Funciones Convexas

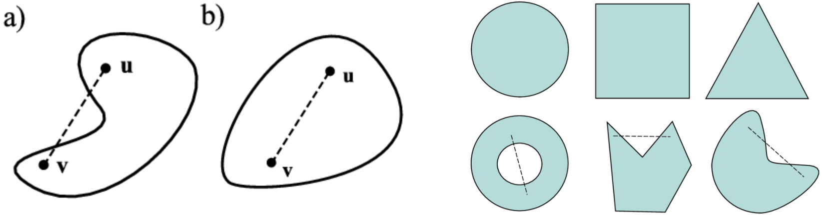
Alan Reyes-Figueroa

Métodos Numéricos II

(Aula 16) 08.septiembre.2022

Definición

Un subconjunto $\Omega \subseteq \mathbb{R}^n$ *es* ***convexo*** *si para todo* $\mathbf{x}, \mathbf{y} \in \Omega$*, el segmento de recta* $[\mathbf{x}, \mathbf{y}] = \{(1-t)\mathbf{x} + t\mathbf{y} : t \in [0, 1]\}$ *está totalmente contenido en* Ω*.*



(a) Conjunto no convexo, (b) Conjunto convexo.

Ejemplos:

- Convexos: esferas, hiperplanos, semiespacios, conos, ...
- No Convexos: conjunto no conexos, uniones de rectas, uniones en general, ...

Definición

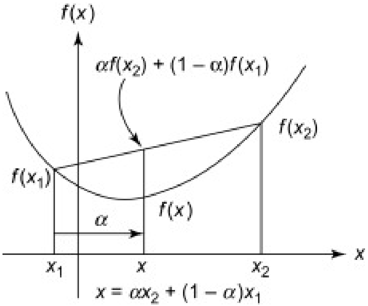
*Una función $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ es **convexa** si $\Omega = \mathrm{dom}\, f$ es un conjunto convexo, y para todo $\mathbf{x}, \mathbf{y} \in \Omega$, y todo $t \in [0, 1]$ vale*

$$f\big((1-t)\mathbf{x} + t\mathbf{y}\big) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y}). \qquad (1)$$

Geométricamente, la desigualdad (1) significa que el segmento de recta entre $(\mathbf{x}, f(\mathbf{x}))$ y $(\mathbf{y}, f(\mathbf{y}))$ está por encima de la gráfica de f.

La función f es **estrictamente convexa** si en (1) vale la desigualdad estricta, siempre que $\mathbf{x} \neq \mathbf{y}$ y $t \neq 0, 1$. Decimos que f es **cóncava** (**estrictamente cóncava**) si $-f$ es convexa (estrictamente convexa).

A la desigualdad (1) se le llama usualmente **desigualdad de Jensen**.



Propiedad

Sea $\Omega \subseteq \mathbb{R}^n$ conjunto convexo. La función $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ es convexa $\iff$ para todo $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \Omega$, y cualesquiera $t_1, \ldots, t_k \in [0, 1]$, con $\sum_{i=1}^k t_i = 1$, se tiene que

$$f\Big(\sum_{i=1}^{k} t_i\,\mathbf{x}_i\Big) \leq \sum_{i=1}^{k} t_i\,f(\mathbf{x}_i). \tag{2}$$

<u>Prueba</u>: $(\Leftarrow)$ Para $k = 2$, tome $\mathbf{x}_1 = \mathbf{x}, \mathbf{x}_2 = \mathbf{y} \in \Omega$, y sean $t_1 = 1 - t, t_2 = t$, con $t \in [0, 1]$. La desigualdad (2) se reduce a $f((1 - t)\mathbf{x} + t\mathbf{y}) \leq (1 - t)f(\mathbf{x}) + tf(\mathbf{y})$, lo que implica que f es convexa.

$(\Rightarrow)$ Mostramos la desigualdad (2) por inducción sobre k.
Para $k = 1$, necesariamente $t_1 = 1$ de modo que $f(\mathbf{x}_1) \leq f(\mathbf{x}_1)$ y (2) se cumple de manera automática. El caso $k = 2$ se cumple a partir de la definición de convexidad (1).

Suponga que (2) se cumple para cualesquiera k puntos $\mathbf{p}_1, \ldots, \mathbf{p}_k \in \Omega$, siempre que se forme una combinación lineal convexa $s_1\mathbf{p}_1 + \ldots + s_k\mathbf{p}_k$, con $0 \leq s_i \leq 1$ y $\sum_{i=1}^k s_i = 1$.

