`AM120 2013/14: III Settimana`

CONVESSITÁ
ed una applicazione della formula di Taylor

Definizione di insieme convesso

Dati due punti $P := (x_1, y_1), Q := (x_2, y_2)$ in $\mathbf{R}^2$, si chiama *segmento congiungente P e Q* l'insieme

$$[P, Q] := \{t(x_2, y_2) + (1-t)(x_1, y_1) := (tx_2 + (1-t)x_1, ty_2 + (1-t)y_1) : \quad t \in [0,1]\}$$

Un sottoinsieme C di $\mathbf{R}^2$ si dice convesso se

$$P, Q \in C \quad \Rightarrow \quad [P, Q] \subset C$$

ovvero *C é convesso se e solo se, contenendo due punti contiene anche il segmento che li unisce.* Notiamo che $C_\alpha, \alpha \in \mathcal{A}$ convessi $\Rightarrow \bigcap_{\alpha \in \mathcal{A}} C_\alpha$ é un convesso.

Definizione di funzione convessa

Sia $f : [a, b] \to \mathbf{R}$.

Sia $\mathcal{G}_f := \{(x, y) : \quad x \in [a, b], \quad y = f(x)\}$ il grafico di f. Si chiama *epi_f=epigrafico di f* l'insieme dei punti che stanno sopra $\mathcal{G}_f$:

$$epi_f := \{(x, y) : \quad x \in [a, b], \quad y \geq f(x)\}$$

Diremo che *f é convessa se epi_f é un insieme convesso.* É chiaro che f é convessa se e solo se

$$P, Q \in \mathcal{G}_f \quad \Rightarrow \quad [P, Q] \subset epi_f \qquad \textit{ovvero}$$

Una caratterizzazione della convessitá

f é convessa se e solo se

$$a \leq x_1 < x_2 \leq b, \ t \in [0,1] \quad \Rightarrow \quad f(tx_2 + (1-t)x_1) \ \leq \ tf(x_2) + (1-t)f(x_1)$$

Il significato geometrico di questa diseguaglianza é chiaramente il seguente:

presi due punti $P = (x_1, f(x_1))$ e $Q = (x_2, f(x_2))$ appartenenti a $\mathcal{G}_f$, il grafico si trova, nell'intervallo $[x_1, x_2]$, interamente al di sotto della retta passante per P e Q.

Per vederlo, scriviamo il segmento $[x_1, x_2]$ in forma parametrica:

$$x_t := tx_2 + (1-t)x_1, \quad t \in [0,1]$$

(chiaramente $x_t \in [x_1, x_2]$ e per ogni $x \in [x_1, x_2]$ esiste un unico $t \in [0,1]$ tale che $x_t = x$; in altre parole $t \to x_t$ é biiezione tra l'intervallo $[0,1]$ e l'intervallo $[x_1, x_2]$) ed osserviamo che la retta passante per $(x_1, f(x_1))$ e $(x_2, f(x_2))$, che ha equazione

$$y = y(x) = f(x_2) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_2) = f(x_1) + \frac{f(x_1) - f(x_2)}{x_1 - x_2}(x - x_1) \quad \forall x \in [x_1, x_2]$$

passa anche per $(x_t, tf(x_2) + (1-t)f(x_1))$, ovvero

$$y(x_t) = tf(x_2) + (1-t)f(x_1)$$

e quindi la convessitá si riscrive $f(x_t) \leq y(x_t) \quad \forall x_1, x_2 \in [a, b], t \in [0, 1]$, ovvero

$$f(x) \leq f(x_2) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_2) = f(x_1) + \frac{f(x_1) - f(x_2)}{x_1 - x_2}(x - x_1) \quad \forall x \in [x_1, x_2]$$

Vogliamo ora stabilire alcune proprietá di regolaritá delle funzioni convesse. Avremo bisogno di due Lemmi.

