

FACULTAD DE CIENCIAS

Introducción al análisis y la geometría de las funciones convexas y aplicaciones: medidas y gradientes

Victor GARZÓN SÁNCHEZ

Tutor: Fernando Sanz Sánchez



Universidad de Valladolid

Máster de Investigación en Matemáticas

Curso 2013/14

1ª Convocatoria. Junio 2014

*R*ecuerda Sa, tenemos todo
el tiempo del mundo... aunque
me parecerá poco a tu lado.
Te amo, ahora y siempre

Índice general

Introducción	VII
1. Funciones convexas reales unidimensionales	1
1.1. Definiciones y primeros resultados	1
1.1.1. Caracterizaciones de la convexidad	3
1.2. Derivabilidad de funciones convexas	6
1.2.1. Derivadas laterales y continuidad	6
1.2.2. Extremos de funciones convexas	9
La derivada simétrica de orden dos	12
1.2.3. Un par de ejemplos: <i>construcción de convexas</i>	15
1.3. La <i>subdiferencial</i> en una dimensión	20
2. Desigualdades notables	22
2.1. Desigualdad de Hardy-Littlewood-Polya	22
2.2. Desigualdad de Jensen-Steffensen	24
2.3. Desigualdad de Hermite-Hadamard	25
2.4. Fórmula <i>integral</i> de la desigualdad de Jensen	26
2.5. Desigualdad de Hardy	29
3. Ejemplos clásicos. Las funciones Γ y β	32
3.1. La clase de las funciones logarítmicamente convexas	32
3.2. Estudio de las funciones Gamma y Beta	33
3.2.1. La función Gamma	33
3.2.2. La función Beta	44
3.2.3. Más expresiones con Gamma y Beta	45
4. Funciones convexas en varias dimensiones	47
4.1. Conjuntos convexos	47
4.1.1. Hiperplanos (soporte) y teoremas de separación	48
4.2. Funciones convexas en dimensión superior	49
4.2.1. Propiedades de regularidad	51
4.3. La subdiferencial en dimensión superior	52
5. Desigualdades de Brunn-Minkovsky e Isoperimétrica	56
5.1. Desigualdad de Brunn-Minkovsky (motivación y enunciado)	57
5.2. Desigualdad Isoperimétrica	58
5.3. Prueba de la desigualdad de Brunn-Minkovsky	60
5.3.1. Primera demostración	60

5.3.2. Desigualdad de Prékopa-Leindler	62
5.3.3. Segunda demostración	64
6. La desigualdad de Lojasiewicz para funciones convexas. Aplicación a sistemas dinámicos	65
6.1. Conjuntos subanalíticos y funciones subanalíticas	65
6.2. La desigualdad de Lojasiewicz	67
6.3. Aplicación a sistemas dinámicos	69
A. Resultados sobre Teoría de la medida	71
A.1. Desigualdades notables en espacios de medida generales	74
A.1.1. La desigualdad de Young	74
A.1.2. La desigualdad de Hölder	78
A.1.3. La desigualdad de Minkovsky	80
Bibliografía	83

Introducción

Los conceptos de función convexa y conjunto convexo son muy naturales y no requieren muchos conocimientos previos. Es por eso que se introducen desde el comienzo de los estudios de matemáticas preuniversitarios. También porque su uso es ubicuo en multitud de ramas de las matemáticas, tanto a nivel teórico como práctico. Así, por ejemplo, aparte de sus aplicaciones más evidentes al análisis matemático, aparecen en la geometría más clásica de polihedros y su uso para la topología algebraica y la teoría de la medida; o también en la geometría algebraica, particularmente en la geometría tórica. En el terreno práctico, la convexidad es la pieza clave en optimización, programación lineal, teoría de control, etc., disciplinas cuyos resultados tienen aplicación directa a la industria, economía y medicina, entre otras.

En este trabajo presentamos una introducción a la teoría de las funciones convexas y damos una muestra, necesariamente muy reducida, pero diversa, de su vasta aplicabilidad. Hemos recogido algunos de los conceptos más básicos y algunas demostraciones de ciertos resultados que dan idea de lo potentes que a veces son los argumentos sencillos cuando se tienen las buenas definiciones. Por mantener una extensión razonable, muchos resultados se citan sin detalles sobre las demostraciones, pero hemos intentado enviar al lector a las referencias (de entre la gran mayoría de libros y textos sobre geometría convexa o temas relacionados que existen) que más adecuadas nos han parecido en cada caso.

Las aplicaciones que aparecen son mayoritariamente en el terreno del análisis matemático, en el que la convexidad y el concepto de *subdiferencial* reemplaza muchas veces la derivabilidad y la diferencial y ofrece resultados importantes en el análisis de funciones “no regulares”, *desigualdades notables* y funciones especiales como las *funciones Gamma y Beta*, a las que dedicamos un capítulo. Por otro lado, presentamos la *desigualdad isoperimétrica* como una aplicación geométrica clásica a la teoría de la medida de Lebesgue. Ésta establece la generalización a dimensión superior de una propiedad geométrica bastante plausible y conocida: el círculo es la figura en el plano de mayor área de entre las que tienen el mismo perímetro.

Por último, recogemos una aplicación mucho más reciente y más específica de los sistemas dinámicos, que entronca con las líneas de investigación del equipo de investigación ECSING de la Universidad de Valladolid al que pertenece el tutor de este trabajo. Se trata de la *desigualdad de Lojasiewicz* para funciones convexas semi-continuas y subanalíticas. Este nombre se debe a que generaliza para las funciones mencionadas una conocida desigualdad debida a Lojasiewicz para funciones analíticas reales en el entorno de un punto crítico (en nuestro caso la desigualdad tiene un carácter global pues las funciones convexas presentan mínimo global).

Precisando un poco más, la desigualdad de Lojasiewicz para una función analíti-

ca, establece que los valores de la función en un entorno del punto crítico, en valor absoluto, están acotados por una cierta potencia de exponente mayor que uno del módulo del gradiente de la función. Entre otras cosas, esta desigualdad permite probar la longitud finita de trayectorias del campo de vectores gradiente de la función analítica. En nuestro caso, esta consecuencia también es válida, siempre que definamos correctamente el sentido del sistema dinámico (generalizado) definido por el (sub)gradiente de una función semicontinua, convexa y subanalítica. A este objetivo empleamos buena parte del último capítulo de esta memoria.

Queremos señalar por último que pensamos que la desigualdad de Lojasiewicz para funciones convexas subanalíticas podría ser útil para tratar el *problema del gradiente de funciones analíticas*, nombre éste que resume ciertas cuestiones abiertas tratadas en el seno del equipo ECSING, como son la existencia de tangente para las trayectorias o su carácter *no-oscilante*. El argumento que sustenta esta esperanza es el siguiente: un campo de vectores gradiente, tras un proceso de reducción de singularidades por explosiones, deja de ser gradiente pues la métrica degenera a lo largo del divisor; bajo ciertas hipótesis, una modificación de la función podría hacer pensar en dicho transformado como un sistema sub-gradiente al que los resultados anteriores pueden aplicarse.

A continuación analizamos con un poco más de detalle los contenidos del trabajo capítulo por capítulo.

En el primer capítulo tratamos en profundidad las funciones convexas de una variable real: operaciones entre ellas, caracterizaciones de la convexidad, puntos extremos, monotonía, a la vez que establecemos unas primeras desigualdades básicas de renombre como son la forma discreta de la desigualdad de Jensen, la desigualdad de las medias (MGA_n) y sus implicaciones o la de Popoviciu. También nos ocupamos de su derivabilidad mediante el lema de Galvani y el posterior teorema de Stolz, que recoge lo necesario a saber acerca de la derivabilidad, y la convergencia de sucesiones de derivadas de funciones convexas. Introducimos la derivada simétrica de orden dos para llegar a probar el test de la segunda derivada como una forma rápida de comprobar la convexidad de una función real. Más adelante, damos dos ejemplos concretos y constructivos de funciones convexas no derivables sobre conjuntos numerables, uno de ellos incluso denso. Acabamos el capítulo introduciendo el concepto de subdiferencial y algunos resultados importantes concernientes a él.

El segundo capítulo, como su propio título indica, versa sobre desigualdades destacables que ponen de manifiesto la importancia de los temas tratados hasta el momento. La primera desigualdad hace uso del concepto de subdiferencial, dándonos una idea de su importancia. La segunda es una generalización clásica de la desigualdad de Jensen que usa combinaciones lineales (no necesariamente convexas) y para la prueba de la cual se necesita de la primera desigualdad del capítulo. La tercera simplemente pondera entre que dos valores se encuentra la media integral en un intervalo de una función convexa. Esta desigualdad se usa más adelante en el texto. La cuarta desigualdad trata sobre la forma integral de la desigualdad de Jensen, usando para ello medias integrales sobre espacios de medida arbitrarios. Para finalizar, mostramos la desigualdad de Hardy como aplicación de la anterior.

El tercer capítulo está dedicado íntegramente a estudiar detalladamente las funciones Gamma y Beta reales, estableciendo, entre otras cosas, varias fórmulas de renombre de aplicación a campos como la estadística o la teoría de números.

En el cuarto capítulo tratamos con las funciones convexas reales en dimensión superior, para lo cual necesitamos recordar “cosas” de conjuntos convexas, teoremas

de separación, hiperplanos soporte, etc. Aquí aparecen también conceptos más sofisticados, como por ejemplo, funciones de valores reales extendidos, semicontinuidad inferior, envolventes de Moreau (suma epigráfica), y como no, la subdiferencial para funciones de varias dimensiones.

En el capítulo cinco se trata la desigualdad de Brunn-Minkovsky en R^n , dándose un par de demostraciones alternativas, y la desigualdad isoperimétrica como aplicación de ésta, de la que ya hemos hablado un poco más arriba. Para ello se usa parte de la teoría de la medida de Lebesgue.

El sexto y último capítulo trata sobre la desigualdad de Lojasiewicz y su posible aplicación a sistemas dinámicos gradientes, como ya hemos indicado antes. Para poder establecerla y probarla se necesita lo básico de conceptos como la subanaliticidad de conjuntos y funciones reales, que tratamos en primer lugar en el capítulo.

El apéndice simplemente es una recopilación de resultados acerca de la teoría de la medida, a los que se hace referencia en algunos momentos a lo largo del trabajo. Son cosas de sobra conocidas que sólo se pretenden recordar brevemente y no están dentro del texto para no estropear la claridad del mismo. Esto no quiere decir que los resultados que aquí aparecen sean sencillos o irrelevantes, de hecho algunos de ellos son de gran importancia y se usan en multitud de áreas de las matemáticas. Es el caso de la desigualdad de Rogers-Hölder o de Minkovsky para espacios de medida arbitrarios.

Capítulo 1

Funciones convexas reales unidimensionales

1.1. Definiciones y primeros resultados

Empecemos tratando las funciones de una variable real. De aquí en adelante, mientras no se diga lo contrario, $I \subseteq \mathbb{R}$ será un intervalo de interior no vacío y f denotará una función de I sobre \mathbb{R} . La siguiente definición engloba los conceptos básicos.

Definición 1.1.1. f es *convexa* si $f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y)$ para cualesquiera $x, y \in I$ y $\lambda \in [0, 1]$; y *estrictamente convexa* si la desigualdad es estricta con $\lambda \in (0, 1)$ y x e y distintos. Se dice f *cóncava* si $-f$ es convexa. Análogamente con *estrictamente cóncava*.

Ejemplos básicos de funciones estrictamente convexas son las funciones elementales e^x , x^2 ó $\log(1/x)$, como puede verse usando el cálculo diferencial elemental.

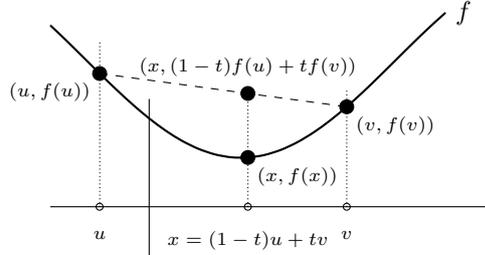
Proposición 1.1.1 (Operaciones con funciones convexas). *Se tienen las siguientes propiedades (la demostración es rutinaria):*

1. La suma o el producto por escalares de funciones convexas (resp. estrictamente convexas) es convexa (resp. estrictamente convexas).
2. La restricción a un subintervalo de una función convexa (resp. estrictamente convexa) es convexa (resp. estrictamente convexa).
3. Si f es convexa (resp. estrictamente convexa) y g es convexa creciente (resp. estrictamente creciente), entonces $g \circ f$ es convexa (resp. estrictamente convexa).
4. Sea f una biyección, si es estrictamente creciente entonces f es convexa (resp. estrictamente convexa) si f^{-1} es cóncava (resp. estrictamente cóncava). Además, si f es biyección decreciente, f y f^{-1} son del mismo tipo de convexidad; es decir, o las dos son convexas o las dos son cóncavas.

Gráficamente, es sencillo ver que f es convexa si y sólo si $\forall u, v \in I, u < v$, la gráfica de $f|_{[u,v]}$ está por debajo (ó coincide con, aunque sólo sea en algunos puntos)

el segmento que une $(u, f(u))$ con $(v, f(v))$. Esto es,

$$f(x) \leq \frac{f(v) - f(u)}{v - u} \cdot (x - u) + f(u) \quad \forall u, v \in I, u < v \quad \forall x \in [u, v]$$



Consecuentemente, las funciones convexas están acotadas en cualquier subintervalo compacto por funciones afines. Así, si una función es a la vez cóncava y convexa entonces es afín.

Teorema 1.

f es convexa si $\forall J \subseteq I$ subintervalo compacto y $\forall \varphi$ función afín, el supremo de $f + \varphi$ en J se alcanza en un punto extremo de J .

Demostración. $[\Rightarrow]$ Defino $F := f + \varphi$, que es convexa porque la suma de convexas lo es. Sea $J = [x, y] \subseteq I$, $z_\lambda := (1 - \lambda)x + \lambda y \in J$ para todo $\lambda \in [0, 1]$. Ocurre que

$$\sup_{z_\lambda \in J} F(z_\lambda) \leq \sup_{\lambda \in [0, 1]} \{(1 - \lambda)F(x) + \lambda F(y)\}$$

Ahora bien, como $[0, 1]$ es un compacto y la función $\psi(\lambda) := \lambda(F(y) - F(x)) + F(x)$ es afín, su máximo es $\max\{\psi(0), \psi(1)\}$ y deducimos entonces $\sup_{z_\lambda \in J} F(z_\lambda) = \max\{F(x), F(y)\}$.

$[\Leftarrow]$ Si $J := [x, y] \subseteq I$, existe una función afín L tal que $f(x) = L(x)$ y $f(y) = L(y)$. Por hipótesis, $\sup_{\lambda \in [0, 1]} \{(f - L)((1 - \lambda)x + \lambda y)\} = 0$; entonces

$$0 \geq f((1 - \lambda)x + \lambda y) - (1 - \lambda)f(x) - \lambda f(y) \quad \forall \lambda \in [0, 1]$$

luego f es convexa. \square

Corolario 1.1.2. *Toda función convexa es acotada en cada subintervalo compacto $[u, v]$ de su intervalo de definición. Aún más, como $f(x) \leq M \quad \forall x \in [u, v]$ con $M = \max\{f(u), f(v)\}$ tenemos que con $|t| \leq (v - u)/2$*

$$f((u + v)/2 + t) \geq 2f((u + v)/2) - f((u + v)/2 - t) \geq 2f((u + v)/2) - M$$

Otro de los resultados básicos es el **caso discreto de la desigualdad de Jensen**, que nos da una caracterización en términos de combinaciones convexas arbitrarias. Para ello necesitamos recordar que los intervalos son cerrados bajo dichas combinaciones.

Proposición 1.1.3 (Desigualdad de Jensen discreta). *f es convexa sii para cualesquiera $x_1, \dots, x_n \in I$, $\lambda_1, \dots, \lambda_n \in [0, 1]$ con $\sum \lambda_k = 1$ se tiene que*

$$f\left(\sum_{k=1}^n \lambda_k x_k\right) \leq \sum_{k=1}^n \lambda_k f(x_k)$$

Demostración. [\Leftarrow] Trivial. Basta tomar $n = 2$.

[\Rightarrow] Inducción. $n = 1$ es trivial, lo suponemos cierto para $n - 1$ y lo probamos para n . Separamos la suma como sigue

$$f\left(\sum_{k=1}^n \lambda_k x_k\right) = f\left(\lambda_1 x_1 + \sum_{k=2}^n \lambda_k x_k\right) = (*)$$

Si $\lambda_1 = 1$ hemos terminado, pues todos los demás son cero y la desigualdad es clara. Mientras que si es distinto de 1

$$\begin{aligned} (*) &= f\left(\lambda_1 x_1 + (1 - \lambda_1) \sum_{k=2}^n \frac{\lambda_k}{1 - \lambda_1} \cdot x_k\right) \\ &\leq \lambda_1 f(x_1) + (1 - \lambda_1) f\left(\sum_{k=2}^n \frac{\lambda_k}{1 - \lambda_1} \cdot x_k\right) \\ &\leq \lambda_1 f(x_1) + (1 - \lambda_1) \sum_{k=2}^n \frac{\lambda_k}{1 - \lambda_1} f(x_k) = \sum \lambda_k f(x_k) \quad \square \end{aligned}$$

1.1.1. Caracterizaciones de la convexidad

Vamos a proponer ahora algunos criterios clásicos para la convexidad de funciones continuas. El primero de ellos se debe a Jensen.

Teorema 2 (JENSEN, 1906).

Si f es continua, entonces es convexa sii es punto-medio convexa; esto es que

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} \quad \forall x, y \in I$$

Demostración. [\Rightarrow] Obvio. Basta tomar $\lambda = 1/2$ en la definición de convexidad.

[\Leftarrow] Reducción al absurdo. Supongamos que f es punto-medio convexa y que no es convexa, esto es que $\exists [a, b] \subseteq I$ tal que la gráfica de $f|_{[a,b]}$ no queda por debajo del segmento que une los puntos $(a, f(a))$ y $(b, f(b))$. Luego la función φ definida por

$$\varphi(x) := f(x) - \frac{f(b) - f(a)}{b - a} \cdot (x - a) - f(a) \quad \forall x \in [a, b]$$

verifica que $\gamma := \sup\{\varphi(x) | x \in [a, b]\} > 0$. Además, φ es continua por serlo f , $\varphi(a) = \varphi(b) = 0$, y es punto-medio convexa. Como se comprueba fácilmente.

Definimos $c := \inf_{[a,b]} \{x : \varphi(x) = \gamma\}$, entonces $\varphi(c) = \gamma$ y necesariamente $c \in (a, b)$. Sea $h > 0$ tal que $c \pm h \in (a, b)$, luego $\varphi(c-h) < \varphi(c)$ y $\varphi(c+h) \leq \varphi(c)$. Así,

$$\varphi(c) > \frac{\varphi(c-h) + \varphi(c+h)}{2}$$

lo que contradice la propiedad de ser punto-medio convexa. Concluimos que f es convexa. \square

Observación. El resultado es falso si quitamos la hipótesis de continuidad, pues se puede dar ejemplos de funciones punto-medio convexas que no son continuas en el interior de su intervalo de definición, y por lo tanto no pueden ser convexas. Pues como veremos, *convexidad implica continuidad*.

Corolario 1.1.4. f convexa sii $f(x+h) + f(x-h) - 2f(x) \geq 0$ para todo $x \in I$ y todo $h > 0$ con $x \pm h \in I$.

Observación. Hay versiones de estos resultados para convexidad estricta, simplemente las desigualdades deben serlo.

EJEMPLO. Usando el corolario anterior podemos dar una demostración alternativa que no usa el cálculo diferencial de que e^x es estrictamente convexa. La conocida desigualdad *media geométrica-media aritmética*

$$\forall a, b > 0, a \neq b, \quad \sqrt{ab} < (a+b)/2 \quad (MGA_2)$$

junto con dicho corolario nos da el resultado. En efecto, si $x \in \mathbb{R}$, $h > 0$, entonces $e^{x+h} \neq e^{x-h}$ y $\frac{1}{2}(e^{x+h} + e^{x-h}) > \sqrt{e^{x+h}e^{x-h}} = e^x$.

Como curiosidad, usando esto podemos demostrar la **desigualdad de las medias** (MGA_n) dada por ROGERS en 1888, una generalización a dimensión n de MGA_2 . A saber, en las condiciones de (1.1.3) con $I = (0, \infty)$ y $\lambda_k \in (0, 1)$ se tiene

$$\sum \lambda_k x_k \geq \prod x_k^{\lambda_k}$$

dándose la igualdad en el caso en que $x_1 = \dots = x_n$. Para probarlo, ponemos $y_k = \log x_k$ y usamos $\prod (e^{y_k})^{\lambda_k} = e^{\sum \lambda_k y_k} < \sum \lambda_k e^{y_k}$. A partir de ésta se obtienen también la *forma ponderada* de la desigualdad media geométrica-media armónica: en las mismas condiciones

$$\left(\sum_{k=1}^n \frac{\lambda_k}{x_k} \right)^{-1} \leq \prod_{k=1}^n x_k^{\lambda_k}$$

dándose la igualdad en el caso en que $x_1 = \dots = x_n$; y la desigualdad más general siguiente: sean a_{ij} , $\alpha_i > 0$ para $i = 1, \dots, m$ y $j = 1, \dots, n$ con $\alpha_1 + \dots + \alpha_m = 1$. Entonces,

$$\sum_{j=1}^n \left(\prod_{i=1}^m a_{ij}^{\alpha_i} \right) \leq \prod_{i=1}^m \left(\sum_{j=1}^n a_{ij} \right)^{\alpha_i}$$

Presentamos ahora un teorema análogo al (2) pero para ternas de puntos en vez de para parejas.

Teorema 3 (Desigualdad de Popoviciu. POPOVICIU, 1965).

Sea f continua. Entonces f es convexa si y sólo si

$$\begin{aligned} \frac{f(x) + f(y) + f(z)}{3} + f\left(\frac{x+y+z}{3}\right) &\geq \\ &\geq \frac{2}{3} \left[f\left(\frac{x+y}{2}\right) + f\left(\frac{x+z}{2}\right) + f\left(\frac{y+z}{2}\right) \right] \end{aligned}$$

para cualesquiera $x, y, z \in I$. En el caso de funciones estrictamente convexas, la desigualdad superior es estricta en todos los casos, salvo cuando $x = y = z$.

Demostración. $[\Rightarrow]$ Ésta implicación no requiere la hipótesis de que f sea continua. Sin pérdida de generalidad, $x \leq y \leq z$. Llevaremos a cabo varios pasos.

1. Supongamos que $y \leq \frac{x+y+z}{3}$, entonces tenemos que

$$\frac{x+y+z}{3} \leq \frac{x+z}{2} \leq z \quad \text{y} \quad \frac{x+y+z}{3} \leq \frac{y+z}{2} \leq z$$

luego existen $s, t \in [0, 1]$ de manera que podemos escribir los puntos intermedios como combinación lineal convexa de los extremos

$$\frac{x+z}{2} = s \cdot \frac{x+y+z}{3} + (1-s)z \quad (1)$$

$$\frac{y+z}{2} = t \cdot \frac{x+y+z}{3} + (1-t)z \quad (2)$$

Multiplicamos ambas por 3 y sumamos, obteniendo

$$\frac{3}{2}(x+y+2z) = (x+y+z)(s+t) + 3z(2-(s+t)) = (s+t)[(x+y+z)-3z] + 6z$$

$$\frac{3}{2}(x+y-2z) = (s+t)(x+y-2z)$$

entonces

$$(x+y-2z)(s+t-3/2) = 0$$

En el caso en que $(x+y-2z) = 0$, $z = \frac{x+y}{2}$. Pero como $x \leq y \leq z$, entonces necesariamente $x = y = z$. Así, la desigualdad de Popoviciu es inmediata.

Mientras que si $(s+t-3/2) = 0$, entonces $s+t = 3/2$. Por otro lado, utilizando que f es convexa y las igualdades (1) y (2) tenemos que

$$f\left(\frac{x+z}{2}\right) \leq s \cdot f\left(\frac{x+y+z}{3}\right) + (1-s)f(z)$$

$$f\left(\frac{y+z}{2}\right) \leq t \cdot f\left(\frac{x+y+z}{3}\right) + (1-t)f(z)$$

$$f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

Multiplcamos todas las desigualdades en ambos lados por $2/3$ y sumándolas obtenemos la desigualdad buscada

$$\begin{aligned} \frac{2}{3}f\left(\frac{x+y}{2}\right) + \frac{2}{3}f\left(\frac{x+z}{2}\right) + \frac{2}{3}f\left(\frac{y+z}{2}\right) &\leq \frac{1}{3}f(x) + \frac{1}{3}f(y) + \\ &+ f\left(\frac{x+y+z}{3}\right)(s+t)\frac{2}{3} + \frac{2}{3}f(z)(2-(s+t)) \\ &= \frac{f(x) + f(y) + f(z)}{3} + f\left(\frac{x+y+z}{3}\right) \end{aligned}$$

2. Para el caso en el que $y > \frac{x+y+z}{3}$ el camino a seguir es una copia del hecho en el punto anterior, por lo que se hace de manera análoga.

[\Leftarrow] Asumiendo que la desigualdad de Popoviciu es cierta, y poniéndonos en el caso en el que $y = z$, obtenemos la siguiente desigualdad que sustituye a la de la definición de función punto-medio convexa

$$\begin{aligned} \frac{f(x) + 2f(y)}{3} + f\left(\frac{x+2y}{3}\right) &\geq \frac{2}{3} \left[f\left(\frac{x+y}{2}\right) + f(z) + f\left(\frac{x+y}{2}\right) \right] \\ &= \frac{4}{3}f\left(\frac{x+y}{2}\right) + \frac{2}{3}f(y) \end{aligned}$$

multiplicando por $3/4$ tenemos que

$$\begin{aligned} f\left(\frac{x+y}{2}\right) + \frac{1}{2}f(y) &\leq \frac{1}{4}f(x) + \frac{1}{2}f(y) + \frac{3}{4}f\left(\frac{x+2y}{3}\right) \\ f\left(\frac{x+y}{2}\right) &\leq \frac{1}{4}f(x) + \frac{3}{4}f\left(\frac{x+2y}{3}\right) \end{aligned}$$

Ahora bien, utilizando la desigualdad superior y siguiendo la prueba hecha en teorema (2), obtenemos la implicación hacia la izquierda. \square

1.2. Derivabilidad de funciones convexas

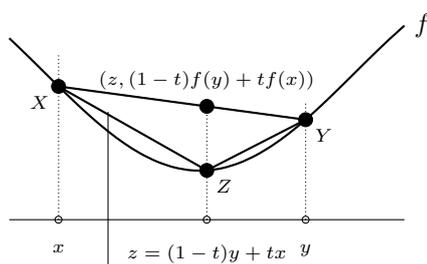
1.2.1. Derivadas laterales y continuidad

Las propiedades de derivabilidad referidas a funciones convexas sobre intervalos se basan principalmente en el estudio de las pendientes de los segmentos inscritos en sus gráficas. El siguiente lema nos ayudará a comparar estas pendientes.

Lema 1.2.1 (Lema de las tres cuerdas. GALVANI, 1916). *Si f es convexa se tiene que*

$$\frac{f(z) - f(x)}{z - x} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(z)}{y - z}$$

para cualesquiera $x < z < y$ de I .



Lema de las tres cuerdas: pendiente $XZ \leq$ pendiente $XY \leq$ pendiente ZY .

Demostración. Sean $x < z < y$ puntos de I , podemos escribir z como combinación lineal convexa de x e y

$$z = \frac{y-z}{y-x}x + \frac{z-x}{y-x}y$$

Como f es convexa, por definición

$$f(z) \leq \frac{y-z}{y-x}f(x) + \frac{z-x}{y-x}f(y)$$

Luego

$$f(z) - f(x) \leq \left(\frac{y-z}{y-x} - 1 \right) f(x) + \frac{z-x}{y-x} f(y) = (f(y) - f(x)) \frac{z-x}{y-x}$$

que es exactamente la desigualdad de la izquierda. La de la derecha se obtiene restando $f(y)$ en vez de $f(x)$. \square

Definición 1.2.1. Dada f convexa y $a \in I$ definimos una nueva función

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

cuyo valor es la pendiente del segmento que une los puntos $(x, f(x))$ y $(a, f(a))$.

Teorema 4.

Sea f convexa. Entonces para cada $a \in I$, s_a es creciente en $I \setminus \{a\}$.

Demostración. Fijado $a \in I$, $x, y \in I \setminus \{a\}$ con $x < y$. Debemos probar que $s_a(x) \leq s_a(y)$. Hay 3 posibilidades para la distribución de x, y, a : o bien $x < y < a$, o bien $x < a < y$, o bien $a < x < y$. Supongamos el primer caso, $x < y < a$. Aplicando el lema de las tres cuerdas con $x = x$, $y = a$ y $z = y$ tenemos que

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

o lo que es lo mismo $s_x(y) \leq s_x(a) \leq s_y(a)$. Esto es, $s_y(x) \leq s_a(x) \leq s_a(y)$. Pues si $\alpha \neq \beta$, $s_\alpha(\beta) = s_\beta(\alpha)$. Los casos restantes se hacen de manera análoga. \square

Vamos ahora con la piedra angular de esta sección, el resultado que recoge lo necesario acerca de la derivabilidad de las funciones convexas.

Teorema 5 (STOLZ, 1893).

Sea f convexa. Entonces, es continua en $\text{int}(I)$. Además, f tiene derivadas finitas por ambos lados en cada punto del $\text{int}(I)$ y se verifica la siguiente cadena de desigualdades para $x < y$ de I

$$f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y)$$

En particular, f'_- y f'_+ son crecientes en $\text{int}(I)$.

Demostración. Sea $a \in \text{int}(I)$, $y \in I$ con $a < y$. Por el teorema anterior, s_a es creciente. Luego $\exists \lim_{x \rightarrow a^-} s_a(x)$ y está acotado superiormente por $s_a(y)$. Esto es que $\exists f'_-(a)$ y satisface la desigualdad $f'_-(a) \leq s_a(y)$. Haciendo $y \rightarrow a^+$ en esta última expresión, vemos por los mismos motivos que antes, que $\exists f'_+(a)$ y verifica $f'_-(a) \leq f'_+(a)$.

Sea ahora $x, y \in \text{int}(I)$. Si $u, v \in \text{int}(I)$ con $x < u \leq v < y$, aplicando el teorema (4) y el lema de Galvani (1.2.1) tenemos que

$$\frac{f(u) - f(x)}{u - x} \leq \frac{f(v) - f(x)}{v - x} \leq \frac{f(v) - f(y)}{v - y}$$

Y además, $s_x(v) = s_v(x) \leq s_v(y) = s_y(v)$. De nuevo, s_x y s_v son crecientes por lo que tienen límite cuando $u \rightarrow x^+$ y $v \rightarrow y^-$ respectivamente. Así, $f'_+(x) \leq f'_-(y)$. Juntándolo todo, si $x < y$ tenemos que $f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y)$.

Por otra parte, como f admite derivadas laterales finitas en cada punto del $\text{int}(I)$, entonces es continua. En efecto, si $a \in \text{int}(I)$

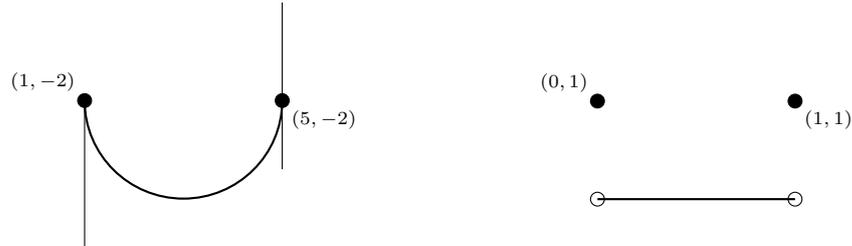
$$\lim_{x \rightarrow a^+} (f(x) - f(a)) = \lim_{x \rightarrow a^+} s_a(x)(x - a) = f'_+(a) \cdot 0 = 0$$

Análogamente, $\exists \lim_{x \rightarrow a^-} (f(x) - f(a)) = 0$. Luego, $\exists \lim_{x \rightarrow a} f(x) = f(a)$, lo que es que f sea continua en a . \square

Observación. *Existencia de derivadas laterales en puntos extremos del intervalo.* Si f es convexa y continua definida en un intervalo cerrado de la forma $[a, b]$, entonces admite derivadas laterales $f'_+(a)$ y $f'_-(b)$ en los puntos extremos del intervalo; aunq ue éstas pueden ser infinitas. Además debe tenerse

$$-\infty \leq f'_+(a) < +\infty \quad \text{y} \quad -\infty < f'_-(b) \leq +\infty$$

Un ejemplo de ésto es la semicircunferencia $y = -(2 + \sqrt{4 - (x - 3)^2})$



Nótese que, tal y como se ilustra en el segundo dibujo de arriba, puede darse el caso de que en los puntos extremos de I la función *salte* hacia arriba (que no hacia abajo, pues contradeciría la convexidad) a pesar de ser convexa. Por ejemplo, la función constante 0 en $(0, 1)$ y 1 en los extremos. Pero ésto puede arreglarse para que la función sea continua sin dejar de ser convexa. La siguiente proposición, que es consecuencia de la que la sigue, nos permite modificar una función convexa para que sea continua en todo su dominio.

Proposición 1.2.1 (Modificación de una función convexa para que sea continua en todo su dominio). Sea $I = [a, b]$, f convexa. Entonces $f(a^+)$ y $f(b^-)$ existen en \mathbb{R} y la función

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in (a, b) \\ f(a^+) & \text{si } x = a \\ f(b^-) & \text{si } x = b \end{cases}$$

es continua y convexa en I .

1.2.2. Extremos de funciones convexas

En este apartado estudiamos la monotonía de una función convexa. El resultado clave se recoge en la siguiente proposición.

Proposición 1.2.2. Sea f convexa, entonces ó bien f es monótona en $\text{int}(I)$, ó bien $\exists \xi \in \text{int}(I)$ tal que f es decreciente en $(-\infty, \xi) \cap I$ y creciente en $[\xi, \infty) \cap I$.