Suponga ahora que $\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{x}_{k+1} \in \Omega$, se combinan para formar un punto

$$\mathbf{x} = t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \ldots + t_{k+1}\mathbf{x}_{k+1} \in \Omega, \quad \sum_{i=1}^{k+1} t_i = 1, \; 0 \leq t_i \leq 1.$$

Definamos $t = t_{k+1}$, $1 - t = \sum_{j=1}^{k} t_j = t_1 + \ldots + t_k$. Ambos coeficientes satisfacen $0 \leq t, 1 - t \leq 1$. En particular, si $\mathbf{p} = \sum_{j=1}^{k} s_j\mathbf{x}_j \in \Omega$, con $\sum_{j=1}^{k} s_j = 1$, podemos escribir

$$\mathbf{x} = (1-t)\mathbf{p} + t\mathbf{x}_{k+1} = (1-t)\sum_{j=1}^{k} s_j\mathbf{x}_j + t_{k+1} \quad \Longrightarrow \quad t_j = (1-t)s_j, \; j = 1, \ldots, k;$$

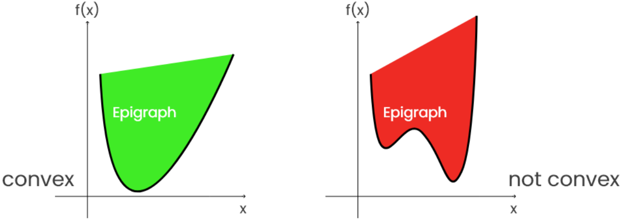
y

$$\begin{aligned}
f\big(\textstyle\sum_{i=1}^{k+1} t_i\,\mathbf{x}_i\big) &= f\big((1-t)\mathbf{p} + t\,\mathbf{x}_{k+1}\big) \;\leq\; (1-t)f(\mathbf{p}) + tf(\mathbf{x}_{k+1}) \\
&\leq (1-t)f\big(\textstyle\sum_{j=1}^{k} s_j\,\mathbf{x}_j\big) + tf(\mathbf{x}_{k+1}) \;\leq\; (1-t)\textstyle\sum_{j=1}^{k} s_j f(\mathbf{x}_j) + tf(\mathbf{x}_{k+1}) \\
&\leq \textstyle\sum_{j=1}^{k} t_j f(\mathbf{x}_j) + tf(\mathbf{x}_{k+1}) \;\leq\; \textstyle\sum_{i=1}^{k+1} t_i f(\mathbf{x}_i). \;\square
\end{aligned}$$

Definición

*Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$. Definimos el **epígrafo** de f, como el conjunto*

$$\mathrm{Epi}(f) = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \; y \geq f(\mathbf{x})\} \subseteq \mathbb{R}^{n+1}.$$



Teorema

f es convexa $\iff$ su epígrafo $\mathrm{Epi}(f)$ es un conjunto convexo.

<u>Prueba</u>: $(\Rightarrow)$. Supongamos que f es convexa, y sean $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_k, y_k) \in \mathrm{Epi}(f)$.

Tomemos cualquier juego de coeficientes $t_1, t_2, \ldots, t_k \in [0, 1]$, tales que $\sum_{i=1}^{k} t_i = 1$. Consideramos el punto

$$(\mathbf{x}, y) = \sum_{i=1}^{k} t_i \, (\mathbf{x}_i, y_i) = \Big(\sum_{i=1}^{k} t_i \, \mathbf{x}_i, \sum_{i=1}^{k} t_i \, y_i \Big) \in \mathbb{R}^{n+1}.$$

Este punto satisface

$$y = \sum_{i=1}^{k} t_i \, y_i \geq \sum_{i=1}^{k} t_i \, f(\mathbf{x}_i) \geq f\Big(\sum_{i=1}^{k} t_i \, \mathbf{x}_i \Big) = f(\mathbf{x}),$$

de modo que $(\mathbf{x}, y) \in Epi(f)$, lo que muestra que $\text{Epi}(f)$ es convexo.

