Analisi II

Samuele Musiani

February 20, 2023 - June 22, 2023

Contents

1 Introduzione agli appunti

Potrebbero esserci degli errori. In generale ho cercato di minimizzare gli errori rileggendo svariate volte, ma purtroppo non credo di essere riuscito nell'intento miracoloso di correggerli tutti. **Non mi assumo nessuna diretta responsabilità riguardo eventuali informazioni errate presenti all'interno di questi appunti**.

1.1 Le varie parti degli appunti

Ci sono vari box colorati che introducono rispettivamente una definizione, una dimostrazione, un lemma e una cosa importante:

Definizione

Questa è una definizione

Dimostrazione

Questa è una dimostrazione

Lemma

Questo è un lemma

Questa è una cosa importante

1.2 Segnalare errori e contribuire

Se per caso si trovano errori **vi prego di segnalarli ad uno dei contatti che seguono**. Se li segnalerete eviterete che altri studenti, nella lettura di questi appunti, imparino una informazione errata. Vi prego quindi di segnalarlo, ci mettete veramente poco tempo. Per quanto mi riguarda farò in modo di fixare quello che mi verrà riportato il prima possibile.

Se volete contribuire agli appunti con aggiunte, modifiche o suggerimenti potete scrivermi o mandare direttamente una *pull request* alla repository (link) sul mio profilo GitHub.

Email: samuele.musiani@studio.unibo.it

Telegram: @Tastier

Github: musianisamuele

Ultima modifica: **June 22, 2023**

2 Integrali

2.1 Calcolo dell'area sottesa ad una curva

Data una funzione $f : [a, b] \to \mathbb{R}$ con $f(x) \geq 0 \; \forall x \in [a, b]$. IL suo **sottografico** è:

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

Il sottografico è quindi un insieme e corrisponde e tutti i punti che soddisfano per le ordinate le disuguaglianza $a \leq x \leq b$ e per le ascisse $0 \leq y \leq f(x)$.

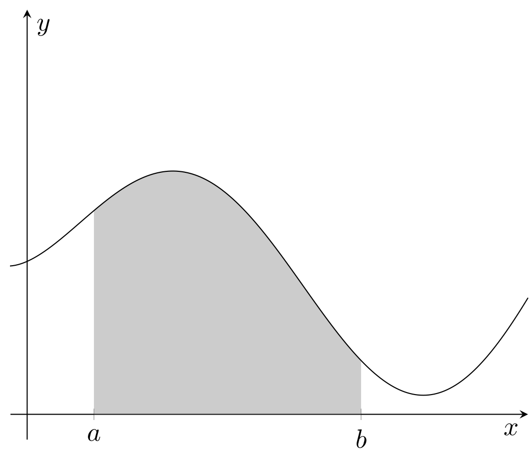

Figure 1: Area del sottografico rappresentata in grigio

Come si calcola il sottografico?

2.2 Somme di Riemann

2.2.1 Scomposizione di un intervallo

Sia dato un interallo $[a, b] \subseteq \mathbb{R}$ lo divido in $n \in \mathbb{N}$ intervalli uguali:

$$\begin{aligned} x_0 &= a \\ x_1 &= x_0 + \frac{b-a}{n} = a + \frac{b-a}{n} \\ x_2 &= x_1 + \frac{b-a}{n} = a + 2 \cdot \frac{b-a}{n} \\ &\vdots \\ x_k &= a + k \cdot \frac{b-a}{n} \end{aligned}$$

Il primo punto corrisponde all'inizio dell'intervallo $x_0 = a$, l'ultimo punto alla fine dell'intervallo $x_n = b$ in quanto:

$$x_n = a + n \cdot \frac{b-a}{n} = a + b - a = b$$

Posso inoltre scegliere dei punti all'interno di questi intervalli:

$$\forall k \in \mathbb{N} : 0 < k \leq n \quad \text{scelgo} \quad \xi_k \in [x_{k-1}, x_k]$$

È importante notare alcune cose:

1. ξ_k è un semplicissimo punto, lo si indica con la lettera greca ξ (xi) per evitare di far confuzione successivamente.
2. La scelta della posizione del punto ξ_k è **totalmente arbitraria**. Può quindi essere il punto medio, coincidere con un estremo o essere completamente casuale purché rispetti la condizione imposta, cioè: $\xi_k \in [x_{k-1}, x_k]$
3. Abbiamo una serie di punti, non solo 1, in quanto questo vale *per ogni k*.

$$\begin{aligned}
\xi_1 &\in [x_0, x_1] = \left[a, a + \frac{b-a}{n}\right] \\
\xi_2 &\in [x_1, x_2] = \left[a + \frac{b-a}{n}, a + 2 \cdot \frac{b-a}{n}\right] \\
&\vdots \\
\xi_n &\in [x_{n-1}, x_n] = \left[a + (n-1) \cdot \frac{b-a}{n}, b\right]
\end{aligned}$$

Definizione

Sia $f : [a, b] \to \mathbb{R}$, continua su $[a, b]$. Sia inoltre $n \in \mathbb{N}$ e siano $x_0, x_1, \cdots x_n$ e $\xi_1, \xi_2, \cdots \xi_n$ i punti introdotti precedentemente. Si definisce **La somma di Riemann *n-esima*** è il numero:

$$S_n = \sum_{k=1}^{n} f(\xi_k) \cdot (x_k - x_{k-1})$$

Notando che il termine $(x_k - x_{k-1})$ è sempre uguale si può riscrivere la somma come segue:

$$S_n = \frac{b-a}{n} \cdot \sum_{k=1}^{n} f(\xi_k)$$

La somma dipende dalla scelta dei punti ξ_k, non è quindi sempre la stessa. Rappresenta inoltre la somma delle aree dei rettangoli che approssimano il sottografico della funzione f nell'intervallo $[a, b]$.

2.3 Integrale dalle somme di Riemann

Teorema

Sia $f : [a, b] \to \mathbb{R}$ continua su $[a, b]$ e $(S_n)_{n \in \mathbb{N}}$ una famiglia di somme di Riemann[a].

$$\exists \lim_{n \to +\infty} S_n \in \mathbb{R}$$

Inoltre **il valore NON dipende dalla scelta dei punti** ξ_k. Tale limite si chiama **integrale di** f:

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{n \to +\infty} S_n \in \mathbb{R}$$

[a] Si noti che questa famiglia in realtà è una successione

Note importanti:

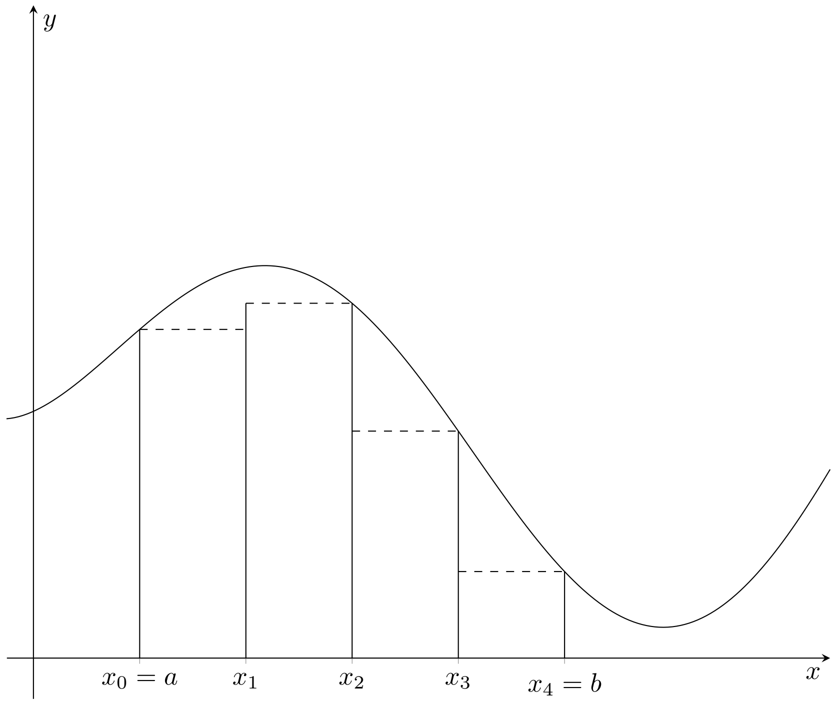

Figure 2: Somme di Riemann

- Il valore del limite, essendo in $\mathbb{R}$ è finito.
- a e b si chiamo **estremi di integrazione**.
- La varibile dentro la funzione è una *variabile muta*. Non indica effettivamente nulla. Le seguenti notazioni sono equivalenti:

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b f(t)\,\mathrm{d}t = \int_a^b f$$

  Si adotta generalmente la variabile muta è il termine $\mathrm{d}x$ che NON ha alcune definzione o qualsivoglia introduzione matematica per comodità.

Vediamo come è definita questa somma in alcuni esempi particolari:

- $f(x) \leq 0 \ \ \forall x \in [a, b]$ come in figura 3:

$$\int_a^b f(x)\,\mathrm{d}x = -(\text{area del sottografico}) = -textarea(A)$$

- Se la funzione assume sia valori positivi sia valori negativi (come in figura 4) allora l'intagrale risulta:

$$\int_a^b f(x)\,\mathrm{d}x = \text{area}(A_1) - \text{area}A_2$$

- Se il sottografico è formato da due aree positive ma separate (come in figura 5):

$$\int_a^b f(x)\,\mathrm{d}x = \text{area}(A_1) + \text{area}A_2$$

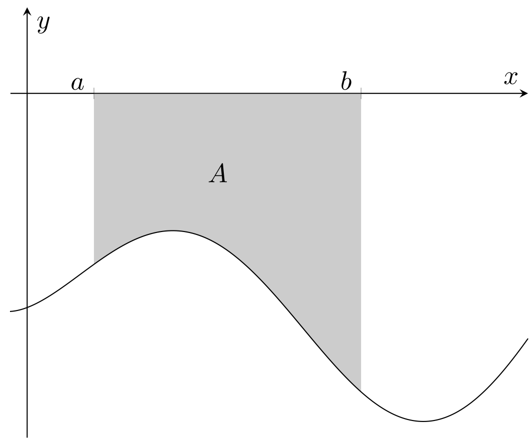

Figure 3: Visualizzazione di un integrale su una funzione negativa

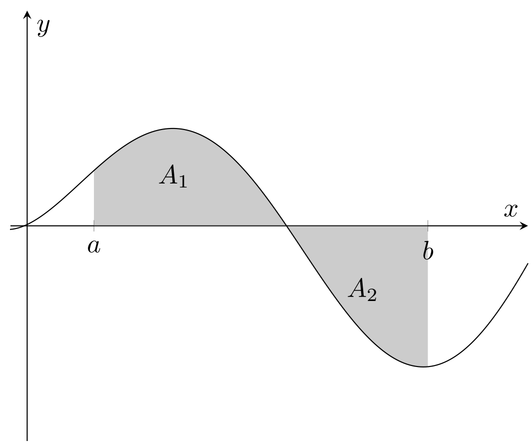

Figure 4: Visualizzazione di un integrale su una funzione che cambia segno

- Se $f(x) = k \;\; \forall x \in [a, b]$, cioè in pratica la funzione è costante (come in figura 6). Consideriamo prima la somma di Riemann:
$$S_n = \sum_{i=0}^{n} f(\xi_i) \cdot \left(\frac{b-a}{n}\right) = \sum_{i=0}^{n} k \cdot \left(\frac{b-a}{n}\right) = n \cdot k \cdot \frac{b-a}{n} = k \cdot (b-a)$$
Ne consegue che:
$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to+\infty} S_n = k \cdot (b-a)$$
Coindice con infatti l'area di un rettangolo di base $b-a$ e di altezza k.
- Se $a = b$ allora:
$$\int_a^a f(x)\,\mathrm{d}x = 0$$
In quanto:
$$S_n = \sum_{i=0}^{n} f(\xi_i) \cdot \left(\frac{b-a}{n}\right) = \sum_{i=0}^{n} f(\xi_i) \cdot \left(\frac{a-a}{n}\right) = \sum_{i=0}^{n} f(\xi_i) \cdot 0 = 0$$

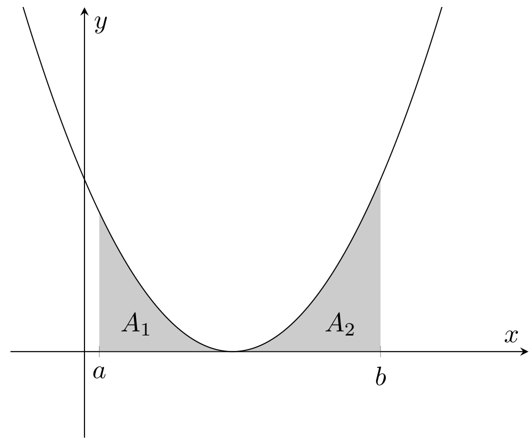

Figure 5: Visualizzazione di un integrale su una funzione che tocca l'asse delle ascisse

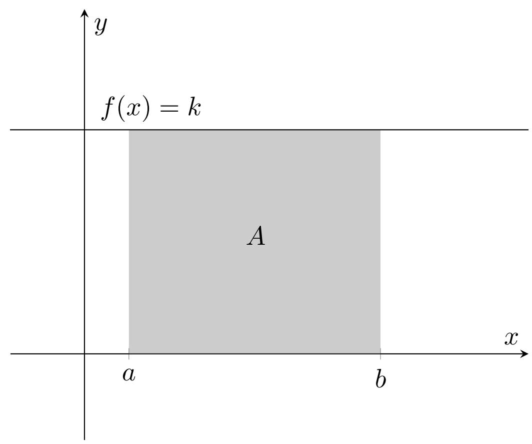

Figure 6: Visualizzazione di un integrale su una funzione costante

2.3.1 Proprietà dell'integrale

1. **Linearità**: Date due funzioni $f, g : [a, b] \to \mathbb{R}$ continue su $[a, b]$. Deti inoltre due punti $c_1, c_2 \in \mathbb{R}$

$$\int_a^b (c_1 \cdot f(x) + c_2 \cdot g(x)) \, \mathrm{d}x = \int_a^b c_1 \cdot f(x) \, \mathrm{d}x \int_a^b c_2 \cdot g(x) \, \mathrm{d}x =$$

2. **Additività** Data una funzione $f : [a, b] \to \mathbb{R}$ e dato $c \in [a, b]$:

$$\int_a^b f(x) \, \mathrm{d}x = \int_a^c f(x) \, \mathrm{d}x + \int_c^b f(x) \, \mathrm{d}x$$

In generale si tende ad adottare la **convenzione** che se $b < a$:

$$\int_a^b f(x) \, \mathrm{d}x = -\int_b^a f(x) \, \mathrm{d}x$$

Ne consegue quindi che si può **generalizzare la seconda proprietà** a: $f : \mathbb{R} \to \mathbb{R}$, $\forall a, b, c \in \mathbb{R}$

implica:

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$$

3. Con $f : [a,b] \to \mathbb{R}$ e $f(x) \geq 0 \ \ \forall x \in [a,b]$ allora:

$$\int_a^b f(x)\,\mathrm{d}x \geq 0$$

Si può **generalizzare**[1]: data un'ulteriore funzione $g : [a,b] \to \mathbb{R}$, se $f(x) \leq g(x) \quad \forall x \in [a,b]$ allora:

$$\int_a^b f(x)\,\mathrm{d}x \geq \int_a^b g(x)\,\mathrm{d}x$$

2.4 Media integrale

Teorema

Se $f : [a,b] \to \mathbb{R}$ continua su $[a,b]$, allora:

$$\exists c \in [a,b] : \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x = f(c)$$

Il valore di $f(c)$ si definisce **media integrale di f in** $[a,b]$.

Il nome media deriva dal datto che:

$$\frac{1}{b-a}S_n = \sum_{k=1}^{n} \frac{f(\xi_k)}{n}$$

Che non è altro che una media aritmetica.

Dimostrazione

Dimostraimo il teorema della media integrale: per il teorema di Weierstrass $\exists x_1, x_2$ punti di minimo e massimo assoluti:

$$f(x_1) \leq f(x) \leq f(x_2) \quad \forall x \in [a,b]$$

Usando la proprietà della monotonia degli integrali:

$$\int_a^b f(x_1)\,\mathrm{d}x \leq \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b f(x_2)\,\mathrm{d}x$$

Ed essendo $f(x_1)$ e $f(x_2)$ valori costanti:

$$(b-a)f(x_1) \leq \int_a^b f(x)\,\mathrm{d}x \leq (b-a)f(x_2)$$

Che quindi dividendo per $(b-a)$

$$f(x_1) \leq \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x \leq f(x_2)$$

[1] Questo si dimostra con il caso non generalizzato e una funzione ausiliaria $h(x) = f(x) - g(x)$

Dal teorema dei valori intermedi[a]:

$$\exists x \in [a, b] : f(x) = y$$

Cioè:

$$f(c) = \frac{1}{b-a} \int_a^b f(x) \, \mathrm{d}x$$

Qed.