Demostración. Como $\text{int}(I) \subseteq \mathbb{R}$ es un abierto, sin pérdida de generalidad podemos suponer que I también es abierto, $\text{int}(I) = I$. Supongamos que f no es monótona en I , entonces existen $a < b < c$ puntos de I de manera que ocurre alguno de los siguientes casos

1. $f(b) < f(a)$ y $f(b) < f(c)$
2. $f(a) < f(b)$ y $f(c) < f(b)$

pero como $f(x) \leq s_u(v)(x-u) + f(u) \quad \forall [u, v] \subseteq I$ tal y como se probó al comienzo de la sección (1.1), el segundo supuesto queda descartado ya que con la elección $u = a, v = c$ y $x = b$ no se cumple dicha desigualdad. Estamos en el caso primero.

Como f es continua en $[a, c]$ por ser convexa, el mínimo en $[a, c]$ se alcanza. Esto es $\exists \xi \in (a, c)$ tal que $f(\xi) = \inf_{[a, c]} \{f(x)\} = \min_{[a, c]} \{f(x)\}$. Veamos que, de hecho, es absoluto, $f(\xi) = \min_{x \in I} \{f(x)\}$. En efecto, sea $x < a$ por el teorema (4), s_ξ es creciente, $s_\xi(x) \leq s_\xi(a)$. Usando que ξ es mínimo y que $a < \xi$ tenemos por definición que $s_\xi(a) \leq 0$ y, por tanto, $s_\xi(x) \leq 0$. De aquí obtenemos, al ser $x < \xi$, que $f(x) \geq f(\xi)$ para todo $x < a$ y por lo tanto ξ es mínimo en $(-\infty, a] \cap I$. Análogamente, se prueba que ξ es mínimo en $[c, \infty) \cap I$. Y así será un mínimo global.

Consideremos ahora $u < v < \xi$ puntos de I , entonces $s_u(\xi) = s_\xi(u) \leq s_\xi(v) \leq 0$ ya que ξ es mínimo. Luego $s_u(v) \leq s_u(\xi) \leq 0$, pero por como es la definición de $s_u(v)$ y recordando que $v > u, f(v) - f(u) \leq 0$. Así, f es decreciente en $I \cap (-\infty, \xi]$. Similarmente, f es creciente en $I \cap [\xi, \infty)$. \square

Corolario 1.2.3. Toda función convexa f que no es monótona en $\text{int}(I)$ tiene un mínimo global en $\text{int}(I)$.

Los últimos resultados ya nos dan una idea de como se comportan las funciones convexas. En la línea del corolario anterior damos a conocer algunas características de las mismas sin utilizar más que conceptos básicos.

Proposición 1.2.4. *Sea f convexa, entonces ocurre que*

1. *Si tiene un máximo global en $\text{int}(I)$, entonces es constante.*
2. *Si tiene un mínimo local, éste es global.*

Demostración. 1. f tiene un máximo global en $c \in \text{int}(I)$. Sean $x < c < y$, entonces $f(x), f(y) \leq f(c)$. Como f es convexa, por el teorema (5), $f'_-(c) \leq f'_+(c)$. Pensemos ahora en como son estas derivadas laterales, por definición de derivada $f'_\epsilon(c) = \lim_{x \rightarrow c^\epsilon} s_c(x)$ y teniendo en cuenta la posición que ocupa x respecto a c obtenemos que $0 \leq f'_-(c) \leq f'_+(c) \leq 0$. Luego f es derivable en c con derivada nula. Tenemos así la siguiente relación entre las pendientes de los segmentos inscritos en la gráfica de f correspondientes a los puntos x, c e y , $s_c(x) \leq 0 \leq s_c(y)$. Pero como $x < c$, entonces $f(x) - f(c) \geq 0$. Que con la desigualdad que ya teníamos $f(x) = f(c) \quad \forall x < c$. Similarmente, $f(y) = f(c) \quad \forall y > c$. Así, f es constante.

2. f tiene un mínimo local en $c \in I$. Sea $x > c$, entonces $\exists y$ con $c < y < x$ y $f(y) \geq f(c)$. Por el lema de las tres cuerdas, $0 \leq s_c(y) \leq s_c(x)$. Luego $f(x) \geq f(c) \quad \forall x > c$. Análogamente si tomamos el punto $x < c$. Así, el mínimo es global. \square

Proposición 1.2.5. *Si $I = \mathbb{R}$ y f es convexa y acotada superiormente, entonces es constante.*

Demostración. Sea M la cota, $x < y$ puntos de I y $\theta > 0$, gracias al lema de las tres cuerdas y a que M es cota superior, sabemos que

$$\frac{f(x) - M}{\theta} \leq \frac{f(x) - f(x - \theta)}{\theta} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(y + \theta) - f(y)}{\theta} \leq \frac{M - f(y)}{\theta}$$

y tomando límite cuando $\theta \rightarrow \infty$

$$0 = \lim_{\theta \rightarrow \infty} \frac{f(x) - M}{\theta} \leq \frac{f(y) - f(x)}{y - x} \leq \lim_{\theta \rightarrow \infty} \frac{M - f(y)}{\theta} = 0$$

Luego, $s_x(y) = 0$. Así, $f(y) = f(x) \quad \forall x, y \in \mathbb{R}$ con $x < y$, esto es que es constante en todo \mathbb{R} . \square

Observación. *Puntos de no derivabilidad de una función convexa.* Ahora que gracias al teorema (5) sabemos que para una función convexa las derivadas laterales existen en los puntos de $\text{int}(I)$, nos preguntamos: ¿Cuánto no derivable puede ser una función convexa? La respuesta es que **a lo más no es derivable sobre una cantidad numerable** de puntos de I , tal y como muestra la siguiente proposición.

Proposición 1.2.6. *Sea f convexa, entonces es derivable en todo I salvo a lo más en un conjunto numerable.*

Demostración. Sea $I_{nd} := \{x \in \text{int}(I) \mid f'_-(x) < f'_+(x)\}$ y a cada $x \in I_{nd}$ le asociamos $r_x \in \mathbb{Q} \cap (f'_-(x), f'_+(x))$. Tenemos así la función $\varphi : I_{nd} \rightarrow \mathbb{Q} \mid x \mapsto r_x$. Se sigue del teorema (5), que si $x, y \in I_{nd}$ con $x < y$ entonces

$$f'_-(x) < r_x < f'_+(x) \leq f'_-(y) < r_y < f'_+(y)$$

Así, $r_x < r_y$. Lo que muestra que el conjunto de puntos de $\text{int}(I)$ (y también de I) para los cuales f no es derivable, es numerable. Nótese que este argumento depende del *axioma de elección*. \square

Proposición 1.2.7. *Sea f convexa, entonces es Lipschitz en cualquier subintervalo compacto contenido en el interior de I .*

Demostración. Sea $[a, b] \subseteq \text{int}(I)$. Por el teorema (5) sabemos que se cumple que

$$f'_+(a) \leq f'_+(x) \leq \frac{f(y) - f(x)}{y - x} \leq f'_-(y) \leq f'_-(b) \quad \forall x, y \in [a, b] \text{ con } x < y$$

Basta con tomar la constante $L := \max\{|f'_+(a)|, |f'_-(b)|\}$, y darse cuenta de que $f|_{[a, b]}$ es Lipschitz con ésta constante. \square

Proposición 1.2.8. *Para cada $n \in \mathbb{N}$, sea f_n convexa en I . Supongamos que la sucesión $(f_n)_{n=1}^\infty$ converge puntualmente a una función f en I . Entonces, f es convexa. Además, la convergencia es uniforme sobre cualquier subintervalo compacto contenido en $\text{int}(I)$, y la sucesión de derivadas (f'_n) converge a f' en todo I salvo quizás en una cantidad numerable de puntos.*

Demostración. Sean $x, y \in I$ y $\lambda, \mu \in [0, 1]$ con $\lambda + \mu = 1$, para la primera parte del corolario simplemente

$$\begin{aligned} f(\lambda x + \mu y) &= \lim_{n \rightarrow \infty} f_n(\lambda x + \mu y) \leq \lim_{n \rightarrow \infty} (\lambda f_n(x) + \mu f_n(y)) \\ &= \lambda \lim_{n \rightarrow \infty} f_n(x) + \mu \lim_{n \rightarrow \infty} f_n(y) = \lambda f(x) + \mu f(y) \end{aligned}$$

luego f es convexa. Definamos para cada natural n , el conjunto A_n de puntos de I donde f_n no es derivable, y $A := \{x \in I : f \text{ no es derivable en } x\}$, ambos son numerables por (1.2.6). Llamamos $B := (\bigcup_{n \in \mathbb{N}} A_n) \cup A$, numerable, y por lo tanto $I - B \neq \emptyset$ es infinito no numerable. Sea ahora $a \in \text{int}(I) - B$ un punto fijo, y tomemos otros dos puntos $x, y \in I$ tales que $x < a < y$. Por las desigualdades que aparecen en la demostración de la proposición anterior, para un natural cualquiera n se tiene que como f_n es convexa y derivable en a

$$\frac{f_n(x) - f_n(a)}{x - a} \leq f'_n(a) \leq \frac{f_n(y) - f_n(a)}{y - a}$$

como hay convergencia puntual de la sucesión (f_n) a f , tomando límite cuando $n \rightarrow \infty$ tenemos

$$\frac{f(x) - f(a)}{x - a} \leq \liminf_n f'_n(a) \leq \limsup_n f'_n(a) \leq \frac{f(y) - f(a)}{y - a}$$

Para abreviar, llamo $\alpha = \liminf_n f'_n(a)$ y $\beta = \limsup_n f'_n(a)$. Tenemos lo siguiente

$$\frac{f(x) - f(a)}{x - a} \leq \alpha \quad \forall x < a$$

luego, tomando límite cuando $x \rightarrow a^-$ y teniendo en cuenta que f es derivable en a

$$\lim_{x \rightarrow a^-} \frac{f(x) - f(a)}{x - a} = f'_-(a) = f'(a) \leq \alpha$$

Análogamente,

$$\lim_{y \rightarrow a^+} \frac{f(y) - f(a)}{y - a} = f'_+(a) = f'(a) \geq \beta$$

Nos queda que

$$f'(a) \leq \alpha \leq \beta \leq f'(a)$$

y entonces $\alpha = \beta$. Por lo tanto, $\exists \lim_{n \rightarrow \infty} f'_n(a) = f'(a)$. Luego $f'_n \rightarrow f'$ en $I \setminus B$. Supongamos ahora $[a, b] \subseteq \text{int}(I)$ con $a, b \in I \setminus B$ y veamos que f_n converge a f uniformemente en dicho intervalo (para la convergencia uniforme en cualquier subintervalo compacto $[r, s] \subseteq \text{int}(I)$, considerar un $[\tilde{r}, \tilde{s}] \subseteq \text{int}(I)$ con los extremos pertenecientes a $I \setminus B$ tal que contenga al intervalo $[r, s]$). Como f_n es convexa, entonces $|f_n(x) - f_n(y)| \leq k_n|x - y| \quad \forall n \quad \forall x, y \in [a, b]$ donde $k_n = \max\{|f'_n(a)|, |f'_n(b)|\}$ por la demostración de la proposición (1.2.7). Luego, como $k_n \rightarrow k = \max\{|f'(a)|, |f'(b)|\}$, $\exists C > 0$ tal que

$$|f(x) - f(y)| \leq C|x - y|, \text{ y } |f_n(x) - f_n(y)| \leq C|x - y| \quad \forall n \quad \forall x, y \in [a, b]$$

(nos basta con tomar como C una cota de la sucesión $(k_n)_1^\infty$). Sea $\epsilon > 0$. Tomemos $E \subseteq [a, b]$ un conjunto finito tal que para todo $x \in [a, b]$ exista $z \in E$ que verifica $|x - z| \leq \epsilon/(3C)$. Entonces, existe $M \in \mathbb{N}$ tal que

$$|f_n(z) - f(z)| < \epsilon/3 \quad \forall n \geq M \quad \forall z \in E$$

Sea $n \geq M$. Si $x \in [a, b]$, sea $z \in E$ tal que $|z - x| < \epsilon/(3C)$, entonces

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(x) - f_n(z)| + |f_n(z) - f(z)| + |f(z) - f(x)| \\ &\leq C|x - z| + \epsilon/3 + C|x - z| \leq \epsilon \end{aligned}$$

Por tanto, $\sup_{x \in [a, b]} \{|f_n(x) - f(x)|\} \leq \epsilon$ si $n \geq M$. Luego (f_n) converge uniformemente a f en $[a, b]$. \square

La derivada simétrica de orden dos

Es bien conocido el criterio de la positividad de la segunda derivada para la convexidad de una función que es al menos dos veces diferenciable. Sin embargo, puesto que la derivada primera de una función convexa puede no existir en un subconjunto denso de puntos, debemos ser más laxos en el concepto de derivada segunda. Por este motivo damos la siguiente definición.

Definición 1.2.2 (Derivada simétrica (superior e inferior) de orden dos). Si f es una función, definimos la *derivada simétrica (superior e inferior, respectivamente) de orden dos* de f en un punto $x \in I$ como

$$\overline{D}^2 f(x) := \limsup_{h \rightarrow 0^+} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}$$

$$\underline{D}^2 f(x) := \liminf_{h \rightarrow 0^+} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}$$

No es difícil ver que si f es dos veces derivable, $f''(x) = \overline{D}^2 f(x) = \underline{D}^2 f(x)$. Sin embargo, tal y como están definidas las derivadas simétricas, éstas pueden existir incluso en los puntos de discontinuidad de f : basta considerar la función signo que no es continua en el cero y por lo tanto no es derivable en dicho punto, pero sí existen las derivadas simétricas de orden dos en el origen y valen cero.

Vamos con un resultado concerniente a estos últimos conceptos y que nos permitirá establecer una importante y sencilla caracterización de las funciones convexas, el **test de la segunda derivada**.

Teorema 6.

Sea I un intervalo abierto y f una función definida sobre I . Entonces, f es convexa sii es continua y $\overline{D}^2 f(x) \geq 0 \quad \forall x \in I$.

En particular, si una función f es convexa en un entorno de cada punto de I , entonces es convexa en todo I .

Demostración. $[\Rightarrow]$ Como $I = \text{int}(I)$ y f es convexa, por (5) es continua en I . Además, gracias a (1.1.4), se verifica

$$f(x+h) - f(x-h) - 2f(x) \geq 0 \quad \forall x \in I \quad \forall h > 0 \quad \text{con } x \pm h \in I$$

Ahora bien, debido a la definición de $\overline{D}^2 f$ y $\underline{D}^2 f$, tenemos que

$$\overline{D}^2 f(x) \geq \underline{D}^2 f(x) \geq 0 \quad \forall x \in I$$

donde la última desigualdad se debe a que, por ser f convexa, los cocientes que intervienen en la definición de $\underline{D}^2 f(x)$ son positivos.

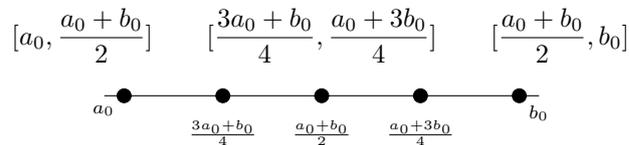
$[\Leftarrow]$ Para esta implicación consideramos dos casos. En el primero supongamos que $\overline{D}^2 f(x) > 0 \quad \forall x \in I$ y veamos que f es convexa. Si no es así, entonces existe un subintervalo de I , $I_0 = [a_0, b_0]$, tal que (aquí se usa el teorema (2))

$$f\left(\frac{a_0 + b_0}{2}\right) > \frac{1}{2}(f(a_0) + f(b_0))$$

Ahora, podemos elegir un nuevo intervalo I_1 más pequeño en longitud que I_0 y contenido en éste de manera que si denominamos $I_1 = [a_1, b_1]$ se siga verificando que

$$f\left(\frac{a_1 + b_1}{2}\right) > \frac{1}{2}(f(a_1) + f(b_1))$$

y sea de longitud la mitad de la de I_0 , $b_1 - a_1 = \frac{b_0 - a_0}{2}$. Para dicho nuevo intervalo nos valdría uno de los tres siguientes



En efecto, si suponemos que ninguno de los anteriores nos valiera, como todos tienen longitud la mitad que la del de partida, sería porque ocurre lo siguiente

$$f\left(\frac{a_0 + b_0}{2}\right) \leq \frac{1}{2}\left(f\left(\frac{3a_0 + b_0}{4}\right) + f\left(\frac{a_0 + 3b_0}{4}\right)\right) = (\diamond)$$

y

$$f\left(\frac{a_0 + 3b_0}{4}\right) \leq \frac{1}{2}\left(f\left(\frac{a_0 + b_0}{2}\right) + f(b_0)\right)$$

y

$$f\left(\frac{3a_0 + b_0}{4}\right) \leq \frac{1}{2}\left(f\left(\frac{a_0 + b_0}{2}\right) + f(a_0)\right)$$

entonces

$$\begin{aligned} (\diamond) &\leq \frac{1}{4}\left[f\left(\frac{a_0 + b_0}{2}\right) + f(b_0) + f\left(\frac{a_0 + b_0}{2}\right) + f(a_0)\right] \\ &= \frac{1}{4}\left[2f\left(\frac{a_0 + b_0}{2}\right) + f(a_0) + f(b_0)\right] \end{aligned}$$

luego

$$f\left(\frac{a_0 + b_0}{2}\right) \leq \frac{1}{2}(f(a_0) + f(b_0))$$

lo cual es una contradicción, pues entra en conflicto con la primera desigualdad estricta obtenida en esta implicación de la demostración.

Si ahora repetimos el proceso por inducción, obtenemos una sucesión de intervalos encajados y no vacíos cuya longitud tiende a cero. Luego, la intersección de todos ellos es un punto $x_0 \in \bigcap_{n=0}^{\infty} I_n$. En dicho punto debemos tener $\overline{D}^2 f(x_0) \leq 0$ por definición. Esto contradice la hipótesis de partida.

En el segundo caso, supongamos que $\overline{D}^2 f(x) \geq 0 \quad \forall x \in I$. Consideremos la sucesión de funciones $f_n(x) := f(x) + \frac{1}{n}x^2$. Veamos que $\overline{D}^2 f_n(x) > 0 \quad \forall x \in I$ y aplicando el caso primero podemos concluir que f_n es convexa para cualquier $n \in \mathbb{N}$. En efecto, si $x \in I$, $h > 0$ es tal que $x \pm h \in I$ y $n \in \mathbb{N}$, entonces

$$f_n(x \pm h) = f(x \pm h) + \frac{1}{n}(x \pm h)^2$$

Por tanto

$$\frac{f_n(x+h) + f_n(x-h) - 2f_n(x)}{h^2} = \frac{2}{n} + \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}$$

luego

$$\overline{D}^2 f_n(x) = \frac{2}{n} + \overline{D}^2 f(x) > 0 \quad \forall x \in I \quad \forall n \in \mathbb{N}$$

Ahora, puesto que $f_n(x) \rightarrow f(x) \quad \forall x \in I$, la convexidad de f la obtenemos de manera inmediata gracias a la proposición (1.2.8). \square

Corolario 1.2.9 (El test de la segunda derivada). *Sea f una función dos veces derivable definida sobre un intervalo abierto. Entonces,*

1. f es convexa $\Leftrightarrow f'' \geq 0$
2. f es estrictamente convexa $\Leftrightarrow f'' \geq 0$ y el conjunto de puntos donde f'' se anula no contiene intervalos de longitud positiva. Es decir, si $N := \{x \in I : f''(x) = 0\}$ entonces $[r, s] \not\subseteq N \quad \forall r, s \in N$.

Demostración. 1. f es dos veces derivable y convexa \Leftrightarrow continua y además ocurre que $f''(x) = \overline{D}^2 f(x) = \underline{D}^2 f(x) \geq 0 \quad \forall x \in I$.

2. En cuanto a la convexidad estricta se sigue de la siguiente manera: por reducción al absurdo, supongamos que $f''(x) = 0$ sobre $(a, b) \subseteq N \subseteq I$. Entonces, sobre (a, b) se tiene que $f'(x) = k \in \mathbb{R}$ y por lo tanto que $f(x) = kx + k'$ con $k' \in \mathbb{R}$ sobre (a, b) ; es decir, f es una función afín sobre el intervalo (a, b) , y evidentemente una función afín no es una función estrictamente convexa (es convexa y cóncava a la vez) tal y como vimos en la primera sección. Recíprocamente, busquemos también una contradicción. Supongamos que f no es estrictamente convexa, entonces $\exists x, y \in I$ tales que

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \quad \forall \lambda, \mu \in [0, 1] : \lambda + \mu = 1$$

entonces derivando dos veces

$$f''(\lambda x + \mu y) = \lambda f''(x) + \mu f''(y) = 0 \quad \forall \lambda, \mu \in [0, 1] : \lambda + \mu = 1$$

luego $f''(z) = 0 \quad \forall z \in [x, y] \Rightarrow [x, y] \subseteq N$. Lo cual es una contradicción, pues N no contenía intervalos. \square

1.2.3. Un par de ejemplos: construcción de convexas

Para acabar esta sección daremos dos ejemplos de funciones convexas no derivables en un subconjunto numerable; en el segundo ejemplo el subconjunto será además denso, y nos proporcionará un método para construir de manera general funciones convexas a partir de crecientes, además no necesariamente derivables.

EJEMPLO 1.2.1 (Función convexa no derivable en un subconjunto numerable).
Sea la función

$$f(x) := \sum_{n=0}^{\infty} \frac{|x - n|}{2^n} \quad \forall x \in \mathbb{R}$$

entonces, f es convexa en \mathbb{R} y es derivable en $\mathbb{R} \setminus \mathbb{N}$ (no siéndolo en los naturales).

Demostración. La convexidad es inmediata sin más que dejarse llevar partiendo de la definición. Luego es continua. Veamos que es derivable en $\mathbb{R} \setminus \mathbb{N}$. Sea $x \in (-\infty, 0)$ la expresión de f se reduce a

$$f(x) = \sum_{n=0}^{\infty} \frac{n - x}{2^n}$$

y no habría problemas de derivabilidad. En efecto, de manera sencilla, en este caso

$$f(x) = \sum_{n=0}^{\infty} \frac{n - x}{2^n} = \sum_{n=0}^{\infty} \frac{n}{2^n} - x \sum_{n=0}^{\infty} \frac{1}{2^n} = A(1/2) - 2x$$

donde $A(1/2) := \sum_{n=0}^{\infty} \frac{n}{2^n} = 2$. Así, f es una función afín y por lo tanto derivable, con derivada -2 para todo $x \in (-\infty, 0)$. Veamos ahora que es derivable en los reales positivos que no son un natural. Sea $a \in \mathbb{R}^+ \setminus \mathbb{N}$, entonces $a \in ([a], [a] + 1)$ donde el

corchete representa la función parte entera. En dicho intervalo podemos escribir

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{|x-n|}{2^n} = \sum_{n=0}^{[a]} \frac{x-n}{2^n} + \sum_{n=[a]+1}^{\infty} \frac{n-x}{2^n} \\ &= x \sum_{n=0}^{[a]} \frac{1}{2^n} - \sum_{n=0}^{[a]} \frac{n}{2^n} + \sum_{n=[a]+1}^{\infty} \frac{n}{2^n} - x \sum_{n=[a]+1}^{\infty} \frac{1}{2^n} \\ &= x \left(\sum_{n=0}^{[a]} \frac{1}{2^n} - \sum_{n=[a]+1}^{\infty} \frac{1}{2^n} \right) + \left(\sum_{n=[a]+1}^{\infty} \frac{n}{2^n} - \sum_{n=0}^{[a]} \frac{n}{2^n} \right) \end{aligned}$$

donde los términos entre paréntesis son escalares que no dependen de la variable x . Luego la función f es, de nuevo, una función afín sobre el intervalo $([a], [a] + 1)$ y por lo tanto derivable, con derivada constante, en dicho intervalo. Explícitamente,

$$f'(x) = \sum_{n=0}^{[a]} \frac{1}{2^n} - \sum_{n=[a]+1}^{\infty} \frac{1}{2^n} = \frac{1 - (1/2)^{[a]+1}}{1 - (1/2)} - \frac{(1/2)^{[a]+1}}{1 - (1/2)} = 2 - \frac{1}{2^{[a]-1}}$$

para cualquier $x \in ([a], [a] + 1)$. Ya sólo resta probar que f no es derivable sobre los naturales. Sean entonces $m \in \mathbb{N}$, y $x, y \in \mathbb{R} \setminus \mathbb{N}$ no fijos tales que $m - 1 < y < m < x < m + 1$. Tal y como hemos visto antes, tratando primero con la x , tenemos

$$f(x) = x \left(\sum_{n=0}^m \frac{1}{2^n} - \sum_{n=m+1}^{\infty} \frac{1}{2^n} \right) + \left(\sum_{n=m+1}^{\infty} \frac{n}{2^n} - \sum_{n=0}^m \frac{n}{2^n} \right) \quad \forall x \in (m, m+1)$$

y para la y

$$f(y) = y \left(\sum_{n=0}^{m-1} \frac{1}{2^n} - \sum_{n=m}^{\infty} \frac{1}{2^n} \right) + \left(\sum_{n=m}^{\infty} \frac{n}{2^n} - \sum_{n=0}^{m-1} \frac{n}{2^n} \right) \quad \forall y \in (m-1, m)$$

Derivando en estos dos casos, nos queda

$$f'(x) = 2 - \frac{1}{2^{m-1}} \quad \forall x \in (m, m+1) \quad f'(y) = 2 - \frac{1}{2^{m-2}} \quad \forall y \in (m-1, m)$$

luego, como f es continua en m

$$f'_-(m) = \lim_{y \rightarrow m^-} f'(y) = 2 - \frac{1}{2^{m-2}} \neq 2 - \frac{1}{2^{m-1}} = \lim_{x \rightarrow m^+} f'(x) = f'_+(m)$$

ya así f no es derivable en m . \square

Antes de mostrar el segundo ejemplo que anunciábamos, vamos a dar unos retozos generales sobre funciones construidas a partir de la integración adecuada de otras, pues es así como estará definida la función de dicho ejemplo. Los resumimos en la siguiente proposición.

Proposición 1.2.10. Sean I un intervalo abierto y una función $\varphi : I \rightarrow \mathbb{R}$ creciente. Entonces, para todo $c \in I$, la primitiva $f : I \rightarrow \mathbb{R}$ definida por

$$f(x) := \int_c^x \varphi(t) dt \quad \forall x \in I$$

es continua, convexa y derivable en cada punto x de I en el que φ es continua, cumpliéndose en tales puntos que $f'(x) = \varphi(x)$.

Demostración. Como φ es creciente y acotada sobre intervalos cerrados y acotados, obtenemos fácilmente que f es localmente Lipschitziana y, por lo tanto, continua. Para ver que es convexa, puesto que ya sabemos que es continua, por el teorema (2) basta ver que f es punto-medio convexa. En efecto, sean $x \leq y$ puntos de I , entonces

$$\begin{aligned} \frac{f(x) + f(y)}{2} - f\left(\frac{x+y}{2}\right) &= \frac{1}{2} \int_c^x \varphi + \frac{1}{2} \int_c^y \varphi - \frac{1}{2} \int_c^{(x+y)/2} \varphi \\ &= \frac{1}{2} \int_c^x \varphi + \frac{1}{2} \left(\int_c^x \varphi + \int_x^{(x+y)/2} \varphi + \int_{(x+y)/2}^y \varphi \right) - \left(\int_c^x \varphi + \int_x^{(x+y)/2} \varphi \right) \\ &= \frac{1}{2} \left(\int_{(x+y)/2}^y \varphi - \int_x^{(x+y)/2} \varphi \right) \geq 0 \end{aligned}$$

donde la positividad se obtiene porque φ es función creciente. La tercera propiedad es consecuencia de la definición de f y del teorema fundamental del cálculo. \square

De hecho, vamos a necesitar un poquito más que la última propiedad enunciada en la proposición anterior relativa a la derivabilidad de la primitiva de una función: la derivabilidad lateral.

Proposición 1.2.11. *Sea ϕ una función integrable Riemann en $[0, 1]$. Sea g la primitiva definida por $g(x) := \int_0^x \phi(t) dt$. Sea $a \in [0, 1]$ y supongamos que existe el $\lim_{x \rightarrow a^+} \phi(x) = R$. Entonces la función g tiene derivada por la derecha en a y $g'_+(a) = R$. Análogo resultado se tiene para las derivadas laterales por la izquierda.*

Demostración. Dado $\epsilon > 0$, sea $\delta > 0$ tal que si $0 < h < \delta$ y $a + h \in [0, 1]$ entonces $|\phi(a + h) - R| < \epsilon$. Observemos lo siguiente

$$\left| \frac{g(a+h) - g(a)}{h} - R \right| = \left| \frac{1}{h} \int_a^{a+h} (\phi - R) \right| = (\spadesuit)$$

Luego

$$(\spadesuit) \leq \frac{1}{h} \int_a^{a+h} |\phi(t) - R| dt < \frac{1}{h} \cdot \epsilon \cdot h = \epsilon$$

luego $\exists \lim_{h \rightarrow 0^+} \frac{g(a+h) - g(a)}{h} = R$, entonces $g'_+(a) = R$. \square

EJEMPLO 1.2.2 (Función convexa no derivable sobre un subconjunto numerable denso). *Sea $(q_n)_{n \geq 1}$ una enumeración de los racionales de $[0, 1]$. Definimos la función $\varphi : [0, 1] \rightarrow \mathbb{R}^+$ como*

$$\varphi(t) := \sum_{\{k: q_k \leq t\}} \frac{1}{2^k}$$

y a partir de ella tomamos la primitiva $f : [0, 1] \rightarrow \mathbb{R}^+$ definida de la siguiente manera

$$f(x) := \int_0^x \varphi(t) dt \quad \forall x \in [0, 1]$$

Entonces, f es convexa en $[0, 1]$ y no derivable en los racionales de $[0, 1]$.

Prueba del ejemplo. Se trata de probar que f no es derivable sobre los elementos de la sucesión (q_n) y sí lo es en los demás puntos de $[0, 1]$. Utilizamos las propiedades que acabamos de ver. Estamos en las condiciones adecuadas, pues la función φ es creciente. En efecto, antes de nada, para abreviar en notación, si $t \in [0, 1]$, al conjunto $\{k : q_k \leq t\}$ lo llamaré A_t . Sean entonces $r, s \in [0, 1]$ con $r < s$, luego

$$\varphi(r) = \sum_{k \in A_r} 1/2^k$$

y

$$\varphi(s) = \sum_{k \in A_r} 1/2^k + \sum_{\{k: r < q_k \leq s\}} 1/2^k$$

y como los racionales son densos en $[0, 1]$, en particular, habrá racionales en $(r, s]$, por lo que el sumatorio último de la segunda expresión no será nulo. Así, $\varphi(r) < \varphi(s)$, y entonces φ es función estrictamente creciente.

Por lo tanto la función f es convexa y continua en $[0, 1]$. Además f es derivable en los puntos de continuidad de φ . Esos puntos de continuidad de φ son exactamente los de $[0, 1] \setminus \{q_n\}_{n \geq 1}$ (probamos más abajo una igualdad cuya demostración es análoga a la que se precisa para la prueba de esta afirmación). Sin embargo, veamos que para cada $n \in \mathbb{N}$, φ tiene una discontinuidad de salto (finito) en q_n . Fijemos un $n \in \mathbb{N}$ y llamemos $q = q_n$. Tenemos

$$\varphi(q^-) := \lim_{t \rightarrow q^-} \varphi(t) = \lim_{t \rightarrow q^-} \sum_{k \in A_t} 1/2^k = \sum_{\{k: q_k < q\}} 1/2^k$$

y

$$\varphi(q^+) := \lim_{t \rightarrow q^+} \varphi(t) = \lim_{t \rightarrow q^+} \sum_{k \in A_t} 1/2^k = \sum_{\{k: q_k \leq q\}} 1/2^k$$

entonces, $\varphi(q^+) = \varphi(q^-) + \frac{1}{2^n}$, y por lo tanto φ tiene una discontinuidad de salto finito en q .

Antes de continuar con la demostración, probemos las igualdades (en realidad una, la otra se hace de manera análoga) no obvias de arriba. Por ejemplo, hagamos la primera. Se trata de ver que

$$\lim_{t \rightarrow q_n^-} \varphi(t) = \sum_{\{k: q_k < q_n\}} 1/2^k$$

equivalentemente

$$\left| \sum_{\{k: q_k < q_n\}} 1/2^k - \sum_{\{k: q_k \leq t\}} 1/2^k \right| \rightarrow 0 \quad \Leftrightarrow \quad \left| \sum_{\{k: t < q_k < q_n\}} 1/2^k \right| \rightarrow 0$$

cuando $t \rightarrow q_n^-$. Para ello utilicemos que la serie $\sum_{m=1}^{\infty} 1/2^m$ es absolutamente convergente a 1. Así, dado $\epsilon > 0 \exists N \in \mathbb{N}$ tal que para todo $n \geq m > N$ se verifica que

$$\sum_{k=m}^n 1/2^k < \epsilon \quad (\star)$$

Se tiene así que

$$\sum_{k \in F} 1/2^k \quad \forall F \subseteq \mathbb{N} \text{ tal que } F \cap \{1, 2, \dots, N\} = \emptyset \quad (\Delta)$$

En efecto, supongamos que $F \cap \{1, 2, \dots, N\} = \emptyset$. Si F es un subconjunto finito de los naturales, entonces es claro, gracias a (\star) , que

$$\sum_{k \in F} 1/2^k \leq \sum_{k=\text{mín } F}^{\text{máx } F} 1/2^k < \epsilon$$

Si $F \subseteq \mathbb{N}$ es infinito llevamos a cabo el siguiente razonamiento. Denotamos $F := \{\eta_1, \eta_2, \dots\} \subseteq \mathbb{N}$, con $\eta_1 < \eta_2 < \dots$. Si llamamos $F_m := \{\eta_1, \dots, \eta_m\}$ y $s_m := \sum_{k \in F_m} 1/2^k$ para cada $m \in \mathbb{N}$, entonces por (\star)

$$s_m \leq \sum_{k=\text{mín } F_m}^{\text{máx } F_m} 1/2^k < \epsilon \quad \forall m \in \mathbb{N}$$

luego, tomando límite cuando $m \rightarrow \infty$ tenemos que

$$s := \sum_{k \in F} 1/2^k = \lim_{m \rightarrow \infty} \sum_{k \in F_m} 1/2^k < \epsilon$$

Para concluir la demostración, definamos

$$\delta := \text{mín}\{|q_n - q_i| : i \in \{1, \dots, N\} \setminus \{n\}\} > 0$$

y sea $t \in (q_n - \delta, q_n)$. Por como hemos definido δ , sabemos que

$$\{k : t < q_k < q_n\} \cap \{1, \dots, N\} = \emptyset$$

Luego, dado $\epsilon > 0$ hemos encontrado $\delta > 0$ tal que si $t \in (q_n - \delta, q_n)$ entonces

$$\sum_{\{k: t < q_k < q_n\}} 1/2^k < \epsilon$$

al verificarse la condición de la expresión (Δ) . Por lo tanto,

$$\lim_{t \rightarrow q_n^-} \varphi(t) = \sum_{k: q_k < q_n} 1/2^k$$

lo cual es lo que buscábamos.

