*Ejercicios - Conjuntos Convexos*

*Universidad Nacional de Colombia, Sede Bogotá*
Introducción a la Optimización
2022-I

Sebastian Leonardo Molina Diaz
smolinad@unal.edu.co

# 2 Convex sets

## Exercises

**2.1** Let $C \subseteq \mathbb{R}^n$ be a convex set, with $x_1, \ldots, x_k \in C$, and let $\theta_1, \ldots, \theta_k \in \mathbb{R}$ satisfy $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$. Show that $\theta_1 x_1 + \cdots + \theta_k x_k \in C$. (The definition of convexity is that this holds for $k = 2$; you must show it for arbitrary $k$). *Hint*. Use induction on $k$.

*Solución.*

Sea $C$ un conjunto convexo. Por definición de convexidad, se tiene el caso base con $k = 2$. Suponga que para $k$ arbitrario se cumple que $\sum_{i=1}^{k} \theta_i x_i \in C$, con $\sum_{i=1}^{k} \theta_i = 1$, $\theta_i \geq 0$ y $x_i \in C$ para $1 \leq i \leq k$. Sean $\mu_j \geq 0$ y $x_j \in C$ con $1 \leq j \leq k+1$. Para el paso inductivo, observe que si $\mu_{k+1} = 0$, obtenemos el mismo caso de la hipótesis de inducción. Si $\mu_{k+1} = 1$, estamos considerando el caso trivial en el que $\mu_{k+1} x_{k+1} = 1 \cdot x_{k+1} = x_{k+1} \in C$. Consideremos el caso en el que $\mu_{k+1} \neq 0$ y $\mu_{k+1} \neq 1$. Observe que

$$
\begin{aligned}
\mathbf{y} &= \sum_{j=1}^{k+1} \mu_j x_j \\
&= \left( \frac{1-\mu_{k+1}}{1-\mu_{k+1}} \right) (\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_k x_k) + \mu_{k+1} x_{k+1} \\
&= (1-\mu_{k+1}) \cdot \left( \frac{\mu_1}{1-\mu_{k+1}} x_1 + \frac{\mu_2}{1-\mu_{k+1}} x_2 + \cdots + \frac{\mu_k}{1-\mu_{k+1}} x_k \right) + \mu_{k+1} x_{k+1} \\
&= (1-\mu_{k+1}) \cdot (\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k) + \mu_{k+1} x_{k+1},
\end{aligned}
$$

donde $\lambda_l = \frac{\mu_l}{1-\mu_{k+1}}$ con $1 \leq l \leq k$. De lo anterior, podemos notar que

$$
\sum_{l=1}^{k} \lambda_l = \frac{\sum_{l=1}^{k} \mu_l}{1-\mu_{k+1}} = \frac{1-\mu_{k+1}}{1-\mu_{k+1}} = 1.
$$

Si denotamos $\mathbf{x} = (\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k)$, por hipótesis de inducción $\mathbf{x} \in C$, ya que $\sum_{l=1}^{k} \lambda_l = 1$ y $\lambda_l \geq 0$ para $1 \leq l \leq k$. Por lo tanto, como $C$ es convexo, $\mathbf{y} = (1-\mu_{k+1}) \cdot \mathbf{x} + \mu_{k+1} x_{k+1} \in C$.

□

**2.3** *Midpoint convexity.* A set $C$ is midpoint convex if whenever two points $a, b$ are in $C$, the average or midpoint $(a+b)/2$ is in $C$. Obviously a convex set is midpoint convex. It can be proved that under mild conditions midpoint convexity implies convexity. As a simple case, prove that if $C$ is closed and midpoint convex, then $C$ is convex. Considere el siguiente número binario, el cual codifica recursivamente

*Solución.*

Sean $a, b \in C$. Queremos ver que la combinación convexa $\theta + (1-\theta)b \in C$. Considere el coeficiente $\theta_k$, resultado de aplicar la convexidad de punto medio $k$ veces a los puntos $a$ y $b$, recursivamente. Entonces, podemos expresar a $\theta_k$ como

$$
\theta_k = \frac{\alpha_1}{2^1} + \frac{\alpha_2}{2^2} + \cdots + \frac{\alpha_k}{2^k},
$$

donde

$$\alpha_i = \begin{cases} 0 & |\theta| < |1-\theta| \\ 1 & |1-\theta| < |\theta| \end{cases}.$$

Observe que $\lim_{k\to\infty} \theta_k a + (1-\theta_k)b = \theta a + (1-\theta)b \in C$, ya que como $C$ es cerrado, entonces contiene todos sus puntos límite. Por lo tanto, $C$ es convexo.

□

**2.4** Show that the convex hull of a set $S$ is the intersection of all convex sets that contain $S$. (The same method can be used to show that the conic, or affine, or linear hull of a set $S$ is the intersection of all conic sets, or affine sets, or subspaces that contain $S$).