[a]Non mi ricordo di averlo fatto ne di averlo copiato. È sugli appunti del 2023-02-20. CONTROLLARE! :)

2.5 Primitiva

Definizione

Sia $f :]a, b[\to \mathbb{R}$. Una funzione $F :]a, b[\to \mathbb{R}$ si dice **primitiva di f su** $]a, b[$ se vale:

$$F'(x) = f(x) \qquad \forall x \in]a, b[$$

Osservazione: Se F è primitiva di f su $]a, b[$, allora:

$$\forall k \in \mathbb{R} \quad F(x) + k \text{ è una primitiva di } f \text{ su }]a, b[$$

Ci sono infinite primitive di una funzione

Teorema

Teorema di caratterizzazione delle primitive su un intervallo: Se $F :]a, b[\to \mathbb{R}$ e $G :]a, b[\to \mathbb{R}$ sono primitive di $f :]a, b[\to \mathbb{R}$, allora:

$$\exists k \in \mathbb{R} : F(x) - G(x) = k \qquad \forall x \in]a, b[$$

Dimostrazione

Vogliamo dimostrare il teorema appena enunciato. Consideriamo la funzione ausiliaria:

$$H(x) = F(x) - G(x) \qquad \forall x \in]a, b[$$

Se faccio la derivata:

$$H'(x) = F'(x) - G'(x) \qquad \forall x \in]a, b[$$

Che per ipotesi:

$$H'(x) = f(x) - f(x) = 0 \qquad \forall x \in]a, b[$$

Da teorema di Lagrange (in particolare dal suo corollario), se una funzione continua su un intervallo ha derivata sempre nulla allora la funzione è costante. Ne consegue:

$$\exists k \in \mathbb{R} : H(x) = k \qquad \forall x \in]a, b[$$

E quindi:

$$F(x) - G(x) = k \qquad \forall x \in]a, b[$$

Qed.

È importante che la funzione sia definita su un intervallo perché è facile prendere un esempio di funzione non definita su un intervallo e fare vedere che il teorema non funziona. Per sempio presa la funzione $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}, f(x) = \frac{1}{x^2}$ ha come primitive:

$$F(x) = -\frac{1}{x} \qquad \forall x \neq 0$$

$$G(x) = \begin{cases} -\dfrac{1}{x} & x < 0 \\ -\dfrac{1}{x} + 1 & x > 0 \end{cases}$$

Nonostante ciò:

$$F(x) - G(x) \text{ non è costante}$$

Proprio perché non stiamo considerando un intervallo.

2.6 Funzioni integrali

Definizione

Sia $f :]a, b[\to \mathbb{R}$ continua e sia $c \in]a, b[$. Definziamo la **funzione integrale di f con punto base c** come:

$$I_c :]a, b[\to \mathbb{R}, \qquad I_c(x) = \int_c^x f(t)\,\mathrm{d}t$$

Osservazione 1:

$$I_c(c) = \int_c^c f(t)\,\mathrm{d}t = 0$$

Osservazione 2: Consideriamo la funzione $f :]a, b[\to \mathbb{R}$ continua e un punto $c_1, c_2 \in]a, b[$. Le funzioni:

$$I_{c_1}(x) = \int_{c_1}^x f \qquad \text{e} \qquad I_{c_2}(x) = \int_{c_2}^x f$$

Hanno differenza costante:

$$I_{c_1}(x) - I_{c_2}(x) = \int_{c_1}^x f - \int_{c_2}^x f = \int_{c_1}^x f + \int_x^{c_2} f = \int_{c_1}^{c_2} f$$

2.6.1 Teoremi fondamentali del calcolo integrale

Il seguente teorema è spesso indicato come secondo, ma il prof ha deciso di farlo per primo.

Teorema

Sia $f :]a, b[\to \mathbb{R}$ continua e sia $c \in]a, b[$. Allora vale:

$$I_c'(x) = f(x) \qquad \forall x \in]a, b[$$

Cioè:

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_c^x f(t)\,\mathrm{d}t = f(x) \qquad \forall x \in]a, b[$$

Esempio: Data $f : \mathbb{R} \to \mathbb{R}, f(x) = e^{x^2}$ se

$$I(x) = \int_0^x e^{t^2}\,\mathrm{d}t$$

Allora:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^x e^{t^2}\,\mathrm{d}t = e^{x^2} \qquad \forall x \in \mathbb{R}$$

Oppure un altro esempio:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_2^x \sqrt{1+t^2}\cdot e^{-t}\sin(t)\,\mathrm{d}t = \sqrt{1+x^2}\cdot e^{-x}\sin(x)$$

Non importa quindi sapere il risultato dell'integrale per sapere la sua derivata.

Perché l'estremo inferiore dell'integrale è ininfluente[2]? Se sostiutiamo l'estremo inferiore a con un qualsiasi $k \in \mathbb{R}$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t = \frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^k f(t)\,\mathrm{d}t + \int_k^x f(t)\,\mathrm{d}t\right) = \frac{\mathrm{d}}{\mathrm{d}x}\int_a^k f(t)\,\mathrm{d}t + \frac{\mathrm{d}}{\mathrm{d}x}\int_k^x f(t)\,\mathrm{d}t$$

L'integrale $\int_a^k f$ è una semplice costante, ne consegue che dopo la derivata si annulla, quindi rimane solo l'integrale:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_k^x f(t)\,\mathrm{d}t = f(x)$$

Ricordiamoci che abbiamo scelto a caso la costante k, ne consegue che è ininfluente sotto l'effetto della derivata.

Dimostrazione

Vogliamo dimostrare il teorema appena enunciato. Supponiamo di avere una funzione $f :]a, b[\to \mathbb{R}$ e un punto $c \in]a, b[$. Vogliamo dimostrare che:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_c^x f(t)\,\mathrm{d}t = f(x) \qquad \forall x \in]a, b[$$

Per la defինzione di derivata dovremmo dimostrare che esiste sia il limite destro che il sinistro e che il loro valore coincide. Ci limitiamo a dimostrare il teorema per il limite destro in quanto è analogo nell'altro caso. Calcoliamo il limite:

$$\lim_{h\to 0^+}\frac{\displaystyle\int_c^{x+h} f - \int_c^x f}{h} = \lim_{h\to 0^+}\frac{1}{h}\left(\int_c^{x+h} f + \int_x^c f\right)$$

Possiamo quindi usare la proprietà di additività degli integrali e ridurci a dimostrare:

$$\lim_{x\to 0^+}\frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t = f(x)$$

Usiamo ora il teorema della media integrale, dove in particolare gli estremi $a = x$ e $b = x + h$ e quindi $b - a = x + h - x = h$:

$$\exists c(h) \in \ ]x, x+h[\ : \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t = f(c(h))$$

[2]Questa è una mia considerazione, il prof non l'ha detta

In questo caso abbiamo scritto $c(h)$ e non semplicemente c per sottolianere che il punto dipende da h. Essendo però $c(h) \in \,]x, x+h[$ diventa che:

$$x < c(h) < x + h$$

Facendo il limite per $h \to 0^+$ e dal teorema del confronto:

$$c(h) \xrightarrow[h \to 0^+]{} x$$

E quindi:

$$f(c(h)) \xrightarrow[h \to 0^+]{} f(x)$$

Qed.

Il seguente viene spesso indicato come il primo teorema fondamentale del calcolo integrale.

Teorema

Sia $f :]a_0, b_0[\to \mathbb{R}$ continua. Sia inoltre $F :]a_0, b_0[\to \mathbb{R}$ la primitiva di f su $]a_0, b_0[$, allora:

$$\forall [a, b] \subseteq]a_0, b_0[\quad \int_a^b f(x)\,\mathrm{d}x = F(b) - F(a) = [F(x)]_a^b = F(x)|_{x=a}^{x=b}$$

Esempio: $f(x) = x^2$ e $F(x) = \dfrac{x^3}{3}$ è primitiva di f.

$$\int_2^3 x^2\,\mathrm{d}x = \left[\frac{x^3}{3}\right]_{x=2}^{x=3} = \frac{3^3}{3} - \frac{2^3}{3} = \frac{19}{3}$$

Dimostrazione

Vogliamo dimostrare il teorema appena enunciato. Prendiamo una funzione $f :]a_0, b_0[\to \mathbb{R}$ continua. Assumiamo inoltre che $F :]a_0, b_0[\to \mathbb{R}$ sia la primitiva di f su $]a_0, b_0[$. Prendiamo ora un punto $c \in]a_0, b_0[$. La funzione integrale di f con punto base c è definita come segue:

$$I_c(x) = \int_c^x f(t)\,\mathrm{d}t \qquad \forall x \in]a_0, b_0[$$

Per il "secondo"[a] teorema del calcolo integrale vale che:

$$I_c'(x) = f(x) \qquad \forall x \in]a_0, b_0[$$

Abbiamo quindi due primitive di f in $]a_0, b_0[$: la prima è F per ipotesi e la seconda è I_c per quanto appena dimostrato. Per il teorema di caratterizzazione delle derivate (Sezione: 2.5):

$$\exists k \in \mathbb{R} : F(x) = I_c(x) + k \qquad \forall x \in]a_0, b_0[$$

Ricordiamoci che vogliamo dimostrare che $\forall [a, b] \subseteq]a_0, b_0[$:

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a) = I_c(b) + k - I_c(a) - k = I_c(b) - I_c(a) =$$

$$= \int_c^b f(t)\,\mathrm{d}t - \int_c^a f(t)\,\mathrm{d}t = \int_c^b f(t)\,\mathrm{d}t + \int_a^c f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t$$

Ed essendo la variabile interna all'integrale una varibile muta (cioè è indifferente come la chiamiamo). Qed.

[a] Generalmente nei libri è indicato come secondo, ma il prof lo ha fatto per primo quindi è il primo teorema che si trova in questa sezione.

Generalizziamo ora il teorema fondamentale del calcolo integrarle:

Teorema

Siano $A, I \subseteq \mathbb{R}$ intervalli aperti. Sia inoltre $f : A \to \mathbb{R}$ continua e $h : I \to A$ derivabile con derivata continua. Se $c \in A$ allora:

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_c^{h(x)} f(t)\,\mathrm{d}t = f(h(x)) \cdot h'(x)$$

Esempio:

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_0^{\sin(x)} e^{t^2}\,\mathrm{d}t = e^{\sin^2(x)} \cdot \cos(x)$$

Dimostrazione

Dimostraimo il teorema fondamentale del calcolo generalizzato. Chiamiamo in particolare:

$$G(x) = \int_c^{h(x)} f(t)\,\mathrm{d}t$$

Ora definiamo una funzione integrale $I_c(x)$:

$$I_c(x) = \int_c^{z} f(t)\,\mathrm{d}t$$

Per il teorema fondamentale del calcolo integrale vale che:

$$I_c'(z) = f(z)$$

La nostra funzione G possiamo riscriverla come:

$$G(x) = I_c(h(x))$$

Facendo quindi la derivata:

$$G'(x) = I_c'(h(x)) \cdot h'(x) = f(h(x)) \cdot h'(x)$$

Qed.

2.7 Tabella primitive elementari

Di seguito riporto la tabella delle derivate al contrario, che serve appunto per calcolare gli integrali immediati. Gli integrali sono riportati volutamente senza estremi. Non è inoltre riportata la costante C che si mette per gli integrale indefiniti in quanto non mi interessa indicare l'infinità delle primitive, ma soltanto una di esse (in quanto caso con $C = 0$).

1. $\int x^n \, \mathrm{d}x = \frac{x^{n+1}}{n+1}$ con $n \in \mathbb{R} \setminus \{-1\}$. Se per caso $n < 0$ bisogna controllare che $x \neq 0$.
2. $\int \frac{1}{x} \, \mathrm{d}x = \ln|x|$
3. $\int e^x \, \mathrm{d}x = e^x$
4. $\int \sin x \, \mathrm{d}x = -\cos x$
5. $\int \cos x \, \mathrm{d}x = \sin x$
6. $\int \frac{1}{\cos^2(x)} \, \mathrm{d}x = \int 1 + \tan^2(x) \, \mathrm{d}x = \tan x$
7. $\int \frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x = \arcsin x$
8. $\int -\frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x = \arccos x$
9. $\int \frac{1}{1+x^2} \, \mathrm{d}x = \arctan x$

2.8 Tecniche di integrazione

2.8.1 Integrazione per parti

Consideriamo due funzioni $F, g : [a,b] \to \mathbb{R}$ dove g è continua, g' continua, F è derivabile, $F'(x) = f(x) \; \forall x \in [a,b]$, f continua. Dalla regola di derivazione del prodotto di due funzioni ci ricordiamo che:

$$\frac{\mathrm{d}}{\mathrm{d}x} [F(x) \cdot g(x)] = f(x)g(x) + F(x)g'(x)$$

Vogliamo ora calcolare il seguente integrale:

$$\int_a^b \frac{\mathrm{d}}{\mathrm{d}x} [F(x) \cdot g(x)] \, \mathrm{d}x = \int_a^b [f(x)g(x) + F(x)g'(x)] \, \mathrm{d}x$$

Usando il teorema fondamentale del calcolo integrale e spezzando l'integrale di destra:

$$[F(x) \cdot g(x)]_a^b = \int_a^b f(x)g(x) \, \mathrm{d}x + \int_a^b F(x)g'(x) \, \mathrm{d}x$$

Ne consegue quindi che:

$$\int_a^b f(x)g(x) \, \mathrm{d}x = [F(x) \cdot g(x)]_a^b - \int_a^b F(x)g'(x) \, \mathrm{d}x$$

Questa formula ci permette quindi di semplificare alcuni tipi di integrali del prodotto di funzioni. Se infatti per una di esse riusciamo a trovare una primitiva, allora possiamo riscrivere il nostro integrale usando quella primitiva e la derivata dell'altra funzione. È importante notare che abbiamo due funzioni e di una di esse dobbiamo trovare una primitiva, mentre per l'altra la sua derivata. La scelta è generalmente arbitraria su quale derivare e quale invece riscrivere come primitiva, ma di solito solo uno modo porta alla soluzione. Non ci sono particolari consigli da dare in questo caso se non fare molti esercizi.

Un classico esempio che si porta di solito per far vedere l'utilità di questa formula (o in generale tecnica) è il calcolo del seguente integrale:

$$\int_a^b e^x \cdot x \, \mathrm{d}x$$

In questo caso scegliamo di derivare x e di trovare una primitiva di e^x. È facile notare che la primitiva di quest'ultima è la funzione stessa e^x. Possiamo quindi riscrivere il nostro integrale nel seguente modo:

$$\int_a^b e^x \cdot x \, \mathrm{d}x = [e^x \cdot x]_a^b - \int_a^b e^x \cdot 1 \, \mathrm{d}x = [e^x \cdot x]_a^b - [e^x]_a^b$$

Notiamo che se avessimo fatto la scelta inversa, cioè di derivare e^x e di trovare una primitiva di x l'integrale si sarebbe complicato e non saremmo riusciti a risolverlo:

$$\int_a^b e^x \cdot x \, \mathrm{d}x = \left[e^x \cdot \frac{x^2}{2}\right]_a^b - \int_a^b e^x \cdot \frac{x^2}{2} \, \mathrm{d}x$$

Alcuni integrali possono richiedere **un'applicazione ripetuta** della formula di integrazione per parti. Un esempio è il seguente[3]:

$$\begin{aligned}\int x^2 \cdot e^x \, \mathrm{d}x &= e^x \cdot x^2 - \int e^x \cdot 2x \, \mathrm{d}x = e^x \cdot x^2 - 2 \int e^x \cdot x \, \mathrm{d}x = \\ &= e^x \cdot x^2 - 2 \left(e^x \cdot x - \int e^x \, \mathrm{d}x \right) = e^x \cdot x^2 - 2(e^x \cdot x - e^x) + C\end{aligned}$$

Come è facile notare questo tipo di integrali hanno in comune le così dette **funzioni circolari**, cioè quelle che hanno come derivata o se stesse oppure una funzione molto simile (tipo sin e cos). Infatti il seguente integrale si calcola sempre con la formula appena introdotta:

$$\int x \sin(x) \, \mathrm{d}x = -\cos(x) \cdot x - \int -\cos(x) \, \mathrm{d}x = -\cos(x) \cdot x + \sin(x) + C$$

Un altro tipo particolare di integrali che si fa con questa tecnica sono i seguenti:

$$\int \ln(x) \, \mathrm{d}x$$

In particolare possiamo osservare che apparentemente non ci sono due funzioni, quindi sembrerebbe che la formula non si possa applicare. In realtà possiamo riscrivere l'integrale come segue:

$$\int 1 \cdot \ln(x) \, \mathrm{d}x$$

In questo caso la funzione che dobbiamo integrare $f(x) = 1$, cioè la funzione costante. Una delle sue primitive è x, qindi possiamo applicare la formula di integrazione per parti:

$$\int 1 \cdot \ln(x) \, \mathrm{d}x = x \cdot \ln(x) - \int x \cdot \frac{1}{x} \, \mathrm{d}x = x \cdot \ln(x) - \int 1 \, \mathrm{d}x = x \ln(x) - x + C$$

Un ulteriore intgrale che si fa con la stessa tecnica è il seguente:

$$\begin{aligned}\int \arctan(x) \, \mathrm{d}x &= x \cdot \arctan x - \int x \cdot \frac{1}{1+x^2} \, \mathrm{d}x = x \cdot \arctan(x) - \frac{1}{2} \int \frac{2x}{1+x^2} \, \mathrm{d}x = \\ &= x \arctan(x) - \frac{1}{2} \ln(1+x^2) + C\end{aligned}$$

[3] Non ripoterò gli estremi di integrazione ma farò solo l'integrale indefinito per semplicità. In pratica mi riduco solo a trovare le primitive della funzione. Nel caso in cui si voglia trovare l'integrale della funzione tra due estremi basta prendere una delle primitive finali trovate e fare la classica formula data nel teorema fondamentale del calcolo integrale.

