F

Funciones convexas

F.1. Caracterización de las funciones convexas de una variable

En este epígrafe repasamos y ampliamos los resultados básicos sobre las funciones convexas de una variable que se suelen estudiar en el curso de Análisis Matemático I. Como novedad, merece la pena destacar la caracterización integral de las funciones convexas F.5. Esta sencilla caracterización, que es una mejora sustancial de la habitual en el contexto de las funciones derivables F.7, no se suele mencionar en los textos habituales dedicados al cálculo diferencial e integral de las funciones de una variable.

Si $-\infty \leq a < b \leq +\infty$, en lo que sigue denotaremos por $|a, b| \subset \mathbb{R}$ un intervalo genérico de extremos a, b (un intervalo acotado de la forma (a, b), $(a, b]$, $[a, b)$, $[a, b]$, con $-\infty < a < b < +\infty$, o un intervalo no acotado de la forma $(a, +\infty)$, $[a, +\infty)$, $(-\infty, b)$, $(-\infty, b]$, $(-\infty, +\infty)$).

Las funciones convexas se suelen definir geométricamente en la siguiente forma: *La función $f : |a, b| \to \mathbb{R}$ es convexa cuando cada par de puntos distintos de su gráfica determinan un segmento que queda por encima de la gráfica.*

Dada una función $f : |a, b| \to \mathbb{R}$ y un intervalo $[u, v] \subset |a, b|$, denotaremos por

$$m(u, v) = \frac{f(v) - f(u)}{v - u}$$

la pendiente de la recta que pasa por los puntos $(u, f(u))$, $(v, f(v))$. Por lo tanto, la ecuación de esta recta se puede escribir en la forma $r(x) = f(u) + m(u, v)(x - u)$, y también en la forma $r(x) = f(v) + m(u, v)(x - v)$.

La definición geométrica de función convexa admite las formulaciones analíticas que recoge la siguiente proposición:

Proposición F.1 *Una función $f : |a, b| \to \mathbb{R}$ es convexa si y sólo si cumple alguna de las condiciones equivalentes*
$[C1]$: $u < x < v \Rightarrow m(u, x) \leq m(u, v)$ en cada $[u, v] \subset |a, b|$.

[$C2$]: $u < x < v \Rightarrow m(u,v) \leq m(x,v)$ *en cada* $[u,v] \subset |a,b|$.
[$C3$]: $u < x < v \Rightarrow m(u,x) \leq m(x,v)$ *en cada* $[u,v] \subset |a,b|$.
[$C4$]: $\alpha + \beta = 1,\ \alpha > 0,\ \beta > 0\ \Rightarrow f(\alpha u + \beta v) \leq \alpha f(u) + \beta f(v)$.

DEM: Si f es convexa, para cada $[u,v] \subset |a,b|$ y cada $x \in (u,v)$ se cumple que $f(x) \leq r(x)$ donde $r(x) = f(u) + m(u,v)(x-u)$ es la función afín cuya gráfica es la recta que pasa por los puntos $(u, f(u))$, $(v, f(v))$. Como $x - u > 0$, la desigualdad $f(x) \leq r(x)$ adopta la forma $m(u,x) \leq m(u,v)$ luego f es convexa si y sólo si cumple [$C1$].

Análogamente, usando la expresión $r(x) = f(v) + m(u,v)(x-v)$, como $x - v < 0$, se llega a que la desigualdad $f(x) \leq r(x)$ adopta la forma $m(u,v) \leq m(x,v)$, luego f es convexa si y sólo si cumple [$C2$]. Si f es convexa, combinando [$C1$] con [$C2$], se obtiene que f cumple [$C3$].

Por otra parte, si f cumple [$C3$], también cumple [$C4$]: Sea $x = \alpha u + \beta v$, donde α, β son números positivos que cumplen $\alpha + \beta = 1$. Según la hipótesis

$$0 \leq m(x,v) - m(u,x) = \frac{f(v) - f(x)}{v - x} - \frac{f(u) - f(x)}{u - x} =$$

$$= \frac{f(v) - f(x)}{\alpha(v-u)} - \frac{f(x) - f(u)}{\beta(v-u)} = \frac{\alpha f(u) + \beta f(v) - f(x)}{\alpha\beta(v-u)}$$

es decir, $f(\alpha u + \beta u) = f(x) \leq \alpha f(u) + \beta f(v)$.