Como ya dije antes, la segunda igualdad (cuando $t \rightarrow q_n^+$) se lleva a cabo de manera análoga; es más, no es necesario hacer toda la demostración de nuevo, pues gran parte de la misma es reutilizar la de arriba. Incluso para probar que φ es continua en $[0, 1] \setminus \mathbb{Q}$ se reutiliza la de arriba parcialmente.

Se trata ahora de aplicar el resultado (1.2.11). Desde luego φ es integrable Riemann en $[0, 1]$ por ser monótona. Sea $n \in \mathbb{N}$, como hemos visto que φ tiene límites laterales en el punto q_n y son finitos, la función f tendrá derivadas laterales en el punto q_n y valdrán lo que los límites laterales de φ antes mencionados. Es decir,

$$\exists f'_-(q_n) = \varphi(q_n^-) \quad \exists f'_+(q_n) = \varphi(q_n^+)$$

pero habíamos visto que φ tiene discontinuidades de salto (finito) en q_n . Luego $f'_-(q_n) \neq f'_+(q_n)$ y, por tanto, f no es derivable en q_n . Consecuentemente, f no es derivable en $\{q_n : n \in \mathbb{N}\}$.

1.3. La subdiferencial en una dimensión

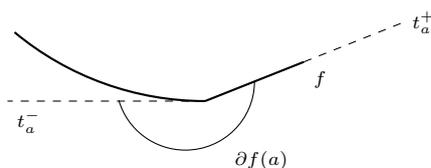
En esta sección trabajaremos con funciones convexas no necesariamente diferenciables. Traduciendo literalmente del inglés, hablamos de funciones *no suaves*; coloquialmente, aquellas cuyas gráficas tienen picos. Puesto que en los puntos donde la función no es derivable no existe recta tangente, necesitamos manejar conceptos más laxos para el estudio de la convexidad. Introducimos los conceptos de **recta soporte** y **subdiferencial**.

Definición 1.3.1. Sea f una función, $x \in I$. Se dice que f admite una recta soporte en x si $\exists \lambda \in \mathbb{R}$ tal que

$$f(y) \geq f(x) + \lambda(y - x) \quad \forall y \in I$$

Definición 1.3.2. Se llama *subdiferencial de f en x* al conjunto de todos los $\lambda \in \mathbb{R}$ que cumplen la condición de la definición (1.3.1). Se denota por $\partial f(x)$.

Pensemos un momento en lo que significa geoméricamente, $\partial f(x)$ está formada por las pendientes de las rectas soporte de f en x . El siguiente dibujo lo aclara



La subdiferencial. Las rectas t_a^+ , t_a^- son las tangentes a la gráfica de la función por la derecha y la izquierda resp. El arco representa el ángulo en el que nos podemos mover para que las rectas que pasan por $(a, f(a))$ sean rectas soporte.

En el mismo las rectas t_a^+ y t_a^- son *rectas límite* para las rectas soporte de f en a . Es evidente que si f es derivable en a , $t_a^- = t_a^+$ es la tangente a f en a y, por lo tanto, $\partial f(a) = \{f'(a)\}$. En general, la subdiferencial es siempre un conjunto convexo, aunque puede ser vacío. Y si $x \in \text{int}(I)$, $\partial f(x) = [f'_-(x), f'_+(x)]$.

Proposición 1.3.1. Siendo $f : I \rightarrow \mathbb{R}$, f es convexa si y sólo si $\partial f(x) \neq \emptyset \quad \forall x \in \text{int}(I)$. Además, cada función $\varphi : I \rightarrow \mathbb{R}$ que cumple que $\varphi(x) \in \partial f(x) \quad \forall x \in \text{int}(I)$ es creciente en el interior de I y verifica que

$$f'_-(x) \leq \varphi(x) \leq f'_+(x)$$

Demostración. Sean $u, v \in I$, $u \neq v$ y sea $t \in (0, 1)$, de modo que $(1 - t)u + tv \in \text{int}(I)$. Podemos considerar entonces $\lambda \in \partial f((1 - t)u + tv)$ y tendremos, por la definición (1.3.2), que

$$\begin{aligned} f(u) &\geq f((1 - t)u + tv) + \lambda[u - (1 - t)u - tv] \\ &= f((1 - t)u + tv) + \lambda(u - v)t \end{aligned}$$

y que

$$f(v) \geq f((1 - t)u + tv) - (1 - t)(u - v)\lambda$$

Ahora bien, si multiplicamos la primera desigualdad por $1 - t > 0$ y la segunda por $t > 0$ y las sumamos obtendremos que

$$f(u)(1 - t) + f(v)t \geq f((1 - t)u + tv)$$

Luego, f es función convexa. Sea ahora $a \in \text{int}(I)$, $x \in I$ con $x > a$, por ser f convexa $s_a(z) \leq s_a(x)$ para todo $z \in (a, x)$. Luego si $z \rightarrow a^+$, tenemos $f'_+(a) \leq s_a(x)$. O lo que es lo mismo, $f(x) \geq f(a) + f'_+(a)(x - a)$. Si ahora $x < a$, similarmente, usando que f es convexa, y que $f'_-(a) \leq f'_+(a)$, obtenemos que $f(x) \geq f(a) + f'_+(a)(x - a)$. Por tanto, $f'_+(a) \in \partial f(a)$. De manera análoga, $f'_-(a) \in \partial f(a)$. Consecuentemente, una función φ como la del enunciado, verifica que $f'_-(a) \leq \varphi(a) \leq f'_+(a)$. Que φ es creciente en el interior de I es consecuencia inmediata del teorema (5). Encadenando desigualdades, sean $x, y \in \text{int}(I)$ con $x < y$. Por dicho teorema, $f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y)$ que junto con la desigualdad del enunciado nos da

$$f'_-(x) \leq \varphi(x) \leq f'_+(x) \leq f'_-(y) \leq \varphi(y) \leq f'_+(y)$$

Luego φ es creciente. □

Observación. La subdiferencial en los puntos extremos de I puede ser vacía. Un ejemplo de ello es la función $1 - \sqrt{1 - x^2}$ con $x \in [-1, 1]$.

Proposición 1.3.2. Sea f continua y convexa y $\varphi : I \rightarrow \mathbb{R}$ una función tal que $\varphi(x) \in \partial f(x) \quad \forall x \in \text{int}(I)$. Entonces,

$$f(z) = \sup_{x \in \text{int}(I)} \{f(x) + (z - x)\varphi(x)\} \quad \forall z \in I$$

Demostración. Por la definición (1.3.1) para puntos $z \in \text{int}(I)$ la igualdad del enunciado se da. Veamos qué pasa entonces para puntos extremos: sea z un punto extremo de I , digamos, por ejemplo, el de la izquierda, entonces si $t > 0$ es suficientemente pequeño tenemos que

$$f(z + t) - f(z) \leq t\varphi(z + t) \leq f(z + 2t) - f(z + t)$$

Obtenemos que $\exists \lim_{t \rightarrow 0^+} t\varphi(z + t) = 0$. Así, dado $\epsilon > 0 \quad \exists \delta > 0$ tal que si $0 < t < \delta$ se verifica que

$$|f(z) - f(z + t)| < \epsilon/2 \quad \text{y} \quad |t\varphi(z + t)| > \epsilon/2$$

Luego, si $0 < t < \delta$ entonces $f(z + t) - t\varphi(z + t) < f(z) + \epsilon$. Identificando $z + t$ con x y t con $z - x$ conseguimos el resultado buscado. □

Proposición 1.3.3. Sea f convexa. Entonces tiene un mínimo global en a si y sólo si $0 \in \partial f(a)$.

Demostración. Que f tenga un mínimo global en a significa que $f(x) \geq f(a) \quad \forall x \in I$. O sea, $f(x) \geq f(a) + 0 \cdot (x - a) \quad \forall x \in I$. Es decir, $0 \in \partial f(a)$. □

Capítulo 2

Desigualdades notables

En este capítulo mostraremos algunas desigualdades para funciones convexas que ponen de manifiesto la importancia de los temas tratados hasta ahora. Las dos primeras son discretas, mientras que las siguientes son integrales, relativas tanto a la recta real como a espacios de medida arbitrarios.

2.1. Desigualdad de Hardy-Littlewood-Polya

El siguiente resultado mostrará la importancia del concepto de *subdiferencial* introducido en el capítulo 1.

Teorema 7 (La desigualdad de Hardy-Littlewood-Polya).

Sea f convexa, y consideremos dos familias finitas de puntos de I , x_1, \dots, x_n e y_1, \dots, y_n tales que verifican

$$\sum_{k=1}^m x_k \leq \sum_{k=1}^m y_k \quad \forall m \in \{1, \dots, n\} \quad \text{y} \quad \sum_{k=1}^n x_k = \sum_{k=1}^n y_k$$

Entonces, si $x_1 \geq \dots \geq x_n$ tenemos que

$$\sum_{k=1}^n f(x_k) \leq \sum_{k=1}^n f(y_k)$$

Mientras que, si $y_1 \leq \dots \leq y_n$, tenemos la desigualdad contraria.

Demostración. Inducción sobre n . Si $n = 1$ es evidente gracias a la segunda hipótesis. Asumimos para familias de $n - 1$ puntos y lo probamos para las de n . Antes de esto, démonos cuenta de que podemos hacer una reducción del problema: debido a las hipótesis tenemos que $x_1, \dots, x_n \in [\text{mín } y_k, \text{máx } y_k]$, luego podemos suponer que $\text{mín } y_k < x_n \leq \dots \leq x_1 < \text{máx } y_k$. Así, es claro que $x_1, \dots, x_n \in \text{int}(I)$. Luego por (1.3.1), existe una $\varphi : \text{int}(I) \rightarrow \mathbb{R}$ creciente tal que $\varphi(x) \in \partial f(x) \quad \forall x \in \text{int}(I)$.

Aplicando la proposición (1.3.2) a f y φ , tenemos

$$f(y_k) \geq f(x_k) + (y_k - x_k)\varphi(x_k) \quad \forall k \in \{1, \dots, n\}$$

Así,

$$\begin{aligned} \sum_{k=1}^n f(y_k) - \sum_{k=1}^n f(x_k) &\geq \sum_{k=1}^n (f(x_k) + (y_k - x_k)\varphi(x_k) - f(x_k)) \\ &= \sum_{k=1}^n (y_k - x_k)\varphi(x_k) = \varphi(x_1)(y_1 - x_1) + \sum_{m=2}^n \varphi(x_m)(y_m - x_m) \\ &= \varphi(x_1)(y_1 - x_1) + \sum_{m=2}^n \varphi(x_m) \left[\sum_{k=1}^m (y_k - x_k) - \sum_{k=1}^{m-1} (y_k - x_k) \right] \\ &= \varphi(x_k) \sum_{k=1}^n (y_k - x_k) + \sum_{k=1}^{n-1} [(\varphi(x_m) - \varphi(x_{m+1})) \sum_{k=1}^m (y_k - x_k)] = (\diamond) \end{aligned}$$

donde la última igualdad se obtiene de la siguiente manera

$$\begin{aligned} &\sum_{k=1}^{n-1} [(\varphi(x_m) - \varphi(x_{m+1})) \sum_{k=1}^m (y_k - x_k)] = \\ &= \sum_{k=1}^{n-1} \varphi(x_m) \sum_{k=1}^m (y_k - x_k) - \sum_{k=1}^{n-1} \varphi(x_{m+1}) \sum_{k=1}^m (y_k - x_k) \\ &= \varphi(x_{n-1})(y_{n-1} - x_{n-1}) + \dots + \varphi(x_1)(y_1 - x_1) - \varphi(x_n) \sum_{k=1}^{n-1} (y_k - x_k) \end{aligned}$$

ya que los términos se van cancelando. Luego

$$(\diamond) = \sum_{m=1}^{n-1} [(\varphi(x_m) - \varphi(x_{m+1})) \sum_{k=1}^m (y_k - x_k)] \geq 0$$

pues φ es creciente, $x_m \leq x_{m+1}$ para todo $m \in \{1, \dots, n-1\}$ e $(y_k - x_k) \geq 0$ para todo $k \in \{1, \dots, n\}$. Resumiendo, $\sum_{k=1}^n f(y_k) - \sum_{k=1}^n f(x_k) \geq 0$ y hemos acabado la inducción. La segunda conclusión sale de la primera sin más que cambiar f por $\tilde{f} : I \rightarrow \mathbb{R}$, donde $\tilde{f}(x) := f(-x) \quad \forall x \in \tilde{I}$, siendo $\tilde{I} := \{-x : x \in I\}$. \square

Observación. Ya vimos en la primera sección la *desigualdad de Popoviciu*, teorema (3), la cual se obtenía como la generalización a ternas de puntos de la desigualdad que aparecía en el concepto de función punto-medio convexa, y cuya demostración estaba basada en la de éste último. Ahora bien, la desigualdad de Popoviciu puede obtenerse a partir del teorema (7). Sean $x, y, z \in I$, sin pérdida de generalidad podemos asumir que $x \geq y \geq z$, pues renombramos las variables si fuera necesario. Entonces, ocurre que

$$\frac{x+y}{2} \geq \frac{x+z}{2} \geq \frac{z+y}{2} \quad \text{y que} \quad x \geq \frac{x+y+z}{3} \geq z$$

Si $x \geq (x+y+z)/3 \geq y \geq z$, se sigue de aplicar el teorema a los puntos $x_1 = x$, $x_2 = x_3 = x_4 = \frac{x+y+z}{3}$, $x_5 = y$, $x_6 = z$ e $y_1 = y_2 = \frac{x+y}{2}$, $y_3 = y_4 = \frac{x+z}{2}$, $y_5 = y_6 = \frac{z+y}{2}$. Mientras que si $x \geq y \geq (x+y+z)/3 \geq z$ se aplica a $x_1 = x$, $x_2 = y$, $x_3 = x_4 = x_5 = \frac{x+y+z}{3}$, $x_6 = z$ e $y_1 = y_2 = \frac{x+y}{2}$, $y_3 = y_4 = \frac{x+z}{2}$, $y_5 = y_6 = \frac{z+y}{2}$. Usamos en ambos casos la segunda conclusión.

2.2. Desigualdad de Jensen-Steffensen

Veamos ahora otra aplicación de la desigualdad de Hardy-Littlewood-Pólya. Se trata de una *generalización clásica de la desigualdad de Jensen* de la primera sección que utiliza combinaciones lineales (no necesariamente convexas).

Teorema 8 (desigualdad de Jensen-Steffensen).

Sean $x_n \leq x_{n-1} \leq \dots \leq x_1$ puntos de un intervalo $[a, b]$, y sean $p_1, \dots, p_n \in \mathbb{R}$ que verifican que sus sumas parciales, $s_k = \sum_{i=1}^k p_i$ con $k \in \{1, \dots, n\}$, están relacionadas como sigue

$$0 \leq s_k \leq s_n \quad \forall k \in \{1, \dots, n\}, \quad y \quad s_n > 0$$

Entonces, cada función convexa f definida sobre el intervalo $[a, b]$ verifica la desigualdad siguiente

$$f\left(\frac{1}{s_n} \sum_{k=1}^n p_k x_k\right) \leq \frac{1}{s_n} \sum_{k=1}^n p_k f(x_k)$$

Demostración. Definamos $\bar{x} := (\sum_{k=1}^n p_k x_k) s_n^{-1}$ y $\bar{s}_k := s_n - s_{k-1} = \sum_{i=k}^n p_i \geq 0 \quad \forall k \in \{2, \dots, n\}$. Ocurren las dos propiedades siguientes

$$\begin{aligned} 1. \quad s_n(x_1 - \bar{x}) &= \left(\sum_{i=1}^n p_i\right) \left(x_1 - \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i}\right) = \sum_{i=1}^n p_i x_1 - \sum_{i=1}^n p_i x_i \\ &= \sum_{i=1}^n p_i (x_1 - x_i) = \sum_{j=2}^n (x_{j-1} - x_j) \bar{s}_j \geq 0, \quad \text{entonces} \\ &\text{se tiene que } x_1 - \bar{x} \geq 0 \end{aligned}$$

$$\begin{aligned} 2. \quad s_n(\bar{x} - x_n) &= \left(\sum_{i=1}^n p_i\right) \left(\frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n p_i} - x_n\right) = \sum_{i=1}^n p_i x_i - \sum_{i=1}^n p_i x_n \\ &= \sum_{i=1}^{n-1} p_i (x_i - x_n) = \sum_{i=1}^{n-1} (x_i - x_{i+1}) s_i \geq 0, \quad \text{entonces} \\ &\text{se tiene que } \bar{x} - x_n \geq 0 \end{aligned}$$

Así, $x_n \leq \bar{x} \leq x_1$. Reduciéndonos al caso en el que f es continua y convexa en $[a, b]$ y los puntos x_1, \dots, x_n pertenecen a (a, b) , lo podemos hacer gracias a proposición (1.2.1), estamos en condiciones de aplicar la proposición (1.3.1). Obtenemos entonces una función $\varphi : (a, b) \rightarrow \mathbb{R}$ tal que $\varphi(x) \in \partial f(x) \quad \forall x \in (a, b)$. Aplicando la proposición (1.3.2), si $c \in (a, b)$, como φ es creciente, entonces

$$f(z) - f(y) \geq \varphi(c)(z - y) \quad \text{si } z \geq y \geq c \quad (\star)$$

y

$$f(z) - f(y) \leq \varphi(c)(z - y) \quad \text{si } c \geq z \geq y \quad (\star\star)$$

Por otro lado, $\exists m \in \{1, \dots, n\}$ tal que $\bar{x} \in [x_{m+1}, x_m]$. Así, si $c = \bar{x}$

$$\begin{aligned} f\left(s_n^{-1} \sum_{i=1}^n p_i x_i\right) - s_n^{-1} \sum_{i=1}^n p_i f(x_i) &= \\ &= \sum_{i=1}^{m-1} (\varphi(\bar{x})(x_i - x_{i+1}) - f(x_i) + f(x_{i+1})) \frac{s_i}{s_n} + \\ &+ [\varphi(\bar{x})(x_m - \bar{x}) - f(x_m) + f(\bar{x})] \frac{s_m}{s_n} + \\ &+ [f(\bar{x}) - f(x_{m+1}) - \varphi(\bar{x})(\bar{x} - x_{m+1})] \frac{\bar{s}_{m+1}}{s_n} + \\ &+ \sum_{i=m+1}^{n-1} [f(x_i) - f(x_{i+1}) - \varphi(\bar{x})(x_i - x_{i+1})] \frac{\bar{s}_{i+1}}{s_n} \end{aligned}$$

luego, aplicando las desigualdades (*) y (**) en los sumandos superiores según corresponda, obtenemos que la suma de arriba es no positiva. Así,

$$f\left(s_n^{-1} \sum_{i=1}^n p_i x_i\right) \leq s_n^{-1} \sum_{i=1}^n p_i f(x_i) \quad \square$$

2.3. Desigualdad de Hermite-Hadamard

Proposición 2.3.1 (La desigualdad de Hermite-Hadamard). *Sea f convexa definida en $[a, b]$, entonces*

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b f(x) dx \leq \frac{f(a) + f(b)}{2}$$

Demostración. Como f convexa entonces $f(x) \leq f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$ para cualquier $x \in [a, b]$. Por proposición (1.2.1), f puede ser modificada en los extremos del intervalo para que sea convexa y continua en $[a, b]$, luego es integrable. Así,

$$\begin{aligned} \int_a^b f(x) dx &\leq f(a)(b-a) + \frac{f(b)-f(a)}{b-a} \int_a^b (x-a) dx \\ &= (b-a)[f(a) + (f(b)-f(a))(1/2)] \end{aligned}$$

que es la desigualdad de la derecha. En cuanto a la de la izquierda, simplemente subdividiendo la integral

$$\frac{1}{b-a} \int_a^b f(x) dx = \frac{1}{b-a} \left(\int_a^{\frac{a+b}{2}} f(x) dx + \int_{\frac{a+b}{2}}^b f(x) dx \right)$$

y juntándolas posteriormente en una sola con un cambio de variable apropiado

$$\frac{1}{2} \int_0^1 \left[f\left(\frac{a+b-t(b-a)}{2}\right) + f\left(\frac{a+b+t(b-a)}{2}\right) \right] dt \geq f\left(\frac{a+b}{2}\right)$$

donde la última desigualdad es gracias a la convexidad punto-medio de f (ver teorema (2)). \square

2.4. Fórmula *integral* de la desigualdad de Jensen

En este apartado mostramos una especie de generalización de la desigualdad de Hermite-Hadamard en la que intervienen espacios de medida más generales que un intervalo. Un *espacio de medida* es una terna (X, Σ, μ) con X un conjunto no vacío, Σ una σ -álgebra sobre X y μ una medida sobre (X, Σ) (en el apéndice se recuerdan estos conceptos brevemente).

Lo análogo a la media aritmética en el contexto de los espacios de medida finita (X, Σ, μ) , recordemos que ésto es que $\mu(X) < \infty$, es la *media aritmética integral* (ó simplemente media aritmética si no hay confusión). Se define para una función μ -integrable $f : X \rightarrow \mathbb{R}$ como

$$M_1(f; \mu) = \frac{1}{\mu(X)} \int_X f \, d\mu \in \mathbb{R}$$

También se denota como $M_1(f)$ por simplicidad.

Uno de los resultados básicos acerca de la media aritmética es la **desigualdad de Jensen**.

Teorema 9 (Desigualdad de Jensen. JENSEN, 1906).

Sea (X, Σ, μ) un espacio de medida finita y $g : X \rightarrow \mathbb{R}$ una función μ -integrable. Si f es convexa e I contiene a la imagen de g , entonces $M_1(g) \in I$, y además la siguiente desigualdad ocurre

$$f(M_1(g)) \leq M_1(f \circ g)$$

siempre que la composición sea μ -integrable

Demostración. Mostremos que $M_1(g) \in I$ por reducción al absurdo. Si no fuera así, la función $h := M_1(g) - g$ sería estrictamente positiva o estrictamente negativa. Supongamos, por ejemplo, que h es estrictamente positiva, tendríamos que

$$\int_X h = \int_X M_1(g) - \int_X g = M_1(g)\mu(X) - M_1(g)\mu(X) = 0$$

lo que no puede ser, pues la integral de una función estrictamente positiva no puede ser nula cuando es evaluada en todo X .

Por la proposición (1.3.1), como f es convexa, podemos elegir una función $\varphi : I \rightarrow \mathbb{R}$ tal que $\varphi(x) \in \partial f(x) \quad \forall x \in \text{int}(I)$. Ahora bien, si $M_1(g) \in \text{int}(I)$, como $g(x) \in I \quad \forall x \in X$ y $\varphi(M_1(g)) \in \partial f(M_1(g))$, utilizando la definición de la subdiferencial sabemos que

$$f(g(x)) \geq f(M_1(g)) + (g(x) - M_1(g))\varphi(M_1(g)) \quad \forall x \in X$$

integrando

$$\begin{aligned} M_1(f \circ g) \cdot \mu(X) &= \int_X f(g(x)) \, d\mu \\ &\geq \mu(X)f(M_1(g)) + \varphi(M_1(g)) \int_X (g(x) - M_1(g)) \, d\mu \\ &= \mu(X)f(M_1(g)) + \varphi(M_1(g)) \int_X g(x) \, d\mu - \\ &\quad - \varphi(M_1(g)) \cdot M_1(g) \cdot \mu(X) = \mu(X) \cdot f(M_1(g)) \end{aligned}$$

luego $f(M_1(g)) \leq M_1(f \circ g)$ pues $0 \leq \mu(X) < \infty$. Mientras que si $M_1(g)$ fuera un punto extremo de I , entonces se sigue directamente que $g = M_1(g)$ en casi todo punto. En efecto, digamos que $M_1(g) = a_0$, entonces a partir de la definición de media aritmética integral de la función g conseguimos que

$$\int_X g \, d\mu = a_0 \mu(X) \Rightarrow \int_X (g - a_0) \, d\mu = 0$$

pero como a_0 es un punto extremo de I , y la imagen de g está contenida en dicho intervalo, no queda otra que el integrando sea nulo en c.t.p.; es decir, $g - a_0 = 0$ en c.t.p. Luego, $g = M_1(g)$ en c.t.p. Así,

$$M_1(f \circ g) = \frac{1}{\mu(X)} \int_X f \circ g \, d\mu = \frac{1}{\mu(X)} \int_X f(M_1(g)) \, d\mu = f(M_1(g))$$

luego se da la igualdad y por lo tanto se cumple la desigualdad. \square

Observación. Si ponemos $X = I$ con la medida de Lebesgue usual y tomamos $g : I \rightarrow \mathbb{R}$ la inclusión natural, la desigualdad de Jensen proporciona la primera de las desigualdades de Hermite-Hadamard.

Observación. De la demostración del teorema se sigue que la *igualdad* en el mismo se tiene cuando $M_1(g)$ es un punto extremo de I . Además, nótese que la hipótesis de que $f \circ g$ es μ -integrable es necesaria. Basta tomar las funciones, $f(x) = -x^{1/2}$ y $g(x) = x^{-2}$ en $(1, \infty)$, de modo que f es convexa, y g es positiva e integrable con integral 1. La composición $(f \circ g)(x) = -x^{-1}$ no es integrable en $(1, \infty)$.

El teorema (9) nos permite establecer la **forma integral de la desigualdad media aritmética geométrica armónica**, cuya versión discreta hemos recordado y abordado más arriba en el primer capítulo. Si (X, Σ, μ) es un espacio de medida finita, f una función de $L^1(\mu)$ tal que $f \geq 0$, y definimos $\log 0 := -\infty$ y $e^{-\infty} := 0$, entonces definimos la *media geométrica (integral) de f en X* como

$$M_0(f; \mu) = \exp(\mu(X)^{-1} \int_X \log(f(x)) \, d\mu)$$

La desigualdad media aritmética-media geométrica es entonces

$$M_0(f) \leq M_1(f)$$

Para ver esta desigualdad, usamos la desigualdad de Jensen en su forma integral aplicada a la función logaritmo (que es cóncava, por lo que ésta cambia de sentido) nos da que

$$\frac{1}{\mu(X)} \int_X \log(f(x)) \, d\mu \leq \log \left(\frac{1}{\mu(X)} \int_X f(x) \, d\mu \right)$$

que tras tomar exponenciales, nos proporciona $M_0(f; \mu) \leq M_1(f; \mu)$, como queríamos.

Ahondando más en este tema, si admitimos que 0 e ∞ son recíprocos el uno del otro ($1/\infty = 0$, $1/0 = \infty$), podemos introducir también la *media armónica de f en su forma integral* como

$$M_{-1}(f; \mu) = \left(\frac{1}{\mu(X)} \int_X \frac{1}{f(x)} d\mu \right)^{-1}$$

Así, es claro que $M_{-1}(f) = (M_1(1/f))^{-1}$ y que

$$\begin{aligned} (M_0(1/f))^{-1} &= \exp \left(\frac{1}{\mu(X)} \int_X -\log f(x) d\mu \right)^{-1} \\ &= \exp \left(\frac{1}{\mu(X)} \int_X \log(f(x)) d\mu \right) = M_0(f) \end{aligned}$$

y por lo que acabamos de ver, $M_{-1}(f) = (M_1(1/f))^{-1} \leq (M_0(1/f))^{-1} = M_0(f)$.
Recopilando, conseguimos la cadena de desigualdades siguiente.

Proposición 2.4.1. *Si $f : X \rightarrow \mathbb{R}$ es no negativa con $f \in L^1(\mu)$, entonces se tiene que $M_{-1}(f) \leq M_0(f) \leq M_1(f)$.*

Teorema 10 (Forma completa de la desigualdad de Jensen).

Sea (X, Σ, μ) un espacio de medida finita y sea $g : X \rightarrow \mathbb{R}$ una función μ -integrable. Si f es convexa definida sobre un intervalo I que incluye la imagen de g y $\varphi : I \rightarrow \mathbb{R}$ es una función que verifica que

1. $\varphi(x) \in \partial f(x) \quad \forall x \in I$
2. $\varphi \circ g$ y $g(\varphi \circ g)$ son μ -integrables

entonces, tienen lugar las siguientes desigualdades

$$0 \leq M_1(f \circ g) - f(M_1(g)) \leq M_1(g(\varphi \circ g)) - M_1(g)M_1(\varphi \circ g)$$

Demostración. Puesto que estamos en unas condiciones más fuertes que en el caso del teorema (9), y la primera desigualdad del teorema (10) es ésta misma, ya sabemos que se cumple. Vayamos con la segunda desigualdad. Como antes $M_1(g)$, $g(x) \in I$, y $\varphi(g(x)) \in \partial f(g(x))$ para cualquier $x \in X$, utilizando la definición (1.3.2)

$$f(M_1(g)) \geq f(g(x)) + (M_1(g) - g(x))\varphi(g(x)) \quad \forall x \in X$$

e integrando sobre X

$$\begin{aligned} f(M_1(g)) &\geq \frac{1}{\mu(X)} \left(\int_X (f \circ g) + M_1(g) \int_X (\varphi \circ g) - \int_X g(\varphi \circ g) \right) \\ &= M_1(f \circ g) + M_1(g)M_1(\varphi \circ g) - M_1(g(\varphi \circ g)) \end{aligned}$$

luego

$$M_1(f \circ g) - f(M_1(g)) \leq M_1(g(\varphi \circ g)) - M_1(g)M_1(\varphi \circ g)$$

que es la segunda desigualdad. □

Corolario 2.4.2 (Versión discreta del teorema (10)). Sea f convexa y $\varphi : I \rightarrow \mathbb{R}$ una función con $\varphi(x) \in \partial f(x) \quad \forall x \in I$. Entonces,

$$\begin{aligned} 0 &\leq \sum_{k=1}^n \lambda_k f(x_k) - f\left(\sum_{k=1}^n \lambda_k x_k\right) \\ &\leq \sum_{k=1}^n \lambda_k x_k \varphi(x_k) - \left(\sum_{k=1}^n \lambda_k x_k\right) \left(\sum_{k=1}^n \lambda_k \varphi(x_k)\right) \end{aligned}$$

para cualesquiera $x_1, \dots, x_n \in I, \lambda_1, \dots, \lambda_n \in [0, 1]$ con $\sum_{k=1}^n \lambda_k = 1$.

Demostración. Basta tomar como espacio el asociado a la medida de contar (ver Apéndice). \square

Corolario 2.4.3 (Aplicación a triángulos del plano). Sean $\alpha_1, \alpha_2, \alpha_3 \in (0, \pi/2)$ ángulos de un triángulo acutángulo. Entonces, tenemos que

$$\frac{3\sqrt{3}}{2} - \sum_{i=1}^3 \left(\frac{\pi}{3} - \alpha_i\right) \cos \alpha_i \leq \sum_{i=1}^3 \operatorname{sen} \alpha_i \leq \frac{3\sqrt{3}}{2}$$

Demostración. Consideremos $f(x) := \operatorname{sen} x \quad \forall x \in (-\pi/2, 0)$, es convexa; y tomamos $\varphi(x) := \cos x$. Como f es derivable, $\partial f(x) = \{f'(x)\} \quad \forall x \in (-\pi/2, 0)$. Así, $\varphi(x) \in \partial f(x) = \{\cos x\} \quad \forall x \in (-\pi/2, 0)$. Luego, estamos en condiciones de aplicar el corolario anterior a los puntos $x_i = -\alpha_i$

$$\begin{aligned} 0 &\leq \sum_{i=1}^3 \frac{1}{3} \operatorname{sen}(-\alpha_i) - f\left(\sum_{i=1}^3 (-\alpha_i)\right) \\ &= -\frac{1}{3} \sum_{i=1}^3 \operatorname{sen}(\alpha_i) - f(-\pi/3) = \frac{\sqrt{3}}{2} - \frac{1}{3} \sum_{i=1}^3 \operatorname{sen}(\alpha_i) \\ &\leq \sum_{i=1}^3 \frac{1}{3} (-\alpha_i) \varphi(-\alpha_i) - \left(\sum_{i=1}^3 \frac{1}{3} (-\alpha_i)\right) \left(\sum_{i=1}^3 \frac{1}{3} \varphi(-\alpha_i)\right) \\ &= -\frac{1}{3} \sum_{i=1}^3 \alpha_i \cos(\alpha_i) + \frac{\pi}{3} \left(\frac{1}{3} \sum_{i=1}^3 \cos \alpha_i\right) = \frac{1}{3} \sum_{i=1}^3 \left(\frac{\pi}{3} - \alpha_i\right) \cos \alpha_i \quad \square \end{aligned}$$

2.5. Desigualdad de Hardy

Ahora vamos a usar la forma integral de la desigualdad de Jensen para probar otro importante y clásico resultado, conocido como **desigualdad de Hardy**. Pero antes necesitamos un pequeño lema.

