Analisi (M2)
Appunti

Alex Bastianini

Contents

Chapter 1

Introduzione agli appunti

1.1 Le varie parti degli appunti

Diversi box colorati per indicare diverse parti degli appunti:

Definition 1.1.1: Definizione

Theorem 1.1.1 Teorema

Dimostrazione: 😄

Corollary 1.1.1 Corollario

Note:

Una roba importante

Chapter 2

Integrali

2.1 Motivazioni

Motivazioni:

- Calcolo di aree di figure curvilinee
- Lunghezze di curve (non lo faremo)

Le nostre figure curvilinee sono sottografici di funzioni.

2.2 Area sottesa a una curva

Definition 2.2.1: Area sottesa

Data $f : [a, b] \longrightarrow \mathbb{R}$
$$A = \{(x, y) \in \mathbb{R}^2 | x \in [a, b]. 0 \leqslant y \leqslant f(x)\}$$

2.2.1 Costruzione integrale di Riemann

Speziamo un intervallo [a,b] in $n \in \mathbb{N}$ sottointervalli uguali. L'ampiezza di ciuascun intervallo e' di $\frac{b-a}{n}$.

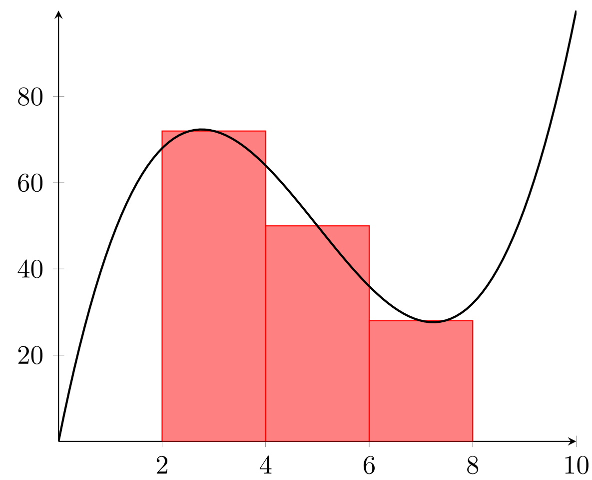
- $x_0 = a$
- $x_1 = a + \frac{b-a}{n}$
- $x_n = b$

In ogni intervallo fisso un punto arbirario ϵ_n

Definition 2.2.2: Somma di Riemann associata a una scomposizione

Data una funzione $f : [a, b] \longrightarrow \mathbb{R}$, fatta la costruzione precedente (spezzettamento), $\forall n \in \mathbb{N}$ la somma di Riemann n-esima di f e' il numero seguente:

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



(2.1)

Esempio di somma di Riemann di una funzione $f : [2, 8] \to \mathbb{R}$ con n=3:

Theorem 2.2.1 Integrabilita' delle funzioni continue

Sia $f : [a, b] \to \mathbb{R}$ continua. Sia (S_n) la successione delle somme di Riemann, allora:

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

E non dipende da quale ϵ_n scegliamo per ogni segmento.

Definition 2.2.3: Integrale

$$\int_a^b f(x)dx = \lim_{n \to +\infty} S_n$$

La x e' una variabile **muta**.

Note:

- Se $\forall x \in [a, b]. f(x) \geqslant 0$, allora $\int_a^b f$ = Area del sottografico.
- $\int_a^a f = 0$ (poiche $\forall n \in \mathbb{N}. S_n = 0$)
- $f(x) = k$(funzione costante) $\implies \int_a^b f = (b - a)k$(Area del rettangolo)

2.2.2 Proprieta' dell'integrale

Linearita'

Se abbiamo due funzioni f, g continue su [a,b], $A, \mu \in \mathbb{R}$

$$\int_a^b (Af(x) + \mu g(x))dx = A \int_a^b f(x)dx + \mu \int_a^b g(x)dx$$

Additivita

Data $f : [a, b] \to \mathbb{R}$ continua, dato $c \in [a, b]$, vale:

$$\int_a^b f = \int_a^c f + \int_c^b f$$

Note:

Convenzione su integrali con estremi "rovesciati":
Dato $f : [a,b] \to \mathbb{R}$
$$\int_a^b f = -\int_b^a f$$

In questo modo possiamo generalizzare la proprieta' addittiva togliendo dall' ipotesi la restrizione sul valore di c.

Monotonia

Se $f : [a,b] \to \mathbb{R}$ e $\forall x \in [a,b].f(x) \geqslant 0$, allora:
$$\int_a^b f \geqslant 0$$

2.2.3 Media Integrale

Premessa 1

Theorem 2.2.2 Valori intermedi
Sia $f : [a,b] \to \mathbb{R}$ continua, $x_1, x_2, \in [a,b].f(x_1) \leqslant f(x_2)$, allora:
$$\forall y \in [f(x_1), f(x_2)].\exists c \in [x1, x2].f(c) = y$$

Premessa 2

Theorem 2.2.3 Weierstrass
Sia $f : [a,b] \to \mathbb{R}$ continua:
$$\exists x_1, x_2 \in [a,b].\forall x \in [a,b].f(x_1) \leqslant f(x) \leqslant f(x_2)$$

Quindi una funzione continua su dominio chiuso e limitato ha massimo e minimo assoluti.

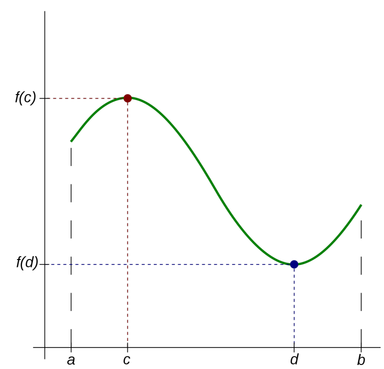
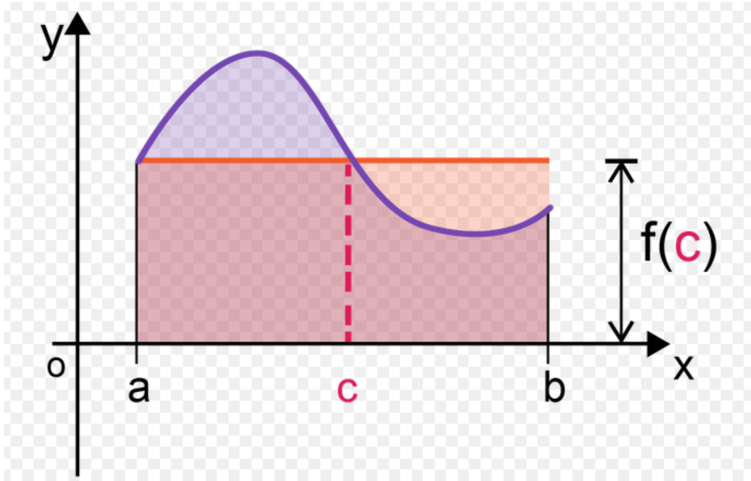


Figure 2.1: Esempio di Weierstrass

Theorem 2.2.4 Media integrale
Sia $f : [a,b] \to \mathbb{R}$ continua, allora:
$$\exists c \in [a,b].\frac{1}{b-a}\int_a^b f(x)dx = f(c)$$



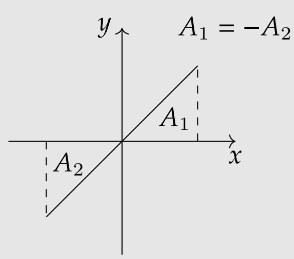
Quindi esiste un punto c in [a,b] t.c. il rettangolo che ha come base b-a e come altezza f(c) ha la stessa area dell'integrale di f da a a b.

Dimostrazione della media integrale: Sia f continua su [a,b]. Per Weierstrass abbiamo che $\exists x_1, x_2 \in [a,b].\forall x \in [a,b].f(x_1) \leqslant f(x) \leqslant f(x_2)$. Per la proprieta' di monotonia risulta $\int_a^b f(x_1)dx \leqslant \int_a^b f(x)dx \leqslant \int_a^b f(x_2)dx$, ovvero $f(x_1) \leqslant \frac{1}{b-a}\int_a^b f(x)dx \leqslant f(x_2)$. Quindi per il teorema dei valori intermedi $\exists c \in [a,b].f(c) = \frac{1}{b-a}\int_a^b f(x)dx$. 😄

Note:

La continuita' di f e' **necessaria**. Ex:

$$f : [-1,1] \to \mathbb{R}, f(x) = x \implies \frac{1}{2}\int_{-1}^{1} x dx = 0 = f(0)(c = 0)$$



Se considerassi $g : [-1,1] \to \mathbb{R}$

$$g(x) = \begin{cases} x & x \neq 0 \\ 2 & x = 0 \end{cases}$$

Si dimostra che g e' intagarbile, e che vale

$$\int_{-1}^{1} g(x)dx = 0$$

Pero' non esiste $c \in [-1,1]$ tale che g(c) = 0, quindi non soddisfa la media integrale.

2.3 Primitive di una funzione

Definition 2.3.1: Primitiva di f

Sia $f : A \to \mathbb{R}, A \subseteq \mathbb{R}$
Una funcione $F : A \to \mathbb{R}$ si dice primitiva di f su A se vale

$$\forall x \in A. F'(x) = f(x)$$

Example 2.3.1 (Primitiva)
$f : \mathbb{R} \to \mathbb{R}$, $f(x) = \cos x \implies F(x) = sinx$ e' una primitiva di f su $\mathbb{R}$. Infatti $\forall x. \frac{d}{dx} sinx = cosx$

Note:
$\forall k \in \mathbb{R}$, la funzione $G(x) = sin(x) + k$ e' anchessa primitiva di f. Quindi se F e' primitiva di f su A allora ci sono infinite primitive di f su A (una per ogni $k \in \mathbb{R}$).

Domanda: sono tutte le possibili primitive?

Proposition 2.3.1 Caratterizzazione delle primitive di una funzione su un intervallo
Sia $f :]a, b[\to \mathbb{R}$. Siano $F :]a, b[\to \mathbb{R}$ e $G :]a, b[\to \mathbb{R}$ due primitive di f su $]a, b[$.
Allora $\exists k \in \mathbb{R}$ tale che:

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

Ovvero F e G "differiscono per una costante".

Dimostrazione: Considero $H :]a, b[\to \mathbb{R}, H(x) = G(x) - F(x)$. Sappiamo che F'(x)=f(x) e G'(x)=f(x) (def. primitiva). $\frac{d}{dx} H(x) = \frac{d}{dx} G(x) - \frac{d}{dx} F(x) = f(x) - f(x) = 0$. Dunque H ha derivata nulla su $]a, b[$, quindi (per coroll. Lagrange) H e' costante. 😄

2.4 Funzioni integrali

D'ora in poi $A =]a, b[$.

Definition 2.4.1: Funzione Integrale

Sia $f : A \to \mathbb{R}$ continua.
Sia $c \in A$. Introduco $I_c : A \to \mathbb{R}$:

$$I_c(x) = \int_c^x f(t)dt$$

Nota: I_c e' ben definita essendo f continua.

Note:

1. $f : A \to \mathbb{R}, c \in A, I_c(x) = \int_c^x f \implies I_c(c) = \int_c^c f(t)dt = 0$.
2. Dati $c, c' \in A, f : A \to \mathbb{R} \implies I_c(x) - I_{c'}(x) =$ costante. Infatti:

$$I_c(x) - I_{c'}(x) = \int_c^x f(t)dt - \int_{c'}^x f(t)dt = \int_c^x f(t)dt + \int_x^{c'} f(t)dt = \int_c^{c'} f(t)dt = k$$

2.4.1 Teoremi fondamentali del calcolo integrale

Theorem 2.4.1 Fondamentale del calcolo sulla derivata della funzione integrale

Sia $f : A \to \mathbb{R}$ continua, $c \in A$. Sia I_c la funzione integrale, allora:

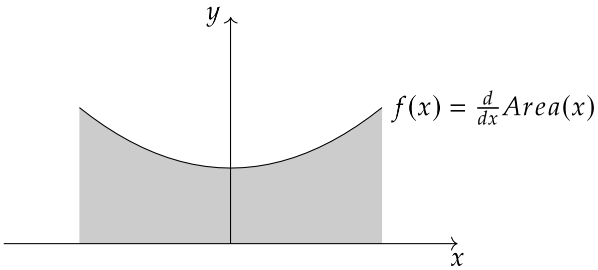
$$I_c \text{ e' derivabile in ogni punto } x \in A \text{ e } I_c'(x) = f(x)$$

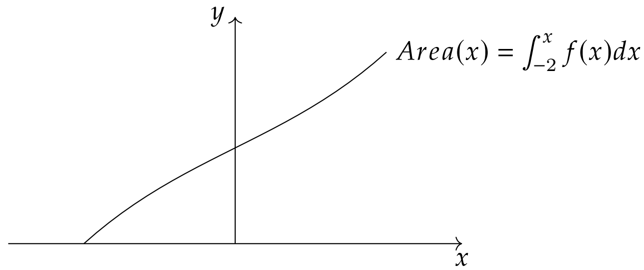
Cioe' $\frac{d}{dx}\int_c^x f(t)dt = f(x), \forall x \in A$, quindi I_c e' **primitiva** di $f(x)$.

