525401 Análisis Funcional I
Capítulo 2: Dualidad[1]

Leonardo E. Figueroa

CI^{2}MA y Departamento de Ingeniería Matemática
Universidad de Concepción

Semestre 2022-1

[1]Casi todos los contenidos tomados del libro **Gabriel N. Gatica:** Introducción al Análisis: Teoría y Aplicaciones, Editorial Reverté, Barcelona, 2014.

2.1 Preliminares

Sea $(V, \|\cdot\|)$ un espacio vectorial normado (EVN) sobre un cuerpo $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$). Toda aplicación $F\colon V \to \mathbb{K}$ se denomina *funcional*.

Definición 2.1

Un funcional $F\colon V \to \mathbb{K}$ se dice *lineal* si

$$(\forall\, \alpha, \beta \in \mathbb{K})\ (\forall\, v, w \in V) \quad F(\alpha v + \beta w) = \alpha F(v) + \beta F(w).$$

Definición 2.2

Un funcional lineal $F\colon V \to \mathbb{K}$ se dice *acotado* respecto a una norma dada $\|\cdot\|$ de V si existe una constante $M > 0$ tal que

$$(\forall\, v \in V) \quad |F(v)| \leq M\, \|v\|\,.$$

Ejemplo 2.1

Ejemplos de funcionales:

1. Para cualquier EVN $(V, \|\cdot\|)$, la norma $\|\cdot\| : V \to \mathbb{R}$ es un funcional nolineal; en efecto, dado cualquier $x \in V \setminus \{0\}$ y $\alpha < 0$, $\|\alpha x\| = |\alpha| \, \|x\| \neq \alpha \, \|x\|$.
2. Dado cualquier espacio de Hilbert $(V, \langle\cdot,\cdot\rangle)$ y un $x \in V$, la aplicación $V \ni v \mapsto \langle v, x\rangle \in \mathbb{K}$ es un funcional lineal. Por la desigualdad de Cauchy–Schwarz es acotado. Después probaremos que todos los funcionales lineales y acotados sobre un espacio de Hilbert poseen esta forma.

El conjunto de todos los funcionales lineales y acotados sobre un EVN V se llama *dual* de V y se denota V'. Sobre V' se definen las operaciones suma y multiplicación por escalar de la forma usual:

$$\begin{aligned} +\colon V' \times V' &\to V' \\ (F, G) &\to F + G, \quad (\forall\, v \in V)\ (F+G)(v) := F(v) + G(v). \end{aligned} \tag{2.1}$$

$$\begin{aligned} \cdot\colon \mathbb{K} \times V' &\to V' \\ (\lambda, F) &\to \lambda F, \quad (\forall\, v \in V)\ (\lambda F)(v) := \lambda F(v). \end{aligned} \tag{2.2}$$

Es fácil probar que $(V', +, \cdot)$ es un espacio vectorial sobre $\mathbb{K}$ cuyo elemento neutro para la adición es el funcional nulo $0_{V'} : V \to \mathbb{K}$, definido por $0_{V'}(v) = 0$ para todo $v \in V$ (usualmente escribiremos 0 en lugar de $0_{V'}$).

Definición 2.3

Dado $F \in V'$ definimos $\|V\|_{V'}$ como el ínfimo de todas las constantes $M > 0$ que satisfacen la condición dada en la Definición 2.2; esto es,

$$\|F\|_{V'} := \text{ínf}\{M \geq 0\colon (\forall\, v \in V)\ |F(v)| \leq M\,\|v\|_V\}.$$

Se sigue que

$$(\forall\, F \in V') \quad \|F\|_{V'} = \sup_{v \in V \setminus \{0\}} \frac{|F(v)|}{\|v\|_V}. \tag{2.3}$$

Equivalentemente,

$$\|F\|_{V'} = \sup_{\substack{v \in V \\ \|v\| \leq 1}} |F(v)|$$

y

$$\|F\|_{V'} = \sup_{\substack{v \in V \\ \|v\| = 1}} |F(v)|\,.$$

Es fácil probar que, como sugiere su notación, $\|\cdot\|_{V'}$ es una norma sobre V', con lo cual $(V', \|\cdot\|_{V'})$ constituye un EVN. Todavía más, este EVN hereda de $\mathbb{K}$ la completitud, siendo por lo tanto un espacio de Banach (probaremos un resultado más general más adelante). Si no causa confusión, para todo $v \in V$ y $F \in V'$ escribiremos $\|v\|$ en lugar de $\|v\|_V$ y $\|F\|$ en lugar de $\|F\|_{V'}$.

2.2 Teorema de Representación De Riesz

Teorema 2.1 (Primer teorema de mejor aproximación)

Sea $U \neq \emptyset$ un subespacio cerrado de un espacio de Hilbert $(V, \langle\cdot,\cdot\rangle)$. Entonces, para todo $x \in V$ existe un único $z \in U$ tal que

$$\|x - z\| = \inf_{u\in U} \|x - u\| . \tag{2.5}$$

el cual se llama la *mejor aproximación* de x por elementos de U. Además, $z \in U$ está caracterizado por la relación de ortogonalidad

$$(\forall\, u \in U) \quad \langle x - z, u\rangle = 0. \tag{2.6}$$

Demostración: Sea $d := \inf_{u\in U} \|x - u\|$. Entonces, para todo $n \in \mathbb{N}$, existe $z_n \in U$ tal que $d \leq \|x - z_n\| < d + 1/n$. Por lo tanto, $\|x - z_n\| \xrightarrow{n\to\infty} d$. Dados $m, n \in \mathbb{N}$, aplicando la ley del paralelógramo a $x - z_n$ y a $x - z_m$, se tiene

$$\|(x - z_n) + (x - z_m)\|^2 + \|(x - z_n) - (x - z_m)\|^2 = 2\left(\|x - z_n\|^2 + \|x - z_m\|^2\right),$$

o bien

$$4\left\|x - \frac{z_n + z_m}{2}\right\|^2 + \|z_n - z_m\|^2 = 2\|x - z_n\|^2 + 2\|x - z_m\|^2 ,$$

de donde

$$\|z_n - z_m\|^2 = 2\|x - z_n\|^2 + 2\|x - z_m\|^2 - 4\left\|x - \frac{z_n + z_m}{2}\right\|^2 .$$

Como $\frac{z_n+z_m}{2} \in U$, resulta que $d \leq \left\|x - \frac{z_n+z_m}{2}\right\|$, por lo que

$$\|z_n - z_m\|^2 \leq 2\|x - z_n\|^2 + 2\|x - z_m\|^2 - 4d^2.$$

Así, $\|z_n - z_m\| \xrightarrow{m,n\to\infty} 0$, lo que indica que $\{z_n\}_{n\in\mathbb{N}}$ es de Cauchy en V. Como U es cerrado, inferimos que existe $z \in U$ tal que $z_n \xrightarrow{n\to\infty} z$ y, por lo tanto,

$$d = \lim_{n\to\infty} \|x - z_n\| = \|x - z\|.$$

Para la unicidad, sean $z_1, z_2 \in U$ tales que $\|x - z_1\| = \|x - z_2\| = d$. Aplicando de nuevo los mismos cálculos anteriores se obtiene

$$\begin{aligned} 0 \leq \|z_1 - z_2\|^2 = 2\|x - z_1\|^2 + 2\|x - z_2\|^2 - 4\left\|x - \frac{z_1+z_2}{2}\right\|^2 \\ = 4d^2 - 4\left\|x - \frac{z_1+z_2}{2}\right\|^2 \leq 0, \end{aligned}$$

por lo que $\|z_1 - z_2\| = 0$.

