Tema 2
Conjuntos convexos

José R. Berrendero

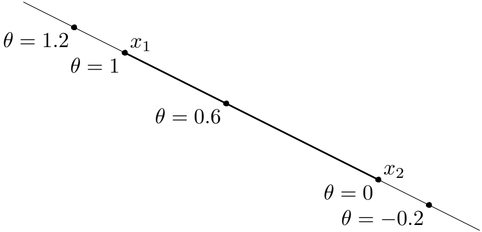
Departamento de Matemáticas
Universidad Autónoma de Madrid

Contenidos del tema 2

- Conjuntos convexos. Propiedades básicas y ejemplos.
- Cierre e interior de un conjunto convexo.
- Teorema de la proyección.
- Teoremas de separación.
- Caracterización de puntos extremos y direcciones extremas de $S = \{x : Ax = b, x \geq 0\}$.
- Teorema de representación.

Conjuntos afines y convexos

$$y = \theta x_1 + (1-\theta)x_2 = x_2 + \theta(x_1 - x_2), \quad x_1, x_2 \in \mathbb{R}^n.$$



- **Conjunto afín**: si $x_1, x_2 \in S$, entonces $\theta x_1 + (1-\theta)x_2 \in S$, para todo $\theta \in \mathbb{R}$.
- **Conjunto convexo**: si $x_1, x_2 \in S$, entonces $\theta x_1 + (1-\theta)x_2 \in S$, para todo $\theta \in [0,1]$.

Ejemplos de conjuntos convexos

- **Hiperplanos**: $S = \{x : \ p^\top x = \alpha\}$, donde $p \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$.
- **Semiespacios**: $S = \{x : \ p^\top x \leq \alpha\}$, donde $p \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$.
- **Intersección arbitraria** de convexos: Si S_i es convexo para todo $i \in I$, entonces $S = \bigcap_{i=1}^{I} S_i$ es un conjunto convexo.
- Un **poliedro** (intersección finita de semiespacios) es un conjunto convexo. Por ejemplo, $S = \{x : \ Ax \leq b, \ x \geq 0\}$ es un conjunto convexo.
- Una **bola** $B(\bar{x}, r) = \{x \in \mathbb{R}^n : \ \|x - \bar{x}\| < r\}$ es un conjunto convexo (para cualquier norma).

Combinaciones afines

Combinación afín de $x_1, \ldots, x_k \in \mathbb{R}^n$:

$$y = \lambda_1 x_1 + \cdots + \lambda_k x_k,$$

donde $\lambda_1 + \cdots + \lambda_k = 1$.

Cierre afín de S:

$$\operatorname{afin}(S) = \left\{ \sum_{i=1}^{k} \lambda_i x_i : \, x_i \in S, \, \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

- Un conjunto es afín si y solo $S = \operatorname{afin}(S)$.
- Un conjunto es afín si y solo si es la traslación de un subespacio vectorial (único)
- La dimensión afín de un conjunto es la dimensión de su cierre afín (que a su vez es la dimensión del correspondiente subespacio vectorial).

Combinaciones convexas y cierre convexo

Combinación convexa de $x_1, \ldots, x_k \in \mathbb{R}^n$:

$$y = \lambda_1 x_1 + \cdots + \lambda_k x_k,$$

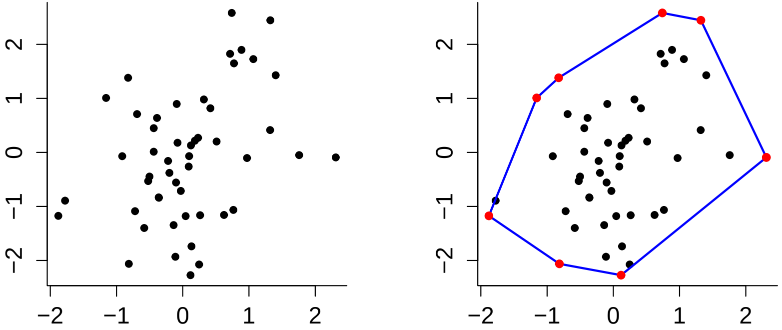
donde $\lambda_1 + \cdots + \lambda_k = 1$ y $\lambda_i \geq 0$, para todo $i = 1, \ldots, n$.

