)\end{array}\right) =\\
&= \det\left(\begin{array}{c|c}\frac{\partial g_{(1,\dots,N-1)}}{\partial\tilde{y}_1}(x_0,y_0)+\frac{\partial g_{(1,\dots,N-1)}}{\partial y_N}(x_0,y_0)\frac{\partial\eta_1}{\partial\tilde{y}_1}(x_0,\tilde{y}_1^0) & \frac{\partial g_{(1,\dots,N-1)}}{\partial y_N}(x_0,y_0)\\ \hline \frac{\partial g_N}{\partial\tilde{y}_1}(x_0,y_0)+\frac{\partial g_N}{\partial y_N}(x_0,y_0)\frac{\partial\eta_1}{\partial\tilde{y}_1}(x_0,\tilde{y}_1^0) & \frac{\partial g_N}{\partial y_N}(x_0,y_0)\end{array}\right)\\
&= \det\left(\frac{\partial g}{\partial\tilde{y}_1}(x_0,y_0)+\frac{\partial g}{\partial y_N}(x_0,y_0)\frac{\partial\eta_1}{\partial\tilde{y}_1}(x_0,\tilde{y}_1^0)\;\middle|\;\frac{\partial g}{\partial y_N}(x_0,y_0)\right)\\
&= \det\left(\frac{\partial g}{\partial\tilde{y}_1}(x_0,y_0)\;\middle|\;\frac{\partial g}{\partial y_N}(x_0,y_0)\right) = \det\frac{\partial g}{\partial y}(x_0,y_0)\neq 0\,.
\end{aligned}$$

Abbiamo quindi

$$\phi(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0)=\boldsymbol{0}\,,\qquad \det\frac{\partial\phi}{\partial\tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0,\tilde{\boldsymbol{y}}_1^0)\neq 0\,.$$

Per l'ipotesi induttiva, esistono un intorno aperto U di $\boldsymbol{x}_0$, un intorno aperto V_1 di $\tilde{\boldsymbol{y}}_1^0$ e una funzione $\eta_2 : U \to V_1$ di classe $\mathcal{C}^1$ tali che $U \times V_1 \subseteq \tilde{U} \times \tilde{V}_1$, per cui si abbia: per ogni $\boldsymbol{x} \in U$ e $\tilde{\boldsymbol{y}}_1 \in V_1$,

$$\phi(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = \mathbf{0} \quad \Leftrightarrow \quad \tilde{\boldsymbol{y}}_1 = \eta_2(\boldsymbol{x}) .$$

In conclusione, per $\boldsymbol{x} \in U$ e $\boldsymbol{y} = (\tilde{\boldsymbol{y}}_1, y_N) \in V_1 \times V_2$, si ha:

$$\begin{aligned}
g(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0} \quad & \Leftrightarrow \quad \begin{cases} g_{(1,\dots,N-1)}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = \mathbf{0} \\ g_N(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = 0 \end{cases} \\
& \Leftrightarrow \quad \begin{cases} g_{(1,\dots,N-1)}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = \mathbf{0} \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
& \Leftrightarrow \quad \begin{cases} \phi(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = \mathbf{0} \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
& \Leftrightarrow \quad \begin{cases} \tilde{\boldsymbol{y}}_1 = \eta_2(\boldsymbol{x}) \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
& \Leftrightarrow \quad \boldsymbol{y} = (\eta_2(\boldsymbol{x}), \eta_1(\boldsymbol{x}, \eta_2(\boldsymbol{x}))) .
\end{aligned}$$

Ponendo $V = V_1 \times V_2$, resta pertanto definita la funzione $\eta : U \to V$:

$$\eta(\boldsymbol{x}) = (\eta_2(\boldsymbol{x}), \eta_1(\boldsymbol{x}, \eta_2(\boldsymbol{x}))) .$$

Tale funzione è di classe $\mathcal{C}^1$, siccome lo sono sia η_1 che η_2. Siccome $g(\boldsymbol{x}, \eta(\boldsymbol{x})) = 0$ per ogni $\boldsymbol{x} \in U$, se ne deduce che che

$$\frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x})) + \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}, \eta(\boldsymbol{x})) J\eta(\boldsymbol{x}) = \mathbf{0} ,$$

da cui la formula per $J\eta(\boldsymbol{x})$. ■

Ed ecco l'enunciato simmetrico.

