SUMA 36
febrero 2001, pp. 51-56

Conjuntos convexos, funciones convexas y desigualdades clásicas

Antonio Álvarez Álvarez

ARTÍCULOS

En este artículo se presenta una propuesta para introducir el concepto de función convexa de un modo diferente al habitual, complementario a éste, que se apoya en la relación entre convexidad de funciones y conjuntos convexos, y que no requiere que la función sea derivable. Además, permite obtener, de forma sencilla y unificada, las desigualdades numéricas clásicas a partir de la convexidad de ciertas funciones

LA TEORÍA DE FUNCIONES y conjuntos convexos incorpora elementos del análisis, la geometría, el álgebra, etc., y tiene numerosas aplicaciones en programación lineal, problemas clásicos de extremos, teoría de juegos... por lo que es de interés especialmente en matemática aplicada.

Una de las ideas fundamentales de esta teoría es la conexión natural que existe entre la noción de función convexa (que en el Bachillerato se estudia generalmente como una noción puramente analítica[1], y bajo el requisito más o menos explícito de derivabilidad) y la de conjunto convexo (que suele aparecer asociada a la de polígono en el estudio de la geometría plana que se hace en la ESO). Utilizando esta conexión, es posible –como se muestra en este artículo– introducir la convexidad de funciones de una forma distinta, complementaria a la habitual, que permite resaltar el carácter geométrico de la convexidad.

En segundo lugar, veremos cómo las desigualdades clásicas –tan útiles en la resolución de problemas, especialmente en los de Olimpiadas Matemáticas– pueden obtenerse de forma muy sencilla a partir de la convexidad de funciones, entendida del modo al que aquí nos referimos. Es más, algunas de esas desigualdades (por ejemplo, la que relaciona la media aritmética con la media cuadrática de una colección finita de números positivos) no son más que una traducción del hecho de ser convexa una determinada función elemental (en el ejemplo citado, la función x^2, en $\mathbb{R}^+$).

Sólo se pretende con este artículo mostrar algunas ideas que, tratadas adecuadamente, son susceptibles de ser empleadas en el aula en la etapa de Bachillerato –también en niveles superiores–; por ello la exposición se limita al ámbito natural de trabajo en la Enseñanza Secundaria (geometría plana, funciones reales de variable real), renunciando a todo tipo de generalizaciones posibles.

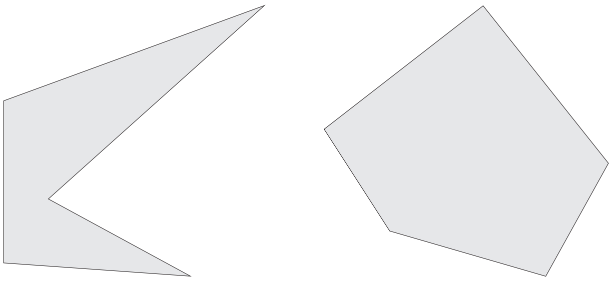
Conjuntos convexos

Intuitivamente, un conjunto C convexo es aquél que contiene el segmento que une dos puntos cualesquiera de C.

Una definición formal, bien conocida, es:

Un subconjunto C de $\mathbb{R}^2$ decimos que es *convexo* si para cualesquiera x, $y \in C$, y para todo número real λ con $0 < \lambda < 1$ se tiene:

$$(1 - \lambda x) + \lambda y \in C.$$



Conjunto no convexo   Conjunto convexo

Figura 1

Hagamos algunas observaciones sobre esta definición:

- Nada cambia si a λ exigimos $0 ″ \lambda ″ 1$.
- La definición puede reescribirse así:

  C es *convexo* si $\alpha x + \beta y \in C$ para cualesquiera $x, y \in C$, y para todo $\alpha \geq 0$, $\beta \geq 0$, con $\alpha + \beta = 1$.

El empleo de herramientas informáticas puede ser muy útil para hacer ver a los alumnos cómo esta definición responde a la noción intuitiva.