Queda por demostrar que si f cumple [$C4$] entonces f es convexa:
Todo $x \in (u,v)$, se puede representar en la forma $x = \alpha u + \beta v$, donde

$$\alpha = (v-x)/(v-u) > 0,\ \beta = (x-u)/(v-u) > 0$$

son números positivos que cumplen $\alpha + \beta = 1$. Según la hipótesis se cumple la desigualdad $f(\alpha u + \beta v) \leq \alpha f(u) + \beta f(v)$ que escrita en la forma:

$$f(x) \leq f(u) + \beta(f(v) - f(u)) = f(u) + m(u,v)\beta(v-u) = f(u) + m(u,v)(x-u) = r(x)$$

pone de manifiesto que, para cada $[u,v] \subset |a,b|$, la recta que pasa por los puntos $(u, f(u))$, $(v, f(v))$, queda por encima de la gráfica de f, luego f es convexa. ■

OBSERVACIÓN: La condición [$C4$] es equivalente a la que resulta suponiendo $\alpha \geq 0$, $\beta \geq 0$, y $\alpha + \beta = 1$, ya que la desigualdad se convierte en una igualdad cuando $\alpha = 0$ ó $\beta = 0$. Por otra parte, en la formulación de [$C4$] basta suponer $u \neq v$, aunque no hay inconveniente en admitir que $u < v$.

Proposición F.2 *Sea* $f : |a,b| \to \mathbb{R}$ *una función convexa y* $x \in |a,b|$.
[$P1$] *Si* $x < b$ *la función* $t \to m(x,t)$ *es creciente en el intervalo* $(x,b|$.
[$P2$] *Si* $a < x$, *la función* $t \to m(x,t)$ *es creciente en el intervalo* $|a,x)$.
[$P3$] *Si* $[s,t] \subset |a,b|$, *y* $s < x < t$, *se cumple* $m(s,x) \leq m(s,t) \leq m(x,t)$.

DEM: La propiedad [$P1$] es consecuencia de [$C1$], y la propiedad [$P2$] es consecuencia de [$C2$] teniendo en cuenta que $m(x,t) = m(t,x)$. La propiedad [$P3$] es consecuencia de las propiedades [$P1$] y [$P2$] aplicadas en los intervalos $(s,b|$, y $|a,t)$, respectivamente (teniendo en cuenta que $m(s,t) = m(t,s)$ y $m(t,x) = m(t,x)$). ■ .

Proposición F.3 *Si $f : |a,b| \to \mathbb{R}$ es convexa se cumple:*

a) En cada $x \in (a,b)$ existen las derivadas laterales, $f_i'(x) \leq f_d'(x)$, y las funciones f_i', f_d' son crecientes en (a,b).

b) Las rectas tangentes por la izquierda y por la derecha a la gráfica de f en un punto genérico $(x, f(x))$, con $x \in (a,b)$, de ecuaciones

$$r_i(t) = f(x) + f_i'(x)(t-x), \quad r_d(t) = f(x) + f_d'(x)(t-x),$$

quedan por debajo de la gráfica de f, es decir $r_i(t) \leq f(t)$, y $r_d(t) \leq f(t)$ para todo $t \in (a,b)$.

DEM: Según F.2, la función $t \to m(x,t)$ es creciente en $(x,b|$ y eligiendo un punto $a < s < x$ obtenemos que $m(s,x)$ es una cota inferior de $\{m(x,t) : x < t < b\}$, luego existe y es finito el límite

$$f_d'(x) = \lim_{t \to x+} m(x,t) = \inf\{m(x,t) : x < t < b\} \geq m(s,x)$$

Como la desigualdad $m(s,x) \leq f_d'(x)$ es cierta para todo $s \in (a,x)$ y la función $s \to m(x,s) = m(s,x)$ es creciente en $|a,x)$, se sigue que existe y es finito el límite

$$f_i'(x) = \lim_{s \to x-} m(x,s) = \sup\{m(x,s) : a < s < x\} \leq f_d'(x)$$

Por otra parte, si $a < x < y < b$, en virtud de lo que acabamos de establecer,

$$f_i'(x) \leq f_d'(x) \leq m(x,y) = m(y,x) \leq f_i'(y) \leq f_d'(y)$$

y por lo tanto las funciones f_i', f_d' son crecientes en (a,b).

