Universidad Nacional de Colombia
Introducción a la Optimización 2022-I
Ejercicios Conjuntos Convexos

29 de abril de 2022
Suárez Montilla, Laura Natalia.

Ejercicio 1 (2.1) Sea $C \subseteq \mathbb{R}^n$ un conjunto convexo, con $x_1, \ldots, x_k \in C$, y $\theta_1, \ldots, \theta_k \in \mathbb{R}$ satisfaciendo $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$. Muestre que $\theta_1 x_1 + \cdots + \theta_k x_k \in C$. (La definición de convexidad es es que esto es válido para $k = 2$; muestre esto para k arbitrario.) Pista. Use inducción sobre k.

Solución. De acuerdo con la pista, haremos inducción sobre k. Así, sea $C \subseteq \mathbb{R}^n$ un conjunto convexo, con $x_1, \ldots, x_k \in C$, y $\theta_1, \ldots, \theta_k \in \mathbb{R}$ satisfaciendo $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$, entonces:

- *Paso base:* Tomemos $k = 3$.
  Bajo las condiciones anteriores se tiene que a lo más uno de los θ_i puede ser igual a 1. Así, sin pérdida de generalidad, supongamos $\theta_1 \neq 1$ y tomemos z como sigue:

  $$z = \theta_1 x_1 + (1 - \theta_1)\left(a_2 x_2 + a_3 x_3\right); \quad a_i = \frac{\theta_i}{1 - \theta_1}, \text{ con } i = 2, 3.$$

  Como $\theta_1, \theta_2, (1 - \theta_1) \geq 0$, entonces $a_2, a_3 \geq 0$; además, $a_1 + a_2 = \frac{\theta_2 + \theta_3}{1 - \theta_1} = \frac{1 - \theta_1}{1 - \theta_1} = 1$. Con esto y el hecho de que C es convexo y $x_2, x_3 \in C$, se tiene que $y = a_2 x_2 + a_3 x_3 \in C$. Así, ya que y y x_1 están en C, concluimos que $z \in C$.

- *Hipótesis de inducción:* Si $k = m$, entonces $\theta_1 x_1 + \cdots + \theta_k x_k \in C$, con $m \in \mathbb{N}$.

- *Paso inductivo:* Si $k = m + 1$, nuevamente se tiene que a lo más uno de los θ_i puede ser igual a 1. Así, sin pérdida de generalidad, supongamos $\theta_1 \neq 1$ y tomemos z como sigue:

  $$z = \theta_1 x_1 + (1 - \theta_1)\left(\sum_{i=2}^{m+1} a_i x_i\right); \quad a_i = \frac{\theta_i}{1 - \theta_1}, \text{ con } i = 2, 3, \ldots, m + 1.$$

  Como $(1 - \theta_1) \geq 0$ y $\theta_i \geq 0$ para todo $i = 2, 3, \ldots m + 1$, entonces $a_i \geq 0$ para todo $i = 2, 3, \ldots m + 1$. Además,

  $$\sum_{i=2}^{m+1} a_i = \frac{\sum_{i=2}^{m+1} \theta_i}{1 - \theta_1} = \frac{1 - \theta_1}{1 - \theta_1} = 1.$$

  De este modo, ya que C es convexo y $x_i \in C$ para todo $i = 2, 3, \ldots, m + 1$, usando la hipótesis de inducción se tiene que $y = \sum_{i=2}^{m+1} a_i x_i \in C$. Así, puesto que y y x_1 están en C, concluimos que $z \in C$.

  De lo anterior, concluimos que si la afirmación es válida para $k = m$, se cumple para $k = m + 1$, por lo que es válida para todo $k \in \mathbb{N}$, como se quería.

■

Ejercicio 2 (2.3) *Convexidad punto medio.* Un conjunto C es punto medio convexo si dados dos puntos a, b en C, el punto medio $(a+b)/2$ está en C. Obviamente un conjunto convexo es punto medio convexo. Puede probarse que bajo condiciones adecuadas se tiene el recíproco.Como un caso sencillo, pruebe que si C es cerrado y punto medio convexo, entonces C es convexo.

Solución. Sean $x_1, x_2 \in C$. Como C es punto medio convexo, si tomamos los siguientes puntos

$$\begin{aligned} a_1 &= x_1(2^{-1}) + x_2(2^{-1}) \\ a_2 &= x_1(2^{-1}) + a_1(2^{-1}) = x_1(2^{-1}+2^{-2}) + x_2(2^{-2}) \\ a_3 &= x_1(2^{-1}) + a_2(2^{-1}) = x_1(2^{-1}+2^{-2}+2^{-3}) + x_2(2^{-3}) \\ a_i &= x_1(2^{-1}) + a_{i-1}(2^{-1});\ i \geq 4 \end{aligned}$$

tenemos que todos están en C. Esta idea se puede formalizar de la siguiente manera: Si tomamos $\theta \in [0,1]$ y denotamos

$$\theta^{(k)} = c_1 2^{-1} + c_2 2^{-2} + \cdots + c_k 2^{-k},\ c_i \in \{0,1\}$$

el número binario de longitud k más cercano a θ entonces, al aplicar la convexidad punto medio k veces de forma similar a como se hizo con los a_i, se tiene que $\theta^{(k)}x_1 + \left(1-\theta^{(k)}\right)x_2 \in C$. Por otra parte, como C es cerrado, se satisface

$$\lim_{k\to\infty} \left(\theta^{(k)}x_1 + \left(1-\theta^{(k)}\right)x_2\right) = \theta x_1 + (1-\theta)x_2 \in C.$$

De este modo, para todo $\theta \in [0,1]$ y todo $x_1, x_2 \in C$ se tiene que $\theta x_1 + (1-\theta)x_2 \in C$; es decir, C es convexo. ∎

Ejercicio 3 (2.4) Muestre que la envolvente convexa de un conjunto S es la intersección de todos los conjuntos convexos que contienen a S.

