La convessità

Dati $x, y \in \mathbb{R}^n$, si definisce "combinazione convessa" di x, y l'insieme

$$\{ x(\lambda) \mid x(\lambda) = \lambda x + (1-\lambda) y, \ 0 \le \lambda \le 1 \}$$

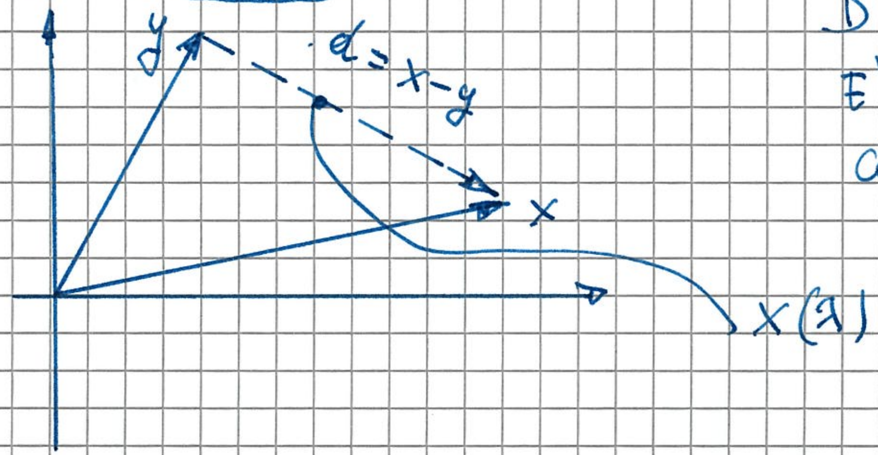
l'insieme dei punti ottenibili come combinazione lineare a coefficienti non negativi e a somma 1 $\left(\lambda + (1-\lambda) = 1 \right)$

Questo insieme è anche detto "segmento di estremi x e y"

$(x(0) = y \ ; \ x(1) = x)$.

$x(\lambda)$ può essere anche scritto $x(\lambda) = y + \lambda \underbrace{(x-y)}_{d}, \ 0 \le \lambda \le 1$

$x(\lambda) = y + \lambda d, \ 0 \le \lambda \le 1$



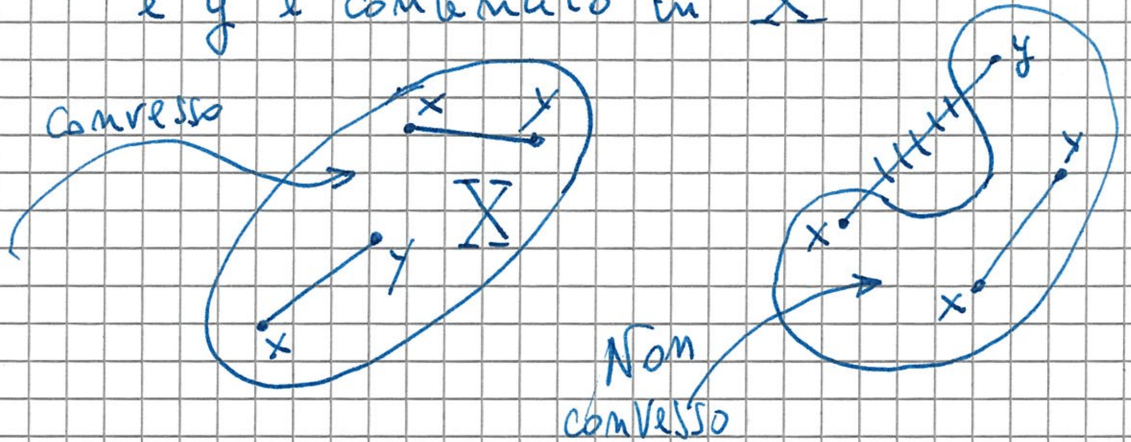
Dati m vettori $x^{(i)} \in \mathbb{R}^n$, $1 \le i \le m$

E' definita la loro "combinazione convessa":