Teorema 13 *Siano $\Omega \subseteq \mathbb{R}^M \times \mathbb{R}^N$ un aperto, $g : \Omega \to \mathbb{R}^M$ una funzione di classe $\mathcal{C}^1$ e $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ un punto di Ω per cui si abbia:*

$$g(\boldsymbol{x}_0, \boldsymbol{y}_0) = \mathbf{0} , \qquad \det \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0 .$$

Allora esistono un intorno aperto U di $\boldsymbol{x}_0$, un intorno aperto V di $\boldsymbol{y}_0$ e una funzione $\eta : V \to U$ di classe $\mathcal{C}^1$ tali che $U \times V \subseteq \Omega$ e, presi $\boldsymbol{x} \in U$ e $\boldsymbol{y} \in V$, si ha:

$$g(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0} \quad \Leftrightarrow \quad \boldsymbol{x} = \eta(\boldsymbol{y}) .$$

Inoltre, la funzione η è di classe $\mathcal{C}^1$ e vale la formula

$$J\eta(\boldsymbol{y}) = -\left(\frac{\partial g}{\partial \boldsymbol{x}}(\eta(\boldsymbol{y})), \boldsymbol{y} \right)^{-1} \frac{\partial g}{\partial \boldsymbol{y}}(\eta(\boldsymbol{y})), \boldsymbol{y}) .$$

Vediamo ora un'importante conseguenza del teorema della funzione implicita.

Definizione 4 *Dati A e B, due aperti di $\mathbb{R}^N$, una funzione $\varphi : A \to B$ sè un "diffeomorfismo" se è di classe $\mathcal{C}^1$, biiettiva e la sua inversa $\varphi^{-1} : B \to A$ è anch'essa di classe $\mathcal{C}^1$.*

Enunciamo il **teorema di inversione locale**.

Teorema 14 *Siano A e B due aperti di $\mathbb{R}^N$ e $\varphi : A \to B$ una funzione di classe $\mathcal{C}^1$. Se per un certo $\boldsymbol{x}_0 \in A$ si ha che $\det J\varphi(\boldsymbol{x}_0) \neq 0$, allora esistono un intorno aperto U di $\boldsymbol{x}_0$ contenuto in A, e un intorno aperto V di $\varphi(\boldsymbol{x}_0)$ contenuto in B, tali che la restrizione $\varphi_{|U} : U \to V$ sia un diffeomorfismo.*

Dimostrazione. Consideriamo la funzione $g : A \times B \to \mathbb{R}^N$ definita da

$$g(\boldsymbol{x}, \boldsymbol{y}) = \varphi(\boldsymbol{x}) - \boldsymbol{y}\,.$$

Posto $\boldsymbol{y}_0 = \varphi(\boldsymbol{x}_0)$, si ha che

$$g(\boldsymbol{x}_0, \boldsymbol{y}_0) = \boldsymbol{0}\,, \qquad \text{e} \qquad \det \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}_0, \boldsymbol{y}_0) = \det J\varphi(\boldsymbol{x}_0) \neq 0\,.$$

Per il teorema della funzione implicita, esistono un intorno aperto V di $\boldsymbol{y}_0$, un intorno aperto U di $\boldsymbol{x}_0$ e una funzione $\eta : V \to U$ di classe $\mathcal{C}^1$ tali che, presi $\boldsymbol{y} \in V$ e $\boldsymbol{x} \in U$, si ha:

$$\varphi(\boldsymbol{x}) = \boldsymbol{y} \quad \Leftrightarrow \quad g(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0} \quad \Leftrightarrow \quad \boldsymbol{x} = \eta(\boldsymbol{y})\,.$$

Quindi, $\eta = \varphi_{|U}^{-1}$ e la dimostrazione è così completa. ■

11 Le M-superfici

Indichiamo con I un rettangolo di $\mathbb{R}^M$, dove $1 \leq M \leq N$.

