11 FUNCIONES CONVEXAS

11.1 DEFINICIÓN

Sea I un intervalo y sean a, $b \in I$ con $a < b$. Si $x \in [a, b]$, x se puede escribir como *combinación convexa* de a y b, esto es,

$$x = (1 - t)a + tb$$

para algún $t \in [0, 1]$. Podemos calcular el valor de t:

$$x = (1 - t)a + tb \iff t = \frac{x - a}{b - a}$$

con lo que

$$x = \frac{b - x}{b - a} \cdot a + \frac{x - a}{b - a} \cdot b.$$

Definición 11.1.1. Sea I un intervalo y sea $f \colon I \to \mathbb{R}$.

1) La función f es *convexa* si para cualesquiera a, $b \in I$ con $a < b$ y cualquier $t \in [0, 1]$ se cumple que

$$f\big((1 - t)a + tb\big) \leqslant (1 - t)f(a) + tf(b).$$

2) La función f es *estrictamente convexa* si para cualesquiera a, $b \in I$ con $a < b$ y cualquier $t \in]0, 1[$, se cumple que

$$f\big((1 - t)a + tb\big) < (1 - t)f(a) + tf(b).$$

3) La función f es *cóncava* (resp. estrictamente cóncava) si $-f$ es convexa (resp. estrictamente cóncava).

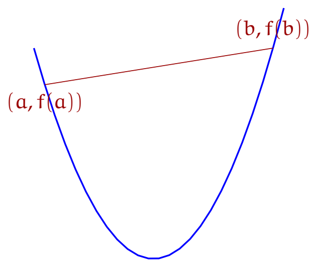


Figura 29.: El segmento que une dos puntos arbitrarios de la gráfica de una función convexa está por encima o coincide con la gráfica de la función

Geométricamente, la convexidad se traduce el segmento que une las imágenes de dos puntos arbitrarios está por encima de la gráfica de la función.

Observación 11.1.2. En la definición de función convexa no es necesario suponer que $a < b$. Si $b < a$, los papeles de t y $1 - t$ se intercambian, pero la definición no cambia. Por tanto, f es convexa si, y sólo si,

$$f((1 - t)a + tb) \leqslant (1 - t)f(a) + tf(b), \quad \forall\, a, b \in I, \ \forall\, t \in [0, 1].$$

Ejemplo 11.1.3. 1) Cualquier función afín, $f(x) = mx + n$, es convexa (más detalles en el ejercicio 11.6).

2) La función valor absoluto es convexa: si x, $y \in \mathbb{R}$ y $t \in [0, 1]$, por la desigualdad triangular

$$|(1 - t)x + ty| \leqslant (1 - t)\,|x| + t\,|y|\,.$$

No es estrictamente convexa: si x, $y > 0$, entonces

$$|(1 - t)x + ty| = (1 - t)\,|x| + t\,|y|$$

para cualquier $t \in]0, 1[$.

3) La función $f(x) = x^2$ es convexa en todo $\mathbb{R}$. En efecto, si a, $b \in \mathbb{R}$ y $t \in [0,1]$,

$$\begin{aligned} f\left((1-t)a+tb\right) &= (1-t)^2a^2+t^2b^2+2t(1-t)ab, \\ (1-t)f(a)+tf(b) &= (1-t)a^2+tb^2. \end{aligned}$$

Para ver cuál es mayor, calculamos la diferencia:

$$\begin{aligned} &(1-t)f(a)+tf(b)-\big(f\left((1-t)a+tb\right)\big) \\ &\qquad = (1-t)a^2+tb^2-\Big((1-t)^2a^2+t^2b^2+2t(1-t)ab\Big) \\ &\qquad = t(1-t)a^2+t(1-t)b^2-2t(1-t)ab \\ &\qquad = t(1-t)\left(a-b\right)^2 \geqslant 0. \end{aligned}$$

Por tanto, la función es convexa.

4) La función $f(x) = x^n$ es convexa en $[0,+\infty[$ para cualquier $n \in \mathbb{N}$.