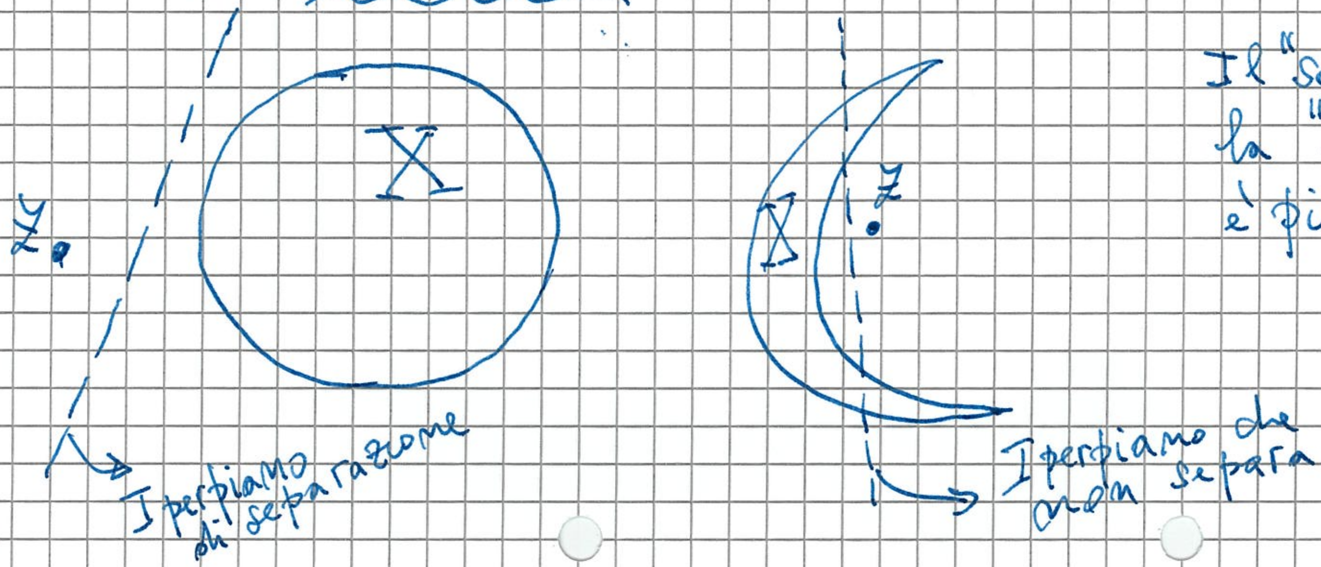
$$\left\{ x \in \mathbb{R}^n \mid x = \sum_{i=1}^{m} \lambda_i x^{(i)}, \ \sum_{i=1}^{m} \lambda_i = 1, \ \lambda_i \ge 0 \right\}$$

"Partendo da y mi sposto lungo la semiretta $d = x - y$ fino a raggiungere il punto x"

Definizione. Un insieme $X \subseteq \mathbb{R}^n$ è convesso se comunque si prendano due punti $x, y \in X$, l'intero segmento di estremi x e y è contenuto in X.



Un'importante proprietà degli insiemi convessi. Dato un insieme convesso chiuso X ed un punto $z \notin X$, esiste un iperpiano che separa strettamente z da X.



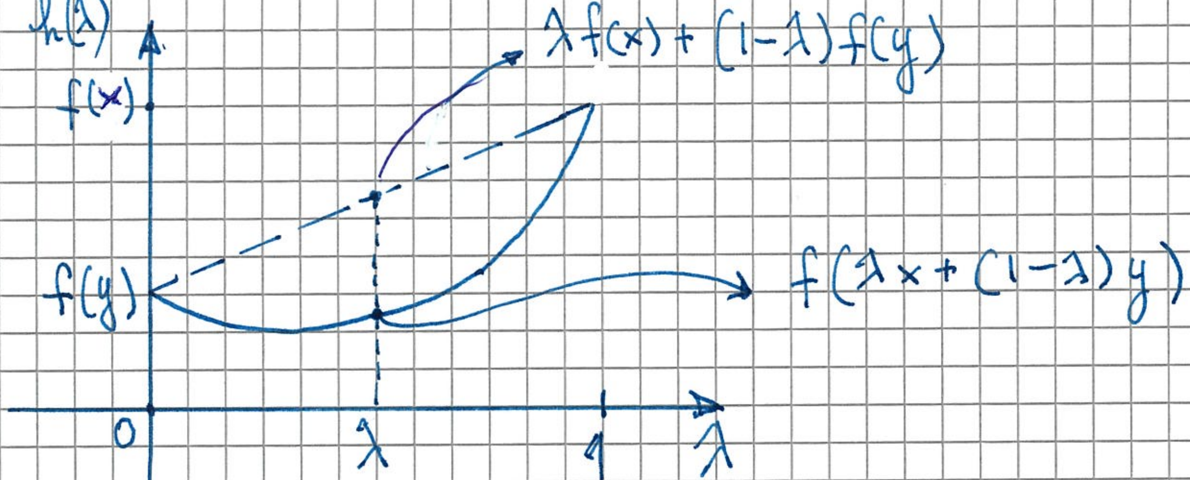
Il "sole" è convesso ma la "luna", quando non è piena, non lo è!