Lema 2.5.1. Sean $0 < b < +\infty$ y $-\infty \leq a < c \leq +\infty$, si u es una función convexa positiva sobre (a, c) entonces se verifica la desigualdad siguiente

$$\int_0^b u\left(\frac{1}{x} \int_0^x h(t) dt\right) \frac{dx}{x} \leq \int_0^b u(h(x)) \left(1 - \frac{x}{b}\right) \frac{dx}{x}$$

para cualquier función integrable $h : (0, b) \rightarrow (a, c)$.

Demostración. Sea $0 < x < b$, por la desigualdad de Jensen, $u(M_1(h)) \leq M_1(u \circ h)$, equivalentemente, restringiendo el dominio de la función h al intervalo $(0, x)$

$$u\left(\frac{1}{x} \int_0^x h(t) dt\right) \leq \frac{1}{x} \int_0^x u(h(t)) dt$$

Integrando entre 0 y b

$$\int_0^b u\left(\frac{1}{x} \int_0^x h(t) dt\right) \frac{dx}{x} \leq \int_0^b \int_0^b \frac{1}{x^2} u(h(t)) \chi_{[0,x]}(t) dt dx = (\spadesuit)$$

Si llamamos $f(x, t) := \frac{1}{x^2} u(h(t)) \chi_{[0,x]}(t)$ definida en $(0, b)^2$, nos encontramos trabajando en un subconjunto de medida finita de \mathbb{R}^2 con la sigma-álgebra de los conjuntos medibles Lebesgue y la medida de Lebesgue y $f \geq 0$. Luego, gracias al teorema de Fubini-Tonelli¹

$$\begin{aligned} (\spadesuit) &= \int_0^b \int_0^b \frac{1}{x^2} u(h(t)) \chi_{[0,x]}(t) dx dt = \int_0^b u(h(t)) \int_0^b \frac{1}{x^2} \chi_{[0,x]}(t) dx dt \\ &= \int_0^b u(h(t)) \int_t^b \frac{1}{x^2} dx dt = \int_0^b u(h(t)) \left(1 - \frac{t}{b}\right) \frac{dt}{t} \quad \square \end{aligned}$$

Teorema 11 (Desigualdad de Hardy).

Sea $1 < p < \infty$, $f \in L^p(0, \infty)$ con $f \geq 0$. Si definimos la función F como

$$F(x) := \frac{1}{x} \int_0^x f(t) dt \quad \forall x > 0$$

entonces se tiene que $\|F\|_p \leq \frac{p}{p-1} \|f\|_p$. Con igualdad sii $f = 0$ en ctp.

Demostración. Consideremos $p \in (1, \infty)$ y la función $u(x) := |x|^p$, que es convexa en todo \mathbb{R} (por ejemplo usando el criterio de la segunda derivada). Consideremos ahora nuestra función u , y sean $\lambda, \mu \in [0, 1]$ tales que su suma da uno, y $x, y \in \mathbb{R}$. Entonces, utilizando la desigualdad triangular y que la función anterior es convexa en $[0, \infty)$ tenemos que

$$u(\lambda x + \mu y) = |\lambda x + \mu y|^p \leq (\lambda|x| + \mu|y|)^p \leq \lambda|x|^p + \mu|y|^p$$

por lo que u es convexa en todo \mathbb{R} . Apliquemos ahora a ésta u el lema anterior, obtenemos que

$$\int_0^\alpha \left| \frac{1}{x} \int_0^x f(t) dt \right|^p dx \leq \left(\frac{p}{p-1} \right)^p \int_0^\alpha |f(x)|^p \left(1 - \left(\frac{x}{\alpha} \right)^{\frac{p-1}{p}} \right) dx$$

donde $\alpha = b^{p/(p-1)}$ y $f(x) = h(x^{1-1/p})x^{-1/p}$. Así hemos conseguido lo análogo a la desigualdad de Hardy para funciones $f \in L^p(0, \alpha)$, con $0 < \alpha < +\infty$, y basta con hacer tender α a infinito. En cuanto a la igualdad, si se da, entonces F y f son proporcionales; es decir, f es de la forma Cx^r con C una constante. Como $f \in L^p(0, \infty)$, entonces ésto sólo es posible si la constante es nula. Y por lo tanto $f = 0$ en ctp. \square

¹ En la pág. 102 de [Cer00] aparece este resultado para espacios de medida arbitrarios

Observación. La constante que aparece en la desigualdad de Hardy no puede ser mejorada. Para ello, encontremos una sucesión (f_n) de $L^p(0, \infty)$ de manera que la sucesión de números reales $(\|F_n\|_p/\|f_n\|_p)$ converja a $p/(p-1)$ cuando $n \rightarrow \infty$. Sea $n \in \mathbb{N}$ y definamos

$$f_n(t) := t^{\frac{1}{n}-\frac{1}{p}} \chi_{[0,1]}(t) \quad \forall t > 0$$

Como

$$F_n(x) := \frac{1}{x} \int_0^x f_n(t) dt = \frac{1}{x} \int_0^x t^{\frac{1}{n}-\frac{1}{p}} \chi_{[0,1]}(t) dt$$

entonces,

$$\begin{aligned} \|F_n\|_p^p &= \int_0^\infty \frac{1}{x^p} \left(\int_0^x t^{\frac{1}{n}-\frac{1}{p}} \chi_{[0,1]}(t) dt \right)^p dx \\ &= \int_0^1 \frac{1}{x^p} \left(\frac{x^{\frac{p-1}{p}+\frac{1}{n}}}{\frac{p-1}{p}+\frac{1}{n}} \right)^p dx + \int_1^\infty \frac{1}{x^p} \left(\frac{1}{\frac{p-1}{p}+\frac{1}{n}} \right)^p dx = (\star) \end{aligned}$$

ya que, de manera general, si $a \in (0, \infty)$, como $1 + 1/n - 1/p > 0$, tenemos que

$$\int_0^a t^{\frac{1}{n}-\frac{1}{p}} dt = \frac{t^{\frac{p-1}{p}+\frac{1}{n}}}{\frac{p-1}{p}+\frac{1}{n}} \Big|_0^a = \frac{a^{\frac{p-1}{p}+\frac{1}{n}}}{\frac{p-1}{p}+\frac{1}{n}}$$

Calculemos ahora las dos integrales de la expresión (\star) . Empecemos por la última, para una mayor sencillez denotemosla por II y definamos $\alpha_n := \left(\frac{p-1}{p} + \frac{1}{n}\right)^{-p}$. Entonces,

$$II = \alpha_n \int_1^\infty \frac{1}{x^p} dx = \alpha_n \cdot \frac{1}{1-p}$$

En cuanto a la primera integral, denotada por I

$$I = \alpha_n \int_0^1 \frac{1}{x^p} \left(x^{\frac{p-1}{p}+\frac{1}{n}} \right)^p dx = \alpha_n \int_0^1 x^{\frac{p}{n}-1} dx = \alpha_n \cdot \frac{x^{p/n}}{p/n} \Big|_0^1 = \alpha_n \cdot \frac{n}{p}$$

Luego,

$$(\star) = \alpha_n \left(\frac{1}{1-p} + \frac{n}{p} \right)$$

Pasemos ahora a calcular la norma de f_n ,

$$\|f_n\|_p^p = \int_0^\infty \left(t^{\frac{1}{n}-\frac{1}{p}} \chi_{[0,1]}(t) \right)^p dt = \int_0^1 t^{\frac{p}{n}-1} dt = \frac{n}{p}$$

Tenemos así, puesto que $\alpha_n \rightarrow ((p-1)/p)^{-p}$ cuando $n \rightarrow \infty$, que la sucesión de números reales $(\|F_n\|_p/\|f_n\|_p)$ converge a $\left(\frac{p}{p-1}\right)^p$. En efecto,

$$\frac{\|F_n\|_p^p}{\|f_n\|_p^p} = \frac{\alpha_n \left(\frac{1}{1-p} + \frac{n}{p} \right)}{\frac{n}{p}} = \alpha_n \left(\frac{p}{1-p} \cdot \frac{1}{n} + 1 \right) \rightarrow \left(\frac{p}{p-1} \right)^p$$

cuando $n \rightarrow \infty$. Por lo tanto, $\|F_n\|_p/\|f_n\|_p \rightarrow p/(p-1)$ cuando $n \rightarrow \infty$. Lo cual era justamente lo que buscábamos.

Capítulo 3

Ejemplos clásicos. Las funciones Γ y β

En este capítulo nos centraremos en el estudio detallado de las funciones Gamma y Beta: definiciones, características, propiedades, estimaciones, relaciones entre ellas, ... Puesto que estas dos funciones pertenecen a una clase muy importante, la clase de las funciones logarítmicamente convexas (log-convexas), la primera sección del capítulo irá dedicada a una pequeña introducción acerca de esta clase de funciones.

3.1. La clase de las funciones logarítmicamente convexas

Sea I un intervalo, se dice que una función $f : I \rightarrow (0, \infty)$ es **logarítmicamente convexa** (ó log-convexa) si cumple que

$$f((1 - \lambda)x + \lambda y) \leq f(x)^{1-\lambda} f(y)^\lambda \quad x, y \in I \quad \forall \lambda \in [0, 1]$$

Una primera y sencilla observación que debemos hacer es que si una función es log-convexa, entonces es convexa. En efecto, utilizando la desigualdad media armónica-media geométrica, vista en el primer capítulo, tenemos que

$$f((1 - \lambda)x + \lambda y) \leq f(x)^{1-\lambda} f(y)^\lambda \leq (1 - \lambda)f(x) + \lambda f(y)$$

para cualesquiera $x, y \in I$ y $\lambda \in [0, 1]$. Por tanto f es una función convexa. Nótese que al revés no ocurre, un contraejemplo de ello es la función $g(x) = e^x - 1$ definida en $(0, +\infty)$, que es convexa y no es log-convexa, pues basta tomar los valores concretos $x = 1/2, y = 1$ y $\lambda = 1/2$ para que ocurra que $g((1 - \lambda)x + \lambda y) > g(x)^{1-\lambda} g(y)^\lambda$.

Por otra parte, a partir de la definición se tiene que f es log-convexa si y sólo si $\log f$ es convexa. De ahí la utilización de la expresión y notación log-convexa. Veamos esta equivalencia: tomando logaritmos en la definición de log-convexidad llegamos a que, dados $x, y \in I$ con $\lambda \in [0, 1]$

$$\begin{aligned} \log(f((1 - \lambda)x + \lambda y)) &\leq \log(f(x)^{1-\lambda} f(y)^\lambda) = \log(f(x)^{1-\lambda}) + \log(f(y)^\lambda) \\ &= (1 - \lambda) \log(f(x)) + \lambda \log(f(y)) \end{aligned}$$

pues \log es función estrictamente creciente. Nótese que tomar logaritmos tiene sentido debido a que la función f tiene como espacio de llegada el intervalo abierto $(0, \infty)$ por pertenecer a la clase de las funciones log-convexas. Para el recíproco, si partimos de que $\log \circ f$ es una función convexa, entonces, por definición, se cumple que

$$\begin{aligned} \log(f((1-\lambda)x + \lambda y)) &\leq (1-\lambda)\log(f(x)) + \lambda\log(f(y)) \\ &= \log(f(x)^{1-\lambda}) + \log(f(y)^\lambda) \\ &= \log(f(x)^{1-\lambda} \cdot f(y)^\lambda) \end{aligned}$$

luego, utilizando de nuevo que \log es creciente, deducimos que f es convexa.

3.2. Estudio de las funciones Gamma y Beta.

3.2.1. La función Gamma

Definición 3.2.1 (La función Gamma). Ésta¹ es una aplicación de los reales no negativos con valores en \mathbb{R} , definida y denotada como²

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt \quad \text{para todo } x \in (0, \infty)$$

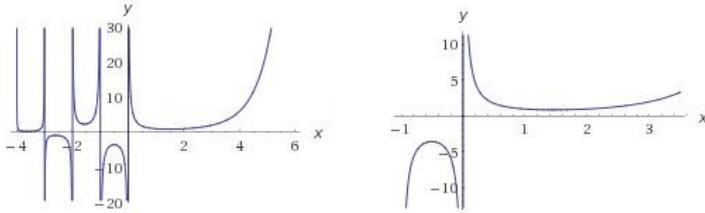


Figura 3.1: Función Γ a diferente escala y zoom

Comprobemos que *está bien definida*; es decir, que la integral impropia converge. Sea $x > 0$, dividimos el dominio de integración en dos partes, a saber, $[0, 1]$ y $[1, \infty]$, y estudiamos la convergencia de cada integral asociada a estos subdominios por separado. La correspondiente al subdominio $[0, 1] \ni t$, como $0 < e^{-t} < 1$, se puede mayorar por $\int_0^1 t^{x-1} dt = \frac{1}{x}$, luego converge. Para la otra parte distinguimos dos casos.

Caso 1. Si $x \leq 1$, entonces $t^{x-1} \leq 1$ y podemos mayorar la correspondiente integral por $\int_1^\infty e^{-t} dt = e^{-1}$.

¹ Las imágenes de la función Γ han sido hechas *online* con el software *WolframAlpha*[®]: *Computational Knowledge Engine*.

² Puesto que a lo largo del trabajo no entramos en el *mundo complejo*, para mantener la coherencia con el resto del texto, aquí únicamente tratamos la función Γ real y para probar sus características nos ceñimos a argumentos de variable real. Aunque, si definimos dicha función en los complejos obtenemos propiedades más completas e interesantes, con mayores aplicaciones. Ésto, sin embargo, conlleva un mayor esfuerzo; pues es necesario saber acerca de la teoría de productos infinitos de funciones holomorfas, entre otras cosas. Un muy buen libro de variable compleja para ver todo ésto en detalle es el de John B. Conway: *Functions of one complex variable I*, Graduate Texts in mathematics, 2^a edición, Springer (vol. 11), referencia [Con].

Caso 2. Si $x > 1$, usamos el criterio de comparación por paso al límite con la función $g(t) = t^{-2} > 0$. Tenemos que $t^{x-1}e^{-t}/g(t) = t^{x+1}e^{-t} \rightarrow 0$ cuando $t \rightarrow \infty$. En efecto, como $[x] + 2 > x + 1$ y $t \geq 1$, entonces $t^{[x]+2} \geq t^{x+1} \geq 0$, donde $\lim_{t \rightarrow \infty} e^{-t} \cdot t^{[x]+2} = 0$ sin más que aplicar la *regla de L'Hôpital* $[x] + 2$ veces. Así, $\exists \lim_{t \rightarrow \infty} e^{-t} \cdot t^{x+1} = 0$, y puesto que la integral de g sobre el intervalo $[1, \infty)$ converge, también lo hará la de $t^{x-1}e^{-t}$.

Ahora, vamos a tratar de establecer diferentes expresiones que involucren a la función Gamma y el teorema de la unicidad de Γ como extensión logarítmicamente convexa de la función factorial, así como sus propiedades. Para ello necesitamos tener en mente el **teorema de derivación bajo el signo integral** (que se obtiene como consecuencia del **teorema de la convergencia dominada de Lebesgue**) (ver A). Éste teorema nos permitirá ver que Γ es derivable; es más, nos dirá que Γ es de clase $C^\infty(0, \infty)$.

Aplicando directamente el teorema de derivación bajo el signo integral a la función Γ , resulta que es de clase C^∞ y que

$$\Gamma^{(n)}(x) = \int_0^\infty t^{x-1} e^{-t} (\log t)^n dt$$

En efecto, tomemos como función $f(t, x)$ el integrando de Γ , definido en $(0, \infty)^2$. Nótese que, para cada $t > 0$ la función de una variable $f(t, \cdot) = f_t(\cdot)$ es trivialmente continua y derivable en $(0, \infty)$; y que si $0 < \alpha < 1 < \beta < +\infty$ y $x \in [\alpha, \beta]$ se tienen las acotaciones

$$0 < t < 1 : \quad f(t, x) \leq t^{x-1} \leq t^{\alpha-1} \Rightarrow \left| \frac{\partial^n f}{\partial x^n}(t, x) \right| \leq t^{\alpha-1} |\log t|^n, \quad \forall n \in \mathbb{N}$$

$$1 < t < +\infty : \quad f(t, x) \leq e^{-t} t^{\beta-1} \Rightarrow \left| \frac{\partial^n f}{\partial x^n}(t, x) \right| \leq e^{-t} t^{\beta-1} (\log t)^n, \quad \forall n \in \mathbb{N}$$

En efecto, sean $(t, x) \in (0, \infty)^2$, por ejemplo, como $\frac{\partial f}{\partial x}(t, x) = e^{-t} \frac{\partial}{\partial x}(t^{x-1}) = f(t, x) \log(t)$, gracias a las cotas de arriba obtenemos el caso para $n = 1$ cuando $0 < t < 1$. El caso n -ésimo se demuestra sin dificultad por inducción sobre n . Simplemente hay que hacer notar que las sucesivas derivadas parciales n -ésimas aumentan en una unidad el exponente del logaritmo. Análogamente se tiene el caso en el que $1 < t < \infty$. Y puesto que las funciones, $t^{\alpha-1} |\log t|^n \geq 0 \quad \forall t \in (0, 1)$ y $e^{-t} t^{\beta-1} (\log t)^n \geq 0 \quad \forall t \in (1, +\infty)$, son integrables, podemos aplicar a cada caso el th. de derivación bajo el signo integral obteniendo lo buscado.

Ahora que ya conocemos la regularidad de la función Gamma, vamos a establecer algunas de sus propiedades y características, así como diferentes fórmulas y expresiones que la involucran. Partimos con tres de las más sencillas.

Teorema 12 (Propiedades elementales de la función Gamma).

La función Γ verifica que

1. $\Gamma(x + 1) = x\Gamma(x)$ para todo $x > 0$.
2. $\Gamma(1) = 1$.
3. Γ es log-convexa.

Demostración. Las dos primeras propiedades son de comprobación inmediata. En cuanto a la tercera, tomemos $x, y > 0$ y sean $\lambda, \mu \geq 0$ tales que $\lambda + \mu = 1$. Así

$$\Gamma(\lambda x + \mu y) = \int_0^\infty t^{\lambda x + \mu y} e^{-t} dt = \int_0^\infty (t^{x-1} e^{-t})^\lambda (t^{y-1} e^{-t})^\mu dt \quad (\star)$$

Si consideramos λ y μ distintos de cero (pues en otro caso la desigualdad que garantiza que Γ es log-convexa sería trivial) y los expresamos como $1/p$ y $1/q$ respectivamente y aplicamos la desigualdad de Hölder a las funciones $(t^{x-1} e^{-t})^\lambda$ y $(t^{y-1} e^{-t})^\mu$, obtenemos que

$$(\star) \leq \left(\int_0^\infty (t^{x-1} e^{-t})^{\lambda p} dt \right)^{1/p} \left(\int_0^\infty (t^{y-1} e^{-t})^{\mu q} dt \right)^{1/q} = \Gamma(x)^\lambda \Gamma(y)^\mu$$

lo que garantiza la log-convexidad de la función Γ . □

Corolario 3.2.1. $\Gamma(n + 1) = n!$ para todo $n \in \mathbb{N} \cup \{0\}$.

Corolario 3.2.2. La función Gamma es convexa y $\lim_{x \rightarrow 0^+} x\Gamma(x) = 1$.

Corolario 3.2.3 (Mínimo absoluto de la función Gamma). La función Gamma tiene un mínimo absoluto en $(0, +\infty)$.

Demostración. Sabemos que $\lim_{n \rightarrow \infty} \Gamma(n) = +\infty$, y que $\Gamma(x) = x\Gamma(x)/x \rightarrow \infty$ cuando $x \rightarrow 0^+$. Luego, Γ no es monótona en $(0, \infty)$. Entonces, por las propiedades de las convexas obtenemos que tiene un mínimo absoluto en $(0, \infty)$. □

Como curiosidad cabe señalar que fue el gran Carl-Friedrich Gauss el primero en averiguar que la función Gamma alcanza su mínimo absoluto en $x = 1,461632145\dots$

Siguiendo con las propiedades, hay una especialmente relevante: la función Gamma es la única extensión logarítmicamente convexa de la función factorial.

Teorema 13 (H. BOHR y J. MOLLERUP, 1922).

Supongamos que la función $f : (0, \infty) \rightarrow \mathbb{R}$ cumple las tres siguientes propiedades

1. $f(x+1) = xf(x)$ para todo $x > 0$.
2. $f(1) = 1$.
3. f es log-convexa.

Entonces f es la función Γ .

Demostración. Al igual que hicimos con la función Γ , podemos obtener a partir de las dos primeras propiedades que $f(n+1) = n!$ para todo $n \in \mathbb{N}$. Sea ahora, $x \in (0, 1]$ y $n \in \mathbb{N}$. Entonces, gracias a (3) y a (1)

$$\begin{aligned} f(n+1+x) &= f((1-x)(n+1) + x(n+2)) \leq (f(n+1))^{1-x} (f(n+2))^x \\ &= (f(n+1))^{1-x} (n+1)^x (f(n+1))^x = (n+1)^x f(n+1) \\ &= (n+1)^x n! \end{aligned}$$

Además, también tenemos que

$$\begin{aligned} n! &= f(n+1) = f(x(n+x) + (1-x)(n+1+x)) \\ &\leq (f(n+x))^x (f(n+x+1))^{1-x} \\ &= (f(n+x+1))^x (n+x)^{-x} (f(n+x+1))^{1-x} \\ &= (n+x)^{-x} f(n+x+1) \end{aligned}$$

Por otra parte

$$\begin{aligned} f(n+x+1) &= (n+x)f(n+x) = (n+x)(n-1+x)f(n-1+x) \\ &= \cdots = (n+x)(n-1+x) \cdots xf(x) \end{aligned}$$

Luego, uniendo convenientemente todas las desigualdades

$$\left(1 + \frac{x}{n}\right)^x \leq \frac{(n+x)(n-1+x) \cdots xf(x)}{n! \cdot n^x} \leq \left(1 + \frac{1}{n}\right)^x$$

conseguimos, $f(x) = \lim_{n \rightarrow \infty} \frac{n! \cdot n^x}{(n+x)(n-1+x) \cdots x}$ cuando $x \in (0, 1]$.

Ahora probaremos que la fórmula de arriba no sólo es válida para los $x \in (0, 1]$, sino para todo $x > 0$. Así, como dicha fórmula ha sido obtenida sólo de las condiciones (1), (2) y (3) del enunciado, la función f quedará inequívocamente determinada por esas condiciones. Consecuentemente, como Γ cumple las tres hipótesis antes mencionadas, entonces $f = \Gamma$. Luego, **no hay más que una función log-convexa que extienda a la función factorial, y ésta es la función Gamma.**

Supongamos que $x > 0$ y elegimos un natural m tal que $0 < x - m \leq 1$. Así,

gracias a (1) y a la fórmula de arriba, tenemos que

$$\begin{aligned}
 f(x) &= (x-1)f(x-1) = (x-1)(x-2)\cdots(x-m)f(x-m) \\
 &= (x-1)\cdots(x-m) \cdot \lim_{n \rightarrow \infty} \frac{n! \cdot n^{x-m}}{(n+x-m)(n-1+x-m)\cdots(x-m)} \\
 &= \lim_{n \rightarrow \infty} \left(\frac{n! \cdot n^x}{(n+x)(n-1+x)\cdots x} \cdot \frac{(n+x)(n+x-1)\cdots(n+x-m+1)}{n^m} \right) \\
 &= \lim_{n \rightarrow \infty} \left(\frac{n! \cdot n^x}{(n+x)(n-1+x)\cdots x} \right) \cdot \lim_{n \rightarrow \infty} \left(\left(1 + \frac{x}{n}\right) \cdots \left(1 + \frac{x-m+1}{n}\right) \right) \\
 &= \lim_{n \rightarrow \infty} \left(\frac{n! \cdot n^x}{(n+x)(n-1+x)\cdots x} \right), \quad \text{que es la expresión buscada.}
 \end{aligned}$$

Pues, el límite de la derecha es 1. \square

Corolario 3.2.4. Para cada $x > 0$ se tiene que

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt = \lim_{n \rightarrow \infty} \left(\frac{n! \cdot n^x}{(n+x)(n-1+x)\cdots x} \right)$$

Relación trigonométrica de Gamma. Una vez vistas las propiedades básicas de la función Γ , la siguiente parte del capítulo corresponde al intento de establecer una relación entre dicha función y la función sen . Pero antes debemos saber escribir $\text{sen}(x)$ como un producto infinito.

Teorema 14 (L. EULER).

Sea $x \in \mathbb{R}$, entonces

$$\text{sen}(x) = x \cdot \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2 \pi^2} \right)$$

Demostración. En primer lugar veamos que fijado $x \in \mathbb{R}$, el producto infinito converge. Para ello utilicemos la caracterización que afirma que si $p_k \geq 0$ para todo $k \in \mathbb{N}$, un producto infinito de la forma $\prod_{k=1}^{\infty} (1 - p_k)$ converge si y solo si la serie $\sum_{k=1}^{\infty} p_k$ converge. En nuestro caso, $p_k = \frac{x^2}{k^2 \pi^2} \geq 0$ y $\sum_{k=1}^{\infty} p_k = \frac{x^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} < +\infty$.

Gracias a la *fórmula de D'Moivre*, $(\cos(x) + i \cdot \text{sen}(x))^n = \cos(nx) + i \cdot \text{sen}(nx) \quad \forall n \in \mathbb{N}$, obtenemos que $\text{sen}(2n+1)\theta$ es un polinomio de grado $2n+1$ en la variable $\text{sen} \theta$ para cada $n \in \mathbb{N}$ fijado. En efecto, sea $n \in \mathbb{N}$ fijo, gracias al desarrollo ofrecido por el binomio de Newton y la fórmula de D'Moivre

$$\begin{aligned}
 (\cos x + i \text{sen} x)^{2n+1} &= \sum_{k=0}^{2n+1} \binom{2n+1}{k} \cos^k x (i \text{sen} x)^{2n+1-k} \\
 &= \cos(2n+1)x + i \text{sen}(2n+1)x
 \end{aligned}$$

Como $i^2 = -1$, el sumatorio de arriba es igual a

$$\begin{aligned} \sum_{j=0}^n \binom{2n+1}{2j} \cos^{2j} x (-1)^{n-j} i \operatorname{sen}^{2(n-j)+1} x + \\ + \sum_{j=0}^n \binom{2n+1}{2j+1} \cos^{2j+1} x (-1)^{n-j} \operatorname{sen}^{2n-2j} x \end{aligned}$$

Igualando las partes imaginarias de la expresión primera y usando que para cada $j \in \{0, \dots, n\}$, $\cos^{2j} x = (\cos^2 x)^j = (1 - \operatorname{sen}^2 x)^j = \sum_{l=0}^j \binom{j}{l} (-1)^{j-l} \operatorname{sen}^{2(j-l)} x$

$$\begin{aligned} \operatorname{sen}(2n+1)x &= \sum_{j=0}^n \binom{2n+1}{2j} \cos^{2j} x (-1)^{n-j} \operatorname{sen}^{2(n-j)+1} x \\ &= \sum_{j=0}^n \sum_{l=0}^j \binom{2n+1}{2j} \binom{j}{l} (-1)^{n-l} \operatorname{sen}^{2(n-l)+1} x \end{aligned}$$

luego $\operatorname{sen}(2n+1)\theta$ es un polinomio en la variable $\operatorname{sen} \theta$ de grado $2n+1$, pues el coeficiente que acompaña al término de mayor orden es $(-1)^n \sum_{j=0}^n \binom{2n+1}{2j}$, que no es nulo. Lo cual era justamente lo que buscábamos. Además, puesto que $\operatorname{sen} \theta$ es una función impar y se anula cuando $\theta = k\pi$ con $k \in \mathbb{Z}$, entonces nuestro polinomio tendrá raíces en los valores $\pm \operatorname{sen}(k\pi/(2n+1))$ con $k = 0, \dots, n$. Nótese que hay exactamente $2n+1$ raíces, pues son 2 posibilidades para el signo (que sale gracias a que el seno es función impar) y $(n+1)$ para los valores que puede tomar k , pero como cuando $k = 0$ el signo no determina nada, entonces en total habrá $2n+1$ posibilidades. Además, es evidente que todas las raíces son distintas ya que la función sen es estrictamente creciente en $[-\frac{\pi}{2}, \frac{\pi}{2}]$ y

$$\begin{aligned} 0 < \frac{\pi}{2n+1} < \frac{2\pi}{2n+1} < \dots < \frac{(n-1)\pi}{2n+1} < \frac{n\pi}{2n+1} < \frac{\pi}{2} \\ 0 > -\frac{\pi}{2n+1} > -\frac{2\pi}{2n+1} > \dots > -\frac{(n-1)\pi}{2n+1} > -\frac{n\pi}{2n+1} > -\frac{\pi}{2} \end{aligned}$$

luego están todas. Si denotamos $\operatorname{sen}(2n+1)\theta = P(\theta)$, siendo P un polinomio de grado $2n+1$, y si α representa al coeficiente director del polinomio P , podemos escribir

$$\begin{aligned} P(X) &= \alpha X \left(X^2 - \operatorname{sen}^2 \frac{\pi}{2n+1} \right) \dots \left(X^2 - \operatorname{sen}^2 \frac{n\pi}{2n+1} \right) \\ &= \alpha X \prod_{k=1}^n \left(X^2 - \operatorname{sen}^2 \frac{k\pi}{2n+1} \right) \end{aligned}$$

Luego

$$\begin{aligned} \operatorname{sen}(2n+1)\theta &= P(\operatorname{sen} \theta) = \alpha \operatorname{sen} \theta \prod_{k=1}^n \left(\operatorname{sen}^2 \theta - \operatorname{sen}^2 \frac{k\pi}{2n+1} \right) \\ &= (-1)^n \alpha \operatorname{sen} \theta \prod_{k=1}^n \left(\operatorname{sen}^2 \frac{k\pi}{2n+1} \right) \left(1 - \frac{\operatorname{sen}^2 \theta}{\operatorname{sen}^2 \frac{k\pi}{2n+1}} \right) \\ &= \left((-1)^n \alpha \prod_{k=1}^n \operatorname{sen}^2 \frac{k\pi}{2n+1} \right) \operatorname{sen} \theta \prod_{k=1}^n \left(1 - \frac{\operatorname{sen}^2 \theta}{\operatorname{sen}^2 \frac{k\pi}{2n+1}} \right) \end{aligned}$$

Si llamamos β al primer factor entre paréntesis, su valor lo calculamos haciendo el límite siguiente

$$\begin{aligned} 1 &= \lim_{\theta \rightarrow 0} \frac{\text{sen}(2n+1)\theta}{(2n+1)\theta} = \lim_{\theta \rightarrow 0} \left[\frac{\beta}{2n+1} \frac{\text{sen } \theta}{\theta} \prod_{k=1}^n \left(1 - \frac{\text{sen}^2 \theta}{\text{sen}^2 \frac{k\pi}{2n+1}} \right) \right] \\ &= \frac{\beta}{2n+1} \cdot \lim_{\theta \rightarrow 0} \prod_{k=1}^n \left(1 - \frac{\text{sen}^2 \theta}{\text{sen}^2 \frac{k\pi}{2n+1}} \right) = \frac{\beta}{2n+1} \end{aligned}$$

Así

$$\text{sen}(2n+1)\theta = (2n+1) \text{sen } \theta \prod_{k=1}^n \left(1 - \frac{\text{sen}^2 \theta}{\text{sen}^2 \frac{k\pi}{2n+1}} \right)$$

Supongamos ahora que $x > 0$, y sean m y n dos enteros tales que $x < m < n$. Tomando $\theta = \frac{x}{2n+1}$ en la expresión de arriba, tenemos que

$$\frac{\text{sen } x}{(2n+1) \text{sen} \frac{x}{2n+1}} = \prod_{k=1}^n \left(1 - \frac{\text{sen}^2 \frac{x}{2n+1}}{\text{sen}^2 \frac{k\pi}{2n+1}} \right)$$

Si denotamos por a_k al término k -ésimo del producto superior, como $2\theta/\pi < \text{sen } \theta < \theta$ cuando $0 < \theta < \pi/2$, entonces

$$0 < 1 - \frac{x^2}{4k^2} < a_k < 1 \quad \text{si } m < k \leq n$$

En efecto, la siguiente cadena de implicaciones muestra que $0 < a_k < 1$

$$\begin{aligned} 0 < x < m < k \leq n &\Rightarrow 0 < x < k\pi \leq n\pi \Rightarrow \\ &\Rightarrow 0 < \frac{x}{2n+1} < \frac{k\pi}{2n+1} \leq \frac{n\pi}{2n+1} < \frac{\pi}{2} \\ &\Rightarrow 0 < \frac{\text{sen} \frac{x}{2n+1}}{\text{sen} \frac{k\pi}{2n+1}} < 1 \Rightarrow 0 < \frac{\text{sen}^2 \frac{x}{2n+1}}{\text{sen}^2 \frac{k\pi}{2n+1}} < 1 \end{aligned}$$

Para ver que $0 < 1 - \frac{x^2}{4k^2} < a_k$, como

$$0 < \text{sen} \frac{x}{2n+1} < \frac{x}{2n+1} \Rightarrow 0 < \text{sen}^2 \frac{x}{2n+1} < \frac{x^2}{(2n+1)^2}$$

y

$$\text{sen} \frac{k\pi}{2n+1} > \frac{2k\pi}{\pi(2n+1)} = \frac{2k}{2n+1} \Rightarrow \text{sen}^2 \frac{k\pi}{2n+1} > \frac{4k^2}{(2n+1)^2}$$

entonces

$$0 < \frac{\text{sen}^2 \frac{x}{2n+1}}{\text{sen}^2 \frac{k\pi}{2n+1}} < \frac{\frac{x^2}{(2n+1)^2}}{\frac{4k^2}{(2n+1)^2}} = \frac{x^2}{4k^2} < 1$$

que era lo que buscábamos. Continuando con la demostración, afirmamos que

$$1 > a_{m+1} \cdot a_{m+2} \cdots a_n > \prod_{k=m+1}^n \left(1 - \frac{x^2}{4k^2} \right) > 1 - \frac{x^2}{4} \sum_{k=m+1}^n \frac{1}{k^2} > 1 - \frac{x^2}{4m}$$

donde la penúltima desigualdad estricta se prueba fácilmente por inducción y la última por métodos de integración. En cuanto a la primera, como si $0 \leq b_k < 1$ para todo $k \in \{1, \dots, n\}$ entonces $\prod_{k=1}^n (1 - b_k) \geq 1 - \sum_{k=1}^n b_k$, aplicamos ésto con $b_k = x^2/(4k^2)$ para todo $k \in \{m+1, \dots, n\}$. Para ver la segunda desigualdad consideremos la función $f(x) := 1/x^2$ definida en el intervalo $[1, \infty)$ y razonemos de la siguiente manera. Para cada $k \in \{m+1, \dots, n\}$, el número $1/k^2 > 0$ coincide exactamente con el área del rectángulo de base el segmento de \mathbb{R}^2 : $[k-1, k] \times \{0\}$, y altura $f(k)$. Así, si k recorre el conjunto $\{m+1, \dots, n\}$, puesto que los rectángulos asociados a cada k son consecutivos y subyacentes a la gráfica de la función f en el intervalo $[m, \infty)$, se tiene que

$$\sum_{k=m+1}^n \frac{1}{k^2} < \int_m^n f(x) dx = -\frac{1}{x} \Big|_m^n = -\frac{1}{n} + \frac{1}{m} < \frac{1}{m}$$

lo cual era justamente lo que se estaba buscando. Así,

$$\frac{\operatorname{sen} x}{(2n+1) \operatorname{sen} \frac{x}{2n+1}} = \prod_{k=1}^n a_k = \left(\prod_{k=1}^m a_k \right) \left(\prod_{k=m+1}^n a_k \right)$$

pertenece al intervalo

$$\left[\left(1 - \frac{x^2}{4m} \right) \prod_{k=1}^m \left(1 - \frac{\operatorname{sen}^2 \frac{x}{2n+1}}{\operatorname{sen}^2 \frac{k\pi}{2n+1}} \right), \prod_{k=1}^m \left(1 - \frac{\operatorname{sen}^2 \frac{x}{2n+1}}{\operatorname{sen}^2 \frac{k\pi}{2n+1}} \right) \right]$$

y haciendo $n \rightarrow \infty$ y teniendo en cuenta que

$$\lim_{n \rightarrow \infty} (2n+1) \operatorname{sen} \frac{x}{2n+1} = \lim_{n \rightarrow \infty} x \cdot \frac{\operatorname{sen} \frac{x}{2n+1}}{\frac{x}{2n+1}} = x$$

y que

$$\lim_{n \rightarrow \infty} \frac{\operatorname{sen} \frac{x}{2n+1}}{\operatorname{sen} \frac{k\pi}{2n+1}} = \lim_{n \rightarrow \infty} \frac{x}{k\pi} \cdot \frac{\left(\operatorname{sen} \frac{x}{2n+1} \right) / \left(\frac{x}{2n+1} \right)}{\left(\operatorname{sen} \frac{k\pi}{2n+1} \right) / \left(\frac{k\pi}{2n+1} \right)} = \frac{x}{k\pi}$$

nos queda que

$$\frac{\operatorname{sen} x}{x} \in \left[\left(1 - \frac{x^2}{4m} \right) \prod_{k=1}^m \left(1 - \frac{x^2}{k^2 \pi^2} \right), \prod_{k=1}^m \left(1 - \frac{x^2}{k^2 \pi^2} \right) \right]$$

y haciendo tender m a infinito, finalmente obtenemos que

$$\prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2 \pi^2} \right) \leq \frac{\operatorname{sen} x}{x} \leq \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2 \pi^2} \right) \quad \square$$

Teorema 15 (Relación trigonométrica de Γ).