Ahora probamos la equivalencia entre (2.5) y (2.6). Sea $z \in U$ tal que

$$d := \|x - z\| = \inf_{u \in U} \|x - u\| .$$

Entonces, para todo $u \in U$, $d^2 \leq \|x - u\|^2$ y, en particular, dado que U es subespacio, en el caso en que $\mathbb{K} = \mathbb{R}$,

$$(\forall\, \alpha \in \mathbb{R})\ (\forall\, u \in U) \quad d^2 \leq \|x - (z - \alpha u)\|^2 = \|(x - z) + \alpha u\|^2 .$$

Desarrollando la desigualdad anterior,

$$d^2 \leq \|x - z\|^2 + 2\alpha\langle x - z, u\rangle + |\alpha|^2 \|u\|^2 ,$$

de donde,

$$(\forall\, \alpha \in \mathbb{R})\ (\forall\, u \in U) \quad 0 \leq 2\alpha\langle x - z, u\rangle + |\alpha|^2 \|u\|^2 .$$

Restringiéndonos a los casos particulares en los que $\alpha = t\langle x - z, u\rangle$ con $t < 0$,

$$(\forall\, t < 0)\ (\forall\, u \in U) \quad 0 \leq 2t\langle x - z, u\rangle^2 + t^2\langle x - z, u\rangle^2 \|u\|^2 .$$

Dividiendo por el número negativo t y tomando el límite lateral $t \to 0^-$,

$$(\forall\, u \in U) \quad 0 \geq 2\langle x - z, u\rangle^2 \implies \langle x - z, u\rangle = 0.$$

¿Cómo sería el caso complejo?

Recíprocamente, dado $z \in U$ tal que $\langle x - z, u \rangle = 0$ para todo $u \in U$. Entonces,

$$(\forall\, u \in U) \quad \|x - u\|^2 = \|(x - z) - (u - z)\|^2 = \|x - z\|^2 + \|u - z\|^2 ,$$

de donde $\|x - z\| \leq \|x - u\|$ para todo $u \in U$ y $\|x - z\| = \|x - u\|$ si y solo si $u = z$. □

Es importante notar que exactamente la misma demostración utilizada para probar la existencia y unicidad de la mejor aproximación $z \in U$ es válida en el caso en que U es meramente un subconjunto convexo y cerrado del Hilbert V. En efecto, los únicos usos en esa parte de la demostración de la hipótesis de que U fuese un espacio vectorial fueron para afirmar que el promedio de miembros de U seguía perteneciendo a U, lo que todavía ocurre si U es convexo.

Sin embargo, la convexidad de U no es suficiente para probar la equivalencia entre (2.5) y (2.6).

De acuerdo a lo anterior, podemos establecer el siguiente resultado:

Teorema 2.2 (Segundo teorema de mejor aproximación)

Sea $U \neq \emptyset$ un subconjunto convexo y cerrado de un Hilbert $(V, \langle \cdot, \cdot \rangle)$. Entonces, para todo $x \in V$ existe un único $z \in U$ tal que

$$\|x - z\| = \inf_{u \in U} \|x - u\| ,$$

el cual se llama la *mejor aproximación* de x por elementos de U.

Ejemplo 2.2

Este ejemplo ilustra el caso en dimensión finita del Teorema 2.1.

1. Si U es un subespacio de dimensión finita de V, entonces el cálculo de la mejor aproximación se reduce a un sistema de ecuaciones lineales. En efecto, si $\{u_1, \ldots, u_N\}$ es una base de U, entonces (2.6) equivale a

$$(\forall\, j \in \{1, \ldots, N\}) \quad \langle x - z, u_j \rangle = 0.$$

Así, si $z = \sum_{i=1}^{N} \alpha_i u_i$, entonces podemos escribir

$$(\forall\, j \in \{1, \ldots, N\}) \quad \langle x, u_j \rangle = \langle z, u_j \rangle = \sum_{i=1}^{N} \alpha_i \langle u_i, u_j \rangle.$$

Equivalentemente, se tiene el sistema de Gram,

$$\boldsymbol{A\alpha} = \boldsymbol{b}, \tag{2.7}$$

donde $\boldsymbol{A} = (a_{i,j})_{j,i=1}^{N}$ con $a_{i,j} = \langle u_i, u_j \rangle$, $\boldsymbol{\alpha} = (\alpha_i)_{i=1}^{N}$ y $\boldsymbol{b} = (b_j)_{j=1}^{N}$ con $b_j = \langle x, u_j \rangle$.

Ejemplo 2.2 (cont.)

2. Si la base $\{u_1, \ldots, u_N\}$ de U es ortogonal (esto es, si $\langle u_i, u_j \rangle = 0$ para $i \neq j$), entonces la resolución de (2.7) es directa, obteniéndose en tal caso

$$(\forall\, i \in \{1, \ldots, N\}) \quad \alpha_i = \frac{b_i}{\|u_i\|^2} = \frac{\langle x, u_i \rangle}{\|u_i\|^2}. \tag{2.8}$$

3. Un ejemplo particular de 2 lo constituye la serie finita de Fourier. En efecto, en este caso consideramos la variante real de $V = L^2(-\pi, \pi)$ provista de su producto escalar usual $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x) g(x)\,\mathrm{d}x$ y escogemos a U como el subespacio generado por las funciones trigonométricas $\{\varphi_0, \ldots, \varphi_N, \psi_1, \ldots, \psi_N\}$, donde $\varphi_j(x) = \cos(jx)$ y $\psi_j(x) = \operatorname{sen}(jx)$. Así, dada $f \in L^2(-\pi, \pi)$, su mejor aproximación por elementos de U queda dada por

$$f_N(x) = \sum_{i=0}^{N} \alpha_i \cos(ix) + \sum_{i=1}^{N} \beta_i \operatorname{sen}(ix),$$

donde los escalares α_i y β_i, llamados *coeficientes de Fourier*, están definidos de acuerdo a (2.8); esto es,

$$\alpha_i := \frac{\int_{-\pi}^{\pi} f(x) \cos(ix)\,\mathrm{d}x}{\int_{-\pi}^{\pi} \cos(ix)^2\,\mathrm{d}x} \quad \text{y} \quad \beta_i := \frac{\int_{-\pi}^{\pi} f(x) \operatorname{sen}(ix)\,\mathrm{d}x}{\int_{-\pi}^{\pi} \operatorname{sen}(ix)^2\,\mathrm{d}x}.$$

Definición 2.4

Dado un subconjunto S de un espacio de Hilbert $(V, \langle\cdot,\cdot\rangle)$, se define el *ortogonal* de S, y se denota $S^\perp$, como

$$S^\perp := \{x \in V : (\forall\, v \in S) \quad \langle x, v\rangle = 0\},$$

el cual es un subespacio cerrado de V. Notar que $S \cap S^\perp = \{0\}$ para todo $S \subseteq V$ que satisfaga $0 \in S$.