Cierre convexo de S:

$$\operatorname{conv}(S) = \left\{ \sum_{i=1}^{k} \lambda_i x_i : \; x_i \in S, \; \lambda_i \geq 0, \; \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

- Un conjunto S es convexo si y solo si $S = \operatorname{conv}(S)$.
- $\operatorname{conv}(S)$ es el menor conjunto convexo que contiene a S.

Cierre convexo



Teorema de Carathéodory

Cualquier punto del cierre convexo de un conjunto de $\mathbb{R}^n$ se puede poner como combinación lineal convexa de a lo más $n+1$ puntos del conjunto.

Teorema: Sea $S \subset \mathbb{R}^n$. Si $x \in \text{conv}(S)$, entonces $x = \sum_{i=1}^{n+1} \lambda_i x_i$, con $\sum_{i=1}^{n+1} \lambda_i = 1$, $\lambda_i \geq 0$, $x_i \in S$, para todo $i = 1, \ldots, n+1$.

Demostración: Sea $x \in \text{conv}(S)$, entonces $x = \sum_{i=1}^{k} \lambda_i x_i$, con $\sum_{i=1}^{k} \lambda_i = 1$, $\lambda_i > 0$, $x_i \in S$.

1. Si $k \leq n+1$, hemos terminado.
2. Si $k > n+1$, $x_2 - x_1, \ldots, x_k - x_1$ son linealmente dependientes.
3. Existen μ_i (alguno estrictamente positivo) tales que $\sum_{i=1}^{k} \mu_i x_i = 0$.
4. Tenemos que $x = \sum_{i=1}^{k} (\lambda_i - \alpha \mu_i) x_i$, donde

$$\alpha = \min\{\lambda_j / \mu_j : \mu_j > 0\} := \lambda_r / \mu_r > 0.$$

5. Hemos conseguido expresar x como combinación convexa de $k-1$ elementos. Volvemos al paso 1. En un número finito de iteraciones habremos terminado.

Interior relativo

Sea $S \subset \mathbb{R}^n$ un convexo cuyo cierre afín es $\mathrm{afin}(S)$. El **interior relativo** de S se define como el conjunto de puntos $x \in S$ tales que existe $r > 0$ con $B(x, r) \cap \mathrm{afin}(S) \subset S$.

- ¿Cuál es el interior relativo de $\{x \in \mathbb{R}^3 : -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1, x_3 = 0\}$?

Teorema: Un conjunto convexo no vacío en $\mathbb{R}^n$ tiene interior relativo no vacío.

Teorema: El interior de un conjunto convexo en $\mathbb{R}^n$ es vacío si y solo si el conjunto está contenido en un hiperplano de $\mathbb{R}^n$.

Cierre e interior de un conjunto convexo

Lema: Sea $S \subset \mathbb{R}^n$ un conjunto convexo con $\operatorname{int}(S) \neq \emptyset$. Sea $x_1 \in \bar{S}$, $x_2 \in \operatorname{int}(S)$. Entonces, $\theta x_1 + (1-\theta)x_2 \in \operatorname{int}(S)$, para todo $\theta \in [0,1)$.

Demostración: Sea $y = \theta x_1 + (1-\theta)x_2$.

1. Existe $\epsilon > 0$ tal que $B(x_2, \epsilon) \subset S$.
2. Sea $\tilde{y} \in B(y, \eta)$, con $\eta = \epsilon(1-\theta)$.
3. Como $x_1 \in \bar{S}$, existe $\tilde{x}_1 \in S$ tal que
$$\|x_1 - \tilde{x}_1\| < \frac{\eta - \|\tilde{y} - y\|}{\theta} \Leftrightarrow \|y - \tilde{y}\| + \theta\|x_1 - \tilde{x}_1\| < \eta.$$
4. Sea $\tilde{x}_2 = (\tilde{y} - \theta\tilde{x}_1)/(1-\theta) \Leftrightarrow \tilde{y} = \theta\tilde{x}_1 + (1-\theta)\tilde{x}_2$.
5. Se verifica
$$\|x_2 - \tilde{x}_2\| = \frac{\|y - \theta x_1 - \tilde{y} + \theta\tilde{x}_1\|}{1-\theta} \leq \frac{1}{1-\theta}(\|y - \tilde{y}\| + \theta\|x_1 - \tilde{x}_1\|) < \frac{\eta}{1-\theta} = \epsilon.$$
6. Por 1 y 5, $\tilde{x}_2 \in S$. Por 4, $\tilde{y} \in S$. Como $B(y, \eta) \subset S$, $y \in \operatorname{int}(S)$.

