Optimización Semidefinida

Clase 05 - SDP / Preliminares

Segundo Cuatrimestre 2021

Facultad de Ciencias Exactas y Naturales, UBA

Conos

Un subconjunto $\mathcal{C}$ de un espacio vectorial real V es un *cono convexo* si es cerrado por combinaciones lineales positivas. Es decir,

1. Si $\boldsymbol{x} \in \mathcal{C}$ y $a \geq 0$, entonces $a\boldsymbol{x} \in \mathcal{C}$.
2. Si $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}$, entonces $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{C}$.

En este apunte siempre que hablemos de *conos* vamos a referirnos a conos convexos.

Conos puntiagudos y cono dual

Decimos que un cono es *puntiagudo* si

$$\boldsymbol{x} \in \mathcal{C} \text{ y } -\boldsymbol{x} \in \mathcal{C} \Rightarrow \boldsymbol{x} = 0.$$

El *cono dual* de un cono $\mathcal{C}$ es

$$\mathcal{C}^* = \{\boldsymbol{y} \in \mathbb{R}^n \mid \boldsymbol{x}^T \boldsymbol{y} \geq 0 \;\; \forall \boldsymbol{x} \in \mathcal{C}\}.$$

Orden parcial en un cono

Dado un cono puntiagudo $\mathcal{C}$ en $\mathbb{R}^n$, podemos definir un orden parcial en $\mathbb{R}^n$ por

$$\boldsymbol{x} \succeq \boldsymbol{y} \iff \boldsymbol{x} - \boldsymbol{y} \in \mathcal{C}$$

para $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Este orden satisface las siguientes propiedades:

- **reflexividad:** $\forall \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \succeq \boldsymbol{x}$
- **antisimetría:** $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{x} \succeq \boldsymbol{y}, \boldsymbol{y} \succeq \boldsymbol{x} \Rightarrow \boldsymbol{x} = \boldsymbol{y}$
- **transitividad:** $\forall \boldsymbol{x}, \boldsymbol{y}, z \in \mathbb{R}^n : \boldsymbol{x} \succeq \boldsymbol{y}, \boldsymbol{y} \succeq \boldsymbol{z} \Rightarrow \boldsymbol{x} \succeq \boldsymbol{z}$
- **homogeneidad:** $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}_{\geq 0} : \boldsymbol{x} \succeq \boldsymbol{y} \Rightarrow \alpha \boldsymbol{x} \succeq \alpha \boldsymbol{y}$
- **aditividad:** $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}', \boldsymbol{y}' \in \mathbb{R}^n : \boldsymbol{x} \succeq \boldsymbol{y}, \boldsymbol{x}' \succeq \boldsymbol{y}' \Rightarrow \boldsymbol{x} + \boldsymbol{x}' \succeq \boldsymbol{y} + \boldsymbol{y}'$

Desigualdad estricta

En general, nos van a interesar conos convexos de dimensión completa, es decir conos con interior no vacío. En ese caso, podemos definir la desigualdad estricta

$$\boldsymbol{x} \succ \boldsymbol{y} \iff \boldsymbol{x} - \boldsymbol{y} \in \operatorname{int} \mathcal{C}.$$

Separación

Tenemos el siguiente resultado de separación.

Lema

Sean $\mathcal{C} \subset \mathbb{R}^n$ un cono convexo cerrado y $\boldsymbol{x} \in \mathbb{R}^n \setminus \mathcal{C}$ un punto fuera de $\mathcal{C}$. Existe un hiperplano que separa a $\{\boldsymbol{x}\}$ de $\mathcal{C}$. Más aún, existe un vector no nulo $c \in \mathbb{R}^n$ tal que

$$\forall \boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{c} \cdot \boldsymbol{y} \geq 0 > \boldsymbol{c} \cdot \boldsymbol{x}.$$

Ejemplo: cono generado

Cono generado por un conjunto de puntos. El cono generado por un conjunto de puntos $A = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \mathbb{R}^n$ es el menor cono que contiene a A. Es el cono

$$\operatorname{cono} A = \left\{ \sum_{i=1}^{N} \alpha_i \boldsymbol{x}_i \mid \alpha_1, \ldots, \alpha_N \in \mathbb{R}_{\geq 0} \right\}.$$

Ejemplos.

