Tema 8

Versión geométrica del teorema de Hahn-Banach

Abordamos ahora la interpretación geométrica del teorema de Hahn-Banach, que consistirá en encontrar condiciones suficientes para separar dos subconjuntos de un espacio vectorial. Empezaremos aclarando en qué consiste esta *separación* y qué tipo de resultados cabe esperar. Obtendremos un teorema general de separación de conjuntos convexos en espacios vectoriales, equivalente a la versión analítica del teorema de Hahn-Banach, del que deduciremos una serie de consecuencias interesantes para espacios normados.

8.1. Motivación

En términos muy genéricos, podríamos decir que el estudio de la dualidad pretende obtener información sobre un espacio a partir de su dual. Hemos visto ya algunos casos en los que esto es posible. Por ejemplo, dado un espacio normado X, y puntos $x,y \in X$ con $x \neq y$, sabemos que existe $f \in X^*$ con $\|f\| = 1$ y $f(x-y) = \|x-y\|$, luego en particular, $\operatorname{Re} f(y) < \operatorname{Re} f(x)$. Se dice que el funcional f "separa" y de x, en el sentido de que distingue dichos puntos, nos hace ver que son distintos. Como esto es válido para cualesquiera $x,y \in X$ con $x \neq y$, aunque el funcional $f \in X^*$ depende de dichos puntos, decimos que X^* separa los puntos de X.

Para poner otro ejemplo, igualmente conocido pero más relevante, dados un subespacio cerrado M del espacio normado X, y un punto $x \in X \setminus M$, existe $f \in M^\perp$ con $\|f\| = 1$, que verifica $f(x) = d(x,M) \in \mathbb{R}^+$. En particular, $\operatorname{Re} f(y) < \operatorname{Re} f(x)$ para todo $y \in M$. De nuevo diremos que f "separa" el subespacio M del punto x, pues pone de manifiesto que $x \notin M$.

La idea de separación que vamos a estudiar generaliza en varios sentidos lo que ocurre en los ejemplos anteriores. Por una parte, trabajamos en un espacio vectorial X, en el que no usamos de entrada ninguna norma. Por otra, en vez de puntos o subespacios, utilizamos subconjuntos convexos de X. Además, sólo exigimos desigualdades no estrictas, que se manejan con más comodidad, sin excluir que, en casos concretos, nos interese obtener desigualdades estrictas, que describan tipos de separación más eficaces.

Sean pues A y B subconjuntos no vacíos y convexos de un espacio vectorial X. Dado un funcional lineal no nulo $f : X \to \mathbb{K}$, decimos que f **separa** A de B, cuando

$$\operatorname{Re} f(a) \leqslant \operatorname{Re} f(b) \qquad \forall\, a \in A\,,\ \forall\, b \in B \tag{1}$$

o equivalentemente,

$$\sup \operatorname{Re} f(A) \leqslant \operatorname{ínf} \operatorname{Re} f(B) \tag{2}$$

En tal caso, tomando $\alpha \in \mathbb{R}$ de forma que $\sup f(A) \leqslant \alpha \leqslant \operatorname{ínf} f(B)$, tenemos claramente

$$\operatorname{Re} f(a) \leqslant \alpha \leqslant \operatorname{Re} f(b) \qquad \forall\, a \in A\,,\ \forall\, b \in B \tag{3}$$

Recíprocamente, si existe $\alpha \in \mathbb{R}$ que verifica (3), deducimos obviamente (1) y (2). Conviene hacer varias aclaraciones acerca de esta definición.

En primer lugar, los conjuntos A y B no están en situación simétrica porque, en todas las desigualdades anteriores, hemos elegido situar A en el primer miembro y B en el segundo. Sin embargo, si el funcional f separa A de B, está claro que $-f$ separa B de A. Por tanto, a la hora de *separar* los conjuntos A y B, es decir, de encontrar un funcional que separe uno del otro, los papeles de ambos conjuntos sí son simétricos.

Por otra parte, si X es un espacio vectorial complejo, recordamos que las partes reales de los funcionales lineales en X no son, ni más ni menos que, los funcionales lineales en el espacio vectorial real $X_{\mathbb{R}}$ subyacente a X. Un funcional lineal $f \neq 0$ separa A de B en el espacio complejo X si, y sólo si, el funcional lineal $\operatorname{Re} f \neq 0$ separa A de B en el espacio real $X_{\mathbb{R}}$. Así pues, el estudio de todas las cuestiones referentes a la separación de conjuntos convexos se puede siempre reducir al caso real.

Cuando los conjuntos A y B son disjuntos, un funcional lineal $f \neq 0$ que separe A de B puede no ponerlo de manifiesto, porque (1) no implica que los conjuntos $f(A)$ y $f(B)$ tengan que ser disjuntos. Esto se debe a que hemos preferido usar en (1) desigualdades no estrictas, para tener un tipo de separación más general. De esta forma, en ocasiones podremos separar conjuntos que no llegan a ser disjuntos, lo que puede ser útil, como veremos más adelante. A este respecto, conviene resaltar que la condición $f \neq 0$, exigida en la definición anterior, es obligada para evitar trivialidades. Si f es idénticamente nulo, se verifica la condición (1), cualesquiera que sean los conjuntos A y B, lo que obviamente carece de interés.

Para entender geométricamente la separación de dos conjuntos convexos, si $f \neq 0$ es un funcional lineal en X y $\alpha \in \mathbb{R}$, consideremos el conjunto $H = \{x \in X \,:\, \operatorname{Re} f(x) = \alpha\}$, que es un *hiperplano afín* en $X_{\mathbb{R}}$, pues se obtiene trasladando el núcleo de un funcional lineal. Da lugar a los *semiespacios* $H^- = \{x \in X : \operatorname{Re} f(x) \leqslant \alpha\}$ y $H^+ = \{x \in X : \operatorname{Re} f(x) \geqslant \alpha\}$. Dados dos conjuntos convexos $A, B \subset X$, la condición (3) nos dice que f separa A de B si, y sólo si, existe $\alpha \in \mathbb{R}$ verificando que $A \subset H^-$ y $B \subset H^+$. Esto significa que el conjunto A está a un lado del hiperplano H mientras que B está al otro, y en este sentido podemos decir que el hiperplano H *separa* los conjuntos A y B. Ahora la situación sí es simétrica, pues al sustituir f por $-f$ y α por $-\alpha$, el hiperplano H no varía, pero los semiespacios H^- y H^+ se intercambian.

