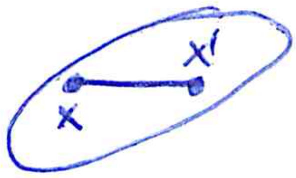
ANALISI 1 | ma 9/4/19 |

Convessità

Definizione Sia $E \subseteq \mathbb{R}^N$. Diciamo che E è <u>convesso</u> se per ogni $\underline{x}, \underline{x}' \in E$ si ha $(1-t)\underline{x} + t\underline{x}' \in E \quad \forall t \in [0,1]$.

Cioè, E è convesso se per ogni $\underline{x}, \underline{x}' \in E$ l'intero <u>segmento</u> di estremi $\underline{x}, \underline{x}'$ è contenuto in E.



Sono convessi:

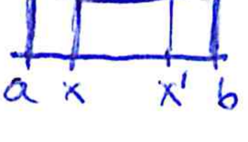
le palle (chiuse o aperte) in $\mathbb{R}^N$, i piani, le rette, i semispazi in $\mathbb{R}^N$, gli intervalli (limitati o no) in $\mathbb{R}$.

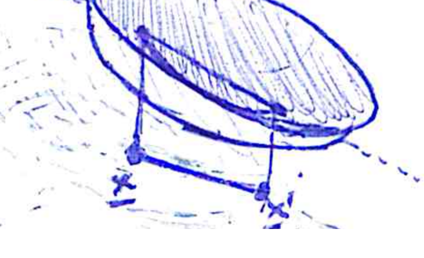
Definizione Sia $K \subseteq \mathbb{R}^N$ convesso. Dico che $f: K \to \mathbb{R}$ è <u>convessa</u> se per ogni $\underline{x}, \underline{x}' \in K$ e per ogni $t \in [0,1]$ si ha

$$f\big((1-t)\underline{x} + t\underline{x}'\big) \leq (1-t) f(\underline{x}) + t f(\underline{x}')$$

Diciamo che f è <u>concava</u> se $-f$ è convessa.

L'interpretazione geometrica è: $(1-t)\underline{x} + t\underline{x}'$, per $t \in [0,1]$, sono i punti del segmento di estremi $\underline{x}, \underline{x}'$; il punto $(\underline{u}, y) \in \mathbb{R}^{N+1}$ di coordinate $\underline{u} = (1-t)\underline{x} + t\underline{x}'$, $y = (1-t) f(\underline{x}) + t f(\underline{x}')$ sta nel segmento di estremi $(\underline{x}, f(\underline{x}))$ e $(\underline{x}', f(\underline{x}'))$. La condizione di convessità ci dice che $f(\underline{u}) \leq y$, cioè che il grafico della restrizione di f al segmento di estremi $\underline{x}, \underline{x}'$ sta <u>sotto</u> il segmento di estremi $(\underline{x}, f(\underline{x}))$ e $(\underline{x}', f(\underline{x}'))$.





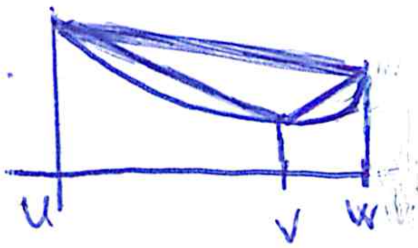
Esercizio. Provare che $f:K\to\mathbb{R}$ è convessa se e solo se il suo epigrafo E,

$$E=\{(\underline{x},y)\in\mathbb{R}^{N+1}:\ \underline{x}\in K,\ f(\underline{x})\le y\}$$

è convesso in $\mathbb{R}^{N+1}$.