Esistono anche integrali che sono il prodotto di due funzioni circolari, ad esempio:

$$\int e^x \cdot sin(x)\,\mathrm{d}x$$

In questo caso si nota che se si prova ad applicare la formula di intgrazione per parti vista sopra si entra in un ”loop”. Per risolvere questo tipo di integrali bisogna applicarla due volte e fare qualche manipolazione algebrica. In questi casi è indifferente la scelta della funzione da derivare e quella da integrare:

$$\int e^x \cdot \sin(x)\,\mathrm{d}x = e^x \cdot \sin(x) - \int e^x \cos(x)\,\mathrm{d}x = e^x \cdot \sin(x) - \left(e^x \cdot \cos(x) - \int e^x \cdot - \sin(x)\,\mathrm{d}x\right) =$$

$$= e^x \cdot \sin(x) - e^x \cdot \cos(x) - \int e^x \cdot \sin(x)\,\mathrm{d}x$$

Notiamo ora che $\int e^x \sin(x)\,\mathrm{d}x$ compare da entrambe le parti dell'uguaglianza, quindi per comodità chiamiamo il nostro integrale I e riscriviamo:

$$I = e^x \cdot \sin(x) - e^x \cdot \cos(x) - I$$

Risolvendo per I:

$$I = \frac{1}{2} e^x \cdot (\sin(x) - \cos(x)) + C$$

In questo caso bisogna aggiungere $+C$ in quanto stiamo sempre trattando con tutte le primitive della funzione. Però possiamo finalmente concludere che:

$$\int e^x \cdot \sin(x)\,\mathrm{d}x = \frac{1}{2} e^x \cdot (\sin(x) - \cos(x)) + C$$

2.8.2 Integrazione per sostituzione

Gli integrali per sostituzione ci permettono di sotituire delle funzioni all'interno dell'integrale con delle funzioni ”più semplici”, in generale per semplificare notevolmente i calcoli. Per sempio se abbiamo il seguente integrale:

$$\int \frac{\sin(x)}{\cos(x)}\,\mathrm{d}x$$

Possiamo fare una sostituzione tipo $t = \cos(x)$ e l'integrale diventa:

$$\int \frac{-1}{t}\,\mathrm{d}t = -\ln(t) + C = -\ln(\cos(x)) + C$$

Non importa capire adesso come ha fatto la funzione $\sin(x)$ ha sparire completamente, ma piuttosto apprezzare la semplificazione dei calcoli data dalla sostituzione.

Teorema

Teorema del cambio di variabile: Siano $I, A \subseteq \mathbb{R}$ intervalli aperti. Data $h : I \to A$ continua e con derivata continua e $f : A \to \mathbb{R}$ continua, vale che:

$$\forall \alpha, \beta \in I \qquad \int_{h(\alpha)}^{h(\beta)} f(x)\,\mathrm{d}x = \int_{\alpha}^{\beta} f(h(t)) \cdot h'(t)\,\mathrm{d}t$$

Dimostrazione

Siano h ed f come da enunciato. Definisco $F : I \to \mathbb{R}$ dove:

$$F(z) = \int_{h(\alpha)}^{h(z)} f(x)\,\mathrm{d}x$$

Inoltre definisco $G : I \to \mathbb{R}$ dove:

$$G(z) = \int_{\alpha}^{z} f(h(t)) \cdot h'(t)\,\mathrm{d}t$$

Vogliamo quindi provare che sono la stessa funzione. Per il teorema di Lagrange abiamo che se due funzione hanno la stessa derivata allora differiscono per una costante. Se riusciamo a provare che hanno la stessa derivata e che la costante che le differenzia è 0 allora sono la stessa funzione[a].
Dimostriamo che hanno la stessa derivata, cioè:

$$F'(x) = G'(x)$$

Usando il teorema del calcolo integrale e quello generalizzato:

$$f(h(z)) \cdot h'(z) = f(h(z)) \cdot h'(z)$$

Dimostriamo ora che la costante per cui differiscono è 0. In particolare scelto un punto a caso la loro differenza deve essere 0. Scegliamo per comodità il punto $z = \alpha$:

$$F(\alpha) - G(\alpha) = \int_{h(\alpha)}^{h(\alpha)} \cdots - \int_{\alpha}^{\alpha} \cdots = 0 - 0 = 0$$

Non ho riportato il contenuto degli integrali volutamente perché non era necessario per far vedere che erano entrambi 0.

Qed.

[a] Sarebbe da chiedere meglio al prof perché non lo ha specificato nella prova ed è una mia supposizione che funzioni così.

Facciamo degli esempi:

$$\int_{3}^{5} e^{\sqrt{x}}\,\mathrm{d}x$$

Scegliamo $t = \sqrt{x}$, quindi $x = t^2$ e quindi $\mathrm{d}x = 2t\,\mathrm{d}t$:

$$\int_{\sqrt{3}}^{\sqrt{5}} e^t \cdot 2t\,\mathrm{d}t = 2\left([e^t \cdot t]_{\sqrt{3}}^{\sqrt{5}} - \int_{\sqrt{3}}^{\sqrt{5}} e^t\,\mathrm{d}t \right) = 2 \cdot [e^t(t-1)]_{\sqrt{3}}^{\sqrt{5}}$$

Un altro esempio potrebbe essere (lo faccio indefinito per evitare di portarmi dietro gli estremi di integrazione per tutto il calcolo):

$$\int \sin(x^{\frac{1}{3}})\,\mathrm{d}x$$

Facciamo la sostituzione $t = x^{\frac{1}{3}}$, quindi $x = t^3$ e quindi $\mathrm{d}x = 3t^2\,\mathrm{d}t$:

$$\int \sin(x^{\frac{1}{3}})\,\mathrm{d}x = \int \sin(t)\cdot 3t^2\,\mathrm{d}t = 3\int t^2\sin(t)\,\mathrm{d}t = 3\left(-\cos(t)\cdot t^2 + 2\int t\cdot\cos(t)\,\mathrm{d}t\right) =$$
$$= 3\left(-\cos(t)\cdot t^2 + 2\left(t\cdot\sin(t) - \int \sin(t)\,\mathrm{d}t\right)\right) =$$
$$= 3(-\cos(t)\cdot t^2 + 2(t\cdot\sin(t) + \cos(t))) + C = 3(-\cos(t)\cdot t^2 + 2t\cdot\sin(t) + 2\cos(t)) + C =$$
$$= -3\cos(t)\cdot t^2 + 6t\cdot\sin(t) + 6\cos(t) + C = -3\cos(x^{\frac{1}{3}})\cdot x^{\frac{2}{3}} + 6x^{\frac{1}{3}}\cdot\sin(x^{\frac{1}{3}}) + 6\cos(x^{\frac{1}{3}}) + C$$

Un ulteriore esempio che può essere considerato ”utile” è il calcolo dell’area del cerchio. In particolare, essendo la circonferenza unitaria definita dall’equazione $x^2 + y^2 = 1$, possiamo considerare solo la porzione nel primo quadrante del piano, che ha equazione $\sqrt{1-x^2}$ e successivamente moltiplicare per 4 l’area ottenuta. In particolare dobbiamo calcolare:

$$4\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x$$

In questo caso la sostituzione da fare è $x = \cos(t)$, quindi $\mathrm{d}x = -\sin(t)\,\mathrm{d}t$:

$$4\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x = 4\int_{\arccos(0)}^{\arccos 1} \sqrt{1-\cos^2(t)}\cdot -\sin(t)\,\mathrm{d}t =$$
$$= -4\int_{\frac{\pi}{2}}^{0} \sin(t)\cdot\sin(t)\,\mathrm{d}t = 4\int_0^{\frac{\pi}{2}} \sin^2(t)\,\mathrm{d}t =$$
$$= 4\int_0^{\frac{\pi}{2}} \frac{1-\cos(2t)}{2}\,\mathrm{d}t = 2\int_0^{\frac{\pi}{2}} 1\,\mathrm{d}t - 2\int_0^{\frac{\pi}{2}} \cos(2t)\,\mathrm{d}t =$$
$$= 2[t]_0^{\frac{\pi}{2}} - 2\left[\frac{1}{2}\sin(2t)\right]_0^{\frac{\pi}{2}} = 2\left[\frac{\pi}{2} - 0\right] - [\sin(\pi) - \sin(0)] =$$
$$= \pi - 0 = \pi$$

2.9 Integrali su intervalli non limitati (integrali generalizzati)

Definizione

Data $f : [a, +\infty[\to \mathbb{R}$ continua, si dice che f è integrabile su $[a, +\infty[$ se esiste **finito**:

$$\lim_{z\to+\infty}\int_a^z f(x)\,\mathrm{d}x =: \int_a^{+\infty} f(x)\,\mathrm{d}x$$

Il limite infatti può divergere o anche non esistere. È però richiesto che se la funzione è integrabile in quell’intervallo allora il limite è finito.

Esempio:

$$f(x) = \frac{1}{x^2} \qquad x \geq 1$$

Diventa quindi che:

$$\int_1^{+\infty} \frac{1}{x^2}\,\mathrm{d}x = \lim_{z\to+\infty}\left[-\frac{1}{x}\right]_1^z = \lim_{z\to+\infty}\left[-\frac{1}{z} + \frac{1}{1}\right] = -0 + 1 = 1$$

Un altro esempio:

$$\int_1^{+\infty} \frac{1}{x}\,\mathrm{d}x = \lim_{z\to+\infty} [\ln|x|]_1^z = \lim_{z\to+\infty} [\ln|z| - 0] = +\infty$$

In questo caso il limite non converge. In generale:

$$\int_1^{+\infty} \frac{1}{x^p}\,\mathrm{d}x = \begin{cases} \text{converge} & \text{se } p > 1 \\ \text{diverge} & \text{se } p \leq 1 \end{cases}$$

Infatti per $p = 1$ lo abbiamo già dimostrato che diverge, mentre per $p \neq 1$:

$$\int_1^{+\infty} \frac{1}{x^p} = \lim_{z\to+\infty} \left[\frac{x^{-p+1}}{-p+1}\right]_1^z = \lim_{z\to+\infty} \frac{z^{1-p}}{1-p} - \frac{1}{1-p} = \begin{cases} -\dfrac{1}{1-p} & \text{se } p > 1 \\ +\infty & \text{se } p < 1 \end{cases}$$

Definizione

Data $f :]a,b] \to \mathbb{R}$ continua, si dice che f è integrabile su $]a,b]$ se esiste **finito**:

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{z\to a^+} \int_z^b f(x)\,\mathrm{d}x$$

Esempio:

$$\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = \lim_{z\to 0} [2\sqrt{x}]_z^1 = \lim_{z\to 0} [2\sqrt{1} - 2\sqrt{z}] = 2$$

Altro esepio:

$$\int_0^1 \frac{1}{x} = \lim_{z\to 0} [2\ln(x)]_z^1 = \lim_{z\to 0} [2\ln(1) - 2\ln(z)] = +\infty$$

2.10 Altro

Il prof ha fatto le seguenti proposizioni, io le riporto per completezza ma mi sembra veramente inutile.
Proposizione: Se f è dispari e continua in $\mathbb{R}$:

$$\forall a, b \in \mathbb{R} \qquad \int_a^b f(x)\,\mathrm{d}x = -\int_{-b}^{-a} \mathrm{d}x$$

Dimostrazione:

$$\int_a^b f(x)\,\mathrm{d}x$$

Sostituiamo $t = -x$, cioè $x = -t$ e quindi $\mathrm{d}x = -\mathrm{d}t$.

$$\int_a^b f(x)\,\mathrm{d}x = \int_{-a}^{-b} f(-t) \cdot -1\,\mathrm{d}t = -\int_{-a}^{-b} -f(t)\,\mathrm{d}t = -\int_{-b}^{-a} f(t)\,\mathrm{d}t$$

Proposizione: Se f è pari:

$$\forall a, b \in \mathbb{R} \qquad \int_a^b f(x)\,\mathrm{d}x = \int_{-b}^{-a} \mathrm{d}x$$

3 Lo spazio $\mathbb{R}^n$

3.1 La struttura lineare

Definizione

$\mathbb{R}^n$ è l'insieme delle n-uple ordinate di numeri:

$$\mathbb{R}^n = \{x = (x_1, x_2, \cdots, x_n) \mid x_1, x_2, \cdots, x_n \in \mathbb{R}\}$$

Gli elementi di $\mathbb{R}^n$ si dicono vettori. In generale si possono pensare i sui elementi come vettori applicati nell'origine. Lo spazio $\mathbb{R}^n$ ha una struttra lineare. In esso sono definite due operazioni di base: la somma tra vettori e il prodotto di un vettore per uno scalare.

3.1.1 Somma tra vettori

Definizione

Presi due vettori $x, y \in \mathbb{R}^n$ dove $x = (x_1, \cdots, x_n)$ e $y = (y_1, \cdots, y_n)$ definiamo la somma tra vettori:

$$x + y := (x_1 + y_1, \cdots, x_n + y_n) \in \mathbb{R}^n$$

Se ci si riduce ad $\mathbb{R}^2$ la somma tra vettori è la stessa della *regola del parallelogramma* che spesso si studia in fisica.

3.1.2 Prodotto di un vettore per uno scalare

Definizione

Dato un vettore $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ e uno scalare $\lambda \in \mathbb{R}$ definiamo il prodotto di un vettore per uno scalare:

$$\lambda x := (\lambda \cdot x_1, \cdots, \lambda \cdot x_n)$$

I numeri in $\mathbb{R}$ vengono chiamati scalari perché se moltiplicati per un vettore visivamente lo scalano proprio del loro valore. Cioè se $\lambda = 2$ allora il vettore che viene moltiplicato per λ è il doppio del vettore originale.

3.1.3 Vettori coordinati

Definizione

In $\mathbb{R}^n$ i **vettori coordinati** sono:

$$\begin{aligned} e_1 &= (1, 0, 0, \cdots, 0) \\ e_2 &= (0, 1, 0, \cdots, 0) \\ &\vdots \\ e_n &= (0, 0, 0, \cdots, 1) \end{aligned}$$

I vettori coordinati sono ortogonali a coppie (per la defінzione di ortogonalità fare riferimento alla sezione 3.3):

$$\forall j \neq k, \quad j, k = 1 \cdots n \quad \langle e_j, e_k \rangle = 0$$

3.2 Prodotto scalare euclideo

Definizione

Dati $x, y \in \mathbb{R}^n$ dove $x = (x_1, \cdots, x_n)$ e $y = (y_1, \cdots, y_n)$ definiamo il prodotto scalare euclideo:

$$\langle x, y \rangle = x_1 \cdot y_1 + \cdots + x_n \cdot y_n = \sum_{k=1}^{n} x_k \cdot y_k \in \mathbb{R}$$

È importante notare che questo prodotto restituisce **un numero in** $\mathbb{R}$. Esiste anche un altra notazione per scrivere il prodotto scalare:

$$\langle x, y \rangle = x \cdot y$$

Esempio:

$$\langle (2, 3), (-1, 4) \rangle = 2 \cdot -1 + 3 \cdot 4 = 10$$

Altro esempio:

$$\langle (2, 1, 3), (-1, 4, 1) \rangle = 2 \cdot -1 + 1 \cdot 4 + 3 \cdot 1 = 5$$

Proprietà:

1. $\langle x, y \rangle = \langle y, x \rangle \quad \forall x, y \in \mathbb{R}^n$
2. Dati $x, y, z \in \mathbb{R}^n$ e $\lambda, \mu \in \mathbb{R}$
   $$\langle \lambda x + \mu y, z \rangle = \lambda \langle x, y \rangle + \mu \langle y, z \rangle$$
   Ovviamente per la proprietà 1 vale:
   $$\langle z, \lambda x + \mu y \rangle = \lambda \langle z, x \rangle + \mu \langle z, y \rangle$$
3. $\forall x \in \mathbb{R}^n \quad \langle x, x \rangle \geq 0$
4. $\langle x, x \rangle = 0 \iff x = \underline{0} = (0, \cdots, 0)$

Se ci riduciamo allo spazio $\mathbb{R}^2$ spesso viene data un'altra definzione di prodotto scalare euclideo. La sua definzione si basa sulla norma (Sezione: 3.4), mentre la sua dimostrazione di equivalenza si basa sulle coordinate polari (Sezione: 3.4.3) che sono entrambi concetti che vengono introdotti in seguito. Nonostante ciò viene riportata qui la definzione per mantenere coerente la suddivisione per argomenti.

Se abbiamo due vettori $x, y \in \mathbb{R}^2$ e chiamiamo *l'angolo compreso tra i due* α allora:

$$\langle x, y \rangle = \|x\| \cdot \|y\| \cdot \cos \alpha$$

Per la dimostrazione scriviamo x e y come:

$$x = (x_1, y_1) = (r_1 \cos \theta_1, r_1 \sin \theta_1)$$
$$y = (x_2, y_2) = (r_2 \cos \theta_2, r_2 \sin \theta_2)$$

Ne consegue che:

$$\langle x, y \rangle = \langle (x_1, y_1), (x_2, y_2) \rangle = r_1 r_2 \cos \theta_1 \cos \theta_2 + r_1 r_2 \sin \theta_1 \sin \theta_2 =$$

Ricordandoci le formule di addizione e sottrazione del coseno possiamo riscrivere quanto segue nella forma:

$$r_1 r_2 \cos \theta_2 - \theta_1$$

Essendo che $\theta_2 - \theta_1$ non è altro che l'angolo tra i due vettori che noi abbiamo chiamato α possiamo riscrivere come:

$$r_1 r_2 \cos\alpha$$

Nelle coordiante polari abbiamo chiamato $r_1 = \|(x_1, y_1)\|$, di conseguenza possiamo riscrivere come:

$$\|(x_1, y_1)\| \cdot \|(x_2, y_2)\| \cdot \cos\alpha = \|x\| \cdot \|y\| \cdot \cos\alpha$$

Abbiamo quindi concluso la dimostrazione.