Los alumnos no tendrán problema en reconocer que rectas, segmentos, polígonos regulares o círculos son conjutos convexos del plano. No obstante, es fácil probar que la intersección de una colección arbitraria de conjuntos convexos es un conjunto convexo. Dos consecuencias destacables de este resultado son:

a) El conjunto de las soluciones de cualquier sistema de ecuaciones e inecuaciones lineales con dos (n) incógnitas es un convexo de $\mathbb{R}^2$ ($\mathbb{R}^n$).

b) Tiene sentido definir la envoltura convexa de un subconjunto cualquiera M de $\mathbb{R}^2$ como la intersección de todas las partes convexas de $\mathbb{R}^2$ que contienen a M. Es, obviamente, el conjunto convexo más pequeño que contiene a M.

De a) se deduce que semiplanos, rectas, segmentos, triángulos, polígonos regulares, son conjuntos convexos del plano. De b) se tiene, por ejemplo, que todo círculo es también convexo por ser envoltura convexa de la correspondiente circunferencia.

El empleo de herramientas informáticas puede ser muy útil para hacer ver a los alumnos cómo esta definición responde a la noción intuitiva.

1 Durante la elaboración de este artículo he encontrado la definición de función convexa que aquí se analiza, primero en el libro de texto *Matemáticas I*, de COU, de J. R. Vizmanos y M. Anzola, editado por SM, y, ya al final, en *Matemáticas aplicadas a las Ciencias Sociales 2*, del Equipo Arrixaca, editado por Santillana.

Funciones convecas

Podemos ya presentar la convexidad de una función. Diremos que una función es *convexa* cuando la región de los puntos del plano que se encuentran en la gráfica de la función o por encima de ella es un conjunto convexo.

O, con un mayor grado de formalismo:

Sea I un intervalo. Sea f: $I \to \mathbb{R}$. Llamamos *epígrafo de f*, y lo denotamos *epi*(f), al subconjunto de $\mathbb{R}^2$:

$$epi(f) = \{(x, z): x \in I, f(x) ″ z\}$$

De una función f: $I \to \mathbb{R}$. decimos que es *convexa* si *epi*(f) es un conjunto convexo de $\mathbb{R}^2$. Y decimos que es *cóncava* si $-f$ es convexa.

Por tanto, de ciertas funciones sencillas, cuyas gráficas son bien conocidas (si disponemos de calculadoras gráficas, esto sirve para todas), podemos decir inmediatamente si son convexas o no. Por ejemplo, $f(x) = ax + b$ es convexa y cóncava.

Puede probarse como ejercicio que toda función convexa y cóncava es afín.

Las siguientes funciones son convexas (y no cóncavas):

- $f(x) = |x|$.
- $f(x) = x^n$, con n par.
- $f(x) = a^x$, con $0 < a < 1$ o $a > 1$.
- $f(x) = -\log_a(x)$, con $a > 1$.

Las siguientes funciones no son convexas en su dominio ($\mathbb{R}$), pero sí lo son en los intervalos que se citan:

- $f(x) = \text{sen}(x)$, en $I = [(2k - 1)\pi, 2k\pi]$, con k entero.
- $f(x) = x^n$ (n impar), en $I = [0,+\infty)$.

El dominio de la función $f(x) = 1/x$ no es un intervalo, pero la restricción de esta función a $\mathbb{R}^+$ es convexa.

Como se observa, la definición de función convexa que hemos utilizado es de naturaleza geométrica y no recurre al

