

Teorema de Hanh-Banach y de Mazur-Orlicz

El objetivo principal de esta sección es demostrar tanto una versión no tan conocida del clásico teorema de Hanh-Banach como el teorema de Mazur-Orlicz. Para ello, iremos utilizando una serie de lemas previos que nos facilitarán el proceso. Así, una vez vistos estos resultados, habremos construido la base para el lema de Simons

En primer lugar vamos a recordar la definición de funcional sublineal sobre un espacio vectorial V . Notar que todos los espacios vectoriales que vamos a usar son reales. Del mismo modo, los conjuntos que usaremos asumiremos que son no vacíos

Definición 1.1. Sea V un espacio vectorial distinto de cero. Decimos que el $P : V \rightarrow \mathbb{R}$ es sublineal si cumple las siguientes condiciones:

- P es subaditiva: $x_1, x_2 \in V \implies P(x_1 + x_2) \leq P(x_1) + P(x_2)$
- P es positivamente homogénea: $x_1 \in V$ y $\lambda > 0 \implies P(\lambda x) = \lambda P(x)$

Por ejemplo, toda seminorma sobre V es un funcional sublineal. También, si $V = \mathbb{R}$ y definimos $P(x) = \max\{0, x\}$, $\forall x \in \mathbb{R}$ obtenemos un funcional sublineal sobre \mathbb{R} .

El lema que exponemos a continuación, y que generalizaremos posteriormente en el lema 1.2, nos servirá para demostrar el teorema de Hanh-Banach.

Lema 1.1. Sea V un espacio vectorial distinto de cero y $P : V \rightarrow \mathbb{R}$ un funcional sublineal. Fijamos un elemento $y \in V$. Para todo $x \in V$ tomamos

$$P_y(x) := \inf_{\lambda > 0} [P(x + \lambda y) - \lambda P(y)]$$

Entonces, $P_y : V \rightarrow \mathbb{R}$, P_y es sublineal, $P_y \leq P$ y $P_y(-y) \leq -P(y)$.

Demostración. Fijamos $y \in V$. Sea $x \in V$ y $\lambda > 0$, como P es sublineal tenemos:

$$\lambda P(y) = P(\lambda y) = P(\lambda y + x - x) \leq P(x + \lambda y) + P(-x)$$

Por lo tanto, se obtiene que $P(x + \lambda y) - \lambda P(y) \geq -P(-x)$. Tomando el ínfimo sobre $\lambda > 0$ llegamos a $P_y(x) \geq -P(-x) > -\infty$. Por consiguiente, $P_y : V \rightarrow \mathbb{R}$.

Probaremos ahora que P_y es sublineal. Empezamos viendo la subaditividad. Tomamos $x_1, x_2 \in V$ y sean $\lambda_1, \lambda_2 > 0$ arbitrarios. Entonces:

$$\begin{aligned} [P(x_1 + \lambda_1 y) - \lambda_1 P(y)] + [P(x_2 + \lambda_2 y) - \lambda_2 P(y)] \\ \geq [P(x_1 + x_2(\lambda_1 + \lambda_2)y)] - (\lambda_1 + \lambda_2)P(y) \\ \geq P_y(x_1 + x_2) \end{aligned}$$

Tomando ínfimo sobre λ_1 y λ_2 , $P_y(x_1) + P_y(x_2) \geq P_y(x_1 + x_2)$. Así, P_y es subaditiva. Para comprobar que es positivamente homogénea tomamos $x \in V$ y $\mu > 0$. Entonces:

$$\begin{aligned} P_y(\mu x) &= \inf_{\lambda > 0} [P(\mu x + \lambda y) - \lambda P(y)] = \mu \inf_{\lambda > 0} [P(\mu x + (\lambda/\mu)y) - (\lambda/\mu)P(y)] \\ &= \mu \inf_{v > 0} [P(\mu x + v y) - v P(y)] = \mu P_y(x) \end{aligned}$$

Obtenemos que P_y es positivamente homogénea y como consecuencia sublineal.

