Tema **10**

Funciones convexas

Los resultados obtenidos en el desarrollo del cálculo diferencial nos permiten estudiar con facilidad una importante familia de funciones reales de variable real definidas en intervalos, las *funciones convexas*. Haremos una discusión breve de esta familia de funciones, empezando por la definición de función convexa y su interpretación geométrica. Para funciones que sean derivables, o dos veces derivables, en un intervalo, obtenemos una útil caracterización de la convexidad.

10.1. Definición de función convexa

Sea I un intervalo, como siempre no vacío y no reducido a un punto, y $f : I \to \mathbb{R}$ una función. Se dice que f es *convexa* cuando verifica la siguiente condición:

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

Conviene resaltar que en (1) aparecen tres variables a,b,t, y la desigualdad ha de verificarse para todos los valores indicados de dichas variables: $a,b \in I$ con $a < b$ y $0 \leqslant t \leqslant 1$.

Se dice que la función $f : I \to \mathbb{R}$ es *cóncava* cuando $-f$ es convexa, es decir, cuando se verifica que $f\big((1-t)a+tb\big) \geqslant (1-t)\,f(a) + t\,f(b)$ para cualesquiera $a,b \in I$ con $a < b$ y $t \in [0,1]$. En lo que sigue trabajaremos solamente con funciones convexas, pues todo el estudio que hagamos se traduce inmediatamente para funciones cóncavas sin más que sustituir una función por su opuesta.

Para empezar a entender el significado de convexidad de una función, observemos que si para $t \in [0,1]$ escribimos $x = (1-t)a+tb = a+t(b-a)$, tenemos evidentemente $x \in [a,b] \subset I$. Recíprocamente, para cada $x \in [a,b]$, basta tomar

$$t = \frac{x-a}{b-a}, \qquad 1-t = \frac{b-x}{b-a}$$

para tener claramente $t \in [0,1]$ y $x = (1-t)a+tb$. Por tanto, la condición (1) se expresa equivalentemente de la siguiente forma:

$$a,b \in I,\ a < b \implies f(x) \leqslant \frac{b-x}{b-a}\,f(a) + \frac{x-a}{b-a}\,f(b) \quad \forall x \in [a,b] \tag{2}$$

Si ahora $g : [a,b] \to \mathbb{R}$ es la función que aparece en la última desigualdad, es decir,

$$g(x) = \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) \quad \forall x \in [a,b]$$

la condición (2) se reformula diciendo que $f(x) \leqslant g(x)$ para todo $x \in [a,b]$, y no conviene olvidar que esto debe verificarse para cualesquiera $a,b \in I$ con $a < b$, ya que la función g depende claramente de ellos. Observamos que g es una función polinómica de grado menor o igual que 1. Puesto que evidentemente $g(a) = f(a)$ y $g(b) = f(b)$, la gráfica de g es el segmento en el plano que une los puntos $\big(a, f(a)\big)$ y $\big(b, f(b)\big)$, un segmento de recta secante a la gráfica de f.

Por tanto, la función f es convexa si, y sólo si, para cualquier intervalo cerrado y acotado $[a,b] \subset I$, la gráfica de f en dicho intervalo se mantiene "por debajo" del segmento que une los puntos $\big(a, f(a)\big)$ y $\big(b, f(b)\big)$. Al variar a y b obtenemos una idea muy clara e intuitiva de la "forma" que debe tener la gráfica de una función convexa.

Conviene comentar que la noción de función convexa es propia del Álgebra Lineal. Más concretamente, si X es un espacio vectorial sobre el cuerpo $\mathbb{R}$, se dice que un conjunto $E \subset X$ es *convexo* cuando $(1-t)a + tb \in E$ para cualesquiera $a,b \in E$ y $t \in [0,1]$. Es natural decir que el conjunto $\{(1-t)a + tb : t \in [0,1]\}$ es el segmento de extremos a y b en el espacio X, luego el conjunto E es convexo cuando contiene al segmento que une dos cualesquiera de sus puntos. En el caso $X = \mathbb{R}$, es evidente que los subconjuntos convexos de $\mathbb{R}$ son precisamente los intervalos, admitiendo en este momento todos los intervalos. También es fácil imaginar la forma que debe tener un subconjunto convexo de $\mathbb{R}^2$ o de $\mathbb{R}^3$.

Pues bien, si ahora tenemos una función $f : E \to \mathbb{R}$, donde E es un subconjunto convexo no vacío de nuestro espacio vectorial X, podemos decir que f es convexa cuando verifique que

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

La interpretación geométrica sería esencialmente la misma que hemos hecho en el caso $X = \mathbb{R}$. Queda claro que las nociones de conjunto convexo y de función convexa sólo involucran la estructura de espacio vectorial de X, luego son nociones de Álgebra Lineal. Sin embargo la *Convexidad*, es decir, el estudio de los conjuntos convexos y de las funciones convexas, se considera hoy día como toda una rama de la Matemática, con multitud de aplicaciones en otras ciencias.

