Introduzione

La convessità occupa una parte importante della matemetica e viene sviluppata in diversi campi tra i quali la geometria e l'analisi.

Già negli *"Elementi"* di Eucide ci sono dei riferimenti al termine *convesso* che oggi è diventato familiare grazie anche alla geometria insegnata nelle scuole.

Le prime porprietà di questi insiemi furono studiate nel 1887 da Brunn, riprese poi a cavallo tra gli anni "800 e "900 da Minkowski. Nel corso del 1900 autori come Aleksandrov ([4]) ampliarono le conoscenze su questi insiemi sviluppando così proprietà anche nell'analisi convessa che prende in considerazione le funzioni.

L'intento, con questa tesi, è quello di raccogliere le principali caratteristiche di tali funzioni confrontandole con le funzioni studiate nell'analisi matematica. Un esempio di questa diversità lo si può trovare nella ricerca di massimi e di minimi: ogni punti stazionario è anche punto di minimo per le funzioni convesse. Mentre per le funzioni questa è una condizione necessaria ma non sufficiente, per le funzioni convesse è anche condizione sufficiente.

Noteremo come le funzioni convesse siano continue, localmente Lipschitziane in ogni punto interno al loro dominio e quindi grazie la teorema di Rademacher quasi ovunque differenziabili nei punti interni ([4], [1]). Inoltre rispetto alle funzioni localmente Lipshitziane esse hanno la caratteristica di essere due vole quasi ovunque differenziabili grazie al teorema di Aleksandrov ([5]). Dopo una prima parte introduttiva, dove definiremo insiemi e funzioni convesse, elencandone le prime proprietà e fornendo vari esempi, procederemo con lo studio delle proprietà in dimensione uno, estendendo tale studio in un secondo momento alle funzioni di più variabili.

La tesi si concluderà con le proprietà delle successioni di funzioni e quindi con il confronto tra varie nozioni di convergenza. Abbiamo in particolare studiato il legame tra le convergenze puntuale, uniforme, in norma e la convergenza dei grafici.

Capitolo 1

Preliminari

Ricordiamo in questo capitolo alcune nozioni e definizioni preliminari che ci serviranno nel prosieguo della tesi. Cercheremo in questo capitolo di restare in un contesto il più generale possibile, particolareggiando le definizioni nei capitoli successivi.

Iniziamo con la definizione di insieme convesso.

Definizione 1.1. Sia V uno spazio vettoriale. Allora $\Omega \subseteq V$ viene detto convesso se per ogni coppia di punti $x, y \in \Omega$ si ha

$$[x, y] := \{tx + (1 - t)y; t \in [0, 1]\} \subseteq \Omega.$$

Si veda la figura (1.1).

Una prima e semplice conseguenza della convessità di insieme è la connessione; l'insieme convesso è inoltre semplicemente connesso essendo un caso particolare di insieme stellato avente come centro ogni suo punto $x_0 \in \Omega$.

Una volta data la nozione di insieme convesso, possiamo passare alla definizione di funzione convessa.

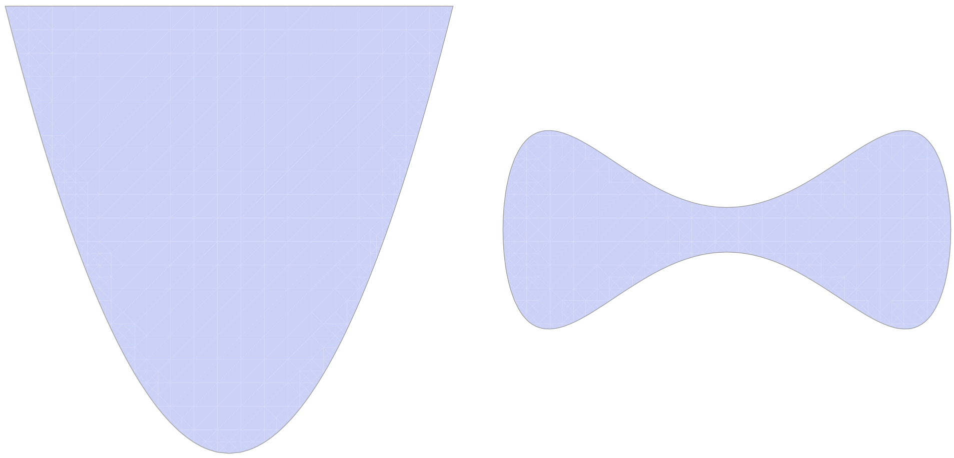
Definizione 1.2. Dato un dominio $\Omega \subset V$ convesso, diremo che una funzione $f\colon \Omega \to \mathbb{R}$ é convessa se per ogni coppia di punti $x, y \in \Omega$

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) \qquad \forall t \in [0, 1]. \tag{1.1}$$

Una funzione si dice *strettamente convessa* se nella formula (1.1) si ha la disuguaglianza stretta per ogni $t \in (0, 1)$. Una funzione si dice concava (rispettivamente strettamente concava) se $-f$ è convessa (rispettivamente strettamente convessa); ciò equivale a dire che per ogni coppia di punti $x, y \in \Omega$

$$f(tx + (1 - t)y) \geq tf(x) + (1 - t)f(y) \qquad \forall t \in [0, 1],$$

con disuguaglianza stretta nel caso di stretta concavità.



(a) *Insieme convesso.* (b) *Insieme non convesso.*

Figura 1.1: Esempi di insieme convesso e non

Osservazione 1.1. *1. La definizione di convessità di una funzione ha una semplice interpretazione geometrica se guardiamo al grafico della funzione convessa; infatti, la convessità equivale alla richiesta che per ogni scelta dei punti $x, y \in \Omega$, allora scelto comunque un punto $x_t \in [x, y]$ appartenente al segmento che congiunge x con y, il punto $(x_t, f(x_t))$ si trova al di sotto del segmento che congiunge i punti $(x, f(x))$ e $(y, f(y))$.*

2. Possiamo notare come nella matematica le funzioni non sono divise in funzioni convesse e funzioni concave; esistono infatti esempi di funzioni non convesse né concave, come ad esempio la funzione x^3 in dimensione 1. Ci sono anche funzioni che non sono né concave né convesse in nessun punto del dominio, ad esempio, in dimensione 2, la funzione $f(x, y) = x^2 - y^2$.

Vediamo subito che esiste una stretta relazione tra insiemi e funzioni convesse; tale relazione passa attraverso la nozione di epi– e sotto–grafico.

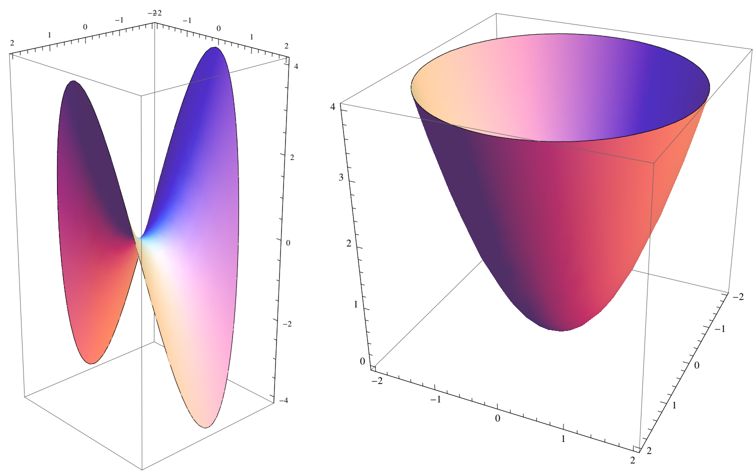
Definizione 1.3. Si chiama *epi–grafico* della funzione f, e si inidca con $Epi(f)$, l'insieme così definito

$$Epi(F) = \{(x, \alpha) \in \Omega \times \mathbb{R}, \alpha \geq f(x)\}. \tag{1.2}$$

Si chiama *sotto–grafico* di f, e si indica con $Sub(f)$, l'insieme così definito

$$Sub(f) = \{(x, \alpha) \in \Omega \times \mathbb{R}, \alpha \leq f(x)\}. \tag{1.3}$$

Abbiamo quindi la seguente proposizione.



(a) *Funzione non convessa* (b) *Funzione convessa*

Figura 1.2: Esempi di funzioni convesse e non in dimensione 2

.

Proposizione 1.1. *La funzione $f : \Omega \to \mathbb{R}$ è convessa se e solo se il suo epi–grafico è convesso; essa è concava se e e solo se il suo sotto–grafico è convesso.*

Dimostrazione. Per procedere con la dimostrazione ci rfaremo al disegno (1.3). Supponiamo f convessa e siano $(x_1, a), (x_2, b) \in Epi(f)$ e r il segmento che li congiunge. Dobbiamo dimostrare che questo segmento appartiene a $Epi(f)$. Sia quindi (x, y) un punto qualsiasi di tale segmento; avremo cioè un $\lambda \in (0, 1)$ per cui

$$x = x_1 + \lambda(x_2 - x_1) = \lambda x_2 + (1 - \lambda)x_1$$
$$y = a + \lambda(b - a) = \lambda b + (1 - \lambda)a.$$

Dato che $a \geq f(x_1)$ e $b \geq f(x_2)$, otteniamo, grazie alla convessità di f, che

$$\begin{aligned} y = \lambda b + (1 - \lambda)a \geq & \lambda f(x_2) + (1 - \lambda)f(x_1) \\ \geq & f(\lambda x_2 + (1 - \lambda)x_1) = f(x) \end{aligned}$$

cioè $y \geq f(x)$ e quindi $(x, y) \in Epi(f)$.

Viceversa, mostriamo anzitutto che se $Epi(f)$ è convesso, allora Ω è convesso; a tal fine, fissiamo $x, y \in \Omega$ e $\alpha \geq \max\{f(x), f(y)\}$. La convessità di $Epi(f)$ implica che il segmento

$$[(x, \alpha), (y, \alpha)] = [x, y] \times \{\alpha\} \subset Epi(f),$$

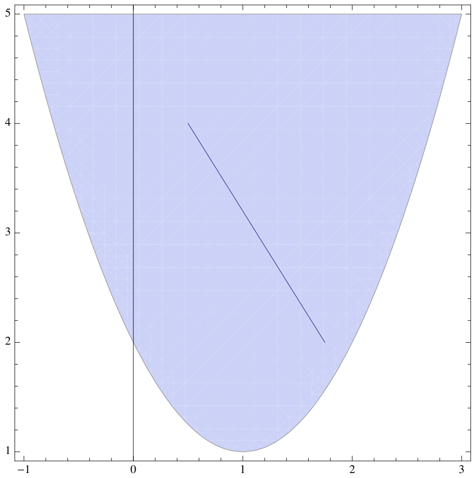


Figura 1.3: Un esempio di funzione convessa

e quindi in particolare $[x, y] \subset \Omega$. Se si considerano ora i punti

$$(x_1, f(x_1)), (x_2, f(x_2)) \in Epi(f),$$

allora la convessità di $Epi(f)$ implica che $(tx_1 + (1-t)x_2, tf(x_1) + (1-t)f(x_2)) \in Epi(f)$, cioè

$$tf(x_1) + (1-t)f(x_2) \geq f(t(x_1) + (1-t)f(x_2)),$$

cioè la convessità di f.

Per dimostrare la seconda parte basta solo ricordare che f concava equivale a dire che $-f$ è convessa e che un punto $(x, \alpha) \in Sub(f)$ se e solo se $(x, -\alpha) \in Epi(-f)$. Da qui la proposizione. $\square$

Proponiamo la seguente osservazione, che stabilisce una ulteriore connessione tra funzioni ed insiemi convessi.

