Conjuntos Convexos

Sergio Alejandro Vargas | savargasqu@unal.edu.co
Introducción a la Optimización — 2022-I
Universidad Nacional de Colombia

29 de abril de 2022

2.1 *Let $C \subseteq \mathbf{R}^n$ be a convex set, with $x_1, \ldots, x_k \in C$, and let $\theta_1, \ldots, \theta_k \in \mathbf{R}$ 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.*

Por definición sabemos que la afirmación se cumple para $k = 2$. Ahora, supongamos que $\theta_1 x_1 + \cdots + \theta_k x_k \in C$ se cumple para algún k arbitrario. Veamos si se cumple para $k + 1$.

Tomemos $x_1, \ldots x_k, x_{k+1} \in C$, y $\theta_1, \ldots, \theta_k, \theta_{k+1} \in \mathbf{R}$ tal que $\theta_1 + \cdots + \theta_k + \theta_{k+1} = 1$. Podemos reescribir la condición de la suma como

$$\sum_{i=1}^{k+1} \theta_i = 1$$

$$\sum_{i=1}^{k} \theta_i = 1 - \theta_{k+1}$$

$$\frac{1}{1 - \theta_{k+1}} \sum_{i=1}^{k} \theta_i = 1$$

$$\sum_{i=1}^{k} \frac{\theta_i}{1 - \theta_{k+1}} = 1$$

Si definimos $\eta_i = \frac{\theta_i}{1-\theta_{k+1}}$ tenemos $\sum_{i=1}^{k} \eta_i = 1$.

Así, podemos reescribir la combinación $\theta_1 x_1 + \ldots \theta_k x_k + \theta_{k+1} x_{k+1}$ como $(1 - \theta_{k+1})(\eta_1 x_1 + \cdots + \eta_k x_k) + \theta_{k+1} x_{k+1}$. Por hipótesis, si $x_1, \ldots, x_k \in C$ y $\sum \eta_i = 1$ entonces $(\eta_1 x_1 + \cdots + \eta_k x_k) \in C$. Llamemos al resultado de esta combinación y. Como el conjunto C es convexo y $x, y \in C$ entonces $(1-\theta_{k+1})y + \theta_{k+1}x_{k+1} \in C$. Por el principio de inducción matemática la afirmación se cumple para todo k.

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.*

Recordemos que un conjunto C es cerrado si y solo si contiene sus punto límite, es decir, el límite de cualquier sucesión de puntos de C converge en un punto en C.

Sean $a, b \in C$. Como C es punto-medio convexo, entonces $(a+b)/2$. Definimos la sucesión $s_n = (s_{n-1} + s_{n-2})/2$ Si tomamos $s_0 = x, s_1 = y$ obtenemos:

$$s_2 = (y+x)/2 s_3 = (s_3 + y)/2 \vdots$$

Vemos que cada iteración obtendremos un punto más cerca de y.

Como todos los puntos de la sucesión pertenecen a C, podemos tomar dos puntos cualesquiera producidos por la sucesión como los valores iniciales de otra sucesión producida por la misma regla. Así podemos ver que "llenar" el segmento de recta entre x, y, probando que el conjunto es convexo.

2.4 *Show that the convex hull of a set S is the intersection of all convex sets that contain S.*

Sea $C = \{A \mid S \subseteq A, \text{tal que } A \text{ es convexo}\}$. Por definición $\mathbf{conv}\,(S)$ contiene a S y es convexa. entonces $\mathbf{conv}\,(S)$ es un elemento de C, por lo que $\bigcap\limits_{A \in C} A \subseteq conv(S)$.

Ahora, sea $x \in \mathbf{conv}\,(S)$. El punto x debe ser la combinación convexa de puntos $x_1, \ldots, x_n$ en S. Si tomamos un conjunto convexo A tal que $S \subseteq A$, entonces $x_1, \ldots, x_n \in A$. Como A es convexo entonces también se tiene que $x \in A$. Si $x \in A$ para todo $A \in C$ tenemos que $x \in \bigcap\limits_{A \in C} A$ luego $\mathbf{conv}\,(S) \subseteq \bigcap\limits_{A \in C} A$.

La envolvente convexa contiene la intersección, y la intersección contiene la envolvente convexa. Por lo tanto $\mathbf{conv}\,(S) = \bigcap\limits_{A \in C} A$.