*Solución.*

Usando doble contenencia, probemos que $\textbf{conv}\, S = I$, donde $I = \bigcap\{C : S \subseteq C, C \text{ conjunto convexo}\}$. Sea $\mathbf{x} \in \textbf{conv}\, S$, luego $\mathbf{x} = \sum_{i=1}^n \theta_i x_i$, donde $\sum_{i=1}^n \theta_i = 1$, $x_i \in S$ y $x_i \in S$ para todo $1 \le i \le n$. Sea $D$ un conjunto convexo arbitrario tal que $S \subseteq D$. Como $D$ es convexo, cualquier combinación convexa de sus elementos está dentro de $D$, en particular $\mathbf{x}$. Como $D$ es un conjunto arbitrario y $\mathbf{x} \in I$ es un elemento arbitrario, entonces $\textbf{conv}\, S \subseteq I$. Ahora, note que $S \subseteq \textbf{conv}\, S$ y $\textbf{conv}\, S$ es convexo por definición. Así, $\textbf{conv}\, S = C$ para algún $C$ en la construcción de $I$. Luego $I \subseteq \textbf{conv}\, S$, y por lo tanto, $I = \textbf{conv}\, S$.

□

**2.5** What is the distance between two parallel hyperplanes $\{x \in \mathbb{R}^n : a^T x = b_1\}$ and $\{x \in \mathbb{R}^n : a^T x = b_2\}$?

*Solución.*

Sean $\mathcal{H}_1 = \{x \in \mathbb{R}^n : a^T x = b_1\}$ y $\mathcal{H}_2 = \{x \in \mathbb{R}^n : a^T x = b_2\}$ dos hiperplanos paralelos. La distancia entre $\mathcal{H}_1$ y $\mathcal{H}_2$ es la distancia entre dos puntos $x_1 \in \mathcal{H}_1$ y $x_2 \in \mathcal{H}_2$, los cuales están sobre la recta normal a los hiperplanos que pasa por el origen, como se puede ver en la Figura 1.

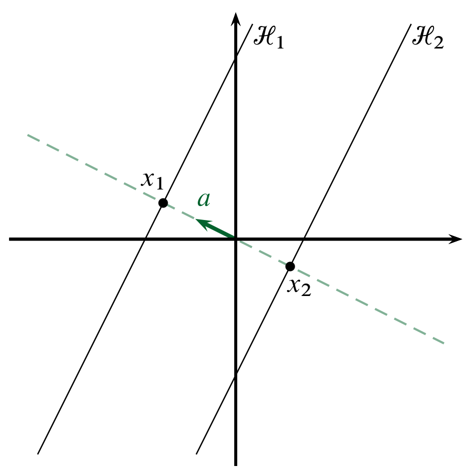


Figura 1: Distancia entre hiperplanos paralelos.

Por lo tanto, si $a$ es el vector normal a los hiperplanos, $x_1$ y $x_2$ son multiplicaciones escalares de $a$. Luego, $x_1 = c_1 \cdot a$, con $c_1 \in \mathbb{R}$. Como $x_1 \in \mathcal{H}_1$, entonces $a^T x_1 = a^T(c_1 \cdot a) = b_1$. Despejando, obtenemos que $c_1 = b_1/a^T a = b_1/\|a\|^2$, por lo que $x_1 = (b_1/\|a\|^2)a$. Similarmente, $x_2 = (b_2/\|a\|^2)a$. Así, la distancia entre $\mathcal{H}_1$ y $\mathcal{H}_2$ es

$$\begin{aligned} \|x_1 - x_2\| &= \left\| \frac{b_1}{\|a\|^2} a - \frac{b_2}{\|a\|^2} a \right\| = \left\| (b_1 - b_2) \frac{a}{\|a\|^2} \right\| \\ &= \frac{1}{\|a\|^2} |b_1 - b_2| \|a\| \\ &= \frac{|b_1 - b_2|}{\|a\|}. \end{aligned}$$

□

**2.8** Which of the following sets $S$ are polyhedra? If possible, express $S$ in the form $S = \{x : Ax \preceq b, Fx = g\}$.

*a)* $S = \{y_1 a_1 + y_2 a_2 : -1 \leq y_1 \leq 1, -1 \leq y_2 \leq 1\}$, where $a_1, a_2 \in \mathbb{R}^n$.

*Solución.*
Observe que en los casos extremos en los que $y_1 = \pm 1$ y $y_2 = \pm 1$, obtenemos cuatro puntos: $a_1 + a_2$, $a_1 - a_2$, $-a_1 + a_2$ y $-a_1 - a_2$. Como $-1 \leq y_1, y_2 \leq 1$, $S$ define al paralelogramo con vértices en los cuatro puntos mencionados anteriormente, por lo que $S$ es un poliedro.

□

*b)* $S = \left\{x \in \mathbb{R}^n : x \succeq 0,\ \mathbf{1}^T x = 1,\ \sum_{i=1}^n x_i a_i = b_1,\ \sum_{i=1}^n x_i a_i^2 = b_2\right\}$, where $a_1, \ldots, a_n \in \mathbb{R}$ and $b_1, b_2 \in \mathbb{R}$.

*Solución.*
Sea $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Claramente $S$ es un poliedro ya que es la intersección de las igualdades $\mathbf{1}^T x = 1$, $\sum_{i=1}^n x_i a_i = b_1, \sum_{i=1}^n x_i a_i^2 = b_2$ y las desigualdades $x_j \geq 0$, donde $x_j$ con $1 \leq j \leq n$ son las entradas de $x$.