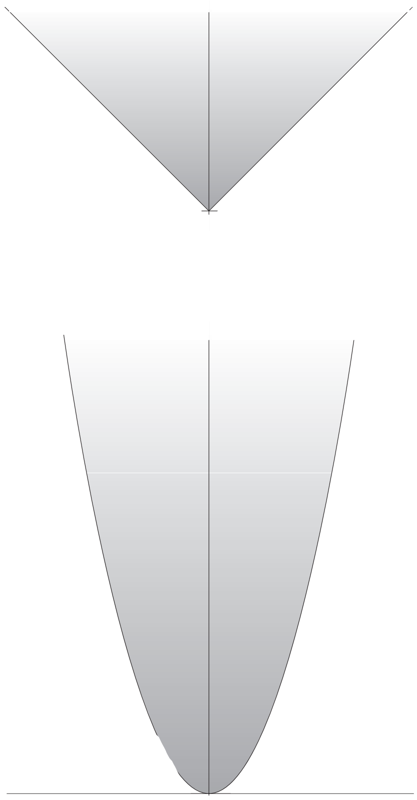


Figura 2. Dos funciones convexas:
$|x|$, x^2

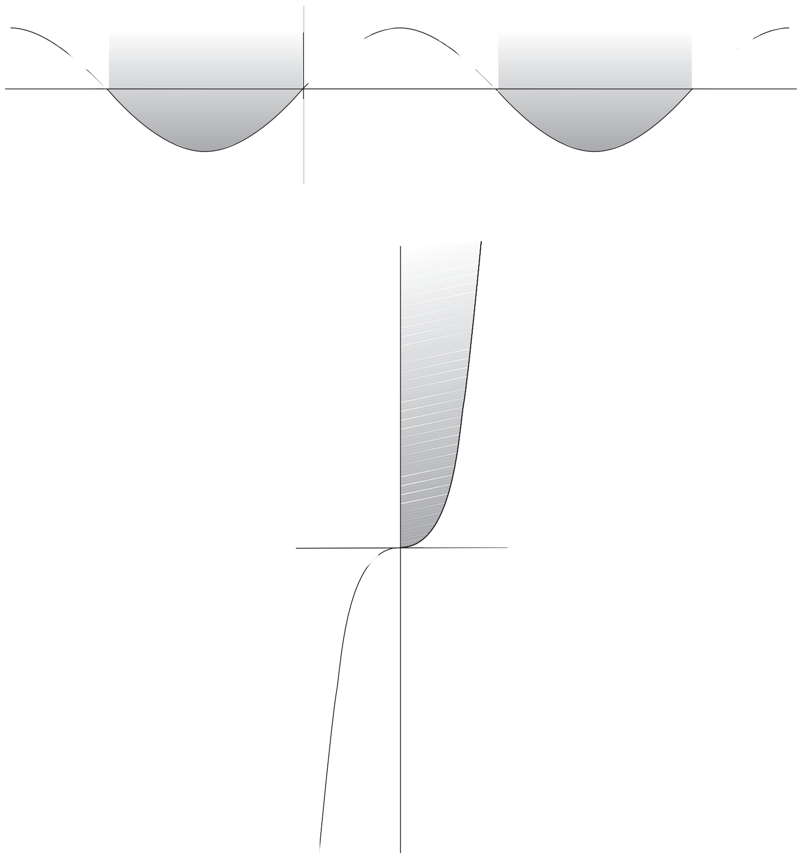


Figura 4. sen x es convexa en [(2k–1)≠, 2k≠], con k entero
x^3 es convexa en [0, + °]