Ejemplo 11.1.4. Es fácil dar ejemplos de funciones convexas que no son continuas: si aumentamos el valor de una función en los extremos del intervalo la convexidad se mantiene, pero la continuidad no. Por ejemplo, la función $f\colon [-1,1] \to \mathbb{R}$ definida como

$$f(x) = \begin{cases} x^2, & \text{si } x \in]-1,1[\\ 2, & \text{si } x = \pm 1 \end{cases}$$

es convexa y no es continua. Más adelante veremos que esta es la única forma de construir funciones convexas discontinuas.

11.1.1 El conjunto de las funciones convexas

Proposición 11.1.5. *Sean* f, $g\colon I \to \mathbb{R}$ *dos funciones convexas y sea* $\alpha > 0$. *Entonces*

1) $f+g$ *es convexa,*

2) αf *es convexa, y*

3) $\max\{f,g\}$ *es convexa.*

Demostración. 1) Sean a, $b \in I$, y sea $t \in [0,1]$, entonces

$$\begin{aligned} (f+g)\left((1-t)a+tb\right) &= f\left((1-t)a+tb\right)+g\left((1-t)a+tb\right) \\ &\leqslant (1-t)f(a)+tf(b)+(1-t)g(a)+tg(b) \\ &= (1-t)(f+g)(a)+t(f+g)(b). \end{aligned}$$

2) Sean a, $b \in I$, y sea $t \in [0,1]$, entonces

$$\begin{aligned} (\alpha f)\left((1-t)a+tb\right) &= \alpha f\left((1-t)a+tb\right) \\ &\leqslant \alpha(1-t)f(a)+\alpha t f(b) \\ &= (1-t)(\alpha f)(a)+t(\alpha f)(b). \end{aligned}$$

3) Sean a, $b \in I$, y sea $t \in [0,1]$, entonces

$$\begin{aligned} \max\{f,g\}\left((1-t)a+tb\right) &= \max\left\{f\left((1-t)a+tb\right), g\left((1-t)a+tb\right)\right\} \\ &\leqslant \max\left\{(1-t)f(a)+tf(b), (1-t)g(a)+tg(b)\right\} \\ &= (1-t)\max\{f,g\}(a)+t\max\{f,g\}(b). \quad \square \end{aligned}$$

Observación 11.1.6. 1) El producto de funciones convexas puede no serlo. Por ejemplo, las funciones $f(x) = x$ y $g(x) = x^2$ son convexas y su producto no lo es.

2) El mínimo de dos funciones convexas puede no serlo. Por ejemplo,

$$\min\{x, -x\} = -|x|$$

no es una función convexa. Sí se cumple, cambiando signos, que el mínimo de dos funciones cóncavas es una función cóncava.

Incluso con ejemplos sencillos, la comprobación de la convexidad puede ser complicada. Las caracterizaciones de la convexidad para funciones derivables o dos veces derivables serán de gran utilidad en la práctica.

11.1.2 Convexidad y acotación

Aunque no sean continuas, sí podemos decir algo sobre la acotación.

Proposición 11.1.7. *Sea* $f\colon [a,b] \to \mathbb{R}$ *una función convexa. Entonces* f *está acotada.*

Demostración. Sea $M = \max\{f(a), f(b)\}$. Si $x = (1-t)a + tb \in [a,b]$, entonces

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

Veamos la acotación inferior. Podemos escribir cualquier elemento $x \in [a,b]$ de la forma $x = (a+b)/2 + y$, con $|y| \leqslant |b-a|/2$. Como

$$\frac{a+b}{2} = \frac{1}{2}\left(\frac{a+b}{2} + y\right) + \frac{1}{2}\left(\frac{a+b}{2} - y\right),$$

usando que f es convexa, se tiene que

$$f\left(\frac{a+b}{2}\right) \leqslant \frac{1}{2} f\left(\frac{a+b}{2} + y\right) + \frac{1}{2} f\left(\frac{a+b}{2} - y\right).$$

Despejando en la desigualdad anterior,

$$\begin{aligned} f(x) = f\left(\frac{a+b}{2} + y\right) &\geqslant 2f\left(\frac{a+b}{2}\right) - f\left(\frac{a+b}{2} - y\right) \\ &\geqslant 2f\left(\frac{a+b}{2}\right) - M, \end{aligned}$$

obtenemos una cota inferior como queríamos. □

En realidad, hemos demostrado un poco más: hemos demostrado que la función f tiene máximo absoluto y este se alcanza en a o en b. ¿Tiene mínimo absoluto?