Solución.

Por definición, $\text{conv}S = \{\theta_1 x_1 + \theta_2 x_2 + \ldots \theta_k x_k \mid x_i \in S, \theta_i \geq 0, i = 1, 2, \ldots, k, \sum_{i=1}^{k} = 1\}$. Sea $\mathcal{C} = \bigcap\{C \mid C \text{ convexo},\ S \subseteq C\}$. Veamos que $\mathcal{C} = \text{conv}S$, así:

$\subseteq$) Sabemos que la envolvente convexa de un conjunto es un conjunto convexo, por lo que $\text{conv}S$ es convexo y $S \subseteq \text{conv}S$. Esto es, $S = C$ para algún C involucrado en $\mathcal{C}$. Luego, $\mathcal{C} \subseteq \text{conv}S$.

$\supseteq$) Sea $x \in \text{conv}S$, es decir, x es una combinación convexa de puntos $x_1, \ldots, x_j \in S$. Si consideramos ahora C un conjunto convexo con $S \subseteq C$, entonces $x_1, \ldots, x_j \in C$. Como C es convexo, contiene toda combinación convexa de puntos en él. En particular, como x es una combinación convexa de $x_1, \ldots, x_j$, se tiene que $x \in C$. Ya que C es arbitrario, x está en cualquier conjunto convexo tal que $S \subseteq C$; es decir, $x \in \mathcal{C}$.

De $\subseteq$) y $\supseteq$, tenemos que $\mathcal{C} = \text{conv}S$, como se quería. ∎

Ejercicio 4 (2.5) ¿Cuál es la distancia entre dos hiperplanos paralelos $H_1 = \left\{x \in \mathbb{R}^n \mid a^T x = b_1\right\}$ y $H_2 = \{x \in \mathbb{R}^n \mid a^T x = b_2\}$?

Solución. La distancia entre dos hiperplanos paralelos es igual a la distancia de uno de ellos a un punto cualquiera del otro hiperplano. Así, sea $x = (x_1, x_2, \ldots, x_n)^T \in H_1$; esto es, $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b_1$. Como a^T es el vector normal de H_2, la recta definida por $x + ta$ es perpendicular a H_2 y pasa por x. Luego, podemos encontrar $z = (z_1, z_2, \ldots, z_n)$ en la intersección de dicha recta y H_2. Para esto basta resolver la ecuación

$$a_1\left(x_1 + ta_1\right) + a_2\left(x_2 + ta_2\right) + \cdots + a_n\left(x_n + ta_n\right) = b_2.$$

Despejando t de la ecuación anterior se tiene

$$t = \frac{b_2 - a_1x_1 - a_2x_2 - \cdots - a_nx_n}{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Así,

$$\begin{aligned}
z &= (z_1, z_2, \ldots, z_n)\\
&= \left(x_1 + \frac{a_1\left(b_2 - a_1x_1 - \cdots - a_nx_n\right)}{a_1^2 + a_2^2 + \cdots + a_n^2}, \cdots, x_n + \frac{a_n\left(b_2 - a_1x_1 - \cdots - a_nx_n\right)}{a_1^2 + a_2^2 + \cdots + a_n^2}\right)\\
&= \left(x_1 + \frac{a_1(b_2 - b_1)}{a_1^2 + \cdots + a_n^2}, \cdots, x_n + \frac{a_n(b_2 - b_1)}{a_1^2 + \cdots + a_n^2}\right).
\end{aligned}$$

Finalmente, la distancia entre H_1 y H_2 será la distancia entre x y H_2, que es la distancia entre x y z; es decir

$$\begin{aligned}
d(H_1, H_2) &= d(x, H_2)\\
&= \|x - z\|\\
&= \sqrt{\left(x_1 - z_1\right)^2 + \cdots + \left(x_n - z_n\right)^2}\\
&= \sqrt{\left(\frac{a_n(b_2 - b_1)}{a_1^2 + \cdots + a_n^2}\right)^2 + \cdots + \left(\frac{a_n(b_2 - b_1)}{a_1^2 + \cdots + a_n^2}\right)^2}\\
&= \frac{|b_2 - b_1|}{\sqrt{a_1^2 + \cdots + a_n^2}}\\
&= \frac{|b_2 - b_1|}{\|a\|_2}.
\end{aligned}$$

■

Ejercicio 5 (2.8) ¿Cuáles de los siguientes conjuntos S son poliedros? Si es posible, exprese S en la forma $S = \{x \mid Ax \preceq b, Fx = g\}$.

(a) $S = \{y_1a_1 + y_2a_2 \mid -1 \leq y_1 \leq 1, -1 \leq y_2 \leq 1\}$, donde $a_1, a_2 \in \mathbb{R}^n$.