Il primo Lemma riguarda una proprietá di 'stabilitá' della nozione di convessitá. Dato $E \subset \mathbf{R}$ ed un insieme di indici $\mathcal{A}$, siano $\quad f_\alpha : E \to \mathbf{R}, \quad \alpha \in \mathcal{A}\quad$ tali che $\sup\{f_\alpha(x) : \quad \alpha \in \mathcal{A}\} < +\infty \quad \forall x \in E;\quad$ resta allora definita in E la funzione

$$\sup_{\alpha \in \mathcal{A}} f_\alpha : x \to \sup\{f_\alpha(x) : \quad \alpha \in \mathcal{A}\}$$

Ad esempio,

$$f^+(x) := max\{f(x), 0\}$$

é la funzione f messa uguale a zero nei punti in cui é negativa. Analogamente

$$f^-(x) := max\{-f(x), 0\}. \qquad \text{Nota che} \qquad f = f^+ - f^-, \quad |f| = f^+ + f^-$$

Lemma 1

$$f_\alpha \quad \text{convesse} \quad \Rightarrow \quad F := \sup_{\alpha \in \mathcal{A}} f_\alpha \quad \text{é convessa}$$

Prova.

$$f_\alpha(tx_2+(1-t)x_1) \le tf_\alpha(x_2)+(1-t)f_\alpha(x_1) \le t\sup_{\alpha\in\mathcal{A}} f_\alpha(x_2)+(1-t)\sup_{\alpha\in\mathcal{A}} f_\alpha(x_1) \Rightarrow$$

$$F(tx_2+(1-t)x_1) = \sup_{\alpha\in\mathcal{A}} f_\alpha(tx_2+(1-t)x_1) \le$$

$$\le t\sup_{\alpha\in\mathcal{A}} f_\alpha(x_2)+(1-t)\sup_{\alpha\in\mathcal{A}} f_\alpha(x_1) = tF(x_2)+(1-t)F(x_1)$$

Prova alternativa: epi_{f_α} convesso per ogni α implica F convessa perché

$$epi_{\sup\limits_{\alpha\in\mathcal{A}} f_\alpha} = \bigcap_{\alpha\in\mathcal{A}} epi_{f_\alpha}$$

ESEMPI. 1. Siano $f(x)=x^2$. Sia $\mathcal{A}=\mathbf{R}$ e, per $\alpha\in\mathbf{R}$, sia

$$f_\alpha(x) := f(\alpha)+f'(\alpha)(x-\alpha) = \alpha^2+2\alpha(x-\alpha) = 2\alpha x-\alpha^2$$

la retta tangente al grafico di f nel punto (α,α^2). Dal Lemma 1 segue che $F := \sup\limits_{\alpha\in\mathcal{A}} f_\alpha$ é convessa. Calcoliamo esplicitamente tale funzione.
Per Weierstrass, quale che sia $x\in\mathbf{R}$, risulta

$$F(x) := \sup_{\alpha\in\mathcal{A}} f_\alpha(x) = \sup_{\alpha\in\mathbf{R}}[2\alpha x-\alpha^2] < +\infty \qquad \text{ed é realizzato in qualche} \quad \alpha=\alpha(x)$$

Per calcolare $F(x) = f_{\alpha(x)}(x)$, occorre cercare tale $\alpha(x)$ tra gli zeri della derivata (rispetto ad α!) della funzione $\alpha\to f_\alpha(x) := 2\alpha x-\alpha^2$. Troviamo che

$$\frac{d}{d\alpha}[2\alpha x-\alpha^2] = 2(x-\alpha) = 0 \quad \Leftrightarrow \quad \alpha = x \quad \textit{ovvero} \quad \alpha(x)=x$$

A tale punto stazionario corrisponde il massimo valore di $\alpha\to f_\alpha(x)$ che é appunto dato da $f_x(x)=x^2$. Dunque

$$F(x) = \sup_{\alpha\in\mathcal{A}} f_\alpha(x) = f_x(x) = x^2 = f(x)$$

Tale procedimento, partito da f, ci ha riportato ad f! Come vedremo, ció é conseguenza di una proprietá che hanno tutte (e sole!) le funzioni convesse, se derivabili.