Cierre e interior de un conjunto convexo

1. Si $S \subset \mathbb{R}^n$ es convexo, entonces tanto $\text{int}(S)$ como $\bar{S}$ son conjuntos convexos.

2. Si $S \subset \mathbb{R}^n$ es convexo, entonces $\text{int}(S) = \text{int}(\bar{S})$. Si además $\text{int}(S) \neq \emptyset$, entonces $\bar{S} = \overline{\text{int}(S)}$.

3. Si $S \subset \mathbb{R}^n$ es convexo, entonces $\partial S = \partial \bar{S}$.

Teorema de la proyección

Teorema: Sea $S \subset \mathbb{R}^n$ un conjunto convexo, no vacío y cerrado. Sea $y \in \mathbb{R}^n$. Existe un **único** $\bar{x} \in S$ (la proyección de y sobre S) tal que

$$\|y - \bar{x}\| \leq \|y - x\|, \quad \text{para todo } x \in S.$$

Además, $\bar{x}$ es la proyección de y sobre S si y solo si

$$(y - \bar{x})^\top (x - \bar{x}) \leq 0. \tag{1}$$

- ¿Cómo queda la condición (**??**) cuando S es un conjunto afín?
- La aplicación $P : \mathbb{R}^n \to S$, que a cada y le hace corresponder su proyección $P(y)$ sobre S, es continua.

Teorema del hiperplano separador

Teorema: Sea $S \subset \mathbb{R}^n$ un conjunto convexo, no vacío y cerrado. Sea $y \notin S$. Entonces existe $p \in \mathbb{R}^n$, $p \neq 0$, y $\alpha \in \mathbb{R}$ tales que

- $p^\top x \leq \alpha$, para todo $x \in S$.
- $p^\top y > \alpha$

Demostración:

1. Sea $\bar{x}$ la proyección de y sobre S.
2. Se verifica $(y - \bar{x})^\top (x - \bar{x}) \leq 0$, para todo $x \in S$.
3. Definimos $p = y - \bar{x}$ y $\alpha = p^\top \bar{x}$. Comprobamos que cumplen las condiciones requeridas:
    - Si $x \in S$, $p^\top x = (y - \bar{x})^\top (x - \bar{x}) + \alpha \leq \alpha$.
    - $p^\top y = (y - \bar{x})^\top (y - \bar{x}) + \alpha = \|y - \bar{x}\|^2 + \alpha > \alpha$.

Teorema del hiperplano soporte

Teorema: Sea $S \subset \mathbb{R}^n$ un conjunto convexo ~~con interior no vacío~~. Sea $\bar{x} \in \partial S$, la frontera de S. Entonces existe $p \in \mathbb{R}^n$, $p \neq 0$, tal que $p^\top (x - \bar{x}) \leq 0$, para todo $x \in S$.

Demostración:

1. $\bar{x} \notin \operatorname{int}(S) = \operatorname{int}(\bar{S})$.
2. Para todo $k \in \mathbb{N}$, existe $y_k \in B(\bar{x}, 1/k) \cap \bar{S}^c$
3. Existe $p_k \in \mathbb{R}^n$, $\|p_k\| = 1$, con $p_k^\top y_k > p_k^\top x$ para todo $x \in \bar{S}$.
4. Tenemos $y_k \to \bar{x}$. Para una subsucesión, $p_k \to p$ con $\|p\| = 1$.
5. Tomando límites en 3, $p^\top (x - \bar{x}) \leq 0$, para todo $x \in \bar{S}$.