□

*c)* $S = \left\{x \in \mathbb{R}^n : x \succeq 0,\ x^T y \leq 1 \text{ for all } y \text{ with } \|y\|_2 = 1\right\}$.

*Solución.*
Si bien $S$ es la intersección de semiespacios, $S$ no es un poliedro ya que es la intersección de infinitos semiespacios definidos por los $x \succeq 0$ tales que $\|x\|_2 \leq 1$, lo cual define la intersección de la bola unitaria en $\mathbb{R}^n$ con $\mathbb{R}^n_+$.

□

*d)* $S = \left\{x \in \mathbb{R}^n : x \succeq 0,\ x^T y \leq 1 \text{ for all } y \text{ with } \sum_{i=1}^n |y_i| = 1\right\}$.

*Solución.*
Por una idea similar a la del literal anterior, $S$ es la intersección de $\mathbb{R}^n_+$ con el conjunto $\{x : |x_k| \leq 1, k \in \{1, \ldots, n\}\}$. Observe que si $x \succ 0$ tal que $x^T y \leq 1$ para $y$ con $\sum_{i=1}^n |y_i| = 1$, entonces podemos elegir a $y$ tal que si $|x_k| = \max\{x_i\}$ con $1 \leq k \leq n$, entonces

$$y_k = \begin{cases} 1 & x_k > 0, \\ -1 & x_k < 0, \\ 0 & \text{en otro caso.} \end{cases}$$

Con lo anterior obtenemos que

$$x^T y = \sum_{i=0}^{n} x_i y_i = y_k x_k = |x_k| = max\{x_i\} \leq 1.$$

Por lo tanto, $S$ es la intersección de $\mathbb{R}^n_+$ con el conjunto finito $\{x : -1 \preceq x \preceq 1\}$. Así, $S$ es un poliedro. □

**2.9** *Voronoi sets and polyhedral decomposition.* Let $x_0, \ldots, x_K \in \mathbb{R}^n$. Consider the set of points that are closer (in Euclidean norm) to $x_0$ than the other $x_i$, i.e.,

$$V = \{x \in \mathbb{R}^n : \|x - x_0\|_2 \leq \|x - x_i\|_2, i = 1, \ldots, K\}$$

$V$ is called the Voronoi region around $x_0$ with respect to $x_1, \ldots, x_K$.

*a)* Show that $V$ is a polyhedron. Express $V$ in the form $V = \{x : Ax \preceq b\}$.

*Solución.*
Sean $x, x_0, \ldots, x_K \in \mathbb{R}^n$. Observe que $x$ está mas cerca con la norma euclidiana a $x_0$ que $x_i$, donde $1 \leq i \leq K$, si

$$\begin{aligned}
\|x - x_0\|_2 \leq \|x - x_i\|_2 &\therefore \|x - x_0\|_2^2 \leq \|x - x_i\|_2^2 \\
&\therefore (x - x_0)^T (x - x_0) \leq (x - x_i)^T (x - x_i) \\
&\therefore x^T x - x^T x_0 - x_0^T x + x_0^T x_0 \leq x^T x - x^T x_i - x_i^T x + x_i^T x_i \\
&\therefore x^T x - 2x_0^T x - x_0^T x_0 \leq x^T x - 2x_i^T x + x_i^T x_i \\
&\therefore 2x_i^T x - 2x_0^T x \leq x_i^T x_i - x_0^T x_0 \\
&\therefore 2(x_i - x_0)^T x \leq x_i^T x_i - x_0^T x_0,
\end{aligned}$$

lo cual define un poliedro o un sistema de desigualdades lineales, donde $V = \{x : Ax \preceq b\}$ con

$$A = 2\begin{bmatrix} x_1 - x_0 \\ x_1 - x_0 \\ \vdots \\ x_1 - x_K \end{bmatrix}^T, \quad b = \begin{bmatrix} x_1^T x_1 - x_0^T x_0 \\ x_2^T x_2 - x_0^T x_0 \\ \vdots \\ x_K^T x_K - x_0^T x_0 \end{bmatrix}.$$

Intuitivamente, como se puede ver en la Figura 2, $V$ es la intersección de los semiespacios definidos por los hiperplanos ortogonales a $x_i - x_0$, que pasan por el punto medio $m_i = \frac{x_0 + x_i}{2}$, con $1 \leq i \leq K$.

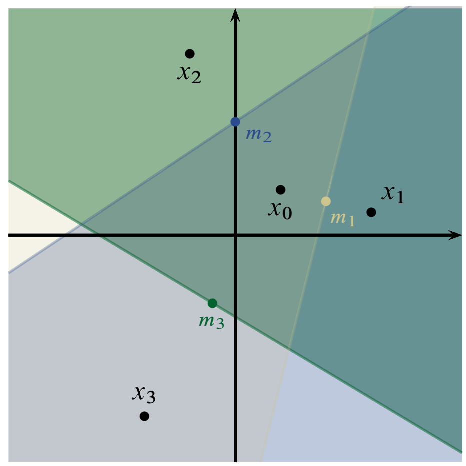


Figura 2: Región de Voronoi alrededor de $x_0$ como intersección de semiespacios.