Para demostrar que $P_y \leq P$, sea $x \in V$ y tomando $\lambda = 1$:

$$P_y(x) \leq P(x + y) - P(y) \leq P(x) + P(y) - P(y) = P(x) \implies P_y \leq P$$

Finalmente, actuando de manera similar al caso anterior:

$$P_y(-y) \leq P(-y + y) - P(y) = -P(y)$$

■

Ahora procedemos a probar el teorema de Hanh-Banach para funcionales sublineales, el cual es uno de los resultados más importantes del análisis funcional.

Teorema 1.1 (Hanh-Banach). *Sea V un espacio vectorial distinto de cero y $P : V \rightarrow \mathbb{R}$ un funcional sublineal. Entonces existe un funcional lineal L en V tal que $L \leq P$.*

Demostración. Sea \mathcal{Q} el conjunto de funcionales sublineales Q en V tales que $Q \leq P$. Primero probaremos que todo subconjunto \mathcal{T} totalmente ordenado de \mathcal{Q} tiene una cota inferior en \mathcal{Q} . Para $T_1, T_2 \in \mathcal{Q}$ tenemos la relación de orden usual, es decir:

$$T_1 \leq T_2 \iff T_1(x) \leq T_2(x) \quad \forall x \in V$$

Definimos $Q(x) := \inf\{T(x) : T \in \mathcal{T}\}$. Si $x \in V$ y $T \in \mathcal{T}$, como T es subaditiva obtenemos la siguiente relación $0 = T(0) = T(x - x) \leq T(x) + T(-x) \implies T(x) \geq -T(-x)$ (1). Como $T \in \mathcal{Q} \implies T(x) \leq P(x) \implies -T(x) \geq -P(x)$ (2). Usando (1), (2) y tomando ínfimo sobre T llegamos a $Q(x) \geq -P(x) \geq -\infty$. Por lo tanto $Q : V \rightarrow \mathbb{R}$.

Ahora probaremos que Q es subaditiva. Para ello, tomamos $x_1, x_2 \in V$. Sean $T_1, T_2 \in \mathcal{T}$ arbitrarios. Si $T_1 \geq T_2$ (el caso de $T_2 \geq T_1$ es análogo.):

$$T_1(x_1) + T_2(x_2) \geq T_2(x_1) + T_2(x_2) \geq T_2(x_1 + x_2) \geq Q(x_1 + x_2)$$

Concluimos que ambos casos $T_1(x_1) + T_2(x_2) \geq Q(x_1 + x_2)$. Tomando ínfimo en T_1 y T_2 obtenemos que $Q(x_1) + Q(x_2) \geq Q(x_1 + x_2)$. Así, Q es sublineal. Que sea positivamente homogénea es consecuencia de que T también lo es. Dado $\mu > 0$:

$$\begin{aligned} Q(\mu x) &= \inf\{T(\mu x) : T \in \mathcal{T}\} \\ &= \inf\{\mu T(x) : T \in \mathcal{T}\} \\ &= \mu \inf\{T(x) : T \in \mathcal{T}\} \\ &= \mu Q(x) \end{aligned}$$

De este modo, Q es sublineal y como es claro que $Q \leq P \implies Q \in \mathcal{Q}$. Así, es directo que Q es el elemento minimal de \mathcal{T} en \mathcal{Q} .

El lema de Zorn nos proporciona entonces un elemento minimal de \mathcal{Q} que llamaremos L . Tomamos ahora $y \in V$. Con la notación del lema anterior, $L_y : V \rightarrow \mathbb{R}$ es sublineal, $L_y \leq L$ (como consecuencia $L_y \in \mathcal{Q}$) y $L_y(-y) \leq L(-y)$. De hecho, como L es minimal en \mathcal{Q} , $L_y = L$ y por ello $L(-y) \leq L(-y)$. Por otro lado, como L es subaditiva, $L(-y) \geq -L(y)$. Combinando ambas desigualdades, $L(-y) = -L(y)$. Tomamos $x \in V$ y $\lambda < 0$, usando la igualdad anterior llegamos a:

$$L(\lambda x) = L(-(-\lambda)x) = -L(-\lambda x) = -(-\lambda)L(x) = \lambda L(x)$$

obteniendo que L es homogénea. Si $x_1, x_2 \in V$, la subaditividad de L nos da $L(-x_1 - x_2) \leq L(-x_1) + L(-x_2)$. Usando la homogeneidad de L :

$$\begin{aligned} L(x_1 + x_2) &= L(-(-x_1 - x_2)) = -L(-x_1 - x_2) \\ &\geq -L(-x_1) - L(-x_2) = L(x_1) + L(x_2) \geq L(x_1 + x_2) \end{aligned}$$

Por ello, $L(x_1 + x_2) = L(x_1) + L(x_2)$ y concluimos que L es lineal. ■

El siguiente resultado importante que demostraremos será el teorema de Mazur-Orlicz. Primero, veamos un lema previo.