3.3 Vettori ortogonali

Definizione

I vettori $x, y \in \mathbb{R}^n$ si dicono ortogonali (o perpendicolari) se $\langle x, y \rangle = 0$

Esempi:

- Se un vettore è il vettore nullo $x = \underline{0}$ allora tutti i vettori sono ortogonali a lui:

$$\langle \underline{0}, y \rangle = 0 \quad \forall y \in \mathbb{R}^n$$

- Se prendiamo $(x, y) = (\cos\theta, \sin\theta)$ e $(\overline{x}, \overline{y}) = \Big(\cos(\theta + \frac{\pi}{2}), \sin(\theta + \frac{\pi}{2})\Big) = (-\sin\theta, \cos\theta)$

$$\langle (x, y), (\overline{x}, \overline{y}) \rangle = \langle (\cos\theta, \sin\theta), (-\sin\theta, \cos\theta) \rangle = -\cos\theta \cdot \sin\theta + \sin\theta \cdot \cos\theta = 0$$

- $(a, b) \in \mathbb{R}^2, (-b, a) \in \mathbb{R}^2$

$$\langle (a, b), (-b, a) \rangle = -ba + ba = 0$$

Un esempio meno astratto e che in realtà conosciamo già e trovare la retta perpendicolare ad un'altra. Se infatti abbiamo due rette generiche $r_1 : y = mx$ e $r_2 : y = px$ possiamo chiederci per quali valori di p la seconda sia perpendicolare (o ortogonale) alla prima. Assumiamo per semplicità $m \neq 0$ e prendiamo due punti generici, uno su r_1 e l'altro su r_2:

$$(1, m) \in r_1 \qquad (1, p) \in r_2$$

Ora imponiamo che siano perpendicolari, in particolare:

$$\langle (1, m), (1, p) \rangle = 0$$

$$1 + mp = 0$$

$$p = -\frac{1}{m}$$

Abbiamo quindi ottenuto la condizione di perpendicolarità che ci è familiare nel piano.

3.4 Norma di un vettore

Definizione

Dato $x \in \mathbb{R}^n$, definiamo[a] la norma come:

$$\|x\| := \sqrt{\langle x, x\rangle} = \sqrt{\sum_{k=1}^{n} x_k^2}$$

[a] Il prof la indica senza le doppie barre ma come un semplice valore assoluto. Secondo me fa molta confusione quindi io userò sempre le doppie barrette.

La norma viene spesso chiamata **lunghezza** perché in effetti rappresenta la lunghezza del vettore su cui è calcolata. In $\mathbb{R}$ è il modulo del vettore stesso, in $\mathbb{R}^2$ è il semplice teorema di pitagora, mentre in spazi di dimensione più grade è già difficile da visualizzare.

Proprietà:

1. $\forall \lambda \in \mathbb{R}$, $\forall x \in \mathbb{R}^n$ vale: $\|\lambda x\| = |\lambda| \cdot \|x\|$. Questo perché:

$$\|\lambda x\| = \sqrt{(\lambda x_1)^2 + \cdots + (\lambda x_n)^2} = \sqrt{\lambda^2(x_1^2 + \cdots + x_n^2)} = |\lambda|\sqrt{x_1^2 + \cdots x_n^2} = |\lambda| \cdot \|x\|$$

2. $\forall x \in \mathbb{R}^n \quad \|x\| \leq 0$
3. $\|x\| = 0 \iff x = \underline{0}$
4. $\forall x, y \in \mathbb{R}^n$ vale: $\|x + y\| \leq \|x\| + \|y\|$

E possibile dimostrare che vale la **disuguaglianza di Cauchy-Schwarz** [4]:

$$\|x, y\| \leq \|x\| \cdot \|y\| \qquad \forall x, y \in \mathbb{R}^n$$

3.4.1 Teorema di pitagora generalizzato

Teorema

Dati $x, y \in \mathbb{R}^n$, se $x \perp y$ allora

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Dimostrazione

Prendo $x, y \in \mathbb{R}^n$ tali che $x \perp y$. Dalla definzione di norma:

$$\|x + y\|^2 = \langle x + y, x + y, \rangle = \langle x, x + y, \rangle + \langle y, x + y, \rangle =$$

$$= \langle x, x, \rangle + \langle x, y, \rangle + \langle y, x, \rangle + \langle y, y, \rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle$$

Per ipotesi però $x \perp y \implies \langle x, y\rangle = 0$ dalla definzione di vettori ortogonali, quindi:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Qed.

[4] Per introdurla il prof ha usato il fatto che si può esprimere il prodotto scalare tra due vettori x, y come $\|x\| \cdot \|y\| \cdot \cos\alpha$, e quindi essendo $\cos\alpha \in [-1, 1]$ $\forall \alpha \in \mathbb{R}$ si riesce a raggiungere questa disuguaglianza. Il fatto però è che secondo questo ragionamento dovrebbe valere solatanto in $\mathbb{R}^2$ ma in realtà vale in $\mathbb{R}^n$

3.4.2 Vettore normalizzato

Il vettore normalizzato nasce dall'idea che preso un vettore $x \in \mathbb{R}^n : x \neq \underline{0}$, si vuole trovare un numero $r > 0$ tale che:

$$\|rx\| = 1$$

Ricordandoci le proprietà della norma:

$$\|rx\| = |r| \cdot \|x\| = 1$$

E quindi:

$$r = \frac{1}{\|x\|}$$

Definizione

Dato un vettore $x \in \mathbb{R}^n : x \neq \underline{0}$ si chiama **normalizzato di x** il vettore:

$$\frac{x}{\|x\|}$$

E vale:

$$\left\| \frac{x}{\|x\|} \right\| = 1$$

Esempio: Troviamo il normalizzato di $x = (2, 3)$. Calcoliamo prima la sua norma:

$$\|x\| = \|(2,3)\| = \sqrt{4+9} = \sqrt{13}$$

Quindi il normalizzato:

$$\frac{x}{\|x\|} = \frac{(2,3)}{\sqrt{13}} = \left(\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}} \right)$$

3.4.3 Coordinate polari

Facciamo un'osservazione abbastanza banale che però si rivelerà estremamente utile nel caso particolare di $\mathbb{R}^2$: qualsiasi vettore (tranne $\underline{0}$) può essere scritto nella forma seguente:

$$x = \|x\| \cdot \frac{x}{\|x\|}$$

Notiamo che $\|x\|$ è un *numero maggiore di 0* mentre $\frac{x}{\|x\|}$ è un *vettore di lunghezza unitaria* [5]. Ne consegue che posso scrivere qualsiasi vettore con solo questi due elementi, cioè un vettore di lunghezza unitaria e uno scalare. Se riduciamo però il nostro studio allo spazio $\mathbb{R}^2$:

$$(x, y) = \|(x, y)\| \cdot \frac{(x, y)}{\|(x, y)\|}$$

Possiamo notare che il nostro vettore normalizzato $\frac{(x,y)}{\|(x,y)\|}$ è in realtà un vettore che sta sulla circonferenza goniometrica in quanto siamo sul piano e ha lunghezza 1. Questo implica che possiamo scriverlo come segue:

$$\frac{(x, y)}{\|(x, y)\|} = (\cos\theta, \sin\theta) \qquad \theta \in [0, 2\pi]$$

[5] La lunghezza è unitaria perché abbiamo definito la lunghezza di un vettore come la sua norma, e per definzione il vettore normalizzato ha norma uguale a 1

Ed essendo $\|(x,y)\|$ un semplice numero maggiore di 0 **possiamo riscrivere qualsiasi vettore in $\mathbb{R}^2$ nella forma:**

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

La coppia (r,θ) sono le **coordinate polari** di (x,y)

Con questo sistema di coordinate si possono inoltre descrivere porzioni di piano che prima erano estremamente difficili da scrivere. Un esempio banale è il semipiano nel *primo quadrante* delimitato dalla retta $y = x$, esso indatti è descrivibile da:

$$A = \left\{(x,y) = (r\cos\theta, r\sin\theta) \;\middle|\; r > 0, \theta \in \left[0, \frac{\pi}{2}\right]\right\}$$

Altri esempio più complessi sono i settori circolari. Se infatti volessimo descrivere un settore circolare che dista r_1 dall'origine ed è largo d ci basterebbe scrivere:

$$B = \left\{(x,y) = (r\cos\theta, r\sin\theta) \;\middle|\; \theta \in \left[0, \frac{\pi}{2}\right], r_1 < r < r_1 + d \;\text{ con }\; r_1 > 0, d > 0\right\}$$

3.4.4 Distanza in $\mathbb{R}^n$

Definizione

La distanza tra due punti $x, y \in \mathbb{R}^n$ è definita come il numero

$$\|x - y\|$$

4 Successioni in $\mathbb{R}^n$

Definizione

Una successione $(x_k)_{k\in\mathbb{N}} \in \mathbb{R}^n$ è una collezione di successioni in $\mathbb{R}$:

$$x_k = (x_k^1,\ x_k^2,\ x_k^3,\ \cdots,\ x_k^n) \in \mathbb{R}^n$$

Da notare che in questo caso i numeri in apice non vogliono indicare un elevamento a potenza ma semplicemente una distinzione tra le varie serie.

Definizione

Successione convergente: Data $(x_k)_{k\in\mathbb{N}} \in \mathbb{R}^n$ e dato $a = (a_1, a_2, \cdots, a_n) \in \mathbb{R}^n$ si dice che:

$$x_k \xrightarrow[k\to+\infty]{} a \iff \begin{cases} \lim_{k\to+\infty} x_k^1 = a_1 \\ \vdots \\ \lim_{k\to+\infty} x_k^n = a_n \end{cases}$$

In pratica una successione in $\mathbb{R}^n$ per essere convergente deve essere composta da sole successioni convergenti. Per sempio x_k è una successione convergente mentre y_k no.

$$x_k = \left(\frac{1}{k}, \frac{k+2}{k-1}\right) \qquad y_k = \left((-1)^k, \frac{1}{k}\right)$$

5 Funzioni in $\mathbb{R}^n$

Per definire le funzione abbiamo bisogni di definire gli intorni di un punto. In $\mathbb{R}^1$ abbiamo definito gli intorni sferici di un punto, mentre in $\mathbb{R}^n$ li definiamo come segue:

Definizione

Dato un **centro** $x \in \mathbb{R}^n$ e un **raggio** $r > 0$ poniamo

$$\mathcal{B}(x, r) = \{y \in \mathbb{R}^n \mid \|y - x\| < r\}$$

$\mathcal{B}$ si chiama **disco di centro x e raggio r**

In pratica sarebbe l'insieme di punti che hanno distanza da x minore o uguale a r. Se ci spostiamo un attimo in $\mathbb{R}^2$ e proviamo a prendere $x = (0, 0)$ e $r > 0$:

$$\mathcal{B}((0, 0), r) = \{(x, y) \in \mathbb{R}^2 \mid \|(x, y) - \underline{0}\| < r\}$$

Se si espande la definizione di norma si può notare che la condizione di appartenenza a questo insieme è:

$$\sqrt{x^2 + y^2} < r$$

Che corrisponde a tutti i punti contenuti in una circonferenza di raggio r con centro nell'origine. Facendo invece lo stesso ragionamento ma con un punto $x = (x_0, y_0)$ generico si arriva alla condizione:

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} < r$$

Dove anche questa volta l'equazione è una circonferenza di raggio r ma con il centro spostato in (x_0, y_0).

Definizione

Il **grafico** di $f : A \to \mathbb{R}^q$ con $A \subseteq \mathbb{R}^m$ è:

$$\text{Graf}(f) = \{(x, f(x)) \in A \times \mathbb{R}^q\}$$

5.1 Funzioni scalari, affini, radiali e cilindriche

Esistono vari tipi di funzioni in $\mathbb{R}^n$. In particolare quelle che hanno codomino $\mathbb{R}^1$ sono dette **funzioni scalari** in quanto appunto restituiscono uno scalare. Più precisamente:

Definizione

Una funzione scalare è del tipo $f : A \to \mathbb{R}$ dove $A \subseteq \mathbb{R}^n$.

Esempio: $f : \mathbb{R}^2 \to \mathbb{R}$ dove $f(x, y) = x^2 + y^2 = \|(x, y)\|^2$. Se riprendiamo la definizione di grafico risulta che:

$$\text{Graf}(f) = \{(x, y, x^2 + y^2) \mid x, y \in \mathbb{R}^2\}$$

Si può notare che $(x, y, x^2 + y^2) \in \mathbb{R}^3$. Se lo intersechiamo con il piano $\pi = \{(0, y, z) \mid y, z \in \mathbb{R}\}$ il grafico diventa:

$$\text{Graf}(f) \bigcap \pi = \{(0, y, y^2) \mid y \in \mathbb{R}^2\}$$

Che corrisponde all'insieme descritto dalle equazioni:

$$\begin{cases} x = 0 \\ z = y^2 \end{cases}$$

Che non è altro che una parabola. Se riscriviamo il nostro punto (x, y) in coordinate polari e ricalcoliamo il valore della funzione:

$$f(x,y) = f(r\cos x, r\sin x) = r^2\cos^2(x) + r^2\sin^2(x) = r^2(\cos^2(x) + \sin^2(x)) = r^2$$

Prendendo ora $\pi_\theta = \{(r\cos\theta, r\sin\theta, z) | r \geq 0, z \in \mathbb{R}\}$ e lo interseco con il grafico di f:

$$\text{Graf}(f) \bigcap \pi_\theta = \{(r\cos\theta, r\sin\theta, r^2) \mid r \geq 0\}$$

Un ulteriore esempio sono le **funzionni radiali**:

Definizione

Una fuzione si dice **radiale** se è esprimibile nella forma:

$$f : \mathbb{R}^2 \to \mathbb{R} \quad f(x,y) = g(\|(x,y)\|)$$

Con g funzione *opportuna*[a] definita in $g : [0, +\infty[\to \mathbb{R}$

[a] In questo caso con funzione opportuna si intende una funzione sufficientemente regolare. Non siamo andati troppo nei dettagli in questo caso quindi riporto quello detto dal prof.

Queste funzioni sono molto particolare perché il loro grafico è dato da rotazioni intorno all'asse z. Possiamo infatti considerare la funzione presa come esempio prima:

$$f(x,y) = x^2 + y^2 = g(\|(x,y)\|) \quad \text{con} \;\; g(r) = r^2$$

Il grafico come abbiamo visto è una rotazione intorno all'asse z di una parabola. L'equazione della parabola infatti risiede proprio in $g(r) = r^2$. È infatti questo il grafico che "ruota".

Se prendiamo un'altra funzione, tipo $g(r) = 1 - r$ e dove quindi:

$$f(x,y) = g(\|x,y\|) = 1 - \|x,y\| = 1 - \sqrt{x^2 + y^2}$$

Il grafico viso nella forma completa sembra complesso da disegnare, ma essendo questa una funzione radiale ci basta osservare la funzione g, disegnare il suo grafico e farlo ruotare intorno all'asse z. Il grafico risulta infatti un cono.

Definizione

Si definisce **funzione affine** una funzione nella forma:

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad f(x,y) = ax + by + c \;\; \text{con} \;\; a, b, c \in \mathbb{R}$$

Il grafico di queste funzioni risulta essere un piano in $\mathbb{R}^3$:

$$\text{Graf}(f) = \{(x, y, ax + by + c) | (x,y) \in \mathbb{R}^2\}$$

Prendiamo come esempio la funzione $f(x, y) = -y$: se la intersechiamo con il piano $\pi = \{(0, y, z) \mid y, z \in \mathbb{R}\}$ otteniamo un grafico del tipo:

$$\text{Graf}(f) \bigcap \pi = \{(0, y, -y) | y \in \mathbb{R}\}$$

Che risulta un sempilce $z = -y$. Essendo che non dipende da x possiamo intersecare tale grafico con quealsisi piano in cui la coordinata x abbia un valore fissato. Ne cosegue che il grafico completo si ottiene traslando lungo l'asse x la retta $z = -y$.

Definizione

Si definisce **funzione cilindrica** una funzione in due variabili che dipende da solo una delle due:

$$f(x,y) = h(y) \qquad \text{oppure} \qquad f(x,y) = g(x)$$

Con h e g opportune.

Il grafico di queste funzioni si ottiene disegnando la funzione $z = h(y)$ (o nel caso $z = g(x)$) e successivamente traslando tale grafico sull'asse da cui non dipende la funzione. Per esempio la funzione presa prima in cosiderazione $f(x,y) = -y$ è una funzione cilindrica e si ottiene appunto tracciando la retta $z = -y$ e translado tale retta sull'asse x, cioè quello da cui non dipende la funzione.

Il fatto di traslare sull'asse da cui non dipende la funzione è semplicemente dato dal fatto che in $\mathbb{R}^3$ una grafico del tipo $\text{Graf}(f) = \{(x, y, h(y)) \mid x, y \in \mathbb{R}\}$ non impone nessun vincolo sul paramentro x, ne consegue che vale per tutti i suoi valori. Non siamo infatti constretti a disegnare tale funzione in $x = 0$, ma possiamo disegnarla per un qualsiasi valore di x e poi traslare tutto il grafico lungo l'asse.