- El cono generado por $\{(1,0),(0,1)\} \subset \mathbb{R}^2$ es el conjunto de puntos
$$\{(x,y) \in \mathbb{R}^2 : x \geq 0 \text{ e } y \geq 0\}.$$
- El cono generado por $\{(0,1),(0,-1)\} \subset \mathbb{R}^2$ es la recta
$$\{(x,y) \in \mathbb{R}^2 : x = 0\}.$$

El ortante no-negativo

El ortante no-negativo que utilizamos en programación lineal, definido por

$$\mathbb{R}^n_{\geq 0} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1, \ldots, x_n \geq 0\}$$

es un cono puntiagudo convexo cerrado de dimensión completa.

El cono de matrices semidefinidas positivas.

Proposición

El conjunto $\mathcal{S}_+^n$ de matrices positivas semidefinidas es un cono convexo puntiagudo.

Demostración. Podemos escribir a $\mathcal{S}_+^n$ como intersección de infinitos semiespacios:

$$\mathcal{S}_+^n = \{\boldsymbol{A} \in \mathcal{S}^n \mid \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} \geq 0 \ \ \forall \boldsymbol{x} \in \mathbb{R}^n\} = \bigcap_{\boldsymbol{x} \in \mathbb{R}^n} \{\boldsymbol{A} \in \mathcal{S}^n \mid \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} \geq 0\},$$

y por lo tanto es un conjunto convexo cerrado.

El cono de matrices semidefinidas positivas.

Para ver que forman un cono, si $\boldsymbol{A}, \boldsymbol{B} \succeq 0$ y $a \geq 0$, es claro que $a\boldsymbol{A} \succeq 0$ y también que

$$\boldsymbol{x}^T(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} \geq 0$$

para todo $\boldsymbol{x} \in \mathbb{R}^n$.

Finalmente, para ver que es un cono puntiagudo, si $\boldsymbol{A} \in \mathcal{S}_+^n$ y $-\boldsymbol{A} \in \mathcal{S}_+^n$, todos los autovalores de $\boldsymbol{A}$ son 0. Como $\boldsymbol{A}$ es simétrica, el teorema espectral para matrices simétricas implica que $\boldsymbol{A}$ es la matriz nula.

En particular: la suma de matrices semidefinidas positivas es semidefinida positiva y una combinación convexa también.

El interior del cono de matrices semidefinidas positivas

Proposición

El interior del cono $\mathcal{S}_+^n$ de matrices semidefinidas positivas es $\mathcal{S}_{++}^n$, el conjunto de matrices definidas positivas.
En particular, $\mathcal{S}_+^n$ es un cono de dimensión completa.

Demostración. Ver [Fawzi, 2018, Teorema 3.1].

Desigualdades lineales matriciales

Definición

Una desigualdad lineal matricial (LMI por su nombre en inglés) tiene la forma

$$\boldsymbol{A}_0 + \sum_{i=1}^{m} \boldsymbol{A}_i y_i \succeq 0,$$

con $\boldsymbol{A}_i \in \mathcal{S}^n$ matrices simétricas dadas.

Un conjunto de desigualdades matriciales definen un espectrahedro.

Espectrahedros

Definición

Un conjunto $S \subset \mathbb{R}^m$ es un espectrahedro si es de la forma

$$S = \left\{ (y_1, \ldots, y_m) \in \mathbb{R}^m : \boldsymbol{A}_0 + \sum_{i=1}^{m} \boldsymbol{A}_i y_i \succeq 0 \right\},$$

para matrices simétricas $\boldsymbol{A}_0, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathcal{S}^n$.