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

Sea $H_1 = \{x \in \mathbf{R}^n \mid a^T x = b_1\}$ y sea $H_2 = \{x \in \mathbf{R}^n \mid a^T x = b_2\}$. Sea x_1 un punto cualquiera en H_1 y sea L la recta que pasa por x_1 ortogonal a H_1 (paralela a a), es decir, $L = \{x_1 + at\}$ para todo $t \in \mathbf{R}$.

Como L es ortogonal a H_1, y además H_2 es paralelo a H_1 entonces L intersecta a H_2 en un punto $x_2 = x_1 + at$ para algún t. La distancia entre x_1 y x_2 es la distancia entre los dos hiperplanos. Para hallar x_2, primero hallamos t

$$\begin{aligned} a^T x_2 &= b_2 \\ a^T(x_1 + at) &= b_2 \\ a^T x_1 + a^T at &= b_2 \\ a^T at &= b_2 - a^T x_1 \\ t &= \frac{b_2 - a^T x_1}{a^T a} \end{aligned}$$

Luego $x_2 = x_1 + a\frac{b_2 - a^T x_1}{a^T a}$. Y para hallar la distancia vemos que

$$\begin{aligned} x_2 &= x_1 + a\frac{b_2 - a^T x_1}{a^T a} \\ x_2 &= x_1 + a\frac{b_2 - b_1}{a^T a} \\ x_2 - x_1 &= a\frac{b_2 - b_1}{a^T a} \\ \|x_2 - x_1\| &= \left\| a\frac{b_2 - b_1}{a^T a} \right\| \\ \|x_2 - x_1\| &= \|a\|\frac{|b_2 - b_1|}{a^T a} \\ \|x_2 - x_1\| &= \|a\|\frac{|b_2 - b_1|}{\|a\|^2} \\ \|x_2 - x_1\| &= \frac{|b_2 - b_1|}{\|a\|} \end{aligned}$$

2.8 *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) $S = \{y_1 a_1 + y_2 a_2 \mid -1 \le y_1 \le 1, -1 \le y_2 \le 1\}$, where $a_1, a_2 \in \mathbf{R}^n$.

Si a_1, a_2 son vectores colineales entonces S es una recta, no un poliedro. Si no son colineales, entonces S tendrá la forma de un "paralelogramo" con esquinas

$$a_1 + a_2, \quad a_1 - a_2, \quad -a_1 + a_2, \quad -a_1 - a_2$$

Podemos describir este paralelogramo con el hiper-plano definido por a_1, a_2 (de dimensión $n-1$) y por cuatro semi-espacios (de dimensión $n-2$) sobre el primer hiper-plano (correspondientes a los lados del "paralelogramo").

El plano definido por a_1, a_2 se puede definir por medio de $n-2$ ecuaciones de la forma $v_k^T x$ para $k = 1, \ldots, n-2$ donde cada v_k es un vector ortogonal a a_1 y a a_2. Llamemos a este plano P.

Los semi-espacios se pueden definir como:

$$\begin{aligned}
S_1 &= \{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = a_2^T z = 0, y_1 \leq 1\} \\
S_2 &= \{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = a_2^T z = 0, -1 \leq y_1\} \\
S_3 &= \{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = a_2^T z = 0, y_2 \leq 1\} \\
S_4 &= \{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = a_2^T z = 0, -1 \leq y_2\}
\end{aligned}$$

Sea c_1 un vector en P y ortogonal a a_2. Por ejemplo, podemos tomar

$$c_1 = a_1 - \frac{a_1^T a_2}{\|a_2\|^2} a^2$$

Entonces $x \in S_1$ si

$$c_1^T x \leq \left| c_1^T a_1 \right|$$

Podemos definir otros tres vectores de manera similar para obtener, junto con la restricción del plano las siguientes desigualdades:

$$\begin{aligned}
v_k^T x &\leq 0, \text{ para } k = 1, \ldots n-2 \\
-v_k^T x &\leq 0, \text{ para } k = 1, \ldots n-2 \\
c_1^T x &\leq \left| c_1^T a_1 \right| \\
c_2^T x &\leq \left| c_2^T a_1 \right| \\
c_3^T x &\leq \left| c_3^T a_2 \right| \\
c_4^T x &\leq \left| c_4^T a_2 \right|
\end{aligned}$$

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

S es un poliedro. Ya está definido por una desigualdad y por tres igualdades.

