

Versión analítica del teorema de Hahn-Banach

Los tres teoremas considerados como "Principios Fundamentales del Análisis Funcional" llevan el nombre de Stefan Banach: el *teorema de Hahn-Banach*, el de *Banach-Steinhaus* y el de la *aplicación abierta*, también conocido como *teorema de Banach-Schauder*. Vamos a probar aquí el primero de ellos, que es la pieza clave para el estudio de la dualidad en espacios normados. Como cualquier resultado importante en Matemáticas, pero muy especialmente en este caso, el teorema de Hahn-Banach admite numerosas versiones equivalentes, que se aplican en campos muy diversos. En este tema vemos la "versión analítica", que nos permitirá avanzar en el estudio de la dualidad. Más adelante veremos una "versión geométrica", que se caracteriza precisamente por eso, por tener clara interpretación geométrica.

6.1. Enunciado y demostración del teorema

Antes de enunciar el teorema, el siguiente razonamiento nos proporciona una motivación. Dado un espacio normado $X \neq \{0\}$, sin más información, cabe preguntar si existen funcionales lineales continuos no nulos en X. En los muchos casos concretos que conocemos, siempre hemos podido describir el espacio dual X^* , al menos para comprobar que $X^* \neq \{0\}$. Pero la cuestión es si podemos asegurar que $X^* \neq \{0\}$ para todo espacio normado $X \neq \{0\}$.

Para abordar esta pregunta, podemos usar un subespacio de dimensión finita $M \subset X$, pues abundan los funcionales lineales no nulos en M, y todos ellos son continuos. Tomando $g \in M^*$ con $\|g\| = 1$, tenemos $\|g(y)\| \le \|y\|$ para todo $y \in M$, Si pudiésemos extender g a todo X, manteniendo su linealidad de forma que la extensión siga verificando la misma desigualdad, tendríamos un funcional lineal continuo en X que no es idénticamente nulo, porque extiende a g. Pues bien, eso es lo que, con hipótesis bastante más generales, nos garantiza la versión analítica del teorema de Hahn-Banach.

6.1.1. Enunciado del teorema

En vez de una norma, usaremos una aplicación con propiedades bastante menos restrictivas. Concretamente, un **funcional sublineal** en un espacio vectorial X es una aplicación $\phi: X \to \mathbb{R}$ que verifica la desigualdad triangular y es homogénea por homotecias, es decir, verifica las siguientes dos condiciones:

$$\varphi(x+y) \leqslant \varphi(x) + \varphi(y)$$
 y $\varphi(rx) = r\varphi(x)$ $\forall x, y \in X, \ \forall r \in \mathbb{R}^+$

Observamos que toda seminorma en X es un funcional sublineal, pero el recíproco es falso. Por ejemplo, si f es un funcional lineal en X, definiendo $\varphi(x)=\operatorname{Re} f(x)$ para todo $x\in X$, vemos que φ es un funcional sublineal en X, que sólo es una seminorma cuando f=0, pues en otro caso, φ toma valores negativos. Por otra parte, en el caso $X=\mathbb{R}$, definiendo $\varphi(x)=\max\{0,x\}$ para todo $x\in\mathbb{R}$, se obtiene un funcional sublineal en \mathbb{R} , que no es lineal y tampoco es una seminorma. Enunciamos ya la versión analítica del teorema de Hahn-Banach:

Teorema (Hahn 1927, Banach 1929). Sea X un espacio vectorial $y \varphi$ un funcional sublineal en X. Sea M un subespacio de X y g un funcional lineal en M, que está dominado por φ , en el siguiente sentido:

$$\operatorname{Re} g(y) \leqslant \varphi(y) \qquad \forall y \in M$$

Entonces existe un funcional lineal f en X, que extiende a g y está dominado por ϕ , es decir,

$$f(y) = g(y) \quad \forall y \in M$$
 y Re $f(x) \leq \varphi(x) \quad \forall x \in X$

Si φ es una seminorma, se tiene de hecho

$$|f(x)| \leqslant \varphi(x) \qquad \forall x \in X$$

Al parecer, la versión demostrada por Hahn suponía que φ es una norma en X; esta versión más general es la aportación de Banach y resulta esencial como se verá, para establecer las versiones geométricas del teorema. Por otra parte, usar la función parte real es lo que permite hacer un enunciado común para el caso real y el caso complejo; naturalmente, en el caso $\mathbb{K} = \mathbb{R}$ dicha función es la identidad. Dividiremos la demostración en tres etapas.