Separación de dos convexos disjuntos

Teorema: Sean $S_1, S_2 \subset \mathbb{R}^n$ conjuntos convexos, no vacíos, tales que $S_1 \cap S_2 = \emptyset$. Entonces existe $p \in \mathbb{R}^n$, $p \neq 0$, tal que

$$\inf\{p^\top x : x \in S_1\} \geq \sup\{p^\top x : x \in S_2\}.$$

Demostración: La idea es considerar

$$S = \{x \in \mathbb{R}^n : x = x_1 - x_2, \; x_1 \in S_1, \; x_2 \in S_2\}$$

y aplicar los teoremas de separación.

Teoremas de la alternativa

Teorema (Gordan): Sea A una matriz $m \times n$. Uno y solo uno de los sistemas siguientes tiene solución:

- $Ax < 0$ para algún $x \in \mathbb{R}^n$.
- $A^\top p = 0$, $p \geq 0$ y $p \neq 0$ para algún $p \in \mathbb{R}^m$.

Demostración:

- (1) tiene solución $\Rightarrow$ (2) no la tiene.

  Sea $\bar{x}$ la solución de (1) y supongamos que $\bar{p}$ fuese una solución de (2). Consideramos $\bar{p}^\top A\bar{x}$.

- (1) no tiene solución $\Rightarrow$ (2) sí la tiene.

  Los conjuntos $S_1 = \{z \in \mathbb{R}^m : \ z = Ax, \ x \in \mathbb{R}^n\}$ y $S_2 = \{z \in \mathbb{R}^m : \ z < 0\}$ son convexos, no vacíos y disjuntos. Existe $p \neq 0$ con $p^\top Ax \geq p^\top z$ para $x \in \mathbb{R}^n$ y $z < 0$. Se demuestra que p resuelve (2).

Teoremas de la alternativa

Ejemplo: ¿Existen x_1, x_2 y x_3 tales que $x_1 + x_2 + x_3 < 0$, $x_1 > 0$ y $2x_1 - x_2 - x_3 < 0$?

Teorema (Farkas): Sea A una matriz $m \times n$ y $c \in \mathbb{R}^n$. Uno y solo unos de los sistemas siguientes tiene solución:

1. $Ax \leq 0$ y $c^\top x > 0$ para algún $x \in \mathbb{R}^n$.
2. $A^\top y = c$, $y \geq 0$ para algún $y \in \mathbb{R}^m$.

Demostración: Ejercicio

Probar que (1) no tiene solución cuando (2) sí la tiene es muy fácil. Cuando (2) no tiene solución considera $S = \{x \in \mathbb{R}^n : \ x = A^\top y, \ y \geq 0\}$ y observa que $c \notin S$.

Teorema de representación de poliedros

Los resultados hasta el final del tema se refieren al poliedro $S = \{x : Ax = b, x \geq 0\}$, el conjunto factible de un problema de optimización lineal (en forma estándar).

Suponemos que A es una matriz $m \times n$ $(m < n)$ con rango $r(A) = m$.

Vamos a representar los puntos de S en términos de los puntos extremos (vértices) de S y sus direcciones extremas.

Cualquier punto de S se puede expresar como una combinación convexa de sus puntos extremos más una combinación lineal positiva de sus direcciones extremas.

Soluciones factibles básicas

Podemos dividir las columnas de A es dos grupos B y N, donde B es $m \times m$ con $r(B) = m$.

$$Ax = b \Leftrightarrow (B, N)\begin{pmatrix} x_B \\ x_N \end{pmatrix} = b \Leftrightarrow Bx_B + Nx_n = b$$

Si hacemos $x_N = 0$, $x_B = B^{-1}b$ y se verifica $B^{-1}b \geq 0$, obtenemos unos puntos especiales de S que se llaman **soluciones factibles básicas**.

Algunas soluciones del sistema compatible indeterminado $Ax = b$, con $r(A) = r(A, b) = m < n$ se consiguen fijando $n - m$ incógnitas como 0 y despejando las m incógnitas restantes. Si estas soluciones son no negativas, son soluciones factibles básicas.