Si A y B son subconjuntos convexos, no vacíos y disjuntos, de un espacio vectorial X, la intuición geométrica sugiere que debe ser posible separar A y B, pero en general no es así, como vamos a comprobar.

Ejemplo: *Dos subconjuntos convexos, no vacíos y disjuntos, de un espacio vectorial, que no se pueden separar.* Sea X el espacio vectorial real de todas las sucesiones de soporte finito, en el que tenemos una base algebraica formada por los vectores unidad $\{e_n : n \in \mathbb{N}\}$. Consideremos el conjunto $A \subset X$ formado por las sucesiones cuyo último término no nulo es positivo, es decir,

$$A = \left\{ \sum_{k=1}^{n} \alpha_k e_k : n \in \mathbb{N}, \ \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}, \ \alpha_n > 0 \right\}$$

Dados $u, v \in A$ y $t \in]0,1[$, veamos que $w \in A$, donde $w = (1-t)u + tv$. Para ello, escribimos $u = \sum_{k=1}^{n} \alpha_k e_k$ y $v = \sum_{k=1}^{m} \beta_k e_k$, con $n, m \in \mathbb{N}$, $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_m \in \mathbb{R}$, siendo $\alpha_n > 0$ y $\beta_m > 0$. Si $m < n$, se tiene $w(n) = (1-t)u(n) + tv(n) = (1-t)\alpha_n > 0$, mientras que $w(k) = 0$ para todo $k \in \mathbb{N}$ con $k > n$, luego $w \in A$. Análogo razonamiento prueba que $w \in A$ cuando $n < m$. Finalmente, si $n = m$, se tiene también $w(n) = (1-t)\alpha_n + t\beta_n > 0$ y $w(k) = 0$ para todo $k \in \mathbb{N}$ con $k > n$. En cualquier caso $w \in A$, luego A es convexo.

Observamos ahora que, si $f \neq 0$ es un funcional lineal en X, se tiene $f(A) = \mathbb{R}$. En efecto, existe $n \in \mathbb{N}$ tal que $f(e_n) = \lambda \neq 0$ y supondremos de momento que $\lambda > 0$. Para todo $\rho \in \mathbb{R}^+$ tenemos $\rho e_n \in A$, luego $\rho\lambda \in f(A)$, y esto prueba que el conjunto $f(A)$ no está mayorado. También para todo $\rho \in \mathbb{R}^+$, tenemos $e_{n+1} - \rho e_n \in A$, luego $f(e_{n+1}) - \rho\lambda \in f(A)$, y esto prueba que $f(A)$ no está minorado. Como A es convexo, vemos que $f(A)$ es un intervalo, que no está mayorado ni minorado, luego $f(A) = \mathbb{R}$. En el caso $\lambda < 0$, el razonamiento anterior se aplica al funcional $-f$, obteniendo que $-f(A) = \mathbb{R}$, luego también $f(A) = \mathbb{R}$.

Queda claro que A no se puede separar de ningún otro conjunto convexo, pues para ello algún funcional lineal no nulo tendría que estar mayorado o minorado en A. Por ejemplo, si tomamos $B = \{0\}$, vemos que A y B son subconjuntos convexos de X, no vacíos y disjuntos, pero no existe ningún funcional lineal no nulo en X que separe uno del otro. ■

Resaltemos que el espacio vectorial X del ejemplo anterior tiene dimensión infinita. Más adelante veremos que, en un espacio vectorial de dimensión finita, no es posible encontrar un ejemplo como el anterior. Para probarlo usaremos un teorema general de separación, mucho más importante. De hecho, la versión analítica del teorema de Hahn-Banach dará una condición suficiente para separar dos subconjuntos convexos, no vacíos y disjuntos, de un espacio vectorial cualquiera, con una hipótesis adicional muy poco restrictiva sobre uno de ellos.

8.2. Separación en espacios vectoriales

Para entender mejor la hipótesis que nos llevará al principal resultado sobre separación de conjuntos convexos, introducimos el siguiente concepto. Se dice que un subconjunto U de un espacio vectorial X es **absorbente** cuando $X = \mathbb{R}^+ U$, es decir, cuando para todo $x \in X$, puede encontrarse $\rho \in \mathbb{R}^+$ tal que $x \in \rho U$. Es claro que entonces $0 \in U$, pero además U debe contener otro punto en cada semirrecta que parta del origen, luego podemos decir que 0 está enteramente "rodeado" por puntos de U. La bola unidad de un espacio normado, y de hecho cualquier bola centrada en el origen con radio estrictamente positivo, es un claro ejemplo de conjunto absorbente.

Si U es un subconjunto convexo y absorbente de un espacio vectorial X, para cada $x \in X$ tenemos un $\rho \in \mathbb{R}^+$ tal que $x/\rho \in U$, con lo que el segmento de extremos 0 y x/ρ estará contenido en U, luego U contiene, no sólo un punto, sino todo un segmento no trivial en cada una de las direcciones del espacio X, si bien la longitud de dicho segmento puede depender de la dirección. Esto nos lleva a pensar que 0 es una especie de "punto interior" de U, en un sentido puramente algebraico. La misma idea se aplica, salvo una traslación, a cualquier punto del espacio: si A es un conjunto convexo y $a_0 \in A$, el hecho de que $A - a_0$ sea absorbente significa que a_0 es un punto interior de A, en el mismo sentido algebraico.

Podemos ya enunciar el principal resultado sobre separación de conjuntos convexos, que puede entenderse como la versión geométrica del teorema de Hahn-Banach, pues veremos que es equivalente a la versión analítica de dicho teorema.