Proposizione. Sia I un intervallo di $\mathbb{R}$. Sia $f: I\to\mathbb{R}$. Allora f è convessa in I se e solo se per ogni $u,v,w\in I$ con $u<v<w$ si ha

$$\frac{f(v)-f(u)}{v-u}\le\frac{f(w)-f(u)}{w-u}\le\frac{f(w)-f(v)}{w-v}.$$



Il disegno è già una dimostrazione! Comunque:

dim ($\Rightarrow$). Sia f convessa e siano $u<v<w$. Allora $\exists\lambda\in]0,1[$ tali che $v=(1-\lambda)u+\lambda w$ e si ha precisamente $\lambda=\frac{v-u}{w-u}$, per cui $1-\lambda=\frac{w-v}{w-u}$. Per ipotesi

$$f(v)\le\frac{w-v}{w-u}f(u)+\frac{v-u}{w-u}f(w),$$

da cui

$$f(v)-f(u)\le\left(\frac{w-v}{w-u}-1\right)f(u)+\frac{v-u}{w-u}f(w)=\frac{v-u}{w-u}[f(w)-f(u)],$$

e dividendo per $v-u$ si ha la prima relazione. Similmente

$$f(v)-f(w)\le\frac{w-v}{w-u}f(u)+\left[\frac{v-u}{w-u}-1\right]f(w)=\frac{w-v}{w-u}(f(u)-f(w))$$

ovvero

$$[f(w)-f(u)]\frac{w-v}{w-u}\le f(w)-f(v);$$

dividendo per $w-v$ si ha la seconda relazione.

($\Longleftarrow$) Dalla 1ª relazione segue

$$f(v) \le \left[-\frac{v-u}{w-u}+1\right] f(u) + \frac{v-u}{w-u} f(w) = \frac{w-v}{w-u} f(u) + \frac{v-u}{w-u} f(w),$$

che è la tesi se si pone $\lambda = \frac{v-u}{w-u}$. □

Osservazione Se $f: [a,b] \to \mathbb{R}$ è convessa, allora f è continua in $]a,b[$.

dim. Sia $x_0 \in]a,b[$. Se ad esempio $x_0 < x < b$, esistono unici $t,s \in]0,1[$ tali che

$$x = (1-t)x_0 + tb, \qquad \text{con } t = \frac{x-x_0}{b-x_0},$$

$$x_0 = (1-s)x + sa, \qquad \text{con } s = \frac{x_0-x}{a-x}.$$

Per convessità

$$f(x) \le (1-t) f(x_0) + t f(b), \qquad f(x_0) \le (1-s) f(x) + s f(a)$$

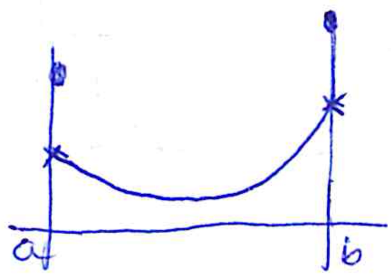
da cui

$$f(x) - f(x_0) \le t[f(b) - f(x_0)] = (x-x_0) \frac{f(b)-f(x_0)}{b-x_0},$$

$$f(x_0) - f(x) \le s[f(a) - f(x)] = (x_0-x) \frac{f(a)-f(x)}{a-x} \le (x_0-x) \frac{f(b)-f(a)}{b-a}.$$

Per $x \to x_0$, si ha la tesi. □

Si noti che una funzione convessa in $[a,b]$ può essere discontinua negli estremi.



Teorema. Sia $A \subseteq \mathbb{R}^N$ un aperto convesso e sia $f : A \to \mathbb{R}$ differenziabile. La f è convessa in A se e solo se

$$f(\underline{x}) \geq f(\underline{x}_0) + \langle \nabla f(\underline{x}_0), \underline{x}-\underline{x}_0 \rangle_N \qquad \forall\, \underline{x}, \underline{x}_0 \in A .$$

dim. ($\Leftarrow$) Siano $\underline{x}_1, \underline{x}_2 \in A$ e poniamo $\underline{x}_0 = (1-t)\underline{x}_1 + t\underline{x}_2$, con $t \in [0,1]$. Dall'ipotesi si ha