Puntos extremos

Definición: Sea $S \subset \mathbb{R}^n$ un conjunto convexo no vacío. Se dice que $x \in S$ es un **punto extremo** de S si $x = \lambda x_1 + (1 - \lambda)x_2$, con $x_1, x_2 \in S$, $\lambda \in (0, 1)$, implica que $x = x_1 = x_2$.

Es decir, x es un punto extremo si no está en el interior (relativo) del segmento definido por otros dos puntos del conjunto.

Piensa ejemplos de conjuntos convexos: con un único punto extremo, con un número finito mayor que uno de puntos extremos, con infinitos puntos extremos, sin puntos extremos.

Direcciones extremas

Definición: Sea $S \subset \mathbb{R}^n$ un conjunto convexo no vacío. Se dice que $d \in \mathbb{R}^n$, $d \neq 0$, es una **dirección** de S si para todo $x \in S$ y para todo $\lambda \geq 0$, se cumple $x + \lambda d \in S$.

¿Qué condiciones caracterizan las direcciones de $S = \{x : Ax = b, x \geq 0\}$?

Definición: Sea $S \subset \mathbb{R}^n$ un conjunto convexo no vacío. Se dice que $d \in \mathbb{R}^n$, $d \neq 0$, es una **dirección extrema** de S si $d = \lambda_1 d_1 + \lambda_2 d_2$, con d_1, d_2 direcciones de S, $\lambda_1, \lambda_2 > 0$, implica que $d_1 = \alpha d_2$, para algún $\alpha > 0$.

Piensa ejemplos de conjuntos convexos: con dos direcciones extremas, con una única dirección extrema, sin direcciones extremas.

Caracterización de puntos extremos

Teorema: x es un punto extremo de S si y solo si x es una solución factible básica.

Demostración: ($\Leftarrow$) Ejercicio. Demostramos ($\Rightarrow$).

1. Reordenando las coordenadas $x = (x_1, \ldots, x_k, 0, \ldots, 0)^\top$, con $x_i > 0$. Spg. $k \geq 1$ (el caso $x = 0$ es trivial).

2. Reordenamos igual las columnas de A: $A = (a_1, \ldots, a_k, \text{resto de columnas})$.

3. Veamos que $a_1, \ldots, a_k$ son linealmente independientes. Si no fuera así existiría un vector $\lambda \neq 0$ tal que $A\lambda = 0$.

4. Para $\alpha > 0$ suficientemente pequeño, $\bar{x}_1 = x + \alpha\lambda$ y $\bar{x}_2 = x - \alpha\lambda$ pertenecen a $S \Rightarrow x = (\bar{x}_1 + \bar{x}_2)/2$ no es punto extremo.

5. Si $k < m$, añadimos columnas de manera que las columnas de $B = (a_1, \ldots, a_k, a_{k+1}, \ldots, a_m)$ formen una base.

Caracterización de puntos extremos

La demostración implica que un punto extremo no puede tener más de m coordenadas estrictamente positivas. El recíproco no es cierto (veremos algún ejemplo más adelante en los ejercicios).

Si una solución factible básica tiene $k < m$ coordenadas estrictamente positivas se llama **solución factible básica degenerada**.

En el caso degenerado puede haber dos bases distintas B y B' que representen el mismo punto extremo (¿por qué?).

Caracterización de puntos extremos

Teorema: El número de puntos extremos de S es finito
(¿Por qué?)

Teorema: S tiene al menos un punto extremo.
(La demostración es muy parecida a la del teorema de Carathéodory.)

Ejemplo: Obtener las soluciones factibles básicas del problema:

$$\begin{array}{ll} \text{minimizar} & -x_1 + 2x_2 \\ & \\ \text{s.a.} & x_1 + 2x_2 \leq 2 \\ & 2x_1 + x_2 \leq 1 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array}$$

Caracterización de direcciones extremas

Es fácil demostrar que d es una dirección de S si y solo si $Ad = 0$ y $d \geq 0$. El resultado siguiente caracteriza las direcciones extremas:

Teorema: $d \in \mathbb{R}^n$ es una dirección extrema de S si y solo si $A = (B, N)$, donde B es una matriz $m \times m$ básica, de manera que $B^{-1}a_j \leq 0$ para alguna de las columnas de N y $d = \alpha\binom{-B^{-1}a_j}{e_j}$ para algún $\alpha > 0$, donde $e_j \in \mathbb{R}^{n-m}$ es un vector de ceros salvo un 1 en la posición correspondiente a a_j.