Osservazione 1.2. *Notiamo che se f è convessa allora i sotto–livelli di f, cioè gli insiemi*

$$E_\alpha(f) = \{f < \alpha\} = \{x \mid f(x) < \alpha\}$$

sono convessi. Infatti se presi $x_1, x_2 \in E_\alpha$, cioè se $f(x_1) < \alpha$, $f(x_2) < \alpha$, allora

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) < \alpha \tag{1.4}$$

e quindi $tx_1 + (1-t)x_2 \in E_\alpha(f)$. Non è vero il viceversa, cioè che se i sotto-livelli sono convessi, allora la funzione è convessa: si prenda ad esempio la funzione

$$f(x) = \sqrt{\|x\|} \quad per\ x \in \mathbb{R}^n.$$

La nozione di convessità equivale alla richiesta che per ogni scelta di punti $x_1, \dots, x_m \in \Omega$ e per ogni scelta $\lambda_1, \dots, \lambda_m \in [0,1]$ tali che

$$\sum_{i=1}^{m} \lambda_i = 1,$$

allora

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i). \tag{1.5}$$

Infatti quet'ultima formula diventa la definizione di convessità di f nel caso $m = 2$, mentre la disuaglianza precedente si ottiene dalla convessità di f mediante induzione. Infatti, se fissiamo $x_1, \dots, x_m \in \Omega$ e $\lambda_1, \dots, \lambda_m$ come sopra, supponendo $\lambda_m \neq 0, 1$, possiamo porre $t = 1 - \lambda_m$ e

$$y_1 = \frac{1}{1-\lambda_m}(\lambda_1 x_1 + \dots \lambda_{m-1} x_{m-1}), \qquad y_2 = x_m,$$

in modo da ottenere che

$$\begin{aligned} f(\lambda_1 x_1 + \dots \lambda_{m-1} x_{m-1} + \lambda_m x_m) = & f(t y_1 + (1-t) y_2) \\ \leq & t f(y_1) + (1-t) f(y_2) \\ = & \frac{1}{1-\lambda_m} f(\lambda_1 x_1 + \dots \lambda_{m-1} x_{m-1}) + \lambda_m f(x_m). \end{aligned}$$

L'induzione si conclude notando che se ridefinendo

$$\tilde{\lambda}_i = \frac{1}{1-\lambda_m} \lambda_i, \qquad i = 1, \dots, m-1,$$

allora

$$\sum_{i=1}^{m-1} \tilde{\lambda}_i = 1.$$

Riscrivere la definizione di convessità nel precedente modo ha la sua rilevanza nel seguente importante Teorema.

Teorema 1.1 (Disuguaglianza di Jensen)**.** *Sia $(X, \mathcal{F}, \mu)$ uno spazio di probabilità, cioè X insieme, $\mathcal{F}$ σ–algebra su X e μ misura di probabilità su X. Sia poi $g \in L^1(X, \mu)$ una funzione integrabile e $f : \mathbb{R} \to \mathbb{R}$ una funzione convessa. Allora*

$$f\left(\int_X g(x) d\mu(x)\right) \leq \int_X f(g(x)) dx.$$

Dimostrazione. Basta ricordare che la definizione di integrale di g viene data mediante approssimazione con funzioni semplici; cioè il valore

$$\int_X g(x)d\mu$$

viene ottenuto come limite dei valori

$$\int_X s(x)d\mu(x) = \sum_{i=1}^{m} c_i\mu(A_i)$$

dove

$$s(x) = \sum_{i=1}^{m} c_i\chi_{A_i}(x)$$

è una funzione semplice che approssima in modo opportuno la funzione g. Gli insiemi A_i formano una partizione di X, cioè

$$X = \bigcup_{i=1}^{m} A_i,$$

e quindi, dato che la misura μ è di probabilità,

$$1 = \mu(X) = \sum_{i=1}^{m} \mu(A_i);$$

possiamo usare la (1.5) con $\lambda_i = \mu(A_i)$ per ottenere che

$$\begin{aligned} f\left(\int_X s(x)d\mu(x)\right) =& f\left(\sum_{i=1}^{m} c_i\mu(A_i)\right) = f\left(\sum_{i=1}^{m} \lambda_i c_i\right) \leq \sum_{i=1}^{m} \lambda_i f(c_i) \\ =& \sum_{i=1}^{m} f(c_i)\mu(A_i) = \int_X f(s(x))d\mu(x). \end{aligned}$$

Come vedremo nella Proposizione 2.1, le funzioni convesse di una variabile reale sono continue, quindi approssimando g mediante funzioni semplici s si ottiene il risultato. □

Osservazione 1.3. *1. L'ipotesi su μ di essere misura di probabilità può essere indebolita richiedendo che μ sia una misura positiva finita nel senso che*

$$\mu(X) < +\infty.$$

Infatti, in tal caso si può definire la nuova misura

$$\tilde{\mu} = \frac{1}{\mu(X)}\mu$$

in modo da ottenere una misura di probabilità. In questo caso la disuguaglianza di Jensen diventa quindi

$$f\left(\frac{1}{\mu(X)}\int_X g(x)d\mu(x)\right) \leq \frac{1}{\mu(X)}\int_X f(g(x))d\mu(x)$$

per ogni funzione $g \in L^1(X,\mu)$.

2. *La disuguaglianza di Jensen fornisce una caratterizzazione equivalente delle funzioni convesse di una variabile; infatti, se* f *è convessa, abbiamo dimostrato il Teorema 1.1. Viceversa, se si suppone che* f *sia una funzione per la quale il Teorema 1.1 sia valido per ogni scelta della misura di probabilità* μ, *allora* f *deve necessariamente essere convessa. Basta infatti considerare la misura*

$$\mu = \sum_{i=1}^{m} \lambda_i \delta_{x_i}$$

dove δ_{x_i} *si indica la misura delta di* Dirac *concentrata in* x_i *definita per ogni* $A \subset \mathbb{R}$ *ponendo*

$$\delta_{x_i}(A) = \begin{cases} 1 & se\ x_i \in A \\ 0 & se\ x_i \notin A. \end{cases}$$

Questo implica che per ogni funzione $g : \mathbb{R} \to \mathbb{R}$

$$\int_{\mathbb{R}} g(x)d\mu(x) = \sum_{i=1}^{m} \lambda_i g(x_i);$$

quindi se vale il Teorema1.1, presa la funzione $g(x) = x$, *dato che* $g \in L^1(\mathbb{R},\mu)$, *ritroviamo la disuguaglianza* (1.5) *in quanto*

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) = f\left(\int_{\mathbb{R}} x d\mu(x)\right) \leq \int_{\mathbb{R}} f(x)d\mu(x) = \sum_{i=1}^{m} \lambda_i f(x_i),$$

da cui la convessità di f.

3. *Il Teorema 1.1 è stato dimostrato per funzioni convesse di una variabile reale, ma con la stessa dimostrazione può essere esteso a funzioni convesse in* $\mathbb{R}^n$; *un pò più delicata è l'estensione a funzioni definite in spazi vettoriali* V *generici, in quanto in tal caso bisognerebbe considerare lo spazio* $L^1(X,\mu,V)$ *delle funzioni integrabili a valori in* V.

1.1 Alcuni esempi di funzioni convesse

Dopo aver dato un quadro generale sulle funzioni convesse possiamo elencare quelle funzioni, utilizzate spesso nella matematica, che sono funzioni convesse.

- Una delle prime funzioni che andiamo a considerare è la funzione norma. Infatti, dalla disuguaglianza triangolare
$$\|x+y\| \leq \|x\| + \|y\|$$
e dalla proprietà
$$\|\lambda x\| = |\lambda| \|x\|,$$
segue facilmente la convessità della funzione norma.

- Esempi importanti di norme sono le norme degli spazi $L^p(X, \mu)$, con X spazio misurabile e μ misura su X. Ricordiamo che la norma in $L^p(X, \mu)$ e $p \in [1, +\infty)$ è definita da
$$\|f\|_{p,X} = (\int_X |f(x)|^p)^{\frac{1}{p}} d\mu(x).$$
Il fatto che tale funzione sia una norma e quindi una funzione convessa segue dalla disuguaglianza di Minkowski: date $f, g \in L^p(X, \mu)$
$$\|f+g\|_{p,X} \leq \|f\|_{p,X} + \|g\|_{p,X}.$$
Tali norme definiscono funzioni strettamente convesse se e solo se $p > 1$, mentre la norma $\|\cdot\|_{1,X}$ non è strettamente convessa, per lo stesso motivo per cui la funzione valore assoluto su $\mathbb{R}$ è convessa ma non strettamente convessa.

- Dato $(V, \|\cdot\|)$ spazio normato, un altro esempio di funzione convessa è la funzione distanza da un sottoinsieme convesso $\Omega \subseteq V$,
$$f(x) = \operatorname{dist}(x, \Omega) = \inf\{\|x-y\|, y \in \Omega\}.$$
Infatti se $x, y \in V$ e $\lambda \in (0,1)$ posto
$$z = (1-\lambda)x + \lambda y$$
e fissato $\varepsilon > 0$ siano $c, d \in \Omega$ tali che
$$\|x-c\| \leq \operatorname{dist}(x, \Omega) + \varepsilon \qquad \|y-d\| \leq \operatorname{dist}(y, \Omega) + \varepsilon,$$

allora preso $e = (1-\lambda)c + \lambda d \in \Omega$

$$\begin{aligned}\mathrm{dist}(z,\Omega) &\leq \|z-e\| \leq (1-\lambda)\|x-c\| + \lambda\|y-d\| \\ &\leq (1-\lambda)(\mathrm{dist}(x,\Omega)+\varepsilon) + \lambda(\mathrm{dist}(y,\Omega)+\varepsilon) \\ &= (1-\lambda)\mathrm{dist}(x,\Omega) + \lambda\mathrm{dist}(y,\Omega) + \varepsilon.\end{aligned}$$

Dall'arbitrarietà di ε si ottiene la convessità della funzione.

- Un' altra importante funzione è la funzione indicatrice. Sia V uno spazio vettoriale e sia $\Omega \subseteq V$. Si definisce *funzione identicatrice* la seguente
$$\delta_\Omega(x) = \begin{cases} 0 & \text{se } x \in \Omega \\ +\infty & \text{se } x \notin \Omega. \end{cases}$$
Notiamo come questa definizione sia ben posta usando la relazione
$$\chi_\Omega = \frac{1}{1+\delta_\Omega}.$$
Infatti con questa si ha
$$\chi_\Omega(x) = \begin{cases} 1 & \text{per } x \in \Omega \\ 0 & \text{per } x \notin \Omega, \end{cases}$$
che equivale proprio alla definizione di funzione identicatrice. Questa funzione è convessa se e solo se l'insieme Ω è conveso.

  Osservazione 1.4. *L'introduzione di questa particolare funzione è lecita se tra i valori ammissibili per una funzione convessa viene permesso anche* $+\infty$.

1.2 Massimi e minimi

In questo paragrafo si prenderà in considerazione il problema della ricerca del massimo e del minimo di una funzione convessa su di un insieme convesso Ω. La dimostrazione dell'esistenza di massimi o minimi ha dei legami con la topologia dello spazio ambiente, come mostrano i seguenti risultati; alcune condizioni sulla topologia implicano l'esistenza del massimo o del minimo, ma in generale non si possono dare condizioni necessarie per l'esistenza, come ci si può facilmente rendere conto cambiando ad esempio la topologia di $\mathbb{R}$.

Ricordiamo prima di iniziare la ricerca dei minimi la seguente:

Definizione 1.4. Sia $\mathcal{D}$ il dominio della funzione $f : \mathcal{D} \to \mathbb{R}$. Un punto $\bar{x}$ si dice *minimo globale* se per ogni x appartenente a $\mathcal{D}$ si ha $f(\bar{x}) \leq f(x)$.

Prima di passare al prossimo teorema, ricordiamo la definizione di *insieme sequenzialmente compatto.*

Definizione 1.5. Sia C sottoinsieme di uno spazio topologico. Diremo che l'insieme C è *sequenzialmente compatto* se ogni successione di punti di C ammette una sottosuccessione convergente in C. In formula matematica

$$\begin{aligned} \forall (x_n)_{n\in\mathbb{N}} \subseteq C \quad & \text{successione contenuta in } C \\ \exists x_{k_n} \quad & \text{sottosuccesione e } x \in C \text{ tale che} \\ & x_{k_n} \to x \quad \text{per } n \to \infty. \end{aligned}$$

Enunciamo dunque il seguente.