Lema 1.2. *Sea V un espacio vectorial distinto de cero y $P : V \rightarrow \mathbb{R}$ un funcional sublineal. Sea D un subconjunto no vacío y convexo de V y $\beta := \inf_D P \in \mathbb{R}$. Para todo $x \in V$ tomamos*

$$Q(x) := \inf_{d \in D, \lambda > 0} [P(x + \lambda d) - \lambda \beta]$$

Entonces, $Q : V \rightarrow \mathbb{R}$, Q es sublineal, $Q \leq P$ y $\forall d \in D, -Q(-d) \geq \beta$.

Demostración. Si $x \in V$, $d \in D$ y $\lambda > 0$ entonces

$$P(x + \lambda d) - \lambda \beta \geq -P(-x) + \lambda P(d) - \lambda \beta \geq -P(-x) \geq -\infty$$

La primera igualdad se deduce de la linealidad de P ya que:

$$\lambda P(d) = P(\lambda d) = P(\lambda d + x - x) \leq P(x + \lambda d) + P(-x) \implies -P(-x) \leq P(x + \lambda d)$$

Y la segunda a que como $\beta = \inf_D P \implies \lambda P(d) \geq \lambda \beta \implies \lambda P(d) - \lambda \beta \geq 0$. Tomando el ínfimo sobre $d \in D$ y $\lambda > 0$ llegamos a $Q(x) \geq -P(-x) > -\infty$. Por consiguiente, $Q : V \rightarrow \mathbb{R}$. Es relativamente fácil probar que Q es positivamente homogénea por lo que para ver que es sublineal solo queda ver la subaditividad. Para ello, tomamos $x_1, x_2 \in V$. Sean $d_1, d_2 \in D$ y $\lambda_1, \lambda_2 > 0$ arbitrarios. Para simplificar la notación llamamos $x := x_1 + x_2$, $\lambda := \lambda_1 + \lambda_2$ y $d := (\lambda_1/\lambda)d_1 + (\lambda_2/\lambda)d_2$. Notar que $d \in D$ al ser este convexo. Entonces:

$$\begin{aligned} [P(x_1 + \lambda_1 d_1) - \lambda_1 \beta] + [P(x_2 + \lambda_2 d_2) - \lambda_2 \beta] &\geq P(x + \lambda_1 d_1 + \lambda_2 d_2) - \lambda \beta \\ &= P(x + \lambda d) - \lambda \beta \\ &\geq Q(x) = Q(x_1 + x_2) \end{aligned}$$

Tomando ínfimo sobre d_1, d_2, λ_1 y λ_2 , $Q(x_1) + Q(x_2) \geq Q(x_1 + x_2)$. Así, Q es subaditiva y como consecuencia sublineal. Fijamos $d \in D$. Sea $x \in V$ arbitrario. Entonces, $\forall \lambda > 0$, $Q(x) \leq P(x) + \lambda [P(d) - \beta]$. Tomando $\lambda \rightarrow 0$, $Q(x) \leq P(x)$ y como consecuencia $Q \leq P$. Finalmente, sea $d \in D$ arbitrario y tomando $\lambda = 1$:

$$Q(-d) \leq Q(-d + d) - \beta = -\beta \implies -Q(-d) \geq \beta$$

■

Visto este lema, estamos preparados para ver el resultado que nos interesa:

Teorema 1.2 (Mazur-Orlicz). *Sea V un espacio vectorial distinto de cero y $P : V \rightarrow \mathbb{R}$ un funcional sublineal. Sea D un subconjunto no vacío y convexo de V . Entonces existe un funcional lineal L sobre V tal que $L \leq P$ e $\inf_D L = \inf_D P$*