Una possibile interpretazione di questo teorema e' quello della derivata dell'area sottesa che e' uguale alla funzione stessa, ovvero:

$$\forall x. f(x) \geqslant 0 \implies \frac{d}{dx}\text{Area} = f(x)$$

$A =]-2, 2[, f(x) = \frac{x^2}{5} + 1$:





Note:
Il teorema assicura che ogni funzione $f : A \to \mathbb{R}$ continua ammette primitive.

Dimostrazione: $f : A \to \mathbb{R}, c \in A, I_c : A \to \mathbb{R}$. Devo calcolare la derivata di I_c. Calcolo $\lim_{h\to 0^+} \frac{I_c(x+h) - I_c(x)}{h}$, che equivale a $\lim_{h\to 0^+} \frac{1}{h}\int_x^{x+h} f(t)dt$. Per teo. media integrale sappiamo che $\exists c \in [x, x+h]. f(c) = \frac{1}{h}\int_x^{x+h} f(t)dt$, quindi possiamo riscrivere la formula come $\lim_{h\to 0^+} f(c_h)$. Dato che $h \to 0^+$ e $x \leqslant c \leqslant x+h$, per il teo. dei carabinieri $c = x$, quindi (grazie alla continuita' di f), $\lim_{h\to 0+} f(c_h) = f(x)$. ☺

Theorem 2.4.2 Fondamentale del calcolo per integrali definiti

Sia $f : A \to \mathbb{R}$ continua su A. Sia $F : A \to \mathbb{R}$ primitiva di f su A. Dati $a, b \in A$, vale:

$$\int_a^b f(x)dx = [F(b) - F(a)] = [F(x)]_a^b$$

Dimostrazione: Sia $c \in A$, $I_c : A \to \mathbb{R}$ la funzione integrale ($I_c(x) = \int_c^x f$). Per il teo. fond. del calc. sulla derivata di I_c, I_c e' una primitiva di f su A. Per le proprieta' delle primitive, $\exists k \in \mathbb{R}. \forall x \in A. F(x) = I_c(x) + k$.

$$F(b) - F(a) = (I_c(b) + k) - (I_c(a) + k) = I_c(b) - I_c(a) = \int_c^b f - \int_c^a f = \int_a^c f + \int_c^b f = \int_a^b f$$

☺

2.4.2 Integrazione per parti

Si parte dalla regola del prodotto delle derivate ($\frac{d}{dx}f(x) \cdot g(x) = f'(x)g(x) + f(x)g'(x)$) per trovare una regola di integrazione.

Theorem 2.4.3 Integrazione per parti

Dati $f, g : A \to \mathbb{R}$, A intervallo aperto e sia F primitiva di f su A con F,f,g continue, g derivabile e g' continua:

$$\int_a^b \frac{d}{dx}(F(x)g(x))dx = \int_a^b f(x)g(x)dx + \int_a^b F(x)g'(x)dx$$

Quindi usando il teorema fondamentale:

$$[F(x)g(x)]_a^b = \int_a^b f(x)g(x)dx + \int_a^b F(x)g'(x)dx$$

$$\int_a^b f(x)g(x)dx = [F(x)g(x)]_a^b - \int_a^b F(x)g'(x)dx$$

2.4.3 Cambio di variabile

Theorem 2.4.4 Formula del cambio di variabile

Date $h : I \to A$, $f : A \to \mathbb{R}$, $I, A \subseteq \mathbb{R}$ e $\exists (f \circ h) : I \to \mathbb{R}. (f \circ h)(t) = f(h(t))$. f continua, h derivabile e h' continua. Presi $\alpha, \beta \in I$, vale:

$$\int_{h(\alpha)}^{h(\beta)} f(x)dx = \int_\alpha^\beta f(h(t))h'(t)dt$$

Questa e' la versione generalizzata del teo. fond. del calcolo

Dimostrazione: Date due funzioni $G, H : I \to \mathbb{R}. G(z) = \int_\alpha^z f(h(t))h'(t)dt, H(z) = \int_{h(\alpha)}^{h(z)} f(x)dx$, dobbiamo dimostrare che $G(z) = H(z)$. Ci riduciamo a dimostrare che:

1. $G(\alpha) = H(\alpha)$: ovvio perche' integrali su intervallo degenere ($G(\alpha) = H(\alpha) = 0$)
2. $\forall z \in I. G'(z) = H'(z)$:
   - $G'(z) = \frac{d}{dz} \int_\alpha^z f(h(t))h'(t)dt = f(h(z))h'(z)$
   - $H'(z) = \frac{d}{dz} \int_{h(\alpha)}^{h(z)} f(x)dx = f(h(z))h'(z)$ (generalizzazione del teorema fondamentale del calcolo)

😄

Note:

1. t integrata in α, β, allora x sara' integrata in $h(\alpha), h(\beta)$.
2. dx si e' trasformato in $h'(t)dt$ ($\frac{d}{dt}h(t) = h'(t) \implies dh(t) = h'(t)dt$)

2.5 Integrali generalizzati (impropri)

Definition 2.5.1: Integrali generalizzati su intervalli illimitati

Sia $f : [a, +\infty[\to \mathbb{R}$. Si dice che l'integrale generalizzato $\int_a^{+\infty} f(x)dx$ e' **convergente** se e' finito il limite $\lim_{r\to+\infty} \int_a^r f(x)dx := \int_a^{+\infty} f(x)dx$. Altrimenti se il limite diverge o oscilla e' detto **divergente** (o oscillante).
La definizione e' analoga per $\int_{-\infty}^a f(x)dx$.

Definition 2.5.2: Integrali generalizzati su intervalli limitati

Sia $f :]a, b] \to \mathbb{R}$. Si dice che l'integrale $\int_a^b f(x)dx$ e' **convergente** se il limite $\lim_{r \to a^+} \int_r^b f(x)dx$ e' finito.
Altrimenti se il limite diverge o oscilla e' detto **divergente** (o oscillante).
La definizione e' analoga per $f : [a, b[\to \mathbb{R}$.

Chapter 3

Spazio euclideo $\mathbb{R}^n$

(Spazio **euclideo** = c'e' il prodotto scalare)
$\mathbb{R}^n = \{x = (x_1, x_2, ..., x_n) | \forall j \in \{1, 2, ..., n\}.x_j \in \mathbb{R}\} = \mathbb{R} \times \mathbb{R} \times ...\mathbb{R}$ (n volte).

- $n = 1$ retta reale
- $n = 2$ piano cartesiano
- $n = 3$ spazio ordinario

3.1 Operazioni su $\mathbb{R}^n$

- Somma di vettori: $x = (x_1, ..., x_n), y = (y_1, ..., y_n)$. Definiamo $x + y = (x_1 + y_1, ..., x_n + y_n) \in \mathbb{R}^n$
- Prodotto di $x \in \mathbb{R}^n$ per uno scalare $\lambda \in \mathbb{R}$. Dato $x = (x_1, ..., x_n), \lambda \in \mathbb{R}$, definiamo $\lambda x = (\lambda x_1, ..., \lambda x_n)$.

Definition 3.1.1: Prodotto scalare

Dati $x, y \in \mathbb{R}_n$, definiamo il prodotto scalare $< x, y >= \sum_{k=1}^{n} x_k y_k = x_1 y_1 + ... + x_n y_n \in \mathbb{R}^n$.
Notazione alternativa: $< x, y >= x \cdot y$.

Note:
Il prodotto scalare non da' un nuovo vettore, ma solo un valore scalare!

3.1.1 Proprieta' del prodotto scalare (euclideo)

- $\forall x, y \in \mathbb{R}^n. < x, y >=< y, x >$ (simmetria)
- $\forall x, y, z \in \mathbb{R}^n, \lambda, \mu \in \mathbb{R}. < \lambda x + \mu y, z >= \lambda < x, z > + \mu < y, z >$ (linearita' nel primo argomento)
  Per simmetria vale $< z, \lambda x + \mu y >= \lambda < z, x > + \mu < z, y >$ (linearita' nel secondo argomento)
- $\forall x \in \mathbb{R}^n. < x, x >\geqslant 0$, inoltre $< x, x >= 0 \iff x = (0, 0, ..., 0) = \underline{0}$ (vettore nullo).

3.2 Ortogonalita'

Definition 3.2.1: Vettori ortogonali

Due vettori $x, y \in \mathbb{R}^n$ si dicono **ortogonali** se vale:

$$< x, y >= 0$$

3.3 Norma euclidea

Sinonimi: modulo, lunghezza

Definition 3.3.1

Dato $x \in \mathbb{R}^n$, allora

$$\|x\| = \sqrt{< x, x >} \geqslant 0$$

Rappresenta la "lunghezza" del vettore usando il teorema di Pittagora.
Notazione alternativa: $|x|$

3.3.1 Proprieta' della norma

- $\forall x \in \mathbb{R}^n . \|\lambda x\| = |\lambda| \cdot \|x\|$
- $\forall x \in \mathbb{R}^n . \|x\| = 0 \iff x = 0$
- $\forall x, y \in \mathbb{R}^n . |x + y| \leqslant |x| + |y|$ (Disuguaglianza triangolare)

3.4 Vettore normalizzato

Definition 3.4.1: Normalizzato

Dato $x \neq 0 \in \mathbb{R}^n$, cerco $r > 0$ t.c.

$$|rx| = 1$$

Visto che $r > 0$, $r|x| = 1$ quindi $r = \frac{1}{|x|}$.
Il vettore $\frac{x}{|x|}$ ha norma 1 e si dice **normalizzato** di $x \in \mathbb{R}^n \setminus \{\underline{0}\}$.
$\frac{x}{|x|}$ si dice **vettore unitario**.

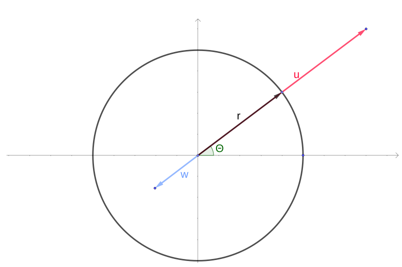
Note:
Possiamo scrivere $x = |x| \cdot \frac{x}{|x|}$ se $x \neq 0$.

3.5 Coordinate polari

In $\mathbb{R}^2$, ogni $(x, y) \neq (0, 0)$ si scrive nella forma $|(x, y)| \cdot (\frac{x}{|(x,y)|}, \frac{y}{|(x,y)|})$. L'insieme di coordinate $\{(\frac{x}{|(x,y)|}, \frac{y}{|(x,y)|}) | x, y \in \mathbb{R}\}$ forma una **circonferenza unitaria** (dato che il loro modulo e' sempre 1), quindi $\exists \theta \in [0, 2\pi[$ tale che $(cos\theta, sin\theta) = (\frac{x}{|(x,y)|}, \frac{y}{|(x,y)|})$ (θ si chiama **"argomento"** di (x, y)). Ponendo $r = |(x, y)| =$ modulo, scriviamo:

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

dove $r > 0$ e $\theta \in [0, 2\pi[$ si chiamano **coordinate polari** di (x,y).



- r = vettore unitario

3.5.1 Prodotto scalare in coordinate polari

Considero due vettori $(x,y) = (rcos\theta, rsin\theta)$ e $(n,o) = (\rho cos\gamma, \rho sin\gamma)$. Il loro prodotto scalare $<(x,y),(n,o)>$ diventa $r\rho cos\theta cos\gamma + r\rho sin\theta sin\gamma$, che puo' essere riscritto come $r\rho cos(\gamma - \theta)$, ovvero:

$$|(x,y)| \cdot |(n,o)| \cdot cos(\gamma - \theta)$$

Proposition 3.5.1 Disuguaglianza di Cauchy-Schwarz

Dati $x, y \in \mathbb{R}^n$, vale:

$$| < x, y > | \leqslant \|x\| \cdot \|y\|$$

L'uguaglianza vale solo se x e y sono linearmente dipendenti.