□

*b)* Conversely, given a polyhedron $P$ with nonempty interior, show how to find $x_0, \ldots, x_K$ so that the polyhedron is the Voronoi region of $x_0$ with respect to $x_1, \ldots, x_K$.

*Solución.*
Sea $V = \{Ax \preceq b\}$ un poliedro con interior no vacío. Intuitivamente, tomando como referencia la Figura 2, para construir la región de Voronoi tome cualquier punto dentro del poliedro como $x_0$. Luego, tome el punto $x_i$, simétrico respecto a los hiperplanos definidos por $v_i^T x = b_i$. Para ello, debemos tomar a $x_i = x_0 + \delta a_i$ tal que su distancia al hiperplano es igual a la distancia de $x_0$ al hiperplano. Es decir

$$a_i^T x_i - b_i = b_i - a_i^T x_0.$$

Reemplazando $x_i$ en la igualdad anterior obtenemos que

$$\begin{aligned} a_i^T(x_0 + \delta a_i) - b_i = b_i - a_i^T x_0 &\therefore 0 = 2b_i - a_i^T x_0 - a_i^T(x_0 + \delta a_i) \\ &\therefore 0 = 2b_i - 2a_i^T x_0 - \delta a_i^T a_i \\ &\therefore 2b_i - 2a_i^T x_0 = \delta a_i^T a_i \\ &\therefore \frac{2(b_i - a_i^T x_0)}{a_i^T a_i} = \frac{2(b_i - a_i^T x_0)}{\|a\|^2} = \delta. \end{aligned}$$

Así, $x_i = x_0 + \frac{2(b_i - a_i^T x_0)}{\|a\|^2} a_i$. □

*c)* We can also consider the sets

$$V_k = \{x \in \mathbb{R}^n : \|x - x_k\|_2 \leq \|x - x_i\|_2, i \neq k\}$$

The set $V_k$ consists of points in $\mathbb{R}^n$ for which the closest point in the set $\{x_0, \ldots, x_K\}$ is $x_k$. The sets $V_0, \ldots, V_K$ give a polyhedral decomposition of $\mathbb{R}^n$. More precisely, the sets $V_k$ are polyhedra, $\bigcup_{k=0}^K V_k = \mathbb{R}^n$, and int $V_i \cap$ int $V_j = \emptyset$ for $i \neq j$, i.e., $V_i$ and $V_j$ intersect at most along a boundary. Suppose that $P_1, \ldots, P_m$ are polyhedra such that $\bigcup_{i=1}^m P_i = \mathbb{R}^n$, and int $P_i \cap$ int $P_j = \emptyset$ for $i \neq j$. Can this polyhedral decomposition of $\mathbb{R}^n$ be described as the Voronoi regions generated by an appropriate set of points?

*Solución.*
Considere dos hiperplanos $\mathcal{H}_1, \mathcal{H}_2 \subset \mathbb{R}^2$, los cuales no son paralelos. Considere que el ángulo entre los hiperplanos $\theta$ es tal que $0 < \theta < \frac{\pi}{3}$. Sean $x_1 \in P_1$ y $x_2 \in P_2$ dos puntos arbitrarios. Note que el punto $\widehat{x_1}$, el cual es el reflejo de $x_1$ respecto a $\mathcal{H}_1$ está en $\widehat{P_1} \subset P_4$. Similarmente, el punto $\widehat{x_2}$, el cual es el reflejo de $x_2$ respecto a $\mathcal{H}_2$ está en $\widehat{P_2} \subset P_4$. Sin embargo, no es posible construir el punto $\widehat{x_3} \in P_4$, el cual debe ser la reflexión de $x_1$ y $x_2$ respecto a $\mathcal{H}_1$ y $\mathcal{H}_2$ simultáneamente, ya que $\widehat{P_1} \cap \widehat{P_2} = \emptyset$.

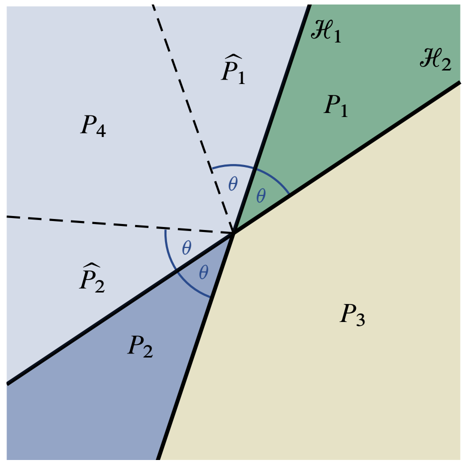


Figura 3: Descomposición de $\mathbb{R}^2$ en cuatro poliedros.

□

**2.10** *Solution set of a quadratic inequality.* Let $C \subseteq \mathbb{R}^n$ be the solution set of a quadratic inequality,

$$C = \left\{ x \in \mathbb{R}^n : x^T Ax + b^T x + c \leq 0 \right\}$$

with $A \in \mathbf{S}^n$, $b \in \mathbb{R}^n$, and $c \in \mathbb{R}$. Are the converses of these statements true?