Teorema general de separación. *Sean A y B subconjuntos no vacíos, convexos y disjuntos, de un espacio vectorial X. Supongamos que existe un punto $a_0 \in A$ tal que $A - a_0$ es absorbente. Entonces existe un funcional lineal no nulo f en X que separa A de B, es decir, que verifica:*

$$\operatorname{Re} f(a) \leqslant \operatorname{Re} f(b) \qquad \forall\, a \in A\,, \ \forall\, b \in B$$

Demostración. Basta considerar el caso real, pues en el caso complejo se usa el espacio real subyacente a X, como ya hemos comentado.

Fijado $b_0 \in B$, consideramos el punto $x_0 \in X$ y el conjunto $U \subset X$ dados por

$$x_0 = b_0 - a_0 \qquad \text{y} \qquad U = A - B + x_0 = \big(A - a_0\big) - \big(B - b_0\big)$$

y como se verá al final, para separar A de B bastará separar U del punto x_0.

Por ser A y B convexos, vemos fácilmente que U también lo es. Para $u_1, u_2 \in U$, tenemos

$$u_1 = a_1 - b_1 + x_0 \qquad \text{y} \qquad u_2 = a_2 - b_2 + x_0$$

donde $a_1, a_2 \in A$ y $b_1, b_2 \in B$. Para todo $t \in [0,1]$ tenemos entonces

$$a = (1-t)\,a_1 + t\,a_2 \in A \qquad \text{y} \qquad b = (1-t)\,b_1 + t\,b_2 \in B$$

de donde deducimos claramente que $(1-t)\,u_1 + t\,u_2 = a - b + x_0 \in U$.

Como $A - a_0 \subset U$ y, por hipótesis, $A - a_0$ es absorbente, deducimos que U también lo es. Por último, de $A \cap B = \emptyset$ se deduce que $x_0 \notin U$, pues en otro caso, existirían $a \in A$ y $b \in B$ tales que $x_0 = a - b + x_0$, de donde $a = b \in A \cap B$, una contradicción.

Para entender mejor el resto de la demostración, imaginemos que U fuese la bola unidad de un espacio normado X. Tomando entonces $f \in X^*$ con $\|f\| = 1$ y $f(x_0) = \|x_0\| > 1$, es claro que f separa U de x_0. Además, la norma de X se calcula fácilmente a partir de U, puesto que, dados $x \in X$ y $\rho \in \mathbb{R}^+$, se tiene $x \in \rho U$ cuando $\rho > \|x\|$ mientras que $x \notin \rho U$ para $\rho < \|x\|$, luego $\|x\| = \inf \big\{ \rho \in \mathbb{R}^+ : x \in \rho U \big\}$. En nuestro caso mucho más general, el segundo miembro de esta igualdad sigue teniendo sentido para todo $x \in X$, y permite definir una aplicación de X en $\mathbb{R}_0^+$, que ya no tiene por qué ser una norma en X, pero sí será un funcional sublineal, que nos permita aplicar la versión analítica del teorema de Hahn-Banach.

Así pues, usando que U es absorbente, definimos $\varphi : X \to \mathbb{R}_0^+$, escribiendo

$$\varphi(x) = \inf\{\rho \in \mathbb{R}^+ \,:\, x \in \rho U\} \qquad \forall x \in X$$

y vamos a comprobar que φ es un funcional sublineal en X.

Para $x \in X$ y $r \in \mathbb{R}^+$, se tiene $\{\lambda \in \mathbb{R}^+ : rx \in \lambda U\} = \{r\rho : \rho \in \mathbb{R}^+,\ x \in \rho U\}$, de donde deducimos claramente que

$$\varphi(rx) = \inf\{\lambda \in \mathbb{R}^+ : rx \in \lambda U\} = r \inf\{\rho \in \mathbb{R}^+ : x \in \rho U\} = r\,\varphi(x)$$

y esto prueba que φ es homogéneo por homotecias.

Para la desigualdad triangular, fijemos $x, y \in X$ y $\rho, \lambda \in \mathbb{R}^+$ tales que $x \in \rho U$ e $y \in \lambda U$. Entonces, por ser U convexo, tenemos

$$x + y \in \rho U + \lambda U = (\rho + \lambda)\left(\frac{\rho}{\rho+\lambda} U + \frac{\lambda}{\rho+\lambda} U\right) \subset (\rho + \lambda) U$$

de donde deducimos que $\varphi(x+y) \leqslant \rho + \lambda$. Como esta desigualdad es válida para todo $\lambda \in \mathbb{R}^+$ tal que $y \in \lambda U$, obtenemos que $\varphi(x+y) \leqslant \rho + \varphi(y)$, pero a su vez esta desigualdad es válida para todo $\rho \in \mathbb{R}^+$ tal que $x \in \rho U$, luego $\varphi(x+y) \leqslant \varphi(x) + \varphi(y)$. Por tanto, φ es un funcional sublineal en X, como queríamos comprobar.

Para $u \in U$ se tiene obviamente que $\varphi(u) \leqslant 1$. En sentido opuesto, vamos a comprobar que, para $x \in X$ con $\varphi(x) < 1$, se tiene $x \in U$. En efecto, existe $\rho \in]0, 1[$ tal que $x \in \rho U$, pero por ser U convexo con $0 \in U$, se tiene $\rho U \subset \rho U + (1-\rho) U \subset U$, luego $x \in U$ como se quería. Puesto que $x_0 \notin U$, deducimos que se ha de tener $\varphi(x_0) \geqslant 1$.

