

Note del corso di Analisi matematica 2 per il corso di laurea in ingegneria industriale a.a. 2015-2016

Le presenti note sono una sintetica descrizione degli argomenti svolti a lezione. Non sostituiscono, bensì integrano gli appunti ed il libro di testo. Si ringraziano Emanuele Bottazzi ed Enrico Pagani per interessanti discussioni e suggerimenti. Si ringrazia inoltre lo studente Muhammad Masood Asghar per l'aiuto nella correzione del testo relativo all'anno accademico 2014-15.

Funzioni a valori vettoriali

Definizione 1 Sia $I \subset \mathbb{R}$ un intervallo della retta reale. Una funzione $f: I \rightarrow \mathbb{R}^n$ è una mappa che ad ogni elemento $t \in I$ associa un unico vettore $f(t) \in \mathbb{R}^n$.

L'insieme $I \subset \mathbb{R}$ è detto *dominio* della funzione f , mentre l'insieme $f(I) := \{x \in \mathbb{R}^n : x = f(t), t \in I\}$ è detto *immagine* di f .

Le funzioni $f_i: I \rightarrow \mathbb{R}$, con $i = 1, \dots, n$, definite da $f_i(t) := f(t) \cdot \hat{e}_i$, dove \hat{e}_i è l' i -esimo vettore della base canonica di \mathbb{R}^n , sono dette *componenti* di f .

Esempio 1 Si consideri la funzione $f: \mathbb{R} \rightarrow \mathbb{R}^4$ definita da

$$f(t) := (t^2, \cos t, \sin t, \exp(t)), \quad t \in \mathbb{R}.$$

Le componenti di f sono le funzioni $f_i: \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, \dots, 4$, definite da:

$$f_1(t) = t^2, \quad f_2(t) = \cos t, \quad f_3(t) = \sin t, \quad f_4(t) = \exp(t).$$

Definizione 2 Sia $f : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ e sia $t_0 \in I$ un punto interno all'intervallo I (cioé tale che esiste $\delta > 0$ tale per cui $(t_0 - \delta, t_0 + \delta) \subset I$). Un vettore $l \in \mathbb{R}^n$ è detto limite di f in t_0 se $\lim_{t \rightarrow t_0} \|f(t) - l\| = 0$. In tal caso si scriverà $\lim_{t \rightarrow t_0} f(t) = l$.

Con il simbolo $\|\cdot\|$ indichiamo la norma in \mathbb{R}^n . Notiamo che $\|f(t) - l\|$ rappresenta la distanza di $f(t)$ da l . Notiamo che la definizione 2 basa il concetto di limite di una funzione a valori vettoriali a quello di una funzione a valori reali, cioè alla funzione che alla variabile reale $t \in I$ associa la distanza $\|f(t) - l\|$. Esplicitando la definizione di limite (finito) per funzioni $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ appresa nel corso di analisi matematica 1, possiamo riformulare la definizione 2 nel modo seguente:

Definizione 3 Sia $f : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ e sia $t_0 \in I$ un punto interno all'intervallo I . Un vettore $l \in \mathbb{R}^n$ è detto limite di f in t_0 $\forall \epsilon > 0$ esiste $\delta > 0$ tale che

$$\|f(t) - l\| < \epsilon, \quad \forall t \in I, t \neq t_0, |t - t_0| < \delta.$$

In tal caso si scriverà $\lim_{t \rightarrow t_0} f(t) = l$.

La seguente proposizione permette di ridurre il calcolo del limite di una funzione a valori vettoriali al calcolo del limite delle sue componenti.

Proposizione 1 Sia $f : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$, $t_0 \in I$ punto interno ad I e $l \in \mathbb{R}^n$. Allora $\lim_{t \rightarrow t_0} f(t) = l$ se e solo se per ogni $i = 1, \dots, n$ si ha che $\lim_{t \rightarrow t_0} f_i(t) = l_i$.

Dimostrazione: Dobbiamo mostrare che, per $t \rightarrow t_0$, $\|f(t) - l\| \rightarrow 0$ se e solo se $|f_i(t) - l_i| \rightarrow 0$ per ogni $i = 1, \dots, n$. Dato che, per ogni $i = 1, \dots, n$ si ha che

$$0 \leq |f_i(t) - l_i|^2 \leq \|f(t) - l\|^2 = \sum_{j=1}^n |f_j(t) - l_j|^2,$$

dalla prima diseuguaglianza e dal teorema del confronto segue che $\|f(t) - l\| \rightarrow 0$ implica $|f_i(t) - l_i| \rightarrow 0$. Dalla seconda uguaglianza e dal fatto che il limite di una somma è la somma dei limiti segue che se $|f_i(t) - l_i| \rightarrow 0$ per ogni $i = 1, \dots, n$, allora $\|f(t) - l\| \rightarrow 0$. \square

Esempio 2 Si calcoli il limite per $t \rightarrow 0$ della funzione $f: \mathbb{R} \rightarrow \mathbb{R}^2$ definita da

$$f(t) := \begin{cases} \left(\frac{\sin(t)}{t}, \frac{\cos(t)-1}{t^2} \right) & t \neq 0 \\ (0, -1/2) & t = 0 \end{cases}$$

Dato che $\lim_{t \rightarrow 0} \frac{\sin(t)}{t} = 1$ e $\lim_{t \rightarrow 0} \frac{\cos(t)-1}{t^2} = -1/2$, per la proposizione 1 abbiamo che $\lim_{t \rightarrow 0} f(t) = (1, -1/2)$.

Esempio 3 Si calcoli il limite per $t \rightarrow 0$ della funzione $f: \mathbb{R} \rightarrow \mathbb{R}$ definita da

$$f(t) := \begin{cases} \left(\frac{\sin(t)}{t}, \frac{\cos(t)-1}{t^3} \right) & t \neq 0 \\ (0, -1/2) & t = 0 \end{cases}$$

Dato che $\lim_{t \rightarrow 0} \frac{\sin(t)}{t} = 0$ ma $\lim_{t \rightarrow 0} \frac{\cos(t)-1}{t^3} = \infty$, per la proposizione 1 non esiste il limite per $t \rightarrow 0$ della funzione f .

Definizione 4 Sia $f: I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ e sia $t_0 \in I$. f è continua in t_0 se esiste $\lim_{t \rightarrow t_0} f(t)$ e $\lim_{t \rightarrow t_0} f(t) = f(t_0)$.

Notiamo che la funzione dell'esempio 2 è continua in $t = 0$, mentre la funzione dell'esempio 3 non è continua in $t = 0$.

Notiamo inoltre che, per la proposizione 1, una funzione $f: I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ è continua in un punto $t_0 \in I$ se e solo se lo sono tutte le sue funzioni componenti.

Curve nel piano e nello spazio

Definizione 5 Sia $I \subset \mathbb{R}$ un intervallo della retta reale. Una funzione continua $\alpha: I \rightarrow \mathbb{R}^n$ è detta parametrizzazione di curva continua. La sua immagine γ , ovvero il sottoinsieme di \mathbb{R}^n definito da

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

è detta sostegno della curva.

Se $n = 2$ allora parleremo di curve nel piano, mentre se $n = 3$ parleremo di curve nello spazio.

Esempio 4 Siano $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2)$ due punti di \mathbb{R}^3 .

Il segmento che congiunge P_1 e P_2 è parametrizzabile tramite la funzione $\alpha: [0, 1] \rightarrow \mathbb{R}^3$ definita da

$$\alpha(t) = P_1 + t(P_2 - P_1) = (x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1), z_1 + t(z_2 - z_1)), \quad t \in [0, 1].$$

La retta passante per P_1 e P_2 è parametrizzabile tramite la funzione $\alpha: \mathbb{R} \rightarrow \mathbb{R}^3$ definita da

$$\alpha(t) = P_1 + t(P_2 - P_1) = (x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1), z_1 + t(z_2 - z_1)), \quad t \in (-\infty, +\infty).$$

La semiretta che parte da P_1 e passante per P_2 è parametrizzabile tramite la funzione $\alpha: [0, +\infty) \rightarrow \mathbb{R}^3$ definita da

$$\alpha(t) = P_1 + t(P_2 - P_1) = (x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1), z_1 + t(z_2 - z_1)), \quad t \in [0, +\infty).$$

Esempio 5 La circonferenza nel piano con centro $C = (x_c, y_c)$ e raggio R è parametrizzabile tramite la funzione $\alpha : [0, 2\pi] \rightarrow \mathbb{R}^2$ data da:

$$\alpha(t) = (x_c + R \cos t, y_c + R \sin t), \quad t \in [0, 2\pi].$$

Esempio 6 L'elica cilindrica di raggio R e passo h è parametrizzabile dalla funzione $\alpha : \mathbb{R} \rightarrow \mathbb{R}^3$ data da

$$\alpha(t) = \left(R \cos t, R \sin t, \frac{h}{2\pi} t \right)$$

Una curva di parametrizzazione $\alpha : [a, b] \rightarrow \mathbb{R}^n$ è detta *chiusa* se $\alpha(a) = \alpha(b)$, mentre è detta *semplice* se per ogni $t_1, t_2 \in (a, b)$ con $t_1 \neq t_2$ si ha che $\alpha(t_1) \neq \alpha(t_2)$. Notiamo che secondo questa definizione una curva chiusa può essere semplice.

Notiamo che una stessa curva $\gamma \subset \mathbb{R}^n$ ammette differenti parametrizzazioni, esistono cioè differenti funzioni $\alpha : I \rightarrow \mathbb{R}^n$ che hanno lo stesso sostegno. Ad esempio la semicirconferenza $\gamma = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = R^2, y \geq 0\}$ può essere parametrizzata tramite la funzione $\alpha : [0, \pi] \rightarrow \mathbb{R}^2$ data da

$$\alpha(t) = (R \cos t, R \sin t), \quad t \in [0, \pi],$$

oppure con la funzione $\beta : [-R, R] \rightarrow \mathbb{R}^2$ data da

$$\beta(t) = (t, \sqrt{R^2 - t^2}), \quad t \in [-R, R].$$

È importante avere ben presente la differenza fra la *parametrizzazione* di una curva, che è una funzione di variabile reale a valori vettoriali, che descrive al variare di un parametro $t \in \mathbb{R}$ la posizione dei punti della curva, e il *sostegno*, che è un insieme di punti (del piano o dello spazio). Per capire la differenza è utile tenere presente la seguente analogia cinematica: possiamo pensare alla funzione $\alpha : I \rightarrow \mathbb{R}^n$, la parametrizzazione, come alla legge oraria del moto di un punto materiale, mentre il sostegno $\gamma = \alpha(I)$ è la traiettoria.

Curve regolari

Vogliamo ora introdurre da un punto di vista matematicamente rigoroso il concetto di "curva liscia", ovvero di curva senza spigoli. Dobbiamo trovare le proprietà che deve possedere la parametrizzazione $\alpha : I \rightarrow \mathbb{R}^n$ di una curva γ affinché γ ammetta tangente in ogni punto e la direzione della retta tangente vari con continuità. Introduciamo prima di tutto il concetto di funzione derivabile.

Definizione 6 Sia $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ e sia $t_0 \in I$. α è detta derivabile in t_0 se esiste il limite $\lim_{t \rightarrow t_0} \frac{\alpha(t) - \alpha(t_0)}{t - t_0}$. In tal caso tale vettore viene indicato con $\alpha'(t_0)$.

Notiamo che se $n = 2$ o $n = 3$ possiamo pensare al vettore $\alpha'(t_0)$ come al vettore che indica la direzione della retta tangente alla curva γ nel punto $\alpha(t_0)$. Infatti vale la seguente proposizione.

Proposizione 2 Sia $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$ derivabile in t_0 e sia $\alpha'(t_0) \neq 0$. Si consideri la retta r passante per $\alpha(t_0)$ e diretta nella direzione del vettore $v \in \mathbb{R}^3$, con $v \neq 0$, parametrizzata tramite la funzione $\beta_v : \mathbb{R} \rightarrow \mathbb{R}^3$:

$$\beta_v(t) = \alpha(t_0) + (t - t_0)v, \quad t \in \mathbb{R}$$

Allora abbiamo che $\alpha(t) - \beta_v(t)$ è un infinitesimo di ordine superiore a $t - t_0$ per $t \rightarrow t_0$, ovvero $\lim_{t \rightarrow t_0} \frac{\alpha(t) - \beta_v(t)}{t - t_0} = 0$, se e solo se $v = \alpha'(t_0)$.

Infatti basta notare che

$$\lim_{t \rightarrow t_0} \frac{\alpha(t) - \beta_v(t)}{t - t_0} = \lim_{t \rightarrow t_0} \frac{\alpha(t) - \alpha(t_0)}{t - t_0} - v$$

e per la derivabilità di α in t_0 abbiamo che $\lim_{t \rightarrow t_0} \frac{\alpha(t) - \alpha(t_0)}{t - t_0} - v = 0$ se e solo se $v = \alpha'(t_0)$.

La proposizione 2 ci permette, nell'ipotesi in cui $\alpha'(t_0) \neq 0$, di identificare $\alpha'(t_0)$ come vettore tangente alla curva in $\alpha(t_0)$ in quanto la retta $\beta_{\alpha'(t_0)}$ è la "migliore retta" che approssima l'arco di curva in un intorno del punto $\alpha(t_0)$.

Esprimiamo ora matematicamente la condizione che la direzione della retta tangente vari con continuità. Dato che al variare del parametro $t \in I$ la direzione della retta tangente nel punto di coordinate $\alpha(t)$ è data dal vettore $\alpha'(t)$, allora possiamo pensare che la continuità della funzione $\alpha' : I \rightarrow \mathbb{R}^n$ data da

$$t \mapsto \alpha'(t)$$

sia sufficiente per assicurare l'assenza di punti angolosi nella curva $\gamma = \alpha(I)$.

Definizione 7 Una funzione $\alpha : I \rightarrow \mathbb{R}^n$ è detta di classe $C^1(I)$ se

- α è continua,
- α è derivabile,
- la funzione $t \mapsto \alpha'(t)$ è continua.

Notiamo che la condizione di continuità della derivata prima è più forte della derivabilità della funzione. Possiamo fornire esempi di funzioni derivabili che hanno derivata prima non continua, come ad esempio la funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ definita da

$$f(t) = \begin{cases} t^2 \sin(1/t) & t \neq 0 \\ 0 & t = 0 \end{cases}$$

Infatti f è derivabile e

$$f'(t) = \begin{cases} 2t \sin(1/t) - \cos(1/t) & t \neq 0 \\ 0 & t = 0 \end{cases}$$

in quanto

$$f'(0) = \lim_{t \rightarrow 0} \frac{f(t) - f(0)}{t} = \lim_{t \rightarrow 0} \frac{t^2 \sin(1/t)}{t} = 0$$

ma non esiste il limite

$$\lim_{t \rightarrow 0} f'(t) = \lim_{t \rightarrow 0} 2t \sin(1/t) - \cos(1/t).$$

Diamo ora le condizioni che assicurano che una funzione $\alpha : I \rightarrow \mathbb{R}^n$ sia la parametrizzazione di una curva $\gamma = \alpha(I)$ liscia, ovvero tale per cui la retta tangente è ben definita in ogni punto e la sua direzione varia con continuità.

Definizione 8 Una mappa $\alpha : I \rightarrow \mathbb{R}^n$ è detta parametrizzazione di curva regolare se

- α è di classe $C^1(I)$;
- $\|\alpha'(t)\| \neq 0$ per ogni $t \in I$.

La seconda condizione assicura che in tutti i punti della curva conosciamo la direzione della retta tangente. Non potremmo essere in possesso di tale informazione nei punti in cui $\alpha'(t) = 0$. I seguenti esempi mostrano come tale condizione sia importante.

Esempio 7 Si consideri la curva piana γ di parametrizzazione $\alpha : [-1, 1] \rightarrow \mathbb{R}^2$ data da

$$\alpha(t) = \begin{cases} (0, t^2) & t \in [-1, 0] \\ (t^2, 0) & t \in (0, 1] \end{cases}$$

Il sostegno di tale curva è la linea spezzata congiungente i punti $(0, 1)$, $(0, 0)$ e $(1, 0)$, che ha un angolo in $(0, 0)$, però la mappa α è di classe C^1 . Di fatto il vettore $\alpha'(t)$ si annulla in $t = 0$ e quindi in $\alpha(0)$ non abbiamo il controllo sulla direzione della retta tangente.

Esempio 8 Si consideri la curva piana γ di parametrizzazione $\alpha : \mathbb{R} \rightarrow \mathbb{R}^2$ data da

$$\alpha(t) = (t^3, t^2), \quad t \in \mathbb{R}$$

La funzione α è di classe C^1 , ma $\alpha'(0) = 0$, quindi in $\alpha(0) = (0, 0)$ non possiamo controllare la direzione della tangente. Il sostegno $\gamma = \alpha(\mathbb{R})$ può essere parametrizzato anche come il grafico della funzione $y = x^{2/3}$, ovvero tramite la parametrizzazione $\beta : \mathbb{R} \rightarrow \mathbb{R}^2$ data da

$$\beta(t) = (t, t^{2/3}), \quad t \in \mathbb{R}$$

Utilizzando tale parametrizzazione è possibile verificare che γ presenta una cuspide in $(0, 0)$.

Definizione 9 Un insieme $\gamma \subset \mathbb{R}^n$ è detto curva regolare se esiste una parametrizzazione regolare $\alpha : I \rightarrow \mathbb{R}^n$ tale che $\gamma = \alpha(I)$.

Osserviamo che se γ è regolare, allora la direzione della retta tangente è individuata punto per punto dal *versore tangente* T definito da

$$T(t) := \frac{\alpha'(t)}{\|\alpha'(t)\|},$$

che è un vettore tangente alla curva e di norma costante pari a 1. L'inclinazione di tale vettore, quindi l'inclinazione della retta tangente, può essere descritta tramite gli angoli che tale vettore individua con gli assi coordinati (l'asse x, y, z se $n = 3$). Ad esempio, se indichiamo con $\theta_x(t)$ l'angolo compreso fra $T(t)$ e l'asse delle x abbiamo che

$$\cos \theta_x(t) = T(t) \cdot \hat{i}, \quad \hat{i} = (1, 0, 0),$$

analogamente se indichiamo con $\theta_y(t)$ (rispettivamente con $\theta_z(t)$) l'angolo compreso fra $T(t)$ e l'asse delle y (risp. delle z) abbiamo che $\cos \theta_y(t) = T(t) \cdot \hat{j}$, dove $\hat{j} = (0, 1, 0)$, (risp. $\cos \theta_z(t) = T(t) \cdot \hat{k}$, dove $\hat{k} = (0, 0, 1)$). Dalla continuità della funzione $t \mapsto \alpha'(t)$ possiamo dedurre quella delle funzioni $t \mapsto \theta_x(t)$, $t \mapsto \theta_y(t)$, $t \mapsto \theta_z(t)$.

È importante anche introdurre il concetto di *curva regolare a tratti*, per descrivere curve rappresentabili come unione di curve regolari.

Definizione 10 Una funzione $\alpha : [a, b] \rightarrow \mathbb{R}^n$ è detta parametrizzazione di curva regolare a tratti se esiste una partizione \mathcal{P} dell'intervallo $[a, b]$ in N sottointervalli:

$$t_0 = a \leq t_1 \leq t_2 \leq \dots \leq t_N = b$$

tale che

- $\alpha : [a, b] \rightarrow \mathbb{R}^n$ è continua

- Per ogni $i = 1, \dots, N$ la funzione $\alpha|_{[t_{i-1}, t_i]}$ ristretta all'intervallo $[t_{i-1}, t_i]$ è regolare.

Definizione 11 Un insieme $\gamma \subset \mathbb{R}^n$ è detto curva regolare a tratti se esiste una parametrizzazione regolare a tratti $\alpha : I \rightarrow \mathbb{R}^n$ tale che $\gamma = \alpha(I)$.

Esercizio 1 Fornire una parametrizzazione per l'arco di curva γ appartenente alla superficie di equazione $z = \sqrt{2y^2 - x}$ che ha come proiezione sul piano xy la curva di equazione $x = y^2$, con $y \in [-1, 1]$. Si specifichi poi se γ è regolare.

Soluzione: Indichiamo con $\alpha : [t_0, t_1] \rightarrow \mathbb{R}^3$ la parametrizzazione cercata. La funzione α avrà la forma

$$t \mapsto \alpha(t) = (x(t), y(t), z(t)), \quad t \in [t_0, t_1]$$

descrivendo le coordinate dei punti della curva γ al variare del parametro t . La condizione di appartenenza di γ alla superficie di equazione $z = \sqrt{2y^2 - x}$ si traduce nella seguente relazione fra la funzione $t \mapsto z(t)$ e le funzioni $t \mapsto x(t)$ e $t \mapsto y(t)$:

$$z(t) = \sqrt{2y(t)^2 - x(t)}, \quad \forall t \in [t_0, t_1]. \quad (1)$$

In altre parole, una volta che abbiamo determinato le funzioni $t \mapsto x(t)$ e $t \mapsto y(t)$, allora abbiamo anche la funzione $t \mapsto z(t)$.

Sfuttiamo ora la seconda informazione, ovvero la conoscenza della proiezione di γ sul piano xy . Se indichiamo con $\tilde{\gamma}$ tale curva, e se $t \mapsto (x(t), y(t), z(t))$ è la parametrizzazione di γ , allora $t \mapsto (x(t), y(t), 0)$ sarà la parametrizzazione di $\tilde{\gamma}$. Ora tale curva è la parabola di equazione $x = y^2$, con $y \in [-1, 1]$, che può essere parametrizzata nel seguente modo:

$$t \mapsto (t^2, t, 0), \quad t \in [-1, 1].$$

Possiamo quindi concludere che $x(t) = t^2$, $y(t) = t$ e, dalla relazione (1), $z(t) = \sqrt{2y(t)^2 - x(t)} = |t|$. La parametrizzazione di γ ottenuta è dunque:

$$\alpha(t) = (t^2, t, |t|), \quad t \in [-1, 1].$$

Verifichiamo ora se α è una parametrizzazione regolare, ovvero se α è continua e derivabile, ed inoltre se α' è una funzione continua tale che $\|\alpha'(t)\| \neq 0$ per ogni $t \in [-1, 1]$.

Possiamo subito dire che $\alpha : [-1, 1] \rightarrow \mathbb{R}^3$ è una funzione continua in quanto le sue componenti

$$t \mapsto x(t) = t^2, \quad t \mapsto y(t) = t, \quad t \mapsto z(t) = |t|$$

sono funzioni continue. Verifichiamo ora la derivabilità.

Se $t < 0$, allora $\alpha(t) = (t^2, t, -t)$ e quindi, per $t \in (-1, 0)$, abbiamo che $\alpha'(t) = (2t, 1, -1)$.

Se $t > 0$, allora $\alpha(t) = (t^2, t, t)$ e quindi, per $t \in (0, 1)$, abbiamo che $\alpha'(t) = (2t, 1, 1)$. Per $t = 0$ calcoliamo direttamente $\alpha'(0)$ come limite (se esiste) del rapporto incrementale:

$$\alpha'(0) = \lim_{h \rightarrow 0} \frac{\alpha(h) - \alpha(0)}{h} = \lim_{h \rightarrow 0} \frac{(h^2, h, |h|)}{h} = \lim_{h \rightarrow 0} \left(h, 1, \frac{|h|}{h} \right)$$

Dato che

$$\lim_{h \rightarrow 0^+} \left(h, 1, \frac{|h|}{h} \right) = (0, 1, 1),$$

mentre

$$\lim_{h \rightarrow 0^-} \left(h, 1, \frac{|h|}{h} \right) = (0, 1, -1),$$

dal fatto che limite destro e limite sinistro non coincidono, possiamo dedurre che non esiste $\lim_{h \rightarrow 0} \frac{\alpha(h) - \alpha(0)}{h}$ e quindi α non è derivabile in 0.

Lunghezza di una curva

Sia $\gamma \subset \mathbb{R}^n$ e sia $\alpha : [a, b] \rightarrow \mathbb{R}^n$ una sua parametrizzazione continua. Per calcolare la lunghezza dell'arco di curva γ approssimiamola con la lunghezza di poligoni inscritte.

Indichiamo con $\mathcal{P} = \{t_0 = a \leq t_1 \leq t_2 \leq \dots \leq b = t_N\}$ una generica partizione dell'intervallo $[a, b]$ e consideriamo la spezzata formata dai segmenti di estremi $\alpha(t_i)$ e $\alpha(t_{i+1})$, con $i = 0, \dots, N - 1$. La lunghezza di tale curva, che indichiamo con $l(\mathcal{P}) = \sum_{i=0}^{N-1} \|\alpha(t_{i+1}) - \alpha(t_i)\|$ approssima per difetto la lunghezza di γ . Inoltre, migliorando l'approssimazione, ovvero aggiungendo dei punti intermedi alla partizione \mathcal{P} e ottenendo così una nuova partizione \mathcal{P}' , con $\mathcal{P} \subset \mathcal{P}'$, abbiamo che $l(\mathcal{P}') \geq l(\mathcal{P})$. Di fatto, man mano che miglioriamo l'approssimazione, otteniamo dei valori sempre maggiori per le lunghezze delle spezzate approssimanti. Definiremo lunghezza della curva l'estremo superiore dell'insieme di questi valori.

Definizione 12 *Un curva continua $\gamma \subset \mathbb{R}^n$ è detta rettificabile se*

$$\sup_{\mathcal{P}} \{l(\mathcal{P})\} < +\infty$$

In tal caso tale valore viene definito lunghezza di γ e indicato con $l(\gamma)$.

Non tutte le curve sono rettificabili, alcune hanno lunghezza infinita, ad esempio la curva piana γ , grafico della funzione $f : [0, 1/\pi] \rightarrow \mathbb{R}$, definita da

$$\begin{cases} f(x) = 0 & x = 0 \\ f(x) = x \sin(1/x) & x \neq 0 \end{cases} \quad y = x \sin(1/x)$$

parametrizzata da

$$\alpha(t) = (t, f(t)), \quad t \in [0, 1/\pi].$$

Costruiamo infatti la successione di poligoni inscritti in γ associate alle partizioni $\{\mathcal{P}_n\}_n$, con $\mathcal{P}_n = \{0, 1/\pi, \frac{1}{\frac{\pi}{2}+2k\pi}, \frac{1}{\frac{3\pi}{2}+2k\pi}, k = 1, \dots, n\}$. Se indichiamo con l_0 la lunghezza del segmento congiungente l'origine $(0, 0) = \alpha(0)$ con il punto $\alpha\left(\frac{1}{\frac{3\pi}{2}+2n\pi}\right) = \left(\frac{1}{\frac{3\pi}{2}+2n\pi}, -\frac{1}{\frac{3\pi}{2}+2n\pi}\right)$, con l_f la lunghezza del segmento congiungente il punto $\alpha\left(\frac{1}{\frac{\pi}{2}+2\pi}\right) = \left(\frac{1}{\frac{\pi}{2}+2\pi}, \frac{1}{\frac{\pi}{2}+2\pi}\right)$ con il punto finale $\alpha(1/\pi) = (1/\pi, 0)$, abbiamo che la lunghezza della poligonale associata alla partizione \mathcal{P}_n può essere stimata da:

$$\begin{aligned} l(\mathcal{P}_n) &= l_0 + l_f + \sum_{k=1}^n \left\| \alpha\left(\frac{1}{\frac{\pi}{2} + 2k\pi}\right) - \alpha\left(\frac{1}{\frac{3\pi}{2} + 2k\pi}\right) \right\| + \\ &\quad + \left\| \alpha\left(\frac{1}{\frac{3\pi}{2} + 2k\pi}\right) - \alpha\left(\frac{1}{\frac{\pi}{2} + 2(k+1)\pi}\right) \right\| \\ &\geq \sum_{k=1}^n \left(\frac{1}{\frac{\pi}{2} + 2k\pi} + \frac{1}{\frac{3\pi}{2} + 2k\pi} \right) \\ &= \sum_{k=1}^n \frac{2\pi + 4k\pi}{\left(\frac{\pi}{2} + 2k\pi\right)\left(\frac{3\pi}{2} + 2k\pi\right)} \end{aligned}$$

Dall'ultima linea della catena di disequazioni possiamo dedurre che $\lim_{n \rightarrow \infty} l(\mathcal{P}_n) = +\infty$ perché tale valore è maggiore della somma della serie divergente $\sum_k \frac{2\pi+4k\pi}{\left(\frac{\pi}{2}+2k\pi\right)\left(\frac{3\pi}{2}+2k\pi\right)}$.

Diamo ora delle condizioni sufficienti sulla parametrizzazione che assicurino non solo che una curva è rettificabile, cioè ha lunghezza finita, ma anche che ci permettano di calcolarne la lunghezza.

Teorema 1 *Sia $\gamma \subset \mathbb{R}^n$ una curva regolare a tratti e sia $\alpha : [a, b] \rightarrow \mathbb{R}^n$ una sua parametrizzazione regolare a tratti. Allora γ è rettificabile e*

$$l(\gamma) = \int_a^b \|\alpha'(t)\| dt$$

Abbiamo visto che una curva γ ammette infinite diverse parametrizzazioni. La seguente proposizione ci assicura che la lunghezza di γ non dipende dalla parametrizzazione, se quest'ultima è una funzione iniettiva.

Proposizione 3 Sia $\gamma \subset \mathbb{R}^n$ una curva regolare a tratti e siano $\alpha : [a, b] \rightarrow \mathbb{R}^n$ e $\beta : [c, d] \rightarrow \mathbb{R}^n$ due sue parametrizzazioni regolari a tratti e iniettive. Allora

$$\int_a^b \|\alpha'(t)\| dt = \int_c^d \|\beta'(t)\| dt = l(\gamma).$$

Trattiamo ora in dettaglio alcune particolari classi di curve piane.

Curve piane grafico di funzioni $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$

Consideriamo ora una funzione $f : [a, b] \rightarrow \mathbb{R}$ continua. Il suo grafico $G_f \subset \mathbb{R}^2$ definito da:

$$G_f := \{(x, y) \in \mathbb{R}^2 : y = f(x), x \in [a, b]\}$$

è una particolare curva piana γ . Una sua parametrizzazione è :

$$\alpha(t) = (t, f(t)), \quad t \in [a, b].$$

Notiamo che una curva di questo tipo non è mai chiusa ed è sempre semplice.

Se f è derivabile allora

$$\alpha'(t) = (1, f'(t))$$

Se la funzione $t \mapsto f'(t)$ è continua, allora la parametrizzazione α è regolare. Inoltre la lunghezza di γ è data da

$$l(\gamma) = \int_a^b \sqrt{1 + (f'(t))^2} dt \tag{2}$$

Esempio 9 Calcolare la lunghezza del grafico della funzione $f : [0, 1] \rightarrow \mathbb{R}$, con $f(x) = \sqrt{x^3}$.

Utilizzando la formula (2) otteniamo

$$l(\gamma) = \int_a^b \sqrt{1 + (f'(t))^2} dt = \int_0^1 \sqrt{1 + \frac{9}{4}t} dt = \frac{13\sqrt{13} - 8}{27}.$$

Curve piane espresse in coordinate polari

Ricordiamo che nel piano xy le coordinate cartesiane xy sono collegate alle coordinate polari ρ, θ dalla relazione

$$(x, y) = (\rho \cos \theta, \rho \sin \theta)$$

dove $\rho = \sqrt{x^2 + y^2}$ è la distanza del punto (x, y) dall'origine $(0, 0)$, mentre $\theta \in [0, 2\pi)$ è l'angolo di inclinazione del vettore (x, y) rispetto all'asse delle x , individuato in maniera univoca, se $\sqrt{x^2 + y^2} \neq 0$, da:

$$\cos \theta = \frac{x}{\sqrt{x^2 + y^2}}, \quad \sin \theta = \frac{y}{\sqrt{x^2 + y^2}}$$

Consideriamo ora delle curve piane parametrizzabili tramite una funzione $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ della forma:

$$\alpha(\theta) = (g(\theta) \cos \theta, g(\theta) \sin \theta), \quad \theta \in I,$$

dove $g : I \rightarrow \mathbb{R}_+$ è una funzione continua che esprime la variazione della coordinata ρ dei punti della curva al variare di θ . Se g è una funzione derivabile e la sua derivata $g'(\theta)$ è continua, allora la funzione α sarà di classe C^1 . Il vettore tangente sarà :

$$\alpha'(\theta) = (g'(\theta) \cos \theta - g(\theta) \sin \theta, g'(\theta) \sin \theta + g(\theta) \cos \theta), \quad \theta \in I,$$

$$\|\alpha'(\theta)\| = \sqrt{(g(\theta))^2 + (g'(\theta))^2}.$$

La funzione α sarà regolare se non ci sono valori di θ in cui g e g' si annullano entrambe.

Se θ varia in un intervallo $[\theta_0, \theta_1]$ allora la lunghezza della curva è calcolabile tramite l'integrale:

$$l(\gamma) = \int_{\theta_0}^{\theta_1} \sqrt{(g(\theta))^2 + (g'(\theta))^2} d\theta. \quad (3)$$

Esempio 10 Calcolare la lunghezza della curva piana di parametrizzazione

$$\alpha(\theta) = (\exp(\theta) \cos \theta, \exp(\theta) \sin \theta), \quad \theta \in [0, 2\pi],$$

Utilizzando la formula (3) con $g(\theta) = \exp(\theta)$ otteniamo

$$l(\gamma) = \int_0^{2\pi} \sqrt{2e^{2\theta}} d\theta = \sqrt{2}(e^{2\pi} - 1).$$

Esercizio 2 Calcolare la lunghezza dell'arco di curva γ intersezione delle superfici $y = x^2$ e $3z = 2xy$ dal punto $(0, 0, 0)$ al punto $(2, 4, 16/3)$.

Soluzione: Indichiamo con $\alpha : [t_0, t_1] \subset \mathbb{R}^3$ la parametrizzazione cercata. La funzione α avrà la forma

$$t \mapsto (x(t), y(t), z(t)), \quad t \in [t_0, t_1]$$

La condizione di appartenenza di γ alla superficie di equazione $y = x^2$ si traduce nella seguente relazione fra la funzione $t \mapsto x(t)$ e la funzione $t \mapsto y(t)$:

$$y(t) = x(t)^2, \quad \forall t \in [t_0, t_1]. \quad (4)$$

La condizione di appartenenza di γ alla superficie di equazione $3z = 2xy$ si traduce nella seguente relazione fra la funzione $t \mapsto z(t)$ e le funzioni $t \mapsto x(t)$ e $t \mapsto y(t)$:

$$z(t) = \frac{2}{3}x(t)y(t), \quad \forall t \in [t_0, t_1]. \quad (5)$$

La relazione (4) ci permette di determinare la forma della funzione $t \mapsto y(t)$, nota quella della funzione $t \mapsto x(t)$, mentre la relazione (5) ci permette di determinare la forma della funzione $t \mapsto z(t)$, nota quella delle funzioni $t \mapsto x(t)$ e $t \mapsto y(t)$.

Scegliamo ora di parametrizzare i punti di γ in funzione della loro coordinata x , ovvero di porre $t = x$. Abbiamo quindi

$$x(t) = t, \quad y(t) = t^2, \quad z(t) = \frac{2}{3}t^3,$$

e la parametrizzazione di γ ottenuta è :

$$\alpha(t) = \left(t, t^2, \frac{2}{3}t^3 \right) \quad t \in [t_0, t_1].$$

L'intervallo di variazione del parametro t può essere determinato imponendo che il punto iniziale sia $(0, 0, 0)$:

$$\alpha(t_0) = (0, 0, 0), \quad \Rightarrow t_0 = 0,$$

e il punto finale sia $(2, 4, 16/3)$:

$$\alpha(t_1) = (2, 4, 16/3), \quad \Rightarrow t_1 = 2.$$

In conclusione

$$\alpha(t) = \left(t, t^2, \frac{2}{3}t^3 \right) \quad t \in [0, 2].$$

La lunghezza di γ è data da

$$l(\gamma) = \int_0^2 \|\alpha'(t)\| dt = \int_0^2 \sqrt{1 + 4t^2 + 4t^4} dt = \int_0^2 (1 + 2t^2) dt = \frac{22}{3}$$

Il parametro d'arco e l'ascissa curvilinea

Consideriamo ora un arco di curva regolare γ e sia $\alpha : [a, b] \rightarrow \mathbb{R}^n$ una sua parametrizzazione regolare e iniettiva. Per ogni valore $t \in [a, b]$ consideriamo la lunghezza dell'arco che ha per estremi i punti $\alpha(a)$ e $\alpha(t)$. Indichiamo con $s(t)$ tale valore, che è calcolabile tramite l'integrale:

$$s(t) = \int_a^t \|\alpha'(u)\| du. \quad (6)$$

Notiamo che la funzione $s : [a, b] \rightarrow [0, l(\gamma)]$, $t \mapsto s(t)$ definita dall'equazione (6), è strettamente crescente in quanto

$$s'(t) = \frac{d}{dt} s(t) = \frac{d}{dt} \int_a^t \|\alpha'(u)\| du = \|\alpha'(t)\| > 0$$

Esiste quindi la funzione inversa che ad un dato valore di s , che ricordiamo essere la lunghezza dell'arco che collega il punto iniziale della curva con un punto $P(s)$ che giace sulla curva, associa quello della variabile t :

$$s \mapsto t(s).$$

la variabile s è detta *parametro d'arco* o *ascissa curvilinea*. Possiamo quindi riparametrizzare la curva γ tramite una nuova parametrizzazione $\beta : [0, l(\gamma)] \rightarrow \mathbb{R}^n$ che ad ogni valore dell'ascissa curvilinea s associa le coordinate del punto $P(s)$ tale per cui la lunghezza dell'arco che ha per estremi il punto iniziale (di coordinate $\alpha(a) = \beta(0)$) e il punto $P(s)$.

Esempio 11 la circonferenza nel piano xy di centro $(0,0)$ e raggio R è parametrizzabile con la funzione $\alpha : [0, 2\pi] \rightarrow \mathbb{R}^2$ data da

$$\alpha(t) = (R \cos(t), R \sin(t)), \quad t \in [0, 2\pi],$$

L'ascissa curvilinea è data da

$$s(t) = \int_0^t \|\alpha'(u)\| du = Rt.$$

La relazione tra s e t può essere facilmente invertita ottenendo la variabile t in funzione di s :

$$t(s) = s/R, \quad s \in [0, 2\pi R].$$

La nuova parametrizzazione $\beta : [0, 2\pi R] \rightarrow \mathbb{R}^2$ che esprime le coordinate dei punti della curva in funzione del parametro d'arco s è:

$$\beta(s) = \alpha(t(s)) = (R \cos(s/R), R \sin(s/R)).$$

Esempio 12 (Elica cilindrica) Consideriamo la curva parametrizzata da

$$\alpha(t) = (R \cos t, R \sin t, ht), \quad t \in [0, 2\pi]$$

Il punto iniziale della curva è $\alpha(0) = (1, 0, 0)$. La lunghezza dell'arco $s(t)$ che ha come primo estremo $(1, 0, 0)$ e come secondo estremo il punto $\alpha(t) = (R \cos t, R \sin t, ht)$ è data da:

$$s(t) = \int_0^t \|\alpha'(u)\| du = \int_0^t \sqrt{R^2 + h^2} du = \sqrt{R^2 + h^2} t.$$

Anche in questo caso la dipendenza di $s(t)$ dalla variabile t è lineare e quindi è immediato invertirla, esplicitando la dipendenza del parametro t da s :

$$s(t) = \sqrt{R^2 + h^2} t, \quad t(s) = s / \sqrt{R^2 + h^2}.$$

La parametrizzazione di γ in funzione del parametro d'arco s è quindi:

$$\begin{aligned} \alpha(s) &= (R \cos t(s), R \sin t(s), ht(s)) \\ &= \left(R \cos \left(\frac{s}{\sqrt{R^2 + h^2}} \right), R \sin \left(\frac{s}{\sqrt{R^2 + h^2}} \right), \frac{hs}{\sqrt{R^2 + h^2}} \right), \quad s \in [0, 2\pi\sqrt{R^2 + h^2}] \end{aligned}$$

Alcune riflessioni per capire meglio:

- Cercate di disegnare la curva γ , comprendendo il significato geometrico dei parametri R e h .

Esempio 13 (Il grafico della funzione $y = \cosh x$) Consideriamo la curva piana parametrizzata da

$$\alpha(t) = (t, \cosh t), \quad t \in [0, 1]$$

Il punto iniziale della curva è $\alpha(0) = (0, 1)$. La lunghezza dell'arco $s(t)$ che ha come primo estremo $(0, 1)$ e come secondo estremo il punto $\alpha(t) = (t, \cosh t)$ è data da:

$$s(t) = \int_0^t \|\alpha'(u)\| du = \int_0^t \sqrt{1 + (\sinh u)^2} du = \int_0^t \cosh u du = \sinh t.$$

In questo caso la dipendenza di $s(t)$ dalla variabile t è data da

$$s(t) = \sinh t$$

Tale relazione può essere invertita utilizzando la funzione inversa del seno iperbolico:

$$t(s) = \operatorname{arcsinh}(s) = \log(s + \sqrt{s^2 + 1})$$

La parametrizzazione di γ in funzione del parametro d'arco s è quindi:

$$\alpha(s) = (\log(s + \sqrt{s^2 + 1}), \sqrt{s^2 + 1}), \quad s \in [0, \sinh 1].$$

Alcune riflessioni per capire meglio:

- Disegnate la curva γ
- Per alcuni valori della variabile t , calcolate il valore $s(t)$.

Integrali curvilinei di prima specie

Dato un arco di curva regolare a tratti γ , parametrizzato da $\alpha : [a, b] \rightarrow \mathbb{R}^3$, nelle lezioni precedenti abbiamo imparato come calcolare la lunghezza di γ . In particolare la lunghezza dell'arco compreso fra i punti $\alpha(t_0)$ e $\alpha(t_1)$ è data da:

$$\int_{t_0}^{t_1} \|\alpha'(t)\| dt$$

A livello infinitesimo abbiamo la relazione $ds = \|\alpha'(t)\| dt$, che esprime la lunghezza ds di un arco infinitesimo di curva compreso fra i punti $\alpha(t)$ e $\alpha(t+dt)$. Immaginiamo ora che la curva γ modelli un oggetto unidimensionale pesante (ad es. una fune, un filo...) e supponiamo che punto per punto lungo la curva sia definita una funzione ρ densità lineare di massa, tale cioè che la massa di una porzione infinitesima di filo di lunghezza ds centrata nel punto di coordinate (x, y, z) è data da

$$dm = \rho(x, y, z) ds.$$

Se descriviamo le coordinate dei punti della curva γ tramite la parametrizzazione α , allora la massa di una porzione infinitesima di filo compresa fra i punti $\alpha(t)$ e $\alpha(t+dt)$ è data da:

$$dm = \rho(\alpha(t)) \|\alpha'(t)\| dt$$

Sommando i contributi di tutti questi archi infinitesimi di curva, otteniamo che la massa totale è data da:

$$M = \int_a^b \rho(\alpha(t)) \|\alpha'(t)\| dt.$$

Abbiamo così introdotto in maniera intuitiva una particolare procedura di integrazione di una funzione scalare ρ lungo un arco di curva γ , che definiamo ora in maniera rigorosa:

Definizione 13 Sia $\alpha : [a, b] \rightarrow \mathbb{R}^n$ una parametrizzazione regolare a tratti di curva, sia $\gamma = \alpha([a, b])$ il suo sostegno e sia $f : \gamma \rightarrow \mathbb{R}$ una funzione continua. Si definisce integrale di linea (di prima specie) della funzione f su γ il numero

$$\int_a^b f(\alpha(t)) \|\alpha'(t)\| dt$$

Tale definizione generalizza in modo opportuno l'integrale che definisce la lunghezza della curva, infatti se f è la funzione costante pari a 1, abbiamo che

$$\int_a^b f(\alpha(t))\|\alpha'(t)\|dt = \int_a^b \|\alpha'(t)\|dt = l(\gamma).$$

Inoltre sotto l'ipotesi di iniettività della parametrizzazione, abbiamo che integrale di linea (di prima specie) della funzione f dipende solo dal sostegno γ e non dalla parametrizzazione.

Proposizione 4 *Sia $\gamma \subset \mathbb{R}^n$ una curva regolare a tratti e siano $\alpha : [a, b] \rightarrow \mathbb{R}^n$ e $\beta : [c, d] \rightarrow \mathbb{R}^n$ due sue parametrizzazioni regolari a tratti e iniettive. Allora se $f : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione continua, si ha*

$$\int_a^b f(\alpha(t))\|\alpha'(t)\|dt = \int_c^d f(\beta(t))\|\beta'(t)\|dt$$

Questo risultato permette di definire l'integrale di linea di una funzione continua $f : \mathbb{R} \rightarrow \mathbb{R}$ su una curva regolare a tratti γ , che viene indicato con il simbolo

$$\int_{\gamma} f ds$$

definito da

$$\int_{\gamma} f ds := \int_a^b f(\alpha(t))\|\alpha'(t)\|dt, \quad (7)$$

dove $\alpha : [a, b] \rightarrow \mathbb{R}^n$ è una qualsiasi parametrizzazione regolare a tratti e iniettiva di γ .

Esempio 14 *Si calcoli $\int_{\gamma} f ds$, dove $f(x, y, z) = 1+x/3$ e γ è la curva di parametrizzazione $\alpha : [0, 1] \rightarrow \mathbb{R}^n$, data da*

$$\alpha(t) = (3t, 3t^2, 2t^3), \quad t \in [0, 1],$$

Utilizzando la formula (7) otteniamo

$$\int_{\gamma} f ds = \int_0^1 (1+t)3(1+2t^2)dt = 8.$$

Un'interessante applicazione fisica di questo concetto è il calcolo delle coordinate del baricentro di oggetti "unidimensionali", ovvero tali per cui una delle tre dimensioni è preponderante rispetto alle altre due. Tali oggetti (ad es. funi, travi, ...) possono

venire rappresentati matematicamente tramite un arco di curva regolare $\gamma \subset \mathbb{R}^3$, di parametrizzazione $\alpha : [a, b] \rightarrow \mathbb{R}^n$. Supponendo che la densità lineare di massa sia espressa da una funzione $\rho : \gamma \rightarrow \mathbb{R}$ continua, allora la massa totale verrà calcolata come:

$$M = \int_{\gamma} \rho ds = \int_a^b \rho(\alpha(t)) \|\alpha'(t)\| dt, \quad (8)$$

mentre le coordinate (x_G, y_G, z_G) del baricentro G dell'oggetto sono calcolabili come

$$x_G = \frac{1}{M} \int_{\gamma} x \rho ds := \frac{1}{M} \int_a^b x(t) \rho(\alpha(t)) \|\alpha'(t)\| dt, \quad (9)$$

$$y_G = \frac{1}{M} \int_{\gamma} y \rho ds := \frac{1}{M} \int_a^b y(t) \rho(\alpha(t)) \|\alpha'(t)\| dt, \quad (10)$$

$$z_G = \frac{1}{M} \int_{\gamma} z \rho ds := \frac{1}{M} \int_a^b z(t) \rho(\alpha(t)) \|\alpha'(t)\| dt, \quad (11)$$

dove $\alpha(t) = (x(t), y(t), z(t))$.

Esempio 15 Calcolate le coordinate (x_G, y_G) di una semicirconferenza omogenea (cioè con densità lineare di massa ρ costante), parametrizzata da

$$\alpha(t) = (R \cos t, R \sin t), \quad t \in [0, \pi]$$

Soluzione: Se la densità ρ è una funzione costante, ovvero $\rho(x, y, z) = c$, allora

$$M = \int_{\gamma} \rho ds = \int_0^{\pi} c R dt = c \pi R,$$

$$x_G = \frac{1}{c \pi R} \int_0^{\pi} c R^2 \cos t dt = 0, \quad y_G = \frac{1}{c \pi R} \int_0^{\pi} c R^2 \sin t dt = 2R/\pi$$

Esempio 16 Determinare la massa e le coordinate del baricentro di un filo a forma di elica cilindrica, parametrizzata da

$$\alpha(t) = (\cos t, \sin t, t), \quad t \in [0, 2\pi],$$

e la densità lineare di massa è data dalla funzione

$$\rho(x, y, z) = z$$

Soluzione:

$$\begin{aligned} M &= \int_{\gamma} \rho ds = \int_0^{2\pi} \rho(\alpha(t)) \|\alpha'(t)\| dt \\ &= \int_0^{2\pi} z(t) \sqrt{2} dt \\ &= \sqrt{2} \int_0^{2\pi} t dt = 2\sqrt{2}\pi^2 \end{aligned}$$

$$\begin{aligned} x_G &= \frac{1}{M} \int_{\gamma} x \rho ds = \frac{1}{M} \int_0^{2\pi} x(t) \rho(\alpha(t)) \|\alpha'(t)\| dt, \\ &= \frac{1}{2\sqrt{2}\pi^2} \int_0^{2\pi} \sqrt{2} t \cos t dt = 0 \end{aligned}$$

$$\begin{aligned} y_G &= \frac{1}{M} \int_{\gamma} y \rho ds = \frac{1}{M} \int_0^{2\pi} y(t) \rho(\alpha(t)) \|\alpha'(t)\| dt, \\ &= \frac{1}{2\sqrt{2}\pi^2} \int_0^{2\pi} \sqrt{2} t \sin t dt = -\frac{1}{\pi} \end{aligned}$$

$$\begin{aligned} z_G &= \frac{1}{M} \int_{\gamma} z \rho ds = \frac{1}{M} \int_0^{2\pi} z(t) \rho(\alpha(t)) \|\alpha'(t)\| dt, \\ &= \frac{1}{2\sqrt{2}\pi^2} \int_0^{2\pi} \sqrt{2} t^2 dt = \frac{4}{3}\pi \end{aligned}$$

Elementi di geometria differenziale delle curve

Consideriamo un arco di curva regolare γ , descritta analiticamente da una parametrizzazione regolare $\alpha : I \rightarrow \mathbb{R}^3$, con I intervallo in \mathbb{R} :

$$t \mapsto \alpha(t) = (x(t), y(t), z(t))$$

Supponiamo inoltre che la funzione $\alpha : I \rightarrow \mathbb{R}^3$ sia almeno di classe $C^3(I)$ (chiediamo che α sia tre volte derivabile con tutte le derivate, fino alla terza, continue).