(b) $S = \left\{x \in \mathbb{R}^n \mid x \succeq 0, \mathbf{1}^T x = 1, \sum_{i=1}^n x_ia_i = b_1, \sum_{i=1}^n x_ia_i^2 = b_2\right\}$, donde $a_1, \ldots, a_n \in \mathbb{R}$ y $b_1, b_2 \in \mathbb{R}$.

(c) $S = \left\{x \in \mathbb{R}^n \mid x \succeq 0, x^T y \leq 1 \text{ para todo } y \text{ con } \|y\|_2 = 1\right\}$.

(d) $S = \left\{x \in \mathbb{R}^n \mid x \succeq 0, x^T y \leq 1 \text{ para todo } y \text{ con } \sum_{i=1}^n |y_i| = 1\right\}$.

Solución.

(a) S sí es un poliedro. Se puede expresar como la intersección de los tres conjuntos siguientes:

- S_1 : el plano definido por a_1 and a_2.
- $S_2 = \left\{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = a_2^T z = 0, -1 \leq y_1 \leq 1\right\}$. Una "loza"paralela a a_2 y ortogonal a S_1.
- $S_3 = \left\{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = a_2^T z = 0, -1 \leq y_2 \leq 1\right\}$. Una "loza"paralela a a_1 y ortogonal a S_1.

(b) S es un poliedro. $S = \{x \in \mathbb{R}^n \mid x \succeq 0, \mathbf{1}^T x = 1, a^T x = b_1, (a^{(2)})^T x = b_2\}$, donde $a = (a_1, \ldots, a_n)^T$ y $a^{(2)} = (a_1^2, \ldots, a_n^2)^T$.

(c) S no es un poliedro. Por la desigualdad de Cauchy-Schwarz, se tiene

$$x^T y \leq 1 \text{ para todo } y \text{ con } \|y\|_2 = 1 \iff \|x\|_2 \leq 1.$$

Así, S es la intersección de la bola unitaria $\{x \mid \|x\|_2 \leq 1\}$ y el octante no negativo de $\mathbb{R}^n$. Como se necesitan infinitos semiespacios intersectados, no es un poliedro.

(d) S sí es un poliedro. Para verlo, consideremos lo siguiente:

- Supongamos $x \succeq 0$. Si $x^T y \leq 1$ para todo y con $\sum_{i=1}^n |y_i| = 1$, en particular lo satisface con $\hat{y}$, donde $\hat{y}_j = 1$, con j satisfaciendo $|x_j| = \text{máx}_i |x_i|$, y $\hat{y_i = 0}$ si $i \neq j$. Así,

$$x^T \hat{y} = \sum_i x_i \hat{y}_i = \hat{y}_k x_k = |x_k| = \max_i |x_i|,$$

por lo que $|x_i| \leq 1$ para todo i.

- Supongamos ahora que $|x_i| \leq 1$ para todo i. Entonces, si $\sum_i |y_i| = 1$ se tiene

$$x^T y = \sum_i x_i y_i \leq \sum_i |x_i| |y_i| \leq \sum_i |y_i| = 1$$

.

Así, $x \in S$ si y solo si $x \succeq 0$ y $x \preceq 1$. Esto es, $S = \{x \in \mathbb{R}^n \mid x \succeq 0, x \preceq 1\}$

■

Ejercicio 6 (2.9) Sean $x_0, \ldots, x_K \in \mathbb{R}^n$. Considere el conjunto de puntos que son más cercanos a x_0 que a los otros x_i, i.e,

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

V es llamada la región Voronoi alrededor de x_0 con respecto a $x_1, \ldots, x_K$.

(a) Muestre que V es un poliedro. Exprese V en la forma $V = \{x \mid Ax \preceq b\}$.

(b) Dado un poliedro P con interior no vacío, muestre cómo encontrar $x_0, \ldots, x_K$ tal que el poliedro sea la región Voronoi de x_0 con respecto a $x_1, \ldots, x_K$.

(c) Suponga que $P_1, \ldots, P_m$ son poliedros tales que $\bigcup_{i=1}^m P_i = \mathbb{R}^n$, e int $P_i \cap$ int $P_j = \emptyset$ para $i \neq j$. ¿Puede esta descomposición en poliedros de $\mathbb{R}^n$ ser descritas como las regiones Voronoi generadas por un conjunto apropriado de puntos?

Solución.

(a) Sea $x \in \mathbb{R}^n$, entonces

$$\begin{aligned} x \in V &\iff \|x - x_0\|_2 \leq \|x - x_i\|_2 \\ &\iff (x - x_0)^T (x - x_0) \leq (x - x_i)^T (x - x_i) \\ &\iff \quad x^T x - 2x_0^T x + x_0^T x_0 \leq x^T x - 2x_i^T x + x_i^T x_i \\ &\iff 2(x_i - x_0)^T x \leq x_i^T x_i - x_0^T x_0 \end{aligned}$$

Así, V es un semiespacio y por tanto un poliedro, que podemos expresar como $V = \{x \mid Ax \preceq b\}$ con