Definizione 5 *Chiameremo* **M-superficie in $\mathbb{R}^N$** *una funzione $\sigma : I \to \mathbb{R}^N$ di classe $\mathcal{C}^1$. Se $M = 1$, σ si dirà anche* **curva***; se $M = 2$, si dirà semplicemente* **superficie***. L'insieme $\sigma(I)$ è detto* **supporto** *della M-superficie σ. Diremo che la M-superficie σ è* **regolare** *se, per ogni $\boldsymbol{u} \in \mathring{I}$, la matrice jacobiana $\sigma'(\boldsymbol{u})$ ha rango M.*

Consideriamo da vicino il caso $N = 3$. Una curva in $\mathbb{R}^3$ è una funzione $\sigma : [a, b] \to \mathbb{R}^3$, $\sigma = (\sigma_1, \sigma_2, \sigma_3)$. La curva è regolare se, per ogni $t \in]a, b[$, il vettore $\sigma'(t) = (\sigma_1'(t), \sigma_2'(t), \sigma_3'(t))$ è non nullo. In tal caso, si definisce il seguente **versore tangente** nel punto $\sigma(t)$:

$$\tau_\sigma(t) = \frac{\sigma'(t)}{||\sigma'(t)||}\,.$$

Esempio. La curva $\sigma : [0, 2\pi] \to \mathbb{R}^3$ definita da

$$\sigma(t) = (R\cos(2t), R\sin(2t), 0)$$

ha come supporto la circonferenza

$$\{(x,y,z) : x^2 + y^2 = R^2, z = 0\}$$

(che viene percorsa due volte). Essendo $\sigma'(t) = (-2R\sin(2t), 2R\cos(2t), 0)$, si tratta di una curva regolare, e si ha:

$$\tau_\sigma(t) = (-\sin(2t), \cos(2t), 0)\,.$$

Una superficie in $\mathbb{R}^3$ è una funzione $\sigma : [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}^3$. La superficie è regolare se, per ogni $(u, v) \in \,]a_1, b_1[\, \times \,]a_2, b_2[\,$, i vettori $\frac{\partial\sigma}{\partial u}(u, v)$, $\frac{\partial\sigma}{\partial v}(u, v)$ sono linearmente indipendenti. In tal caso, essi individuano un piano, detto **piano tangente** alla superficie nel punto $\sigma(u, v)$, e si definisce il seguente **versore normale**:

$$\nu_\sigma(u,v) = \frac{\frac{\partial\sigma}{\partial u}(u,v) \times \frac{\partial\sigma}{\partial v}(u,v)}{||\frac{\partial\sigma}{\partial u}(u,v) \times \frac{\partial\sigma}{\partial v}(u,v)||}\,.$$

Esempi. 1. La superficie $\sigma : [0, \pi] \times [0, \pi] \to \mathbb{R}^3$ definita da

$$\sigma(\phi, \theta) = (R\sin\phi\cos\theta, R\sin\phi\sin\theta, R\cos\phi)$$

ha come supporto la semisfera

$$\{(x,y,z) : x^2 + y^2 + z^2 = R^2, y \geq 0\}\,.$$

Essendo

$$\frac{\partial\sigma}{\partial\phi}(\phi,\theta) = (R\cos\phi\cos\theta, R\cos\phi\sin\theta, -R\sin\phi)\,,$$
$$\frac{\partial\sigma}{\partial\theta}(\phi,\theta) = (-R\sin\phi\sin\theta, R\sin\phi\cos\theta, 0)\,,$$

si tratta di una superficie regolare, e si ha:

$$\nu_\sigma(\phi,\theta) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)\,.$$

2. La superficie $\sigma : [r, R] \times [0, 2\pi] \to \mathbb{R}^3$, con $0 \leq r < R$, data da

$$\sigma(u,v) = (u\cos v, u\sin v, 0)\,,$$

ha come supporto un cerchio se $r = 0$, una corona circolare se $r > 0$. È una superficie regolare.