2. Sia $f(x)=x^4-x^2$ e, come sopra,

$$f_\alpha(x) := \alpha^4-\alpha^2+(4\alpha^3-2\alpha)(x-\alpha) = \alpha^2-3\alpha^4+2\alpha x(2\alpha^2-1)] \qquad \alpha\in\mathbf{R}$$

il fascio delle rette tangenti al grafico di f. Anche qui, come sopra, $F(x) := \sup\limits_{\alpha\in\mathcal{A}} f_\alpha(x) < +\infty$ per ogni $x\in\mathbf{R}$ ed é realizzato (per Weierstrass). Qui, fissato, per ragioni di simmetria, $x\ge 0$, troviamo che

$$\frac{d}{d\alpha} f_\alpha(x) = 2(6\alpha^2-1)(x-\alpha) = 0 \quad \Leftrightarrow \quad \alpha = x \quad \textit{oppure} \quad \alpha = \pm\sqrt{\frac{1}{6}}$$

Diciamo $\quad \alpha_1 = -\sqrt{\frac{1}{6}}, \quad \beta = x, \quad \gamma = \sqrt{\frac{1}{6}},$ i tre punti stazionari. Per sapere quale é di minimo, quale é di massimo, basta studiare il segno della derivata seconda

$$\frac{d^2}{d\alpha^2} f_\alpha(x) = 12\alpha(x-\alpha) - (6\alpha^2 - 1)$$

Visto tuttavia che f_α é un polinomio di quarto grado (in α) con coefficente del termine di grado quattro negativo, il piú piccolo ed il piú grande tra i punti stazionari (quando non ve ne sono di coincidenti) sono punti di massimo, mentre l'altro é di minimo. Dunque, fissato $x \geq 0, \quad \alpha_1$ é punto di massimo locale e

$$f_{-\sqrt{\frac{1}{6}}}(x) = \frac{3}{36} + \frac{4}{3}\sqrt{\frac{1}{6}}x \qquad \text{é un massimo locale per} \quad \alpha \to f_\alpha(x)$$

($y = \frac{3}{36} + \frac{4}{3}\sqrt{\frac{1}{6}}x$ é la retta tangente al grafico di f nel punto, di inflessione per f, $(-\sqrt{\frac{1}{6}}, -\frac{5}{36})$), $\quad$ e $\quad f_{\sqrt{\frac{1}{6}}}(x) = \frac{3}{36} - \frac{4}{3}\sqrt{\frac{1}{6}}x < f_{-\sqrt{\frac{1}{6}}}(x) \quad \forall x > 0.$ Infine, $\beta = x \geq 0$ é di massimo se e solo se $x > \sqrt{\frac{1}{6}}$ e tale massimo vale $x^4 - x^2$. Dunque $F(x) = f_{-\sqrt{\frac{1}{6}}}(x)$ sicuramente per $x \in [0, \sqrt{\frac{1}{6}}]$. Per confrontare i due massimi (locali) per $x > \sqrt{\frac{1}{6}}$, osserviamo che lo sviluppo di Taylor centrato in $-\sqrt{\frac{1}{6}}$ fornisce

$$x^4 - x^2 - [\frac{3}{36} + \frac{4}{3}\sqrt{\frac{1}{6}}x] = (x + \sqrt{\frac{1}{6}})^3 (x - \frac{3}{\sqrt{6}})$$

e quindi

$$f_x(x) = x^4 - x^2 \;\geq\; \frac{3}{36} + \frac{4}{3}\sqrt{\frac{1}{6}}x = f_{\alpha_1}(x) \qquad \Leftrightarrow \qquad x \geq \frac{3}{\sqrt{6}}$$

ove $\frac{3}{\sqrt{6}}$ é l'ascissa del punto intersezione del grafico di f con la retta tangente al grafico di f nel punto (di flesso) $(-\sqrt{\frac{1}{6}}, -\frac{5}{36})$. Quindi

$$\textit{in} \quad [0, \frac{3}{\sqrt{6}}), \qquad F(x) := \sup_{\alpha \in \mathcal{A}} f_\alpha(x) = \quad \frac{3}{36} + \frac{4}{3}\sqrt{\frac{1}{6}}x$$

$$\textit{in} \quad x \geq \frac{3}{\sqrt{6}} \qquad F(x) := \sup_{\alpha \in \mathcal{A}} f_\alpha(x) = f_x(x) = x^4 - x^2$$