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

(b) Consideremos $V = \{x \mid Ax \preceq b\}$ con $A \in \mathbb{R}^{K \times n}$ y $b \in \mathbb{R}^K$. Podemos elegir algún $x_0 \in \{x \mid Ax \prec b\}$, y construir K puntos x_i de la forma $x_i = x_0 + \lambda a_i$, donde λ es tal que la distancia de x_i al hiperplano definido por $a_i^T x = b_i$ es igual a la distancia de x_0 al hiperplano, esto es

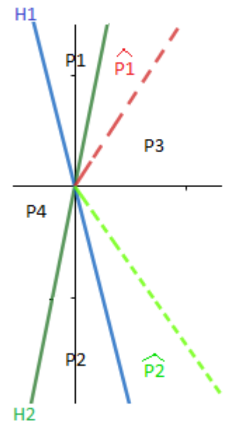
$$\begin{aligned} b_i - a_i^T x_0 = a_i^T x_i - b_i &\Rightarrow 2b_i = a_i^T (x_0 + \lambda a_i) + a_i^T x_0 \\ &\Rightarrow 2b_i = 2a_i^T x_0 + \lambda \|a_i\|^2 \\ &\Rightarrow \lambda = 2\left(b_i - a_i^T x_0\right) / \|a_i\|_2^2 \end{aligned}$$

y

$$x_i = x_0 + \frac{2\left(b_i - a_i^T x_0\right)}{\|a_i\|^2} a_i.$$

(c) Al tomar los hiperplanos $H_1 = \{x \in \mathbb{R}^2 \mid x_2 + 4x_1 = 0\}$ y $H_2 = \{x \in \mathbb{R}^2 \mid x_2 - 5x_1 = 0\}$, como en la figura, tenemos un contraejemplo en $\mathbb{R}^2$. Observemos que $\mathbb{R}^2$ está descompuesto en los poliedros P_1, P_2, P_3 y P_4. Si elegimos arbitrariamente $x_1 \in P_1$ y $x_2 \in P_2$, entonces $x_3 \in P_3$ debe ser el reflejo de x_1 y x_2 con respecto a H_2 y H_1, respectivamente. Sin embargo, tales reflejos se encuentran en $\hat{P}_1$ y $\hat{P}_2$, respectivamente. De este modo, es imposible encontrar x_3 que satisfaga los requerimientos. Por tanto, una descomposición en poliedros de $\mathbb{R}^n$ no siempre puede ser descrita como regiones Voronoi generadas por un conjunto de puntos $\{x_1, \ldots, x_m\}$.

∎



Ejercicio 7 (2.10) Sea $C \subseteq \mathbb{R}^n$ el conjunto solución de una desigualdad cuadrática,

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

con $A \in \mathbf{S}^n, b \in \mathbb{R}^n$, y $c \in \mathbb{R}$.

(a) Muestre que C es convexo si $A \succeq 0$.

(b) Muestre que la intersección de C y el hiperplano definido por $g^T x + h = 0$ (donde $g \neq 0$) es convexo si $A + \lambda gg^T \succeq 0$ para algún $\lambda \in \mathbb{R}$.

¿Son ciertos los recíprocos de las afirmaciones anteriores?

Solución. Un conjunto es convexo sii su intersección con una línea $\{j + tv \mid t \in \mathbb{R}\}$ es convexo.

(a) Si un punto está en la intersección entre la línea anterior y C, entonces

$$\begin{aligned}(j+tv)^T A(j+tv) + b^T(j+tv) + c &= v^T Avt^2 + (b^T v + 2j^T Av)t + c + b^T j + j^T Aj \\ &= \alpha t^2 + \beta t + \gamma \\ &\leq 0;\end{aligned}$$

con

$$\alpha = v^T Av,\ \ \beta = b^T v + 2j^T Av \text{ y } \gamma = c + b^T j + j^T Aj$$

Así, los puntos en tal intersección son los puntos en el conjunto

$$I = \{j + tv \mid \alpha t^2 + \beta t + \gamma \leq 0\}.$$

Observemos que si $\alpha > 0$, entonces $\alpha t^2 + \beta t + \gamma$ será similar a una parábola que abre hacia arriba y el conjunto I será convexo. Igualmente, si $\alpha = 0$, entonces $\alpha t^2 + \beta t + \gamma$ será una línea y el conjunto I será convexo. Es decir, si $\alpha \geq 0$, I es convexo, por lo que C también lo será. Esto se tiene si $v^T Av \geq 0$ para todo v, es decir, si $A \succeq 0$.

El recíproco no es cierto. Si tomamos $A = -2, b = 0, c = -1$ se tiene

$$x^T Ax + b^T x + c = -\|x\|_2^2 - 1 \leq 0 \text{ para todo } x \in \mathbb{R}.$$

Entonces, $C = \mathbb{R}$ es convexo pero $A \not\succeq 0$.

(b) Sea $H = \left\{x \mid g^T x + h = 0\right\}$. Tomemos α, β y γ como en la parte (a), y $\delta = g^T v$. Sin pérdida de generalidad, podemos asumir $j \in H$. La intersección de $C \cap H$ con la línea definida por j y v es

$$\left\{j + tv \mid \alpha t^2 + \beta t + \gamma \leq 0, \delta t = 0\right\}.$$

Si $\delta = g^T v \neq 0$, la intersección es $\{j\}$, si $\gamma \leq 0$, o es vacía. En cualquiera de los casos es un conjunto convexo. Si $\delta = g^T v = 0$, tal conjunto es

$$\left\{j + tv \mid \alpha t^2 + \beta t + \gamma \leq 0\right\},$$

que como vimos en (a), es convexo si $\alpha \geq 0$. Entonces $C \cap H$ es convexo si

$$g^T v = 0 \Longrightarrow v^T A v \geq 0.$$

Tal implicación se tiene si $A + \lambda g g^T \succeq 0$ para algún $\lambda \in \mathbb{R}$ puesto que

$$v^T A v = v^T \left(A + \lambda g g^T\right) v \geq 0$$

para todo v tal que $g^T v = 0$.

