

ENTREGA A

Christian Limbert Paredes Aguilera

Latex Source: <https://n9.c1/ua8c3>

CAPÍTULO 2

Ejercicio 1. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ dos funciones convexas cuyo dominio es \mathbb{R}^n . Se define su convolución ínfima (infimal convolution) como

$$h(x) = (f \diamond g)(x) := \inf \{f(y) + g(z) : y + z = x\}.$$

Demuestra que el epígrafo de h es un conjunto convexo.

Demostración.- Para demostrar que el epígrafo de h es un conjunto convexo, necesitamos mostrar que para cualquier par de puntos en el epígrafo de h , la línea que los conecta también está en el epígrafo de h .

Dado que f y g son funciones convexas, sus epígrafos son conjuntos convexas¹. Esto significa que para cualquier par de puntos en el epígrafo de f o g , la línea que los conecta también está en el epígrafo de f o g .

Comenzamos considerando dos puntos arbitrarios (x_1, t_1) y (x_2, t_2) en el epígrafo² de h . Entonces, si (x_1, t_1) y (x_2, t_2) están en el epígrafo de h , esto significa que t_1 y t_2 son mayores o iguales que el valor de h en x_1 y x_2 , respectivamente. Es decir,

$$t_1 \geq h(x_1) = \inf \{f(y) + g(z) : y + z = x_1\}.$$

$$t_2 \geq h(x_2) = \inf \{f(y) + g(z) : y + z = x_2\}.$$

Además, h se define como la convolución ínfima de f y g , es decir,

$$h(x) = (f \diamond g)(x) := \inf \{f(y) + g(z) : y + z = x\}.$$

Esto significa que para cada punto x , buscamos el valor mínimo de

$$f(y) + g(z) \text{ donde } y + z = x.$$

Entonces, si (x_1, t_1) y (x_2, t_2) están en el epígrafo de h , esto significa que existen y_1, z_1, y_2, z_2 tal que

$$y_1 + z_1 = x_1, \quad y_2 + z_2 = x_2, \quad f(y_1) + g(z_1) \leq t_1 \quad \text{y} \quad f(y_2) + g(z_2) \leq t_2.$$

¹Teorema: f es convexa sii $\text{Epi}(f)$ es convexo.

²Recordemos que el epígrafo de una función es el conjunto de puntos que están por encima del gráfico de la función.

En otras palabras, podemos encontrar puntos y_1, z_1, y_2, z_2 tal que la suma de sus imágenes bajo f y g , respectivamente, es menor o igual que t_1 y t_2 . (Este paso será crucial para la demostración porque nos permite relacionar los puntos en el epígrafo de h con los puntos en los epígrafos de f y g , lo cual es necesario para demostrar que el epígrafo de h es un conjunto convexo).

Ahora consideremos un punto (x, t) en la línea que conecta (x_1, t_1) y (x_2, t_2) . Podemos escribir (x, t) como una combinación convexa de (x_1, t_1) y (x_2, t_2) . Es decir,

$$x = \lambda x_1 + (1 - \lambda)x_2 \quad y \quad t = \lambda t_1 + (1 - \lambda)t_2$$

para algún $\lambda \in [0, 1]$.

Dado que los epígrafos de f y g son conjuntos convexos, existen y y z tal que

$$y = \lambda y_1 + (1 - \lambda)y_2, \quad f(y) \leq \lambda f(y_1) + (1 - \lambda)f(y_2).$$

$$z = \lambda z_1 + (1 - \lambda)z_2, \quad g(z) \leq \lambda g(z_1) + (1 - \lambda)g(z_2).$$

Entonces, tenemos que

$$y + z = x \quad y \quad f(y) + g(z) \leq t.$$

Por lo tanto, (x, t) está en el epígrafo de h . Así, el epígrafo de h es un conjunto convexo. ■

Ejercicio 2. Un punto extremal de un conjunto convexo C es aquel punto $x \in C$ que cumple que si existen y, z en C tales que $x = (y + z)/2$ entonces $y = z$. Demuestra que si $C \subseteq \mathbb{R}^n$ es convexo, cerrado y acotado. Entonces, C es la envoltura convexa del conjunto $E(C)$ de sus puntos extremales. [Indicación: Si tienes problemas para hacer la demostración, puedes asumir que $E(C)$ es no vacío].