Una conseguenza del Lemma 1 riguarda le funzioni di classe C^1. Se $f \in C^1([a,b])$ e $x_o \in (a,b)$ allora il grafico di f é dotato di retta tangente in $(x_o, f(x_o))$, data dall'equazione

$$y = T_{x_o}(x) := f(x_o) + f'(x_o)(x - x_o) \qquad x_o \in (a,b),\ x \in \mathbf{R}$$

Proposizione 1 Sia $f \in C^1([a,b])$. Se

$$f(x) \geq T_{x_o}(x) := f(x_o) + f'(x_o)(x - x_o) \qquad \forall x, x_o \in [a,b]$$

allora f é convessa in $[a,b]$

Prova. Dall'ipotesi segue infatti che

$$f(x) := \sup_{x_o \in [a,b]} T_{x_o}(x) \qquad (\sharp)$$

e quindi f é convessa perché le T_{x_o} sono ovviamente convesse, e l'estremo superiore di una famiglia di funzioni convesse é una funzione convessa.
Per mostrare ($\sharp$), osserviamo che, da una parte, fissato $x \in [a,b]$, si ha

$$\sup_{x_o \in [a,b]} T_{x_o}(x) \geq T_x(x) = f(x)$$

D'altra parte, l'ipotesi dice che $f(x) \geq T_{x_o}(x) \quad \forall x, x_o \in [a,b]$ e quindi, fissato x,

$$f(x) \geq \sup_{x_o \in [a,b]} T_{x_o}(x) \quad \forall x \in [a,b]$$

Il secondo Lemma riguarda la monotonia del rapporto incrementale di una funzione convessa,

Lemma 2 Sia f convessa in (a,b), $x_1, x_2 \in (a,b)$. Allora

$$x_1 < x < x_2 \quad \Rightarrow \quad \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}$$

In particolare, $x \to \frac{f(x)-f(x_1)}{x-x_1}$ e $x \to \frac{f(x_2)-f(x)}{x_2-x}$ sono non decrescenti in (x_1, x_2)

Prova del Lemma.

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \quad \Leftrightarrow \quad f(x) \leq f(x_1) + \left(\frac{f(x_2) - f(x_1)}{x_2 - x_1}\right)(x - x_1)$$

e la seconda diseguaglianza é vera perché $y = f(x_1) + \left(\frac{f(x_2)-f(x_1)}{x_2-x_1}\right)(x - x_1)$ é la retta passante per i punti $(x_1, f(x_1))$ e $(x_2, f(x_2)$ ed f é convessa. Analogamente,

$$f(x) \leq f(x_2) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_2) \quad \Leftrightarrow \quad (x_2 - x)\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f(x_2) - f(x).$$

Una importante conseguenza del Lemma 2 é il seguente

Teorema 1

Sia f convessa in (a,b). Allora

(i) per ogni $x_0 \in (a,b)$ esistono finiti $f'_-(x_0)$ ed $f'_+(x_0)$ e $f'_-(x_0) \le f'_+(x_0)$.

(ii) f é continua in (a,b)

(iii) f é derivabile in (a,b) al di fuori di un insieme di punti al piú numerabile.

Prova di (i). Siano $x_1 < x_0 < x_2$.
Dalla monotonia di $\frac{f(x)-f(x_0)}{x-x_0}$ a sinistra e a destra di x_0 e da (ii) segue che

$$\frac{f(x_1)-f(x_0)}{x_1-x_0} \le f'_-(x_0) := \lim_{x\to x_0^-} \frac{f(x)-f(x_0)}{x-x_0} = \sup_{x<x_0} \frac{f(x)-f(x_0)}{x-x_0} \le$$

$$\le \inf_{x>x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{x\to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} = f'_+(x_0) \le \frac{f(x_2)-f(x_0)}{x_2-x_0}$$

Prova di (ii). Da 'esiste finito il limite, al tendere di x a x_0 da sinistra (destra) di $\frac{f(x)-f(x_0)}{x-x_0}$', segue che il numeratore $f(x)-f(x_0)$ deve tendere a zero, e quindi $f(x)$ tende, al tendere di x a x_0 da sinistra (da destra) a $f(x_0)$. Dunque $f(x)$ tende a $f(x_0)$ al tendere di x a x_0.