Cosa accade se X è convesso ma non chiuso?

Definizione. Una funzione $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$, con X convesso è detta convessa se $\forall\, x, y \in X$ vale

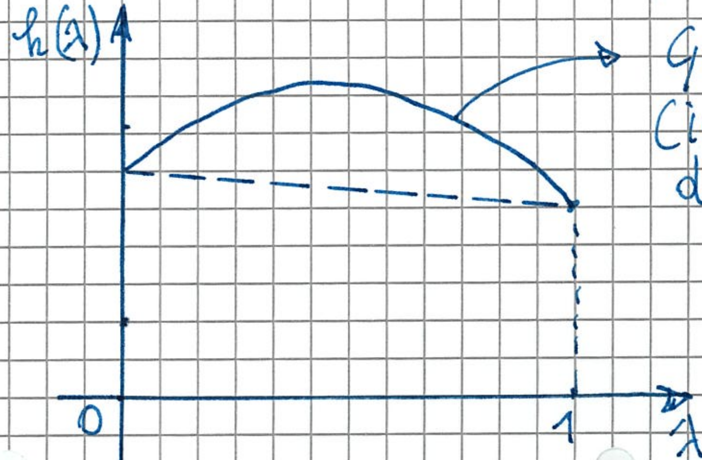
$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y), \quad \forall \lambda \in [0,1]$$

Fissati x, y $\quad f(\lambda x + (1-\lambda) y) = h(\lambda)$ con $h(0) = f(y)$ e $h(1) = f(x)$



"Sul segmento di estremi x e y il grafo della restrizione di f è tutto al di sotto della corda"

Def. f è concava se $-f$ è convessa



Grafo sopra la corda (inversione del $\leq$ nelle definizione)

Nota La definizione di convessità (e concavità) prescinde dalle proprietà di differenziabilità di f.

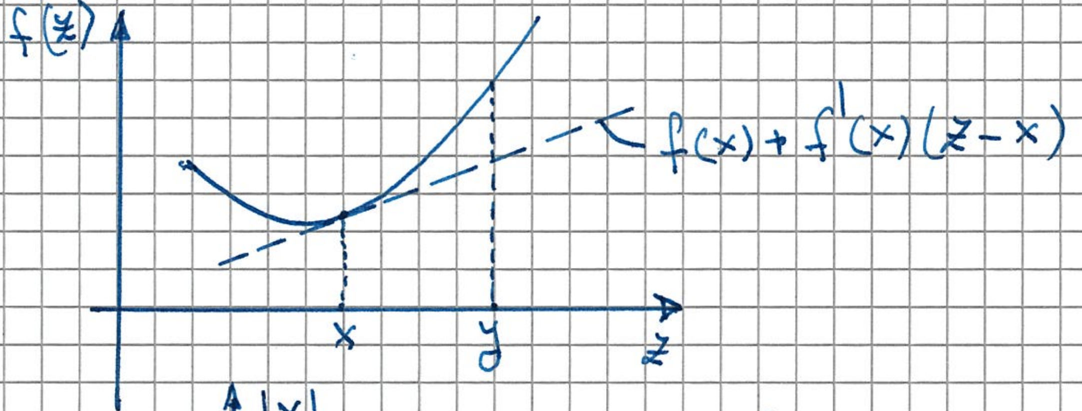
Teorema. Sia $f: \mathbb{R}^n \to \mathbb{R}$ differenziabile

$$f \text{ è convessa} \iff f(y) - f(x) \geq \nabla f(x)^T (y - x) \quad \forall (x,y)$$

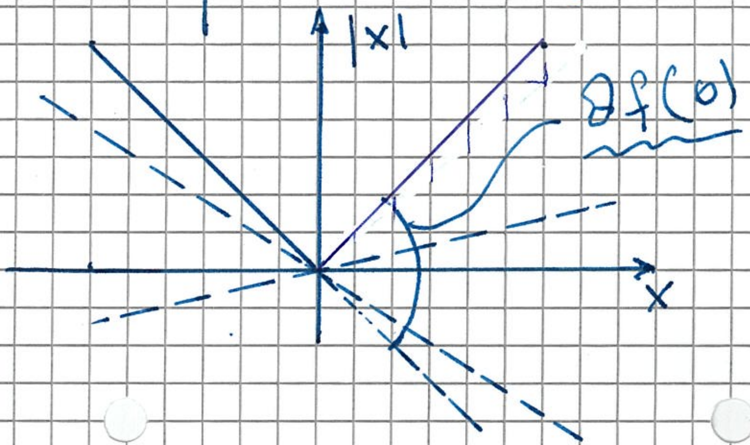
Dimostrazione (vedi dispensa)

Significato geometrico. Fissato x la relazione precedente può scriversi $\quad f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall y$

"Il grafico della funzione è sempre al di sopra di qualunque retta ad esso tangente"



La linearizzazione in un punto è sempre un'approssimazione per difetto di una funzione convessa.