Abbiamo visto che per ogni $\bar{t} \in I$ il vettore $\alpha'(\bar{t})$ è tangente alla curva γ nel punto $\alpha(\bar{t})$. La conoscenza di $\alpha'(\bar{t})$ permette quindi di scrivere l'equazione della retta tangente a γ in $\alpha(\bar{t})$. In forma parametrica questa è infatti data da

$$r(t) = \alpha(\bar{t}) + t\alpha'(\bar{t}), \quad t \in \mathbb{R}.$$

Esempio

Descrivere la retta tangente alla *spirale di Archimede*, parametrizzata da

$$\alpha(t) = (t \cos t, t \sin t), \quad t \in [0, 2\pi]$$

nel punto $\alpha(\pi/2)$.

Soluzione: Il punto corrispondente al valore del parametro $t = \pi/2$ ha coordinate $\alpha(\pi/2) = (0, \pi/2)$. Il vettore tangente alla curva è dato da $\alpha'(t) = (\cos t - t \sin t, \sin t + t \cos t)$. In $t = \pi/2$ abbiamo $\alpha'(\pi/2) = (-\pi/2, 1)$. La retta passante per il punto $(0, \pi/2)$ e diretta lungo il vettore $(-\pi/2, 1)$ ha equazione parametrica

$$\begin{aligned} r(t) &= \left(0, \frac{\pi}{2}\right) + t \left(-\frac{\pi}{2}, 1\right) \\ &= \left(-\frac{\pi}{2}t, \frac{\pi}{2} + t\right), \quad t \in (-\infty, +\infty) \end{aligned}$$

Versore tangente, versore normale e curvatura

Introduciamo ora un parametro reale positivo che esprime la rapidità con cui varia la direzione della retta tangente.

Il primo passo è la definizione del *versore tangente* alla curva γ nel punto $\alpha(t)$. Viene indicato con $T(t)$ e si calcola dividendo il vettore tangente $\alpha'(t)$ per la sua norma:

$$T(t) := \frac{\alpha'(t)}{\|\alpha'(t)\|}$$

Per costruzione $T(t)$ è un vettore di norma 1 tangente punto per punto alla curva γ . Notiamo che, per l'ipotesi di regolarità della curva, $\|\alpha'(t)\| \neq 0$ per ogni t , e quindi $T(t)$ è sempre ben definito.

Il passo successivo è lo studio della variazione di $T(t)$ lungo la curva, ovvero del vettore $\frac{d}{dt}T(t)$. Tale vettore ha le seguenti caratteristiche:

1. La direzione di $\frac{d}{dt}T(t)$ è sempre ortogonale a quella di $T(t)$, infatti, dato che $\|T(t)\| = 1$ per ogni t abbiamo:

$$0 = \frac{d}{dt} (T(t) \cdot T(t)) = 2T(t) \cdot \frac{d}{dt}T(t).$$

Se $\|\frac{d}{dt}T(t)\| \neq 0$, possiamo definire il *versore normale* alla curva γ nel punto $\alpha(t)$ come il vettore di norma 1 con la direzione ed il verso di $\frac{d}{dt}T(t)$. Tale versore, ortogonale a $T(t)$, viene indicato con $N(t)$ ed è dato da

$$N(t) := \frac{\frac{d}{dt}T(t)}{\|\frac{d}{dt}T(t)\|}$$

2. Per quanto riguarda la norma di $\frac{d}{dt}T(t)$, intuitivamente questa esprimerà la velocità di variazione della direzione della retta tangente. In particolare, se $\|\frac{d}{dt}T(t)\| = 0$ per ogni t , allora la costanza del versore tangente ci dirà che la curva γ è una retta ¹. Viceversa, quanto maggiore sarà $\|\frac{d}{dt}T(t)\|$, tanto maggiore sarà la rapidità con cui γ si allontana dalla sua retta tangente in $\alpha(t)$. Sulla base di queste considerazioni definiamo la *curvatura* di γ nel punto $\alpha(t)$ come il numero reale positivo dato da

$$k(t) := \frac{\left\| \frac{d}{dt}T(t) \right\|}{\|\alpha'(t)\|} \quad (12)$$

Indipendenza dalla parametrizzazione

Il parametro $k(t)$ definito dall'equazione (12) è una caratteristica geometrica della curva γ nel punto $\alpha(t)$ e non dipende dalla particolare parametrizzazione $\alpha : [a, b] \rightarrow \mathbb{R}^3$ che la descrive. In altre parole, se parametrizziamo γ con un nuova parametrizzazione $\beta : I \rightarrow \mathbb{R}^3$, il valore della curvatura non cambia.

In particolare, introduciamo il parametro d'arco s , definito tramite la funzione $s : [a, b] \rightarrow [0, l(\gamma)]$:

$$s(t) = \int_a^t \|\alpha'(u)\| du, \quad \text{con } s'(t) = \|\alpha'(t)\|.$$

Indicata con $t : [0, l(\gamma)] \rightarrow [a, b]$ la funzione inversa di s :

$$s \in [0, l(\gamma)] \mapsto t(s), \quad \text{con } t'(s) = \frac{d}{ds}t(s) = \frac{1}{\frac{d}{dt}s(t(s))} = \frac{1}{\|\alpha'(t)\|},$$

introduciamo la parametrizzazione $\beta : [0, l(\gamma)] \rightarrow \mathbb{R}^3$, definita da

$$s \mapsto \beta(s) = (x(s), y(s), z(s)) := \alpha \circ t(s) = \alpha(t(s)).$$

Le formule introdotte finora assumono una forma particolarmente semplice, che mette in evidenza il significato geometrico del parametro k . In particolare il versore tangente è dato da:

$$\beta'(s) = \frac{d}{ds}\beta(s) = \frac{d}{ds}\alpha \circ t(s) = \alpha'(t(s))t'(s) = \frac{\alpha'(t(s))}{\|\alpha'(t(s))\|},$$

mentre, dal fatto che $\|\beta'(s)\| = 1$, otteniamo

$$\frac{d}{ds}T(s) = k(s)N(s), \quad k(s) = \left\| \frac{d}{ds}T(s) \right\|$$

¹Provate per esercizio a dimostrare questa affermazione

Raggio di curvatura e cerchio osculatore

Definiamo *raggio di curvatura* il valore $\rho(t) = \frac{1}{k(t)}$. Tale numero reale positivo rappresenta il raggio del *cerchio osculatore* della curva γ nel punto $\alpha(t)$. Il cerchio passante per $\alpha(t)$ che meglio approssima la curva γ in un intorno di $\alpha(t)$.

Tale cerchio giace sul piano individuato dai versori $T(t)$ e $N(t)$, detto *piano osculatore*. Il centro è il punto C ottenuto partendo dal punto $\alpha(t)$ e muovendosi di una distanza pari a $\rho(t)$ nella direzione del versore $N(t)$:

$$C = \alpha(t) + \rho(t)N(t)$$

Versore binormale e torsione

In tutti i punti in cui $k \neq 0$, sono ben definiti i versori tangente T e normale N alla curva. Definiamo ora il *versore binormale* B definito dal prodotto vettoriale di T e N :

$$B = T \wedge N$$

Per costruzione B è un vettore di norma 1 ortogonale a T e N . La terna T, N, B forma dunque una base ortonormale di \mathbb{R}^3 . Inoltre il vettore B fornisce la direzione ortogonale al piano osculatore. Se calcoliamo $\frac{d}{ds}B(s)$ otterremo un vettore che esprime la variazione della pendenza del piano osculatore.

In particolare vale la seguente formula

Teorema 2

$$\frac{d}{ds}B(s) = \tau(s)N(s)$$

Dimostrazione: Dato che $\|B(s)\| = 1$, abbiamo che $B(s) \cdot \frac{d}{ds}B(s) = 0$, infatti:

$$0 = \frac{d}{ds}(1) = \frac{d}{ds}(B(s) \cdot B(s)) = 2B(s) \cdot \frac{d}{ds}B(s).$$

Inoltre, dato che $T(s) \cdot B(s) = 0$, deduciamo che

$$T(s) \cdot \frac{d}{ds}B(s) = -B(s) \cdot \frac{d}{ds}T(s) = -B(s) \cdot k(s)N(s) = 0,$$

infatti:

$$0 = \frac{d}{ds}(1) = \frac{d}{ds}(T(s) \cdot B(s)) = T(s) \cdot \frac{d}{ds}B(s) + B(s) \cdot \frac{d}{ds}T(s).$$

□

In altre parole il vettore $\frac{d}{ds}B(s)$ è proporzionale al versore normale e il coefficiente di proporzionalità viene chiamato *torsione della curva* e indicato col simbolo $\tau(s)$. Tale numero va ad indicare di quanto varia la direzione ortogonale al piano osculatore, infatti:

Teorema 3 *Sia γ arco di curva regolare e tale che per ogni $P \in \gamma$ è ben definita la terna $\{T(s), N(s), B(s)\}$.*

Allora la curva γ giace su un piano se e solo se $\tau(s) = 0$ per ogni $s \in [0, l(\gamma)]$.

Dimostrazione: È immediato verificare che se una curva γ , giace su un piano π , allora $\tau = 0$, infatti se $\gamma \subset \pi$, allora $B(s)$ sarà un vettore unitario costante ortogonale a π e quindi $B'(s) = 0$.

Viceversa, nell'ipotesi in cui $\tau(s) = 0$ e quindi $B'(s) = 0$, possiamo mostrare che γ , di parametrizzazione $\beta : [0, l(\gamma)] \rightarrow \mathbb{R}^3$, giace sul piano π passante per il punto iniziale della curva $\beta(0)$ e ortogonale al versore binormale $B(0)$. I punti $\beta(s)$, $s \in [0, l(\gamma)]$ soddisfano infatti l'equazione che caratterizza il piano π :

$$(\beta(s) - \beta(0)) \cdot B(0) = 0; \quad s \in [0, l(\gamma)],$$

come si può dimostrare notando che la funzione $f : [0, l(\gamma)] \rightarrow \mathbb{R}$ definita da $f(s) := (\beta(s) - \beta(0)) \cdot B(s)$ è la funzione costante uguale a 0 in quanto $f(0) = 0$ e $f'(s) = 0$ perché:

$$f'(s) = \frac{d}{ds} ((\beta(s) - \beta(0)) \cdot B(s)) = T(s) \cdot B(s) + (\beta(s) - \beta(0)) \cdot B'(s) = 0.$$

□

Esempio: elica cilindrica

Calcolare punto per punto il versore tangente, normale, binormale, curvatura e la torsione dell'arco di elica cilindrica parametrizzato da:

$$\alpha(t) = (R \cos t, R \sin t, ht), \quad t \in [0, 2\pi]$$

Soluzione:

$$\alpha'(t) = (-R \sin t, R \cos t, h), \quad \|\alpha'(t)\| = \sqrt{R^2 + h^2},$$

Il parametro d'arco è dato da

$$s(t) = \int_0^t \sqrt{R^2 + h^2} dt = \sqrt{R^2 + h^2} t$$

questa relazione può essere invertita fornendo la variabile t in funzione di s :

$$t = \frac{s}{\sqrt{R^2 + h^2}}$$

e la parametrizzazione $\beta : [0, \sqrt{R^2 + h^2}2\pi] \rightarrow \mathbb{R}^3$ della curva γ in funzione del parametro d'arco è :

$$\beta(s) = \alpha(t(s)) = \left(R \cos \left(\frac{s}{\sqrt{R^2 + h^2}} \right), R \sin \left(\frac{s}{\sqrt{R^2 + h^2}} \right), \frac{hs}{\sqrt{R^2 + h^2}} \right)$$

il versore tangente è dato da:

$$T(s) = \left(-\frac{R}{\sqrt{R^2 + h^2}} \sin \left(\frac{s}{\sqrt{R^2 + h^2}} \right), \frac{R}{\sqrt{R^2 + h^2}} \cos \left(\frac{s}{\sqrt{R^2 + h^2}} \right), \frac{h}{\sqrt{R^2 + h^2}} \right)$$

Inoltre

$$\frac{d}{ds}T(s) = \left(-\frac{R}{R^2 + h^2} \cos \left(\frac{s}{\sqrt{R^2 + h^2}} \right), -\frac{R}{R^2 + h^2} \sin \left(\frac{s}{\sqrt{R^2 + h^2}} \right), 0 \right)$$

quindi

$$k(s) = \left\| \frac{d}{ds}T(s) \right\| = \frac{R}{R^2 + h^2}$$

$$N(s) = \left(-\cos \left(\frac{s}{\sqrt{R^2 + h^2}} \right), -\sin \left(\frac{s}{\sqrt{R^2 + h^2}} \right), 0 \right)$$

Il versore binormale è dato da

$$B(s) = T(s) \wedge N(s)$$

$$= \left(\frac{h}{\sqrt{R^2 + h^2}} \sin \left(\frac{s}{\sqrt{R^2 + h^2}} \right), -\frac{h}{\sqrt{R^2 + h^2}} \cos \left(\frac{s}{\sqrt{R^2 + h^2}} \right), \frac{R}{\sqrt{R^2 + h^2}} \right)$$

e la sua derivata

$$\frac{d}{ds}B(s) = \left(\frac{h}{R^2 + h^2} \cos \left(\frac{s}{\sqrt{R^2 + h^2}} \right), \frac{h}{R^2 + h^2} \sin \left(\frac{s}{\sqrt{R^2 + h^2}} \right), 0 \right)$$

$$= -\frac{h}{R^2 + h^2} \left(-\cos \left(\frac{s}{\sqrt{R^2 + h^2}} \right), -\sin \left(\frac{s}{\sqrt{R^2 + h^2}} \right), 0 \right)$$

da cui possiamo concludere che

$$\tau(s) = -\frac{h}{R^2 + h^2}$$

Formule di Frenet-Serret

Data una curva regolare $\gamma \subset \mathbb{R}^3$, parametrizzata in termini del parametro d'arco con la parametrizzazione $\beta : [0, l(\gamma)] \rightarrow \mathbb{R}^3$ di classe $C^3([0, l(\gamma)])$, valgono le seguenti formule, dette formule di *Frenet-Serret* per le derivate dei tre versori della terna $\{T(s), N(s), B(s)\}$:

$$\begin{cases} \frac{d}{ds}T(s) = k(s)N(s) \\ \frac{d}{ds}N(s) = -k(s)T(s) - \tau(s)B(s) \\ \frac{d}{ds}B(s) = \tau(s)N(s) \end{cases} \quad (13)$$

Tali formule implicano in particolare che la conoscenza delle due funzioni scalari $k(s)$ e $\tau(s)$ è sufficiente per determinare univocamente la curva γ , a meno di rototraslazioni. Utilizzando infatti la teoria dei sistemi di equazioni differenziali ordinarie e riguardando (13) come un sistema di 9 equazioni nelle 9 funzioni incognite $T_1(s), T_2(s), T_3(s), N_1(s), N_2(s), N_3(s), B_1(s), B_2(s), B_3(s)$ (le componenti dei versori $T(s), N(s), B(s)$), possiamo concludere che, nota la configurazione iniziale, cioè le componenti dei versori $T(0), N(0), B(0)$, la soluzione di (13) esiste ed è unica.

Analoga cinematica

Se guardiamo la parametrizzazione $\alpha : I \rightarrow \mathbb{R}^3$ della curva γ come la legge oraria del moto di un punto materiale, allora il parametro t viene interpretato come la variabile tempo, mentre il vettore $\alpha(t)$ viene interpretato come il vettore posizione del punto materiale all'istante t . Analogamente $\alpha'(t)$ va interpretato come il vettore velocità istantanea $\vec{v}(t)$ e $\alpha''(t)$ come il vettore accelerazione $\vec{a}(t)$. Proviamo a scrivere tali vettori mettendo in evidenza T, N e k . Scriviamo il vettore velocità $\vec{v}(t)$ mettendo in evidenza la sua norma:

$$\alpha'(t) \equiv \vec{v}(t) = \|\vec{v}(t)\|T(t)$$

Il vettore accelerazione quindi è dato da

$$\begin{aligned} \vec{a}(t) &= \frac{d}{dt}\vec{v}(t) = \frac{d}{dt}(\|\vec{v}(t)\|T(t)) \\ &= \frac{d}{dt}(\|\vec{v}(t)\|)T(t) + \|\vec{v}(t)\|\frac{d}{dt}T(t) \\ &= \frac{d}{dt}(\|\vec{v}(t)\|)T(t) + \|\vec{v}(t)\|^2k(t)N(t) \end{aligned}$$

Nell'ultima riga si vede come il vettore accelerazione istantanea viene scomposto in 2 componenti: una tangente al moto (diretta lungo il versore T), la componente

tangenziale, e una ortogonale al moto (diretta lungo il versore N), la componete *centripeta*. Quest'ultima aumenta all'aumentare della curvatura e del modulo della velocità istantanea. In altre parole, affinché la traiettoria del punto descriva la curva γ , è necessaria una forza in grado di produrre, oltre alla componente tangenziale dell'accelerazione, anche quella centripeta. Quest'ultima sarà tanto più grande quanto maggiore sarà la curvatura di γ e la velocità del punto materiale.

Formule alternative per il calcolo di T, N, B, k, τ

A volte, per semplificare i calcoli, è utile utilizzare le seguenti formule, che valgono in tutti i punti in cui $\alpha'(t) \wedge \alpha''(t) \neq 0$:

$$T(t) = \frac{\alpha'(t)}{\|\alpha'(t)\|}, \quad (14)$$

$$B(t) = \frac{\alpha'(t) \wedge \alpha''(t)}{\|\alpha'(t) \wedge \alpha''(t)\|}, \quad (15)$$

$$N(t) = B(t) \wedge T(t), \quad (16)$$

$$k(t) = \frac{\|\alpha'(t) \wedge \alpha''(t)\|}{\|\alpha'(t)\|^3}, \quad (17)$$

$$\tau(t) = -\frac{(\alpha'(t) \wedge \alpha''(t)) \cdot \alpha'''(t)}{\|\alpha'(t) \wedge \alpha''(t)\|^2} \quad (18)$$

Esercizio 3 Dimostrare le formule (15), (17) e (18).

Esempio: Calcolare la curvatura della curva γ nel piano xy grafico della funzione $y = x^2$.

Soluzione: Parametizziamo γ nel modo seguente

$$\alpha(t) = (t, t^2, 0), \quad t \in \mathbb{R}.$$

Abbiamo allora:

$$\alpha'(t) = (1, 2t, 0)$$

$$\alpha''(t) = (0, 2, 0)$$

$$\alpha'(t) \wedge \alpha''(t) = (0, 0, 2)$$

$$K(t) = \frac{\|\alpha'(t) \wedge \alpha''(t)\|}{\|\alpha'(t)\|^3} = \frac{2}{(1 + 4t^2)^{3/2}}$$

Notiamo che $k(t)$ assume il valore massimo per $t = 0$, ovvero nel vertice $(0, 0)$ della parabola $y = x^2$.

Esercizio per casa: Ricavate il versore tangente ed il versore normale in tutti i punti della curva γ .

Funzioni reali di più variabili

Studiamo ora funzioni $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ definite su un sottoinsieme di A di \mathbb{R}^n , il *dominio* della funzione, a valori reali, ovvero delle mappe che ad ogni vettore $x = (x_1, \dots, x_n) \in A$ associano un unico numero reale $f(x_1, \dots, x_n)$. Possiamo costruire facilmente esempi di tali funzioni.

Se $n = 5$ consideriamo ad esempio la funzione $f : \mathbb{R}^5 \rightarrow \mathbb{R}$ definita da $f(x_1, x_2, x_3, x_4, x_5) = \sin(x_1 x_2) * \cos(x_3) + x_4 x_5$.

Se $n = 2$ consideriamo la funzione di 2 variabili $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, con dominio $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$, definita da $f(x, y) := \sqrt{1 - x^2 - y^2}$.

Nel seguito ci concentreremo sui casi particolari $n = 2$ (dove f è definita su sottoinsiemi del piano) e $n = 3$.

Rappresentazione grafica

Nel corso di analisi 1 abbiamo imparato a studiare funzioni reali di variabile reale, ovvero funzioni $f : A \subset \mathbb{R} \rightarrow \mathbb{R}$ definite su sottoinsiemi della retta reale. In tal caso il comportamento della funzione può essere rappresentato graficamente tramite il *grafico di f* , cioè tramite la curva nel piano xy definita da

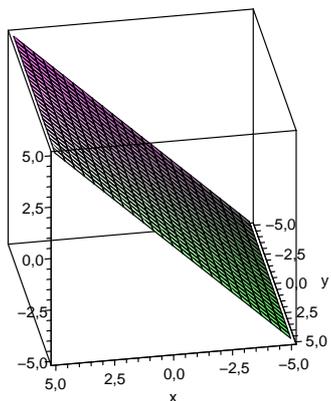
$$G_f := \{(x, y) \in \mathbb{R}^2 : y = f(x), x \in A\}$$

Nel caso in cui consideriamo funzioni $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, è ancora possibile rappresentarne il comportamento tramite la rappresentazione del grafico di f , che in tal caso è una particolare insieme² dello spazio tridimensionale, definito da

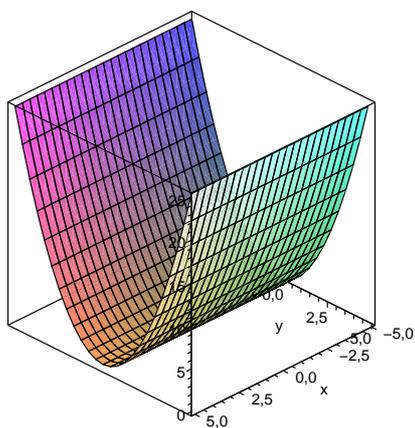
$$G_f := \{(x, y, z) \in \mathbb{R}^3 : z = f(x, y), (x, y) \in A\}$$

Diamo qui di seguito alcuni esempi

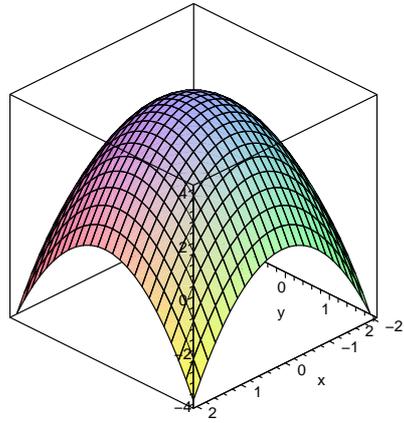
²Vedremo presto che, sotto ipotesi di regolarità sulla funzione f , il grafico di f è una superficie.



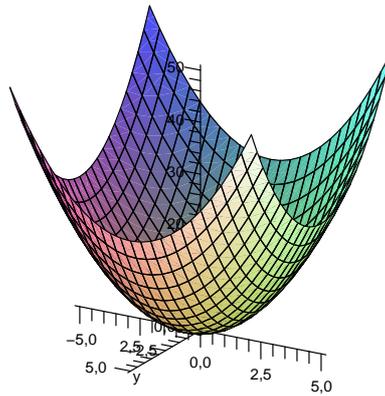
il grafico della funzione $f(x, y) = x$ è un piano



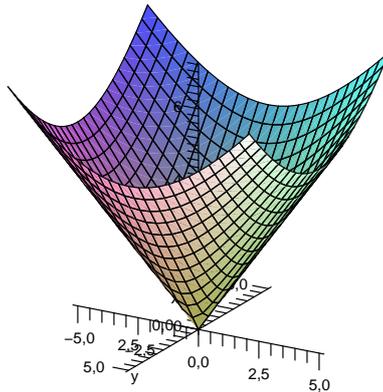
il grafico della funzione $f(x, y) = y^2$



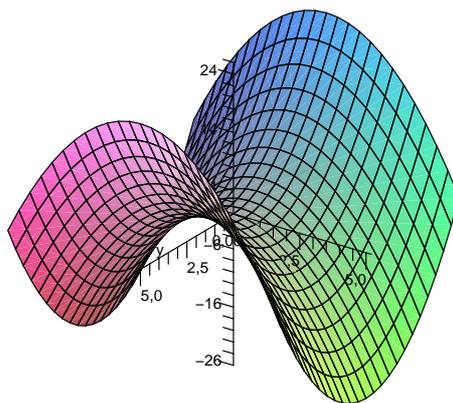
il grafico della funzione $f(x, y) = 4 - x^2 - y^2$



il grafico della funzione $f(x, y) = x^2 + y^2$ è detto paraboloido



il grafico della funzione $f(x, y) = \sqrt{x^2 + y^2}$ è un cono



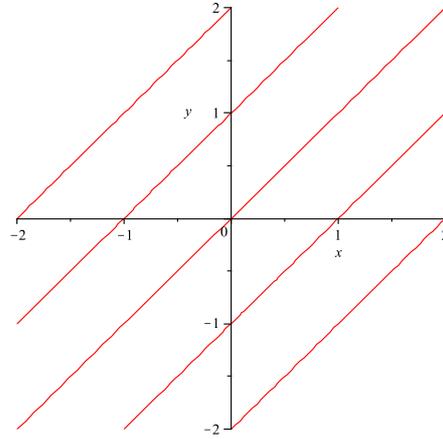
il grafico della funzione $f(x, y) = x^2 - y^2$ è detto paraboloido iperbolico.

Un metodo alternativo per la rappresentazione grafica del comportamento di una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ è lo studio delle *linee di livello*. Fissato un valore $c \in \mathbb{R}$, si definisce linea o curva di livello di f relativa al livello c il sottoinsieme di \mathbb{R}^2 definito da

$$l_f^c := \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}.$$

C'è un'interessante relazione tra linee di livello e grafico di f . La linea di livello di f relativa al livello c è ottenibile come la proiezione sul piano xy dell'intersezione tra la superficie G_f e il piano (orizzontale) di equazione $z = c$.

Studiamo come esempio le curve di livello della funzione $f(x, y) = x - y$. Gli insiemi di punti del piano che soddisfano l'equazione $x - y = c$ sono rette parallele alla bisettrice del primo e terzo quadrante e di equazione $y = x - c$.



alcune curve di livello della funzione $f(x, y) = x - y$

È importante sottolineare come le curve di livello non sono sempre delle curve. A seconda dell'espressione della funzione f e del valore del livello c , gli insiemi l_f^c possono essere differenti. Consideriamo ad esempio la funzione $f(x, y) = x^2 + y^2$:

- Se $c > 0$ allora $l_f^c := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = c\}$ è la circonferenza di centro $(0,0)$ e raggio \sqrt{c} .
- Se $c = 0$ allora $l_f^c := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 0\}$ contiene solo il punto $(0,0)$
- Se $c < 0$ allora $l_f^c := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 0\}$ è l'insieme vuoto.

Nel casi in cui vogliamo rappresentare graficamente il comportamento di funzioni di tre variabili reali, ovvero nel caso in cui $f : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, allora non è più possibile fare un disegno del grafico di f in quanto questo è un sottoinsieme di \mathbb{R}^4 definito da

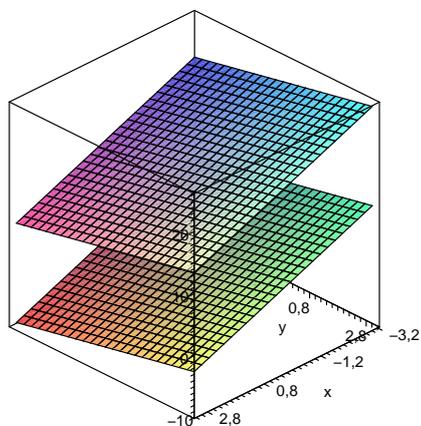
$$G_f = \{(x, y, z, w) \in \mathbb{R}^4 : w = f(x, y, z), (x, y, z) \in A\}$$

È invece ancora possibile rappresentare graficamente gli insiemi di livello, che in questo caso sono detti *superfici di livello*³ e sono definite da

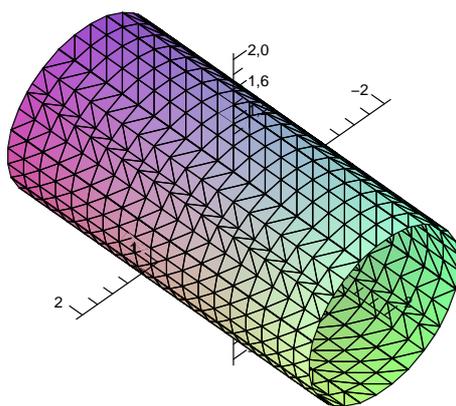
$$S_f^c := \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = c\}$$

³Vedremo più avanti nel corso delle condizioni sufficienti sulla funzione f che assicurino che gli insiemi di livello siano effettivamente delle superfici

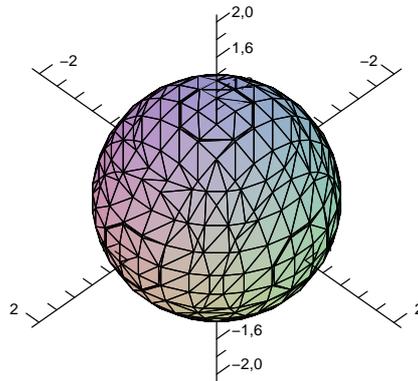
diamo qui di seguito alcuni esempi:



Alcune superfici di livello della funzione $f(x, y, z) = 2x - y + z$



superficie di livello della funzione $f(x, y, z) = x^2 + z^2$



superficie di livello della funzione $f(x, y, z) = e^{x^2+y^2+z^2}$

Limiti per funzioni da $\mathbb{R}^n \rightarrow \mathbb{R}$

Il concetto di limite è la base per lo studio di proprietà "locali" delle funzioni. Premettiamo dunque la definizione di intorno sferico di un punto $x_0 \in \mathbb{R}^n$, che generalizza il concetto di intervallo (aperto) di \mathbb{R} centrato in un punto.

Definizione 14 Sia $x_0 \in \mathbb{R}^n$ si definisce intorno sferico di x_0 (di raggio δ) l'insieme:

$$U_\delta(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| < \delta\}$$

L'insieme $U_\delta(x_0)$ contiene tutti i punti di \mathbb{R}^n la cui distanza da x_0 è minore di δ .

Se $n = 1$ allora $U_\delta(x_0)$ è l'intervallo $(x_0 - \delta, x_0 + \delta)$.

Se $n = 2$ allora $U_\delta(x_0)$ è il cerchio (pieno) di centro x_0 e raggio δ

Se $n = 3$ allora $U_\delta(x_0)$ è la sfera (piena) di centro x_0 e raggio δ .

Introduciamo inoltre alcuni elementi di topologia di \mathbb{R}^n .

Dato $A \subset \mathbb{R}^n$ un generico sottoinsieme $A \subset \mathbb{R}^n$ di \mathbb{R}^n , indicheremo con A^c l'insieme complementare di A , definito da

$$A^c = \{x \in \mathbb{R}^n : x \notin A\}$$

Un punto $x \in \mathbb{R}^n$ è detto:

- *interno* se esiste un intorno sferico $B_\delta(x)$ di x tale che $B_\delta(x) \subset A$,
- *esterno* se esiste un intorno sferico $B_\delta(x)$ di x tale che $B_\delta(x) \subset A^c$,

- di *frontiera* se per ogni intorno sferico $B_\delta(x)$ di x si ha che $B_\delta(x) \cap A \neq \emptyset$ e $B_\delta(x) \cap A^c \neq \emptyset$.

Si definisce *interno* di A , e si indica con $\overset{\circ}{A}$, l'insieme dei punti interni di A .

Si definisce *frontiera* (o *bordo*) di A , e si indica con ∂A , l'insieme dei punti di frontiera di A .

Si definisce *chiusura* di A , e si indica con \bar{A} , l'insieme $\bar{A} := \overset{\circ}{A} \cup \partial A$.

Un insieme A è detto *aperto* se tutti i suoi punti sono interni, cioè $A = \overset{\circ}{A}$. Un insieme A è detto *chiuso* se coincide con la sua chiusura, cioè $A = \bar{A}$.

Esempio 17 Sia $A := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. Abbiamo che $\overset{\circ}{A} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$, $\partial A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ e $\bar{A} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. Possiamo concludere che l'insieme A è chiuso.

Esempio 18 Sia $A := \{(x, y) \in \mathbb{R}^2 : 0 < x^2 + y^2 \leq 1\}$. Abbiamo che $\overset{\circ}{A} = \{(x, y) \in \mathbb{R}^2 : 0 < x^2 + y^2 < 1\}$, $\partial A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \cup \{(0, 0)\}$ e $\bar{A} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. Possiamo concludere che l'insieme A non è né aperto né chiuso.

Per il concetto di limite di funzione è importante quello di limite di una successione. Una successione di punti $\{x_k\}_{k \in \mathbb{N}}$ è un'applicazione che ad ogni $k \in \mathbb{N}$ associa un vettore $x_k \in \mathbb{R}^n$. Il vettore $\bar{x} \in \mathbb{R}^n$ è detto limite della successione $\{x_k\}_{k \in \mathbb{N}}$ se

$$\lim_{k \rightarrow \infty} \|x_k - \bar{x}\| = 0.$$

Vediamo ora due differenti definizioni di limite, di fatto tra loro equivalenti, di limite di una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Definizione 15 Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $x_0 \in \mathbb{R}^n$ tale che esiste almeno un intorno aperto $U_\delta(x_0)$ (eccetto al più il punto x_0 stesso) tutto contenuto in A . Sia $l \in \mathbb{R}$. Si dice che $\lim_{x \rightarrow x_0} f(x) = l$ se per ogni $\epsilon > 0$ esiste un valore $\delta > 0$ tale che $U_\delta(x_0) \setminus \{x_0\} \subset A$ e

$$|f(x) - l| < \epsilon, \quad \forall x \in U_\delta(x_0) \setminus \{x_0\}$$

La seguente definizione è equivalente ed è utile nei casi in cui si vuole dimostrare la non esistenza del limite.

Definizione 16 Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $x_0 \in \mathbb{R}^n$ tale che esiste almeno un intorno aperto $U_\delta(x_0)$ (eccetto al più il punto x_0 stesso) tutto contenuto in A . Sia $l \in \mathbb{R}$. Si dice che $\lim_{x \rightarrow x_0} f(x) = l$ se per ogni successione $\{x_k\}_{k \in \mathbb{N}} \subset A$ tale che

- $x_k \neq x_0$ per ogni $k \in \mathbb{N}$
- $\lim_{k \rightarrow \infty} x_k = x_0$

si ha che

$$\lim_{k \rightarrow \infty} f(x_k) = l$$

Da qui in avanti ci occuperemo solamente di funzioni di due variabili, in quanto in tal caso abbiamo un'interpretazione grafica di numerosi concetti.

Sia ora $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ tale che esiste un intorno aperto di (x_0, y_0) tutto contenuto in A , eccetto al più il punto (x_0, y_0) stesso. La definizione 16 è particolarmente utile se vogliamo mostrare che non esiste $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$. A tal fine infatti è sufficiente trovare due successioni $(x_k, y_k)_{k \in \mathbb{N}}$ e $(x'_k, y'_k)_{k \in \mathbb{N}}$ tali che

- $(x_k, y_k) \neq (x_0, y_0)$ e $(x'_k, y'_k) \neq (x_0, y_0)$ per ogni $k \in \mathbb{N}$
- $\lim_{k \rightarrow \infty} (x_k, y_k) = \lim_{k \rightarrow \infty} (x'_k, y'_k) = (x_0, y_0)$.

ma tali che $\lim_{k \rightarrow \infty} f(x_k, y_k) \neq \lim_{k \rightarrow \infty} f(x'_k, y'_k)$.

Esempio 19 Si dimostri che non esiste $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$, dove

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

È sufficiente considerare ad esempio le successioni $(x_k, y_k) = (1/k, 1/k)$ e $(x'_k, y'_k) = (1/k, 2/k)$. Abbiamo che $\lim_{k \rightarrow \infty} (x_k, y_k) = \lim_{k \rightarrow \infty} (x'_k, y'_k) = (0, 0)$, ma

$$\lim_{k \rightarrow \infty} f(x_k, y_k) = 1/2, \quad \lim_{k \rightarrow \infty} f(x'_k, y'_k) = 2/5$$

Una variante di tale metodo è la seguente. È sufficiente trovare due curve continue γ_1, γ_2 di parametrizzazioni continue

$$\alpha_1 : I \subset \mathbb{R} \rightarrow \mathbb{R}^2, \quad t \mapsto \alpha_1(t),$$

$$\alpha_2 : I \subset \mathbb{R} \rightarrow \mathbb{R}^2, \quad s \mapsto \alpha_2(s),$$

tali che $\alpha_1(t_0) = \alpha_2(s_0) = (x_0, y_0)$ e

$$\lim_{t \rightarrow t_0} f(\alpha_1(t)) \neq \lim_{s \rightarrow s_0} f(\alpha_2(s)).$$

Ad esempio, per la funzione dell'esempio 19, è sufficiente considerare le rette γ_1, γ_2 di parametrizzazioni

$$\alpha_1 : I \subset \mathbb{R} \rightarrow \mathbb{R}^2, \quad t \mapsto \alpha_1(t) = (t, t),$$

$$\alpha_2 : I \subset \mathbb{R} \rightarrow \mathbb{R}^2, \quad s \mapsto \alpha_2(s) = (s, 2s).$$

Abbiamo che $\alpha_1(0) = \alpha_2(0) = (0, 0)$ e

$$\lim_{t \rightarrow 0} f(\alpha_1(t)) = 1/2, \quad \lim_{s \rightarrow 0} f(\alpha_2(s)) = 2/5.$$

Per dimostrare invece che $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y) = l$, possiamo esprimere le coordinate cartesiane (x, y) in funzioni delle coordinate polari ρ, θ centrate nel punto (x_0, y_0) :

$$x = x_0 + \rho \cos \theta, \quad y = y_0 + \rho \sin \theta$$

e cercare di ricavare un diseuguaglianza della forma seguente:

$$|f(x, y) - l| = |f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) - l| \leq g(\rho)$$

dove $g : \mathbb{R}^+ \rightarrow \mathbb{R}$ è una funzione della variabile ρ tale che $\lim_{\rho \rightarrow 0} g(\rho) = 0$.

Esempio 20 Si dimostri che $\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y}{x^2 + y^2} = 0$.

Dato che

$$|f(x, y) - l| = |f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) - l| = \rho |\sin \theta| \cos^2 \theta \leq \rho$$

abbiamo $g(\rho) = \rho$ e $\lim_{\rho \rightarrow 0} g(\rho) = 0$.

La definizione di continuità per funzioni di più variabili non presenta differenze rispetto a quella per funzioni di una singola variabile reale.

Definizione 17 Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ tale che esiste un intorno aperto di (x_0, y_0) tutto contenuto in A . Si dice che f è continua in (x_0, y_0) se

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = f(x_0, y_0)$$

Derivate parziali e derivate direzionali

Nel caso di funzioni $f : \mathbb{R} \rightarrow \mathbb{R}$ di variabile reale, se la funzione f è derivabile in un punto x_0 del suo dominio, allora la derivata prima $f'(x_0)$ esprime il tasso di incremento di f in x_0 . Per funzioni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ di due variabili il tasso di incremento della funzione in un punto (x_0, y_0) dipenderà dalla direzione in cui ci spostiamo partendo da (x_0, y_0) . Introduciamo dunque il concetto di derivata direzionale.

Definizione 18 Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ un punto interno ad A . Fissato un vettore $v \in \mathbb{R}^2$, con $\|v\| = 1$, si definisce derivata direzionale di f in (x_0, y_0) rispetto al versore v il limite del rapporto incrementale:

$$\lim_{h \rightarrow 0} \frac{f((x_0, y_0) + hv) - f(x_0, y_0)}{h}$$

se tale limite esiste finito. In tal caso viene indicato con il simbolo $D_v f(x_0, y_0)$.

Dalla definizione precedente notiamo che $D_v f(x_0, y_0)$ esprime il tasso di incremento di f in (x_0, y_0) nella direzione del versore v .

Particolarmente importanti sono le derivate direzionali rispetto ai versori $\hat{i} = (1, 0)$ e $\hat{j} = (0, 1)$ della base canonica di \mathbb{R}^2 che esprimono il tasso di incremento di f in (x_0, y_0) lungo direzioni parallele agli assi coordinati. Si dice *derivata parziale di f rispetto a x in (x_0, y_0)* il limite, se esiste finito:

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

e viene indicato con $\frac{\partial}{\partial x} f(x_0, y_0)$. Analogamente si dice *derivata parziale di f rispetto a y in (x_0, y_0)* il limite, se esiste finito:

$$\lim_{h \rightarrow 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}$$

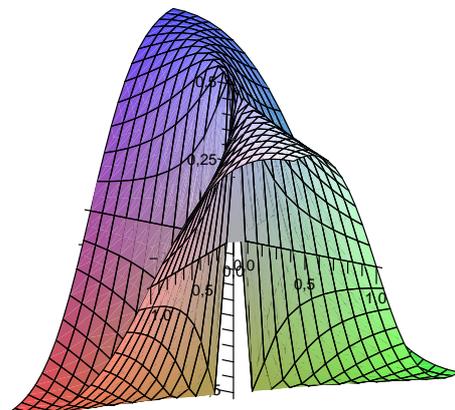
e viene indicato con $\frac{\partial}{\partial y} f(x_0, y_0)$.

Definizione 19 Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ un punto interno ad A . Si dice che f è derivabile in (x_0, y_0) se esistono finite tutte le derivate parziali di f in (x_0, y_0) . In tal caso il vettore $\left(\frac{\partial}{\partial x} f(x_0, y_0), \frac{\partial}{\partial y} f(x_0, y_0) \right) \in \mathbb{R}^2$ viene detto *gradiente di f in (x_0, y_0)* e indicato con il simbolo $\nabla f(x_0, y_0)$.

Notiamo che, a differenza di quanto abbiamo visto nel corso di analisi 1 per funzioni $f : \mathbb{R} \rightarrow \mathbb{R}$, la derivabilità di una funzione non assicura la sua continuità. Ad esempio la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

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

non è continua nell'origine $(0, 0)$, ma è derivabile in quel punto e $\left(\frac{\partial}{\partial x} f(0, 0), \frac{\partial}{\partial y} f(0, 0) \right) = (0, 0)$ (infatti la funzione f è identicamente $= 0$ sugli assi x, y).



il grafico della funzione $f(x, y) = \frac{xy}{x^2+y^2}$

Nella prossima sezione vedremo di introdurre una condizione piú forte della derivabilità, che generalizzi al caso di funzioni di piú variabili il concetto di derivabilità per funzioni di una sola variabile.

Differenziabilità

Data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e un punto $(x_0, y_0) \in A$ interno all'insieme A , occupiamoci ora di risolvere i seguenti problemi:

1. trovare l'equazione del piano tangente al grafico di f nel punto $(x_0, y_0, f(x_0, y_0))$;
2. trovare il "miglior" polinomio di primo grado nelle variabili (x, y) , ovvero una funzione della forma

$$(x, y) \mapsto Ax + By + C,$$

con A, B, C coefficienti reali, che approssimi f in un intorno di (x_0, y_0) .

Vedremo tra breve che i due problemi sono strettamente collegati.

Ricordiamo innanzitutto l'equazione di un piano passante per un punto $P_0 = (x_0, y_0, z_0)$ e ortogonale ad un vettore $N = (\alpha, \beta, \gamma)$. Le coordinate (x, y, z) di un generico punto P appartenente a tale piano devono soddisfare l'equazione

$$(x - x_0, y - y_0, z - z_0) \cdot N = 0.$$

Tale equazione esprime l'ortogonalità fra il vettore $P - P_0$ e N . L'equazione può essere scritta anche nella forma:

$$\alpha(x - x_0) + \beta(y - y_0) + \gamma(z - z_0) = 0$$

e se $\gamma \neq 0$

$$z = z_0 - \frac{\alpha(x - x_0) + \beta(y - y_0)}{\gamma}$$

Ad esempio l'equazione del piano passante per $P_0 = (1, 1, 5)$ e ortogonale a $N = (2, -1, 1)$ è :

$$z = 5 - 2(x - 1) + (y - 1)$$

Consideriamo ora la superficie G_f grafico della funzione f , e un suo punto $P_0 = (x_0, y_0, f(x_0, y_0))$. Il piano tangente a G_f nel punto $(x_0, y_0, f(x_0, y_0))$ (se esiste) dovrà essere il piano contenente tutte le rette tangenti a G_f in $(x_0, y_0, f(x_0, y_0))$. In altre parole, se $\gamma \subset G_f$ è una generica curva regolare passante per $(x_0, y_0, f(x_0, y_0))$ e che giace tutta sulla superficie G_f , allora la retta tangente a γ in $(x_0, y_0, f(x_0, y_0))$ dovrà essere contenuta nel piano.

Consideriamo ora due particolari curve γ_1 e γ_2 appartenenti alla superficie G_f . Sia γ_1 l'intersezione tra G_f e il piano verticale (parallelo al piano yz) di equazione $x = x_0$. Tale curva può essere parametrizzata tramite la funzione $\alpha_1 : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$, dove I è un intervallo contenente il punto $t = 0$:

$$\alpha_1(t) = (x_0, y_0 + t, f(x_0, y_0 + t)), \quad t \in I$$

Abbiamo che $\alpha_1(0) = (x_0, y_0, f(x_0, y_0))$ e $\alpha_1'(0) = \left(0, 1, \frac{\partial}{\partial y} f(x_0, y_0)\right)$ è il vettore che esprime la direzione della retta tangente a γ_1 in $(x_0, y_0, f(x_0, y_0))$ (e che dovrà appartenere al piano tangente).

Analogamente consideriamo la curva γ_2 intersezione tra G_f e il piano verticale (parallelo al piano xz) di equazione $y = y_0$, parametrizzata tramite la funzione $\alpha_2 : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$, dove I è un intervallo contenente il punto $t = 0$:

$$\alpha_2(t) = (x_0 + t, y_0, f(x_0 + t, y_0)), \quad t \in I.$$

Abbiamo che $\alpha_2(0) = (x_0, y_0, f(x_0, y_0))$ e $\alpha_2'(0) = \left(1, 0, \frac{\partial}{\partial x} f(x_0, y_0)\right)$ è il vettore che esprime la direzione della retta tangente a γ_2 in $(x_0, y_0, f(x_0, y_0))$ (e che dovrà appartenere al piano tangente).

Due vettori linearmente indipendenti individuano un piano, ortogonale al prodotto vettoriale dei due vettori. Quindi, se il piano tangente è ben definito, questo sarà il piano passante per $P_0 = (x_0, y_0, f(x_0, y_0))$ e ortogonale al vettore

$$N := \left(1, 0, \frac{\partial}{\partial x} f(x_0, y_0)\right) \wedge \left(0, 1, \frac{\partial}{\partial y} f(x_0, y_0)\right) = \left(-\frac{\partial}{\partial x} f(x_0, y_0), -\frac{\partial}{\partial y} f(x_0, y_0), 1\right)$$

L'equazione del piano sarà dunque:

$$z = f(x_0, y_0) + \frac{\partial}{\partial x} f(x_0, y_0)(x - x_0) + \frac{\partial}{\partial y} f(x_0, y_0)(y - y_0) \quad (19)$$

che può essere scritta in forma sintetica come:

$$z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0) \quad (20)$$

Quindi, se il piano tangente è ben definito, sarà descritto dall'equazione (19), o, equivalentemente, (20). Non è detto però che tale piano tangente sia ben definito. Questo è abbastanza intuitivo se pensiamo a superfici come il grafico della funzione $f(x, y) = \sqrt{x^2 + y^2}$, che è un cono e nel punto $(0,0,0)$ non ammette piano tangente. In questo caso la funzione f non è derivabile in $(0,0)$.

Si possono presentare anche situazioni interessanti in cui, anche se la funzione f è derivabile in (x_0, y_0) , l'equazione (20) descrive un piano che non contiene tutti i vettori tangenti a G_f in $(x_0, y_0, f(x_0, y_0))$.

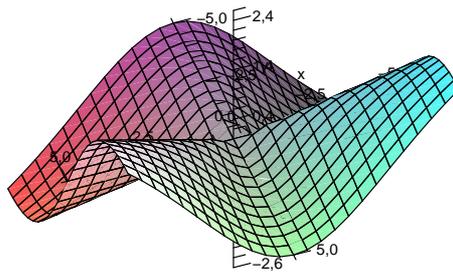
Esempio 21 Consideriamo la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

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

Si ha che f è derivabile in $(0,0)$ e $\left(\frac{\partial}{\partial x} f(0,0), \frac{\partial}{\partial y} f(0,0)\right) = (0,0)$. L'equazione del piano tangente dovrebbe essere quindi $z = 0$ (che descrive il piano xy). Consideriamo però la curva $\gamma \subset G_f$ parametrizzata da

$$\alpha(t) = (t, t, f(t, t)) = (t, t, \frac{t}{2}), \quad t \in I$$

Abbiamo che $\alpha(0) = (0,0,0)$ e $\alpha'(0) = (1, 1, \frac{1}{2})$. Il vettore $(1, 1, \frac{1}{2})$ è tangente a γ in $(0,0,0)$, ma non appartiene al piano $z = 0$.



il grafico della funzione $f(x, y) = \frac{x^2 y}{x^2 + y^2}$

Vediamo ora come questo problema è collegato al punto 2, ovvero all'approssimazione di f in un intorno di (x_0, y_0) con un polinomio di primo grado nelle variabili x, y . Notiamo che la funzione

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

il cui grafico rappresenta il piano candidato ad essere tangente a G_f in $(x_0, y_0, f(x_0, y_0))$, è un polinomio di primo grado nelle variabili x, y . Tale polinomio è una "buona" approssimazione per f in un intorno di (x_0, y_0) se la funzione f possiede una proprietà di regolarità detta *differenziabilità*, che andiamo ora a definire.

Definizione 20 Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ un punto interno ad A . Si dice che f è differenziabile in (x_0, y_0) se esiste un vettore $a = (a_x, a_y) \in \mathbb{R}^2$ tale che

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x, y) - f(x_0, y_0) - a \cdot (x - x_0, y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0 \quad (21)$$

L'equazione (21) può essere scritta nella forma equivalente

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(x_0 + h, y_0 + k) - f(x_0, y_0) - a \cdot (h, k)}{\sqrt{h^2 + k^2}} = 0 \quad (22)$$

Inoltre l'equazione (21) esprime il fatto che, in un intorno di (x_0, y_0) , la funzione f può essere approssimata dal polinomio

$$(x, y) \mapsto P(x, y) = f(x_0, y_0) + a \cdot (x - x_0, y - y_0) = f(x_0, y_0) + a_x(x - x_0) + a_y(y - y_0),$$

e la differenza $f(x, y) - P(x, y)$ tende a 0 per $(x, y) \rightarrow (x_0, y_0)$ più velocemente di $\sqrt{(x - x_0)^2 + (y - y_0)^2}$. Torneremo tra poco sul significato geometrico di tale condizione. Intanto vediamo come la differenziabilità sia una condizione più forte della sola derivabilità, infatti vale il seguente teorema.

Teorema 4 Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ un punto interno ad A . Se f è differenziabile in (x_0, y_0) allora f è derivabile in (x_0, y_0) e il vettore a è dato da $a = \nabla f(x_0, y_0)$.