El siguiente teorema es una consecuencia del Teorema 2.1:

Teorema 2.3 (Teorema de descomposición ortogonal)

Sea $U \neq \emptyset$ un subespacio cerrado de un Hilbert $(V, \langle\cdot,\cdot\rangle)$. Entonces,

$$V = U \oplus U^\perp. \tag{2.9}$$

Demostración: Dado $x \in V$, podemos escribir $x = z + (x - z)$, donde $z \in U$ es la mejor aproximación de x por elementos de U y, por la relación de ortogonalidad (2.6), $x - z \in U^\perp$. □

Como $U^\perp$ también es un subespacio cerrado del Hilbert V, observamos que $z \in U$ es la mejor aproximación de x por elementos de U si y solo si $x - z \in U^\perp$ es la mejor aproximación de x por elementos de $U^\perp$. Equivalentemente, si P_S es el operador que a cada $x \in V$ le asigna su mejor aproximación por elementos de un subespacio cerrado S de V, entonces (2.9) se reescribe como

$$I = P_U + P_{U^\perp},$$

donde I denota a la aplicación identidad de V en V. Aprovechamos de mencionar, en virtud de la relación de ortogonalidad (2.6), que las aplicaciones $P_U : V \to U$ y $P_{U^\perp} : V \to U^\perp$ se llaman *proyecciones ortogonales* sobre U y $U^\perp$, respectivamente. De acuerdo a esto, el Teorema 2.1 también percibe el nombre Teorema de proyección ortogonal.

Ejemplo 2.3

Los siguientes ejemplos ilustral el Teorema de descomposición ortogonal.

1. Sea Ω un abierto acotado de $\mathbb{R}^2$ y consideremos el espacio de Lebesgue $L^2(\Omega)$ sobre $\mathbb{R}$ con el producto escalar usual $\langle f, g\rangle := \int_\Omega fg$. Definimos el subespacio

$$U := \{\varphi\colon \Omega \to \mathbb{R}\colon \varphi \text{ es constante}\} = \operatorname{span}(\{\varphi_0\}),$$

donde $\varphi_0 \in V$ es tal que $\varphi_0(x) := 1$ para todo $x \in \Omega$. Se sigue que

$$U^\perp = \{f \in V\colon \langle f, \varphi_0\rangle = 0\} = \left\{f \in V\colon \int_\Omega f = 0\right\}.$$

Entonces, dada $f \in V$, su mejor aproximación por elementos de U se reduce a la función constante definida por

$$(\forall\, x \in \Omega) \quad f_U(x) := \frac{\langle f, \varphi_0\rangle}{\langle \varphi_0, \varphi_0\rangle}\varphi_0(x) = \frac{1}{|\Omega|}\int_\Omega f.$$

De este modo, cada $f \in V$ puede descomponerse de manera única como

$$f = f_U + f_{U^\perp}, \quad \text{con} \quad f_U \in U \quad \text{y} \quad f_{U^\perp} := f - \frac{1}{|\Omega|}\int_\Omega f \in U^\perp.$$

Ejemplo 2.3 (cont.)

2. Sea $V := \mathbb{R}^{n\times n}$, el espacio de matrices cuadradas de orden n con coeficientes reales, provisto del producto escalar $\langle A, B\rangle := \operatorname{tr}(B^{\mathtt{t}}A)$. Entonces, definimos los subespacios

$$U_1 := \{\kappa\mathbb{I} \colon \kappa \in \mathbb{R}\} = \operatorname{span}(\{\mathbb{I}\}) \quad \text{y} \quad U_2 := \left\{A \in V \colon A = A^{\mathtt{t}}\right\},$$

donde $\mathbb{I}$ denota a la matriz identidad. Se sigue que

$$U_1^{\perp} = \{A \in V \colon \operatorname{tr}(A) = 0\} \quad \text{y} \quad U_2^{\perp} = \{A \in V \colon A^{\mathtt{t}} = -A\}.$$

Entonces, dada $A \in V$, se prueba que sus mejores aproximaciones por elementos de U_1 y U_2 se reducen, respectivamente, a

$$A_1 := \frac{1}{n}\operatorname{tr}(A)\mathbb{I} \quad \text{y} \quad A_2 := \frac{1}{2}\left(A + A^{\mathtt{t}}\right).$$

Luego, cada $A \in V$ puede descomponerse de manera única como

$$A = A_1 + A_1^{\perp}, \quad \text{con} \quad A_1 \in U_1 \quad \text{y} \quad A_1^{\perp} := A - \frac{1}{n}\operatorname{tr}(A)\mathbb{I} \in U_1^{\perp}$$

y también como

$$A = A_2 + A_2^{\perp}, \quad \text{con} \quad A_2 \in U_2 \quad \text{y} \quad A_2^{\perp} := \frac{1}{2}\left(A - A^{\mathtt{t}}\right) \in U_2^{\perp}.$$

Ejemplo 2.3 (cont.)

2. (cont.) Un caso de esta segunda descomposición, de gran interés en mecánica de medios continuos, se aplica al vector de desplazamientos $\boldsymbol{u}\colon \Omega \to \mathbb{R}^3$ de un sólido elástico lineal. En tal caso, usualmente se escribe

$$\nabla \boldsymbol{u} := \boldsymbol{e}(\boldsymbol{u}) + \boldsymbol{\gamma}(\boldsymbol{u}),$$

donde $\boldsymbol{e}(\boldsymbol{u}) = \frac{1}{2}\left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathtt{t}}\right)$ es el tensor de pequeñas deformaciones y $\boldsymbol{\gamma}(\boldsymbol{u}) := \frac{1}{2}\left(\nabla \boldsymbol{u} - (\nabla \boldsymbol{u})^{\mathtt{t}}\right)$ es el tensor de rotaciones.