6.1.2. Caso real, primera extensión

Empezamos considerando el caso $\mathbb{K} = \mathbb{R}$ y sólo extendemos el funcional g al subespacio que se obtiene sumando una recta a M. Esta es la etapa más importante, pero no es difícil.

Fijado $x \in X \setminus M$, consideramos el subespacio $Z = M \oplus \mathbb{R} x$, y queremos definir un funcional lineal h en Z que extienda a g y siga dominado por φ . Obviamente debemos definir

$$h(y + \lambda x) = g(y) + \lambda \alpha \qquad \forall y \in M, \ \forall \lambda \in \mathbb{R}$$

para conveniente $\alpha \in \mathbb{R}$. Cualquiera que sea α , está claro que h es lineal y extiende a g, luego el problema es encontrar $\alpha \in \mathbb{R}$, de forma que h esté dominado por φ , es decir, que verifique:

$$g(y) + \lambda \alpha \leqslant \varphi(y + \lambda x) \qquad \forall y \in M, \ \forall \lambda \in \mathbb{R}$$
 (1)

Para $\lambda \in \mathbb{R}^+$, dividiendo por λ los dos miembros de (1) y usando que φ es homogéneo por homotecias, vemos que (1) toma la forma

$$g\left(\frac{y}{\lambda}\right) + \alpha \leqslant \phi\left(\frac{y}{\lambda} + x\right)$$

Observamos ahora que $u = y/\lambda$ es un vector de M tan arbitrario como y. Por tanto, $\alpha \in \mathbb{R}$ verifica la desigualdad (1) para todo $\lambda \in \mathbb{R}^+$ y todo $y \in M$ si, y sólo si,

$$\alpha \leqslant \varphi(u+x) - g(u) \qquad \forall u \in M$$
 (2)

Para $\lambda \in \mathbb{R}^-$ dividimos por $-\lambda$ ambos miembros de (1) y usando de que ϕ es homogéneo por homotecias, (1) toma la forma

$$g\left(-\frac{y}{\lambda}\right) - \alpha \leqslant \varphi\left(-\frac{y}{\lambda} - x\right)$$

Como $w = -y/\lambda$ es un vector de M tan arbitrario como y, vemos que $\alpha \in \mathbb{R}$ verifica la desigualdad (1) para todo $\lambda \in \mathbb{R}^-$ y todo $y \in M$ si, y sólo si,

$$\alpha \geqslant g(w) - \varphi(w - x) \qquad \forall w \in M \tag{3}$$

Obsérvese por último que, para $\lambda=0$, la desigualdad (1) se cumple por hipótesis, cualquiera que sea $\alpha\in\mathbb{R}$. En resumen, hemos probado que esta etapa de la demostración estará concluida si encontramos $\alpha\in\mathbb{R}$ verificando (2) y (3).

Ahora bien, para cualesquiera $u, w \in M$, tenemos por hipótesis:

$$g(u) + g(w) = g(u+w) \leqslant \varphi(u+w) \leqslant \varphi(u+x) + \varphi(w-x)$$

donde se ha usado que φ verifica la desigualdad triangular. Así pues, tenemos

$$g(w) - \varphi(w - x) \leq \varphi(u + x) - g(u)$$
 $\forall w, u \in M$

de donde deducimos claramente que

$$\sup \left\{ g\left(w\right) - \varphi\left(w - x\right) : w \in M \right\} \leqslant \inf \left\{ \varphi\left(u + x\right) - g\left(u\right) : u \in M \right\} \tag{4}$$

Si ahora α es cualquier número real comprendido entre los dos miembros de (4), está bien claro que α verifica (2) y (3), luego también (1), como queríamos.

Nótese que si (4) es una igualdad, sólo hay una elección posible de α . El razonamiento anterior permite en la práctica discutir la posible unicidad del funcional f cuya existencia afirma el teorema, un asunto del que no nos vamos a ocupar.

6.1.3. Caso real, extensión definitiva

Hablando intuitivamente, lo que haremos para concluir la demostración en el caso $\mathbb{K} = \mathbb{R}$, es iterar la extensión realizada en la primera etapa, añadiendo en cada paso una recta al subespacio obtenido en el paso anterior, hasta llegar a X.