Inoltre vale il seguente risultato

Teorema 5 Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ un punto interno ad A . Se f è differenziabile in (x_0, y_0) allora f è continua in (x_0, y_0) .

Abbiamo dunque che, se f è differenziabile in (x_0, y_0) allora

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x,y) - f(x_0,y_0) - \nabla f(x_0,y_0) \cdot (x-x_0, y-y_0)}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} = 0$$

Tale equazione esprime il fatto che in un intorno di (x_0, y_0) lo scarto fra il valore della funzione in un punto (x, y) e la quota z del piano $z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0)$ (candidato ad essere tangente a G_f in $(x_0, y_0, f(x_0, y_0))$) tende a 0 quando $(x, y) \rightarrow (x_0, y_0)$ più velocemente di $\sqrt{(x - x_0)^2 + (y - y_0)^2}$, la distanza fra (x, y) e (x_0, y_0) .

Notiamo che la funzione dell'esempio 21 non è differenziabile, infatti non esiste

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x,y) - f(x_0,y_0) - \nabla f(x_0,y_0) \cdot (x-x_0, y-y_0)}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} = \lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y}{(x^2 + y^2)^{3/2}}$$

come si può verificare studiando il limite lungo rette differenti passanti per $(0, 0)$.

Il seguente teorema fornisce una condizione sufficiente per la differenziabilità, facile da verificare nella maggior parte dei casi di interesse.

Teorema 6 *Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ un punto interno ad A . Se le derivate parziali di f esistono in un intorno di (x_0, y_0) e sono continue in (x_0, y_0) allora f è differenziabile in (x_0, y_0) . In particolare se le derivate parziali di f esistono e sono continue in un intorno di (x_0, y_0) allora f è differenziabile in (x_0, y_0)*

Se le derivate parziali di una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ esistono e sono continue in tutti i punti di un insieme A si dice che $f \in C^1(A)$.

Ad esempio la funzione $f(x, y) = x^2 + y^2$ è di classe $C^1(\mathbb{R}^2)$, infatti le sue due derivate parziali $\frac{\partial}{\partial x} f(x, y) = 2x$ e $\frac{\partial}{\partial y} f(x, y) = 2y$ sono continue. Per il teorema precedente f è differenziabile.

Il seguente teorema permette, nel caso in cui f sia differenziabile in un punto (x_0, y_0) , di calcolare tutte le derivate direzionali tramite una formula sintetica.

Teorema 7 (Formula del gradiente) *Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ un punto interno ad A . Se f è differenziabile in (x_0, y_0) , allora f ammette tutte le derivate direzionali in (x_0, y_0) e:*

$$D_v f(x_0, y_0) = \nabla f(x_0, y_0) \cdot v \tag{23}$$

Il precedente risultato permette di identificare il gradiente come quel vettore che indica la direzione di massima crescita della funzione, infatti se $\nabla f(x_0, y_0) \neq 0$ allora la derivata direzionale assume il valore massimo se $v = \frac{\nabla f(x_0, y_0)}{\|\nabla f(x_0, y_0)\|}$ e $D_v f(x_0, y_0) = \|\nabla f(x_0, y_0)\|$, mentre $D_v f(x_0, y_0) = 0$ se v è un vettore ortogonale a $\nabla f(x_0, y_0)$. Concludiamo con un teorema che generalizza il teorema di derivazione di funzione composta al caso di funzioni di più variabili.

Teorema 8 (derivazione di funzione composta) Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $(x_0, y_0) \in \mathbb{R}^2$ un punto interno ad A . Sia $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ tale che $\gamma \equiv \alpha(I) \subset A$ e $\alpha(t_0) = (x_0, y_0)$, con $t_0 \in I$. Allora se α è derivabile in t_0 e se f è differenziabile in $\alpha(t_0) = (x_0, y_0)$, allora la funzione composta $f \circ \alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}$, definita da

$$f \circ \alpha(t) = f(\alpha(t)), \quad t \in I$$

è derivabile in t_0 e

$$(f \circ \alpha)'(t_0) = \nabla f(\alpha(t_0)) \cdot \alpha'(t_0).$$

Tale risultato può essere applicato per mostrare che se f è differenziabile in (x_0, y_0) , allora il piano di equazione $z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0)$ è effettivamente tangente al grafico di f in $(x_0, y_0, f(x_0, y_0))$ in quanto contiene tutte le rette tangenti a curve (regolari) che giacciono sulla superficie G_f e passanti per $(x_0, y_0, f(x_0, y_0))$.

Consideriamo infatti un arco di curva regolare $\gamma \subset G_f$ contenuta nella superficie G_f e passante per $(x_0, y_0, f(x_0, y_0))$ e mostriamo che la retta tangente a γ in $(x_0, y_0, f(x_0, y_0))$ appartiene al piano $z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0)$. Se indichiamo con $T \in \mathbb{R}^3$ il vettore tangente a γ in $(x_0, y_0, f(x_0, y_0))$, la parametrizzazione $r : (-\infty, +\infty) \rightarrow \mathbb{R}^3$ della retta passante per $(x_0, y_0, f(x_0, y_0))$ e diretta nella direzione indicata dal vettore T è :

$$r(s) = (x_0, y_0, f(x_0, y_0)) + sT, \quad s \in (-\infty, +\infty).$$

Calcoliamo il vettore tangente T . Sia $P\gamma$ la curva ottenuta proiettando γ sul piano xy e sia $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ una sua parametrizzazione:

$$t \mapsto \alpha(t) = (x(t), y(t)), \quad t \in I,$$

inoltre esiste un valore $t_0 \in I$ tale che $\alpha(t_0) = (x_0, y_0)$. La curva γ sarà allora parametrizzata tramite la mappa $\beta : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$:

$$t \mapsto \beta(t) = (x(t), y(t), f(x(t), y(t))), \quad t \in I,$$

con $\beta(t_0) = (x_0, y_0, f(x_0, y_0))$. Il vettore tangente a γ in $\beta(t_0)$ sarà quindi il vettore

$$T = \beta'(t_0) = (x'(t_0), y'(t_0), (f \circ \alpha)'(t_0)) = (x'(t_0), y'(t_0), \nabla f(\alpha(t_0)) \cdot \alpha'(t_0))$$

Le coordinate $(x(s), y(s), z(s))$ dei punti della retta tangente a γ in $(x_0, y_0, f(x_0, y_0))$ sono date quindi da

$$\begin{aligned} r(s) &= (x(s), y(s), z(s)) \\ &= (x_0, y_0, f(x_0, y_0)) + sT \\ &= (x_0, y_0, f(x_0, y_0)) + s(x'(t_0), y'(t_0), \nabla f(\alpha(t_0)) \cdot \alpha'(t_0)) \\ &= (x_0 + sx'(t_0), y_0 + sy'(t_0), f(x_0, y_0) + s\nabla f(\alpha(t_0)) \cdot \alpha'(t_0)) \end{aligned}$$

ed è possibile verificare direttamente che tale retta appartiene al piano $z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0)$, ovvero che le funzioni $x(s), y(s), z(s)$ soddisfano l'equazione:

$$z(s) = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x(s) - x_0, y(s) - y_0), \quad \forall s \in \mathbb{R}.$$

Ortogonalità del gradiente alle curve di livello

Un'ulteriore applicazione del teorema 8 è la dimostrazione che il vettore gradiente di una funzione $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ è ortogonale, punto per punto, alle linee di livello di g . Sia infatti $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione differenziabile e $\gamma = \{(x, y) \in \mathbb{R}^2 : g(x, y) = c\}$ una sua linea di livello. Sia $\alpha : I \rightarrow \mathbb{R}^2$ una sua parametrizzazione regolare. Allora, considerata la funzione composta $g \circ \alpha$, si ha che $g \circ \alpha(t) = g(\alpha(t)) = c$ per ogni $t \in I$. Quindi si ha che $\frac{d}{dt}g \circ \alpha(t) = 0$. D'altra parte, per il teorema di derivazione di funzione composta, $\frac{d}{dt}g \circ \alpha(t) = \nabla g(\alpha(t)) \cdot \alpha'(t)$. Otteniamo dunque:

$$\nabla g(\alpha(t)) \cdot \alpha'(t) = 0, \quad \forall t \in I.$$

Tale relazione esprime l'ortogonalità fra il vettore ∇g valutato nel punto della curva γ di coordinate $\alpha(t)$ e il vettore $\alpha'(t)$, tangente alla curva γ nel punto $\alpha(t)$.

Differenziabilità per funzioni $f : \mathbb{R}^n \rightarrow \mathbb{R}$

I risultati delle sezioni precedenti, introdotti per funzioni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, possono essere estesi al caso generale di funzioni $f : \mathbb{R}^n \rightarrow \mathbb{R}$, anche se per $n > 2$ non è possibile rappresentare il grafico della funzione f e quindi non riusciamo a dare un'interpretazione geometrica diretta di tutti i concetti introdotti.

In generale, data una funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e un punto $x_0 \in \mathbb{R}^n$ interno ad A , si definisce *derivata direzionale* di f in x_0 nella direzione di un versore $v \in \mathbb{R}^n$ il limite (se esiste finito) del rapporto incrementale

$$D_v f(x_0) := \lim_{h \rightarrow 0} \frac{f(x_0 + hv) - f(x_0)}{h}.$$

Nel caso particolare in cui $v = \hat{e}_i$, l' i -esimo vettore della base canonica di \mathbb{R}^n , con $i = 1, \dots, n$, si definisce *derivata parziale di f* in x_0 rispetto alla i -esima variabile x_i il limite (se esiste finito) del rapporto incrementale

$$\frac{\partial}{\partial x_i} f(x_0) := \lim_{h \rightarrow 0} \frac{f(x_0 + h\hat{e}_i) - f(x_0)}{h}.$$

Se esistono tutte le n derivate parziali di f in x_0 si definisce *gradiente* di f in x_0 il vettore $\nabla f(x_0) \in \mathbb{R}^n$:

$$\nabla f(x_0) \in \mathbb{R}^n := \left(\frac{\partial}{\partial x_1} f(x_0), \dots, \frac{\partial}{\partial x_n} f(x_0) \right).$$

La funzione f è detta *differenziabile* in x_0 se esiste un vettore $a \in \mathbb{R}^n$ tale che

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - a \cdot (x - x_0)}{\|x - x_0\|} = 0.$$

Il teoremi 4, 5,6,7, enunciati al caso di funzioni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, rimangono validi anche nel caso di funzioni $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Il teorema 8 assume la forma seguente.

Teorema 9 (Derivazione di funzione composta) Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $x_0 \in \mathbb{R}^n$ un punto interno ad A . Sia $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ tale che $\gamma \equiv \alpha(I) \subset A$ e $\alpha(t_0) = x_0$, con $t_0 \in I$. Allora se α è derivabile in t_0 e se f è differenziabile in $\alpha(t_0) = x_0$, allora la funzione composta $f \circ \alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}$, definita da

$$f \circ \alpha(t) = f(\alpha(t)), \quad t \in I$$

è derivabile in t_0 e

$$(f \circ \alpha)'(t_0) = \nabla f(\alpha(t_0)) \cdot \alpha'(t_0).$$

Derivate di ordine superiore

Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione di due variabili definita su un insieme aperto A . Supponiamo che f sia derivabile in A , cioè che siano definite le funzioni $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$, derivate parziali di f :

$$(x, y) \mapsto \frac{\partial f}{\partial x}(x, y) \quad \text{e} \quad (x, y) \mapsto \frac{\partial f}{\partial y}(x, y).$$

Se queste funzioni sono a loro volta derivabili, allora possiamo definire le rispettive derivate parziali:

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) &\equiv \frac{\partial^2 f}{\partial x^2}, & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) &\equiv \frac{\partial^2 f}{\partial y \partial x}, \\ \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) &\equiv \frac{\partial^2 f}{\partial x \partial y}, & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) &\equiv \frac{\partial^2 f}{\partial y^2}. \end{aligned}$$

Queste vengono chiamate *derivate parziali seconde di f* e indicate con i simboli

$$\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y^2}.$$

Nel caso generale di una funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di n variabili definita su un insieme aperto $A \subset \mathbb{R}^n$ e derivabile in tutti i suoi punti, considerando le n funzioni derivate parziali di f

$$\frac{\partial f}{\partial x_i}, \quad i = 1, \dots, n,$$

si definiscono le n^2 derivate parziali seconde

$$\frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) \equiv \frac{\partial^2 f}{\partial x_j \partial x_i}, \quad i, j = 1, \dots, n,$$

ottenute calcolando la derivata parziale della funzione $\frac{\partial f}{\partial x_i}$ rispetto alla j -esima coordinata x_j .

In numerose applicazioni risulta utile costruire una matrice $n \times n$, detta *matrice Hessiana* di f e indicata col simbolo H_f in cui l'elemento $(H_f)_{i,j}$ (situato nella i -esima riga e nella j -esima colonna) è la derivata parziale seconda $\frac{\partial^2 f}{\partial x_i \partial x_j}$. In particolare, nel caso di una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di due variabili, la matrice Hessiana di f in un punto $(x, y) \in A$ (in cui sono definite tutte le 4 derivate parziali seconde) è della forma

$$H_f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Il seguente teorema mostra come, sotto opportune ipotesi di regolarità sulla funzione f , la matrice Hessiana è simmetrica.

Teorema 10 (di Schwarz) *Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione di n variabili e $x_0 \in A$ un punto interno ad A . Si supponga che per certi indici $i, j \in \{1, \dots, n\}$ esistano le derivate parziali seconde $\frac{\partial^2 f}{\partial x_j \partial x_i}$ e $\frac{\partial^2 f}{\partial x_i \partial x_j}$ in un intorno di x_0 e che queste siano entrambe continue in x_0 . Allora si ha che*

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0).$$

In particolare se le derivate seconde miste $\frac{\partial^2 f}{\partial x_j \partial x_i}$ e $\frac{\partial^2 f}{\partial x_i \partial x_j}$ esistono e sono continue in A , queste coincidono su tutto A .

Analogamente alle derivate parziali seconde, è possibile definire derivate parziali miste di ogni ordine iterando opportunamente l'operazione di derivazione parziale. Il teorema di Schwarz si generalizza nello stesso modo, dimostrando l'indipendenza dall'ordine di derivazione delle derivate parziali miste, nell'ipotesi che queste siano continue. Introduciamo la definizione di funzione di classe C^k

Definizione 21 *Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione di n variabili definita su un insieme aperto $A \subset \mathbb{R}^n$. f è detta di classe $C^k(A)$ se tutte le derivate parziali*

$$\frac{\partial^h f}{\partial x_{i_h} \dots \partial x_{i_1}}$$

esistono e sono continue in A , per ogni $1 \leq h \leq k$, $i_1, \dots, i_h \in \{1, \dots, n\}$

Polinomio di Taylor per funzioni di 2 variabili

Nella sezione precedente abbiamo introdotto il concetto di differenziabilità di una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ in un punto $(x_0, y_0) \in \mathbb{R}^2$ tramite la relazione

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x,y) - P_1(x,y)}{\|(x-x_0, y-y_0)\|} = 0, \quad (24)$$

dove $P_1(x, y) = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0)$. In virtù di (24), il polinomio P_1 è il miglior polinomio di primo grado nelle variabili x, y che approssima la funzione f in un intorno del punto (x_0, y_0) , in quanto per $(x, y) \rightarrow (x_0, y_0)$ la differenza fra $f(x, y)$ e $P_1(x, y)$ tende a 0 più velocemente della distanza fra il punto (x, y) e il punto (x_0, y_0) .

Cerchiamo ora di migliorare l'approssimazione di f in un intorno del punto (x_0, y_0) tramite un polinomio P_2 di secondo grado nelle variabili x, y . Vale il seguente risultato

Teorema 11 (Formula di Taylor) *Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe $C^2(A)$. Per ogni $(x_0, y_0) \in A$ vale la formula*

$$f(x, y) = P_2(x, y) + o(\|(x - x_0, y - y_0)\|^2)$$

dove $P_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ è il polinomio di secondo grado nelle variabili x, y , dato da:

$$\begin{aligned} P_2(x, y) &= f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0) + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(x_0, y_0) (x - x_0)^2 + \\ &\quad + \frac{\partial^2}{\partial x \partial y} f(x_0, y_0) (x - x_0)(y - y_0) + \frac{1}{2} \frac{\partial^2}{\partial y^2} f(x_0, y_0) (y - y_0)^2 \\ &= f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0) + \\ &\quad + \frac{1}{2} (x - x_0, y - y_0) H_f(x_0, y_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \end{aligned}$$

dove $H(x_0, y_0)$ è una matrice quadrata 2×2 contenente le derivate parziali doppie di f in (x_0, y_0) :

$$H_f(x_0, y_0) = \begin{pmatrix} \frac{\partial^2}{\partial x^2} f(x_0, y_0) & \frac{\partial^2}{\partial x \partial y} f(x_0, y_0) \\ \frac{\partial^2}{\partial x \partial y} f(x_0, y_0) & \frac{\partial^2}{\partial y^2} f(x_0, y_0) \end{pmatrix}$$

mentre la scrittura $f(x, y) = P_2(x, y) + o(\|(x - x_0, y - y_0)\|^2)$ significa che:

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x,y) - P_2(x,y)}{\|(x-x_0, y-y_0)\|^2} = 0$$

la matrice $H_f(x_0, y_0)$ è detta *matrice Hessiana* di f nel punto (x_0, y_0) .

Spesso è conveniente esprimere la formula di Taylor in funzione delle componenti (h, k) del vettore $(x - x_0, y - y_0)$, dove $x = x_0 + h$, $y = y_0 + k$. In tal caso abbiamo:

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (h, k) + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(x_0, y_0) h^2 + \\ + \frac{\partial^2}{\partial x \partial y} f(x_0, y_0) h k + \frac{1}{2} \frac{\partial^2}{\partial y^2} f(x_0, y_0) k^2 + o(|(h, k)|^2)$$

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(x_0 + h, y_0 + k) - \nabla f(x_0, y_0) \cdot (h, k) - \frac{1}{2} (h, k) H_f(x_0, y_0) \begin{pmatrix} h \\ k \end{pmatrix}}{h^2 + k^2} = 0$$

Tale risultato è generalizzabile al caso di funzioni $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di n variabili reali.

Teorema 12 (Formula di Taylor) *Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe $C^2(A)$. Per ogni $x_0 \in A$ vale la formula*

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2} (x - x_0) H_f(x_0) (x - x_0) + o(\|x - x_0\|^2)$$

Massimi e minimi locali

In questa sezione descriveremo dei criteri per individuare punti di massimo e punti di minimo locale per una funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di n variabili reali. Tratteremo in dettaglio il caso $n = 2$ perché in questo caso proprietà analitiche della funzione hanno un'interpretazione grafica, mentre faremo solo dei cenni al caso $n > 2$.

Definizione 22 *Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$.*

Un punto $x_0 \in A$ interno ad A è detto punto di massimo relativo (o di massimo locale) per f se esiste un intorno sferico $B_\delta(x_0) \subset A$ tale che per ogni $x \in B_\delta(x_0)$ si ha $f(x) \leq f(x_0)$. Nel caso in cui per ogni $x \in B_\delta(x_0)$ con $x \neq x_0$ si ha $f(x) < f(x_0)$, allora $x_0 \in A$ è detto punto di massimo locale stretto.

Analogamente un punto $x_0 \in A$ interno ad A è detto punto di minimo relativo (o di minimo locale) per f se esiste un intorno sferico $B_\delta(x_0) \subset A$ tale che per ogni $x \in B_\delta(x_0)$ si ha $f(x) \geq f(x_0)$. Nel caso in cui per ogni $x \in B_\delta(x_0)$ con $x \neq x_0$ si ha $f(x) > f(x_0)$, allora $x_0 \in A$ è detto punto di minimo locale stretto.

Il seguente teorema fornisce un criterio per individuare i punti candidati ad essere di massimo o minimo relativo.

Teorema 13 (di Fermat) Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e sia $x_0 \in A$ un punto interno ad A . Se x_0 è un punto di massimo o di minimo locale e f è derivabile in x_0 , allora si ha che $\nabla f(x_0) = 0$.

Definizione 23 Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Un punto $x_0 \in A$ interno ad A tale che $\nabla f(x_0) = 0$ è detto punto critico o punto stazionario per la funzione f .

Notiamo che, secondo il teorema di Fermat, i punti di massimo o di minimo locale vanno ricercati tra

- i punti in cui f è derivabile e $\nabla f = 0$,
- i punti in cui f non è derivabile.

Ad esempio :

1. la funzione $f(x, y) = x^2 + y^2$, ovunque derivabile con $\nabla f = (2x, 2y)$, ha un punto di minimo locale in $(0,0)$, unico punto in cui si annulla il gradiente;
2. la funzione $f(x, y) = \sqrt{x^2 + y^2}$, ovunque derivabile tranne in $(0,0)$ con $\nabla f(x, y) = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right)$ in tutti i punti $(x, y) \neq (0, 0)$, ha un punto di minimo locale in $(0,0)$;
3. la funzione $f(x, y) = x^2 - y^2$, ovunque derivabile con $\nabla f = (2x, -2y)$, non ha né punti di minimo locale né punti di massimo locale. Infatti l'unico punto candidato è $(0,0)$ in cui $\nabla f = (0,0)$, ma non può essere minimo locale in quanto in ogni intorno sferico $B_\delta(0,0)$ ci sono dei punti $(x, y) \in B_\delta(0,0)$ tali che $f(x, y) \leq f(0,0)$ (è sufficiente considerare i punti dell'asse delle y) e non può neppure essere massimo locale in quanto in ogni intorno sferico $B_\delta(0,0)$ ci sono dei punti $(x, y) \in B_\delta(0,0)$ tali che $f(x, y) \geq f(0,0)$ (è sufficiente considerare i punti dell'asse delle x).

L'ultimo caso illustrato è un tipico esempio di un *punto di sella* (la forma del grafico delle funzione $f(x, y) = x^2 - y^2$ giustifica questo nome).

Definizione 24 Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Un punto critico $x_0 \in A$ è detto punto di sella se per ogni intorno sferico $B_\delta(x_0)$ esistono due punti $x, x' \in B_\delta(x_0) \cap A$ tali che $f(x) > f(x_0)$ e $f(x') < f(x_0)$.

Una volta individuati i punti candidati ad essere massimo locale, minimo locale o sella, il secondo problema è quello di determinare di che tipo sono. In altre parole in un intorno di x_0 , dobbiamo studiare il segno di $f(x) - f(x_0) \equiv \Delta(x)$. in particolare

- se esiste un intorno $U_\delta(x_0)$ tale che per ogni $x \in U_\delta(x_0)$ si ha $\Delta(x) \geq 0$ allora x_0 è un punto di minimo locale;
- se esiste un intorno $U_\delta(x_0)$ tale che per ogni $x \in U_\delta(x_0)$ si ha $\Delta(x) \leq 0$ allora x_0 è un punto di massimo locale;
- se per ogni intorno $U_\delta(x_0)$ esistono almeno due punti $x, x' \in U_\delta(x_0)$ tali che $\Delta(x) > 0$ e $\Delta(x') < 0$ allora x_0 è un punto di sella.

Occupiamoci ora del caso particolare in cui $n = 2$ e consideriamo una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe $C^2(A)$ e un punto (x_0, y_0) tale che $\nabla f(x_0, y_0) = (0, 0)$. Utilizzando ora la formula di Taylor abbiamo che

$$\Delta(x, y) = f(x, y) - f(x_0, y_0) = \frac{1}{2}(x-x_0, y-y_0)H_f(x_0, y_0) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + o(\|(x-x_0, y-y_0)\|^2)$$

Notiamo che a causa di questa formula, che esprime un'approssimazione di f in un intorno di (x_0, y_0) , tanto migliore quanto più il punto (x, y) è vicino a (x_0, y_0) , il comportamento locale di f , in particolare il segno dello scarto $\Delta(x, y)$, dipende dal segno del termine $\frac{1}{2}(x-x_0, y-y_0)H_f(x_0, y_0) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$. Per alleggerire la notazione è utile introdurre le componenti h, k del vettore $(x-x_0, y-y_0)$, in modo tale che la formula precedente diviene

$$\Delta(x_0+h, y_0+k) = f(x_0+h, y_0+k) - f(x_0, y_0) = \frac{1}{2}(h, k)H_f(x_0, y_0) \begin{pmatrix} h \\ k \end{pmatrix} + o(\|(h, k)\|^2)$$

Se la matrice hessiana ha coefficienti $H_f(x_0, y_0) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ (ricordiamo che per il teorema di Schwartz la matrice Hessiana di una funzione di classe C^2 è simmetrica) allora lo scarto Δ può essere scritto nella forma

$$\Delta = \frac{1}{2}(h, k) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = \frac{1}{2}(a h^2 + 2b h k + c k^2).$$

Ricordiamo che data una generica matrice simmetrica M , esiste una matrice unitaria $U = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$, con $U^T = U^{-1}$, tale che $M = U^T D U$, dove D è la matrice diagonale $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, dove $\lambda_1, \lambda_2 \in \mathbb{R}$ sono i due autovalori della matrice M . Questo significa che se introduciamo in \mathbb{R}^2 delle nuove coordinate h', k' , definite da

$$\begin{pmatrix} h' \\ k' \end{pmatrix} = U \begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} \cos \theta h + \sin \theta k \\ -\sin \theta h + \cos \theta k \end{pmatrix}$$

collegate alle coordinate cartesiane (x, y) dalla relazione

$$x = x_0 + \cos \theta h' - \sin \theta k', \quad y = y_0 + \sin \theta h' + \cos \theta k',$$

lo studio del segno dello scarto Δ in un intorno del punto (x_0, y_0) più semplice in quanto abbiamo:

$$f(x, y) - f(x_0, y_0) = f(x_0 + \cos \theta h' - \sin \theta k', y_0 + \sin \theta h' + \cos \theta k') - f(x_0, y_0) = \lambda_1 (h')^2 + \lambda_2 (k')^2.$$

Inoltre abbiamo che la distanza tra il punto (x_0, y_0) ed il punto (x, y) è data da:

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = \sqrt{(h')^2 + (k')^2}$$

Se entrambi gli autovalori λ_1, λ_2 della matrice M sono non nulli, allora siamo in grado di controllare con le informazioni in nostro possesso il comportamento locale della funzione f in un intorno del punto critico (x_0, y_0) , in particolare:

- se $\lambda_1 > 0$ e $\lambda_2 > 0$, allora il comportamento locale della funzione è quello illustrato in fig. 1. In questo caso il punto (x_0, y_0) è un punto di minimo locale.

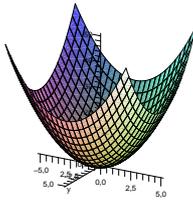


fig.1: punto di minimo locale

- se $\lambda_1 < 0$ e $\lambda_2 < 0$, allora il comportamento locale della funzione è quello illustrato in fig. 2. In questo caso il punto (x_0, y_0) è un punto di massimo locale.

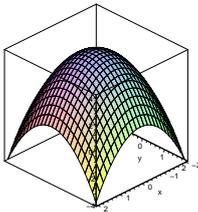


fig.2: punto di massimo locale

- se $\lambda_1 > 0$ e $\lambda_2 < 0$, allora il comportamento locale della funzione è quello illustrato in fig. 3. In questo caso il punto (x_0, y_0) è un punto di sella.

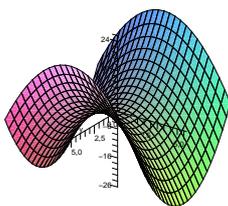


fig.3: punto di sella.

Se almeno uno dei due autovalori è nullo, allora non siamo in grado di controllare, solo tramite lo studio della matrice Hessiana, il comportamento della funzione f in un intorno di (x_0, y_0) *nella direzione dell'autovettore relativo all'autovalore nullo*. In linea di principio dovremmo andare a studiare le derivate di f di ordine superiore a 2. Questo è il limite intrinseco di questo tipo di tecnica.

Generalizziamo ora questo tipo di discussione al caso di funzioni $f : \mathbb{R}^n \rightarrow \mathbb{R}$ di n variabili reali. Supponiamo che f sia di classe C^2 e sia $x_0 \in \mathbb{R}^n$ un punto critico, ovvero $\nabla f(x_0) = 0$. Grazie alla formula di Taylor è possibile stimare lo scarto $\Delta(x) := f(x) - f(x_0)$ in termini delle derivate seconde di f :

$$f(x) - f(x_0) = \frac{1}{2}(x - x_0)H_f(x_0)(x - x_0) + o(\|x - x_0\|^2).$$

- Se la matrice hessiana è *definita positiva*, ovvero se tutti i suoi autovalori sono strettamente positivi, allora x_0 è un punto di minimo locale
- Se la matrice hessiana è *definita negativa*, ovvero se tutti i suoi autovalori sono strettamente negativi, allora x_0 è un punto di massimo locale
- Se la matrice hessiana è *indefinita*, ovvero se possiede almeno due autovalori di segno discorde, allora x_0 è un punto di sella
- Se la matrice hessiana è *semidefinita positiva*, ovvero se tutti i suoi autovalori sono positivi ed è presente almeno un autovalore nullo, allora non è possibile stabilire solo tramite l'analisi dell'hessiana la natura del punto stazionario x_0
- Se la matrice hessiana è *semidefinita negativa*, ovvero se tutti i suoi autovalori sono negativi ed è presente almeno un autovalore nullo, allora non è possibile stabilire solo tramite l'analisi dell'hessiana la natura del punto stazionario x_0

Il seguente criterio permette di verificare se una matrice simmetrica $n \times n$ è definita positiva o definita negativa, senza dover calcolare esplicitamente gli autovalori.

Teorema 14 (criterio dei minori nord-ovest) Sia M una matrice simmetrica $n \times n$ e indichiamo con M_k , con $k = 1, \dots, n$, le n sottomatrici composte mediante le prime k righe e k colonne di M (tali matrici sono chiamate minori principali di nord-ovest). Allora

- i. M è definita positiva se e solo se $\det M_k > 0$ per ogni $k = 1, \dots, n$.
- ii. M è definita negativa se e solo se $(-1)^k \det M_k > 0$ per ogni $k = 1, \dots, n$.

Massimi e minimi assoluti

In questa sezione descriveremo dei criteri per individuare punti di massimo e di minimo assoluto per una funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ su un sottoinsieme $A \subset \mathbb{R}^n$.

Definizione 25 Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $A \subset \mathbb{R}^n$.

Un punto $x_0 \in A$ è detto punto di massimo assoluto per f sull'insieme A se per ogni $x \in A$ si ha $f(x) \leq f(x_0)$. Analogamente un punto $x_0 \in A$ è detto punto di minimo assoluto per f sull'insieme A se per ogni $x \in A$ si ha $f(x) \geq f(x_0)$.

Non è detto in generale che data una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e un insieme $A \subset \mathbb{R}^n$, esistano i punti di massimo e/o minimo assoluti di f su A . Ad esempio, nel caso in cui $n = 1$, la funzione $f(x) = x$ sull'insieme $A = \mathbb{R}$ non ha punti di massimo o minimo assoluto. Lo stesso avviene, ad esempio, per la funzione $f(x) = \tan x$ sull'insieme $A = (-\pi/2, \pi/2)$.

Diamo qui di seguito delle condizioni sufficienti per l'esistenza dei punti di massimo e di minimo assoluto.

Teorema 15 (di Weierstrass) Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione continua e A un insieme chiuso e limitato. Allora f ammette massimo e minimo assoluto su A .

Ricordiamo che un sottoinsieme $A \subset \mathbb{R}^n$ è detto chiuso se coincide con la sua chiusura, ed è limitato se esiste un $R > 0$ tale che $A \subset B_R(0)$ (cioè la norma di ogni elemento $x \in A$ è minore di R).

Secondo il teorema di Weierstrass il problema della ricerca dei punti di massimo e di minimo assoluto di una funzione continua f su un insieme chiuso e limitato A ha sempre soluzione. Diamo qui di seguito un metodo per l'individuazione di tali punti.

1. Cominciamo con la ricerca dei punti di massimo e minimo assoluto all'interno dell'insieme A . Se x_0 è un punto di massimo (o minimo) assoluto per f su A e $x_0 \in \overset{\circ}{A}$, allora x_0 è anche un punto di massimo (rispettivamente minimo) locale di f , quindi si possono presentare solo due possibilità :

- se f è derivabile in x_0 allora $\nabla f(x_0) = 0$,
 - altrimenti f non è derivabile in x_0 .
2. Come secondo passo andiamo a cercare i punti di massimo e di minimo assoluto di f sulla frontiera ∂A di A . Vedremo in dettaglio dei metodi per tale ricerca nel caso $n = 2$.

Alla fine si confrontano i valori di f nei punti individuati al passo 1 ed al passo 2. Si seleziona il valore maggiore, che sarà il valore massimo di f su A . Il punto x_0 in cui f assume tale valore sarà il punto di massimo assoluto di f su A . Analogamente si seleziona il valore minore, che sarà il valore minimo di f su A . Il punto x_0 in cui f assume tale valore sarà il punto di minimo assoluto di f su A .

Estremi vincolati e moltiplicatori di Lagrange

Analizziamo ora in dettaglio il caso in cui $n = 2$ e vogliamo cercare i punti di massimo e minimo assoluto di una funzione f di classe C^1 su una curva regolare $\gamma \subset \mathbb{R}^2$.

Se $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ è una parametrizzazione regolare di γ , con $\alpha(t) = (x(t), y(t))$, allora dobbiamo studiare la funzione composta $f \circ \alpha : I \subset \mathbb{R}$, definita da

$$f \circ \alpha(t) = f(\alpha(t)) = f(x(t), y(t)), \quad t \in I$$

che dà i valori che la funzione f assume nei punti della curva γ . Se $I = [a, b]$, allora i punti di massimo o minimo assoluto di f su γ vanno cercati tra

1. i punti $\alpha(t')$, con $t' \in (a, b)$ tale che $\frac{d}{dt}(f \circ \alpha)(t') = 0$.
2. i punti $\alpha(a)$ e $\alpha(b)$

Alla fine si confrontano i valori di f nei punti individuati al passo 1 ed al passo 2 e si selezionano il maggiore ed il minore.

Se non abbiamo a disposizione una parametrizzazione $\alpha : I \rightarrow \mathbb{R}^2$ per la curva γ , che invece viene descritta come insieme di livello di una funzione $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe C^1 :

$$\gamma = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$$

possiamo utilizzare il metodo dei *moltiplicatori di Lagrange*.

Teorema 16 *Siano $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ e $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ due funzioni di classe C^1 . Se (x_0, y_0) è un punto di massimo o di minimo assoluto di f sull'insieme $\gamma = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$ e se $\nabla g(x_0, y_0) \neq 0$, allora esiste $\lambda \in \mathbb{R}$ tale che*

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

In altre parole, secondo tale teorema, i punti di estremo per f su γ vanno ricercati fra quei punti della curva γ in cui il vettore gradiente di f è proporzionale al vettore gradiente di g .

Per dimostrare tale affermazione supponiamo di avere in mano una parametrizzazione regolare $\alpha : I \rightarrow \mathbb{R}^2$ per la curva γ . Allora i punti candidati ad essere di massimo o minimo per f su γ sono quei punti $\alpha(t')$, con t' tale che $\frac{d}{dt}f \circ \alpha(t') = 0$. Per la formula di derivazione di funzione composta $\frac{d}{dt}f \circ \alpha(t') = \nabla f(\alpha(t')) \cdot \alpha'(t')$. Dobbiamo dunque cercare quei punti $\alpha(t')$ della curva γ in cui $\nabla f(\alpha(t'))$ è ortogonale al vettore tangente $\alpha'(t')$. Ricordiamo che, dato che γ è descritta come curva di livello della funzione g , il vettore ∇g è ortogonale a γ in tutti i punti della curva. Quindi, nel caso in cui $\nabla g \neq 0$, la condizione di ortogonalità di ∇f a γ può essere espressa in termini della condizione $\nabla f = \lambda \nabla g$.

La ricerca dei punti dell'insieme γ in cui $\nabla f = \lambda \nabla g$ può essere effettuata introducendo la funzione $\mathcal{L} : \mathbb{R}^3 \rightarrow \mathbb{R}$ di tre variabili (x, y, λ) definita da

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

e cercando i punti critici di tale funzione, ovvero i punti (x, y, λ) tali che $\nabla \mathcal{L}(x, y, \lambda) = (0, 0, 0)$. Questo porta a risolvere un sistema di tre equazioni in tre incognite della forma

$$\begin{cases} \frac{\partial}{\partial x} \mathcal{L} = \frac{\partial}{\partial x} f(x, y) - \lambda \frac{\partial}{\partial x} g(x, y) = 0 \\ \frac{\partial}{\partial y} \mathcal{L} = \frac{\partial}{\partial y} f(x, y) - \lambda \frac{\partial}{\partial y} g(x, y) = 0 \\ \frac{\partial}{\partial \lambda} \mathcal{L} = -g(x, y) = 0 \end{cases}$$

Le prime due equazioni esprimono la proporzionalità fra ∇f e ∇g , mentre la terza equazione impone che il punto (x, y) che stiamo cercando appartenga all'insieme $\{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$.

Tale risultato è generalizzabile al caso di funzioni di n variabili e in questi casi si rivela particolarmente utile.

Teorema 17 (Moltiplicatori di Lagrange) *Siano $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}$ due funzioni di classe C^1 . Se $x_0 \in \mathbb{R}^n$ è un punto di massimo o di minimo assoluto di f sull'insieme $\Omega = \{x \in \mathbb{R}^n : g(x) = 0\}$ e se $\nabla g(x_0) \neq 0$, allora esiste $\lambda \in \mathbb{R}$ tale che*

$$\nabla f(x_0) = \lambda \nabla g(x_0)$$

Metodo dei minimi quadrati

Supponiamo di sapere che due grandezze, che indicheremo con x e y , siano collegate da una relazione di tipo lineare affine della forma $y = \alpha + \beta x$, con $\alpha, \beta \in \mathbb{R}$.

Esempi di tale situazione sono: tempo trascorso e spazio percorso da un punto materiale che si muove di moto rettilineo uniforme; altezza e peso degli individui di una popolazione; numero di pezzi prodotti e costo di produzione in un processo industriale. Supponiamo di aver raccolto un certo numero n di dati sperimentali (x_i, y_i) con $i = 1, \dots, n$, ovvero di osservazioni congiunte delle due variabili, che rappresentiamo in un sistema di assi cartesiani x, y e cerchiamo la retta che passi "più vicino possibile" ai punti (x_i, y_i) . Supponiamo che $n > 2$ e che i valori x_1, x_2, \dots, x_n non siano tutti uguali fra loro. Data una generica retta di equazione $y = \alpha + \beta x$, stimiamo la "distanza" fra tale retta e i punti tramite la seguente quantità :

$$E(\alpha, \beta) := \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$$

Cerchiamo valori dei parametri $(\alpha, \beta) \in \mathbb{R}^2$ che minimizzino la funzione E . In altre parole cerchiamo il punto di minimo assoluto della funzione $E : \mathbb{R}^2 \rightarrow \mathbb{R}$.

Troviamo i valori di (α, β) tali per cui $\nabla E(\alpha, \beta) = (0, 0)$. Questi sono le soluzioni del sistema lineare di 2 equazioni in 2 incognite della forma:

$$\begin{cases} \frac{\partial E}{\partial \alpha} = \sum_{i=1}^n 2(\beta x_i + \alpha - y_i) = 0 \\ \frac{\partial E}{\partial \beta} = \sum_{i=1}^n 2x_i(\beta x_i + \alpha - y_i) = 0 \end{cases}$$

cioé

$$\begin{cases} \beta \sum_{i=1}^n x_i + n\alpha = \sum_{i=1}^n y_i \\ \beta \sum_{i=1}^n x_i^2 + \alpha \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i \end{cases}$$

Notiamo che il sistema ammette una e un'unica soluzione in quanto la matrice dei coefficienti ha determinante non nullo⁴, infatti:

$$\det \begin{pmatrix} \sum_{i=1}^n x_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \end{pmatrix} = \left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2 = -n \sum_{i=1}^n \left(x_i - \frac{\sum_{j=1}^n x_j}{n} \right)^2 < 0$$

La soluzione è data da:

$$\begin{aligned} \bar{\alpha} &= \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \\ \bar{\beta} &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \end{aligned}$$

⁴di fatto l'unico caso in cui il determinante della matrice dei coefficienti può annullarsi è quello in cui la coordinata x di tutti i punti è costante (cioé $x_i = \bar{X}$ per ogni $i = 1..n$). In questo caso tutti gli n punti sperimentali sono allineati su una retta verticale. Lavoriamo nell'ipotesi che si possa escludere tale situazione.

Notiamo che i valori trovati di α e β sono effettivamente di minimo assoluto per la funzione E , infatti sviluppando tale funzione in serie di Taylor attorno a $(\alpha, \beta) = (\bar{\alpha}, \bar{\beta})$ otteniamo

$$E(\alpha, \beta) = E(\bar{\alpha}, \bar{\beta}) + \frac{1}{2}(\alpha - \bar{\alpha}, \beta - \bar{\beta})H_E(\bar{\alpha}, \bar{\beta}) \begin{pmatrix} \alpha - \bar{\alpha} \\ \beta - \bar{\beta} \end{pmatrix}$$

Dato che E è un polinomio di 2 grado nelle variabili (α, β) , il secondo membro non è un'approssimazione locale per la funzione E , ma un'espressione esatta che rappresenta E in funzione di $(\alpha - \bar{\alpha}, \beta - \bar{\beta})$. Inoltre notiamo che la matrice Hessiana di E , data da

$$H_E = \begin{pmatrix} 2n & 2 \sum_{i=1}^n x_i \\ 2 \sum_{i=1}^n x_i & 2 \sum_{i=1}^n x_i^2 \end{pmatrix}$$

è definita positiva in quanto $2n > 0$ e $\det H_E = 4(n \sum_{i=1}^n x_i - (\sum_{i=1}^n x_i)^2) > 0$. Possiamo dunque concludere che $(\bar{\alpha}, \bar{\beta})$ è un punto di minimo assoluto per E . I valori ottenuti $\bar{\alpha}$ e $\bar{\beta}$ sono i coefficienti della retta di regressione lineare.

Un'altra interessante applicazione del metodo dei minimi quadrati è la seguente. Sia data una funzione $f : [a, b] \rightarrow \mathbb{R}$ e supponiamo di voler cercare il polinomio P di grado k della forma

$$P(x) = A_k x^k + A_{k-1} x^{k-1} + \dots + A_1 x + A_0$$

con A_0, \dots, A_k coefficienti reali, in modo da minimizzare la seguente quantità :

$$E(A_0, \dots, A_k) := \int_a^b (f(x) - P(x))^2 dx = \int_a^b (f(x) - A_k x^k - A_{k-1} x^{k-1} - \dots - A_1 x - A_0)^2 dx.$$

$E : \mathbb{R}^{k+1} \rightarrow \mathbb{R}$ è una funzione di $k + 1$ variabili reali A_0, \dots, A_k . Cerchiamo i valori di tali coefficienti in modo da minimizzare E . Imponendo che $\nabla E = 0$, otteniamo un sistema di $k + 1$ equazioni in $k + 1$ incognite

$$\begin{cases} -2 \int_a^b (f(x) - A_k x^k - A_{k-1} x^{k-1} - \dots - A_1 x - A_0) dx = 0 \\ -2 \int_a^b (f(x) - A_k x^k - A_{k-1} x^{k-1} - \dots - A_1 x - A_0) x dx = 0 \\ \dots \\ \dots \\ -2 \int_a^b (f(x) - A_k x^k - A_{k-1} x^{k-1} - \dots - A_1 x - A_0) x^k dx = 0 \end{cases}$$

la cui soluzione A_0, \dots, A_k fornisce i coefficienti del polinomio di ordine k "più vicino" a f nell'intervallo $[a, b]$ nel senso che minimizza la funzione E . Di fatto per mostrare

che i valori di (A_0, \dots, A_k) ottenuti forniscono effettivamente il punto di minimo assoluto di E su \mathbb{R}^{k+1} è sufficiente notare che la funzione E è un polinomio di secondo grado nelle variabili (A_0, \dots, A_k) . Ha un unico punto critico e la matrice Hessiana H è una matrice simmetrica con elementi $H_{ij} = 2 \int_a^b x^{i+j} dx$. Possiamo verificare che H è definita positiva mostrando che per ogni vettore $v \in \mathbb{R}^{k+1}$ si ha che $v \cdot Hv \geq 0$ e $v \cdot Hv = 0$ se e solo se $v = 0$. Se indichiamo con $v_i, i = 0, 1, \dots, k$, le componenti di un generico vettore $v \in \mathbb{R}^{k+1}$, abbiamo:

$$v \cdot Hv = \sum_{i,j} v_i H_{ij} v_j = \int_a^b \sum_{i,j} v_i v_j x^{i+j} dx = \int_a^b \sum_{i,j} x^i v_i x^j v_j dx = \int_a^b (\sum_i x^i v_i)^2 dx$$

l'integrale al secondo membro della funzione polinomiale positiva $g(x) := (\sum_i x^i v_i)^2$ è sempre maggiore di 0 e si annulla se e solo se il polinomio $\sum_i x^i v_i$ è uguale a 0 per ogni $x \in [a, b]$, cioè se $v_i = 0$ per ogni $i = 0, 1, \dots, k$. Possiamo dunque concludere che il punto stazionario (A_0, \dots, A_k) ottenuto è un punto di minimo assoluto per la funzione E su \mathbb{R}^{k+1} .

Esempio 22 Si consideri la funzione $f(x) = \cos(x)$ sull'intervallo $[-\pi/2, \pi/2]$. Determinare il miglior polinomio di 2 grado $P(x) = ax^2 + bx + c$ che approssimi la f su $[-\pi/2, \pi/2]$, ovvero quello che minimizzi

$$E(a, b, c) := \int_{-\pi/2}^{\pi/2} (\cos x - ax^2 - bx - c)^2 dx.$$

Imponendo che $\nabla f(a, b, c) = (0, 0, 0)$ otteniamo un sistema di tre equazioni nelle tre incognite a, b, c :

$$\begin{cases} \int_{-\pi/2}^{\pi/2} (\cos(x) - ax^2 - bx - c) x^2 dx = 0 \\ \int_{-\pi/2}^{\pi/2} (\cos(x) - ax^2 - bx - c) x dx = 0 \\ \int_{-\pi/2}^{\pi/2} (\cos(x) - ax^2 - bx - c) dx = 0 \end{cases}$$

ovvero

$$\begin{cases} a \int_{-\pi/2}^{\pi/2} x^4 dx + b \int_{-\pi/2}^{\pi/2} x^3 dx + c \int_{-\pi/2}^{\pi/2} x^2 dx = \int_{-\pi/2}^{\pi/2} x^2 \cos x dx \\ a \int_{-\pi/2}^{\pi/2} x^3 dx + b \int_{-\pi/2}^{\pi/2} x^2 dx + c \int_{-\pi/2}^{\pi/2} x dx = \int_{-\pi/2}^{\pi/2} x \cos x dx \\ a \int_{-\pi/2}^{\pi/2} x^2 dx + b \int_{-\pi/2}^{\pi/2} x dx + c \int_{-\pi/2}^{\pi/2} 1 dx = \int_{-\pi/2}^{\pi/2} \cos x dx \end{cases}$$

dato che:

$$\int_{-\pi/2}^{\pi/2} 1 dx = \pi, \int_{-\pi/2}^{\pi/2} x dx = 0, \int_{-\pi/2}^{\pi/2} x^2 dx = \frac{\pi^3}{12}, \int_{-\pi/2}^{\pi/2} x^3 dx = 0, \int_{-\pi/2}^{\pi/2} x^4 dx = \frac{\pi^5}{80},$$

$\int_{-\pi/2}^{\pi/2} \cos x dx = 2$, $\int_{-\pi/2}^{\pi/2} x \cos x dx = 0$, $\int_{-\pi/2}^{\pi/2} x^2 \cos x dx = \frac{\pi^2}{2} - 4$
il sistema diviene:

$$\begin{cases} \frac{\pi^5}{80}a + \frac{\pi^3}{12}c = \frac{\pi^2}{2} - 4 \\ \frac{\pi}{12}b = 0 \\ \frac{\pi^3}{12}a + \pi c = 2 \end{cases}$$

La soluzione è $a = \frac{60(\pi^2-12)}{\pi^5}$, $b = 0$ e $c = -3\frac{\pi^2-20}{\pi^3}$. Il polinomio cercato è dunque

$$P(x) = \frac{60(\pi^2 - 12)}{\pi^5}x^2 - 3\frac{\pi^2 - 20}{\pi^3}$$

Funzioni definite implicitamente

Nella sezione precedente abbiamo introdotto il metodo dei moltiplicatori di Lagrange, che fornisce un criterio per la ricerca dei punti di massimo e minimo assoluto di una funzione su un insieme $\Omega \subset \mathbb{R}^n$ definito come insieme di livello di una funzione $g : \mathbb{R}^n \rightarrow \mathbb{R}$, cioè $\Omega = \{x \in \mathbb{R}^n : g(x) = 0\}$, nell'ipotesi che $g \in C^1(\mathbb{R}^n)$ e $\nabla g \neq 0$. In questa sezione cerchiamo di capire come le proprietà della funzione g ci permettono di caratterizzare l'insieme Ω , identificandolo nel caso $n = 2$ con una curva regolare (localmente grafico di funzione $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$) e nel caso $n = 3$ con una superficie (localmente grafico di una funzione $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$).

Cominciamo considerando il caso in cui $n = 2$

Teorema 18 (di Dini, della funzione implicita) *Sia $g : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, con A insieme aperto, una funzione di classe $C^1(A)$. Si consideri l'insieme $\Omega = \{(x, y) \in A : g(x, y) = 0\}$ e sia $(x_0, y_0) \in \Omega$ un punto tale che $\nabla g(x_0, y_0) \neq (0, 0)$. Allora*

- Se $\frac{\partial}{\partial y}g(x_0, y_0) \neq 0$, allora esiste un intorno I di x_0 e un'unica funzione $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ tale che $y_0 = f(x_0)$ e $g(x, f(x)) = 0$ per ogni $x \in I$. Inoltre f è una funzione di classe $C^1(I)$ e

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

- Se $\frac{\partial}{\partial x}g(x_0, y_0) \neq 0$, allora esiste un intorno J di y_0 e un'unica funzione $h : J \subset \mathbb{R} \rightarrow \mathbb{R}$ tale che $x_0 = h(y_0)$ e $g(h(y), y) = 0$ per ogni $y \in J$. Inoltre h è una funzione di classe $C^1(J)$ e

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

Il teorema del Dini fornisce quindi delle condizioni sufficienti per identificare gli insiemi di livello di una funzione $g \in C^1(A)$ con archi di curve regolari. Non è detto che questo sia possibile in un intorno dei punti critici di g , cioè i punti in cui $\nabla g = (0, 0)$.