3. Sea Ω un abierto acotado de $\mathbb{R}^n$ y consideremos el espacio de Hilbert $V := [L^2(\Omega)]^{n\times n}$ sobre el cuerpo $\mathbb{R}$ con el producto escalar $\langle \boldsymbol{\sigma}, \boldsymbol{\tau}\rangle := \int_\Omega \boldsymbol{\sigma} : \boldsymbol{\tau}$, donde $\boldsymbol{\sigma} : \boldsymbol{\tau} = \sum_{i,j=1}^{n} \sigma_{i,j}\tau_{i,j}$ para todo $\boldsymbol{\sigma} = (\sigma_{i,j})_{i,j=1}^{n}$ y $\boldsymbol{\tau} = (\tau_{i,j})_{i,j=1}^{n}$. Definamos el subespacio

$$U := \{\kappa\mathbb{I} \colon \kappa \in \mathbb{R}\} = \operatorname{span}(\{\mathbb{I}\}),$$

donde $\mathbb{I}$ es el tensor identidad entendido como función constante en V. Se sigue que

$$U^{\perp} := \left\{\boldsymbol{\sigma} \in V \colon \int_\Omega \operatorname{tr}(\boldsymbol{\sigma}) = 0\right\}.$$

Ejemplo 2.3 (cont.)

3. (cont.) Luego, dada $\boldsymbol{\sigma} \in V$, sus mejores aproximaciones por elementos de U y $U^\perp$ se reducen, respectivamente, a

$$\boldsymbol{\sigma}_U := \frac{1}{n\,|\Omega|}\int_\Omega \operatorname{tr}(\boldsymbol{\sigma})\mathbb{I} \quad \text{y} \quad \boldsymbol{\sigma}_{U^\perp} := \boldsymbol{\sigma} - \frac{1}{n\,|\Omega|}\int_\Omega \operatorname{tr}(\boldsymbol{\sigma})\mathbb{I},$$

con lo cual, cada $\boldsymbol{\sigma} \in V$ se descompone de manera única como la suma de un tensor $\boldsymbol{\sigma}_U$, múltiplo escalar de la identidad, y un tensor $\boldsymbol{\sigma}_{U^\perp}$, cuya traza tiene media nula en Ω. Esta descomposición también es de interés en formulaciones variacionales mixtas de algunos problemas de elasticidad.

4. Sea Ω un abierto acotado de $\mathbb{R}^n$ y consideremos el espacio de Sobolev $V := H^1(\Omega)$ sobre el cuerpo $\mathbb{R}$, provisto de su producto escalar usual $\langle w, v\rangle := \int_\Omega (\nabla w \cdot \nabla v + w\,v)$. Además, sea $U := H_0^1(\Omega)$ la clausura de $C_0^\infty(\Omega)$ respecto a la norma inducida por $\langle\cdot,\cdot\rangle$. Entonces, es fácil ver que

$$U^\perp = \{v \in V:\ -\Delta v + v = 0 \text{ en } \mathcal{D}'(\Omega)\}.$$

En consecuencia, dado $w \in V$, su mejor aproximación w_U por elementos de U será la única solución débil del problema de valores de contorno:

$$-\Delta w_U + w_U = f := -\Delta w + w \text{ en } \Omega, \quad w_U = 0 \text{ en } \partial\Omega.$$

El siguiente lema es un corolario directo del teorema de descomposición ortogonal (Teorema 2.3):

Lema 2.1

Sea $U \neq \emptyset$ un subespacio cerrado propio de un Hilbert V. Entonces, existe $\tilde{x} \in V$, $\tilde{x} \neq 0$, tal que $\tilde{x} \in U^{\perp}$.

Lema extra sin cargo adicional

Un funcional lineal $f \colon E \to \mathbb{R}$ es continuo si y solo si es acotado.

DEMOSTRACIÓN: Supongamos que f es acotado. Entonces, dada una sucesión $\{x_n\}_{n\in\mathbb{N}} \subset E$ que converge a un límite x en E,

$$|f(x) - f(x_n)| \leq \|f\| \, \|x - x_n\| \xrightarrow{n\to\infty} 0,$$

lo que prueba la continuidad secuencial de f, que en espacios normados como E es lo mismo que la continuidad.

Supongamos ahora que f es continuo y por reducción al absurdo asumamos que f *no* es acotado. Entonces, para todo $M > 0$ existe $x_M \in E$ tal que $|f(x_M)| > M \|x_M\|$. Para todo $n \in \mathbb{N}$ definamos $y_n := n^{-1} \|x_n\|^{-1} x_n$. Por un lado, $\|y_n\| = 1/n$, de donde $y_n \xrightarrow{n\to\infty} 0$. Pero por otro lado, como f es lineal

$$f(y_n) = n^{-1} \|x_n\|^{-1} f(x_n) > n^{-1} \|x_n\|^{-1} n \|x_n\| = 1,$$

por lo que $\{f(y_n)\}_{n\in\mathbb{N}}$ no converge a $0 = f(0)$, contradiciendo la continuidad de f. □

Con los resultados anteriores podemos demostrar el teorema principal de esta sección.

Teorema 2.4 (Teorema de representación de Riesz)

Sea $(V, \langle\cdot,\cdot\rangle)$ un espacio de Hilbert y sea $F \in V'$. Entonces, existe un único $x \in V$ tal que $F(v) = \langle v, x\rangle$ para todo $v \in V$ y, además, $\|F\| = \|x\|$.

DEMOSTRACIÓN: Dado $F \in V'$, definamos $U := \{v \in V : F(v) = 0\}$. En el caso especial en el que $U = V$, entonces F era el funcional nulo sobre V y en tal caso basta tomar $x = 0$.

Supongamos ahora que $U \neq V$. Como U es un subespacio cerrado de V, del Lema 2.1 se deduce que existe $\tilde{x} \in V \setminus \{0\}$ tal que $\tilde{x} \in U^\perp$. Es claro entonces que $\tilde{x} \notin U$, y por lo tanto $F(\tilde{x}) \neq 0$. Además, para cada $v \in V$ se observa que $F(\tilde{x})v - F(v)\tilde{x} \in U$, y dado que $\tilde{x} \in U^\perp$ se deduce que

$$(\forall\, v \in V) \quad 0 = \langle F(\tilde{x})v - F(v)\tilde{x}, \tilde{x}\rangle.$$

Esto es,
$$\langle F(\tilde{x})v, \tilde{x}\rangle = \langle F(v)\tilde{x}, \tilde{x}\rangle = F(v)\,\|\tilde{x}\|^2\,,$$
de donde, escogiendo $x = \dfrac{\overline{F(\tilde{x})}\tilde{x}}{\|\tilde{x}\|^2}$,
$$(\forall\, v \in V) \quad F(v) = \langle v, x\rangle.$$
Para probar la unicidad de x, supongamos que existe otro $z \in V$ tal que
$$(\forall\, v \in V) \quad F(v) = \langle v, z\rangle.$$
Entonces, $x - z \in V^{\perp} = \{0\}$, por lo que $z = x$.