Está claro que necesitamos tantos pasos como indique la dimensión de X/M, que puede ser infinita, e incluso no numerable. Hemos de hacer por tanto una "inducción transfinita", que se formaliza usando el Lema de Zorn.

Consideramos el conjunto \mathcal{E} de todos los pares ordenados de la forma (Z,h), donde Z es un subespacio de X, que contiene a M, y h un funcional lineal en Z, que extiende a g y está dominado por φ . Para (Z_1,h_1) , $(Z_2,h_2) \in \mathcal{E}$ definimos

$$(Z_1, h_1) \leqslant (Z_2, h_2) \qquad \iff \qquad Z_1 \subset Z_2 \quad \text{y} \quad h_2 \big|_{Z_1} = h_1$$
 (5)

y es evidente que de esta forma se obtiene una relación de orden en \mathcal{E} .

Sea $\mathcal{E}_0 = \big\{ (Z_j, h_j) : j \in J \big\}$ un subconjunto totalmente ordenado de \mathcal{E} , y veamos que \mathcal{E}_0 tiene un mayorante. Observamos que $Z = \bigcup_{j \in J} Z_j$ es un subespacio de X, que evidentemente contiene a Z_j para todo $j \in J$, y en particular contiene a M. En efecto, para $u, v \in Z$ y $\lambda \in \mathbb{R}$ tomamos $j, k \in J$ tales que $u \in Z_j$ y $v \in Z_k$. Como \mathcal{E}_0 está totalmente ordenado, (5) nos dice que $Z_j \subset Z_k$ o $Z_k \subset Z_j$, y deducimos que se ha de tener, o bien $u, v \in Z_k$ y $\lambda u + v \in Z_k \subset Z$, o bien $u, v \in Z_j$ y $\lambda u + v \in Z_j \subset Z$.

Definimos ahora $h: Z \to \mathbb{R}$ como sigue. Dado $u \in Z$, tomamos $j \in J$ tal que $u \in Z_j$ y comprobamos que podemos definir $h(u) = h_j(u)$. En efecto, si también $u \in Z_k$ con $k \in J$, de nuevo por estar \mathcal{E}_0 totalmente ordenado, (5) nos dice que, o bien h_j extiende a h_k , o bien h_k extiende a h_j , luego en ambos casos $h_j(u) = h_k(u)$. Por definición de h, tenemos $h \mid_{Z_j} = h_j$ para todo $j \in J$, y esto permitirá comprobar que h es lineal. Si $u, v \in Z$ y $\lambda \in \mathbb{R}$, hemos visto que existe $j \in J$ tal que $u, v \in Z_j$, con lo que $\lambda u + v \in Z_j$, y tenemos

$$h(\lambda u + v) = h_j(\lambda u + v) = \lambda h_j(u) + h_j(v) = \lambda h(u) + h(v)$$

Esto prueba que h es lineal, y es obvio que h extiende a g, pues para cualquier $j \in J$, sabemos que h extiende a h_j , que a su vez extiende a g. Finalmente, para $u \in Z$, basta tomar $j \in J$ tal que $u \in Z_j$, para tener $h(u) = h_j(u) \leqslant \varphi(u)$, luego h está dominado por φ . En resumen, tenemos $(Z,h) \in \mathcal{E}$ que verifica $(Z_j,h_j) \leqslant (Z,h)$ para todo $j \in J$, luego \mathcal{E}_0 tiene un mayorante, como queríamos comprobar.

El Lema de Zorn nos proporciona un elemento maximal $(Z,f) \in \mathcal{E}$, y sólo queda comprobar que Z=X, pues entonces f será el funcional que buscamos, un funcional lineal en X que extiende a g y está dominado por φ . Ahora bien, si $Z \neq X$, podemos tomar $x \in X \setminus Z$, y la primera etapa de la demostración nos da un funcional lineal \widehat{f} , en el subespacio $\widehat{Z}=Z\oplus \mathbb{R} x$, que extiende a f y está dominado por φ . Entonces $(\widehat{Z},\widehat{f})\in\mathcal{E}$ verifica que $(Z,f)\leqslant(\widehat{Z},\widehat{f})$, con $\widehat{Z}\neq Z$, lo que contradice la maximalidad de (Z,f). Esto concluye la demostración del teorema, en el caso real.