5.2 Continuità funzioni scalari

Definizione

Data una funzione $f : A \to \mathbb{R}$ con $A \subseteq \mathbb{R}^n$ e un punto $x_0 \in \mathbb{A}$, si dice che f è **continua in $\mathbf{x_0}$** se:

$$\forall (x_k)_{k\in\mathbb{N}} \subseteq A \quad \text{(cioè di punti in } A\text{)}$$

vale:

$$x_k \xrightarrow[k\to+\infty]{} x_0 \implies f(x_k) \xrightarrow[k\to+\infty]{} f(x_0)$$

Oppure equivalentemente: f è continua in $x_0 \in A$ se:

$$\forall \epsilon > 0, \exists \delta > 0 : \begin{cases} x \in A \\ x \in \mathcal{B}(a, \delta) \end{cases} \implies |f(x) - f(x_0)| < \epsilon$$

Dove $x \in B(x_0, \delta)$ è il disco di centro x_0 e raggio δ.

Tutte le funzioni "elementari" sono **continue** nei loro domini.

Una funzione quindi del tipo

$$f(x,y) = \frac{\ln(y - x^2)}{\sqrt{1 + \sin^2(xy)}}$$

è continua nel suo dominio, cioè:

$$\text{Dom}(f) = \{(x,y) \in \mathbb{R}^2 \mid y - x^2 > 0\}$$

5.3 Insieme di livello

Definizione

Dato $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$ e $b \in \mathbb{R}$ si definisce **insieme di livello**:

$$L_b = \{x \in A \mid f(x) = b\} = f^{-1}(b)$$

$f^{-1}(b)$ è la *controimmagine* di b. Si noti che la controimmagine non richiede che la funzione sia invertibile perché appunto è l'insime di tutti i punti che presi dalla funzione restituiscono b. Se la funzione fosse invertibile questo insieme sarebbe composto da un elemento, cosa che però non è richiesta.

Gli insiemi di livello vengono largamente usati nelle cartine geografiche in quanto rappresenta una linea lungo la quale l'altidudine rimane costante.

Facciamo un esempio in $\mathbb{R}^2$: Prendiamo la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita come $f(x, y) = x^2 + y^2$. Proviamo a calcolare qualche insieme di livello:

1. Se prendiamo come numero $b < 0$ notiamo che l'insieme di livello risulta vuoto:

$$L_b = \{x \in \mathbb{R}^2 \mid x^2 + y^2 = b\} = \{x \in \mathbb{R}^2 \mid x^2 + y^2 < 0\} = \emptyset$$

   Questo perché non esiste nessun vettore in $\mathbb{R}^2$ che soddisfi l'equazione $x^2 + y^2 < 0$.

2. Se prendiamo invece come numero $b = 0$ l'insieme di livello risulta:

$$L_b = \{x \in \mathbb{R}^2 \mid x^2 + y^2 = 0\} = \{(0, 0)\}$$

   In quanto $x^2 + y^2 = 0 \iff (x, y) = (0, 0)$.

3. Se prendiamo invece come numero $b > 0$ l'insieme di livello risulta:

$$L_b = \{x \in \mathbb{R}^2 \mid x^2 + y^2 = b\}$$

   L'equazione $x^2 + y^2 = b$ descrive una circonferenza di raggio $\sqrt{b}$. Possiamo osservare la funzione f "dall'alto" e notare che su queste circonferenze la funzione è costante. È come se tagliassimo con un piano parallelo al piao $z = 0$ la funzione.

5.4 Curve parametrizzate

Definizione

Una curva parametrizzata è una funzione $f : \mathbb{R} \to \mathbb{R}^n$ definita come:

$$t \mapsto r(t) = (r_1(t), \cdots, r_n(t))$$

Sono quindi funzioni che prendono uno scalare e restituiscono un vettore. Un esempio protrebbe essere una legge orari in fisica:

$$r(t) = (x_0, y_0) + t(v_1, v_2)$$

Quest'ultima infatti, preso il tempo che è uno scalare, ci restituisce un vettore in $\mathbb{R}^2$ rappresena una posizione. Scritta nella foma della defnzione verebbe:

$$t \mapsto r(t) = (r_1(t), r_2(t)) = (x_0 + tv_1, y_0 + tv_2)$$

Un ulteriore esempio potrebbe essere la cura:

$$r(t) = (\cos(t), \sin(t))$$

Notiamo che essa descrive una circonferenza, ma anche la seguente descrive la stessa circonferenza:

$$f(t) = (\cos(t^3), \sin(t^3))$$

La differenza è che percorre la circonferenza in maniera diversa. Come formalizziamo questa diversità? Con il concetto di velocità di una curva.

6 Derivabilità e differenziabilità

6.1 Derivabilità di funzioni scalari

In generale quando ci si trova in $\mathbb{R}^1$ si definisce la derivata di una funzione su un intervallo aperto del tipo $]a, b[$. Questo perché se si definisce la derivata su un intervallo chiuso si avrebbero dei problemi in quanto potrebbe esistere solo la derivata destra o sinistra in un punto. Definiamo quindi l'equivalente degli intervalli aperti in $\mathbb{R}^n$.

Definizione

Dato $A \subseteq \mathbb{R}^n$, dico che **A è aperto** se:

$$\forall x_0 \in A, \exists \epsilon > 0 : \mathcal{B}(x_0, \epsilon) \subseteq A$$

Se ci mettiamo in $\mathbb{R}^2$ un insieme di questo tipo è dato da un figura con i bordi tratteggiati, cioè dove tutti i punti sul bordo non appartengono all'insieme A.

Teorema

Individuazione degli intervalli aperti: Data $f : \mathbb{R}^n \to \mathbb{R}$ continua, $\forall b \in \mathbb{R}$ l'insieme $\{x \in \mathbb{R}^n \mid f(x) > b\}$ è aperto. L'apertura è data dal fatto che l'insieme è definito tramite ugualianza stretta.

6.1.1 Derivate parziali di funzioni scalari

Definizione

Dato $A \subseteq \mathbb{R}^2$ aperto e sia $(x_0, y_0) \in A$. Si dice che f **è derivabile rispetto a x in** $(\mathbf{x_0}, \mathbf{y_0})$ se esiste **finito**:

$$\frac{\partial f}{\partial x}(x_0, y_0) := \lim_{t \to 0} = \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t} \in \mathbb{R}$$

Questo limite si indica con il simbolo di derivata parziale (∂). Di seguito alcune notazioni equivalenti:

$$\frac{\partial f}{\partial x} = \partial_x f = D_x f$$

Equivalentemente possiamo riscrivere il limite come:

$$\frac{\partial f}{\partial x}(x_0, y_0) := \lim_{x \to x_0} = \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} \in \mathbb{R}$$

Vale ovviamente che la derivata rispetto a y in (x_0, y_0) risulta:

$$\frac{\partial f}{\partial y}(x_0, y_0) := \lim_{t \to 0} = \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{t} \in \mathbb{R}$$

Opppure equivalentemente:

$$\frac{\partial f}{\partial y}(x_0, y_0) := \lim_{y \to y_0} = \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0} \in \mathbb{R}$$

Esempio: prendiamo la funzione $f(x, y) = xy^2$. Le sue derivate parziali sono:

$$\frac{\partial f}{\partial x}(x, y) = y^2 \qquad \frac{\partial f}{\partial y}(x, y) = 2xy$$

L'idea per calcolare queste derivate è quella di considerarle come una funzioni normale (cioè come quelle viste in $\mathbb{R}^1$) e di considerare come variabile quella che bisogna derivare, mentre tutte le altre vengono considerate costanti. Nel primo caso infatti, dove abbiamo derivato rispetto alla x, ci è bastato considerare la y costante. Quindi è come se avessimo fatto la derivata in $\mathbb{R}^1$ della semplice funzione:

$$f(x) = a^2 \cdot x \implies f'(x) = a^2$$

Dove però al posto della a ci sarebbe la y. Nel secondo caso invece la derivata è rispetto alla y, quindi la x viene considerata come semplice costante: $f(y) = a \cdot y^2 \implies f'(y) = 2ay$. Dove ovviamente al posto della a ci sarebbe la x.

Se si vuole **generalizzare la definzione di derivata parziale** è necessario riprendere la definzione di vettori coordinati (Sezione: 3.1.3). In generale quindi data $f : \mathbb{R}^n \to \mathbb{R}$ e un punto $\overline{x} \in \mathbb{R}^n$ si definisce la i-esima derivata parziale come:

$$\frac{\partial f}{\partial x_i}(\overline{x}) = \lim_{t \to 0} \frac{f(\overline{x} + t \cdot e_i) - f(\overline{x})}{t}$$

6.1.2 Gradiente

Le derivate parziali si possono aggregare in una funzione vettoriale (cioè che restituisce un vettore) chiamata gradiente:

Definizione

Data una funzione $f : A \in \mathbb{R}$ con $A \subseteq \mathbb{R}^n$, se esistono tutte le derivate pariziali di f in ogni punto possiamo definire **il gradiente di f** come:

$$\nabla f(x_1, x_2, \cdots) = \left(\frac{\partial f}{\partial x_1}(x_1, x_2, \cdots), \frac{\partial f}{\partial x_2}(x_1, x_2, \cdots), \cdots\right)$$

Nel caso particolare di $\mathbb{R}^2$:

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y)\right)$$

Il simbolo del gradiente (∇ si lege *nabla*).

Esempio: data la funzione $f(x, y) = \sin(x + y^2)$ il suo gradiente risulta:

$$\nabla f(x, y) = (\cos(x + y^2), \cos(x + y^2) \cdot 2y)$$

6.1.3 Derivabilità e continuità di funzioni scalari

In $\mathbb{R}^1$ vale la relazione:

$$f \text{ derivabile in } x \implies f \text{ continua in } x$$

Questa affermazione vale ache in generale? Cioè se per esempio prendiamo una funzione scalare con le derivate parziali definiti, allora questa funzione è continua? Proviamo a verificarlo[6]:

[6] La seguente non è una dimostrazione di un teorema. Ho comunque voluto evidenziarla perché potrebbe risultare effettivamente importante

Dimostrazione

Vogliamo provare che presa una funzione scalare con dominio maggiore di $\mathbb{R}$ se esitono tutte le sue derivate parziali in un punto non vuol dire che quella funzione è derivabile. In particoalare data $f : \mathbb{R}^2 \to \mathbb{R}$:

$$\exists \partial_x f(x,y), \partial_y f(x,y) \not\Longrightarrow f \text{ continua}$$

Per dimostrare che qualcosa non è vero basta portare un controesempio: prendiamo la funzione definita:

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

Mostriamo quindi che esistono le derivate parziali in $(0,0)$ ma la funzione è discontinua in $(0,0)$.

$$\partial_x f(0,0) = \lim_{t\to 0} \frac{f(t,0) - f(0)}{t} = \lim_{t\to 0} \frac{\dfrac{t\cdot 0}{t^2+0} - 0}{t} = \lim_{t\to 0} \frac{0-0}{t} = 0$$

Da notare che questa non è una forma indeterminata perché dalla definzione di limite $t \neq 0 \; \forall t \in \mathbb{R}$. Il numeratore invece è esattamente 0. Per la derivata parziale rispetto a y si fa esattamente la stessa cosa in quanto la funzione è simmetrica. Le derivate parziali quindi esistono:

$$\partial_x f(0,0) = 0$$
$$\partial_y f(0,0) = 0$$

Mostriamo ora che non è continua: prendiamo due successioni che tendono a $(0,0)$:

$$(u_n, v_n) = \left(0, \frac{1}{n}\right) \xrightarrow[n\to+\infty]{} (0,0)$$
$$(x_n, y_n) = \left(\frac{1}{n}, \frac{1}{n}\right) \xrightarrow[n\to+\infty]{} (0,0)$$

Per essere continua devono valere i seguenti limiti:

$$f(u_n, v_n) \xrightarrow[n\to+\infty]{} f(0,0) = 0$$
$$f(x_n, y_n) \xrightarrow[n\to+\infty]{} f(0,0) = 0$$

Mostriamo che solo il primo vale:

$$f(u_n, v_n) = \frac{0\cdot\dfrac{1}{n}}{0^2 + \dfrac{1}{n^2}} \xrightarrow[n\to+\infty]{} 0$$

$$f(x_n, y_n) = \frac{\dfrac{1}{n}\cdot\dfrac{1}{n}}{\dfrac{1}{n^2}+\dfrac{1}{n^2}} = \frac{\dfrac{1}{n^2}}{2\cdot\dfrac{1}{n^2}} = \frac{1}{2} \; \forall n$$

Ne consegue che la funzione è discontinua. Qed.

Abbiamo visto che non vale la relazione:

$$f \text{ derivabile in } x \implies f \text{ continua in } x$$

Quella che però vale è:

$$f \text{ differenziabile in } x \implies f \text{ continua in } x$$

6.2 Differenziabilità di funzioni scalari

In $\mathbb{R}^1$ preso un punto $x_0 \in \mathbb{R}$ vale che:

$$\exists f'(x_0) \in \mathbb{R} \iff \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0)$$

Possiamo riscrivere il limite come segue (mantenendo sempre il *se e solo se*):

$$\iff \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - h \cdot f'(x_0)}{h} = 0$$

$$\iff f(x_0 + h) - f(x_0) - f'(x_0) \cdot h = o(h)$$

$$\iff f(x_0 + h) = f(x_0) + f'(x_0) \cdot h + o(h)$$

Questa è la formula di Taylor di ordine 1 per $h \to 0$. Se facciamo una sostituzione del tipo $x_0 + h = x$ possiamo riscrivere la formula per $x \to x_0$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$$

Definizione

Notazioni per o-piccoli: Sia $f : \mathbb{R}^2 \to \mathbb{R}$ si scrive che (con $p > 0$):

$$f(x, y) = o(\|(x, y)\|^p)$$

se:

$$\forall \epsilon > 0, \exists \delta > 0, \forall (x, y) : \|(x, y)\| < \delta \implies |f(x, y)| \leq \epsilon \cdot \|(x, y)\|^p$$

Se $(x, y) \neq (0, 0)$ si può riscrivere come:

$$\cdots \implies \frac{f(x, y)}{\|(x, y)\|^p} \leq \epsilon$$

Equivalentemente[a]:

$$\forall (x_n, y_n) \to (0, 0) \quad \text{vale} \quad \lim_{n \to +\infty} \frac{f(x_n, y_n)}{\|(x_n, y_n)\|^p} = 0$$

[a] In realtà non è perfettamente equivalente perché il rapporto non sarebbe ben definito se (x_n, y_n) assumesse il valore $(0, 0)$. In generale è quindi preferibile la prima defnzione.

Esempio: verifichiamo che la funzione $f(x, y) = x^2$ è o-piccolo di $\|(x, y)\|$ per $(x, y) \to (0, 0)$. In questo caso $p = 1$. Deve quindi valere il limite:

$$\lim_{(x,y) \to (0,0)} \frac{f(x, y)}{\|(x, y)\|^p} = \lim_{(x,y) \to (0,0)} \frac{x^2}{\sqrt{x^2 + y^2}} = 0$$

Osserviamo che $|x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2} = \|(x,y)\|$. Ne consegue che:

$$0 \leq \frac{x^2}{\|(x,y)\|} \leq \frac{\|(x,y)\|^2}{\|(x,y)\|} = \|(x,y)\| \xrightarrow[(x,y)\to(0,0)]{} 0$$

Vale quindi il limite. Una cosa che non dimostriamo è che:

$$f(x,y) = \text{polinomio omogeneo dei grado p} \implies f(x,y) = o(\|(x,y)\|^p)$$

Un polinomio omogeneo è un polinomio che ha tutti i termini di grado esattamente p.

Definizione

Funzione differenziabile: Dato $A \subseteq \mathbb{R}^n$ aperto, $x_0 \in A$, $h \in \mathbb{R}^n$ e $f : A \to \mathbb{R}$ si dice che f è **differenziabile in** x_0 se:

1. $\exists \partial_1 f, \partial_2 f, \cdots, \partial_n f \in \mathbb{R}$ nel punto x_0
2. Vale la formula:

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + o(\|h\|) \quad h \to \underline{0}$$

La 2 si può anche riscrivere con $x = x_0 + h$:

$$f(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + o(\|x - x_0\|) \quad x \to x_0$$

Ovviamente la notazione con il prodotto euclideo si può espandere:

$$\langle \nabla f(x_0), h \rangle = \partial_1 f \cdot h_1 + \cdots + \partial_n f \cdot h_n = \sum_{k=1}^{n} \partial_k f \cdot h_k$$

Se $f : \mathbb{R}^n \to \mathbb{R}$ è differenziabile in x_0 il polinomio

$$T_1(x) = f(x_0) + \sum_{k=1}^{n} \partial_k f(x_0) \cdot (x - x_0)$$

si chiama **polinomio di Taylor di grado 1 di punto inziale $\mathbf{x_0}$**. Inoltre in $\mathbb{R}^2$ il piano di equazione

$$z = T_1(x,y)$$

risulta essere il **piano tangente al grafico di f** nel punto $(x_0, y_0, f(x_0, y_0))$.

Teorema

Differenziabile totale: Data una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ continua, se esistono le sue derivate parziali e sono continue in ogni punto allora f è differenziabile in ogni punto. Cioè:

$$\exists \partial_x f, \partial_y f \text{ continue } \forall x_0 \in \mathbb{R}^2 \implies f \text{ differenziabile } \forall y_0 \in \mathbb{R}^2$$

Tutte le funzioni elementari sono differenziabili in quanto soddisfano la formula di Taylor.

Il seguente lemma è necessario alla dimostrazione che segue.

