Tema 14

Funciones convexas

Los resultados obtenidos en el desarrollo del cálculo diferencial nos permiten estudiar con facilidad una importante familia de funciones definidas en intervalos, las *funciones convexas*. Haremos una discusión breve de estas funciones, empezando por su definición, que se basa en una sencilla idea geométrica. Para funciones que sean derivables, o dos veces derivables, en un intervalo, obtendremos útiles caracterizaciones de la convexidad.

14.1. Definición de función convexa

La noción de función convexa es muy intuitiva y fácil de entender geométricamente. Si I es un intervalo no trivial, una función $f : I \to \mathbb{R}$ será convexa cuando la gráfica de la restricción de f a cualquier intervalo cerrado y acotado $[a,b] \subset I$, queda siempre por debajo del segmento que une los puntos $\big(a, f(a)\big)$ y $\big(b, f(b)\big)$. Enseguida convertimos esta sencilla idea geométrica en una definición concreta.

Para cualesquiera $a,b \in I$ con $a < b$, la recta secante a la gráfica de f que pasa por los puntos $\big(a, f(a)\big)$ y $\big(b, f(b)\big)$ tiene ecuación

$$y - f(a) = \frac{f(b)-f(a)}{b-a}\,(x-a)$$

Por tanto f será convexa cuando, para cualesquiera $a,b \in I$ con $a < b$, se tenga que

$$f(z) \leqslant f(a) + \frac{f(b)-f(a)}{b-a}\,(z-a) \quad \forall z \in [a,b]$$

o equivalentemente,

$$f(z) \leqslant \frac{b-z}{b-a} f(a) + \frac{z-a}{b-a} f(b) \quad \forall z \in [a,b] \tag{1}$$

Obsérvese que esta desigualdad es obvia para $z = a$ y para $z = b$, de hecho en ambos casos se da la igualdad. Bastaría por tanto exigirla para todo $z \in]a,b[$. Pero vamos a expresarla de forma más cómoda y fácil de recordar.

Dado $t \in [0,1]$ podemos tomar $z = a + t\,(b-a) \in [a,b]$ y (1) nos dice que

$$f\big((1-t)\,a + t\,b\big) \leqslant (1-t)\,f(a) + t\,f(b) \quad \forall t \in [0,1] \tag{2}$$

Pero recíprocamente, dado $z \in [a,b]$ podemos tomar

$$t = \frac{z-a}{b-a}\,, \quad \text{que verifica} \quad t \in [0,1] \quad \text{y} \quad 1-t = \frac{b-z}{b-a}$$

con lo que $(1-t)a+tb = z$ y, al aplicar (2), obtenemos directamente (1). Nótese que ahora, en (2) se da obviamente la igualdad para $t=0$ y $t=1$.

En resumen, f será convexa cuando verifique (2) para cualesquiera $a,b \in I$ con $a<b$ y para todo $t \in [0,1]$. Ahora bien, observamos que cuando $a=b$ se tiene siempre la igualdad en (2). Pero es que además, en el caso $a>b$, podemos aplicar (2) intercambiando los papeles de a y b, pero sustituyendo t por $1-t$, con lo que obtenemos exactamente la misma desigualdad. Así pues, en (2) podemos tomar como a y b dos puntos cualesquiera del intervalo I. Hemos llegado así a la definición cómoda de función convexa que buscábamos.

Si I es un intervalo no trivial, se dice que una función $f: I \to \mathbb{R}$ es *convexa* cuando verifica:

$$f\big((1-t)\,x + t\,y\big) \leqslant (1-t)\,f(x) + t\,f(y) \quad \forall x,y \in I\,,\ \forall t \in [0,1] \tag{3}$$

Aunque esta es la forma más conveniente de expresar la convexidad de una función, hay algo en la discusión anterior que no conviene olvidar: para probar (3), no se pierde generalidad suponiendo que $x<y$ y que $0<t<1$. Tampoco conviene olvidar la interpretación geométrica de la convexidad, con la que hemos iniciado la discusión.

Como ejemplo, la función valor absoluto es convexa, pues evidentemente,

$$\big|(1-t)\,x + t\,y\big| \leqslant (1-t)\,|x| + t\,|y| \quad \forall x,y \in \mathbb{R}\,,\ \forall t \in [0,1]$$

Conviene comentar que lo ocurrido en este ejemplo es poco usual. Rara vez se prueba que una función es convexa usando la definición, pues incluso para funciones muy sencillas, no suele ser fácil comprobar (3). Lo habitual es, sabiendo que una función es convexa gracias a alguna de las caracterizaciones que vamos a estudiar, usar (3) para obtener desigualdades nada evidentes.