$$f(\underline{x}_1) \geq f(\underline{x}_0) + \langle \nabla f(\underline{x}_0), \underline{x}_1-\underline{x}_0 \rangle_N = f(\underline{x}_0) + \langle \nabla f(\underline{x}_0), t(\underline{x}_1-\underline{x}_2) \rangle_N$$

$$f(\underline{x}_2) \geq f(\underline{x}_0) + \langle \nabla f(\underline{x}_0), \underline{x}_2-\underline{x}_0 \rangle_N = f(\underline{x}_0) + \langle \nabla f(\underline{x}_0), (1-t)(\underline{x}_1-\underline{x}_2) \rangle_N ,$$

da cui

$$(1-t) f(\underline{x}_1) + t f(\underline{x}_2) \geq f(\underline{x}_0) + 0 .$$

($\Rightarrow$) Siano $\underline{x}, \underline{x}_0 \in A$ e sia $\underline{h} = \underline{x} - \underline{x}_0$. Per convessità,

$$f(\underline{x}_0 + t\underline{h}) = f((1-t)\underline{x}_0 + t\underline{x}) \leq (1-t) f(\underline{x}_0) + t f(\underline{x}) =$$
$$= f(\underline{x}_0) + t\left(f(\underline{x}_0+\underline{h}) - f(\underline{x}_0)\right) .$$

Ne segue

$$f(\underline{x}_0 + t\underline{h}) - f(\underline{x}_0) - t\langle \nabla f(\underline{x}_0), \underline{h} \rangle_N \leq$$
$$\leq t \left[f(\underline{x}_0+\underline{h}) - f(\underline{x}_0) - \langle \nabla f(\underline{x}_0), \underline{h} \rangle_N \right]$$

e dividendo per t, e applicando la definizione di differenziabilità segue

$$0 \leq f(\underline{x}_0+\underline{h}) - f(\underline{x}_0) - \langle \nabla f(\underline{x}_0), \underline{h} \rangle_N . \qquad \square$$

Per provare gli ultimi due risultati, è utile introdurre un metodo per "trasformare" funzioni convesse di N variabili in funzioni convesse di 1 variabile.

<u>Lemma</u> Sia $A \subseteq \mathbb{R}^N$ un aperto convesso, sia $f: A \to \mathbb{R}$.
Per $\underline{x}, \underline{x}' \in A$, poniamo

$$\phi_{\underline{x},\underline{x}'}(\lambda) = \phi(\lambda) = f((1-\lambda)\underline{x} + \lambda\underline{x}'), \quad \lambda \in [0,1].$$

Allora f è convessa in A se e solo se ϕ è convessa in $[0,1]$. Inoltre f è di classe C^k se e solo se ϕ è di classe C^k.

<u>dim</u>. Se ϕ è convessa in $[0,1]$, allora f è convessa in A: infatti se $\underline{x}, \underline{x}' \in A$ e $\lambda \in [0,1]$, allora

$$f((1-\lambda)\underline{x} + \lambda\underline{x}') = \phi_{\underline{x},\underline{x}'}(\lambda) \leq (1-\lambda)\phi_{\underline{x},\underline{x}'}(0) + \lambda\phi_{\underline{x},\underline{x}'}(1) =$$

$$= (1-\lambda) f(\underline{x}) + \lambda f(\underline{x}').$$

Se f è convessa in A, proviamo che ϕ è convessa in $[0,1]$.
Siano $\lambda_1, \lambda_2 \in [0,1]$ e poniamo $\lambda = (1-t)\lambda_1 + t\lambda_2$. Allora

$$\phi_{\underline{x},\underline{x}'}(\lambda) = \phi_{\underline{x},\underline{x}'}((1-t)\lambda_1 + t\lambda_2) =$$

$$= f((1-(1-t)\lambda_1 - t\lambda_2)\underline{x} + ((1-t)\lambda_1 + t\lambda_2)\underline{x}') =$$

$$= f\Big(\big(1-(1-t)\lambda_1\big)\underline{x} + (1-t)\lambda_1 \underline{x}' - t\lambda_2 \underline{x} + t\lambda_2 \underline{x}' \Big) =$$