Note:

Vale in ogni $\mathbb{R}^n \forall n \in \mathbb{N}$

Proposition 3.5.2 Quadrato di binomio

$$|x+y|^2 = |x|^2 + |y|^2 + 2 < x, y >$$

Generalizzazione di Pitagora (In due dimensioni diventa il teorema di Carneau).

Proposition 3.5.3 Disuguaglianza Triangolare

$\forall x, y \in \mathbb{R}^n$ si ha che:

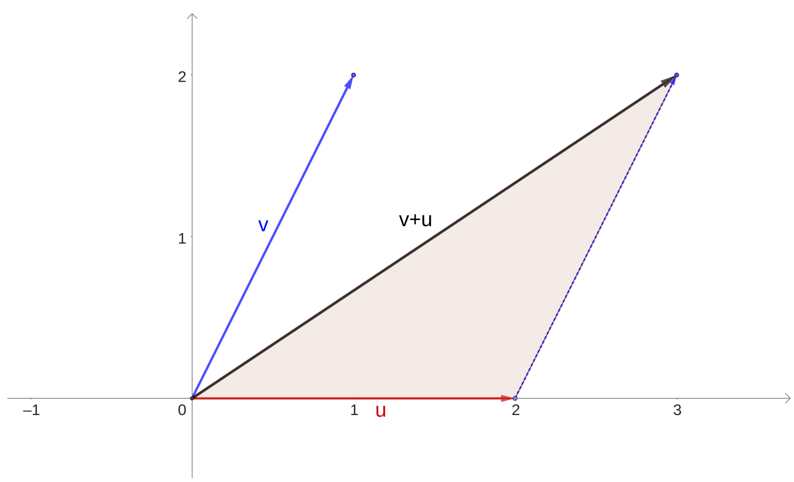
$$|x+y| \leqslant |x| + |y|$$

: Dimostriamo il quadrato della disuguaglianza per poter usare il quadrato di binomio:

$$|x+y|^2 = |x|^2 + |y|^2 + 2 < x,y > \leqslant |x|^2 + |y|^2 + 2| < x,y > | \leqslant |x|^2 + |y|^2 + 2|x||y| = (|x|+|y|)^2$$

$$\implies |x+y|^2 \leqslant (|x|+|y|)^2$$

😊



3.6 Distanza tra punti

Definition 3.6.1: Distanza fra due punti

Dati $x, y \in \mathbb{R}^n$, la distanza fra x e y e' $|x-y|$

3.6.1 Punto di minima distanza da una retta

Problema $\mathbb{R}^n, v \neq 0, x \in \mathbb{R}^n$. Considero le linee $l_v = \{tv | t \in \mathbb{R}\}$, cerco fra tutti i punti di l_v quello che ha minima distanza da x. Devo minimizzare la funzione $h : \mathbb{R} \to \mathbb{R}, h(t) = |x - tv| =$ distanza fra x e tv, che (essendo non negativa), equivale a minimizzare la funzione alla seconda $g(t) = |x - tv|^2$. Usando il quadrato di binomio, otteniamo $g(t) = |x|^2 + t^2|v|^2 - 2t\langle x, v\rangle$, ovvero una parabola con concavita' verso l'alto. Quindi il punto di minimo e' l'unico punto stazionario:

$$g'(t) = 2t|v|^2 - 2\langle x, v\rangle = 0 \iff t = \frac{\langle x, v\rangle}{|v|^2} = \bar{t}$$

Quindi basta moltiplicare $\bar{t}$ per v per trovare il punto di minima distanza:

Proposition 3.6.1 Punto di minima distanza da una retta

Dati $v \neq 0, x \in \mathbb{R}^n$ e la retta $l_v = \{tv | t \in \mathbb{R}\}$, il punto $\bar{t}v$ di minima distanza da x e':

$$\bar{t}v = \frac{\langle x, v\rangle}{|v|^2} v$$

E la distanza vale $\sqrt{|x|^2 - \frac{\langle x,v\rangle^2}{|v|^2}}$.

Inoltre, come possiamo intuire, il vettore della minima distanza da l_v a x e' ortogonale a v:

Proposition 3.6.2 Ortogonalita' distanza minima

Dati $v \neq 0$ e $x \in \mathbb{R}^n$, il punto di minima distanza $\bar{t}v$ soddisfa:

$$< x - \bar{t}v, v >= 0$$

Quindi il vettore che parte dal punto di minima distanza e arriva al punto x e' perpendicolare alla retta l_v.

Verifica: (Basta svolgere i calcoli) 😃

3.7 Intorni

Generalizziamo in $\mathbb{R}^n$ il concetto di intorno di un punto e di insiemi aperti/chiusi, che ci saranno utili per definire derivate in n dimensioni:

Definition 3.7.1: Intorno sferico di un punto

Dato $x \in \mathbb{R}^n$, $r > 0 \in \mathbb{R}$, poniamo

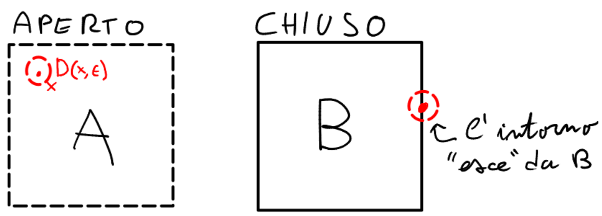
$$D(x, r) = \{y \in \mathbb{R}^n | |x - y| < r\}$$

$D(x, r)$ si dice disco di centro x e raggio r.

Definition 3.7.2: Insiemi aperti

$A \subseteq \mathbb{R}^n$ si dice aperto se:

$$\forall x \in A . \exists \epsilon > 0 . D(x, \epsilon) \subseteq A$$



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

Definition 3.8.1: Successione

Una funzione $x(k) : \mathbb{N} \longrightarrow \mathbb{R}^n$, indicata come x_k, si chiama successione in n dimensioni ed e' un vettore di k successioni:

$$x_k = (x_k^1, x_k^2, ..., x_k^n) \text{ con } k \in \mathbb{N}$$

Come in una dimensione, una successione in n dimensioni e' convergente se:

Definition 3.8.2: Successione convergente

Sia x_k una successione in $\mathbb{R}^n$ e $x \in \mathbb{R}^n$. Si dice che x_k **converge** a x se $\lim_{k \to +\infty} x_k = x$, ovvero se si verifica:

$$\begin{cases} \lim_{k \to +\infty} x_k^1 = x^1 \\ \vdots \\ \lim_{k \to +\infty} x_k^n = x^n \end{cases}$$

Proposition 3.8.1

Sia x_k una successione in $\mathbb{R}^n$ e sia $x \in \mathbb{R}^n$, si ha che:

$$\lim_{k \to +\infty} x_k = x \iff \lim_{k \to +\infty} \|x_k - x\| = 0$$

Note:

Per vedere se una successione in $\mathbb{R}^n$ e' convergente, basta studiare la convergenza di una successione in $\mathbb{R}$ $(\|x_k - x\|_k)$

Chapter 4

Funzioni a piu' variabili

Dati $A \subseteq \mathbb{R}^n, B \subseteq \mathbb{R}^q$, consideriamo funzioni $f : A \to B$ (A = dominio, B = codominio).
Casi modello:

- $f : \mathbb{R}^n \to \mathbb{R}$ (funzioni scalari)
- $f : \mathbb{R} \to \mathbb{R}^q$ (cammini in $\mathbb{R}^q$)

4.1 Insiemi di livello

Definition 4.1.1

$A \subseteq \mathbb{R}^n, f : A \to \mathbb{R}, b \in \mathbb{R}$. L'insieme di livello b di f e':

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

Se cammino lungo l'insieme di livello, la funzione corrispondente non cambia

4.2 Continuita'

Definition 4.2.1: Funzioni continue

$A \subseteq \mathbb{R}^n, f : A \to \mathbb{R}, k \in A$. Si dice che f e' continua in $k \in A$ se $\forall (x_n)_{n\in\mathbb{N}} \in \mathbb{R}^n$ vale:

$$\begin{cases} \forall n \in \mathbb{N}. x_n \in A \\ x_n \xrightarrow{n\to+\infty} k \end{cases} \implies f(x_n) \xrightarrow{n\to+\infty} f(k)$$

Proposition 4.2.1

Sia $A \subseteq \mathbb{R}^n, f : A \to \mathbb{R}, k \in A$. Allora f e' continua in k se:

$$\forall \epsilon > 0 \exists \delta > 0. \forall x \in A : |x - k| < \delta \implies |f(x) - f(k)| < \epsilon$$

Ovvero, $\lim_{x\to k} f(x) = f(k)$.

4.3 Derivata parziale

Definition 4.3.1: Derivata parziale

Data $f : A \to \mathbb{R}, A \subseteq \mathbb{R}^2$. Dati $(\overline{x}, \overline{y}) \in A$ f si dice derivabile parzialmente rispetto a $\overline{x}$ se:

$$\exists \lim_{h\to 0} \frac{f(\overline{x}+h, \overline{y}) - f(\overline{x}, \overline{y})}{h} = \frac{\partial f}{\partial x}(\overline{x}, \overline{y})$$

In modo analogo per $\frac{\partial f}{\partial y}(\overline{x}, \overline{y})$.

Definition 4.3.2: Gradiente

Se $f : \mathbb{R}^2 \to \mathbb{R}$ ammette derivate parziali $\forall (\overline{x}, \overline{y}) \in \mathbb{R}^2$, definiamo il **gradiente di f** come:

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

$\nabla f : \mathbb{R}^2 \to \mathbb{R}^2$ (funzione vettoriale)

Note:

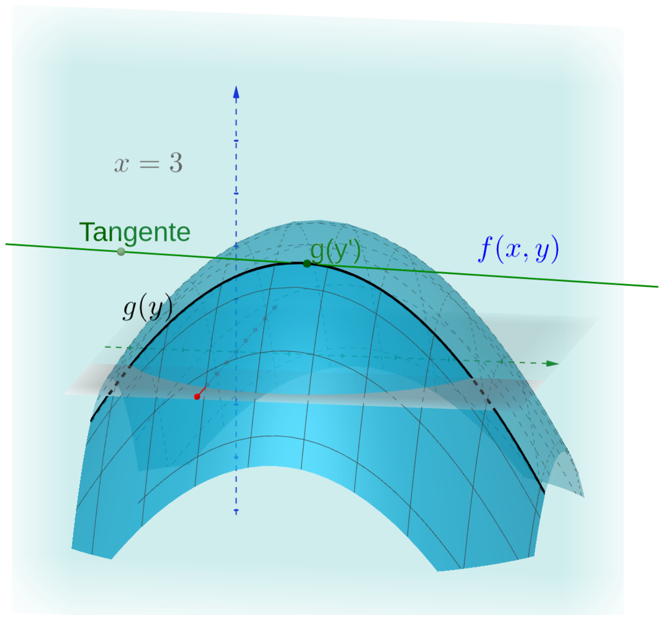
$f : \mathbb{R}^2 \to \mathbb{R}$, $(\overline{x}, \overline{y})$ fissato:

$$\partial_x f(\overline{x}, \overline{y}) = \lim_{h\to 0} \frac{f(\overline{x}+h, \overline{y}) - f(\overline{x}, \overline{y})}{h} = \lim_{x\to \overline{x}} \frac{f(x, \overline{y}) - f(\overline{x}, \overline{y})}{x - \overline{x}}$$

Introduco una funzione $g : \mathbb{R} \to \mathbb{R}$, $g(x) = f(x, \overline{y})$, in modo che, facendo la normale derivata di g:

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

Abbiamo quindi trasformato una derivata parziale in una derivata "normale" fissando tutti i parametri tranne uno.



4.3.1 Derivate parziali in $\mathbb{R}^n$

In $\mathbb{R}^n$, data $f : \mathbb{R}^2 \to \mathbb{R}$ possiamo riscrivere la derivata parziale cosi:

$$\frac{\partial f}{\partial x}(\overline{x}, \overline{y}) = \lim_{t \to 0} \frac{f((\overline{x}, \overline{y}) + t(1,0)) - f(\overline{x}, \overline{y})}{t}$$

Possiamo usare quindi le basi canoniche ($e_1 = (1, ..., 0), ..., e_n(0, ..., 1)$) per indicare per quale dei valori del vettore passato come parametro vogliamo derivare.