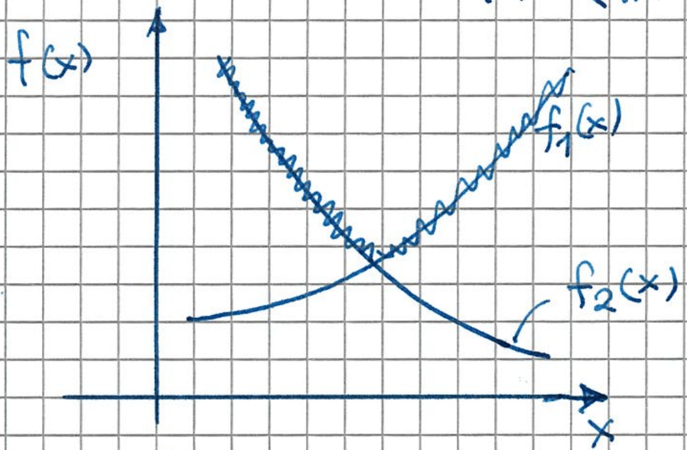


$f(x)$ nello 0 non è derivabile, tuttavia esistono molte rette che passano per l'origine e lasciano il grafo "tutto da una parte": sono tutte le rette di pendenza compresa tra -1 e $+1$

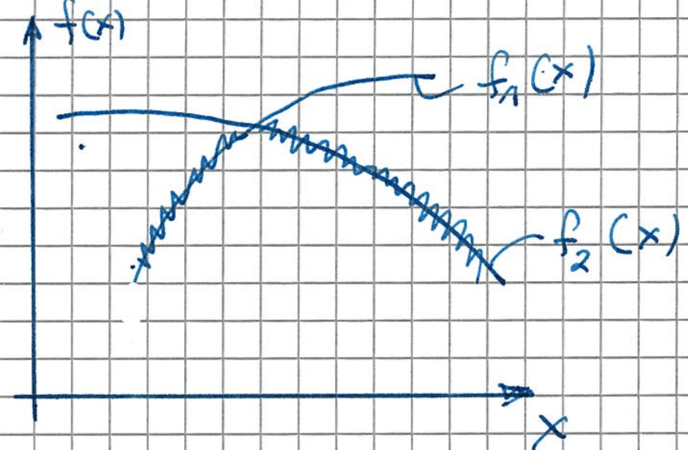
L'intervallo $[-1, 1]$ è il subdifferenziale di $|x|$ in 0.

Esercizio. Date m funzioni <u>convesse</u> $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, <u>dimostrare</u> che: $f(x) = \max_{1 \le i \le m} f_i(x)$ è convessa

Date m funzioni <u>concave</u> $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, <u>dimostrare</u> che: $f(x) = \min_{1 \le i \le m} f_i(x)$ è concava



$f(x) = \max_{1 \le i \le 2} f_i(x)$



$f(x) = \min_{1 \le i \le 2} f_i(x)$

Osservazione: Il <u>minimo</u> di funzioni <u>convesse</u> <u>non</u> è concava. Il <u>massimo</u> di funzioni <u>concave</u> <u>non</u> è convessa.

Nota. Se le m funzioni sono <u>affini</u>: $\underbrace{f_i(x) = c^{(i)T} x + a_i}_{\text{sia concava che convessa}}$ allora

$\max_{1 \le i \le m} f_i(x)$ è convessa

$\min_{1 \le i \le m} f_i(x)$ è concava

Richiami sulle matrici

Una matrice quadrata simmetrica A di dimensione $n \times n$ è *definita positiva* se $x^T A x > 0$, $\forall x \in \mathbb{R}^n$, $x \neq 0$; è *semidefinita positiva* se $x^T A x \geq 0$, $\forall x \in \mathbb{R}^n$.

Definizioni analoghe valgono per le matrici *definite negative* e *semidefinite negative*.