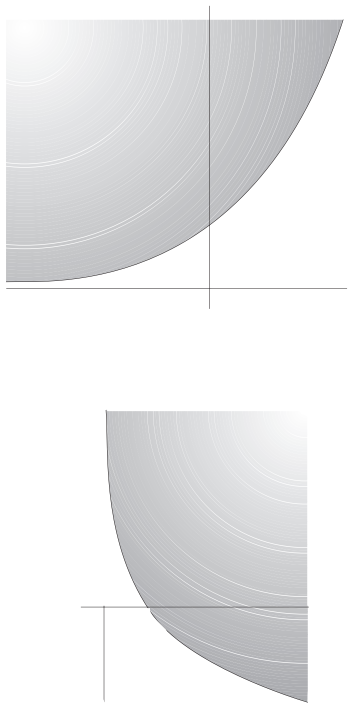


Figura 3. 2^x y $-\log_2 x$ son convexas

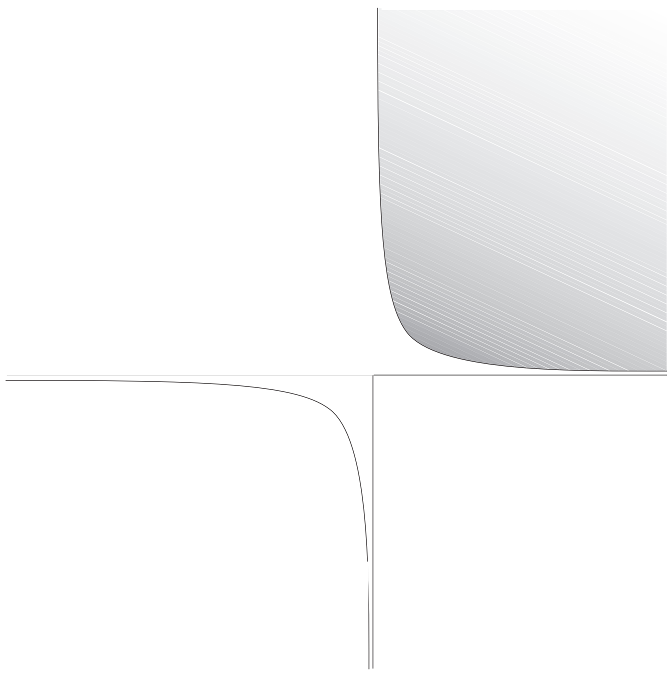


Figura 5. $1/x$ es convexa en $\mathbb{R}^+$

concepto de derivada. De hecho, una función puede ser convexa y no ser derivable en algún punto del intervalo de definición I, como le ocurre a la función valor absoluto en $x = 0$. Incluso, una función convexa puede no ser continua en los extremos de I. Es el caso de:

$$f(x) = \begin{cases} 0, & a < x < b \\ 1, & x = a \text{ o } x = b \end{cases}$$

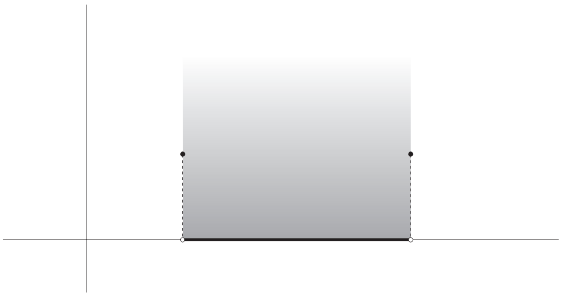


Figura 6. Función convexa no continua

Puede probarse, eso sí, que una función convexa definida en I es continua en todo punto de I que no sea extremo de I (Aparicio y Payá, 1985).

Así pues, el concepto de función convexa que se propone es una generalización del que se usa habitualmente (que requiere la existencia de derivada) y, sin embargo, no sólo no pierde la posibilidad de ser interpretado geométricamente, sino que su naturaleza geométrica es manifiesta. Además, las nociones de conjunto convexo y de función convexa aparecen, así, relacionadas de forma natural.

El siguiente teorema es una útil caracterización de las funciones convexas, que algunos autores adoptan como definición. Como puede apreciarse en la figura 7, tiene una clara interpretación geométrica, que puede servir para justificar el teorema si se desea obviar su demostración, que, no obstante, aquí se expone.

Teorema 1

Sea $f\colon I \to \mathbb{R}$, donde I es un intervalo. Entonces f es convexa en I si y sólo si

$$f((1 - \lambda x) + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y),\ 0 < \lambda < 1,$$

para todo x, y de I.

Demostración

$\Rightarrow$) Sean $x, y \in I$ arbitrarios. Obviamente, $(x, f(x))$ e $(y, f(y))$ pertenecen a $epi(f)$, que es convexo. Por tanto,

$$(1 - \lambda)(x, f(x)) + \lambda(y, f(y)) =$$
$$= ((1 - \lambda)x + \lambda y,\ (1 - \lambda)f(x) + \lambda f(y))$$

pertenece a $epi(f)$, para todo λ comprendido entre 0 y 1, de donde se tiene la propiedad de arriba.