Lemma

Lagrange per derivate parziali: Se $f : \mathbb{R}^2 \to \mathbb{R}$ è continua e ha derivate parizali continue, allora:

$$\forall (a,b) \in \mathbb{R}^2, \forall h \in \mathbb{R},\ \exists \theta \in [0,1]\ :\ f(a+h,b) - f(a,b) = \partial_x f(a+\theta h, b)h$$

Inoltre:

$$\forall (a,b) \in \mathbb{R}^2, \forall k \in \mathbb{R},\ \exists \theta \in [0,1]\ :\ f(a,b+k) - f(a,b) = \partial_y f(a, b+\theta k)k$$

Dimostrazione:
Definiamo una funzione ausiliaria $g : \mathbb{R} \to \mathbb{R}$ dove $g(x) = f(x,b)\ \forall x \in \mathbb{R}$. Applichiamo ora il teorema di Lagrange per g nell'intervallo di estremi $[a, a+h]$:

$$\exists \theta \in [0,1] : g(a+h) - g(a) = g'(a+\theta h)(a+h-a)$$

Che per definzione di derivata parziale:

$$g'(a+\theta h) \cdot h = \partial_x f(a+\theta h, b) \cdot h$$

Qed.

Questo lemma richiama il teorema di Lagrange visto per le funzioni definite in $\mathbb{R}^1$, solo che è scritto diversamente. Prendiamo il primo caso, quello con la derivata parziale in x e proviamo a riscriverlo: Il punto $a+h$ è semplicemente un punto generico che quindi possiamo indicare con $A = a+h$. Il punto $a+\theta h$ invece, in quanto $\theta \in [0,1]$, è semplicemente un altro modo di scrivere un punto nell'intervallo $[a, a+h]$. Possiamo quindi riscriverlo come $c \in [a, a+h]$, cioè per la sostituzione di prima $c \in [a, A]$. Possiamo quindi riscrivere l'enunciato come segue:

$$\forall (a,b) \in \mathbb{R}^2, \forall A \in \mathbb{R},\ \exists c \in [a, A] : f(A,b) - f(a,b) = \partial_x f(c,b) \cdot (A-a)$$

In questo caso abbiamo riscritto $h = A - a$. Se inoltre assumiamo che $A \neq a \implies h \neq 0$, quindi possiamo dividere per h:

$$\partial_x f(c,b) = \frac{f(A,b) - f(a,b)}{A-a}$$

Che se lo confrontiamo con il teorema di Lagrange per le funzioni ad una variabile ci assomiglia effettivamente molto:

$$f'(c) = \frac{f(b) - f(a)}{b-a}$$

Dimostriamo ora il teorema del differenziale totale:

Dimostrazione

Data $f : \mathbb{R}^2 \to \mathbb{R}$ continua di cui esistono le derivate parziali e sono anche loro continue devo dimostrare che f è differenziabile. Cioè per definzione di funzione differenziabile mi riduco a dimostrare:

$$f(x_0+h, y_0+k) = f(x_0,y_0) + \langle \nabla f(x_0,y_0), (h,k) \rangle + o(\|(h,k)\|)$$

Per dimostrare una formula di questo tipo in pratica devo dimostrare che l'o-piccolo è corretto.

$$f(x_0+h, y_0+k) - f(x_0,y_0) = \langle \nabla f(x_0,y_0), (h,k) \rangle + o(\|(h,k)\|)$$

Aggiungiamo e sottraiamo alla parte sinistra dell'uguaglianza il fattore $f(x_0+h, y_0)$:

$$f(x_0+h, y_0+k) - f(x_0+h, y_0) + f(x_0+h, y_0) - f(x_0,y_0) = \langle \cdots \rangle + o(\cdots)$$

Possiamo raggruppare a due a due i termini identificandoli per comodità nel seguente modo:

1. $f(x_0 + h, y_0 + k) - f(x_0 + h, y_0)$
2. $f(x_0 + h, y_0) - f(x_0, y_0)$

Possiamo ora applicare Langrange per le derivate parziali (cioè il lemma enunciato sopra):
Per 1:

$$\exists \theta_1 \in [0, 1] \, : \, f(x_0 + h, y_0) - f(x_0, y_0) = \partial_x f(x_0 + \theta_1 h, y_0)h$$

Per 2:

$$\exists \theta_2 \in [0, 1] \, : \, f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) = \partial_y f(x_0 + h, y_0 + \theta_2 k)k$$

Riscrviamo la formula iniziale:

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \langle \nabla f(x_0, y_0), (h, k) \rangle + o(\|(h, k)\|)$$

Sotituendo 1 e 2 dopo Lagrange ed espandendo il prodotto scalare euclideo:

$$\begin{aligned} &\partial_x f(x_0 + \theta_1 h, y_0)h + \partial_y f(x_0 + h, y_0 + \theta_2 k)k \\ =&\partial_x f(x_0, y_0)h + \partial_y f(x_0, y_0)k + o(\|(h, k)\|) \end{aligned}$$

Portando il prodotto scale espanso a destra e raccogliendo h e k:

$$[\partial_x f(x_0 + \theta_1 h, y_0) - \partial_x f(x_0, y_0)]h + [\partial_y f(x_0 + h, y_0 + \theta_2 k) - \partial_y f(x_0, y_0)]k = o(\|(h, k)\|)$$

Ci basta quindi far vedere che tutto quello nelle parentesi quadre è effettivamente un o-piccolo della norma di (h, k). Per la definzione di o-piccolo devo dimostrare che (vale la stessa cosa per l'altra parentesi quadra):

$$\frac{[\partial_x f(x_0 + \theta_1 h, y_0) - \partial_x f(x_0, y_0)]h}{\|(h, k)\|} \xrightarrow[(h,k)\to(0,0)]{} 0$$

In quanto:

$$\left| \frac{h}{\|(h, k)\|} \right| < \frac{\|(h, k)\|}{\|(h, k)\|} = 1$$

Mi basta quindi mostrare che:

$$[\partial_x f(x_0 + \theta_1 h, y_0) - \partial_x f(x_0, y_0)] \xrightarrow[(h,k)\to(0,0)]{} 0$$

Prendiamo una successione $(h_n, k_n) \to (0, 0)$ e calcoliamo il limite:

$$\lim_{n\to+\infty} \partial_x f(x_0 + \theta_1 h_n, y_0) - \partial_x f(x_0, y_0) =$$

Essedo per ipotesi $\partial_x f$ continua $(x_0 + \theta_1 h_n, y_0) \xrightarrow[h_n \to 0]{} (x_0, y_0)$:

$$= \partial_x f(x_0, y_0) - \partial_x f(x_0, y_0) = 0$$

Qed.

Il teorema appena enunciato vale anche con l'implicazione contraria. Questa però è una proposizione più che un vero e proprio teorema:

PROP: Se $f : \mathbb{R}^2 \to \mathbb{R}$ è differenziabile in (x_0, y_0) allora f è continua in (x_0, y_0).

Dimostrazione

Dimostro la proprozione appena enunciata: Suppongo che f sia differenziabile in (x_0, y_0) e dimostro che f è continua in (x_0, y_0). Cioè che presa una successione $(x_n, y_n) \xrightarrow[n\to+\infty]{} (x_0, y_0)$ deve valere il limite:

$$\lim_{n\to+\infty} f(x_n, y_n) = f(x_0, y_0)$$

Essendo che per ipotesi f è diffirenziabile in (x_0, y_0) vale la formula di Taylor:

$$f(x_n, y_n) = f(x_0, y_0) + \langle \nabla f(x_0, y_0), (x_n - x_0, y_n - y_0)\rangle + o(\|(x_n - x_0, y_n - y_0)\|)$$

Sostituendo nel limite:

$$\begin{aligned} &\lim_{n\to+\infty} f(x_0, y_0) + \langle \nabla f(x_0, y_0), (x_n - x_0, y_n - y_0)\rangle + o(\|(x_n - x_0, y_n - y_0)\|) = \\ &= \lim_{n\to+\infty} f(x_0, y_0) + \langle \nabla f(x_0, y_0), (x_0 - x_0, y_0 - y_0)\rangle + o(\|(x_0 - x_0, y_0 - y_0)\|) = \\ &= \lim_{n\to+\infty} f(x_0, y_0) + \langle \nabla f(x_0, y_0), (0,0)\rangle + o(\|(0,0)\|) = \lim_{n\to+\infty} f(x_0, y_0) = f(x_0, y_0) \end{aligned}$$

Qed.

6.3 Derivate direzionali di funzioni scalari

Definizione

Sia $f : \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n$ e $v \in \mathbb{R}^n$ con $\|v\| = 1$. **La derivata direzionale** di f in x_0 nella direzione v è il limite (se esiste finito):

$$\frac{\partial f}{\partial v}(x_0) := \lim_{t\to 0} \frac{f(x_0 + t \cdot v) - f(x_0)}{t}$$

Un altro per indicare la derivata dirazionale di f in x_0 lungo la direzione v è: $\partial_v f(x_0)$

Notiamo che le derivate parziali sono un caso speciale della derivata direzionale: esse infatti si formano usando come vettore v un vettore coordinato della base di $\mathbb{R}^n$. Se infatti scegliamo $v = (1, 0, \cdots, 0) = e_1$:

$$\lim_{t\to 0} \frac{f((x_0) + t \cdot (1, 0, \cdots, 0)) - f(x_0)}{t} = \frac{\partial f}{\partial x_1}(x_0)$$

Data $f : \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n$ e $v \in \mathbb{R}^n \setminus \underline{0}$ con $\|v\| = 1$. Se prendiamo la funzione $g : \mathbb{R} \to \mathbb{R}$ definita come:

$$g(t) = f(x_0 + t \cdot v)$$

Possiamo osservare che:

$$g'(0) = \frac{\partial f}{\partial v}(x_0)$$

Infatti:

$$g'(0) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = \lim_{t\to 0} \frac{f(x_0 + t \cdot v) - f(x_0)}{t} = \partial_v f(x_0)$$

Questa piccola osservazione aiuta i calcoli, questo perché per calcolare una derivata direzionale di una funzione in $\mathbb{R}^n$ ci riduciamo semplicemente al calcolo della derivata di una funzione in $\mathbb{R}^1$. Esempio: Data la funzione $f(x, y) = xy^2$ vogliamo calcolare $\partial_v(1, 2)$ con v generico:

$$g(t) = f(1 + tv_1, 2 + tv_2) = (1 + tv_1)(2 + tv_2)^2$$

Quindi la derivata ora:

$$g'(t) = v_1(2 + tv_2)^2 + (1 + tv_1) \cdot 2 \cdot (2 + tv_2) \cdot v_2$$

Per calcolare quindi $\partial_v f(1, 2)$ ci basta semplicemente calcolare $g'(0)$:

$$g'(0) = v_1 \cdot 2^2 + 1 \cdot 2 \cdot 2 \cdot v_2 = 4v_1 + 4v_2 = 4(v_1 + v_2)$$

Teorema

Teorema delle derivate direzionali: Data $f : \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n$ e $v \in \mathbb{R}^n$ con $\|v\| = 1$. Se f è differenziabile in x_0 allora vale:

$$\frac{\partial f}{\partial v}(x_0) = \langle \nabla f(x_0), v \rangle = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) \cdot v_i$$

Questo implica che la derivata direzionale **è sempre lineare rispetto a v**.

Dimostrazione

Dalla defnzione di derivata direzionale:

$$\frac{\partial f}{\partial v}(x_0) = \lim_{t \to 0} \frac{f(x_0 + t \cdot v) - f(x_0)}{t}$$

Essendo la funzione differenziabile in x_0 per ipotesi vale la formula di Taylor:

$$f(x_0 + t \cdot v) - f(x_0) = \langle \nabla f(x_0), t \cdot v \rangle + o(\|t \cdot v\|)$$

Notiamo che:

$$o(\|t \cdot v\|) = o(|t| \cdot \|v\|)$$

Ed essendo per ipotesi $\|v\| = 1$ l'o-piccolo diventa $o(t)$. Riportiamo la formula di Taylor sopra il limite:

$$\lim_{t \to 0} \frac{\langle \nabla f(x_0), t \cdot v \rangle + o(t)}{t} = \lim_{t \to 0} \left(\langle \nabla f(x_0), v \rangle + \frac{o(t)}{t} \right) = \langle \nabla f(x_0), v \rangle$$

Qed.

6.3.1 Direzione di massima crescita

Avendo definito la derivata direzionale per una funzione in un punto, ora possiamo derivare tale funzione in infinite direzioni date dal versore v. Chiediamoci qual'è la direzione che massimizza la derivata, cioè la direzione che massimizza la crescita delle funzione. Prendiamo quindi una funzione generica $f : \mathbb{R}^2 \to \mathbb{R}$ differenziabile nel punto $(x_0, y_0) \in \mathbb{R}^2$. Per evitare banalità $\nabla f(x_0, y_0) \neq \underline{0}$ in quanto se il gradiente fosse nullo $\partial_v f = 0 \; \forall v$. Volgiamo quindi massimizzare la quantità:

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

Per farlo passiamo alle coordinate polari. Il gradiente risulta quindi:

$$\nabla f(x_0, y_0) = (r\cos(\phi), r\sin(\phi)) \quad \text{dove} \quad r = \|\nabla f(x_0, y_0)\| \quad q \in [0, 2\pi]$$

Il versore v, in quanto unitario per la definzione di derivata direzionale:

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

Calcoliamo quindi la derivata direzionale, che per il teorema delle derivate direzionali risulta:

$$\begin{aligned}\frac{\partial f}{\partial v}(x_0, y_0) &= \langle \nabla f(x_0, y_0), (v_1, v_2)\rangle = \langle (r\cos(\phi), r\sin(\phi)), (\cos(\theta), \sin(\theta))\rangle = \\ &= r\cos(\phi)\cos(\theta) + r\sin(\phi)\sin(\theta) = r\cos(\phi - \theta)\end{aligned}$$

Per rendere massima quindi l'espressione $r\cos(\phi - \theta)$ il coseno deve essere uguale a 1. Quindi:

$$\theta = \phi$$

Questo implica che il versore di massima crescia ha la stessa direzione del gradiente[7]. Inoltre essendo che $\theta = \phi$:

$$\frac{\partial f}{\partial v}(x_0, y_0) = r\cos(\phi - \theta) = r\cos(\phi - \phi) = r = \|\nabla f(x_0, y_0)\|$$

Tornando un attimo alla formula:

$$\frac{\partial f}{\partial v}(x_0, y_0) = r\cos(\phi - \theta)$$

Notiamo che se $\phi - \theta = \frac{\pi}{2}$, cioè in pratica il versore v è perpendicolare[8] al gradiente $\nabla f(x_0.y_0)$, si ha che la derivata è sempre nulla:

$$\frac{\partial f}{\partial v}(x_0, y_0) = r\cos\frac{\pi}{2} = 0$$

Essendo la derivata sempre nulla significa che la funzione è costante lungo quella direzione: ci troviamo quindi lungo una linea di livello. Ne consegue che, essendo v lungo la linea di livello ed essendo $\nabla f(x_0, y_0) \perp v$, allora il gradiente è perpendicolare alla linea di livello. Facendo un riassunto:

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile in un punto $x_0 \in \mathbb{R}^n$. Il versore v che massimizza $\partial_v f(x_0)$ lo denotimo con $v_{\max}$ ed è il normalizzato del gradiente nel punto x_0:

$$v_{\max} = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$$

Inoltre vale che:

$$\frac{\partial f}{\partial v_{\max}}(x_0) = \|\nabla f(x_0)\|$$

E il gradiente è sempre ortogonale all'insieme di livello della funzione che comprende il punto $b = f(x_0)$. Cioè l'insieme:

$$L_b = \{x \in \mathbb{R}^n | f(x) = b = f(x_0)\}$$

6.4 Derivata di una curva

[7] Questo fatto vale anche in $\mathbb{R}^n$. Per dimostrarlo però non si può fare uso delle coordinate polari e bisogna prendere un'altra strada

[8] Questo perché appunto $r\cos(\phi - \theta)$ è dato dal prodotto scalare euclideo tra il gradiente $\nabla f(x_0, y_0)$ e il versore v. Ne consegue che per definizione di ortogonalità il gradiente e il versore v sono perpendicolari in quanto questo prodotto è nullo.

Definizione

Velocità di una curva: Data $r :]a, b[\to \mathbb{R}^n$ definita come:

$$t \mapsto r(t) = (r_1(t), \cdots, r_n(t))$$

La velocità del punto t_0 è:

$$r'(t_0) = (r'_1(t_0), \cdots, r'_n(t_0))$$

Quest'ultimo si chiama vettore velocità e rappresenta la tangente alla curva nel punto t_0. Possiamo anche evitare la sostituzione del punto t_0 e quindi ottenere la funzione velocità.

Tornando infatti alle funzioni di prima le velocità sono diverse:

$$\begin{aligned} r'(t) =& (-\sin(t), \cos(t)) \\ f(t) =& (-\sin(t^3) \cdot 3t^2, \cos(t^3) \cdot 3t^2) \end{aligned}$$

Mentre infatti $r(t)$ e $f(t)$ descrivono lo stesso insieme di punti al variare di t, le loro velocità no.