3. La superficie $\sigma : [r, R] \times [0, 2\pi] \to \mathbb{R}^3$, con $0 < r < R$, definita da

$$\begin{aligned}\sigma(u,v) = \Bigg(&\left(\frac{r+R}{2} + \left(u - \frac{r+R}{2}\right)\cos\left(\frac{v}{2}\right)\right)\cos v, \\ &\left(\frac{r+R}{2} + \left(u - \frac{r+R}{2}\right)\cos\left(\frac{v}{2}\right)\right)\sin v, \\ &\left(u - \frac{r+R}{2}\right)\sin\left(\frac{v}{2}\right)\Bigg)\,,\end{aligned}$$

ha comes supporto un nastro di Möbius. È anch'essa una superficie regolare.

4. La superficie $\sigma : [0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3$ definita da

$$\sigma(u, v) = ((R + r \cos u) \cos v, (R + r \cos u) \sin v, r \sin u)$$

dove $0 < r < R$, ha come supporto l'anello toroidale o "toro"

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

Si può verificare che anche in questo caso si tratta di una superficie regolare.

Una 3-superficie in $\mathbb{R}^3$ si dice anche **volume**.

Esempio. La funzione $\sigma : [0, R] \times [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3$ definita da

$$\sigma(\rho, \phi, \theta) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi)$$

ha come supporto la palla chiusa

$$\{(x, y, z) : x^2 + y^2 + z^2 \leq R^2\} .$$

In questo caso, $\det \sigma'(\rho, \phi, \theta) = \rho^2 \sin \phi$ e pertanto si tratta di un volume regolare.

12 Analisi locale delle M-superfici

Sia $1 \leq M < N$. Identificando $\mathbb{R}^N$ con $\mathbb{R}^M \times \mathbb{R}^{N-M}$, ogni vettore $\boldsymbol{x} = (x_1, \ldots, x_N)$ di $\mathbb{R}^N$ si scriverà nella forma $\boldsymbol{x} = (\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}})$, con $\hat{\boldsymbol{x}} = (x_1, \ldots, x_M)$ e $\tilde{\boldsymbol{x}} = (x_{M+1}, \ldots, x_N)$.

Useremo inoltre la seguente notazione: dato $\hat{\boldsymbol{x}} = (x_1, \ldots, x_M) \in \mathbb{R}^M$ e $r > 0$,

$$B[\hat{\boldsymbol{x}}, r] = [x_1 - r, x_1 + r] \times \cdots \times [x_M - r, x_M + r] \subseteq \mathbb{R}^M .$$

Per semplicità, scriveremo $B[r]$ invece di $B[\mathbf{0}, r]$.

Teorema 15 *Siano Ω un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di Ω e $g : \Omega \to \mathbb{R}^{N-M}$ una funzione di classe $\mathcal{C}^1$, tale che*

$$g(\boldsymbol{x}_0) = \mathbf{0} , \quad e \quad Jg(\boldsymbol{x}_0) \text{ ha rango } N - M.$$

Allora esistono un intorno U di $\boldsymbol{x}_0$ e una M-superficie regolare e iniettiva $\sigma : B[r] \to \mathbb{R}^N$, per un certo $r > 0$, tali che $\sigma(\mathbf{0}) = \boldsymbol{x}_0$ e

$$\{\boldsymbol{x} \in U : g(\boldsymbol{x}) = \mathbf{0}\} = \sigma(\mathbf{B}[\mathbf{r}]) .$$

Dimostrazione. Supponiamo, per esempio, che sia invertibile

$$\frac{\partial g}{\partial \tilde{\boldsymbol{x}}}(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial g_1}{\partial p_{M+1}}(\boldsymbol{x}_0) & \cdots & \frac{\partial g_1}{\partial p_N}(\boldsymbol{x}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_{N-M}}{\partial p_{M+1}}(\boldsymbol{x}_0) & \cdots & \frac{\partial g_{N-M}}{\partial p_N}(\boldsymbol{x}_0) \end{pmatrix} .$$