Demostración.- Supongamos que existe un punto $x \in C$ que no está en $\text{co}(E(C))$ ³. En otras palabras, si x no está en $\text{co}(E(C))$ estos dos conjuntos son disjuntos. Así, por el teorema de separación de Hiperplanos⁴, podemos encontrar un hiperplano, que separe a x de $\text{co}(E(C))$.

Como C es cerrado (contiene todos sus puntos límite)⁵, y como C es acotado (tiene un tamaño finito). Estas dos propiedades garantizan que C debe tener al menos un punto extremal, es decir, $E(C)$ no es vacío.

Dado que $E(C)$ no es vacío y C es cerrado⁵ y acotado, debe existir al menos un punto extremal en C que esté en el mismo lado del hiperplano que x . Llamamos a este punto y .

Según la definición de un punto extremal, no puede existir un hiperplano que separe C e y . Sin embargo, encontramos un hiperplano que separa x e y . Puesto que, x está en C , esto significa que hemos encontrado un hiperplano que separa C e y , lo cual contradice la definición de un punto extremal. Por lo tanto, nuestra suposición inicial de que existe un punto x en C que no está en $\text{co}(E(C))$ debe ser falsa.

Esto implica que todos los puntos en C deben estar en $\text{co}(E(C))$, y por lo tanto, C es igual a $\text{co}(E(C))$. Lo que completa la demostración. ■

³Se llama **envoltura convexa** de S al menor conjunto convexo que contienen a S , denotado por $\text{co}(S)$. También es equivalente a decir que: $\text{co}(S) = \{\text{Combinación convexa de puntos de } S\}$.

⁴El **teorema de separación de hiperplanos** establece que si tenemos dos conjuntos convexos disjuntos, entonces existe un hiperplano que los separa. Es decir, sean A y B conjuntos convexos con $A \cap B = \emptyset$ y $A^i \neq \emptyset$. Entonces, existe $f \in E^*$ lineal, tal que $f(a^i) < \inf \{f(b), b \in B\}$.

⁵Decimos que A es **cerrado** cuando $A = \overline{A}$ donde $\{x \in \mathbb{R}^n : x \text{ está en el cierre de } A\}$.

Ejercicio 4. Prueba que si A es un conjunto cerrado de \mathbb{R}^n , con interior no vacío, y tiene un hiperplano soporte en cada punto de su frontera, entonces A es convexo.

Demostración.- Sea A es un conjunto cerrado con interior no vacío y tiene un hiperplano soporte en cada punto de su frontera ⁶. Esta función f representa un hiperplano soporte en el punto x_0 .

Ahora, sea A no es convexo. De donde, supondremos dos puntos $x, y \in A$ tal que el segmento de línea que los une no está completamente en A . Es decir, existe un $t \in [0, 1]$ tal que

$$tx + (1 - t)y \notin A.$$

(Observemos que este punto $tx + (1 - t)y$ está en la línea entre x e y , pero no está en A , lo cual es una contradicción a la convexidad).

Luego, dado que $tx + (1 - t)y \notin A$, debe existir un hiperplano soporte⁶ en $tx + (1 - t)y$. Este hiperplano divide \mathbb{R}^n en dos semiespacios cerrados, uno de los cuales contiene a A . Este hiperplano está representado por una función $f \in E^\#$ tal que

$$f(tx + (1 - t)y) = \sup \{f(a) : a \in A\} \geq f(a), \forall a \in \bar{A}.$$

Sin embargo, tanto x como y están en A , por lo que deben estar en el mismo semiespacio que A . Esto se debe a que la función f que representa el hiperplano soporte satisface

$$f(x) \geq f(a) \quad y \quad f(y) \geq f(a)$$

para todo $a \in \bar{A}$. Pero, todos los puntos en el segmento que une x e y también deben estar en el mismo semiespacio que A , ya que este segmento es un subconjunto del hiperplano.

Esto contradice nuestra suposición de que

$$tx + (1 - t)y \notin A.$$

Por lo tanto, la suposición inicial, de que A no es convexo, debe ser incorrecta. Así, si A es un conjunto cerrado con interior no vacío y tiene un hiperplano soporte en cada punto de su frontera, entonces A debe ser convexo. Esta demostración es válida independientemente de si x y y están en el interior o en la frontera de A ⁷. ■

⁶**Hiperplano soporte** significa que para cada $x_0 \in \partial A$, existe una función $f \in E^\#$ tal que $f(x_0) = \sup \{f(a) : a \in A\} \geq f(a), \forall a \in \bar{A}$.