Las ecuaciones de las rectas tangentes, por la izquierda y por la derecha, a la gráfica de f en un punto genérico $(x, f(x))$, con $a < x < b$, se escriben así:

$$r_i(u) = f(x) + f_i'(x)(u-x); \quad r_d(u) = f(x) + f_d'(x)(u-x).$$

Si $a < s < x < t < b$, como $t - x > 0$ y $s - x < 0$, las desigualdades

$$f_d'(x) \leq m(x,t) = (f(t) - f(x))/(t-x); \quad f_i'(t) \leq f_d'(t)$$

$$f_i'(x) \geq m(x,s) = (f(s) - f(x)/(s-x); \quad f_i'(s) \leq f_d'(s)$$

permiten afirmar que

$$t \in (x,b) \Rightarrow f(t) \geq f(x) + f_d'(x)(t-x) = r_d(t) \geq r_i(t)$$

$$s \in (a,x) \Rightarrow f(s) \geq f(x) + f_i'(x)(s-x) = r_i(s) \geq r_d(s)$$

y queda establecido que $r_i(u) \leq f(u)$ y $r_d(u) \leq f(u)$ para cada $u \in (a,b)$. ∎

OBSERVACIÓN: En las condiciones de la proposición anterior si x es uno de los extremos del intervalo sólo se puede garantizar la existencia de la correspondiente

derivada lateral en sentido amplio (con valores en $[-\infty, +\infty]$):
Si $x = a \in |a, b|$ no podemos asegurar que $\{m(a,t) : a < t < b\}$ esté acotado inferiormente, pero debido al carácter creciente de la función $t \to m(a,t)$, existe el límite en la recta real ampliada

$$f'_d(a) = \lim_{t \to a+} m(a,t) = \inf\{m(a,t) : a < t < b\} \geq -\infty$$

Análogamente, si $x = b \in |a, b|$, existe la derivada en sentido amplio $f'_i(b) \leq +\infty$.

Según esto, haciendo intervenir derivadas laterales con valores infinitos en los extremos del intervalo que le pertenezcan, podemos afirmar que f'_i es creciente en $(a, b|$ y f'_d es creciente en $|a, b)$. Además, cuando a (resp. b) pertenece al intervalo $|a, b|$ y la derivada lateral $f'_d(a)$ (resp. $f'_i(b)$) es finita, también se cumple que la tangente lateral en $(a, f(a))$, (resp. $(b, f(b))$ queda por debajo de la gráfica de f.

Por consiguiente, si f es convexa y derivable en $|a, b|$ (con derivabilidad lateral en los extremos del intervalo $|a, b|$ que estén en el mismo) se cumple que la derivada f' es creciente en $|a, b|$ y para todo $x \in |a, b|$ la gráfica de f queda por encima de su tangente en el punto $(x, f(x))$.

Corolario F.4 *Toda función convexa $f : |a, b| \to \mathbb{R}$ es continua en $(a, b) \subset \mathbb{R}$.*

Dem: La función es continua por la izquierda y por la derecha en cada $x \in (a, b)$, porque, según la proposición F.3, existen las derivadas laterales $f'_i(x)$,$f'_d(x)$ ■

En las condiciones del corolario F.4 no se puede asegurar la continuidad en un extremo del intervalo: Es inmediato que la función $f : [a, b) \to \mathbb{R}$, definida por $f(a) = 1$ y $f(t) = 0$ si $a < t < b$, es convexa en $[a, b)$, pero no es continua en a.