Ejemplo 11.1.8. La función $f\colon [0,1] \to \mathbb{R}$ definida como

$$f(x) = \begin{cases} 1, & \text{si } x = 0 \\ x^2, & \text{si } x \in]0,1] \end{cases}$$

es convexa, pero no tiene mínimo absoluto aunque está acotada inferiormente.

11.2 CONTINUIDAD Y DERIVABILIDAD DE FUNCIONES CONVEXAS

La convexidad tiene consecuencias sobre el comportamiento de la función. Veremos que, salvo en los extremos del intervalo, es suficiente para obtener continuidad. De hecho, veremos un poco más, la convexidad conduce a la existencia de derivadas laterales en todos los puntos del interior del intervalo.

Lema 11.2.1 (de las tres secantes (L. Galvani))**.** *Sea* I *un intervalo y sea* $f\colon I \to \mathbb{R}$ *una función convexa. Sean* $x_1, x_2, x_3 \in I$ *tales que* $x_1 < x_2 < x_3$*. Se cumple que*

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leqslant \frac{f(x_3)-f(x_1)}{x_3-x_1} \leqslant \frac{f(x_3)-f(x_2)}{x_3-x_2}.$$

Demostración. Comenzamos con la primera desigualdad:

$$\begin{aligned}
\frac{f(x_2)-f(x_1)}{x_2-x_1} \leqslant \frac{f(x_3)-f(x_1)}{x_3-x_1} &\iff f(x_2)-f(x_1) \leqslant (x_2-x_1)\,\frac{f(x_3)-f(x_1)}{x_3-x_1}\\
&\iff f(x_2) \leqslant f(x_1) + (x_2-x_1)\,\frac{f(x_3)-f(x_1)}{x_3-x_1}\\
&= \frac{f(x_1)(x_3-x_1) + (x_2-x_1)\,(f(x_3)-f(x_1))}{x_3-x_1}\\
&= \frac{x_3-x_2}{x_3-x_1}\cdot f(x_1) + \frac{x_2-x_1}{x_3-x_1}\cdot f(x_3)
\end{aligned}$$

y está última desigualdad es cierta aplicando la definición de convexidad a la terna x_1, x_2, x_3.

La segunda desigualdad se puede probar de la misma forma. □

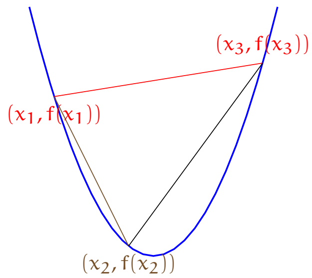


Figura 30.: Fijado uno cualquiera de los extremos, la pendiente de las rectas secantes es una función creciente

Lema 11.2.2. *Sea* I *un intervalo,* $a \in I$ *y* $f\colon I \to \mathbb{R}$ *una función convexa. Entonces la función* $f_a\colon I \setminus \{a\} \to \mathbb{R}$ *definida por*

$$f_a(x) = \frac{f(x)-f(a)}{x-a}, \quad \forall x \in I \setminus \{a\},$$

es creciente.

Demostración. Sean $x, y \in I \setminus a$ con $x < y$. Para comprobar que $f_a(x) \leqslant f_a(y)$ sólo hace falta aplicar el lema anterior tomando como x_1, x_2 y x_3, los puntos x, y y a en el orden que corresponda:

- si $x < y < a$, tomamos $x_1 = x$, $x_2 = y$, $x_3 = a$;
- si $x < a < y$, tomamos $x_1 = x$, $x_2 = a$, $x_3 = y$;
- si $a < x < y$, tomamos $x_1 = a$, $x_2 = x$, $x_3 = y$. □

Teorema 11.2.3 (O. Stolz)**.** *Sea* I *un intervalo abierto y sea* $f\colon I \to$ *una función convexa.*

1) *La función* f *tiene derivadas laterales en todos los puntos. En particular,* f *es continua.*