10.2. Continuidad y derivabilidad

Para obtener propiedades importantes de las funciones convexas, conviene expresar las condiciones (1) o (2) de forma más simétrica con respecto a las tres variables que en ellas intervienen. Si I es un intervalo y tomamos $x_1, x_2, x_3 \in I$ tales que $x_1 < x_2 < x_3$, podemos escribir

$$x_2 = (1-t)x_1 + t x_3 \quad \text{con} \quad t = \frac{x_2 - x_1}{x_3 - x_1} \in [0,1] \quad \text{y} \quad (1-t) = \frac{x_3 - x_2}{x_3 - x_1}$$

Por tanto, si $f : I \to \mathbb{R}$ es una función convexa, se tendrá:

$$f(x_2) \leqslant (1-t)\, f(x_1) + t\, f(x_3) \tag{3}$$

Restando en ambos miembros $f(x_1)$ obtenemos

$$f(x_2) - f(x_1) \leqslant t\,\big(f(x_3) - f(x_1)\big), \quad \text{es decir,} \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant \frac{f(x_3) - f(x_1)}{x_3 - x_1}$$

Por otra parte, si en (3) cambiamos de signo ambos miembros invirtiendo la desigualdad, y sumamos $f(x_3)$ a ambos, obtenemos

$$f(x_3) - f(x_2) \geqslant (1-t)\,\big(f(x_3) - f(x_1)\big) \quad \text{es decir,} \quad \frac{f(x_3) - f(x_2)}{x_3 - x_2} \geqslant \frac{f(x_3) - f(x_1)}{x_3 - x_1}$$

Obsérvese que las dos desigualdades obtenidas pueden enlazarse. En resumen hemos probado que si $f : I \to \mathbb{R}$ es convexa, para cualesquiera $x_1, x_2, x_3 \in I$ con $x_1 < x_2 < x_3$, se tiene

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

Podemos ya probar fácilmente la primera propiedad clave de las funciones convexas, que puede expresarse diciendo que sus cocientes incrementales son funciones crecientes.

- *Sea I un intervalo y $f : I \to \mathbb{R}$ una función convexa. Para cada $a \in I$ consideremos la función $f_a : I \setminus \{a\} \to \mathbb{R}$ dada por*

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

  Entonces f_a es una función creciente, para todo $a \in I$.

En efecto, dados $a \in I$ y $x, y \in I \setminus \{a\}$ con $x < y$, consideramos los tres casos posibles, para probar siempre que $f_a(x) \leqslant f_a(y)$.

– Si $a < x < y$ usamos la primera desigualdad de (4) con $x_1 = a$, $x_2 = x$ y $x_3 = y$, obteniendo directamente que $f_a(x) \leqslant f_a(y)$.

– Si $x < y < a$ usamos la segunda desigualdad de (4) con $x_3 = a$, $x_1 = x$ y $x_2 = y$, obteniendo

$$\frac{f(a) - f(x)}{a - x} \leqslant \frac{f(a) - f(y)}{a - y} \quad \text{es decir,} \quad f_a(x) \leqslant f_a(y)$$

– Finalmente, si $x < a < y$ usamos la desigualdad entre el primer y último miembro de (4) con $x_2 = a$, $x_1 = x$ y $x_3 = y$, obteniendo de nuevo $f_a(x) \leqslant f_a(y)$. ■

Del resultado anterior deducimos ahora que toda función convexa admite derivadas laterales, y por tanto es continua, en todo punto interior de su intervalo de definición:

- *Sea I un intervalo 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,*

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

Fijado $a \in I^\circ$, la demostración de (5) es bien sencilla, usando solamente el crecimiento de la función f_a. 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 $\{f_a(x) : x \in I,\ x < a\}$ es no vacío (de nuevo porque $a \in I^\circ$) y mayorado, sea s_a su supremo y veamos 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 $\delta = a - x_0 > 0$, para $a - \delta < x < a$ tendremos $s_a - \varepsilon < f_a(x_0) \leqslant f_a(x) \leqslant s_a$, de donde $|f_a(x) - s_a| < \varepsilon$. Esto prueba que $\lim\limits_{x \to a-} f_a(x) = s_a$, 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, ya que

$$|(1-t)a + tb| \leqslant (1-t)|a| + t|b| \qquad \forall a, b \in \mathbb{R},\ \ \forall t \in [0,1]$$

pero no es derivable en 0.

Tampoco podemos asegurar que una función convexa sea continua en los extremos de su intervalo de definición. 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.