Definition 4.3.3

$f : \mathbb{R}^n \to \mathbb{R}$, $x = (x_1, ..., x_n), \overline{x} \in \mathbb{R}^n$. Per $k = \{1, ..., n\}$:

$$\frac{\partial f}{\partial x_k}(\overline{x}) = \lim_{t \to 0} \frac{f(\overline{x} + te_k) - f(\overline{x})}{t}$$

4.4 Derivabilita' e continuita'

In $\mathbb{R}$, se una funzione era derivabile in un punto allora era anche continua, pero' se in piu' variabili non e' cosi:(guardare es slide)

4.5 Differenziabilita'

In una dimensione, possiamo dire che:

$$\exists f'(\overline{x}) \in \mathbb{R} \iff f(\overline{x} + h) = f(\overline{x}) + f'(\overline{x})h + o(h)$$

Quindi una funzione e' derivabile in un punto sse vale lo sviluppo di Taylor. Infatti, se sostituiamo x a $\overline{x} + h$, dove $x \to \overline{x}$ otteniamo il polinomio di Taylor di primo grado nel punto $\overline{x}$: $f(x) = f(\overline{x}) + f'(\overline{x})(x - \overline{x}) + o(x - \overline{x})$. Come vedremo, questa prposizione non vale quando aumentiamo le dimensioni. Infatti, solo in una dimensione differenziabilita' e derivabilita' coincidono.

4.5.1 o piccolo in piu' variabili

Definition 4.5.1: o piccolo

Dati $A \subseteq \mathbb{R}^n, h \in A, g : A \to \mathbb{R}$, assumendo che $0 \in A$, si dice che g e' **o piccolo** di $|h|^p$ (con $p \geqslant 0$) se:

1. $g(0) = 0$
2. $\forall \epsilon > 0. \exists \delta > 0. \forall h \neq 0. |h| < \delta. \frac{|g(h)|}{||h||^p} < \epsilon$

Example 4.5.1

Verifica le seguenti uguaglianze:

- $g(h,k) = h^2 + k^2 = o(|h + k|)$

  $$h^2 + k^2 = |(h,k)|^2$$

  Quindi:

  $$\frac{|(h,k)|^2}{|(h,k)|} = |(h,k)|$$

  Dobbiamo dimostrare che $\forall \epsilon > 0. \exists \delta > 0. \forall 0 < |(h,k)| < \delta. |(h,k)| < \epsilon$, che possiamo fare mettendo $\delta = \epsilon$ per ogni $\epsilon > 0$.

Possiamo riscrivere questa definizione usando le successioni:

$$g(h) = o(|h|^p) \iff \forall (h_j)_{j\in\mathbb{N}}. \begin{cases} h_j \neq 0 & \forall j \in \mathbb{N} \\ \lim_{j\to+\infty}(h_j) = 0 \end{cases} \implies \lim_{j\to+\infty} \frac{g(h_j)}{||h_j||^p} = 0$$

Definition 4.5.2: Differenziabilita' in due variabili

$(x, y), (h, x) \in \mathbb{R}^2$. Sia $f : \mathbb{R}^2 \to \mathbb{R}$. Sia $(\overline{x}, \overline{y}) \in \mathbb{R}^n$. Si dice che f e' differenziabile in $(\overline{x}, \overline{y})$ se:

1. $\exists \partial_x f(\overline{x}, \overline{y}), \partial_y f(\overline{x}, \overline{y})$
2. Vale lo sviluppo:

$$f((\overline{x}, \overline{y}) + (h, k)) = f(\overline{x}, \overline{y}) + < \nabla f(\overline{x}, \overline{y}), (h, k) > + o(|(h, k)|) =$$

$$f(\overline{x}, \overline{y}) + \partial_x f(\overline{x}, \overline{y})h + \partial_y f(\overline{x}, \overline{y})k + o(|(h, k)|)$$

Per $(h, k) \to (0, 0)$

4.6 Continuita' di una funzione differenziabile

Sappiamo che l'esistenza delle derivate parziali non implica la continuita' della funzione. Mostreremo pero' che se sappiamo che una funzione e' differenziabile, allora sara' sicuramente continua.

Proposition 4.6.1 Differenziabilita' implica continuita'

Data $f : \mathbb{R}^2 \to \mathbb{R}$, se f e' differenziabile in $(\overline{x}, \overline{y}) \in \mathbb{R}^2$, allora f e' continua in $(\overline{x}, \overline{y})$.

: Verifichiamo che la proposizione valga. Per ipotesi, sappiamo che f e' differenziabile in $(\overline{x}, \overline{y}) \in \mathbb{R}^2$. Quindi $f(\overline{x} + h, \overline{y} + k) = f(\overline{x}, \overline{y}) + \langle \nabla f(\overline{x}, \overline{y}), (h, k)\rangle + o(|h, k|)$ (TODO: capire perche' questa vale solo quando (h,k) tende a (0,0)). Dobbiamo dimostrare che f sia continua in $(\overline{x}, \overline{y})$, ovvero che (usando la definizione di continuita' per successioni): $\forall (h_j, k_j) : \mathbb{N} \to \mathbb{R}^2 . (h_j, k_j) \xrightarrow{j\to+\infty} (0, 0). \lim_{j\to+\infty} f(\overline{x} + h_j, \overline{y} + k_j) = f(\overline{x}, \overline{y})$. Quindi usando la nostra ipotesi possiamo ridurci a dimostrare che $\lim_{j\to+\infty}\langle \nabla f(\overline{x}, \overline{y}), (h_j, k_j)\rangle + o(|h_j, k_j|) = 0$. Dato che (h_j, k_j) tende a $(0, 0)$ per j che tende a $+\infty$, il prodotto scalare diventa nullo. Per quanto riguarda l'o-piccolo, possiamo usare la sua proprieta' per cui $\lim_{j\to+\infty} \frac{o(|h_j,k_j|)}{|h_j,k_j|} = 0$ riscrivendolo come $\frac{o(|h_j,k_j|)}{|h_j,k_j|}|h_j, k_j|$. Con $j \to +\infty$ sia la frazione che il fattore moltiplicativo tendono a 0. 😃

4.7 Condizioni sufficenti per la differenziabilita'

Theorem 4.7.1

Sia $f : \mathbb{R}^2 \to \mathbb{R}$. Assumo che esistano continue $\partial_x f(\overline{x}, \overline{y}), \partial_y f(\overline{x}, \overline{y})$ per ogni $(\overline{x}, \overline{y}) \in \mathbb{R}^2$, allora:

$$f \text{ e' differenziabile in ogni punto di } \mathbb{R}^2$$

Note:

Questo teorema vale anche in $\mathbb{R}^n$. Inoltre le funzioni elementari soddisfano sempre le ipotesi nel loro dominio, quindi sono differenziabili.

Per dimostrare questo teorema, ci serve prima una proposizione che equivale al teorema di Lagrange, pero' in piu' dimensioni. Usando le derivate parziali, ci riduciamo a due casi (uno dove ci si muove lungo x, e uno lungo y) dove ci riduciamo ad una funzione $\mathbb{R} \to \mathbb{R}$ che sappiamo gia' dimostrare.

Proposition 4.7.1 Lagrange per derivate parziali

Sia $f : \mathbb{R}^2 \to \mathbb{R}$ una funzione con derivate parziali $\partial_x f, \partial_y f$ continue, $\forall (\overline{x}, \overline{y}), (h, k) \in \mathbb{R}^2 . \exists \delta, \overline{\delta} \in]0, 1[$ tali

che:
$$\begin{cases} \frac{f(\overline{x}+h,\overline{y})-f(\overline{x},\overline{y})}{h} = \partial_x f(\overline{x}+\delta h,\overline{y}) \\ \frac{f(\overline{x},\overline{y}+k)-f(\overline{x},\overline{y})}{k} = \partial_y f(\overline{x},\overline{y}+\overline{\delta} k) \end{cases}$$

Ora che abbiamo Lagrange per le derivate parziali, possiamo dimostrare il teorema:

Dimostrazione teo. differenziabilita': Sia $(\overline{x},\overline{y}) \in \mathbb{R}^2$ e siano le derivate parziali di f continue, dobbiamo dimostrare che:
$$f(\overline{x}+h,\overline{y}+k) - f(\overline{x},\overline{y}) =< \nabla f(\overline{x},\overline{y}),(h,k) > +o(|(h,k)|)$$

Riscriviamo la parte sinistra:

$$f(\overline{x}+h,\overline{y}+k) - f(\overline{x},\overline{y}) = [f(\overline{x}+h,\overline{y}+k) - f(\overline{x}+h,\overline{y})]_1 + [f(\overline{x}+h,\overline{y}) - f(\overline{x},\overline{y})]_2$$

Ci riduciamo a dimostrare che:

$$[]_1 = \partial_y f(\overline{x},\overline{y})k + o(|(h,k)|)$$
$$[]_2 = \partial_x f(\overline{x},\overline{y})h + o(|(h,k)|)$$

Analizziamo il caso $[]_2$: usiamo Lagrange, che ci dice che $\exists\theta \in]1,0[.[]_2 = \partial_x f(\overline{x}+\theta h,\overline{y})h$, quindi dimostriamo che questo equivale a $\partial f(\overline{x},\overline{y})h + o(|(h,k)|)$:

$$[\partial f(\overline{x}+\theta h,\overline{y}) - \partial f(\overline{x},\overline{y})]h = o(|(h,k)|)$$
$$g(h,k) = o(|(h,k)|)$$

Sappiamo che $g(0,0) = 0$, quindi data una successione $(h_n,k_n) \xrightarrow{n\longrightarrow\infty} (0,0)$ con $(h_n,k_n) \neq \underline{0} \forall n \in \mathbb{N}$ dimostriamo che:
$$\lim_{n\to+\infty} \left| \frac{[\partial f(\overline{x}+\theta h_n,\overline{y}) - \partial f(\overline{x},\overline{y})]h_n}{|(h_n,k_n)|} \right| = 0$$

Grazie al valore assoluto sappiamo che la funzione sara' sempre $\geqslant 0$, e se la riscriviamo in questo modo:

$$|[\partial f(\overline{x}+\theta h_n,\overline{y}) - \partial f(\overline{x},\overline{y})]| \frac{|h_n|}{|(h_n,k_n)|}$$

possiamo dire che sara' sempre $\leqslant |\partial f(\overline{x}+\theta h_n,\overline{y}) - \partial f(\overline{x},\overline{y})|$, dato che $\frac{|h_n|}{|(h_n,k_n)|} \leqslant 1$, $\forall n \in \mathbb{N}, \theta \in]0,1[$. Usiamo quindi il teorema dei carabinieri e ci riduciamo a dimostrare che questo upper-bound tenda a 0 per n che tende a infinito. Grazie alla continuita' della derivata parziale, possiamo dire che:

$$\lim_{n\to+\infty} \partial f(\overline{x}+\theta h_n,\overline{y}) = \partial f(\overline{x},\overline{y})$$

Quindi la loro differenza e' 0. Si possono fare passaggi simili per dimostrare $[]_1$. ☺

4.8 Derivate direzionali

Fino ad ora abbiamo visto il valore di crescita della funzione solo lungo le assi principali x, y (le derivate parziali), vediamo ora cosa succede se ci allontaniamo da un punto cambiando entrambe le sue coordinate:

Definition 4.8.1: Derivata direzionale

Sia $f : \mathbb{R}^2 \to \mathbb{R}$ e $(\overline{x},\overline{y}) \in \mathbb{R}^2$. Dato un vettore unitario $v = (v_1,v_2)$ ($|v| = 1$), f si dice derivabile lungo la direzione v nel punto $(\overline{x},\overline{y})$ se esiste finito:

$$\frac{\partial f}{\partial v}(\overline{x},\overline{y}) = \lim_{h\to 0} \frac{f((\overline{x},\overline{y})+hv) - f(\overline{x},\overline{y})}{h}$$

Theorem 4.8.1 Formula del gradiente

Sia $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ differenziabile in $\overline{x}$, allora:

$$\forall v = (v_1, ..., v_n) \neq \underline{0}. |v| = 1 :$$

$$\frac{\partial f}{\partial v}(\overline{x}) =< \nabla f(\overline{x}), v >$$

: Dobbiamo dimostrare che $\lim_{t \to 0} \frac{f(\overline{x}+tv)}{t} =< \nabla f(\overline{x}), v >$. Essendo f differenziabile possiamo dire che $f(\overline{x}+tv) = f(\overline{x})+ < \nabla f(\overline{x}), tv > +o(|vt|)$. Dato che $|vt| = |v||t| = 1|t| = |t|$, possiamo riscrivere l'o-piccolo come $o(|t|)$. Sostituendo, rimane $\lim_{t \to 0} \frac{<\nabla f(\overline{x}), vt> + o(|t|)}{t}$. Possiamo dividere il limite, ottenendo $< \nabla f(\overline{x}), v >$ (abbiamo diviso per t) e $\lim_{t \to 0} \frac{o(|t|)}{t}$ che per definizione e' 0. L'uguaglianza quindi vale. ☺

Note:

Lineare in $v_1, ..., v_n$ con coefficenti le derivate parziali, quindi tutte le derivate direzionali $\partial_v f$ si possono scrivere conoscendo solo le derivate parziali.