Ejemplo.

$$S = \left\{ y \in \mathbb{R} : \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \succeq 0 \right\},$$

define el conjunto

Espectrahedros

Definición

Un conjunto $S \subset \mathbb{R}^m$ es un espectrahedro si es de la forma

$$S = \left\{ (y_1, \ldots, y_m) \in \mathbb{R}^m : \boldsymbol{A}_0 + \sum_{i=1}^{m} \boldsymbol{A}_i y_i \succeq 0 \right\},$$

para matrices simétricas $\boldsymbol{A}_0, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathcal{S}^n$.

Ejemplo.

$$S = \left\{ y \in \mathbb{R} : \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \succeq 0 \right\},$$

define el conjunto $S = [-1, 1]$.

Escritura alternativa

Observamos que podemos escribir una desigualdad matricial lineal mediante una sola matriz

$$\boldsymbol{A}(y_1, \dots, y_m) = \boldsymbol{A}_0 + \sum_{i=1}^{m} \boldsymbol{A}_i y_i \succeq 0,$$

donde las coordenadas de $\boldsymbol{A}(y_1, \dots, y_m)$ son expresiones lineales afines en las variables $y_1, \dots, y_m$.

En el ejemplo anterior obtenemos

$$S = \left\{ y \in \mathbb{R} : \begin{pmatrix} 1 & y \\ y & 1 \end{pmatrix} \succeq 0 \right\}.$$

Poliedros y espectrahedros

Ejemplo. Todos los poliedros son espectrahedros.

Para una matriz $\boldsymbol{X} \in \mathbb{R}^{n \times n}$ diagonal,

$$\boldsymbol{X} \succeq 0 \text{ si y solo si } x_{ii} \geq 0 \text{ para todo } 1 \leq i \leq n.$$

Por lo tanto, podemos describir al poliedro $S = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$ como

Poliedros y espectrahedros

Ejemplo. Todos los poliedros son espectrahedros.

Para una matriz $\boldsymbol{X} \in \mathbb{R}^{n \times n}$ diagonal,

$$\boldsymbol{X} \succeq 0 \text{ si y solo si } x_{ii} \geq 0 \text{ para todo } 1 \leq i \leq n.$$

Por lo tanto, podemos describir al poliedro $S = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$ como

$$S = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : \begin{pmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{x} - b_1 & & \\ & \ddots & \\ & & \boldsymbol{a}_m \cdot \boldsymbol{x} - b_m \end{pmatrix} \succeq 0 \right\}.$$

Sin embargo, existen espectrahedros que no son poliedros.

Ejemplo: el disco unidad

El disco unitario $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ es un espectrahedro.

Podemos definirlo mediante la desigualdad matricial

Ejemplo: el disco unidad

El disco unitario $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ es un espectrahedro.

Podemos definirlo mediante la desigualdad matricial

$$D = \left\{(x, y) \in \mathbb{R}^2 \mid \begin{pmatrix} 1 - x & y \\ y & 1 + x \end{pmatrix} \succeq 0 \right\}.$$

Utilizando el criterio de los menores principales, vemos que $\begin{pmatrix} 1 - x & y \\ y & 1 + x \end{pmatrix} \succeq 0$ si y solo si:

$$\begin{aligned} 1 - x &\geq 0, \\ 1 + x &\geq 0, \\ (1 - x)(1 + x) - y^2 = 1 - x^2 - y^2 &\geq 0. \end{aligned}$$

Geometría de los espectrahedros

Geométricamente, un espectrahedro está definido por la intersección del cono $\mathcal{S}_+$ de matrices semidefinidas positivas con un espacio afín: el espacio generado por las matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ trasladado a $\boldsymbol{A}_0$.

Por lo tanto, un espectrahedro es siempre un conjunto cerrado y convexo.

Determinar qué conjuntos convexos son espectrahedros y cuáles no mediante un criterio sencillo es un problema abierto de investigación.