($\Leftarrow$) Tomamos $(\mathbf{x}_1, f(\mathbf{x}_1)), \ldots, (\mathbf{x}_k, f(\mathbf{x}_k)) \in \text{Epi}(f)$. Como $\text{Epi}(f)$ es convexo, entonces se cumple que

$$\sum_{i=1}^{k} t_i \, (\mathbf{x}_i, f(\mathbf{x}_i)) = \Big(\sum_{i=1}^{k} t_i \, \mathbf{x}_i, \sum_{i=1}^{k} t_i \, f(\mathbf{x}_i) \Big) \in \text{Epi}(f).$$

Esto implica que $f\big(\sum_{i=1}^{k} t_i \, \mathbf{x}_i \big) \leq \sum_{i=1}^{k} t_i \, f(\mathbf{x}_i)$, y portanto f es convexa. $\square$

Funciones Convexas

Observaciones:

- Directamente de la definición, tenemos que f es convexa $f\big|_{[\mathbf{a},\mathbf{b}]}$ es convexa, cuando se restringe a cualquier segmento $[\mathbf{a},\mathbf{b}]$, con $\mathbf{a},\mathbf{b} \in \Omega$.
  De ahí que f es convexa para todo $\mathbf{x} \in \Omega$, y para todo $\mathbf{h} \in \mathbb{R}^n$, la función $g(t) = (\mathbf{x} + t\mathbf{h})$ es convexa en el dominio $\{t \in \mathbb{R} : \ \mathbf{x} + t\mathbf{h} \in \Omega\}$.

- En ocasiones conviene extender una función convexa $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ a valores en la recta extendida $\widehat{f} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, por

$$\widehat{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \text{si } \mathbf{x} \in \Omega; \\ +\infty, & \text{si } \mathbf{x} \notin \Omega. \end{cases}$$

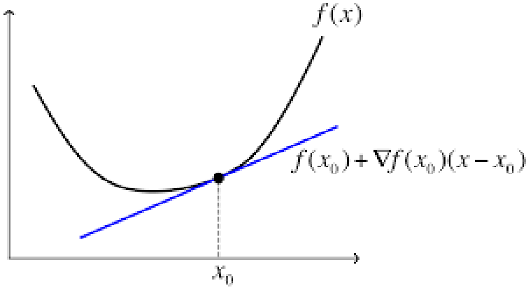
  Claramente $\widehat{f}\big|_{\Omega} = f$, y se tiene que f es convexa $\iff \widehat{f}$ es convexa.

- En el caso de $\widehat{f}$, la convexidad sigue siendo definida por la desigualdad (1), con la diferencia que se usa aritmética extendida.

Teorema (Condición de 1er Orden)

Suponga que $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable y que Ω es convexo. Entonces f es convexa si, y sólo si, para todo $\mathbf{x}, \mathbf{x}_0 \in \Omega$ vale

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0).$$



Si f es convexa, el plano tangente a f en $\mathbf{x}_0$ está por debajo del grafo de f.

Condiciones de Optimalidad

<u>Prueba</u>: ($\Rightarrow$) Como Ω es convexo, para $\mathbf{x}, \mathbf{x}_0 \in \Omega$ se tiene que $(1-t)\mathbf{x}_0 + t\mathbf{x} \in \Omega$, $\forall 0 \leq t \leq 1$. Si f es convexa, de la desigualdad de Jensen (1), tenemos

$$f\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big) = f\big((1-t)\mathbf{x}_0 + t\mathbf{x}\big) \leq (1-t)f(\mathbf{x}_0) + tf(\mathbf{x}) = f(\mathbf{x}_0) + t\big(f(\mathbf{x}) - f(\mathbf{x}_0)\big),$$

para todo $0 < t < 1$. Luego $f\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big) - f(\mathbf{x}_0) \leq t\big(f(\mathbf{x}) - f(\mathbf{x}_0)\big)$, y

$$\frac{f\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big) - f(\mathbf{x}_0)}{t} \leq f(\mathbf{x}) - f(\mathbf{x}_0), \quad \text{para } \; 0 < t < 1.$$

Como f es diferenciable, tomando el límite cuando $t \to 0^+$, obtenemos

$$\nabla f(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) = \lim_{t \to 0^+} \frac{f\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big) - f(\mathbf{x}_0)}{t} \leq f(\mathbf{x}) - f(\mathbf{x}_0),$$

lo que produce $f(\mathbf{x}) \geq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0)$.