Teorema Una matrice A è definita positiva $\Longleftrightarrow$ I determinanti di tutti i suoi *minori principali* sono positivi

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n-1,1} & & & a_{n-1,n} \\ a_{n1} & & & a_{nn} \end{pmatrix}$$

Teorema A semidefinita positiva $\Longrightarrow$ I determinanti di tutti i suoi minori principali sono *non negativi*

Nota. $\Longleftarrow$ Non vale!!! $\begin{pmatrix} 0 & 0 \\ 0 & -3 \end{pmatrix}$ Tutti i determinanti dei minori principali sono ≥ 0, ma la matrice *non* è semidefinita positiva.

Teorema. Gli autovalori di una matrice definita positiva sono positivi

Prova. Sia $Av = \lambda v$ con $\lambda \leq 0 \Rightarrow v^T \underbrace{Av}_{\lambda v} = \lambda \|v\|^2 \leq 0$. Contraddiz.

* Teorema A definita positiva, λ_{min} e λ_{max} siano il più piccolo e il più grande degli autovalori $\Rightarrow \forall x$, $\lambda_{min}\|x\|^2 \leq x^T A x \leq \lambda_{max}\|x\|^2$

Prova. E' sufficiente ricordare che gli autovettori costituiscono una base ortonormale in $\mathbb{R}^n$

Teorema $f: \mathbb{R}^n \to \mathbb{R}$; $\underbrace{f \in C^{(2)}}$. Se in un punto x $\nabla^2 f(x)$ è definita positiva, allora f è convessa nell'intorno di x

Prova $f \in C^{(2)} \Rightarrow f(x+d) - f(x) = \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(x) d + O(\|d\|^2)$

$\underset{(Th. *)}{\Rightarrow}$ $f(x+d) - f(x) \geq \nabla f(x)^T d + \underbrace{\frac{1}{2}\lambda_{min}\|d\|^2 + O(\|d\|^2)}$

Il segno di una data quantità si mantiene se la stessa quantità viene divisa per un numero positivo. Quindi il segno di $\frac{1}{2}\lambda_{min}\|d\|^2 + O(\|d\|^2)$ è lo stesso di $\frac{1}{2}\lambda_{min} + \frac{O(\|d\|^2)}{\|d\|^2}$, somma di una quantità positiva e di un infinitesimo.

Quindi, per valori piccoli di $\|d\|$ (cioè nell'intorno di x) il segno di $\frac{1}{2}\lambda_{min}\|d\|^2 + o(\|d\|^2)$ è positivo $\Longrightarrow$

$$f(x+d) - f(x) \geq \nabla f(x)^T d \Longrightarrow \text{convessità (locale) di } f$$

(Th. fond. convessità)

Teorema $f \in C^{(2)}$

f è convessa $\Longleftrightarrow$ $\nabla^2 f(x)$ è semidefinita positiva $\forall x$

Prova (vedi dispensa)

Corollario Una funzione quadratica $f(x) = \frac{1}{2}x^T A x + b^T x$ è convessa se e solo se A è semidefinita positiva

Minimi (e massimi) di funzioni di più variabili

Non vincolato
$$\begin{cases} \min\limits_{x \in \mathbb{R}^n} f(x) \; . \quad x^* = \arg\min\limits_{x \in \mathbb{R}^n} f(x) \implies \exists \text{ un intorno } I_\varepsilon(x^*), \\ I_\varepsilon(x^*) = \{x \mid \|x - x^*\| \leq \varepsilon\} \text{ tale che } f(x^*) \leq f(x) \; \forall x \in I_\varepsilon(x^*) \end{cases}$$

Vincolato
$$\begin{cases} \min\limits_{x \in X \subset \mathbb{R}^n} f(x) \; . \quad x^* = \arg\min\limits_{x \in X \subset \mathbb{R}^n} f(x) \implies \exists \text{ un intorno } I_\varepsilon(x^*), \\ I_\varepsilon(x^*) = \{x \mid \|x - x^*\| \leq \varepsilon\} \text{ tale che } f(x^*) \leq f(x) \; \forall x \in \underbrace{I_\varepsilon(x^*) \cap X} \end{cases}$$

Ci riferiremo sempre a minimi locali e non globali.

(Minimo globale di f : x^* tale che $f(x^*) \leq f(x) \; \forall x \in \mathbb{R}^n$ (Caso non vincolato)

$f(x^*) \leq f(x) \; \forall x \in X$ (Caso vincolato)

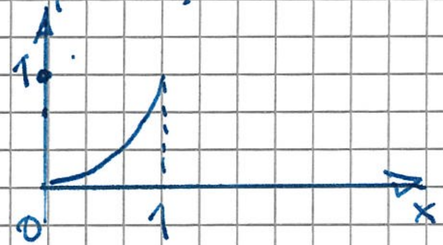
Ottimizzazione locale Ottimizzazione globale

Esistenza di punti di minimo.