*a)* Show that $C$ is convex if $A \succeq 0$.

*Solución.*
Un conjunto es convexo si su intersección con una línea arbitraria $y + tv$ es convexa. Considere entonces lo siguiente:

$$\begin{aligned}
(y+tv)^T A(y+tv) + b^T(y+tv) + c &= (y^T A + tv^T A)(y+tv) + b^T y + b^T tv + c \\
&= y^T Ay + 2y^T Atv + tv^T Atv + b^T y + b^T tv + c \\
&= (v^T Av)t^2 + (2y^T Atv + b^T tv) + (b^T y + y^T Ay + c) \\
&= (v^T Av)t^2 + (2y^T Av + b^T v)t + (b^T y + y^T Ay + c) \\
&= \gamma t^2 + \mu t + \phi.
\end{aligned}$$

Note que la intersección del conjunto solución $C$ y la línea $y + tv$ es el conjunto $\{y + tv : \gamma t^2 + \mu t + \phi \leq 0\}$, el cual es convexo si $\gamma \geq 0$, lo cual se tiene ya que $A \succeq 0$. La afirmación recíproca no se tiene, ya que se pueden construir ejemplos en los que $A \preceq 0$, sin embargo $C$ es convexo.

□

*b)* Show that the intersection of $C$ and the hyperplane defined by $g^T x + h = 0$ (where $g \neq 0$) is convex if $A + \lambda gg^T \succeq 0$ for some $\lambda \in \mathbb{R}$.

*Solución.*
Sea $B = A + \lambda gg^T$, entonces $A = B - \lambda gg^T$. Reemplazando en la desigualdad cuadrática:

$$\begin{aligned}
x^T Ax + b^T x + c \leq 0 &\therefore x^T(B - \lambda gg^T)x + b^T x + c \leq 0 \\
&\therefore x^T Bx - \lambda x^T gg^T x + b^T x + c \leq 0 \\
&\therefore x^T Bx + b^T x + (c - \lambda x^T gg^T x) \leq 0.
\end{aligned}$$

Por el literal anterior sabemos que el conjunto solución de la desigualdad obtenida es convexo, ya que $B \succeq 0$. Justamente, esta desigualdad representa la intersección entre $C$ y el hiperplano, el cual también es convexo.

$\square$

**2.13** *Conic hull of outer products.* Consider the set of rank-$k$ outer products, defined as $\{XX^T : X \in \mathbb{R}^{n\times k}, \textbf{rank}\, X = k\}$. Describe its conic hull in simple terms.

*Solución.*
Sea $\mathcal{O} = \{XX^T : X \in \mathbb{R}^{n\times k}, \textbf{rank}\, X = k\}$. Observe que $XX^T$ es semidefinida positiva, ya que para un vector real $y$, $y^T XX^T y = z^T z = \|z\|^2 \geq 0$, con $z = X^T y$. Además, sabemos que $\textbf{rank}\, XX^T = \textbf{rank}\, X^T X = \textbf{rank}\, X^T = \textbf{rank}\, X = k$. Sean $A, B \in \mathcal{O}$ matrices semidefinidas positivas tal que $\textbf{rank}\, A = \textbf{rank}\, B = k$. Sea $v \in \ker(A + B)$, entonces

$$(A+B)v = 0 \quad \therefore \quad v^T(A+B)v = 0 \quad \therefore \quad v^T A v + v^T B v = 0,$$

por lo que $v^T A v = v^T B v = 0$, y en particular $Av = Bv = 0$. Por lo tanto, $v$ está en los espacios nulos de $A$ y $B$, es decir, $v \in \ker A$ y $v \in \ker B$. Así, $\dim(\ker A + B) \leq \dim(\ker A), \dim(\ker B)$, y por el teorema de rango-nulidad, cualquier combinación $\gamma A + \delta B$ con $\gamma, \delta \geq 0$ debe tener rango mayor o igual que $k$, ya que al disminuir la nulidad, aumenta el rango. Por lo tanto, $\textbf{cone}\, \mathcal{O} = \{A : A \text{ es semidefinida positiva}, \textbf{rank}\, A \geq k\} \cup \{0_{n\times n}\}$.

$\square$

**2.14** *Expanded and restricted sets.* Let $S \subseteq \mathbb{R}^n$, and let $\|\cdot\|$ be a norm on $\mathbb{R}^n$.

*a)* For $a \geq 0$ we define $S_a$ as $\{x : \mathrm{dist}(x, S) \leq a\}$, where $\mathrm{dist}(x, S) = \inf_{y\in S} \|x - y\|$. We refer to $S_a$ as $S$ expanded or extended by $a$. Show that if $S$ is convex, then $S_a$ is convex.