La demostración anterior encierra una idea que merece la pena resaltar: de la función f_a sólo hemos usado que es creciente. Por tanto, exactamente la misma demostración nos da la siguiente propiedad de las funciones crecientes, que obviamente comparten las decrecientes:

- *Sea I un intervalo y $g : I \to \mathbb{R}$ una función monótona. Entonces g tiene límite por la izquierda y por la derecha en todo punto $a \in I^\circ$. Como consecuencia, g es continua, o tiene una discontinuidad de salto, en todo punto de I°.*

Observemos que, como consecuencia de lo anterior, *el conjunto de puntos de discontinuidad de una función monótona en un intervalo es numerable*. En efecto, si Δ es el conjunto de puntos de I° en los que una función creciente $g : I \to \mathbb{R}$ no es continua, bastará evidentemente probar que Δ es numerable. Para $a \in \Delta$ se tiene $g(a-) < g(a+)$ y la densidad de $\mathbb{Q}$ en $\mathbb{R}$ nos proporciona $r_a \in \mathbb{Q}$ tal que $g(a-) < r_a < g(a+)$. El crecimiento de g asegura que para $a, b \in \Delta$ con $a < b$, se tiene $g(a+) \leqslant g(b-)$ y por tanto $r_a < r_b$. Así pues, la aplicación $a \mapsto r_a$ nos hace ver que Δ es equipotente a un subconjunto de $\mathbb{Q}$, luego Δ es numerable.

Del mismo modo, podemos asegurar que *toda función convexa (o cóncava) en un intervalo es derivable salvo, a lo sumo, en un conjunto numerable de puntos del intervalo.*

10.3. Caracterizaciones de las funciones convexas

Podemos ya caracterizar la convexidad de funciones que sean derivables en un intervalo:

- *Sea I un intervalo 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 $f(x) \geqslant f(a) + f'(a)(x-a)$.*

$(i) \Rightarrow (ii)$. Sean $a,b \in I$ con $a < b$ y tomemos $c \in]a,b[\subset I^\circ$. Para $a < x < c < y < b$, usando que f_a y f_b son crecientes, junto con las expresiones de la derivada por la derecha y por la izquierda en c calculadas anteriormente, tenemos:

$$f_a(x) \leqslant f_a(c) = f_c(a) \leqslant f'(c) \leqslant f_c(b) = f_b(c) \leqslant f_b(y)$$

de donde deducimos claramente que

$$f'(a) = \lim_{x \to a} f_a(x) \leqslant f'(c) \leqslant \lim_{y \to b} f_b(y) = f'(b)$$

$(ii) \Rightarrow (iii)$. Sean $a,x \in I$ con $a \neq x$ (si $a = x$ no hay nada que demostrar). Si $a < x$, aplicando el Teorema del Valor Medio encontramos $c \in]a,x[$ verificando que

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

donde hemos usado que $f'(c) \geqslant f'(a)$ porque f' es creciente. Si $x < a$ será $c \in]x,a[$, pero llegamos a la misma desigualdad, pues ahora $f'(c) \leqslant f'(a)$ pero $x - a < 0$.

$(iii) \Rightarrow (i)$. Para ver que f es convexa, fijamos $x,y \in I$ con $x < y$, junto con $t \in]0,1[$, y tomando $a = (1-t)x + ty$ bastará comprobar que $f(a) \leqslant (1-t)f(x) + t f(y)$, desigualdad que es evidente para $t = 0$ y $t = 1$. Anotemos para uso posterior que

$$t = \frac{a-x}{y-x} \quad \text{y} \quad 1-t = \frac{y-a}{y-x}$$

Aplicando (iii) tenemos $f(x) \geqslant f(a) + f'(a)(x-a)$, así como $f(y) \geqslant f(a) + f'(a)(y-a)$. Teniendo en cuenta que $y - a > 0$ y $x - a < 0$, enlazamos ambas desigualdades:

$$\frac{f(x)-f(a)}{x-a} \leqslant f'(a) \leqslant \frac{f(y)-f(a)}{y-a}$$

Basta ya operar con la desigualdad anterior para obtener

$$f(a) \leqslant \frac{y-a}{y-x} f(x) + \frac{a-x}{y-x} f(y) = (1-t) f(x) + t f(y) \qquad \blacksquare$$

Nótese la clara interpretación geométrica de la condición (iii) anterior: para cada punto $a \in I$, la gráfica de la función f se mantiene siempre por encima de la recta tangente a dicha gráfica en el punto $\big(a, f(a)\big)$. Es ya evidente lo que ocurre si disponemos de la segunda derivada:

- *Sea I un intervalo y $f \in D^2(I)$. Las siguientes afirmaciones son equivalentes:*

  (i) *f es convexa.*
  (ii) *$f''(x) \geqslant 0$ para todo $x \in I$.*

10.4. Ejemplos

Antes de mostrar varios ejemplos de funciones convexas observemos por comodidad que, si I es un intervalo y $f : I \to \mathbb{R}$ una función convexa, la desigualdad

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

se verifica para cualesquiera $a,b \in I$. Cuando $a < b$, ello ocurre por definición de función convexa, cuando $a = b$ es evidente y cuando $a > b$ basta intercambiar los papeles de a y b, intercambiando también t con $1-t$.