Per il teorema della funzione implicita, esistono un intorno aperto $\hat{U}$ di $\hat{\boldsymbol{x}}_0$, un intorno aperto $\widetilde{U}$ di $\tilde{\boldsymbol{x}}_0$ e una funzione $\eta : \hat{U} \to \widetilde{U}$, tali che $\hat{U} \times \widetilde{U} \subseteq \Omega$ e, se $\hat{\boldsymbol{x}} \in \hat{U}$ e $\tilde{\boldsymbol{x}} \in \widetilde{U}$, si ha:

$$g(\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) = \mathbf{0} \quad \Leftrightarrow \quad \tilde{\boldsymbol{x}} = \eta(\hat{\boldsymbol{x}}) \, .$$

Preso $r > 0$ tale che $B[\hat{\boldsymbol{x}}_0, r] \subseteq \hat{U}$, sia $U = B[\hat{\boldsymbol{x}}_0, r] \times \widetilde{U}$ e $\sigma : B[r] \to \mathbb{R}^N$ definita da $\sigma(\boldsymbol{u}) = (\boldsymbol{u} + \hat{\boldsymbol{x}}_0, \eta(\boldsymbol{u} + \hat{\boldsymbol{x}}_0))$. Si verifica che la matrice jacobiana $J\sigma(\boldsymbol{u})$ ha come sottomatrice la matrice $M \times M$ identità, per cui σ è regolare. Si vede facilmente che σ è iniettiva, in quanto lo è la prima componente $\boldsymbol{u} \mapsto \boldsymbol{u} + \hat{\boldsymbol{x}}_0$. Inoltre, se $\boldsymbol{x} = (\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) \in U$,

$$g(\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) = \mathbf{0} \quad \Leftrightarrow \quad \tilde{\boldsymbol{x}} = \eta(\hat{\boldsymbol{x}}) \quad \Leftrightarrow \quad (\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}}) = \sigma(\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}_{\mathbf{0}}) \, ,$$

da cui la tesi. Nel caso in cui la sottomatrice considerata non sia invertibile, basterà operare degli scambi nelle colonne della matrice $Jg(\boldsymbol{x}_0)$ per ricondursi alla situazione precedente. ■

La M-superficie σ individuata dal teorema precedente è detta “M-parametrizzazione locale”.

Vediamo tre casi di particolare interesse. Iniziamo con una curva in $\mathbb{R}^2$ (caso $M = 1$, $N = 2$).

Corollario 1 *Siano Ω un sottoinsieme aperto di $\mathbb{R}^2$, (x_0, y_0) un punto di Ω e $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$, tale che*

$$g(x_0, y_0) = 0 \qquad e \qquad \nabla g(x_0, y_0) \neq 0 \, .$$

Allora esistono un intorno U di (x_0, y_0) e una curva regolare e iniettiva $\sigma : [-r, r] \to \mathbb{R}^2$, per un certo $r > 0$, tali che $\sigma(0) = (x_0, y_0)$ e

$$\{(x, y) \in U : g(x, y) = 0\} = \sigma([-r, r]) \, .$$

Vediamo ora il caso di una superficie in $\mathbb{R}^3$ (caso $M = 2$, $N = 3$).

Corollario 2 *Siano Ω un sottoinsieme aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω e $g : \Omega \to \mathbb{R}$ una funzione di classe $\mathcal{C}^1$, tale che*

$$g(x_0, y_0, z_0) = 0 \qquad e \qquad \nabla g(x_0, y_0, z_0) \neq 0 \, .$$

Allora esistono un intorno U di (x_0, y_0, z_0) e una superficie regolare e iniettiva $\sigma : [-r, r] \times [-r, r] \to \mathbb{R}^3$, per un certo $r > 0$, tali che $\sigma(0, 0) = (x_0, y_0, z_0)$ e

$$\{(x, y, z) \in U : g(x, y, z) = 0\} = \sigma([-r, r] \times [-r, r]) \, .$$

Infine, vediamo il caso di una curva in $\mathbb{R}^3$ (caso $M = 1$, $N = 3$).