Para la igualdad de normas deseada, por la desigualdad de Cauchy–Schwarz,
$$\|F\| = \sup_{v\in V\setminus\{0\}} \frac{|F(v)|}{\|v\|} = \sup_{v\in V\setminus\{0\}} \frac{|\langle v, x\rangle|}{\|v\|} \leq \|x\|\,.$$
Además,
$$\|F\| = \sup_{v\in V\setminus\{0\}} \frac{|F(v)|}{\|v\|} \geq \frac{|F(x)|}{\|x\|} = \frac{\langle x, x\rangle}{\|x\|} = \|x\|\,.$$
□

Es muy importante observar que el Teorema de representación de Riesz (Teorema 2.4; abreviado por TRR) induce la definición del *operador o aplicación de Riesz*

$$\begin{aligned} \mathcal{R}\colon V' &\to V \\ F &\to \mathcal{R}(F) \end{aligned}$$

donde $\mathcal{R}(F)$ es el único elemento en V, llamado *representante* de F, que satisface

$$(\forall\, v \in V) \quad F(v) = \langle v, \mathcal{R}(F) \rangle.$$

Notar que $\mathcal{R}$ es una biyección isométrica.

2.3 Teorema de Hahn–Banach

Sea $(V, \|\cdot\|)$ un EVN no trivial ($V \neq \{0\}$) sobre $\mathbb{R}$ cualquiera. ¿Existe necesariamente algún elemento no nulo del dual V'?

Para explorar esta pregunta tomemos un elemento arbitrario $x_0 \in V \setminus \{0\}$ y definamos el subespacio de unidimensional de V dado por

$$V_0 := \{\alpha x_0 \colon \alpha \in \mathbb{R}\} = \operatorname{span}(\{x_0\}).$$

Definimos el funcional

$$\begin{aligned} f\colon V_0 &\to \mathbb{R} \\ \alpha x_0 &\to f(\alpha x_0) := \alpha \end{aligned}$$

Claramente f es lineal. Además, dado un miembro genérico $x = \alpha x_0 \in V_0$, se tiene

$$|f(x)| = |\alpha| = \frac{|\alpha| \, \|x_0\|}{\|x_0\|} = \frac{1}{\|x_0\|} \|x\| ,$$

lo cual prueba que f es acotado y $\|f\|_{V_0'} = 1/\|x_0\|$.

Este funcional, así como está, no nos sirve porque solamente está definido y es acotado sobre el subespacio V_0. La respuesta a nuestra pregunta sería afirmativa si fuese posible **extender** a f sobre todo V **preservando** su condición de acotado. La existencia de semejante extensión precisamente está garantizada por el Teorema de Hahn–Banach.

- Definición 2.5 (Conjunto parcialmente ordenado)
- Ejemplo 2.4
- Definición 2.6 (Elemento maximal)
- Lema 2.2 (Lema de Zorn)

Omitidos

Definición 2.7

Sea V un espacio vectorial sobre $\mathbb{R}$. Un funcional $p\colon V \to \mathbb{R}$ se dice *sublineal* si

- $(\forall\, x, y \in V) \quad p(x+y) \leq p(x) + p(y)$ y
- $(\forall\, x \in V)\ (\forall\, \alpha > 0) \quad p(\alpha x) = \alpha\, p(x)$.

Teorema 2.5 (Forma analítica del teorema de Hahn–Banach)

Sea V un espacio vectorial sobre $\mathbb{R}$ y sea $p\colon V \to \mathbb{R}$ un funcional sublineal. Sea U un subespacio de V y sea $f\colon U \to \mathbb{R}$ un funcional lineal que satisface

$$(\forall\, x \in U) \quad f(x) \leq p(x).$$

Entonces, existe un funcional lineal $F\colon V \to \mathbb{R}$ tal que

$$(\forall\, x \in U) \quad F(x) = f(x) \qquad \text{y} \qquad (\forall\, x \in V) \quad F(x) \leq p(x).$$

DEMOSTRACIÓN: Omitida.

Teorema 2.6 (Teorema de Hahn–Banach en espacios normados)

Sea $(V, \|\cdot\|)$ un EVN sobre $\mathbb{R}$, U un subespacio de V y $f \in U'$. Entonces, existe $F \in V'$ tal que

$$(\forall\, x \in U) \quad F(x) = f(x) \qquad \text{y} \qquad \|F\|_{V'} = \|f\|_{U'}\,.$$

Demostración: Sea $p\colon V \to \mathbb{R}$ el funcional definido por $p(x) = \|f\|_{U'}\,\|x\|$ para todo $x \in V$, el cual claramente es sublineal. Aplicando la forma analítica del teorema de Hahn–Banach (Teorema 2.5) a $f \in U'$ y a p se deduce que existe $F\colon V \to \mathbb{R}$ lineal tal que

$$(\forall\, x \in U) \quad F(x) = f(x) \qquad \text{y} \qquad (\forall\, x \in V) \quad F(x) \le \|f\|_{U'}\,\|x\|\,.$$

Se sigue que

$$(\forall\, x \in V) \quad -F(x) = F(-x) \le \|f\|_{U'}\,\|-x\| = \|f\|_{U'}\,\|x\|\,,$$

lo que permite concluir que

$$(\forall\, x \in V) \quad |F(x)| \le \|f\|_{U'}\,\|x\|\,.$$

Esto prueba que $F \in V'$ y $\|F\|_{V'} \le \|f\|_{U'}$. Por otro lado,

$$\|F\|_{V'} = \sup_{v \in V\setminus\{0\}} \frac{|F(v)|}{\|v\|} \ge \sup_{v \in U\setminus\{0\}} \frac{|F(v)|}{\|v\|} = \sup_{v \in U\setminus\{0\}} \frac{|f(v)|}{\|v\|} = \|f\|_{U'}\,.$$

□

A continuación presentamos algunas consecuencias de uso frecuente del teorema de Hahn–Banach en espacios normados (Teorema 2.6).

Teorema 2.7

Sea $(V, \|\cdot\|)$ un EVN sobre $\mathbb{R}$ y sea $x_0 \in V \setminus \{0\}$. Entonces, existe $F \in V'$ tal que $\|F\| = 1$ y $F(x_0) = \|x_0\|$.

DEMOSTRACIÓN: Sea $S = \operatorname{span}(\{x_0\}) = \{\alpha\, x_0 \colon \alpha \in \mathbb{R}\}$ y definamos el funcional $f \colon S \to \mathbb{R}$ que a cada miembro genérico $x = \alpha\, x_0 \in S$ asocia $f(x) := \alpha\, \|x_0\|$. El funcional f es claramente lineal. Además, para todo $\alpha\, x_0 \in S$, $|f(\alpha\, x_0)| = |\alpha|\, \|x_0\| = \|\alpha\, x_0\|$, lo que muestra que f es acotado y $\|f\|_{S'} = 1$. Aplicando el Teorema 2.6 se concluye que existe $F \in V'$ tal que $F(x) = f(x)$ para todo $x \in S$ y $\|F\|_{V'} = \|f\|_{S'} = 1$. En particular, para $x_0 \in S$ se tiene $F(x_0) = f(x_0) = \|x_0\|$. □

Lema 2.3

Sea $(V, \|\cdot\|)$ un EVN sobre $\mathbb{R}$ y sea $v \in V$ tal que $F(v) = 0$ para todo $F \in V'$. Entonces, $v = 0$.

