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, estaremos preparados para enunciar los conocidos como teoremas de la alternativa necesarios en el campo de la optimización convexa.

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 subconjuntos que usaremos asumiremos que son no vacíos

Definición 0.1. *Sea V un espacio vectorial distinto de cero. Decimos que el $P : V \to \mathbb{R}$ es sublineal si cumple las siguientes condiciones:*

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

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

El lema que exponemos a continuación, y que generalizaremos posteriormente en 0.2, nos servirá para demostrar el teorema de Hanh-Banach.

Lema 0.1. *Sea V un espacio vectorial distinto de cero y $P : V \to \mathbb{R}$ un funcional sublineal. Fijamos un elemento $y \in V$. Para todo $x \in V$ tomamos*

$$P_y(x) := \text{ínf}_{\lambda>0}\,[P(x+\lambda y) - \lambda P(y)]$$

Entonces, $P_y : V \to \mathbb{R}$, 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, $P_y : V \to \mathbb{R}$.

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:

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

Obtenemos que P_y es positivamente homogénea y como consecuencia sublineal.

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

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

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-Banch para funcionales sublineales, el cual es uno de los resultados más importantes del análisis funcional.

Teorema 0.1 (Hanh-Banach). *Sea V un espacio vectorial distinto de cero y $P : V \to \mathbb{R}$ 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 $T_1, T_2 \in \mathcal{Q}$ definimos:

$$T_1 \leq T_2 \Longleftrightarrow T_1(x) \leq T_2(x) \quad \forall x \in \mathrm{V}$$

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

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 $T_1 \geq T_2$:

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

El caso de $T_2 \geq T_1$ es análogo. 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 $\mu > 0$:

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

De este modo, Q es sublineal y como es claro que $Q \leq P \Longrightarrow Q \in \mathcal{Q}$. Así, 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$\longrightarrow \mathbb{R}$ 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 0.2. *Sea V un espacio vectorial distinto de cero y $P : V \to \mathbb{R}$ un funcional sublineal. Sea D un subconjunto no vacío y convexo de V y $\beta := \text{ínf}_D P \in \mathbb{R}$. Para todo $x \in V$ tomamos*

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

Entonces, $Q : V \to \mathbb{R}$, 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:

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

Y la segunda a que como $\beta = \text{ínf}_D P \Longrightarrow \lambda P(d) \geq \lambda\beta \Longrightarrow \lambda P(d) - \lambda\beta \geq 0$. Tomando el ínfimo sobre $d \in D$ y $\lambda > 0$ llegamos a $Q(x) \geq -P(-x) > -\infty$. Por consiguiente, $Q : V \to \mathbb{R}$. 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 $\lambda \longrightarrow 0$, $Q(x) \leq P(x)$ y como consecuencia $Q \leq P$. Finalemente, sea $d \in D$ arbitrario y tomando $\lambda = 1$:

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

■

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

Teorema 0.2 (Mazur-Orlicz). *Sea V un espacio vectorial distinto de cero y $P : V \to \mathbb{R}$ 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 $\operatorname{ínf}_D L = \operatorname{ínf}_D P$*

Demostración. Sea $\beta := \operatorname{ínf}_D P$. En el caso de que $\beta = -\infty$ por el teorema 0.1 tenemos que $\exists L$ sobre V tal que es lineal y $L \leq P$. Así:

$$L \leq P \Longrightarrow inf_D L \leq \operatorname{ínf}_D P = -\infty \Longrightarrow inf_D L = \operatorname{ínf}_D P$$

Supongamos entonces que $\beta \in \mathbb{R}$. Definimos el funcional auxiliar Q tal y como en el lema 0.2. Del teorema 0.1 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 $d \in D$:

$$\operatorname*{ínf}_D L \geq \beta = \operatorname*{ínf}_D P$$

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

$$\operatorname*{ínf}_D L \leq \operatorname*{ínf}_D P$$

Juntando ambas desigualdades obtenemos $\operatorname{ínf}_D L = \operatorname{ínf}_D P$ ■

Lema de Simons y Teorema de la Alternativa de Gordan

Enunciamos ahora el siguiente lema en que destacamos la ausencia de hipótesis topológicas, lo que será importante posteriormente [SH08].

Lema 0.3. *Sea C un subconjunto no vacío y convexo de un espacio vectorial. Dadas $f_1, ..., f_N$ funciones sobre C reales y convexas, entones existen $\lambda_1, ..., \lambda_N \geq 0$ tales que $\lambda_1 + ... + \lambda_N = 1$ y*

$$\inf_C [máx\{f_1, ..., f_N\}] = \inf_C [\lambda_1 f_1 + ... + \lambda_N f_N]$$

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

$$S(x_1, ..., x_N) := máx\{x_1, ..., x_N\}$$

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

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

Por ello, S es sublineal. Tomamos el subconjunto:

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

Veamos que 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 $f_i(c_y) \leq y_i \quad \forall i = 1, ..., N$. 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:

$$f_i(c) = f_i((1-\lambda)c_x + \lambda c_y) \leq (1-\lambda)f(c_x) + \lambda f(c_y) \leq (1-\lambda)x_i + \lambda y_i \quad, \forall i = 1, ..., N$$

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

Al ser L lineal, existen $\lambda_1, .., \lambda_N \in \mathbb{R}$ tales que:

$$L(x) = \lambda_1 x_1 + .. + \lambda_N x_N, \forall x \in V$$

Como $L \leq S$ tenemos que $\lambda_1 x_1 + .. + \lambda_N x_N \leq máx\{x_1, ..., x_N\}$ por lo que también se tiene que cumplir que $\lambda_1, ..., \lambda_N \geq 0$ y $\lambda_1 + ... + \lambda_N = 1$.

Finalmente:

$$\underset{D}{\text{ínf}}\, L = \underset{c \in C}{\text{ínf}}\, [\lambda_1 f(c) + .. + \lambda_N f(c)] = \underset{C}{\text{ínf}}\, [\lambda_1 f + .. + \lambda_N f]$$

y

$$\underset{D}{\text{ínf}}\, S = \underset{c \in C}{\text{ínf}}\, [máx\{f_1(c), ..., f_N(c)\}] = \underset{C}{\text{ínf}}\, [máx\{f_1, ..., f_N\}]$$

por lo que

$$\underset{C}{\text{ínf}}\, [\lambda_1 f + .. + \lambda_N f] = \underset{C}{\text{ínf}}\, [máx\{f_1, ..., f_N\}]$$

$\blacksquare$

Bibliografía

[SH08] Stephen Simons y Hola. *From Hanh-Banach to Monotonicity*. Springer, 2 edición, 2008.