Corollario 3 *Siano Ω un sottoinsieme aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω e $g_1, g_2 : \Omega \to \mathbb{R}$ due funzioni di classe $\mathcal{C}^1$, tali che*

$$g_1(x_0, y_0, z_0) = g_2(x_0, y_0, z_0) = 0 \qquad e \qquad \nabla g_1(x_0, y_0, z_0) \times \nabla g_2(x_0, y_0, z_0) \neq 0 \, .$$

Allora esistono un intorno U di (x_0, y_0, z_0) e una curva regolare e iniettiva $\sigma : [-r, r] \to \mathbb{R}^3$, per un certo $r > 0$, tali che $\sigma(0) = (x_0, y_0, z_0)$ e

$$\{(x, y, z) \in U : g_1(x, y, z) = g_2(x, y, z) = 0\} = \sigma([-r, r]) \, .$$

12.1 I moltiplicatori di Lagrange

Siano Ω un aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di Ω e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in $\boldsymbol{x}_0$. Vogliamo cercare eventuali punti di massimo o di minimo per la funzione ristretta a un “vincolo”, che sarà descritto da un'altra funzione, in generale a valori vettoriali.

Teorema 16 *Sia $g : \Omega \to \mathbb{R}^{N-M}$ una funzione di classe $\mathcal{C}^1$ tale che*

$$g(\boldsymbol{x}_0) = \boldsymbol{0}\,, \quad e \quad Jg(\boldsymbol{x}_0) \text{ ha rango } N - M.$$

Posto

$$S = \{\boldsymbol{x} \in \Omega : g(\boldsymbol{x}) = \boldsymbol{0}\}\,,$$

se $\boldsymbol{x}_0$ è un punto di minimo o massimo locale per $f|_S$, allora esistono $(N-M)$ numeri reali $\lambda_1, \dots, \lambda_{N-M}$ tali che

$$\nabla f(\boldsymbol{x}_0) = \sum_{j=1}^{N-M} \lambda_j \nabla g_j(\boldsymbol{x}_0)\,.$$

I numeri $\lambda_1, \dots, \lambda_{N-M}$ si chiamano **moltiplicatori di Lagrange**.

Dimostrazione. Per il teorema precedente, esistono un intorno U di $\boldsymbol{x}_0$, un $r > 0$ e una M-superficie regolare $\sigma : B[r] \to \mathbb{R}^N$ tali che $\sigma(\boldsymbol{0}) = \boldsymbol{x}_0$ e

$$S \cap U = \sigma(B[r])\,.$$

Considerata la funzione $F : B[r] \to \mathbb{R}$, definita da $F(\boldsymbol{u}) = f(\sigma(\boldsymbol{u}))$, si ha che $\boldsymbol{0}$ è un punto di minimo o massimo locale per F. Quindi, $\nabla F(\boldsymbol{0}) = \boldsymbol{0}$, per cui

$$0 = JF(\boldsymbol{0}) = Jf(\boldsymbol{x}_0)J\sigma(\boldsymbol{0})\,.$$

Ne segue che

$$\nabla f(\boldsymbol{x}_0) \ \text{ è ortogonale a } \ \frac{\partial \sigma}{\partial u_1}(\boldsymbol{0}), \dots, \frac{\partial \sigma}{\partial u_M}(\boldsymbol{0})\,.$$

Inoltre, essendo $g(\sigma(\boldsymbol{u})) = \boldsymbol{0}$ per ogni $\boldsymbol{u} \in B[r]$, si ha che

$$Jg(\boldsymbol{x}_0)J\sigma(\boldsymbol{0}) = 0\,.$$

Quindi anche i vettori

$$\nabla g_1(\boldsymbol{x}_0), \dots, \nabla g_{N-M}(\boldsymbol{x}_0) \ \text{ sono tutti ortogonali a } \ \frac{\partial \sigma}{\partial u_1}(\boldsymbol{0}), \dots, \frac{\partial \sigma}{\partial u_M}(\boldsymbol{0})\,.$$