Al cambiar de signo una función convexa se obtiene una función cóncava. Por tanto, si I es un intervalo no trivial, se dice que una función $f: I \to \mathbb{R}$ es *cóncava* cuando $-f$ es convexa, es decir, cuando se verifica que $f\big((1-t)x+ty\big) \geqslant (1-t)\,f(x) + t\,f(y)$ para cualesquiera $x,y \in I$ y para todo $t \in [0,1]$.

Obsérvese que la convexidad o concavidad de una función, como ocurre con la monotonía, es bastante restrictiva. Es frecuente que una función no sea convexa ni cóncava, pero que su restricción a un cierto intervalo, contenido en su intervalo de definición, sí tenga una de esas propiedades. De hecho, podemos también considerar funciones definidas en un conjunto A que no es un intervalo, pero que pueden restringirse a un intervalo $J \subset A$. Para tratar cómodamente ambas situaciones, si tenemos una función $f: A \to \mathbb{R}$ y un intervalo no trivial $J \subset A$, diremos que f es *convexa en* J cuando $f|_J$ sea convexa. Lógicamente, f será *cóncava en* J cuando $f|_J$ sea cóncava, es decir, cuando $-f$ sea convexa en J.

En lo sucesivo trabajaremos preferentemente con las funciones convexas, pues cualquier resultado que obtengamos dará información sobre una función cóncava, sin más que aplicarlo a la función opuesta.

14.2. Continuidad y derivabilidad

Para obtener propiedades importantes de las funciones convexas, conviene deducir de la definición de función convexa, o más directamente de la condición (1), una doble desigualdad:

Lema *(de las tres secantes). Sea I un intervalo no trivial y $f : I \to \mathbb{R}$ una función convexa. Entonces, para cualesquiera $x_1, x_2, x_3 \in I$, con $x_1 < x_2 < x_3$, se tiene:*

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leqslant \frac{f(x_3) - f(x_2)}{x_3 - x_2} \tag{3}$$

La demostración no tiene dificultad. Aplicamos (1) con $a = x_1$, $z = x_2$ y $b = x_3$ para obtener

$$f(x_2) \leqslant \frac{x_3 - x_2}{x_3 - x_1} f(x_1) + \frac{x_2 - x_1}{x_3 - x_1} f(x_3) \tag{4}$$

Ahora restamos $f(x_1)$ en ambos miembros:

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

y al dividir por $x_2 - x_1 > 0$ tenemos la primera desigualdad de (3). Para la otra, cambiamos de signo ambos miembros de (4), con lo que la desigualdad se invierte, y sumamos en ambos $f(x_3)$, obteniendo

$$f(x_3) - f(x_2) \geqslant \frac{x_2 - x_3}{x_3 - x_1} f(x_1) + \frac{x_3 - x_2}{x_3 - x_1} f(x_3) = (x_3 - x_2) \frac{f(x_3) - f(x_1)}{x_3 - x_1}$$

con lo que basta dividir por $x_3 - x_2 > 0$. ■

A poco que se piense, el lema anterior nos da una relación entre las pendientes de tres rectas secantes a la gráfica de una función convexa, que tiene una interpretación geométrica muy clara. La traducción analítica de esta idea es la propiedad clave de las funciones convexas que nos permitirá estudiar su derivabilidad:

- *Sea I un intervalo no trivial y $f : I \to \mathbb{R}$ una función convexa. Entonces, para cada $a \in I$ la función $f_a : I \setminus \{a\} \to \mathbb{R}$, dada por $f_a(x) = \dfrac{f(x) - f(a)}{x - a}$ para todo $x \in I \setminus \{a\}$, es creciente.*

En efecto, dados $a \in I$ y $x, y \in I \setminus \{a\}$ con $x < y$, distinguimos los tres casos posibles, para probar siempre que $f_a(x) \leqslant f_a(y)$. Si $a < x < y$, usamos la primera desigualdad de (3) con $x_1 = a$, $x_2 = x$, $x_3 = y$, obteniendo directamente que $f_a(x) \leqslant f_a(y)$. Si $x < y < a$, usamos la segunda desigualdad de (3) con $x_1 = x$, $x_2 = y$, $x_3 = a$, obteniendo la misma conclusión. Finalmente, si $x < a < y$, usamos la desigualdad entre el primer y último miembro de (3) con $x_1 = x$, $x_2 = a$, $x_3 = y$, obteniendo de nuevo $f_a(x) \leqslant f_a(y)$. ■