Demostración. Sea $\beta := \inf_D P$. En el caso de que $\beta = -\infty$ por el teorema de Hanh-Banach tenemos que $\exists L$ sobre V tal que es lineal y $L \leq P$. Así:

$$L \leq P \implies \inf_D L \leq \inf_D P = -\infty \implies \inf_D L = \inf_D P$$

Supongamos entonces que $\beta \in \mathbb{R}$. Definimos el funcional auxiliar Q como en el lema 1.2. Del teorema de Hanh-Banach obtenemos que existe un funcional lineal L sobre V tal que $L \leq Q$ (como $Q \leq P$ tenemos que $L \leq P$). Sea $d \in D$, entonces:

$$L(d) = -L(-d) \geq -Q(-d) \geq \beta$$

Tomando ínfimo sobre $d \in D$:

$$\inf_D L \geq \beta = \inf_D P$$

Por otro lado, como $L \leq P$:

$$\inf_D L \leq \inf_D P$$

Juntando ambas desigualdades obtenemos $\inf_D L = \inf_D P$ ■

Lema de Simons y Teorema de la Alternativa de Gordan

Este capítulo se centra en lema de Simons, el cual será de gran ayuda para demostrar el Teorema de la Alternativa de Gordan.

Antes de comenzar, necesitamos hacer la siguiente definición.

Definición 2.1. Dado $N \in \mathbb{N}$ con $N > 1$ llamamos *símplex unitario* de \mathbb{R}^N al conjunto conexo y compacto de \mathbb{R}^N dado por:

$$\Delta_N := \left\{ \mathbf{t} \in \mathbb{R}^N : \sum_{i=1}^N t_i = 1, t_1, \dots, t_N \geq 0 \right\}$$

Por ejemplo, si $N = 2$, tenemos que $\Delta_2 = \text{co}\{(1, 0), (0, 1)\}$, es decir es la envolvente convexa de los vectores de la base usual de \mathbb{R}^2 . De hecho en general, $\Delta_N = \text{co}\{e_1, \dots, e_N\}$.

Enunciamos este lema provio que usaremos posteriormente el la demostración del de Simons.

Lema 2.1. Sea $N \in \mathbb{N}$, con $N \geq 1$ y $L : \mathbb{R}^N \rightarrow \mathbb{R}$ definida por $S(x) := \max\{x_1, \dots, x_N\}$. Entonces, S es sublineal. Además, si $L : \mathbb{R}^N \rightarrow \mathbb{R}$ es un funcional lineal tal que $L \leq S$ entonces L es de la forma $L(x) = t_1 x_1 + \dots + t_N x_N$ con $(t_1, \dots, t_N) \in \Delta_N$. De hecho, el recíproco también es cierto, es decir, si $L = (t_1, \dots, t_N) \in \Delta_N$ entonces $L \leq S$.

Demostración. Claramente, S es positivamente homogénea. También es subaditiva ya que dados $x, y \in V$:

$$\begin{aligned} S(x + y) &= \max\{x_1 + y_1, \dots, x_N + y_N\} \\ &\leq \max\{x_1, \dots, x_N\} + \max\{y_1, \dots, y_N\} = S(x) + S(y) \end{aligned}$$

Por ello, S es sublineal. Para terminar veamos que

$$\{L : \mathbb{R}^N \rightarrow \mathbb{R} : L \text{ lineal y } L \leq S\} = \Delta_N$$

\supseteq) Sea $\mathbf{t} \in \Delta_N$ definimos $L : \mathbb{R}^N \rightarrow \mathbb{R}$ como $L(x) := \langle \mathbf{t}, x \rangle$. Es evidente que L es lineal al serlo el producto escalar. Dado $x \in \mathbb{R}^N$:

$$L(x) = \sum_{i=1}^N t_i x_i \leq \sum_{i=1}^N t_i S(x) = S(x) \sum_{i=1}^N t_i = S(x)$$

donde la primera desigualdad se debe a que $x_i \leq S(x)$ para todo x_i con $i = 1, \dots, N$ y a que $t_i \geq 0$ ya que $\mathbf{t} \in \Delta_N$. Esto también justifica la que última igualdad ya que $\sum_{i=1}^N t_i = 1$.