El recíproco no es cierto. Si tomamos $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $c = -1$ y $g = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, se tiene

$$x^T A x + b^T x + c = -\|x\|_2^2 - 1 \leq 0 \text{ para todo } x \in \mathbb{R}.$$

Entonces, $C = \mathbb{R}$ es convexo y como H es convexo, $C \cap H$ también lo es. Sin embargo,

$$A + \lambda g g^T = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} + \lambda \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 + \lambda \end{pmatrix},$$

por lo que no existe $\lambda \in \mathbb{R}$ tal que $A + \lambda g g^T \succeq 0$.

■

Ejercicio 8 (2.13) Considere el conjunto de productos externos de rango k, definido como $\left\{XX^T \mid X \in \mathbb{R}^{n \times k}, \text{ rango } X = k\right\}$. Describa su envolvente cónica.

Solución. De la descripción del conjunto, tenemos que $XX^T \succeq 0$ y rango $\left(XX^T\right) = k$. Sean A y B matrices semidefinidas positivas de rango k y sea $v \in \mathcal{N}(A + B)$, entonces

$$(A + B)v = 0 \Rightarrow v^T (A + B) v = 0 \Rightarrow v^T A v + v^T B v = 0.$$

Como $A, B \succeq 0$, entonces $v^T A v = 0$ y $v^T B v = 0$. Así,

$$v^T A v = v^T (XX^T) v = (X^T v)^T X^T v = 0 \Rightarrow X^T v = 0 \Rightarrow XX^T v = Av = 0,$$

y de manera análoga, $v^T B v = 0 \Rightarrow Bv = 0$. Es decir, $v \in \mathcal{N}(A) \cap \mathcal{N}(B)$. Por tanto, $\mathcal{N}(A + B) \subseteq \mathcal{N}(A) \cap \mathcal{N}(B)$.

De este modo, $\dim \mathcal{N}(A + B) \leq \dim \mathcal{N}(A)$ y $\dim \mathcal{N}(A + B) \leq \dim \mathcal{N}(B)$. Entonces, una combinación cónica de matrices semidefinidas positivas solo puede aumentar o mantener el rango. De aquí que la envolvente cónica del conjunto enunciado sea el conjunto de matrices semidefinidas positivas, de rango mayor o igual a k, junto con la matriz nula. ■

Ejercicio 9 (2.14) Sea $S \subseteq \mathbb{R}^n$, y sea $\|\cdot\|$ una norma sobre $\mathbb{R}^n$.

(a) Para $a \geq 0$ definimos $S_a = \{x \mid \operatorname{dist}(x, S) \leq a\}$, donde $\operatorname{dist}(x, S) = \operatorname{ínf}_{y\in S} \|x - y\|$. Nos referimos a S_a como S expandido o extendido por a. Muestre que si S es convexo, entonces S_a es convexo.

(b) Para $a \geq 0$ definimos $S_{-a} = \{x \mid B(x, a) \subseteq S\}$, donde $B(x, a)$ es la bola centrada en x, con radio a. Nos referimos a S_{-a} como S encogido o restricto por a, ya que S_{-a} consiste de todos los puntos que están al menos a una distancia a de $\mathbb{R}^n \backslash S$. Muestre que si S es convexo, entonces S_{-a} es convexo.

Solución.

(a) Sean $x_1, x_2 \in S_a$. Para $0 \leq \theta \leq 1$ se tiene

$$\begin{aligned}
\operatorname{dist}(\theta x_1 + (1-\theta)x_2, S) &= \operatorname*{ínf}_{y\in S} \|\theta x_1 + (1-\theta)x_2 - y\| \\
&= \operatorname*{ínf}_{y_1,y_2\in S} \|\theta x_1 + (1-\theta)x_2 - \theta y_1 - (1-\theta)y_2\| \\
&= \operatorname*{ínf}_{y_1,y_2\in S} \|\theta (x_1 - y_1) + (1-\theta)(x_2 - y_2)\| \\
&\leq \operatorname*{ínf}_{y_1,y_2\in S} (\theta \|x_1 - y_1\| + (1-\theta)\|x_2 - y_2\|) \\
&= \theta \operatorname*{ínf}_{y_1\in S} \|x_1 - y_1\| + (1-\theta) \operatorname*{ínf}_{y_2\in s} \|x_2 - y_2\| \Big) \\
&\leq a,
\end{aligned}$$

entonces $\theta x_1 + (1-\theta)x_2 \in S_a$, es decir, S_a es convexo.