Teorema F.5 *Una condición necesaria y suficiente para que $f : (a, b) \to \mathbb{R}$ sea convexa es que exista una función creciente $\varphi : (a, b) \to \mathbb{R}$ tal que*

$$f(y) - f(x) = \int_x^y \varphi(t)dt \quad \textit{para todo intervalo} \quad [x, y] \subset (a, b)$$

Dem: Necesidad: Demostraremos que si f es convexa la función creciente $\varphi = f'_d$ cumple la condición del enunciado.

Si $p = (t_0 < t_1 < \cdots t_n)$ con $t_0 = x$, $t_n = y$, es una subdivisión arbitraria del intervalo $[x, y] \subset (a, b)$, como f'_d es creciente, se tiene

$$s(f'_d, p) = \sum_{j=1}^{n} f'_d(t_{j-1})(t_j - t_{j-1}); \qquad S(f'_d, p) = \sum_{j=1}^{n} f'_d(t_j)(t_j - t_{j-1});$$

Por otra parte, según se ha visto en la demostración de la proposición F.3

$$f'_d(t_{j-1}) \leq m(t_{j-1}, t_j) \leq f'_d(t_j)$$

Multiplicando por $(t_j - t_{j-1}) > 0$, y sumando para $j = 1 \cdots n$, obtenemos

$$s(f'_d, p) \leq \sum_{j=1}^{n} m(t_j, t_{j-1})(t_j - t_{j-1}) \leq S(f'_d, p)$$

Teniendo en cuenta que

$$\sum_{j=1}^{n} m(t_j, t_{j-1})(t_j - t_{j-1}) = \sum_{j=1}^{n}(f(t_j) - f(t_{j-1}) = f(y) - f(x)$$

llegamos a que, para todo $p \in \mathcal{P}([x, y])$, se cumple

$$s(f'_d, p) \leq f(y) - f(x) \leq S(f'_d, p)$$

Como f'_d es integrable Riemann en $[x, y]$ (por ser creciente) se sigue que

$$\int_x^y f'_d(t)dt = f(y) - f(x)$$

Queda demostrado que $\varphi = f'_d$ cumple la condición del enunciado (con un razonamiento similar se puede ver que f'_i también la cumple).

Suficiencia: Sea $\varphi : (a, b) \to \mathbb{R}$ una función creciente tal que para todo $[x, y] \subset (a, b)$

$$f(y) - f(x) = \int_x^y \varphi(t)dt$$

Si $s = \alpha x + \beta y$, con $\alpha > 0, \beta > 0, \alpha + \beta = 1$, teniendo en cuenta que el máximo valor de φ en $[x, s]$ es $\varphi(s)$ y el mínimo valor de φ en $[s, y]$ es $\varphi(s)$, se obtienen las desigualdades

$$m(x, s) = \frac{1}{s - x}\int_x^s \varphi(t)dt \leq \varphi(s) \leq \frac{1}{y - s}\int_s^y \varphi(t)dt = m(s, y)$$

Queda establecido así que en todo intervalo $[x, y] \subset (a, b)$ cumple la condición $[C3]$ de la proposición F.1, y por lo tanto f es convexa. ■

Corolario F.6 *Si $f : (a, b) \to \mathbb{R}$ es convexa el conjunto de puntos donde no es derivable, $\{x \in (a, b) : f'_i(x) < f'_d(x)\}$, es numerable.*

Dem: Es bien conocido que el conjunto de los puntos de discontinuidad $D(\varphi)$ de una función creciente $\varphi : (a, b) \to \mathbb{R}$ es numerable. Si φ es la función creciente que interviene en el teorema F.5, en virtud del teorema fundamental del cálculo podemos asegurar que la función $f(s) = f(x) + \int_x^s \varphi(t)dt$ es derivable en cada $s \in (a, b) \setminus D(\varphi)$. ■

En el siguiente teorema, cuando $a \in |a, b|$, (resp. $b \in |a, b|$) la derivabilidad de la función en a (resp. en b) significa derivabilidad por la derecha (resp. por la izquierda) y la tangente a la gráfica en $(a, f(a))$ (resp. $(b, f(b))$ es la correspondiente tangente lateral.