Esempio 23 *L'insieme di livello $\Omega = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$ della funzione $g(x, y) = x^2 + y^2 - 1$ è una circonferenza. Le ipotesi del teorema del Dini sono verificate in tutti i punti di Ω .*

Notiamo che in un intorno del punto $(x_0, y_0) = (1, 0)$ o del punto $(x_0, y_0) = (-1, 0)$, in cui $\frac{\partial}{\partial y}g(x_0, y_0) = 0$ ma $\frac{\partial}{\partial x}g(x_0, y_0) \neq 0$, è possibile rappresentare una parte dell'insieme Ω come grafico di funzione $x = h(y)$, dove $h : (-1, 1) \rightarrow \mathbb{R}$ è la funzione definita da $h(y) = \sqrt{1 - y^2}$ nel caso in cui consideriamo il punto $(1, 0)$, mentre $h(y) = -\sqrt{1 - y^2}$ nel caso in cui consideriamo il punto $(-1, 0)$.

In un intorno del punto $(x_0, y_0) = (0, 1)$ o del punto $(x_0, y_0) = (0, -1)$, in cui $\frac{\partial}{\partial x}g(x_0, y_0) = 0$ ma $\frac{\partial}{\partial y}g(x_0, y_0) \neq 0$, è possibile rappresentare una parte dell'insieme Ω come grafico di funzione $y = f(x)$, dove $f : (-1, 1) \rightarrow \mathbb{R}$ è la funzione definita da $f(x) = \sqrt{1 - x^2}$ nel caso in cui consideriamo il punto $(0, 1)$, mentre $f(x) = -\sqrt{1 - x^2}$ nel caso in cui consideriamo il punto $(0, -1)$.

In tutto gli altri punti (x_0, y_0) della circonferenza entrambe le derivate parziali di g sono non nulle ed è possibile rappresentare un sottoinsieme di Ω contenente (x_0, y_0) sia come grafico di funzione $y = f(x)$, sia come grafico di funzione $x = h(y)$.

Esempio 24 *L'insieme di livello $\Omega = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$ della funzione $g(x, y) = x^2 - y^2$ è formato da due rette (la retta $y = x$ e la retta $y = -x$). Le ipotesi del teorema del Dini sono verificate in tutti i punti di Ω , tranne che nell'origine, in cui $\nabla g(0, 0) = (0, 0)$. Di fatto in un intorno di $(0, 0)$ l'insieme Ω non è una curva regolare.*

Trattiamo ora in dettaglio il caso in cui $n = 3$.

Teorema 19 (Dini, caso 3-dimensionale) *Sia $g : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, con A insieme aperto, una funzione di classe $C^1(A)$. Si consideri l'insieme $\Omega = \{(x, y, z) \in A : g(x, y, z) = 0\}$ e sia $(x_0, y_0, z_0) \in \Omega$ un punto tale che $\nabla g(x_0, y_0, z_0) \neq (0, 0, 0)$. Allora*

- *Se $\frac{\partial}{\partial z}g(x_0, y_0, z_0) \neq 0$, allora esiste un intorno $U \subset \mathbb{R}^2$ di (x_0, y_0) e un'unica funzione $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che $z_0 = f(x_0, y_0)$ e $g(x, y, f(x, y)) = 0$ per ogni $(x, y) \in U$. Inoltre f è una funzione di classe $C^1(U)$ le cui derivate parziali sono date da:*

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

- Se $\frac{\partial}{\partial y}g(x_0, y_0, z_0) \neq 0$, allora esiste un intorno $V \subset \mathbb{R}^2$ di (x_0, z_0) e un'unica funzione $h : V \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che $y_0 = h(x_0, z_0)$ e $g(x, h(x, z), z) = 0$ per ogni $(x, z) \in V$. Inoltre h è una funzione di classe $C^1(V)$ le cui derivate parziali sono date da:

$$\frac{\partial}{\partial x}h(x, z) = -\frac{\frac{\partial g}{\partial x}(x, h(x, z), z)}{\frac{\partial g}{\partial y}(x, h(x, z), z)}, \quad \frac{\partial}{\partial z}h(x, z) = -\frac{\frac{\partial g}{\partial z}(x, h(x, z), z)}{\frac{\partial g}{\partial y}(x, h(x, z), z)}$$

- Se $\frac{\partial}{\partial x}g(x_0, y_0, z_0) \neq 0$, allora esiste un intorno $W \subset \mathbb{R}^2$ di (y_0, z_0) e un'unica funzione $k : W \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che $x_0 = k(y_0, z_0)$ e $g(k(y, z), y, z) = 0$ per ogni $(y, z) \in W$. Inoltre k è una funzione di classe $C^1(W)$ le cui derivate parziali sono date da:

$$\frac{\partial}{\partial y}k(y, z) = -\frac{\frac{\partial g}{\partial y}(k(y, z), y, z)}{\frac{\partial g}{\partial x}(k(y, z), y, z)}, \quad \frac{\partial}{\partial z}k(y, z) = -\frac{\frac{\partial g}{\partial z}(k(y, z), y, z)}{\frac{\partial g}{\partial x}(k(y, z), y, z)}$$

Esempio 25 L'insieme di livello $\Omega = \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 0\}$ della funzione $g(x, y, z) = x^2 + y^2 + z^2 - 1$ è la superficie sferica di centro $(0, 0, 0)$ e raggio 1. Le ipotesi del teorema del Dini sono verificate in tutti i punti di Ω .

Notiamo che in un intorno del punto $(x_0, y_0, z_0) = (0, 0, 1)$ in cui $\frac{\partial}{\partial y}g(x_0, y_0, z_0) = \frac{\partial}{\partial x}g(x_0, y_0, z_0) = 0$ ma $\frac{\partial}{\partial z}g(x_0, y_0, z_0) \neq 0$, è possibile rappresentare una parte dell'insieme Ω come grafico di funzione $z = f(x, y)$, dove $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ è la funzione definita da $f(x, y) = \sqrt{1 - x^2 - y^2}$ e $U = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$.

In un intorno del punto $(x_0, y_0, z_0) = (0, 1, 0)$ in cui $\frac{\partial}{\partial x}g(x_0, y_0, z_0) = \frac{\partial}{\partial z}g(x_0, y_0, z_0) = 0$ ma $\frac{\partial}{\partial y}g(x_0, y_0, z_0) \neq 0$, è possibile rappresentare una parte dell'insieme Ω come grafico di funzione $y = h(x, z)$, dove $h : V \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ è la funzione definita da $h(x, z) = \sqrt{1 - x^2 - z^2}$ e $V = \{(x, z) \in \mathbb{R}^2 : x^2 + z^2 < 1\}$.

Esempio 26 L'insieme di livello $\Omega = \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 0\}$ della funzione $g(x, y, z) = x^2 + y^2 - z^2$ è formato da due coni con vertice nell'origine (descritti rispettivamente dalle equazioni $z = \sqrt{x^2 + y^2}$ e $z = -\sqrt{x^2 + y^2}$). Le ipotesi del teorema del Dini sono verificate in tutti i punti di Ω , tranne che nell'origine, in cui $\nabla g(0, 0, 0) = (0, 0, 0)$. Di fatto in un intorno di $(0, 0, 0)$ l'insieme Ω non può essere rappresentato come grafico di un'unica funzione.

Integrali doppi

Ricordiamo che data una funzione continua $f : [a, b] \rightarrow \mathbb{R}$ l'integrale di f sull'intervallo $[a, b]$ è dato dal limite delle *somme di Cauchy-Riemann*, ovvero

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i) \frac{(b-a)}{n}$$

dove la somma $s_n = \sum_{i=1}^n f(\xi_i) \frac{(b-a)}{n}$ fornisce un'approssimazione del valore dell'integrale $\int_a^b f(x)dx$ ed è costruita suddividendo l'intervallo $[a, b]$ in n sottointervalli di ampiezza $(b-a)/n$ e di estremi $[x_i, x_{i+1}]$, dove $x_i = a + i(b-a)/n$, $i = 0, \dots, n$. In ogni sottointervallo $[x_i, x_{i+1}]$ si sceglie un punto $\xi_i \in [x_i, x_{i+1}]$ e si costruisce il rettangolo con base data dall'intervallo $[x_i, x_{i+1}]$ e con altezza pari a $f(\xi_i)$. La somma $s_n = \sum_{i=1}^n f(\xi_i) \frac{(b-a)}{n}$ rappresenta la somma delle aree dei rettangoli (con segno). In particolare, se $f \geq 0$, l'integrale $\int_a^b f(x)dx$ rappresenta l'area sottesa al grafico di f , ovvero l'area dell'insieme $B \subset \mathbb{R}^2$ definito da

$$B := \{(x, y) \in \mathbb{R}^2 : x \in [a, b], 0 \leq y \leq f(x)\},$$

Più in generale, data una funzione $f : [a, b] \rightarrow \mathbb{R}$ *limitata* (cioè tale che esistono due valori $m, M \in \mathbb{R}$ per cui $m \leq f(x) \leq M$ per ogni $x \in [a, b]$) si dice che f è integrabile secondo Riemann su $[a, b]$ se esiste finito $\lim_{n \rightarrow \infty} s_n$ e inoltre tale limite non dipende dalla scelta dei punti $\xi_i \in [x_i, x_{i+1}]$. In tal caso $\int_a^b f(x)dx := \lim_{n \rightarrow \infty} s_n$.

Generalizziamo ora tale definizione al caso di funzioni di due variabili reali. Consideriamo un rettangolo $R \subset \mathbb{R}^2$, con $R = [a, b] \times [c, d]$, e una funzione $f : R \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ limitata. Suddividiamo il rettangolo R in n^2 rettangoli di area $(b-a)(d-c)/n^2$, ottenuti dividendo ognuno degli intervalli $[a, b]$ e $[c, d]$ in n sottointervalli di estremi rispettivamente $[x_i, x_{i+1}]$, con $x_i = a + i(b-a)/n$ e $i = 0, \dots, n$, e $[y_k, y_{k+1}]$, con $y_k = c + k(d-c)/n$ e $k = 0, \dots, n$. Indichiamo, per ogni $i = 1, \dots, n$ ed ogni $k = 1, \dots, n$, con R_{ik} il sottorettangolo $R_{ik} := [x_{i-1}, x_i] \times [y_{k-1}, y_k]$ e con $\xi_{ik} \in R_{ik}$ un generico punto scelto all'interno di R_{ik} . Costruiamo per ogni $i = 1, \dots, n$ e $k = 1, \dots, n$ il parallelepipedo di base R_{ik} ed altezza $f(\xi_{ik})$ e calcoliamo la somma dei volumi (con segno) degli n^2 parallelepipedi, ottenendo:

$$s_n := \sum_{i,k=1}^n f(\xi_{ik}) \frac{(b-a)(d-c)}{n^2}$$

Se esiste finito il limite per $n \rightarrow \infty$ di tali somme e tale valore non dipende dalla scelta dei punti $\xi_{ik} \in R_{ik}$, allora tale valore definisce l'integrale di f sul rettangolo R .

Definizione 26 Una funzione $f : R \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ limitata è detta integrabile su R se esiste finito

$$\lim_{n \rightarrow \infty} \sum_{i,k=1}^n f(\xi_{ik}) \frac{(b-a)(d-c)}{n^2} \quad (25)$$

e tale valore non dipende dalla scelta di $\xi_{ik} \in R_{ik}$. In tal caso viene indicato con il simbolo $\int_R f(x,y) dx dy$

Notiamo che, nel caso in cui $f \geq 0$ il valore $\int \int_R f(x,y) dx dy$ ha il significato geometrico di volume sotteso al grafico di f , cioè il volume dell'insieme $B \subset \mathbb{R}^3$:

$$B = \{(x,y,z) \in \mathbb{R}^3 : x \in [a,b], y \in [c,d], 0 \leq z \leq f(x,y)\}. \quad (26)$$

Esempio 27 La seguente funzione definita sul quadrato $[0,1] \times [0,1]$ non è integrabile:

$$f(x,y) := \begin{cases} 1 & \text{se } x \in \mathbb{Q} \\ 0 & \text{altrimenti} \end{cases}$$

infatti il valore del limite (25) non è indipendente dalla scelta del punto $\xi_{ik} \in R_{ik}$. Infatti se scegliamo per ogni $i = 1, \dots, n$ e $k = 1, \dots, n$ un punto $\xi_{ik} \equiv (x_{ik}, y_{ik}) \in R_{ik}$ tale che $x_{ik} \in \mathbb{Q}$ allora

$$s_n = \sum_{i,k=1}^n f(\xi_{ik}) \frac{(b-a)(d-c)}{n^2} = \sum_{i,k=1}^n \frac{(b-a)(d-c)}{n^2} = 1$$

e $\lim_{n \rightarrow \infty} s_n = 1$. Se invece scegliamo per ogni $i = 1, \dots, n$ e $k = 1, \dots, n$ un punto $\xi_{ik} \equiv (x_{ik}, y_{ik}) \in R_{ik}$ tale che $x_{ik} \notin \mathbb{Q}$ allora

$$s_n = \sum_{i,k=1}^n f(\xi_{ik}) \frac{(b-a)(d-c)}{n^2} = 0$$

e $\lim_{n \rightarrow \infty} s_n = 0$.

Vale il seguente teorema, che assicura la possibilità di integrare una funzione continua su un dominio di integrazione rettangolare.

Teorema 20 Sia $f : [a,b] \times [c,d] \rightarrow \mathbb{R}$ una funzione continua. Allora f è integrabile su $R = [a,b] \times [c,d]$.

Descriviamo ora una tecnica che permette di ricondurre il calcolo di un integrale doppio, ovvero di un integrale di una funzione di due variabili reali, al calcolo di due integrali semplici iterati. Scelto un valore $\bar{x} \in [a,b]$, Indichiamo con $S_{\bar{x}}$ la

sezione dell'insieme B , definito in (26) con il piano verticale parallelo al piano yz di equazione $x = \bar{x}$:

$$S_{\bar{x}} = \{(x, y, z) \in \mathbb{R}^3 : x = \bar{x}, y \in [c, d], 0 \leq z \leq f(\bar{x}, y)\},$$

la cui area (con segno) è data da $A(S_{\bar{x}}) = \int_c^d f(\bar{x}, y)dy$. Analogamente, scelto un valore $\bar{y} \in [c, d]$, Indichiamo con $S_{\bar{y}}$ la sezione dell'insieme B , definito in (26) con il piano verticale parallelo al piano xz di equazione $y = \bar{y}$:

$$S_{\bar{y}} = \{(x, y, z) \in \mathbb{R}^3 : y = \bar{y}, x \in [a, b], 0 \leq z \leq f(x, \bar{y})\},$$

la cui area (con segno) è data da $A(S_{\bar{y}}) = \int_a^b f(x, \bar{y})dx$.

Teorema 21 (metodo di riduzione, per un rettangolo) Sia $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ una funzione continua. Allora le funzioni

$$x \mapsto A(S_x) = \int_c^d f(x, y)dy, \quad x \in [a, b]$$

$$y \mapsto A(S_y) = \int_a^b f(x, y)dx, \quad y \in [c, d]$$

sono continue e

$$\begin{aligned} \int \int_{[a,b] \times [c,d]} f(x, y) dx dy &= \int_a^b A(s_x) dx = \int_a^b \left(\int_c^d f(x, y) dy \right) dx \\ &= \int_c^d A(s_y) dy = \int_c^d \left(\int_a^b f(x, y) dx \right) dy \end{aligned}$$

In particolare se la funzione $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ è fattorizzabile, cioè della forma

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

dove $g : [a, b] \rightarrow \mathbb{R}$ e $h : [c, d] \rightarrow \mathbb{R}$ sono continue, allora:

$$\int \int_{[a,b] \times [c,d]} f(x, y) dx dy = \int \int_{[a,b] \times [c,d]} g(x)h(y) dx dy = \left(\int_{[a,b]} g(x) dx \right) \left(\int_{[c,d]} h(y) dy \right)$$

Estendiamo ora la definizione di integrale doppio al caso in cui il dominio di integrazione è un generico sottoinsieme limitato $\Omega \subset \mathbb{R}^2$.

Definizione 27 Sia $\Omega \subset \mathbb{R}^2$ un insieme limitato e $f : \Omega \rightarrow \mathbb{R}$ una funzione limitata. Considerato un rettangolo $R = [a, b] \times [c, d]$ contenente Ω , si definisca la funzione $\tilde{f} : R \rightarrow \mathbb{R}$ come:

$$\tilde{f}(x, y) := \begin{cases} f(x, y) & (x, y) \in \Omega \\ 0 & (x, y) \in R \setminus \Omega \end{cases}$$

Si dice che f è integrabile su Ω se la funzione \tilde{f} è integrabile su R . In tal caso l'integrale $\int \int_R \tilde{f}(x, y) dx dy$ si definisce integrale (doppio) di f su Ω e si indica con $\int \int_\Omega f(x, y) dx dy$:

$$\int \int_\Omega f(x, y) dx dy := \int \int_R \tilde{f}(x, y) dx dy$$

Di fatto si dimostra che la definizione appena data non è ambigua in quanto l'integrabilità di una funzione f su un insieme limitato Ω non dipende dalla scelta del rettangolo R che lo contiene.

Notiamo che, nella definizione appena data, l'esistenza dell'integrale $\int \int_\Omega f(x, y) dx dy$ dipende dalle proprietà sia della funzione f che dell'insieme Ω . In particolare se Ω non gode di buone proprietà (che tra breve andremo a specificare), è possibile che anche se $f : \Omega \rightarrow \mathbb{R}$ è continua, l'integrale $\int \int_\Omega f(x, y) dx dy$ non è ben definito. Ad esempio si consideri come Ω il sottoinsieme di \mathbb{R}^2 definito da:

$$\Omega = \{(x, y) \in [0, 1] \times [0, 1] : x \in \mathbb{Q}\},$$

e la funzione $f : \Omega \rightarrow \mathbb{R}$ definita da

$$f(x, y) = 1, \quad (x, y) \in \Omega.$$

Allora scelto il rettangolo $R = [0, 1] \times [0, 1]$ contenente Ω e definita $\tilde{f} : R \rightarrow \mathbb{R}$ tramite

$$\tilde{f} = \begin{cases} 1 & (x, y) \in \Omega \\ 0 & (x, y) \in R \setminus \Omega \end{cases}$$

abbiamo che $\int \int_\Omega f(x, y) dx dy$ esiste se \tilde{f} è integrabile su R . D'altra parte, come visto nell'esempio 27, l'integrale di \tilde{f} su R non è ben definito e, di conseguenza, neppure quello di f su Ω .

Definizione 28 Un insieme $\Omega \subset \mathbb{R}^2$ limitato è detto misurabile (secondo Peano-Jordan) se la funzione $f : \Omega \rightarrow \mathbb{R}$ definita da $f(x, y) = 1$ per ogni $(x, y) \in \Omega$ è integrabile su Ω . In tal caso l'integrale $\int \int_\Omega 1 dx dy$ viene chiamato area (o misura) di Ω .

La definizione precedente generalizza a particolari sottoinsiemi del piano il concetto di area. Introduciamo anche la definizione di insieme di misura o area nulla.

Definizione 29 Un insieme $\Omega \subset \mathbb{R}^2$ limitato è detto di misura nulla se è misurabile e la sua area è pari a 0.

Analizzando la definizione di integrale possiamo notare che un insieme limitato Ω ha area nulla se e solo se considerato un rettangolo R che lo contiene e diviso R in n^2 rettangoli uguali, la somma delle aree dei rettangoli che hanno intersezione non vuota con Ω tende a 0 quando $n \rightarrow \infty$. Analizzando infatti la definizione 28 abbiamo che Ω è misurabile e ha area nulla se e solo se $\lim_{n \rightarrow \infty} \sum_{i,k=1}^n Area(R_{ik})\tilde{f}(\xi_{ik}) = 0$, dove $\xi_{ik} \in R_{ik}$ e la funzione $\tilde{f} : R \rightarrow \mathbb{R}$ è definita come $\tilde{f}(x, y) = 1$ se $(x, y) \in \Omega$ e $\tilde{f}(x, y) = 0$ se $(x, y) \notin \Omega$ (di fatto \tilde{f} è la funzione caratteristica dell'insieme Ω). In particolare quindi, se per determinati valori degli indici i, k il rettangolo R_{ik} non contiene punti di Ω , allora, comunque viene scelto $\xi_{ik} \in R_{ik}$ avremo $\tilde{f}(\xi_{ik}) = 0$. Possiamo dunque, nella somma $\sum_{i,k} Area(R_{ik})\tilde{f}(\xi_{ik})$ su tutti gli n^2 sottorettangoli in cui viene suddiviso il rettangolo R , contare solo i sottorettangoli che contengono dei punti di Ω , perché gli altri danno contributo nullo, scrivendo:

$$\sum_{i,k} Area(R_{ik})\tilde{f}(\xi_{ik}) = \sum_{i,k \text{ tali che } R_{ik} \cap \Omega \neq \emptyset} Area(R_{ik})\tilde{f}(\xi_{ik}) \quad (27)$$

Quindi ogni somma di Cauchy-Riemann per la funzione \tilde{f} assume la forma al secondo membro di (27).

Se Ω ha area nulla, allora il limite per $n \rightarrow \infty$ delle somme di Cauchy-Riemann (27) vale 0 e tale valore non dipende dalla scelta dei punti $\xi_{ik} \in R_{ik}$. In particolare quindi scegliendo, negli addendi della somma al secondo membro di (27), $\xi_{ik} \in \Omega$ abbiamo $\tilde{f}(\xi_{ik}) = 1$ e quindi

$$0 = \lim_{n \rightarrow \infty} \sum_{i,k \text{ tali che } R_{ik} \cap \Omega \neq \emptyset} Area(R_{ik})\tilde{f}(\xi_{ik}) = \lim_{n \rightarrow \infty} \sum_{i,k \text{ tali che } R_{ik} \cap \Omega \neq \emptyset} Area(R_{ik}).$$

Il secondo membro mostra proprio che la somma delle aree dei rettangoli che hanno intersezione non vuota con Ω tende a 0 quando $n \rightarrow \infty$.

Viceversa, supponendo che la somma delle aree dei rettangoli che hanno intersezione non vuota con Ω tende a 0 quando $n \rightarrow \infty$, è possibile mostrare che Ω è misurabile e ha area nulla in quanto, ogni somma di Cauchy Riemann è limitata da:

$$0 \leq \sum_{i,k \text{ tali che } R_{ik} \cap \Omega \neq \emptyset} Area(R_{ik})\tilde{f}(\xi_{ik}) \leq \lim_{n \rightarrow \infty} \sum_{i,k \text{ tali che } R_{ik} \cap \Omega \neq \emptyset} Area(R_{ik})$$

Dato che per ipotesi il secondo membro tende a 0, abbiamo che, indipendentemente dalla scelta di $\xi_{ik} \in R_{ik}$

$$\lim_{n \rightarrow \infty} \sum_{i,k} \text{Area}(R_{ik}) \tilde{f}(\xi_{ik}) = 0$$

e quindi Ω è misurabile e ha misura nulla.

In particolare, utilizzando la proprietà di continuità uniforme delle funzioni continue sugli insiemi chiusi e limitati, si può dimostrare il seguente risultato.

Teorema 22 *Sia $\Omega \subset \mathbb{R}^2$ il grafico di una funzione continua $g : [a, b] \rightarrow \mathbb{R}$, ovvero un insieme della forma:*

$$\Omega = \{(x, y) \in \mathbb{R}^2 : x \in [a, b], y = g(x)\}.$$

Allora Ω ha misura nulla.

Enunciamo ora alcune proprietà fondamentali dell'integrale. Se $\Omega \subset \mathbb{R}^2$ è un insieme limitato e misurabile e $f_1, f_2 : \Omega \rightarrow \mathbb{R}$ due funzioni integrabili su Ω , $C_1, C_2 \in \mathbb{R}$ due costanti reali, allora vale la seguente proprietà di linearità :

$$\int \int_{\Omega} (C_1 f_1(x, y) + C_2 f_2(x, y)) dx dy = C_1 \int \int_{\Omega} f_1(x, y) dx dy + C_2 \int \int_{\Omega} f_2(x, y) dx dy \quad (28)$$

Inoltre se $\Omega_1 \subset \mathbb{R}^2$ e $\Omega_2 \subset \mathbb{R}^2$ sono insiemi limitati e misurabili tali che $\Omega_1 \cap \Omega_2 = \emptyset$ e f è integrabile su $\Omega_1 \cup \Omega_2$ allora vale la proprietà di *additività rispetto al dominio di integrazione*

$$\int \int_{\Omega_1 \cup \Omega_2} f(x, y) dx dy = \int \int_{\Omega_1} f(x, y) dx dy + \int \int_{\Omega_2} f(x, y) dx dy \quad (29)$$

Andiamo ora a descrivere un particolare tipo di insiemi misurabili sui quali può ancora essere applicato il metodo di riduzione per il calcolo dell'integrale di una funzione continua.

Definizione 30 *Un insieme $\Omega \subset \mathbb{R}^2$ è detto *y*-semplice se è della forma:*

$$\Omega = \{(x, y) \in \mathbb{R}^2 : x \in [a, b], g_1(x) \leq y \leq g_2(x)\},$$

dove $g_1 : [a, b] \rightarrow \mathbb{R}$ e $g_2 : [a, b] \rightarrow \mathbb{R}$ sono due funzioni continue.

In particolare un insieme Ω y -semplice gode della seguente proprietà : scelto un valore $\bar{x} \in [a, b]$ l'intersezione tra Ω e la retta (parallela all'asse delle y) di equazione $x = \bar{x}$ è un segmento (di estremi $[g_1(\bar{x}), g_2(\bar{x})]$) e tale segmento varia con continuità al variare di $\bar{x} \in [a, b]$.

Definizione 31 Un insieme $\Omega \subset \mathbb{R}^2$ è detto x -semplice se è della forma:

$$\Omega = \{(x, y) \in \mathbb{R}^2 : y \in [c, d], h_1(y) \leq x \leq h_2(y), \}$$

dove $h_1 : [c, d] \rightarrow \mathbb{R}$ e $h_2 : [c, d] \rightarrow \mathbb{R}$ sono due funzioni continue.

In particolare un insieme Ω x -semplice gode della seguente proprietà : scelto un valore $\bar{y} \in [c, d]$ l'intersezione tra Ω e la retta (parallela all'asse delle x) di equazione $y = \bar{y}$ è un segmento (di estremi $[h_1(\bar{y}), h_2(\bar{y})]$) e tale segmento varia con continuità al variare di $\bar{y} \in [c, d]$.

Definizione 32 Un insieme limitato $\Omega \subset \mathbb{R}^2$ è detto regolare se può essere scomposto come unione di un numero finito di insiemi semplici con intersezione di misura nulla.

Vale la seguente proprietà

Teorema 23 Sia $\Omega \subset \mathbb{R}^2$ un insieme regolare e $f : \Omega \rightarrow \mathbb{R}$ una funzione continua. Allora f è integrabile su Ω .

In particolare, da un punto di vista operativo, per la proprietà di additività rispetto al dominio di integrazione 29, l'integrale di f su un insieme regolare può essere scomposto nella somma di integrali di f su domini semplici. Questi ultimi possono venire calcolati tramite il metodo di riduzione:

Teorema 24 (metodo di riduzione, per domini semplici) Sia $f : \Omega \rightarrow \mathbb{R}$ una funzione continua e $\Omega \subset \mathbb{R}^2$ un insieme y -semplice della forma:

$$\Omega = \{(x, y) \in \mathbb{R}^2 : x \in [a, b], g_1(x) \leq y \leq g_2(x), \}$$

Allora

$$\int \int_{\Omega} f(x, y) dx dy = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) dy \right) dx$$

Se invece $\Omega \subset \mathbb{R}^2$ è un insieme x -semplice della forma:

$$\Omega = \{(x, y) \in \mathbb{R}^2 : y \in [c, d], h_1(y) \leq x \leq h_2(y), \}$$

allora

$$\int \int_{\Omega} f(x, y) dx dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x, y) dx \right) dy$$

Trasformazioni di coordinate e formula di cambiamento di variabili negli integrali doppi

Ricordiamo una formula appresa nel corso di analisi 1, associata al metodo di *integrazione per sostituzione*.

$$\int_a^b f(x)dx = \int_{a'}^{b'} f(T(u))T'(u)du, \quad (30)$$

dove $T : T^{-1}([a, b]) \rightarrow [a, b]$ è una funzione derivabile e invertibile dall'intervallo $T^{-1}([a, b])$ all'intervallo $[a, b]$. In particolare $a' = T^{-1}(a)$ e $b' = T^{-1}(b)$. Di solito la formula (30) risulta utile se la ricerca della primitiva della funzione $u \mapsto f(T(u))T'(u)$ risulta più semplice della ricerca della primitiva della funzione $x \mapsto f(x)$. La mappa $T : T^{-1}([a, b]) \rightarrow [a, b]$ che alla variabile reale $u \in T^{-1}([a, b])$ associa la variabile reale $x = T(u) \in [a, b]$ viene riguardata come un *cambiamento di variabili*.

Lo scopo di questa sezione è la generalizzazione della formula (30) al calcolo di integrali doppi, ovvero lo sviluppo di una procedura che permetta, una volta introdotto un opportuno cambiamento di variabili in \mathbb{R}^2 , ovvero una mappa $T : \Omega' \subset \mathbb{R}^2 \rightarrow \Omega \subset \mathbb{R}^2$ che rimpiazzì le "vecchie" variabili $(x, y) \in \mathbb{R}^2$ con delle "nuove" variabili $(u, v) \in \Omega' \subset \mathbb{R}^2$, di calcolare un integrale di una funzione $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ su un insieme misurabile $\Omega \subset \mathbb{R}^2$, cioè

$$\int \int_{\Omega} f(x, y)dx dy,$$

in un integrale sull'insieme $\Omega' := T^{-1}(\Omega)$ rispetto alle variabili (u, v) .

Indicheremo con $T_1 : \Omega' \rightarrow \mathbb{R}$ e con $T_2 : \Omega' \rightarrow \mathbb{R}$ le due funzioni componenti di T :

$$(x, y) = (T_1(u, v), T_2(u, v)).$$

Affinché la mappa $T : \Omega' \subset \mathbb{R}^2 \rightarrow \Omega \subset \mathbb{R}^2$, con Ω, Ω' aperti di \mathbb{R}^2 , possa svolgere il ruolo di un "buon" cambiamento di coordinate, è necessario richiedere che possenga particolari proprietà.

1. Innanzitutto T dovrà essere *invertibile*, cioè iniettiva e suriettiva in modo tale che ogni punto $(x, y) \in \Omega$ sia l'immagine di un unico punto $(u, v) \in \Omega'$. In particolare quindi sarà possibile definire l'inversa $T^{-1} : \Omega \rightarrow \Omega'$ di T , cioè la mappa che ad ogni $(x, y) \in \Omega$ fa corrispondere l'unico punto (u, v) tale che $(x, y) = T(u, v)$
2. Vogliamo inoltre che una funzione *continua* $f : \Omega \rightarrow \mathbb{R}$ delle variabili (x, y) continui a possedere questa proprietà se "espressa nelle nuove variabili" (u, v) .

In altre parole chiediamo che $f \circ T : \Omega' \rightarrow \mathbb{R}$ sia continua per ogni $f : \Omega \rightarrow \mathbb{R}$ continua. Questa condizione è assicurata se $T : \Omega' \rightarrow \Omega$ è continua (cioè se le sue funzioni componenti $T_1 : \Omega' \rightarrow \mathbb{R}$ e con $T_2 : \Omega' \rightarrow \mathbb{R}$ sono continue).

3. Inoltre chiediamo che per ogni $f : \Omega \rightarrow \mathbb{R}$ di classe C^1 , la funzione $f \circ T : \Omega' \rightarrow \mathbb{R}$ sia anch'essa di classe C^1 e, viceversa, per ogni $g : \Omega' \rightarrow \mathbb{R}$ di classe C^1 , la funzione $g \circ T^{-1} : \Omega \rightarrow \mathbb{R}$ sia anch'essa di classe C^1 . Queste condizioni sono assicurate se $T : \Omega' \rightarrow \Omega$ e $T^{-1} : \Omega \rightarrow \Omega'$ sono di classe C^1 (cioè se le loro componenti sono di classe C^1).

Una mappa $T : \Omega' \rightarrow \Omega$ con le proprietà 1., 2. e 3. è detta *diffeomorfismo globale*.

Esempio 28 (Traslazioni) Consideriamo la mappa $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita da

$$(x, y) = T(u, v) = (u + a, v + b),$$

con $a, b \in \mathbb{R}$. T è invertibile con inversa

$$(u, v) = T^{-1}(x, y) = (x - a, y - b).$$

È immediato verificare che T e la sua inversa T^{-1} sono di classe C^1 .

Esempio 29 (Trasformazioni lineari) Consideriamo la mappa $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita da

$$\begin{pmatrix} x \\ y \end{pmatrix} = M \begin{pmatrix} u \\ v \end{pmatrix},$$

con M matrice 2×2 a coefficienti reali. T è invertibile se e solo se $\det(M) \neq 0$ con inversa

$$\begin{pmatrix} u \\ v \end{pmatrix} = M^{-1} \begin{pmatrix} x \\ y \end{pmatrix}.$$

È immediato verificare che T e la sua inversa T^{-1} sono di classe C^1 .

Notiamo che T trasforma rette in rette, infatti la retta $r \in \mathbb{R}^2$ descritta da

$$r := \{(x, y) \in \mathbb{R}^2 : y = \alpha x + \beta\}$$

viene mandata attraverso la trasformazione T^{-1} nell'insieme

$$\begin{aligned} T^{-1}(r) &= \{(u, v) \in \mathbb{R}^2 : (u, v) = T^{-1}(x, y), (x, y) \in r\} \\ &= \{(u, v) \in \mathbb{R}^2 : m_{21}u + m_{22}v = \alpha(m_{11}u + m_{12}v) + \beta\} \end{aligned}$$

dove

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}.$$

Il quadrato di area unitaria $Q = \{(u, v) \in \mathbb{R}^2 : u \in [0, 1], v \in [0, 1]\}$ viene trasformato nel parallelogramma di vertici $(0, 0), (m_{11}, m_{21}), (m_{12}, m_{22}), (m_{11} + m_{12}, m_{21} + m_{22})$, di area pari a $|(m_{11}, m_{21}) \wedge (m_{12}, m_{22})| = |\det M|$. In particolare, se M è una matrice unitaria, allora la trasformazione T preserva le distanze, cioè:

$$\|T(u_1, v_1) - T(u_2, v_2)\| = \|(u_1, v_1) - (u_2, v_2)\|$$

e le aree in quanto, per le matrici unitarie, $\det M = 1$.

Esempio 30 (coordinate polari nel piano) Consideriamo la mappa $T : \mathbb{R}^+ \times [0, 2\pi] \rightarrow \mathbb{R}^2$ definita da

$$(x, y) = T(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta),$$

Si verifica facilmente che T è di classe C^1 . A rigore T non è un diffeomorfismo globale da $\mathbb{R}^+ \times [0, 2\pi]$ a \mathbb{R}^2 in quanto non è iniettiva (tutti i punti con $\rho = 0$ e $\theta \in [0, 2\pi]$ vengono mandati nell'origine $(0, 0)$). T diventa un diffeomorfismo globale se restringiamo il suo dominio e definiamo T da $(0, +\infty) \times (0, 2\pi)$ in \mathbb{R}^2 privato della semiretta $y = 0, x \geq 0$.

Notiamo che la mappa T manda il segmento $\rho = R, \theta \in [\alpha, \beta]$ in un arco di circonferenza di raggio R . Inoltre, ad esempio, T manda il rettangolo $\{(\rho, \theta) \in \mathbb{R}^+ \times [0, 2\pi] : \rho \in [R_1, R_2], \theta \in [0, 2\pi]\}$ nella corona circolare $\{(x, y) \in \mathbb{R}^2, R_1^2 \leq x^2 + y^2 \leq R_2^2\}$.

In generale, data una funzione $T : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, definita su un aperto $A \subset \mathbb{R}^n$, indicheremo con T_1, \dots, T_m , con $T_i : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ per $i = 1, \dots, m$, le sue componenti. Diremo che T è continua in un punto $x \in A$ se tutte le funzioni componenti lo sono. Analogamente diremo che T è differenziabile in $x \in A$ se lo sono tutte le sue funzioni componenti. In tal caso è utile introdurre la *matrice Jacobiana* di T in $x \in A$, indicata con $JT(x)$, cioè la matrice $m \times n$, i cui m vettori riga sono i gradienti $\nabla T_1, \dots, \nabla T_m$ delle funzioni componenti calcolate in x :

$$JT(x) = \begin{pmatrix} \frac{\partial T_1}{\partial x_1}(x) & \frac{\partial T_1}{\partial x_2}(x) & \dots & \frac{\partial T_1}{\partial x_n}(x) \\ \frac{\partial T_2}{\partial x_1}(x) & \frac{\partial T_2}{\partial x_2}(x) & \dots & \frac{\partial T_2}{\partial x_n}(x) \\ \dots & \dots & \dots & \dots \\ \frac{\partial T_m}{\partial x_1}(x) & \frac{\partial T_m}{\partial x_2}(x) & \dots & \frac{\partial T_m}{\partial x_n}(x) \end{pmatrix}$$

Notiamo che se la funzione T è differenziabile in un punto x_0 interno al suo dominio, allora la matrice Jacobiana $JT(x_0)$ fornisce la matrice associata alla migliore funzione lineare affine che approssima T in un intorno di x_0 . In altre parole vale la formula $T(x) = T(x_0) + JT(x_0)(x - x_0) + o(\|x - x_0\|)$.

Introducendo le variabili $(h_1, \dots, h_n) \in \mathbb{R}^n$ come $h_i := x_i - x_{0,i}$, e il vettore $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ possiamo anche scrivere:

$$T(x_0 + h) - T(x_0) = JT(x_0)h + o(\|h\|).$$

Consideriamo ora una generica mappa $T : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ differenziabile in ogni punto del suo dominio A (che supporremo essere un insieme aperto). e cerchiamo delle condizioni che assicurino l'invertibilità di A . Nell'ipotesi che $T \in C^1(A)$ (cioè che tutte le sue componenti siano funzioni di classe $C^1(A)$), il seguente teorema mostra che dato un punto $x_0 \in A$, il non annullarsi del determinante di $JT(x_0)$ (condizione che assicura l'invertibilità della migliore mappa lineare affine che approssima T in un intorno di x_0), assicura l'invertibilità *locale* di T . Più precisamente abbiamo:

Teorema 25 (della funzione inversa) *Sia $T : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, con A aperto, funzione di classe $C^1(A)$. Sia $x_0 \in A$ tale che $\det JT(x_0) \neq 0$. Allora esiste un intorno U di x_0 ed un intorno V di $T(x_0)$ tali che la mappa $T : U \rightarrow V$ è biettiva. Inoltre la sua inversa $T^{-1} : V \rightarrow U$ è ancora di classe $C^1(V)$ e $JT^{-1}(T(x)) = (JT(x))^{-1}$ per ogni $x \in U$.*

Ad esempio, nel caso particolare in cui $n = 2$, consideriamo la funzione $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita da $T(u, v) = (e^u \cos v, e^u \sin v)$. La matrice Jacobiana è data da

$$JT(u, v) = \begin{pmatrix} e^u \cos v & -e^u \sin v \\ e^u \sin v & e^u \cos v \end{pmatrix}$$

il cui determinante è uguale a $\det(JT(u, v)) = e^{2u}$ che è sempre strettamente positivo. Ogni punto $(u_0, v_0) \in \mathbb{R}^2$ ammette un intorno in cui la funzione T è invertibile localmente. ma se consideriamo tutto il dominio di definizione di T , abbiamo che T non è invertibile *globalmente* in quanto non è iniettiva. Ad esempio infatti i punti (u_0, v_0) e $(u_0, v_0 + 2\pi)$ hanno la stessa immagine.

Oltre alla locale invertibilità della mappa T , il determinante della matrice Jacobiana fornisce anche un interessante relazione che connette nel caso in cui $n = 2$ la misura (area) di un rettangolo di lati infinitesimi nel dominio di T con la misura della sua immagine.

Consideriamo infatti un rettangolo nel piano uv di vertici (u, v) , $(u + du, v)$, $(u, v + dv)$ e $(u + du, v + dv)$ viene trasformato attraverso la mappa T in un parallelogramma infinitesimo di vertici $(T_1(u, v), T_2(u, v))$, $(T_1(u + du, v), T_2(u + du, v))$, $(T_1(u, v + dv), T_2(u, v + dv))$ e $(T_1(u + du, v + dv), T_2(u + du, v + dv))$, i cui lati adiacenti sono dati dai vettori

$$(T_1(u + du, v), T_2(u + du, v)) - (T_1(u, v), T_2(u, v)) \sim \left(\frac{\partial}{\partial u} T_1(u, v) du, \frac{\partial}{\partial u} T_2(u, v) du \right)$$

$$(T_1(u, v + dv), T_2(u, v + dv)) - (T_1(u, v), T_2(u, v)) \sim \left(\frac{\partial}{\partial v} T_1(u, v) dv, \frac{\partial}{\partial v} T_2(u, v) dv \right)$$

L'area di tale parallelogramma è data da

$$\left| \det \begin{pmatrix} \frac{\partial}{\partial u} T_1(u, v) & \frac{\partial}{\partial v} T_1(u, v) \\ \frac{\partial}{\partial u} T_2(u, v) & \frac{\partial}{\partial v} T_2(u, v) \end{pmatrix} \right| dudv = |\det(JT(u, v))| dudv$$

Diamo qui di seguito la formula di cambiamento di variabili negli integrali doppi.

Teorema 26 Sia $\Omega \subset \mathbb{R}^2$ un insieme regolare, $f : \Omega \rightarrow \mathbb{R}$ una funzione continua e $T : \Omega' \rightarrow \Omega$ un diffeomorfismo globale che trasforma Ω' in Ω , con:

$$(x, y) = T(u, v) = (T_1(u, v), T_2(u, v)).$$

Allora

$$\int \int_{\Omega} f(x, y) dx dy = \int \int_{\Omega'} f((T_1(u, v), T_2(u, v))) |\det JT(u, v)| dudv, \quad (31)$$

dove

$$JT(u, v) = \begin{pmatrix} \frac{\partial}{\partial u} T_1(u, v) & \frac{\partial}{\partial v} T_1(u, v) \\ \frac{\partial}{\partial u} T_2(u, v) & \frac{\partial}{\partial v} T_2(u, v) \end{pmatrix}$$

è la matrice Jacobiana della trasformazione di coordinate

Notiamo che il termine $|\det JT(u, v)|$ gioca il ruolo di fattore di cambiamento di scala per le aree nel passaggio dalle coordinate (u, v) alle coordinate (x, y) .

Integrali doppi generalizzati

Finora abbiamo trattato solo l'integrazione di funzioni limitate su insiemi limitati. Vediamo ora un esempio notevole di integrale doppio generalizzato.

Sappiamo dal corso di analisi 1 che la funzione $x \mapsto e^{-x^2}$ non ha una primitiva esprimibile in termini di funzioni elementari. Vediamo ora come calcolare in maniera esatta l'integrale generalizzato $\int_{-\infty}^{+\infty} e^{-x^2} dx$, cioè il limite

$$\lim_{R \rightarrow +\infty} \int_{-R}^{+R} e^{-x^2} dx.$$

Consideriamo in \mathbb{R}^2 il quadrato $Q_R = [-R, R] \times [-R, R]$ e l'integrale doppio

$$\int \int_{Q_R} e^{-x^2 - y^2} dx dy$$

Notiamo che

$$\lim_{R \rightarrow +\infty} \int \int_{Q_R} e^{-x^2-y^2} dx dy = \lim_{R \rightarrow +\infty} \int_{-R}^{+R} e^{-x^2} dx \int_{-R}^{+R} e^{-y^2} dy = \left(\lim_{R \rightarrow +\infty} \int_{-R}^{+R} e^{-x^2} dx \right)^2$$

Considerati i cerchi $B_R = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}$ e $B_{\sqrt{2}R} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 2R^2\}$, per ogni valore di R valgono le inclusioni $B_R \subset Q_R \subset B_{\sqrt{2}R}$. Dato che la funzione $f(x, y) = e^{-x^2-y^2}$ è positiva, allora:

$$\int \int_{B_R} f(x, y) dx dy \leq \int \int_{Q_R} f(x, y) dx dy \leq \int \int_{B_{\sqrt{2}R}} f(x, y) dx dy \quad (32)$$

Inoltre, utilizzando le coordinate polari nel piano

$$\int \int_{B_R} f(x, y) dx dy = \int_0^R \int_0^{2\pi} e^{-\rho^2} \rho d\rho d\theta = \pi(1 - e^{-R^2})$$

e analogamente $\int \int_{B_{\sqrt{2}R}} f(x, y) dx dy = \pi(1 - e^{-2R^2})$. Dato che

$$\lim_{R \rightarrow \infty} \int \int_{B_R} f(x, y) dx dy = \lim_{R \rightarrow \infty} \int \int_{B_{\sqrt{2}R}} f(x, y) dx dy = \pi,$$

per la disuguaglianza (32) ed il teorema dei carabinieri, abbiamo

$$\lim_{R \rightarrow \infty} \int \int_{Q_R} e^{-x^2-y^2} dx dy = \pi$$

e quindi:

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

Integrali tripli (o di volume)

Una buona parte delle definizioni ed dei risultati dei capitoli precedenti, relativi al calcolo di integrali di funzioni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, si generalizzano al caso di funzioni $f : \mathbb{R}^3 \rightarrow \mathbb{R}$.

Consideriamo un parallelepipedo $P \subset \mathbb{R}^3$, con $P = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$, e una funzione $f : P \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ limitata. Definiamo l'integrale triplo (o di volume) di f sull'insieme P , indicato col simbolo

$$\int \int \int_P f(x, y, z) dx dy dz.$$

Suddividiamo il parallelepipedo P in n^3 parallelepipedi di volume $(b_1 - a_1)(b_2 - a_2)(b_3 - a_3)/n^3$, ottenuti dividendo ognuno degli intervalli $[a_1, b_1]$, $[a_2, b_2]$ e $[a_3, b_3]$ in n sottointervalli di estremi rispettivamente $[x_i, x_{i+1}]$, con $x_i = a_1 + i(b_1 - a_1)/n$, e $i = 0, \dots, n$, $[y_j, y_{j+1}]$, con $y_j = a_2 + j(b_2 - a_2)/n$ e $j = 0, \dots, n$, e $[z_k, z_{k+1}]$, con $z_k = a_3 + k(b_3 - a_3)/n$ e $k = 0, \dots, n$. Indichiamo per ogni terna (i, j, k) , dove $i = 1, \dots, n$, $j = 1, \dots, n$ e $k = 1, \dots, n$, con P_{ijk} il parallelepipedo $P_{ijk} := [x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k]$ e con $\xi_{ijk} \in P_{ijk}$ un generico punto scelto all'interno di P_{ijk} . Costruiamo le somme di Cauchy-Riemann come:

$$s_n := \sum_{i,j,k=1}^n f(\xi_{ijk}) \frac{(b_1 - a_1)(b_2 - a_2)(b_3 - a_3)}{n^3}$$

Se esiste finito il limite per $n \rightarrow \infty$ di tali somme e tale valore non dipende dalla scelta dei punti $\xi_{ijk} \in P_{ijk}$, allora tale valore definisce l'integrale di f sul parallelepipedo P .

Definizione 33 Una funzione $f : P \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ limitata è detta integrabile su $P = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ se esiste finito

$$\lim_{n \rightarrow \infty} \sum_{i,j,k=1}^n f(\xi_{ijk}) \frac{(b_1 - a_1)(b_2 - a_2)(b_3 - a_3)}{n^3} \quad (33)$$

e tale valore non dipende dalla scelta di $\xi_{ijk} \in P_{ijk}$. In tal caso viene indicato con il simbolo $\int \int \int_P f(x, y, z) dx dy dz$

Notiamo che, nel caso in cui $f \geq 0$, interpretando la funzione f come la densità di massa del solido P , il valore $\int \int \int_P f(x, y, z) dx dy dz$ può essere interpretato come la massa totale del corpo.

Il teorema 20 si generalizza al caso di integrali tripli:

Teorema 27 Sia $f : [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] \rightarrow \mathbb{R}$ una funzione continua. Allora f è integrabile su $P = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$.

Il metodo di riduzione vale anche per gli integrali tripli e permette di ricondurre il calcolo di un integrale di una funzione di tre variabili reali al calcolo di tre integrali semplici iterati. nel caso in cui il dominio di integrazione sia un parallelepipedo l'ordine di integrazione è irrilevante e di fatto ce ne sono 6 possibili:

$$\int \int \int_{[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]} f(x, y, z) dx dy dz = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\int_{a_3}^{b_3} f(x, y, z) dz \right) dy \right) dx$$

$$\begin{aligned}
&= \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} \left(\int_{a_3}^{b_3} f(x, y, z) dz \right) dx \right) dy \\
&= \int_{a_3}^{b_3} \left(\int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x, y, z) dy \right) dx \right) dz \\
&= \int_{a_1}^{b_1} \left(\int_{a_3}^{b_3} \left(\int_{a_2}^{b_2} f(x, y, z) dy \right) dz \right) dx \\
&= \int_{a_2}^{b_2} \left(\int_{a_3}^{b_3} \left(\int_{a_1}^{b_1} f(x, y, z) dx \right) dz \right) dy \\
&= \int_{a_3}^{b_3} \left(\int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x, y, z) dx \right) dy \right) dz
\end{aligned}$$

Come nel caso degli integrali doppi, possiamo estendere la la definizione di integrale triplo al caso in cui il dominio di integrazione è un generico sottoinsieme limitato $\Omega \subset \mathbb{R}^3$.

Definizione 34 Sia $\Omega \subset \mathbb{R}^3$ un insieme limitato e $f : \Omega \rightarrow \mathbb{R}$ una funzione limitata. Considerato un parallelepipedo $P = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ contenente Ω , si definisca la funzione $\tilde{f} : P \rightarrow \mathbb{R}$ come:

$$\tilde{f}(x, y, z) := \begin{cases} f(x, y, z) & (x, y, z) \in \Omega \\ 0 & (x, y, z) \in P \setminus \Omega \end{cases}$$

Si dice che f è integrabile su Ω se la funzione \tilde{f} è integrabile su P . In tal caso l'integrale $\int \int \int_P \tilde{f}(x, y, z) dx dy dz$ si definisce integrale (triplo) di f su Ω e si indica con $\int \int \int_{\Omega} f(x, y, z) dx dy dz$:

$$\int \int \int_{\Omega} f(x, y, z) dx dy dz := \int \int \int_P \tilde{f}(x, y, z) dx dy dz$$

Un insieme $\Omega \subset \mathbb{R}^3$ limitato è detto *misurabile* se la funzione $f : \Omega \rightarrow \mathbb{R}$ definita da $f(x, y, z) = 1$ per ogni $(x, y, z) \in \Omega$ è integrabile su Ω . In tal caso l'integrale $\int \int \int_{\Omega} 1 dx dy dz$ viene chiamato *volume* o *misura* di Ω .