Teorema 1.2. *Sia Ω uno spazio topologico e sia $f \colon \Omega \to \mathbb{R}$ una funzione continua tale che esiste un $c \in \mathbb{R}$ con*

$$\mathcal{C} = \{x \in \Omega \, \textit{tale che} \, f(x) \leq c\} \subseteq \Omega \qquad \mathcal{C} \neq \emptyset$$

sequenzialmente compatto. Allora esiste il

$$\min_{\Omega} f.$$

Dimostrazione. Poniamo $\alpha = \inf_{\Omega} f$. Dimostreremo che $\alpha = \min_{\Omega} f$.
Prima di tutto notiamo che α non può essere uguale a $-\infty$. Infatti, se supponiamo per assurdo che lo sia, allora per ogni $n \in \mathbb{N}$ esiste x_n tale che

$$f(x_n) \leq -n.$$

Dato che esisterà un n_0 tale che $-n \leq c$ per ogni $n \geq n_0$,

$$x_n \in \{f \leq -n\} \subseteq \{f \leq c\} = \mathcal{C}, \quad \forall n \geq n_0.$$

Poichè abbiamo detto essere sequenzialmente compatto, esiste una sottosuccessione x_{k_n} e un $x \in \mathcal{C}$ tale che

$$x_{n_n} \to x.$$

Inoltre f continua implica

$$f(x) = \lim_{n\to\infty} f(x_{k_n}) = -\infty,$$

assurdo perchè $f(x) \in \mathbb{R}$.
Supponiamo dunque che $\alpha > -\infty$. Se $\alpha = c$, allora la dimostrazione è conclusa. Se Invece $\alpha < c$, dato che $\alpha = \inf_\Omega f$, possiamo costruire una successione x_n per chi $f(x_n) \leq \alpha + 1/n$. Esisterà poi $n_0 \in \mathbb{N}$ tale che $\alpha + 1/n < c$ per ogni $n \geq n_0$. Quindi, per ogni $n \geq n_0$

$$\alpha \leq f(x_n) \leq \alpha + \frac{1}{n} \leq c$$

cioè la successione x_n è contenuta in $\mathcal{C}$, quindi, dalla sequenziale compattezza, esiste una sottosuccession

$$x_{k_n} \to x \in C \quad \text{per } n \to \infty.$$

Sempre per la continuità di f

$$f(x) = \lim_{n\to\infty} f(x_{n_n}) = \alpha.$$

Abbiamo così dimostrato che α è proprio il minimo di f nell'insieme Ω. □

Il risultato precedente può essere esteso anche al caso in cui esista un sotto-livello compatto e non necessariamente sequenzialmente compatto. A tal fine è necessario ricordare il seguente risultato di topologia.

Lemma 1.1. *Sia $(\Omega, \mathcal{T})$ uno spazio topologico T2 e siano $(K_n)_{n\in\mathbb{N}}$ una successione di insiemi compatti non vuoti tali che $K_{n+1} \subset K_n$; allora*

$$\bigcap_{n\in\mathbb{N}} K_n \neq \emptyset.$$

Dimostrazione. Supponiamo per assurdo che

$$\bigcap_{n\in\mathbb{N}} K_n \neq \emptyset;$$

ciò equivale a dire che gli insiemi, che sono aperti grazie al fatto che la topologia è T2,

$$A_n = \Omega \setminus K_n$$

sono un ricoprimento aperto di Ω, cioè

$$\Omega = \bigcup_{n\in\mathbb{N}} A_n.$$

Fissiamo $\bar{n} \in \mathbb{N}$; dato che $K_{\bar{n}} \subset \Omega$ è ricoperto dagli $(A_n)_{n\in\mathbb{N}}$, per compattezza esistono $n_1 < \ldots < n_h \in \mathbb{N}$ tali che

$$K_{\bar{n}} \subset \bigcup_{i=1}^{h} A_{n_i} = \Omega \setminus \bigcap_{i=1}^{h} K_{n_i}.$$

Dato che gli insiemi K_{n_i} contengono tutti l'insieme K_{n_h}, otteniamo quindi che

$$K_{\bar{n}} \subset \Omega \setminus K_{n_h},$$

cioè

$$K_{\bar{n}} \cap K_{n_h} = \emptyset,$$

che è un assurdo in quanto uno dei due deve essere contenuto nell'altro. □

Possiamo quindi ripetere il teorema di esistenza del minimo.

Teorema 1.3. *Sia Ω uno spazio topologico T2 e sia $f : \Omega \to \mathbb{R}$ una funzione continua tale che esiste un $c \in \mathbb{R}$ con*

$$\mathcal{C} = \{x \in \Omega \; \textit{tale che} \, f(x) \leq c\} \subseteq V \qquad \mathcal{C} \neq 0$$

compatto. Allora esiste il

$$\min_{\Omega} f.$$

Dimostrazione. Basta ripetere la dimostrazione del Teorema precedente; per dimostrare che

$$\alpha = \inf_{\Omega} f$$

non sia $-\infty$, bisognerà considerare gli insiemi

$$K_n = \{f \leq -n\}.$$

Tali insiemi saranno compatti se $-n \leq c$ e soddisfano le ipotesi del Lemma 1.1; quindi

$$\bigcap_{n\in\mathbb{N}} K_n \neq \emptyset,$$

e quindi esiste un $x \in \Omega$ con $x \in K_n$ per ogni $n \in \mathbb{N}$, cioè

$$f(x) \leq -n, \qquad \forall n \in \mathbb{N},$$

cosa che non può essere.

A questo punto, dato che $\alpha > -\infty$, basterà considerare gli insiemi

$$K_n = \left\{ f \leq \alpha + \frac{1}{n} \right\}$$

ed applicare ancora il Lemma 1.1 per trovare che esiste $x \in \Omega$ tale che

$$\alpha \leq f(x) \leq \alpha + \frac{1}{n}, \qquad \forall n \in \mathbb{N},$$

e quindi $f(x) = \alpha$, cioè x è un punto di minimo. $\square$

Proposizione 1.2. *Supponiamo che $f : \Omega \to \mathbb{R}$ sia una funzione strettamente convessa con $\Omega \subset V$ insiemi convesso; allora se il minimo di f su Ω esiste, allora il punto di minimo è unico.*

Dimostrazione. Supponiamo che esistano due punti di minimo x_1 e x_2 con $x_1 \neq x_2$.
Sia $\alpha = \min_\Omega f$: allora poichè x_1 e x_2 sono due punti di minimo si ha che

$$\alpha = f(x_1) \qquad \alpha = f(x_2).$$

Se ne deduce, dalla stretta convessità, che Poichè la funzione è strettamente convessa si ha

$$\alpha \leq f(tx_1 + (1-t)x_2) < tf(x_1) + (1-t)f(x_2) = \alpha,$$

il che è un assurdo. Allora $x_1 = x_2$. $\square$

Per quanto riguarda il problema della ricerca dei massimi di una funzione convessa, il meglio che in generale si possa dire è dato dal seguente Teorema. Supporremo qui che lo spazio vettoriale in considerazione sia uno spazio normato, cioè che sia definita una norma $\|\cdot\|$ su V; in questo modo la topologia indotta da tale norma è T2 ed inoltre la restrizione di tale norma sui sottospazi finito dimensionali di V definisce una norma che sarà equivalente alla norma Euclidea di tale sottospazio. Useremo in particolare questa osservazione per avere che la topologia sulle rette contenute in V è la topologia Euclidea e quindi, in particolare, ogni insieme compatto contenuto in una retta è dato dall'unione finita di intervalli.

Teorema 1.4. *Sia V uno spazio normato e $f : \Omega \to \mathbb{R}$ una funzione convessa con $\Omega \subset V$ insieme aperto e convesso; allora, per ogni compatto $K \subset \Omega$*

$$\sup_\Omega f = \sup_{\Omega \setminus K} f.$$

Dimostrazione. Supponiamo per assurdo che

$$\sup_K f > \sup_{\Omega \setminus K} f;$$

esiste quindi $\varepsilon > 0$ e $x_0 \in K$ tali che

$$f(x_0) - \varepsilon \geq f(x), \qquad \forall x \in \Omega \setminus K.$$

Fissiamo un $x \in \Omega \setminus K$ e consideriamo quindi la retta r parametrizzata da

$$x_t = x_0 + t(x - x_0);$$

definiamo poi

$$t_0 = \inf\{t < 0 : x_t \in \Omega\}, \quad t_1 = \min\{t \leq 0 : x_t \in K\}.$$

Preso quindi $t_0 < t < t_1 \leq 0$, avremo che il punto x_0 è un punto intermedio tra x_t e x, cioè

$$x_0 = \lambda x_t + (1 - \lambda)x, \qquad \lambda = \frac{1}{1-t} \in (0, 1).$$

Quindi

$$f(x_0) \leq \lambda f(x_t) + (1 - \lambda)f(x) = \frac{1}{1-t} f(x_t) - \frac{t}{1-t} f(x),$$

da cui

$$f(x_t) \geq f(x_0) + t(f(x) - f(x_0)) \geq f(x_0) - t\varepsilon > f(x_0),$$

da cui l'assurdo dato che $x_t \in \Omega \setminus K$. $\square$

Il risultato appena dimostrato asserisce in pratica che la ricerca del massimo della funzione f va ricercato verso il bordo di Ω; saremo più precisi in tal senso nel prossimo capitolo.

Capitolo 2

Funzioni convesse uni–dimensionali

In questo capitolo vogliamo elencare le proprietà più importanti delle funzioni convesse partendo con il caso più semplice, senza porre delle particolari condizioni di regolarità, fino ad arrivare a proprietà di funzioni regolari. Per le dimostrazioni e le proposizioni sono state rielaborate informazioni tratte da varie opere [8], [1], [6] [5].

Premettiamo il seguente Lemma.

Lemma 2.1. *Siano $x_0 < x_1 < x_2$ punti di I e sia $f : I \to \mathbb{R}$ convessa; allora i rapporti incrementali di f sono monotoni, cioè*

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_2) - f(x_0)}{x_2 - x_0}. \tag{2.1}$$

Dimostrazione. Possiamo esprimere x_1 come combinazione convessa di x_0 e x_2, cioè

$$x_1 = x_0 + t(x_2 - x_0), \qquad t = \frac{x_1 - x_0}{x_2 - x_0}.$$

Quindi la convessità di f implica che

$$f(x_1) = f(tx_2 + (1-t)x_0) \leq tf(x_2) + (1-t)f(x_0) = f(x_0) + \frac{x_1 - x_0}{x_2 - x_0} f(x_2),$$

che equivale a

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_2) - f(x_0)}{x_2 - x_0}.$$

□

La monotonia dei rapporti incrementali permette di dimostrare il seguente risultato di derivabilità di f.

Teorema 2.1. *Sia I un intervallo di $\mathbb{R}$ e sia $f : I \to \mathbb{R}$ convessa. Esistono allora per ogni x punto interno di I le derivate destre e sinistre*

$$f'(x^-) = \lim_{y \to x^-} \frac{f(y) - f(x)}{y - x}, \quad f'(x^+) = \lim_{y \to x^+} \frac{f(y) - f(x)}{y - x}.$$

Inoltre, le funzioni

$$x \mapsto f'(x^-), \quad x \mapsto f'(x^+)$$

sono monotone crescenti e

$$f'(x^-) \leq f'(x^+).$$

Dimostrazione. Fissiamo x punto interno di I e consideriamo $x < y_1 < y_2$ punti di I; grazie al Lemma 2.1 con $x_0 = x$, $x_1 = y_1$ e $x_2 = y_2$ sappiamo che

$$\frac{f(y_1) - f(x)}{y_1 - x} \leq \frac{f(y_2) - f(x)}{y_2 - x},$$

cioè la funzione

$$y \mapsto \frac{f(y) - f(x)}{y - x}$$

è monotona crescente, da cui il fatto che

$$\exists \lim_{y \to x^+} \frac{f(y) - f(x)}{y - x} = f'(x^+).$$

Analogamente, prendendo $y_2 < y_1 < x$ e $x_0 = y_2$, $x_1 = y_1$ $x_2 = x$ nel Lemma 2.1 si dimostra l'esistenza di

$$f'(x^-) = \lim_{y \to x^-} \frac{f(y) - f(x)}{y - x}.$$

Se poi si fissa $h > 0$ e si applica ancora il Lemma 2.1 con $x_0 = x - h$, $x_1 = x$, $x_2 = x + h$, si arriva a

$$-\frac{f(x - h) - f(x)}{h} \leq \frac{f(x + h) - f(x)}{h},$$

da cui, passando al limite per $h \to 0$, il fatto che

$$f'(x^-) \leq f'(x^+).$$

Infine se si considera $x < y$ e $h > 0$ tale che $x < x + h < y < y + h \in I$, iterando la stima di monotonia (2.1), si trova che

$$\frac{f(x + h) - f(x)}{h} \leq \frac{f(y) - f(x + h)}{y - x - h} \leq \frac{f(y + h) - f(y)}{h},$$

da cui passando al limite per $h \to 0$ nel primo e nel terzo termine, il fatto che

$$f(x^+) \leq f'(y^+).$$

In modo analogo si dimostra la monotonia di $f'(x^-)$.