⁷En la demostración, se supone que x e y son puntos arbitrarios en A . Esto significa que pueden ser cualquier punto en A , ya sea en el interior de A o en la frontera de A . La demostración no depende de la ubicación específica de x e y dentro de A .

Además, la definición de un hiperplano soporte implica que para cada punto en la frontera de A , existe un hiperplano que "soporta" A en ese punto. Esto significa que el hiperplano no intersecta A en ese punto y todos los puntos de A están en un lado del hiperplano. Esta propiedad es válida independientemente de si estamos considerando un punto en el interior de A o en la frontera de A .

Por lo tanto, la demostración de que A es convexo si tiene un hiperplano soporte en cada punto de su frontera es válida independientemente de si x e y están en el interior o en la frontera de A .

CAPÍTULO 3

Ejercicio 1. Sea f convexa y x_0 en el interior del dominio de f . Demuestra que:

- (a) El subdiferencial $\partial f(x_0)$ es distinto del vacío, convexo y cerrado.

Demostración.-

- **No vacío:** Por definición, si f es convexa y x_0 está en el interior del dominio de f , entonces existe al menos un subgradiente⁸ en x_0 . Esto significa que el conjunto subdiferencial⁹ $\partial f(x_0)$ no puede ser vacío.
- **Convexo:** Supongamos que $g, h \in \partial f(x_0)$ y $\theta \in [0, 1]$. Demostraremos que

$$\theta g + (1 - \theta)h \in \partial f(x_0).$$

Por la definición de subgradiente⁸, tenemos que:

$$f(x) \geq f(x_0) + g^T(x - x_0)$$

y

$$f(x) \geq f(x_0) + h^T(x - x_0)$$

Multiplicando la primera desigualdad por θ y la segunda por $(1 - \theta)$ y sumándolas, obtenemos:

$$f(x) \geq f(x_0) + (\theta g + (1 - \theta)h)^T(x - x_0)$$

Esto demuestra que $\theta g + (1 - \theta)h$ es un subgradiente de f en x_0 , por lo que $\partial f(x_0)$ es convexo.

- **Cerrado:** Para demostrar que $\partial f(x_0)$ es cerrado, necesitamos demostrar que el límite de cualquier secuencia convergente de subgradientes también pertenece a $\partial f(x_0)$. Supongamos que tenemos una secuencia $\{g_k\}$ de subgradientes tal que $g_k \rightarrow g$ cuando $k \rightarrow \infty$ ¹⁰. Por la definición de subgradiente⁸, para cada k tenemos que:

$$f(x) \geq f(x_0) + g_k^T(x - x_0)$$

Tomando el límite cuando $k \rightarrow \infty$, obtenemos:

$$f(x) \geq f(x_0) + g^T(x - x_0).$$

Lo que estamos diciendo es que si g_k se acerca a g cuando k tiende a infinito, entonces la desigualdad también debe mantenerse en el límite cuando k tiende a infinito.

Esto demuestra que g es un subgradiente de f en x_0 , por lo que $\partial f(x_0)$ es cerrado.

■

⁸Si f es convexa y x_0 en $\text{int}(\text{dom}(f))$, encontramos el subestimador afín del resultado anterior. Al vector que define unívocamente este subestimador, se le llama **subgradiente** de f en el punto. Esto significa que g satisface la desigualdad $f(x) \geq f(x_0) + g^T(x - x_0)$, para todo x en el dominio de f . Esta desigualdad dice que la función f está por encima de la línea que pasa por el punto $(x_0, f(x_0))$ con pendiente g , lo que significa que esta línea subestima a f en x_0 . Por lo tanto, g se llama un subgradiente de f en x_0 .

⁹Se llama **subdiferencial** de f en x_0 al conjunto de todos los subgradientes de f en x_0 : $\partial f(x_0)$.

¹⁰Esto significa que para cada número positivo ϵ , existe un número entero N tal que para todo $k > N$, la distancia entre g_k y g es menor que ϵ . En otras palabras, a medida que k se hace más y más grande, g_k se acerca cada vez más a g .

- (b) El subdiferencial $\partial f(x_0)$ es un conjunto unipuntual si y sólo si f es diferenciable en x_0 . En dicho caso $\partial f(x_0) = \nabla f(x_0)$.

Demostración.- \Leftarrow] Si f es diferenciable en x_0 , entonces existe un único vector, el gradiente de f en x_0 , que satisface la definición de subgradiente. Esto se debe a que la función tangente lineal a f en x_0 proporciona el mejor subestimador afín de f en x_0 . Por lo tanto, el subdiferencial $\partial f(x_0)$ es un conjunto unipuntual, y $\partial f(x_0) = \nabla f(x_0)$.