(b) Sean $x_1, x_2 \in S_{-a}$; esto es, para todo u con $\|u\| \leq a$, $x_1 + u \in S$ y $x_2 + u \in S$. Para $0 \leq \theta \leq 1$ y $\|u\| \leq a$,

$$\theta x_1 + (1-\theta)x_2 + u = \theta (x_1 + u) + (1-\theta)(x_2 + u) \in S,$$

porque S es convexo. Así, $\theta x_1 + (1-\theta)x_2 \in S_{-\alpha}$; es decir, $S_{-\alpha}$ es convexo.

■

Ejercicio 10 (2.15) Sea x una variable aleatoria real con $\operatorname{prob}(x = a_i) = p_i, i = 1, \ldots, n$, donde $a_1 < a_2 < \ldots < a_n$. Claramente $p \in \mathbf{R}^n$ se encuentra en la probabilidad estándar $P = \{p \mid 1^T p = 1, p \geq 0\}$. ¿Cuáles de las siguientes condiciones son convexas en p?

(a) $\alpha \leq \mathbf{E}f(x) \leq \beta$, donde $\mathbf{E}f(x)$ es el valor esperado de $f(x)$, i.e., $\mathbf{E}f(x) = \sum_{i=1}^n p_i f(a_i)$.

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

(c) $\mathbf{E}|x^3| \leq \alpha \mathbf{E}|x|$.

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

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

(f) $\operatorname{var}(x) \leq \alpha$, donde $\operatorname{var}(x) = \mathbf{E}(x - \mathbf{E}x)^2$ es la varianza de x.

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

(h) quartile $(x) \geq \alpha$, donde quartile $(x) = \inf\{\beta \mid \text{prob}(x \leq \beta) \geq 0{,}25\}$.

(i) quartile $(x) \leq \alpha$.

Solución.

(a) Reemplazando $\mathbf{E}f(x)$ se tiene $\alpha \leq \sum_{i=1}^{n} p_i f(a_i) \leq \beta$. Observemos que si p y q cumplen tal condición, entonces para todo $\theta \in [0,1]$

$$\begin{aligned} \alpha = \theta\alpha + (1-\theta)\alpha &\leq \sum_{i=1}^{n}(p_i + q_i) f(a_i) = \sum_{i=1}^{n} \theta p_i f(a_i) + (1-\theta)\sum_{i=1}^{n} q_i f(a_i) \\ &= \theta \sum_{i=1}^{n} p_i f(a_i) + (1-\theta)\sum_{i=1}^{n} q_i f(a_i) \leq \theta\beta + (1-\theta)\beta = \beta. \end{aligned}$$

Luego, la condición es convexa.

(b) La proposición equivale a $\sum_{i:a_i \geq \alpha} p_i \leq \beta$. Así, si p y q cumplen la condición, para todo $\theta \in [0,1]$ se tiene

$$\sum_{i:a_i \geq \alpha}(\theta p_i + (1-\theta)q_i) = \theta \sum_{i:a_i \geq \alpha} p_i + (1-\theta)\sum_{i:a_i \geq \alpha} q_i \leq \theta\beta + (1-\theta)\beta = \beta.$$

Luego, la condición es convexa.

(c) La proposición equivale a $\sum_{i=1}^{n} p_i|a_i^3| \leq \alpha \sum_{i=1}^{n} p_i|a_i| \iff \sum_{i=1}^{n} p_i(|a_i^3| - \alpha|a_i|) \leq 0$. Así, si p y q satisfacen la condición, se tiene que para todo $\theta \in [0,1]$

$$\begin{aligned} \sum_{i=1}^{n}(\theta p_i + (1-\theta)q_i)(|a_i^3| - \alpha|a_i|) &= \theta \sum_{i=1}^{n} p_i(|a_i^3| - \alpha|a_i|) + (1-\theta)\sum_{i=1}^{n} q_i(|a_i^3| - \alpha|a_i|) \\ &\leq 0 + 0 = 0. \end{aligned}$$

Luego, la condición es convexa.

(d) La proposición equivale a $\sum_{i=1}^{n} p_i a_i^2 \leq \alpha$. Así, si p y q satisfacen la condición, para todo $\theta \in [0,1]$ se tiene

$$\sum_{i=1}^{n}(\theta p_i + (1-\theta)q_i)(a_i^2) = \theta \sum_{i=1}^{n} p_i a_i^2 + (1-\theta)\sum_{i=1}^{n} q_i a_i^2 \leq \theta\alpha + (1-\theta)\alpha = \alpha.$$

Luego, la condición es convexa.

(e) Bajo un razonamiento análogo al del literal anterior, la proposición es convexa.

(f) La proposición equivale a $\sum_{i=1}^{n} p_i a_i^2 - \left(\sum_{i=1}^{n} p_i a_i\right)^2 \leq \alpha$. Si tomamos $n = 2$, $a_i = 0$, $a_2 = 1$ y $\alpha = 1/5$, tenemos que $p = (0,1)$ y $q = (1,0)$ satisfacen la condición. Sin embargo, para $\frac{1}{2}p + \frac{1}{2}q$ se tiene

$$\sum_{i=1}^{n}\left(\frac{1}{2}p_i + \frac{1}{2}q_i\right)a_i^2 - \left(\sum_{i=1}^{n}\left(\frac{1}{2}p_i + \frac{1}{2}q_i\right)a_i\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4} \not\leq \frac{1}{5} = \alpha,$$

por lo que la condición no es convexa.