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\operatorname{sen} \pi x}, \quad \forall x \in (0, 1)$$

Demostración. Tenemos que

$$\Gamma(x)\Gamma(1-x) = \lim_{n \rightarrow \infty} \frac{n!n^x n!n^{1-x}}{(n+x)(n-1+x) \cdots x(n+1-x)(n-x) \cdots (1-x)} = (\star)$$

Pero démonos cuenta que simplemente reordenando

$$\begin{aligned} (n+x)(n-1+x) \cdots x \cdot (n+1-x) \cdots (1-x) &= \\ &= (n+x)(n-x)(n-1+x)(n-1-x) \cdots (1+x)(1-x) \cdot x \cdot (n+1-x) \\ &= \left(\prod_{k=1}^n (k^2 - x^2) \right) \cdot x \cdot (n+1-x) = x \cdot (n+1-x) \prod_{k=1}^n \left(k^2 \left(1 - \frac{x^2}{k^2} \right) \right) \\ &= (n+1-x)(n!)^2 x \prod_{k=1}^n \left(1 - \frac{x^2}{k^2} \right) \end{aligned}$$

Y como $\lim_{n \rightarrow \infty} n/(n+1-x) = 1$, obtenemos

$$(\star) = \lim_{n \rightarrow \infty} \frac{n}{(n+1-x)x \prod_{k=1}^n \left(1 - \frac{x^2}{k^2} \right)} = \frac{1}{x \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2} \right)} = \frac{\pi}{\operatorname{sen} \pi x} \quad \square$$

Corolario 3.2.5. *La función Γ en $1/2$ toma el valor $\sqrt{\pi}$.*

Nótese que para llegar a éste corolario no hubiera sido necesario utilizar los teoremas anteriores. Simplemente, haciendo uso de las herramientas del cálculo integral básico, defino la función real φ en todo \mathbb{R} como $\varphi(x) = e^{-x^2/2}$. Nuestro objetivo es calcular la integral $I = \int_{-\infty}^{\infty} \varphi(x) dx$, que existe, pues $0 < e^{-x^2/2} < e^{-|x|+1}$ si $-\infty < x < \infty$ y $\int_{-\infty}^{\infty} e^{-|x|+1} dx = 2e$. Hacemos un cambio de variable a coordenadas polares, entonces

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} \varphi(x) dx \cdot \int_{-\infty}^{\infty} \varphi(y) dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)} dx dy \\ &= \int_0^{2\pi} \int_0^{\infty} e^{-\frac{1}{2}r^2} r d\theta dr = \int_0^{2\pi} \left[e^{-\frac{1}{2}r^2} \right]_0^{\infty} d\theta = \int_0^{2\pi} 1 d\theta = 2\pi \end{aligned}$$

luego $I = \sqrt{2\pi}$. Ahora bien, como $\Gamma(1/2) = \int_0^{\infty} t^{-1/2} e^{-t} dt$, si hacemos el cambio de variable $t = u^2/2$, $dt = u du$ nos queda que

$$\Gamma(1/2) = \int_0^{\infty} \frac{u^{-1}}{2^{-1/2}} \cdot e^{-u^2/2} u du = \sqrt{2} \cdot \frac{I}{2} = \sqrt{\pi}$$

Hemos demostrado así, de forma directa, sin pasar por el teorema de Euler, la igualdad superior. Una variante de la misma es

$$\frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-t^2/2} dt = 1$$

Teorema 16 (Fórmula de duplicación de Gauss-Legendre).

Sea $x > 0$, entonces se tiene que

$$\Gamma\left(\frac{x}{2}\right) \cdot \Gamma\left(\frac{x+1}{2}\right) = \frac{\sqrt{\pi}}{2^{x-1}} \cdot \Gamma(x)$$

Demostración. Simplemente, como la función f definida a partir de la fórmula del enunciado verifica las hipótesis del teorema (13), el mismo nos asegura que $f = \Gamma$. \square

El siguiente paso que daremos será el de probar la *fórmula de Stirling*, la cual tiene importantes aplicaciones en la teoría analítica de números. Pues, para $n \in \mathbb{N}$ grande, permite aproximar el natural $n!$ y puesto que $\Gamma(n+1) = n!$, también permitirá aproximar la función Gamma en estos valores. Después del resultado aparecerá un corolario que mostrará de manera explícita la importancia de la fórmula de Stirling en relación con la función Γ . Necesitaremos de una proposición previa para facilitar la demostración de la misma.

Proposición 3.2.6. *La sucesión $(a_n)_{n \in \mathbb{N}}$, donde $a_n = \log n! - (n + \frac{1}{2}) \log n + n$, converge.*

Demostración. La convergencia de (a_n) estará garantizada si vemos que es decreciente y que está acotada inferiormente. Sea $n \in \mathbb{N}$, entonces

$$\begin{aligned} a_n - a_{n+1} &= \log \frac{n!}{(n+1)!} - (n+1/2) \log n + (n+1+1/2) \log(n+1) - 1 \\ &= (n+1/2)(\log(n+1) - \log n) - 1 \\ &= (n+1/2) \log \left(1 + \frac{1}{n}\right) - 1 \end{aligned}$$

Ahora bien, utilizando la desigualdad de Hermite-Hadamard, proposición (2.3.1), aplicada a la función $1/x$ sobre el intervalo $[n, n+1]$, obtenemos que $(n + \frac{1}{2})^{-1} \leq \int_n^{n+1} 1/x \, dx = \log(1 + \frac{1}{n})$. Luego

$$a_n - a_{n+1} = (n+1/2) \log \left(1 + \frac{1}{n}\right) - 1 \geq (n+1/2) \left(n + \frac{1}{2}\right)^{-1} - 1 = 0$$

y así la sucesión (a_n) es decreciente. Veamos ahora que es acotada inferiormente por $1/2$. Usamos la siguiente desigualdad, obtenida a partir de la proposición (2.3.1), si $g(x) := \log(x)$

$$g\left(\frac{r-1/2+r+1/2}{2}\right) = g(r) = \log(r) \geq \int_{r-1/2}^{r+1/2} g(x) \, dx$$

Tenemos que

$$\begin{aligned} \int_1^n \log x \, dx &= \int_1^{1+1/2} \log x \, dx + \int_{1+1/2}^{2+1/2} \log x \, dx + \cdots + \int_{n-1+1/2}^n \log x \, dx \\ &\leq \frac{1}{2} \log(3/2) + \log(2) + \cdots + \log(n-1) + \frac{1}{2} \log(n) \\ &< \frac{1}{2} + \log(n!) - \frac{1}{2} \log(n) \end{aligned}$$

Ahora bien, como $\int_1^n \log x \, dx = n \log n - n + 1$, entonces

$$a_n = \log n! - (n+1/2) \log n + n > \int_1^n \log x \, dx - \frac{1}{2} - n \log n + n = 1/2$$

la arbitrariedad de n nos garantiza que (a_n) está acotada inferiormente por $1/2$. \square

Teorema 17 (Fórmula de Stirling).

$$\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi} \cdot n^{n+1/2} \cdot e^{-n}} = 1$$

Demostración. Antes de nada, en esta demostración todas las convergencias que aparecen son cuando $n \rightarrow \infty$. Utilizando la notación de la proposición anterior, dado $n \in \mathbb{N}$ defino el término n -ésimo de la sucesión $(b_n)_{n \in \mathbb{N}}$ como

$$b_n := e^{a_n} = e^{\log n! - (n+\frac{1}{2}) \log n + n} = n! \cdot n^{-(n+\frac{1}{2})} \cdot e^n = \frac{n!}{n^{n+\frac{1}{2}} \cdot e^{-n}}$$

Ahora bien, debido a la definición de b_n , si digamos que $a_n \rightarrow a_0$ entonces, $b_n \rightarrow b = e^{a_0} > 0$. Luego,

$$\frac{b_n^2}{b_{2n}} = \frac{(n!)^2 \cdot [e^{-2n} \cdot n^{2n+1}]^{-1}}{(2n)! \cdot [(2n)^{2n+\frac{1}{2}} \cdot e^{-2n}]^{-1}} = \frac{(n!)^2 \cdot 2^{2n+\frac{1}{2}}}{(2n)! \cdot n^{1/2}} \rightarrow \frac{b^2}{b} = b$$

Esta conclusión la utilizaremos más adelante. Ahora definamos una nueva sucesión, $(c_n)_{n \in \mathbb{N}}$ que nos ayudará a encontrar el valor real de b . Dado $n \in \mathbb{N}$

$$c_n := \frac{n! \cdot n^{1/2}}{(n + \frac{1}{2}) \cdot \dots \cdot \frac{3}{2} \cdot \frac{1}{2}}$$

Luego, $\lim_{n \rightarrow \infty} c_n = \Gamma(1/2) = \sqrt{\pi}$. Así, utilizando lo que anunciábamos antes

$$\frac{b_n^2}{b_{2n}} = c_n \left(1 + \frac{1}{2n}\right) \sqrt{2} \rightarrow \sqrt{\pi} \sqrt{2} = \sqrt{2\pi} \quad (*)$$

Por lo que $b = \sqrt{2\pi}$ ya que habíamos probado que $b_n^2/b_{2n} \rightarrow b > 0$. Consecuentemente

$$\frac{b_n}{\sqrt{2\pi}} = \frac{n!}{\sqrt{2\pi} n^{n+\frac{1}{2}} \cdot e^{-n}} \rightarrow \frac{b}{2\pi} = 1$$

Para completar la demostración deberemos probar la igualdad no obvia en la expresión (*). Para ello démonos cuenta de lo siguiente. Escrito de manera expandida para una mejor comprensión

$$\begin{aligned} \frac{(2n)!}{2^{2n}} &= \frac{1}{2} \cdot \frac{2}{2} \cdot \frac{3}{2} \cdot \frac{4}{2} \cdot \frac{5}{2} \cdot \dots \cdot \frac{n}{2} \cdot \frac{n+1}{2} \cdot \frac{n+2}{2} \cdot \dots \cdot \frac{2n-2}{2} \cdot \frac{2n-1}{2} \cdot \frac{2n}{2} \\ &= \left(\frac{1}{2} \cdot \frac{3}{2} \cdot \frac{5}{2} \cdot \dots \cdot \left(n - \frac{1}{2}\right) \right) \cdot (1 \cdot 2 \cdot 3 \cdot \dots \cdot n) \end{aligned}$$

pues lo único que hemos hecho ha sido reordenar los factores, agrupando primero los que ocupaban una posición impar en la expresión expandida de $\frac{(2n)!}{2^{2n}}$ y en segundo lugar los que ocupaban una posición par. Nótese que también hemos simplificado las expresiones fraccionarias. Teniendo ésto en mente es más fácil comprobar la igualdad deseada. Simplemente

$$\begin{aligned} \frac{b_n^2}{b_{2n}} &= \frac{(n!)^2 \cdot 2^{2n+\frac{1}{2}}}{(2n)! \cdot n^{1/2}} = \frac{n! \cdot n^{1/2} \cdot 2^{2n}}{(2n)!} \cdot \frac{n!}{n} \cdot \sqrt{2} \\ &= \frac{n! \cdot n^{1/2}}{\frac{1}{2} \cdot \frac{3}{2} \cdot \dots \cdot \left(n - \frac{1}{2}\right)} \cdot \frac{n!}{n \cdot 1 \cdot 2 \cdot \dots \cdot n} \cdot \sqrt{2} = c_n \left(1 + \frac{1}{2n}\right) \cdot \sqrt{2} \quad \square \end{aligned}$$

3.2.2. La función Beta

Definición 3.2.2 (La función Beta). Es una función real de 2 variables que viene definida a partir de la expresión

$$\beta(x, y) := \int_0^1 t^{x-1}(1-t)^{y-1} dt$$

para cualesquiera $x, y > 0$. Nótese que β sólo está definida en el primer cuadrante de \mathbb{R}^2 (sin incluir los ejes)³.

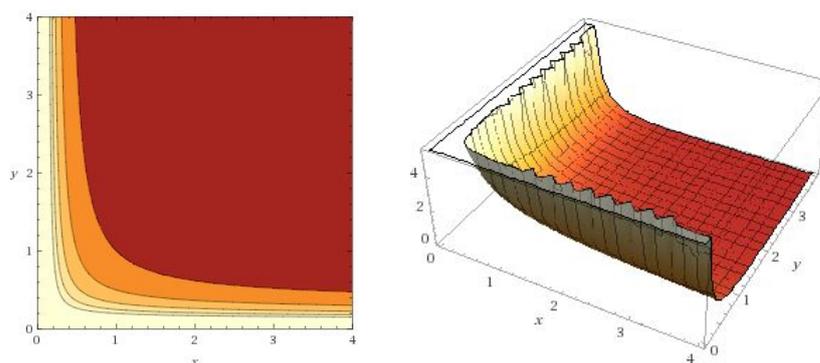


Figura 3.2: Función β con vista cenital

Lo primero que haremos será comprobar que la función Beta *está bien definida*. Para ello separamos el dominio de integración en dos obteniendo dos nuevas integrales. Por el criterio de comparación por paso al límite, dichas integrales convergen para $x, y > 0$ ya que las integrales $\int_0^{1/2} t^{x-1} dt$ y $\int_{1/2}^1 (1-t)^{y-1} dt$ hacen lo propio.

El siguiente resultado reúne las propiedades más relevantes acerca de la función Beta, y su relación con Γ .

Teorema 18 (Propiedades de β).

La función Beta tiene las siguientes propiedades:

1. β es simétrica; es decir, $\beta(x, y) = \beta(y, x) \quad \forall x, y > 0$
2. $\beta(x+1, y) = \frac{x}{x+y} \beta(x, y) \quad \forall x, y > 0$
3. Fijado $y > 0$, $\beta(\cdot, y)$ es log-convexa
4. Relación con Γ ,

$$\beta(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \quad \forall x, y > 0$$

³ Al igual que las gráficas de la función gamma, las de β también están hechas *online* con *WolframAlpha*[®]: *Computational Knowledge Engine*.

Demostración. Para la primera, simplemente hacemos el cambio de variable $u = 1-t$, $du = -dt$. La segunda se obtiene a partir de la definición e integrando por partes; mientras que la tercera sólo hay que aplicar la desigualdad de Rogers-Hölder, ver teorema (36). Hagamos con algo más de detalle la cuarta.

Fijemos un $y > 0$ y definamos la función

$$\psi_y(x) := \frac{\Gamma(x+y)\beta(x,y)}{\Gamma(y)} \quad \forall x > 0$$

Notesé que ψ_y está bien definida y que $\Gamma(y) \in \mathbb{R}^+$ es un valor fijo. Así, puesto que ψ_y es producto de funciones logarítmicamente convexas, ella también lo será. En efecto, de manera general, si tenemos dos funciones log-convexas, $u, v : I \rightarrow \mathbb{R}$, y dos puntos $x, y \in I$, y tomamos dos escalares $\lambda, \mu \geq 0$ tales que su suma valga uno

$$\begin{aligned} (uv)(\lambda x + \mu y) &= u(\lambda x + \mu y) \cdot v(\lambda x + \mu y) \leq u(x)^\lambda u(y)^\mu v(x)^\lambda v(y)^\mu \\ &= (uv)^\lambda(x) \cdot (uv)^\mu(y) \end{aligned}$$

Continuando con la demostración, observemos que por las propiedades ya probadas de las funciones Gamma y Beta

$$\psi_y(x+1) = \frac{\Gamma(x+y+1)\beta(x+1,y)}{\Gamma(y)} = \frac{(x+y)\Gamma(x+y)x\beta(x,y)}{\Gamma(y)(x+y)} = x \cdot \psi_y(x)$$

Además, también se tiene que

$$\psi_y(1) = \frac{\Gamma(1+y)\beta(1,y)}{\Gamma(y)} = \frac{y\Gamma(y)\beta(1,y)}{\Gamma(y)} y \int_0^1 t^0(1-t)^{y-1} dt = 1$$

Consecuentemente, ψ_y tiene las tres propiedades que caracterizan a la función Gamma y que sólo ella tiene, luego por el teorema (13) concluimos que $\psi_y(x) = \Gamma(x)$ para todo $x > 0$. Así $\psi_y(x) = \Gamma(x) \quad \forall x > 0$ Como el $y > 0$ fijado es arbitrario, deducimos (4). \square

3.2.3. Algunas otras importantes expresiones que involucran a las funciones Gamma y Beta

A partir de los resultados anteriores podemos obtener las siguientes expresiones.

PRIMERA. Valor que toma Γ a medio camino entre un natural y el siguiente

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)! \sqrt{\pi}}{n! 4^n} \quad \forall n \in \mathbb{N}$$

Para obtener este resultado simplemente aplicamos la fórmula de duplicación de Gauss-Legendre

SEGUNDA. Expresión de la función β en términos de sen y cos:

$$\beta(x,y) = 2 \int_0^{\pi/2} \text{sen}^{2x-1} t \cdot \text{cos}^{2y-1} t dt \quad \forall x, y > 0$$

Simplemente hacemos un cambio de variable en la integral dada por la definición de la propia β , $t = \text{sen}^2 u$, $dt = 2 \text{sen } u \cos u \, du$.

TERCERA. Se tiene la siguiente igualdad para el cálculo de integrales Riemman

$$\int_0^{\pi/2} \text{sen}^{2n} t \, dt = \frac{(2n)! \sqrt{\pi}}{2^{2n+1} (n!)^2}$$

CUARTA (Fórmula de Euler para integración). Si $0 < x < 1$, entonces ocurre que

$$\int_0^{\infty} \frac{t^{x-1}}{1+t} \, dt = \frac{\pi}{\text{sen}(\pi x)}$$

QUINTA (Fórmula de Weierstrass para la función Γ). Si $\gamma_n = 1 + 1/2 + \dots + 1/n - \log n$, llamamos constante de Euler a $\gamma := \lim_{n \rightarrow \infty} \gamma_n \approx 0,5772\dots$. Entonces, se tiene la siguiente expresión de la función Gamma

$$\Gamma(x) = \frac{e^{-\gamma x}}{x} \prod_{n=1}^{\infty} \left(1 + \frac{x}{n}\right)^{-1} e^{x/n}$$

para todo x estrictamente positivo.

SEXTA. Para todo $x > 0$ se tiene que

$$\frac{d^2}{dx^2} (\log(\Gamma(x))) = \sum_{n=0}^{\infty} \frac{1}{(x+n)^2}$$

Como consecuencia de la forma de expresar Γ como producto infinito, podemos calcular fácilmente la función $\log \circ \Gamma$

$$\log(\Gamma(x)) = -\gamma x - \log x + \sum_{n=1}^{\infty} \left[\frac{x}{n} - \log \left(1 + \frac{x}{n}\right) \right]$$

y derivando

$$\begin{aligned} \frac{d}{dx} (\log(\Gamma(x))) &= \frac{\Gamma'(x)}{\Gamma(x)} = -\gamma - \frac{1}{x} + \sum_{n=1}^{\infty} \left[\frac{1}{n} - \frac{1}{n} \cdot \frac{1}{1+x/n} \right] \\ &= -\gamma - \frac{1}{x} + \sum_{n=1}^{\infty} \frac{x}{n(n+x)} \end{aligned}$$

Además, también hemos obtenido otra fórmula para la derivada de Gamma

$$\Gamma'(x) = \Gamma(x) \left[-\gamma - \frac{1}{x} + \sum_{n=1}^{\infty} \frac{x}{n(n+x)} \right] \quad (x > 0)$$

Estamos ya en condiciones de abordar la sexta cuestión, que sale de manera inmediata gracias a la deducción que acabamos de hacer.

SÉPTIMA. Para cada $x > 0$ se tiene que

$$\Gamma(x/3)\Gamma((x+1)/3)\Gamma((x+2)/3) = 2\pi\sqrt{3} \Gamma(x) 3^{-x}$$

Capítulo 4

Funciones convexas en varias dimensiones

Una vez visto en profundidad el caso de una variable, pasamos ahora a tratar las funciones convexas en varias variables reales, para lo cual necesitamos primero recordar algunas cosas de los conjuntos convexas, pues es allí donde están definidas estas funciones.

4.1. Conjuntos convexas

En lo que sigue E^1 será un espacio lineal real mientras no se diga lo contrario. Un conjunto $C \subseteq E$ es *convexo* si para cada dos puntos $x, y \in C$, el conjunto C contiene al segmento lineal convexo²

$$[x, y] := \{(1 - \lambda)x + \lambda y : \lambda \in [0, 1]\}$$

Dados $A, B \subseteq E$ y $\lambda, \mu \in \mathbb{R}$, se define la *suma lineal de conjuntos* como $\lambda A + \mu B := \{\lambda x + \mu y : x \in A, y \in B\}$. Esta suma es convexa si A y B lo son y los escalares λ, μ son no negativos. Un conjunto A se dice *afín* si $(1 - \lambda)x + \lambda y \in A$ para cualesquiera $x, y \in A$ y $\lambda \in \mathbb{R}$. Dados $x_1, \dots, x_n \in E$, una *combinación afín* de ellos es

$$x = \sum_{i=1}^n \lambda_i x_i \quad \text{con} \quad \lambda_i \in \mathbb{R} \quad \text{y} \quad \sum_{i=1}^n \lambda_i = 1$$

Se dice que la combinación es *convexa* si $\lambda_i \geq 0 \quad \forall i$.

Sea ahora $A \subseteq E$, se define la *envolvente convexa* de A , $co(A)$, como la intersección de todos los subconjuntos convexas de E que contienen a A . Así, $co(A)$ es convexo y es el menor convexo que contiene a A . Se tiene que $co(A)$ es el conjunto de todas las combinaciones convexas de elementos de A . La variación afín de esta construcción se llama *envolvente afín*, $aff(A)$. Se define la *dimensión* de un conjunto convexo $A \subseteq E$ como $dim(A) := dim(aff(A))$. La demostración del siguiente teorema que involucra todos estos nuevos conceptos se encuentra en [NP06].

¹ Para la mayoría de las aplicaciones que describimos aquí, E será de dimensión finita, pero las definiciones y primeros resultados pueden establecerse en dimensión cualquiera

² Para ver una teoría más detallada y explicativa acerca de los conjuntos convexas ir a [Web94]

Teorema 19 (Teorema de Carathéodory).

Sea $A \subseteq E$, $\dim(\text{co}(A)) = \dim(\text{aff}(\text{co}(A))) = m$. Entonces cada punto $x \in \text{co}(A)$ es la combinación convexa de a lo más $m + 1$ puntos de A .

Se llama *polítopos*³ a cualquier conjunto de la forma $C = \text{co}(\{x_0, \dots, x_n\})$. Si los puntos $x_1 - x_0, \dots, x_n - x_0$ son independientes entonces C es un n -símplice con vértices los puntos x_0, \dots, x_n . Se tiene también que $\dim(C) = n$. Para cualquier punto x de un n -símplice existen $\lambda_i \geq 0$ (únicos) tales que $x = \sum \lambda_i x_i$ es combinación convexa y dichos λ_i se denominan *coordenadas baricéntricas* de x .

4.1.1. Hiperplanos (soporte) y teoremas de separación

Los hiperplanos⁴ son la extrapolación a dimensión superior del concepto de línea en \mathbb{R}^2 y de plano en \mathbb{R}^3 . Son usados para separar el espacio en dos partes llamadas (*sub*)espacios mitad. Un *hiperplano* de E es un conjunto de la forma

$$H := \{x \in E : h(x) = \alpha\}$$

con $h : E \rightarrow \mathbb{R}$ un funcional lineal no nulo y $\alpha \in \mathbb{R}$. A los conjuntos $\{x \in E : h(x) \geq \alpha\}$ y $\{x \in E : h(x) \leq \alpha\}$ se les llama *semi-espacios* asociados (determinados) por H . Son adjetivados de abiertos ó cerrados según sea la desigualdad estricta ó no. Si $h \in E^*$, es decir, h es continua, entonces H y los dos semi-espacios determinados por él son conjuntos cerrados de E . Decimos simplemente que H es cerrado.

Teorema 20 (Teorema de separación).

Sean $U, V \subseteq E$ convexos contenidos en un espacio lineal normado con $\text{int}(U) \neq \emptyset$, $V \cap \text{int}(U) = \emptyset$. Entonces existe un hiperplano cerrado que separa U y V .

Teorema 21 (Teorema de separación fuerte).

Sean $K, C \subseteq E$ convexos no vacíos con $K \cap C = \emptyset$, K compacto y C cerrado. Entonces existe un hiperplano cerrado que separa K y C estrictamente.

Idea de la prueba del teorema (21) en el caso $\dim(E) < \infty$. Se prueba primero que existen puntos $x \in K$ e $y \in C$ de modo que la distancia de K a C se realiza por la distancia de x a y . Después es fácil ver que un hiperplano que pasa por un punto del segmento $[x, y]$ y es ortogonal a $[x, y]$ separa estrictamente K y C .

³ En el libro *Convexity*, referencia [Web94], se da una buena teoría acerca de los polítopos: características, propiedades, usos en otros ambientes para facilitar la prueba de determinados resultados, volumen de los mismos, *área de su frontera*, etc. De ésta última hablaremos más adelante cuando tratemos la desigualdad de Brunn-Minkovsky

⁴ La teoría de teoremas de separación e hiperplanos soporte aparece en multitud de libros. Aquí se ha seguido la recopilación de resultados que aparece en el apéndice A de [NP06]. También puede verse algo más detallado en [Web94]

Definición 4.1.1 (Hiperplano soporte). Se dice que un hiperplano H *soporta* A , un convexo de E , en un punto $a \in A$ si dicho punto también está en H y A está contenido en uno de los semi-espacios determinados por H ⁵.

Definición 4.1.2 (Punto extremo). Un $z \in U \subseteq E$ convexo es un *punto extremo* si no existen $x, y \in U$ distintos ni $\lambda \in (0, 1)$ tales que $z = (1 - \lambda)x + \lambda y$.

Así, por ejemplo, los puntos extremos de un símlice son sus vértices, y en un disco cerrado como $D_R(0) = \{(x, y) : x^2 + y^2 \leq R^2\}$ lo son los puntos de $\partial D_R(0)$. El semiplano $y \geq 0$ de \mathbb{R}^2 no tiene puntos extremos.

Teorema 22 (Minkovsky).

Cada $K \subseteq \mathbb{R}^n$ no vacío, convexo y compacto es la envolvente convexa de sus puntos extremos.

Demostración. Por inducción sobre la dimensión m de K . Si $m = 0$ ó $m = 1$; esto es, cuando K es un punto o un segmento cerrado la afirmación es obvia. Asumimos cierto el resultado para $m \leq n - 1$ y tomamos K de dimensión $m + 1$ embebido en un subespacio lineal E de dimensión $m + 1$.

Tenemos dos casos, si $z \in \partial K$, entonces existe un $H \subseteq E$ hiperplano soporte para K en z . Así, como $K \cap H$ es compacto, convexo y de dimensión menor o igual que m , por la hipótesis de inducción, z es una combinación convexa de puntos extremos de $K \cap H$. Ahora bien, cualquier punto extremo $e \in K \cap H$ es también un punto extremo de K . En efecto, sea

$$H = \{t \in E : \varphi(t) = \alpha\}$$

podemos suponer que $K \subseteq \{\varphi \leq \alpha\}$. Si $e = (1 - \lambda)x + \lambda y$, $x \neq y$ en K , $\lambda \in (0, 1)$ entonces $\varphi(x) = \varphi(y) = \alpha$. Esto es que $x, y \in K \cap H$, en contradicción con la elección de e .

Si $z \in \text{int}(K)$, cada línea a través de z intersecta a K en un segmento cuyos puntos finales están en ∂K . Consecuentemente, z es una combinación convexa de puntos frontera que son combinaciones convexas de puntos extremos, como acabamos de ver. \square

Observación. Versión más fina del teorema (22). *Cada punto en un compacto convexo de \mathbb{R}^n es la combinación convexa de a lo más $n + 1$ puntos extremos⁶. Su demostración puede hacerse usando el teorema (19).*

4.2. Funciones convexas en dimensión superior

Sea en lo que sigue $U \subseteq E$ un convexo de un espacio lineal real E .

Definición 4.2.1 (Función convexa). Se dice que $f : U \rightarrow \mathbb{R}$ es *convexa* si

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y) \quad \forall x, y \in U \quad \forall \lambda \in [0, 1]$$

⁵ Usaremos el siguiente resultado explícitamente más adelante: si A es un conjunto convexo y x es un punto de la frontera de A entonces, por el teorema (20), existe un hiperplano soporte de A en x

⁶ Ver su enunciado completo y demostración en [RW98], teorema 2.29, pág. 55

Podemos reducir la convexidad en varias variables a ver la convexidad en un sola variable, así una tal f como la de arriba sería convexa sii para todo $x, y \in U$, la función $\varphi(t) := f((1-t)x + ty)$ en $[0, 1]$ también lo es. Además nótese que la convexidad aquí definida es más fuerte que ser convexa en cada variable por separado: la función $f(x, y) = xy$ en \mathbb{R}^2 es convexa cuando fijamos un valor cualquiera a una de las variables (es una función afín), sin embargo no es convexa en \mathbb{R}^2 como se comprueba al restringirla a la recta $y = -x$. Ejemplos de funciones estrictamente convexas serían

$$f(x_1, \dots, x_n) = \sum_{i=1}^n \varphi(x_i) \quad g(x_1, \dots, x_n) = \sum_{i < j} c_{ij} (x_i - x_j)^2$$

con φ una función estrictamente convexa de \mathbb{R} y $c_{ij} > 0$.

Definición 4.2.2 (Epigrafo). Se denomina *epigrafo* de una función $f : U \rightarrow \mathbb{R}$ al conjunto $\text{epi}(f) = \{(x, y) : x \in U, y \in \mathbb{R}, f(x) \leq y\}$ ⁷

Se tiene la caracterización de que f es convexa sii $\text{epi}(f) \subseteq E \times \mathbb{R}$ es un conjunto convexo. Así podemos subordinar el estudio de las funciones convexas al de los conjuntos convexas. Si ahora U es un abierto convexo del espacio lineal normado E y f es continua y convexa, entonces $\text{epi}(f)$ tiene interior no vacío y cada $(a, f(a))$ es un punto frontera para $\text{epi}(f)$. Luego, $\exists H \subseteq E \times \mathbb{R}$ hiperplano cerrado que contiene a $(a, f(a))$ y $\text{epi}(f)$ cae en uno de los subespacios mitad determinados por H . Se dice en este caso que H es un *hiperplano soporte* de f en a .