Para demostrar esto de manera algebraica, consideremos la definición de subgradiente⁸. Un vector g es un subgradiente de f en x_0 si satisface la desigualdad

$$f(x) \geq f(x_0) + g^T(x - x_0).$$

para todo x en el dominio de f . Si f es diferenciable en x_0 , entonces su gradiente

$$\nabla f(x_0)$$

satisface esta desigualdad, porque la función tangente lineal

$$f(x_0) + \nabla f(x_0)^T(x - x_0)$$

es la mejor¹¹ aproximación lineal¹² a f en x_0 . Por lo tanto, si f es diferenciable en x_0 , entonces el gradiente de f en x_0 es el único vector que satisface la definición de subgradiente. En otras palabras, el gradiente de f en x_0 es el único subgradiente de f en x_0 . Por lo que, el subdiferencial de f en x_0 , que es el conjunto de todos los subgradietes de f en x_0 , es un conjunto unipuntual que contiene solo el gradiente de f en x_0 . Así,

$$\partial f(x_0) = \nabla f(x_0).$$

\Rightarrow] Si el subdiferencial $\partial f(x_0)$ es un conjunto unipuntual, entonces existe un único subgradiente en x_0 . Esto implica que la función tangente lineal a f en x_0 es única, lo cual es posible sólo si f es diferenciable en x_0 .

Para demostrar esto de manera algebraica, supongamos que $\partial f(x_0)$ es un conjunto unipuntual, es decir,

$$\partial f(x_0) = \{g\}$$

para algún vector g . Por la definición de subgradiente, sabemos que

$$f(x) \geq f(x_0) + g^T(x - x_0)$$

para todo x en el dominio de f . Ahora, si f no fuera diferenciable en x_0 , entonces existiría otro vector h tal que

$$f(x) \geq f(x_0) + h^T(x - x_0)$$

para todo x en el dominio de f . Pero esto contradiría nuestra suposición de que $\partial f(x_0)$ es un conjunto unipuntual. Por lo tanto, f debe ser diferenciable en x_0 , y su gradiente

¹¹"Mejor" aquí significa que la función tangente lineal es la función lineal que más se acerca a f en el punto x_0 .

¹²Cuando decimos que una función f es diferenciable en un punto x_0 , esto significa que la función tiene una derivada en ese punto. En términos geométricos, esto significa que existe una línea tangente a la gráfica de la función en el punto $(x_0, f(x_0))$. Esta línea tangente es la mejor aproximación lineal a la función cerca del punto x_0 , en el sentido de que la distancia entre la función y la línea tangente es mínima en comparación con cualquier otra línea que pase por el punto $(x_0, f(x_0))$.

La ecuación de esta línea tangente es $f(x_0) + \nabla f(x_0)^T(x - x_0)$, donde $\nabla f(x_0)$ es el gradiente de f en x_0 . Esta ecuación nos da el valor de la línea tangente para cualquier valor de x cerca de x_0 .

De esta manera, si f es diferenciable en x_0 , entonces tiene una única mejor aproximación lineal en x_0 , que es la función tangente lineal. Esto es lo que nos permite definir el subgradiente y el subdiferencial de f en x_0 .

en x_0 es el único subgradiente, es decir, $\nabla f(x_0) = g$.

Así, demostramos que f es diferenciable en x_0 si y sólo si el subdiferencial $\partial f(x_0)$ es un conjunto unipuntual. En dicho caso, $\partial f(x_0) = \nabla f(x_0)$. ■

- (c) Si f y g son dos funciones convexas, demuestra que $\partial(f+g)(x) = \partial f(x) + \partial g(x)$. ¿Se tiene la igualdad en todos los puntos del dominio? ¿sólo en los puntos del interior del dominio? ¿Es esencial la hipótesis de que las funciones sean convexas?