(h) La proposición es equivalente a $\operatorname{prob}(x \leq \beta) < 0{,}25$ para todo $\beta < \alpha$. Si $\alpha \leq a_1$, se tiene pues $\operatorname{prob}(x \leq \beta) = 0 < 0{,}25$. En caso contrario, sea $k = \max\{i \mid a_i < \alpha\}$. La condición se cumple si y solo si

$$\operatorname{prob}(x \leq a_k) = \sum_{i=1}^{k} p_i < 0{,}25.$$

Esta es una desigualdad lineal estricta en p, que define un semiespacio, por lo que es convexa.

(i) La proposición es equivalente a $\operatorname{prob}(x \leq \beta) \geq 0{,}25$ para todo $\beta \geq \alpha$. Si consideramos k como en el literal anterior, o $k = 0$ si $\alpha \leq a_1$ la condición puede ser expresada como

$$\sum_{i=k+1}^{n} p_i \geq 0{,}25,$$

que al igual que antes, define un semiespacio, por lo que es convexa.

■

Ejercicio 11 (2.16) Muestre que si S_1 y S_2 son conjuntos convexos en $\mathbb{R}^{m\times n}$, entonces lo es su suma parcial

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

Solución. Sean $(a, b_1 + b_2), (c, d_1 + d_2) \in S$; esto es, $(a, b_1), (c, d_1) \in S_1$ y $(a, b_2), (c, d_2) \in S_2$. Como S_1 y S_2 son convexos, para $0 \leq \theta \leq 1$ se tiene

$$\theta(a, b_1) + (1-\theta)(c, d_1) = (\theta a + (1-\theta)c, \theta b_1 + (1-\theta)d_1) \in S_1$$

y

$$\theta(a, b_2) + (1-\theta)(c, d_2) = (\theta a + (1-\theta)c, \theta b_2 + (1-\theta)d_2) \in S_2.$$

Luego,

$$\theta(a, b_1 + b_2) + (1-\theta)(c, d_1 + d_2) = (\theta a + (1-\theta)c, (\theta b_1 + (1-\theta)d_1) + (\theta b_2 + (1-\theta)d_2)) \in S;$$

es decir, S es convexo. ■

Ejercicio 12 (2.17) Estudiamos la imagen de hiperplanos, semiespacios y poliedros bajo la función perspectiva $P(x, t) = x/t$, con $\operatorname{dom} P = \mathbb{R}^n \times \mathbb{R}_{++}$. Para cada uno de los siguientes conjuntos C, de una simple descripción de

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

(a) El poliedro $C = \operatorname{conv}\{(v_1, t_1), \ldots, (v_K, t_K)\}$ donde $v_i \in \mathbb{R}^n$ y $t_i > 0$.

(b) El hiperplano $C = \left\{(v, t) \mid f^T v + gt = h\right\}$ (con f y g no ambas nulas).

(c) El semiespacio $C = \left\{(v, t) \mid f^T v + gt \leq h\right\}$ (con f y g no ambas nulas).

(d) El poliedro $C = \{(v,t) \mid Fv + gt \preceq h\}$.

Solución.

(a) El poliedro $P(C) = \operatorname{conv}\{v_1/t_1, \ldots, v_K/t_K\}$.

$\subseteq$) Sea $x = (v,t) \in C$, esto es

$$v = \sum_{i=1}^{K} \theta_i v_i, \quad t = \sum_{i=1}^{K} \theta_i t_i$$

donde $\theta \succeq 0, 1^T\theta = 1$. La imagen de x mediante P es

$$P(x) = v/t = \frac{\sum_{i=1}^{K} \theta_i v_i}{\sum_{i=1}^{K} \theta_i t_i} = \sum_{i=1}^{K} \alpha_i v_i/t_i;\ \alpha_i = \frac{\theta_i t_i}{\sum_{k=1}^{K} \theta_k t_k}, \quad i = 1, \ldots, K.$$

Se tiene que $\alpha \succeq 0$ y

$$\sum_{i=1}^{K} a_i = \sum_{i=1}^{K} \frac{\theta_1 t_i}{\sum_{k=1}^{K} \theta_k t_k} = \frac{\sum_{i=1}^{K} \theta_i t_i}{\sum_{k=1}^{K} \theta_k t_k} = 1,$$

por lo que $P(x) \in \operatorname{conv}\{v_1/t_1, \ldots, v_K/t_K\}$ para todo $x \in C$.

$\supseteq$) Sea $z \in \operatorname{conv}\{v_1/t_1, \ldots, v_K/t_K\}$, esto es

$$z = \sum_{i=1}^{K} \alpha_i v_i/t_i; \quad \text{con } \alpha \succeq 0, 1^T\alpha = 1.$$

Tomemos

$$\theta_i = \frac{\alpha_i}{t_i \sum_{j=1}^{K} \alpha_j/t_j}, \quad i = 1, \ldots, K.$$

Se tiene que $\theta \succeq 0$ y

$$\sum_{i=1}^{K} \theta_i = \sum_{i=1}^{K} \frac{\alpha_i}{t_i \sum_{j=1}^{K} \alpha_j/t_j} = \frac{\sum \alpha_i/t_i}{\sum_{j=1}^{K} \alpha_j} = 1.$$

Además, $z = P(v,t)$ donde

$$t = \sum_i \theta_i t_i = \frac{\sum_i \alpha_i}{\sum_j \alpha_j/t_j} = \frac{1}{\sum_j \alpha_j/t_j}, \quad v = \sum_i \theta_i v_i,$$

es decir, $(v,t) \in C$, por lo que $z \in P(C)$.