Razonamos ahora con φ como hicimos en otro momento con una norma. Consideramos el subespacio $M = \mathbb{R}x_0 \subset X$, en el que definimos $g(\lambda x_0) = \lambda \varphi(x_0)$ para todo $\lambda \in \mathbb{R}$, obteniendo un funcional lineal $g : M \to \mathbb{R}$. Para $\lambda \in \mathbb{R}^+$ tenemos $g(\lambda x_0) = \varphi(\lambda x_0)$, puesto que φ es homogéneo por homotecias. Pero si $\lambda \in \mathbb{R}_0^-$, tenemos claramente $g(\lambda x_0) \leqslant 0 \leqslant \varphi(\lambda x_0)$, ya que φ no toma valores negativos. Por tanto, g está dominado por φ, y la versión analítica del teorema de Hahn-Banach nos da un funcional lineal $f : X \to \mathbb{R}$ que verifica

$$f(x) \leqslant \varphi(x) \quad \forall x \in X \qquad \text{y} \qquad f(\lambda x_0) = g(\lambda x_0) = \lambda \varphi(x_0) \quad \forall \lambda \in \mathbb{R}$$

En particular, podemos tomar $x = u \in U$ y $\lambda = 1$ para obtener

$$f(u) \leqslant \varphi(u) \leqslant 1 \leqslant \varphi(x_0) = f(x_0)$$

Por tanto $f \neq 0$ y, dados $a \in A$ y $b \in B$, tomamos $u = a - b + x_0 \in U$ para obtener que

$$f(a) - f(b) + f(x_0) = f(u) \leqslant f(x_0)$$

es decir, $f(a) \leqslant f(b)$. Esto prueba que f separa A de B, concluyendo la demostración. ■

8.3. Equivalencia de la versión geométrica con la analítica

Ha quedado claro que el teorema de separación recién demostrado es consecuencia directa de la versión analítica del teorema de Hahn-Banach. Recíprocamente, del teorema de separación vamos a obtener un resultado formalmente más general que dicha versión analítica, con lo que al final tendremos tres enunciados equivalentes.

Teorema. *Sea X un espacio vectorial, $\varphi : X \to \mathbb{R}$ una función convexa, M un subespacio de X y g un funcional lineal en M que está dominado por φ, es decir,*

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

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

$$f(y) = g(y) \quad \forall y \in M \qquad \text{y} \qquad \operatorname{Re} f(x) \leqslant \varphi(x) \quad \forall x \in X$$

Demostración. Basta considerar el caso real, pues en el caso complejo, φ también es una función convexa en $X_{\mathbb{R}}$, con lo que basta usar los espacios reales $X_{\mathbb{R}}$ y $M_{\mathbb{R}}$, exactamente igual que hicimos en su momento, para probar la versión analítica del teorema de Hahn-Banach.

En el espacio vectorial producto $X \times \mathbb{R}$ consideramos entonces los conjuntos

$$A = \big\{ (x,t) \in X \times \mathbb{R} \,:\, \varphi(x) < t \big\} \qquad \text{y} \qquad B = \big\{ \big(y, g(y)\big) \,:\, y \in M \big\}$$

a los que pretendemos aplicar el teorema general de separación. Para $x \in X$ arbitrario, basta tomar $t \in \mathbb{R}$ con $t > \varphi(x)$, para tener $(x,t) \in A$, luego $A \neq \emptyset$, y es aún más obvio que $B \neq \emptyset$.

Para cualesquiera $(x_1,t_1),(x_2,t_2) \in A$ y $r \in [0,1]$, la convexidad de φ nos dice que

$$\varphi\big((1-r)x_1 + r x_2\big) \leqslant (1-r)\varphi\big(x_1\big) + r\varphi\big(x_2\big) < (1-r)t_1 + r t_2$$

luego $(1-r)(x_1,t_1) + r(x_2,t_2) \in A$ y esto prueba que A es convexo. Por otra parte, como g es lineal, vemos que B es un subespacio de $X \times \mathbb{R}$ y, en particular, un conjunto convexo.

Para todo $y \in M$ tenemos por hipótesis que $g(y) \leqslant \varphi(y)$, luego $\big(y, g(y)\big) \notin A$ y esto prueba que $A \cap B = \emptyset$. Por tanto, A y B son subconjuntos no vacíos, convexos y disjuntos, de $X \times \mathbb{R}$.

Para un punto arbitrario $a_0 \in A$, probaremos ahora que $A - a_0$ es un conjunto absorbente. Dado $w \in X \times \mathbb{R}$, encontraremos $\rho \in \mathbb{R}^+$ tal que $a_0 + \rho w \in A$, con lo que $w \in (1/\rho)(A - a_0)$, y por tanto $w \in \mathbb{R}^+(A - a_0)$. Si $a_0 = (x_0,t_0)$ y $w = (x,t)$ con $x_0, x \in X$ y $t_0, t \in \mathbb{R}$, se trata de encontrar $\rho \in \mathbb{R}^+$ que verifique

$$\big(x_0 + \rho x,\, t_0 + \rho t\big) \in A, \qquad \text{es decir,} \qquad \varphi\big(x_0 + \rho x\big) < t_0 + \rho t \tag{4}$$

Ahora bien, para todo $\rho \in]0,1[$, usando de nuevo la convexidad de φ, tenemos

$$\begin{aligned}
\varphi\big(x_0 + \rho x\big) - t_0 - \rho t &= \varphi\big((1-\rho)x_0 + \rho\,(x_0 + x)\big) - t_0 - \rho t \\
&\leqslant (1-\rho)\,\varphi(x_0) + \rho\,\varphi(x_0 + x) - t_0 - \rho t \\
&= \varphi(x_0) - t_0 + \rho\,\big(\varphi(x_0 + x) - \varphi(x_0) - t\big)
\end{aligned}$$

Como $a_0 \in A$, tenemos $\varphi(x_0) - t_0 < 0$, y la desigualdad anterior muestra claramente que podemos elegir $\rho \in]0,1[$ de forma que se cumpla (4).