DEMOSTRACIÓN: Se sigue directamente del Teorema 2.7. □

Este importante lema manifiesta que la norma de un vector puede expresarse en forma análoga a como se expresa la norma de un funcional acotado.

Lema 2.4

Sea $(V, \|\cdot\|)$ un EVN. Entonces,

$$(\forall\, v \in V) \quad \|v\| = \sup_{F \in V' \setminus \{0\}} \frac{|F(v)|}{\|F\|} = \max_{F \in V' \setminus \{0\}} \frac{|F(v)|}{\|F\|}.$$

Demostración: Si $v = 0$ el resultado es trivial. Si $v \neq 0$, del Teorema 2.7 se infiere que existe $\tilde{F} \in V'$ tal que $\|\tilde{F}\| = 1$ y $\tilde{F}(v) = \|v\|$, con lo cual

$$\sup_{F \in V' \setminus \{0\}} \frac{|F(v)|}{\|F\|} \geq \frac{|\tilde{F}(v)|}{\|\tilde{F}\|} = \|v\|\,.$$

Por otro lado, es claro que

$$\sup_{F \in V' \setminus \{0\}} \frac{|F(v)|}{\|F\|} \leq \sup_{F \in V' \setminus \{0\}} \frac{\|F\|\,\|v\|}{\|F\|} = \|v\|\,.$$

$\square$

Teorema 2.8

Sea S un subespacio de un EVN $(V, \|\cdot\|)$ y sea $x_0 \in V$ tal que

$$\operatorname{dist}(x_0, S) := \inf_{x \in S} \|x_0 - x\| > 0.$$

Entonces, existe $F \in V'$ tal que $\|F\| = 1$, $F(x_0) = \operatorname{dist}(x_0, S)$ y $F|_S \equiv 0$.

DEMOSTRACIÓN: Notemos primero que $S \cap \operatorname{span}(\{x_0\}) = \{0\}$; de lo contrario, existiría un $\tilde{x} \in S \setminus \{0\}$ de la forma $\tilde{x} = \tilde{\alpha}\, x_0$ con algún $\tilde{\alpha} \in \mathbb{R} \setminus \{0\}$, de donde $x_0 = \tilde{\alpha}^{-1}\tilde{x} \in S$, lo que contradice que $d := \operatorname{dist}(x_0, S) > 0$.

Sea $S_1 := S + \operatorname{span}(\{x_0\})$. Por lo tanto, todo $z \in S_1$ puede representarse en la forma $z = x + \alpha\, x_0$, donde $x \in S$ y $\alpha \in \mathbb{R}$ y por lo anterior, aquella representación es única. Podemos entonces definir sin ambigüedades el funcional

$$\begin{aligned} f\colon S_1 &\to \mathbb{R} \\ z &\to f(z) := \alpha\, d, \quad \text{donde } z = x + \alpha\, x_0,\ x \in S,\ \alpha \in \mathbb{R}. \end{aligned}$$

El funcional f es claramente lineal y satisface $f|_S \equiv 0$. Además, dado algún $z \in S_1 \setminus S$, lo podemos descomponer como $z = x + \alpha\, x_0$ con $\alpha \neq 0$ y $x \in S$, de donde $-\alpha^{-1}x \in S$ y, consiguientemente,

$$|f(z)| = |\alpha\, d| = |\alpha|\, d \leq |\alpha| \left\|x_0 + \alpha^{-1}x\right\| = \|\alpha\, x_0 + x\| = \|z\|,$$

lo que muestra que f es acotado y $\|f\|_{S_1'} \leq 1$.

Por otro lado, dado cualquier $\varepsilon > 0$, existe $x_\varepsilon \in S$ tal que $\|x_0 - x_\varepsilon\| < d + \varepsilon$. Se sigue que $f(x_0 - x_\varepsilon) = d$. Luego,

$$\frac{|f(x_0 - x_\varepsilon)|}{\|x_0 - x_\varepsilon\|} > \frac{d}{d+\varepsilon},$$

con lo cual

$$(\forall\, \varepsilon > 0) \quad \|f\|_{S_1'} := \sup_{z \in S_1 \setminus \{0\}} \frac{|f(z)|}{\|z\|} \geq \frac{|f(x_0 - x_\varepsilon)|}{\|x_0 - x_\varepsilon\|} > \frac{d}{d+\varepsilon},$$

de donde, tomando el límite $\varepsilon \to 0^+$, resulta $\|f\|_{S_1'} \geq 1$. Así, $\|f\|_{S_1'} = 1$ y aplicando el teorema de Hahn–Banach en espacios normados (Teorema 2.6) se deduce que existe $F \in V'$ tal que $\|F\|_{V'} = \|f\|_{S_1'} = 1$ y, para todo $z \in S_1$, $F(z) = f(z)$. En particular, $F(x_0) = d$ y, para todo $x \in S$, $F(x) = 0$. $\square$

Ahora procuraremos algunas así llamadas formas geométricas del teorema de Hahn–Banach, que tienen que ver con la separación de conjuntos convexos. En lo que sigue, E es un EVN sobre $\mathbb{R}$.

Definición 2.8

Dado $\alpha \in \mathbb{R}$ y un funcional lineal $f\colon E \to \mathbb{R}$, no nulo y no necesariamente acotado, se define un *hiperplano afín* como

$$H := \{x \in E\colon f(x) = \alpha\},$$

y se dice que H es el hiperplano de ecuación $[f = \alpha]$.

Lema 2.5

El hiperplano $[f = \alpha]$ es cerrado si y solo si f es acotado.

Demostración: Si f es acotado, también es continuo. Entonces, como $\{\alpha\}$ es cerrado, $[f = \alpha] = f^{-1}(\{\alpha\})$ también es cerrado.

Supongamos ahora que $[f = \alpha]$ es cerrado. Se sigue que $E \setminus [f = \alpha]$ es abierto y no vacío, dado que, por definición, f es no nulo. Sea entonces $x_0 \in E \setminus [f = \alpha]$ y supongamos, sin pérdida de generalidad, que $f(x_0) < \alpha$. Sea $r > 0$ un radio tal que $B(x_0, r) \subset E \setminus [f = \alpha]$. Afirmamos que $f(x) < \alpha$ para todo $x \in B(x_0, \alpha)$.