Un insieme $\Omega \subset \mathbb{R}^3$ limitato è detto *di misura nulla* se è misurabile e il suo volume è pari a 0. Analogamente al caso bidimensionale, è possibile dimostrare che un insieme limitato $\Omega \subset \mathbb{R}^3$ ha volume nullo se considerato un parallelepipedo P che lo contiene e diviso P in n^3 parallelepipedi uguali, la somma dei volumi dei parallelepipedi che hanno intersezione non vuota con Ω tende a 0 quando $n \rightarrow \infty$.

Le proprietà fondamentali dell'integrale doppio, in particolare la linearità (Eq. (28)) e l'additività rispetto al dominio di integrazione (Eq. (29)) valgono anche per gli integrali tripli.

Non ci addenteremo nei problemi relativi alla misurabilità dei sottoinsiemi $\Omega \subset \mathbb{R}^3$. Ci limiteremo a considerare integrali di funzioni continue su particolari sottoinsiemi $\Omega \subset \mathbb{R}^3$ detti "semplici", dove è possibile ridurre il calcolo di un integrale triplo al calcolo iterato di un integrale semplice e di un integrale doppio.

Definizione 35 Un sottoinsieme $\Omega \subset \mathbb{R}^3$ è detto semplice per fili se è della forma:

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, g_1(x, y) \leq z \leq g_2(x, y)\},$$

dove $D \subset \mathbb{R}^2$ è un sottoinsieme regolare del piano xy , mentre $g_1 : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $g_2 : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ sono due funzioni continue.

In altre parole un insieme Ω è semplice per fili se può essere rappresentato come quella parte dello spazio \mathbb{R}^3 compresa fra i grafici di due funzioni continue $g_1 : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ e $g_2 : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. L'insieme $D \subset \mathbb{R}^2$ può essere riguardato come la proiezione di Ω sul piano xy

Su tali insiemi l'integrale $\int \int \int_{\Omega} f(x, y, z) dx dy dz$ di una funzione continua $f : \Omega \rightarrow \mathbb{R}$ può essere calcolato tramite la tecnica di *integrazione per fili*. Fissato un punto $(\bar{x}, \bar{y}) \in D$ consideriamo il segmento verticale di estremi $g_1(\bar{x}, \bar{y})$ e $g_2(\bar{x}, \bar{y})$:

$$\{(x, y, z) \in \mathbb{R}^3 : x = \bar{x}, y = \bar{y}, z \in [g_1(\bar{x}, \bar{y}), g_2(\bar{x}, \bar{y})]\}$$

ottenuto intersecando l'insieme Ω con la retta verticale passante per il punto $(\bar{x}, \bar{y}) \in D$. Consideriamo la funzione f ristretta a tale segmento e il suo integrale, rispetto alla variabile z , su tale segmento, ovvero $\int_{g_1(\bar{x}, \bar{y})}^{g_2(\bar{x}, \bar{y})} f(\bar{x}, \bar{y}, z) dz$. Consideriamo, al variare di $(\bar{x}, \bar{y}) \in D$, la funzione di due variabili definita sull'insieme D del piano xy :

$$(x, y) \mapsto \int_{g_1(x, y)}^{g_2(x, y)} f(x, y, z) dz$$

e integriamo tale funzione sull'insieme regolare D . Allora, utilizzando il metodo di integrazione per fili, si ha che:

$$\int \int \int_{\Omega} f(x, y, z) dx dy dz = \int \int_D \left(\int_{g_1(x, y)}^{g_2(x, y)} f(x, y, z) dz \right) dx dy \quad (34)$$

Esempio 31 Calcolare l'integrale della funzione $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definita da $f(x, y, z) = x^2 z$ sull'insieme $\Omega = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1, z \geq 0\}$.

Possiamo rappresentare l'insieme Ω , ovvero la semisfera di raggio 1 situata nel semispazio $z \geq 0$, come un insieme "semplice per fili" $\Omega = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, g_1(x, y) \leq z \leq g_2(x, y)\}$, dove $D \subset \mathbb{R}^2$, $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ è il

cerchio di centro $(0,0)$ e raggio 1, mentre $g_1(x,y) = 0$ e $g_2(x,y) = \sqrt{1-x^2-y^2}$. Integrando "per fili":

$$\begin{aligned} \int \int \int_{\Omega} f(x,y,z) dx dy dz &= \int \int_D \left(\int_0^{\sqrt{1-x^2-y^2}} x^2 z dz \right) dx dy \\ &= \int \int_D x^2 \frac{1-x^2-y^2}{2} dx dy \\ &= \int_0^1 \int_0^{2\pi} \rho^2 \cos^2 \theta \frac{1-\rho^2}{2} \rho d\rho d\theta = \frac{\pi}{24} \end{aligned}$$

Un altro metodo alternativo è quello di integrazione "per strati".

Definizione 36 Un sottoinsieme $\Omega \subset \mathbb{R}^3$ è detto semplice per strati se è della forma:

$$\Omega = \{(x,y,z) \in \mathbb{R}^3 : z \in [h_1, h_2], (x,y) \in \Omega(z)\},$$

dove per ogni $\bar{z} \in [h_1, h_2]$, l'insieme $\Omega(\bar{z})$ ottenuto intersecando Ω con il piano orizzontale di equazione $z = \bar{z}$ è un insieme regolare.

Se consideriamo per ogni $\bar{z} \in [h_1, h_2]$ la funzione f ristretta all'insieme $\Omega(\bar{z})$ e ne calcoliamo l'integrale su tale insieme, ovvero il valore $\int \int_{\Omega(\bar{z})} f(x,y,\bar{z}) dx dy$, costruiamo la funzione

$$z \mapsto \int \int_{\Omega(z)} f(x,y,z) dx dy, \quad z \in [h_1, h_2]$$

Secondo il metodo di integrazione "per strati", l'integrale di tale funzione sull'intervallo $z \in [h_1, h_2]$ fornisce il valore dell'integrale di f su Ω :

$$\int \int \int_{\Omega} f(x,y,z) dx dy dz = \int_{h_1}^{h_2} \left(\int \int_{\Omega(z)} f(x,y,z) dx dy \right) dz \quad (35)$$

Esempio 32 L'integrale dell'esempio 31 può essere calcolato anche per strati. Possiamo infatti rappresentare l'insieme Ω , ovvero la semisfera di raggio 1 situata nel semispazio $z \geq 0$, come un insieme "semplice per strati" $\Omega = \{(x,y,z) \in \mathbb{R}^3 : z \in [0,1], (x,y) \in \Omega(z)\}$, dove $\Omega(z)$ è il cerchio di centro $(0,0)$ e raggio $\sqrt{1-z^2}$. Integrando "per strati":

$$\begin{aligned} \int \int \int_{\Omega} f(x,y,z) dx dy dz &= \int_0^1 \left(\int \int_{\Omega(z)} z x^2 dx dy \right) dz \\ &= \int_0^1 z \left(\int_0^{\sqrt{1-z^2}} \int_0^{2\pi} \rho^2 \cos^2 \theta \rho d\rho d\theta \right) dz \\ &= \int_0^1 \pi z \frac{(1-z^2)^2}{4} dz = \frac{\pi}{24} \end{aligned}$$

Si definisce un sottoinsieme $\Omega \subset \mathbb{R}^3$ *regolare* se può essere ottenuto come unione di un numero finito di insiemi semplici (per fili o per strati) la cui intersezione ha misura nulla:

$$\Omega = \bigcup_i^N \Omega_i, \quad Vol(\Omega_i \cap \Omega_j) = 0 \quad \forall i \neq j.$$

Formula di cambiamento di variabili negli integrali di volume

La formula di cambiamento di variabili negli integrali doppi si generalizza al caso degli integrali di volume.

Teorema 28 *Sia $\Omega \subset \mathbb{R}^3$ un insieme regolare, $f : \Omega \rightarrow \mathbb{R}$ una funzione continua e $T : \Omega' \subset \mathbb{R}^3 \rightarrow \Omega$ un diffeomorfismo globale che trasforma $\Omega' \subset \mathbb{R}^3$ in $\Omega \subset \mathbb{R}^3$, con:*

$$(x, y, z) = T(u, v, w) = (T_1(u, v, w), T_2(u, v, w), T_3(u, v, w)).$$

Allora

$$\begin{aligned} \int \int \int_{\Omega} f(x, y, z) dx dy dz &= \\ &= \int \int \int_{\Omega'} f((T_1(u, v, w), T_2(u, v, w), T_3(u, v, w)) |\det JT(u, v, w)| du dv dw, \end{aligned} \quad (36)$$

dove

$$JT(u, v, w) = \begin{pmatrix} \frac{\partial}{\partial u} T_1(u, v, w) & \frac{\partial}{\partial v} T_1(u, v, w) & \frac{\partial}{\partial w} T_1(u, v, w) \\ \frac{\partial}{\partial u} T_2(u, v, w) & \frac{\partial}{\partial v} T_2(u, v, w) & \frac{\partial}{\partial w} T_2(u, v, w) \\ \frac{\partial}{\partial u} T_3(u, v, w) & \frac{\partial}{\partial v} T_3(u, v, w) & \frac{\partial}{\partial w} T_3(u, v, w) \end{pmatrix}$$

è detta matrice Jacobiana della funzione T

Analogamente al caso bidimensionale, il termine $|\det JT|$ svolge il ruolo di un fattore di cambiamento di scala dei volumi nel passaggio dalle coordinate (u, v, w) alle coordinate (x, y, z) attraverso la trasformazione T .

Descriviamo ora due tra i cambiamenti di variabili più utilizzati in \mathbb{R}^3 .

Coordinate sferiche in \mathbb{R}^3

Consideriamo la trasformazione $T : [0, +\infty) \times [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{R}^3$ descritta da

$$\begin{cases} x(\rho, \phi, \theta) = \rho \sin \phi \cos \theta \\ y(\rho, \phi, \theta) = \rho \sin \phi \sin \theta \\ z(\rho, \phi, \theta) = \rho \cos \phi \end{cases}$$

dove $\rho = \sqrt{x^2 + y^2 + z^2}$ è la distanza fra (x, y, z) e $(0, 0, 0)$, ϕ è l'angolo tra i vettori $(0, 0, 1)$ e (x, y, z) , mentre θ è l'angolo tra i vettori $(1, 0, 0)$ e $(x, y, 0)$.

A rigore T non è un diffeomorfismo globale, in quanto non è iniettiva: ad esempio l'immagine di tutti punti $(\rho, \phi, \theta) \in [0, +\infty) \times [0, \pi] \times [0, 2\pi]$ con $\rho = 0$ è l'origine $(0, 0, 0)$. T diviene un diffeomorfismo globale andando a restringere il suo dominio all'aperto $(0, +\infty) \times (0, \pi) \times (0, 2\pi)$ e l'immagine attraverso T di tale insieme è

$$T((0, +\infty) \times (0, \pi) \times (0, 2\pi)) = \mathbb{R}^3 \setminus \{(x, y, z) \in \mathbb{R}^3 : y = 0, x \geq 0\}$$

Di fatto, dato che l'insieme $\{(x, y, z) \in \mathbb{R}^3 : y = 0, x \geq 0\}$ ha volume nullo, questo problema non compromette la possibilità di applicare le coordinate sferiche nella formula (36).

La matrice Jacobiana della trasformazione è :

$$JT(\rho, \phi, \theta) = \begin{pmatrix} \sin \phi \cos \theta & \rho \cos \phi \cos \theta & -\rho \sin \phi \sin \theta \\ \sin \phi \sin \theta & \rho \cos \phi \sin \theta & \rho \sin \phi \cos \theta \\ \cos \phi & -\rho \sin \phi & 0 \end{pmatrix}$$

è il valore assoluto del suo determinante è :

$$|\det JT(\rho, \phi, \theta)| = \rho^2 \sin \phi$$

Esempio 33 *L'integrale dell'esempio 31 può essere calcolato anche utilizzando la formula di cambiamento di variabili. La semisfera $\Omega = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1, z \geq 0\}$ è l'immagine attraverso il cambiamento di variabili T dell'insieme $\Omega' = \{(\rho, \phi, \theta) \in [0, +\infty) \times [0, \pi] \times [0, 2\pi] : 0 \leq \rho \leq 1, 0 \leq \phi \leq \pi/2, 0 \leq \theta \leq 2\pi\}$. Abbiamo quindi*

$$\begin{aligned} \int \int \int_{\Omega} f(x, y, z) dx dy dz &= \int_0^1 \int_0^{\pi/2} \int_0^{2\pi} \rho^2 \sin^2 \phi \cos^2 \theta \rho \cos \phi \rho^2 \sin \phi d\rho d\phi d\theta \\ &= \int_0^1 \rho^5 d\rho \int_0^{\pi/2} \sin^3 \phi \cos \phi d\phi \int_0^{2\pi} \cos^2 \theta d\theta = \frac{\pi}{24} \end{aligned}$$

Coordinate cilindriche in \mathbb{R}^3

Consideriamo la trasformazione $T : [0, +\infty) \times [0, 2\pi] \times \mathbb{R} \rightarrow \mathbb{R}^3$ descritta da

$$\begin{cases} x(\rho, \theta, t) = \rho \cos \theta \\ y(\rho, \theta, t) = \rho \sin \theta \\ z(\rho, \theta, t) = t \end{cases}$$

dove $\rho = \sqrt{x^2 + y^2}$ è la distanza fra $(x, y, 0)$ e $(0, 0, 0)$, θ è l'angolo tra i vettori $(1, 0, 0)$ e $(x, y, 0)$, mentre t coincide con la variabile z .

A rigore T non è un diffeomorfismo globale, in quanto non è iniettiva: ad esempio l'immagine di tutti i punti $(\rho, \theta, t) \in [0, +\infty) \times [0, 2\pi] \times \mathbb{R}$ con $\rho = 0$ e $t = 0$ è l'origine $(0, 0, 0)$. T diviene un diffeomorfismo globale andando a restringere il suo dominio all'aperto $(0, +\infty) \times (0, 2\pi) \times \mathbb{R}$ e l'immagine attraverso T di tale insieme è \mathbb{R}^3 privato del semipiano $\{(x, y, z) \in \mathbb{R}^3 : y = 0, x \geq 0\}$. Dato che l'insieme $\{(x, y, z) \in \mathbb{R}^3 : y = 0, x \geq 0\}$ ha volume nullo, è comunque possibile utilizzare coordinate cilindriche nella formula (36).

La matrice Jacobiana della trasformazione è :

$$JT(\rho, \theta, t) = \begin{pmatrix} \cos \theta & -\rho \sin \theta & 0 \\ \sin \theta & \rho \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

è il valore assoluto del suo determinante è :

$$|\det JT(\rho, \theta, t)| = \rho$$

Esempio 34 Calcolare il volume del cono $\Omega = \{(x, y, z) \in \mathbb{R}^3 : \sqrt{x^2 + y^2} \leq z \leq h\}$. Utilizzando le coordinate cilindriche, l'insieme Ω è l'immagine attraverso il cambiamento di variabili T dell'insieme $\Omega' = \{(\rho, \theta, t) \in [0, +\infty) \times [0, 2\pi] \times \mathbb{R} : 0 \leq \rho \leq h, \rho \leq t \leq h\}$. Abbiamo quindi

$$\begin{aligned} \int \int \int_{\Omega} 1 dx dy dz &= \int_0^h \int_0^{2\pi} \int_{\rho}^h \rho dt d\rho d\theta \\ &= \int_0^h \int_0^{2\pi} \rho(h - \rho) d\rho d\theta = \frac{\pi h^3}{3} \end{aligned}$$

Campi vettoriali

In questa sezione studieremo funzioni $F : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, con $n = 2, 3$ ovvero applicazioni che associano a punti di \mathbb{R}^n un vettore con n componenti. Tali funzioni permettono di modellizzare il concetto fisico di campo vettoriale.

Consideriamo un sottoinsieme $\Omega \subset \mathbb{R}^n$, con $n = 2, 3$, aperto e connesso⁵. Definiamo *campo vettoriale* su Ω una funzione $F : \Omega \rightarrow \mathbb{R}^n$. Tratteremo in dettaglio il caso in cui $n = 3$, ma le definizioni ed i risultati che andremo ad enunciare si adattano facilmente al caso $n = 2$.

⁵Ricordiamo che un insieme $\Omega \subset \mathbb{R}^n$ è connesso se scelti due punti qualsiasi $P_1, P_2 \in \Omega$, esiste un arco di curva continua γ contenuto in Ω che collega P_1 e P_2 .

Dato un campo vettoriale $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$, definiamo le sue funzioni *componenti* $F_i : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, con $i = 1, 2, 3$ tramite la relazione

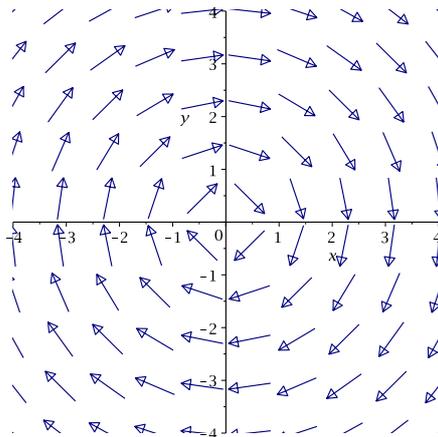
$$\begin{aligned} F(x, y, z) &= (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z)) \\ &= F_1(x, y, z)\mathbf{i} + F_2(x, y, z)\mathbf{j} + F_3(x, y, z)\mathbf{k} \quad (x, y, z) \in \Omega \end{aligned}$$

Nel caso in cui $n = 2$ le componenti $F_i : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, con $i = 1, 2$ sono le funzioni definite da:

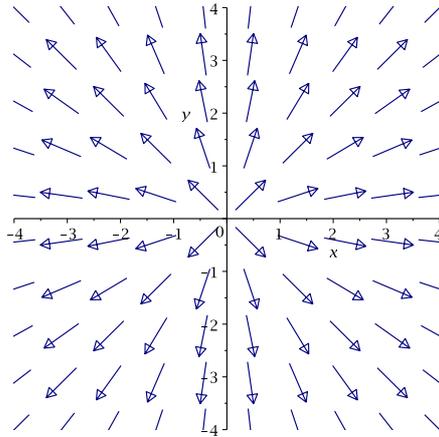
$$\begin{aligned} F(x, y) &= (F_1(x, y), F_2(x, y)) \\ &= F_1(x, y)\mathbf{i} + F_2(x, y)\mathbf{j} \quad (x, y) \in \Omega \end{aligned}$$

In questa sezione supporremo sempre che il campo F sia di classe $C^1(\Omega)$, cioè che le sue funzioni componenti siano derivabili e che le derivate parziali siano funzioni continue in tutti i punti di Ω .

Un campo vettoriale può venire rappresentato graficamente scegliendo un certo numero di punti $P_i \in \Omega$ e disegnando in ognuno di tali punti il vettore $F(P_i)$ applicato nel punto P_i .



Il campo vettoriale $F(x, y) = \left(-\frac{y}{\sqrt{x^2+y^2}}, \frac{x}{\sqrt{x^2+y^2}} \right)$



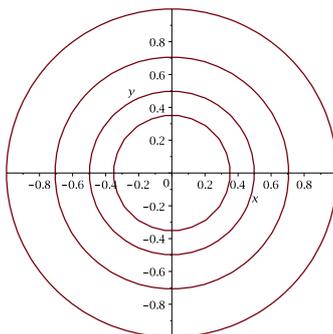
Il campo vettoriale $F(x, y) = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}} \right)$

Una rappresentazione grafica alternativa dell'andamento di un campo vettoriale può essere ottenuta tramite le *linee di campo*. Dato un campo vettoriale $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ di classe $C^1(\Omega)$, si definisce linea di campo di F ogni curva regolare tangente in ogni punto ad F . Più precisamente, se $\gamma \subset \Omega$ è una linea di campo di F e $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$ è una sua parametrizzazione, allora esiste una funzione scalare $\lambda : I \subset \mathbb{R} \rightarrow \mathbb{R}$ tale che

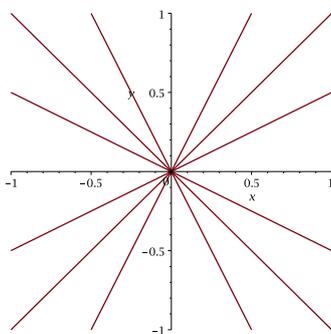
$$\alpha'(t) = \lambda(t)F(\alpha(t)), \quad \forall t \in I$$

In particolare se $\alpha(t) = (x(t), y(t), z(t))$ allora

$$\begin{cases} x'(t) = \lambda(t)F_1(x(t), y(t), z(t)) \\ y'(t) = \lambda(t)F_2(x(t), y(t), z(t)) \\ z'(t) = \lambda(t)F_3(x(t), y(t), z(t)) \end{cases}$$



linee di campo di $F(x, y) = \left(-\frac{y}{\sqrt{x^2+y^2}}, \frac{x}{\sqrt{x^2+y^2}} \right)$



linee di campo di $F(x, y) = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}} \right)$

Gradiente, rotore e divergenza

Consideriamo una funzione $f : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ di classe $C^1(\Omega)$. Il vettore gradiente di f definisce un campo vettoriale continuo su Ω :

$$(x, y, z) \mapsto \nabla f(x, y, z) = \left(\frac{\partial}{\partial x} f(x, y, z), \frac{\partial}{\partial y} f(x, y, z), \frac{\partial}{\partial z} f(x, y, z), \right)$$

Se indichiamo con $C^k(\Omega, \mathbb{R})$ l'insieme delle funzioni $f : \Omega \rightarrow \mathbb{R}$ di classe C^k e con $C^{k'}(\Omega, \mathbb{R}^3)$ l'insieme dei campi vettoriali di classe $C^{k'}$, allora possiamo definire l'operazione di calcolo del gradiente come una mappa $\nabla : C^k(\Omega, \mathbb{R}) \rightarrow C^{k-1}(\Omega, \mathbb{R}^3)$ (dove $k \geq 1$):

$$f \mapsto \nabla f$$

Inoltre, definendo in $C^k(\Omega, \mathbb{R})$ e in $C^{k'}(\Omega, \mathbb{R}^3)$ le operazioni di somma e di prodotto per scalare come:

$$(f_1 + f_2)(x, y, z) := f_1(x, y, z) + f_2(x, y, z), \quad (\lambda f)(x, y, z) := \lambda f(x, y, z),$$

$$(F_1 + F_2)(x, y, z) := F_1(x, y, z) + F_2(x, y, z), \quad (\lambda F)(x, y, z) := \lambda F(x, y, z),$$

dove $\lambda \in \mathbb{R}$, $f, f_1, f_2 \in C^k(\Omega, \mathbb{R})$, $F, F_1, F_2 \in C^{k'}(\Omega, \mathbb{R}^3)$, è facile verificare che $C^k(\Omega, \mathbb{R})$ e $C^{k'}(\Omega, \mathbb{R}^3)$ sono spazi vettoriali e $\nabla : C^k(\Omega, \mathbb{R}) \rightarrow C^{k-1}(\Omega, \mathbb{R}^3)$ è un operatore lineare.

Introduciamo ora l'operatore rotore. Dato un campo vettoriale $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ di classe $C^1(\Omega)$, definiamo *rotore di F*, indicato con $\text{rot } F$ oppure $\nabla \times F$, il campo vettoriale

$$\text{rot } F = \left(\frac{\partial}{\partial y} F_3 - \frac{\partial}{\partial z} F_2, \frac{\partial}{\partial z} F_1 - \frac{\partial}{\partial x} F_3, \frac{\partial}{\partial x} F_2 - \frac{\partial}{\partial y} F_1 \right)$$

Da un punto di vista astratto, possiamo vedere il rotore come un operatore lineare $\text{rot} : C^k(\Omega, \mathbb{R}^3) \rightarrow C^{k-1}(\Omega, \mathbb{R}^3)$.

Anche se l'operazione di calcolo del rotore è stata definita per campi $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, è possibile definire il rotore di un campo piano $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, con $F(x, y) = F_1(x, y)\mathbf{i} + F_2(x, y)\mathbf{j}$ come $\text{rot } F(x, y) = \left(\frac{\partial}{\partial x} F_2 - \frac{\partial}{\partial y} F_1 \right) \mathbf{k}$.

Un campo $F \in C^1(\Omega, \mathbb{R}^3)$ tale che $\text{rot } F = (0, 0, 0)$ è detto *irrotazionale*.

Introduciamo infine la *divergenza* di un campo vettoriale $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ di classe $C^1(\Omega)$, come la funzione $\text{div } F : \Omega \rightarrow \mathbb{R}$ definita da:

$$\text{div } F(x, y, z) := \frac{\partial}{\partial x} F_1(x, y, z) + \frac{\partial}{\partial y} F_2(x, y, z) + \frac{\partial}{\partial z} F_3(x, y, z)$$

Da un punto di vista astratto, possiamo vedere la divergenza come un operatore lineare $\text{div} : C^k(\Omega, \mathbb{R}^3) \rightarrow C^{k-1}(\Omega, \mathbb{R})$, dove $k \geq 1$.

Un campo con divergenza nulla è detto *solenoidale*.

Integrali di linea di campi vettoriali (integrali curvilinei di seconda specie)

Definizione 37 Sia $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un campo vettoriale continuo e $\gamma \subset \Omega$ un arco di curva regolare a tratti parametrizzata dalla mappa $r : [a, b] \rightarrow \mathbb{R}^3$, con $r(t) = (x(t), y(t), z(t))$. Definiamo integrale di linea (o lavoro) di F lungo γ l'integrale

$$\int_{\gamma} F \cdot dr := \int_a^b F(r(t)) \cdot r'(t) dt.$$

L'integrale di linea appena definito permette di generalizzare il concetto di lavoro fatto da un campo di forze per spostare il punto di applicazione da $r(a)$ a $r(b)$ lungo il percorso γ . In particolare, se F è un campo costante e γ è il segmento che congiunge A e B , allora $\int_{\gamma} F \cdot dr = F \cdot (B - A)$.

L'integrale $\int_{\gamma} F \cdot dr$ può essere rappresentato come l'integrale di linea di prima specie della componente del campo F tangente alla curva γ :

$$\int_{\gamma} F \cdot dr = \int_{\gamma} F \cdot T ds,$$

infatti lungo i tratti regolari dell'arco di curva γ , abbiamo:

$$\int_a^b F(r(t)) \cdot r'(t) dt = \int_a^b F(r(t)) \cdot \frac{r'(t)}{\|r'(t)\|} \|r'(t)\| dt = \int_{\gamma} F \cdot T ds$$

Se la curva γ è semplice e chiusa, allora l'integrale $\int_{\gamma} F \cdot dr$ viene indicato con il simbolo $\oint_{\gamma} F \cdot dr$ e chiamato *circuitazione* di F .

L'integrale di linea di seconda specie $\int_{\gamma} F \cdot dr$ dipende dall'orientazione della curva γ , ovvero dal suo verso di percorrenza. Se indichiamo con γ^* la curva γ percorsa in senso opposto (partendo da $r(b)$ e terminando in $r(a)$) allora il segno dell'integrale cambia:

$$\int_{\gamma} F \cdot dr = - \int_{\gamma^*} F \cdot dr$$

L'integrale $\int_{\gamma} F \cdot dr$ è inoltre invariante per riparametrazioni. Più precisamente, se $r_1 : [a, b] \rightarrow \mathbb{R}^3$ e $r_2 : [c, d] \rightarrow \mathbb{R}^3$ sono due parametrizzazioni differenti di γ , allora

$$\int_a^b F(r_1(t)) \cdot r_1'(t) dt = C \int_c^d F(r_2(t)) \cdot r_2'(t) dt,$$

dove $C = 1$ se r_1 e r_2 inducono la stessa orientazione di γ (cioè se $r_1(a) = r_2(c)$ e $r_1(b) = r_2(d)$), mentre $C = -1$ se r_1 e r_2 inducono orientazioni differenti di γ (cioè se $r_1(a) = r_2(d)$ e $r_1(b) = r_2(c)$).

Campi conservativi

In generale il lavoro di un campo F lungo una curva γ , ovvero l'integrale di linea $\int_{\gamma} F \cdot dr$, dipende sia dal campo che dal percorso γ . In generale quindi se γ_1 e γ_2 sono due curve differenti, anche se hanno lo stesso punto iniziale e lo stesso punto finale, non ci possiamo aspettare che $\int_{\gamma_1} F \cdot dr = \int_{\gamma_2} F \cdot dr$. Esiste però una particolare classe di campi vettoriali per i quali l'integrale di linea lungo un arco γ dipende solo dagli estremi di γ .

Definizione 38 Un campo vettoriale $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ è detto conservativo se esiste una funzione $U : \Omega \rightarrow \mathbb{R}$ di classe $C^2(\Omega)$ tale che $F = \nabla U$:

$$F(x, y, z) = \left(\frac{\partial}{\partial x} U(x, y, z), \frac{\partial}{\partial y} U(x, y, z), \frac{\partial}{\partial z} U(x, y, z), \right)$$

La funzione U è detta potenziale del campo F .

Lemma 1 Sia $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un campo conservativo, con $F = \nabla U$ e γ un arco di curva regolare a tratti di parametrizzazione $\alpha : [a, b] \rightarrow \mathbb{R}^3$. Indichiamo con $P_0 \equiv \alpha(a)$ e con $P_1 \equiv \alpha(b)$ i punti estremi di γ . Allora:

$$\int_{\gamma} F \cdot dr = U(P_1) - U(P_0)$$

In particolare, se γ è una curva chiusa, allora $\oint_{\gamma} F \cdot dr = 0$

Dimostrazione:

$$\int_{\gamma} F \cdot dr = \int_a^b F(\alpha(t)) \cdot \alpha'(t) dt = \int_a^b \nabla U(\alpha(t)) \cdot \alpha'(t) dt = \int_a^b (U \circ \alpha)'(t) dt = U(\alpha(b)) - U(\alpha(a))$$

dove nel terzo passaggio è stata applicata la formula di derivazione di funzione composta. \square

In particolare se F è un campo conservativo e γ è una qualsiasi curva chiusa, per il lemma precedente abbiamo che la circuitazione di F è nulla:

$$\oint_{\gamma} F \cdot dr = 0$$

La seguente proposizione mostra come i campi conservativi sono i soli a possedere queste proprietà.

Proposizione 5 Sia $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un campo vettoriale di classe $C^1(\Omega)$. Le seguenti tre affermazioni sono equivalenti:

1. F è conservativo,
2. per ogni curva chiusa e semplice $\gamma \subset \Omega$

$$\oint_{\gamma} F \cdot dr = 0,$$

3. per ogni coppia di curve $\gamma_1, \gamma_2 \subset \Omega$ regolari a tratti e aventi lo stesso punto iniziale e lo stesso punto finale (e per il resto disgiunte) si ha

$$\int_{\gamma_1} F \cdot dr = \int_{\gamma_2} F \cdot dr.$$

Dimostrazione: 1. \Rightarrow 2. Se F è conservativo, per il lemma 1, abbiamo $\oint_{\gamma} F \cdot dr = U(P_1) - U(P_0)$. Dato che γ è chiusa, allora $P_1 = P_0$ e $\oint_{\gamma} F \cdot dr = 0$.

2. \Rightarrow 3. Consideriamo due curve $\gamma_1, \gamma_2 \subset \Omega$ regolari a tratti e aventi lo stesso punto iniziale e lo stesso punto finale (e per il resto disgiunte). Indichiamo con γ_2^* la curva ottenuta cambiando l'orientazione di γ_2 e con γ la curva chiusa semplice regolare a tratti ottenuta dall'unione di γ_1 e γ_2^* : $\gamma = \gamma_1 \cup \gamma_2^*$. Per il punto 2 abbiamo

$$0 = \oint_{\gamma} F \cdot dr = \int_{\gamma_1} F \cdot dr + \int_{\gamma_2^*} F \cdot dr = \int_{\gamma_1} F \cdot dr - \int_{\gamma_2} F \cdot dr,$$

da cui segue la condizione 3. Nella seconda uguaglianza abbiamo sfruttato l'addittività dell'integrale di linea rispetto al dominio di integrazione, mentre nella terza abbiamo utilizzato il cambiamento del segno dell'integrale per cambiamento di orientazione della curva.

3. \Rightarrow 1. Supponiamo ora che sia valida l'affermazione 3. e dimostriamo che F è conservativo, cioè che esiste una funzione $U : \Omega \rightarrow \mathbb{R}$ di classe $C^2(\Omega)$ tale che $F = \nabla U$. L'esistenza di tale funzione U si dimostra costruendola esplicitamente. Fissiamo un punto $(x_0, y_0, z_0) \in \Omega$ e definiamo il potenziale U in un generico punto $(x, y, z) \in \Omega$ come il lavoro di F lungo un cammino regolare a tratti γ da (x_0, y_0, z_0) a (x, y, z) :

$$U(x, y, z) := \int_{\gamma} F \cdot dr$$

La funzione U è definita in modo non ambiguo, perché, per la validità dell'affermazione 3, scelto un differente cammino $\tilde{\gamma}$ che congiunge (x_0, y_0, z_0) con (x, y, z) , il valore del lavoro da (x_0, y_0, z_0) a (x, y, z) lungo $\tilde{\gamma}$ è identico al valore del lavoro da (x_0, y_0, z_0) a (x, y, z) lungo γ . Mostriamo ora che $F = \nabla U$, cioè che $(F_1, F_2, F_3) = (\frac{\partial}{\partial x} U, \frac{\partial}{\partial y} U, \frac{\partial}{\partial z} U)$. Partiamo dalla dimostrazione che $F_1(x, y, z) = \frac{\partial}{\partial x} U(x, y, z)$. Dobbiamo calcolare

$$\frac{\partial}{\partial x} U(x, y, z) = \lim_{\Delta x \rightarrow 0} \frac{U(x + \Delta x, y, z) - U(x, y, z)}{\Delta x}$$

Se indichiamo con γ il cammino da (x_0, y_0, z_0) a (x, y, z) e con Γ il segmento da (x, y, z) a $(x + \Delta x, y, z)$ abbiamo che

$$U(x + \Delta x, y, z) = \int_{\gamma \cup \Gamma} F \cdot dr = \int_{\gamma} F \cdot dr + \int_{\Gamma} F \cdot dr = U(x, y, z) + \int_{\Gamma} F \cdot dr$$

Per il calcolo del lavoro $\int_{\Gamma} F \cdot dr$ parametrizziamo il segmento Γ congiungente (x, y, z) con $(x + \Delta x, y, z)$ tramite la funzione $\alpha : [0, 1] \rightarrow \mathbb{R}^3$:

$$\alpha(t) = (x, y, z) + t(\Delta x, 0, 0) = (x + t\Delta x, y, z), \quad t \in [0, 1].$$

Abbiamo dunque

$$\begin{aligned} \int_{\Gamma} F \cdot dr &= \int_0^1 (F_1(x + t\Delta x, y, z), F_2(x + t\Delta x, y, z), F_3(x + t\Delta x, y, z)) \cdot (\Delta x, 0, 0) dt \\ &= \Delta x \int_0^1 F_1(x + t\Delta x, y, z) dt \end{aligned}$$

e quindi

$$\begin{aligned} \frac{\partial}{\partial x} U(x, y, z) &= \lim_{\Delta x \rightarrow 0} \frac{U(x + \Delta x, y, z) - U(x, y, z)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \int_0^1 F_1(x + t\Delta x, y, z) dt \\ &= F_1(x, y, z) \end{aligned}$$

La dimostrazione che $\frac{\partial}{\partial y} U(x, y, z) = F_2(x, y, z)$ e che $\frac{\partial}{\partial z} U(x, y, z) = F_3(x, y, z)$ è del tutto analoga e viene lasciata per esercizio. \square

La proposizione precedente fornisce un criterio per caratterizzare i campi conservativi: sono tutti e soli quelli con circuitazione nulla. È quindi sufficiente trovare una curva chiusa γ per cui $\oint_{\gamma} F \cdot dr \neq 0$ per poter dire che F non può essere conservativo. D'altra parte, per riconoscere se F è conservativo, la proposizione precedente prescrive di verificare che la circuitazione di F è nulla su ogni curva chiusa semplice γ , richiedendo quindi, in linea di principio, un numero infinito di verifiche. Cerchiamo quindi ora un criterio alternativo che permetta di riconoscere se un campo è conservativo tramite un numero finito di calcoli.

Il seguente teorema dà una condizione necessaria affinché F sia conservativo.

Teorema 29 *Sia $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, un aperto connesso e F un campo vettoriale di classe $C^1(\Omega)$ e conservativo. Allora F è irrotazionale.*

Dimostrazione: Scriviamo $F = \nabla U$ e calcoliamo $\text{rot } F = \text{rot } \nabla U$, ottenendo

$$\text{rot } F = \left(\frac{\partial U}{\partial y \partial z} - \frac{\partial U}{\partial z \partial y}, \frac{\partial U}{\partial z \partial x} - \frac{\partial U}{\partial x \partial z}, \frac{\partial U}{\partial x \partial y} - \frac{\partial U}{\partial y \partial x} \right)$$

Per il teorema di Schwartz si annullano tutte le componenti del vettore al secondo membro. \square

L'implicazione opposta non è sempre vera, come mostra il seguente esempio

Esempio 35 Il campo vettoriale $F : \Omega \subset \mathbb{R}^2$, con $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$ e $F(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$ verifica la condizione $\text{rot } F(x,y) = (0,0)$ per ogni $(x,y) \in \Omega$, ma non è conservativo. Se calcoliamo infatti la circuitazione di F lungo la circonferenza γ di centro $(0,0)$ e raggio 1, otteniamo

$$\oint_{\gamma} F \cdot dr = 2\pi \neq 0$$

Affinché la condizione $\text{rot } F = 0$ sia sufficiente per garantire che $F = \nabla U$, è necessario aggiungere un'ipotesi sulle proprietà dell'insieme Ω di definizione del campo vettoriale F . Premettiamo la seguente definizione.

Definizione 39 Un insieme aperto $\Omega \subset \mathbb{R}^n$ è detto semplicemente connesso se è connesso e inoltre se ogni curva γ continua chiusa e semplice, interamente contenuta in Ω , può essere ridotta ad un punto mediante una deformazione continua senza mai uscire da Ω

Ad esempio, in \mathbb{R}^2 sono semplicemente connessi cerchi, poligoni, semipiani e \mathbb{R}^2 stesso, mentre non sono semplicemente connessi tutti questi insiemi se privati di un punto. Inoltre non è semplicemente connessa la corona circolare. In \mathbb{R}^3 sono semplicemente connessi le sfere, i poligoni, i semispazi e \mathbb{R}^3 stesso, anche se privati di un numero finito di punti. Non sono invece semplicemente connessi il toro, la sfera privata di un diametro e \mathbb{R}^3 privato di una retta.

Teorema 30 Sia $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, un aperto semplicemente connesso e F un campo vettoriale di classe $C^1(\Omega)$. Allora F è conservativo se e solo se F è irrotazionale.

Campi centrali

In questa sezione parleremo di campi centrali cioè di campi vettoriali $F : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ (dove $\Omega = \mathbb{R}^3$ oppure $\Omega = \mathbb{R}^3 \setminus \{(0,0,0)\}$) della forma $F(x,y,z) = f(x,y,z)\hat{e}_r$, dove $f : \Omega \rightarrow \mathbb{R}$ è una funzione di classe $C^1(\Omega)$, mentre \hat{e}_r è il versore

$$\hat{e}_r = \left(\frac{x}{\sqrt{x^2+y^2+z^2}}, \frac{y}{\sqrt{x^2+y^2+z^2}}, \frac{z}{\sqrt{x^2+y^2+z^2}} \right) = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r} \right)$$

Abbiamo introdotto la variabile $r = \sqrt{x^2+y^2+z^2}$, che esprime la distanza del punto (x,y,z) dall'origine $(0,0,0)$. Notiamo che $\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2+z^2}}$, $\frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2+y^2+z^2}}$ e $\frac{\partial r}{\partial z} = \frac{z}{\sqrt{x^2+y^2+z^2}}$, quindi possiamo rappresentare un campo centrale nella forma:

$$F(x,y,z) = f(x,y,z) \left(\frac{\partial r}{\partial x}, \frac{\partial r}{\partial y}, \frac{\partial r}{\partial z} \right)$$

Un campo vettoriale centrale è detto *a simmetria sferica* se la funzione $f(x, y, z)$ dipende solo da r , cioè se esiste una funzione $\tilde{f} : [0, +\infty) \rightarrow \mathbb{R}$ tale che $f(x, y, z) = \tilde{f}(\sqrt{x^2 + y^2 + z^2})$.

Non è complicato verificare che un campo vettoriale centrale a simmetria sferica è conservativo. Se infatti $g : [0, +\infty) \rightarrow \mathbb{R}$ è una primitiva della funzione \tilde{f} , allora la funzione $U : \Omega \rightarrow \mathbb{R}$, definita da $U(x, y, z) = g(\sqrt{x^2 + y^2 + z^2})$ è un potenziale per il campo vettoriale F .

Viceversa, supponiamo che F sia un campo vettoriale centrale e conservativo e dimostriamo che F è a simmetria sferica. Introducendo le coordinate polari sferiche (r, ϕ, θ) ed esprimendo il potenziale U del campo F tramite tali coordinate, cioè introducendo la funzione $\tilde{U} : [0, +\infty) \times [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{R}$ definita da

$$\tilde{U}(r, \phi, \theta) = U(x(r, \phi, \theta), y(r, \phi, \theta), z(r, \phi, \theta)) = U(r \sin \phi \cos \theta, r \sin \phi \sin \theta, r \cos \phi),$$

Vogliamo mostrare che \tilde{U} dipende solo da r , cioè che $\frac{\partial \tilde{U}}{\partial \phi} = 0$ e $\frac{\partial \tilde{U}}{\partial \theta} = 0$. Notiamo che, viceversa, il potenziale U in coordinate cartesiane è legato al potenziale \tilde{U} in coordinate polari dalla relazione $U(x, y, z) = \tilde{U}(r(x, y, z), \phi(x, y, z), \theta(x, y, z))$, abbiamo quindi:

$$\begin{aligned} \frac{\partial U}{\partial x} &= \frac{\partial \tilde{U}}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial \tilde{U}}{\partial \phi} \frac{\partial \phi}{\partial x} + \frac{\partial \tilde{U}}{\partial \theta} \frac{\partial \theta}{\partial x} = f(r, \phi, \theta) \frac{\partial r}{\partial x} \\ \frac{\partial U}{\partial y} &= \frac{\partial \tilde{U}}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial \tilde{U}}{\partial \phi} \frac{\partial \phi}{\partial y} + \frac{\partial \tilde{U}}{\partial \theta} \frac{\partial \theta}{\partial y} = f(r, \phi, \theta) \frac{\partial r}{\partial y} \\ \frac{\partial U}{\partial z} &= \frac{\partial \tilde{U}}{\partial r} \frac{\partial r}{\partial z} + \frac{\partial \tilde{U}}{\partial \phi} \frac{\partial \phi}{\partial z} + \frac{\partial \tilde{U}}{\partial \theta} \frac{\partial \theta}{\partial z} = f(r, \phi, \theta) \frac{\partial r}{\partial z} \end{aligned}$$

Le tre equazioni messe a sistema possono essere scritte nella forma

$$\begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial \phi}{\partial x} & \frac{\partial \theta}{\partial x} \\ \frac{\partial r}{\partial y} & \frac{\partial \phi}{\partial y} & \frac{\partial \theta}{\partial y} \\ \frac{\partial r}{\partial z} & \frac{\partial \phi}{\partial z} & \frac{\partial \theta}{\partial z} \end{pmatrix} \begin{pmatrix} \frac{\partial \tilde{U}}{\partial r} - f(r, \phi, \theta) \\ \frac{\partial \tilde{U}}{\partial \phi} \\ \frac{\partial \tilde{U}}{\partial \theta} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Tale sistema omogeneo, dato che la matrice dei coefficienti $\begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial \phi}{\partial x} & \frac{\partial \theta}{\partial x} \\ \frac{\partial r}{\partial y} & \frac{\partial \phi}{\partial y} & \frac{\partial \theta}{\partial y} \\ \frac{\partial r}{\partial z} & \frac{\partial \phi}{\partial z} & \frac{\partial \theta}{\partial z} \end{pmatrix}$ ha determinante non nullo, ha come unica soluzione

$$\frac{\partial \tilde{U}}{\partial r} - f(r, \phi, \theta) = 0, \quad \frac{\partial \tilde{U}}{\partial \phi} = 0, \quad \frac{\partial \tilde{U}}{\partial \theta} = 0$$

da cui possiamo dedurre che \tilde{U} è una funzione solo della variabile radiale r e quindi il campo vettoriale F è a simmetria centrale.

La formula di Gauss-Green nel piano

Introduciamo ora la prima di una serie di formule che mettono in relazione integrali su insiemi $D \subset \mathbb{R}^n$, con $n = 2, 3$ con integrali sulla frontiera ∂D . Di fatto, nel corso di analisi 1 è già stata introdotta una formula di questo tipo. Il teorema fondamentale del calcolo integrale, che per una funzione $f \in C^1([a, b])$ afferma

$$\int_a^b f'(x)dx = f(b) - f(a), \quad (37)$$

non è solamente una formula che riduce l'integrazione sulla retta reale alla ricerca di primitive. Può essere anche vista come una relazione che collega l'integrale sull'intervallo $I = [a, b]$ di f' alla valutazione di f sulla frontiera ∂I dell'intervallo I , formata dai soli due punti a, b . È in quest'ottica che possiamo lavorare per generalizzare la formula (37) al caso in cui I è sostituito da un sottoinsieme $D \subset \mathbb{R}^n$, con $n > 1$ e f è una funzione di n variabili reali.

Consideriamo un sottoinsieme $D \subset \mathbb{R}^2$ limitato e semplice rispetto ad entrambi gli assi, in modo tale che possiamo rappresentare D come:

$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\} \quad (38)$$

dove $g_1, g_2 : [a, b] \rightarrow \mathbb{R}$ sono continue, oppure come:

$$D = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\} \quad (39)$$

dove $h_1, h_2 : [c, d] \rightarrow \mathbb{R}$ sono continue. Supponiamo che la frontiera ∂D di D sia una curva regolare a tratti e orientiamola in modo tale che percorrendo tale curva l'insieme D sia sempre a sinistra. Indichiamo con $\partial^+ D$ la curva orientata in tal modo. Consideriamo due funzioni $P, Q : D \rightarrow \mathbb{R}$ di classe $C^1(D)$. Allora vale il seguente risultato

Lemma 2 1. Se D è y -semplice (cioè della forma (38)), e $F : D \rightarrow \mathbb{R}^2$ è il campo vettoriale $F(x, y) = (P(x, y), 0)$, allora

$$\oint_{\partial^+ D} F \cdot dr = - \int \int_D \frac{\partial P}{\partial y}(x, y) dx dy \quad (40)$$

2. Se D è x -semplice (cioè della forma (39)), e $F : D \rightarrow \mathbb{R}^2$ è il campo vettoriale $F(x, y) = (0, Q(x, y))$, allora

$$\oint_{\partial^+ D} F \cdot dr = \int \int_D \frac{\partial Q}{\partial x}(x, y) dx dy \quad (41)$$

Dimostrazione: Dimostriamo la parte 1. del lemma. La dimostrazione della parte 2. è analoga.

Utilizzando la rappresentazione di D come insieme y -semplice, abbiamo:

$$\begin{aligned} - \int \int_D \frac{\partial P}{\partial y}(x, y) dx dy &= - \int_a^b \left(\int_{g_1(x)}^{g_2(x)} \frac{\partial P}{\partial y}(x, y) dy \right) dx \\ &= \int_a^b (P(x, g_1(x)) - P(x, g_2(x))) dx \end{aligned}$$

Usando ancora la rappresentazione (38) per l'insieme D abbiamo che il bordo $\partial^+ D$ è composto da 4 curve:

- la curva γ_1 , grafico della funzione $g_1 : [a, b] \rightarrow \mathbb{R}$, di parametrizzazione $\alpha_1(x) = (x, g_1(x))$, $x \in [a, b]$, percorsa dal punto $(a, g_1(a))$, al punto $(b, g_1(b))$. Lungo tale curva orientata il lavoro del campo F è $\int_{\gamma_1} F \cdot dr = \int_a^b P(x, g_1(x)) dx$.
- il segmento γ_2 dal punto $(b, g_1(b))$ al punto $(b, g_2(b))$. Lungo tale curva orientata il lavoro del campo F è nullo perché il campo F è ortogonale al vettore tangente alla curva
- la curva γ_3 , grafico della funzione $g_2 : [a, b] \rightarrow \mathbb{R}$, di parametrizzazione $\alpha_2(x) = (x, g_2(x))$, $x \in [a, b]$, percorsa dal punto $(b, g_2(b))$, al punto $(a, g_2(a))$. Lungo tale curva orientata il lavoro del campo F è $\int_{\gamma_3} F \cdot dr = - \int_a^b P(x, g_2(x)) dx$.
- il segmento γ_4 dal punto $(a, g_2(a))$ al punto $(a, g_1(a))$. Lungo tale curva orientata il lavoro del campo F è nullo perché il campo F è ortogonale al vettore tangente alla curva.

Complessivamente quindi $\oint_{\partial^+ D} F \cdot dr = \int_a^b (P(x, g_1(x)) - P(x, g_2(x))) dx$ e abbiamo verificato l'uguaglianza (40). \square

Teorema 31 (Formula di Gauss-Green nel piano) *Sia $D \subset \mathbb{R}^2$ semplice rispetto ad entrambi gli assi e $F = (P, Q)$ un campo vettoriale di classe $C^1(D)$. Allora*

$$\oint_{\partial^+ D} F \cdot dr = \int \int_D \left(\frac{\partial Q}{\partial x}(x, y) - \frac{\partial P}{\partial y}(x, y) \right) dx dy \quad (42)$$

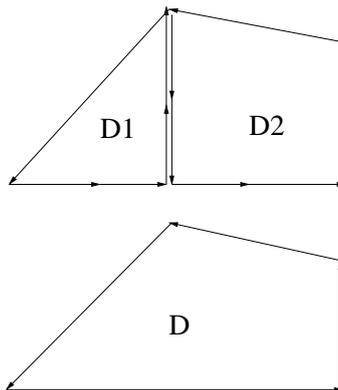
dove $\partial^+ D$ è la frontiera di D orientata in modo tale che percorrendo tale curva l'insieme D rimane sempre a sinistra della curva.