Teorema F.7 *Para una función derivable $f : |a, b| \to \mathbb{R}$ son equivalentes* *a)* f *es convexa.*

b) La gráfica de f *queda por encima de su tangente en cualquier punto* $(x, f(x))$*, con* $x \in |a, b|$*.*

c) La derivada f' *es creciente en* $|a, b|$*.*

Dem: a) $\Rightarrow$ b) es consecuencia directa de la proposición F.3, teniendo en cuenta las observaciones que le siguen.
b $\Rightarrow$ c): Si $x \in |a, b|$, la ecuación de la recta tangente a la gráfica de f en el punto $(x, f(x))$ es $r(t) = f(x) + f'(x)(t - x)$, y si suponemos que $r(t) \leq f(t)$ para todo $t \in |a, b|$ se obtiene que

$$f'(x)(t - x) \leq f(t) - f(x)$$

Análogamente, cambiando los papeles de t y x podemos escribir

$$f'(t)(x - t) \leq f(x) - f(t)$$

Si $a \leq x < t < b$ se sigue que $f'(x)(t - x) \leq f(t) - f(x) \leq f'(t)(t - x)$, luego $f'(x) \leq f'(t)$. Análogamente, si $a < t < x \leq b$ resulta $f'(t) \leq f'(x)$.
c) $\Rightarrow$ a): Suponemos ahora que la derivada f' es creciente en $|a, b|$. Dado un intervalo $[u, v] \subset |a, b|$, si consideramos la recta que pasa por $(u, f(u))$, y $(v, f(v))$, de ecuación $r(x) = f(u) + m(u, v)(x - u)$, basta demostrar que $\varphi(x) = r(x) - f(x) \geq 0$ para todo $x \in [u, v]$. La función φ es continua y derivable en $[u, v]$ con derivada decreciente

$$\varphi'(x) = r'(x) - f'(x) = m(u, v) - f'(x)$$

Como $\varphi(u) = \varphi(v) = 0$, en virtud del teorema de Rolle, existe $\eta \in (u, v)$ tal que $\varphi'(\eta) = 0$. Como φ' es decreciente podemos afirmar que $\varphi'(s) \geq 0$ si $u \leq s \leq \eta$, y $\varphi'(t) \leq 0$ si $\eta \leq t \leq v$, luego φ es creciente en $[u, \eta]$ y decreciente en $[\eta, v]$, de donde se sigue que $\varphi(x) \geq 0$ para todo $x \in [u, v]$.
■

observación: Para un intervalo abierto (a, b), la implicación c) $\Rightarrow$ a) en el teorema F.7 también se puede obtener utilizando la caracterización integral de las funciones convexas (F.5): Si f' es creciente entonces es integrable en cada intervalo $[x, y] \subset (a, b)$ y por lo tanto

$$f(y) - f(x) = \int_x^y f'(t)dt \quad \text{para cada } \ [x, y] \subset (a, b)$$

Corolario F.8 *Una función dos veces derivable* $f : |a, b| \to \mathbb{R}$ *es convexa si y sólo si* $f''(t) \geq 0$ *para cada* $t \in |a, b|$*.*

Dem: Es consecuencia directa de la equivalencia a) $\Leftrightarrow$ c) en el teorema F.7. ■

nota: Si en la definición de función convexa se cambian las desigualdades $\leq$ por desigualdades estrictas $<$ se obtiene la noción de función estrictamente convexa. En este caso las funciones $t \to m(x, t)$ son estrictamente crecientes en los intervalos donde están definidas. Para una función derivable $f : (a, b) \to \mathbb{R}$ son equivalentes

a) f es estrictamente convexa.

b) f' es estrictamente creciente.

c) La gráfica de f queda estrictamente por encima de su tangente en un punto genérico $(x, f(x))$ (excepto en dicho punto).

Si $f : (a, b) \to \mathbb{R}$ es derivable dos veces en (a, b) y $f''(x) > 0$ para todo $x \in (a, b)$ entonces f es estrictamente convexa porque su derivada es estrictamente creciente, pero el recíproco no es cierto (la función $f(x) = x^4$ es estrictamente convexa en $\mathbb{R}$ pero $f''(0) = 0$).