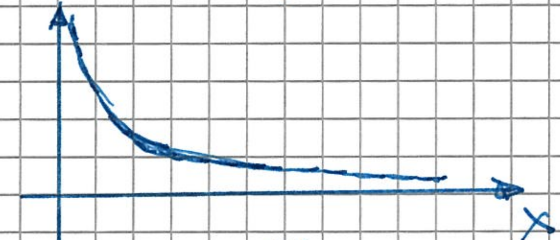
Teorema di Weierstrass. Una funzione $f: \mathbb{R}^n \to \mathbb{R}$ continua su un insieme chiuso e limitato ammette un minimo (e un massimo) globale.

Osservazioni.

a) f deve essere continua: $f(x)=\begin{cases} 1 & \text{se } x=0 \\ x^2 & \text{per } x \neq 0 \end{cases}$ non ammette minimo in $[0,1]$



b) L'insieme deve essere limitato. $f(x)=\frac{1}{x}$ non ammette un minimo in $(0,\infty)$

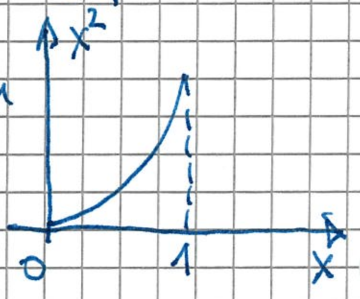


Nota $\inf\limits_{x\in(0,\infty)} f(x) = 0$

$$\left[f_{inf} = \inf_{x\in X} f(x) \Rightarrow \forall \varepsilon, \exists\, x \in X : f(x) < f_{inf} + \varepsilon \right]$$

c) L'insieme deve essere chiuso: $f(x)=x^2$ non ammette un minimo in $(0,1]$

aperto a sinistra

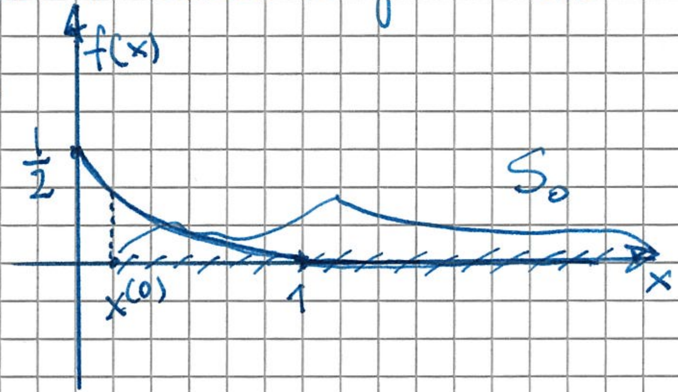


Corollario. f continua e $\exists\, x^{(0)}$ tale che $\{x \mid f(x) \le f(x^{(0)})\} = S_0$ è chiuso e limitato $\Rightarrow$ f ammette un minimo globale.

Prova. Il teorema di Weierstrass assicura che f ammette un punto di minimo globale su S_0, sia $x^* \in S_0$. Poiché $\forall\, x \notin S_0$ vale $f(x) \ge f(x^{(0)}) \ge f(x^*)$ segue che x^* è punto di minimo globale su $\mathbb{R}^n$

Nota. Il teorema definisce una condizione sufficiente ma non necessaria di esistenza di un punto di minimo globale.

Es $f(x) = \begin{cases} \frac{1}{2}(x-1)^2 & \text{se } x \in [0,1] \\ 0 & \text{se } x > 1 \end{cases}$



f ammette infiniti punti di minimo globale (tutta la semiretta $[1, \infty)$)

Def Una funzione $f: \mathbb{R}^n \to \mathbb{R}$ è coerciva se

$$\lim_{\|x\| \to \infty} f(x) = \infty$$

$$\left(\forall\, K \;\exists\, r(K) \text{ tale che } f(x) \ge K \;\; \forall\, x : \|x\| > r(K)\right) \text{ a } 1.$$

Teorema. $f: \mathbb{R}^n \to \mathbb{R}$ continua e coerciva $\Rightarrow$ Tutti gli insiemi di livello sono chiusi e limitati