Definizione

Velocità scalare: Data $f : \mathbb{R} \to \mathbb{R}^n$ si definisce la velocità scalare al tempo t_0 lo scalare:

$$\|r'(t_0)\|$$

Possiamo definire la **formula di Taylor** per le curve: se $t \in]a, b[$

$$\begin{cases} r_1(t+h) = r_1(t) + r'_1(t)h + o_1(h) \\ \vdots \\ r_n(t+h) = r_n(t) + r'_n(t)h + o_n(h) \end{cases}$$

Possiamo anche scrivere in modo più compatto:

$$r(t+h) = r(t) + r'(t)h + o(h)$$

Definizione

Lunghezza di una curva: Data $r : \mathbb{R} \to \mathbb{R}^n$ dove $t \to r(t) \in \mathbb{R}^n$ sia $[a, b] \in\subseteq R$, supponiamo che $r'(t) \neq \underline{0}\ \forall t$. La lunghezza del tratto di curva tra $r(a)$ e $r(b)$ vale:

$$\text{Length}(a, b) = \int_a^b \|r'(t)\| \,\mathrm{d}t$$

Esempio 1: Prediamo la funzione $r(t) = (\cos(t), \sin(t))$ e calcoliamo la lunghezza nell'intervallo $[0, 2\pi]$. Essendo una circonferenza il valore che ci aspettimo è proprio la lunghezza di quest'ultima:

$$\int_0^{2\pi} \|r'(t)\| \,\mathrm{d}t = \int_0^{2\pi} \|(-\sin(t), \cos(t))\| \,\mathrm{d}t = \int_0^{2\pi} \sqrt{\sin^2(t) + \cos^2(t)} \,\mathrm{d}t = \int_0^{2\pi} 1 \,\mathrm{d}t = [t]_0^{2\pi} = 2\pi$$

Esempio 2: Prediamo la funzione $r(t) = (\cos(t^2), \sin(t^2))$ e calcoliamo la lunghezza nell'intervallo $[0, \sqrt{2\pi}]$. La funzione nell'intervallo ristretto descrive sempre una circonferenza, quello che cambia è però la sua velocità:

$$\int_0^{\sqrt{2\pi}} \|r'(t)\| \,\mathrm{d}t = \int_0^{\sqrt{2\pi}} \|(-\sin(t^2)\cdot 2t, \cos(t^2)\cdot 2t)\| \,\mathrm{d}t = \int_0^{\sqrt{2\pi}} \sqrt{4t^2(\sin^2(t)+\cos^2(t))} \,\mathrm{d}t =$$

$$= \int_0^{\sqrt{2\pi}} 2t \,\mathrm{d}t = [t^2]_0^{\sqrt{2\pi}} = 2\pi$$

Notiamo però che nonostante la velocita sia diversa la lunghezza rimane la stessa. Questo significa che **la lunghezza non dipende dalla parametrizzazione**[9].

6.5 Derivata funzioni composte

Teorema

Derivata lungo una curva: Data $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile e $r : \mathbb{R} \to \mathbb{R}^n$ derivabile. Allora è definita la funzione:

$$(f \circ r)(t) = f(r(t)) \quad \forall(t) \in \mathbb{R}$$

Tale funzione è derivabile $\forall t \in \mathbb{R}$ e vale:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(r(t)) = \langle \nabla f(r(t)), r'(t) \rangle$$

Questo implica che la derivata **è lineare rispetto alla velocità**, cioè rispetto al vettore $r'(t)$.

Notiamo che se prendiamo una funzione generica $f : \mathbb{R}^n \to \mathbb{R}$ e una funzione $r(t) = x + tv$ dove $x, v \in \mathbb{R}^n$ e $\|v\| = 1$ possiamo calcolare la derivata in 0. La compisizione delle due funzioni risulta:

$$(f \circ r)(t) = f(x + tv)$$

Dal teorema appena enunciato la derivata in 0 risulterebbe quindi:

$$(f \circ r)'(0) = \langle \nabla f(r(0)), r'(0) \rangle = \langle \nabla f(0), v \rangle$$

Se invece usassimo la defինzione di derivata per la funzione $g(t) = f(x + tv)$ verebbe:

$$g'(0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = \lim_{h \to 0} \frac{f(x + hv) - f(x)}{h} = \frac{\partial f}{\partial v}(x) = \langle \nabla f(0), v \rangle$$

Notiamo che il risultato è esattamente lo stesso.

Dimostrazione

Assumiamo che f sia differenziabile e r derivabile nei loro rispettivi domini. Vogliamo dimostrare che:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(r(t)) = \langle \nabla f(r(t)), r'(t) \rangle$$

Iniziamo con espandere la definzione di derivata:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(r(t)) = \lim_{h \to 0} \frac{f(r(t+h)) - f(r(t))}{h}$$

Ricodandoci le formule di Taylor per le curve:

$$r(t+h) = r(t) + r'(t)h + o(h) \quad h \to 0$$

[9] Il prof ha detto che non è molto rigorosa come affermazione, ma per noi va più che bene

E quindi:
$$r(t+h) - r(t) = r'(t)h + o(h) \quad h \to 0$$
Ricordandoci inoltre Taylor per le funzioni in più variabili:
$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + o(\|h\|)$$
Ora applichiamo la formula di Taylor al numeratore della nostra derivata:
$$f(r(t+h)) - f(r(t)) = \langle \nabla f(r(t)), r(t+h) - r(t) \rangle + o(\|r(t+h) - r(t)\|) =$$
Sostituendo ora l'espressione $r(t+h) - r(t)$ con la formula di Taylor scritta sopra:
$$= \langle \nabla f(r(t)), r'(t)h + o(h) \rangle + o(\|r'(t)h + o(h)\|) =$$
Usando la linearità del prodotto scalare:
$$= \langle \nabla f(r(t)), r'(t) \rangle \cdot h + \langle \nabla f(r(t)), o(h) \rangle + o(\|r'(t)h + o(h)\|)$$
Riprendiamo il limite che dobbiamo calcolare e sostituiamo il numeratore:
$$\lim_{h\to 0} \frac{\langle \nabla f(r(t)), r'(t) \rangle \cdot h + \langle \nabla f(r(t)), o(h) \rangle + o(\|r'(t)h + o(h)\|)}{h}$$
Nel primo fattore il termine h si semplifica mentre negli alti due il rapporto con h manda banalmente tutto a zero:
$$= \lim_{h\to 0} \frac{\langle \nabla f(r(t)), r'(t) \rangle \cdot h}{h} + \frac{\langle \nabla f(r(t)), o(h) \rangle}{h} + \frac{o(\|r'(t)h + o(h)\|)}{h} =$$
$$= \lim_{h\to 0} \langle \nabla f(r(t)), r'(t) \rangle + 0 + 0 = \langle \nabla f(r(t)), r'(t) \rangle$$

Qed.

Gradiente e insiemi di livello: Se prendiamo una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ e un punto $b \in \mathbb{R}$ l'insieme di livello associeto risutla:
$$L_b = f^{-1}(b) = \{(x,y) \in \mathbb{R} | f(x,y) = b\}$$
Supponiamo per comodità $\nabla f(x,y) \neq \underline{0} \quad \forall (x,y)$. Supponiamo inoltre di poter costruire una curva $r : \mathbb{R} \to \mathbb{R}^2$ t.c. $r(t) \in L_b$ $\forall t$ e $r'(t) \neq \underline{0}$. Se ora applichiamo la nostra funzione iniziale f alla curva oteniamo che:
$$f(r(t)) = b \quad \forall t \in \mathbb{R}$$
Questo significa che è costante e quindi:
$$\frac{\mathrm{d}}{\mathrm{d}t} f(r(t)) = 0 \quad \forall t$$
Per il teorema precedente:
$$\langle \nabla f(r(t)), r'(t)h \rangle = 0$$
Ne consegue quindi che $\nabla f(r(t))$ è ortogonale a $r'(t)$.

6.6 Derivate parziali seconde

Come nel caso di $\mathbb{R}^1$ in cui si poteva considerare la derivata prima come una funzione e quindi derivarla nuovamente, anche in $\mathbb{R}^2$ è possibile considerare le derivate parziali come funzioni e quindi applicare nuovamente l'operatore di derivata. Per esempio se abbiamo la funzione $f = x^2 + xy^3$ possiamo calcolare le sue derivate parziali:

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) = 2x + y^3 \\ \dfrac{\partial f}{\partial y} f(x,y) = 3xy^2 \end{cases}$$

Ora se consideriamo le derivate prime come delle funzioni possiamo calcolare le derivate seconde, anche questa volta parziali:

$$\begin{cases} \dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial x}(x,y) = 2 \\ \dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial y}(x,y) = 3y^2 \\ \dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial x}(x,y) = 3y^2 \\ \dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial y}(x,y) = 6xy \end{cases}$$

Le derivate rispetto alla stessa variabile vengono generalemente indicate come **derivate quadrate** o pure, mentre le altre come semplici **derivate miste**. Si usano di solito varie convenzioni per evitare di scrivere due volte la frazione per le derivate seconde:

$$\begin{cases} \dfrac{\partial^2 f}{\partial x^2}(x,y) := \dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial x}(x,y) \\ \dfrac{\partial^2 f}{\partial x\, \partial y}(x,y) := \dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial y}(x,y) \\ \dfrac{\partial^2 f}{\partial y\, \partial x}(x,y) := \dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial x}(x,y) \\ \dfrac{\partial^2 f}{\partial y^2}(x,y) := \dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial y}(x,y) \end{cases}$$

Come possiamo notare dall'esempio riportato sopra le derivate miste hanno lo stesso risultato. Questo non è un semplice caso:

Teorema

Teorema di Schwarz: Dato $A \subseteq \mathbb{R}^n$ e $f : A \to \mathbb{R}^n$. Se f ha tutte le derivate seconde continue, allora vale che:

$$\frac{\partial^2 f}{\partial x_j\, \partial x_k} = \frac{\partial^2 f}{\partial x_k\, \partial x_j} \quad \forall j,k \in \{1, \cdots, n\}$$

Per le derivate miste quindi **non conta l'ordine di derivazione**.

Se prendiamo quindi una funzione $f(x,y,z) = x^3 y e^{5z}$ e proiamo a calcolare due derivate miste:

$$\frac{\partial f}{\partial x} = 3x^2 y e^{5z} \qquad \frac{\partial^2 f}{\partial x\, \partial z} = 15x^2 y e^{5z}$$

$$\frac{\partial f}{\partial z} = 5x^3 y e^{5z} \qquad \frac{\partial^2 f}{\partial z\, \partial x} = 15x^2 y e^{5z}$$

Definizione

Matrice Hessiana: Dato $A \subseteq \mathbb{R}^n$ e $f : A \to \mathbb{R}^n$ con derivate seconde continue, possiamo raggruppare le derivate seconde in una matrice quadrata detta matrice Hessiana e indicato con:

$$\mathbf{H}f(x) \in M_{n\times n}(\mathbb{R}) = \mathbb{R}^{n\times n}$$

E i cui elementi sono:

$$(\mathbf{H}f(x))_{jk} = \frac{\partial^2 f}{\partial x_j \, \partial x_k}(x)$$

In forma leggermente più esplicita risulta:

$$\mathbf{H}f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \, \partial x_1} & \frac{\partial^2 f}{\partial x_1 \, \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \, \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \, \partial x_1} & \frac{\partial^2 f}{\partial x_2 \, \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \, \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \, \partial x_1} & \frac{\partial^2 f}{\partial x_n \, \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \, \partial x_n} \end{bmatrix}$$

Per il teorema di Schwarz la matrice è **simmetrica**.

Definizione

Una matrice $\mathbf{A} \in \mathbb{R}^{n\times n}$ si dice **simmetrica** se è uguale alla sua trasposta:

$$\mathbf{A} = \mathbf{A}^T$$

In $\mathbb{R}^2$ la matrice Hessiana assume la seguente forma:

$$\mathbf{H}f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \, \partial y} \\ \frac{\partial^2 f}{\partial y \, \partial x} & \frac{\partial^2 f}{\partial x^2} \end{bmatrix}$$

Quindi per la funzione considera all'inzio di questa sezione, cioè $f = x^2 + xy^3$ la matrice Hessiana risulta essere:

$$\mathbf{H}f(x) = \begin{bmatrix} 2 & 3y^2 \\ 3y^2 & 6xy \end{bmatrix}$$

E come abbiamo visto dal teorema di Schwarz è simmetrica.

6.6.1 Forme quadratiche

Definizione

Forma quadratica: Sia $\mathbf{A} \in \mathbb{R}^{n\times n}$ e $\mathbf{A} = \mathbf{A}^T$. La forma quadratica associata ad $\mathbf{A}$ è la funzione $q : \mathbb{R}^n \to \mathbb{R}$ definita come:

$$q(h) = \langle \mathbf{A}h, h \rangle \qquad \forall h \in \mathbb{R}^n$$

Esempio in $\mathbb{R}^2$: Prendiamo la matrice $\mathbf{A}$ che risulta essere (in quanto simmetrica):

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

La forma quadratica associata risulta quindi essere:

$$\begin{aligned} q(h_1, h_2) =& \langle \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \rangle = \langle \begin{bmatrix} ah_1 + bh_2 \\ bh_1 + ch_2 \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \rangle = ah_1^2 + bh_1h_2 + bh_1h_2 + ch_2^2 \\ =& ah_1^2 + 2bh_1h_2 + ch_2^2 \end{aligned}$$

Il risultato è un polinomio omogeneo di grado 2 che assomiglia molto ad un quadrato di binomio (sarà da qui che viene il nome?). Nel caso di $\mathbb{R}^3$ siulta invece:

$$\begin{aligned} q(h_1, h_2, h_3) =& \langle \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \rangle = \langle \begin{bmatrix} ah_1 + bh_2 + ch_3 \\ bh_1 + dh_2 + eh_3 \\ ch_1 + eh_2 + fh_3 \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \rangle = \\ &= ah_1^2 + bh_1h_2 + ch_1h_3 + bh_1h_2 + dh_2^2 + eh_2h_3 + ch_1h_3 + eh_2h_3 + fh_3^2 = \\ &= ah_1^2 + dh_2^2 + fh_3^2 + 2bh_1h_2 + 2ch_1h_3 + 2eh_2h_3 \end{aligned}$$

Possiamo anche riscrivere la forma quadratica tramite la sommatoria:

$$\langle \mathbf{A}h, h \rangle = \sum_{j=1}^{n} (\mathbf{A}h)_j \cdot h_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_{jk} h_k \right)_j \cdot h_j = \sum_{j,k=1}^{n} a_{jk} h_k h_j$$

In questo caso si nota ancora di più il fatto che il polinomio è omogeneo e ha grado 2.

Definizione

Segno di forme quadratiche: Sia $A = A^T \in \mathbb{R}^{n \times n}$ simmetrica e la sua forma quadratica associata q. Si dice che:

- A è *definita positiva* ($A > 0$) se vale $\langle Ah, h \rangle > 0 \ \ \forall h \in \mathbb{R}^n \setminus \{\underline{0}\}$
- A è *definita negativa* ($A < 0$) se vale $\langle Ah, h \rangle < 0 \ \ \forall h \in \mathbb{R}^n \setminus \{\underline{0}\}$
- A è *indefinita* se esistono $h_-, h_+ \in \mathbb{R}^n$ tali che $\langle Ah_-, h_- \rangle < 0 < \langle Ah_+, h_+ \rangle$

Da notare che le disiguaglianze sono strette.

Teorema

Classificazione forme quadratiche non degeneri 2x2: Se $A = A^T = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathbb{R}^2$, allora:

- $A > 0$ sse $a > 0$ e $\det(A) > 0$
- $A < 0$ sse $a < 0$ e $\det(A) > 0$
- A è indefinita sse $\det(A) < 0$

Per quando si faranno gli autovalori in algebra:

- $A > 0$ sse tutti gli autovalori sono positivi.
- $A < 0$ sse tutti gli autovalori sono negativi.
- A è indefinita sse un autovalore è positivo e l'altro è negativo.

6.7 Taylor di ordine 2 con resto secondo Lagrange

La forma classica della formula di Taylor al secondo ordine è:

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h + o(h)$$

Il resto in questo caso è espresso dall'o-piccolo, ma esiste anche un' altra formula dove il resto viene espresso in maniera differente:

Teorema

Taylor ordine 2 con resto secondo Lagrange in $\mathbb{R}^1$: Sia $I \subseteq \mathbb{R}$ e sia $f : I \to \mathbb{R}$ con derivata prima e seconda continue sull'intervallo aperto I, vale che:

$$\forall x_0 \in I, x_0 + h \in I, \exists \theta \in]0,1[: \; f(x_0 + h) = f(x_0) + f'(x_0)h + f''(x_0 + \theta h)\frac{h^2}{2}$$

Da notare che in questo caso il resto non si trova nell'o-piccolo (anche perché è assente), ma bensì nel termine di secondo grado. In particolare nel fatto che la matrice Hessiana non viene valutata nel punto esatto x_0, ma in una sorta di punto medio tra x_0 e h, cioè $x_0 + \theta h$.