2) f'_+ *y* f'_- *son funciones crecientes.*

Demostración. Si $a \in I$, fijemos $u \in I$ con $a < u$, entonces

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(u)-f(a)}{u-a}, \quad (x \in I, x < a)$$

con lo que f_a es una función acotada superiormente. Por el lema 11.2.2, sabemos que f_a es creciente. Usando el teorema 5.8.2, se obtiene que f tiene derivadas laterales en todos los puntos de I. por lo que existe

$$\lim_{x\to a^-} \frac{f(x)-f(a)}{x-a} = \sup\left\{\frac{f(x)-f(a)}{x-a} : x \in I, x < a\right\} \leqslant \frac{f(u)-f(a)}{u-a}$$

para cualquier $u > a$. De forma análoga se obtiene que existe el límite por la derecha. En general, si $x < y$, obtenemos que

$$f'_-(x) \leqslant f'_+(x) \leqslant f'_-(y) \leqslant f'_+(y).$$ □

Corolario 11.2.4. *Sea* I *un intervalo y sea* $f\colon I \to \mathbb{R}$ *una función convexa. El conjunto de puntos donde* f *no es derivable es numerable.*

Demostración. f es derivable en un punto $a \in I$ si, y sólo si, $f'_-(a) = f'_+(a)$. Podemos razonar como en el teorema 5.8.3. Si $a \neq b$,

$$]f'_-(a), f'_+(a)[\cap]f'_-(b), f'_+(b)[= \emptyset,$$

con lo que eligiendo un número racional en cada uno de los intervalos, tenemos definida una aplicación inyectiva del conjunto de puntos donde la aplicación no es derivable en $\mathbb{Q}$ y, por tanto, es numerable. □

Corolario 11.2.5. *Sea* I *un intervalo y sea* $f\colon I \to \mathbb{R}$ *una función convexa. Supongamos que* f *es derivable, entonces* f' *es continua o, lo que es lo mismo,* $f \in C^1(I)$.

Demostración. f' es una función monótona creciente y su imagen, por el teorema del valor intermedio para las derivadas, es un intervalo. Por el teorema 4.3.5, f' es continua. □

11.3 CARACTERIZACIONES DE LA CONVEXIDAD

Teorema 11.3.1. *Sea* I *un intervalo y* $f\colon I \to \mathbb{R}$ *una función derivable. Las siguientes afirmaciones son equivalentes.*

1) f *es convexa.*

2) f' *es creciente.*

3) *Se cumple que* $f(x) \geqslant f(a) + f'(a)(x-a)$ *para todo* $a, x \in I$.

Demostración. 1) $\Rightarrow$ 2) Sean $a, b \in I$ con $a < b$. Tenemos que probar que $f'(a) \leqslant f'(b)$. Elegimos $x \in]a, b[$. Entonces

$$f'(a) \leqslant f_a(x) = f_x(a) \leqslant f_x(b) = f_b(x) \leqslant f'(b).$$

2) $\Rightarrow$ 3) Distinguimos dos casos: $x < a$ y $x > a$. En ambos casos, el teorema del valor medio nos dice que existe c entre a y x tal que

$$f(x) = f(a) + f'(c)(x-a).$$

Usando que f' es creciente se obtiene la desigualdad buscada.

3) $\Rightarrow$ 1) Sean $x, y \in I$, $t \in [0,1]$ y sea $a = (1-t)x + ty$. Tenemos que demostrar que
$$f(a) \leqslant (1-t)f(x) + tf(y).$$
Para ello usamos el apartado anterior aplicado a x e y:
$$\begin{aligned}(1-t)f(x) + tf(y) &\geqslant (1-t)\left(f(a) + f'(a)(x-a)\right)\\ &\quad + t\left(f(a) + f'(a)(y-a)\right)\\ &= f(a)\,(1-t+t)\\ &\quad + f'(a)\,((1-t)(x-a) + t(y-a))\\ &= f(a).\end{aligned}$$
□

Corolario 11.3.2. *Sea* I *un intervalo y* $f\colon I \to \mathbb{R}$ *una función convexa dos veces derivable. Entonces* f *es convexa si, y sólo si,* $f''(x) \geqslant 0$ *para todo* $x \in I$.

Demostración (1.ª versión). Es consecuencia directa del teorema 11.3.1. □

Demostración (2.ª versión). Podemos suponer sin pérdida de generalidad que el intervalo I es abierto.

