Lección n°8: Envolturas convexas y puntos extremales

EPN, verano 2012

Recordemos para empezar que una *combinación convexa* de vectores es una combinación lineal en dónde todos los coefficientes son positivos y de suma uno. Es decir que, dada una colección de vectores $x_1, ..., x_n$, su combinación convexa es un vector x definido por la expresión

$$x = \sum_{i=1}^{n} \alpha_i x_i \qquad \text{con } \alpha_i \geq 0 \text{ y } \sum_{i=1}^{n} \alpha_i = 1.$$

Definición 1 (Envoltura convexa) *Sea E un $\mathbb{K}$-espacio vectorial y sea $A \subset E$ un subconjunto de E. La* envoltura convexa *de A, notada $co(A)$ es la intersección de todos los conjuntos convexos de E que contienen A. De forma equivalente, $co(A)$ es el conjunto de todas las combinaciones convexas finitas de los elementos de A.*

Nótese que se tiene que $A \subset co(A)$: así por ejemplo, en el plano, si se consideran tres puntos a, b, c distintos, se tiene que $co(\{a, b, c\})$ es el triángulo que se obtiene tomando estos puntos como vértices. Esto muestra que la envoltura convexa de un conjunto puede ser muy diferente del conjunto inicial y que sus distintas propiedades pueden resultar muy particulares. Evidentemente, si A es un conjunto convexo, se tiene $A = co(A)$.

Hemos visto en secciones anteriores las relaciones existentes entre conjuntos acotados y conjuntos compactos, en particular vimos que todo conjunto compacto es un conjunto acotado. Uno de lo objetivos de esta lección es estudiar las relaciones entre conjuntos compactos y sus envolturas convexas, pues estos ingredientes intervienen en el enunciado del teorema de Krein-Milman, de manera que para llevar a cabo esta misión es necesario precisar un poco la noción de conjunto acotado.

Definición 2 (Conjuntos Totalmente acotados)

1) *Sea E un $\mathbb{K}$-espacio vectorial topológico. Un subconjunto A de E es* totalmente acotado *en E si para cada vecindad V del origen de E corresponde un conjunto finito F tal que $A \subset F + V$.*

2) *Sea (E, d_E) un espacio métrico. Un subconjunto A de E es* totalmente acotado *en E si A está contenido en la unión finita de bolas abiertas de radio ε para todo $\varepsilon > 0$.*

Si E es un espacio vectorial topológico metrizable, entonces estas dos condiciones coinciden si nos restringimos a las distancias compatibles con la estructura topológica.

El siguiente lema nos indica una primera relación entre los conjuntos compactos y los conjuntos totalmente acotados en los espacios métricos completos.

Lema 1 *Sea K un subconjunto cerrado de un espacio métrico completo (E, d_E); entonces los tres puntos siguientes son equivalentes*

1) *K es compacto,*

2) *si A es conjunto infinito de K, entonces cada sucesión de A posee un punto límite en K,*

3) *K es totalmente acotado.*

Prueba. Empezamos suponiendo que se tiene 1) y supongamos que ningún punto de K es el límite de una sucesión de puntos de A, entonces existe un recubrimiento abierto $(V_i)_{i \in I}$ de K tal que cada V_i contiene al menos un punto de A. Esto implica que de $(V_i)_{i\in I}$ no se puede extraer un subrecubrimiento finito lo cual es una contradicción con el hecho que K es compacto, de donde se tiene que 1) $\Longrightarrow$ 2).

Supongamos ahora que se tiene 2). Fijemos $\epsilon > 0$ y sea $x_1 \in K$. Supongamos que hemos fijado $x_1, ..., x_n \in K$ puntos tales que $d_E(x_i, x_j) \geq \epsilon$ si $i \neq j$. De ser posible, fijamos $x_{n+1} \in K$ tal que $d_E(x_i, x_{n+1}) \geq \epsilon$ para todo $1 \leq i \leq n$. Dado que hemos supuesto que se tiene el punto 2), este proceso debe terminarse después de un número finito de pasos, de lo contrario no se cumpliría la condición $d_E(x_i, x_j) \geq \epsilon$. Se tiene entonces que las bolas centradas en los puntos $x_1, ..., x_n$ de radio ϵ forman un recubrimiento de K de manera que se tiene que 2) $\Longrightarrow$ 3).