Demostración.- Para que se entienda mejor, supongamos que $h = f + g$, y que x está en el interior del dominio de h . Entonces, por la definición de subdiferencial⁹, para cualquier $v \in \partial h(x)$, tenemos que

$$h(y) \geq h(x) + v^T(y - x)$$

para todo y en el dominio de h . Ahora, si podemos descomponer h en f y g . De lo que nos da

$$f(y) + g(y) \geq f(x) + g(x) + v^T(y - x)$$

para todo y en el dominio de h . Luego, supongamos que

$$u \in \partial f(x) \quad y \quad w \in \partial g(x).$$

Entonces, por la definición de subdiferencial⁹, tenemos que

$$f(y) \geq f(x) + u^T(y - x) \quad y \quad g(y) \geq g(x) + w^T(y - x) \quad (1)$$

Sumando estas dos desigualdades nos da

$$f(y) + g(y) \geq f(x) + g(x) + (u + w)^T(y - x)$$

para todo y en el dominio de h . Por último, comparando esta desigualdad con (1), vemos que si elegimos $v = u + w$, entonces la desigualdad se mantiene. Por lo tanto,

$$u + w \in \partial h(x).$$

Es decir,

$$\partial f(x) + \partial g(x) \subseteq \partial(f+g)(x).$$

Demostremos la inclusión. Es decir,

$$\partial(f+g)(x) \subseteq \partial f(x) + \partial g(x).$$

Supongamos que

$$v \in \partial(f+g)(x).$$

Por la definición de subgradiente⁸, esto significa que para todo y en el dominio de $f + g$, tenemos que

$$(f+g)(y) \geq (f+g)(x) + v^T(y - x).$$

Descomponiendo $(f+g)(y)$ y $(f+g)(x)$ en f y g , nos da

$$f(y) + g(y) \geq f(x) + g(x) + v^T(y - x) \quad (2)$$

para todo y en el dominio de $f + g$. Ahora, supongamos que existen

$$u \in \partial f(x) \quad y \quad w \in \partial g(x).$$

Entonces, por la definición de subgradiente⁸ tenemos que

$$f(y) \geq f(x) + u^T(y - x) \quad y \quad g(y) \geq g(x) + w^T(y - x)$$

para todo y en el dominio de $f + g$. Sumamos estas dos desigualdades, nos da

$$f(y) + g(y) \geq f(x) + g(x) + (u + w)^T(y - x)$$

para todo y en el dominio de $f + g$. Comparamos esta desigualdad con (2), vemos que si elegimos

$$v = u + w,$$

entonces la desigualdad se mantiene. Por lo tanto,

$$v \in \partial f(x) + \partial g(x);$$

es decir,

$$\partial(f + g)(x) \subseteq \partial f(x) + \partial g(x).$$

De esta manera, demostramos que

$$\partial(f + g)(x) = \partial f(x) + \partial g(x). \blacksquare$$

En cuanto a la primera pregunta, no necesariamente se tiene la igualdad en todos los puntos del dominio. La igualdad se mantiene en los puntos del interior del dominio donde ambas funciones son diferenciables. En los puntos de la frontera del dominio o en los puntos donde las funciones no son diferenciables, la igualdad puede no mantenerse.

En cuanto a la segunda pregunta, sí es esencial la hipótesis de que las funciones sean convexas. La convexidad de las funciones garantiza la existencia de subgradientes en cada punto del interior de su dominio. Si las funciones no son convexas, entonces pueden no existir subgradientes en algunos puntos, y la igualdad $\partial(f + g)(x) = \partial f(x) + \partial g(x)$ puede no mantenerse.

Ejercicio 2. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función convexa y K un compacto contenido en el interior del dominio de f . Demuestra que f es Lipschitz en K , es decir, que existe una constante $L > 0$ tal que

$$|f(x) - f(y)| \leq L\|x - y\|,$$

para todo $x, y \in K$.

Demostración.- Dado que K es compacto, sabemos que es cerrado y acotado. Por lo tanto, existe un $M > 0$ tal que

$$\|x\| \leq M \text{ para todo } x \in K.$$

Consideremos dos puntos arbitrarios $x, y \in K$. Podemos aplicar la definición de función convexa¹³ para obtener

$$f(y) \geq f(x) + \nabla f(x)^T(y - x)^{14}.$$

¹³Una función f es convexa si para todo x, y en su dominio y para todo $t \in [0, 1]$, se cumple que $f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y)$.

¹⁴Se asume que la función f es diferenciable. Esto se evidencia para el uso del gradiente $\nabla f(x)$ en la desigualdad de Jensen.