F.2. Continuidad de las funciones convexas de varias variables

Al lector interesado le presentamos una demostración de la continuidad de las funciones convexas de varias variables. Este atractivo resultado con un enunciado tan simple no suele figurar en los textos usuales de cálculo para funciones de varias variables reales. Requiere algunos lemas preliminares de carácter técnico.

Si E es un espacio vectorial (sobre el cuerpo $\mathbb{R}$), para cada par de puntos $\mathbf{x}, \mathbf{y} \in E$ sea $[\mathbf{x}, \mathbf{y}] = \{(1-t)\mathbf{x} + t\mathbf{y} : 0 \leq t \leq 1\}$ el segmento que los une. Recordemos que un conjunto $A \subset E$ es convexo cuando $[\mathbf{x}, \mathbf{y}] \subset A$ para cada par de puntos $\mathbf{x}, \mathbf{y} \in A$. Las siguientes propiedades se obtienen mediante comprobaciones rutinarias que se dejan al cuidado del lector:

a) La intersección de cualquier familia (finita o no) de conjuntos convexos es un conjunto convexo.

b) Si $A, B \subset E$ son convexos y $\lambda \in \mathbb{R}$ entonces también son convexos

$$A + B = \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in A,\ \mathbf{b} \in B\}; \quad \lambda A = \{\lambda \mathbf{a} : \mathbf{a} \in A\}.$$

Si E es un espacio vectorial sobre $\mathbb{R}$, la envoltura convexa de $B \subset E$, denotada $\mathrm{co}(B)$ es la intersección de todos los conjuntos convexos que contienen a B. Según a), $\mathrm{co}(B)$ es convexo y por lo tanto es el mínimo convexo que contiene a B.

Lema F.9 *Si $A \subset E$ es convexo y $t_1 \geq 0, t_2 \geq 0, \cdots t_m \geq 0$ entonces*

$$t_1 A + t_2 A + \cdots + t_m A = \left(\sum_{i=1}^{m} t_i \right) A$$

DEM: Basta demostrar la inclusión $t_1 A + t_2 A + \cdots + t_m A \subset \left(\sum_{i=1}^{m} t_i\right) A$, ya que la otra inclusión $\supset$ es inmediata. Lo haremos por inducción sobre el número de sumandos. El resultado es trivial para $m = 1$. Cuando $m = 2$ podemos suponer $t_1 + t_2 > 0$ y se obtiene la inclusión

$$t_1A + t_2A = (t_1 + t_2)\left(\frac{t_1}{t_1 + t_2}A + \frac{t_2}{t_1 + t_2}A\right) \subset (t_1 + t_2)A$$

como consecuencia de la definición de conjunto convexo, ya que

$$\frac{t_1}{t_1 + t_2} + \frac{t_1}{t_1 + t_2} = 1.$$

Si suponemos cierta la inclusión para $m = k - 1$, tenemos

$$t_1A + t_2A + \cdots + t_{k-1}A + t_kA \subset \left(\sum_{i=1}^{k-1} t_i\right) A + t_kA \subset \left(\sum_{i=1}^{k} t_i\right) A$$

donde la primera inclusión se cumple por hipótesis de inducción y la segunda por lo que ha sido demostrado en el caso $m = 2$. ■

El resultado que se acaba de establecer en el lema F.9 es falso cuando A no se supone convexo: Si $E = \mathbb{R}$ el conjunto $A = \{0, 1\}$ verifica $2A = \{0, 2\} \subsetneq \{0, 1, 2\} = A+A$.

Cada expresión de la forma $\sum_{i=1}^m t_i\mathbf{b}_i$ donde $t_i \geq 0$ y $\sum_{i=1}^m t_i = 1$ se dice que es una combinación convexa de los vectores $\mathbf{b}_1, \mathbf{b}_2, \cdots \mathbf{b}_m$. El siguiente lema se puede enunciar diciendo que la envoltura convexa de $B \subset E$ está formada por el conjunto de las combinaciones convexas de elementos de B.