Supongamos en primer lugar que f es convexa. Si $a \in \mathring{I}$, por el ejercicio 10.1
$$f''(a) = \lim_{h\to 0} \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}.$$
Como $a = \frac{1}{2}(a+h) + \frac{1}{2}(a-h)$, por la convexidad se tiene que
$$f(a) \leqslant \frac{1}{2}f(a+h) + \frac{1}{2}f(a-h),$$
y, por tanto, $f(a+h) - 2f(a) + f(a-h) \geqslant 0$. Como $h^2 > 0$, se tiene que $f''(a) \geqslant 0$.

En el otro sentido, supongamos que $f''(x) \geqslant 0$, para cualquier $x \in I$. Sean $x_1, x_2 \in I$ con $x_1 < x_2$ y $t \in]0,1[$. Si $x_0 = (1-t)x_1 + tx_2$, la fórmula de Taylor con resto de Lagrange aplicada la función f y $a = x_0$, nos dice que existen $c_1 \in]x_1, x_0[$ y $c_2 \in]0, x_2[$ tales que
$$\begin{aligned}f(x_1) &= f(x_0) + f'(x_0)(x_1 - x_0) + \frac{1}{2}f''(c_1)(x_1-x_0)^2, \quad \text{y}\\ f(x_2) &= f(x_0) + f'(x_0)(x_2 - x_0) + \frac{1}{2}f''(c_2)(x_2-x_0)^2.\end{aligned}$$
Como $f''(x) \geqslant 0$, para cualquier $x \in I$,
$$\begin{aligned}(1-t)f(x_1) + tf(x_2) &= f(x_0) + f'\left((1-t)x_1 + tx_2 - x_0\right)\\ &\quad + \frac{1}{2}(1-t)f''(c_1)(x_1-x_0)^2 + \frac{1}{2}tf''(c_2)(x_2-x_0)^2\\ &\geqslant f(x_0) = f\left((1-t)x_1 + tx_2\right).\end{aligned}$$
Por tanto, f es convexa en I. □

Ejemplos

Usando la caracterización de la convexidad en términos del signo de la segunda derivada es sencillo estudiar la convexidad de las funciones elementales.

1) Las funciones afines, polinomios de grado uno, son funciones convexas y cóncavas (véase el ejercicio 11.6).

2) La función exponencial es convexa y el logaritmo es cóncava.

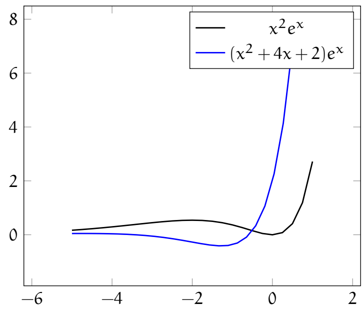
3) Si $\alpha \in \mathbb{R}$, la función $f\colon \mathbb{R}^+ \to \mathbb{R}$ definida como $f(x) = x^\alpha$, $f'(x) = \alpha x^{\alpha-1}$, $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$. Por tanto, es convexa si $\alpha(\alpha-1) \geqslant 0$ lo que ocurre si $\alpha \leqslant 0$ o $\alpha \geqslant 1$.

4) $f(x) = x^3$ es convexa en $[0, +\infty[$ y cóncava en $]-\infty, 0]$.

5) $f(x) = \arctan(x)$, $f''(x) = -2x/(1+x^2)^2$. La función arcotangente es convexa en $]-\infty, 0]$ y cóncava en $[0, +\infty[$.

Ejemplo 11.3.3. Para estudiar la concavidad y convexidad de la función $f\colon \mathbb{R} \to \mathbb{R}$ definida por $f(x) = x^2e^x$, estudiamos el signo de la segunda derivada.

$$f'(x) = 2xe^x + x^2e^x,$$
$$f''(x) = e^x\left(x^2 + 4x + 2\right).$$



Las raíces de la segunda derivada son

$$f''(x) = 0 \iff x^2 + 4x + 2 = 0 \iff x = -2 \pm \sqrt{2}.$$

Como la segunda derivada es continua, podemos averiguar el signo evaluando en un punto arbitrario: por ejemplo, $f''(4) > 0$, $f''(-1) < 0$ y $f''(-10) > 0$ de donde se deduce que

- f es convexa en $]-2+\sqrt{2}, +\infty[$,
- f es cóncava en $]-2-\sqrt{2}, -2+\sqrt{2}[$, y
- f es convexa en $]-\infty, -2-\sqrt{2}[$.