4.8.1 Direzione di massima crescita

Problema: trovare la direzione di massima crescita di una funzione f in un assegnato punto.
Per rispondere a questa domanda, dobbiamo trovare il vettore unitario v la cui derivata direzionale e' massima, ovvero dobbiamo massimizzare la funzione $< \nabla f(\overline{x}, \overline{y}), (v_1, v_2) >= \partial_x f(\overline{x}, \overline{y})v_1 + \partial_y f(\overline{x}, \overline{y})v_2$. Riscrivendo il gradiente e v usando coordinate polari, otteniamo $< r(cos\theta, sin\theta), (cos\gamma, sin\gamma) >= r(cos\theta cos\gamma + sin\theta sin\gamma)$, che per la formula del coseno di una differenza diventa $rcos(\theta - \gamma)$, che raggiunge il suo valore massimo r quando $\theta = \gamma$. Questo vuol dire che la direzione di massima crescita in un punto e' quella dove giace il vettore gradiente, e il valore di questa crescita e' dato dal modulo dello stesso vettore. Possiamo generalizzare questa scoperta in n dimensioni (usando la disuguaglianza di Cauchy-Schwarz 3.5.1):

Proposition 4.8.1

Sia $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ differenziabile in $x \in \mathbb{R}^n$ e sia $\nabla f(x) \neq \underline{0}$, allora:

- $v_{\max} = \frac{\nabla f(x)}{|\nabla f(x)|}$
- $\frac{\partial f}{\partial v}(x) = |\nabla f(x)|$

4.9 Curve in $\mathbb{R}^n$

Una curva in $\mathbb{R}^n$ e' una "funzione" (non sempre) r (oppure s, γ):

$$r : (a, b) \longrightarrow \mathbb{R}^n$$

Che prende quindi un numero reale e ci da un punto in $\mathbb{R}^n$, ed e' definita da un parametro $t \in (a, b)$, quindi si puo' scrivere come $r(t) = (r_1(t), ..., r_n(t))$. Esempio:

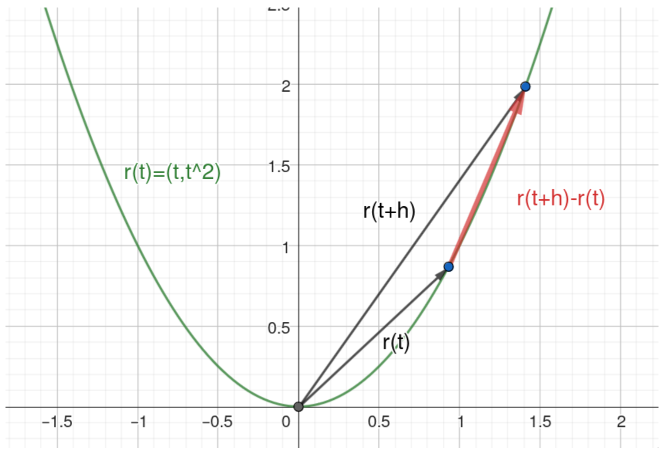
$$r : \begin{cases} x = t \\ y = 2 - t \\ z = 2t \end{cases} \quad \text{equazione parametrica di retta in } \mathbb{R}^3$$

Oppure come intersezione di due piani:

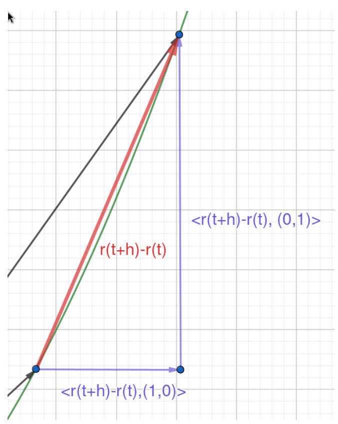
$$r = \pi + \sigma \implies \begin{cases} y = 2 - x \\ z = 2x \end{cases} \implies \begin{cases} x + y - 2 = 0 \\ 2x - z = 0 \end{cases}$$

4.9.1 Derivata di una curva

Supponiamo di avere una curva di classe $C^1((a,b))$, quindi esistono tutte le derivate prime e sono continue (in (a,b)). Il significato di derivata e' sempre lo stesso: e' il cambiamento del valore della funzione quando si incrementa l'input di un valore infinitesimo. Vediamo graficamente cosa succede (con $r : \mathbb{R} \to \mathbb{R}^2$ per semplicita'):



Vediamo che la differenza fra i valori di r e' rappresentata da un vettore $r(t+h)-r(t)$, che puo' essere scomposto come la somma del cambiamento rispetto alle direzioni della base ortonormale di riferimento, in questo caso le direzioni delle assi:



Il valore del cambiamento lungo l'asse x e' dato dalla derivata di $r_1(t)$, mentre il valore lungo l'asse y e' dato dalla derivata di $r_2(t)$. Quindi possiamo applicare la definizione di limite e calcolare il rapporto incrementale in questo modo:

$$\lim_{h\to 0} \frac{r(t+h)-r(t)}{h} = \left(\lim_{h\to 0} \frac{r_1(t+h)-r_1(t)}{h}, \lim_{h\to 0} \frac{r_2(t+h)-r_2(t)}{h} \right) = (r_1'(t), r_2'(t))$$

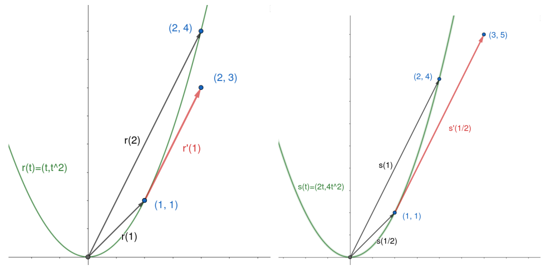
Ovvero, la derivata un punto t di una curva $r : \mathbb{R} \to \mathbb{R}^n$ e' data dalla somma delle derivate delle n funzioni parametriche di r moltiplicate per il corrispondente vettore base e_k. Quindi diamo la definizione:

Definition 4.9.1: Velocita di una curva

La velocita' (o derivata) di una curva e' il vettore tangente alla curva data:

$$r'(t) = (r_1'(t), ..., r_n'(t))$$

Quindi la direzione del vettore velocita' e' tangente alla curva nel punto dato, ma la sua lunghezza cosa indica? Prendiamo due curve $r(t) = (t, t^2), r' = (2t, 4t^2)$ e consideriamo il parametro t come il tempo trascorso. Notiamo che le due curve tracciate sono identiche, ma che una ci mette la meta' del tempo per percorrerla rispetto all'altra:



Inoltre, la lunghezza del vettore velocita' nello stesso punto sulla curva e' doppia nella funzione piu' "veloce". Quindi possiamo intuire che la norma della derivata ci indica quanto "velocemente" ci stiamo muovendo nel punto dato della curva, chiamata anche **velocita' instantanea**:

Definition 4.9.2: Velocita' scalare (instantanea)

Non e' altro che la norma del vettore tangente, cioe':

$$||v(t)|| = ||r'(t)|| = \sqrt{(r_1'(t))^2 + ... + (r_n'(t))^2}$$

4.9.2 Altre definizioni

Definition 4.9.3: Curva regolare

Una curva e' detta **regolare** in $t_0 \in (a, b)$ se:

$$r'(t_0) \neq \underline{0}$$

Note:
Una curva regolare $r : (a, b) \longrightarrow \mathbb{R}^n$ e' anche detta essere una **immersione** di (a, b) in $\mathbb{R}^n$.

Definition 4.9.4: Punto singolare

Se $t_0 \in (a, b)$ e' tale per cui

$$r'(t_0) = \underline{0}$$

Allora t_0 e' chiamato **punto singolare**

Definition 4.9.5: Curve semplici

Una curva e' detta **semplice** se e' iniettiva. (Cioe' la curva non si interseca)

4.9.3 Curvatura

Nelle curve almeno C^2, si definiscono:

Definition 4.9.6: Curvatura

Si definisce **curvatura** di $r : (a, b) \longrightarrow \mathbb{R}^n$ il modulo della accelerazione

$$k = ||a(t)|| = ||v'(t)|| = |||r''(t)|$$

Definition 4.9.7: Raggio di curvatura

Si chiama **raggio di curvatura**

$$R = \frac{1}{k}$$

Bisogna considerarlo nella retta reale estesa $\overline{\mathbb{R}} = [-\infty, +\infty]$, dato che le rette hanno il raggio di curvatura e ha senso che questo sia ∞

4.9.4 Derivata di una funzione scalare lungo una curva

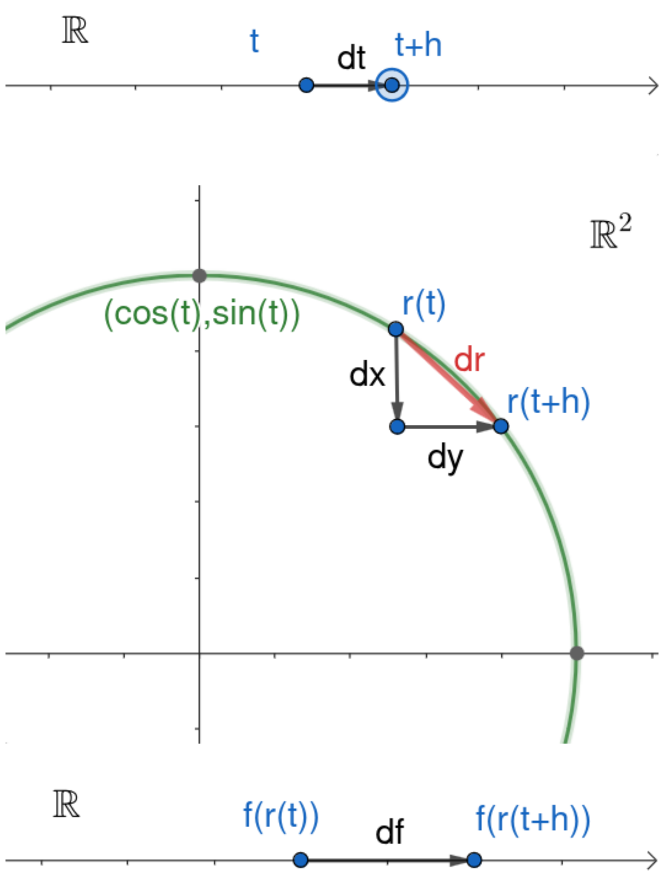
Prendiamo una funzione scalare $f : \mathbb{R}^2 \to \mathbb{R}$, definita come $f(x, y) = x^2y$. Ora al posto delle due variabili di input x e y mettiamo due funzioni $x(t) = cos(t)$ e $y(t) = sin(t)$. Abbiamo ottenuto ora una funzione che dipende da un solo parametro (t), quindi e' possibile calcolare la sua derivata rispetto a questo parametro. Nel nostro esempio, possiamo espandere la nostra funzione in questo modo:

$$f(cos(t), sin(t)) = cos^2(t)sin(t)$$

Quindi per calcolare la sua derivata basta applicare le classiche regole di derivazione:

$$\frac{d}{dt}f(cos(t), sin(t)) = -2cos(t)sin^2(t) + cos^3(t)$$

Ma esiste un modo piu' generale per descrivere questo procedimento? Possiamo iniziare considerando una curva $r : \mathbb{R} \to \mathbb{R}^2$, che per seguire l'esempio sopra puo' essere definito come $r(t) = (cos(t), sin(t))$. Quindi ora possiamo usare il vettore dato dalla curva r come input alla nostra funzione scalare f, creando quindi una composizione $f \circ r : \mathbb{R} \to \mathbb{R}$ di cui vogliamo studiare la derivata. Analizziamo la situazione:



Le tre figure ci mostrano, dall'alto verso il basso, come si propaga un aumento (infinitesimo) del parametro iniziale dt, prima nella curva r, poi dal valore spostato di r in f. Alla fine otteniamo lo spostamento totale, chiamato df che ora vediamo come calcolare:

- Prima di tutto, dobbiamo calcolare il valore del vettore $r(t+h) - r(t)$, ovvero dobbiamo trovare lo "spostamento" causato da un aumento infinitesimo del parametro di input di r. Se conosciamo la derivata della curva in quel punto, basta moltiplicarla per dt e otteniamo il vettore dr:

$$dr = r'(t)dt = (\frac{dx}{dt}dt, \frac{dy}{dt}dt)$$

- Ora che abbiamo il cambiamento in r, dobbiamo vedere come calcolare il cambiamento che questo causa in f. Puo' essere utile considerare le due componenti di dr, in modo da avere il delta del primo e del secondo parametro di f separatamente. In modo analogo al primo passo moltiplichiamo le derivate parziali con gli spostamenti dei parametri corrispondenti, sommandoli per trovare df:

$$df = \frac{\partial f}{\partial x}\frac{dx}{dt}dt + \frac{\partial f}{\partial y}\frac{dy}{dt}dt$$

- Ci rimane soltanto dividere l'equazione per dt per trovare la derivata:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

  Che puo' essere riscritta come un prodotto scalare:

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

Theorem 4.9.1

Sia $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile, allora data una curva $r : (a,b) \to \mathbb{R}^n$ (derivabile) si ha che:

$$f \circ r : (a,b) \to \mathbb{R}$$

E' derivabile, e la sua derivata e' la **derivata lungo la curva** r, e si ha:

$$(f \circ r)'(t) =< \nabla f(r(t)), r'(t) >$$

Prima di dare una dimostrazione formale di questo teorema, dobbiamo definire lo sviluppo di Taylor al primo ordine di una curva:

Definition 4.9.8

Data $r :]a,b[\to \mathbb{R}^n$ derivabile, sia $k \in \{1, ..., n\}$. Lo sviluppo di Taylor per $r_k :]a,b[\to \mathbb{R}$ in $t \in]a,b[$ e':

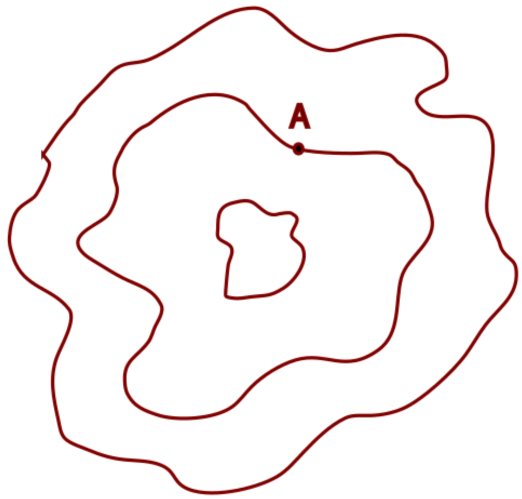
$$r_k(t+s) = r_k(t) + r_k'(t)s + o_k(s)$$

Quindi $r(t+s) = r(t) + r'(t)s + o(s)$ (dove $o(s) = (o_1(s), ..., o_n(s))$).

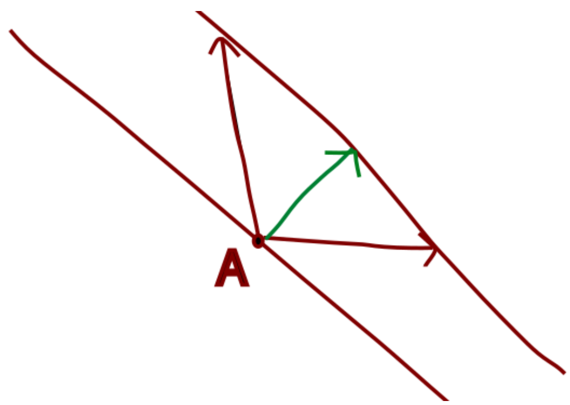
Dimostrazione formale: Usando la definizione di derivata, $(f \circ r)'(t) = \lim_{s \to 0} \frac{f(r(t+s)) - f(r(t))}{s}$. Dato che f e' differenziabile, possiamo sostituire il numeratore con $\langle f'(r(t)), r(t+s) - r(t) \rangle + o(|r(t+s) - r(t)|)$. Essendo derivabile, possiamo usare lo sviluppo di Taylor al primo ordine della curva r per sostituire $r(t+s) - r(t)$ con $r'(t)s + o(s)$. Usando la proprieta' distributiva e raccogliendo il fattore s, il limite diventa $\langle f(r(t)), r'(t) \rangle + \lim_{s \to 0} \frac{\langle f(r(t)), o(s) \rangle + o(|r'(t)s + o(s)|)}{s}$. Dividendo la frazione, nel primo addendo possiamo spostare il divisore all'interno del prodotto scalare che diventa $\langle f(r(t)), \frac{o(s)}{s} \rangle$, che per definizione di o-piccolo diventa nullo. Per il secondo addendo, con $s \to 0$ abbiamo che $o(|r'(t)s + o(s)|) = o(s)$, quindi sempre per definizione di o-piccolo anche il secondo addendo si annulla e il limite si azzera. 😊

4.9.5 Ortogonalita' gradiente-insiemi di livello

Possiamo utilizzare questo nuovo teorema per dimostrare la relazione di ortogonalita' fra il gradiente di una funzione calcolata in un punto e la curva di livello che passa per quel punto.
Come esempio, prendiamo una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ e tracciamone delle curve di livello:



Pensiamo ora di "zoomare" sul punto A e consideriamo la prossima curva di livello. Se ci avviciniamo abbastanza, possiamo considerare le due sezioni di curva come due linee parallele, quindi il segmento piu' corto che parte dal punto A e arriva alla prossima curva di livello e' quello perpendicolare alla curva di livello di A:



Dato che la direzione perpendicolare alla curva di livello in A coincide con la direzione che ci porta piu' velocemente alla prossima curva di livello (direzione di massima crescita), possiamo intuire che il vettore gradiente in A e' perpendicolare alla curva che passa per A:

Proposition 4.9.1

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile, si ha che:

$$\forall \underline{x} \in \mathbb{R}^n . \nabla f(\underline{x}) \perp L_{f(\underline{x})}$$

Dove $L_{f(\underline{x})}$ e' la curva di livello $\{x \in \mathbb{R}^n | f(x) = f(\underline{x})\}$

Dimostrazione formale: Assumiamo di avere una funzione $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile e un punto $\overline{x} \in \mathbb{R}^n$. Supponiamo di poter costruire una curva $r :]-1, 1[\to \mathbb{R}^n$ tale che: $\forall t \in]-1, 1[. r(t) \in L_b$, $r(0) = \overline{x}$ e $r'(0) \in \mathbb{R}^n \neq \underline{0}$ (direzione tangente a $L_{f(\overline{x})}$ nel punto $\overline{x}$). Dato che tutti i punti della curva r fanno parte dello stesso insieme di livello, $\forall t \in]-1, 1[. (f \circ r)(t) = f(\overline{x})$. Essendo una funzione costante, sul suo dominio la derivata e' nulla, ovvero $\forall t \in]-1, 1[. (f \circ r)'(t) = 0$, che puo' essere scritta come $\langle \nabla f(r(t)), r'(t) \rangle = 0$ (usando il teorema di derivata lungo una curva). In altre parole, sappiamo che il prodotto scalare fra il gradiente di f in un punto e la derivata della curva di livello nello stesso punto e' sempre nullo. Quindi il gradiente in $\overline{x}$ e' ortogonale alla tangente della curva di livello in $\overline{x}$. 😄

Chapter 5

Punti di massimo e minimo in piu' dimensioni

La definizione di punti di massimo/minimo in piu' dimensioni e' molto simile a quella in $\mathbb{R}$:

Definition 5.0.1: Maxima e minima

Data $f : A \to \mathbb{R}$, con $A \subseteq \mathbb{R}$, $\underline{x} \in A$ si dice punto di massimo/minimo se:

$$\exists \delta > 0. \forall x \in D(\underline{x}, \delta) \cap A. f(x) \geqslant (\leqslant) f(\underline{x})$$

Come facciamo a trovare questi punti? Guardiamo i metodi usati in una dimensione per vedere se sono generalizzabili:

- Studio del segno della derivata prima: se $f(\underline{x}) = 0$ e in un intorno di $\underline{x}$ si ha che $f(x) > (<)0$, allora $\underline{x}$ e' un punto di massimo (minimo)
- Utilizziamo il segno della derivata seconda per trovare la concavita' del grafico nei punti di stazionamento:

$$\begin{cases} f'(\underline{x}) = 0 \\ f''(\underline{x}) > (<)0 \end{cases} \iff \underline{x} \text{ e' punto di massimo (minimo)}$$

Quando $n \geqslant 2$, $f'(x)$ viene sostituito dal gradiente $\nabla f(x)$, che e' un vettore e non uno scalare. Dato che non ha senso chiedersi se un vettore e' $\geqslant 0$, il concetto di "crescente" e "decrescente" non puo' essere applicato in piu' dimensioni (perche' puo' essere diversa in direzioni diverse). Quindi dobbiamo utilizzare il secondo metodo, che ci da' due condizioni da rispettare (una sulla derivata di primo grado, una sulla seconda). Vediamo come generalizzare queste condizioni in piu' dimensioni:

5.1 Condizioni del primo ordine

Guardiamo la prima condizione relativa alla derivata di primo grado. Questa ci dice che e' condizione necessaria (ma non sufficente) per un punto di massimo/minimo avere la derivata prima nulla, vediamo se cio' vale anche in piu' dimensioni:

Theorem 5.1.1 Fermat

Data $f : A \to \mathbb{R}$ differenziabile, con $A \subseteq \mathbb{R}^n$ aperto, si ha che se un punto $\underline{x} \in A$ e' massimo o minimo, allora:

$$\nabla f(\underline{x}) = 0$$

Dimostrazione: Presa una funzione $f : A \to \mathbb{R}$ ($A \subseteq \mathbb{R}^n$ aperto) con punto minimo in $\overline{x}$, devo dimostrare che le derivate parziali in quel punto siano nulle. Semplicemente basta guardare la funzione sui piani paralleli alle direzioni canoniche passanti per $\overline{x}$ e usare Fermat per le funzioni semplici. Per fare cio' creiamo una funzione

$h_k :]-\delta, \delta[\to \mathbb{R}$, definita $\forall k \in \{1, \dots, n\}$ come $h_k(t) = f(\overline{x} + te_k)$. Il δ scelto e' quello garantito dalla definizione di punto minimo, quindi si ha che $\forall x$ che dista meno di δ da $\overline{x}$ si ha che $f(x) \leqslant (\geqslant) f(\overline{x})$. Quindi h ha max/min con $t = 0$, ed essendo una funzione semplice abbiamo che $h'_k(0) = 0$. Ora dimostriamo che cio' implica che tutte le derivate parziali sono nulle: $h'_k(0) = \lim_{t \to 0} \frac{h_k(t)+h_k(0)}{t} = \lim_{t \to 0} \frac{f(\overline{x}+te_k)+f(\overline{x})}{t} = \frac{\partial f}{\partial x_k}(\overline{x}) = 0$. 😊

Pero' questa e' solo una condizione necessaria, non sufficente per essere un punto di massimo/minimo. Vedremo infatti che esistono dei punti "di sella" che sono stazionari ma che non sono ne massimi ne minimi (sono una generalizzazione dei punti di flesso a tg orizzontale). Per distinguere i tipi di punti stazionari, dobbiamo guardare la derivata seconda:

5.2 Condizioni di secondo ordine

5.2.1 In una variabile

Diamo un attimo una intuizione al perche' la condizione di secondo ordine funziona in una variabile. Per fare cio', prendiamo una funzione f e scriviamone il polinomio di Taylor nel punto stazionario x_0 fino al secondo grado (dato che stiamo facendo solo un ragionamento informale, l'errore possiamo anche tralasciarlo):

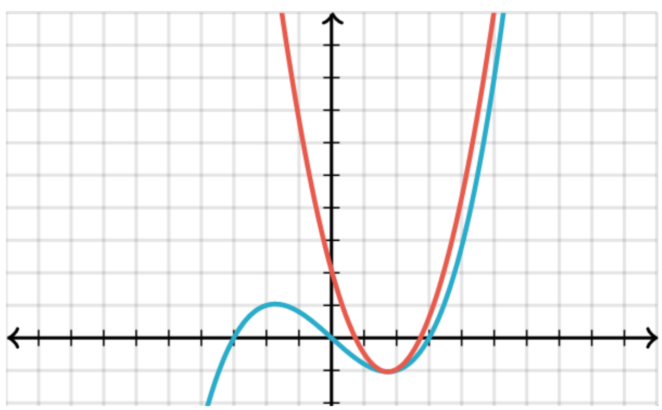
$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

Dato che siamo interessati ai punti stazionari, $f'(x_0) = 0$, quindi possiamo semplificare l'approssimazione cosi':

$$f(x) \approx f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

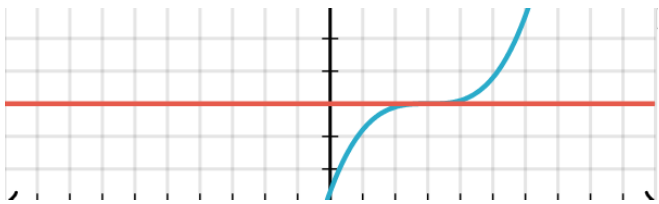
Dato che questa e' un'approssimazione di un piccolo intorno centrato nel punto x_0, se vediamo che x_0 e' un max/min nell'approssimazione, allora significa che nella funzione effettiva f esiste un intervallo contenente x_0 dove questo e' maggiore/minore di tutti i punti dell'intervallo. Guardiamo i vari casi:

- Se $f''(x_0) > 0$, allora x_0 e' un punto di minimo, dato che $(x - x_0)^2$ e' sempre positivo tranne quando $x = x_0$, che ci darebbe il valore del minimo $f(x_0) \approx f(x_0) + 0$.