$\Leftarrow$) Sean $(x, z), (y, t) \in epi(f)$. Entonces $f(x) \leq z$, $f(y) \leq t$.

Así se tiene que:

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) +$$
$$+ \lambda f(y) \leq (1 - \lambda)z + \lambda t,\ \forall \lambda \in (0, 1).$$

Y entonces:

$$((1 - \lambda)x + \lambda y,\ (1 - \lambda)z + \lambda t) =$$
$$= (1-\lambda)(x, z) + \lambda(y, t) \in epi(f),$$

para todo λ con $0 < \lambda < 1$; es decir, $epi(f)$ es convexo.

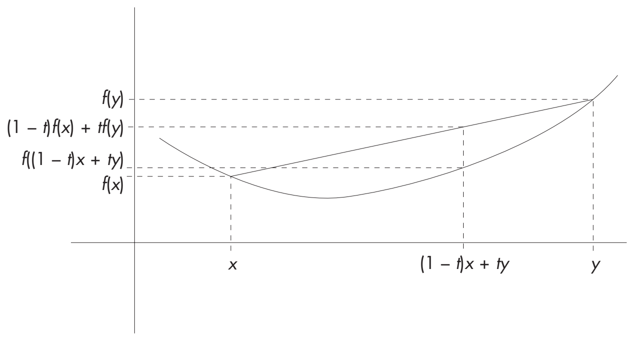


Figura 7

…las nociones de conjunto convexo y de función convexa aparecen, así, relacionadas de forma natural.

El teorema anterior sugiere hacer la siguiente definición:

Sea $f\colon I \to \mathbb{R}$, donde I es un intervalo. Decimos que f es estrictamente convexa si:

$$f((1 - \lambda x) + \lambda y) < (1 - \lambda)f(x) + \lambda f(y)$$

con $0 < \lambda < 1$, para todo x, y de I, con $x \neq y$.

Pero, como se ha dicho, esta forma de presentar la convexidad de funciones es complementaria de la usual: obviamente no podemos prescindir del potente criterio de la derivada segunda. Volvamos a él en este punto:

Sea $f\colon I \to \mathbb{R}$ una función dos veces derivable en I. Entonces f es convexa si y sólo si $f''(x) \geq 0$ para todo x de I.

La demostración de esta propiedad, utilizando el teorema 1, sobrepasa el nivel de un curso de Bachillerato. Puede consultarse la demostración en Aparicio y Payá (1985) o en Webster (1994). En Rockafellar (1970) se encuentra una demostración distinta de un resultado algo más restrictivo (exige continuidad a la derivada segunda).

De forma análoga, puede demostrarse que si $f''(x) > 0$ para todo x de I, entonces f es estrictamente convexa en I. Sin embargo, el recíproco no es cierto. Por ejemplo, la función $f(x) = x^4$ es estrictamente convexa y $f''(0) = 0$.

Estas dos propiedades habrán de justificarse, por tanto, apelando a la intuición de los alumnos, como, por otra parte, es lo habitual.

Convexidad y desigualdades clásicas

Un aspecto muy interesante de la introducción a la convexidad que aquí se ha propuesto es que permite mostrar de un modo muy sencillo la íntima relación que existe entre las desigualdades numéricas clásicas y la convexidad de ciertas funciones elementales, relación que suele pasar desapercibida.

Un aspecto muy interesante de la introducción a la convexidad que aquí se ha propuesto es que permite mostrar de un modo muy sencillo la íntima relación que existe entre las desigualdades numéricas clásicas y la convexidad de ciertas funciones elementales, relación que suele pasar desapercibida.

La pieza clave es el teorema 1, del que se pueden deducir de forma sencilla y unificada algunas de las desigualdades clásicas; por ejemplo, las que relacionan las distintas medias –que obtendremos aquí–, y otras como la desigualdad de Minkowski, la de Hölder, etc.