(c) $S = \{x \in \mathbf{R}^n \mid x \succeq 0, x^T y \preceq 1 \text{ for all } y \text{ with } \|y\|_2 = 1\}$.

S no es un poliedro. La condición $\|y\| = 1$ forma una bola unitaria, que no se puede definir por un número finito de semi-espacios e hiper-planos.

2.9 *Voronoi sets and polyhedral decomposition. Let $x_0, \dots, x_K \in \mathbf{R}^n$. Consider the set of points that are closer (in Euclidean norm) to x_0 than the other x_i, i.e., $V = \{x \in \mathbf{R}^n \mid \|x - x_0\|_2 \leq \|x - x_i\|_2, i = 1, \dots, K\}$. V is called the Voronoi region around x_0 with respect to $x_1, \dots, x_K$.*

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

Consideremos un caso particular con dos puntos x_1 y x_2 en $\mathbf{R}^2$. Intuitivamente, podemos ver que las regiones de Voronoi de estos dos puntos están separadas por una linea que pasa por el punto medio entre x_1 y x_2 y es ortogonal a la linea que conecta x_1 y x_2. Generalizando esta idea a K puntos en $\mathbf{R}^n$ se tiene que un punto x está en la región de x_0 si

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

Esta desigualdad define un semi-spacio. Si tomamos la desigualdad para cada punto x_i tenemos un poliedro que podemos expresar como

$$V = \{x \mid 2 \begin{bmatrix} x_1 - x_0 \\ \vdots \\ x_K - x_0 \end{bmatrix} \preceq \begin{bmatrix} x_1^T x_1 - x_0^T x_0 \\ \vdots \\ x_K^T x_K - x_0^T x_0 \end{bmatrix} \}$$

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

Tomemos un punto x_0 tal que $x_0 \in P$. Sabemos que cada hiper-plano $H_i = \{x \mid a_i^T x = b_i\}$ está a la misma distancia de x_0 que de otro punto x_i, es decir, la distancia de x_0 a x_i es 2 veces la distancia de x_0 al hiper-plano. Si tomamos x_0 y le sumamos 2 veces su distancia al plano en dirección del vector normal al plano a_i entonces tendremos x_i.

$$x_i = x_0 + 2a_i \frac{|a_i^T + b_i|}{\|a_i\|}$$

2.10 *Solution set of a quadratic inequality. Let $C \subseteq \mathbf{R}^n$ be the solution set of a quadratic inequality, $C = \{x \in \mathbf{R}^n \mid x^T Ax + b^T x + c \leq 0\}$, with $A \in \S^n, b \in \mathbf{R}^n$, and $c \in R$.*

(a) Show that C is convex if $A \succeq 0$. Un conjunto es convexo si y solo su su intersección con una linea (conjunto afín) arbitraria $L = \{\hat{x} + tv \mid t \in \mathbf{R}\}$ es convexa.

Tomando $x = \hat{x} + tv$ tenemos

$$
\begin{aligned}
(\hat{x}+tv)^T A(\hat{x}+tv) + b^T(\hat{x}+tv) + c(\hat{x}^T + tv^T)(A\hat{x} + Atv) + b^T\hat{x} + b^T tv + c \\
(\hat{x}^T A\hat{x} + \hat{x}^T Atv + tv^T A\hat{x} + tv^T Atv) + b^T\hat{x} + b^T tv + c \\
(\hat{x}^T A\hat{x} + 2\hat{x}^T Avt + v^T Avt^2) + b^T\hat{x} + b^T vt + c \text{ } (A \text{ es simétrica}) \\
v^T Avt^2 + 2\hat{x}^T Avt + b^T vt + b^T\hat{x} + \hat{x}^T A\hat{x} + c \\
v^T Avt^2 + (2\hat{x}^T Av + b^T v)t + b^T\hat{x} + \hat{x}^T A\hat{x} + c
\end{aligned}
$$

Sea $\alpha = v^T Av, \beta = (2x^T Av + b^T v)$ y $\gamma = b^T x + x^T Ax + c$. Podemos escribir la intersección de C con L como

$$C \cap L = x^T + tv \mid \alpha t^2 + \beta t + \gamma \leq 0$$

La intersección es convexa si $\alpha \geq 0$. Si A es semi-definida positiva entonces $\alpha = v^T Av \geq 0$ para todo v.

La recíproca no es cierta. Que C sea convexo no implica que $A \succeq 0$.