En efecto, *si hubiese* algún $x_1 \in B(x_0, r)$ tal que $f(x_1) \geq \alpha$, en primer lugar tendríamos que $f(x_1) > \alpha$ (pues $x_1 \in B(x_0, r) \subset E \setminus [f = \alpha]$). En segundo lugar, todo el segmento $\{(1-t)x_1 + tx_0 \colon t \in [0,1]\}$ estaría contenido en $B(x_0, r)$ y, por lo tanto,

$$(\forall\, t \in [0,1]) \quad f((1-t)x_1 + tx_0) \neq \alpha.$$

Sin embargo, como $f(x_0) < \alpha < f(x_1)$, se deduce que $t^* := \dfrac{f(x_1) - \alpha}{f(x_1) - f(x_0)} \in (0,1)$ y

$$f((1-t^*)x_1 + t^* x_0) = (1-t^*)f(x_1) + t^* f(x_0) = f(x_1) - t^*(f(x_1) - f(x_0)) = \alpha,$$

lo que contradice que $B(x_0, r) \in E \setminus [f = \alpha]$. Así,

$$(\forall\, x \in B(x_0, r)) \quad f(x) < \alpha \iff (\forall\, z \in B(0,1)) \quad f(x_0 + rz) < \alpha.$$

En particular,

$$(\forall\, x \in E \setminus \{0\})\, (\forall\, \varepsilon \in (0,1)) \quad f\left(x_0 \pm r\varepsilon \frac{x}{\|x\|}\right) < \alpha,$$

que podemos reescribir en la forma

$$(\forall\, x \in E \setminus \{0\})\, (\forall\, \varepsilon \in (0,1)) \quad |f(x)| < \frac{\alpha - f(x_0)}{r\varepsilon} \|x\|\,.$$

Tomando el límite $\varepsilon \to 1^-$ e incluyendo la desigualdad obvia del caso $x = 0$, obtenemos

$$(\forall\, x \in E) \quad |f(x)| \leq \frac{\alpha - f(x_0)}{r} \|x\| \,.$$

□

Definición 2.9

Sean $A, B \subseteq E$. Se dice que el hiperplano $[f = \alpha]$ *separa en el sentido amplio* a A y a B si se tiene

$$(\forall\, x \in A) \quad f(x) \leq \alpha \qquad \text{y} \qquad (\forall\, x \in B) \quad f(x) \geq \alpha.$$

A su vez, se dice que este hiperplano *separa en el sentido estricto* a A y a B si existe $\varepsilon > 0$ tal que

$$(\forall\, x \in A) \quad f(x) \leq \alpha - \varepsilon \qquad \text{y} \qquad (\forall\, x \in B) \quad f(x) \geq \alpha + \varepsilon.$$

Lema 2.6 (*Gauge* de un convexo)

Sea $S \subseteq E$ convexo, abierto, con $0 \in S$. Definimos el *funcional de Minkowski* $p\colon E \to \mathbb{R}$ por

$$(\forall\, x \in E) \quad p(x) := \text{ínf}\{\alpha > 0\colon \alpha^{-1}x \in S\}.$$

Entonces, p es sublineal y existe $M \geq 0$ tal que

$$(\forall\, x \in E) \quad 0 \leq p(x) \leq M\,\|x\|\,.$$

Además,

$$S = \{x \in E\colon p(x) < 1\}.$$

Demostración: Como $0 \in S$, existe $r > 0$ tal que $\overline{B(0,r)} \subseteq S$. Luego, para cada $x \in E \setminus \{0\}$, $\frac{r}{\|x\|}x \in S$, lo cual prueba que el escalar $\|x\|/r$ pertenece al conjunto cuyo ínfimo es $p(x)$. Por lo tanto, $p(x)$ es un número bien definido y $p(x) \leq \|x\|/r$. Esta desigualdad junto el hecho que $p(0) = 0$ implican la desigualdad requerida con $M = 1/r$.

La homogeneidad positiva de p emana de

$$(\forall\, \lambda > 0)\ (\forall\, x \in E) \quad \begin{aligned} p(\lambda x) &= \operatorname{ínf}\{\alpha > 0\colon \alpha^{-1}\lambda x \in S\} \\ &= \operatorname{ínf}\{\alpha > 0\colon (\alpha/\lambda)^{-1} x \in S\} \\ &= \lambda \operatorname{ínf}\{\alpha > 0\colon \alpha^{-1} x \in S\} = \lambda\, p(x). \end{aligned}$$

Nuestro siguiente paso será probar la caracterización $S = \{x \in E\colon p(x) < 1\}$. Si $x \in S$, por ser S abierto, existe $\varepsilon > 0$ tal que todavía $(1+\varepsilon)x \in S$, por lo que $(1+\varepsilon)^{-1} \in \{\alpha > 0\colon \alpha^{-1} x \in S\}$, de donde

$$p(x) \leq \frac{1}{1+\varepsilon} < 1.$$

Para la otra inclusión, sea $x \in E$ con $p(x) < 1$. Se sigue que existe $\alpha \in (0,1)$ tal que $\alpha^{-1} x \in S$. Como S es convexo, $\alpha(\alpha^{-1}x) + (1-\alpha)0 = x \in S$.

Para probar la desigualdad triangular —que es lo que nos falta para que p satisfaga nuestra definición de funcional sublineal (Definición 2.7)—, tomemos $x, y \in E$, arbitrarios. Dado $\varepsilon > 0$ observamos que

$$p\left(\frac{x}{p(x)+\varepsilon}\right) = \frac{p(x)}{p(x)+\varepsilon} < 1 \qquad \text{y} \qquad p\left(\frac{y}{p(y)+\varepsilon}\right) = \frac{p(y)}{p(y)+\varepsilon} < 1,$$

lo cual indica que $\dfrac{x}{p(x)+\varepsilon}, \dfrac{y}{p(y)+\varepsilon} \in S$. Sea $t := \dfrac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon} \in (0,1)$. Entonces, por la convexidad de S

$$\frac{x+y}{p(x)+p(y)+2\varepsilon} = t\frac{x}{p(x)+\varepsilon} + (1-t)\frac{y}{p(y)+\varepsilon} \in S.$$

Por lo tanto,

$$1 > p\left(\frac{x+y}{p(x)+p(y)+2\varepsilon}\right) = \frac{p(x+y)}{p(x)+p(y)+2\varepsilon}.$$

Como esto ocurre para todo $\varepsilon > 0$, concluimos que $p(x+y) \leq p(x) + p(y)$. $\square$

Lema 2.7

Sea $S \subsetneq E$ convexo, abierto y no vacío y sea $x_0 \in E$ tal que $x_0 \notin S$. Entonces existe $f \in E' \setminus \{0\}$ tal que

$$(\forall\, x \in S) \quad f(x) < f(x_0);$$

esto es, el hiperplano $[f = f(x_0)]$ separa a $\{x_0\}$ y a S en el sentido amplio.