Finalmente, suponemos que se tiene el punto 3). Sea Λ un recubrimiento abierto de K y supongamos que ninguna subfamilia de Λ recubre K. Sabemos por hipótesis que K es la unión finita de conjuntos cerrados de diámetro $r \leq 1$. Se tiene entonces que uno de estos conjuntos, digamos C_0, no puede ser recubierto por medio de una familia finita de elementos de Λ. Podemos repetir este proceso tomando C_0 en vez de K y el resultado es una sucesión de conjuntos cerrados C_i tales que $K \supset C_0 \supseteq C_1 \supseteq C_2 ...$, tales que $diam(C_n) \leq 1/n$ y tales que ningún C_n puede ser recubierto por una familia finita de elementos de Λ. Fijemos ahora un punto $x_n \in C_n$. Dado que el espacio métrico E es completo y que cada conjunto C_n es cerrado se obtiene que la sucesión $(x_n)_{n\in\mathbb{N}}$ es una sucesión de Cauchy que converge hacia un punto $x \in \bigcap_{n\in\mathbb{N}} C_n$. por lo tanto se tiene que $x \in V$ para alguna vecindad $V \in \Lambda$. Dado que $C_n \subset V$ si n es suficientemente grande se obtiene una contradicción y se concluye que 3) $\Longrightarrow$ 1). ■

Necesitaremos el resultado que sigue pues clarifica lo que sucede en el caso de la dimensión finita:

Lema 2 *Si $A \subset \mathbb{R}^n$ y si $x \in co(A)$, entonces x pertenece a la envoltura convexa de algún subconjunto de A que contiene a lo mucho $n + 1$ puntos.*

Prueba. Es suficiente mostrar que si $k > n$ y si $x = \sum_{i=1}^{k+1} t_i x_i$ es una combinación convexa de vectores $x_i \in \mathbb{R}^n$, entonces x es en realidad una combinación convexa de k de estos vectores. Para ello asumimos, sin pérdida de generalidad que $t_i > 0$ para todo $1 \leq i \leq k + 1$. Observamos ahora que el núcleo de la aplicación

$$(a_1, ..., a_{k+1}) \longmapsto (\sum_{i=1}^{k+1} a_i x_i, \sum_{i=1}^{k+1} a_i)$$

que envía $\mathbb{R}^{k+1}$ en $\mathbb{R}^n \times \mathbb{R}$ tiene dimensión positiva puesto que $k > n$. Existe por lo tanto un vector $(a_1, ..., a_{k+1})$ no nulo tal que $\sum_{i=1}^{k+1} a_i x_i = 0$ y $\sum_{i=1}^{k+1} a_i = 0$. Dado que $t_i > 0$, existe una constante $c > 0$ tal que $|ca_i| \leq t_i$ para todo i y $ca_j = t_j$ para al menos un cierto j. Definimos ahora $b_i = t_i - ca_i$ y obtenemos que $x = \sum_{i=1}^{k+1} b_i x_i$ y que al menos uno de estos b_i es nulo. Notamos también que $\sum_{i=1}^{k+1} b_i = \sum_{i=1}^{k+1} t_i = 1$ y que $b_i \geq 1$ para todo i. Hemos verificado que es posible reducir el número de vectores de $k+1$ a k en la representación convexa del vector x. ■

Con estos lemas obtenemos la proposición siguiente que explica, en diversos casos, algunas relaciones entre los conjuntos compactos y su envoltura convexa.

Proposición 1

1) *Sea E un espacio vectorial topológico y sean $A_1, ..., A_n$ subconjuntos convexos compactos. Entonces $co(\bigcup_{i=1}^n A_i)$ es un conjunto compacto.*

2) *Sea E un espacio vectorial topológico localmente convexo separado. Si $A \subset E$ es totalmente acotado, entonces $co(A)$ es un conjunto totalmente acotado.*

3) *Si $(E, (p_n)_{n\in\mathbb{N}})$ es un espacio de Fréchet y si $K \subset E$ es un conjunto compacto, entonces $\overline{co}(K)$ es un conjunto compacto.*

4) *Si K es un subconjunto compacto de $\mathbb{R}^n$, entonces $co(K)$ es un conjunto compacto.*

Demostración.