- Se $f''(x_0) < 0$, allora x_0 e' un punto di massimo, dato che siamo nella situazione precedente ma $(x - x_0)^2$ e' moltiplicato per una costante negativa. Quindi il polinomio descrive una parabola "all'ingiu'", con vertice in x_0.

- Se $f''(x_0) = 0$, allora siamo in un caso particolare dove la derivata seconda non riesce a darci informazioni aggiuntive rispetto alla natura del punto stazionario.



Quindi, per riassumere: quando $f'(x_0) = 0$, sappiamo che se il termine quadratico dell'approssimazione di Taylor $\frac{1}{2}f''(x_0)(x - x_0)^2$ e' sempre positivo/negativo (includendo lo 0), allora x_0 e' un punto di massimo/minimo.

5.3 Approssimazione quadratica in piu' dimensioni

Proviamo a seguire lo stesso procedimento per funzioni in una sola variabile. La prima cosa da fare e' scrivere un'approssimazione quadratica (Taylor al secondo ordine) di una funzione in piu' dimensioni. Partiamo prima con la definizione delle derivate seconde in n dimensioni:

5.3.1 Derivate seconde e hessiana

Introduciamo le derivate seconde in n dimensioni, ovvero il calcolo due volte di fila delle derivate (parziali). La differenza questa volta e' che ci sono diverse derivate parziali, che quindi possiamo applicare in sequenza in n^2 modi diversi. Introduciamo la notazione:

Definition 5.3.1: Derivate seconde in n dimensioni

Sia $f : \mathbb{R}^n \to \mathbb{R}$, assumiamo che $\forall k \in \{1, ..., n\}.\exists \frac{\partial f}{\partial x_k}$ in ogni punto di $\mathbb{R}^n$. Consideriamo una nuova variabile x_j, con $j \in \{1, ..., n\}$, si ha che:

- Se $j \neq k$ (ovvero se stiamo applicando due derivate parziali diverse) scriviamo:

$$\frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_k}) = \frac{\partial^2 f}{\partial x_j \partial x_k}$$

- Nel caso $j = k$ (applichiamo due volte la stessa derivata parziale) scriviamo:

$$\frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_k}) = \frac{\partial^2 f}{\partial x_k^2}$$

Seppure sembrano tante, in verita' si puo' dimostrare che l'ordine in cui applichiamo le derivate non cambia il valore finale:

Theorem 5.3.1 Schwarz

Sia $f : \mathbb{R}^n \to \mathbb{R}$ tale che per qualche (almeno uno) intorno centrato in $p \in \mathbb{R}^n$ la derivata seconda (non mista!) di f e' continua, allora:

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

Come abbiamo visto sopra, con n derivate parziali diverse esistono n^2 combinazioni di derivate seconde. Queste derivate seconde le possiamo mettere in una matrice (detta "hessiana") quadrata $Hf(\overline{x}) \in M_n(\mathbb{R})$ definita cosi':

$$Hf(\overline{x})_{jk} = \frac{\partial^2 f}{\partial x_j \partial x_k}$$

Example 5.3.1

Data una funzione in due dimensioni $f : \mathbb{R}^2 \to \mathbb{R}$, la sua matrice hessiana associata nel punto x_0 e':

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

Ho omesso l'esplicitazione del punto da derivare nella matrice perche' spesso si fa cosi'.

5.3.2 Taylor secondo Lagrange

Torniamo un attimo in una dimensione per dare una nuova formula di approssimazione, che come noteremo da' informazioni aggiuntive rispetto a Taylor con il resto di Peano:

Theorem 5.3.2 Taylor secondo Lagrange

(n=1) $f : \mathbb{R} \to \mathbb{R}$ assumo f, f', f'' continue. Dati $\overline{x}$ e $\overline{x} + h \in \mathbb{R}$, esiste $\theta \in (0, 1)$.

$$f(\overline{x} + h) = f(\overline{x}) + f'(\overline{x})h + \frac{1}{2} f''(\overline{x} + \theta h)h^2$$

Note:

$\overline{x} + \theta h$ appartiene all'intervallo di estremi $\overline{x}$ e $\overline{x} + h$ (aperto)

Note:

Ricordiamo la formula nota: $f(\overline{x} + h) = f(\overline{x}) = f'(\overline{x})h + \frac{1}{2} f''(\overline{x})h^2 + o(h^2)$. Per h che non sono vicini a 0, la formula col resto di Peano non ci da molte informazioni riguardo all'errore, mentre con il resto di Lagrange abbiamo sempre una informazione quantitativa per qualunque h scelto.

Note:

Osserviamo che dalla formula di Lagrange possiamo ricondurci a quella sopra, quindi possiamo dire che contiene piu' "informazioni". Infatti sottraendo la seconda dalla prima rimane $\frac{h^2}{2}(f''(\overline{x}) - f''(\overline{x} + \theta h)) = o(h^2)$. Per verificarlo dividiamo la parte sinistra per h^2 e facciamo il limite per $h \to 0$, e si nota che effettivamente il limite e' 0 (usando la continuita' di f'').

5.3.3 Forma quadratica

Per classificare i punti stazionari dobbiamo analizzare la parte quadratica dello sviluppo di Taylor, che come poi dimostreremo ha la seguente forma:

$$ax^2 + 2bxy + cy^2$$

Un'equazione di questo tipo e' chiamata una **forma quadratica**, che puo' essere rappresentata anche tramite il prodotto scalare fra una matrice e due vettori, che e' utile in quanto una sola scrittura vale per qualsiasi numero di variabili. Vediamo che in questo modo l'equazione diventa:

$$\left\langle \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle$$

Notiamo che la matrice associata alle forme quadratiche e' sempre una matrice simmetrica. Vediamo ora la definizione di forma quadratica usando questa nuova scrittura:

Definition 5.3.2: Forma quadratica

Sia $A \in M_n(\mathbb{R})$ simmetrica ($A = A^T$) e $h \in \mathbb{R}^n$ (vettore colonna $n \times 1$), definiamo la funzione $q_A : \mathbb{R}^n \to \mathbb{R}$:

$$q_A(h) = \langle Ah, h \rangle \in \mathbb{R}$$

Note:

$p(h) = 3h_1^2 + \frac{1}{3}h_1h_2 - 7h_2^2$, quindi $p = q_A$ con $A = \begin{pmatrix} 3 & 1/6 \\ 1/6 & -7 \end{pmatrix}$

Segno

Per generalizzare la condizione $f''(\overline{x}) > 0$ in $\mathbb{R}$ dobbiamo dare una nozione di segno di una forma quadratica:

Definition 5.3.3

Sia $A = A^T \in M_n(\mathbb{R})$. Allora si dice che:

- A e' definita positiva se $\langle Ah, h \rangle > 0 \ \forall h \neq 0$
- A e' definita negativa se $\langle Ah, h \rangle < 0 \ \forall h \neq 0$
- A e' indefinita se $\exists h, t \in \mathbb{R}^n . \langle Ah, h \rangle < 0 < \langle At, t \rangle$

Note:

Tutte le disuguaglianze sono strette

Per ora restringiamoci a solo due variabili. E' possibile sapere il segno di una forma quadratica conoscendo solo la sua matrice? Si, si puo' dimostrare trattando inizialmente y come una costante, chiamiamola y_0, e calcolando gli zeri:

$$y_0 \left(\frac{-b \pm \sqrt{b^2 - ac}}{a} \right)$$

Vediamo subito che se $y_0 = 0$, allora abbiamo entrambe gli zeri nel punto $(0, 0)$, mentre se $y_0 \neq 0$ il numero di zeri dipende dal delta (in realta' studiamo il meno-delta per poter usare il determinante dopo). La concavita' della parabola, invece, dipende come al solito dal segno di a:

- se $ac - b^2 < 0$, allora esistono due zeri distinti $\forall y_0 \neq 0$ e la parabola e' sia negativa che positiva, quindi in questo caso la forma quadratica e' indefinita.
- se $ac - b^2 > 0$, non esistono zeri reali e quindi la parabola e' tutta positiva se $a > 0$ o tutta negativa se $a < 0$. Quindi il segno della forma quadratica dipende dal segno di a.
- se $ac - b^2 = 0$, allora esistono punti $(x, y) \neq (0, 0)$ per cui il valore della forma quadratica e' 0 e il resto dell'immagine e' tutta positiva o tutta negativa. Per questo motivo non riusciamo a dire che sia sempre positiva o negativa, ne che sia indefinita.

Il valore $ac - b^2$ in realta' corrisponde al determinante della matrice associata alla forma quadratica, quindi possiamo riscrivere i casi sopra in modo piu' breve:

Proposition 5.3.1 Regola dei segni delle forme quadratiche

Sia $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Allora:

- $A > 0 \iff \begin{cases} a > 0 \\ detA > 0 \end{cases}$
- $A < 0 \iff \begin{cases} a < 0 \\ detA > 0 \end{cases}$
- A e' indefinita $\iff$ $detA < 0$ (punto di sella)

Forma quadratica Hessiana

Essendo una matrice simmetrica, anche la matrice hessiana, che abbiamo definito insieme alle derivate seconde in n dimensioni, ha una forma quadratica associata:

$$q(h) = \langle H_{f(\overline{x})} h, h \rangle$$

dove $H_{f(\overline{x})}$ e la matrice hessiana di ordine n e h e' un qualunque vettore di $\mathbb{R}^n$.

5.3.4 Taylor di secondo ordine in n dimensioni

Theorem 5.3.3 Formula di Taylor con resto di Lagrange di secondo ordine in n dimensioni

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ con tutte le derivate seconde continue, si ha che $\forall x_0, h \in \mathbb{R}^n$, $\exists \delta \in (0, 1)$ tale che:

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle H_{f(x_0 + \delta h)} h, h \rangle$$

Note:

Come avevamo visto in una dimensione, questa formula e' "globale", ovvero vale anche per h "lontani" da x_0.

: Usiamo una funzione $h : \mathbb{R} \to \mathbb{R}$ che ha derivata seconda continua (?). Per Taylor in una variabile sappiamo che: $\forall t_0, t_0 + t \in \mathbb{R}. \exists \theta \in]0, 1[$: $h(t_0 + t) = h(t_0) + h'(t_0)t + \frac{1}{2} h''(t_0 + \theta t)t^2$. Prendendo come f la funzione del teorema e dati $x_0, x_0 + h \in \mathbb{R}^n$, definisco la funzione $h(t) = f(x_0 + th)$ (stare attenti a non confondere h-funzione con h-variabile). Scriviamo Taylor di h ponendo $t_0 = 0$ e $t = 1$: $h(1) = h(0) + h'(0) + \frac{1}{2} h''(\theta)$. Usando la definizione di $h(t)$, otteniamo $h(1) = f(x_0 + h)$, $h(0) = f(x_0)$, $h'(0) = \langle \nabla f(x_0), h \rangle$ (derivata di f lungo la curva $v(t) = x_0 + th$) e $h''(\theta) = \langle H_{f(x_0 + \theta h) h}, h \rangle$. Sostituendo otteniamo la tesi del teorema. 😃

Corollary 5.3.1 Taylor di secondo ordine in n dimensioni con resto di Peano

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ con tutte le derivate seconde continue, si ha che $\forall x_0, h \in \mathbb{R}^n$ si ha che:

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

Dove $\forall \epsilon > 0. \exists \delta > 0. \forall 0 < \|h\| < \delta \implies \left| \frac{o(\|h\|^2)}{\|h\|^2} \right| < \epsilon$

: Scriviamo Taylor nelle due versioni:

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

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

Sostituiamo il membro a sinistra del secondo con il membro di destra del primo, e semplificando otteniamo:

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

Moltiplichiamo per 2 e scriviamo le forme quadratiche per esteso raccogliendo i termini con le stesse derivate seconde. Otteniamo una somma di tre forme di questo tipo (dove $j, k \in \{1, ..., n\}$):

$$h_j h_k(\partial_{jk}f(x_0 + \theta h) - \partial_{jk}f(x_0))$$

Dobbiamo mostrare che questa somma sia un o-piccolo di $|h|^2$, quindi basta controllare che $\forall j, k \in \{, ..., n\}$:

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

Per analisi asintotica, $\frac{h_j h_k}{|h|^2}$ tende a un numero reale, mentre il fattore a destra tende a 0. 😊

5.4 Classificazione di punti stazionari

Diamo prima di tutto una definizione formale di punto di sella:

Definition 5.4.1: Punto di sella

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ con derivate seconde continue, il punto critico $\overline{x} \in \mathbb{R}^n$ $(\nabla f(\overline{x}) = \underline{0})$ si dice di **sella** se:

$$\forall\delta > 0.\exists\overline{x}^-, \overline{x}^+.f(\overline{x}^-) < f(\overline{x}) < f(\overline{x}^+)$$

Ci serve anche una proprieta' delle forme quadratiche:

Proposition 5.4.1

Sia $A = A^T$ definita positiva di ordine n, allora $\exists\lambda \in \mathbb{R} > 0$ tale che $\forall h \in \mathbb{R}^n$:

$$\langle Ah, h\rangle \geqslant \lambda\|h\|$$

Possiamo anche scrivere $\langle A\frac{h}{\|h\|}, \frac{h}{\|h\|}\rangle \geqslant \lambda$.