$\square$

Osserviamo che come corollario del risultato precedente, se

$$\alpha = \inf I, \qquad \beta = \sup I,$$

allora

$$\exists \lim_{x \to \alpha^+} f'(x^-) = \lim_{x \to \alpha^+} f'(x^+), \quad \exists \lim_{x \to \beta^-} f'(x^+) = \lim_{x \to \beta^-} f'(x^-);$$

ha quindi senso parlare di $f'(\alpha^+)$ e $f'(\beta^-)$ se si accettano come valori anche $-\infty$ per la prima e $+\infty$ per la seconda.

Osserviamo anche che dal corollario precedente e dalla monotonia dei rapporti incrementali segue che per ogni $x_0 \in I$ e per ogni $x \geq x_0$

$$f(x) \geq f(x_0) + f'(x_0^+)(x - x_0).$$

Analogamente, per ogni $x \leq x_0$

$$f(x) \geq f(x_0) + f'(x_0^-)(x - x_0).$$

Piú in generale quindi si ha che per ogni $m \in [f'(x_0^-), f'(x_0^+)]$ e per ogni $x \in I$

$$f(x) \geq f(x_0) + m(x - x_0).$$

L'intervallo

$$\partial f(x_0) = [f'(x_0^-), f'(x_0^+)]$$

prende il nome di sottodifferenziale della funzione f e si riduce ad un solo punto se e solo se f è derivabile in x_0.

Passiamo ora alla dimostrazione di alcune proprietà di regolarità delle funzioni convesse; iniziamo con la continuità.

Proposizione 2.1. *Le funzioni convesse nell'intervallo I contenuto in $\mathbb{R}$ sono continue nei punti interni di tale intervallo.*

Dimostrazione. Sia $\bar{x}$ un punto interno a I. Dal teorema precedente si sa che esistono le derivate laterali. Prendiamo ad esempio quella destra: esiste quindi

$$\lim_{x \to \bar{x}^+} \frac{f(x) - f(\bar{x})}{x - \bar{x}} = f'(\bar{x}^+) \in \mathbb{R}.$$

Ciò equivale a dire che

$$f(x) - f(\bar{x}) = f'(\bar{x}^+)(x - \bar{x}) + q(x)(x - \bar{x})$$

con

$$\lim_{x \to \bar{x}^+} q(x) = 0.$$

Passando al limite per $x \to \bar{x}^+$ si ha

$$\exists \lim_{x \to \bar{x}^+} f(x) = f(\bar{x}).$$

In modo del tutto analogo, usando la derivata sinistra, si dimostra che

$$\exists \lim_{x \to \bar{x}^-} f(x) = f(\bar{x}),$$

e quindi la continutà di f in $\bar{x}$. □

Osservazione 2.1. 1. *A priori nulla si può dire sul comportamento di f nei punti estremi di I, cioè nei punti*

$$\alpha = \inf I, \qquad \beta = \sup I,$$

neanche nei casi in cui eventualmente f sia definita nei punti α e β. Basta considerare i seguenti esempi; $f(x) = \tan x$ per $x \in (-\pi/2, \pi/2)$ dove f non è definita negli estremi, così come la funzione $f(x) = -\sqrt{x}$ con $x \in [0, 1]$ che è definita in $x = 0$ ma

$$f'(0^+) = -\infty,$$

ed infine la sua modificazione

$$f(x) = \begin{cases} 1 & \text{se } x = 0 \\ -\sqrt{x} & \text{se } x \in (0, 1] \end{cases}$$

che è ancora convessa ma non continua in 0.

2. *Si osservi che il fatto di avere le derivate laterali non ci permette di dire che la funzione f è derivabile in tutti i punti interni a I. Un esempio può essere $f(x) = |x|$. Questa funzione infatti non è derivabile nel punto $x = 0$.*

Per quanto riguarda la derivabilità di f abbiamo però il seguente risultato.

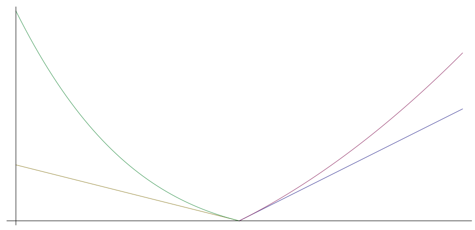


Figura 2.1: Non derivabilità in 0

Proposizione 2.2. *Le funzioni $f'(x^-)$ e $f'(x^+)$ hanno al più una quantità numerabile di punti di discontinuità. Cioè definito*

$$\mathcal{N}^{\pm} = \{x \text{ tali che } f'(x^{\pm}) \text{ è discontinua}\}$$

o equivalentemente

$$\mathcal{N} = \mathcal{N}^{+} \cup \mathcal{N}^{-} = \{x \text{ tali che } f \text{ non sia derivabile in } x\},$$

questo insieme ha cardinalità finita o al più numerabile.

Dimostrazione. Sappiamo per le proprietà prima elencate che dato $x < y$

$$f'(x^-) \leq f'(x^+) \leq f'(y^-) \leq f'(y^+)$$

Una maniera anaoga per descrivere l'insieme $\mathcal{N}$ è la seguente

$$\mathcal{N}^{\pm} = \{x \in I : f'(x^-) < f'(x^+)\}.$$

Ad ogni $x \in \mathcal{N}^{\pm}$ posso associare l'intervallo non vuoto $\mathcal{J}_x = (f'_-(x), f'_+(x))$. Questi intervalli sono a due a due disgiunti e ogni intervallo contiene un razionale. Dunque

$$\{x : \mathcal{J}_x \neq \emptyset\} = \mathcal{N}$$

è al più numerabile. □

Osservazione 2.2. *Poichè f è derivabile in ogni punto $x \in I \backslash (\mathcal{N}^{+} \cup \mathcal{N}^{-})$ la funzione f è derivabile quasi ovunque nel senso della misura di Lebesgue.*

Proposizione 2.3. *La funzione derivata destra $x \mapsto f'(x^+)$ è continua da destra mentre la derivata sinistra $x \mapsto f'(x^-)$ è continua da sinistra. Inoltre tali funzioni sono localmente limitate e localmente a variazione limitata.*

Prima di incominciare con la dimostrazione della proposizione ricordiamo la definizione di *funzione a variazione limitata* e un importante risultato.

Definizione 2.1. Sia $f\colon [a,b] \to \mathbb{R}$; data una partizione $\pi : a = t_0 < t_1 < \cdots < t_n = b$ di $[a,b]$ definiamo *variazione totale* la seguente

$$Var(f,[a,b]) = \sup_{\pi} \left\{ \sum_{i=1}^{n} |f(t_{i+1}) - f(t_i)| \right\}.$$

Diciamo poi che f è una funzione a vaiazione limitata in $[a,b]$ e scriviamo $f \in BV([a,b])$ se

$$Var(f,[a,b]) < +\infty.$$

Possiamo prendere come esempio

$$f(x) = \begin{cases} 0 & \text{se } x = 0 \\ x^2 \sin(\frac{1}{x}) & \text{se } x \neq 0 \end{cases}$$

funzione a variazione limitata in $[0,1]$. Mentre non è a variazione limitata la seguente funzione

$$f(x) = \begin{cases} 0 & \text{se } x = 0 \\ x \sin \frac{1}{x} & \text{se } x \neq 0. \end{cases}$$

Osservazione 2.3. *Se f è monotona in $[a,b]$ allora $f \in BV([a,b])$.*
Infatti

$$Var(f,[a,b]) = \sup \sum_{i=1}^{n} |f(t_{i+1}) - f(t_i)| = |f(b) - f(a)| < +\infty$$

Dimostrazione. (Proposizione 2.3) Iniziamo a dimostrare la prima parte della proposizione. Se la derivata destra non è continua a destra in un punto $x \in I$, allora per la monotonia della funzione derivata destra e dei rapporti incrementali di f segue che esiste $\varepsilon > 0$ tale che per ogni $y > x$ e $h > 0$ tali che $y + h \in I$,

$$f'(x^+) + \varepsilon \leq f'(y^+) \leq \frac{f(y+h) - f(y)}{h}.$$

Dato che la funzione f è anche continua, allora passando al limite per $y \to x^+$ ricaviamo che

$$f'(x^+) + \varepsilon \leq \frac{f(x+h) - f(x)}{h}.$$

A questo punto si ottiene l'assurdo passando al limite $h \to 0^+$. Allo stesso modo si ragiona con la derivata sinistra.

Dimostriamo ora la seconda parte. Poichè le funzioni $f'(x^-)$ e $f'(x^+)$ sono monotone allora hanno limite e dunque localmente limitate in I. Cioè per ogni $[a,b] \subseteq I$

$$\sup_{[a,b]} |f(x^\pm)| \le c_{a,b}$$

dove $c_{a,b}$ è una costante dipendente da a, b ed è così definita

$$c_{a,b} = \max\{|f(a^+)|, |f(b^-)|\}$$

Poichè le funzioni $f'(x^\pm)$ sono monotone crescenti si ha, per quanto detto nell'Osservazione 2.3, che sono a variazione limitata. $\square$

Proposizione 2.4. *Le funzioni convesse sono localmente Lipschitziane nei punti interni del loro dominio I.*

Dimostrazione. Sia x_0 un punto interno all'intervallo I. Poichè è interno esiste un $\delta > 0$ tale che $x_0 \pm \delta \in I$

$$\begin{aligned} f'((x_0-\delta)^+) \le & \frac{f(x_0-\delta)-f(x_0)}{(x_0-\delta)-x_0} \le \frac{f(x)-f(x_0)}{x-x_0} \\ \le & \frac{f(x_0+\delta)-f(x_0)}{(x_0+\delta)-(x_0)} \le f'((x_0+\delta)^-) \end{aligned}$$

per ogni $x \in (x_0-\delta, x_0+\delta) \setminus \{x_0\}$ e dunque :

$$|f(x)-f(x_0)| \le K|x-x_0|$$

con

$$K = \max\left\{|f'((x_0-\delta)^+)|, |f'((x_0+\delta)^-)|\right\},$$

dunque è localmente Lipschitziana. $\square$

Il problema della ricerca dei massimi e minimi per funzioni convesse uni–dimensionali può essere particolareggiato, dal Teorema 1.4 e 1.3 dal seguente risultato. In particolare, si nota come i problemi della ricerca dei massimi di una funzione convessa uni–dimensionale sia un problema banale; in effetti, la parte interessante della seguente proposizione non è tanto l'esistenza del massimo, che è già una conseguenza del Teorema di Weierstrass essendo una funzione convessa continua in ogni intervallo chiuso e limitato contenuto in I, ma più che altro la dislocazione del punto di massimo.

Proposizione 2.5. *Sia $f : I \to \mathbb{R}$ una funzione convessa e $[a,b] \subset I$; allora f ammette massimo in $[a,b]$ e*

$$\max_{[a,b]} f = \max\{f(a), f(b)\}.$$

Dimostrazione. Sia $M = \max\{f(a), f(b)\}$. Per la prima definizione di funzione convessa sappiamo che preso un $t \in [0, 1]$ si ha

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

e cioè nel nostro caso

$$f(ta + (1-t)b) \leq tf(a) + (1-t)f(b) \leq tM + (1-t)M = M.$$

□

Osservazione 2.4.

1. *Si osservi che una funzione convessa definita in un intervallo chiuso e limitato ha sempre massimo ma potrebbe non avere minimo. Per esempio la funzione*

$$f(x) = \begin{cases} 2 & \textit{per } x = 0 \\ x^2 & \textit{per } (0, 1]. \end{cases}$$

2. *La ricerca del massimo sul generico intervallo I dipende dalla struttura dell'intervallo I e dal comportamento di f negli estremi di tale intervallo; posto*

$$\alpha = \inf I, \qquad \beta = \sup I,$$

per poter parlare di massimo si deve avere necessariamente almeno uno tra α e β finito. Se infatti $\alpha = -\infty$ e $\beta = +\infty$ allora f non ammette massimo. Nel caso in cui ad esempio α sia finito, si deve andare a vedere cosa succede in α; se $\alpha \in I$ allora bisogna considerare il valore $f(\alpha)$ come candidato massimo, mentre se $\alpha \notin I$ bisogna studiare

$$\lim_{x \to \alpha^+} f(x).$$

Tale limite si può dimostrare che esiste e se è finito, significa che si può estendere f con continuità in α e si torna al punto precedente.