Comprobadas todas sus hipótesis, el teorema general de separación nos da un funcional lineal no nulo $F : X \times \mathbb{R} \to \mathbb{R}$, que separa A de B. Definimos entonces

$$h(x) = F(x,0) \quad \forall x \in X \qquad \text{y} \qquad \lambda = F(0,-1)$$

Obteniendo un funcional lineal $h : X \to \mathbb{R}$, y un $\lambda \in \mathbb{R}$, tales que

$$F(x,t) = h(x) - \lambda t \qquad \forall (x,t) \in X \times \mathbb{R}$$

Como $F \neq 0$, se ha de tener $h \neq 0$ o $\lambda \neq 0$ y, como F separa A de B, tendremos

$$h(x) - \lambda t \leqslant h(y) - \lambda g(y) \qquad \forall (x,t) \in A, \ \forall y \in M \tag{5}$$

Tomando $x = y = 0$, y $t \in \mathbb{R}^+$ con $t > \varphi(0)$, obtenemos que $-\lambda t \leqslant 0$, de donde $\lambda \geqslant 0$, pero veamos que de hecho se tiene $\lambda > 0$. Si fuese $\lambda = 0$, usando que $h \neq 0$ tenemos $x \in X$ tal que $h(x) > 0$, y tomamos $t \in \mathbb{R}$ con $t > \varphi(x)$ para tener $(x,t) \in A$. De (5) con $y = 0$ deducimos entonces que $h(x) \leqslant 0$, una contradicción.

Por otra parte, si h_0 es la restricción de h a M, usamos (5) con $(x,t) \in A$ fijado, obteniendo que el funcional lineal $h_0 - \lambda g$ está minorado en M, lo que sólo es posible si $h_0 - \lambda g = 0$. Tomando $f = h/\lambda$, tenemos un funcional lineal en X cuya restricción a M es $h_0/\lambda = g$, es decir, f extiende a g.

Sólo queda ver que f está dominado por φ. Para ello, fijado $x \in X$, para todo $\varepsilon > 0$ tenemos que $\big(x, \varphi(x) + \varepsilon\big) \in A$. Usando (5) con $y = 0$, por ser $\lambda > 0$, obtenemos que

$$h(x) \leqslant \lambda(\varphi(x) + \varepsilon) \qquad \text{de donde} \qquad f(x) \leqslant \varphi(x) + \varepsilon$$

Como $\varepsilon > 0$ era arbitrario, tenemos $f(x) \leqslant \varphi(x)$, como se quería. ■

La versión analítica del teorema de Hahn-Banach es caso particular del resultado anterior, porque todo funcional sublineal φ en un espacio vectorial X es una función convexa. En efecto, basta observar que, para cualesquiera $x, y \in X$ y $t \in]0,1[$ se tiene claramente

$$\varphi\big((1-t)x + ty\big) \leqslant \varphi\big((1-t)x\big) + \varphi\big(ty\big) = (1-t)\varphi\big(x\big) + t\varphi\big(y\big)$$

8.4. Separación en espacios normados

El teorema general de separación de conjuntos convexos se usa con especial comodidad cuando trabajamos en un espacio normado X. Entonces, todo entorno de cero en X contiene una bola centrada en cero con radio estrictamente positivo, luego es un conjunto absorbente. Denotando por $\operatorname{int}(A)$ al interior de un conjunto $A \subset X$, para $a_0 \in \operatorname{int}(A)$ se tiene que $A - a_0$ es entorno de cero, luego es un conjunto absorbente. Por tanto, la hipótesis que se requiere para poder aplicar el teorema de separación, se consigue fácilmente suponiendo que uno de los conjuntos convexos que queremos separar tiene interior no vacío. Por otra parte, vamos a comprobar dos propiedades básicas de los conjuntos convexos en espacios normados, que permitirán mejorar el teorema de separación en varios aspectos.

- *Sea A un subconjunto convexo de un espacio normado X, y supongamos que A tiene interior no vacío. Entonces* $\operatorname{int}(A)$ *es convexo y verifica que* $A \subset \overline{\operatorname{int}(A)}$.

Si $x_1, x_2 \in \operatorname{int}(A)$, existen $r_1, r_2 \in \mathbb{R}^+$ con $(x_1 + r_1 U) \cup (x_2 + r_2 U) \subset A$ donde U es la bola abierta unidad de X. Si $x = (1-t)x_1 + t x_2$ con $t \in [0,1]$, tomando $r = (1-t)r_1 + t r_2 \in \mathbb{R}^+$, tenemos claramente que $rU \subset (1-t)r_1 U + t r_2 U$, y deducimos que

$$x + rU \subset (1-t)(x_1 + r_1 U) + t(x_2 + r_2 U) \subset (1-t)A + tA \subset A$$

donde hemos usado que A es convexo. Esto prueba que $x \in \operatorname{int}(A)$, luego $\operatorname{int}(A)$ es convexo.

Para $a \in A$ debemos probar ahora que $a \in \overline{\operatorname{int}(A)}$. Para ello fijamos $x \in \operatorname{int}(A)$ y $r \in \mathbb{R}^+$ tal que $x + rU \subset A$. Para todo $t \in]0,1[$, se tiene entonces

$$(1-t)x + ta + (1-t)rU = (1-t)(x+rU) + ta \subset (1-t)A + tA \subset A$$

Como $(1-t)r \in \mathbb{R}^+$, vemos que $(1-t)x + ta \in \operatorname{int}(A)$ para todo $t \in]0,1[$. Basta ahora observar que $a = \lim\limits_{t \nearrow 1} \big((1-t)x + ta\big)$, para concluir que $a \in \overline{\operatorname{int}(A)}$, como se quería. ■

Pasamos ya a probar la versión del teorema de separación más adecuada para trabajar en espacios normados.

Teorema de separación en espacios normados. *Sean A y B subconjuntos convexos de un espacio normado X, verificando que* $\operatorname{int}(A) \neq \emptyset$ *y* $B \neq \emptyset$, *pero* $\operatorname{int}(A) \cap B = \emptyset$. *Entonces existen $f \in X^*$ y $\alpha \in \mathbb{R}$ tales que:*

$$\operatorname{Re} f(a) \leqslant \alpha \leqslant \operatorname{Re} f(b) \qquad \forall a \in A, \ \forall b \in B \tag{6}$$

De hecho se tiene que $\operatorname{Re} f(x) < \alpha$ *para todo* $x \in \operatorname{int}(A)$.