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

El producto exterior XX^T tiene que ser una matriz cuadrada semi-definida positiva, además $\mathbf{rank}\, XX^T = \mathbf{rank}\, X^T = \mathbf{rank}\, X = k$.

Ahora, sean A, B matrices semi-definidas de rango k, se tiene que $\mathbf{rank}\, A + B \leq k$. Sea $v \in \mathcal{N}(A+B)$, entonces

$$(A+B)v = 0 \iff v^T(A+B)v = 0 \iff v^T Av + v^T Bv = 0$$

Luego también se tiene que

$$v^T Av = 0 \iff Av = 0, v^T Bv = 0 \iff Bv = 0.$$

Por lo que $v \in \mathcal{N}(A)$ y $v \in \mathcal{N}(B)$. Es decir, la dimensión del núcleo de $A + B$ debe ser mayor o igual a la de los núcleos de A o B. Entonces el rango de $A+B$ no puede ser menor al de A o al de B.

Como la multiplicación por escalar no afecta el rango, y la suma de matrices semi-definidas no disminuye el rango, tenemos que la envolvente cónica es el conjunto de matrices semi-definidas con un rango mayor o igual a k.

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

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

Sean $x_1, x_2 \in S_a, \theta \in \mathbf{R}$ tal que $0 \leq \theta \leq 1$. Si probamos que **dist** es una función convexa, probamos que S_a es convexa. Usando la definición de **dist** tenemos

$$\mathbf{dist}(\theta x_1 + (1-\theta)x_2, S) = \inf_{y \in S} \|\theta x_1 + (1-\theta)x_2 - y\|$$

S es convexo, entonces podemos reemplazar y

$$\begin{aligned} &= \inf_{y_1, y_2 \in S} \|\theta x_1 + (1-\theta)x_2 - (\theta y_1 + (1-\theta)y_2)\| \\ &= \inf_{y_1, y_2 \in S} \|\theta(x_1 - y_1) + (1-\theta)(x_2 - y_2)\| \end{aligned}$$

Por desigualdad triangular

$$\begin{aligned} \mathbf{dist}(\theta x_1 + (1-\theta)x_2, S) &\leq \inf_{y_1, y_2 \in S} (\theta\|x_1 - y_1\| + (1-\theta)\|x_2 - y_2\|) \\ &\leq \theta \inf_{y_1, y_2 \in S} \|x_1 - y_1\| + (1-\theta) \inf \|x_2 - y_2\| \end{aligned}$$

x_1 y x_2 están en S_a entonces $\mathbf{dist}(\theta x_1 + (1-\theta)x_2, S) \leq a$. La combinación convexa entre x_1 y x_2 está en S_a, por lo tanto S_a es convexa.

(b) *For $a \geq 0$ we define $S_{-a} = \{x \mid B(x, a) \subseteq S\}$, 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 $\mathbf{R}^n \backslash S$. Show that if S is convex, then S_{-a} is convex.*

Tomemos dos puntos $x_1, x_2 \in S_{-a}$ y $\theta \in \mathbf{R}$ tal que $1 \leq \theta \leq 1$. Para todo u tal que $\|u\| \leq a$ se tiene que $x_1 + u \in S$ y que $x_2 + u \in S$. Como S es convexo, si sumamos u a la combinación convexa de x_1 con x_2 se tiene

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

Por lo tanto $\theta x_1 + (1-\theta)x_2 \in s_a$.

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

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

Sean $(x_1, y_1) \in S_1, (x_1, y_2) \in S(x_2, z_1) \in S_1$ y $(x_2, z_2) \in S$. Consideremos los puntos $(x_1, y_1 + y_2), (x_2, z_1, z_2)$. Para $0 \leq \theta \leq 1$ se tiene

$$\theta(x_1, y_1 + y_2) + (1 - \theta)(x_2, z_1 + z_2) = (\theta x_1 + (1 - \theta)x_1, (\theta y_1 + (1 - \theta)y_1) + (\theta z_1 + (1 - \theta)z_1)$$

S_1 es convexo entonces $(\theta x_1 + (1 - \theta)x_1, (\theta y_1 + (1 - \theta)y_1) \in S_1$. Igualmente S_2 es convexo entonces $(\theta x_1 + (1 - \theta)x_1, (\theta z_1 + (1 - \theta)z_1) \in S_2$. Luego $\theta(x_1, y_1 + y_2) + (1 - \theta)(x_2, z_1 + z_2) \in S$. Por lo tanto S es convexo.