Prova di (iii). Da (i), vediamo che la f é derivabile a destra e a sinistra in ogni punto e che, se f non é derivabile in x_0, resta definito l'intervallo non vuoto $(f'_-(x_0), f'_+(x_0))$.
Inoltre, se $x_1 > x_0$ é un altro punto in cui f non é derivabile, allora $(f'_+(x_0) \le f'_-(x_1))$ e quindi $(f'_-(x_0), f'_+(x_0))$ e $(f'_-(x_1), f'_+(x_1))$ sono disgiunti (analogamente se $x_1 < x_0$). Dunque la famiglia dei punti di non derivabilitá di f individua una famiglia di intervalli aperti disgiunti. Siccome ogni intervallo aperto individua (almeno) un razionale, e quindi una famiglia di intervalli disgiunti individua una famiglia di razionali (distinti tra loro!), tale famiglia puó essere, al pari di $\mathbf{Q}$, al piú numerabile.

In classi di funzioni derivabili la convessitá ha semplici caratterizzazioni. Per vederlo, cominciamo con un'altra semplice conseguenza del Lemma 2: se f é convessa e $x_1 < x < x_2$, si ha

$$f'_+(x_1) = \lim_{x\to x_1} \frac{f(x)-f(x_1)}{x-x_1} = \inf_{x>x_1} \frac{f(x)-f(x_1)}{x-x_1} \le$$

$$\leq \sup_{x<x_2} \frac{f(x_2)-f(x)}{x_2-x} = \lim_{x\to x_2} \frac{f(x_2)-f(x)}{x_2-x} = f'_-(x_2)$$

e quindi, se $f \in C^1$, f' é non decrescente.

Se, viceversa, f' é non decrescente in (a,b), allora

$$T_{x_o}(x) := f(x_o) + f'(x_o)(x-x_o) \leq f(x) \qquad \forall x, x_o \in (a,b)$$

perché, se no, esisterebbe $x_1 \in (a,b)$, diciamo $x_1 > x_o$, tale che $f(x_1) < T_{x_o}(x_1)$ e quindi otterremmo, per Lagrange

$$\exists\, \xi \in (x_o, x_1): \qquad f'(\xi) = \frac{f(x_1)-f(x_o)}{x_1-x_o} < \frac{T_{x_o}(x_1)-f(x_o)}{x_1-x_o} = f'(x_o)$$

Riassumendo, da quanto sopra, e dalla Proposizione 1, deduciamo due formulazioni, equivalenti, della convessitá per funzioni di classe C^1 (qui, come sopra, $T_{x_o}(x) := f(x_o) + f'(x_o)(x-x_o)$ denota la retta tangente al grafico di f in $(x_o, f(x_o))$):

Teorema 2 Sia $f \in C^1((a,b))$. Allora le seguenti affermazioni sono tra loro equivalenti:

(i) f é convessa

(ii) f' é non decrescente

(iii) $T_{x_o}(x) \leq f(x) \quad \forall x, x_o \in (a,b)$

Otteniamo in particolare una caratterizzazione della convessitá per funzioni C^2:

Teorema 3

Sia $f \in C^2([a,b])$. Allora f é convessa se e solo se $f'' \geq 0$.

Notiamo che la non negativitá di f'' implica subito, via Taylor, la convessitá:

$$\forall x, \exists \xi(x): \quad f(x) = f(x_o) + f'(x_o)(x-x_o) + \frac{1}{2} f''(\xi)(x-x_o)^2 =$$

$$T_{x_o}(x) + \frac{1}{2} f''(\xi)(x-x_o)^2 \geq T_{x_o}(x)$$