*Solución.*
Sean $x, z \in S_a$. Queremos ver que la combinación convexa $\theta x + (1-\theta)z \in S_a$. Note que

$$\mathrm{dist}(\theta x + (1-\theta)z, S) = \inf_{y\in S} \|\theta x + (1-\theta)z - y\|.$$

Como $y \in S$ y $S$ es convexo, considere el caso particular $y = \theta y_1 + (1-\theta)y_2$. Luego,

$$\begin{aligned}
\mathrm{dist}(\theta x + (1-\theta)z, S) &= \inf_{y_1,y_2\in S} \|\theta x + (1-\theta)z - \theta y_1 - (1-\theta)y_2\| \\
&= \inf_{y_1,y_2\in S} \|\theta(x - y_1) + (1-\theta)(z - y_2)\| \\
&\leq \theta \inf_{y_1\in S} \|x - y_1\| + (1-\theta) \inf_{y_2\in S} \|x - y_2\| \\
&\leq \big(\theta + (1-\theta)\big)a \\
&\leq a.
\end{aligned}$$

$\square$

*b)* For $a \geq 0$ we define $S_{-a} = \{x : B(x, a) \subseteq S\}$, where $B(x, a)$ is the ball (in the norm $\|\cdot\|$), centered at $x$, with radius $a$. We refer to $S_{-a}$ as $S$ shrunk or restricted by $a$, since $S_{-a}$ consists of all points that are at least a distance $a$ from $\mathbb{R}^n \setminus S$. Show that if $S$ is convex, then $S_{-a}$ is convex.

*Solución.*
Sean $x, y \in S_{-a}$, y sea $z$ tal que $|z| \leq a$. Entonces $x + z, y + z \in S$, por lo que para $0 \leq \theta \leq 1$ se tiene que $\theta x + (1-\theta)y + z \in S$, ya que $S$ es convexo. Por lo tanto $\theta x + (1-\theta)y \in S_{-a}$.

$\square$

**2.15** *Some sets of probability distributions.* Let $x$ be a real-valued random variable with $\textbf{prob}\,(x = a_i) = p_i, i = 1, \ldots, n$, where $a_1 < a_2 < \cdots < a_n$. Of course $p \in \mathbb{R}^n$ lies in the standard probability simplex $P = \{p : \mathbf{1}^T p = 1, p \succeq 0\}$. Which of the following conditions are convex in $p$ ? (That is, for which of the following conditions is the set of $p \in P$ that satisfy the condition convex?)

*a)* $\alpha \leq \mathbf{E}\, f(x) \leq \beta$, where $\mathbf{E}\, f(x)$ is the expected value of $f(x)$, i.e., $\mathbf{E}\, f(x) = \sum_{i=1}^n p_i f(a_i)$. (The function $f : \mathbb{R} \to \mathbb{R}$ is given.)

*Solución.*
Esta condición es convexa en $p$, ya que es equivalente a la desigualdad lineal $\alpha \leq \sum_{i=1}^{n} p_i f(a_i) \leq \beta$. $\square$

*b)* $\mathbf{prob}(x > \alpha) \leq \beta$.

*Solución.*
Esta condición es convexa en $p$, ya que es equivalente a la desigualdad lineal $\sum_{i\,:\,a_i \geq \alpha} p_i \leq \beta$. $\square$

*c)* $\mathbf{E}\left|x^3\right| \leq \alpha\, \mathbf{E}\,|x|$.

*Solución.*
Esta condición es convexa en $p$, ya que es equivalente a la desigualdad lineal

$$\sum_{i=1}^{n} p_i |a_i|^3 \leq \alpha \sum_{i=1}^{n} p_i |a_i| = \sum_{i=1}^{n} p_i \left(|a_i|^3 - \alpha |a_i|\right) \leq 0.$$

$\square$

*d)* $\mathbf{E}\,x^2 \leq \alpha$.

*Solución.*
Esta condición es convexa en $p$, ya que es equivalente a la desigualdad lineal $\sum_{i=1}^{n} p_i (a_i)^2 \leq \alpha$. $\square$

*e)* $\mathbf{E}\,x^2 \geq \alpha$.

*Solución.*
Esta condición es convexa en $p$, ya que es similar a la del literal anterior. $\square$

*f)* $\mathbf{var}(x) \leq \alpha$, where $\mathbf{var}(x) = \mathbf{E}(x - \mathbf{E}\,x)^2$ is the variance of $x$.

*Solución.*
Considere la distribución de probabilidad $\{\mathbf{prob}(x = a_1 = 0) = 0, \mathbf{prob}(x = a_2 = 1) = 1\}$. Sea $\alpha = \frac{1}{8}$. Entonces

$$\mathbf{var}(x) = (0 \cdot 0^2 + 1 \cdot 1^2) - (0 \cdot 0 + 1 \cdot 1)^2 = 1 - 1 = 0 \leq \frac{1}{8}.$$

Sin embargo, al considerar la distribución $\{\mathbf{prob}(x = a_1 = 0) = \frac{1}{3}, \mathbf{prob}(x = a_2 = 1) = \frac{2}{3}\}$, podemos ver que $\mathbf{var}(x)$ no es convexa en general, ya que

$$\left(\frac{1}{3} \cdot 0^2 + \frac{2}{3} \cdot 1^2\right) - \left(\frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1\right)^2 = \frac{2}{9} \not\leq \frac{1}{8}.$$

$\square$

*g)* $\mathbf{var}(x) \geq \alpha$.