Es fácil ya conseguir el principal resultado sobre la derivabilidad de funciones convexas:

Teorema. *Sea I un intervalo no trivial y $f : I \to \mathbb{R}$ una función convexa. Entonces f es derivable por la izquierda y por la derecha, y por tanto es continua, en todo punto $a \in I^\circ$. De hecho, se tiene*

$$\begin{aligned} f'(a-) &= \sup\{f_a(x) : x \in I,\ x < a\} \\ f'(a+) &= \inf\{f_a(x) : x \in I,\ x > a\} \end{aligned} \tag{5}$$

Demostración. Fijado $a \in I^\circ$, la comprobación de (5) es bien sencilla, usando solamente que f_a es una función creciente. Tomando $b \in I$ con $b > a$, que existe porque $a \in I^\circ$, para $x \in I$ con $x < a$ se tiene $f_a(x) \leqslant f_a(b)$, luego el conjunto $\{f_a(x) : x \in I,\ x < a\}$ está mayorado y, llamando s_a a su supremo, veremos enseguida que $f'(a-) = s_a$.

Dado $\varepsilon > 0$, por definición de supremo existirá $x_0 \in I$ con $x_0 < a$ tal que $f_a(x_0) > s_a - \varepsilon$. Tomando $\delta = a - x_0 > 0$, tenemos

$$a - \delta < x < a \implies s_a - \varepsilon < f_a(x_0) \leqslant f_a(x) \leqslant s_a < s_a + \varepsilon$$

luego $\lim\limits_{x \to a-} f_a(x) = s_a$, como se quería. El cálculo de la derivada por la derecha es análogo. ■

Merece la pena resaltar que en general no podemos asegurar que una función convexa sea derivable en todos los puntos interiores de su intervalo de definición. Por ejemplo, la función valor absoluto es convexa pero no es derivable en 0. Cuando el intervalo de definición tiene mínimo o máximo, tampoco podemos asegurar que una función convexa sea continua en tales puntos. Por ejemplo, tomando $f(x) = 0$ para todo $x \in]0,1[$ y $f(0) = f(1) = 1$, obtenemos una función convexa $f : [0,1] \to \mathbb{R}$ que no es continua en 0 ni en 1.

Nótese también que, en la demostración del teorema anterior, sólo hemos usado que la función f_a es creciente. Por tanto, con la misma demostración, obtendríamos lo siguiente:

- *Sea I un intervalo no trivial y $g : I \to \mathbb{R}$ una función creciente. Entonces g tiene límite por la izquierda y por la derecha en todo punto $a \in I^\circ$.*

Obviamente, este resultado también es válido para funciones decrecientes, al igual que el teorema anterior tiene su correspondiente versión para funciones cóncavas.

14.3. Caracterizaciones de las funciones convexas

Para funciones derivables, vamos ya a obtener una útil caracterización de la convexidad, pero conviene aclarar un hecho que facilitará la demostración.

Sea I un intervalo no trivial y $f \in D^1(I)$ una función convexa. Para $a \in I^\circ$, la primera parte del teorema anterior no nos dice nada que no sepamos, pero las igualdades (5) nos dan dos expresiones de $f'(a)$ que son útiles. Cuando $a \in I$ es un extremo del intervalo, una de esas expresiones no tiene sentido, porque el conjunto que en ella aparece es vacío, pero la otra sigue siendo cierta. De hecho, si $a = \max I$, el conjunto $\{f_a(x) : x \in I,\ x < a\}$ ha de estar mayorado, pues en otro caso, f_a divergería en el punto a y f no sería derivable en a. Entonces, se puede razonar como hemos hecho en el teorema anterior, para probar que $f'(a)$ es el supremo de dicho conjunto. Análogamente, si $a = \min I$, también tendremos $f'(a) = \inf\{f_a(x) : x \in I,\ x > a\}$.