1) Sea S el subconjunto de $\mathbb{R}^n$ formado por los vectores $s = (s_1, ..., s_n)$ con $s_i \geq 0$ y tal que $s_1 + \cdots + s_n = 1$.

Sea $A = \prod_{i=1}^n A_i$ y definamos la función f por medio de la expresión

$$\begin{array}{rcl} f : S \times A & \longrightarrow & E \\ (s,x) & \longmapsto & f(s,x) = \displaystyle\sum_{i=1}^n s_i x_i \end{array}$$

Definamos ahora $K = f(S \times A)$. Dado que todos los conjuntos A_i son compactos, se tiene que K es compacto y se tiene por construcción que $K \subset co(\bigcup_{i=1}^n A_i)$. Para demostrar que $co(\bigcup_{i=1}^n A_i)$ es un conjunto compacto, vamos a demostrar que $co(\bigcup_{i=1}^n A_i) \subset K$.
Vemos que si (s,x) y (t,y) son elementos de $S \times A$ y si $\alpha, \beta \geq 0$ con $\alpha + \beta = 1$ son dos escalares, entonces

$$\alpha f(s,x) + \beta f(t,y) = f(r,z)$$

en donde $r = \alpha s + \beta t$ y $z \in A$ puesto que se tiene

$$z_i = \frac{\alpha s_i x_i + \beta t_i y_i}{\alpha s_i + \beta t_i} \in A_i \qquad (1 \leq i \leq n)$$

Esto muestra que K es un conjunto convexo. Como se tiene $A_i \subset K$ para cada i, la convexidad de K implica que $co(\bigcup_{i=1}^n A_i) \subset K$ y así terminamos la demostración del primer punto.

2) Sea U una vecindad del origen de E y fijemos una vecindad convexa V del origen de E tal que $V + V \subset U$. Entonces se tiene que $A \subset F + V$ para algún conjunto finito $F \subset E$. Por lo tanto $A \subset co(F) + V$ y observamos que el conjunto $co(F) + V$ es convexo, por lo tanto $co(A) \subset co(F) + V$.
Por el primer punto, se tiene que $co(F)$ es compacto y por lo tanto $co(F) \subset F_1 + V$ para algún conjunto finito $F_1 \subset E$. Se tiene entonces $co(A) \subset F_1 + V + V \subset F_1 + U$, dado que la vecindad U era arbitraria se tiene que $co(A)$ es totalmente acotado.

3) Las cerraduras de conjuntos totalmente acotados son conjuntos totalmente acotados en cada espacio métrico y usando el lema 1 vemos que estos conjuntos son compactos en todo espacio métrico completo. Por lo tanto, si K es un conjunto compacto en un espacio de Fréchet entonces K es totalmente acotado y entonces $co(K)$ es totalmente acotado. Aplicamos entonces el punto anterior para obtener que $\overline{co}(K)$ es un conjunto compacto.

4) Sea S el subconjunto de $\mathbb{R}^{n+1}$ determinado por todos los vectores $t = (t_1, ..., t_{n+1})$ tales que $t_i \geq 0$ y $\sum_{i=1}^{n+1} t_i = 1$. Sea $K \subset \mathbb{R}^n$ un conjunto compacto. Por el lema 2 se tiene que $x \in co(K)$ si y solo si

$$x = \sum_{i=1}^{n+1} t_i x_i$$

para algún $t \in S$ y $x_i \in K$. Dicho de otra manera $co(K)$ es la imagen de $S \times K^{n+1}$ en $\mathbb{R}^n$ bajo la aplicación continua $(t, x_1, ..., x_{n+1}) \longmapsto \sum_{i=1}^{n+1} t_i x_i$ y por lo tanto $co(K)$ es un conjunto compacto. ■

Cuando el espacio dual E' de un espacio vectorial separa los puntos se tiene el resultado a continuación que es una versión particular del teorema de separación de convexos.