El esquema que seguiremos para obtener desigualdades es muy simple: habiendo comprobado que una determinada función es convexa, aplicaremos el teorema 1 y obtendremos la desigualdad correspondiente. Para saber si una función es convexa podemos recurrir, si no conocemos su gráfica, al criterio usual que caracteriza a las funciones convexas dos veces derivables.

Vamos a obtener ya las desigualdades anunciadas:

a) La función $-\ln$: $\mathbb{R}^+ \to \mathbb{R}$ es estrictamente convexa. Es decir, para todo $x, y > 0$ se verifica:

$$-\ln((1-\lambda)x + \lambda y) \le (1-\lambda)(-\ln x) + \lambda(-\ln y),\ 0 < \lambda < 1.$$

Y la igualdad sólo es cierta si x = y.

Por ser ln (estrictamente) creciente, esto equivale a:

$$x^{1-\lambda} \cdot y^{\lambda} \le (1-\lambda)x + \lambda y,\ 0 < \lambda < 1. \qquad [1]$$

Tenemos así que la conocida desigualdad entre la media geométrica y la media aritmética ponderadas de dos números positivos x e y equivale a la convexidad de $-\ln x$.

En particular, para $\lambda = 1/2$:

$$\sqrt{xy} \le \frac{x+y}{2} \qquad [2]$$

donde la desigualdad es estricta salvo que $x = y$.

b) La función x^2 es (estrictamente) convexa. Esto es:

$$((1-\lambda)x + \lambda y)^2 \le (1-\lambda)x^2 + \lambda y^2,\ 0 < \lambda < 1,$$

para todo $x, y \in \mathbb{R}$. Como es (estrictamente) creciente en $\mathbb{R}^+$:

$$(1-\lambda)x + \lambda y \le ((1-\lambda)x^2 + \lambda y^2)^{1/2},\ 0 < \lambda < 1, \qquad [3]$$

para todo $x, y > 0$.

Es decir, la convexidad de x^2 en $\mathbb{R}^+$ se *traduce* en la desigualdad entre las medias aritmética y cuadrática ponderadas de dos números positivos.

En particular, para $\lambda = 1/2$:

$$\frac{x+y}{2} \le \sqrt{\frac{x^2+y^2}{2}} \qquad [4]$$

donde sólo se alcanza la igualdad si $x = y$.

c) Razonando del mismo modo, la convexidad de $1/x$ en $\mathbb{R}^+$ nos dice que la media armónica de dos números positivos es menor o igual que la media aritmética. No obstante, aplicando [2] a $1/x$ y $1/y$, podemos probar algo mejor: que la media armónica es también menor o igual que la geométrica:

$$\frac{1}{\frac{1}{2}\left(\frac{1}{x}+\frac{1}{y}\right)} \le \frac{1}{\sqrt{\frac{1}{x}\frac{1}{y}}} = \sqrt{xy} \qquad [5]$$

Al igual que antes, la igualdad es cierta si y sólo si $x = y$.

d) Llamemos $h(x, y)$, $g(x, y)$, $m(x, y)$ y $q(x, y)$ a las medias armónica, geométrica y aritmética y cuadrática, respectivamente, de dos números positivos x, y.

Sean $0 < x \le y$. Es inmediato comprobar que

$$x \le h(x, y)\ ,\ q(x, y) \le y$$

En resumen:

$$x \leq \frac{1}{\frac{1}{2}\left(\frac{1}{x}+\frac{1}{y}\right)} \leq \sqrt{xy} \leq \frac{x+y}{2} \leq \sqrt{\frac{x^2+y^2}{2}} \leq y \qquad [6]$$

esto es,

$$x \leq h(x, y) \leq g(x,y) \leq m(x,y) \leq q(x,y) \leq y$$

Y todas las desigualdades son estrictas a menos que $x = y$.