$$= f\Big((1-t)\big((1-\lambda_1)\underline{x} + \lambda_1 \underline{x}'\big) + t\big((1-\lambda_2)\underline{x} + \lambda_2 \underline{x}'\big) \Big) \leq$$

$$\leq (1-t) f\big((1-\lambda_1)\underline{x} + \lambda_1 \underline{x}'\big) + t f\big((1-\lambda_2)\underline{x} + \lambda_2 \underline{x}'\big) =$$

$$= (1-t)\,\phi_{\underline{x},\underline{x}'}(\lambda_1) + t\,\phi_{\underline{x},\underline{x}'}(\lambda_2).$$

Infine è chiaro che se f è di classe C^k, tale è ϕ, essendo la composizione di f con un polinomio in λ di grado 1. $\square$

<u>Teorema</u> Sia $A \subseteq \mathbb{R}^N$ un aperto convesso e sia $f: A \to \mathbb{R}$ differenziabile. Allora f è convessa se e solo se ∇f è <u>monotono</u>, vale a dire

$$\langle \nabla f(\underline{x}) - \nabla f(\underline{x}'), \underline{x} - \underline{x}' \rangle_N \geq 0 \quad \forall \underline{x}, \underline{x}' \in A.$$

Si noti che, se $N=1$, l'enunciato è:

$$f \text{ è convessa in }]a,b[\iff f' \text{ è crescente in }]a,b[.$$

<u>dim</u>. Se $N=1$, e f è convessa, allora se $\xi, \eta \in]a,b[$ con $\xi < \eta$, si ha, per la crescenza dei rapporti incrementali,

$$f'(\xi) = \lim_{h \to 0^+} \frac{f(\xi+h) - f(\xi)}{h} \leq \frac{f\left(\frac{\xi+\eta}{2}\right) - f(\xi)}{\frac{\eta-\xi}{2}} \leq \frac{f(\eta) - f\left(\frac{\xi+\eta}{2}\right)}{\frac{\eta-\xi}{2}} \leq$$

$$\leq \lim_{k \to 0^-} \frac{f(\eta) - f(\eta-k)}{-k} = f'(\eta).$$

dunque f' è crescente.

Viceversa, se f' è crescente, siano $s,t \in]a,b[$ con $s<t$, e sia $x=(1-\lambda)s+\lambda t$, $\lambda \in]0,1[$. Allora esistono $\xi \in]s,x[$, $\eta \in]x,t[$ tali che

$$f(x)-f(s) = f'(\xi)(x-s) = f'(\xi)\,\lambda\,(t-s),$$
$$f(t)-f(x) = f'(\eta)(t-x) = f'(\eta)\,(1-\lambda)(t-s).$$

Dunque, per la crescenza di f',

$$\frac{f(x)-f(s)}{\lambda(t-s)} = f'(\xi) \leq f'(\eta) = \frac{f(t)-f(x)}{(1-\lambda)(t-s)}$$

e pertanto

$$f(x)\left(\frac{1}{\lambda}+\frac{1}{1-\lambda}\right) \leq \frac{f(s)}{\lambda}+\frac{f(t)}{1-\lambda}.$$

Moltiplicando per $\lambda(1-\lambda)$ si ottiene $f(x) \leq (1-\lambda)f(s)+\lambda f(t)$.

Sia ora $N>1$. Se f è convessa e $\underline{x},\underline{x}' \in A$, allora per il teorema di pag. 275

$$f(\underline{x}) \geq f(\underline{x}') + \langle \nabla f(\underline{x}'), \underline{x}-\underline{x}' \rangle_N$$
$$f(\underline{x}') \geq f(\underline{x}) + \langle \nabla f(\underline{x}), \underline{x}'-\underline{x} \rangle_N$$

e sommando si ricava, dopo semplificazione,

$$0 \geq \langle \nabla f(\underline{x}')-\nabla f(\underline{x}), \underline{x}-\underline{x}' \rangle_N \,;$$

Moltiplicando per -1 si ha la monotonia di ∇f.