Esto implica que

$$f(y) - f(x) \geq \nabla f(x)^T (y - x).$$

Tomando el valor absoluto de ambos lados, obtenemos

$$|f(y) - f(x)| \geq |\nabla f(x)^T (y - x)|.$$

Usando la desigualdad de Cauchy-Schwarz ¹⁵,

$$|\nabla f(x)^T (y - x)| \leq \|\nabla f(x)\| \|y - x\|.$$

Por lo tanto,

$$|f(y) - f(x)| \geq \|\nabla f(x)\| \|y - x\|.$$

Luego, necesitamos demostrar que $\|\nabla f(x)\|$ es acotado en K . Dado que f es convexa, su gradiente satisface la condición de monotonía del gradiente. Además, dado que K es compacto, el gradiente de f debe ser acotado en K . Por lo tanto, existe una constante $L > 0$ tal que

$$\|\nabla f(x)\| \leq L \text{ para todo } x \in K.$$

Finalmente, combinando estas dos desigualdades, obtenemos

$$|f(y) - f(x)| \leq L \|y - x\|.$$

Así, la función f es Lipschitz en K con constante de Lipschitz L . ■

Ejercicio 3. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ dos funciones convexas cuyo dominio es \mathbb{R}^n . Se define su convolución infima (*infimal convolution*) como

$$h(x) = (f \diamond g)(x) := \inf \{f(y) + g(z) : y + z = x\}.$$

Demuestra que $h(y)$. ¿Es suficiente esta identidad para demostrar que si f y g son convexas, entonces h es convexa?

Demostración.- Necesitaremos demostrar que para cualquier par de puntos $x, y \in \mathbb{R}^n$ y cualquier $t \in [0, 1]$, se cumple que

$$h(tx + (1 - t)y) \leq th(x) + (1 - t)h(y).$$

Dado que

$$h(x) = (f \diamond g)(x) = \inf \{f(u) + g(v) : u + v = x\},$$

podemos elegir (u_1, v_1) y (u_2, v_2) tal que

$$u_1 + v_1 = x, \quad h(x) = f(u_1) + g(v_1)$$

$$u_2 + v_2 = y, \quad h(y) = f(u_2) + g(v_2).$$

Ahora, consideremos

$$z = tx + (1 - t)y = tu_1 + (1 - t)u_2 + tv_1 + (1 - t)v_2.$$

¹⁵La **desigualdad de Cauchy-Schwarz** establece que para cualquier espacio vectorial complejo con un producto escalar, y para cualquier par de vectores x y y en ese espacio, se cumple la siguiente desigualdad: $|x \cdot y| \leq \|x\| \|y\|$, donde $x \cdot y$ denota el producto escalar de x y y , y $\|x\|$ y $\|y\|$ son las normas (o longitudes) de los vectores x y y , respectivamente. La igualdad se cumple si y solo si los vectores x y y son linealmente dependientes. Esto significa que uno de los vectores puede ser escrito como un múltiplo escalar del otro.

Donde, notamos que

$$tu_1 + (1-t)u_2 \quad \text{y} \quad tv_1 + (1-t)v_2$$

son combinaciones convexas de u_1, u_2 y v_1, v_2 respectivamente. Como f y g son convexas, tenemos que

$$f(tu_1 + (1-t)u_2) \leq tf(u_1) + (1-t)f(u_2),$$

$$g(tv_1 + (1-t)v_2) \leq tg(v_1) + (1-t)g(v_2).$$

Sumando estas dos desigualdades, obtenemos

$$\begin{aligned} f(tu_1 + (1-t)u_2) + g(tv_1 + (1-t)v_2) &\leq tf(u_1) + (1-t)f(u_2) + tg(v_1) + (1-t)g(v_2) \\ &= th(x) + (1-t)h(y). \end{aligned}$$

Dado que el lado izquierdo de esta desigualdad es una cota superior para $h(z)$, tenemos que

$$h(z) = h(tx + (1-t)y) \leq th(x) + (1-t)h(y),$$

lo que demuestra que h es convexa. Por lo tanto, si f y g son convexas, entonces su convolución ínfima h también es convexa. Esta identidad es suficiente para demostrar la convexidad de h . Ya que, si f es convexa, entonces para cualquier x, y en el dominio de f y cualquier t en el intervalo $[0, 1]$, se cumple que

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

En la demostración dada, mostramos que para cualquier par de puntos x, y en el dominio de h y cualquier t en el intervalo $[0, 1]$, se cumple que

$$h(tx + (1-t)y) \leq th(x) + (1-t)h(y).$$

Por lo tanto, hemos demostrado que h satisface la definición de una función convexa, lo cual es suficiente para demostrar que h es convexa. En otras palabras, hemos demostrado que si f y g son convexas, entonces su convolución ínfima h también es convexa. ■