- *Sea I un intervalo no trivial y $f \in D^1(I)$. Las siguientes afirmaciones son equivalentes:*
  (i) *f es convexa.*
  (ii) *f' es creciente.*
  (iii) *Para cualesquiera $a,x \in I$ se tiene que $f(x) \geqslant f(a) + f'(a)(x-a)$.*

$(i) \Rightarrow (ii)$. Dados $a,b \in I$ con $a < b$, deberemos probar que $f'(a) \leqslant f'(b)$. Para ello basta tomar $x \in]a,b[$ y escribir

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

donde hemos usado que f_x es una función creciente y las expresiones de $f'(a)$ y $f'(b)$ dadas por (5), que como hemos comentado, son válidas también cuando $a = \text{mín}\, I$ o $b = \text{máx}\, I$.

$(ii) \Rightarrow (iii)$. Para $a,x \in I$ con $a \neq x$, el teorema del valor medio nos da

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

donde c es un punto intermedio entre a y x, luego bastará ver que $f'(c)(x-a) \geqslant f'(a)(x-a)$.

En efecto, si $a < c < x$, al ser f' creciente, tendremos $f'(c) \geqslant f'(a)$, y basta multiplicar ambos miembros por $x - a > 0$. En otro caso será $x < c < a$, luego $f'(c) \leqslant f'(a)$, pero esta desigualdad se invierte al multiplicar ambos miembros por $x - a < 0$.

$(iii) \Rightarrow (i)$. Dados $x,y \in I$ y $t \in [0,1]$, tomando $a = (1-t)x + ty \in I$, deberemos probar que $(1-t)f(x) + tf(y) \geqslant f(a)$. Aplicando dos veces (iii), tenemos:

$$\begin{aligned}(1-t)f(x) + t\,f(y) &\geqslant (1-t)\big(f(a) + f'(a)(x-a)\big) + t\big(f(a) + f'(a)(y-a)\big)\\ &= f(a) + f'(a)\big((1-t)(x-a) + t(y-a)\big)\end{aligned}$$

y basta observar que

$$\begin{aligned}(1-t)(x-a) + t(y-a) &= \big((1-t)x + ty - (1-t)a - ta\\ &= a - (1-t)a - ta = 0\end{aligned}$$

■

Nótese la clara interpretación geométrica de la condición (iii) anterior: la gráfica de f se mantiene siempre por encima de la recta tangente en cualquiera de sus puntos.

Por otra parte, si disponemos de la segunda derivada, podemos usarla para caracterizar el crecimiento de f', obteniendo:

- *Sea I un intervalo no trivial y $f \in C^1(I) \cap D^2(I^\circ)$. Entonces f es convexa si, y sólo si, $f''(x) \geqslant 0$ para todo $x \in I^\circ$.*

Por supuesto, los dos resultados anteriores se traducen inmediatamente para tener sendas caracterizaciones de la concavidad.

14.4. Ejemplos

Veamos que las caracterizaciones anteriores permiten encontrar fácilmente intervalos de convexidad o de concavidad para diversas funciones.

Empecemos con un ejemplo obvio: todo polinomio de primer orden P define una función que es simultáneamente convexa y cóncava en $\mathbb{R}$, porque su primera derivada es constante, o porque su segunda derivada es idénticamente nula. En realidad esto se puede comprobar directamente con las definiciones: es claro que $P\big((1-t)x+ty\big) = (1-t)P(x)+tP(y)$ para cualesquiera $x,y,t \in \mathbb{R}$. Recíprocamente, es fácil comprobar que si I es un intervalo no trivial y $f : I \to \mathbb{R}$ es a la vez cóncava y convexa, entonces f es un polinomio de primer orden.

La función definida por un polinomio de segundo grado, $P(x) = ax^2 + bx + c$ para todo $x \in \mathbb{R}$, es convexa si $a > 0$ y cóncava si $a < 0$, ya que $P''(x) = 2a$ para todo $x \in \mathbb{R}$.

Si $f \in C^\infty(\mathbb{R}^+)$ es la función potencia de exponente $\alpha \in \mathbb{R}$, como $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$ para todo $x \in \mathbb{R}^+$, tenemos:

- *La función potencia de exponente* $\alpha \in \mathbb{R}$ *es convexa cuando* $\alpha \leqslant 0$ *o* $\alpha \geqslant 1$*, y cóncava cuando* $0 \leqslant \alpha \leqslant 1$.