\subseteq) Sea $L = (t_1, \dots, t_N) \in \mathbb{R}^N$ lineal tal que $\forall x \in \mathbb{R}^N$ cumple $\sum_{i=1}^N t_i x_i \leq \max\{x_1, \dots, x_N\}$. Así, si tomamos $e_i \in \mathbb{R}^N$ donde e_i representa el i -ésimo elemento de la base usual de \mathbb{R}^N con $i = 1, \dots, N$. Entonces:

$$L(-e_i) = -t_i \leq 0 \implies t_i \geq 0 \quad \forall i = 1, \dots, N$$

Si ahora llamamos $e = \sum_{i=1}^N e_i$. Obtenemos:

$$\left. \begin{aligned} L(e) &= \sum_{i=1}^N t_i \leq \max\{1, \dots, 1\} = 1 \\ L(-e) &= -\sum_{i=1}^N t_i \leq \max\{-1, \dots, -1\} = -1 \end{aligned} \right\} \implies \sum_{i=1}^N t_i = 1$$

Concluimos entonces que $L = (t_1, \dots, t_N) \in \Delta_N$.

■

Enunciamos ahora lema de Simons[Sim08] en que destacamos la ausencia de hipótesis topológicas, lo que será importante posteriormente.

Lema 2.2 (Lema de Simons). *Sea C un subconjunto no vacío y convexo de un espacio vectorial. Dadas f_1, \dots, f_N ($N \geq 1$) funciones sobre C reales y convexas, entonces existen $(t_1, \dots, t_N) \in \Delta_N$ que cumplen*

$$\inf_C [\max\{f_1, \dots, f_N\}] = \inf_C [t_1 f_1 + \dots + t_N f_N]$$

Demostración. Sea $V = \mathbb{R}^N$ con $N \in \mathbb{N}$. Definimos $S : V \rightarrow P$ como

$$S(x_1, \dots, x_N) := \max\{x_1, \dots, x_N\}$$

. Por el lema 2.1 S es sublineal. Tomamos el subconjunto:

$$D = \{(x_1, \dots, x_N) \in V : \exists c \in C \text{ tal que } \forall i = 1, \dots, N, \quad f_i(c) \leq x_i\}$$

D es un subconjunto convexo de V . Sean $x, y \in D$, por ello, existen $c_x, c_y \in C$ tales que $f_i(c_x) \leq x_i$ y $f_i(c_y) \leq y_i \quad \forall i = 1, \dots, N$. Dado $\lambda \in [0, 1]$, llamamos $c := (1 - \lambda)c_x + \lambda c_y$ que pertenece a C por ser este convexo.

Veamos que c es el elemento necesario de C para que cualquier combinación convexa de x e y esté en D . Así, para todo $i = 1, \dots, N$:

$$f_i(c) = f_i((1 - \lambda)c_x + \lambda c_y) \leq (1 - \lambda)f(c_x) + \lambda f(c_y) \leq (1 - \lambda)x_i + \lambda y_i$$

donde la primera desigualdad se debe a que las f_i son convexas y la segunda a que $x, y \in D$. Por ello, $(1 - \lambda)x_i + \lambda y_i \in D$, $\forall \lambda \in [0, 1]$ por lo que D es convexo. Aplicando el Teorema de Mazur-Orlitz, existe L funcional lineal sobre V tal que $L \leq S$ e $\inf_D L = \inf_D S$.

Nuevamente, por el lema 2.1 tenemos que $L = (t_1, \dots, t_N) \in \Delta_N$.

Finalmente:

$$\inf_D L = \inf_{c \in C} [t_1 f(c) + \dots + t_N f(c)] = \inf_C [t_1 f + \dots + t_N f]$$

y

$$\inf_D S = \inf_{c \in C} [\max\{f_1(c), \dots, f_N(c)\}] = \inf_C [\max\{f_1, \dots, f_N\}]$$

por lo que

$$\inf_C [t_1 f + \dots + t_N f] = \inf_C [\max\{f_1, \dots, f_N\}]$$

■

Enunciamos ahora el Teorema de la alternativa de Gordan en su versión convexa.