Todo el razonamiento anterior es bastante rutinario, simplemente ocurre que el lema de Zorn es el instrumento que permite formalizar la idea intuitiva consistente en iterar un proceso hasta concluirlo, sin importar el número de iteraciones que sean necesarias. Conviene resaltar que, cuando usamos el lema de Zorn para formalizar un proceso infinito, nuestro razonamiento no es constructivo, no controlamos el resultado del proceso. En nuestro caso, hemos probado la existencia de un funcional f, que no conocemos explícitamente. Esto contrasta claramente con lo ocurrido en la primera etapa de la demostración, en la que el funcional extendido se pudo definir explícitamente.

6.1.4. Fin de la demostración

Para completar la prueba del teorema nos queda considerar el caso complejo y comprobar la última afirmación, referente al caso en que φ es una seminorma.

Si Z es un espacio vectorial complejo, el producto por escalares de X es una aplicación definida en $\mathbb{C} \times X$, que podemos restringir a $\mathbb{R} \times X$. Es claro que de esta forma se obtiene un espacio vectorial real, llamado **espacio real subyacente** a Z, que se denota por $Z_{\mathbb{R}}$. Vamos ahora a aclarar la relación entre los funcionales lineales en ambos espacios. Si h es un funcional lineal en Z, es evidente que Re h es un funcional lineal en $Z_{\mathbb{R}}$. Además h queda determinado por su parte real, ya que

$$\operatorname{Im} h(z) = \operatorname{Re} (-ih(z)) = -\operatorname{Re} h(iz) \quad \forall z \in \mathbb{Z}$$

Recíprocamente, si h_0 es un funcional lineal en $Z_{\mathbb{R}}$, definimos

$$h(z) = h_0(z) - i h_0(iz)$$
 $\forall z \in Z$

y comprobamos que h es un funcional lineal en Z, que obviamente verifica Re $h=h_0$. En efecto, es claro que h(u+v)=h(u)+h(v) y $h(\alpha u)=\alpha h(u)$ para $u,v\in Z$ y $\alpha\in\mathbb{R}$. Pero además, para todo $z\in Z$ se tiene

$$h(iz) = h_0(iz) - ih_0(-z) = i(h_0(z) - ih_0(iz)) = ih(z)$$

de donde se deduce ya claramente que h es lineal. Podemos por tanto enunciar:

■ Si Z es un espacio vectorial complejo, los funcionales lineales en $Z_{\mathbb{R}}$ son las partes reales de los funcionales lineales en Z.

Es ya fácil probar la versión analítica del teorema de Hahn-Banach en el caso complejo. Sea pues X un espacio vectorial complejo, provisto de un funcional sublineal $\varphi: X \to \mathbb{R}$, que también es un funcional sublineal en $X_{\mathbb{R}}$. Sea también M un subespacio de X y $g: M \to \mathbb{C}$ un funcional lineal dominado por φ , es decir, que verifica $\operatorname{Re} g(y) \leqslant \varphi(y)$ para todo $y \in M$. Entonces el espacio vectorial real $M_{\mathbb{R}}$ es subespacio de $X_{\mathbb{R}}$, y tomando $g_0 = \operatorname{Re} g$, tenemos que g_0 es un funcional lineal en $M_{\mathbb{R}}$ dominado por φ , esto es, $g_0(y) \leqslant \varphi(y)$ para todo $y \in M$. Como el teorema en el caso real ya está demostrado, existe un funcional lineal $f_0: X_{\mathbb{R}} \to \mathbb{R}$ que extiende a g_0 y está dominado por φ . Existe ahora un (único) funcional lineal $f: X \to \mathbb{C}$, tal que $f_0 = \operatorname{Re} f$. Entonces f está dominado por φ , puesto que se tiene $\operatorname{Re} f(x) = f_0(x) \leqslant \varphi(x)$ para todo $x \in X$. Concluimos viendo que f extiende a g, ya que tanto $f|_{M}$ como g quedan determinados por sus respectivas partes reales. Concretamente, para todo $y \in M$, se tiene

$$f(y) = f_0(y) - i f_0(iy) = g_0(y) - i g_0(iy) = g_0(y)$$

Sólo queda probar la última afirmación del teorema, pero esto es muy sencillo: si φ es de hecho una seminorma, para cada $x \in X$ podemos escribir $|f(x)| = \lambda f(x)$ con $\lambda \in \mathbb{K}$ y $|\lambda| = 1$. Se tiene entonces

$$|f(x)| = \lambda f(x) = f(\lambda x) = \text{Re } f(\lambda x) \le \varphi(\lambda x) = |\lambda| \varphi(x) = \varphi(x)$$

donde hemos usado que, evidentemente $f(\lambda x) \in \mathbb{R}$, y que φ es una seminorma.