Lema F.10 *Si E es un espacio vectorial sobre $\mathbb{R}$ y $B \subset E$, se verifica*

$$co(B) = \left\{\sum_{i=1}^m t_i\mathbf{b}_i : \ \mathbf{b}_i \in B, \ t_i \geq 0, \ \sum_{i=1}^m t_i = 1, \ m \in \mathbb{N}\right\}$$

DEM: Según F.9 cada convexo $A \supset B$ contiene a todas las combinaciones convexas de elementos de B pues

$$\sum_{i=1}^m t_i\mathbf{b}_i \in t_1A + t_2A + \cdots t_mA = (\sum_{i=1}^m t_i)A = A$$

Por lo tanto basta comprobar que el conjunto de las combinaciones convexas de B es convexo. Dadas dos combinaciones convexas de elementos de B

$$\mathbf{x} = \sum_{i=1}^m t_i\mathbf{b}_i; \quad \mathbf{y} = \sum_{i=1}^m t'_i\mathbf{b}'_i$$

es claro que, para cada $t \in [0, 1]$, el vector $(1 - t)\mathbf{x} + t\mathbf{y}$ también se puede expresar como combinación convexa de elementos de B ya que

$$(1 - t)\mathbf{x} + t\mathbf{y} = \sum_{i=1}^m (1 - t)t_i\mathbf{b}_i + \sum_{i=1}^m tt'_i\mathbf{b}'_i$$

donde

$$(1-t)t_i \geq 0; \quad tt'_i \geq 0; \quad \sum_{i=1}^{m}(1-t)t_i + \sum_{i=1}^{m} tt_i = 1$$

■

Lema F.11 *Para $r > 0$ sea $B_n(r) = [-r, r]^n \subset \mathbb{R}^n$ la bola cerrada de centro 0 y radio $r > 0$ para la norma $\| \|_\infty$ de $\mathbb{R}^n$. Se cumple que $B_n(r) = co(V_n(r))$ donde*

$$V_n(r) = \{(v_1, v_2 \cdots v_n) : |v_1| = r, |v_2| = r, \cdots |v_n| = r\}$$

es el conjunto de los vértices de $B_n(r)$.

DEM: El resultado es inmediato para $n = 1$ y para hacer la demostración por inducción sobre la dimensión, lo suponemos cierto en $\mathbb{R}^{n-1}$. En lo que sigue cada vector $\mathbf{x} \in \mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$ lo representamos en la forma $\mathbf{x} = (\mathbf{y}, t)$ donde $\mathbf{y} \in \mathbb{R}^{n-1}$ y $t \in \mathbb{R}$.

Dado $\mathbf{x} = (\mathbf{y}, t) \in B_n(r) = B_{n-1}(r) \times [-r, r]$, por la validez del resultado para $n = 1$ podemos escribir $t = \alpha(-r) + \beta r$ con $\alpha \geq 0$, $\beta \geq 0$, y $\alpha + \beta = 1$, luego $\mathbf{x} = \alpha\mathbf{p} + \beta\mathbf{q}$ donde $\mathbf{p} = (\mathbf{y}, -r)$, $\mathbf{q} = (\mathbf{y}, r)$. Según la hipótesis de inducción $\mathbf{y} \in \mathbb{R}^{n-1}$ se puede expresar como una combinación convexa

$$\mathbf{y} = \sum_{\mathbf{v} \in V_{n-1}(r)} t(\mathbf{v})\mathbf{v}$$

de vectores de $V_{n-1}(r)$.