Proposición 2 *Sea E un $\mathbb{K}$-espacio vectorial topológico tal que E' separa sus puntos. Si A, B son dos conjuntos disjuntos, no vacíos, convexos y compactos de E, entonces existe una forma lineal $T \in E'$ tal que*

$$\sup_{x \in A} \Re e(T(x)) < \inf_{y \in B} \Re e(T(y))$$

Prueba. Notamos, por comodidad, E_σ el espacio E dotado de su topología débil $\sigma(E, E')$. Vemos que los conjuntos A y B son compactos en E_σ y son cerrados en E_σ puesto que E_σ es un espacio separado. Como E_σ es localmente convexo separado podemos aplicar el teorema de separación de conjuntos convexos utilizando el espacio E_σ en vez de E: de esta manera se obtiene que existe una forma lineal $T \in (E_\sigma)'$ que satisface las desigualdades buscadas. Pero sabemos, que $(E_\sigma)' = E'$, lo que termina la prueba. ■

Pasamos ahora a presentar una noción que será de gran utilidad.

Definición 3 (Conjuntos extremales y puntos extremales) *Sea E un $\mathbb{K}$-espacio vectorial y sea A un subconjunto de E.*

1) *Diremos que el conjunto $S \subset A$ es un* conjunto extremal *de A si para todo $x, y \in A$ y todo $0 < t < 1$ tal que $(1-t)x + ty \in S$ se tiene que $x, y \in S$.*

2) *Los* puntos extremales *de A son los conjuntos extremales que están formados por un solo punto. Dicho de otra manera, $x \in A$ es un* punto extremal *si para todo $y, z \in A$ y para todo $t \in [0, 1]$ se tiene que la identidad $x = ty + (1-t)z$ implica que $x = y$ ó $x = z$.*

El conjunto de todos los puntos extremales de A será notado $\mathcal{E}(A)$.

Observación 1 Para verificar que un punto x de un conjunto A no es extremal, basta exhibir dos puntos $y, z \in A$ tal que $x \in A$ pertenece al segmento formado por estos dos puntos. En particular, toda parte del interior de un espacio convexo no puede ser extremal; esto significa que los puntos extremales están de alguna manera "en el borde".

Veamos un ejemplo. Consideremos el cuadrado $A = [0, 1] \times [0, 1]$ en el plano $\mathbb{R}^2$. Por la observación anterior vemos que todo punto en el interior de A no puede ser extremal, de manera que hay que buscarlos en el borde de A. Sin embargo no todos los puntos del borde del cuadrado son puntos extremales: si fijamos $x_1 \in]0, 1[$ y definimos un punto de $\mathbb{R}^2$ por $(x_1, 0) = x_1(1, 0)$ vemos que este punto no es un punto extremal, a pesar de estar en el borde del cuadrado. Se puede ver sin mayor problema que en este ejemplo los puntos extremales están dados por los vértices del cuadrado.

La noción de borde utilizada anteriormente merece ser precisada. Recordemos que la frontera de un subconjunto A de un espacio topológico está definida por $\partial A = \overline{A} \setminus \overset{\circ}{A}$. Esta notación permite enunciar un resultado que será de gran utilidad, pero antes una definición.

Definición 4 (Norma estrictamente convexa) *Sea E un espacio vectorial normado. Una norma $\|\cdot\|_E$ definida sobre E es una norma* estrictamente convexa *si para todo $\|x\|_E = \|y\|_E = 1$ con $x \neq y$, con $0 < t < 1$, entonces $\|tx + (1-t)y\|_E < 1$.*
Diremos entonces que el espacio $(E, \|\cdot\|_E)$ es estrictamente convexo.

Es muy importante notar que esta propiedad no se mantiene al pasar a otras normas equivalentes. Ver el ejercicio **??** para más detalles. Demos ahora un ejemplo de aplicación de esta noción de norma estrictamente convexa relativo a los puntos extremales.

Proposición 3 *Si E es un espacio vectorial normado que admite una norma $\|\cdot\|_E$ estrictamente convexa y si notamos B_E la bola unidad de E, entonces se tiene la identidad*

$$\mathcal{E}(B_E) = \partial B_E.$$

Prueba. Tenemos por la observación anterior que $\mathcal{E}(B_E) \subset \partial B_E$, pues los puntos extremales no pueden ser puntos interiores en los conjuntos convexos. Debemos pues verificar la inclusión recíproca para terminar la demostración, para ello mostraremos que todo punto de la frontera ∂B_E es un punto extremal de la bola unidad y procedemos por el absurdo. Sea $x \in \partial B_E$ un punto que no es extremal, entonces existen dos puntos $y, z \in \partial B_E$ distintos de x y $t \in]0, 1[$ tales que $x = ty + (1-t)z$. Dado que $x \in \partial B_E$ y que la norma es estrictamente convexa, se tiene

$$1 = \|x\|_E = \|ty + (1-t)z\|_E < 1,$$

de donde se obtiene una contradicción: los elementos de la frontera ∂B_E son puntos extremales de la bola unidad.■

Podemos ahora sí enunciar y demostrar uno de los resultados más importantes de esta subsección. Como lo veremos un poco más adelante, este teorema permitirá verificar si un cierto espacio normado posee o no un espacio predual.