2.1 Funzioni convesse di classe C^1 e C^2

Abbiamo fino ad ora visto le proprietà generali di una funzione convessa. Supponiamo ora che oltre ad essere convessa la nostra funzione sia pure derivabile in un intervallo $[a, b]$. Allora vale il seguente

Teorema 2.2. *Sia $f \in C^1$. Allora f è convessa se e solo se*

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0) \textit{ per ogni } x, x_0 \in I. \tag{2.2}$$

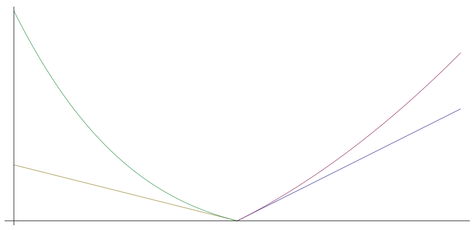


Figura 2.2: Grafico sopra le tangenti

Osservazione 2.5. *Con questa proposizione si indica la proprietà delle funzioni convesse di avere il grafico sopra la retta tangente.*

Dimostrazione. Abbiamo visto, in una dimostrazione precedente, che

$$\frac{f(x_1)-f(x_0)}{x_1-x_0} \leq \frac{f(x_2)-f(x_1)}{x_2-x_1}.$$

Passando ora al limite per $x_1 \to x_0^+$ si ha

$$f'(x_0)(x_2-x_0) \leq f(x_2)-f(x_0).$$

Se ora poniamo $x_2 = x$ si ottiene

$$f(x) \geq f(x_0)+f'(x_0)(x-x_0).$$

Se prendiamo invece la disuguaglianza

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \geq \frac{f(x_2)-f(x_0)}{x_2-x_0}$$

e facciamo il limite per $x_1 \to x_2^-$ e ponendo poi $x_2 = x$ si ottiene che

$$\begin{aligned} f'(x_2)(x_2-x_0) &\geq f(x_2)-f(x_0) \\ f'(x_0)(x_0-x) &\geq f(x_0)-f(x) \end{aligned} \tag{2.3}$$

e da qui segue la tesi. Vediamo ora la seconda implicazione. Sia $x_1 < x_2$ e $x_0 \in (x_1, x_2)$ dove x_1 e x_2 sono punti interni ad I. Sia $t(x)$ la retta tangente al grafico nel punto x_0 e cioè

$$t(x) = f(x_0)+f'(x_0)(x-x_0).$$

Sia $s(x)$ la secante passante per i due estremi x_1, x_2 così definita

$$s(x) = f(x_1)+\frac{f(x_2)-f(x_1)}{x_2-x_1}(x-x_1).$$

Allora dalla costruzione si nota che $s(x_1) = f(x_1) \geq t(x_1)$ poichè abbiamo supposto che
$$f(x) \geq f(x_0) + f'(x_0)(x - x_0).$$
Ripetiamo lo stesso ragionamento con x_2 e si otterrà $s(x_2) \geq t(x_2)$ Ricordando ora che $s(x)$ e $t(x)$ sono delle rette e utilizzando i risultati appena ottenuti si può concludere che $s(x) \geq t(x)$ per ogni $x \in [x_1, x_2]$. In particolare varrà per x_0 e dunque
$$s(x_0) \geq t(x_0 = f(x_0),$$
dunque $f(x_0) \leq s(x_0)$, cioè il grafico della funzione in x_0 si trova sempre sotto la secante. Poichè abbiamo scelto un x_0 arbitrario la proprietà appena enunciata vale per ogni punto interno a (x_1, x_2) e dunque la proposizione è dimostrata. □

Proposizione 2.6. *La funzione $f\colon I \to \mathbb{R}$ è convessa se e solo se $f'(x)$ è monotona crescente*

Dimostrazione. Poichè la nostra funzione è continua in I e derivabile nello stesso intervallo si può applicare il teorema di Lagrange : esiste un $\xi \in (x_0, x)$ con $x_0 < x$ tale che
$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi) \geq f'(x_0)$$
Dove nell'ultimo passaggio abbiamo sfruttato il fatto che la funzione f' è crescente. Da questa si ricava che
$$f(x) \geq f(x_0) + f'(x_0)(x - x_0)$$
ma abbiamo appena dimostrato che la formula ottenuta equivale alla definizione di convessità. Assumiamo ora che $x < x_0$. Procedendo come prima si ottiene lo stesso risultato, dimostrando così la prima implicazione.

Supponiamo ora che la funzione sia convessa. Dobbiamo dimostrare che la funzione f' è crescente. Come prima per dimostrarlo utilizziamo la proposizione precedente. Consideriamo due punti $x_1\, x_2$ tali che $x_1 < x_2$. Applichiamo la proprietà ed otteniamo
$$f(x_1) \geq f(x_2) + f'(x_2)(x_1 - x_2)$$
$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1)$$
Ora sommando membro a membro si ottiene
$$0 \geq (x_2 - x_1)(f'(x_1) - f'(x_2))$$

Poichè $x_1 < x_2$ si ha

$$f'(x_1) < f'(x_2)$$

□

Teorema 2.3. *Sia $f : I \to \mathbb{R}$ e derivabile due volte in I. Allora f è convessa se e solo se $f''(x) \geq 0$ per ogni $x \in I$.*

Dimostrazione. Grazie alla proposizione precedente sappiamo che se f è una funzione convessa e derivabile allora f' è crescente. D'altra parte f' è crescente se e solo se la sua derivata è maggiore di zero e cioè $(f'(x))' \geq 0$ o equivalentemente $f''(x) \geq 0$. □

Capitolo 3

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

Per sviluppare questo capitolo sono stati introdotti concetti rivisitati di varie opere, si veda la bibliografia. [6], [8] , [5], [7], [3]. Prima di iniziare vogliamo richiamare alcune notazioni, risultati e definizioni che verranno poi utilizzate nel corso del capitolo. Iniziamo dunque con la seguente

Definizione 3.1 (Derivata parziale)**.** Sia $f\colon A \to \mathbb{R}$ con $A \subseteq \mathbb{R}^n$ insieme aperto, una data funzione e sia $x_0 \in A$. La quantità, se esiste

$$\frac{\partial f}{\partial x_i}(x_0) = \lim_{h\to 0} \frac{f(x_0 + he_i) - f(x_0)}{h} \tag{3.1}$$

dove con e_i è il vettore $(0, 0, \dots, 1, 0, \dots, 0)$ con 1 nella i–ma posizione, viene chiamata *derivata parziale di f rispetto a x_i* e viene indicata con f_x.
Se invece di considerare e_i considerassimo un versore v qualunque appartenente a $\mathbb{R}^n$ avremmo una relazione simile a quella precedente

$$\frac{\partial f}{\partial v}(x_0) = \lim_{h\to 0} \frac{f(x_0 + hv) - f(x_0)}{h}. \tag{3.2}$$

Questa viene chiamata *derivata parziale di f nella direzione v*. Se la funzione $\frac{\partial f}{\partial x_i}$ è ulteriormente derivabile in x_0, allora possiamo definire le *derivate seconde* che indicheremo con

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

Una volta introdotte le derivate parziali, possiamo dare le seguenti definizioni.

Definizione 3.2 (Gradiente)**.** Il vettore che ha per componenti i–esima le derivate parziali di f calcolate in x_0 viene chiamato *gradiente* di f in x_0 e viene indicato con $\nabla f(x_0)$.

Definizione 3.3 (Funzione differenziabile e differenziale)**.** Dato $A \subset \mathbb{R}^n$ insieme aperto, una funzione $f\colon A \to \mathbb{R}$ si dice *differenziabile* in $x_0 \in A$ se esiste una funzione lineare $L\colon \mathbb{R}^n \to \mathbb{R}$ tale che

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} = 0. \tag{3.3}$$

La funzione lineare L prende il nome di *differenziale* di f in x_0 e viene denotato con $d_{x_0}f$ o con $df(x_0)$.

Ricordando che il differenziale, se esiste, è unico, si vede abbastanza facilmente che esso è rappresentabile dal vettore gradiente di f, cioè può essere scritto nella forma

$$L(x) = \nabla f(x_0) \cdot x. \tag{3.4}$$

Osservazione 3.1. *La nozione di differenziale permette di definire in modo appropriato il piano tangente al grafico di f; esso sarà dato dal piano appartenente ad $\mathbb{R}^{n+1}$ parallelo al grafico di L e passante per il punto $(x_0, f(x_0))$, e che quindi sarà descritto dall'equazione*

$$x_{n+1} = f(x_0) + \nabla f(x_0) \cdot (x - x_0).$$

Definizione 3.4 (Piano tangente al grafico di f)**.** Definiamo quindi *piano tangente al grafico di f*, e lo indichiamo con $\mathcal{T}_{x_0}\Gamma(f)$ il piano contenuto in $\mathbb{R}^{n+1}$ descritto dall'equazione

$$\mathcal{T}_{x_0}\Gamma(f) = \left\{ x \in \mathbb{R}^{n+1} : x_{n+1} = f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0,i}) \right\}. \tag{3.5}$$

Concludiamo questa sezione introduttiva con la definizione di matrice Hessiana.

Definizione 3.5 (Matrice Hessiana)**.** Supponiamo che la funzione $f\colon A \to \mathbb{R}$ sia derivabile due volte nel punto x_0; si definisce matrice Hessiana di f in x_0 la matrice così definita:

$$\mathcal{H}f(x_0) = \begin{pmatrix} f_{xx}(x_0) & f_{xy}(x_0) \\ f_{yx}(x_0) & f_{yy}(x_0) \end{pmatrix}.$$

Questo capitolo vuole mettere in evidenza molte delle proprietà già prese in esame, delle funzioni convesse definite in $\mathbb{R}^n$.

Un importante teorema che ci ripropone un risultato già visto nel caso uni–dimensionale è il seguente.

Teorema 3.1. *Sia $\Omega \in \mathbb{R}^n$ un insieme convesso e aperto e sia $f\colon \Omega \to \mathbb{R}$ una funzione convessa. Allora f è differenziabile quasi ovunque rispetto alla misura di Lebesgue.*

Dimostrazione. Introduciamo le seguenti notazioni; per $v \in \mathbb{S}^{n-1}$, denotiamo con

$$\Omega_v = \{y \in v^\perp : \exists t \text{ tale che } y + tv \in \Omega\}$$

la proiezione di Ω sullo spazio $v^\perp$. Inoltre, per ogni $y \in \Omega_v$, denotiamo con

$$\Omega_v^y = \{t \in \mathbb{R} : y + tv \in \Omega\};$$

la convessità di Ω implica che Ω_v^y è un interavallo. Inoltre, la precedente costruzione divide Ω in una componente ortogonale a v ed in una parallela a v e, grazie al Teorema di Fubini–Tonelli

$$\mathcal{L}^n(\Omega) = \int_{\Omega_v} \mathcal{L}^1(\Omega_v^y)dy.$$

Dato che le restrizioni uni–dimensionali

$$t \mapsto f(y + tv), \qquad t \in \Omega_v^y$$

sono funzioni convesse per ogni $v \in \mathbb{S}^{n-1}$, e per ogni $y \in \Omega_v$, se ne deduce che la derivata parziale in direzione v esiste per quasi ogni $t \in \Omega_v^y$. Se denotiamo con N_v^y l'insieme dove, fissato $y \in \Omega_v$, non esiste $\partial f/\partial v$ e

$$N_v = \{x = y + tv \in \Omega : y \in \Omega_v, t \in N_v^y\},$$

dal Teorema di Fubini–Tonelli avremo che

$$\mathcal{L}^n(N_v) = \int_{\Omega_v} \mathcal{L}^1(N_v^y)dy = 0,$$

quindi la derivata parziale in direzione v esiste per quasi ogni $x \in \Omega$.