Como los hiperplanos cerrados H están asociados a funcionales lineales continuos no nulos en $E \times \mathbb{R}$ y $(E \times \mathbb{R})^*$ está constituido por los pares (h, λ) de funcionales continuos y lineales de E y números reales, entonces un hiperplano soporte de f en a está determinado por un par (h, λ) y un $\alpha \in \mathbb{R}$ tales que

$$h(a) + \lambda f(a) = \alpha, \quad h(x) + \lambda y \geq \alpha \quad \forall y \geq f(x) \quad \forall x \in U$$

en realidad, $\lambda > 0$ ⁸ y así de la existencia de hiperplano soporte concluimos que $\exists h$ funcional lineal continuo tal que

$$f(x) \geq f(a) + h(x - a) \quad \forall x \in U$$

al que llamamos *soporte* de f en a .

En lo que sigue, vamos a necesitar tener funciones que pueden tomar valor infinito en algún punto del dominio. Hablaremos así de *funciones de valores extendidos* $f : U \rightarrow \mathbb{R}^c := \mathbb{R} \cup \{\infty\}$. Se define para ellas el *dominio*, denotado por $\text{dom}(f)$, como el conjunto de puntos donde f toma valor finito. El *epigrafo* de f , denotado por $\text{epi}(f)$, sería como en la definición (4.2.2), esto es $\text{epi}(f) = \{(x, y) : x \in U, y \geq f(x)\}$.

Definición 4.2.3 (Suma epigráfica). Se define la *suma epigráfica* de dos funciones de valores reales extendidos $f, g : U \rightarrow \mathbb{R}^c$ como la función $h : U \rightarrow \mathbb{R}^c$ tal que

$$U \ni u \mapsto h(u) := \inf \{f(v) + g(v - u) : v \in \mathbb{R}^n\} \in [-\infty, +\infty]$$

⁷ Ver [RW98] para una muy buena exposición sobre epigrafos y envolventes y teoremas de separación

⁸ $\lambda \neq 0$, pues de otra manera se tendría que $h(x) \geq h(a)$ para todo $x \in B_r(a)$, lo que implicaría que $h = 0$. Un momento de reflexión muestra que λ es positivo

Recordamos ahora otra definición común que vamos a necesitar.

Definición 4.2.4 (Función semicontinua inferior). $f : U \rightarrow \mathbb{R} \cup \{\infty\} = \mathbb{R}^c$ se dice *semicontinua inferior* si $f(x) = \liminf_{y \rightarrow x} f(y)$ para todo $x \in \mathbb{R}^n$.

Antes de pasar a hablar del concepto de subdiferencial para funciones de varias variables vamos a probar un resultado importante que hace ver la rigidez que proporciona el concepto de convexidad a una función, el *principio del máximo*.

Teorema 23 (Principio del máximo).

Si f es convexa y tiene un máximo global en $\text{int}(U)$ entonces es constante. Localizando, si f es convexa y tiene un máximo local en $s \in \text{int}(U)$ entonces f es localmente constante en a .

Demostración. Por RA. Supongamos f no constante con un máximo global en $a \in \text{int}(U)$. Tomamos $x \in U$ y $\epsilon \in (0, 1)$ tales que $f(x) < f(a)$ e $y = a + \epsilon(a - x) \in U$. Entonces, $a = \frac{y}{1+\epsilon} + \frac{\epsilon x}{1+\epsilon}$, lo que da una contradicción ya que

$$f(a) \leq \frac{f(y)}{1+\epsilon} + \frac{\epsilon f(x)}{1+\epsilon} < \frac{f(a)}{1+\epsilon} + \frac{\epsilon f(a)}{1+\epsilon} = f(a) \quad \square$$

Terminamos este párrafo con una consecuencia importante que es un principio básico de la teoría de optimización.

Corolario 4.2.1. *Toda función convexa y continua f en un conjunto compacto convexo K alcanza su máximo en un punto extremo.*

Demostración. Sea $a \in K$ un punto donde f alcanza el valor máximo. Siendo a combinación convexa de puntos extremos $e_1, \dots, e_r \in K$ (por el teorema (22)), tenemos $a = \sum \lambda_i e_i$ con $\lambda_i \geq 0$ y $\sum \lambda_i = 1$. Al ser f convexa tenemos que $f(a) \leq \sum \lambda_i f(e_i)$, lo que obliga a que $f(a) = f(e_{i_0})$ para algún e_{i_0} . \square

4.2.1. Propiedades de regularidad

Aquí solo mostramos algunos resultados claves relativos a la regularidad de funciones convexas de dimensión superior⁹. Para funciones suficientemente regulares tenemos, como en el caso de funciones de una variable, criterios sencillos para la convexidad. Más explícitamente.

Proposición 4.2.2 (Test de la derivada para dimensión superior¹⁰). *Sea f una función de clase C^2 definida en un abierto convexo $U \subseteq \mathbb{R}^n$, entonces f es convexa si alguno de los siguientes ocurre*

1. $\langle y - x, \nabla f(y) - \nabla f(x) \rangle \geq 0 \quad \forall x, y \in U$
2. $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \quad \forall x, y \in U$

⁹ Aquí solo damos lo necesario que usaremos en relación a las funciones convexas en dimensión superior. Para ver una teoría más profunda sobre ellas, por ejemplo entre otras cosas, la continuidad y diferenciabilidad en términos de las derivadas parciales; así como otros resultados más sofisticados que involucran conceptos mucho más refinados como la *subdiferencial*, ver [NP06], [Web94] y [RW98]

¹⁰ Ver demostración en [RW98], pág. 47

3. $D^2 f(x)$ es semidefnida positiva para todo $x \in U$

La relación convexidad/continuidad que establecimos para funciones de una variable también pasa a funciones en varias variables, si bien la demostración requiere pasar por varias etapas (básicamente mostrar que una función convexa es localmente acotada y localmente lipschitz). Resumimos en el siguiente enunciado estos resultados.

Teorema 24 (Continuidad de funciones convexas).

Toda función convexa $f : U \rightarrow \mathbb{R}$ en un abierto convexo $U \subset E$ es localmente lipschitz y, por tanto, continua.

Toda función convexa $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ es continua en $\text{int}(\text{dom}(f))$.

La demostración de la primera parte del teorema está hecha en [NP06]. Por otro lado, en la pág. 59 y ss. de [RW98] aparece demostrada la segunda parte como teorema independiente y la primera como corolario de ésta. Siguiendo este método, basta aplicar la segunda parte del teorema a la función g definida como f en U y ∞ en el resto; así, $\text{int}(\text{dom}(g)) = U$. Una demostración alternativa para la parte de funciones finitas, sin usar la generalización a funciones de valores reales extendidos viene detallada en [Cer00], pág. 173.

EJEMPLO. En \mathbb{R}^2 , la función

$$f(x, y) = \begin{cases} \frac{x^2}{2y} & \text{si } y > 0 \\ 0 & \text{si } x = y = 0 \\ \infty & \text{en el resto} \end{cases}$$

es semicontinua inferior y convexa. Además, no es continua en el compacto y convexo $\{(x, y) : x^4 \leq y \leq 1\} \cap \text{dom}(f)$ ¹¹.

La semicontinuidad inferior de f se deduce sin mucho problema analizando por separado el punto $(0, 0)$ y los puntos $(x, 0)$ del dominio donde la función no es continua. Veamos, por otro lado, la convexidad en el semiplano abierto $H = \{(x, y) : y > 0\}$. Usando la proposición (4.2.2), el hessiano sería

$$D^2 f(x, y) = \frac{1}{y} \begin{pmatrix} 1 & -x/y \\ -x/y & x^2/y^2 \end{pmatrix}$$

que es semidefnido positivo, pues el primer menor principal es 1 y $\det(D^2 f(x, y)) = 0$. La convexidad en todo \mathbb{R}^2 (como función extendida) se obtiene a partir de esta propiedad usando un resultado mostrado en [RW98]¹².

4.3. La subdiferencial en dimensión superior

Tratamos ahora la generalización a varias variables del concepto de subdiferencial de una función convexa y recogemos algunas de sus propiedades básicas. Para ver

¹¹ Ir a [RW98], pág. 61, para ver este ejemplo más detallado con algunas propiedades adicionales que aquí no enunciamos

¹² Ver [RW98], teorema 2.35 y ss.

una exposición más extensa remitimos al lector, como en otras ocasiones, a los libros [NP06], [RW98], y en menor medida [Web94].

Sea f una función en un abierto $U \subseteq E$, dado $a \in U$, decir que f tiene un hiperplano soporte en a equivale, por lo que hemos visto en la sección anterior, a que $\exists h : E \rightarrow \mathbb{R}$ funcional lineal y continuo con

$$f(x) \geq f(a) + h(x - a) \quad \forall x \in U$$

El conjunto de los funcionales h de esta forma es la *subdiferencial* de f en a , denotado por $\partial f(a)$. Por lo tanto, haciendo uso de este concepto podemos caracterizar las funciones convexas como sigue.

Proposición 4.3.1. *Supongamos que U es un subconjunto convexo de E y sea $f : U \rightarrow \mathbb{R}$. Entonces, f convexa ssi $\partial f(a) \neq \emptyset \quad \forall a \in U$.*

Ahora bien, si $E = \mathbb{R}^n$ ó es un espacio de Hilbert h puede ser representada de manera única como $h(x) = \langle x, z \rangle$ para algún $z \in E$. Así, la inecuación que define los elementos de la subdiferencial queda como

$$f(x) \geq f(a) + \langle z, x - a \rangle \quad \forall x \in U \quad (*)$$

y la subdiferencial se vería como el conjunto de los z (comúnmente llamados *subgradiantes*) que verifican (*).

Observación. Nótese que si f es de clase C^1 en $a \in U$, la subdiferencial coincide con el concepto usual de gradiente. Esto es que $\partial f(a) = \{\nabla f(a)\}$.

En el caso de funciones no diferenciables, la subdiferencial proporciona una herramienta para poder realizar un cálculo diferencial, sobre todo en el caso de funciones convexas. Por ejemplo, recordemos que las derivadas laterales direccionales de una función f en un punto $a \in U$ según la dirección v se definen como

$$f'_+(a, v) := \lim_{t \rightarrow 0^+} \frac{f(a + tv) - f(a)}{t} \quad f'_-(a, v) := \lim_{t \rightarrow 0^-} \frac{f(a + tv) - f(a)}{t}$$

Cuando f es convexa, por el teorema de Stolz para una variable, la derivada direccional por la derecha es mayor ó igual que la de por la izquierda. Además, tenemos fácilmente de la definición, $z \in \partial f(a)$ ssi $f'_+(a, v) \geq \langle z, v \rangle \quad \forall v \in \mathbb{R}^n$.

Queremos analizar también algunos aspectos del cálculo diferencial de funciones extendidas convexas, no necesariamente regulares. Fijemos entonces ahora, $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ semicontinua inferior y convexa. La subdiferencial de f se define igual (como el conjunto de subgradiantes) para puntos $a \in \text{dom}(f)$ y para un punto $a \notin \text{dom}(f)$ como $\partial f(a) = \emptyset$. Primero, vemos que la fórmula (*) puede también enunciarse como sigue.

Proposición 4.3.2. *Siendo $a \in \text{dom}(f)$, se tiene que a es un mínimo global de la función f si y sólo si $0 \in \partial f(a)$.*

También, el siguiente resultado nos permite ver la subdiferencial de una función tal y como se usa de manera usual dicho concepto en Análisis Convexo.

Proposición 4.3.3. *Sea f semicontinua inferior y convexa, entonces el grafo de la subdiferencial es cerrado y*

$$\partial f(x) = \{x^* \in \mathbb{R}^n : f(\cdot) - \langle x^*, \cdot \rangle \text{ tiene un mínimo global en } x\} \quad \forall x \in U$$

Vamos a entender la subdiferencial de f , como la aplicación punto a conjunto ∂f que asigna a cada $a \in \mathbb{R}^n$ el conjunto $\partial f(a) \subset \mathbb{R}^n$. Denotamos $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ según la notación usual de aplicaciones punto a conjunto. Recordemos también los siguientes conceptos, si $u, v : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$

- $\text{dom}(u) = \{x : u(x) \neq \emptyset\}$
- $\text{grafo}(u) = \{(x, y) : y \in u(x)\}$
- $u^{-1}(y) = \{x : y \in u(x)\}$
- $u \subseteq v$ si $\text{grafo}(u) \subseteq \text{grafo}(v)$
- u es monótona si $\langle x_1 - x_2, y_1 - y_2 \rangle \geq 0$ para todo $x_1, x_2 \in \mathbb{R}^n$ y todo $y_1 \in u(x_1), y_2 \in u(x_2)$.
- u es maximal si se verifica la implicación $[u \subseteq v, v \text{ monótona}] \Rightarrow v = u$.

Proposición 4.3.4. *Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ es semicontinua inferior convexa, entonces ∂f es maximal monótona y verifica¹³*

$$\text{int}(\text{dom}(f)) \subseteq \text{dom}(\partial f) \subseteq \text{dom}(f)$$

Estudiemos ahora unos conceptos que nos harán falta más abajo para la *desigualdad de Lojasiewicz* de funciones convexas, subanalíticas y semicontinuas inferiormente.

Definición 4.3.1 (Pendiente no suave y puntos críticos). Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ una función de valores extendidos. La *pendiente no regular* (ó *no suave, traducción directa del inglés: non-smooth*) de f en un punto x se define como

$$m_f(x) := \inf\{\|x^*\| : x^* \in \partial f(x)\}$$

cuando $\partial f(x) \neq \emptyset$ y cuando $\partial f(x) = \emptyset$, como $m_f(x) := +\infty$. Se define también el *conjunto de puntos críticos* de f como

$$\text{crit}(f) := \{x \in \mathbb{R}^n : 0 \in \partial f(x)\}$$

Proposición 4.3.5. *Si f es semicontinua inferior y convexa, ó si el $\text{dom}(f)$ es cerrado y $f|_{\text{dom}(f)}$ es continua, entonces $\text{grafo}(\partial f)$ es cerrado. Como consecuencia, $\text{crit}(f)$ es cerrado también.*

Finalmente, necesitaremos el siguiente resultado interesante sobre funciones convexas semicontinuas que describe un procedimiento sistemático para “regularizar” una tal función. No incluimos aquí la demostración, que puede verse con detalle en [RW98] pág. 20 y ss¹⁴.

¹³ Ver la demostración en [NP06], enunciado 3.7.7, pág. 131

¹⁴ A lo largo de dichas páginas se establece también, de forma muy detallada, la teoría de las *envolventes de Moreau*, de las que hablamos más abajo y que nosotros sólo damos unos pequeños retazos

Proposición 4.3.6 (Diferenciabilidad de la suma epigráfica). *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ una función convexa semicontinua inferior tal que $\inf_{\mathbb{R}^n} f \in \mathbb{R}$. Si h es la suma epigráfica de f y $\frac{1}{2}\|\cdot\|^2$*

$$h(x) := \inf \left\{ f(u) + \frac{1}{2}\|x - u\|^2 : u \in \mathbb{R}^n \right\} \quad \forall x \in \mathbb{R}^n$$

entonces h es una función C^1 . Además se tienen las siguientes propiedades¹⁵

- (i) $h \leq f$.
- (ii) $\text{crit}(h) = \text{crit}(f)$.
- (iii) $\inf_{\mathbb{R}^n}(h) = \inf_{\mathbb{R}^n}(f)$.
- (iv) h toma siempre valores finitos, $\text{dom}(h) = \mathbb{R}^n$.

Este resultado forma parte de, la teoría más general, las *envolventes de Moreau*. Aquí daremos sólo unos pequeños retazos de ella.

Definición 4.3.2 (Envolventes de Moreau). Para una función propia, semicontinua inferior $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ y un $\lambda > 0$, la *envolvente de Moreau* es la función definida y denotada como

$$e_\lambda f(x) := \inf_{\mathbb{R}^n} \left\{ f(w) + \frac{1}{2\lambda}\|w - x\|^2 \right\} \leq f(x)$$

Nótese que si $\lambda = 1$ tenemos lo que nosotros hemos definido como la suma epigráfica.

En general, $e_\lambda f$ aproxima f ‘desde abajo’; y cuanto más pequeño sea el valor λ mejor será dicha aproximación. Ahora necesitamos una definición asociada

Definición 4.3.3 (prox-acotada). Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ es *prox-acotada* si existe un $\lambda > 0$ tal que $e_\lambda f(x) > -\infty$ para algún $x \in \mathbb{R}^n$. El supremo del conjunto de tales λ es el *umbral* λ_f de las prox-acotaciones para f .

Obtenemos estas caracterizaciones, que f sea prox-acotada equivale a cualquiera de las siguientes:

1. f mayor a una función cuadrática
2. $f + r/2\|\cdot\|^2$ está acotado por debajo en \mathbb{R}^n para algún $r \in \mathbb{R}$
3. $\liminf_{\|x\| \rightarrow \infty} f(x)/\|x\|^2 > -\infty$

Además se tiene que $e_\lambda f$ es finita y continua para todo $\lambda \in (0, \lambda_f)$.

¹⁵ Tal y como pueden verse con mayor detalle en [RW98]

Capítulo 5

Desigualdades de Brunn-Minkovsky e Isoperimétrica

En este capítulo vamos a desarrollar una aplicación clásica de las funciones convexas en varias variables que es la desigualdad de Brunn-Minkovsky y su corolario geométrico más conocido que es la desigualdad isoperimétrica. Daremos un par de demostraciones de la desigualdad de Brunn-Minkovsky, una usando el famoso *teorema de Prékopa-Leindler* y otra simplemente haciendo consideraciones de carácter general sobre la *medida de Lebesgue*. En realidad, ambas están relacionadas porque comparten parte de la demostración¹.

En lo que sigue, se supone conocida la teoría de integración de Lebesgue en los espacios euclídeos \mathbb{R}^n (si bien en un apéndice recogemos algunos de los resultados más básicos de la teoría de la medida en general, aplicables, claro está, a la medida de Lebesgue). Entonces, si $A \subset \mathbb{R}^n$, se dirá que A es *medible* (o que *tiene volumen*) si es medible para la medida de Lebesgue y en tal caso $\text{vol}(A)$ denotará la medida de Lebesgue del mismo².

¹ Todas estas cuestiones han sido muy estudiadas y aparecen en multitud de libros. Aquí probaremos la desigualdades de Brunn-Minkovsky e isoperimétrica para espacios euclídeos; para la extensión de éstas a los conocidos como *volúmenes mixtos*, la *rotación simétrica de Schwarz* de un conjunto adecuado alrededor de un eje, y la *simetrización de Steiner* ver [Web94]. En el libro de Burago, [BZ88], *Geometric Inequalities*, hay un estudio muy profundo de las desigualdades de Brunn-Minkovsky e isoperimétrica generalizadas a diferentes situaciones y espacios. Por ejemplo: en la *esfera* y en el *esp. de Lobachevsky*, para familias de conjuntos, en *esp. normados finito dimensionales*, para varias definiciones de área diferentes y para *volúmenes mixtos*.

² También representamos, indistintamente, por m a la medida de Lebesgue, por ejemplo en las integrales usamos esta notación

5.1. Desigualdad de Brunn-Minkovsky (motivación y enunciado)

Empecemos con una pequeña motivación. Si $A, B \subseteq \mathbb{R}^n$ son subconjuntos, entonces se define su **suma** de la forma usual como

$$A + B := \{x \in \mathbb{R}^n : x = a + b \text{ con } a \in A, b \in B\}$$

La primera observación que hay que hacer es que aunque A y B sean medibles, ello no implica que su suma también lo sea³. Pero si ambos son cerrados o uno de ellos es abierto entonces $A + B$ es medible. También conviene destacar que si A, B son compactos entonces $A + B$ es compacto y por tanto medible.

EJEMPLO. Sea $A \subset \mathbb{R}^n$ cerrado y sea B_ε la bola cerrada de centro 0 y radio $\varepsilon > 0$. Entonces $A + B_\varepsilon$ es el conjunto de puntos de \mathbb{R}^n que distan del conjunto A a lo más ε , es decir

$$A + B_\varepsilon = \{x \in \mathbb{R}^n : \text{dist}(x, A) \leq \varepsilon\}$$

Nos preguntamos ahora, suponiendo que $A+B$ sea medible, ¿podemos acotar $\text{vol}(A+B)$ superiormente en términos de $\text{vol}(A)$ y $\text{vol}(B)$? La respuesta es NO, pues por ejemplo en \mathbb{R}^2 , basta tomar los subconjuntos (segmentos ambos de \mathbb{R}^2)

$$A = [0, 1] \times \{0\} = \{(x, 0) : x \in [0, 1]\} \quad B = \{0\} \times [0, 1] = \{(0, y) : y \in [0, 1]\}$$

que tienen medida (volumen) cero; y sin embargo, su suma $A + B = [0, 1]^2$ tiene medida 1. Sin embargo, acotar la suma de los volúmenes de A y B por el volumen de $A + B$ si va a ser posible y es precisamente lo que constituye la **desigualdad de Brunn-Minkovsky** en \mathbb{R}^n .

Teorema 25 (Desigualdad de Brunn-Minkovsky).

Sean $A, B \subseteq \mathbb{R}^n$ medibles tal que $A + B$ también lo es. Entonces,

$$\text{vol}(A + B)^{1/n} \geq \text{vol}(A)^{1/n} + \text{vol}(B)^{1/n} \quad (\star)$$

Con respecto a la desigualdad de Brunn-Minkovsky, queremos señalar que se conocen bien *los casos en que se da la igualdad*. Por ejemplo, es fácil ver que si $\text{vol}(A + B) = 0$ (la suma es de medida nula) entonces se da la igualdad y ambos sumandos son de medida nula. También es bastante fácil mostrar que si, por ejemplo, $\text{vol}(A) = 0$ y se da la igualdad $\text{vol}(A + B) = \text{vol}(B)$, entonces A es un punto. Pero también se tiene el siguiente resultado para el caso en que se da la igualdad y ambos sumandos A y B tienen medida positiva.

Proposición 5.1.1. Si A y B son compactos y de medida no nula para los cuales se da la igualdad (\star) del teorema (25), entonces A, B son ambos convexos y uno se obtiene de otro por una homotecia de razón positiva más una traslación⁴.

³ En [Sha05] aparece un ejemplo de esto, concretamente es el ejercicio 13 del capítulo 2, pág. 92. Básicamente consiste en considerar los subconjuntos de \mathbb{R}^2 , $\{0\} \times [0, 1]$ y $\mathcal{N} \times \{0\}$, donde \mathcal{N} es el conjunto de Vitali, que se construye en la pág. 24 del mismo libro

⁴ En [BZ88] puede verse una demostración geométrica muy elegante de esta proposición

5.2. Desigualdad Isoperimétrica

Se trata aquí de obtener una versión de la desigualdad isoperimétrica para convexos y compactos de \mathbb{R}^n como consecuencia de la desigualdad de Brunn-Minkovsky.

Definición 5.2.1 (Área de un conjunto de \mathbb{R}^n). Sea $C \subseteq \mathbb{R}^n$ un subconjunto compacto. Se define el **área** de C ó el **volumen de ∂C** como

$$a(C) := \limsup_{t \rightarrow 0^+} \frac{\text{vol}(C + tB_n) - \text{vol}(C)}{t} =: \text{vol}(\partial C) \in [0, +\infty]$$

siendo B_n la bola unidad euclídea de \mathbb{R}^n .

En algunas referencias de la bibliografía al área se le llama también el *contenido exterior de Minkovsky (outer Minkovsky content)*. Hemos preferido tomar el nombre más “gráfico” de área porque esta definición nos da precisamente el área $(n - 1)$ -dimensional de la frontera de C , ∂C , en el caso en que C es un cuerpo convexo, es decir, un compacto, convexo con interior no vacío. Caso en el que además la existencia del límite (no necesariamente superior) está garantizada. Por otro lado, si C es tal que su frontera es diferenciable a trozos (esto es, C es un *dominio regular*), entonces el límite en la fórmula anterior existe y es igual al área $(n - 1)$ -dimensional de dicho borde⁵.

Fijémonos que en la definición del área, el $\text{vol}((C + tB_n))$ es, según lo que hemos indicado en la sección (6.1), el volumen del conjunto de puntos que distan de C como mucho t . Esto puede justificar gráficamente porqué la definición de área mide el “volumen” de ∂C .

Como curiosidad, para ilustrar esta definición vamos a evaluar el área de la bola unidad cerrada $B_n \subseteq \mathbb{R}^n$. Denotemos por $\omega_n = \text{vol}(B_n)$. Sabemos que, para cualquier $t > 0$, $\text{vol}(B_n + tB_n) = \text{vol}((1 + t)B_n) = \omega_n(1 + t)^n$, así

$$a(B_n) = \lim_{t \rightarrow 0^+} \frac{\omega_n(1 + t)^n - \omega_n}{t} = n\omega_n$$

Esto es, si $n = 2$, $a(B_2) = 2\omega_2 = 2\pi$, el perímetro de un círculo de radio unidad, claro está. Y si, $n = 3$, $a(B_3) = 3\omega_3 = 4\pi$, el área de la superficie de una bola unidad en \mathbb{R}^3 .

Por la proposición (5.1.1), si se da la igualdad de Brunn-Minkovsky para el conjunto $C + tB_n$ empleado en la definición de su área, entonces C es homotético a B_n , y es por tanto también una bola. Estas consideraciones nos pueden servir de motivación para la desigualdad isoperimétrica clásica, que se obtiene directamente como consecuencia de Brunn-Minkovsky.

⁵ Ver [BZ88] para ello, aunque dicho autor envía al lector a otra referencia para la prueba que está en alemán y no hemos sido capaces de encontrar. También aparece un resultado similar en [Web94] para polítopos y después para compactos convexos. Este autor usa la denominación *outer Minkovsky content* que decíamos arriba

Teorema 26 (La desigualdad isoperimétrica).

Sea C un subconjunto convexo y compacto con interior no vacío de \mathbb{R}^n , entonces se tiene que

$$\left(\frac{\text{vol}(C)}{\text{vol}(B_n)}\right)^{1/n} \leq \left(\frac{a(C)}{a(B_n)}\right)^{1/(n-1)}$$

Demostración. Por Brunn-Minkovsky,

$$\begin{aligned} \text{vol}(C + tB_n) &\geq [\text{vol}(C)^{1/n} + t \cdot \text{vol}(B_n)^{1/n}]^n \\ &= \text{vol}(C) + n \cdot t \cdot \text{vol}(B_n)^{1/n} \text{vol}(C)^{(n-1)/n} + O(t^2) \end{aligned}$$

entonces

$$\frac{\text{vol}(C + tB_n) - \text{vol}(C)}{t} \geq n \cdot \text{vol}(B_n)^{1/n} \text{vol}(C)^{(n-1)/n} + O(t)$$

por lo que

$$a(C) \geq n \cdot \text{vol}(B_n)^{1/n} \text{vol}(C)^{(n-1)/n} + \lim_{t \rightarrow 0^+} O(t)$$

Consecuentemente

$$\begin{aligned} \left(\frac{a(C)}{a(B_n)}\right)^{1/(n-1)} &\geq \frac{n^{\frac{1}{n-1}} \text{vol}(B_n)^{\frac{1}{n(n-1)}} \text{vol}(C)^{1/n}}{(n \cdot \text{vol}(B_n))^{1/(n-1)}} = \frac{\text{vol}(C)^{1/n}}{\text{vol}(B_n)^{\frac{1}{n-1} - \frac{1}{n(n-1)}}} \\ &= \left(\frac{\text{vol}(C)}{\text{vol}(B_n)}\right)^{1/n} \end{aligned}$$

y queda así probada ésta versión⁶ de la desigualdad isoperimétrica. \square

Observación. Nótese que $a(B_n) = n \cdot \text{vol}(B_n)$, y que para conjuntos C convexos, compactos y de interior no vacío que tengan área igual a la de B_n , el volumen de ésta es mayor ó igual que el del conjunto C . Esto justifica el nombre de desigualdad isoperimétrica. Más precisamente, tenemos el siguiente resultado como consecuencia inmediata del teorema (26).

Teorema 27 (Frontera de menor volumen).

Sea $\omega_n := \text{vol}(B_n)$. Entonces B_n tiene la frontera con menor volumen; es decir, el área de B_n es la menor de entre todos los subconjuntos de \mathbb{R}^n compactos convexos y con interior no vacío de volumen ω_n .

Demostración. Sea $A \subseteq \mathbb{R}^n$ con $\text{vol}(A) = \omega_n$. Defino $f(t) := \text{vol}(A + tB_n)$ para todo $t \geq 0$. Entonces

$$\left.\frac{d}{dt}\right|_{t=0} (f(t)^{1/n}) = \frac{1}{n} f(0)^{\frac{1}{n}-1} f'(0) = \frac{1}{n} \text{vol}(A)^{1/n-1} a(A)$$

⁶ Ya hemos comentado en una nota al principio del capítulo donde encontrar más resultados sobre este tema. Añadir que, en [Sha05], pág. 134 y ss. viene la desigualdad isoperimétrica para curvas rectificables de \mathbb{R}^n

pero por otra parte

$$\left. \frac{d}{dt} \right|_{t=0} (f(t)^{1/n}) = \lim_{t \rightarrow 0^+} \frac{\text{vol}(A + tB_n)^{1/n} - \text{vol}(A)^{1/n}}{t} \geq \text{vol}(B_n)$$

luego, $a(A) \geq n \text{vol}(B_n)^{1/n} \text{vol}(A)^{1-1/n} = n \cdot \omega_n = a(B_n)$. \square

5.3. Prueba de la desigualdad de Brunn-Minkovsky

5.3.1. Primera demostración

Empecemos por la que no hace uso del teorema (28). Se trata de una demostración geométrica y directa de (25), tal y como puede encontrarse en [BZ88] y en [Sha05]. Pero antes establezcamos la demostración para el caso de compactos de dimensión 1, pues se usa parcialmente en la prueba del teorema (28).

Paso 0. Desigualdad de Brunn-Minkovsky para compactos de dimensión 1. Supongamos primero que A, B son subconjuntos compactos de \mathbb{R} . Sean $a_0 := \inf(A) \in A$ y $b_1 := \sup(B) \in B$. Entonces $(a_0 + B) \cap (b_1 + A) = \{a_0 + b_1\}$, pues $a_0 + b \leq a + b_1$ para todos $a \in A$ y $b \in B$, con igualdad si y sólo si $a = a_0$ y $b = b_1$. Luego, claramente $(a_0 + B) \cup (b_1 + A) \subseteq A + B$ y, por tanto, como la medida de Lebesgue es invariante por traslaciones

$$\text{vol}(A + b_1) + \text{vol}(a_0 + B) = \text{vol}(A) + \text{vol}(B)$$

y como la intersección hecha arriba tiene medida nula, pues es un punto, la expresión superior es igual a

$$\text{vol}((a_0 + B) \cup (b_1 + A)) \leq \text{vol}(A + B)$$

lo que nos da la desigualdad de Brunn-Minkovsky en este caso. Para continuar será necesario el siguiente lema.

Lema 5.3.1. Sean $A \subseteq \mathbb{R}^n$ medible y acotado y $\theta \geq 0$. Entonces, para cada $i = 1, \dots, n$ $\exists \lambda_i \in \mathbb{R}$ tal que

$$\text{vol}(\{(x_1, \dots, x_n) \in A : x_i < \lambda_i\}) = \theta \text{vol}(\{(x_1, \dots, x_n) \in A : x_i > \lambda_i\})$$

Demostración. Como A es acotado, $\exists a > 0$ tal que

$$A \subseteq \{(x_1, \dots, x_n) : -a \leq x_i \leq a \quad \forall i = 1, \dots, n\} \quad (1)$$

Defino $f_i : \mathbb{R} \rightarrow \mathbb{R}$ como $f_i(x) := \text{vol}(\{y \in A : y_i < x\})$; así si $x < y$ usando (1)

$$0 \leq f_i(y) - f_i(x) = \text{vol}(\{(x_1, \dots, x_n) \in A : x \leq x_i \leq y\}) \leq (2a)^{n-1}(y - x)$$

lo que demuestra que cada f_i es continua. Por otro lado, $f_i(-a) = \text{vol}(\emptyset) = 0$ y $f_i(a) = \text{vol}(A \setminus N) = \text{vol}(A)$ con N un conjunto de medida nula. Luego para algún $\lambda_i \in [-a, a]$ tendremos que $f_i(\lambda_i) = \theta(1 + \theta)^{-1} \text{vol}(A)$. Consecuentemente $f_i(\lambda_i)(1 + \theta) = \theta f_i(a)$, y $f_i(\lambda_i) = \theta(f_i(a) - f_i(\lambda_i)) = \theta \text{vol}(\{x \in A : x_i > \lambda_i\})$. \square

Primera demostración

Se lleva a cabo en varios pasos, usando el anterior para el siguiente. Se trata de ir restringiendo cada vez menos el tipo de conjuntos que son A y B .

Paso 1. Para rectángulos⁷. Sean A y B dos rectángulos de \mathbb{R}^n con *lados* de longitudes $\{a_i\}_{i=1}^n$ y $\{b_i\}_{i=1}^n$ respectivamente. Hacemos la reducción siguiente; gracias a que sabemos medir rectángulos y a que la suma de rectángulos da otro rectángulo, la desigualdad (\star) equivale a probar

$$\left(\prod_1^n (a_i + b_i)\right)^{1/n} \geq \left(\prod_1^n a_i\right)^{1/n} + \left(\prod_1^n b_i\right)^{1/n} \quad (\clubsuit)$$

Por homogeneidad, podemos suponer que $a_i + b_i = 1$ para todo $i = 1, \dots, n$. En efecto, si sustituimos las cantidades a_i, b_i por $\alpha_i a_i, \alpha_i b_i$ respectivamente con $\alpha_i > 0$ para todo i , entonces en la desigualdad (\clubsuit) ambos miembros quedan multiplicados por el factor $(\alpha_1 \cdots \alpha_n)^{1/n}$ y nada cambiaría en ella. Así, sólo tenemos que elegir $\alpha_i = \frac{1}{a_i + b_i}$, que lo podemos hacer ya que $a_i, b_i > 0$ por ser longitudes de lados de rectángulos de interior no vacío.