($\Leftarrow$) Tome $\mathbf{x}, \mathbf{x}_0 \in \Omega$, y sea $\mathbf{z} = (1-t)\mathbf{x}_0 + t\mathbf{x} \in \Omega$, con $0 \leq t \leq 1$. Por hipótesis

$$
\begin{aligned}
f(\mathbf{x}_0) &\geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{x}_0 - \mathbf{z}),\\
f(\mathbf{x}) &\geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{x} - \mathbf{z}).
\end{aligned}
$$

Haciendo una combinación convexa de ambas ecuaciones, resulta

$$
\begin{aligned}
(1-t)f(\mathbf{x}_0) + tf(\mathbf{x}) &\geq (1-t)\big[f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{x}_0 - \mathbf{z})\big] + t\big[f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{x} - \mathbf{z})\big]\\
&\geq \big[(1-t)+t\big]f(\mathbf{z}) + \nabla f(\mathbf{z})^T\big[(1-t)(\mathbf{x}_0 - \mathbf{z}) + t(\mathbf{x} - \mathbf{z})\big]\\
&\geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T\big[\underbrace{(1-t)\mathbf{x}_0 + t\mathbf{x} - \mathbf{z}}_{=0}\big] \;=\; f(\mathbf{z})\\
&\geq f\big((1-t)\mathbf{x}_0 + t\mathbf{x}\big).
\end{aligned}
$$

Dado que $\mathbf{x}, \mathbf{x}_0$ son arbitrarios, esto muestra que f es convexa. $\square$

Corolario

Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable y convexa. Si $\nabla f(\mathbf{x}^*) = \mathbf{0}$, para $\mathbf{x}^* \in \Omega$, entonces $\mathbf{x}^*$ es un mínimo global de f.

Condiciones de Optimalidad

Prueba: El teorema anterior implica que $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) = f(\mathbf{x}^*)$, $\forall \mathbf{x} \in \Omega$. Portanto, $\mathbf{x}^*$ es mínimo global de f. $\square$

Teorema (Condición de 2do Orden)

Suponga que $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ es dos veces diferenciable y que Ω es abierto y convexo. Entonces f es convexa si, y sólo si, para todo $\mathbf{x} \in \Omega$, $D^2 f(\mathbf{x}) \succeq 0$.

Prueba: ($\Leftarrow$) Suponga que $D^2 f(\mathbf{x})$ es positiva semidefinida, para todo $\mathbf{x} \in \Omega$. Tomemos $\mathbf{x}, \mathbf{x}_0 \in \Omega$ y definamos $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$. De la Fórmula de Taylor

$$f(\mathbf{x}) = f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T\mathbf{h} + \tfrac{1}{2}\mathbf{h}^T D^2 f(\mathbf{x}_0 + t\mathbf{h})\,\mathbf{h},$$

para algún $0 < t < 1$.
Peero $\mathbf{x}_0 + t\mathbf{h} = \mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0) = (1-t)\mathbf{x}_0 + t\mathbf{x} \in \Omega$, ya que es una conbinación convexa de $\mathbf{x}, \mathbf{x}_0 \in \Omega$. Esto implica que el término $\frac{1}{2}\mathbf{h}^T D^2 f(\mathbf{x}_0 + t\mathbf{h})\,\mathbf{h} \geq 0$, de modo que

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T(\mathbf{x}_{\mathbf{x}0}).$$

De la Condición de optimalidad de 1er. orden, f es convexa.

Condiciones de Optimalidad

($\Rightarrow$) Suponga ahora que f es convexa. Tome $\mathbf{x} \in \Omega$, $\mathbf{p} \in \mathbb{R}^n$. Como Ω es abierto, existe $r > 0$ tal que $\mathbb{D}_r(\mathbf{x}) \subseteq \Omega$. Tomamons $\widetilde{\mathbf{p}} = \frac{1}{k}\mathbf{p}$ un múltiplo suficientemente pequeño de $\mathbf{p}$, de modo que $\mathbf{y} = \mathbf{x} + \widetilde{\mathbf{p}} \in \mathbb{D}_r(\mathbf{x})$. Como f es convexa,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T\, \widetilde{\mathbf{p}}.$$