6.2. Teorema de extensión en espacios normados

El resto de este tema se dedica a obtener las consecuencias más inmediatas del teorema recién demostrado. La primera contesta afirmativamente, a plena generalidad, la pregunta que habíamos planteado como motivación. En el contexto de los espacios normados, el teorema de Hahn-Banach se usa casi siempre de esta forma:

Teorema de extensión Hahn-Banach. Sea X un espacio normado, M un subespacio de X $y \ g \in M^*$. Entonces existe $f \in X^*$ tal que f extiende a g y verifica que ||f|| = ||g||. Se dice que f es una **extensión Hahn-Banach** de g.

Demostración. Definiendo $\varphi(x) = \|g\| \|x\|$ para todo $x \in X$, está bien claro que φ es una seminorma en X, de hecho una norma salvo en el caso trivial g = 0. La continuidad de g nos dice claramente que Re $g(y) \le |g(y)| \le \varphi(y)$ para todo $y \in Y$. Por tanto, la versión analítica del teorema de Hahn-Banach nos da un funcional lineal f en X que verifica:

$$f(y) = g(y) \quad \forall y \in Y \qquad y \qquad |f(x)| \le \varphi(x) = ||g|| ||x|| \quad \forall x \in X$$

Vemos que $f \in X^*$ con $||f|| \le ||g||$ pero, como f extiende a g, también se tiene $||f|| \ge ||g||$, luego ambas normas coinciden.

El hecho de que todo funcional lineal continuo, en cualquier subespacio M de un espacio normado X, se extienda para obtener un funcional lineal continuo en todo el espacio X, ya es bastante importante. Como ya se dijo, nos garantiza por ejemplo la abundancia de funcionales lineales continuos no nulos en cualquier espacio normado $X \neq \{0\}$. Basta tomar un subespacio de dimensión finita $M \subset X$ y un funcional lineal no nulo g en M, que automáticamente es continuo, y el teorema anterior nos da $f \in X^*$ que extiende a g, luego $f \neq 0$.

En general está claro que, al extender un funcional lineal continuo, su norma no puede disminuir, y el teorema anterior nos asegura que siempre existe una extensión Hahn-Banach, es decir, que la extensión puede hacerse sin que la norma aumente. La existencia de este tipo de extensión tendrá más adelante consecuencias importantes, pero por ahora vamos a considerar dos situaciones en las que basta disponer de una extensión continua.

En primer lugar aprovechamos simplemente el hecho de que $X^* \neq \{0\}$ para todo espacio normado $X \neq \{0\}$. Para cualquier otro espacio normado $Y \neq \{0\}$, esto nos permite definir operadores lineales continuos no nulos de X en Y, es decir, podemos ver que $L(X,Y) \neq \{0\}$. En efecto, fijados $y \in Y \setminus \{0\}$ y $f \in X^* \setminus \{0\}$, basta definir

$$T(x) = f(x)y \qquad \forall x \in X$$

Está claro que $T: X \to Y$ es un operador lineal que verifica

$$||T(x)|| = |f(x)| ||y|| \le ||f|| ||x|| ||y|| \quad \forall x \in X$$

luego $T \in L(X,Y)$ con $||T|| \le ||f|| ||y||$. De hecho se verifica la igualdad, pues si B es la bola unidad de X, se tiene

$$||T|| = \sup \{ |f(x)| ||y|| : x \in B \} = ||y|| \sup \{ |f(x)| : x \in B \} = ||y|| ||f||$$

Aprovechando la idea anterior, podemos ahora probar algo que quedó aplazado al estudiar la complitud de los espacios de operadores. Vimos que si Y es un espacio de Banach, el espacio de operadores L(X,Y) es completo, para todo espacio normado X, pero ahora podemos probar el recíproco.