Siccome σ è regolare,

$$\text{lo spazio vettoriale } \mathcal{T} \text{ generato da } \frac{\partial \sigma}{\partial u_1}(\boldsymbol{0}), \dots, \frac{\partial \sigma}{\partial u_M}(\boldsymbol{0}) \text{ ha dimensione } M\,.$$

Quindi lo spazio ortogonale $\mathcal{T}^\perp$ ha dimesnione $N - M$. E siccome, come abbiamo visto,

$$\nabla f(\boldsymbol{x}_0), \nabla g_1(\boldsymbol{x}_0), \dots, \nabla g_{N-M}(\boldsymbol{x}_0) \in \mathcal{T}^\perp\,,$$

questi vettori devono essere linearmente dipendenti. Quindi, essendo che i vettori $\nabla g_1(\boldsymbol{x}_0)$, ..., $\nabla g_{N-M}(\boldsymbol{x}_0)$ linearmente indipendenti, ne segue che $\nabla f(\boldsymbol{x}_0)$ si deve poter esprimere come una loro combinazione lineare. ■

Vediamo anche qui tre casi particolari interessanti.

Corollario 4 *Siano* Ω *un aperto di* $\mathbb{R}^2$, (x_0, y_0) *un punto di* Ω, $g : \Omega \to \mathbb{R}$ *una funzione di classe* $\mathcal{C}^1$ *tale che*

$$g(x_0, y_0) = 0 \qquad e \qquad \nabla g(x_0, y_0) \neq 0,$$

e $f : \Omega \to \mathbb{R}$ *una funzione differenziabile in* (x_0, y_0)*. Posto*

$$S = \{(x, y) \in \Omega : g(x, y) = 0\},$$

se (x_0, y_0) *è un punto di minimo o massimo locale per* $f|_S$*, allora esiste un numero reale* λ *tale che*

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0).$$

Corollario 5 *Siano* Ω *un aperto di* $\mathbb{R}^3$, (x_0, y_0, z_0) *un punto di* Ω, $g : \Omega \to \mathbb{R}$ *una funzione di classe* $\mathcal{C}^1$ *tale che*

$$g(x_0, y_0, z_0) = 0 \qquad e \qquad \nabla g(x_0, y_0, z_0) \neq 0,$$

e $f : \Omega \to \mathbb{R}$ *una funzione differenziabile in* (x_0, y_0, z_0)*. Posto*

$$S = \{(x, y, z) \in \Omega : g(x, y, z) = 0\},$$

se (x_0, y_0, z_0) *è un punto di minimo o massimo locale per* $f|_S$*, allora esiste un numero reale* λ *tale che*

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0).$$

Corollario 6 *Siano* Ω *un aperto di* $\mathbb{R}^3$, (x_0, y_0, z_0) *un punto di* Ω, $g_1, g_2 : \Omega \to \mathbb{R}$ *due funzioni di classe* $\mathcal{C}^1$ *tali che*

$$g_1(x_0, y_0, z_0) = g_2(x_0, y_0, z_0) = 0 \qquad e \qquad \nabla g_1(x_0, y_0, z_0) \times \nabla g_2(x_0, y_0, z_0) \neq 0,$$

e $f : \Omega \to \mathbb{R}$ *una funzione differenziabile in* (x_0, y_0, z_0)*. Posto*

$$S = \{(x, y, z) \in U : g_1(x, y, z) = 0,\ g_2(x, y, z) = 0\},$$

se (x_0, y_0, z_0) *è un punto di minimo o massimo locale per* $f|_S$*, allora esistono due numeri reali* λ_1, λ_2 *tali che*

$$\nabla f(x_0, y_0, z_0) = \lambda_1 \nabla g_1(x_0, y_0, z_0) + \lambda_2 \nabla g_2(x_0, y_0, z_0).$$