Demostración: Supongamos, sin pérdida de generalidad (¿seguro?), que $0 \in S$ y definamos el funcional de Minkowski asociado $p\colon E \to \mathbb{R}$ como en el enunciado del Lema 2.6:

$$(\forall\, x \in E) \quad p(x) := \text{ínf}\{\alpha > 0\colon \alpha^{-1} x \in E\}$$

Definimos el subespacio $G := \operatorname{span}(\{x_0\}) = \{t\, x_0\colon t \in \mathbb{R}\}$ y sea $g\colon G \to \mathbb{R}$ el funcional definido por $g(t\, x_0) := t$ para todo $t\, x_0 \in \mathbb{R}$. Como $x_0 \notin S$, del Lema 2.6 se desprende que $p(x_0) \geq 1$. Entonces,

$$(\forall\, t > 0) \quad g(t\, x_0) = t \leq t\, p(x_0) = p(t\, x_0).$$

A su vez,

$$(\forall\, t \leq 0) \quad g(t\, x_0) = t \leq 0 \leq p(t\, x_0).$$

Consecuentemente,

$$(\forall\, x \in G) \quad g(x) \leq p(x).$$

Aplicando la forma analítica del teorema de Hahn–Banach (Teorema 2.5) se deduce que existe un funcional lineal $f\colon E \to \mathbb{R}$ tal que

$$(\forall\, x \in G) \quad f(x) = g(x) \qquad \text{y} \qquad (\forall\, x \in E) \quad f(x) \le p(x).$$

En particular, $f(x_0) = g(x_0) = 1$, lo que confirma que $f \not\equiv 0$. Además, por el Lema 2.6 sabemos que existe $M \ge 0$ tal que, para todo $x \in E$, $0 \le p(x) \le M\,\|x\|$. Luego, $f(x) \le M\,\|x\|$ para todo $x \in E$. A su vez, $-f(x) = f(-x) \le M\,\|-x\| = M\,\|x\|$, con lo cual

$$(\forall\, x \in E) \quad |f(x)| \le M\,\|x\|\,,$$

probando así que $f \in E'$.

Por último, utilizando la caracterización del convexo S dada en el Lema 2.6, se concluye que

$$(\forall\, x \in S) \quad f(x) \le p(x) < 1 = f(x_0).$$

$\square$

Ahora estamos en condiciones de enunciar y demostrar las dos versiones geométricas del Teorema de Hahn–Banach.

Teorema 2.9 (Primera versión geométrica del teorema de Hahn–Banach)

Sean $A, B \subseteq E$ convexos, no vacíos y disjuntos, tales que A es además abierto. Entonces existe un hiperplano cerrado $[f = \alpha]$, con $f \not\equiv 0$, que separa a A y a B en el sentido amplio.

Demostración: Sea $S = A - B := \{x - y \colon x \in A,\ y \in B\}$. El conjunto S hereda la convexidad de A y de B. Además, S es abierto porque $S = \bigcup_{y \in B}(A - \{y\})$ y cada uno de los conjuntos de la forma $A - \{y\}$ es abierto porque A lo es. También $0 \notin S$ porque $A \cap B = \emptyset$. Aplicando el Lema 2.7 se deduce que existe $f \in E' \setminus \{0\}$ tal que, para todo $z \in S$, $f(z) < f(0) = 0$; esto es,

$$(\forall\, x \in A)\ (\forall\, y \in B) \quad f(x) < f(y).$$

Definiendo $m := \sup_{x \in A} f(x)$ y $M := \text{ínf}_{x \in B} f(y)$ se sigue que, para cualquier $\alpha \in [m, M]$, el hiperplano $[f = \alpha]$ separa a A y a B en el sentido amplio. □

Teorema 2.10 (Segunda versión geométrica del teorema de Hahn–Banach)

Sean $A, B \subseteq E$ convexos, no vacíos y disjuntos, tales que A es cerrado y B es compacto. Entonces, existe un hiperplano cerrado $[f = \alpha]$ con $f \not\equiv 0$, que separa a A y a B en el sentido estricto.

Demostración: Dado $\varepsilon > 0$, definamos los conjuntos no vacíos

$$A_\varepsilon := A + B(0, \varepsilon) \quad \text{y} \quad B_\varepsilon := B + B(0, \varepsilon).$$

Es fácil ver, por la convexidad de A, B y $B(0, \varepsilon)$, que A_ε y B_ε son convexos. Además, por las caracterizaciones $A_\varepsilon = \bigcup_{x \in A}(\{x\} + B(0, \varepsilon))$ y $B_\varepsilon = \bigcup_{x \in B}(\{x\} + B(0, \varepsilon))$, A_ε y B_ε son abiertos.

Probaremos ahora que existe $\varepsilon > 0$ tales que A_ε y B_ε son disjuntos. En efecto, supongamos por contradicción que no es el caso. Entonces, para todo $n \in \mathbb{N}$ el conjunto $A_{1/n} \cap B_{1/n}$ es no vacío, lo que implica que existen $x_n \in A$, $y_n \in B$ y $u_n, v_n \in B(0, 1/n)$ tales que $x_n + u_n = y_n + v_n$. Se sigue que

$$\|x_n - y_n\| = \|u_n - v_n\| < \frac{2}{n}.$$

Como B es compacto, existe una subsucesión $\{y_{\phi(n)}\}_{n \in \mathbb{N}}$ de $\{y_n\}_{n \in \mathbb{N}}$ que converge en E a algún $y \in B$. Por la desigualdad anterior, $\{x_{\phi(n)}\}_{n \in \mathbb{N}}$ también converge a y. Como A es cerrado, necesariamente $y \in A$, lo que contradice la hipótesis de que A y B son disjuntos.

Entonces, fijando $\varepsilon > 0$ como alguno de los que mantiene a A_ε y B_ε disjuntos y aplicándoles la primera versión geométrica del teorema de Hahn–Banach (Teorema 2.9), se deduce que existe un hiperplano cerrado de ecuación $[f = \alpha]$, con $f \not\equiv 0$, que los separa en el sentido amplio; esto es,

$$(\forall\, x \in A)\ (\forall\, y \in B)\ (\forall\, u, v \in B(0,\varepsilon)) \quad f(x+u) \le \alpha \le f(y+v).$$

En particular,

$$\begin{aligned}&(\forall\, x \in A)\ (\forall\, y \in B)\ (\forall\, z \in B(0,1))\\ &\qquad f(x) + \varepsilon f(z) = f(x+\varepsilon z) \le \alpha \le f(y - \varepsilon z) = f(y) - \varepsilon f(z).\end{aligned}$$

Escogiendo a $z^* \in B(0,1)$ tal que $f(z^*) > 0$ se deduce que

$$(\forall\, x \in A)\ (\forall\, y \in B) \quad f(x) \le \alpha - \varepsilon f(z^*) \quad \text{y} \quad f(y) \ge \alpha + \varepsilon f(z^*),$$

lo que prueba que el hiperplano $[f = \alpha]$ separa a A y a B en el sentido estricto. $\square$

La hipótesis de compacidad no se puede relajar. Por ejemplo, en $E = \mathbb{R}^2$ no existe un hiperplano que separe estrictamente a los conjuntos cerrados, convexos y disjuntos $A = \left\{x \in \mathbb{R}^2 \colon x_1 \ge 1,\ x_1^{-1} \le x_2 \le 1\right\}$ y $B = \left\{x \in \mathbb{R}^2 \colon x_1 \ge 1,\ -1 \le x_2 \le -x_1^{-1}\right\}$, ninguno de los cuales es compacto.