Se applichiamo tale ragionamento con $v = e_i$, allora tutte le derivate parziali esistono per quasi ogni $x \in \Omega$, e più precisamente per ogni $x \in \Omega \setminus N$, con

$$N = \bigcup_{i=1}^n N_{e_i}.$$

Fissiamo quindi $x_0 \in \Omega \setminus N$ e sia $x \in \Omega$ qualsiasi; possiamo scrivere

$$x = \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \lambda_i' \mu \frac{\lambda_i}{|\lambda_i|} v_i$$

con

$$\mu = \sum_{i=1}^{n} |\lambda_i| \qquad \lambda_i' = \frac{|\lambda_i|}{\mu} \geq 0.$$

Allora con le notazioni precedenti si ha che

$$\sum_{i=1}^{n} \lambda_i' = 1,$$

potendo quindi usare applicare la disuguaglianza (1.5). Definiamo quindi la funzione convessa

$$F(x) := f(x + x_0) - f(x_0)$$

definita per $x \in \Omega - x_0$; da (1.5) otteniamo quindi che

$$F(x) \leq \sum_{i=1}^{n} \lambda_i' F\left(\mu \frac{\lambda_i}{|\lambda_i|} v_i\right). \tag{3.6}$$

Denotiamo con $L(x) := \nabla f(x_0) \cdot x$; seguendo la definizione, per dimostrare la differenziabilità di f in x_0, dobbiamo dimostrare che

$$\lim_{x \to 0} \frac{f(x + x_0) - f(x_0) - L(x)}{\|x\|} = 0.$$

Sottraiamo dunque ad entrambi i membri della (3.6) la funzione $L(x)$ ottenendo

$$F(x) - L(x) \leq \sum_{i=1}^{n} \left[\lambda_i' F\left(\mu \frac{\lambda_i}{|\lambda_i|} v_i\right) - \lambda_i L(v_i)\right].$$

Raccogliendo a secondo membro la λ_i' e ricordando che L è una funzione lineare possiamo riscrivere la disuguaglianza nella seguente forma

$$F(x) - L(x) \leq \sum_{i=1}^{n} \lambda_i' \left[F\left(\mu \frac{\lambda_i}{|\lambda_i|} v_i\right) - L\left(\mu \frac{\lambda_i}{|\lambda_i|} v_i\right)\right] = o(\|x\|). \tag{3.7}$$

Vediamo inoltre che vale anche la seguente disuguaglianza

$$\frac{F(x) + F(-x)}{2} \geq F(0) = 0$$

da cui si deduce

$$F(x) \geq -F(-x).$$

D'altra parte riconsideriamo la disuguaglianza (3.7) sostituendo $-x$ a x e otteniamo così

$$F(-x) - L(-x) \leq \sum_{i=1}^{n} \lambda_i' \left[F\left(-\mu\frac{\lambda_i}{|\lambda_i|}v_i\right) - L\left(-\mu\frac{\lambda_i}{|\lambda|}v_i\right)\right]. \tag{3.8}$$

Mettendo insieme le due relazioni (3.7) e (3.8) otteniamo

$$\begin{aligned} F(x) &\geq -F(-x) \geq L(x) - o(\|x\|) \\ F(-x) &\leq -L(x) + o(\|x\|) \end{aligned}$$

da cui

$$|F(x) - L(x)| \leq o(\|x\|)$$

e dunque la tesi. □

La differenziabilità quasi ovunque implica ovviamente la continuità quasi ovunque di una funzione convessa; in realtà si può dimostrare che una funzione convessa è continua in tutti i punti interni di Ω e anche localmente Lipschitziana. Non metteremo qui i dettagli della dimostrazione, ma sostanzialmente il risultato si può derivare dal caso uni–dimensionale; infatti, se si fissano due punti $x, y \in \Omega$ con $x \neq y$, allora considerando la funzione convessa uni–dimensionale

$$g_{x,y}(t) = f\left(x + t\frac{y-x}{\|y-x\|}\right),$$

dalla locale Lipschitzianità di $g_{x,y}$ e dal fatto che l'intervallo $[0, \|y-x\|]$ è contenuto nel dominio di $g_{x,y}$, si ricava che

$$|f(x) - f(y)| = |g_{x,y}(0) - g_{x,y}(\|y-x\|)| \leq M\|y-x\|.$$

Dalla locale Lipschitzianità di f segue poi la sua continuità nei punti interni di Ω.

3.1 Proprietà per funzioni di classe C^1, C^2

Dopo aver dimostrato che una funzione convessa è quasi ovunque differenziabile, passiamo a studiare le proprietà delle funzioni con qualche ipotesi aggiuntiva sulla regolartà. Inizieremo, come abbiamo fatto nel capitolo precedente, con il supporre $f \in C^1(\Omega)$ per poi concludere con il caso $f \in C^2(\Omega)$.

Proposizione 3.1. *Sia $f \in C^1(\Omega)$ con Ω insieme aperto e convesso contenuro in $\mathbb{R}^n$: allora f è convessa se e solo se per ogni $x_0 \in \Omega$*

$$f(x) \geq f(x_0) + \nabla f(x_0)(x - x_0). \tag{3.9}$$

Dimostrazione. Dimostriamo per prima cosa che se f è convessa allora vale la (3.9). Fissiamo $x_0 \in \Omega$ e sia $x \in \Omega$ arbitrario: definiamo quindi il punto $x_t \in \Omega$

$$x_t = x_0 + t(x - x_0) = tx + (1 - t)x_0, \qquad t \in [0, 1].$$

Per la convessità di f otteniamo che

$$f(x_t) \leq tf(x) + (1 - t)f(x_0) = f(x_0) + t(f(x) - f(x_0)),$$

che può essere riscritta nella forma

$$\frac{f(x_0 + t(x - x_0)) - f(x_0)}{t} \leq f(x) - f(x_0).$$

Quindi passando al limite per $t \to 0$, dato che

$$\lim_{t \to 0} \frac{f(x_0 + t(x - x_0)) - f(x_0)}{t} = \nabla f(x_0) \cdot (x - x_0),$$

e quindi

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0).$$

Concludiamo ora dimostrando la seconda implicazione. Fissiamo x, $y \in \Omega$ e poniamo

$$x_t = tx + (1 - t)y = y + t(x - y) \in \Omega$$

sostituendo nella (3.9) ad x_0 la x_t e applicando la formula ad x e ad y si ottiene

$$f(x) \geq f(x_t) + \nabla f(x_t)(x - x_t) \tag{3.10}$$

$$f(y) \geq f(x_t) + \nabla f(x_t)(y - x_t). \tag{3.11}$$

Dato poi che

$$\begin{aligned} x - x_t &= x - tx - (1 - t)y = (1 - t)(x - y) \\ y - x_t &= y - tx - (1 - t)y = t(y - x) \end{aligned}$$

sostituendo nelle equazioni (3.10), (3.11) si ottiene

$$\begin{aligned} f(x) &\geq f(x_t) + (1 - t)\nabla f(x_t)(x - y) \\ f(y) &\geq f(x_t) + t\nabla f(x_t)(y - x). \end{aligned}$$

Facendo ora una combinazione si ottiene

$$tf(x) + (1 - t)f(y) \geq f(x_t) = f(tx + (1 - t)y)$$

e quindi la proposizione è dimostrata. □

Osservazione 3.2. *Per le funzioni convesse non differenziabili viene introdotto il sottodifferenziale*

$$\partial f(x_0) = \{m \in \mathbb{R}^n : f(x) \geq f(x_0) + m \cdot (x - x_0)\};$$

se f è differenziabile, abbiamo appena visto che

$$\partial f(x_0) = \{\nabla f(x_0)\},$$

mentre in generale, grazie a quanto abbiamo visto nel caso uni–dimensionale,

$$\partial f(x_0) = \prod_{i=1}^{n} \left[\frac{\partial f}{\partial x_i}(x_0^-), \frac{\partial f}{\partial x_i}(x_0^+) \right],$$

dove abbiamo denotato con

$$\frac{\partial f}{\partial x_i}(x_0^{\pm}) = \lim_{h \to 0^{\pm}} \frac{f(x_0 + he_i) - f(x_0)}{h}.$$

Un importante corollario della precedente proposizione ha applicazione nel problema della ricerca dei punti di minimo per funzioni convesse. Tale risultato mostra la rilevanza della nozione di convessità e ci mette di fronte alla differenza sotanziale tra le funzioni convesse e le funzioni differenziabili in generale. Infatti, mentre solitamente la condizione per x_0 di essere punto stazionario, cioè $\nabla f(x_0) = 0$ è condizione necessaria per la minimalità o massimalità, nel caso di funzioni convesse diventa anche condizione sufficiente.

Proposizione 3.2. *Sia Ω un insieme aperto e convesso e sia $f : \Omega \to \mathbb{R}$ una funzione convessa e differenziabile. Allora ogni punto stazionario di f in Ω è anche punto di minimo globale per f. Se poi la funzione è strettamente convessa, allora f ha al piú un punto stazionario che è anche punto di minimo globale stretto.*

Dimostrazione. Sia f convessa e x_0 punto stazionario. Poichè f è anche differenziabile abbiamo visto grazie alla proposizione (3.9) che vale

$$f(x) \geq f(x_0) + \nabla f(x_0)(x - x_0) \geq f(x_0) \quad \text{per ogni } x \in \Omega$$

dove nell'ultimo passaggio si è sfruttato il fatto che x_0 è punto stazionario. Allora x_0 è proprio il punto di minimo globale.

Se poi la funzione è strettamente convessa convessa, avremo che per $x \neq x_0$

$$f(x) > f(x_0) + \nabla f(x_0)(x - x_0) = f(x_0),$$

cioè $f(x_0)$ è punto di minimo globale stretto e non possono quindi esistere altri punti di minimo. □

Esempio 3.1. Prendiamo ad esempio la funzione

$$f\colon \mathbb{R}^2 \to \mathbb{R} \qquad f(x,y) = x^2$$

Si può vedere, come noteremo tra poco, che la matrice Hessiana è semidefinita positiva. Da qui la proposizione. Infatti con lo studio della matrice Hessiana arriviamo a calcolare il determinante che è uguale a zero e grazie al risultato precedente $(0, y)$ è punto di minimo.

Vediamo ora condizioni che caratterizzino la convessità e che coinvolgono le derivate successive di f. Richiediamo quindi ora che la funzione sia di classe C^2 in modo da ottenere condizioni sulla matrice Hessiana. Il seguente teorema non è altro che una generalizzazione del Teorema 2.3 proposto nel capitolo precedente.

Teorema 3.2. *Sia Ω un insieme aperto e convesso e sia $f\colon \Omega \to \mathbb{R}$ una funzione tale che $f \in C^2(\Omega)$. Allora f è una funzione convessa se e solo se la matrice Hessiana $\mathcal{H}f(x_0)$ è semidefinita positiva per ogni $x_0 \in \Omega$.*

Dimostrazione. Supponiamo che f sia convessa. Siano $x_0 \in \Omega$, $v \in \mathbb{S}^{n-1}$ e $\alpha \in \mathbb{R}$ tale che $x_0 + \alpha v \in \Omega$. Applicando la formula di Taylor al secondo ordine si ottiene

$$\begin{aligned} f(x_0 + \alpha v) &= f(x_0) + \alpha \nabla f(x_0) \cdot v + \frac{1}{2}\alpha^2 \mathcal{H}f(x_0)v \cdot v + o(\alpha^2) \\ &\geq f(x_0) + \alpha \nabla f(x_0) \cdot v \end{aligned}$$

con $o(\alpha^2)$ un infinitesimo tale che

$$\lim_{\alpha \to 0} \frac{o(\alpha^2)}{\alpha^2} = 0$$

e dove nell'ultimo passaggio si è usata la relazione (3.9) dove $x - x_0 = v$. Ora poichè questa disequazione è vera devo avere che

$$\begin{aligned} \frac{1}{2}\mathcal{H}f(x_0)v \cdot v + \frac{o(\alpha^2)}{\alpha^2} \geq 0 \\ \mathcal{H}f(x_0)v \cdot v \geq 0 \end{aligned}$$

e dunque la matrice Hessiana è semidefinita positiva.

Dimostriamo ora la seconda parte del teorema. Supponiamo che la matrice Hessiana sia semidefinita positiva e riscriviamo la formula di Taylor al secondo ordine e fissiamo $x_0 \in \Omega$ ottenendo per ogni $x \in \Omega$

$$\begin{aligned} f(x) =& f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}\mathcal{H}f(\xi)(x - x_0)(x - x_0) \\ \geq& f(x_0) + \nabla f(x_0)(x - x_0). \end{aligned}$$

Abbiamo così concluso la dimostrazione del teorema. $\square$

Osservazione 3.3. *Abbiamo visto che per avere la convessità dobbiamo supporre che la matrice Hessiana sia semidefinita positiva.*
Condizione sufficiente per avere la stretta convessità è invece assumere la matrice Hessiana definita positiva. Per poter avere la condizione necessaria bisogna assumere che ci sia quasi ovunque la stretta definizione positiva. Basti pensare a x^4 *che è strettamente convessa ma in* $x_0 = 0$ *la derivata seconda si annulla.*

Capitolo 4

Successioni di funzioni convesse

Dopo aver elencato le proprietà delle funzioni convesse, in questo capitolo ci vogliamo occupare di famiglie di funzioni convesse.