Todo polinomio de grado menor o igual que uno define una función convexa y cóncava en $\mathbb{R}$, pues su derivada es constante. Tenemos lo que suele denominarse una función *afín*. Una función dada por un polinomio de grado 2, $P(x) = ax^2 + bx + c$ para todo $x \in \mathbb{R}$, es convexa cuando $a > 0$ y cóncava cuando $a < 0$, puesto que $P''(x) = 2a$ para todo $x \in \mathbb{R}$. Tomando $Q(x) = x^3$ para todo $x \in \mathbb{R}$, tenemos una función que no es cóncava ni convexa. De hecho, puesto que $Q''(x) = 6x$ para todo $x \in \mathbb{R}$, Q es convexa en $\mathbb{R}_0^+$ y cóncava $\mathbb{R}_0^-$.

Un nuevo ejemplo salta a la vista, pues la derivada de la función exponencial es creciente, o bien su segunda derivada es siempre positiva. Obtenemos así una desigualdad nada trivial.

- *La función exponencial es convexa, es decir,*

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

Para el logaritmo, podemos igualmente usar que su derivada es decreciente, o que su segunda derivada es negativa en todo punto de $\mathbb{R}^+$:

- *El logaritmo es una función cóncava.*

Deducimos la llamada *desigualdad de Young*:

- *Para* $p \in]1,+\infty[$*, definiendo* $p^* \in]1,+\infty[$ *mediante la igualdad* $\dfrac{1}{p} + \dfrac{1}{p^*} = 1$*, se tiene*

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

En efecto, basta usar la concavidad del logaritmo,

$$\log\left(\frac{a^p}{p} + \frac{b^{p^*}}{p^*}\right) \geqslant \frac{1}{p}\log a^p + \frac{1}{p^*}\log a^{p^*} = \log(ab)$$

y ahora aplicar que la función exponencial es creciente.

Consideremos ahora la función potencia de exponente $\alpha \in \mathbb{R}$, $f(x) = x^\alpha$ para todo $x \in \mathbb{R}^+$. Tenemos $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$ para todo $x \in \mathbb{R}^+$, luego:

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

Concluimos con algunas funciones trigonométricas:

- *Para* $k \in \mathbb{Z}$, *la función seno es convexa en el intervalo* $[(2k-1)\pi, 2k\pi]$ *y cóncava en* $[2k\pi, (2k+1)\pi]$.
- *La función arco tangente es convexa en* $\mathbb{R}_0^-$ *y cóncava en* $\mathbb{R}_0^+$.

10.5. Ejercicios

1. Sean I, J intervalos, $f : I \to J$ una función convexa y $g : J \to \mathbb{R}$ una función convexa y creciente. Probar que entonces $g \circ f$ es convexa. Mostrar con un ejemplo que la hipótesis de que g sea creciente no puede suprimirse.

2. Probar que si I es un intervalo y $f : I \to \mathbb{R}$ una función convexa, la función $h : I \to \mathbb{R}$ definida por
$$h(x) = e^{f(x)} \quad \forall x \in I$$
también es convexa. Mostrar con un ejemplo que el recíproco no es cierto, h puede ser convexa sin que f lo sea.

3. Dar un ejemplo de dos funciones convexas cuyo producto no sea una función convexa. Probar que si I es un intervalo y $f, g : I \to \mathbb{R}^+$ son funciones tales que $\log \circ f$ y $\log \circ g$ son convexas (se dice que f y g son *logarítmicamente convexas*), entonces f, g y fg son funciones convexas.

4. Sea I un intervalo y $f : I \to \mathbb{R}$ una función continua en I, convexa e inyectiva, luego estrictamente monótona. Consideremos la función inversa f^{-1}, definida en el intervalo $J = f(I)$. Probar que, si f es creciente, f^{-1} es cóncava, mientras que, si f es decreciente, entonces f^{-1} es convexa.

5. Encontrar intervalos de convexidad o concavidad para la función $f : \mathbb{R} \to \mathbb{R}$ en cada uno de los siguientes casos:

$$\text{(a)} \quad f(x) = x^5 - 5x^4 + 5x^3 + 10 \quad \forall x \in \mathbb{R}$$
$$\text{(b)} \quad f(x) = \frac{x^2 + 3x + 1}{x^2 + 1} \quad \forall x \in \mathbb{R}$$
$$\text{(c)} \quad f(x) = x^2 \log|x| \quad \forall x \in \mathbb{R}^*, \quad f(0) = 0$$