Ahora bien, una vez hecha la reducción, la desigualdad (\clubsuit) sale debido a la desigualdad media aritmética mayor que media geométrica (ver capítulo 1, pág. 4). Así, en nuestro caso, fácilmente

$$\frac{1}{n} \sum_{i=1}^n a_i \geq \left(\prod_{i=1}^n a_i\right)^{1/n} \quad \text{y} \quad \frac{1}{n} \sum_{i=1}^n b_i \geq \left(\prod_{i=1}^n b_i\right)^{1/n}$$

Sumando ambas expresiones tendremos que

$$\left(\prod_{i=1}^n b_i\right)^{1/n} + \left(\prod_{i=1}^n a_i\right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^n (a_i + b_i) = 1 = \left(\prod_1^n (a_i + b_i)\right)^{1/n}$$

Paso 2. Para uniones finitas de rectángulos con interiores disjuntos dos a dos.

Sean ahora $A = I_1 \cup \dots \cup I_m$ y $B = J_1 \cup \dots \cup J_p$ como en el título del paso. Lo demostraremos por inducción sobre $m + p$. Nótese que el primer paso de la inducción, cuando $m + p = 2$, correspondería al paso 1. Supongamos entonces que se cumple para cuando el número total de rectángulos es menor que $m + p > 2$. Podemos asumir que $m \geq 2$. Como los rectángulos I_1 e I_2 son *casi-disjuntos*, existe algún $i \in \{1, \dots, n\}$ y algún $\mu \in \mathbb{R}$ tal que I_1 cae en el *subespacio mitad cerrado* $x_i \leq \mu$ e I_2 hace lo propio en $x_i \geq \mu$, ó viceversa; es decir existe un hiperplano que separa los rectángulos I_1 e I_2 . Denotamos por A^- y A^+ a las intersecciones de A con los subespacios mitad abiertos $x_i < \mu$ y $x_i > \mu$. Entonces, tanto A^- como A^+ son no vacíos, porque contienen al interior de I_1 ó I_2 ; y además son unión de una cantidad finita de rectángulos casi-disjuntos dos a dos menor que m , pues al menos un rectángulo cae en el otro lado. Es decir, A^- y A^+ tienen al menos un rectángulo menos que A .

Por otra parte, como $A = A^+ \cup A^- \cup \{x_i = \mu\}$ con uniones disjuntas y $\text{vol}(\{x_i = \mu\}) = 0$, $\text{vol}(A) = \text{vol}(A^-) + \text{vol}(A^+)$. Por el lema (5.3.1), existe un $\lambda \in \mathbb{R}$ tal que

⁷ En lo que sigue, los rectángulos son de lados paralelos a los ejes y de interior no vacío. También conocidos como *hiper-rectángulos* o *hipercubos*

el hiperplano $x_i = \lambda$ divide a B (de manera similar a como hemos hecho más arriba con A) en dos conjuntos disjuntos B^- , B^+ verificando que

$$\frac{\text{vol}(B^-)}{\text{vol}(A^-)} = \frac{\text{vol}(B^+)}{\text{vol}(A^+)} = \alpha$$

donde B^- , B^+ son unión de p ó menos rectángulos casi-disjuntos dos a dos y $\text{vol}(B) = \text{vol}(B^-) + \text{vol}(B^+)$. Ahora bien, los conjuntos suma $A^- + B^-$ y $A^+ + B^+$ caen en semiespacios abiertos acotados por el hiperplano $x_i = \lambda + \mu$, y así son disjuntos; además su unión es un subconjunto de $A + B$. Aplicando la HI a los pares (A^-, B^-) y (A^+, B^+) tenemos que

$$\begin{aligned} \text{vol}(A + B) &\geq \text{vol}(A^- + B^-) + \text{vol}(A^+ + B^+) \\ &\geq (\text{vol}(A^-)^{1/n} + \text{vol}(B^-)^{1/n})^n + (\text{vol}(A^+)^{1/n} + \text{vol}(B^+)^{1/n})^n \\ &= (\text{vol}(A^-) + \text{vol}(A^+))(1 + \alpha^{1/n})^n = \text{vol}(A)(1 + \alpha^{1/n})^n \\ &= (\text{vol}(A)^{1/n} + \alpha^{1/n}\text{vol}(A)^{1/n})^n = (\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n})^n \end{aligned}$$

lo que completa la inducción.

Paso 3. Conjuntos medibles con volumen positivo. Una de las propiedades de la medida de Lebesgue nos asegura que existen sucesiones $\{A_i\}_{i \in \mathbb{N}} \subseteq A$, $\{B_i\}_{i \in \mathbb{N}} \subseteq B$ de conjuntos no vacíos formados por uniones finitas de rectángulos con interior no vacío casi-disjuntos dos a dos y de lados paralelos a los ejes, verificando que $\text{vol}(A_i) \rightarrow \text{vol}(A)$ y $\text{vol}(B_i) \rightarrow \text{vol}(B)$. Luego, por el paso 2 y $A_i + B_i \subseteq A + B$

$$\text{vol}(A + B)^{1/n} \geq \text{vol}(A_i + B_i)^{1/n} \geq \text{vol}(A_i)^{1/n} + \text{vol}(B_i)^{1/n}$$

Estas desigualdades se mantienen tomando el límite superior cuando $i \rightarrow \infty$ y puesto que los dos sumandos en el último término tienen límite igual a $\text{vol}(A)^{1/n}$ y $\text{vol}(B)^{1/n}$ respectivamente, tenemos el resultado buscado. \square

Corolario 5.3.1. Sean $A, B \subseteq \mathbb{R}^n$ conjuntos convexos acotados y no vacíos. Entonces la función $f : [0, 1] \rightarrow \mathbb{R}$ con $f(t) := \text{vol}((1-t)A + tB)^{1/n}$ es cóncava.

Demostración. Sean $x, y \in [0, 1]$, $\lambda, \mu \geq 0$ con suma 1. Aplicamos el teorema (25) a los conjuntos $\lambda((1-x)A + xB)$ y $\mu((1-y)A + yB)$

$$\begin{aligned} f(\lambda x + \mu y) &= \text{vol}((1 - (\lambda x + \mu y))A + (\lambda x + \mu y)B)^{1/n} \\ &= \text{vol}(\lambda((1-x)A + xB) + \mu((1-y)A + yB))^{1/n} \\ &\geq \lambda \text{vol}((1-x)A + xB)^{1/n} + \mu \text{vol}((1-y)A + yB)^{1/n} \\ &= \lambda f(x) + \mu f(y) \end{aligned} \quad \square$$

5.3.2. Desigualdad de Prékopa-Leindler

A continuación proponemos el siguiente resultado que tiene un gran interés en sí mismo y que nos servirá para dar una demostración distinta de la desigualdad de Brunn-Minkovsky.

Teorema 28 (Desigualdad de Prékopa-Leindler).

Sean $f, g, \phi : \mathbb{R}^n \rightarrow [0, \infty]$ medibles y tales que, para algún $0 < \lambda < 1$ y para todo $r, s \in \mathbb{R}^n$, se verifica

$$\phi(\lambda r + (1 - \lambda)s) \geq f(r)^\lambda g(s)^{1-\lambda}$$

Entonces, se tiene que para la integral de Lebesgue

$$\int \phi \, dm \geq \left(\int f \, dm \right)^\lambda \left(\int g \, dm \right)^{1-\lambda}$$

Demostración. Se hace la prueba por inducción. Cuando $n = 1$. Sin pérdida de generalidad podemos suponer que $\|f\|_\infty = \|g\|_\infty = 1$ ⁸. Para cada $0 \leq a < 1$ se tiene que

$$\{\phi \geq a\} \supseteq \lambda\{f \geq a\} + (1 - \lambda)\{g \geq a\}$$

En efecto, si x pertenece al conjunto de la derecha entonces x es de la forma $\lambda r + (1 - \lambda)s$, y por lo tanto $\phi(\lambda r + (1 - \lambda)s) \geq f(r)^\lambda g(s)^{1-\lambda} \geq a^\lambda a^{1-\lambda} = a$. Ahora bien, como los dos conjuntos de la derecha son no vacíos, pues $\|f\|_\infty = \|g\|_\infty = 1$, entonces, usando la desigualdad del paso 0

$$\text{vol}(\{\phi \geq a\}) \geq \lambda \text{vol}(\{f \geq a\}) + (1 - \lambda) \text{vol}(\{g \geq a\})$$

e integrando, obtenemos que

$$\int_a^\infty \phi \, dm \geq \lambda \int_a^\infty f \, dm + (1 - \lambda) \int_a^\infty g \, dm \geq \left(\int_a^\infty f \, dm \right)^\lambda \left(\int_a^\infty g \, dm \right)^{1-\lambda}$$

donde en la última desigualdad hemos usado la desigualdad de las medias. Concluimos la desigualdad requerida haciendo $a \rightarrow 0^+$.

Supongamos cierto para $n - 1$ y probémoslo para n . Partimos de $f, g, \phi : \mathbb{R}^n \rightarrow [0, \infty]$ medibles como en el enunciado y tomamos $y \in \mathbb{R}$ fijo. Definimos, $f_y, g_y, \phi_y : \mathbb{R}^{n-1} \rightarrow [0, \infty]$ las respectivas funciones que resultan de fijar el valor y para la primera coordenada. Es claro que si y_0, y_1 son tales que $y = \lambda y_1 + (1 - \lambda)y_0$, se tiene que

$$\phi_y(\lambda r + (1 - \lambda)s) \geq f_{y_1}(r)^\lambda g_{y_0}(s)^{1-\lambda}$$

para todo $r, s \in \mathbb{R}^{n-1}$ sin más que reagrupar y usar la hipótesis del enunciado. Por lo tanto, por la hipótesis de inducción

$$\int_{\mathbb{R}^{n-1}} \phi_y \, dm \geq \left(\int_{\mathbb{R}^{n-1}} f_{y_1} \, dm \right)^\lambda \left(\int_{\mathbb{R}^{n-1}} g_{y_0} \, dm \right)^{1-\lambda}$$

y utilizando el caso $n = 1$ se consigue que

$$\int_{\mathbb{R}^n} \phi \, dm \geq \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} \phi_y \, dm \right) dy \geq \left(\int_{\mathbb{R}^n} f \, dm \right)^\lambda \left(\int_{\mathbb{R}^n} g \, dm \right)^{1-\lambda} \quad \square$$

⁸ Esta norma denota el supremo de los valores de la función

5.3.3. Segunda demostración

No está de más observar que, aunque hemos utilizado la desigualdad de Brunn-Minkovsky para probar la desigualdad de Prekopka-Leindler, en realidad sólo hemos usado dicha desigualdad para conjuntos medibles en dimensión uno, caso que se prueba fácilmente con el paso 0 de (6.3.1) más un argumento como el que se ha hecho en general, que permite pasar de conjuntos compactos a conjuntos medibles.

En este apartado comprobamos que la desigualdad de Prekopka-Leindler permite también dar una demostración de la desigualdad de Brunn-Minkovsky, por lo que *ambas desigualdades son equivalentes*. Aún más, la desigualdad (28) es la forma inversa de la desigualdad de Rogers-Hölder (teorema (36))⁹. La desigualdad de Prekopka-Leindler puede, evidentemente, demostrarse independientemente de la desigualdad de Brunn-Minkovsky. Puede verse una tal demostración en el libro de Constantinu, referencia [NP06], como una consecuencia de una desigualdad más general conocida con el nombre de *desigualdad de Borell-Brascamp-Lieb*¹⁰.

Paso 1. Para compactos n -dimensionales. Aplicamos el teorema (28) a las funciones $f = \chi_A$, $g = \chi_B$ y $\phi = \chi_{\lambda A + (1-\lambda)B}$ con $\lambda > 0$ obtenemos fácilmente que

$$\text{vol}(\lambda A + (1 - \lambda)B) \geq \text{vol}(A)^\lambda \text{vol}(B)^{1-\lambda} \quad (\spadesuit)$$

Pero si hacemos

$$\lambda = \frac{\text{vol}(A)^{1/n}}{\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}} > 0$$

conseguimos rápidamente la desigualdad buscada. Pues, si definimos los conjuntos $A' = \text{vol}(A)^{-1/n}A$ y $B' = \text{vol}(B)^{-1/n}B$ y les aplicamos (\spadesuit), teniendo en cuenta que en \mathbb{R}^n , $\text{vol}(tC) = t^n \text{vol}(C)$, y concluimos la desigualdad usando $\text{vol}(\lambda A' + (1 - \lambda)B') \geq 1$ y como

$$\lambda A' + (1 - \lambda)B' = \frac{1}{\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}}(A + B)$$

Paso 2. Para medibles. Supongamos ahora que A, B son conjuntos medibles con volumen positivo y finito. Para probar la desigualdad de Brunn-Minkovsky para A, B usamos la propiedad que ya hemos usado en el último paso de la primera demostración: existe una sucesión de conjuntos compactos $A_i \subset A$ de modo que $\lim_i \text{vol}(A_i) = \text{vol}(A)$, y lo mismo para B . \square

Corolario 5.3.2 (Forma multiplicativa de la desigualdad BM). Sean $A, B \subseteq \mathbb{R}^n$ medibles no vacíos, $0 < \lambda < 1$ tal que $\lambda A + (1 - \lambda)B$ es medible. Entonces,

$$\text{vol}(\lambda A + (1 - \lambda)B) \geq \text{vol}(A)^\lambda \text{vol}(B)^{1-\lambda}$$

Observación. La demostración es inmediata a partir del teorema (28). De hecho, de la segunda demostración se deduce que este corolario es equivalente al teorema (25).

⁹ Ir al teorema 3.12.1, pág. 158 de [NP06] para verlo

¹⁰ Ver [NP06], teorema 3.12.5, pág. 160 para la demostración. La desigualdad de Prekopka-Leindler, (28), es el caso particular $p = 0$ de ésta

Capítulo 6

La desigualdad de Lojasiewicz para funciones convexas. Aplicación a sistemas dinámicos

En este último capítulo vamos a utilizar algunos de los conceptos introducidos en capítulos anteriores sobre funciones convexas, como la subdiferencial, para probar una generalización, a funciones convexas y subanalíticas, de la desigualdad de Lojasiewicz para funciones analíticas¹. Seguidamente, aplicaremos esta desigualdad al estudio de sistemas dinámicos definidos por el (sub)gradiente de la función convexa en cuestión, probando la longitud finita de las trayectorias (curvas integrales) de dicho sistema.

El contenido de este capítulo está basado en el artículo de investigación de *Bolte-Danilidis-Lewis*, referencia [BDL07].

6.1. Conjuntos subanalíticos y funciones subanalíticas

Aquí sólo pretendemos dar una recopilación de los resultados que nos hacen falta para probar la desigualdad de Lojasiewicz del texto. Para una exposición más detallada de la geometría subanalítica, ver [Bie88] o [Mil96] para generalizaciones a la geometría o-minimal.

Definición 6.1.1 (Subanaliticidad). Englobamos varias definiciones

1. Un conjunto $A \subseteq \mathbb{R}^n$ se dice *semianalítico* si cada punto de \mathbb{R}^n admite una vecindad V para la cual $A \cap V$ se puede escribir como

$$\bigcup_{i=1}^p \bigcap_{j=1}^q \{x \in V : f_{ij}(x) = 0, g_{ij}(x) > 0\}$$

donde las funciones $f_{ij}, g_{ij} : V \rightarrow \mathbb{R}$ son analíticas reales para cada i y j .

¹ Ver referencia [Loj83], *Sur les trajectoires du gradient d'une fonction analytique*

2. A se dice que es *subanalítico* si cada punto de \mathbb{R}^n tiene un entorno V tal que

$$A \cap V = \{x \in \mathbb{R}^n : (x, y) \in B\}$$

donde $B \subseteq V \times \mathbb{R}^m$ es un subconjunto semianalítico acotado para algún $m \geq 1$; en pocas palabras, A es subanalítico si localmente en cada punto es proyección de un conjunto semianalítico acotado.

3. Dados $m, n \in \mathbb{N}^*$, una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ (resp. una aplicación punto a conjunto $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$) se dice *subanalítica* si su grafo es un conjunto subanalítico de $\mathbb{R}^n \times \mathbb{R}$ (resp. de $\mathbb{R}^n \times \mathbb{R}^m$).

La siguiente proposición recoge las propiedades geométricas básicas de los conjuntos subanalíticos.

Proposición 6.1.1. *Se tiene que*

1. La familia de conjuntos subanalíticos de \mathbb{R}^n es cerrada bajo uniones e intersecciones finitas, por paso al complementario y por productos cartesianos.
2. **Teorema de la proyección.** La imagen de un conjunto subanalítico por una proyección lineal $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n$, con $m \geq n$, es subanalítico. También lo es la imagen por una aplicación subanalítica propia (ver [Bie88], lema 6.4).
3. La frontera, clausura e interior de un subanalítico también lo son.
4. La función distancia a un conjunto subanalítico S , $d_S(x) := \inf_{a \in S} \{\|x - a\|\}$ es subanalítica.
5. Un conjunto subanalítico tiene localmente en cada punto un número finito de componentes conexas que son subanalíticas y conexas por caminos subanalíticos (ver [Mil96], 1.10 a 1.12).
6. Los conjuntos subanalíticos verifican el lema de la elección de curva: Si $A \subset \mathbb{R}^n$ es subanalítico y $x \in \partial A$, existe un camino subanalítico continuo $z : (-1, 1) \rightarrow \mathbb{R}^n$ que en el origen pasa por x y tal que $z((0, 1))$ está contenido en A (ver [Bie88], lema 6.3).

Las propiedades 1 y 2 de la proposición anterior pueden axiomatizarse para familias de subconjuntos de \mathbb{R}^n para cada n , dando lugar a las llamadas *estructuras o-minimales* cuando se da la propiedad adicional de finitud (global) del número de componentes conexas. Los conjuntos subanalíticos como los hemos definido no tienen esta propiedad globalmente, sólo localmente como asegura la propiedad 5. Tómese por ejemplo la intersección del grafo de la función $\sin(x)$ en \mathbb{R} con el eje $y = 0$, conjunto subanalítico que consiste en una sucesión infinita de puntos en la recta real. Esta sucesión no tiene puntos de acumulación en la recta, pero sí en el “infinito”, en un entorno del cual deja de ser un conjunto subanalítico.

Para evitar estos fenómenos se definen los conjuntos *subanalíticos globales* que son también subanalíticos “en el infinito” y que tienen todas las propiedades de la proposición anterior y además la finitud del número de componentes conexas.

Definición 6.1.2 (Subanaliticidad global, ver [Mil96] pág. 506). Para cada $n \in \mathbb{N}$, consideremos la aplicación

$$\tau_n(x_1, \dots, x_n) := \left(\frac{x_1}{1+x_1^2}, \dots, \frac{x_n}{1+x_n^2} \right) \in (-1, 1)^n$$

1. Un subconjunto $S \subseteq \mathbb{R}^n$ se dice que es *globalmente subanalítico* si $\tau_n(S) \subseteq \mathbb{R}^n$ es subanalítico.
2. Una función de valores reales extendidos es *globalmente subanalítica* si su grafo lo es.

Necesitamos también el siguiente resultado.

Proposición 6.1.2 (Lema de factorización de Lojasiewicz (ver [Mil96], ejem. 4, pág. 505)). *Sea $K \subseteq \mathbb{R}^n$ un compacto, y $f, g : K \rightarrow \mathbb{R}$ dos funciones continuas subanalíticas. Si $f^{-1}(0) \subseteq g^{-1}(0)$, entonces existe un $c > 0$ y un entero positivo r tales que*

$$|g(x)|^r \leq c|f(x)| \quad \forall x \in K$$

6.2. La desigualdad de Lojasiewicz

Aquí tratamos con funciones extendidas convexas subanalíticas globales² y semi-continuas inferiores de la forma $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ con $\text{dom}(f) \neq \emptyset$. Por la definición de función (extendida) subanalítica, se deduce que el dominio $\text{dom}(f)$, así como el epigrafo $\text{epi}(f)$ son conjuntos subanalíticos (globales, si lo es f).

Recordemos también del apartado 4.3 la definición de la subdiferencial $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ de f , el conjunto $\text{crit}(f) = \{x : 0 \in \partial f(x)\}$ de puntos críticos generalizados y la función de pendiente no regular $m_f : \mathbb{R}^n \rightarrow \mathbb{R}^c$, $m_f(x) = \min\{\|x^*\| : x^* \in \partial f(x)\}$. Gracias a la proposición (4.3.3), en este caso el grafo de la subdiferencial ∂f es cerrado y el conjunto $\text{crit}(f)$ es cerrado y convexo y coincide con el conjunto de puntos que minimizan los valores de f .

Pues bien, de la definición de función subanalítica global y usando las propiedades de los conjuntos subanalíticos, recogidas en la proposición (6.1.1), obtenemos que ∂f es una aplicación punto a conjunto subanalítica global, que m_f es una función (extendida) subanalítica global y que $\text{crit}(f)$ es un conjunto subanalítico global. Para ilustrar los argumentos estándar que se usan para probar estas afirmaciones, vamos a comprobar que la subdiferencial es subanalítica. Para ello recordamos primero que $\partial f(x)$ es el conjunto de puntos $x^* \in \mathbb{R}^n$ tales que la función $f(\cdot) - \langle x^*, \cdot \rangle$ tiene un mínimo global en x (ver proposición (4.3.3)). Así, tendremos que el grafo de ∂f se describe como el conjunto de pares $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ tales que $x \in \text{dom}(f)$ y

$$f(t) - \langle y, t \rangle \geq f(x) - \langle y, x \rangle \quad \forall t \in \text{dom}(f)$$

o, equivalentemente,

$$\nexists t \in \text{dom}(f) \quad \text{con} \quad f(t) - \langle y, t \rangle < f(x) - \langle y, x \rangle$$

Consideramos entonces primero el conjunto A de $\text{dom}(f) \times \mathbb{R}^n \times \text{dom}(f) \subset (\mathbb{R}^n)^3$ definido por $\{(x, y, t) : f(t) - \langle y, t \rangle < f(x) - \langle y, x \rangle\}$, que es subanalítico en $(\mathbb{R}^n)^3$ por serlo la función f y nuestro conjunto $\text{grafo}(\partial f)$ es el complementario

² En realidad, no va a resultar una verdadera restricción suponer que f es subanalítica *global*. Pueden probarse todos los resultados y afirmaciones que siguen (con los cambios necesarios en los enunciados) para funciones subanalíticas, aún no subanalíticas globales si suponemos la propiedad de que f es localmente acotada relativamente en su dominio, ver [BDL07] para los detalles

en $\mathbb{R}^n \times \mathbb{R}^n$ de la imagen de A por la proyección lineal $\pi(x, y, t) = (x, y)$.

Por otro lado, recordemos también del apartado 4.3, que hemos definido para una función semicontinua inferiormente y convexa $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ la suma epigráfica (una de las envolventes de Moreau de f)

$$h(x) := \inf \left\{ f(u) + \frac{1}{2} \|x - u\|^2 : u \in \mathbb{R}^n \right\} \quad \forall x \in \mathbb{R}^n$$

que, por la proposición (4.3.6) es una función de clase C^1 . Usando argumentos parecidos a los anteriores, puede también probarse lo siguiente.

Proposición 6.2.1 (Subanaliticidad de la suma epigráfica). *Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ es semicontinua inferiormente, convexa y subanalítica global, entonces la suma epigráfica h también es subanalítica global.*

Ahora podemos enunciar el resultado principal que queremos presentar en este capítulo.

Teorema 29 (Desigualdad de Lojasiewicz).

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^c$ una función convexa subanalítica y semicontinua inferior con $\text{crit}(f) \neq \emptyset$. Para cada compacto K existe un $\theta \in [0, 1)$ y $C > 0$ tal que la función

$$|f - \text{mín } f|^\theta \leq C \cdot m_f(x) \quad \forall x \in K \cap \text{dom}(f)$$

Demostración. Sea $S := \text{crit}(f)$. Como f es convexa, para cada $a \in S$ y todo par $(x, x^*) \in \text{grafo}(\partial f)$ tenemos que

$$f(a) \geq f(x) + \langle x^*, a - x \rangle$$

lo que implica que $|f(x) - f(a)| \leq \|x^*\| \cdot \|x - a\|$ para todo $(x, x^*) \in \text{grafo}(\partial f)$. Tomando mínimo sobre a

$$|f(x) - \text{mín } f| \leq \|x^*\| \cdot d_S(x) \quad (\star)$$

donde d_S es la función distancia al conjunto S .

Por la proposición (6.2.1), la función g definida antes es subanalítica y continua, además de convexa. Usando (ii) de la proposición (4.3.6) y las propiedades que sabemos del conjunto de puntos críticos, llegamos a que S es cerrado y subanalítico. El contenido

$$\{|g - \text{mín } g| = 0\} \cap K \subseteq \{d_S = 0\} \cap K$$

usado con el lema de factorización de Lojasiewicz (6.1.2) para las funciones subanalíticas y continuas $|g - \text{mín } g|_K$ y $d_S|_K$, nos garantizan la existencia de constantes $r > 1$ y $c > 0$ tales que

$$c \cdot d_S(x)^r \leq |g(x) - \text{mín } g| \quad \forall x \in K$$

Por otra parte, las propiedades (i), (ii), (iii) de g implican que

$$|f(x) - \text{mín } f| \geq |g(x) - \text{mín } g| \quad \forall x \in \mathbb{R}^n$$

luego

$$d_S(x) \leq c^{-1/r} |f(x) - \min f|^{1/r} \quad (\spadesuit)$$

Juntando (\star) y (\spadesuit) , tenemos que para todo $x \in K$ y todo $(x, x^*) \in \text{grafo}(\partial f)$

$$|f(x) - \min f| \leq c^{-1/r} \|x^*\| \cdot |f(x) - \min f|^{1/r}$$

y tomando $\theta = 1 - r^{-1}$, la última desigualdad implica que

$$|f(x) - \min f|^\theta \leq c^{-1/r} m_f(x)$$

para todo $x \in K$. Que era lo que buscábamos. \square

6.3. Algunas ideas sobre su aplicación a sistemas dinámicos

En esta sección únicamente pretendo dar una idea de la posible aplicación actual a sistemas dinámicos de la desigualdad de Lojasiewicz que nosotros hemos tratado, por eso no lo veremos a fondo³. Empezamos asumiendo las hipótesis que necesitaremos

(H1) f es semicontinua inferior y convexa.

(H2) f es finita en algún lugar, esto es $\text{dom}(f) \neq \emptyset$ y acotada por abajo.

Partimos de un *sistema dinámico subgradiente*

$$\dot{x}(t) + \partial f(x(t)) \ni 0$$

Una *trayectoria* es por definición cualquier curva *absolutamente continua* $x : [0, T) \rightarrow \mathbb{R}^n$ satisfaciendo las condiciones

$$(\star) \quad \begin{cases} \dot{x}(t) + \partial f(x(t)) \ni 0 & \text{en c.t.p. de } (0, T) \\ \partial f(x(t)) \neq \emptyset & \forall x \in [0, T) \end{cases}$$

Recordemos que ser absolutamente continua significa ser diferenciable en c.t.p. y que puede ser determinada enteramente, salvo constante, por integración de su derivada clásica. Al igual que en la teoría ordinaria de EDO's, se dice *maximal* si su dominio no puede extenderse verificando las condiciones (\star) .

Por otra parte, las hipótesis (H1) y (H2) nos garantizan⁴, por cada punto $x_0 \in \mathbb{R}^n$, la existencia de una única trayectoria x con la condición inicial

$$(\star\star) \quad x(0) = x_0$$

Usando algunas de las propiedades de la subdiferencial y que $x(t)$ es diferenciable en c.t.p., puede verse que la función $x^* \mapsto \langle \dot{x}(t), x^* \rangle$ es constante en $\partial f(x(t))$. De aquí se deducen las siguientes igualdades (ver los detalles en [BDL07], colorarios 4.1

³ Para más detalles ver el artículo [BDL07]

⁴ Para ver los resultados concernientes a la existencia y unicidad de trayectorias en el caso convexo y convexo con perturbación Lipschitz ver [Bre73] el teorema 3.2 de la página 57 y la proposición 3.12 de la página 106

y 4.2): si x es una trayectoria maximal cumpliendo (\star) y $(\star\star)$, entonces para casi todo $t > 0$ real se tiene que

$$\|\dot{x}(t)\| = m_f(x(t)) \quad \text{y} \quad \frac{d}{dt}(f \circ x)(t) = -[m_f(x(t))]^2$$

Un análisis de este último resultado nos proporciona la siguiente característica de las trayectorias, en analogía con los sistemas de EDO's: si la trayectoria pasa por un punto crítico de f , permanece constante, es decir si $m_f(x(t_0)) = 0$, entonces $x(t) = x(t_0)$ para todo $t \geq t_0$ y tiene así longitud finita $\int_0^{t_0} \|\dot{x}(t)\| dt$.

Sin embargo, aunque la compacidad implica que trayectorias acotadas tengan al menos un punto de acumulación cuando $t \rightarrow +\infty$, éstas pueden no converger a uno de ellos y tener así longitud infinita. Ahora bien, si f satisface la desigualdad de Lojasiewicz, como en el caso analítico o el del teorema (29), ésto no puede pasar.

Teorema 30 (Longitud finita de trayectorias).

Sea f una función subanalítica satisfaciendo (H1) y (H2), entonces cualquier trayectoria maximal acotada de (\star) tiene longitud finita y converge a algún punto crítico de f .

La demostración usa la desigualdad de Lojasiewicz y puede verse en [BDL07], teorema 4.5, que es una adaptación de la prueba hecha en el artículo de Lojasiewicz para funciones analíticas, referencia [Loj83].

Finalizamos con un resultado que describe con más precisión el modo en que una trayectoria de un sistema dinámico subgradiente converge a un punto crítico⁵.

Se dice un número $\theta \in [0, 1)$ es un *exponente de Lojasiewicz* para una función f en un punto a de su dominio, si la desigualdad de Lojasiewicz se conserva alrededor de dicho a .

Teorema 31.

Sean f una función subanalítica cumpliendo (H1) y (H2), y $x(t)$ una trayectoria maximal acotada de (\star) . Entonces, $x(t)$ converge a algún punto crítico $a \in \mathbb{R}^n$ de f . Sea además $\theta \in [0, 1)$ un exponente de Lojasiewicz para este punto, entonces existen $k, k' > 0$ y $t_0 \geq 0$ tales que para todo $t \geq t_0$ las estimaciones siguientes ocurren

1. Si $\theta \in (1/2, 1)$, entonces $\|x(t) - a\| \leq k(t + 1)^{-\frac{1-\theta}{2\theta-1}}$
2. Si $\theta = 1/2$, entonces $\|x(t) - a\| \leq k \exp(-k't)$
3. Si $\theta \in [0, 1/2)$, entonces $x(t)$ converge en tiempo finito.

⁵ Ver la demostración en [BDL07], teorema 4.7

Apéndice A

Resultados sobre Teoría de la medida

Aquí aparecerán resultados ó ideas necesarias para entender el texto, que se suponen conocidas y en la mayoría de los casos sólo se recuerdan brevemente, y no están en el mismo para no estropear la línea argumentativa y claridad de éste.

Empecemos recordando lo que es un *espacio de medida*¹. Tal y como aparece en el texto es una terna (X, Σ, μ) donde X es un conjunto no vacío, Σ una σ -álgebra de conjuntos sobre X y μ una medida sobre el *espacio medible* (X, Σ) . La siguiente definición engloba todos estos conceptos.

Definición A.0.1. Sea X un conjunto no vacío. Una σ -álgebra, Σ , sobre X es un subconjunto de partes de X verificando que $X, \emptyset \in \Sigma$; que si $A, B \in \Sigma$, entonces $A \setminus B \in \Sigma$; y que

$$\text{si } \{A_k\}_{k \in \mathbb{N}} \subseteq \Sigma, \text{ entonces } \bigcap_{k \in \mathbb{N}} A_k, \bigcup_{k \in \mathbb{N}} A_k \in \Sigma$$

Mientras que una *medida* μ sobre (X, Σ) es una función de conjunto, $\mu : \Sigma \rightarrow [0, \infty]$ cumpliendo que $\mu(\emptyset) = 0$ y la σ -aditividad, esto es que $\mu(\bigsqcup_{k \in \mathbb{N}} A_k) = \sum_{k \in \mathbb{N}} \mu(A_k)$ donde la unión es disjunta.

Por otra parte, si (X, Σ) y (Y, Σ') son espacios medibles, una función $g : X \rightarrow Y$ se dice *medible* (o (Σ, Σ') -medible) si $f^{-1}(E) \in \Sigma$ para toda $E \in \Sigma'$.

Un ejemplo de medida que aparece en el texto es *la medida de contar*. La cual, denotada como ν , sobre un conjunto no vacío arbitrario X es la definida sobre la σ -álgebra partes de X por $\nu(A) = \text{card}(A)$ si A es finito, y ∞ en otro caso.

Definición A.0.2 (Función integrable). Sea $f : X \rightarrow (-\infty, +\infty)$ una función medible, se dice *integrable* si cumple que $\int |f| d\mu < +\infty$. Y se dice que f es integrable sobre $E \subseteq X$ si $f \chi_E$ es integrable, y se define entonces $\int_E f d\mu = \int f \chi_E d\mu$.

Definición A.0.3 (Espacios de funciones $L_p(\mu)$). Consideramos *funciones reales integrables* las de la clase de las funciones $f : X \rightarrow \mathbb{R}$ medibles tales que $\|f\|_1 := \int |f| d\mu < +\infty$, clase que se denotará como $L_1(\mu)$.

¹ En el libro de Cerdà, referencia [Cer00], aparece una muy buena exposición sobre Teoría de la medida, con todos sus detalles. Ver también [Sha05]

Su generalización son los espacios de funciones $L_p(\mu)$. Si $p \in [1, +\infty]$, su conjugado es $p' = p/(p-1)$. Se definen las normas

$$\|f\|_p := \left(\int |f|^p d\mu \right)^{1/p} \quad \text{y} \quad \|f\|_\infty := \inf\{M \geq 0 : |f| \leq M \text{ c.t.p.}\}$$

Así, $L_p(\mu) := \{f : \|f\|_p < \infty\}$.