Demostración. Como viene ocurriendo, basta considerar el caso real, pero en este caso conviene aclarar una cuestión. Si X es un espacio normado complejo y f un funcional lineal en X, entonces f es continuo si, y sólo si, lo es su parte real, ya que $\operatorname{Im} f(x) = \operatorname{Re} f(-ix)$ para todo $x \in X$. Por tanto, los elementos de $(X_{\mathbb{R}})^*$ coinciden con las partes reales de los elementos de X^* y ello hace que, una vez demostrado el teorema para $X_{\mathbb{R}}$, se obtenga claramente para X.

Por el resultado anterior sabemos que $\operatorname{int}(A)$ es convexo, luego $\operatorname{int}(A)$ y B son subconjuntos convexos, no vacíos y disjuntos, del espacio vectorial X. Además, para todo $x \in \operatorname{int}(A)$ se tiene que $A - x$ es entorno de cero en X, luego es absorbente. Esto nos permite usar el teorema general de separación, para obtener un funcional lineal no nulo en X, y un $\alpha \in \mathbb{R}$, tales que

$$f(x) \leqslant \alpha \leqslant f(b) \qquad \forall x \in \operatorname{int}(A), \ \forall b \in B \tag{7}$$

Como $f \neq 0$, podemos fijar $u \in X$, con $\|u\| = 1$ y $f(u) > 0$. Para $x \in \operatorname{int}(A)$ podemos entonces tomar $\varepsilon > 0$ de forma que $x + \varepsilon u \in \operatorname{int}(A)$, con lo que de (7) deducimos que

$$f(x) < f(x) + \varepsilon f(u) = f\big(x + \varepsilon u\big) \leqslant \alpha$$

y hemos probado así la última afirmación del enunciado.

Observamos ahora que el hiperplano $H = \{x \in X : f(x) = \alpha\}$ verifica que $H \cap \operatorname{int}(A) = \emptyset$. Como $\operatorname{int}(A)$ es un abierto no vacío de X, vemos que H no es denso en X y, mediante una traslación, deducimos que $\ker f$ tampoco lo es. Por tanto $\ker f$ es cerrado, es decir, $f \in X^*$.

Finalmente, como $f \in X^*$, el semiespacio $S = \{x \in X : f(x) \leqslant \alpha\}$ es cerrado. En (7) tenemos que $\operatorname{int}(A) \subset S$, con lo que podemos usar de nuevo el resultado anterior, para concluir que $A \subset \overline{\operatorname{int}(A)} \subset S$. Esto significa que $f(a) \leqslant \alpha$ para todo $a \in A$, y hemos obtenido (6), concluyendo la demostración. ■

Comparando este resultado con el teorema general de separación, observamos que hemos fortalecido la hipótesis sobre A, exigiendo que $\operatorname{int}(A) \neq \emptyset$. A cambio hemos conseguido algo importante: separar A y B mediante un funcional lineal *continuo*. Además, hemos debilitado la hipótesis de que A y B sean disjuntos, exigiendo solamente $\operatorname{int}(A) \cap B = \emptyset$, e incluso podemos decir que hemos separado "estrictamente" $\operatorname{int}(A)$ y B, ya que los conjuntos $f\big(\operatorname{int}(A)\big)$ y $f(B)$ son disjuntos.

8.5. Funcionales y puntos de soporte

Vamos a destacar un caso particular del último teorema, cuya interpretación geométrica es especialmente interesante. Sea X un espacio normado y A un subconjunto convexo y cerrado de X con interior no vacío. Dado un punto x_0 en la frontera de A, podemos aplicar el teorema anterior tomando $B = \{x_0\}$, y obtenemos $f \in X^* \setminus \{0\}$ que verifica:

$$\operatorname{Re} f(a) \leqslant \operatorname{Re} f(x_0) \quad \forall a \in A. \tag{8}$$

Vemos que la función $\operatorname{Re} f$ tiene máximo en el conjunto A, que se alcanza en el punto x_0.

La interpretación geométrica de (8), que como siempre debe hacerse en el espacio real subyacente a X, es muy clara: el hiperplano afín $\big\{x \in X : \operatorname{Re} f(x) = \operatorname{Re} f(x_0)\big\}$ pasa por el punto x_0 y deja el conjunto A a un lado. Intuitivamente, es muy natural pensar que dicho hiperplano "soporta" al conjunto A en el punto x_0. Por ello, dado un subconjunto no vacío A de un espacio vectorial X, si un funcional lineal no nulo $f : X \to \mathbb{K}$, y un punto $x_0 \in A$, verifican la desigualdad (8), decimos que f es un **funcional de soporte** del conjunto A en el punto x_0, y también que x_0 es un **punto de soporte** de A. Con esta nomenclatura, el resultado obtenido es el siguiente:

- **Abundancia de puntos de soporte.** *Si X es un espacio normado y A un subconjunto convexo y cerrado de X, con interior no vacío, todo punto de la frontera de A es un punto de soporte de A.*

Un caso particular del corolario anterior era ya conocido. Si A es la bola unidad de X y tomamos $x_0 \in X$ con $\|x_0\| = 1$, sabíamos que existe $f \in X^*$ con $\|f\| = 1$ y $f(x_0) = \|x_0\| = 1$. Está claro entonces que f es un funcional de soporte de la bola unidad en el punto x_0. Así pues, sabíamos que todos los puntos de la esfera unidad de un espacio normado son puntos de soporte de la bola unidad. Pero ahora tenemos el mismo resultado, para cualquier conjunto convexo y cerrado con interior no vacío.

8.6. Separación fuerte

La abundancia de puntos de soporte se ha conseguido usando el teorema de separación para dos conjuntos convexos que no llegaban a ser disjuntos. Analizamos ahora la situación opuesta, en la que dichos conjuntos, no sólo son disjuntos, sino que están a distancia positiva. Entonces podremos "fortalecer" la separación, en el sentido que pasamos a explicar.