Por el Teorema de Taylor,

$$f(\mathbf{y}) = f(\mathbf{x} + \widetilde{\mathbf{p}}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T\, \widetilde{\mathbf{p}} + \tfrac{1}{2}\widetilde{\mathbf{p}}^T\, D^2 f(\mathbf{x} + t\widetilde{\mathbf{p}})\, \widetilde{\mathbf{p}}, \quad t \in (0,1).$$

Combinando las dos expresiones anteriores, obtenemos

$$f(\mathbf{x}) + \nabla f(\mathbf{x})^T\, \widetilde{\mathbf{p}} + \tfrac{1}{2}\widetilde{\mathbf{p}}^T\, D^2 f(\mathbf{x} + t\widetilde{\mathbf{p}})\, \widetilde{\mathbf{p}}, \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T\, \widetilde{\mathbf{p}}.$$

$\Rightarrow\ \ \widetilde{\mathbf{p}}^T\, D^2 f(\mathbf{x} + t\widetilde{\mathbf{p}})\, \widetilde{\mathbf{p}} \geq 0.$
Haciendo $t \to 0$, se obtiene que $\widetilde{\mathbf{p}}^T\, D^2 f(\mathbf{x})\, \widetilde{\mathbf{p}} \geq 0$. Y como esto vale para todo $\widetilde{\mathbf{p}} \in \mathbb{D}_r(\mathbf{x})$, se tiene que $D^2 f(\mathbf{x}) \succeq 0$ es positiva semidefinida. $\square$

Condiciones de Optimalidad

Corolario

Si $f : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}$ es convexa y 2-veces diferenciable, cualquier punto estacionario $\mathbf{x}^* \in \Omega$ de f, es un mínimo global.

Prueba: Como $\mathbf{x}^*$ es punto estacionario de f, entonces $\nabla f(\mathbf{x}^*) = \mathbf{0}$. Además, como f es convexa y 2-veces diferenciable, entonces $D^2 f(\mathbf{x}^*) \succeq 0$. De las condiciones de obtimalidad, se obtiene que para todo $\mathbf{x} \in \Omega$

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) + \tfrac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T D^2 f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \geq f(\mathbf{x}^*), \quad \forall \mathbf{x} \in \Omega.$$

Portanto, $\mathbf{x}^*$ es mínimo global de f. $\square$

Proposición

Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable, Ω convexo. Entonces f es convexa $\iff$ $(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y}))^T (\mathbf{x} - \mathbf{y}) \geq 0$, $\forall \mathbf{x}, \mathbf{y} \in \Omega$.

<u>Prueba</u>: Como f es convexa, de la condición de 1er orden tenemos

$$\left.\begin{aligned} f(\mathbf{y}) &\geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x}) \\ f(\mathbf{x}) &\geq f(\mathbf{y}) + \nabla f(\mathbf{y})^T(\mathbf{x}-\mathbf{y}) \end{aligned}\right\} \quad \Rightarrow \quad 0 \geq \big(\nabla f(\mathbf{x})^T - \nabla f(\mathbf{y})^T\big)(\mathbf{y}-\mathbf{x})$$

Portanto, $\big(\nabla f(\mathbf{x})^T - \nabla f(\mathbf{y})^T\big)(\mathbf{x}-\mathbf{y}) \geq 0$.

$(\Rightarrow)$ (pendiente).

Proposición

Si $f : \mathbb{R}^n \to \mathbb{R}$ es diferenciable y convexa, entonces $\mathbf{x}^*$ es un óptimo global de f $\iff$ $\nabla f(\mathbf{x}^*)^T(\mathbf{x}-\mathbf{x}^*) =\geq 0$, para todo $\mathbf{x} \in \mathbb{R}^n$.

<u>Prueba</u>: Haga $\mathbf{x} = -\nabla f(\mathbf{x}^*)^T + \mathbf{x}^*$. Entonces, $\nabla f(\mathbf{x}^*)^T(\mathbf{x}-\mathbf{x}^*) = -||\nabla f(\mathbf{x}^*)||^2 \leq 0$. De ahí que $\nabla f(\mathbf{x}^*)^T(\mathbf{x}-\mathbf{x}^*) \geq 0$, $\forall \mathbf{x} \in \mathbb{R}^n$ implica que $\nabla f(\mathbf{x}^*)^T = \mathbf{0}$, y $\mathbf{x}^*$ es un punto crítico. La convexidad de f implica entonces que $\mathbf{x}^*$ es un mínimo global. $\square$

Ejemplos

Ejemplos de funciones convexas en $\mathbb{R}$:

- (Exponencial): $f(x) = e^{ax}$ es convexa en todo $\mathbb{R}$, para todo $a \in \mathbb{R}$.
  Basta ver que $f'(x) = ae^{ax}$ y $f''(x) = a^2 e^{ax} \geq 0$. Luego f es convexa.
  (De hecho, f es estrictamente convexa para $a \neq 0$).