Dimostrazione: Dato che D è semplice rispetto ad entrambi gli assi sono valide entrambe le formule (40) e (41). Abbiamo quindi

$$\begin{aligned} \oint_{\partial^+ D} F \cdot dr &= \oint_{\partial^+ D} (P, Q) \cdot dr = \oint_{\partial^+ D} (P, 0) \cdot dr + \oint_{\partial^+ D} (0, Q) \cdot dr \\ &= - \int \int_D \frac{\partial P}{\partial y}(x, y) dx dy + \int \int_D \frac{\partial Q}{\partial x}(x, y) dx dy \\ &= \int \int_D \left(\frac{\partial Q}{\partial x}(x, y) - \frac{\partial P}{\partial y}(x, y) \right) dx dy \end{aligned}$$

□

Il risultato precedente è valido anche nel caso in cui D non è semplice rispetto ad entrambi gli assi (una condizione in effetti abbastanza restrittiva). La formula (42) è valida anche se D è un insieme limitato *s-decomponibile*, cioè un insieme che può essere decomposto nell'unione finita di insiemi semplici rispetto ad entrambi gli assi, a due a due disgiunti.



Analizziamo ad esempio il caso in cui D sia rappresentabile come unione di due insiemi semplici rispetto ad entrambi gli assi D_1 e D_2 e disgiunti. Abbiamo che:

$$\int \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int \int_{D_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy + \int \int_{D_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy$$

applicando la formula di Gauss-Green ad ognuno dei due insiemi D_1 e D_2 abbiamo:

$$\int \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \oint_{\partial^+ D_1} F \cdot dr + \oint_{\partial^+ D_2} F \cdot dr$$

dove $\partial^+ D_1$ e $\partial^+ D_2$ indicano rispettivamente la frontiera di D_1 e la frontiera di D_2 percorse in senso antiorario. Sommando gli integrali di linea $\oint_{\partial^+ D_1} F \cdot dr$ e $\oint_{\partial^+ D_2} F \cdot dr$, abbiamo che la parte dell'integrale $\oint_{\partial^+ D_1} F \cdot dr$ lungo la curva che fa da confine tra D_1 e D_2 è uguale e contrario alla parte del lavoro $\oint_{\partial^+ D_2} F \cdot dr$ lungo la stessa curva. Dato che questi due contributi si elidono a vicenda, abbiamo

$$\oint_{\partial^+ D_1} F \cdot dr + \oint_{\partial^+ D_2} F \cdot dr = \oint_{\partial^+ D} F \cdot dr$$

ottenendo in questo modo la formula di Gauss-Green per $D = D_1 \cup D_2$. Lo stesso discorso si generalizza con un procedimento induttivo ad insiemi D formati da un numero arbitrario n di insiemi semplici rispetto ad entrambi gli assi e disgiunti a due a due.

Applicazione al calcolo di aree di figure piane

Consideriamo un insieme D s-decomponibile. L'area di D è definita come l'integrale doppio $\int \int_D 1 dx dy$. Se consideriamo il campo vettoriale $F(x, y) = (-y/2, x/2)$, con $P(x, y) = -y/2$ e $Q(x, y) = x/2$, allora $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$ e

$$A(D) = \int \int_D 1 dx dy = \int \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \oint_{\partial^+ D} F \cdot dr$$

Tale formula riduce il calcolo dell'area di D a quello di un integrale di linea lungo il bordo di D .

Esempio 36 *Calcolare l'area della figura piana racchiusa dalla curva, detta astroide di parametrizzazione*

$$\alpha(\theta) = (\cos^3 \theta, \sin^3 \theta), \quad \theta \in [0, 2\pi]$$

Utilizzando la formula $A(D) = \oint_{\partial^+ D} F \cdot dr$, con $F(x, y) = (-y/2, x/2)$, otteniamo

$$\begin{aligned} A(D) &= \oint_{\partial^+ D} F \cdot dr = \int_0^{2\pi} \left(-\frac{\sin^3 \theta}{2}, \frac{\cos^3 \theta}{2} \right) \cdot (-3 \cos^2 \theta \sin \theta, 3 \sin^2 \theta \cos \theta) d\theta \\ &= \frac{3}{2} \int_0^{2\pi} \cos^2 \theta \sin^2 \theta d\theta = \frac{3}{8} \pi \end{aligned}$$

Superfici e integrali di superficie

Superfici regolari in forma parametrica

Consideriamo un insieme $A \subset \mathbb{R}^2$ limitato chiusura di un aperto connesso e una funzione $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ differenziabile:

$$(u, v) \mapsto r(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in A.$$

la mappa $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ viene detta *parametrizzazione di superficie* e la sua immagine $\Sigma \equiv r(A) = \{(x, y, z) \in \mathbb{R}^3 : (x, y, z) = r(u, v), (u, v) \in A\}$ è un sottoinsieme di \mathbb{R}^3 che viene detto *superficie in forma parametrica*. La mappa $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ descrive le 3 coordinate cartesiane dei punti appartenenti alla superficie Σ in funzione di 2 parametri indipendenti $(u, v) \in A \subset \mathbb{R}^2$. Vediamo qui di seguito alcuni esempi.

1. Una particolare classe di superfici $\Sigma \subset \mathbb{R}^3$ sono le *superfici cartesiane*, cioè le superfici grafici di funzioni $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. Una possibile parametrizzazione per tali superfici è la funzione $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$:

$$r(x, y) = (x, y, f(x, y)), \quad (x, y) \in A$$

Se $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ è differenziabile, allora lo è anche la parametrizzazione $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$.

2. La superficie sferica $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = R^2\}$ può essere parametrizzata tramite la mappa $r : [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{R}^3$:

$$r(\phi, \theta) = (R \sin \phi \cos \theta, R \sin \phi \sin \theta, R \cos \phi), \quad (\phi, \theta) \in [0, \pi] \times [0, 2\pi]$$

che descrive le coordinate dei punti $P = (x, y, z) \in \Sigma$ tramite gli angoli ϕ, θ delle coordinate polari sferiche in \mathbb{R}^3

3. La superficie laterale del cilindro $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = R^2, a \leq z \leq b\}$ può essere parametrizzata tramite la mappa $r : [0, 2\pi] \times [a, b] \rightarrow \mathbb{R}^3$:

$$r(\theta, z) = (R \cos \theta, R \sin \theta, z), \quad (\theta, z) \in [0, 2\pi] \times [a, b]$$

4. La superficie Σ ottenuta dalla rotazione di una curva continua γ appartenente al piano yz di equazione parametrica $\alpha(t) = (y(t), z(t))$, $t \in I \subset \mathbb{R}$, può essere parametrizzata come:

$$r(t, \theta) = (-y(t) \sin \theta, y(t) \cos \theta, z(t)), \quad t \in I, \theta \in [0, 2\pi]$$

Consideriamo ora una generica superficie Σ in forma parametrica, descritta dalla parametrizzazione $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ differenziabile:

$$(u, v) \mapsto r(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in A.$$

L'immagine attraverso la mappa r delle rette nel piano uv parallele rispettivamente all'asse delle u e all'asse delle v sono delle curve che giacciono sulla superficie Σ chiamate *linee coordinate*. In particolare, se consideriamo l'immagine attraverso la mappa $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ della retta di equazione $u = u_0$ (parallela all'asse delle v), più precisamente della curva $\gamma_{u_0} = \{(u, v) \in A, u = u_0\}$, otteniamo la curva in \mathbb{R}^3 di parametrizzazione

$$\alpha_{u_0}(v) = r(u_0, v) = (x(u_0, v), y(u_0, v), z(u_0, v)), \quad v \in I \subset \mathbb{R}.$$

Analogamente, se consideriamo l'immagine attraverso la mappa $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ della retta di equazione $v = v_0$ (parallela all'asse delle u), più precisamente della curva $\gamma_{v_0} = \{(u, v) \in A, v = v_0\}$, otteniamo la curva in \mathbb{R}^3 di parametrizzazione

$$\alpha_{v_0}(u) = r(u, v_0) = (x(u, v_0), y(u, v_0), z(u, v_0)), \quad u \in I \subset \mathbb{R}.$$

Al variare di u_0 e v_0 otteniamo due famiglie di curve che giacciono sulla superficie Σ . Se la parametrizzazione $r : A \rightarrow \mathbb{R}^3$ è iniettiva tali curve non si intersecano.

Se r è differenziabile, allora per il teorema di derivazione delle funzioni composte abbiamo che le funzioni α_{u_0} e α_{v_0} , parametrizzazioni delle linee coordinate sono derivabili e i vettori tangenti alle linee coordinate in un punto $(x, y, z) \in \Sigma$ della forma $(x, y, z) = r(u_0, v_0)$ sono dati da:

$$\begin{aligned} \mathbf{v}_1 &= \frac{\partial}{\partial u} r(u_0, v_0) = \left(\frac{\partial x}{\partial u}(u_0, v_0), \frac{\partial y}{\partial u}(u_0, v_0), \frac{\partial z}{\partial u}(u_0, v_0) \right) \\ \mathbf{v}_2 &= \frac{\partial}{\partial v} r(u_0, v_0) = \left(\frac{\partial x}{\partial v}(u_0, v_0), \frac{\partial y}{\partial v}(u_0, v_0), \frac{\partial z}{\partial v}(u_0, v_0) \right) \end{aligned}$$

Se i due vettori sono linearmente indipendenti allora individuano un piano passante per il punto $r(u_0, v_0)$ e ortogonale al vettore (non nullo) $N(u_0, v_0) := \mathbf{v}_1 \wedge \mathbf{v}_2$. Le coordinate (x, y, z) dei punti di tale piano verificano l'equazione:

$$(x - x(u_0, v_0), y - y(u_0, v_0), z - z(u_0, v_0)) \cdot N(u_0, v_0) = 0 \quad (43)$$

Si può dimostrare che se $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ è differenziabile, allora il piano di equazione (43) contiene non solo le rette tangenti alle due linee coordinate passanti per il punto $r(u_0, v_0)$, ma tutte le rette tangenti a Σ in $r(u_0, v_0)$. L'equazione (43)

descrive quindi il piano tangente alla superficie Σ in $r(u_0, v_0)$. La condizione di non annullamento del prodotto vettoriale $\mathbf{v}_1 \wedge \mathbf{v}_2$ equivale quindi alla possibilità di definire il piano tangente alla superficie. Queste considerazioni portano dunque alla definizione di superficie regolare.

Definizione 40 Una superficie $\Sigma \subset \mathbb{R}^3$ parametrizzata da $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ è detta regolare se la funzione r è differenziabile in A e se per ogni $(u, v) \in A$ è verificata la condizione $\frac{\partial}{\partial u}r(u, v) \wedge \frac{\partial}{\partial v}r(u, v) \neq (0, 0, 0)$.

Area di una superficie regolare e integrali di superficie di funzioni $f : \mathbb{R}^3 \rightarrow \mathbb{R}$

Consideriamo una superficie regolare Σ parametrizzata dalla funzione differenziabile $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$

$$(u, v) \mapsto r(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in A.$$

Supponiamo che A sia s -decomponibile e che la funzione r sia iniettiva sull'interno di A . Definiamo area di Σ il seguente integrale

$$Area(\Sigma) = \int \int_A \|N(u, v)\| dudv \quad (44)$$

La formula 44 è una buona definizione, infatti l'integrale (44) non dipende dalla parametrizzazione di Σ . È motivata dalla considerazione che un rettangolo infinitesimo $Q = [u, u + du] \times [v, v + dv]$ contenuto in A viene trasformato attraverso la funzione r in una porzione di superficie Σ compresa fra due linee coordinate. L'area di tale "parallelogramma curvilineo" è approssimata con quella del parallelogramma individuato dai vettori $\frac{\partial}{\partial u}r(u, v)du$ e $\frac{\partial}{\partial v}r(u, v)dv$, di valore

$$dS = \left| \frac{\partial}{\partial u}r(u, v)du \wedge \frac{\partial}{\partial v}r(u, v)dv \right| = \|N(u, v)\|dudv$$

Esempio 37 Nel caso della superficie sferica $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = R^2\}$ parametrizzata tramite la mappa $r : [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{R}^3$:

$$r(\phi, \theta) = (R \sin \phi \cos \theta, R \sin \phi \sin \theta, R \cos \phi), \quad (\phi, \theta) \in [0, \pi] \times [0, 2\pi]$$

Abbiamo

$$N(u, v) = R^2 \sin \phi (\sin \phi \cos \theta, \sin \phi \sin \theta, \cos \phi), \quad \|N(u, v)\| = R^2 \sin \phi$$

e l'area è data da

$$Area(\Sigma) = \int \int_{[0, \pi] \times [0, 2\pi]} R^2 \sin \phi d\phi d\theta = R^2 \int_0^\pi \sin \phi d\phi \int_0^{2\pi} d\theta = 4\pi R^2$$

Esempio 38 (Esercizio proposto dal dott. E. Bottazzi)

La pseudosfera è la superficie Σ di equazione parametrica

$$r(u, v) = (\cos u \sin v, \sin u \sin v, \cos v + \log(\tan(v/2))) \quad u \in [0, 2\pi], v \in (0, \pi)$$

Ricordiamo che

$$\begin{aligned} \frac{\partial \log(\tan(v/2))}{\partial v} &= \frac{1}{2} \frac{1}{\tan(v/2)} \sec^2(v/2) \\ &= \frac{1}{2} \frac{\cos(v/2)}{\sin(v/2) \cos^2(v/2)} \\ &= \frac{1}{2 \sin(v/2) \cos(v/2)} \\ &= \frac{1}{\sin v} \end{aligned}$$

Quindi

$$\frac{\partial r}{\partial u} = (-\sin u \sin v, \cos u \sin v, 0)$$

$$\frac{\partial r}{\partial v} = \left(\cos u \cos v, \sin u \cos v, -\sin v + \frac{1}{\sin v} \right)$$

$$\frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v} = (-\cos u \sin^2 v + \cos u, -\sin u \sin^2 v + \sin u, -\sin v \cos v)$$

$$\begin{aligned} \left\| \frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v} \right\|^2 &= \cos^2 u \sin^4 v - 2 \cos^2 u \sin^2 v + \cos^2 u + \\ &\quad \sin^2 u \sin^4 v - 2 \sin^2 u \sin^2 v + \sin^2 u + \\ &\quad \sin^2 v \cos^2 v \\ &= \sin^4 v - 2 \sin^2 v + 1 + \cos^2 v \sin^2 v \\ &= \sin^4 v - 2 \sin^2 v + 1 + (1 - \sin^2 v) \sin^2 v \\ &= \cos^2 v \end{aligned}$$

$$A(\Sigma) = \int_0^{2\pi} \int_0^\pi |\cos v| dv du = 4\pi$$

Analogamente, data una funzione continua $f : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, con $\Sigma \subset \Omega$, si definisce *integrale di superficie di f su Σ* e si indica con il simbolo $\int_{\Sigma} f dS$, l'integrale

$$\begin{aligned} \int_{\Sigma} f dS &= \int \int_A f(r(u, v)) \|N(u, v)\| dudv \\ &= \int \int_A f(x(u, v), y(u, v), z(u, v)) \|N(u, v)\| dudv. \end{aligned} \quad (45)$$

La formula (45) è indipendente dalla parametrizzazione, infatti se $r_1 : A_1 \rightarrow \mathbb{R}^3$ e $r_2 : A_2 \rightarrow \mathbb{R}^3$ sono due parametrizzazioni (iniettive rispettivamente sull'interno di A_1 e sull'interno di A_2) allora:

$$\int \int_{A_1} f(r_1(u, v)) \|N(u, v)\| dudv = \int \int_{A_2} f(r_2(u, v)) \|N(u, v)\| dudv$$

Nel caso in cui $f(x, y, z) = 1$, allora $\int_{\Sigma} f dS = Area(\Sigma)$.

Integrali di superficie di campi vettoriali

Consideriamo una superficie regolare Σ parametrizzata dalla funzione differenziabile $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$

$$(u, v) \mapsto r(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in A.$$

La condizione di regolarità permette di definire punto per punto un vettore normale alla superficie $N(u, v) = \frac{\partial}{\partial u} r(u, v) \wedge \frac{\partial}{\partial v} r(u, v)$, e un vettore unitario (versore) $\mathbf{n}(u, v) = \frac{N(u, v)}{\|N(u, v)\|}$ ortogonale a Σ .

Di fatto in ogni punto della superficie è possibile scegliere fra due possibili versori normali

$$\mathbf{n}(u, v) = \frac{\frac{\partial}{\partial u} r(u, v) \wedge \frac{\partial}{\partial v} r(u, v)}{\|\frac{\partial}{\partial u} r(u, v) \wedge \frac{\partial}{\partial v} r(u, v)\|} \quad \text{e} \quad \mathbf{n}'(u, v) = -\mathbf{n}(u, v) = \frac{\frac{\partial}{\partial v} r(u, v) \wedge \frac{\partial}{\partial u} r(u, v)}{\|\frac{\partial}{\partial v} r(u, v) \wedge \frac{\partial}{\partial u} r(u, v)\|}$$

Se è possibile scegliere il verso del vettore normale in modo che quest'ultimo vari con continuità sulla superficie Σ , allora Σ viene detta orientabile.

Definizione 41 Una superficie regolare Σ è detta orientabile se è possibile definire un campo vettoriale continuo $\mathbf{n} : \Sigma \rightarrow \mathbb{R}^3$ che sia ortogonale punto per punto alla superficie Σ e tale che $\|n(x, y, z)\| = 1$ per ogni $(x, y, z) \in \Sigma$.

Non tutte le superfici regolari sono orientabili, ad esempio il "nastro di Möbius", di parametrizzazione

$$r(u, v) = ((R+v \cos(u/2)) \cos u, (R+v \cos(u/2)) \sin u, v \sin(u/2)), \quad u \in [0, 2\pi], v \in [-h, h]$$

non è orientabile.



La scelta di un verso di \mathbf{n} , cioè di un'orientazione di Σ determina una distinzione fra le due "facce" di Σ : si indica come positiva la faccia Σ^+ dalla parte in cui punta \mathbf{n} , negativa la faccia opposta Σ^- .

Introduciamo ora la definizione di flusso di un campo vettoriale attraverso una superficie orientata Σ .

Definizione 42 Sia Σ una superficie regolare orientabile e orientata con campo di vettori normali \mathbf{n} . Sia $F : \Omega \subset \mathbb{R}^3$ un campo vettoriale continuo, con $\Sigma \subset \Omega$. Si definisce integrale di superficie (oppure flusso) di F su Σ l'integrale

$$\int \int_{\Sigma} F \cdot \mathbf{n} dS$$

Se $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $r(u, v) = (x(u, v), y(u, v), z(u, v))$ è una parametrizzazione di Σ , e se il vettore normale $N(u, v)$ rispetta l'orientazione di Σ , in modo tale che $\mathbf{n} = \frac{N(u, v)}{\|N(u, v)\|}$, allora

$$\begin{aligned} \int \int_{\Sigma} F \cdot \mathbf{n} dS &= \int \int_A F(x(u, v), y(u, v), z(u, v)) \cdot \frac{N(u, v)}{\|N(u, v)\|} \|N(u, v)\| dudv \\ &= \int \int_A F(x(u, v), y(u, v), z(u, v)) \cdot N(u, v) dudv \end{aligned} \quad (46)$$

Il teorema della divergenza

In questa sezione e in quella successiva introdurremo delle formule che generalizzano il teorema fondamentale del calcolo integrale al caso multidimensionale.

Consideriamo un insieme $\Omega \subset \mathbb{R}^3$, chiusura di un aperto connesso limitato, semplice rispetto ai tre assi⁶ con frontiera $\partial\Omega$ superficie regolare a pezzi e orientabile. Orientiamo $\partial\Omega$ scegliendo il versore normale \mathbf{n} che punta verso l'esterno di Ω . Indichiamo con \mathbf{n}_e tale campo di versori.

Lemma 3 Sia $F = (F_1, F_2, F_3)$ un campo vettoriale di classe $C^1(\Omega)$.

1. Se Ω è x semplice, cioè della forma (47), allora vale la formula

$$\int \int \int_{\Omega} \frac{\partial F_1}{\partial x} dx dy dz = \int \int_{\partial\Omega} F_1 \mathbf{i} \cdot \mathbf{n}_e dS$$

2. Se Ω è y semplice, cioè della forma (48), allora vale la formula

$$\int \int \int_{\Omega} \frac{\partial F_2}{\partial y} dx dy dz = \int \int_{\partial\Omega} F_2 \mathbf{j} \cdot \mathbf{n}_e dS$$

3. Se Ω è z semplice, cioè della forma (49), allora vale la formula

$$\int \int \int_{\Omega} \frac{\partial F_3}{\partial z} dx dy dz = \int \int_{\partial\Omega} F_3 \mathbf{k} \cdot \mathbf{n}_e dS$$

Dimostrazione: Dimostriamo la parte 3, la dimostrazione di 1. e 2. è analoga e viene lasciata per esercizio.

Utilizzando la formula di integrazione "per fili" paralleli all'asse delle z :

$$\begin{aligned} \int \int \int_{\Omega} \frac{\partial F_3}{\partial z} dx dy dz &= \int \int_{D_3} \int_{g_3(y,z)}^{h_3(y,z)} \frac{\partial F_3}{\partial z} dz dx dy \\ &= \int \int_{D_3} (F_3(x, y, h_3(x, y)) - F_3(x, y, g_3(x, y))) dx dy \end{aligned}$$

D'altra parte, per il calcolo del flusso $\int \int_{\partial\Omega} F_3 \mathbf{k} \cdot \mathbf{n}_e dS$, notiamo che la frontiera di Ω può essere rappresentata come unione di 3 superfici:

⁶cioè che sia rappresentabile sia nella forma

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 : g_1(y, z) \leq x \leq h_1(y, z), (y, z) \in D_1\}, \quad g_1, h_1 \in C^1(\bar{D}_1). \quad (47)$$

sia nella forma

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 : g_2(x, z) \leq y \leq h_2(x, z), (x, z) \in D_2\}, \quad g_2, h_2 \in C^1(\bar{D}_2), \quad (48)$$

sia nella forma

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 : g_3(x, y) \leq z \leq h_3(x, y), (x, y) \in D_3\}, \quad g_3, h_3 \in C^1(\bar{D}_3), \quad (49)$$

dove $D_i \subset \mathbb{R}^2$, $i = 1 \dots 3$, insiemi regolari di \mathbb{R}^2 , limitati e semplicemente connessi

- la superficie $\Sigma_1 = \{(x, y, z) \in \mathbb{R}^3 : z = g_3(x, y), (x, y) \in D_3\}$, grafico della funzione $g_3 : D_3 \rightarrow \mathbb{R}$. Tale superficie può essere parametrizzata tramite $r : D_3 \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$:

$$r(x, y) = (x, y, g_3(x, y)), \quad (x, y) \in D_3$$

In questo caso il vettore normale uscente da Ω è $N(x, y) = \left(\frac{\partial}{\partial x} g_3(x, y), \frac{\partial}{\partial y} g_3(x, y), -1 \right)$ e il flusso di $F_3 \mathbf{k}$ uscente da Σ_1 è :

$$\begin{aligned} \int \int_{\Sigma_1} F_3 \mathbf{k} \cdot \mathbf{n}_e dS &= \int \int_{D_3} (0, 0, F_3(x, y, g_3(x, y))) \cdot \left(\frac{\partial}{\partial x} g_3(x, y), \frac{\partial}{\partial y} g_3(x, y), -1 \right) dx dy \\ &= - \int \int_{D_3} F_3(x, y, g_3(x, y)) dx dy \end{aligned}$$

- la superficie $\Sigma_2 = \{(x, y, z) \in \mathbb{R}^3 : z = h_3(x, y), (x, y) \in D_3\}$, grafico della funzione $h_3 : D_3 \rightarrow \mathbb{R}$. Tale superficie può essere parametrizzata tramite $r : D_3 \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$:

$$r(x, y) = (x, y, h_3(x, y)), \quad (x, y) \in D_3$$

In questo caso il vettore normale uscente da Ω è $N(x, y) = \left(-\frac{\partial}{\partial x} h_3(x, y), -\frac{\partial}{\partial y} h_3(x, y), 1 \right)$ e il flusso di $F_3 \mathbf{k}$ uscente da Σ_2 è :

$$\begin{aligned} \int \int_{\Sigma_2} F_3 \mathbf{k} \cdot \mathbf{n}_e dS &= \int \int_{D_3} (0, 0, F_3(x, y, h_3(x, y))) \cdot \left(-\frac{\partial}{\partial x} h_3(x, y), -\frac{\partial}{\partial y} h_3(x, y), +1 \right) dx dy \\ &= \int \int_{D_3} F_3(x, y, h_3(x, y)) dx dy \end{aligned}$$

- La superficie "laterale" $\Sigma_3 = \{(x, y, z) \in \mathbb{R}^3 : g_3(x, y) \leq z \leq h_3(x, y), (x, y) \in \partial D_3\}$. Se $\alpha(t) = (x(t), y(t))$, $t \in I$, è una parametrizzazione della curva ∂D_3 , allora

$$r(t, z) = (x(t), y(t), z), \quad t \in I, g_3(x(t), y(t)) \leq z \leq h_3(x(t), y(t))$$

è una parametrizzazione di Σ_3 . È immediato verificare che il vettore normale a Σ

$$N(t, z) = \frac{\partial}{\partial t} r(t, z) \wedge \frac{\partial}{\partial z} r(t, z) = (x'(t), y'(t), 0) \wedge (0, 0, 1) = (y'(t), -x'(t), 0)$$

è ortogonale al campo $F_3 \mathbf{k}$ e quindi $\int \int_{\Sigma_3} F_3 \mathbf{k} \cdot \mathbf{n}_e dS = 0$

Concludendo

$$\begin{aligned} \int \int_{\partial\Omega} F_3 \mathbf{k} \cdot \mathbf{n}_e dS &= \int \int_{\Sigma_1} F_3 \mathbf{k} \cdot \mathbf{n}_e dS + \int \int_{\Sigma_2} F_3 \mathbf{k} \cdot \mathbf{n}_e dS + \int \int_{\Sigma_3} F_3 \mathbf{k} \cdot \mathbf{n}_e dS \\ &= \int \int_{D_3} F_3(x, y, h_3(x, y)) dx dy - \int \int_{D_3} F_3(x, y, g_3(x, y)) dx dy \end{aligned}$$

□

Diamo ora il risultato principale di questa sezione.

Teorema 32 (della divergenza, o di Gauss) Sia $\Omega \subset \mathbb{R}^3$ un insieme semplice rispetto ai 3 assi, con frontiera $\partial\Omega$ superficie regolare a pezzi e orientabile, orientata scegliendo il versore normale esterno \mathbf{n}_e a $\partial\Omega$. Se $F = (F_1, F_2, F_3)$ è un campo vettoriale di classe $C^1(\Omega)$, allora

$$\int \int \int_{\Omega} \operatorname{div} F dx dy dz = \int \int_{\partial\Omega} F \cdot \mathbf{n}_e dS \quad (50)$$

Dimostrazione: Scrivendo $F = F_1 \mathbf{i} + F_2 \mathbf{j} + F_3 \mathbf{k}$ e utilizzando il risultato del lemma 3, abbiamo

$$\begin{aligned} \int \int_{\partial\Omega} F \cdot \mathbf{n}_e dS &= \int \int_{\partial\Omega} F_1 \mathbf{i} \cdot \mathbf{n}_e dS + \int \int_{\partial\Omega} F_2 \mathbf{j} \cdot \mathbf{n}_e dS + \int \int_{\partial\Omega} F_3 \mathbf{k} \cdot \mathbf{n}_e dS \\ &= \int \int \int_{\Omega} \frac{\partial F_1}{\partial x} dx dy dz + \int \int \int_{\Omega} \frac{\partial F_2}{\partial y} dx dy dz + \int \int \int_{\Omega} \frac{\partial F_3}{\partial z} dx dy dz \\ &= \int \int \int_{\Omega} \operatorname{div} F(x, y, z) dx dy dz \end{aligned}$$

□

Il teorema precedente si generalizza al caso in cui $\Omega \subset \mathbb{R}^3$ è la chiusura di un aperto connesso limitato decomponibile in un numero finito di insiemi (disgiunti) Ω_i , semplici rispetto ai tre assi, con frontiera $\partial\Omega_i$ superficie regolare a pezzi e orientabile. Infatti, se consideriamo per semplicità in caso in cui $\Omega = \Omega_1 \cup \Omega_2$, con Ω_1, Ω_2 semplici rispetto ai tre assi, allora

$$\begin{aligned} \int \int \int_{\Omega} \operatorname{div} F dx dy dz &= \int \int \int_{\Omega_1} \operatorname{div} F dx dy dz + \int \int \int_{\Omega_2} \operatorname{div} F dx dy dz \\ &= \int \int_{\partial\Omega_1} F \cdot \mathbf{n}_e^1 dS + \int \int_{\partial\Omega_2} F \cdot \mathbf{n}_e^2 dS \end{aligned}$$

dove \mathbf{n}_e^1 indica il versore normale a $\partial\Omega_1$ uscente da Ω_1 , mentre \mathbf{n}_e^2 indica il versore normale a $\partial\Omega_2$ uscente da Ω_2 . Se indichiamo con $\Sigma = \partial\Omega_1 \cap \partial\Omega_2$ allora

$$\begin{aligned} & \int \int_{\partial\Omega_1} F \cdot \mathbf{n}_e^1 dS + \int \int_{\partial\Omega_2} F \cdot \mathbf{n}_e^2 dS \\ = & \int \int_{\partial\Omega_1 \setminus \Sigma} F \cdot \mathbf{n}_e^1 dS + \int \int_{\Sigma} F \cdot \mathbf{n}_e^1 dS + \int \int_{\partial\Omega_2 \setminus \Sigma} F \cdot \mathbf{n}_e^2 dS + \int \int_{\Sigma} F \cdot \mathbf{n}_e^2 dS \end{aligned}$$

Notando che sulla superficie Σ abbiamo $\mathbf{n}_e^2 = -\mathbf{n}_e^1$, abbiamo che il secondo ed il quarto addendo si semplificano e quindi

$$\begin{aligned} \int \int \int_{\Omega} \operatorname{div} F dx dy dz &= \int \int_{\partial\Omega_1 \setminus \Sigma} F \cdot \mathbf{n}_e^1 dS + \int \int_{\partial\Omega_2 \setminus \Sigma} F \cdot \mathbf{n}_e^2 dS \\ &= \int \int_{\partial\Omega} F \cdot \mathbf{n}_e^2 dS \end{aligned}$$

Il teorema della divergenza permette di fornire un'interpretazione della divergenza di un campo vettoriale. Sia infatti F campo vettoriale di classe C^1 e fissiamo un punto $(x_0, y_0, z_0) \in \mathbb{R}^3$. Consideriamo la sfera centrata in (x_0, y_0, z_0) e di raggio $\epsilon > 0$:

$$B_\epsilon(x_0, y_0, z_0) = \{(x, y, z) \in \mathbb{R}^3 : (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 \leq \epsilon\}$$

Allora per il teorema della divergenza

$$\int \int \int_{B_\epsilon(x_0, y_0, z_0)} \operatorname{div} F(x, y, z) dx dy dz = \int \int_{\partial B_\epsilon(x_0, y_0, z_0)} F \cdot \mathbf{n}_e dS$$

Se dividiamo entrambi i membri per il volume di $B_\epsilon(x_0, y_0, z_0)$ e facciamo tendere ϵ a 0, per il teorema della media integrale⁷ abbiamo che

$$\begin{aligned} \operatorname{div} F(x_0, y_0, z_0) &= \lim_{\epsilon \rightarrow 0} \frac{1}{\operatorname{Vol}(B_\epsilon(x_0, y_0, z_0))} \int \int \int_{B_\epsilon(x_0, y_0, z_0)} \operatorname{div} F(x, y, z) dx dy dz \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\operatorname{Vol}(B_\epsilon(x_0, y_0, z_0))} \int \int_{\partial B_\epsilon(x_0, y_0, z_0)} F \cdot \mathbf{n}_e dS \end{aligned}$$

Tale uguaglianza permette di interpretare la divergenza di un campo vettoriale in un punto come la densità di flusso del campo uscente da quel punto per unità di volume.

⁷Se Ω è connesso e $f : \Omega \rightarrow \mathbb{R}$ continua e limitata, allora esiste un punto $(\bar{x}, \bar{y}, \bar{z}) \in \Omega$ tale che $\int \int \int_{\Omega} f(x, y, z) dx dy dz = f(\bar{x}, \bar{y}, \bar{z}) \operatorname{Vol}(\Omega)$

Esempio 39 Consideriamo il campo vettoriale $F(x, y, z) = \frac{k}{(x^2+y^2+z^2)^{3/2}} (x, y, z)$, di classe C^1 su $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$. Abbiamo

$$\operatorname{div} F(x, y, z) = k \left(\frac{(x^2 + y^2 + z^2) - 3x^2}{(x^2 + y^2 + z^2)^{3/2}} + \frac{(x^2 + y^2 + z^2) - 3y^2}{(x^2 + y^2 + z^2)^{3/2}} + \frac{(x^2 + y^2 + z^2) - 3z^2}{(x^2 + y^2 + z^2)^{3/2}} \right) = 0$$

Sia $\Sigma \subset \mathbb{R}^3$ una superficie chiusa regolare a tratti, orientabile ed orientata, frontiera di un insieme $\Omega \subset \mathbb{R}^3$ che verifichi le ipotesi del teorema della divergenza. Se il punto $(0, 0, 0)$ non appartiene all'insieme Ω allora, applicando la formula (50) abbiamo $\int \int_{\Sigma} F \cdot \mathbf{n}_e dS = 0$, mentre se $(0, 0, 0) \in \Omega$, applicando la formula (50) all'insieme $\Omega' = \Omega \setminus B_{\epsilon}(0, 0, 0)$, abbiamo:

$$0 = \int \int_{\partial\Omega'} F \cdot \mathbf{n}_e dS = \int \int_{\Sigma} F \cdot \mathbf{n}_e dS - \int \int_{\partial B_{\epsilon}(0,0,0)} F \cdot \mathbf{n}_e dS$$

Utilizzando ad esempio la parametrizzazione per $\partial B_{\epsilon}(0, 0, 0)$

$$r(\phi, \theta) = (\epsilon \sin \phi \cos \theta, \epsilon \sin \phi \sin \theta, \epsilon \cos \phi), \quad \phi \in [0, \pi], \theta \in [0, 2\pi]$$

abbiamo $\int \int_{\partial B_{\epsilon}(0,0,0)} F \cdot \mathbf{n}_e dS = 4\pi$. Otteniamo dunque che $\int \int_{\Sigma} F \cdot \mathbf{n}_e dS = k4\pi$.

Il teorema del rotore

Definizione 43 Data una superficie $\Sigma \subset \mathbb{R}^3$, un punto $(x, y, z) \in \Sigma$ è detto interno a Σ se esiste un intorno sferico $B_{\epsilon}(x, y, z)$ di (x, y, z) e una parametrizzazione $r : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ di $\Sigma \cap B_{\epsilon}(x, y, z)$ tale che $(x, y, z) = r(u, v)$, con (u, v) punto interno ad A .

L'insieme dei punti $(x, y, z) \in \Sigma$ che non sono interni viene detto bordo di Σ e viene indicato col simbolo $\partial\Sigma$.

Esempio 40 La sfera $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = R^2\}$ non ha bordo, in quanto tutti i suoi punti sono interni. Infatti dato un punto (x, y, z) appartenente alla superficie sferica Σ è possibile mostrare che è interno utilizzando il teorema 19 parametrizzando $\Sigma \cap B_{\epsilon}(x, y, z)$ come grafico di una funzione $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. Di fatto, applicando il teorema del Dini nel caso 3-dimensionale (vedi teorema 19) è possibile dimostrare che ogni superficie $\Sigma \subset \mathbb{R}^3$ della forma $\Sigma := \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 0\}$, dove $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ è una funzione di classe C^1 e tale che $\nabla g(x, y, z) \neq (0, 0, 0)$ per ogni $(x, y, z) \in \Sigma$, è una superficie regolare senza bordo (tutti i suoi punti sono interni).

Esempio 41 Una superficie Σ grafico della funzione $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, con $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ e $f(x, y) = x^2 + y^2$ è la circonferenza $\gamma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, z = 1\}$. Più in generale, se consideriamo una superficie cartesiana Σ di parametrizzazione $r(x, y) = (x, y, f(x, y))$, con $(x, y) \in D$, abbiamo che il bordo della superficie Σ è la curva $\gamma \subset \mathbb{R}^3$ descritta da $\gamma = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in \partial D, z = f(x, y)\}$.

Sia $\Sigma \subset \mathbb{R}^3$ una superficie regolare, parametrizzata da

$$r(u, v) = (x(u, v), y(u, v), z(u, v)); \quad (u, v) \in A \subset \mathbb{R}^2$$

dove $A \subset \mathbb{R}^2$ è un insieme s-decomponibile e r è di classe $C^1(A)$. Supponiamo che $r : \bar{A} \rightarrow \mathbb{R}^3$ sia orientabile⁸ e orientata, con orientazione fissata da un campo continuo di versori normali \mathbf{n} , dove le due orientazioni possibili sono:

$$\mathbf{n}^+ = \frac{\frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v}}{\left\| \frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v} \right\|}$$

$$\mathbf{n}^- = -\frac{\frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v}}{\left\| \frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v} \right\|}$$

Ricordiamo che la scelta di un'orientazione permette la distinzione fra le due "facce" di Σ . Indichiamo con Σ^+ la faccia che si trova dalla parte in cui punta il versore \mathbf{n} . Supponiamo che il bordo $\partial\Sigma$ sia una curva regolare a tratti, e indichiamo con $\partial\Sigma^+$ la curva $\partial\Sigma$ orientata in modo tale che un osservatore in piedi sulla faccia Σ^+ percorra $\partial\Sigma$ tenendo Σ alla propria sinistra. Se si cambia orientazione di Σ , l'orientazione di $\partial\Sigma$ cambia corrispondentemente.

Teorema 33 (del rotore, o di Stokes) Sia $F : \Omega \rightarrow \mathbb{R}^3$ campo vettoriale di classe $C^1(\Omega)$ e $\Sigma \subset \Omega$ superficie regolare orientabile e orientata con versore normale \mathbf{n} , con bordo $\partial\Sigma^+$ curva regolare a tratti orientata (coerentemente con l'orientazione di Σ). Allora

$$\oint_{\partial\Sigma^+} F \cdot dr = \int \int_{\Sigma} \text{rot } F \cdot \mathbf{n} dS$$

Quindi il lavoro di un campo vettoriale F lungo il bordo di una superficie Σ è uguale al flusso del rotore di F attraverso Σ .

Notiamo che la formula di Gauss-Green nel piano (teorema 31) è un caso particolare del teorema di Stokes, applicato a superfici appartenenti al piano xy .

⁸si può dimostrare che se r è iniettiva su \bar{A} e di classe C^1 allora Σ è orientabile

Il teorema di Stokes inoltre permette di dare un'interpretazione del rotore di un campo vettoriale. Sia infatti F campo vettoriale di classe C^1 e fissiamo un punto $(x_0, y_0, z_0) \in \Omega$. Sia $D_\epsilon(x_0, y_0, z_0)$ un cerchio di centro (x_0, y_0, z_0) , raggio ϵ che giace sul piano passante per (x_0, y_0, z_0) e versore normale \mathbf{n} . Se applichiamo a tale insieme il teorema del rotore e dividiamo per l'area del disco otteniamo

$$\frac{1}{\pi\epsilon^2} \int \int_{D_\epsilon(x_0, y_0, z_0)} \text{rot } F \cdot \mathbf{n} \, dS = \frac{1}{\pi\epsilon^2} \oint_{\partial^+ D_\epsilon(x_0, y_0, z_0)} F \cdot dr$$

Se $\epsilon \rightarrow 0$, il primo membro tende, per il teorema della media integrale, a $\text{rot } F(x_0, y_0, z_0) \cdot \mathbf{n}$. Abbiamo quindi:

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\pi\epsilon^2} \oint_{\partial^+ D_\epsilon(x_0, y_0, z_0)} F \cdot dr = \text{rot } F(x_0, y_0, z_0) \cdot \mathbf{n}$$

Possiamo infine utilizzare il teorema del rotore per mostrare che in insiemi semplicemente connessi l'annullarsi del rotore di F equivale alla conservatività di F , cioè all'esistenza di un potenziale U tale che $F = \nabla U$.

Notiamo inoltre che se $F \in C^1(\Omega)$, $\text{rot } F = (0, 0, 0)$ e Ω non è semplicemente connesso, allora in ogni sottoinsieme $\Omega' \subset \Omega$ semplicemente connesso, esiste un potenziale $U_{\Omega'}$ locale per F , cioè una funzione $U_{\Omega'} \in C^2(\Omega')$ tale che $F(x, y, z) = \nabla U_{\Omega'}(x, y, z)$ per ogni $(x, y, z) \in \Omega'$. È possibile costruire un potenziale globale $U \in C^2(\Omega)$ se per ogni coppia di sottoinsiemi $\Omega'_1, \Omega'_2 \subset \Omega$ semplicemente connessi, con $\Omega'_1 \cap \Omega'_2 \neq \emptyset$, è possibile definire i rispettivi potenziali $U_{\Omega'_1}$ e $U_{\Omega'_2}$ in modo tale che coincidano sul dominio comune $\Omega'_1 \cap \Omega'_2$. Se Ω non è semplicemente connesso questa operazione di "incollamento" non è sempre possibile.

Serie di funzioni

Sia $A \subset \mathbb{R}$ e consideriamo per ogni $n \in \mathbb{N}$ una funzione $f_n : A \rightarrow \mathbb{R}$. Otteniamo una successione $\{f_n\}$ di funzioni definite tutte sull'insieme A .

Analogamente, per ogni $x \in A$, consideriamo la serie di numeri reali $\sum_n f_n(x)$. Se tale serie è convergente per ogni $x \in A$, si dice che la serie di funzioni $\sum_n f_n$ *converge puntualmente* sull'insieme A . In questo modo costruiamo, punto per punto nell'insieme A , una funzione $S : A \rightarrow \mathbb{R}$, definita da $S(x) := \lim_{N \rightarrow \infty} \sum_{n=0}^N f_n(x)$ e detta *somma* della serie di funzioni.

Esempio 42 $A = (-1, 1)$, $f_n(x) = x^n$. La serie $\sum_n f_n(x) = \sum_{n=0}^{\infty} x^n$ è la *serie geometrica di ragione x* , con $|x| < 1$, che converge puntualmente alla funzione $S(x) = \frac{1}{1-x}$.

Ci occupiamo prima di tutto di studiare le proprietà delle funzioni della forma $\sum_n f_n$, in particolare la continuità, la derivabilità e l'integrabilità. Sappiamo, dalle proprietà dei limiti, che la somma di un numero finito di funzioni continue è una funzione continua e la somma di un numero finito di funzioni derivabili è una funzione derivabile. Vedremo che questo non è sempre vero quando consideriamo un numero infinito di addendi, cioè una serie di funzioni. Per garantire queste proprietà è necessario introdurre una nozione di convergenza più forte.

Definizione 44 La serie di funzioni $\sum_{n=0}^{\infty} f_n$ è detta *totalmente convergente su A* se:

1. per ogni $n \in \mathbb{N}$ esiste una costante $c_n \in \mathbb{R}$ tale che $|f_n(x)| \leq c_n$ per ogni $x \in A$;
2. la serie $\sum_{n=0}^{\infty} c_n$ è convergente.

È immediato verificare che se la serie $\sum_{n=0}^{\infty} f_n$ è totalmente convergente su A allora è anche semplicemente convergente su A , infatti, per ogni $x \in A$,

$$\sum_n |f_n(x)| \leq \sum_n c_n, \infty.$$

La convergenza totale comunque è una proprietà più forte di quella semplice. Ad esempio la serie dell'esempio 42 converge semplicemente ma non totalmente sull'intervallo $(-1, 1)$. Vale il seguente risultato.

Teorema 34 Se la serie di funzioni $\sum_{n=0}^{\infty} f_n$ è totalmente convergente su A allora la sua somma $S = \sum_n f_n$ è una funzione continua su A .

Ad esempio la serie $\sum_n \frac{\sin(3^n x)}{2^n}$ è totalmente convergente infatti $\left| \frac{\sin(3^n x)}{2^n} \right| \leq \frac{1}{2^n}$ e $\sum_n 1/2^n < \infty$.

La serie definisce una funzione continua $S(x) = \sum_n \frac{\sin(3^n x)}{2^n}$.

Notiamo che, anche se le funzioni $f_n(x) = \frac{\sin(3^n x)}{2^n}$ sono derivabili non è detto che la funzione $S(x) = \sum_n f_n(x)$ sia derivabile e che valga la formula $S'(x) = \sum_n f'_n(x)$. Ad esempio, nel caso appena studiato in cui $f_n(x) = \frac{\sin(3^n x)}{2^n}$, abbiamo $f'_n(x) = \frac{3^n \cos(3^n x)}{2^n}$, e la serie $\sum_n \frac{3^n \cos(3^n x)}{2^n}$ non è convergente, infatti in $x = 0$ abbiamo $\sum_n \frac{3^n}{2^n} = +\infty$. Il seguente teorema fornisce delle condizioni per la validità della formula $(\sum_{n=0}^{\infty} f_n)' = \sum_{n=0}^{\infty} f'_n$.

Teorema 35 Sia $I \subset \mathbb{R}$ intervallo e $f_n : I \rightarrow \mathbb{R}$ derivabili e tali che

- $\sum_{n=0}^{\infty} f_n$ converge puntualmente su I ,
- $\sum_{n=0}^{\infty} f'_n$ converge totalmente su I .

Allora la funzione $s(x) = \sum_{n=0}^{\infty} f_n(x)$ è derivabile e $s'(x) = \sum_{n=0}^{\infty} f'_n(x)$

Vale un analogo teorema per gli integrali.

Teorema 36 Sia $I = [a, b] \subset \mathbb{R}$ intervallo $f_n : I \rightarrow \mathbb{R}$ continue e tali che $\sum_{n=0}^{\infty} f_n(x)$ totalmente convergente su $[a, b]$ ad una funzione $s(x) = \sum_{n=0}^{\infty} f_n(x)$. Allora f è integrabile su $[a, b]$ e

$$\int_a^b s(x) dx = \int_a^b \sum_{n=0}^{\infty} f_n(x) dx = \sum_{n=0}^{\infty} \int_a^b f_n(x) dx$$

Serie di potenze

Fissato un punto $x_0 \in \mathbb{R}$ e una successione $\{a_n\} \subset \mathbb{R}$, consideriamo la serie di funzioni $\sum_{n=0}^{\infty} a_n(x - x_0)^n$. Vale il seguente risultato

Teorema 37 se la serie $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converge in un punto \bar{x} , allora converge in ogni x tale che $|x - x_0| < |\bar{x} - x_0|$. Inoltre per ogni $\delta < |\bar{x} - x_0|$, la serie $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converge totalmente in $[x_0 - \delta, x_0 + \delta]$.

Dimostrazione: Dato che la serie $\sum_{n=0}^{\infty} a_n(\bar{x} - x_0)^n$ è convergente, allora la successione $a_n(\bar{x} - x_0)^n$ è limitata e $|a_n(\bar{x} - x_0)^n| < K$ per ogni $n \in \mathbb{N}$. Scegliendo $x \in \mathbb{R}$ tale che $|x - x_0| < |\bar{x} - x_0|$ abbiamo quindi:

$$\sum_{n=0}^{\infty} |a_n(x - x_0)^n| = \sum_{n=0}^{\infty} |a_n(\bar{x} - x_0)^n| \left(\frac{|x - x_0|}{|\bar{x} - x_0|} \right)^n \leq K \sum_{n=0}^{\infty} \left(\frac{|x - x_0|}{|\bar{x} - x_0|} \right)^n \leq \infty$$

Inoltre se $x \in [x_0 - \delta, x_0 + \delta]$, con $\delta < |\bar{x} - x_0|$, abbiamo:

$$|a_n(x - x_0)^n| \leq K \left(\frac{\delta}{|\bar{x} - x_0|} \right)^n$$

$$\text{e } \sum_n K \left(\frac{\delta}{|\bar{x} - x_0|} \right)^n < \infty.$$

□

Definizione 45 Si definisce raggio di convergenza della serie di potenze $\sum_n a_n x^n$ l'estremo superiore⁹ dell'insieme $\{|x| : \sum_n a_n |x|^n \text{ è convergente}\}$:

$$R := \sup\{|x| : \sum_n a_n |x|^n \text{ è convergente}\}$$

Se $\sum_n a_n x^n$ è convergente per ogni $x \in \mathbb{R}$ allora $R = +\infty$

Ad esempio il raggio di convergenza della serie $\sum_n \frac{x^n}{n!}$ è $R = +\infty$, mentre il raggio di convergenza della serie $\sum_n x^n$ è $R = 1$.

Dal teorema 37 abbiamo il seguente risultato.

Teorema 38 Sia $\sum_n a_n (x - x_0)^n$ una serie di potenze con raggio di convergenza R . Allora

1. Se $|x - x_0| < R$ la serie $\sum_n a_n (x - x_0)^n$ è convergente,
2. Se $|x - x_0| > R$ la serie $\sum_n a_n (x - x_0)^n$ non è convergente,
3. per ogni $\delta < R$, la serie $\sum_n a_n (x - x_0)^n$ è totalmente convergente in $[x_0 - \delta, x_0 + \delta]$. Su tale insieme la somma $s(x) = \sum_n a_n (x - x_0)^n$ è una funzione continua.