Con los mismos coeficientes $t(\mathbf{v})$ podemos expresar $\mathbf{p} = (\mathbf{y}, -r)$, y $\mathbf{q} = (\mathbf{y}, r)$ como combinaciones convexas

$$\mathbf{p} = \sum_{\mathbf{v} \in V_{n-1}(r)} t(\mathbf{v})(\mathbf{v}, -r); \quad \mathbf{q} = \sum_{\mathbf{v} \in V_{n-1}(r)} t(\mathbf{v})(\mathbf{v}, +r)$$

Obsérvese que $V_n(r) = V_n(r)^+ \cup V_n(r)^-$ donde

$$V_n(r)^+ = \{(\mathbf{v}, +r) : \mathbf{v} \in V_{n-1}(r)\}, \quad V_n(r)^- = \{(\mathbf{v}, -r) : \mathbf{v} \in V_{n-1}(r)\}$$

Si para $\mathbf{w} = (\mathbf{v}, \pm r) \in V_n(r)$ definimos $s(\mathbf{w}) = t(\mathbf{v})$, es claro que

$$\mathbf{x} = \alpha\mathbf{p} + \beta\mathbf{q} = \sum_{\mathbf{w} \in V_n(r)^+} \alpha s(\mathbf{w})\mathbf{w} + \sum_{\mathbf{w} \in V_n(r)^-} \beta s(\mathbf{w})\mathbf{w}$$

y así se obtiene $\mathbf{x}$ como combinación convexa de vectores de $V_n(r)$. ■

Teorema F.12 *Toda función convexa $f : \Omega \to \mathbb{R}$ definida en un abierto convexo $\Omega \subset \mathbb{R}^n$ es continua.*

DEM: Comencemos con una sencilla observación: Si $\mathbf{p} = \sum_{j=1}^{m} t_j \mathbf{p}_j$ es una combinación convexa de puntos $\mathbf{p}_j \in \Omega$ se demuestra fácilmente (por inducción sobre m) que

$$f(\mathbf{p}) \leq \sum_{j=1}^{m} t_j f(\mathbf{p}_j)$$

Fijado un punto $\mathbf{a} \in \Omega$, para demostrar que f es continua en $\mathbf{a}$ basta demostrar que es continua en 0 la función convexa $g(\mathbf{x}) = f(\mathbf{x} + \mathbf{a}) - f(\mathbf{a})$, que está definida en el abierto convexo $\Omega_{\mathbf{a}} = -\mathbf{a} + \Omega$. Puesto que $0 \in \Omega_{\mathbf{a}}$ existe $r > 0$ tal que $B_n(r) = [-r, r]^n \subset \Omega_{\mathbf{a}}$. Según el lema F.11 cada $\mathbf{x} \in B_n(r)$ se puede escribir como combinación convexa de los vértices de $B_n(r)$:

$$\mathbf{x} = \sum_{\mathbf{v} \in V_n(r)} t(\mathbf{v})\mathbf{v}$$

Como el conjunto de vértices $V_n(r)$ es finito (con 2^n elementos) podemos considerar el máximo $M_r = \text{máx}\{g(\mathbf{v}) : \mathbf{v} \in V_n(r)\}$, y en virtud de la observación preliminar, aplicada a la función g, resulta

$$g(\mathbf{x}) \leq \sum_{\mathbf{v} \in V_n(r)} t(\mathbf{v})g(\mathbf{v}) \leq \sum_{\mathbf{v} \in V_n(r)} t(\mathbf{v})M_r = M_r$$

luego M_r es una cota superior de g en $B_n(r)$. Si $0 < \epsilon < 1$, y $\mathbf{x}/\epsilon \in B_n(r)$, en virtud de la convexidad de g se obtiene

$$g(\mathbf{x}) \leq (1 - \epsilon)g(\mathbf{0}) + \epsilon g(\mathbf{x}/\epsilon) = \epsilon g(\mathbf{x}/\epsilon) \leq \epsilon M_r$$

Usando otra vez la convexidad de g podemos escribir

$$0 = g(\mathbf{0}) \leq (1/2)g(-\mathbf{x}) + (1/2)g(\mathbf{x})$$

luego, $-g(\mathbf{x}) \leq g(-\mathbf{x}) \leq \epsilon M_r$ y se concluye que $\mathbf{x}/\epsilon \in B_n(r) \Rightarrow |g(\mathbf{x})| \leq \epsilon M_r$ es decir

$$\|\mathbf{x}\|_\infty \leq \epsilon r \Rightarrow |g(\mathbf{x})| \leq \epsilon M_r$$

y así queda demostrado que g es continua en $\mathbf{0}$. ■