Merece la pena comentar las extensiones de la función potencia que podemos hacer para algunos valores del exponente α. La comprobación de las siguientes afirmaciones es la ya comentada, pues la segunda derivada responde siempre a la misma expresión. Sólo hay que tener en cuenta el signo de $x^{\alpha-2}$ cuando $x \in \mathbb{R}^-$.

- *Para* $n \in \mathbb{N}$ *con* $n > 1$*, se tiene:*
  - (i) *Si* n *es par, la función* $x \mapsto x^n$*, es convexa (en* $\mathbb{R}$*), mientras que, si* n *es impar, dicha función es cóncava en* $\mathbb{R}_0^-$ *y convexa en* $\mathbb{R}_0^+$.
  - (ii) *Si* n *es par, la función* $x \mapsto x^{-n}$ *(definida en* $\mathbb{R}^*$*), es convexa en* $\mathbb{R}^+$ *y también en* $\mathbb{R}^-$*, mientras que, si* n *es impar, es cóncava en* $\mathbb{R}^-$ *y convexa en* $\mathbb{R}^+$.
  - (iii) *Si* n *es impar, la función raíz n-ésima es convexa en* $\mathbb{R}_0^-$ *y cóncava en* $\mathbb{R}_0^+$.

Conviene resaltar que algunas de las afirmaciones anteriores, por muy fácil que haya sido comprobarlas, implican desigualdades nada triviales. Por ejemplo, el hecho de que la función $x \mapsto x^4$ sea convexa en $\mathbb{R}$, significa que $\big((1-t)x+ty\big)^4 \leqslant (1-t)x^4 + ty^4$ para cualesquiera $x,y \in \mathbb{R}$ y $t \in [0,1]$. Esta desigualdad se puede comprobar por métodos puramente algebraicos, pero no es del todo fácil. Aún menos trivial es el siguiente resultado, cuya demostración es evidente:

- *La exponencial es una función convexa y el logaritmo es una función cóncava.*

La concavidad del logaritmo se puede reformular equivalentemente, obteniendo el siguiente resultado, que se conoce como *desigualdad de Young*:

- *Para* $p \in \mathbb{R}$ *con* $p > 1$, *definiendo* $p^* \in \mathbb{R}$ *por la igualdad* $\dfrac{1}{p} + \dfrac{1}{p^*} = 1$, *se tiene:*

$$ab \leqslant \frac{a^p}{p} + \frac{b^{p^*}}{p^*} \qquad \forall a, b \in \mathbb{R}^+$$

Tomando $x = a^p$, $y = b^{p^*}$ y $t = 1/p^* \in [0,1]$, tenemos:

$$\begin{aligned} \log\left(\frac{a^p}{p} + \frac{b^{p^*}}{p^*}\right) &= \log\big((1-t)x + t\,y\big) \\ &\geqslant (1-t)\log x + t\log y = \frac{1}{p}\log a^p + \frac{1}{p^*}\log b^{p^*} = \log(ab) \end{aligned}$$

y basta usar que la exponencial es creciente. ■

Encontrar intervalos de concavidad y convexidad para diversas funciones trigonométricas es también un ejercicio sencillo. Mencionamos un ejemplo:

- *El arco-tangente es una función convexa en* $\mathbb{R}_0^-$ *y cóncava en* $\mathbb{R}_0^+$.

14.5. Ejercicios

1. Probar que la suma de dos funciones convexas es una función convexa.

2. Dar un ejemplo de dos funciones convexas cuyo producto no sea una función convexa.

3. Sean I, J intervalos no triviales, $f : I \to J$ una función convexa y $g : J \to \mathbb{R}$ una función convexa y creciente. Probar que entonces $g \circ f$ es convexa.

4. Sea I un intervalo no trivial y $f : I \to \mathbb{R}$ una función convexa. Supongamos además que f es continua e inyectiva, luego estrictamente monótona. Probar que, si f es creciente, f^{-1} es cóncava, mientras que, si f es decreciente, entonces f^{-1} es convexa.

5. En cada uno de los siguientes casos, encontrar intervalos de convexidad o concavidad para la función $f : \mathbb{R} \to \mathbb{R}$ definida, para todo $x \in \mathbb{R}$ por

   (a) $f(x) = x^5 - 5x^4 + 5x^3 + 10$ (b) $f(x) = \dfrac{x^2 + 3x + 1}{x^2 + 1}$

   (c) $f(x) = \log\left(1 + x^2\right)$ (d) $f(x) = \operatorname{sen}^2 x$