Il teorema precedente afferma che la serie di potenze $\sum_n a_n (x - x_0)^n$ converge in un'intervallo $(x_0 - R, x_0 + R)$ centrato in x_0 e di ampiezza pari al raggio di convergenza, ma non dice nulla sul comportamento sul bordo dell'intervallo $(x_0 - R, x_0 + R)$, cioè nei punti $x = x_0 + R$ e $x = x_0 - R$. Non c'è una regola che valga in tutti i

⁹Sia $A \subset \mathbb{R}$. Un numero reale $x \in \mathbb{R}$ è detto *maggiorante di A* se per ogni $y \in A$ si ha $y \leq x$. Se l'insieme dei maggioranti di A è vuoto (cioè se A è un insieme illimitato superiormente) allora si definisce $\sup A = +\infty$. Se l'insieme dei maggioranti di A non è vuoto, allora definiamo estremo superiore di A il più piccolo dei suoi valori, precisamente $x \in \mathbb{R}$ è detto estremo superiore di A se:

1. $x \geq y \forall y \in A$,
2. Se $z \in \mathbb{R}$ è un maggiorante per A allora $x \leq z$.

casi. Ad esempio, la serie $\sum_{n=1}^{\infty} \frac{x^n}{n}$ ha raggio di convergenza $R = 1$. In $x = 1$ la serie diviene $\sum_{n=1}^{\infty} \frac{1}{n}$ che non è convergente, mentre in $x = -1$ la serie diviene $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ che è convergente. La serie $\sum_{n=1}^{\infty} \frac{x^n}{n^2}$ ha ancora raggio di convergenza $R = 1$, in $x = 1$ diviene $\sum_{n=1}^{\infty} \frac{1}{n^2}$ che è convergente, analogamente in $x = -1$ diviene $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$ che è ancora convergente.

La descrizione del comportamento di una serie di potenze sul bordo dell'intervallo di convergenza viene data dal seguente teorema.

Teorema 39 (di Abel) *Sia $\sum_n a_n(x - x_0)^n$ una serie di potenze con raggio di convergenza $0 < R < +\infty$. Se la serie numerica $\sum_n a_n R^n$ è convergente allora*

$$\sum_n a_n R^n = \lim_{x \rightarrow x_0 + R^-} \sum_n a_n (x - x_0)^n,$$

mentre se la serie numerica $\sum_n a_n (-R)^n$ è convergente allora

$$\sum_n a_n (-R)^n = \lim_{x \rightarrow x_0 - R^+} \sum_n a_n (x - x_0)^n.$$

Diamo ora dei metodi per il calcolo del raggio di convergenza.

Teorema 40 *Se esiste $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = l$ allora $R = 1/l$ (con la convenzione che $1/0 = +\infty$ e $1/\infty = 0$)*

Dimostrazione: Per il criterio della radice per la convergenza assoluta delle serie numeriche, applicato alla serie $\sum_n |a_n(x - x_0)^n|$, abbiamo che se esiste

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n(x - x_0)^n|} = |x - x_0| \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1,$$

quindi se $|x - x_0| < \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}}$, allora la serie è convergente, mentre se esiste

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n(x - x_0)^n|} = |x - x_0| \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1,$$

quindi se $|x - x_0| > \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}}$, allora la serie non è convergente. \square

Ad esempio, considerata la serie di potenze $\sum_n \frac{x^n}{2^n}$, con $a_n = 1/2^n$, abbiamo che $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 1/2$ e quindi $R = 2$.

Teorema 41 *Se esiste $\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} = l$ allora $R = 1/l$ (con la convenzione che $1/0 = +\infty$ e $1/\infty = 0$)*

Dimostrazione: Per il criterio del rapporto per la convergenza assoluta delle serie numeriche, applicato alla serie $\sum_n |a_n(x - x_0)^n|$, abbiamo che se esiste

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}| |x - x_0|^{n+1}}{|a_n| |x - x_0|^n} = |x - x_0| \lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} < 1,$$

quindi se $|x - x_0| < \frac{1}{\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|}}$, allora la serie è convergente, mentre se esiste

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}| |x - x_0|^{n+1}}{|a_n| |x - x_0|^n} = |x - x_0| \lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} > 1,$$

quindi se $|x - x_0| > \frac{1}{\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|}}$, allora la serie non è convergente. □

Ad esempio, considerata la serie di potenze $\sum_n \frac{x^n}{n!}$, con $a_n = 1/n!$, abbiamo che

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!} = \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0$$

e quindi $R = +\infty$ (la serie converge per ogni $x \in \mathbb{R}$).

Se invece consideriamo la serie di potenze $\sum_n n!x^n$, con $a_n = n!$, abbiamo

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n \rightarrow \infty} \frac{(n+1)!}{n!} = \lim_{n \rightarrow \infty} n+1 = +\infty$$

e quindi $R = 0$ (la serie converge solo in $x = 0$).

Analizziamo ora il problema della derivabilità della funzione definita, all'interno dell'intervallo $(x_0 - R, x_0 + R)$ dalla serie $\sum_n a_n(x - x_0)^n$.

Proposizione 6 (Proprietà delle serie di potenze) *Sia $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ serie di potenze con raggio di convergenza $R > 0$. Allora*

1. (Continuità) . *Per ogni $\delta < R$ la serie è totalmente convergente in $[x_0 - \delta, x_0 + \delta]$ e definisce una funzione continua $s(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ nell'intervallo $(x_0 - R, x_0 + R)$.*
2. (Derivabilità) . *La funzione $s(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ è derivabile nell'intervallo $(x_0 - R, x_0 + R)$ e la sua derivata è :*

$$s'(x) = \frac{d}{dx} \sum_{n=0}^{\infty} a_n(x - x_0)^n = \sum_{n=0}^{\infty} \frac{d}{dx} a_n(x - x_0)^n = \sum_{n=1}^{\infty} n a_n(x - x_0)^{n-1}$$

Inoltre la serie $\sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}$ ha lo stesso raggio di convergenza della serie $\sum_{n=0}^{\infty} a_n(x-x_0)^n$.

Tale procedimento può essere ripetuto un numero arbitrario di volte, mostrando che per ogni $k \in \mathbb{N}$, la funzione $s(x)$ è k volte derivabile e

$$s^{(k)}(x) = \frac{d^k}{dx^k} \sum_{n=0}^{\infty} a_n(x-x_0)^n = \sum_{n=0}^{\infty} \frac{d^k}{dx^k} a_n(x-x_0)^n = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n(x-x_0)^{n-k}$$

Inoltre la serie $\sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n(x-x_0)^{n-k}$ ha lo stesso raggio di convergenza della serie $\sum_{n=0}^{\infty} a_n(x-x_0)^n$.

3. La funzione $s(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ ammette primitiva nell'intervallo $(x_0 - R, x_0 + R)$ e può essere calcolata termine a termine tramite la formula

$$\int s(x)dx = \int \sum_{n=0}^{\infty} a_n(x-x_0)^n dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x-x_0)^{n+1} + c$$

Inoltre in ogni intervallo $[a, b] \subset (x_0 - R, x_0 + R)$ la funzione $s(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ è integrabile in $[a, b]$ e

$$\int_a^b s(x)dx = \int_a^b \sum_{n=0}^{\infty} a_n(x-x_0)^n dx = \sum_{n=0}^{\infty} a_n \int_a^b (x-x_0)^n dx$$

Tali proprietà possono essere applicate al calcolo dello sviluppo in serie di funzioni che sono primitive o derivate di funzioni di cui si conosce lo sviluppo in serie, nonché al calcolo della somma di alcune serie numeriche notevoli. Consideriamo ad esempio la serie $\sum_{n=0}^{\infty} (-x)^n$, che di fatto è la serie geometrica di ragione $(-x)$, convergente se $|x| < 1$ e non convergente se $|x| > 1$). Il raggio di convergenza è quindi $R = 1$, e per tali valori di x la somma è

$$\sum_{n=0}^{\infty} (-x)^n = \frac{1}{1+x}$$

Calcolando la primitiva di entrambi i membri, otteniamo

$$\log(1+x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{n+1}}{n+1}$$

In $x = 1$, cioè sul bordo dell'intervallo di convergenza, abbiamo che $\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}$ è convergente, quindi per il teorema di Abel abbiamo che

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} = \lim_{x \rightarrow 1^-} \log(x+1) = \log(2)$$

Analogamente, se consideriamo la serie $\sum_{n=0}^{\infty} (-x^2)^n$, cioè la serie geometrica di ragione $(-x^2)$ il cui raggio di convergenza è $R = 1$, e la cui somma è :

$$\sum_{n=0}^{\infty} (-x^2)^n = \sum_{n=0}^{\infty} (-1)^n x^{2n} = \frac{1}{1+x^2}$$

Calcolando la primitiva di entrambi i membri, otteniamo

$$\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}$$

In $x = 1$, cioè sul bordo dell'intervallo di convergenza, abbiamo che $\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$ è convergente, quindi per il teorema di Abel abbiamo che

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \lim_{x \rightarrow 1^-} \arctan x = \pi/4$$

Occupiamoci infine della relazione tra le serie di Taylor e le serie di potenze. Per la proposizione 6 sappiamo che nell'intervallo $(x_0 - R, x_0 + R)$ una serie di potenze centrata in x_0 definisce una funzione infinite volte derivabile. Sappiamo anche dal corso di analisi 1 che una funzione f derivabile infinite volte in $x = x_0$ ammette serie di Taylor in x_0 , con

$$f(x) = \sum_{n=0}^N \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + R_N(x) \quad (51)$$

dove $R_N(x) = f(x) - \sum_{n=0}^N \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$ soddisfa

$$\lim_{x \rightarrow x_0} \frac{R_N(x)}{|x - x_0|^N} = 0$$

Inoltre $R_N(x)$ può essere stimato con la formula

$$R_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!} (x - x_0)^{N+1}$$

dove $c \in (x_0, x)$. Di fatto la formula di Taylor (51) fornisce un'approssimazione di $f(x)$, per N fissato, tanto più precisa quanto più $|x - x_0|$ è "piccolo".

Se invece una funzione f è somma di una serie di potenze di centro x_0 e raggio di convergenza $R > 0$, allora fissato $x \in (x_0 - R, x_0 + R)$ si ha che la serie $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ è convergente e la sua somma ha il valore $f(x)$, quindi

$$f(x) = \sum_{n=0}^N a_n(x - x_0)^n + R_N(x),$$

dove $R_N(x) = f(x) - \sum_{n=0}^N a_n(x - x_0)^n$ soddisfa

$$\lim_{N \rightarrow \infty} R_N(x) = 0$$

In questo caso, fissato $x \in (x_0 - R, x_0 + R)$, la somma parziale $\sum_{n=0}^N a_n(x - x_0)^n$ fornisce un'approssimazione di $f(x)$ tanto migliore quanto più grande è il numero N degli addendi della somma parziale.

Definizione 46 Una funzione $f : (a, b) \rightarrow \mathbb{R}$ è detta analitica in (a, b) se per ogni $x_0 \in (a, b)$ la funzione è sviluppabile in serie di potenze con centro x_0 con raggio di convergenza $R > 0$.

Ad esempio sono analitiche su tutto \mathbb{R} i polinomi, le funzioni e^x , $\sin x$, $\cos x$, $\sinh x$ e $\cosh x$, mentre è analitica in $(-1, +\infty)$ la funzione $\frac{1}{1+x}$.

Osserviamo che se $f = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ è somma di una serie di potenze di centro x_0 e raggio di convergenza $R > 0$, la sua serie di Taylor in x_0 coincide con la serie $\sum_{n=0}^{\infty} a_n(x - x_0)^n$, cioè $a_n = \frac{f^{(n)}(x_0)}{n!}$.

È importante tenere presente che non tutte le funzioni derivabili infinite volte in un punto x_0 , che ammettono quindi polinomio di Taylor di ogni ordine in x_0 , sono analitiche in un intorno di quel punto. Si possono infatti presentare due tipi di problemi

1. La serie di Taylor $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$ può non avere raggio di convergenza positivo (cioè converge solo per $x = x_0$).

Ad esempio la funzione $f(x) = \int_0^{\infty} \frac{e^{-t}}{1+xt} dt$ ha derivata $f^{(k)}(0) = (-1)^k (k!)^2$ e la sua serie di Taylor in $x = 0$ è $\sum_{n=0}^{\infty} (-1)^n n! x^n$, che ha raggio di convergenza $R = 0$.

2. Anche se la serie di Taylor $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$ è convergente, la sua somma in x può non dare il valore della funzione $f(x)$.

Ad esempio la funzione f definita da $f(0) = 0$ se $x = 0$ e $f(x) = e^{-1/x^2}$ per $x \neq 0$ è derivabile infinite volte in 0 con $f^{(k)}(0) = 0$ per ogni $k \in \mathbb{N}$. La serie di Taylor di f in 0 è $\sum_n \frac{f^{(n)}(0)}{n!} x^n = 0$ ha come somma la funzione costante pari a 0.

Applicazione alla soluzione di equazioni differenziali

Consideriamo un'equazione differenziale lineare del secondo ordine a coefficienti non costanti della forma

$$u''(x) + p(x)u'(x) + q(x)u(x) = 0,$$

e cerchiamo una soluzione analitica della forma $u(x) = \sum_n a_n x^n$, o più in generale, della forma $u(x) = \sum_n a_n x^{n+r}$, con $r \in \mathbb{R}$ una costante reale. Tale tecnica viene chiamata *metodo di Frobenius*. Non tratteremo il caso più generale ma vedremo un esempio particolare per mettere in luce una possibile applicazione delle serie di potenze.

Consideriamo l'equazione $u''(x) + \frac{u'(x)}{x} + u(x) = 0$, e cerchiamo una soluzione della forma $u(x) = \sum_{n=0} a_n x^n$. Dato che $u'(x) = \sum_{n=1} n a_n x^{n-1}$ e $u''(x) = \sum_{n=2} n(n-1) a_n x^{n-2}$, sostituendo nell'equazione differenziale otteniamo:

$$\sum_{n=2} n(n-1) a_n x^{n-2} + \sum_{n=1} n a_n x^{n-2} + \sum_{n=0} a_n x^n = 0$$

Moltiplicando per x^2 e riorganizzando le somme, otteniamo

$$\begin{aligned} a_1 x + \sum_{n=2} (n(n-1) + n) a_n x^n + \sum_{n=0} a_n x^{n+2} &= 0 \\ a_1 x + \sum_{n=2} (n^2 a_n + a_{n-2}) x^n &= 0 \end{aligned}$$

Affinché tale serie di potenze sia la funzione identicamente nulla, è necessario che tutti i coefficienti siano 0, da cui otteniamo:

$$a_1 = 0, \quad n^2 a_n + a_{n-2} = 0 \text{ per } n \geq 2,$$

che possono essere riguardate come delle relazioni che devono essere soddisfatte dai coefficiente a_n della soluzione:

$$a_1 = 0, \quad a_n = -\frac{a_{n-2}}{n^2} \text{ per } n \geq 2.$$

Possiamo quindi dedurre che tutti i coefficienti a_n con indice n dispari sono nulli, mentre i coefficienti con indici pari saranno della forma:

$$a_2 = -\frac{a_0}{2^2}, \quad a_4 = -\frac{a_2}{4^2} = \frac{a_0}{2^2 4^2}, \quad \dots \quad a_{2k} = (-1)^k \frac{a_0}{2^{2k} (k!)^2}.$$

La soluzione è dunque un multiplo (con costante moltiplicativa a_0 , della funzione

$$f(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2^{2k} (k!)^2}$$

Serie di Fourier

Funzioni periodiche

Definizione 47 Una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ è detta periodica di periodo T , con $T > 0$, se per ogni $x \in \mathbb{R}$:

$$f(x + T) = f(x).$$

Notiamo che le funzioni periodiche godono delle seguenti proprietà :

1. se f è periodica di periodo T , allora f è periodica di periodo nT per ogni $n \in \mathbb{N}$, con $n \geq 1$.
2. se f è periodica di periodo T , allora la funzione $f_n : \mathbb{R} \rightarrow \mathbb{R}$ definita da $f_n(x) := f(nx)$ è periodica di periodo T .
3. se f, g sono periodiche di periodo T , allora per ogni $\alpha, \beta \in \mathbb{R}$ la funzione $\alpha f + \beta g$ è periodica di periodo T .
4. se f è periodica di periodo T , allora per ogni $a \in \mathbb{R}$

$$\int_a^{a+T} f(x) dx = \int_0^T f(x) dx$$

Infatti

$$\int_a^{a+T} f(x) dx = \int_a^0 f(x) dx + \int_0^T f(x) dx + \int_T^{a+T} f(x) dx$$

Per la formula di cambiamento di variabili e la periodicità di f abbiamo

$$\int_T^{a+T} f(x) dx = \int_0^a f(x+T) dx = \int_0^a f(x) dx = - \int_a^0 f(x) dx$$

e dunque

$$\int_a^{a+T} f(x) dx = \int_a^0 f(x) dx + \int_0^T f(x) dx - \int_a^0 f(x) dx = \int_0^T f(x) dx$$

Esempi notevoli di funzioni periodiche di periodo $T = 2\pi$ sono le funzioni trigonometriche:

$$\sin(x), \cos(x), \sin(nx), \cos(nx)$$

e, per la proprietà 3., le loro combinazioni lineari. In particolare si definisce *Polinomio trigonometrico di ordine N* la funzione 2π -periodica della forma:

$$P_N(x) = \alpha_0 + \sum_{n=1}^N \alpha_n \cos(nx) + \beta_n \sin(nx)$$

In generale, se $T \neq 2\pi$, esempi notevoli di funzioni periodiche di periodo T sono:

$$\sin\left(\frac{2\pi}{T}x\right), \cos\left(\frac{2\pi}{T}x\right), \sin\left(\frac{2\pi n}{T}x\right), \cos\left(\frac{2\pi n}{T}x\right)$$

e, per la proprietà 3., le loro combinazioni lineari.

Introduciamo ora un tipo particolare di serie di funzioni, le *serie trigonometriche* della forma

$$a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \quad (52)$$

ottenute sommando le funzioni trigonometriche di periodo 2π , oppure, più in generale, sommando le funzioni trigonometriche di periodo T

$$a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n}{T}x\right) + b_n \sin\left(\frac{2\pi n}{T}x\right) \quad (53)$$

Osserviamo che se le serie (52) e (53) convergono (puntualmente), allora definiscono rispettivamente una funzione periodica con lo stesso periodo delle funzioni periodiche con cui sono costruite. Infatti, indicata con $s : \mathbb{R} \rightarrow \mathbb{R}$ la funzione definita da

$$\begin{aligned} s(x) &= \lim_{N \rightarrow \infty} \left(a_0 + \sum_{n=1}^N a_n \cos\left(\frac{2\pi n}{T}x\right) + b_n \sin\left(\frac{2\pi n}{T}x\right) \right) \\ &= \lim_{N \rightarrow \infty} P_N(x) \end{aligned}$$

abbiamo che

$$s(x+T) = \lim_{N \rightarrow \infty} P_N(x+T) = \lim_{N \rightarrow \infty} P_N(x) = s(x),$$

dove la prima e la terza disuguaglianza seguono dalla definizione della funzione s come limite (puntuale) della successione dei polinomi trigonometrici $\{P_N\}$, mentre la seconda uguaglianza segue dalla periodicità di tali funzioni.

Occupiamoci ora del problema della convergenza della serie trigonometrica (53).

Teorema 42 Se per ogni $n \geq 1$ i coefficienti a_n, b_n soddisfano la disuguaglianza¹⁰:

$$|a_n| + |b_n| \leq c_n, \quad \text{con} \quad \sum_{n=1}^{\infty} c_n < \infty \quad (54)$$

allora la serie di funzioni

$$a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n}{T}x\right) + b_n \sin\left(\frac{2\pi n}{T}x\right)$$

è totalmente convergente e definisce una funzione continua $s: \mathbb{R} \rightarrow \mathbb{R}$, periodica di periodo T .

Dimostrazione: È sufficiente verificare che per ogni n è verificata la disuguaglianza

$$\left| a_n \cos\left(\frac{2\pi n}{T}x\right) + b_n \sin\left(\frac{2\pi n}{T}x\right) \right| \leq |a_n| + |b_n| \leq c_n,$$

con $\sum_{n=1}^{\infty} c_n < \infty$, ed è quindi verificata la condizione che assicura la convergenza totale della serie. Per il teorema 34, dato che le funzioni trigonometriche sono continue, la somma della serie di funzioni è a sua volta una funzione continua. \square

Il seguente teorema fornisce delle condizioni più deboli sui coefficienti a_n e b_n che assicurano la convergenza puntuale della serie trigonometrica.

Proposizione 7 Sia $\{a_n\}$ successione a valori reali positivi, con $\lim_{n \rightarrow \infty} a_n = 0$. Allora valgono le seguenti proprietà:

1. le serie trigonometriche $\sum_n a_n \sin(nx)$ e $\sum_n a_n \cos(nx)$ convergono puntualmente per ogni $x \in (0, 2\pi)$. In $x = 0$, e in $x = 2\pi$ la serie $\sum_n a_n \sin(nx)$ è identicamente nulla, mentre la serie $\sum_n a_n \cos(nx)$ può essere convergente oppure no.
2. le serie trigonometriche $\sum_n (-1)^n a_n \sin(nx)$ e $\sum_n (-1)^n a_n \cos(nx)$ convergono puntualmente per ogni $x \in [0, 2\pi]$ escluso il punto $x = \pi$, in cui la serie $\sum_n (-1)^n a_n \sin(nx)$ è identicamente nulla, mentre la serie $\sum_n (-1)^n a_n \cos(nx)$ può essere convergente oppure no.

¹⁰Di fatto, affinché valga la tesi, è sufficiente che la disuguaglianza (54) sia valida per tutti gli n maggiori di un certo valore n_0 .

Ad esempio, la proposizione precedente permette di mostrare che le serie trigonometriche

$$\sum_n \frac{\sin(nx)}{\sqrt{n}}, \quad \sum_n \frac{\cos(nx)}{\sqrt{n}}$$

sono convergenti in $(0, 2\pi)$. In $x = 0$ e in $x = 2\pi$ la serie $\sum_n \frac{\sin(nx)}{\sqrt{n}}$ vale 0, mentre la serie $\sum_n \frac{\cos(nx)}{\sqrt{n}}$ non è convergente in $x = 0$ e in $x = 2\pi$.

Calcolo dei coefficienti di Fourier

Occupiamoci ora di stabilire sotto quali condizioni una funzione periodica $f : \mathbb{R} \rightarrow \mathbb{R}$ di periodo T può essere scritta come somma nella forma

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n}{T}x\right) + b_n \sin\left(\frac{2\pi n}{T}x\right)$$

e quali sono i valori dei coefficienti a_n, b_n . Premettiamo le definizioni di funzione continua a tratti e di funzione derivabile a tratti.

Definizione 48 Una funzione $f : [a, b] \rightarrow \mathbb{R}$ è detta continua a tratti sull'intervallo $[a, b]$ se esiste un insieme $D \subset [a, b]$ con un numero finito di punti tale che:

- f è continua in ogni punto $x \in [a, b] \setminus D$
- per ogni $\bar{x} \in D$ esistono e sono finiti il limite destro ed il limite sinistro di f in \bar{x} , che denoteremo rispettivamente con $f(\bar{x}^+)$ e $f(\bar{x}^-)$:

$$\lim_{x \rightarrow \bar{x}^+} f(x) \equiv f(\bar{x}^+), \quad \lim_{x \rightarrow \bar{x}^-} f(x) \equiv f(\bar{x}^-).$$

Definizione 49 Una funzione $f : [a, b] \rightarrow \mathbb{R}$ è detta derivabile a tratti sull'intervallo $[a, b]$ se è continua a tratti su $[a, b]$ e se per ogni $x \in [a, b]$ esistono e sono finiti i limiti:

$$\lim_{h \rightarrow 0^+} \frac{f(x+h) - f(x^+)}{h}, \quad \lim_{h \rightarrow 0^-} \frac{f(x+h) - f(x^-)}{h}$$

Teorema 43 (Convergenza puntuale delle serie di Fourier) Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione periodica di periodo $T > 0$ e derivabile a tratti sull'intervallo $[0, T]$. Definiti i coefficienti:

$$a_n := \frac{2}{T} \int_{-T/2}^{T/2} f(x) \cos\left(\frac{2\pi n}{T}x\right) dx, \quad b_n := \frac{2}{T} \int_{-T/2}^{T/2} f(x) \sin\left(\frac{2\pi n}{T}x\right) dx$$

allora per ogni $x \in \mathbb{R}$ la serie di Fourier di f converge alla media fra il limite destro ed il limite sinistro di f in x :

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n}{T}x\right) + b_n \sin\left(\frac{2\pi n}{T}x\right) = \frac{f(x^+) + f(x^-)}{2}$$

Esempio 43 Consideriamo ad esempio la funzione $f: \mathbb{R} \rightarrow \mathbb{R}$ periodica di periodo 2π definita da

$$f(x) = \begin{cases} +\pi & \text{per } x \in [0, \pi) \\ -\pi & \text{per } x \in [-\pi, 0) \end{cases}$$

I coefficienti di Fourier di f sono

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx = 0$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx = \frac{2}{\pi} \int_0^{\pi} \pi \sin(nx) dx = 2 \frac{1 - \cos(n\pi)}{n} = \begin{cases} \frac{4}{n} & \text{per } n \text{ dispari} \\ 0 & \text{per } n \text{ pari} \end{cases}$$

Otteniamo quindi

$$\frac{f(x^+) + f(x^-)}{2} = \sum_{k=0}^{\infty} \frac{4}{2k+1} \sin((2k+1)x)$$

quindi se $x \in (0, \pi)$ abbiamo $\sum_{k=0}^{\infty} \frac{4}{2k+1} \sin((2k+1)x) = \pi$, mentre se $x \in (-\pi, 0)$ abbiamo $\sum_{k=0}^{\infty} \frac{4}{2k+1} \sin((2k+1)x) = -\pi$, mentre se $x = 0$ e $x = \pi$ otteniamo l'identità banale $0 = 0$.

Applicazioni delle serie di Fourier alla soluzione di equazioni differenziali

La motivazione che condusse Fourier (1768-1830) verso lo studio delle serie trigonometriche fu originariamente l'indagine dei fenomeni di propagazione del calore. L'argomento risultava di particolare interesse per le sue numerose applicazioni: dalla lavorazione dei metalli ai tentativi di determinare la temperatura all'interno della terra.

Illustriamo qui di seguito un modello semplificato che permette comunque di mettere in luce l'utilità delle serie di Fourier.

Consideriamo una sbarra di metallo di lunghezza $L = \pi$, e indichiamo con la variabile x la posizione dei punti della sbarra. Supponiamo che ai 2 estremi, in $x = 0$ e in $x = \pi$, la temperatura della sbarra sia mantenuta costante e pari a 0. Supponiamo inoltre che la distribuzione iniziale della temperatura della sbarra sia descritta

dalla funzione $f : [0, \pi] \rightarrow \mathbb{R}$. L'evoluzione in tempo della distribuzione di temperatura della sbarra, descritta dalla funzione $u(t, x)$, con $t \in \mathbb{R}$ e $x \in [0, \pi]$, è descritta dall'equazione del calore $\frac{\partial}{\partial t}u(t, x) = k\frac{\partial^2}{\partial x^2}u(t, x)$, dove k è una costante positiva collegata alle caratteristiche del materiale. In particolare dobbiamo risolvere il problema

$$\begin{cases} \frac{\partial}{\partial t}u(t, x) = k\frac{\partial^2}{\partial x^2}u(t, x) \\ u(t, 0) = u(t, \pi) = 0 \\ u(0, x) = f(x) \end{cases} \quad (55)$$

Cerchiamo delle soluzioni della forma $u(t, x) = T(t)X(x)$.

Notiamo che, affinché siano soddisfatte le condizioni al bordo $u(t, 0) = u(t, \pi) = 0$, la funzione $X : [0, 2\pi] \rightarrow \mathbb{R}$ dovrà soddisfare le condizioni $X(0) = X(\pi) = 0$.

Sostituendo nell'equazione $\frac{\partial}{\partial t}u(t, x) = k\frac{\partial^2}{\partial x^2}u(t, x)$ otteniamo $T'(t)X(x) = kT(t)X''(x)$, che può essere scritta come $\frac{T'(t)}{kT(t)} = \frac{X''(x)}{X(x)}$. Quest'ultima identità può essere soddisfatta se e solo se entrambi i membri sono uguali ad una costante λ :

$$\begin{cases} \frac{T'(t)}{kT(t)} = \lambda \\ \frac{X''(x)}{X(x)} = \lambda, \quad X(0) = X(\pi) = 0 \end{cases}$$

L'equazione $X''(x) = \lambda X(x)$ con le condizioni al bordo $X(0) = X(\pi) = 0$, ammette una soluzione non banale se $\lambda = -n^2$, $n \geq 1$. In questo caso la soluzione è $X(x) = \sin(nx)$. La soluzione dell'equazione $T'(t) = k\lambda T(t)$ collegata, della forma quindi $T'(t) = -kn^2T(t)$, è $T(t) = e^{-kn^2t}$.

Abbiamo costruito quindi, per ogni $n \in \mathbb{N}$ con $n \geq 1$, una soluzione $u_n(t, x) = e^{-kn^2t} \sin(nx)$ dell'equazione $\frac{\partial}{\partial t}u(t, x) = k\frac{\partial^2}{\partial x^2}u(t, x)$ con condizioni al bordo $u(t, 0) = u(t, \pi) = 0$. Dato che l'equazione del calore è lineare, una qualsiasi combinazione lineare delle sue soluzioni è ancora soluzione¹¹. Cerchiamo quindi la soluzione del problema (55) della forma $\sum_n c_n u_n(t, x)$, con c_n costanti reali.

Imponendo la condizione iniziale $u(0, x) = f(x)$, otteniamo $\sum_n c_n \sin(nx) = f(x)$. Vediamo quindi che riusciamo a risolvere il problema (55) se siamo in grado di esprimere la funzione $f : [0, \pi] \rightarrow \mathbb{R}$, la distribuzione iniziale di temperatura della sbarra, come serie trigonometrica somma solamente di multipli delle funzioni $\sin(nx)$.

A tale scopo consideriamo la funzione dispari $g : \mathbb{R} \rightarrow \mathbb{R}$ periodica di periodo 2π definita da $g(x) = f(x)$ per $x \in [0, \pi]$ e $g(x) = -f(x)$ per $x \in [-\pi, 0]$. Dato che g è dispari, allora il suo sviluppo in serie di Fourier è della forma $g(x) = \sum_{n=1}^{\infty} c_n \sin(nx)$, con $c_n = \frac{2}{\pi} \int_0^{\pi} g(x) \sin(nx) dx$. Inoltre, dall'uguaglianza $g(x) = f(x)$ sull'intervallo

¹¹Notiamo inoltre che, dato che le condizioni al bordo $u(t, 0) = u(t, \pi) = 0$ sono omogenee, una combinazione lineare di soluzioni che soddisfano le condizioni al bordo è ancora una soluzione che soddisfa le stesse condizioni al bordo.

$[0, \pi]$ otteniamo:

$$f(x) = \sum_{n=1}^{\infty} c_n \sin(nx), \quad c_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx) dx.$$

La soluzione del problema (55) è dunque:

$$u(t, x) = \sum_{n=1}^{\infty} c_n e^{-kn^2 t} \sin(nx), \quad c_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx) dx.$$

Ad esempio, se la distribuzione iniziale di temperatura è descritta dalla funzione $f(x) = x(\pi - x)$, con $x \in [0, \pi]$, allora abbiamo che i coefficienti di Fourier c_n valgono:

$$c_n = 0, \quad \text{se } n \text{ è pari}, \quad c_n = \frac{8}{\pi n^3}, \quad \text{se } n \text{ è dispari}.$$

La soluzione del problema (55) è quindi:

$$u(t, x) = \sum_{m=0}^{\infty} \frac{8}{\pi(2m+1)^3} e^{-k(2m+1)^2 t} \sin((2m+1)x).$$

Un modello simile descrive la propagazione del calore attraverso la superficie terrestre. Se indichiamo con la variabile x la profondità e con $u(t, x)$ la temperatura al tempo t e alla profondità x , tale funzione è soluzione dell'equazione $\frac{\partial}{\partial t} u(t, x) = k \frac{\partial^2}{\partial x^2} u(t, x)$. La funzione $u(t, x)$ è nota solo in superficie, quindi in $x = 0$, $u(t, 0) = f_0(t)$. Ipotizzando che $u(t, x)$ sia una funzione periodica di periodo $T = 1$ anno, scriviamo la soluzione come serie di Fourier nella variabile t con dei coefficienti che dipendono dalla profondità x :

$$\begin{aligned} u(t, x) &= u_0(x) + \sum_{n=1}^{\infty} a_n(x) \cos\left(\frac{2\pi n}{T} t\right) + b_n(x) \sin\left(\frac{2\pi n}{T} t\right), \\ &= u_0(x) + \sum_{n=1}^{\infty} u_n(x) \cos\left(\frac{2\pi n}{T} t + \varphi_n(x)\right) \end{aligned}$$

dove $u_0(x) = \frac{1}{T} \int_0^T u(t, x) dt$ è la media temporale della temperatura alla profondità x . Nella seconda riga, la serie di Fourier è stata scritta in forma equivalente con $u_n(x) = \sqrt{a_n(x)^2 + b_n(x)^2}$ e il termine $\varphi_n(x)$ esprime una fase, cioè un angolo definito da $\cos \varphi_n(x) = \frac{a_n(x)}{u_n(x)}$, e $\sin \varphi_n(x) = -\frac{b_n(x)}{u_n(x)}$. Esprimendo inoltre la funzione $f_0(t)$ in serie di Fourier nella forma

$$f_0(t) = A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{2\pi n}{T} t + \beta_n\right)$$

e imponendo che la funzione $u(t, x)$ sia soluzione dell'equazione differenziale con condizione al bordo $u(t, 0) = f_0(t)$ si ottiene:

$$u_n(x) = A_n e^{-\sqrt{\frac{n\pi}{kT}}x}, \quad \varphi_n(x) = \beta_n - \sqrt{\frac{n\pi}{kT}}x$$

La soluzione quindi diviene

$$u(t, x) = A_0 + \sum_{n=1}^{\infty} A_n e^{-\sqrt{\frac{n\pi}{kT}}x} \cos\left(\frac{2\pi n}{T}t + \beta_n - \sqrt{\frac{n\pi}{kT}}x\right)$$

Per valori realistici di K , la costante $\sqrt{\frac{\pi}{kT}}x$ vale approssimativamente $\frac{2\pi}{900}cm^{-1}$. Ad una profondità pari a $x = 450$ cm si ha uno sfasamento dell'onda con frequenza $2\pi/T$, pari a π . Questo significa che a una tale profondità il picco di calore dell'estate e rispettivamente il freddo invernale arrivano con un ritardo di 6 mesi. Inoltre l'ampiezza dell'oscillazione, cioè $u_1(x) = A_1 e^{-\sqrt{\frac{\pi}{kT}}x}$ viene smorzata di un fattore pari a $e^{-\pi} \sim 1/16$. Questo significa che se al livello del suolo c'è una differenza di 40 gradi tra il minimo di temperatura invernale ed il massimo di temperatura estivo, alla profondità $x = 450cm$ tale oscillazione viene smorzata di $1/16$ e vale circa 2 gradi.

Un altro interessante problema risolvibile con tali tecniche è lo studio delle oscillazioni di una corda vibrante. Consideriamo una corda di lunghezza $L = \pi$ vincolata ai due estremi e indichiamo con $u(t, x)$ l'altezza di un punto della corda che a riposo si trova nella posizione x . Questa soddisfa l'equazione delle onde

$$\frac{\partial^2}{\partial t^2}u(t, x) = a^2 \frac{\partial^2}{\partial x^2}u(t, x)$$

Supponiamo di conoscere la forma della corda all'istante iniziale descritta dalla funzione $u(0, x) = f(x)$ e la distribuzione iniziale delle velocità dei punti della corda, descritte dalla funzione $\frac{\partial}{\partial t}u(0, x) = g(x)$. La condizione al bordo $u(t, 0) = u(t, \pi) = 0$ esprime il fatto che i due estremi della corda sono bloccati.

Cercando soluzioni della forma $u(t, x) = T(t)X(x)$ e imponendo le condizione al bordo abbiamo $\frac{T''(t)}{aT(t)} = \frac{X''(x)}{X(x)} = \lambda$, che ammette soluzioni non banali se $\lambda = -n^2$.

Otteniamo $X(x) = \sin(nx)$, e $T(t) = c_n \cos(ant) + d_n \sin(ant)$.

Costruendo la soluzione $u(t, x)$ come combinazione lineare delle soluzioni appena trovate, otteniamo

$$u(t, x) = \sum_{n=1}^{\infty} c_n \cos(ant) \sin(nx) + d_n \sin(ant) \sin(nx)$$

Imponendo le condizioni iniziali

$$u(0, x) = \sum_{n=1} c_n \sin(nx) = f(x), \quad \frac{\partial}{\partial t} u(0, x) = \sum_{n=1} n a d_n \sin(nx) = g(x)$$

vediamo che le costanti c_n e d_n possono essere ricavate tramite lo sviluppo in serie di Fourier di $f(x)$ e $g(x)$.

Ad esempio, se $f(x) = x(\pi - x)$ e $g(x) = 0$, allora abbiamo che $d_n = 0$ e $c_n = 0$ se n è dispari, mentre $c_n = \frac{8}{\pi n^2}$ se n è pari. Abbiamo quindi:

$$u(t, x) = \sum_{m=0} \frac{8}{\pi(2m+1)^2} \cos(a(2m+1)t) \sin((2m+1)x).$$

Elementi di statistica descrittiva

La statistica si occupa dell'organizzazione, dell'elaborazione e dell'interpretazione di insiemi di dati. Si suddivide nella *statistica descrittiva*, che tratta la raccolta, l'organizzazione e la descrizione sintetica dei dati, e la *statistica inferenziale*, che si occupa dell'interpretazione dei dati.

È importante ricordare che i dati raccolti nella maggior parte delle situazioni dipendono da fattori casuali. Ad esempio i dati raccolti in un esperimento scientifico, le misurazioni, sono soggette agli errori di misura e ripetendo l'esperimento si ottengono valori differenti. Un altro esempio che mette in luce questo tipo di problemi è la raccolta dei dati in un sondaggio: in questo caso infatti le risposte raccolte (ad esempio in un sondaggio riguardante le intenzioni di voto degli elettori) sono influenzate da fattori casuali, tra i quali la scelta del campione della popolazione a cui vengono rivolte le domande. L'interpretazione dei dati statistici, cioè l'analisi di problemi di tipo inferenziale, è basata quindi sul calcolo delle probabilità, una teoria i cui primi elementi verranno illustrati nel prossimo capitolo di queste note. Per una trattazione più dettagliata e completa rimandiamo il lettore ai seguenti testi:

P. Baldi, *Introduzione alla probabilità con elementi di statistica*. McGraw-Hill.

M. Boella, *Probabilità e statistica per ingegneria e scienze*. Pearson. Prentice Hall.

W. Navidi, *Probabilità e statistica l'ingegneria e le scienze*. McGraw-Hill.

Un insieme di dati statistici viene detto *campione*. Di fatto i dati statistici possono essere di due tipi: quantitativo e qualitativo. I dati di tipo quantitativo sono dei numeri sui quali ha senso effettuare una serie di operazioni per il calcolo di grandezze significative. Pensiamo ad esempio ai dati relativi all'altezza degli individui di una certa popolazione: è possibile calcolarne la media aritmetica ottenendo un valore significativo (che esprime appunto l'altezza media della popolazione). I dati di tipo qualitativo invece non sono di carattere numerico (pensiamo ad esempio ai dati relativi al gruppo sanguigno degli individui di una certa popolazione) e non è possibile quindi effettuare su di essi operazioni che hanno invece senso per i dati quantitativi.

I possibili valori che può assumere un dato statistico (qualitativo o quantitativo) sono detti *modalità*. Ad esempio, nel caso in cui i dati rappresentano il gruppo sanguigno degli individui di una certa popolazione abbiamo 4 modalità : A, B, AB, 0. Se i dati rappresentano il numero di figli degli individui di una certa popolazione le possibili modalità sono i numeri interi positivi. Se i dati rappresentano l'altezza degli individui abbiamo un continuo di possibili valori e le modalità possibili sono infatti tutti in numeri reali positivi.

Consideriamo ora un insieme di n dati (un campione statistico) $\{x_1, x_2, \dots, x_n\}$ relativi ad una grandezza che ha un numero finito m di modalità z_1, \dots, z_m , che etichetteremo con l'indice $i = 1, \dots, m$. Indichiamo con N_i il numero dei dati del campione che corrispondono all' i -esima modalità :

$$N_i = \#\{j : x_j = z_i\}$$

Dove il simbolo $\#$ indica la cardinalità (il numero di elementi) di un insieme. Si definisce *frequenza relativa* p_i relativa all' i -esima modalità il rapporto tra N_i e n :

$$p_i = \frac{N_i}{n}, \quad i = 1, \dots, m.$$

Se i dati di un campione statistico $\{x_1, x_2, \dots, x_n\}$ sono relativi ad una grandezza che ha un numero infinito di modalità, che supponiamo essere comunque incluse in un intervallo $I \subset \mathbb{R}$, allora è ancora possibile adottare una descrizione in termini di frequenze relative dividendo l'intervallo I in un numero finito di sottointervalli I_1, I_2, \dots, I_m disgiunti e tali che $I = I_1 \cup I_2 \cup \dots \cup I_m$. Si indica quindi con N_i il numero dei dati del campione che cadono all'interno dell' i -esimo intervallo:

$$N_i = \#\{j : x_j \in I_i\}$$

e con p_i il rapporto $p_i = \frac{N_i}{n}$. È chiaro che questi valori dipendono dal tipo di suddivisione I_1, I_2, \dots, I_m dell'intervallo I .

La rappresentazione grafica di tali valori viene effettuata tramite gli *istogrammi*. Nel caso di un numero finito di modalità questi grafici si costruiscono mettendo sull'asse delle ascisse le modalità e disegnando, per ognuna di esse, una colonna la cui altezza è proporzionale alla frequenza relativa di quella modalità. Questo tipo di rappresentazione grafica permette di mettere in evidenza la modalità con frequenza relativa più alta.

Nel caso di una variabile con un numero infinito di modalità incluse in un intervallo $I \subset \mathbb{R}$, una volta suddiviso quest'ultimo in sottointervalli I_1, I_2, \dots, I_m che vengono rappresentati sull'asse delle ascisse, si costruisce per ognuno di essi un rettangolo la cui area è proporzionale alla corrispondente frequenza relativa. L'intervallo I_j corrispondente al rettangolo con altezza maggiore viene detto *moda*.

Media e varianza

Dato un campione $\{x_1, x_2, \dots, x_n\}$ di dati numerici, definiamo *media* la quantità :

$$\bar{x} := \frac{\sum_{i=1}^n x_i}{n}$$

La media possiede le seguenti proprietà :

1. \bar{x} è quel particolare valore che minimizza la quantità

$$s(a) = \sum_{i=1}^n (x_i - a)^2$$

2. Introdotta la nuova variabile y , collegata alla variabile x dalla relazione $y = \alpha x + \beta$, con $\alpha, \beta \in \mathbb{R}$, e considerato il campione $\{y_1, y_2, \dots, y_n\}$, con $y_i = \alpha x_i + \beta$, allora

$$\bar{y} = \alpha \bar{x} + \beta$$

3. Considerati due campioni $\{x_1, x_2, \dots, x_n\}$ e $\{x'_1, x'_2, \dots, x'_m\}$ con media rispettivamente \bar{x} e \bar{x}' , la media del campione $\{x_1, x_2, \dots, x_n, x'_1, x'_2, \dots, x'_m\}$ ottenuto unendo i due campioni è data da

$$\frac{n}{n+m} \bar{x} + \frac{m}{n+m} \bar{x}'$$

La media è un *indice di centralità* : dà delle indicazioni sul valore intorno al quale sono centrati i dati del campione.

Si definisce *varianza* del campione $\{x_1, x_2, \dots, x_n\}$ la quantità, indicata con il simbolo σ_x^2 , calcolata come:

$$\sigma_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Tale quantità possiede le seguenti proprietà :

1. può essere calcolata anche tramite la formula $\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2$.
2. introdotta la nuova variabile y , collegata alla variabile x dalla relazione $y = \alpha x + \beta$, con $\alpha, \beta \in \mathbb{R}$, e considerato il campione $\{y_1, y_2, \dots, y_n\}$, con $y_i = \alpha x_i + \beta$, allora la varianza del nuovo campione è collegata alla varianza del vecchio dalla relazione

$$\sigma_y^2 = \alpha^2 \sigma_x^2$$

Si definisce *deviazione standard* o *scarto quadratico medio* la radice quadrata della varianza. Tale valore si indica con il simbolo σ_x :

$$\sigma_x := \sqrt{\sigma_x^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

La varianza e lo scarto quadratico medio sono degli *indici di dispersione*: indicano di quanto i valori del campione si allontanano dal valore medio.

Mediana e quantili

Dato un campione di dati numerici $\{x_1, x_2, \dots, x_n\}$ ordiniamoli in maniera crescente come

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

Si definisce *mediana* quel valore $x_{(i)}$ tale per cui metà dei dati ha valori minori di $x_{(i)}$ e metà dei dati ha valori maggiori di $x_{(i)}$. Più precisamente, se n è dispari, cioè $n = 2k + 1$, allora la mediana è il valore $x_{(k+1)}$, mentre se n è pari, cioè $n = 2k$, allora la mediana è il valore $x_{(k)}$. In alcuni testi viene definita come mediana il valore intermedio fra $x_{(k)}$ e $x_{(k+1)}$, cioè $\frac{x_{(k)} + x_{(k+1)}}{2}$.

La mediana è un indice di centralità differente, in generale, dalla media. Rispetto a quest'ultima ha il vantaggio di essere "robusta", ovvero poco sensibile alla presenza di dati affetti da errore.

L'ampiezza dell'intervallo $[x_{(1)}, x_{(n)}]$ a cui appartengono i valori del campione, cioè il valore $x_{(n)} - x_{(1)}$ viene chiamato *range*. Rappresenta un indice di dispersione.

Introduciamo ora il concetto di quantile. Fissato un valore $0 < \alpha < 1$, viene chiamato *quantile* di ordine α il valore $q_\alpha \equiv x_{(i)}$, con $i = \lfloor \alpha(n+1) \rfloor$, dove il simbolo $\lfloor \alpha(n+1) \rfloor$ indica il più piccolo intero minore o uguale di $\alpha(n+1)$. Di fatto in letteratura non c'è una definizione univoca, e in alcuni testi si usa il valore $x_{(i)}$, con $i = \lfloor \alpha(n+1) \rfloor + 1$. Di fatto q_α corrisponde a quel valore tale per cui αn dati sono minori di q_α , mentre αn dati sono maggiori di q_α .

Il valore $q_{1/2}$ corrisponde alla mediana.

I valori $q_{1/4}, q_{2/4}, q_{3/4}$ sono chiamati *quartili*. Il particolare l'intervallo $[q_{1/4}, q_{3/4}]$ contiene metà dei dati e la sua ampiezza $q_{3/4} - q_{1/4}$ è una misura della dispersione dei dati.

Correlazione e regressione

Talvolta si considerano dei campioni $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ i cui dati sono delle coppie (x_i, y_i) , che indicano i valori di due grandezze. Consideriamo ad esempio i dati relativi al peso e all'altezza degli individui di una popolazione, oppure i dati sperimentali relativi a 2 grandezze fisiche misurate in un certo esperimento.

Ipotizzando una relazione tra la grandezza x e la grandezza y di tipo:

$$y = ax + b$$

cerchiamo di stimare a partire dai dati del campione i "migliori" valori dei parametri $a, b \in \mathbb{R}$. Rappresentando i punti (x_i, y_i) in un piano cartesiano xy , il problema

equivale a cercare la "migliore retta" di interpolazione dei punti. Applicando il metodo dei minimi quadrati, cerchiamo quei valori a, b che minimizzano la quantità

$$S(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2$$

che esprime la somma dei quadrati delle distanze fra i punti (x_i, y_i) e il punto $(x_i, ax_i + b)$ di ascissa x_i appartenente alla retta $y = ax + b$. Imponendo $\nabla S(a, b) = (0, 0)$ si ottiene:

$$a = \frac{\sigma_{xy}}{\sigma_x^2} \quad b = \bar{y} - a\bar{x} = \bar{y} - \frac{\sigma_{xy}}{\sigma_x^2} \bar{x},$$

dove $\sigma_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$ è detta *covarianza di x e y* . La retta $y = ax + b$ è detta *retta di regressione lineare*. Se $\sigma_{xy} > 0$ allora il coefficiente angolare a della retta di regressione è positivo e un aumento di x comporta un aumento del valore della variabile y . Se $\sigma_{xy} < 0$ allora il coefficiente angolare a della retta di regressione è negativo e un aumento di x comporta una diminuzione del valore della variabile y . Se $\sigma_{xy} \sim 0$ allora il coefficiente angolare a della retta di regressione è approssimativamente nullo e una variazione del valore della variabile x non comporta una variazione del valore della variabile y . In questo caso si dice che le grandezze x e y non sono correlate.

È utile introdurre, al posto della quantità σ_{xy} , la grandezza $\rho_{xy} := \frac{\sigma_{xy}}{\sigma_x \sigma_y}$, detta *coefficiente di correlazione*. Dato che vale la seguente disuguaglianza¹²

$$|\sigma_{xy}| \leq \sigma_x \sigma_y$$

si ha che il coefficiente di correlazione soddisfa la disuguaglianza $-1 \leq \rho_{xy} \leq 1$. Per valori di ρ_{xy} vicini a 1 le variabili x e y hanno una correlazione positiva: un aumento del valore di x comporta un aumento del valore di y . Per valori di ρ_{xy} vicini a -1 le variabili x e y hanno una correlazione negativa: un aumento del valore di x comporta una diminuzione del valore di y . Per valori di ρ_{xy} vicini a 0 le variabili x e y sono scorrelate.

¹²la disuguaglianza $|\sigma_{xy}| \leq \sigma_x \sigma_y$ si dimostra applicando la disuguaglianza di Cauchy-Schwartz

$$|v \cdot w| \leq \|v\| \|w\|$$

(valida in ogni spazio vettoriale con prodotto scalare) ai due vettori di \mathbb{R}^n :

$$v = \frac{1}{\sqrt{n}}(x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}), \quad e \quad w = \frac{1}{\sqrt{n}}(y_1 - \bar{y}, y_2 - \bar{y}, \dots, y_n - \bar{y})$$

Elementi di calcolo delle probabilità

Per un gran numero di fenomeni non è possibile, o non è conveniente, dare una descrizione deterministica, ovvero dare un insieme di leggi che, note certe premesse (condizioni iniziali) permettono di descrivere il fenomeno con certezza e precisione. Questo avviene ad esempio quando non si ha una conoscenza completa di come si sviluppa il fenomeno, ad esempio a causa di un numero eccessivo di variabili da tenere in considerazione, oppure a causa dell'incertezza introdotta dagli apparati di misura delle grandezze in gioco. Per descrivere tali problemi è dunque necessario introdurre un approccio probabilistico, cercando di stimare la probabilità con cui si verifica un dato evento. La branca della matematica che si occupa di modellizzare problemi di predizione e derivare delle “regole di calcolo” che, nota la probabilità di eventi elementari, permettono di calcolare la probabilità di eventi complessi è detta *calcolo delle probabilità*.