*Solución.*
Observe que esta condición la podemos reescribir como

$$\begin{aligned} \sum_{i=1}^{n} p_i (a_i)^2 - \left(\sum_{i=1}^{n} p_i a_i\right)^2 \geq \alpha \therefore -\sum_{i=1}^{n} p_i (a_i)^2 + \left(\sum_{i=1}^{n} p_i a_i\right)^2 &\leq \alpha \\ \therefore b^T p + p^T a a^T p &\leq \alpha, \end{aligned}$$

donde $b = a_i^2$. Esto es un conjunto convexo dado que $aa^T$ es semidefinida positiva. $\square$

*h)* $\mathbf{quartile}(x) \geq \alpha$, where $\mathbf{quartile}(x) = \inf\{\beta : \mathbf{prob}(x \leq \beta) \geq 0.25\}$.

*Solución.*
Observe que si $\alpha \leq a_1$, la condición se cumple trivialmente. En otro caso, ver que $\inf\{\beta : \mathbf{prob}(x \leq \beta) \geq 0.25\} \geq \alpha$ corresponde a ver la condición $\mathbf{prob}(x \leq \beta) < \alpha$ para todo $\beta < \alpha$. Por lo tanto, si definimos a $j = \max\{i : a_i \leq \alpha\}$, entonces

$$\mathbf{prob}(x \leq a_j) = \sum_{i=0}^{j} p_i < 0.25$$

define una desigualdad lineal, que a su vez define un semiespacio abierto. □

*i)* $\mathbf{quartile}(x) \leq \alpha$.

*Solución.*
Por un razonamiento similar al del literal anterior, ver que $\inf\{\beta : \mathbf{prob}(x \leq \beta) \geq 0.25\} \leq \alpha$ corresponde a ver la condición $\mathbf{prob}(x \leq \beta) \geq \alpha$ para todo $\beta \geq \alpha$. Esto es justamente el caso dual al literal anterior, por lo que esta condición corresponde a la desigualdad lineal definida por

$$\mathbf{prob}(x \leq a_{j+1}) = \sum_{i=j+1}^{n} p_i \geq 0.25.$$

□

**2.16** Show that if $S_1$ and $S_2$ are convex sets in $\mathbb{R}^{m+n}$, then so is their partial sum

$$S = \{(x, y_1 + y_2) : x \in \mathbb{R}^m, y_1, y_2 \in \mathbb{R}^n, (x, y_1) \in S_1, (x, y_2) \in S_2\}$$

*Solución.*
Sean $\mathbf{x}_1, \mathbf{x}_2 \in S$ tales que $x_1 = (\breve{x}, \breve{y_1} + \breve{y_2})$, $x_2 = (\widehat{x}, \widehat{y_1} + \widehat{y_2})$ tal que $(\breve{x}, \breve{y_1}), (\widehat{x}, \widehat{y_1}) \in S_1$ y $(\breve{x}, \breve{y_2}), (\widehat{x}, \widehat{y_2}) \in S_2$. Considere la siguiente combinación convexa:

$$\begin{aligned}
\mathbf{x}_1 + \mathbf{x}_2 &= \theta(\breve{x}, \breve{y_1} + \breve{y_2}) + (1-\theta)(\widehat{x}, \widehat{y_1} + \widehat{y_2}) \\
&= (\theta\breve{x}, \theta\breve{y_1} + \theta\breve{y_2}) + \big((1-\theta)\widehat{x}, (1-\theta)\widehat{y_1} + (1-\theta)\widehat{y_2}\big) \\
&= (\theta\breve{x} + (1-\theta)\widehat{x}, (\theta\breve{y_1} + (1-\theta)\widehat{y_1}) + (\theta\breve{y_2} + (1-\theta)\widehat{y_2}))
\end{aligned}$$

Como $S_1$ y $S_2$ son convexos, entonces $(\theta\breve{x} + (1-\theta)\widehat{x}, \theta\breve{y_1} + (1-\theta)\widehat{y_1}) \in S_1$ y $(\theta\breve{x} + (1-\theta)\widehat{x}, \theta\breve{y_2} + (1-\theta)\widehat{y_2}) \in S_2$. Por lo tanto, $\mathbf{x}_1 + \mathbf{x}_2 \in S$.

□

**2.17** *Image of polyhedral sets under perspective function.* In this problem we study the image of hyperplanes, halfspaces, and polyhedra under the perspective function $P(x,t) = x/t$, with $\mathbf{dom}\, P = \mathbb{R}^n \times \mathbb{R}_{++}$. For each of the following sets $C$, give a simple description of

$$P(C) = \{v/t : (v,t) \in C, t > 0\}.$$

*a)* The polyhedron $C = \mathbf{conv}\{(v_1, t_1), \ldots, (v_K, t_K)\}$ where $v_i \in \mathbb{R}^n$ and $t_i > 0$.

*Solución.*
Observe que $P(C) = \mathbf{conv}\left\{\frac{v}{t} : (v,t) \in C, t > 0\right\}$. Esto se obtiene por doble contenencia y por el resultado de la inducción del **Ejercicio 2.1**.