Extensión al caso finito. La desigualdad de Jensen

Hemos obtenido las desigualdades anteriores como aplicación del teorema 1, que caracteriza a las funciones convexas mediante una propiedad (desigualdad) en la que intervienen sólo dos números x, y de un intervalo I. Cabe objetar que podrían haberse obtenido más fácilmente por otros medios. Por ejemplo, la desigualdad [2] es consecuencia inmediata de ser $(x - y)^2 \geq 0$. Sin embargo, como vamos a ver, la desigualdad del teorema 1 equivale a su extensión al caso de una colección finita de números x_1, ..., x_n de I, por lo que casi con el mismo esfuerzo podríamos haber deducido directamente las correspondientes extensiones de las desigualdades anteriores para un número finito de números positivos. (No lo hemos hecho así buscando una mayor simplicidad en la notación y pensando en su posible uso en el aula.)

Teorema 2 (Desigualdad de Jensen)

Sea $f: I \to \mathbb{R}$. Sean $x_1, \ldots, x_n \in I$ y sean $\alpha_1, \ldots, \alpha_n \geq 0$, tales que $\alpha_1 + \ldots + \alpha_n = 1$. Entonces f es convexa en I si y sólo si:

$$f(\alpha_1 x_1 + \ldots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \ldots + \alpha_n f(x_n).$$

La demostración de este teorema es un buen ejemplo de prueba por inducción, que podría ser empleado en el 2.° curso de Bachillerato.

Utilizando la desigualdad de Jensen en lugar de la del teorema 1, se obtienen las correspondientes extensiones de todas las desigualdades anteriores.

Conclusión

Como ya se ha dicho, la forma de introducir la convexidad de funciones que se ha descrito no pretende sustituir a la forma habitual; más bien, se ha buscado aportar un enfoque distinto, complementario, del que podríamos destacar, resumiendo, varios aspectos interesantes:

- Hace hincapié en la naturaleza geométrica del concepto de función convexa.
- Permite mostrar a los alumnos cómo una misma idea o concepto puede ser abordado desde perspectivas diversas.
- Permite un acercamiento puramente intuitivo o una exposición más formalizada.
- Supone un ejemplo simple y elegante de generalización de un concepto.
- Permite mostrar la estrecha relación que existe entre la convexidad y las desigualdades numéricas, así como obtener de forma sencilla y unificada las desigualdades clásicas entre medias.

…la forma de introducir la convexidad de funciones que se ha descrito no pretende sustituir a la forma habitual; más bien, se ha buscado aportar un enfoque distinto, complementario…

Antonio Álvarez
IES José de Mora
Baza (Granada).
Sociedad Andaluza
de Educación Matemática
«Thales»

Bibliografía

APARICIO DEL PRADO, C. y R. PAYÁ ALBERT (1985): *Análisis Matemático I*, Universidad de Granada.

BALBAS, A. y J. A. GIL (1990): *Programación Matemática*, AC, Madrid.

BECKENBACH, E. F. y R. BELLMAN (1965): *Inequalities*, Springer-Verlag, Berlin.

DEMIDOVICH, B. P. (1985): *5000 problemas de Análisis Matemático*, Paraninfo, Madrid.

FERNÁNDEZ VIÑA, J. A. y E. SÁNCHEZ MAÑES (1979): *Ejercicios y complementos de Análisis Matemático I*, Tecnos, Madrid.

HARDY, G., J. E. LITTLEWOOD y G. PÓLYA (1952): *Inequalities*, Cambridge University Press.

PÉREZ DE MADRID, A. y C. GARCÍA ARRIBAS (1986): «Métodos para resolver problemas. Las Olimpiadas Matemáticas. Su preparación», *Epsilon*, n.° 6/7, 100-111.

ROCKAFELLAR, R. T. (1970): *Convex Analysis,* Princeton University Press.

SPIVAK, M. (1981): *Cálculo Infinitesimal*, Reverté, Barcelona.

WEBSTER, R. J. (1994): *Convexity*, Oxford University Press.