Teorema 1 (de Krein-Milman) *Sea E un $\mathbb{K}$-espacio vectorial topológico tal que E' separa sus puntos. Si K es un conjunto no vacío, convexo y compacto en E, entonces K es la cerradura de la envoltura convexa de sus puntos extremales. Es decir*
$$K = \overline{co}(\mathcal{E}(K)).$$

Demostración. Sea $\mathscr{P}$ la colección de todos los conjuntos extremales compactos de K. Como por hipótesis $K \in \mathscr{P}$ se tiene que este conjunto no es vacío. Empezamos por un par de observaciones:

i) La intersección de cualquier familia no vacía de $\mathscr{P}$ pertenece a $\mathscr{P}$ (a menos que la intersección sea ella mismo vacía),

ii) Si $S \in \mathscr{P}$, si $T \in E'$ es una forma lineal continua, si notamos μ el máximo de $\mathfrak{Re}(T(x))$ sobre S y si definimos el conjunto $S_T = \{x \in S : \mathfrak{Re}(T(x)) = \mu\}$, entonces se tiene que $S_T \in \mathscr{P}$.

El primer punto no causa ningún problema, de manera que nos concentramos en el segundo. Si suponemos que $tx + (1-t)y = z \in S_T$ con $x, y \in K$ y $0 < t < 1$, dado que $z \in S$ y que $S \in \mathscr{P}$, se tiene que $x \in S$ y $y \in S$. Por lo tanto $\mathfrak{Re}(T(x)) \leq \mu$ y $\mathfrak{Re}(T(y)) \leq \mu$. Puesto que $\mathfrak{Re}(T(z)) = \mu$ y que T es una forma lineal, se obtiene que $\mathfrak{Re}(T(x)) = \mathfrak{Re}(T(y)) = \mu$ y entonces se tiene que $x, y \in S_T$.

Fijemos ahora un conjunto $S \in \mathscr{P}$ y sea $\mathscr{P}'$ la colección de todos los elementos de $\mathscr{P}$ que son subconjuntos de S. Vemos que este conjunto no es vacío pues $S \in \mathscr{P}'$ y podemos ordenar parcialmente $\mathscr{P}'$ utilizando la inclusión de conjuntos. Sea ahora Ω una subcolección maximal, totalmente ordenada, de $\mathscr{P}'$ y sea M la intersección de todos los miembros de Ω. Como Ω es una colección de conjuntos compactos, se tiene que $M \neq \emptyset$ y además se tiene que $M \in \mathscr{P}'$. La maximalidad de Ω implica que no existe un subconjunto propio de M que pertenece a $\mathscr{P}$. Podemos ver entonces por el punto *ii)* que cada forma lineal continua $T \in E'$ es constante sobre M, como E' separa los puntos de E se obtiene que M está reducido a un solo punto y por lo tanto M es un punto extremal de K.

Hemos demostrado que
$$\mathcal{E}(K) \cap S \neq \emptyset \tag{1}$$
para todo $S \in \mathscr{P}$. Como K es compacto y convexo se tiene que $\overline{co}(\mathcal{E}(K)) \subset K$ y esto muestra que $\overline{co}(\mathcal{E}(K))$ es compacto.