□

*b)* The hyperplane $C = \left\{(v,t) : f^T v + gt = h\right\}$ (with $f$ and $g$ not both zero).

*Solución.*
Como $t > 0$, observe que $P(C) = \left\{w = \frac{v}{t} : \frac{f^T v + gt}{t} = \frac{h}{t}\right\} = \left\{w : f^T w + g = \frac{h}{t}\right\}$. Es decir,

$$P(C) = \begin{cases} \{w : f^T w + g = 0\} & h = 0, \\ \{w : f^T w + g < 0\} & h < 0, \\ \{w : f^T w + g > 0\} & h > 0. \end{cases}$$

□

*c)* The halfspace $C = \{(v,t) : f^T v + gt \leq h\}$ (with $f$ and $g$ not both zero).

*Solución.*
De manera similar al literal anterior, como $t > 0$, observe que

$$P(C) = \left\{ w = \frac{v}{t} \ : \ \frac{f^T v + gt}{t} \leq \frac{h}{t} \right\} = \left\{ w \ : \ f^T w + g \leq \frac{h}{t} \right\}.$$

Es decir,

$$P(C) = \begin{cases} \{w \ : \ f^T w + g \leq 0\} & h = 0, \\ \{w \ : \ f^T w + g < 0\} & h < 0, \\ \{w \ : \ f^T w + g > 0\} = \mathbb{R}^n & h > 0. \end{cases}$$

□

*d)* The polyhedron $C = \{(v,t) : Fv + gt \preceq h\}$.

*Solución.*
Observe que

$$P(C) = \left\{ w = \frac{v}{t} \ : \ \frac{Fv + gt}{t} \preceq \frac{h}{t} \right\} = \left\{ w \ : \ Fw + g \preceq \frac{h}{t} \right\}.$$

Luego, $w \in P(C)$ si cumple los dos primeros casos de la proyección del literal anterior para cualquier vector columna $f_i$ de $F$, y además, si $h_i > 0$ y $h_j < 0$ entonces $\frac{f_i^T w + g_i}{h_i} \leq \frac{f_j^T w + g_j}{h_j}$.

□

**2.18** *Invertible linear-fractional functions.* Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be the linear-fractional function

$$f(x) = (Ax + b)/(c^T x + d), \quad \operatorname{dom} f = \{x : c^T x + d > 0\}$$

Suppose the matrix

$$Q = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix}$$

is nonsingular. Show that $f$ is invertible and that $f^{-1}$ is a linear-fractional mapping. Give an explicit expression for $f^{-1}$ and its domain in terms of $A, b, c$, and $d$. *Hint.* It may be easier to express $f^{-1}$ in terms of $Q$.

*Solución.*
Sea $x \in \mathbb{R}^n$. Considere la función $\mathcal{P}(x) = \{s(x,1) : s \geq 0\}$, la cual envía a vectores en $\mathbb{R}^n$ en rayos de $\mathbb{R}^{n+1}$. Sea $P$ la función perspectiva, entonces $f$ puede ser expresada como

$$\begin{aligned} P(Q\mathcal{P}(x)) &= P(Q \cdot s(x,1)) \\ &= P\left( \begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \begin{pmatrix} sx \\ s \end{pmatrix} \right) \\ &= P\left( \begin{bmatrix} s(Ax + b) \\ s(c^T + d) \end{bmatrix} \right) \\ &= \frac{s(Ax + b)}{s(c^T + d)} = \frac{Ax + b}{c^T + d} \\ &= f(x). \end{aligned}$$

Como $P$, $\mathcal{P}$ y $Q$ tienen inversas —$Q$ tiene inversa ya que es no singular—, entonces la inversa de $f$ es

$$f^{-1}(x) = \mathcal{P}^{-1}(Q^{-1} P^{-1}(x)).$$

□

## Ejercicios adicionales

**2.** Implementar un algoritmo para determinar si un punto está en el cono convexo de un conjunto de puntos en cualquier dimensión.

*Solución.*

La solución está implementada en el notebook de julia en `https://github.com/smolinad/IntroductionToOptimization/tree/main/Convex%20Sets`.

$\square$

**3.** Verificar que el conjunto de soluciones de un conjunto de desigualdades de matrices lineales es convexa.

*Solución.*

Sea $\mathcal{A}(x) = [A_1, A_2, \ldots, A_n]\,(x_1, x_2, \ldots, x_n)^T$. Considere la desigualdad lineal de matrices

$$\mathcal{A}(x) = x_1 A_1 + x_2 A_2 + x_3 A_3 + \cdots + x_n A_n \preceq B.$$

Note que la solución de la desigualdad anterior es el conjunto $\{x : \mathcal{A}(x) \preceq B\} = f^{-1}(\mathbf{S}^m_+)$, donde $f$ es la función afín $f : \mathbb{R}^n \to \mathbf{S}^m$ definida por $f(x) = B - \mathcal{A}(x)$. Que el conjunto $\{x : \mathcal{A}(x) \preceq B\}$ sea convexo sigue de que $f$ es afín.

$\square$

## Referencias


[1] S. Boyd y L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2009.