Teorema 2.1 (Teorema de la Alternativa de Gordan-versión convexa).
Sea C un subconjunto convexo de un espacio vectorial y sean $f_1, \dots, f_N : C \rightarrow \mathbb{R}$ funciones convexas ($N \geq 1$). Entonces una, y solo una, de la siguientes afirmaciones se cumple:

i) $\exists t \in \Delta_N$ tal que $0 \leq \inf_{c \in C} \sum_{i=1}^N t_i f_i(c)$.

ii) $\exists c \in C$ que cumple $\max_{i=1, \dots, N} f_i(c) < 0$.

Demostración. Si aplicamos el lema de Simons, lema 2.2 a las funciones f_1, \dots, f_N obtenemos:

$$\exists t \in \Delta_N : \inf_{c \in C} \left[\max_{i=1, \dots, N} f_i(c) \right] = \inf_{c \in C} \sum_{i=1}^N t_i f_i(c)$$

Supongamos en primer lugar que $\alpha := \inf_{c \in C} [\max_{i=1, \dots, N} f_i(c)] \in \mathbb{R}$. Planteamos la siguiente alternativa, cuyos casos son excluyentes:

a) $0 \leq \alpha$: implica i) ya que $\alpha = \inf_{c \in C} \sum_{i=1}^N t_i f_i(c)$

b) $\alpha < 0$: este caso, por su parte, implica *ii*) ya que:

$$\alpha = \inf_{c \in C} \left[\max_{i=1, \dots, N} f_i(c) \right] < 0 \iff \exists c \in C : \max_{i=1, \dots, N} f_i(c) < 0$$

Para finalizar, veamos cuando $\alpha = -\infty$. En este caso, estamos en la misma situación que en b) por lo que solo se puede dar *ii*). ■

Destacamos las siguientes observaciones:

Observación 1. *Esta versión convexa del teorema implica la versión clásica del mismo.*

Dados $\{x_1, \dots, x_N\}$ con $x_i \in \mathbb{R}^M$, ($M \geq 1$) $i = 1, \dots, N$, la versión clásica del teorema nos aporta las siguientes alternativas:

$$i^*) \exists t \in \Delta_N \text{ tal que } 0 = \sum_{i=1}^N t_i x_i.$$

$$ii^*) \exists y \in \mathbb{R}^M \text{ tal que cumple } \max_{i=1, \dots, N} \langle y, x_i \rangle < 0.$$

Para ello, basta aplicar la versión convexa del teorema a $C := \mathbb{R}^M$ y a las funciones $f_1, \dots, f_N : C \rightarrow \mathbb{R}$ definidas por $f_i(c) := \langle c, x_i \rangle$, $\forall i = 1, \dots, N$. En este caso, la alternativa *ii*) implica *ii**) ya que:

$$\exists c \in C = \mathbb{R}^M : \max_{i=1, \dots, N} \langle c, x_i \rangle = \max_{i=1, \dots, N} f_i(c) < 0$$

Por su parte, la alternativa *i*) nos da:

$$\exists t \in \Delta_N : 0 \leq \inf_{c \in \mathbb{R}^M} \sum_{i=1}^N t_i f_i(c) = \inf_{c \in \mathbb{R}^M} \sum_{i=1}^N t_i \langle c, x_i \rangle = \inf_{c \in \mathbb{R}^M} \langle c, \sum_{i=1}^N t_i x_i \rangle$$

Si para $c \in \mathbb{R}^M$ se cumple que $\langle c, \sum_{i=1}^N t_i x_i \rangle = 0$ entonces podemos deducir que $\sum_{i=1}^N t_i x_i = 0$ ya que $\sum_{i=1}^N t_i x_i \in (\mathbb{R}^M)^\perp = \{0\}$. Así pues, tenemos que

$$\text{Se cumple } i) \iff \exists t \in \Delta_N : 0 = \sum_{i=1}^N t_i x_i$$

Es claro entonces que obtenemos *i**).

Observación 2. *El lema de Simons (lema 2.2) y el Teorema de la Alternativa de Gordan (teorema 2.1) son equivalentes.*

Ya hemos visto que el Lema de Simons implica el Teorema de la Alternativa de Gordan. Veamos que el recíproco también es cierto. COMPLETAR.

Bibliografía

- [Sim08] Stephen Simons. *From Hahn-Banach to Monotonicity*. Springer, 2^a edición, 2008.