Falta verificar la inclusión recíproca. Para ello procedemos por el absurdo. Supongamos pues que existe $x_0 \in K$ tal que $x_0 \notin \overline{co}(\mathcal{E}(K))$. La proposición 2 nos dice entonces que existe una forma lineal continua $T \in E'$ tal que $\mathfrak{Re}(T(x)) < \mathfrak{Re}(T(x_0))$ para cada $x \in \overline{co}(\mathcal{E}(K))$. Si K_T es definido como S_T, entonces se tiene que $K_T \in \mathscr{P}$. Pero por la forma de escoger T se tiene que K_T es disjunto de $\overline{co}(\mathcal{E}(K))$ y esto contradice (1). ■

En el caso en que se trabaje sobre espacios localmente convexos separados se obtiene:

Corolario 1 *Si K es un subconjunto compacto de un espacio localmente convexo separado, entonces $K \subset \overline{co}(\mathcal{E}(K))$ y se tiene además $\overline{co}(K) = \overline{co}(\mathcal{E}(K))$.*

Pasamos ahora a estudiar lo que sucede con los puntos extremales del conjunto $\overline{co}(K)$ en relación al conjunto K.

Teorema 2 (Milman) *Si K es un conjunto compacto de un espacio localmente convexo separado E y si $\overline{co}(K)$ es compacto, entonces todo punto extremal de $\overline{co}(K)$ pertenece a K.*

Demostración. Supongamos que algún punto extremal p de $\overline{co}(K)$ no pertenece a K. Entonces existe una vecindad convexa y equilibrada V del origen tal que
$$(p + \overline{V}) \cap K = \emptyset. \tag{2}$$

Fijemos $x_1, ..., x_n$ puntos de K tales que $K \subset \bigcup_{i=1}^{n}(x_i + V)$. Cada uno de los conjuntos $A_i = \overline{co}(K \cap (x_i + V))$ con $1 \leq i \leq n$ es convexo y compacto puesto que $A_i \subset \overline{co}(K)$. Observamos además que $K \subset A_1 \cup A_2 \cup \cdots \cup A_n$. Entonces, por el punto 1) de la proposición 1 tenemos que

$$\overline{co}(K) \subset \overline{co}(A_1 \cup A_2 \cup \cdots \cup A_n) = co(A_1 \cup A_2 \cup \cdots \cup A_n).$$

Pero la inclusión recíproca también se tiene puesto que $A_i \in \overline{co}(K)$ para cada i de manera que

$$\overline{co}(K) = co(A_1 \cup A_2 \cup \cdots \cup A_n). \tag{3}$$

En particular, $p = t_1 y_1 + \cdots + t_n y_n$ en donde cada y_i pertenece a algún A_i con $t_i > 0$ tal que $\sum_i^n t_i = 1$. Escribimos ahora

$$p = t_1 y_1 + (1 - t_1)\frac{t_2 y_2 + \cdots + t_n y_n}{t_2 + \cdots + t_n} \tag{4}$$

para ver que p es una combinación convexa de dos puntos de $\overline{co}(K)$. Se tiene entonces que p es un punto extremal de $\overline{co}(K)$ y se obtiene utilizando (4) que $y_1 = p$, de manera que, para algún i, se tiene

$$p \in A_i \subset x_i + \overline{V} \subset K + \overline{V}$$

lo cual es una contradicción con (2). ■

A partir del teorema 1 podemos deducir un resultado relativo a la eventual existencia de un espacio predual y de esta manera cumplimos nuestro propósito mostrando cómo el estudio de las envolturas convexas y de los puntos extremales permite tratar este problema.

Tenemos pues el siguiente criterio clásico:

Corolario 2 (Criterio de existencia de un espacio predual) *Sea $(E, \|\cdot\|_E)$ un espacio vectorial normado. Sea $\mathcal{E}(\overline{B}_E)$ el conjunto de puntos extremales de la bola unidad cerrada de E.*

1) Si $\overline{B}_E \neq \overline{co}(\mathcal{E}(\overline{B}_E))$, entonces E no es el espacio dual de ningún espacio.

2) Si el conjunto extremal de la bola unidad B_E es vacío, entonces E no es el espacio dual de ningún espacio.

Prueba.

1) Razonamos por una reducción al absurdo: si E es el predual de algún espacio F, por el teorema de Banach-Aloaglu-Bourbaki, se tendría que la bola unidad cerrada de E sería compacta para la toplogía débil-$*$. Entonces, por el teorema 1 de Krein-Milman esta bola sería la adherencia de la envoltura convexa de sus puntos extremales; de donde se obtendría una contradicción con el hecho $\overline{B}_E \neq \overline{co}(\mathcal{E}(\overline{B}_E))$.

2) Inmediato a partir de 1): la bola unidad cerrada de un espacio normado nunca es vacía pues contiene el vector cero. ■