Si A y B son subconjuntos no vacíos y convexos de un espacio vectorial X, decimos que un funcional lineal $f : X \to \mathbb{K}$ **separa fuertemente** A de B, cuando existen $\alpha, \beta \in \mathbb{R}$ tales que

$$\operatorname{Re} f(a) \leqslant \alpha < \beta \leqslant \operatorname{Re} f(b) \qquad \forall a \in A, \ \forall b \in B \tag{9}$$

Nótese que $f \neq 0$. Veamos, en el espacio real subyacente a X, el significado geométrico de estas desigualdades. Escribiendo $S_\alpha = \{x \in X : \operatorname{Re} f(x) \leqslant \alpha\}$ y $T_\beta = \{x \in X : \operatorname{Re} f(x) \geqslant \beta\}$, vemos que S_α y T_β son semiespacios disjuntos, tales que $A \subset S_\alpha$ y $B \subset T_\beta$. No sólo los hiperplanos que dan lugar a dichos semiespacios separan A y B, sino que A y B están a distintos lados de la *banda* $D = \{x \in X : \alpha < \operatorname{Re} f(x) < \beta\}$.

Del teorema de separación en espacios normados, deducimos el siguiente resultado:

- **Separación fuerte en espacios normados.** *Supongamos que A y B son subconjuntos, no vacíos y convexos, de un espacio normado X, que están a distancia positiva, esto es:*

  $$d(A,B) = \operatorname{ínf}\{\|b-a\| : a \in A, \, b \in B\} = \rho > 0$$

  Entonces, existe un funcional lineal y continuo en X que separa fuertemente A de B. Más concretamente, existen $f \in X^*$, con $\|f\| = 1$, y $\alpha \in \mathbb{R}$, tales que

  $$\operatorname{Re} f(a) \leqslant \alpha < \alpha + \rho \leqslant \operatorname{Re} f(b) \qquad \forall a \in A, \ \forall b \in B \tag{10}$$

Si U es la bola abierta unidad de X, es claro que el conjunto $A + \rho U$ es convexo y abierto. Si algún $b \in B$ verificase que $b = a + \rho u$ con $a \in A$ y $u \in U$, se tendría $\|b - a\| = \rho \|u\| < \rho$, lo cual es imposible, luego $(A + \rho U) \cap B = \emptyset$. Por tanto, $A + \rho U$ y B son dos subconjuntos no vacíos, convexos y disjuntos de X, el primero de los cuales es abierto. Por el teorema de separación, existe $f \in X^*$ verificando que

$$\operatorname{Re} f(a) + \rho \operatorname{Re} f(u) < \operatorname{Re} f(b) \qquad \forall a \in A, \ \forall b \in B, \ \forall u \in U$$

Estas desigualdades se mantienen al dividir f por su norma, con lo que conseguimos $\|f\| = 1$, o lo que es lo mismo, $\sup\{\operatorname{Re} f(u) : u \in U\} = 1$. Para $a \in A$ y $b \in B$, deducimos que

$$\operatorname{Re} f(a) + \rho \leqslant \operatorname{Re} f(b)$$

Si ahora definimos $\alpha = \sup\{\operatorname{Re} f(a) : a \in A\}$, obtenemos claramente (10). En particular tenemos (9) con $\beta = \alpha + \rho$, luego f separa fuertemente A de B. ■

El resultado anterior tiene un aspecto cuantitativo, que conviene resaltar. Recordando la interpretación geométrica de la separación fuerte, hemos visto que los conjuntos A y B están a distintos lados de la banda $D = \{x \in X : \alpha < \operatorname{Re} f(x) < \beta\}$, donde $\beta - \alpha = \rho = d(A,B)$.

La cuestión es que, siendo $\|f\| = 1$, la diferencia $\beta - \alpha = d(A,B)$ es la máxima posible. De manera más intuitiva, la banda D es lo más “ancha” posible. Para comprobarlo, supongamos que A y B verifican (9), con $f \in X^*$ y $\|f\| = 1$. Entonces, para $a \in A$ y $b \in B$ se tiene

$$\|b-a\| = \|f\| \, \|b-a\| \geqslant |f(b-a)| \geqslant \operatorname{Re} f(b) - \operatorname{Re} f(a) \geqslant \beta - \alpha$$

de donde deducimos que $\beta - \alpha \leqslant d(A,B)$, como se quería.

Hay un caso particular del corolario anterior que ya conocíamos: si M es un subespacio cerrado de un espacio normado X y $x \in X \setminus M$, sabíamos que existe $f \in M^\perp$ con $\|f\| = 1$, tal que $f(x) = d(x,M)$. Esto se deduce del resultado anterior, pues tomando $A = M$ y $B = \{x\}$, con lo que $d(A,B) = d(x,M) = \rho > 0$, obtenemos $f \in X^*$, con $\|f\| = 1$, y $\alpha \in \mathbb{R}$, tales que

$$\operatorname{Re} f(y) \leqslant \alpha < \alpha + \rho \leqslant \operatorname{Re} f(x) \qquad \forall y \in M$$

Vemos que $f(M) \neq \mathbb{K}$, luego $f(M) = \{0\}$, es decir, $f \in M^\perp$. Entonces $\alpha \geqslant 0$, y $\operatorname{Re} f(x) \geqslant \rho$, pero al ser $\|f\| = 1$, tenemos fácilmente $|f(x)| \leqslant \rho$, luego $f(x) = \rho$, como se quería.

Sin entrar en cuestiones cuantitativas, la siguiente es una forma natural de asegurarse que dos conjuntos están a distancia positiva, para tener separación fuerte:

- *Sean A y B subconjuntos convexos, no vacíos y disjuntos, de un espacio normado X, y supongamos que A es compacto y B es cerrado. Entonces existe un funcional $f \in X^*$ que separa fuertemente A y B.*

La función continua $x \mapsto d(x,B)$ tiene mínimo en el compacto A, es decir, existe $a_0 \in A$ tal que $d(a_0,B) = \min\{d(a,B) : a \in A\} = d(A,B)$. Si fuese $d(a_0,B) = 0$, por ser B cerrado se tendría $a_0 \in A \cap B$ contra la hipótesis, luego $d(A,B) > 0$ y se aplica el resultado anterior. ■

8.7. Separación en espacios de dimensión finita

Concluimos viendo que, en un espacio vectorial de dimensión finita, no se requiere ninguna hipótesis restrictiva para separar dos conjuntos convexos, no vacíos y disjuntos. Ello se debe a que el caso que no queda cubierto por el teorema general de separación, puede resolverse mediante el siguiente resultado elemental.