Dimostrazione

Per dimostrare il teorema facciamo la sostituzione $x_0 + h = x$. Ci riduciamo quindi a voler dimostrare che:

$$\forall x_0, x \in I, \exists \theta \in]0,1[: \quad f(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0 + \theta(x - x_0))\frac{(x - x_0)^2}{2}$$

Osserviamo che:

$$\exists k \in \mathbb{R} : \; f(x) - f(x_0) - f'(x_0)(x - x_0) - k(x - x_0)^2 = 0$$

Questo perché se $(x - x_0) \neq 0$ (caso banale) è una semplice equazione di primo grado in k. Riduciamoci quindi a dimostrare che:

$$\exists \theta \in]0,1[: k = \frac{f''(x_0 + \theta(x - x_0))}{2}$$

Consideriamo la funzione ausiliaria $g : [x_0, x] \to \mathbb{R}$:

$$g(t) = f(x) - f(t) - f'(t)(x-t) - k(x-t)^2$$

Notiamo che grazie a questa costruzione:

$$g(x) = f(x) - f(x) - f'(x)(x-x) - k(x-x)^2 = -f'(x) \cdot 0 - k \cdot 0 = 0$$

Inoltre:

$$g(x_0) = f(x) - f(x_0) - f'(x_0)(x-x_0) - k(x-x_0)^2 = 0$$

Questo vale proprio per la scelta di k fatta in precedenza. Essendo la funzione derivabile ed essendo nulla agli estremi del dominio possiamo applicare il teorema di Rolle:

$$\exists \theta \in]0,1[: \ g'(x_0 + \theta(x-x_0)) = 0$$

Calcoliamo la derivata:

$$g'(t) = -f'(t) - f''(t)(x-t) - f'(t) \cdot -1 - 2k(x-t) \cdot -1 = [-f''(t) + 2k](x-t)$$

Per il teorema di Rolle appena applicato:

$$[-f''(x_0 + \theta(x-x_0)) + 2k](x - (x_0 + \theta(x-x_0))) = 0$$

Notiamo che il secondo fattore è diverso da zero in quanto:

$$x - (x_0 + \theta(x-x_0)) = x - x_0 - \theta(x-x_0) = (x-x_0)(\theta - 1)$$

Essendo infatti $x \neq x_0$ e $\theta \in]0,1[$. Ne consegue che:

$$-f''(x_0 + \theta(x-x_0)) + 2k = 0 \implies k = \frac{f''(x_0 + \theta(x-x_0))}{2}$$

Qed.

Il teorema vale anche in generale :

Teorema

Taylor ordine 2 con resto secondo Lagrange in $\mathbb{R}^n$: Sia $A \subseteq \mathbb{R}^n$ e sia $f : A \to \mathbb{R}$ con derivate prime e seconde continue, vale che: $\forall x_0 \in A, \forall h \in \mathbb{R}^n$ tale che il segmento $\{x_0 + th : t \in [0,1]\} \subseteq A$, $\exists \theta \in]0,1[$ tale che:

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle \mathbf{H} f(x_0 + \theta h) h, h \rangle$$

Dimostrazione

Sia f, x_0 e h come da ipotesi. Consideriamo ora la funzione ausiliaria $g : [0,1] \to \mathbb{R}$ definita come segue:

$$g(t) = f(x_0 + th)$$

La funzione è derivabile due volte e quindi valgono le formule:

$$g'(t) = \frac{\mathrm{d}}{\mathrm{d}t} f(x_0 + th) = \langle \nabla f(x_0 + th), h \rangle = \sum_{j=1}^{n} \partial_j f(x_0 + th) h_j$$

Si noti che questa derivata è data dal Teorema per le derivate composte (Sezione: 6.5).

$$g''(t) = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{j=1}^{n} \partial_j f(x_0 + th) h_j = \sum_{j,k=1}^{n} \partial_{jk} f(x_0 + th) h_k h_j = \langle \mathbf{H} f(x_0 + th) h, h \rangle$$

Se ora applichiamo la formula di Taylor per il secondo ordine sulla funzione g:

$$\exists \theta \in]0, 1[: \; g(x_0 + h) = g(x_0) + g'(x_0) h + g''(x_0 + \theta h) \frac{h^2}{2}$$

Scegliamo appositamente $x_0 = 0$ e $h = 1$. La formula quindi diventa:

$$g(1) = g(0) + g'(0) + \frac{1}{2} g''(\theta)$$

Facendo ora le sostituzioni con le formule ottenute in precedenza:

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle \mathbf{H} f(x_0 + \theta h) h, h \rangle$$

Qed.

Corollario: (Formula di Taylor con resto in forma di o-piccolo (forma di Peano)). Sia $A \subseteq \mathbb{R}^n$ aperto e sia $f : A \to \mathbb{R}$ con derivate prime e seconde continue, vale che: $\forall x_0 \in A, \exists \delta > 0$ tale che $\forall h \in \mathcal{B}(0, \delta)$:

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle \mathbf{H} f(x_0) h, h \rangle + o(\|h\|^2)$$

Dimostrazione: La dimostrazione segue dalla formula di Taylor con resto secondo Lagrange. Infatti aggiungendo e sottraendo il termine $\frac{1}{2} \langle \mathbf{H} f(x_0) h, h \rangle$:

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle \mathbf{H} f(x_0 + \theta h) h, h \rangle - \frac{1}{2} \langle \mathbf{H} f(x_0) h, h \rangle + \frac{1}{2} \langle \mathbf{H} f(x_0) h, h \rangle$$

Riduciamoci quindi a dimostrare che:

$$\frac{1}{2} \langle \mathbf{H} f(x_0 + \theta h) h, h \rangle - \frac{1}{2} \langle \mathbf{H} f(x_0) h, h \rangle = o(\|h\|^2)$$

Eliminando $\frac{1}{2}$ in quanto costante e usando la linearità del prodotto scalare:

$$\langle (\mathbf{H} f(x_0 + \theta h) - \mathbf{H} f(x_0)) h, h \rangle = o(\|h\|^2)$$

Riscrivendo la forma quadratica come sommatoria:

$$\sum_{j,k=1}^{n} (\partial_{jk} f(x_0 + \theta h) - \partial_{jk} f(x_0)) h_j h_k = o(\|h\|^2)$$

Essendo una sommatoria ci basta dimostare che ogni parte è o-piccolo della norma al quadrato di h. Per farlo prendiamo un termine generico e dimostriamo:

$$(\partial_{jk} f(x_0 + \theta h) - \partial_{jk} f(x_0)) h_j h_k = o(\|h\|^2)$$

Cioè espandendo la defnzione di o-piccolo:

$$\lim_{h \to \underline{0}} \frac{(\partial_{jk} f(x_0 + \theta h) - \partial_{jk} f(x_0)) h_j h_k}{\|h\|^2} = 0$$

Notiamo che $h_j h_k \leq \|h\|^2$ in quanto $h_j h_k \leq h_1^2 + \cdots + h_j^2 + \cdots + h_k^2 + \cdots + h_n^2$. Infatti ci basta provare che $h_j h_k \leq h_j^2 + h_k^2$. Cioè che $h_j^2 + h_k^2 - h_j h_k \geq 0 \implies 2h_j^2 + 2h_k^2 - 2h_j h_k \geq 0 \implies h_j^2 + h_k^2 + (h_j - h_k)^2 \geq 0$. Ovvio quindi in quanto tutti i termini sono di secondo grado e quindi positivi. Possiamo quindi dire che:

$$\lim_{h \to \underline{0}} \frac{(\partial_{jk} f(x_0 + \theta h) - \partial_{jk} f(x_0)) h_j h_k}{\|h\|^2} \leq \lim_{h \to \underline{0}} \frac{(\partial_{jk} f(x_0 + \theta h) - \partial_{jk} f(x_0)) \|h\|^2}{\|h\|^2}$$

$$\lim_{h \to \underline{0}} (\partial_{jk} f(x_0 + \theta h) - \partial_{jk} f(x_0))$$

Essendo la derivata seconda continua:

$$\lim_{h \to \underline{0}} (\partial_{jk} f(x_0 + \theta h) - \partial_{jk} f(x_0)) = 0$$

Qed.

Il **polinomio di Taylor di grado due** con punto iniziale $\mathbf{x_0}$:

$$T_2(x) = f(x_0) + \langle \nabla f(x_0), (x - x_0) \rangle + \frac{1}{2} \langle \mathbf{H} f(x_0)(x - x_0), (x - x_0) \rangle$$

7 Punti critici

Per trovare i punti critici di una funzione non possiamo usare lo stesso approccio usato per le funzioni ad una varibiale in quanto in $\mathbb{R}^n$ non riusciamo a definire una vettore maggiore di zero e quindi non sappiamo distinguere funzioni crescenti e decrescenti. Esiste però un altro approccio per determinare se un punto è di minimo o di massimo che si basa sulle derivata seconda in un punto. Se infatti prendiamo una funzione $f : \mathbb{R} \to \mathbb{R}$ con derivata seconda continua possiamo dire che un punto x_0 è un punto di minimo se:

$$\begin{cases} f'(x_0) = 0 \\ f''(x_0) > 0 \end{cases}$$

Vale invece che se $f''(x_0) < 0$ il punto è di massimo mentre se $f''(x_0) = 0$ il punto è di flesso. Per definire quindi i punti di massimo, di minimo e di sella per le funzioni in più variabili useremo un approccio molto simile a questo.

Definizione

Punto di massimo o minimo: Dato $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$ e un punto $x_0 \in A$. Il punto x_0 si dice di minimo per f se:

$$\exists \delta > 0 : f(x) \geq f(x_0) \;\; \forall x \in A \cap \mathcal{B}(x_0, \delta)$$

Per i massimi la condizione risulta essere: $f(x) \leq f(x_0)$. Inoltre x_0 è un minimo assoluto se:

$$f(x) \geq f(x_0) \quad \forall x \in A$$

Teorema

Condizione necessaria del primo ordine per punti di massimo o di minimo: Dato $A \subseteq \mathbb{R}^n$ aperto e una funzione $f : A \to \mathbb{R}$ differenziabile. Se $x_0 \in A$ è un punto di massimo (o di minimo) locale allora:

$$\nabla f(x_0) = \underline{0} \in \mathbb{R}^n$$

Se vale questa condizione il punto x_0 si dice **punto critico** (o **stazionario**).

Da notare che questa condizione è estremamente simile a quella che si verificava nelle funzioni ad una varibiale in quanto il gradiente in $\mathbb{R}^1$ è semplicemente la derivata prima della funzione.

Dimostrazione

La prova la riduciamo solo al caso di $\mathbb{R}^2$. Prendiamo $A \subseteq \mathbb{R}^2$ aperto e una funzione $f : A \to \mathbb{R}$. Assumiamo inoltre che il punto (x_0, y_0) si un punto di minimo. Dimostriamo ora che:

$$\nabla f(x_0, y_0) = \underline{0} \in \mathbb{R}^n$$

In pratica ci riduciamo a dimostrare che:

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \wedge \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0$$

Riduciamoci a dimostrare solo la derivata rispetto ad x in quanto l'altro caso è analogo. Per farlo consideriamo la funzione ausiliaria:

$$h(t) = f(x_0 + t, y_0)$$

Essa è definita per t in un intorno dell'origine ed è derivabile. Visto che f ha minimo locale in

(x_0, y_0) per la definzione di punto di minimo:

$$h(t) \geq h(0) \quad \forall t \text{ in un intorno di } 0$$

Notiamo infatti che $h(0) = f(x_0, y_0)$. Inoltre:

$$h'(t) = \partial_x f(x_0 + t, y_0)$$

Per il teorema del primo sementre, in quanto h è una funzione ad una varibiale che ha un minimo in 0:

$$h'(0) = 0$$

Ne consegue che:

$$\partial_x f(x_0 + t, y_0) = 0$$

Qed.

7.1 Teorema classificazione punti critici

Teorema

Sia $f : A \to \mathbb{R}$ una funzione con derivate prime e seconde continue sull'insieme aperto $A \subseteq \mathbb{R}^2$ e sia x_0 un punto critico (quindi $\nabla f(x_0) = \underline{0}$). Allora:

- Se $\mathbf{H}f > 0$, allora x_0 è un punto di minimo locale.
- Se $\mathbf{H}f < 0$, allora x_0 è un punto di massimo locale.
- Se $\mathbf{H}f$ è *indefinita*, allora x_0 è un punto di sella.

Prima di della dimostrazione è necesarrio dimostrare un lemma:

Lemma

Se $A = A^T \in \mathbb{R}^{n \times n}$ è definita positiva, allora:

$$\exists m : \quad \langle Ah, h \rangle \geq m \|h\|^2 \quad \forall h \in \mathbb{R}^n$$

Dimostrazione: Fissiamoci in due dimensioni ed usiamo le coordinate polari. In particolare il vettore h si può scrivere nella forma $h = (r \cos\theta, r \sin\theta)$ con $r > 0$ e $\theta \in [0, 2\pi]$. Mettiamo che A sia nella forma $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Risulta quindi:

$$\langle Ah, h \rangle = r^2 (a \cos^2(\theta) + 2b \cos(\theta) \sin(\theta) + c \sin^2(\theta))$$

Notiamo che $r^2 = \|h\|^2$. Definiamo inoltre quanto appena ottenuto come una funzione g rispetto al parametro θ:

$$g(\theta) := (a \cos^2(\theta) + 2b \cos(\theta) \sin(\theta) + c \sin^2(\theta))$$

La funzione $g(\theta) > 0 \quad \forall \theta \in [0, 2\pi]$ ed inoltre è continua. Indicando quindi con $\theta_0 \in [0, 2\pi]$ un suo punto di minimo assoluto e con $m := g(\theta_0) > 0$ il suo minimo assoluto si ottiene:

$$\langle Ah, h \rangle = \|h\|^2 g(\theta) \geq m \|h\|^2$$

Qed.

Dimostriamo ora solo il primo caso, cioè quello del punti di minimo:

Dimostrazione

Sia $f : A \to \mathbb{R}$ con $A \subseteq \mathbb{R}$ aperto. Sia $x_0 \in A$ tale che $\nabla f(x_0) = 0$ e $\mathbf{H}f(x_0) > 0$. Dimostriamo che x_0 è un minimo locale, cioè:

$$\exists \delta > 0 : f(x_0 + h) \geq f(x_0) \quad \forall h \in B(0, \delta)$$

Possiamo usare la fomula di Taylor di secondo ordine ricordandoci che il gradiente è nullo:

$$f(x_0 + h) - f(x_0) = \frac{1}{2}\langle \mathbf{H}f(x_0)h, h\rangle + o(\|h\|^2)$$

Dal lemma:

$$\exists m > 0 : \langle \mathbf{H}f(x_0)h, h\rangle \geq m\|h\|^2 \quad \forall h \in \mathbb{R}^2$$

Dunque:

$$f(x_0 + h) - f(x_0) \geq \frac{m}{2}\|h\|^2 + o(\|h\|^2) \quad \text{per } h \to 0$$

Per defінzione di o-piccolo (con $\epsilon = \frac{m}{4}$):

$$\exists \delta > 0 : \forall \langle h\rangle < \delta \implies |o(\|h\|^2)| \leq \frac{m}{4}\|h\|^2$$

Notiamo che $\forall \langle h\rangle$ è come dire $\forall h \in B(0, \delta)$. Eliminando il valore assoluto prendendo solo la parte negativa:

$$o(\|h\|^2) \geq -\frac{m}{4}\|h\|^2 \quad \forall h \in B(0, \delta)$$

Risostituendo nella formula precedente:

$$f(x_0 + h) - f(x_0) \geq \frac{m}{2}\|h\|^2 - \frac{m}{4}\|h\|^2 \quad \forall h \in B(0, \delta)$$

Che semplificando diventa:

$$f(x_0 + h) - f(x_0) \geq \frac{m}{4}\|h\|^2 \quad \forall h \in B(0, \delta)$$

Essendo tutta la parte di destra positiva, abbiamo dimostrato che x_0 è un minimo. Qed.

8 Integrali in due variabili

8.1 Domini semplici

In più variabili i domini della funzioni possono essere molto irregolari. Nonostante ciò alcune funzioni presentano gradi di regolarità per cui risulta possibile descrivere il loro dominio attraverso intervalli e grafici di funzioni. Si parla quindi di **dominio semplice**. Questi gradi di regolarità sono estremamente utili nella formula di riduzione degli integrali doppi e in generali anche per molti teoremi che però non vengono trattati nel corso.

Definizione

Siano $g_1, g_2 : [a,b] \to \mathbb{R}$ continue con $g_1(x) \leq g_2(x) \forall x \in [a,b]$. Il **dominio x-semplice**[a] (o insieme x-semplice) è definito come:

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

[a]Nota: gli insiemi x-semplici e y-semplici sono estremamente simili, quindi è facile confodersi (anche il prof lo fa). Quindi se in altri appunti vedete le defnzioni con i nome invertiti non preoccupativi anche perché non è importante il nome.

Un insieme y-semplice è estremamente simile, solo che si scambiano le variabili x e y:

Definizione

Siano $h_1, h_2 : [c,d] \to \mathbb{R}$ continue con $h_1(y) \leq h_2(y) \forall y \in [c,d]$. Il **dominio y-semplice** (o insieme y-semplice) è definito come:

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

8.2 Proprietà integrali

Tutte le proprietà degli integrali viste in una variabile si conservano anche in due. Data f continua e A semplice:

1. Linearità:
$$\int_A (\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 \int_A f_1 + \lambda_2 \int_A f_2$$
2. Se A è un insime degenere (cioè una linea e quindi vale che $g_1(x) \leq g_2(x) \forall x$) allora:
$$\int_a f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0$$
3. Se la funzione è costante con valore 1 allora l'integrale è l'area di A:
$$\int_A 1\,\mathrm{d}x\,\mathrm{d}y = \text{Area di } A$$

8.3 Formula di riduzione

Dato $A = \{(x,y) \in \mathbb{R}^2 | x \in [a,b], g_1(x) \leq y \leq g_2(x)\}$. Data inoltre $f : A \to \mathbb{R}$ continua, vale la formula:

$$\int_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

Se l'insieme invece che essere x-semplice era y-semplice valeva la stessa formula con la variabili ivertite, cioè se $A = \{(x, y) \in \mathbb{R}^2 | y \in [c, d], h_1(y) \leq x \leq h_2(y)\}$ allora:

$$\int_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x,y)\,\mathrm{d}x \right) \mathrm{d}y$$