Viceversa, sia ∇f monotono: se $\underline{x},\underline{x}' \in A$, allora $\phi_{\underline{x},\underline{x}'}(\lambda)$ verifica, posto $\underline{x}_\lambda = (1-\lambda)\underline{x}+\lambda\underline{x}'$,

$$\phi'(\lambda) = \langle \nabla f(\underline{x}_\lambda), \underline{x}' - \underline{x} \rangle_N$$

e quindi per $\mu > \lambda$, osservato che $\underline{x}_\mu - \underline{x}_\lambda = (\mu - \lambda)(\underline{x}' - \underline{x})$,

$$\phi'(\mu) - \phi'(\lambda) = \langle \nabla f(\underline{x}_\mu) - \nabla f(\underline{x}_\lambda), \underline{x}' - \underline{x} \rangle_N =$$

$$= \frac{\langle \nabla f(\underline{x}_\mu) - \nabla f(\underline{x}_\lambda), \underline{x}_\mu - \underline{x}_\lambda \rangle_N}{\mu - \lambda} \geq 0$$

Perciò ϕ' è crescente e quindi ϕ è convessa. Pertanto f è convessa. $\square$

<u>Teorema</u> Sia $A \subseteq \mathbb{R}^N$ un aperto convesso e sia $f : A \to \mathbb{R}$ di classe C^2. Allora f è convessa se e solo se

$$\langle H_f(\underline{x}) \underline{v}, \underline{v} \rangle_N \geq 0 \quad \forall \underline{v} \in \mathbb{R}^N, \ \forall \underline{x} \in A,$$

ove $H_f(\underline{x})$ è la matrice Hessiana di f in $\underline{x}$,

$$H_f(\underline{x}) = \{ D_i D_j f(\underline{x}) \}_{1 \leq i \leq N, \ 1 \leq j \leq N} .$$

Si noti che per $N=1$ l'enunciato è "

$$f \text{ è convessa in }]a,b[\iff f'' \geq 0 \text{ in }]a,b[.$$

<u>dim</u>. Se $N=1$, allora

$$f \text{ convessa} \iff f' \text{ crescente} \iff f'' \geq 0 .$$

Sia $N > 1$. Se $\langle H_f(\underline{x}) \underline{v}, \underline{v} \rangle_N \geq 0$ per ogni $\underline{v} \in \mathbb{R}^N$ e $\underline{x} \in A$, siano $\underline{x}, \underline{x}' \in A$. La funzione $\phi_{\underline{x},\underline{x}'}$ è di classe C^2 e, per ogni $\lambda \in [0,1]$,

$$\phi''(\lambda) = \langle H_f(\underline{x}_\lambda)(\underline{x}'-\underline{x}), \underline{x}'-\underline{x} \rangle_N ,$$

ove $\underline{x}_\lambda = (1-\lambda)\underline{x} + \lambda \underline{x}'$. Per ipotesi, $\phi''(\lambda) \geq 0$ per ogni $\lambda \in [0,1]$, ossia ϕ è convessa; quindi anche f è convessa.

Viceversa, sia f convessa. Fissiamo $\underline{x} \in A$ e $\underline{v} \in \mathbb{R}^N$. Allora $\underline{x} + t\underline{v} \in A$ per t piccolo e positivo. La funzione $\phi_{\underline{x}, \underline{x}+t\underline{v}}$ è convessa, essendo f convessa; quindi si ha $\phi'' \geq 0$, ossia

$$\langle H_f(\underline{x}+\lambda t\underline{v})\, t\underline{v}, t\underline{v} \rangle_N \geq 0 \quad \forall \lambda \in [0,1],$$

e per ogni $t > 0$ piccolo. Dividendo per t^2 e mandando $t \to 0^+$, si ricava

$$\langle H_f(\underline{x})\underline{v}, \underline{v} \rangle_N \geq 0 . \quad \square$$