Dimostrazione in $\mathbb{R}^2$: Prendiamo una matrice simmetrica di ordine 2 A definita positiva, con $a, b, c \in \mathbb{R}$. Dobbiamo dimostrare che esiste un $\lambda \in \mathbb{R} > 0$ tale che per ogni vettore unitario $v = (cos\theta, sin\theta)$ si ha che $\langle Av, v\rangle \geqslant \lambda$. Chiamiamo $g : [0, 2\pi] \to \mathbb{R}$ la forma quadratica ottenuta. Essendo g continua su un intervallo chiuso possiamo applicare Weierstrass e dire che esiste un minimo assoluto, ovvero $\exists\overline{\theta} \in [0, 2\pi].\forall\theta \in [0, 2\pi] : g(\theta) \geqslant g(\overline{\theta})$. Inoltre, dato che g e' definita positiva, $\forall\theta \in [0, 2\pi].g(\theta) > 0$, quindi $g(\overline{\theta})$ e' strettamente positiva ed equivale al valore λ cercato. 😊

Note:
λ puo' essere scelto come l'autovalore minore di A.

Ora siamo in grado di dimostrare come il segno della matrice hessiana puo' classificare un punto critico:

Theorem 5.4.1 Classificazione dei punti critici

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ con derivate seconde continue, sia $\overline{x} \in \mathbb{R}^n$ un punto critico, allora:

- Se $H_{f(\overline{x})} > 0$, allora $\overline{x}$ e' un punto di minimo
- Se $H_{f(\overline{x})} < 0$, allora $\overline{x}$ e' un punto di massimo
- Se $H_{f(\overline{x})}$ e' indefinita, allora $\overline{x}$ e' un punto di sella

:

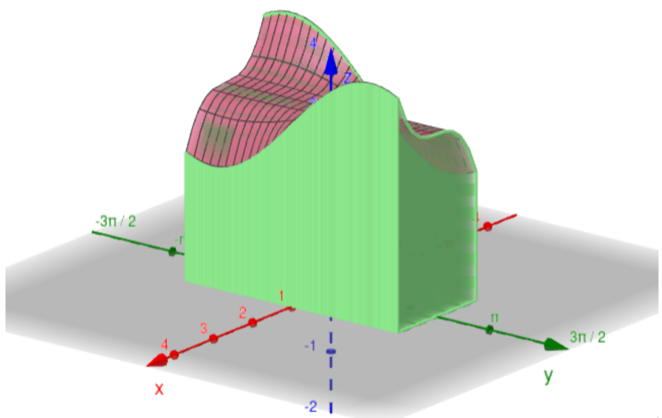
- 1) Per ipotesi sappiamo che $\nabla f(\overline{x}) = 0$ e che $H_{f(\overline{x})} > 0$. Dobbiamo dimostrare che $\exists \delta \in \mathbb{R} > 0. \forall h \in \mathbb{R}^n. \|h\| < \delta : f(\overline{x}+h) - f(\overline{x}) \geqslant 0$ (definizione di minimo). Dato che f ha derivate seconde continue per ipotesi, possiamo scrivere il suo sviluppo di Taylor di secondo ordine: $f(\overline{x}+h) - f(\overline{x}) = \frac{1}{2}\langle H_{f(\overline{x})}h, h\rangle + o(\|h\|^2)$ (dato che il gradiente e' nullo). Per 5.4 sappiamo che $\langle H_{f(\overline{x})}h, h\rangle \geqslant \lambda\|h\|$, quindi possiamo continuare l'equazione sopra aggiungendo $\geqslant \frac{1}{2}\lambda\|h\|^2 + o(\|h\|^2) = \|h\|^2\left(\lambda + \frac{o(\|h\|^2)}{\|h\|^2}\right)$. Usiamo la definizione di o-piccolo, scegliendo come ϵ il valore $\frac{\lambda}{2}$ e attribuiamo al δ che stiamo cercando il valore positivo per cui $\forall h \in \mathbb{R}^n. \|h\| < \delta$: $\left|\frac{o(\|h^2\|)}{\|h\|^2}\right| < \frac{\lambda}{2}$. Controlliamo che $\forall h. \|h\| < \delta$ valga la proprieta' di minimo: $\geqslant \frac{\|h\|^2}{2}\left(\lambda - \frac{\lambda}{2}\right) \geqslant 0$.
- 2) (Uguale a quello sopra)
- 3)

😃

Chapter 6

Integrali su piu' variabili

Vediamo ora come generalizzare il concetto di integrale definito in due dimensioni. Considero $f : A \longrightarrow \mathbb{R}$, $A \subseteq \mathbb{R}^2$, $\forall (x, y) \in A . f(x, y) \geqslant 0$, il nostro scopo e' quello di trovare il volume del sottografico rispetto all'area A, ovvero dell'insieme di punti $\{(x, y, z) \in \mathbb{R}^3 | (x, y) \in A, 0 \leqslant z \leqslant f(x, y)\}$.



Si dimostra che, se A e' **opportuno** ed f e' continua, il colume del sottografico e' dato da:

$$\int_A f(x, y) dx dy = \text{ integrale doppio}$$

Se mettiamo $f = 1$ costante, si trova $\int_A 1 dx dy$ che ha lo stesso valore dell' area di A (come $\int_a^b 1 dx = b - a =$ lunghezza di $[a, b]$).

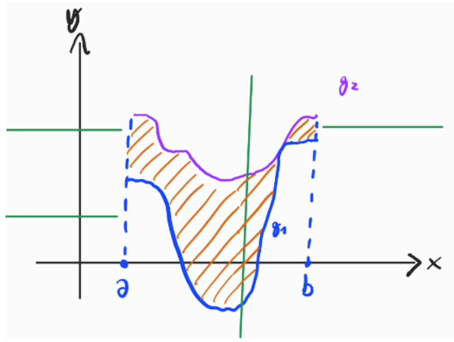
6.1 Insiemi semplici

Definition 6.1.1: Insieme y-semplice nel piano xy

Sia $[a, b] \subseteq \mathbb{R}$, siano $g_1, g_2 : [a, b] \longrightarrow \mathbb{R}$, tali che $g_1(x) \leqslant g_2(x) \forall x \in [a, b]$ (con g_1, g_2 continue).

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

Si chiama y-semplice perche' intersecandolo con una retta verticale otteniamo sempre un segmento continuo:

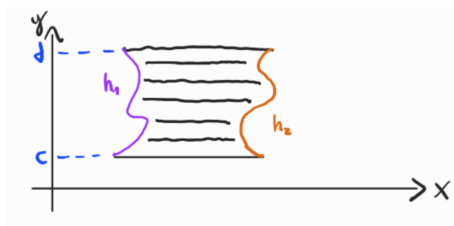


Se abbiamo la stessa proprieta' ma per le rette orizzontali, allora sono insiemi x-semplici:

Definition 6.1.2: Insieme x-semplice nel piano xy

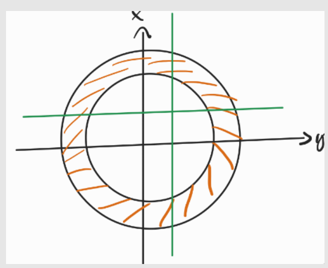
Sia $[c,d] \subseteq \mathbb{R}_y$, siano $h_1, h_2 : [c,d] \to \mathbb{R}$, tali che $h_1(y) \leqslant h_2(y) \forall y \in [c,d]$ (con h_1, h_2 continue).

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



Note:

Esistono insiemi che non sono ne' x, ne' y -semplici (corona circolare):



6.2 Formule di riduzione

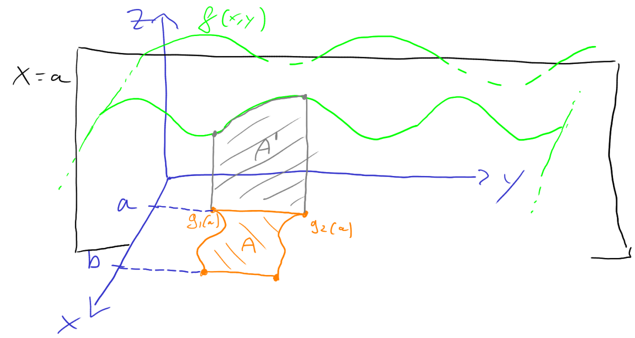
Vediamo ora come le proprieta' degli insiemi semplici possano aiutarci a calcolare il volume di un sottografo in tre dimensioni. L'obbiettivo principale e' quello di calcolare l'area sotto il grafico fissando la x o la y usando un integrale in una variabile, e poi di sommare tutte queste aree lungo la direzione opposta utilizzando ancora un integrale semplice. Vediamo come applicare questa intuizione formalmente:

Proposition 6.2.1 Riduzione y-semplice

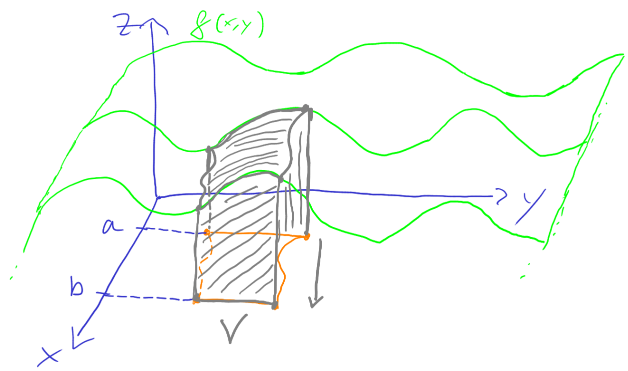
Siano $g_1, g_2 : [a,b] \to \mathbb{R}$ due funzioni continue t.c. $\forall x \in [a,b]. g_1(x) \leqslant g_2(x)$. Sia $A \in \mathbb{R}^2$ un insieme y-semplice e sia $f : A \to \mathbb{R}$ continua, allora vale:

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

Quindi, dato un insieme A y-semplice e una funzione $f : A \to \mathbb{R}$ continua, per trovare il volume del sottografo relativo ad A bisogna prima di tutto trovare la formula che ci da' l'area sotto il grafico dato un valore di x. Quindi consideriamo la x come costante e applichiamo l'integrale in dy:



Da notare come questi integrali sono calcolati sull'intervallo $[g_1(x), g_2(x)]$, che sono sempre gli estremi dell'area dato che A e' y-semplice. Ora ci basta sommare tutte le aree trovate per tutti gli x nell'intervallo $[a, b]$, che possiamo fare con un secondo integrale in dx questa volta:



Se dobbiamo calcolare il volume sopra un'area x-semplice, basta ripetere gli stessi passi trovando prima le aree fissando y e poi sommarle tutte lungo l'asse y:

Proposition 6.2.2 Riduzione x-semplice

Siano $h_1, h_2 : [c, d] \to \mathbb{R}$ due funzioni continue t.c. $\forall y \in [c, d]. h_1(y) \leqslant h_2(y)$. Sia $A \in \mathbb{R}^2$ un insieme x-semplice e sia $f : A \to \mathbb{R}$ continua, allora vale:

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