(b) Observemos que si $x = (v,t) \in C$, entonces, para $P(x) = v/t$ se tendrá

$$f^T\frac{v}{t} + \frac{gt}{t} = \frac{f^Tv + gt}{t} = \frac{h}{t}.$$

Así,

$$P(C) = \left\{z \mid f^Tz + g = h/t \text{ para algún } t > 0\right\} = \begin{cases} \left\{z \mid f^Tz + g = 0\right\} & h = 0 \\ \left\{z \mid f^Tz + g > 0\right\} & h > 0 \\ \left\{z \mid f^Tz + g < 0\right\} & h < 0. \end{cases}$$

(c) Observemos que si $x=(v,t)\in C$, entonces, para $P(x)=v/t$ se tendrá

$$f^T\frac{v}{t}+\frac{gt}{t}=\frac{f^Tv+gt}{t}\leq\frac{h}{t}.$$

Así,

$$P(C)=\left\{z \mid f^Tz+g\leq h/t \text{ para algún } t>0=\begin{cases}\left\{z\mid f^Tz+g\leq 0\right\} & h=0\\ \mathbb{R}^n & h>0\\ \left\{z\mid f^Tz+g<0\right\} & h<0\end{cases}\right.$$

(d) The polyhedron $C=\{(v,t)\mid Fv+gt\preceq h\}$. Observemos que si $x=(v,t)\in C$, entonces, para $P(x)=v/t$ se tendrá

$$\frac{1}{t}f^Tv+\frac{1}{t}gt\preceq\frac{1}{t}h.$$

Así, $P(C)=\{z\mid Fz+g\preceq(1/t)h \text{ for some } t>0\}$.

■

Ejercicio 13 (2.18) *Funciones fraccionarias lineales invertibles.* Sea $f:\mathbb{R}^n\to\mathbb{R}^n$ la función fraccionaria lineal

$$f(x)=(Ax+b)/\left(c^Tx+d\right),\quad \operatorname{dom} f=\left\{x\mid c^Tx+d>0\right\}$$

Suponga que la matriz

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

es no singular. Muestre que f es invertible y que f^{-1} es un mapeo lineal fraccionario. Dé una expresión explicita para f^{-1} y su dominio en términos de A,b,c, y d.

Solución. Como f puede ser expresada como $f(x)=P^{-1}(QP(x))$, su inversa f está dada por

$$f^{-1}(x)=(P^{-1}(QP(x)))^{-1}=P^{-1}\left(Q^{-1}P(x)\right)$$

entonces f^{-1} es la transformación proyectiva asociada con Q^{-1}.

En particular, si $n=1$, entonces $Q=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$, por lo que $Q^{-1}=\begin{pmatrix}-d & b\\ c & -a\end{pmatrix}$. Así, $f^{-1}(x)=\frac{-dx+b}{cx-a}$. Para comprobarlo, observemos lo siguiente:

-
$$\begin{aligned}(f\circ f^{-1})(x)&=f\left(\frac{-dx+b}{cx-a}\right)=\frac{a\left(\frac{-dx+b}{cx-a}\right)+b}{c\left(\frac{-dx+b}{cx-a}\right)+d}\\ &=\frac{\frac{-adx+ba+bcx-ab}{cx-a}}{\frac{-dcx+bc+dcx-ad}{cx-a}}=\frac{-adx+bcx}{bc-ad}\\ &=\frac{x(bc-ad)}{bc-ad}=x.\end{aligned}$$

- $$\begin{aligned}
(f^{-1}\circ f)(x) &= f^{-1}\left(\frac{ax+b}{cx+d}\right) = \frac{-d\left(\frac{ax+b}{cx+d}\right)+b}{c\left(\frac{ax+b}{cx+d}\right)-a} \\
&= \frac{\frac{-dax-bd+bcx+db}{cx+d}}{\frac{cax+bc-acx-ad}{cx+d}} = \frac{-dax+bcx}{bc-ad} \\
&= \frac{x(bc-ad)}{bc-ad} = x.
\end{aligned}$$

■

Ejercicio 14 Verificar que el conjunto de soluciones de un conjunto de desigualdades de matrices lineales es convexa. Es decir, que $S = \{x \mid A(x) = x_1A_1 + \cdots x_nA_n \preceq B\}$(con $A_i, B \in S^p$) es convexo.

Solución. Sean $x, y \in S$ y $\theta \in [0,1]$, entonces

$$\begin{aligned}
A(\theta x + (1-\theta)y) &= (\theta x_1 + (1-\theta)y1)A_1 + \cdots + (\theta x_n + (1-\theta)y_n)A_n \\
&= \theta x_1A_1 + (1-\theta)y_1A_1 + \cdots + \theta x_nA_n + (1-\theta)y_nA_n \\
&= \theta(x_1A_1 + \cdots + x_nA_n) + (1-\theta)(y_1A_1 + \cdots + y_nA_n) \\
&\preceq \theta B + (1-\theta)B = B.
\end{aligned}$$

Es decir, $\theta x + (1-\theta)y \in S$, por lo que S es convexo. ■