- (Potencias): $f(x) = x^a$ es estrictamente convexa sobre $\mathbb{R}^+$, para $a \geq 1$ o $a \leq 0$.
  Basta ver que $f'(x) = ax^{a-1}$ y $f''(x) = a(a-1)x^{a-2} > 0$, cuando $a < 0$ ó $a > 1$.

- (Potencias del valor absoluto): Las funciones $f(x) = |x|^p$ son convexas para $p \geq 1$.

  $f'(x) = p|x|^{p-1} \cdot \frac{d}{dx}|x| = p|x|^{p-1} \cdot \frac{x}{|x|} = p\,x|x|^{p-2}.$

$$\begin{aligned} f''(x) &= p|x|^{p-2} + p(p-2)\,x|x|^{p-3} \cdot \tfrac{d}{dx}|x| \;=\; p|x|^{p-2} + p(p-2)\,x^2|x|^{p-4} \cdot \tfrac{x}{|x|} \\ &= p\,|x|^{p-4}\big(|x|^2 + (p-2)x^2\big) \;=\; p\,|x|^{p-4}\big(x^2 + (p-2)x^2\big) \\ &= p(p-1)\,x^2|x|^{p-4} \;\geq\; 0 \end{aligned}$$

  cuando $p \geq 1$.

Ejemplos

- (Logaritmo negativo): $f(x) = -\log(x)$ es convexa en todo $\mathbb{R}^+$, para todo $a \in \mathbb{R}$. Basta ver que $f'(x) = -\frac{1}{x}$ y $f''(x) = \frac{1}{x^2} > 0$. Luego f es estrictamente convexa.
- (Entropía): $f(x) = x\log(x)$ es estrictamente convexa sobre $\mathbb{R}^+$. Observe que $f'(x) = \log(x) + 1$ y $f''(x) = \frac{1}{x} > 0$.
- (Funciones lienales): Las funciones $f(x) = ax + b$ son siempre convexas y cóncavas, $\forall a, b \in \mathbb{R}$

Ejemplos

Ejemplos de funciones convexas en $\mathbb{R}^n$:

- (Normas): Toda norma en $\mathbb{R}^n$ es convexa.
  De la homogeneidad y la desigualdad triangular, tenemos
  $$||(1-t)\mathbf{x} + t\mathbf{y}|| \leq ||(1-t)\mathbf{x}|| + ||t\mathbf{y}|| = (1-t)||\mathbf{x}|| + t||\mathbf{y}||, \quad \text{para } t \in [0,1].$$
- (Máximos): La función $f(\mathbf{x}) = \max_{1 \leq i \leq n}\{x_i\}$ es convexa.
  Recuerde que $\max_i\{a_i + b_i\} \leq \max_i\{a_i\} + \max_i\{b_i\}$. (¿Por qué?)
  De ahí que
  $$\max_i \big((1-t)x_i + ty_i\big) \leq \max_i (1-t)x_i + \max_i ty_i = (1-t)\max_i x_i + t\,\max_i y_i.$$
- (Log-sum-exp): La función $f(\mathbf{x}) = \log\Big(\sum_{i=1}^{n} e^{x_i}\Big)$ es convexa en $\mathbb{R}^n$. (Hay que calcular la Hessiana y mostrar que es $\succeq 0$ + Cauchy-Schwarz).

Ejemplos

- (Media geométrica): La función $f(\mathbf{x}) = \Big(\prod_{i=1}^{n} x_i\Big)^{1/n}$ es convexa en $\mathbb{R}^n$.

  (De nuevo, calcular la Hessiana y mostrar que es $\succeq 0$ + Cauchy-Schwarz).

- (Log-det): El logaritmo negativo del determinante $f : \mathbb{R}^{n\times n} \to \mathbb{R}$, dado por $f(X) = -\log\det X$ es convexa en el conjunto de matrices positivas definidas en $\mathbb{R}^{n\times n}$.