Mostramos ahora los teoremas que se usan en algún momento en el texto. Empecemos por el de la convergencia dominada. Todos los resultados que de aquí en adelante no estén probados aparecen demostrados en [Cer00].

Teorema 32 (Teorema de la convergencia dominada).

Sean (X, Σ, μ) un espacio de medida, $(f_n)_{n=1}^\infty$ una sucesión de funciones medibles con valores reales, f otra función real medible y $0 \leq g \in L_1(\mu)$. Entonces, si

- i. (f_n) converge puntualmente a f en casi todo punto, y
- ii. $|f_n(x)| \leq g(x)$ en c.t.p. $x \in X$, $\forall n \in \mathbb{N}$

se tiene que $f_n, f \in L_1(\mu)$; aún más, el límite y la integral permutan. Es decir,

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu$$

Éste importante teorema de la teoría de integración de Lebesgue se puede obtener como consecuencia inmediata del **lema de Fatou** sin más que aplicar dicho lema a las funciones $g + f_n$ y $g - f_n$ que son no negativas en c.t.p. Es evidente que f_n y f son integrables, pues el valor absoluto de las primeras está mayorado por una función integrable y el límite puntual del valor absoluto de éstas es, esencialmente, el valor absoluto de la función f ; por lo que también estará mayorado por dicha función g integrable.

Ésto es esencialmente el camino de la demostración del teorema de la convergencia dominada. Para saber hacerla debemos conocer el ya nombrado lema de Fatou, que aunque tenga la categoría de lema no es de menospreciar, pues es muy útil, dice que:

Lema A.0.1 (Lema de Fatou). Si $(f_n)_{n=1}^\infty$ es una sucesión de funciones medibles no negativas, entonces

$$\int \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu$$

La prueba de éste lema se basa principalmente en considerar unas funciones especiales definidas a partir de las f_n y aplicar el **teorema de la convergencia monótona**, que asegura que:

Teorema 33 (Teorema de la convergencia monótona).

Sean $(f_n)_{n=1}^{\infty}$ una sucesión creciente de funciones medibles no negativas y f una función medible ≥ 0 . Si f es límite en c.t.p. de la sucesión (f_n) , entonces la integral de f permuta con el límite; es decir,

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu$$

Puesto que no es objetivo aquí desarrollar una base para la teoría de la medida, no demostraré los importantes resultados anteriores, sino que sólo los enuncio de manera rigurosa y ofrezco una pequeña orientación de como iría la prueba de los mismos. Sin embargo, no haré ésto con el teorema que sigue, pues como lo utilizaré explícitamente en la prueba de una propiedad de Γ , lo enunciaré y probaré formalmente.

Teorema 34 (Teorema de derivación bajo el signo integral).

Sea (X, Σ, μ) un espacio de medida e I un intervalo de \mathbb{R} . Sea además $f(\cdot, t)$ una función definida en X integrable para todo $t \in I$ y definimos

$$F(t) := \int f(x, t) \, d\mu(x)$$

Si $0 \leq g \in L_1(\mu)$, tenemos que

1. Si $f(x, \cdot)$ es continua en t_0 para cada $x \in X$ y $|f(x, t)| \leq g(x) \quad \forall x \in X \quad \forall t \in I$, entonces también es continua en t_0 la función F .
2. Si para todo $x \in X$ la función $f(x, \cdot)$ es derivable en todo $t \in I$ y además ocurre que $\left| \frac{\partial f}{\partial t}(x, t) \right| \leq g(x) \quad \forall x \in X \quad \forall t \in I$, entonces obtenemos que F es derivable; más aún, obtenemos que

$$F'(t) = \int \frac{\partial f}{\partial t}(x, t) \, d\mu(x)$$

Demostración. Antes de nada, nótese que gracias a las hipótesis la función F del enunciado está bien definida. Además, tampoco es necesario considerar dos funciones no negativas integrables distintas en lo referente a las hipótesis 1 y 2 del enunciado; pues si así fuera, digamos g_1 y g_2 para las hipótesis 1 y 2 respectivamente, bastaría considerar la función

$$g(x) := \max\{g_1(x), g_2(x)\} \geq 0 \quad \forall x \in X$$

que es trivialmente integrable. En efecto, sean los subconjuntos de X siguientes:

$$A_1 := \{x \in X : g(x) = g_1(x)\}$$

$$A_2 := \{x \in X : g(x) = g_2(x), \quad g(x) \neq g_1(x)\}$$

entonces es claro que $X = A_1 \cup A_2$ y $A_1 \cap A_2 = \emptyset$. Así,

$$\int_X g \, d\mu = \int_{A_1} g \, d\mu + \int_{A_2} g \, d\mu \leq \int_X g_1 \, d\mu + \int_X g_2 \, d\mu < \infty$$

Hagamos los casos por separado.

1. Para probar este apartado basta tomar una sucesión (t_k) de elementos de I que converja a t_0 y aplicar el teorema de la convergencia dominada de Lebesgue a la sucesión de funciones $f_k := f(\cdot, t_k)$, obteniendo así la continuidad de la función F . Podemos aplicar éste teorema pues estamos en las hipótesis del mismo por haberlo preparado así conscientemente. En efecto, $|f_k| = |f(\cdot, t_k)| \leq g$, y $f_k(x) = f(x, t_k) \rightarrow f(x, t_0) = f_0(x)$ para cada $x \in X$ cuando $k \rightarrow \infty$.
2. Aquí trabajaremos un poco más para conseguir lo buscado. Sea $t \in I$ y (h_m) una sucesión de números reales no nulos que converge a 0, hay que probar que

$$\lim_{m \rightarrow \infty} \left(\int \frac{f(x, t + h_m) - f(x, t)}{h_m} \, d\mu(x) \right) = \int \frac{\partial f}{\partial t}(x, t) \, d\mu(x)$$

pero ésto es fácil, pues gracias al teorema del valor medio conseguimos la siguiente acotación

$$\left| \frac{f(x, t + h_m) - f(x, t)}{h_m} \right| = \left| \frac{\partial f}{\partial t}(x, c_m) \right| \leq g(x)$$

donde $c_m \in (t, t + h_m)$. Luego, dado $t \in I$, ya podemos aplicar el teorema de la convergencia dominada a las funciones

$$f_m(x) := \frac{f(x, t + h_m) - f(x, t)}{h_m} \quad \forall x \in X$$

obteniendo lo buscado. □

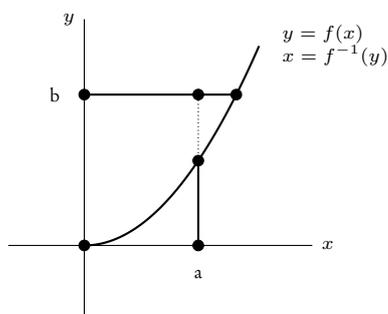
A.1. Desigualdades notables en espacios de medida generales

A.1.1. La desigualdad de Young

Antes de establecer ésta importante desigualdad que da lugar a otras dos también bien conocidas, la de Hölder y la de Minkovsky, necesitamos saber lo que entendemos por par conjugado de números reales. En nuestro caso, diremos que dos números reales p, q pertenecientes al intervalo $[1, \infty]$ forman un **par conjugado** si verifican que $1/p + 1/q = 1$ (entendiéndose que la expresión fraccionaria $1/\infty$ es igual a 0, y viceversa). Podemos ya decir que la **desigualdad de Young** para números reales afirma que *si a y b son dos números reales no negativos, entonces*

$$a \cdot b \leq \frac{a^p}{p} + \frac{b^q}{q}$$

donde p, q son un par conjugado de $(1, \infty)$. La igualdad se da si y sólo si se verifica que $a^p = b^q$.



La Des. de Young. Dibujo en general de la demostración, que representa lo siguiente: *la suma de las áreas de los dos triángulos curvilíneos excede a la del rectángulo de lados a y b , siendo igual cuando $b = f(a)$.*

Para demostrar esto aprovecharemos que ya vimos que la función exponencial es estrictamente convexa. En efecto, sean $p, q \in (1, \infty)$ un par conjugado. Sean $a, b > 0$ tales que $a^p \neq b^q$, entonces

$$\begin{aligned} a \cdot b &= e^{\log(a \cdot b)} = e^{(1/p) \log(a^p) + (1/q) \log(b^q)} \\ &< \frac{1}{p} e^{\log(a^p)} + \frac{1}{q} e^{\log(b^q)} \\ &= \frac{a^p}{p} + \frac{b^q}{q} \end{aligned}$$

Es evidente que si $ab = 0$ la conclusión se cumple.

Nótese que hemos probado, a la vez que la desigualdad, la forma contrapositiva de la implicación hacia la derecha en la equivalencia de la igualdad, sólo nos resta probar la implicación hacia la izquierda. Es decir, que si, con las notaciones del enunciado, $a^p = b^q$ entonces se da la igualdad. Pero ésto es fácil,

$$ab = a(b^q)^{1/q} = aa^{p/q} = a^{p/p} a^{p/q} = a^{p(1/p+1/q)} = a^p = \frac{a^p}{p} + \frac{b^q}{q}$$

Tenemos así demostrada la desigualdad de Young para reales, pero antes de dar la generalización establezcamos otra prueba más gráfica: al igual que antes, sean a y b reales no negativos, y $p, q \in (1, \infty)$ conjugados. Utilizando la función continua $y = f(x) = x^{p-1}$ en $[0, \infty)$, establecemos los siguientes casos

1. Supongamos primero que $a < b^{q-1}$. Sabemos que si $x = b^{q-1}$, entonces $y = b$. Ya que con pares conjugados se verifica que $p(q-1) = q$ y que $q(p-1) = p$. Así, el área del rectángulo cuyos lados tienen longitud a y b es menor o igual que la suma de las áreas S_1 y S_2 , entonces

$$ab \leq S_1 + S_2 = \int_0^a x^{p-1} dx + \int_0^b y^{q-1} dy = \frac{a^p}{p} + \frac{b^q}{q}$$

y evidentemente la igualdad se da cuando $b^{q-1} = a \Leftrightarrow a^p = b^{(q-1)p} = b^q$.

2. En el caso en que $a > b^{q-1}$, aunque el dibujo varíe levemente, se procede de la misma manera.

3. Es evidente que cuando $a = b^{q-1}$ se verifica la desigualdad, pues se da la igualdad, ya que nos encontraríamos en el caso en que la suma de las áreas S_1 y S_2 sería exactamente la del rectángulo de lados a y b .

Teorema 35 (Desigualdad de Young).

Sea $f : [0, \infty) \rightarrow [0, \infty)$ una función continua y estrictamente creciente que verifica que $f(0) = 0$ y existe $\lim_{x \rightarrow \infty} f(x) = \infty$. Entonces,

$$ab \leq \int_0^a f(x) dx + \int_0^b f^{-1}(x) dx$$

para cualesquiera $a, b \geq 0$. Además, la igualdad se da si y sólo si $b = f(a)$.

Demostración. Sean $a, b \in [0, \infty)$. Es evidente que si $a = 0$ ó $b = 0$, entonces la desigualdad del enunciado tiene lugar. Supongamos entonces que tanto a como b son distintos de cero.

Como f es estrictamente creciente, entonces es inyectiva, y por lo tanto tiene inversa. También ocurre, puesto que f es continua, $f(0) = 0$ y $\lim_{x \rightarrow \infty} f(x) = \infty$, que el rango de f es $[0, \infty)$, entonces f^{-1} es continua en su dominio (que es el rango de f). Queda así justificado que la desigualdad establecida en el enunciado tiene sentido. Es fácil convencerse de ella haciendo un dibujo. Definamos las dos funciones siguientes

$$F(x) := \int_0^x f(t) dt \quad \text{y} \quad G(x) := \int_0^x f^{-1}(t) dt$$

para todo $x > 0$. Del párrafo anterior se sigue que están bien definidas.

Como f^{-1} es estrictamente creciente (por serlo f); entonces, gracias a un argumento ya utilizado anteriormente, sabemos que G es estrictamente convexa. Así, el segmento que une cualesquiera dos puntos de la gráfica de G queda por encima (de manera estricta, a excepción de los extremos del segmento) de la propia gráfica de G . Por lo tanto, para cada $c \in (0, \infty)$, con $c \neq b$, tenemos la desigualdad (la veremos más tarde)

$$\begin{aligned} G(b) &> G(c) + G'(c)(b - c) \quad (\star) \\ &= G(c) + f^{-1}(c)(b - c) \end{aligned}$$

Si $c = f(a)$, $f(a) \neq b$

$$F(a) + G(b) > F(a) + G(f(a)) + a(b - f(a)) = ab$$

pues $F(a) + G(f(a)) = af(a)$, y tenemos así probada la desigualdad. Demostremos lo que falta, la igualdad anterior y la primera desigualdad estricta escrita arriba. Empecemos por la igualdad. Definamos, para una mayor claridad, la función

$$y(x) := F(x) + G(x) - xf(x) \quad \forall x \in [0, \infty)$$

Se trata de ver que y es derivable y con derivada nula en todo punto; entonces sería constante, pero como $y(0) = 0$, tendríamos que $y(x) = 0$ para todo $x \geq 0$. Así,

$F(x) + G(x) = xf(x)$ para todo $x \geq 0$. En efecto, sea $d \in [0, \infty)$. Tenemos que ver que

$$\lim_{h \rightarrow 0} \frac{y(d+h) - y(d)}{h} = 0$$

Para ello,

$$\begin{aligned} \left| \frac{1}{h}(y(d+h) - y(d)) \right| &= \left| \frac{1}{h} \left(\int_d^{d+h} f + \int_{f(d)}^{f(d+h)} f^{-1} + df(d) - (d+h)f(d+h) \right) \right| \\ &= \left| \frac{1}{h} \left(\int_d^{d+h} f + \int_{f(d)}^{f(d+h)} f^{-1} - d(f(d+h) - f(d)) - hf(d+h) \right) \right| \\ &= \left| \frac{1}{h} \left(\int_d^{d+h} f + \int_{f(d)}^{f(d+h)} f^{-1} - \int_{f(d)}^{f(d+h)} d - hf(d+h) \right) \right| \\ &= \left| \frac{\int_d^{d+h} f}{h} + \frac{\int_{f(d)}^{f(d+h)} (f^{-1} - d)}{h} - f(d+h) \right| \end{aligned}$$

Ahora bien, como f y $f^{-1} - d$ son continuas (pues d es constante) entonces, la última expresión de arriba tiende cuando $h \rightarrow 0$ a

$$f(d) + (f^{-1}(f(d)) - d) - f(d) = 0$$

Que era lo que buscábamos.

En cuanto a la desigualdad de la expresión (\star) , sea $c \in (0, \infty)$; como G es continua y derivable en $(0, \infty)$ (recuérdese la definición de G y que ésta es convexa), en particular, lo será en $[b, c]$ (supongamos que $c > b$, el caso en el que $c < b$ se hace de manera análoga). Así, existe un $\xi \in (b, c)$ tal que

$$G'(\xi) = \frac{G(c) - G(b)}{c - b}$$

Ahora bien, como $G' = f^{-1}$ y f^{-1} es estrictamente creciente y $c > \xi$, entonces $G'(c) > G'(\xi)$. Consecuentemente,

$$G'(c)(c - b) > G(c) - G(b)$$

O lo que es lo mismo

$$G(b) > G(c) + G'(c)(b - c)$$

Tenemos así establecida la desigualdad estricta que buscábamos.

Notesé que además de haber probado la desigualdad, también queda probada la implicación hacia la derecha en la equivalencia de la igualdad (ésto está hecho por contrarrecíproco: si $b \neq f(a)$, entonces no se da la igualdad). Para completar la demostración falta probar que si $b = f(a)$, entonces la igualdad ocurre. Pero ésto es fácil, simplemente teniendo en mente la interpretación geométrica de la integral

$$ab = af(a) = \int_0^a f + \int_0^{f(a)} f^{-1} = \int_0^a f + \int_0^b f^{-1} \quad \square$$

Ésta desigualdad de Young es muy útil pues tiene un montón de aplicaciones, por ejemplo para la obtención de otras desigualdades. Veamos las desigualdades de Hölder (o de Rogers-Hölder) y de Minkovsky obtenidas a partir de la desigualdad de Young.

Éstas se establecen para funciones de espacios $L^p(\mu)$, con la norma asociada $\|f\|_p = (\int_X |f|^p)^{1/p}$, en el marco de un espacio de medida arbitrario (X, Σ, μ) , donde X es un conjunto no vacío, Σ una sigma-álgebra de subconjuntos de X y μ una medida positiva sobre Σ . En lo que sigue, consideraremos que las funciones de $L^p(\mu)$ verifican que $|f(x)| < +\infty$ para todo $x \in X$ (pues las modificamos sobre un conjunto μ -nulo si fuera necesario). Empecemos por la de Hölder.

A.1.2. La desigualdad de Hölder

Teorema 36 (Desigualdad de Rogers-Hölder).

Sean $p, q \in (1, \infty)$ conjugados, $f \in L^p(\mu)$, $g \in L^q(\mu)$. Entonces, se tienen las siguientes desigualdades:

$$\left| \int_X fg \, d\mu \right| \leq \int_X |fg| \, d\mu$$

y

$$\int_X |fg| \, d\mu \leq \|f\|_p \|g\|_q$$

y

$$\left| \int_X fg \, d\mu \right| \leq \|f\|_p \|g\|_q$$

Cabe señalar que la desigualdad clave aquí, la comúnmente conocida como **desigualdad de Hölder**, es la número dos. Pues a partir de ella, de forma inmediata, sabemos que $fg \in L_1(\mu)$. Y así es claro que se verifica la primera desigualdad. Pues, de forma general, sabemos que si h es una función real integrable (lo que implica que su valor absoluto también lo sea), como $-|h| \leq h \leq |h|$, se obtiene que gracias a las propiedades básicas de la integral de Lebesgue se verifica $-\int |h| \leq \int h \leq \int |h|$; por lo que $|\int h \, d\mu| \leq \int |h| \, d\mu$.

Nótese que una vez que tenemos la primera y segunda desigualdad, la tercera sale como consecuencia de estas sin más que encadenar las desigualdades. Empecemos entonces a probar la parte fundamental de este teorema.

Demostración. Como ya dijimos, obtendremos la desigualdad de Hölder a partir de la de Young.

El caso trivial es cuando f ó g son cero en μ -c.t.p. (μ -casi-todo-punto), esto es que el conjunto de puntos en el cual f ó g no son nulas es un conjunto μ -nulo; es decir, está contenido en un conjunto medible y con medida cero. Más claramente, tomemos f como ejemplo, sea

$$C := \{x \in X | f(x) \neq 0\}, \exists B \in \Sigma \text{ con } C \subseteq B \text{ tal que } \mu(B) = 0$$

Si el espacio de medida en el que nos encontramos trabajando es completo, entonces todo conjunto μ -nulo es medible.

Una vez desechado el caso trivial continuemos con el grueso de la demostración. Llamo

$$A := \left(\int |f|^p \right)^{1/p} \text{ y } B := \left(\int |g|^q \right)^{1/q}$$

Como $f \in L^p$, $g \in L^q$, y hemos desconsiderado el caso trivial, entonces $0 < A < \infty$, $0 < B < \infty$. En efecto, pongamos que $A = 0$, de manera equivalente tendríamos que $|f|^p = 0$ en μ -c.t.p., que es lo mismo que $f = 0$ en μ -c.t.p., pero esto no puede ser, pues hemos quitado el caso trivial.

Sean ahora

$$a_x = \frac{|f(x)|}{A}, \text{ y } b_x = \frac{|g(x)|}{B}$$

los cuales están bien definidos, y son no negativos en todo X . Podemos entonces aplicar la desigualdad de Young

$$\frac{|f(x)|}{A} \cdot \frac{|g(x)|}{B} = a_x \cdot b_x \leq \frac{a_x^p}{p} + \frac{b_x^q}{q} = \frac{1}{p} \cdot \frac{|f(x)|^p}{A^p} + \frac{1}{q} \cdot \frac{|g(x)|^q}{B^q}$$

para todo $x \in X$. Luego, tomando integrales

$$\frac{1}{\|f\|_p \|g\|_q} \cdot \int_X |fg| \, d\mu \leq \frac{1}{p} + \frac{1}{q} = 1$$

que es lo mismo que

$$\int_X |fg| \, d\mu \leq \|f\|_p \|g\|_q$$

quedando ya desmotrada la desigualdad de Hölder. \square

Observación. ¿Cuándo se da la igualdad en la desigualdad de Hölder? Ésta ocurre si y sólo si existen dos constantes α y β no nulas al mismo tiempo y positivas tales que $\alpha|f|^p = \beta|g|^q$ para μ -c.t.p. En efecto, con las notaciones de arriba, durante la demostración de la desigualdad de Hölder aplicábamos en cierto momento la desigualdad de Young a unos ciertos valores reales no negativos dependientes de la variable x , estos eran $a_x = \frac{|f(x)|}{A}$ y $b_x = \frac{|g(x)|}{B}$. Puesto que para esta última desigualdad ya conocemos cuando se da la igualdad, en dicho momento la igualdad se dará si y sólo si $|f(x)|^p \|g\|_q^q = |g(x)|^q \|f\|_p^p$ para casi todo $x \in X$. Recíprocamente, si sabemos que $|f|^p = M|g|^q$ en μ -c.t.p., una sustitución en la desigualdad de Hölder nos proporcionará la igualdad.

A continuación veremos la otra desigualdad anunciada: la de Minkovsky, que se obtiene a partir de la de Hölder y por lo tanto a partir de la de Young. Así:

$$\text{YOUNG} \implies \text{HÖLDER} \implies \text{MINKOVSKY}$$

A.1.3. La desigualdad de Minkovsky

Teorema 37 (Desigualdad de Minkovsky).

Sean $1 \leq p < \infty$, y $f, g \in L^p(\mu)$. Entonces se verifica que

$$\left(\int_X |f + g|^p d\mu \right)^{1/p} \leq \left(\int_X |f|^p d\mu \right)^{1/p} + \left(\int_X |g|^p d\mu \right)^{1/p}$$

equivalentemente, escrito de otra manera

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Demostración. Es evidente que, si $p = 1$ entonces la desigualdad sale de integrar la ya conocida: $|f + g| \leq |f| + |g|$. Supongamos ahora que $1 < p < \infty$. Para que tenga sentido la desigualdad debe ocurrir que $f + g \in L^p(\mu)$. Pero sencillamente como

$$|f + g|^p \leq (|f| + |g|)^p \leq (2 \sup\{|f|, |g|\})^p \leq 2^p(|f|^p + |g|^p)$$

entonces

$$\|f + g\|_p^p \leq 2^p(\|f\|_p^p + \|g\|_p^p) < +\infty$$

y así conseguimos que $f + g \in L^p(\mu)$.

Por otra parte, si llamamos q al conjugado de p , también sabemos que la función $|f + g|^{p-1}$ pertenece al espacio de funciones integrables, $L^q(\mu)$. En efecto, como p y q son conjugados, entonces $(|f + g|^{p-1})^q = |f + g|^{q(p-1)} = |f + g|^p$, luego $\| |f + g|^{p-1} \|_q^q < \infty$.

Partiendo de la desigualdad

$$|f + g|^p = |f + g||f + g|^{p-1} \leq |f||f + g|^{p-1} + |g||f + g|^{p-1}$$

e integrando, obtenemos que

$$\int_X |f + g|^p d\mu \leq \int_X |f||f + g|^{p-1} d\mu + \int_X |g||f + g|^{p-1} d\mu$$

Nótese que puesto que $|f| \in L^p$ y $|f + g|^{p-1} \in L^q$, entonces tal y como probamos durante la demostración de la desigualdad de Hölder, su producto, $|f||f + g|^{p-1}$, pertenecerá a L^1 y así tendrá sentido lo escrito arriba.

Recopilando toda la información que tenemos hasta el momento nos damos cuenta de que estamos en las condiciones adecuadas para poder aplicar la desigualdad de Hölder a las funciones $|f|$ y $|f + g|^{p-1}$. Quedándonos entonces la última expresión de arriba menor ó igual que

$$\begin{aligned} & \left(\int_X |f|^p d\mu \right)^{1/p} \left(\int_X |f + g|^{q(p-1)} d\mu \right)^{1/q} + \\ & + \left(\int_X |g|^p d\mu \right)^{1/p} \left(\int_X |f + g|^{q(p-1)} d\mu \right)^{1/q} \end{aligned}$$

Ahora bien, si se da el caso de que $\int_X |f + g|^{q(p-1)} d\mu = \int_X |f + g|^p d\mu = 0$, entonces la desigualdad de Minkovsky es trivial. Supongamos entonces que dicha integral no es

nula; además tampoco puede ser infinita, pues ya hemos visto que $|f+g|^{p-1} \in L^q(\mu)$. Podemos dividir entonces la última expresión por el real

$$\gamma := \left(\int_X |f+g|^{q(p-1)} d\mu \right)^{1/q} = \left(\int_X |f+g|^p d\mu \right)^{1/q} > 0$$

quedándonos

$$\frac{\int_X |f+g|^p d\mu}{\gamma} \leq \left(\int_X |f|^p d\mu \right)^{1/p} + \left(\int_X |g|^p d\mu \right)^{1/p}$$

pero como

$$\gamma^{-1} \cdot \int_X |f+g|^p d\mu = \left(\int_X |f+g|^p d\mu \right)^{1-1/q} = \left(\int_X |f+g|^p d\mu \right)^{1/p}$$

hemos conseguido la desigualdad buscada. \square

Observación. ¿Cuándo se da la igualdad en la desigualdad de Minkovsky? Aquí la respuesta dependerá del caso en el que nos encontremos.

1. Si $p = 1$ se da la igualdad si y sólo si $|f+g| = |f| + |g|$ en μ -c.t.p., lo que equivale a que exista una función medible h , $h \geq 0$ en μ -c.t.p., tal que $g(x) = h(x)f(x)$ siempre que sea $f(x)g(x) \neq 0$.
2. Si $1 < p < \infty$, la igualdad se da si y sólo si existen dos constantes α y β no simultáneamente nulas tales que $\alpha f = \beta g$ en μ -c.t.p.

Empecemos con el caso más sencillo, $p = 1$. Aquí la desigualdad equivale a que la integral $\int_X (|f| + |g| - |f+g|) d\mu \geq 0$. Pero esta integral es nula, al ser la función que se integra no negativa, si y sólo si dicha función es nula en casi todo punto. O sea, $|f+g| = |f| + |g|$ en μ -c.t.p. Tenemos así la equivalencia que nos da la igualdad. En el caso en que $1 < p < \infty$, una sustitución muestra la suficiencia. Sencillamente, supongamos que $\alpha, \beta \neq 0$ (si alguna lo fuera la función que no va asociada a la nula valdría 0 en casi todo punto y la igualdad sería trivial)

$$\left(\int_X |f + \frac{\alpha}{\beta} f|^p \right)^{1/p} = \left(\frac{\alpha}{\beta} + 1 \right) \left(\int_X |f|^p \right)^{1/p} = \left(\int_X |f|^p \right)^{1/p} + \left(\int_X \left| \frac{\alpha}{\beta} f \right|^p \right)^{1/p}$$

Para la necesidad es necesario trabajar un poco más. Si se tiene la igualdad en la desigualdad de Minkovsky, entonces en su prueba deben darse las igualdades siguientes

$$\int_X |f| |f+g|^{p-1} d\mu = \|f\|_p \| |f+g|^{p-1} \|_q$$

$$\int_X |g| |f+g|^{p-1} d\mu = \|g\|_p \| |f+g|^{p-1} \|_q$$

Por lo que en virtud de la desigualdad de Hölder, salvo conjunto de μ -medida cero, deben existir $\alpha_1, \beta_1, \alpha_2, \beta_2 \geq 0$, no nulos a la vez cada dupla con el mismo subíndice, tales que

$$\alpha_1 |f| = \beta_1 |f+g|^{p-1} \quad (1)$$

$$\alpha_2 |g| = \beta_2 |f+g|^{p-1} \quad (2)$$

Nótese que si, por ejemplo, $\beta_2 = 0$, entonces $g = 0$ en μ -c.t.p., y se cumpliría trivialmente la condición a verificar. Ya que bastaría tomar $\alpha = 0$ y β una constante cualquiera mayor que 0. Así, una vez desechados los casos en los que algún peso de las igualdades (1) y (2) valen 0, podemos llevar a cabo el razonamiento siguiente

$$\frac{\alpha_2}{\beta_2}|g| = |f + g|^{p-1} \Rightarrow \alpha_1|f| = \beta_1 \cdot \frac{\alpha_2}{\beta_2}|g|$$

y tomando como α el valor $\alpha_1\beta_2$ y como β el valor $\beta_1\alpha_2$, tenemos *casi* la condición buscada (aquí todas las igualdades que involucren a funciones, aunque no aparezca explícitamente escrito, son en μ -c.t.p.). Pues la que obtenemos realmente es

$$\alpha|f| = \beta|g| \quad \text{en } \mu\text{-c.t.p.}$$

Si conseguimos eliminar los valores absolutos ya habríamos acabado de probar la necesidad. Para ello veamos que el signo de las dos funciones es el mismo en casi todo punto. Reescribamos la igualdad de arriba como $|g| = \lambda|f|$ en c.t.p., puesto que suponemos que se da la igualdad en la desigualdad de Minkovsky, se obtiene

$$\left(\int_X |s(f)| |f| + \lambda s(g) |f|^p d\mu \right)^{1/p} = (1 + \lambda) \left(\int_X |f|^p d\mu \right)^{1/p}$$

con $f = s(f)|f|$ en c.t.p. y $|s(f)| = 1$ en c.t.p., y $|s(g)| = 1$ en c.t.p. de manera análoga. Tendremos así que $|s(f) + \lambda s(g)| = |1 + \lambda s(g)/s(f)| = 1 + \lambda$ en c.t.p., y se deduce ahora fácilmente que $s(g)/s(f) = 1$ en c.t.p., que era lo que estábamos buscando. Concluimos que $\alpha f = \beta g$ en c.t.p.

Para acabar con estas dos famosas desigualdades demos un ejemplo concreto. Se trata del caso en el que el espacio de medida (X, Σ, ν) es un espacio de medida asociado a la *medida de contar* sobre un conjunto finito X , donde

$$\nu : P(X) \longrightarrow \mathbb{R}^+ \quad \text{tal que} \quad A \subseteq X \Rightarrow \nu(A) = \text{card}(A)$$

es la medida de contar sobre X y $\text{card}(A)$ representa el cardinal de A . Obtenemos las **formas discretas** de las desigualdades antes mencionadas. **Desigualdad de Rogers-Hölder**,

$$\left| \sum_{k=1}^n \eta_k \xi_k \right| \leq \left(\sum_{k=1}^n |\xi_k|^p \right)^{1/p} \cdot \left(\sum_{k=1}^n |\eta_k|^q \right)^{1/q}$$

donde $\eta_k, \xi_k \in \mathbb{R} \quad \forall k = 1, \dots, n$ y $1/p + 1/q = 1$ con $p, q \in (1, \infty)$

Desigualdad de Minkovsky,

$$\left(\sum_{k=1}^n |\eta_k + \xi_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^n |\eta_k|^p \right)^{1/p} + \left(\sum_{k=1}^n |\xi_k|^p \right)^{1/p}$$

donde $\eta_k, \xi_k \in \mathbb{R} \quad \forall k = 1, \dots, n$ y $p \in (1, \infty)$

Desigualdad de Cauchy-Schwarz, se obtiene haciendo $p = 2$ en la de Hölder

$$\left| \sum_{k=1}^n \eta_k \xi_k \right| \leq \left(\sum_{k=1}^n |\xi_k|^2 \right)^{1/2} \cdot \left(\sum_{k=1}^n |\eta_k|^2 \right)^{1/2}$$

donde $\eta_k, \xi_k \in \mathbb{R} \quad \forall k = 1, \dots, n$.

Bibliografía

- [BDL07] Jerome Bolte, Aris Danilidis, y Adrian Lewis. The lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. *Society for Industrial and Applied Mathematics*, 17(4):1205–1223, 2007.
- [Bie88] P. Bierstone, & E. Milman. Semianalytic and subanalytic sets. *IHES Publ. Math.*, (67):5–42, 1988.
- [Bre73] H. Brezis. *Opérateurs maximaux monotones et semi-groupes de contraction dans des espaces de Hilbert*. Número 5 de North-Holland Math. Stud. North-Holland, 1973. Amsterdam.
- [Loj83] S. Lojasiewicz. *Sur les trajectoires du gradient d'une fonction analytique*. Seminari di Geometria, 115-117, 1983. Bologna.
- [BZ88] Yu. D. Burago & V. A. Zalgaller. *Geometric Inequalities*. Número 285 de A Series of Comprehensive Studies in Mathematics. Springer-Verlag, 1 edición, 1988.
- [Cer00] Joan Cerdá. *Análisis Real*. Número 23 de Coleccion UB. Edicions Universitat de Barcelona, 2 edición, 2000.
- [Con] John B. Conway. *Functions of one complex variable I*, volumen 11 de *Graduate Texts in Mathematics*. Springer, 2 edición.
- [Jab14] JabRef Development Team. *JabRef*, Junio 2014.
- [Mil96] L. Van Den Dries & C. Miller. Geometric categories and o-minimal structures. *Duke Math. J.*, (84):497–540, 1996.
- [NP06] Constantin Niculescu & Lars-Erik Persson. *Convex functions and their applications. A contemporary approach*. Número 23 de CMS books in Mathematics. Springer, 1 edición, 2006.
- [RW98] R. Tyrrell Rockafellar & Roger J-B Wets. *Variational Analysis*. Número 317 de A series of Comprehensive Studies in Mathematics. Springer, 1 edición, 1998.
- [Sha05] Elias M. Stein & Rami Shakarchi. *Real Analysis: measure theory, integration, and Hilbert Spaces*, volumen III de *Princeton Lectures in Analysis*. Princeton University Press, 1 edición, 2005.
- [Web94] Roger Webster. *Convexity*. Oxford University Press, New York, 1 edición, 1994.

*Nunca un Siempre fue tan
Corto*