11.3.1 Puntos de inflexión*

Aunque no hay acuerdo total sobre la definición de punto de inflexión, parece que la más popular es un punto donde la función cambia de cóncava a convexa o viceversa. Hay definiciones más restrictivas que exigen derivabilidad. También hay variantes geométricas: un punto es de inflexión si la recta tangente corta a la gráfica de la función. No son definiciones equivalentes.

Definición 11.3.4. Sea I un intervalo y sea $f\colon I \to \mathbb{R}$ una función. Diremos que $a \in \mathring{I}$ es un *punto de inflexión* de la función f si existe $r > 0$ tal que

1) f es convexa en $]a-r, a]$ y cóncava en $[a, a+r[$, o

2) f es cóncava en $]a-r, a]$ y convexa en $[a, a+r[$.

Proposición 11.3.5. *Sea* I *un intervalo y* $f\colon I \to \mathbb{R}$ *una función de clase* $C^2(I)$. *Si* $a \in I$ *es un punto de inflexión de* f, *entonces* $f''(a) = 0$.

Demostración. Por el corolario 11.3.2, la segunda derivada tiene es mayor o igual que cero a un lado de a y menor o igual que cero al otro lado. Por continuidad se deduce que $f''(a) = 0$. □

Ejemplo 11.3.6. El origen es un punto de inflexión de la función $f(x) = x^3$.

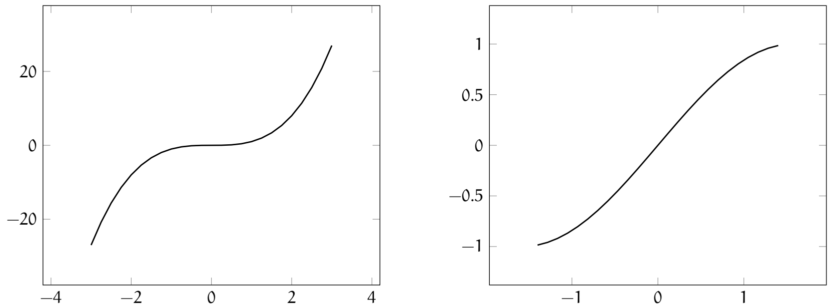


Figura 31.: Las funciones x^3 y $\operatorname{sen}(x)$ tiene un punto de inflexión en el origen

Observación 11.3.7. Si la derivada de una función f se anula en un punto a, pero en dicho punto no se alcanza un extremo relativo ¿tiene la función f un punto de inflexión en a? La respuesta es no. La función $f\colon \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} x^6 \operatorname{sen}\left(\frac{1}{x}\right), & \text{si } x \neq 0, \\ 0, & \text{si } x = 0, \end{cases}$$

verifica que:

1) f es impar y $f(0) = 0$ por lo que no tiene un extremo relativo en 0.

2) La dos primeras derivadas de f son

$$f'(x) = \begin{cases} 6x^5 \operatorname{sen}\left(\frac{1}{x}\right) - x^4 \cos\left(\frac{1}{x}\right), & \text{si } x \neq 0, \\ 0, & \text{si } x = 0, \end{cases}$$

$$f''(x) = \begin{cases} 30x^4 \operatorname{sen}\left(\frac{1}{x}\right) - 10x^3 \cos\left(\frac{1}{x}\right) + x^2 \operatorname{sen}\left(\frac{1}{x}\right), & \text{si } x \neq 0, \\ 0, & \text{si } x = 0. \end{cases}$$

Como $f\left(\frac{1}{n\pi}\right) = \frac{(-1)^{n+1}10}{(n\pi)^3}$, se cumple que no hay ningún entorno de cero donde la segunda derivada tengo signo constante y, por tanto, la función no tiene un punto de inflexión en cero.

11.4 EJERCICIOS

Ejercicio 11.1. Sea $f\colon I \to \mathbb{R}$ una función convexa.

1) Si f tiene un mínimo local en un punto, entonces dicho mínimo es absoluto.

2) Supongamos que f es derivable. Si x_0 es un punto crítico de f, entonces f alcanza su mínimo absoluto en x_0. ¿Es cierto el mismo resultado si sólo suponemos que f es derivable en el punto crítico?

3) Si f es estrictamente convexa, f tiene a lo sumo un mínimo absoluto.