Prova. La chiusura degli insiemi di livello è conseguenza della continuità di f $\left(\{x^{(k)}\} \subset S_\alpha,\ x^{(k)} \to \bar{x};\ f(x^{(k)}) \leq \alpha\ \forall k \right.$

$$\underbrace{\lim_{k\to\infty} f(x^{(k)}) = f(\bar{x})}_{\text{continuità}}. \quad \underbrace{f(\bar{x}) \leq \alpha}_{\text{Th. limiti delle successioni}} \Rightarrow \left. \bar{x} \in S_\alpha \right)$$

Limitatezza. Supponiamo per assurdo che $\exists\ \bar{\alpha}$ tale che $S_{\bar{\alpha}}$ non sia limitato $\Rightarrow \exists\ \{x^{(k)}\} \subset S_{\bar{\alpha}}$ con $\lim_{k\to\infty} \|x^{(k)}\| \to \infty$. Poichè f è coerciva segue che $\lim_{k\to\infty} f(x_k) = \infty$. Ciò contraddice il fatto che tutta la sequenza $\{x^{(k)}\} \subset S_{\bar{\alpha}}$

Condizioni di ottimalità per problemi non vincolati

Condizione necessaria I ordine (Th. Fermat)

Sia $f \in C^{(1)}$ e x^* punto di minimo $\Longrightarrow$ $\nabla f(x^*) = 0$

Prova. Se $\nabla f(x^*)$ fosse $\neq 0$, allora $d = -\nabla f(x^*)$ sarebbe una direzione di discesa. Contraddizione.

Nota La condizione è necessaria ma non è sufficiente

Def. Un punto $\bar{x}$ in cui è $\nabla f(\bar{x}) = 0$ è detto "stazionario"

Un punto stazionario o è un minimo o un massimo oppure un punto di sella (corrispondente al punto di flesso delle funzioni di una sola variabile).

Un punto di minimo nella direzione longitudinale, di massimo in quella trasversale

Condizione del II ordine (necessaria)

Sia $f \in C^{(2)}$. x^* punto di minimo $\Longrightarrow$

a) $\nabla f(x^*) = 0$ (Fermat)

b) $\nabla^2 f(x^*)$ è semidefinita positiva

Prova (b)) Comunque si scelga $d \in \mathbb{R}^n$, poiché $f \in C^{(2)}$, vale

$$f(x^* + td) = f(x^*) + t \underbrace{\nabla f(x^*)^T d}_{0} + \frac{1}{2} t^2 d^T \nabla^2 f(x^*) d + O(t^2 \|d\|^2)$$

$$\Longrightarrow \quad \frac{f(x^* + td) - f(x^*)}{t^2} = \frac{1}{2} d^T \nabla^2 f(x^*) d + \frac{O(t^2 \|d\|^2)}{t^2}$$

Poichè x^* è un punto di minimo, per valori di piccoli t il primo membro è nonnegativo e così il secondo

$$\frac{1}{2} d^T \nabla^2 f(x^*) d + \frac{O(t^2 \|d\|^2)}{t^2} \geq 0$$

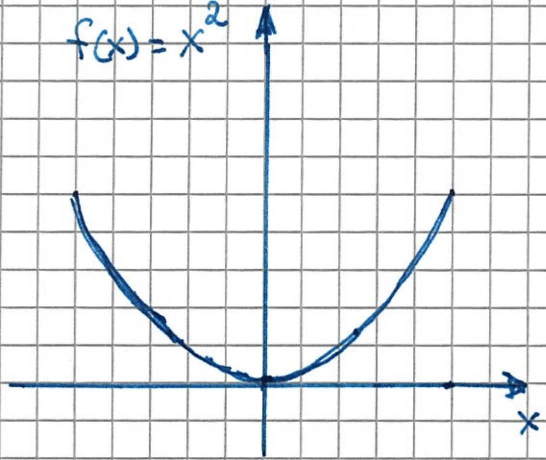
Passando al limite per $t \to 0$ segue che

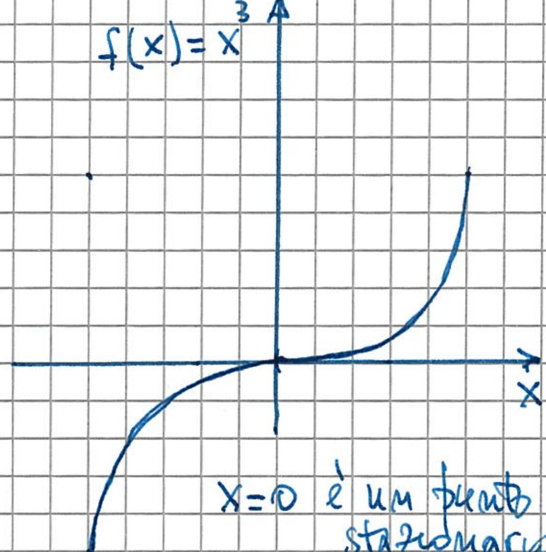
$$\frac{1}{2} d^T \nabla^2 f(x^*) d \geq 0 \quad (*)$$