Demostración:

($\Leftarrow$) Ejercicio.

($\Rightarrow$) No la hacemos. Es totalmente análoga a la del teorema de caracterización de puntos extremos.

Ejemplo: Calcula las direcciones extremas del conjunto

$$S = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 - x_1 \leq 1, x_2 \leq 2, x_1 \geq 0, x_2 \geq 0\}.$$

Teorema de representación

Teorema: Sean $x_1, \ldots, x_k$ los puntos extremos de S y sean $d_1, \ldots, d_\ell$ sus direcciones extremas. Entonces,

$$x \in S \Leftrightarrow x = \sum_{i=1}^{k} \lambda_i x_i + \sum_{j=1}^{\ell} \mu_j d_j,$$

donde $\lambda_i \geq 0$, $\mu_j \geq 0$, $\sum_{i=1}^{k} \lambda_i = 1$.

Consecuencias:

- Si S es acotado, cualquier punto de S es combinación lineal convexa de sus puntos extremos.
- S tiene al menos una dirección extrema si y solo si S no es acotado.

¿Es el número de direcciones extremas siempre un número finito?

Demostración del teorema de representación

$$C = \{x \in \mathbb{R}^n : x = \sum_{i=1}^{k} \lambda_i x_i + \sum_{j=1}^{\ell} \mu_j d_j,\ \lambda_i \geq 0,\ \mu_j \geq 0,\ \sum_{i=1}^{k} \lambda_i = 1\}$$

- $C \neq \emptyset$
- C es convexo y cerrado.
- $C \subset S$

- Para probar $S \subset C$, supongamos que existe $z \in S$ con $z \notin C$.

- Existe $p \in \mathbb{R}^n$ ($p \neq 0$) y $\alpha \in \mathbb{R}$ tales que $p^\top z > \alpha$ y $p^\top x \leq \alpha$ para todo $x \in C$.
- Como consecuencia, $p^\top d_j \leq 0$ y $p^\top x_i \leq \alpha \Rightarrow p^\top z > p^\top x_i,\ i = 1, \ldots, k$.
- Sea $\bar{x}$ el punto extremo tal que $p^\top \bar{x} = \max_i p^\top x_i$.
- Consideramos la partición $A = (B, N)$ en parte básica y no básica correspondiente a $\bar{x}$.
- Se comprueba $0 < p^\top z - p^\top \bar{x} = (p_N^\top - p_B^\top B^{-1} N) z_N$, lo que implica que para algún $j > m$ se tiene $z_j > 0$ y $p_j - p_B^\top B^{-1} a_j > 0$.
- Se demuestra que $y_j := B^{-1} a_j \not\leq 0$. (Por reducción al absurdo ya que si $y_j \leq 0$, entonces $d_j = \binom{-y_j}{e_j}$ es una dirección extrema y por tanto $p_j - p_B^\top B^{-1} a_j \leq 0$.)

Demostración del teorema de representación

- Definimos el vector

$$x = \begin{pmatrix} \bar{b} \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -y_j \\ e_j \end{pmatrix},$$

donde $\bar{b} = B^{-1}b$ y $\alpha = \min \left\{ \frac{\bar{b}_i}{y_{ij}} \, : \, y_{ij} > 0 \right\} := \frac{\bar{b}_r}{y_{rj}} > 0.$

- Comprobamos $x \in S$, $x_r = 0$, $x_j > 0$.
- Las columnas $a_1, a_2, \ldots, a_{r-1}, a_{r+1}, a_j$ son linealmente independientes, es decir, x es un punto extremo de S.

- Pero $p^\top x = p^\top \bar{x} + \alpha(p_j - p_B^\top B^{-1} a_j) > p^\top \bar{x}$.
- Una contradicción ya que habíamos supuesto $p^\top \bar{x} = \max_i p^\top x_i$.