Prima di incominciare ricordiamo alcune definizioni che ci saranno utili nel corso del capitolo ([6]).

Definizione 4.1 (Convergenza puntuale e uniforme)**.** Sia $(f_n)_n \colon I \to \mathbb{R}$ una successione di funzioni e sia $f \colon I \to \mathbb{R}$ una data funzione. Allora diremo che f_n *converge puntualmente* a f nell'insieme I se per ogni $x \in I$ vale

$$\lim_{n\to\infty} f_n(x) = f(x), \tag{4.1}$$

cioè se per ogni $x \in I$ e per ogni $\varepsilon > 0$

$$\exists n_{\varepsilon,x} \quad \text{tale che} \quad \forall n > n_{\varepsilon,x} \quad \text{vale} \quad |f_n(x) - f(x)| < \varepsilon. \tag{4.2}$$

Diremo che f_n *converge uniformemente* ad f in I se

$$\forall\varepsilon, \exists n_\varepsilon \quad \text{tale che} \quad \forall n > n_\varepsilon \quad \text{risulta} \quad |f_n(x) - f(x)| < \varepsilon \forall x \in I, \tag{4.3}$$

cioè se, posto

$$a_n = \sup_{x\in I} |f_n(x) - f(x)|$$

si ha

$$\lim_{n\to\infty} a_n = 0. \tag{4.4}$$

Vogliamo confrontare le convergenze puntuale e uniforme con altri tipi di convergenze. Per questo, diamo la seguente definizione.

Definizione 4.2 (Distanza di Hausdorff)**.** Sia $A \subseteq \mathbb{R}^n$ denoto con $(A)_t$ il seguente insieme

$$(A)_t = \{y \in \mathbb{R}^n \quad \text{tale che} \quad dist(y, A) < t\}. \tag{4.5}$$

Dati $A, B \subseteq \mathbb{R}^n$ si chiama *distanza di Hausdorff* la seguente funzione d'insieme

$$dist_H(A, B) = \inf\{t > 0 \quad \text{tale che} \quad A \subseteq (B)_t, B \subseteq (A)_t\}. \tag{4.6}$$

Una volta introdotta la distanza tra due insiemi, possiamo dare la seguente definizione di convergenza per insiemi.

Definizione 4.3. Diremo che $(A_j)_j$ converge ad A se

$$\lim_{j \to \infty} dist_H(A_j, A) = 0. \tag{4.7}$$

Introduciamo anche la seguente nozione di convergenza di insiemi.

Definizione 4.4. Dati $(A_j)_j$ e $A \subseteq \mathbb{R}^n$, diremo che $(A_j)_j$ *converge secondo Kuratowski* se valgono le seguenti proprietà:

1. per ogni successione $(a_j)_j$, $a_j \in A_j$ tale che $a_j \to a$, allora $a \in A$;
2. per ogni $a \in A$, esiste una successione $(a_j)_j$, $a_j \in A_j$ con $a_j \to a$.

Mostriamo qui alcune relazioni e proprietà tra le nozioni appena introdotte.

Proposizione 4.1. *Valgono i seguenti fatti;*

1. *sugli insiemi chiusi si ha che*

$$dist_H(A, B) = 0 \iff A = B;$$

2. *la convergenza nella metrica di Hausdorff è equivalente sugli insiemi chiusi alla convergenza di Kuratowski.*

Dimostrazione. Per la dimostrazione del primo punto, se $A = B$ allora chiaramente la distanza di Hausdorff è nulla. Viceversa, supponiamo che $A \neq B$; allora possiamo supporre che esista $a \in A \setminus B$, altrimenti dovremmo assumere che esista $b \in B \setminus A$ e ragionare in modo analogo. Dato che B è chiuso, troveremmo che esiste $r > 0$ tale che

$$B(a, r) \cap B = \emptyset,$$

cioè la distanza di a da ogni punto $b \in B$ è maggiore o uguale ad r, e quindi

$$a \notin (B)_r,$$

da cui il fatto che

$$dist_H(A, B) \geq r.$$

Per la seconda parte della proposizione, notiamo che

$$\lim_{j \to +\infty} dist_H(A_j, A) = 0$$

se e solo se per ogni $h \in \mathbb{N}$ esiste $j_0 \in \mathbb{N}$ tali che per ogni $j \geq j_0$

$$A_j \subset (A)_{\frac{1}{h}}, \qquad A \subset (A_j)_{\frac{1}{h}}.$$

Quindi, dall'inclusione di sinistra si deduce che se $(a_j)_j$ è una successione con $a_j \in A_j$ che converge ad un elemento $a \in \mathbb{R}^n$, si deve avere

$$a \in (A)_{\frac{1}{h}}, \qquad \forall h \in \mathbb{N},$$

e quindi, dato che A è chiuso, $a \in A$. Equivalentemente, dall'inclusione di destra, se $a \in A$, allora si deve avere che

$$B\left(a, \frac{1}{h}\right) \cap A_j \neq \emptyset, \forall j \geq j_0,$$

e quindi esiste $a_j \in A_j$ con $\|a - a_j\| < \frac{1}{h}$. In questo modo si costruisce la successione che converge ad a. $\square$

Si noti che dalla dimostrazione della precedente proposizione si evince la seguente proprià della distanza di Hausdorff:

$$dist_H(A, B) = dist_H(\bar{A}, \bar{B}), \qquad \forall A, B \subset \mathbb{R}^n.$$

Ricordiamo ora la definizione degli spazi di Lebesgue.

Definizione 4.5. Sia $f \colon \Omega \to \mathbb{R}$ con $\Omega \subseteq \mathbb{R}^n$ misurabile e sia $p \geq 1$.
Si dice che $f \in \mathcal{L}^p(\Omega)$ se vale

$$\int_\Omega |f(x)|^p dx < \infty. \tag{4.8}$$

Diremo poi che$f \in \mathcal{L}^\infty(\Omega)$ se

$$\sup_{x \in \Omega} |f(x)| < +\infty.$$

Se si definisce inoltre la relazione di equivalenza $f \sim g$ se e solo se $f = g$ quasi ovunque, allora si definisce

$$L^p(\Omega) := \mathcal{L}^p(\Omega)/\sim, \qquad p \in [1, +\infty].$$

Definizione 4.6 (spazi L^p_{loc})**.** Sia $p \geq 1$: si definisce lo spazio delle funzioni localmente sommabili a potenza $p \in [1, +\infty]$ come

$$L^p_{loc}(\Omega) := \{f \colon \Omega \to \mathbb{R} \text{ misurabile:} \exists K \subset \Omega \text{ compatto con } f \in L^p(K)\}. \tag{4.9}$$

Possiamo dunque procedere con l'elencare alcuni importanti risultati.

Proposizione 4.2. *Consideriamo le successioni di funzione convesse $(f_n)_n \colon \Omega \to \mathbb{R}$ tale che f_n converge puntualmente a f in Ω. Allora la funzione f è convessa.*

Dimostrazione. Sappiamo per la definizione (4.2) che per ogni $x \in \Omega$

$$f(x) = \lim_{n \to +\infty} f_n(x).$$

Siano ora $x, y \in \Omega$ arbitrari e fissiamo $t \in [0, 1]$; useremo la convergenza puntuale nei tre punti $tx + (1-t)y$, x e y, cioè il fatto che

$$\begin{aligned} f(tx + (1-t)y) &= \lim_{n \to +\infty} f_n(tx + (1-t)y), \\ f(x) &= \lim_{n \to +\infty} f_n(x), \\ f(y) &= \lim_{n \to +\infty} f_n(y). \end{aligned}$$

Quindi, dato che f_n è convessa per ogni n,

$$f_n(tx + (1-t)y) \leq t f_n(x) + (1-t) f_n(y),$$

da cui, passando al limite, che esiste ad ambo i membri,

$$\begin{aligned} f(tx + (1-t)y) = \lim_{n \to +\infty} f_n(tx + (1-t)t) &\leq \lim_{n \to +\infty} (t f_n(x) + (1-t) f_n(y)) \\ &= t \lim_{n \to +\infty} f_n(x) + (1-t) \lim_{n \to +\infty} f_n(y) = t f(x) + (1-t) f(y). \end{aligned}$$

□

Grazie alla continuità delle funzioni convesse otteniamo che se f è convessa, allora per ogni compatto $K \subset \Omega$ f è limitata su tale insieme. Siccome poi la misura di un compatto è sempre finita, ne deduciamo che

$$\int_K |f(x)^p| dx \leq \max_{x \in K} |f(x)|^p \mathcal{L}^n(K) < +\infty,$$

dove abbiamo denotato con $\mathcal{L}^n$ la misura di Lebesgue in $\mathbb{R}^n$. Quindi, se $f : \Omega \to \mathbb{R}$ è una funzione convessa,

$$f \in L^p_{loc}(\Omega), \qquad \forall p \in [1, +\infty].$$

Dimostriamo ora che sulle funzioni convesse, basta la nozione più debole di convergenza per implicare in realtà, almeno sugli insiemi compatti, quella più forte. Iniziamo col caso uni–dimensionale.

Proposizione 4.3. *Siano $f_n, f \colon I \to \mathbb{R}$ funzioni convesse con $I \subseteq \mathbb{R}$ intervallo aperto tali che f_n converge in $L^1_{loc}(I)$ ad f; allora f_n converge uniformemente ad f sui compatti di I.*

Dimostrazione. Fissiamo i punti $a' < a < b' < b$ in modo che $[a', b'] \subseteq I$ sia chiuso e limitato. Stiamo supponendo che $f_n \to f$ in L^1_{loc} che implica, grazie alla Definizione 4.6, che

$$\lim_{n\to\infty} \int_{a'}^{b'} |f_n(x) - f(x)|\, dx = 0.$$

Supponiamo che f_n non converga uniformemente a f allora

$$\lim_{n\to\infty} \sup_{x\in[a',b']} |f_n(x) - f(x)| \neq 0$$

ma poichè f, f_n sono continue si ha

$$\lim_{n\to\infty} \max_{x\in[a,b]} |f_n(x) - f(x)| \neq 0.$$

Cioè esiste un $\varepsilon > 0$ tale che per ogni $n \in \mathbb{N}$ esiste $k_n \geq n$ e un $x_n \in [a, b]$ con

$$|f_{k_n}(x_n) - f(x_n)| \geq \varepsilon$$

Sia ora $M = M(a', b')$ la costante di Lipschitzianità di f su $[a', b']$ e definiamo

$$\delta = \min \left\{ \frac{\varepsilon}{2M}, a - a', b - b' \right\}.$$

Dunque per come abbiamo scelto δ, preso un $x_n \in [a, b]$ si ha

$$[x_n - \delta, x_n] \subseteq [a', b'].$$

Fissiamo ora un $n \in \mathbb{N}$ ma allora dalla relazione di prima si ha

$$f_{k_n}(x_n) \geq f(x_n) + \varepsilon \qquad f_{k_n}(x_n) \leq f(x_n) - \varepsilon.$$

Prendiamo in considerazione il primo caso (per il secondo si procederà in maniera del tutto analoga). Consideriamo dunque la derivata sinistra ottenendo così due possibili casi

$$f'_{k_n}(x_n^-) \leq 0 \qquad \text{oppure} \qquad f'_{k_n}(x_n^-) \geq 0.$$

Sia $f'_{k_n}(x_n^-) \leq 0$, allora si ha che per $x \leq x_n$

$$f_{k_n}(x) \geq f_{k_n}(x_n)$$

e quindi preso $x \in [x_n - \delta, x_n]$ e ricordando che M è costante di Lipshitzianità e cioè presi x, y vale

$$|f(x) - f(y)| \leq M|x - y|$$

si ha

$$\begin{aligned} f_{k_n}(x) - f(x) &\geq f_{k_n}(x_n) - f(x) \geq f(x_n) + \varepsilon - f(x) \\ &\geq -M|x_n - x| + \varepsilon \geq \frac{\varepsilon}{2} > 0. \end{aligned}$$

Per come abbiamo scelto δ si deduce quindi che

$$\int_{a'}^{b'} |f_{k_n}(x) - f(x)|\, dx \geq \int_{x_n-\delta}^{x_n} |f_{k_n}(x) - f(x)|\, dx \geq \frac{\varepsilon}{2}\delta.$$