- *Dado $N \in \mathbb{N}$, sea U un subconjunto convexo de $\mathbb{R}^N$ tal que $0 \in U$ y $\operatorname{Lin} U = \mathbb{R}^N$. Entonces se tiene $\operatorname{int}(U) \neq \emptyset$ para la topología usual de $\mathbb{R}^N$.*

Como U contiene una base de $\mathbb{R}^N$, salvo una biyección lineal de $\mathbb{R}^N$ sobre sí mismo, que es un homeomorfismo, podemos suponer que U contiene a la base usual $\{e_1, e_2, \ldots, e_N\}$. Para cada $k \in \{1,2,\ldots,N\}$, consideramos entonces el conjunto

$$\Delta_k = \left\{ \sum_{j=1}^{k} \rho_j e_j \ : \ \rho_1, \rho_2, \ldots, \rho_k \in \mathbb{R}^+, \ \sum_{j=1}^{k} \rho_j < 1 \right\}$$

Observamos que los puntos de Δ_1 pertenecen al segmento de extremos 0 y e_1, que está contenido en U por ser U convexo, luego tenemos $\Delta_1 \subset U$. Razonando por inducción finita, suponemos que, para un $k \in \{1,2,\ldots,N-1\}$, se tiene $\Delta_k \subset U$, y probamos que $\Delta_{k+1} \subset U$.

Si $\rho_1,\rho_2,\ldots,\rho_{k+1} \in \mathbb{R}^+$ y $\displaystyle\sum_{j=1}^{k+1} \rho_j < 1$, tenemos $\dfrac{\rho_j}{1-\rho_{k+1}} \in \mathbb{R}^+$ para todo $j \in \{1,2,\ldots,k\}$, así como $\displaystyle\sum_{j=1}^{k} \frac{\rho_j}{1-\rho_{k+1}} < 1$, luego tomando $x = \displaystyle\sum_{j=1}^{k} \frac{\rho_j}{1-\rho_{k+1}} e_j$, vemos que $x \in \Delta_k$, y por la hipótesis de inducción tenemos $x \in U$. Por tanto, $\displaystyle\sum_{j=1}^{k+1} \rho_j e_j = (1-\rho_{k+1})x + \rho_{k+1} e_{k+1} \in U$. Esto prueba que $\Delta_{k+1} \subset U$ como se quería. Por inducción concluimos entonces que $\Delta_N \subset U$.

Comprobemos que Δ_N es un abierto no vacío de $\mathbb{R}^N$, con lo que tendremos $\operatorname{int}(U) \neq \emptyset$. Basta para ello observar que

$$\Delta_N = \left(\bigcap_{k=1}^{N} \{ x \in \mathbb{R}^N : x(k) \in \mathbb{R}^+ \} \right) \bigcap \left\{ x \in \mathbb{R}^N : \sum_{k=1}^{N} x(k) < 1 \right\}$$

lo que muestra a Δ_N como intersección de $N+1$ conjuntos abiertos, luego Δ_N es abierto. Tomando $x(k) = 1/(N+1)$ para todo $k \in \{1,2\ldots,N\}$, se tiene $x \in \Delta_N$, luego $\Delta_N \neq \emptyset$. ■

Podemos ya probar el teorema general de separación para espacios de dimensión finita:

- **Separación en dimensión finita.** *Sean A y B dos subconjuntos no vacíos, convexos y disjuntos de $\mathbb{R}^N$, con $N \in \mathbb{N}$. Entonces existe un funcional lineal no nulo en $\mathbb{R}^N$ que separa A y B. Equivalentemente, existe $c \in \mathbb{R}^N \setminus \{0\}$ tal que*

$$\sum_{k=1}^{N} c(k)\,a(k) \leqslant \sum_{k=1}^{N} c(k)\,b(k) \qquad \forall a \in A,\ \forall b \in B$$

Igual que en el teorema general de separación, veremos que el problema se reduce a separar un conjunto convexo $U \subset \mathbb{R}^N$ tal que $0 \in U$, de un punto $x_0 \in \mathbb{R}^N \setminus U$. En este caso particular, hay dos posibles situaciones. Si $\operatorname{int}(U) \neq \emptyset$, el teorema de separación en espacios normados, aplicado a $\mathbb{R}^N$ con cualquier norma, nos da un funcional lineal no nulo en $\mathbb{R}^N$ que separa U del punto x_0. Si por el contrario $\operatorname{int}(U) = \emptyset$, el resultado anterior nos dice que $\operatorname{Lin} U \neq \mathbb{R}^N$, luego existe un funcional lineal no nulo f en $\mathbb{R}^N$, tal que $\operatorname{Lin} U \subset \ker f$, y en particular $f(u) = 0$ para todo $u \in U$. Sustituyendo si es preciso f por $-f$, conseguimos que se tenga $f(x_0) \geqslant 0$, con lo que f separa U del punto x_0.

En general, fijados $a_0 \in A$ y $b_0 \in B$, tomamos $x_0 = b_0 - a_0$ y $U = A - B + x_0$, con lo que el conjunto U es convexo, con $0 \in U$ pero $x_0 \notin U$. Por lo ya demostrado, existe un funcional lineal no nulo f en $\mathbb{R}^N$, verificando que $f(u) \leqslant f(x_0)$ para todo $u \in U$. Dados $a \in A$ y $b \in B$, basta tomar $u = a - b + x_0 \in U$, para obtener que

$$f(a) - f(b) + f(x_0) = f(u) \leqslant f(x_0)$$

de donde $f(a) \leqslant f(b)$. Esto prueba que f separa A de B. ■