■ Sean X e Y espacios normados, con $X \neq \{0\}$. Si el espacio de operadores L(X,Y) es completo, entonces Y es completo.

Dada una sucesión de Cauchy $\{y_n\}$ en Y, fijamos $f \in X^* \setminus \{0\}$ y definimos

$$T_n(x) = f(x)y_n \quad \forall x \in X, \ \forall n \in \mathbb{N}$$

obteniendo una sucesión $\{T_n\}$ en L(X,Y) que también es de Cauchy, ya que

$$||T_n - T_m|| = ||f|| ||y_n - y_m|| \qquad \forall n \in \mathbb{N}$$

Por ser L(X,Y) completo, la sucesión $\{T_n\}$ converge en L(X,Y) a un operador T. Puesto que

$$||T_n(x) - T(x)|| \le ||T_n - T|| ||x||$$
 $\forall x \in X, \forall n \in \mathbb{N}$

deducimos que la sucesión $\{T_n(x)\}=\{f(x)y_n\}$ converge en Y para todo $x\in X$, luego basta tomar $x\in X$ con f(x)=1, para concluir que $\{y_n\}$ converge.

Como otra consecuencia del teorema de extensión, probamos un resultado sobre subespacios complementados, que es la contrapartida del obtenido al estudiar los espacios normados de dimensión finita. Vimos en su momento que, si M es un subespacio cerrado, de codimensión finita en un espacio normado X, entonces M está complementado en X, y de hecho, todo complemento algebraico de M en X es un complemento topológico. Si Z es cualquiera de estos complementos, está claro que Z tiene dimensión finita. Puede pensarse, por tanto, que *cualquier* subespacio de dimensión finita $Z \subset X$, también va a estar complementado en X. El resultado anterior no permite obtener tal conclusión, pues si M es un complemento algebraico de Z en X, ciertamente M tiene codimensión finita en X, pero M puede no ser cerrado. Recuérdese que, en todo espacio normado de dimensión infinita, existe un funcional lineal que no es continuo, luego existe un subespacio de codimensión 1 que no es cerrado. No obstante, el teorema de Hahn-Banach sí nos va a permitir probar que los subespacios de dimensión finita de cualquier espacio normado, siempre están complementados.

 Si M es un subespacio de dimensión finita de un espacio normado X, entonces M está complementado en X.

En efecto, sea $n \in \mathbb{N}$ la dimensión de M y $\{u_1, u_2, \dots, u_n\}$ una base algebraica de M. Está claro que las coordenadas de cada vector $y \in M$, con respecto a dicha base, dependen linealmente de y, es decir, existen g_1, g_2, \dots, g_n , funcionales lineales en M, tales que

$$y = \sum_{k=1}^{n} g_k(y) u_k \qquad \forall y \in M$$
 (6)

Como M tiene dimensión finita, para cada $k \in \{1, 2, ..., n\}$ tenemos $g_k \in M^*$ y el teorema anterior nos proporciona un funcional $f_k \in X^*$ que extiende a g_k . Definiendo ahora

$$P(x) = \sum_{k=1}^{n} f_k(x) u_k \qquad \forall x \in X$$

tenemos claramente un operador lineal $P: X \to X$, que es continuo, pues se obtiene como suma de funciones continuas. Para todo $x \in X$, se tiene $P(x) \in \text{Lin}\{u_1, u_2, \dots, u_n\} = M$, luego vemos que $P(X) \subset M$. Pero por otra parte, como f_k extiende a g_k para todo $k \in \{1, 2, \dots, n\}$, la igualdad (6) nos dice que P(y) = y para todo $y \in M$. Así pues, $P \circ P = P$ y P(X) = M, es decir, P es una proyección lineal continua de X sobre M.

Resaltamos la diferencia entre este resultado y el que conocíamos para los subespacios de codimensión finita. Si M es un subespacio de codimensión finita en un espacio normado X, suponiendo que M es cerrado en X, tenemos que todo complemento algebraico de M en X es un complemento topológico. En cambio, si M tiene dimensión finita, sabemos que M siempre es cerrado, y el resultado anterior nos da un complemento topológico de M en X, pero no es cierto que todo complemento algebraico de M en X sea un complemento topológico.