Cosa si intende per probabilità di un evento?

L'interesse matematico per questo tipo di problemi ha origine nel 1600 (con lo studio della teoria dei giochi), ma per avere una formalizzazione matematicamente rigorosa del concetto di probabilità di un evento dobbiamo aspettare fino all'inizio del 1900 con l'assiomatizzazione di Kolmogorov.

Da un punto di vista intuitivo possiamo pensare alla probabilità di un evento come il grado di fiducia nel suo verificarsi. Per trattare questo concetto in maniera matematicamente rigorosa il matematico russo A.N. Kolmogorov ha introdotto la definizione di probabilità in termini di una *misura* non negativa sull'insieme dei possibili eventi.

Spazio delle prove ed eventi

Supponiamo di effettuare un esperimento il cui risultato non è prevedibile con certezza, però è noto l'insieme dei possibili risultati. Tale insieme viene detto **spazio delle prove** e viene indicato con Ω . Ad esempio se l'esperimento consiste nel lancio di un dado lo spazio Ω è l'insieme

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Se invece l'esperimento consiste nella misura del tempo impiegato dal un nucleo radioattivo a decadere lo spazio Ω sarà l'insieme di tutti i numeri reali positivi, ovvero $\Omega = [0, +\infty)$.

Definiamo **evento** ogni sottoinsieme¹³. Ad esempio nel caso del lancio di un dado, l'evento "esce il numero 2" è rappresentato dall'insieme $A = \{2\}$, mentre l'evento "esce un numero dispari" è rappresentato dall'insieme $B = \{1, 3, 5\}$. Nell'esperimento sul decadimento del nucleo radioattivo l'evento "il nucleo decade dopo l'istante T_1 e prima dell'istante T_2 " è rappresentato dall'insieme $A = [T_1, T_2]$.

Tale modellizzazione degli eventi come insiemi dello spazio delle prove è utile perchè le operazioni di unione, di intersezione insiemistica e di passaggio al complementare possono venire applicate in questo contesto.

Dati due eventi $A, B \subseteq \Omega$, l'evento **unione** $A \cup B$ di A e B rappresenta l'evento in cui "si verifica l'evento A oppure l'evento B ". L'evento **intersezione** $A \cap B$ di A e B rappresenta l'evento in cui "si verificano sia l'evento A sia l'evento B ", mentre l'evento complementare di A , indicato con $A^c = \Omega \setminus A$ rappresenta l'evento in cui "non si verifica l'evento A ".

Ad esempio, nel caso del lancio di un dado, se $A = \{1, 3, 5\}$ rappresenta l'evento "esce un numero dispari" e $B = \{1, 2, 3, 4\}$ rappresenta l'evento "esce un numero minore o uguale a 4" allora $A \cup B = \{1, 2, 3, 4, 5\}$ rappresenta l'evento "esce un numero dispari o un numero minore o uguale a 4", $A \cap B = \{1, 3\}$ rappresenta l'evento "esce un numero dispari minore di 4", mentre $A^c = \{2, 4, 6\}$ rappresenta l'evento "esce un numero pari".

Se due eventi A, B sono tali che $A \cap B = \emptyset$ allora A e B sono detti *disgiunti* o *incompatibili*.

Probabilità di eventi

Modellizziamo ora la probabilità di un evento. Ad ogni evento assegnamo un numero, compreso tra 0 e 1, che misura quanto è verosimile che tale evento si verifichi.

Definizione 50 *Dato uno spazio delle prove Ω , una probabilità P è un'applicazione che assegna ad ogni sottoinsieme $A \subseteq \Omega$ un numero reale $P(A)$, che soddisfa le seguenti proprietà :*

1. $0 \leq P(A) \leq 1$,
2. $P(\Omega) = 1$,

¹³In generale, da un punto di vista matematicamente rigoroso, si definiscono eventi solo particolari sottoinsiemi dello spazio Ω che appartengono ad una famiglia di insiemi, contenete Ω e l'insieme vuoto, chiusa per unione numerabile e passaggio al complementare, detta σ -algebra dello spazio Ω . In queste note non ci soffermeremo sul concetto di σ -algebra, ma ci limiteremo a considerare casi in cui possiamo considerare "eventi" tutti i sottoinsiemi di Ω .

3. data una successione $\{A_n\}_n$ di eventi a due a due disgiunti si ha che $P(\cup_n A_n) = \sum_n P(A_n)$.

Dall'ultima proprietà segue che $P(\emptyset) = 0$.

Ad esempio nell'esperimento del lancio di un dado, se questo non è truccato è verosimile supporre che tutti i 6 possibili risultati siano equiprobabili, assegnamo quindi

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}$$

La probabilità dei restanti eventi viene calcolata utilizzando la proprietà 3 di P , ad esempio rappresentando l'evento $\{1, 2, 4\}$ come unione di eventi disgiunti $\{1, 2, 4\} = \{1\} \cup \{2\} \cup \{4\}$ abbiamo:

$$P(\{1, 2, 4\}) = P(\{1\}) + P(\{2\}) + P(\{4\}) = \frac{1}{2}.$$

Questo esempio è un caso particolare della *probabilità uniforme*. Dato uno spazio delle prove Ω con un numero finito di elementi n allora la probabilità uniforme su Ω è quella che assegna ad ogni evento semplice di Ω , ovvero ad ogni sottoinsieme $\{\omega\}$ di Ω con un unico elemento, la stessa probabilità p . Il valore p può quindi venire calcolato notando che

$$1 = P(\Omega) = P(\cup_{\omega \in \Omega} \{\omega\}) = \sum_{\omega \in \Omega} P(\{\omega\}) = np,$$

da cui $p = 1/n$. Quindi la probabilità di un sottoinsieme $A \subseteq \Omega$ con m elementi è data da

$$P(A) = m/n = \frac{\#A}{\#\Omega}.$$

In tutti i casi in cui si suppone che su un insieme Ω sia definita una probabilità uniforme, la probabilità di un evento A viene calcolata come rapporto tra il numero di elementi di A ed il numero di elementi di Ω . Per calcolare queste quantità vengono applicate le regole del *calcolo combinatorio*.

Alcune regole di calcolo per P

- $P(A^c) = 1 - P(A)$, infatti A e A^c sono disgiunti e $P(A \cup A^c) = P(\Omega) = 1$.
- Dati due eventi A, B , non necessariamente disgiunti, la probabilità della loro unione è data da:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ad esempio consideriamo l'esperimento in cui si lancia due volte un dado. Lo spazio delle prove è l'insieme Ω delle coppie (n_1, n_2) , dove $n_1, n_2 = 1, 2, 3, 4, 5, 6$. In tutto ha $6 * 6 = 36$ elementi. Se il dado non è truccato, possiamo supporre che ogni singolo evento sia equiprobabile, ovvero che su Ω sia definita la probabilità uniforme. Supponiamo di voler calcolare la probabilità dell'evento "esce almeno una volta il numero 6". Tale evento è rappresentato dall'insieme

$$E = \{(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6), (1, 6), (2, 6), (3, 6), (4, 6), (5, 6)\}.$$

Tale evento può essere rappresentato anche come unione dell'evento A "al primo lancio esce il numero 6" e dell'evento B "al secondo lancio esce il numero 6". Abbiamo che $P(A) = 1/6$, $P(B) = 1/6$, mentre $P(A \cap B) = P(\{(6, 6)\}) = 1/36$. Si ha dunque

$$P(A \cup B) = 1/6 + 1/6 - 1/36 = 11/36.$$

Probabilità condizionata

Sia dato uno spazio delle prove Ω e una probabilità P sugli eventi di Ω . Fissiamo inoltre un evento $E \subseteq \Omega$, tale che $P(E) > 0$. Si definisce **probabilità condizionata** di un evento A dato l'evento E , e si indica con $P(A|E)$, una nuova probabilità che esprime un giudizio su quanto è verosimile che l'evento A si verifichi *sapendo che si è verificato l'evento E* .

Viene calcolata tramite la formula seguente

$$P(A|E) = \frac{P(A \cap E)}{P(E)},$$

che risulta più intuitiva se scritta come $P(A \cap E) = P(E)P(A|E)$. $P(\cdot|E)$ è una nuova misura di probabilità a tutti gli effetti, infatti verifica le proprietà 1,2,3 della definizione 50.

Consideriamo l'esperimento del lancio di due dadi (non truccati). Supponiamo che l'evento E sia "il risultato del primo lancio è 4" e l'evento A "la somma dei risultati dei due lanci è 6". Vogliamo calcolare la probabilità che la somma dei risultati dei due lanci dia 6 sapendo che il risultato del primo lancio è 4, ovvero la probabilità condizionata di A dato E .

Per calcolare tale valore possiamo ragionare nel modo seguente. Dato che sappiamo che il risultato del primo lancio è 4, allora per il lancio dei due dadi si presentano le seguenti possibilità $\{(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6)\}$. Nell'ipotesi che il dado non sia truccato allora questi 6 coppie di risultati sono equiprobabili e la loro probabilità è $1/6$. L'evento "la somma dei risultati dei due lanci è 6" è rappresentato dall'insieme $(4, 2)$ la cui probabilità è appunto $1/6$.

Potevamo però ragionare anche nel modo seguente. La probabilità dell'evento A all'evento E è data da

$$P(A|E) = \frac{P(A \cap E)}{P(E)},$$

dove $P(E) = 1/6$ e $P(A \cap E) = P((4, 2)) = 1/36$, quindi

$$P(A|E) = \frac{1/36}{1/6} = 1/6.$$

Eventi indipendenti

Dato uno spazio delle prove Ω e una probabilità P sui sottoinsiemi di Ω , due eventi $A, B \subseteq \Omega$ sono detti indipendenti se

$$P(A \cap B) = P(A)P(B).$$

In altre parole A e B sono indipendenti se $P(A|B) = P(A)$: il verificarsi dell'evento B non ha influenza sulla possibilità di verificarsi dell'evento A .

Ad esempio, se consideriamo il lancio di due dadi, indichiamo con A l'evento "la somma dei risultati è 6" e con B l'evento "il risultato del lancio del primo dado è 4", allora si ha che

$$P(A \cap B) = P(\{(4, 2)\}) = 1/36$$

mentre $P(B) = 1/6$, $P(A) = P(\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}) = 5/36$, quindi

$$P(A)P(B) = \frac{1}{6} \frac{5}{36} \neq \frac{1}{36} = P(A \cap B)$$

quindi i due eventi non sono indipendenti.

Consideriamo ora l'evento C "la somma dei risultati è 7" e con B l'evento "il risultato del lancio del primo dado è 4", allora si ha che

$$P(C \cap B) = P(\{(4, 3)\}) = 1/36$$

mentre $P(B) = 1/6$, $P(C) = P(\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}) = 6/36 = 1/6$, quindi

$$P(C)P(B) = \frac{1}{6} \frac{1}{6} = \frac{1}{36} = P(C \cap B)$$

quindi i due eventi sono indipendenti.

La definizione di indipendenza può essere estesa ad un numero arbitrario n di eventi. Dati n eventi A_1, A_2, \dots, A_n , questi sono detti indipendenti se per ogni $m \leq n$, scelto un sottoinsieme $A_{1'}, A_{2'}, \dots, A_{m'}$ di A_1, A_2, \dots, A_n si ha che

$$P(A_{1'} \cap A_{2'} \cap \dots \cap A_{m'}) = P(A_{1'})P(A_{2'}) \cdots P(A_{m'}).$$

Il seguente esempio mostra come in certe situazioni tre eventi non sono indipendenti se considerati insieme, anche se sono a due a due indipendenti. Consideriamo l'esperimento dell'estrazione di 4 palline, numerate da 1 a 4, da un'urna. Consideriamo i tre eventi

$$A_1 = \{1, 2\}, \quad A_2 = \{1, 3\}, \quad A_3 = \{1, 4\}.$$

Assumendo una probabilità uniforme su $\Omega = \{1, 2, 3, 4\}$ abbiamo che

$$P(A_1) = 1/2 \quad P(A_2) = 1/2, \quad P(A_3) = 1/2,$$

$$P(A_1 \cap A_2) = P(\{1\}) = 1/4 = P(A_1)P(A_2),$$

$$P(A_1 \cap A_3) = P(\{1\}) = 1/4 = P(A_1)P(A_3),$$

$$P(A_2 \cap A_3) = P(\{1\}) = 1/4 = P(A_2)P(A_3),$$

però abbiamo che

$$P(A_1 \cap A_2 \cap A_3) = P(\{1\}) = 1/4 \neq P(A_1)P(A_2)P(A_3) = 1/8.$$

La formula delle probabilità totali

Una famiglia di eventi A_1, A_2, \dots, A_n è detta una *partizione* di Ω se sono disgiunti e la loro unione è l'insieme Ω :

$$\Omega = \cup_{i=1}^n A_i, \quad A_i \cap A_j = \emptyset, \quad i \neq j$$

dato un generico evento B , questo può essere rappresentato come unione disgiunta $B = \cup_{i=1}^n (A_i \cap B)$ e quindi la probabilità di B può essere ottenuta come

$$P(B) = \sum_{i=1}^n P(A_i \cap B) = \sum_{i=1}^n P(B|A_i)P(A_i). \quad (56)$$

La formula (56) è chiamata *formula delle probabilità totali*.

La formula di Bayes

Dato uno spazio delle prove Ω , probabilità P sui sottoinsiemi di Ω e due eventi $A, B \subset \Omega$ tali che $P(B) \neq 0$ e $P(A) \neq 0$, la formula di Bayes mette in relazione la probabilità condizionata di A dato B con la probabilità condizionata di B dato A :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Tale formula si dimostra facilmente notando che la probabilità dell'evento intersezione di A e B può essere ottenuta sia come $P(A \cap B) = P(B|A)P(A)$, sia come $P(A \cap B) = P(A|B)P(B)$, da cui

$$P(B|A)P(A) = P(A|B)P(B)$$

Il seguente esempio mostra un'applicazione interessante della formula di Bayes. Una popolazione è composta per il 40% da fumatori (F) e per il 60% da non fumatori (N). Il 25% dei fumatori e il 7% dei non fumatori sono affetti da una malattia respiratoria cronica. Qual'è la probabilità che una persona affetta dalla malattia sia fumatore? Qual'è la probabilità che una persona affetta dalla malattia sia non fumatore?

Indichiamo con Ω l'insieme di tutti gli individui, con F l'insieme degli individui che fumano, con N l'insieme degli individui che non fumano e con M l'insieme degli individui affetti dalla malattia. I dati in nostro possesso sono

$$P(F) = 0.4, \quad P(N) = 0.6, \quad P(M|F) = 0.25, \quad P(M|N) = 0.07.$$

Vogliamo calcolare $P(F|M)$ e $P(N|M)$. Per utilizzare la formula di Bayes $P(F|M) = \frac{P(M|F)P(F)}{P(M)}$ dobbiamo prima calcolare $P(M)$. Rappresentiamo l'insieme M come l'unione disgiunta di $M \cap F$ e di $M \cap N$, quindi

$$\begin{aligned} P(M) &= P(M \cap F) + P(M \cap N) \\ &= P(M|F)P(F) + P(M|N)P(N) = \\ &= 0.25 * 0.4 + 0.07 * 0.6 = 0.142 \end{aligned}$$

quindi

$$\begin{aligned} P(F|M) &= \frac{P(M|F)P(F)}{P(M)} = \frac{0.25 * 0.4}{0.142} = 0.704 \\ P(N|M) &= \frac{P(M|N)P(N)}{P(M)} = \frac{0.07 * 0.6}{0.142} = 0.296 \end{aligned}$$

Elementi di calcolo combinatorio

Il calcolo combinatorio fornisce una serie di regole di calcolo per calcolare il numero di elementi di un insieme.

È basato sul *principio di fattorizzazione*:

Se si devono compiere k scelte e per la i -scelta abbiamo n_i possibilità, allora il numero totale di alternative possibili è

$$n_1 * n_2 * \dots * n_k$$

Esempio Se abbiamo 6 camice, 2 pantaloni, 5 paia di calzini e 3 paia di scarpe, il numero totale di modi in cui ci si può vestire è $6 * 2 * 5 * 3 = 180$.

Esempio Considerando le 21 lettere dell'alfabeto italiano il numero totale di parole (sensate o meno) composte da 4 lettere è $21 * 21 * 21 * 21 = (21)^4$

Esempio Un'urna contiene 90 palline, numerate da 1 a 90. Si effettuano 4 estrazioni e dopo ogni estrazione la pallina estratta viene reinserita nell'urna. Allora il numero totale di sequenze di numeri estratti è $90 * 90 * 90 * 90 = (90)^4$.

Disposizioni

Definizione 51 Sia N un insieme con n elementi $N = \{x_1, x_2, \dots, x_n\}$. Una disposizione di k elementi di N è una k -upla ordinata $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ di elementi di N tutti distinti fra loro.

Il numero totale delle possibili disposizioni di k elementi di un insieme con n elementi è

$$n(n-1)(n-2) \dots (n-k+1) = \frac{n!}{(n-k)!}$$

Esempio Considerando le 21 lettere dell'alfabeto italiano il numero totale di parole (sensate o meno) composte da 4 lettere distinte è $21 * 20 * 19 * 18 = \frac{21!}{17!}$

Esempio Un'urna contiene 90 palline, numerate da 1 a 90. Si effettuano 4 estrazioni e dopo ogni estrazione la pallina estratta non viene reinserita nell'urna. Allora il numero totale di sequenze di numeri estratti è $90 * 89 * 88 * 87 = 90!/86!$.

Permutazioni

Definizione 52 Sia N un insieme con n elementi $N = \{x_1, x_2, \dots, x_n\}$. Una disposizione di n elementi di N , ovvero una n -upla ordinata $(x_{i_1}, x_{i_2}, \dots, x_{i_n})$ di elementi distinti di N , è detta permutazione di N .

Il numero totale di permutazioni di un insieme con n elementi è $n!$

Esempio Una classe è composta da 20 studenti. Il numero totale di modi in cui possono venire ordinati (es ordine alfabetico...) è $20!$

Combinazioni

Definizione 53 Sia N un insieme con n elementi $N = \{x_1, x_2, \dots, x_n\}$. Una combinazione di k elementi di N è un sottoinsieme di N composto da k elementi $\{x_{i_1}, x_{i_2}, \dots, x_{i_k}\}$

Il numero totale delle possibili combinazioni di k elementi di un insieme con n elementi è

$$\frac{n!}{(n-k)!k!}$$

Esempio Una classe è composta da 20 studenti. Il numero totale di gruppi formati da tre studenti è $\frac{20!}{3!17!}$.

Notiamo che selezionare un sottoinsieme di k elementi da un insieme N con n elementi equivale a scegliere un particolare modo di dividere N in 2 gruppi, il primo con k elementi ed il secondo con $n - k$ elementi. Il numero totale di suddivisioni possibili è dunque $\frac{n!}{(n-k)!k!}$.

Più in generale, dato un insieme N con n elementi, possiamo chiederci in quanti modi può essere suddiviso in k sottoinsiemi, dove il numero di elementi dell' i -esimo sottoinsieme è n_i , con $i = 1, \dots, k$, e $\sum_{i=1}^k n_i = n$. Possiamo ragionare come segue. Per selezionare il primo sottoinsieme abbiamo $\frac{n!}{(n-n_1)!n_1!}$ possibilità; per selezionare il secondo sottoinsieme di n_2 elementi scelti fra i restanti $n - n_1$, abbiamo $\frac{(n-n_1)!}{(n-n_1-n_2)!n_2!}$; per selezionare il terzo sottoinsieme di n_3 elementi scelti fra i restanti $n - n_1 - n_2$, abbiamo $\frac{(n-n_1-n_2)!}{(n-n_1-n_2-n_3)!n_3!}$; infine, scelti gli elementi dei primi $k - 1$ sottoinsiemi, gli elementi dell'ultimo sottoinsieme sono obbligatoriaemente i restanti n_k elementi e abbiamo quindi un'unica scelta possibile. Complessivamente il numero delle possibili scelte è

$$\begin{aligned} & \frac{n!}{(n-n_1)!n_1!} \frac{(n-n_1)!}{(n-n_1-n_2)!n_2!} \dots \frac{(n-n_1-n_2\dots-n_{k-2})!}{(n-n_1-n_2-\dots-n_{k-1})!n_{k-1}!} \\ = & \frac{n!}{(n-n_1)!n_1!} \frac{(n-n_1)!}{(n-n_1-n_2)!n_2!} \dots \frac{(n-n_1-n_2\dots-n_{k-2})!}{n_k!n_{k-1}!} \\ = & \frac{n!}{n_1!n_2! \dots n_k!} \end{aligned}$$

Un esempio

Calcoliamo la probabilità che in un'aula con k studenti, almeno 2 compiano gli anni lo stesso giorno.

Indichiamo con N l'insieme dei 365 giorni dell'anno (per semplicità escludiamo il 29 febbraio). Numeriamo i giorni dell'anno da 1 a 365:

$$N = \{1, 2, 3, \dots, 365\}$$

Rappresentiamo i compleanni dei k studenti della classe con un k -pla ordinata di elementi di N e indichiamo con Ω l'insieme di tali k -ple ordinate:

$$\Omega = \{\omega = (\omega_1, \omega_2, \dots, \omega_k), \omega_i \in N\}.$$

L'insieme Ω ha $(365)^k$ elementi distinti.

Se supponiamo per semplicità che tutti i giorni dell'anno siano equiprobabili, ovvero che ci sia una distribuzione uniforme delle nascite (cosa di fatto non completamente vera), allora ogni elemento di Ω è equiprobabile e su Ω possiamo considerare la probabilità uniforme P . La probabilità di un generico sottoinsieme $A \subset \Omega$ è quindi

$$P(A) = \frac{\#A}{\#\Omega} = \frac{\#A}{(365)^k}$$

Vogliamo ora calcolare la probabilità dell'evento "almeno due persone compiono gli anni lo stesso giorno", rappresentato dall'insieme $A \subset \Omega$ delle k -ple ordinate $(\omega_1, \omega_2, \dots, \omega_k)$ in cui almeno due elementi sono uguali. Di fatto è più semplice calcolare la probabilità dell'evento complementare A^c e poi trovare la probabilità di A tramite la formula

$$P(A) = 1 - P(A^c).$$

L'evento complementare di A è l'evento "tutte le persone compiono gli anni in giorni diversi", rappresentato dall'insieme $A^c \subset \Omega$ delle k -ple ordinate $(\omega_1, \omega_2, \dots, \omega_k)$ in cui tutti gli elementi sono distinti. A^c è quindi l'insieme delle disposizioni di k elementi dell'insieme N ed il numero totale di tali disposizioni è $n!/(n-k)! = 365!/(365-k)!$. La probabilità di A^c è dunque

$$P(A^c) = \frac{365!}{(365-k)!(365)^k}$$

e la probabilità di A è quindi

$$P(A) = 1 - P(A^c) = 1 - \frac{365!}{(365-k)!(365)^k}$$

Ad esempio, per $k=23$ otteniamo $P(A) = 0.507 > 1/2$.

Variabili aleatorie

Spesso in un esperimento non si è interessati solamente ai possibili risultati, bensì ad una funzione del risultato. Immaginiamo ad esempio il lancio di due dadi e supponiamo che si vinca se la somma dei risultati sia maggiore o uguale a 7, mentre si perda altrimenti. In questo caso la variabile di interesse non è la coppia dei valori ottenuti nei due lanci quanto la loro somma.

Un altro esempio è il gioco in cui lanciando un dado se il risultato è strettamente maggiore di 3 allora si vince una somma pari a S , mentre se il risultato è minore o uguale di 3 si perde una somma pari a $2S$. Supponiamo che ci interessi sapere qual'è la somma che ci aspettiamo di vincere, o perdere, dopo 5 partite.

Questa classe di problemi può essere formalizzata introducendo il concetto di variabile aleatoria.

Definizione 54 Sia Ω uno spazio delle prove e P una probabilità sui sottoinsiemi di Ω . Si chiama *variabile aleatoria* una funzione $X : \Omega \rightarrow \mathbb{R}$.

Si usa distinguere tra *variabili aleatorie discrete*, ovvero variabili aleatorie X che possono assumere solo un numero discreto di valori, e *variabili aleatorie continue* ovvero variabili aleatorie che assumono valori all'interno di un insieme continuo.

Ad esempio nell'esperimento del lancio di un due dadi, scegliendo come variabile aleatoria X la somma dei risultati dei due lanci, abbiamo che X può assumere i valori 2,3,4,5,6,7,8,9,10,11,12 e dunque è una variabile aleatoria discreta.

Nell'esperimento della misura del decadimento di un nucleo radioattivo, scegliendo come variabile aleatoria X l'istante in cui il nucleo decade, abbiamo che, in linea di principio, X può assumere un qualsiasi valore all'interno dell'intervallo $[0, +\infty)$. X è dunque una variabile aleatoria continua.

Data una variabile aleatoria X , possiamo introdurre, dato un insieme $I \subset \mathbb{R}$, la probabilità che la variabile X assuma valori contenuti all'interno di A , ovvero

$$P(\{\omega \in \Omega, : X(\omega) \in I\})$$

L'applicazione, che ad un insieme $I \subset \mathbb{R}$ associa $P(\{\omega \in \Omega, : X(\omega) \in I\})$ è detta *legge* o *distribuzione* della variabile aleatoria X .

Si definisce inoltre la *funzione di distribuzione* $F_X : \mathbb{R} \rightarrow [0, 1]$, data da

$$F_X(t) := P(\{\omega \in \Omega : X(\omega) \leq t\}) = P(X \leq t)$$

per il suo significato, F è monotona crescente (in generale non strettamente). Si ha che

$$\lim_{t \rightarrow -\infty} F_X(t) = 0, \quad \lim_{t \rightarrow +\infty} F_X(t) = 1$$

inoltre

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

Densità di probabilità

Consideriamo una variabile aleatoria discreta X , che può quindi assumere solo un insieme discreto di valori $\{x_1, x_2, \dots\}$. La distribuzione di X è completamente specificata se conosciamo la probabilità degli insiemi

$$\{\omega \in \Omega : X(\omega) = x_i\}, \quad i = 1, 2, \dots$$

Indichiamo con $p(x_i) = P(\{\omega \in \Omega : X(\omega) = x_i\})$. la funzione p è detta *densità di probabilità discreta* e gode delle seguenti proprietà :

- $p(x_i) \geq 0$ per ogni i ,
- $\sum_i p(x_i) = 1$.

Nota la densità discreta p , allora per ogni $I \subseteq \mathbb{R}$, si ha che

$$P(X \in I) = P(\{\omega \in \Omega : X(\omega) \in I\}) = \sum_{x_i \in I} p(x_i).$$

Ad esempio, nel caso dell'esperimento del lancio di due dadi, scegliendo come variabile e aleatoria X la somma dei risultati dei due lanci, abbiamo che la distribuzione di X è data da:

$$P(X = 2) = P(\{(1, 1)\}) = \frac{1}{36} = p(2)$$

$$P(X = 3) = P(\{(1, 2), (2, 1)\}) = \frac{2}{36} = p(3)$$

$$P(X = 4) = P(\{(1, 3), (2, 2), (3, 1)\}) = \frac{3}{36} = p(4)$$

$$P(X = 5) = P(\{(1, 4), (2, 3), (3, 2), (4, 1)\}) = \frac{4}{36} = p(5)$$

$$P(X = 6) = P(\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}) = \frac{5}{36} = p(6)$$

$$P(X = 7) = P(\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}) = \frac{6}{36} = p(7)$$

$$P(X = 8) = P(\{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}) = \frac{5}{36} = p(8)$$

$$P(X = 9) = P(\{(3, 6), (4, 5), (5, 4), (6, 3)\}) = \frac{4}{36} = p(9)$$

$$\begin{aligned}
P(X = 10) &= P(\{(4, 6), (5, 5), (6, 4)\}) = \frac{3}{36} = p(10) \\
P(X = 11) &= P(\{(5, 6), (6, 5)\}) = \frac{2}{36} = p(11) \\
P(X = 12) &= P(\{(6, 6)\}) = \frac{1}{36} = p(12)
\end{aligned}
\tag{57}$$

Quindi, ad esempio,

$$P(4 \leq X \leq 6) = p(4) + p(5) + p(6) = \frac{3}{36} + \frac{4}{36} + \frac{5}{36} = \frac{12}{36}$$

Nel caso di una variabile aleatoria continua X , si ha che $P(X = x) = 0$ e si definisce la *densità di probabilità* f , come quella funzione $f : \mathbb{R} \rightarrow \mathbb{R}^+$ tale che

$$\begin{aligned}
F_X(t) &= \int_{-\infty}^t f(u) du \\
P(X \in [a, b]) &= \int_a^b f(u) du
\end{aligned}$$

Notiamo che, per il teorema fondamentale del calcolo integrale, si ha che:

$$f(t) = F'_X(t)$$

Nel caso dell'esperimento della misura del tempo di decadimento di un nucleo radioattivo si ha che

$$P(X \leq t) = 1 - e^{-\lambda t}, t > 0 \quad P(X \in [T_1, T_2]) = e^{-\lambda T_1} - e^{-\lambda T_2}.$$

dove λ è una costante positiva che caratterizza l'elemento.

La densità di probabilità f è data da

$$\begin{aligned}
f(t) &= \lambda e^{-\lambda t} & t \geq 0 \\
f(t) &= 0 & t < 0
\end{aligned}$$

Media e varianza

Nei problemi concreti, spesso non è conveniente e neppure necessario conoscere l'intera distribuzione della variabile aleatoria X , ma è sufficiente conoscere un numero finito di parametri, che indicano il modo in cui X è distribuita.

Media

Si definisce *media* o *speranza matematica* o *valore atteso* di una variabile aleatoria X , e si indica col simbolo $\mathbb{E}[X]$ o μ_X il numero

$$\mu_X = \mathbb{E}[X] := \begin{cases} \sum_i x_i p(x_i) & \text{per variabili aleatorie discrete} \\ \int_{-\infty}^{+\infty} x f(x) dx & \text{per variabili aleatorie continue} \end{cases}$$

nell'ipotesi la sommatoria (estesa a tutti i possibili valori che possono essere assunti dalla variabile aleatoria discreta) o l'integrale siano convergenti. In caso contrario la media non esiste.

Ad esempio, nel caso dell'esperimento del lancio di un due dadi, scegliendo come variabile aleatoria X la somma dei risultati dei due lanci, la media di X è data da

$$\begin{aligned} \mu_X = \mathbb{E}[X] &= 2p(2) + 3p(3) + 4p(4) + 5p(5) + 6p(6) + 7p(7) + 8p(8) + 9p(9) + \\ &\quad + 10p(10) + 11p(11) + 12p(12) \\ &= 2\frac{1}{36} + 3\frac{2}{36} + 4\frac{3}{36} + 5\frac{4}{36} + 6\frac{5}{36} + 7\frac{6}{36} + 8\frac{5}{36} + 9\frac{4}{36} + 10\frac{3}{36} + \\ &\quad + 11\frac{2}{36} + 12\frac{1}{36} \\ &= \frac{252}{36} = 7 \end{aligned}$$

Nel caso dell'esperimento della misura del tempo di decadimento di un nucleo radioattivo, dove la densità di probabilità f è data da

$$\begin{aligned} f(t) &= \lambda e^{-\lambda t} & t \geq 0 \\ f(t) &= 0 & t < 0 \end{aligned}$$

abbiamo che il valore medio è dato da

$$\mu_X = \mathbb{E}[X] = \int_0^{+\infty} t \lambda e^{-\lambda t} dt = \frac{1}{\lambda}$$

Varianza

La varianza di una variabile aleatoria X , indicata con il simbolo σ_X^2 , è definita come il valore atteso del quadrato dello scarto tra il valore della variabile X e al sua media μ_X :

$$\sigma_X^2 = \mathbb{E}[(X - \mu_X)^2].$$

La varianza fornisce quindi una misura della "dispersione" della variabile X rispetto alla sua media μ_X .

Per una variabile aleatoria discreta si calcola come:

$$\sigma_X^2 = \sum_i (x_i - \mu_X)^2 p(x_i),$$

mentre per una variabile aleatoria continua si calcola come

$$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x) dx,$$

(nell'ipotesi che la sommatoria o l'integrale convergono, in caso contrario la varianza non esiste).

Si dimostra facilmente una formula alternativa per il calcolo della varianza:

$$\sigma_X^2 = \mathbb{E}[X^2] - (\mu_X)^2$$

Per una variabile aleatoria discreta si calcola quindi come:

$$\sigma_X^2 = \sum_i x_i^2 p(x_i) - (\mu_X)^2,$$

mentre per una variabile aleatoria continua si calcola come

$$\sigma_X^2 = \int_{-\infty}^{+\infty} x^2 f(x) dx - (\mu_X)^2.$$

Ad esempio, nel caso dell'esperimento del lancio di un due dadi, scegliendo come variabile aleatoria X la somma dei risultati dei due lanci, abbiamo visto che la media μ_X vale $\mu_X = 7$. Per calcolare la varianza, calcoliamo prima $\mathbb{E}[X^2]$

$$\begin{aligned} \mathbb{E}[X^2] &= 2^2 p(2) + 3^2 p(3) + 4^2 p(4) + 5^2 p(5) + 6^2 p(6) + 7^2 p(7) + 8^2 p(8) + 9^2 p(9) + \\ &\quad + 10^2 p(10) + 11^2 p(11) + 12^2 p(12) \\ &= 4 \frac{1}{36} + 9 \frac{2}{36} + 16 \frac{3}{36} + 25 \frac{4}{36} + 36 \frac{5}{36} + 49 \frac{6}{36} + 64 \frac{5}{36} + 81 \frac{4}{36} + 100 \frac{3}{36} + \\ &\quad + 121 \frac{2}{36} + 144 \frac{1}{36} \\ &= \frac{1974}{36} = 54.8 \end{aligned}$$

Quindi

$$\sigma_X^2 = \mathbb{E}[X^2] - (\mu_X)^2 = 54.8 - (7)^2 = 5.8$$

Nel caso dell'esperimento della misura del tempo di decadimento di un nucleo radioattivo, dove la densità di probabilità f è data da

$$\begin{aligned} f(t) &= \lambda e^{-\lambda t} & t \geq 0 \\ f(t) &= 0 & t < 0 \end{aligned}$$

abbiamo che il valore medio è dato da $\mu_X = 1/\lambda$, mentre

$$\mathbb{E}[X^2] = \int_0^{+\infty} t^2 \lambda e^{-\lambda t} dt = \frac{2}{\lambda^2}.$$

Quindi la varianza è data da

$$\sigma_X^2 = \mathbb{E}[X^2] - (\mu_X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

La seguente diseuguaglianza illustra il significato della varianza

Teorema 44 (Diseuguaglianza di Chebyshev) *Per ogni $\eta > 0$ vale la diseuguaglianza*

$$P(|X - \mathbb{E}[X]| > \eta) \leq \frac{\sigma_X^2}{\eta^2}$$

Dimostrazione: Fissato il valore di $\eta \in \mathbb{R}$, consideriamo la variabile $Y : \Omega \rightarrow \mathbb{R}$ definita da:

$$Y(\omega) := \begin{cases} \eta^2 & \text{se } |X(\omega) - \mathbb{E}[X]| > \eta \\ 0 & \text{se } |X(\omega) - \mathbb{E}[X]| \leq \eta \end{cases}$$

Dato che per ogni $\omega \in \Omega$ abbiamo che $Y(\omega) \leq (X(\omega) - \mathbb{E}[X])^2$ allora

$$\mathbb{E}[Y] \leq \mathbb{E}[(X(\omega) - \mathbb{E}[X])^2] = \sigma_X^2$$

Inoltre $\mathbb{E}[Y] = \eta^2 P(|X - \mathbb{E}[X]| > \eta)$, da cui otteniamo la tesi. □

La legge dei grandi numeri

Consideriamo una variabile aleatoria X e una successione $\{X_n\}_n$ di variabile aleatorie. Si dice che la successione X_n converge in probabilità alla variabile aleatoria X se per ogni $\eta > 0$ fissato si ha :

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \eta) = 0$$

Tale concetto permette di formulare il seguente risultato

Teorema 45 (Legge dei grandi numeri) Sia $\{X_n\}_n$ una successione di variabile aleatorie indipendenti ed equidistribuite, con media μ e varianza σ^2 finita. Definita la successione $\{\bar{X}_n\}_n$, con

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

si ha che \bar{X}_n converge in probabilità a μ per $n \rightarrow \infty$.

Dimostrazione: La variabile aleatoria \bar{X}_n ha media μ e varianza pari a σ^2/n . Applicando la disuguaglianza di Chebyshev a \bar{X}_n abbiamo, per ogni $\eta \in \mathbb{R}$:

$$P(|\bar{X}_n - \mu| > \eta) \leq \frac{\sigma^2}{n\eta^2}$$

Mandando $n \rightarrow +\infty$ si ottiene che per ogni $\eta \in \mathbb{R}$:

$$\lim_{n \rightarrow +\infty} P(|\bar{X}_n - \mu| > \eta) = 0.$$

□

Modelli discreti

Legge binomiale

Supponiamo di lanciare n volte una moneta truccata, tale per cui la probabilità che esca testa è pari a p , con $p \in [0, 1]$, mentre la probabilità che esca croce è $1 - p$. Vogliamo calcolare la probabilità di ottenere k volte testa su n lanci.

Introduciamo lo spazio delle prove $\Omega := \{(\omega_1, \omega_2, \dots, \omega_n) : \omega_i = 0, 1, i = 1, \dots, n\}$, delle n -ple $(\omega_1, \omega_2, \dots, \omega_n)$, dove ω_i , con $i = 1, \dots, n$, può prendere solo i valori 0 e 1. Se $\omega_i = 1$ allora il risultato dell' i -esimo lancio è testa, mentre se $\omega_i = 0$ allora il risultato dell' i -esimo lancio è croce.

Introduciamo la variabile casuale discreta $X : \Omega \rightarrow \mathbb{R}$ che associa ad ogni elemento $(\omega_1, \omega_2, \dots, \omega_n) \in \Omega$ dello spazio delle prove, cioè ad ogni sequenza di risultati dei lanci, il numero delle teste:

$$X(\omega_1, \omega_2, \dots, \omega_n) = \omega_1 + \omega_2 + \dots + \omega_n.$$

Calcoliamo la probabilità $P(X = k)$, con $k = 0, \dots, n$. Nell'ipotesi che gli n lanci siano indipendenti, la probabilità di un evento semplice $(\omega_1, \omega_2, \dots, \omega_n)$ in cui ci sono k teste e $n - k$ croci è pari a $p^k(1 - p)^{n-k}$. La probabilità di ottenere quindi k

teste su n -lanci è pari a $p^k(1-p)^{n-k}$ moltiplicato per il numero di possibili $n-k$ ($\omega_1, \omega_2, \dots, \omega_n$) contenenti k volte 1 e $n-k$ volte 0, cioè $\frac{n!}{k!(n-k)!}$. Otteniamo quindi che la probabilità di ottenere k teste su n lanci è

$$P(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

Questo tipo di ragionamento può essere applicato ad ogni situazione in cui ci sono n prove indipendenti, ognuna delle quali ha solo due possibili risultati: il primo, che chiameremo "successo", si verifica con probabilità p , mentre il secondo, che chiameremo "insuccesso", si verifica con probabilità $1-p$. Introdotta la variabile casuale X che conta in numero dei "successi", si ha che la probabilità di ottenere k successi con n prove è $P(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$.

La densità di probabilità discreta data da

$$p(k) = \begin{cases} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} & k = 0, 1, \dots, n \\ 0 & \text{altrimenti} \end{cases}$$

è detta *legge binomiale*. La media è data da

$$\mathbb{E}[X] = \sum_{k=0}^n k p(k) = \sum_{k=0}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=0}^n \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} = np$$

mentre la varianza $\sigma_X^2 = np(1-p)$.

Esempio 44 *I bulloni prodotti da una ditta sono difettosi con una probabilità del 20%. Vengono messi in commercio in confezioni di 3 pezzi. Qual è la probabilità che in una confezione vi sia al più un bullone difettoso?*

Utilizziamo lo schema successo-insuccesso, con $n = 3$ esperimenti, dove la probabilità di "successo" (=trovare un bullone difettoso) è pari a $p = 0.2$. La probabilità di trovare al più un bullone difettoso è uguale alla somma della probabilità di trovare 0 bulloni difettosi e della probabilità di trovare 1 bullone difettoso, date da

$$\frac{3!}{0!3!} (0.2)^0 (0.8)^3 + \frac{3!}{1!2!} (0.2)^1 (0.8)^2 = 0.896$$

Esempio 45 *Tenendo conto del fatto che circa il 30% dei passeggeri che ha acquistato un biglietto aereo non si presenta alla partenza, una compagnia aerea accetta 28 prenotazioni per un volo con 24 posti. Qual è la probabilità che almeno un passeggero resti a terra?*

Utilizziamo lo schema "successo" (=passeggero si presenta alla partenza)- "insuccesso" (=passeggero non si presenta alla partenza) su 28 "prove" indipendenti, nell'ipotesi semplificativa che le scelte dei passeggeri siano indipendenti. In tal caso la probabilità che si presentino k passeggeri è data dalla legge binomiale $\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$ con $n = 28$ e $p = 0.7$. La probabilità di avere $k \geq 25$ è quindi:

$$\begin{aligned} & p(25) + p(26) + p(27) + p(28) \\ &= \frac{28!}{25!(28-25)!}0.7^{25}0.3^3 + \frac{28!}{26!(28-26)!}0.7^{26}0.3^2 + \frac{28!}{27!(28-27)!}0.7^{27}0.3^1 \\ &+ \frac{28!}{28!(28-28)!}0.7^{28}0.3^0 = 0.0157 \end{aligned}$$

Tempo d'attesa

Riprendiamo l'esempio che ha portato alla costruzione della legge binomiale. Consideriamo una moneta truccata, tale per cui la probabilità che esca testa è pari a p , con $p \in [0, 1]$, mentre la probabilità che esca croce è pari a $1 - p$. Supponiamo che venga lanciata più volte, fino a quando non si ottiene testa e vogliamo calcolare la probabilità che servano esattamente k lanci, con $k \geq 1$.

Introduciamo la variabile aleatoria discreta T , data dal numero dei lanci necessari per ottenere testa. T può assumere tutti i valori interi positivi maggiori o uguali a 1. La densità di probabilità di T è data da:

$$p(k) = P(T = k) = (1 - p)^{k-1}p, \quad k = 1, 2, \dots, +\infty,$$

calcolata usando l'indipendenza dei vari lanci. Infatti la probabilità che si ottenga testa esattamente al k -esimo lancio è data dal prodotto della probabilità che ai primi $k - 1$ lanci esca croce, pari a $(1 - p)^{k-1}$, per la probabilità che al k -esimo lancio esca testa, pari a p . Notiamo che

$$\sum_{k=1}^{\infty} (1 - p)^{k-1}p = p \sum_{k=1}^{\infty} (1 - p)^{k-1} = p \sum_{k=0}^{\infty} (1 - p)^k = p \frac{1}{1 - (1 - p)} = 1$$

Più in generale, T rappresenta il tempo di attesa del primo "successo" nel modello "successo-insuccesso" descritto dalla legge binomiale.

Notiamo che per ottenere la densità di probabilità $p(k)$ potevamo ragionare anche nel seguente modo. La probabilità che per ottenere un successo servano più di $k - 1$ lanci può essere ottenuta come la somma della probabilità che per ottenere un successo servano più di k lanci e della probabilità che al k -esimo lancio si ottenga un successo:

$$P(T > k - 1) = P(T = k) + P(T > k).$$

D'altra parte $P(T > k)$ è uguale alla probabilità di non avere successi nei primi k lanci, cioè $(1 - p)^k$ e, analogamente, $P(T > k - 1) = (1 - p)^{k-1}$. Abbiamo quindi

$$P(T = k) = P(T > k - 1) - P(T > k) = (1 - p)^{k-1} - (1 - p)^k = (1 - p)^{k-1}p$$

La media e la varianza di tale variabile è data da:

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k(1 - p)^{k-1}p = \frac{1}{p}, \quad \sigma_X^2 = \frac{1 - p}{p^2}$$

La distribuzione di Poisson

Consideriamo la variabile aleatoria X descritta dalla legge binomiale, associata al numero di successi in n esperimenti indipendenti:

$$p(k) = \begin{cases} \frac{n!}{k!(n-k)!} p^k (1 - p)^{n-k} & k = 0, 1, \dots, n \\ 0 & \text{altrimenti} \end{cases}$$

+ e studiamo il limite in cui $n \rightarrow \infty$ e $p \rightarrow 0$, mantenendo però costante il valore medio $\mathbb{E}[X] = np \equiv \lambda$, con $\lambda \in \mathbb{R}$, $\lambda > 0$. Ponendo quindi $p = \lambda/n$, si ha che

$$\begin{aligned} \lim_{n \rightarrow \infty} p(k) &= \lim_{n \rightarrow \infty} \frac{n!}{k!(n-k)!} p^k (1 - p)^{n-k} = \lim_{n \rightarrow \infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \lim_{n \rightarrow \infty} \frac{\lambda^k}{k!} \frac{n!}{n^k(n-k)!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &= \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

La densità di probabilità discreta

$$p(k) = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda} & k = 0, 1, \dots, \infty \\ 0 & \text{altrimenti} \end{cases}$$

è detta distribuzione di Poisson di parametro $\lambda > 0$. La media e la varianza sono date da:

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda, \quad \sigma_X^2 = \lambda$$

Modelli continui

Legge esponenziale

Si consideri la funzione di distribuzione di variabile aleatoria continua T descritta dalla funzione di distribuzione

$$F_T(t) = P(T \leq t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

dove $\lambda \in \mathbb{R}$, $\lambda > 0$. La densità di probabilità è la funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ data da

$$f(t) = \frac{d}{dt} F_T(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

La media e la varianza sono date da:

$$\mathbb{E}[T] = \int_0^{\infty} \lambda t e^{-\lambda t} dt = \frac{1}{\lambda}, \quad \sigma_T^2 = \frac{1}{\lambda^2}$$

La variabile casuale con questa distribuzione è adatta a descrivere l'istante in cui un nucleo radioattivo decade, oppure il momento in cui un certo oggetto si rompe. Per questo la costante $\tau := \frac{1}{\lambda}$ è anche detta "vita media".

Legge normale

Fissate due costanti $\mu \in \mathbb{R}$ e $\sigma \in \mathbb{R}^+$, consideriamo la variabile casuale continua X descritta dalla densità di probabilità

$$f(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}, \quad x \in \mathbb{R},$$

detta densità di probabilità Gaussiana (o normale) di parametri μ e σ . La media e la varianza sono date da

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx = \mu, \quad \mathbb{E}[(x-\mu)^2] = \int_{-\infty}^{+\infty} (x-\mu)^2 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx = \sigma^2$$

Inoltre si ha che

$$\int_{\mu-\sigma}^{\mu+\sigma} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx = 0,68 \quad \int_{\mu-2\sigma}^{\mu+2\sigma} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx = 0,95 \quad \int_{\mu-3\sigma}^{\mu+3\sigma} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx = 0,98$$

La legge dei grandi numeri

Variabili aleatorie congiunte

Due variabili aleatorie $X, Y : \Omega \rightarrow \mathbb{R}$ definite sullo stesso spazio Ω vengono dette variabili aleatorie congiunte.

Ad esempio consideriamo l'esperimento del lancio di due dadi (non truccati) e indichiamo con X la somma dei risultati e con Y il valore assoluto della differenza.

Consideriamo l'esperimento che consiste nello scegliere a caso un individuo dalla popolazione di una città e nel misurarne l'altezza X e il peso Y . In tal caso X e Y sono due variabili aleatorie congiunte in quanto definite sullo stesso insieme Ω , l'insieme degli individui della città.

Se le due variabili X e Y sono entrambe discrete, allora ha senso calcolare la probabilità che assumano rispettivamente dei particolari valori x_i e y_j . Si definisce *densità di probabilità congiunta* di X, Y la probabilità

$$P(X = x_i, Y = y_j) \equiv p_{X,Y}(x_i, y_j)$$

che gode delle seguenti proprietà :

- $\sum_{x_i, y_j} p_{X,Y}(x_i, y_j)$.
- Sommando la densità congiunta su tutti i possibili valori di una variabile casuale, si ottiene la densità di probabilità dell'altra variabile, detta anche *densità marginale*:

$$\sum_{y_j} p_{X,Y}(x_i, y_j) = p_X(x_i),$$

$$\sum_{x_i} p_{X,Y}(x_i, y_j) = p_Y(y_j).$$

$p_X(x_i)$ ha il significato di probabilità che la variabile X assuma il valore x_i indipendentemente dal valore assunto dalla variabile Y . Analogamente $p_Y(y_j)$ ha il significato di probabilità che la variabile Y assuma il valore y_j indipendentemente dal valore assunto dalla variabile X .

Nel caso in cui le due variabili aleatorie congiunte sono entrambe continue, allora si definisce la funzione *densità di probabilità congiunta* $f_{X,Y}$ come quella funzione $f_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che

$$P(a \leq X \leq b, c \leq Y \leq d) = \int \int_{[a,b] \times [c,d]} f_{X,Y}(x, y) dx dy$$

Come nel caso di variabili aleatorie discrete, la funzione densità di probabilità congiunta soddisfa le seguenti proprietà

- $\int \int_{\mathbb{R}^2} f_{X,Y}(x, y) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx dy = 1.$
- Integrando la densità congiunta su tutti i possibili valori di una variabile casuale, si ottiene la densità di probabilità dell'altra variabile, detta anche *densità marginale*:

$$\int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy = f_X(x),$$

$$\int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx = f_Y(y).$$

$f_X(x)$ ha il significato di densità di probabilità della variabile X indipendentemente dal valore assunto dalla variabile Y , cioè:

$$P(a \leq X \leq b, Y \in \mathbb{R}) = \int_a^b f_X(x) dx.$$

Analogamente $f_Y(y)$ ha il significato di densità di probabilità della variabile Y indipendentemente dal valore assunto dalla variabile X , cioè:

$$P(X \in \mathbb{R}, Y \in [c, d]) = \int_c^d f_Y(y) dy.$$

Due variabili aleatorie congiunte X, Y sono dette *indipendenti* se:

$$p_{X,Y}(x_i, y_j) = p_X(x_i)p_Y(y_j), \quad \text{se } X \text{ e } Y \text{ sono variabili aleatorie discrete}$$

$$f_{X,Y}(x, y) = f_X(x)f_Y(y), \quad \text{se } X \text{ e } Y \text{ sono variabili aleatorie continue}$$