Se invece supponessimo $f'_{k_n}(x_n^-) \geq 0$ si ha che per monotonia anche $f'_{k_n}(x_n^+) \geq 0$ e inoltre per $x \geq x_n$

$$f_{k_n}(x) \geq f_{k_n}(x_n)$$

e dunque procedendo come prima con $x \in [x_n, x_n + \delta]$

$$\begin{aligned} f_{k_n}(x) - f(x) &\geq f_{k_n}(x_n) - f(x) \geq f(x_n) - f(x) + \varepsilon \\ &\geq -M|x_n - x| + \varepsilon \geq \frac{\varepsilon}{2} > 0 \end{aligned}$$

da cui

$$\int_{a'}^{b'} |f_{k_n}(x) - f(x)|\, dx \geq \int_{x_n}^{x_n+\delta} |f_{k_n}(x) - f(x)|\, dx \geq \frac{\varepsilon}{2}\delta.$$

In entrambi i casi si ottiene che per ogni n

$$\int_{a'}^{b'} |f_{k_n}(x) - f(x)|\, dx \geq \frac{\varepsilon}{2}\delta$$

e quindi

$$\liminf_{n\to\infty} \int_{a'}^{b'} |f_n(x) - f(x)|\, dx \geq \frac{\varepsilon\delta}{2} \neq 0$$

che contraddice la convergenza in $L^1_{loc}(I)$. □

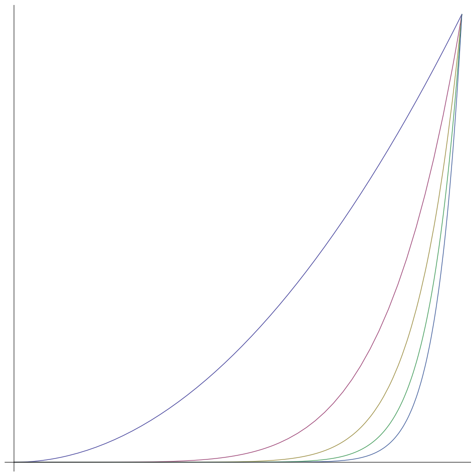


Figura 4.1: Esempio di funzione

.

Osservazione 4.1. *Il fatto che la converganza non possa essere in tutto* Ω *è sottolineato dal fatto che per esempio le funzioni* $f_n(x) = x^n$ *in* $(0, 1)$ *convergono a* 0 *in* L^1 *ma non convergono uniformemente.*

Tali convergenze sono legate anche alla convergenza degli insiemi

$$Epi(f_n) \to Epi(f), Sub(f_n) \to Sub(f) \text{ e } \Gamma(f_n) \to \Gamma(f).$$

Infatti dato che dato che

$$\|f_n - f\|_{L^1(\Omega)} = \int_\Omega |f_n(x) - f(x)|dx = \int_{\Omega\times\mathbb{R}} |\chi_{Epi(f_n)}(x,t) - \chi_{Epi(f)}(x,t)|dxdt,$$

dove abbiamo indicato con $\chi_{Epi(f)}$ la funzione caratteristica dell'epi–grafico di f, se f_n converge a f in $L^1_{loc}(\Omega)$ allora le funzioni caratteristiche di $Epi(f_n)$ convergono alle funzioni caratteristiche di $Epi(f)$ in $L^1_{loc}(\Omega \times \mathbb{R})$. Si ragiona in modo analogo con le funzioni caratteristiche del sotto-grafico.

Per quanto riguarda la convergenza del grafico abbiamo la seguente

Proposizione 4.4. *Siano* $f_j, f\colon K \to \mathbb{R}$ *funzioni continue con* $K \subset \mathbb{R}^n$ *insieme compatto: allora* f_j *converge uniformemente su* K *ad* f *se e solo se*

$$\Gamma(f_j) \quad \textit{converge secondo la convergenza di Kuratowski a} \quad \Gamma(f).$$

Dimostrazione. Dalla definizione di convergenza uniforme (4.4) sappiamo che se f_j è una successione di funzioni qualunque che converge uniformemente ad f, allora il grafico di f_n è compreso tra i grafici delle funzioni $f - \varepsilon$ e $f + \varepsilon$.

Dimostriamo quindi che se f_j convergente uniformemente a f, allora $\Gamma(f_j)$ converge a $\Gamma(f)$ nel senso di Kuratowski; dobbiamo quindi mostrare due cose:

1. per ogni $y_j \in \Gamma(f_j)$ successione con $y_j \to y$, allora $y \in \Gamma(f)$;
2. per ogni $y \in \Gamma(f)$, esiste una successione $y_j \in \Gamma(f_j)$ con $y_j \to y$.

Suponiamo quindi che

$$y_j \in \Gamma(f_j) \quad \text{tale che } y_j \to y$$

e cioè

$$y_j = (x_j, f_j(x_j)) \to y = (x_0, y_0).$$

In particolare questo implica separatamente che

$$x_j \to x_0, \qquad f(x_j) \to y_0.$$

Dobbiamo dimostrare che $f(x_0) = y_0$; dalla continuità di f e dalle precedenti considerazioni, abbiamo che per la continuità di f

$$\lim_{j\to+\infty} |f(x_j) - f(x_0)) = 0,$$

mentre grazie alla convergenza uniforme

$$\lim_{j\to+\infty} |f(x_j) - f_j(x_j)| = 0.$$

Quindi

$$|f(x_0) - y_0| \leq |f(x_0) - f(x_j)| + |f(x_j) - f_j(x_j)| + |f_j(x_j) - y_0| \to 0$$

da cui $f(x_0) = y_0$. Per la seconda verifica della convergenza di Kuratowski, preso $y = (x_0, f(x_0)) \in \Gamma(f)$, basterà considerare i punti $y_j = (x_0, f_j(x_0)) \in \Gamma(f_j)$ per ottenere, grazie al fatto che $f_j(x_0) \to f(x_0)$, che $y_j \to y$.

Dimostriamo ora la seconda parte.

VA ABBASTANZA BENE QUELLO CHE HA SCRITTO: PER CONCLUDERE DEVE SOLO USARE LA CONTINUITA UNIFORME DI f SU K.

Supponiamo che $\Gamma(f_j) \to \Gamma(f)$ secondo Kuratowski allora sappiamo che preso $y_j \in \Gamma(f_j)$ cioè preso un $y_j = (x_j, f_j(x_j))$ con $y_j \to y = (x_0, y_0) \quad x_j \to x_0$ allora per la definizione di convergenza

$$y \in \Gamma(f) \implies y_0 = f(x_0)$$

basta ora dimostrare che $\forall x \in [a, b]$ e $\forall \varepsilon > 0$

$$\exists j_\varepsilon \in \mathbb{N} \quad \text{tale che } \forall j > j_\varepsilon$$
$$|f_j(x_j) - f(x_j)| \leq \varepsilon$$

$$|f_j(x_j) - f(x_j)| \leq |f_j(x_i) - f(x_0)| + |f(x_0) - f(x_j)|$$

Da qui sfruttando il fatto che f è continua su K insieme compatto e quindi uniformemente continua si ha che

$$|f(x_0) - f(x_j)| \quad \text{è sufficientemente piccolo}$$

Inoltre grazie alla convergenza dei grafici secondo Kuratowski

$$|f_j(x_j) - f(x_0)|$$

tende a 0 per $j \to +\infty$. □

Notiamo anzitutto che dato che le funzioni sono continue, allora i grafici $\Gamma(f_j)$ e $\Gamma(f)$ sono insiemi compatti e quindi chiusi. Quindi possiamo sfruttare l'equivalenza tra convergenza di Kuratowski e convergenza nella distanza di Hausdorff. In particolare, per ogni $\varepsilon > 0$ esiste $j_0 \in \mathbb{N}$ tale che per ogni $j \geq j_0$

$$\Gamma(f_j) \subset (\Gamma(f))_\varepsilon.$$

Quindi per ogni $x \in K$, dato che $(x, f_j(x)) \in \Gamma(f_j)$, esiste $x_\varepsilon \in K$ tale che

$$\|(x, f_j(x)) - (x_\varepsilon, f(x_\varepsilon))\| \leq \varepsilon,$$

e quindi in particolare

$$m - \varepsilon \leq f(x_\varepsilon) - \varepsilon \leq f_j(x) \leq f(x_\varepsilon) + \varepsilon \leq M + \varepsilon,$$

dove abbiamo indicato con

$$m = \min_{x \in K} f(x), \qquad M = \max_{x \in K} f(x).$$

Le funzioni f_j sono quindi equi-limitate.

Supponiamo ora che non ci sia convergenza uniforme; abbiamo quindi che esiste $\varepsilon > 0$ tale che per ogni $j \in \mathbb{N}$ esiste $n_j \geq j$ e $x_j \in K$ per cui

$$|f_{n_j}(x_j) - f(x_j)| \geq \varepsilon.$$

La successione $(x_j)_{j \in \mathbb{N}}$ è contenuta nel compatto K e quindi a meno di sottosuccessioni possiamo supporre converga ad un punto $x_0 \in K$. Anche la successione $f_{n_j}(x_j)$ è contenuta in un compatto di $\mathbb{R}$ per l'equi-limitatezza della successione di funzioni, quindi, sempre a meno di sottosuccessioni, tale successione converge ad un $y_0 \in \mathbb{R}$. Per l'ipotesi di convergenza di Kuratowski, si dovrebbe quindi avere $(x_0, y_0) \in \Gamma(f)$, cioè $y_0 = f(x_0)$; questo non può verificarsi in quanto

$$\begin{aligned}|f(x_0) - y_0| =& |f(x_0) - f(x_j) + f(x_j) - f_{n_j}(x_j) + f_{n_j}(x_j) - y_0| \\ \geq& \Big| |f_{n_j}(x_j) - f(x_j)| - |f(x_0) - f(x_j) + f_{n_j}(x_j) - y_0| \Big|\end{aligned}$$

dove abbiamo usato la disuguaglianza $|a-b| \geq ||a| - |b||$ con $a = -f_{n_j}(x_j) + f(x_j)$ e $b = f(x_0) - f(x_j) + f_{n_j}(x_j) - y_0$. D'altronde, dato che

$$|f(x_0) - f(x_j) + f_{n_j} - y_0| \leq |f(x_0) - f(x_j)| + |f_{n_j}(x_j) - y_0| \to 0$$

grazie alla continuità di f e al fatto che $f_{n_j}(x_j)$ tende a y_0, si ricava che

$$|f(x_0) - y_0| \geq \liminf_{j \to +\infty} |f_{n_j}(x_j) - f(x_j)| \geq \varepsilon,$$

da cui l'assurdo.

Osservazione 4.2. *L'ipotesi sulla continuità delle funzioni coinvolte è fondamentale; infatti la successione di funzioni* $f_j : [0,1] \to \mathbb{R}$

$$f_j(x) = \chi_{[1-\frac{1}{j},1]}(x) = \begin{cases} 0 & se\ 0 \leq 1 < \frac{1}{j} \\ \\ 1 & se\ \frac{1}{j} \leq x \leq 1 \end{cases}$$

converge puntualmente ma non uniformemente alla funzione $f : [0,1] \to \mathbb{R}$

$$f(x) = \begin{cases} 0 & se\ 0 \leq x < 1 \\ \\ 1 & se\ x = 1. \end{cases}$$

Si può vedere che in questo caso $\Gamma(f_j)$ *converge secondo Kuratowski a* $\Gamma(f)$.

Bibliografia

[1] R.Tyrrel Rockafellar, *Convex Analysis*, Princeton Univ. Press, Princeton (1972).

[2] R.V.Gamkrelidze, *Analysis II Vol. 14* , Spinger-Verlag, United States of America (1990).

[3] H.G. Eggleston, *Convexity*, Cambidge Univ. Press, Cambridge (1997)

[4] R.Howard, *Alexandrov's theorem on the second derivatives of convex functions via Rademacher's theorem on the first derivates of lipschitz functions* , Univ. of South Carolina Columbia, Columbia.

[5] Ferrari, *Analisi convessa* , Univ. Milano, Milano (2008).

[6] P.Marcellini, C.Sbordone, *Analisi matematica 2* , Liguori, (1998).

[7] P.Acquistapace, *Appunti di Analisi Convessa*

[8] Anonimo, *Note sulle funzioni convesse, concave*, (2008)