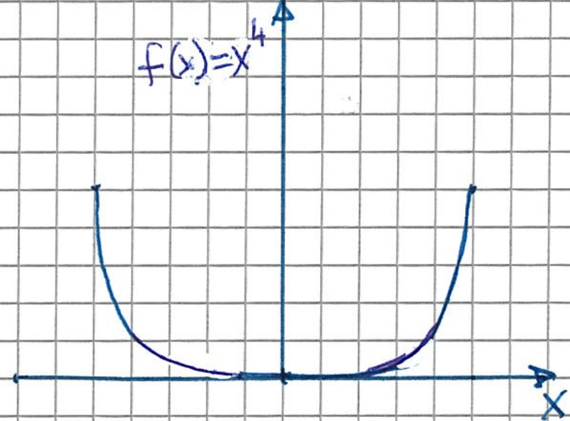
e il teorema segue perchè $(*)$ vale $\forall d$.

Condizione del secondo ordine (sufficiente)

Sia $f \in C^{(2)}$. Se in x^* è $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*)$ positiva definita allora x^* è un punto di minimo.







$x=0$ è un minimo

$f'(x) = 2x$; $f''(x) = 2$

$f'(0) = 0$ $\quad f''(0) = 2$

$\Downarrow$

$x=0$ soddisfa le cond. necessarie e sufficienti del II ordine

$x=0$ è un punto stazionario

$f'(x) = 3x^2$; $f''(x) = 6x$

$f'(0) = 0$; $f''(0) = 0$

Soddisfa solo le condizioni necessarie di minimo. Non è un minimo

$x = 0$ è un punto di minimo

$f'(x) = 4x^3$; $f''(x) = 12x^2$

$f'(0) = 0$; $f''(0) = 0$

Soddisfa le condizioni solo necessarie di minimo. È un minimo

Il soddisfacimento delle condizioni necessarie del II ordine non discrimina tra punto di minimo e stazionario.

Ottimizzazione di funzioni convesse

Teorema $f: \mathbb{R}^n \to \mathbb{R}$ convessa. x^* punto di minimo locale $\Rightarrow$
$\Rightarrow$ x^* è un punto di minimo globale

Prova. Sia x^* un punto di minimo locale $\Rightarrow$
$\Rightarrow \exists\, I_\varepsilon(x^*)$ tale che $f(x) \geq f(x^*)$ $\forall x \in I_\varepsilon(x^*)$
Supponiamo che x^* non sia un minimo globale $\Rightarrow$
$\Rightarrow \exists\, z \in \mathbb{R}^n$ tale che $f(z) < f(x^*)$
Consideriamo il segmento di estremi x^*, z cioè
$\{x \mid x = tz + (1-t)x^*,\ 0 \leq t \leq 1\}$.
f convessa $\Rightarrow$ $f(x) = f(tz + (1-t)x^*) \leq t f(z) + (1-t) f(x^*) <$
$< t f(x^*) + (1-t) f(x^*) = f(x^*)$, $\forall t \in (0, 1)$
Ma, per valori piccoli di t, x è quanto si vuole vicino ad x^*, quindi cade in $I_\varepsilon(x^*)$ e si ottiene quindi una contraddizione.

Teorema. $f: \mathbb{R}^n \to \mathbb{R}$ $\underbrace{\text{convessa}}$ $\overbrace{\sqrt{\ }}^{f \in C^{(1)}}$. $\nabla f(x^*) = 0 \Rightarrow x^*$ è un punto di minimo

Prova. $\forall x$ vale $f(x) \geq f(x^*) + \underbrace{\nabla f(x^*)^T (x - x^*)}_{0} = f(x^*)$

Nota. Per le funzioni convesse differenziabili $\nabla f(x^*) = 0$ è condizione necessaria e sufficiente di ottimalità.

Nota Una funzione convessa può anche avere <u>molti</u> punti di minimo <u>tutti con lo stesso valore di funzione</u>.

Esercizio Dimostrare che l'insieme dei punti di minimo di una funzione convessa è convesso.