Proyección de espectrahedros

Ejemplo. El conjunto $S = \{(x, y) : xy \geq 1, x \geq 0, y \geq 0\}$ es un espectrahedro (¿cómo se puede definir por desigualdades matriciales?).

La proyección de S sobre el eje X:

$$\pi_x(S) = \{x \geq 0 : \exists y \geq 0 \text{ tal que } xy \geq 1\}$$

no es un conjunto cerrado y por lo tanto no es un espectrahedro.

Utilizando proyecciones de espectrahedros podemos ampliar la familia de conjuntos convexos sobre la que podemos resolver problemas SDP.

Conjuntos semi-algebraicos

Una variedad algebraica sobre $\mathbb{C}$ es un subconjunto de $\mathbb{C}^n$ definido por ecuaciones polinomiales:

$$V = \{\boldsymbol{x} \in \mathbb{C}^n : p_1(\boldsymbol{x}) = 0, \ldots, p_s(\boldsymbol{x}) = 0, p_i \in \mathbb{C}[X_1, \ldots, X_n], 1 \leq i \leq s\}.$$

Un conjunto semi-algebraico es un subconjunto de $\mathbb{R}^n$ definido por ecuaciones polinomiales e inecuaciones polinomiales

$$S = \{\boldsymbol{x} \in \mathbb{R}^n : p_i(\boldsymbol{x}) = 0, 1 \leq i \leq s, q_j(\boldsymbol{x}) > 0, 1 \leq j \leq t\},$$

o uniones finitas de estos conjuntos.

Conjuntos semi-algebraicos y espectrahedros

Utilizando el criterio de los menores para matrices semidefinidas positivas, obtenemos:

Propiedad: Los espectrahedros son conjuntos semi-algebraicos.

El teorema de Tarski-Seidenberg asegura que cualquier proyección de un conjunto semi-algebraico es también un conjunto semi-algebraico, pero vimos que esto no ocurre para los espectrahedros.

Pregunta: ¿cualquier conjunto semialgebraico convexo será la proyección de un espectrahedro?

En 2016, C. Scheiderer demostró que no todo conjunto semi-algebraico convexo es la proyección de un espectrahedro.

Matriz por bloques

Dadas matrices simétricas $\boldsymbol{A} \in \mathbb{R}^n$, $\boldsymbol{B} \in \mathbb{R}^m$, definimos la matriz $\boldsymbol{A} \oplus \boldsymbol{B}$ por

$$\boldsymbol{A} \oplus \boldsymbol{B} = \begin{pmatrix} \boldsymbol{A} & 0 \\ 0 & \boldsymbol{B} \end{pmatrix}.$$

Se cumple: $\boldsymbol{A} \succeq 0$ y $\boldsymbol{B} \succeq 0$ si y solo si $\boldsymbol{A} \oplus \boldsymbol{B} \succeq 0$.

Esto nos permite juntar varias desigualdades matriciales en una sola matriz.

Bola de norma-4

Ejercicio

Demostrar que el conjunto

$$D = \{(x, y) \in \mathbb{R}^2 \mid x^4 + y^4 \leq 1\}$$

es un espectrahedro.

Sugerencia:

$$x^4 + y^4 \leq 1 \iff \exists u, v \text{ tales que } x^2 \leq u, y^2 \leq v, u^2 + v^2 \leq 1.$$

Producto interno traza

Para plantear los problemas de programación semi-definida en forma análoga a un problema de programación lineal, introducimos la siguiente notación.

Para matrices $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^{n \times n}$, llamamos *producto interno traza* al producto

$$\boldsymbol{X} \bullet \boldsymbol{Y} = \operatorname{Tr}(\boldsymbol{X}^T \boldsymbol{Y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} y_{ij}, \quad \text{donde } \operatorname{Tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii},$$

que coincide con el producto interno usual pensado a $\boldsymbol{X}$ e $\boldsymbol{Y}$ como vectores de $n \times n$ coordenadas.

Con este producto interno, $\mathbb{R}^{n \times n}$ resulta un espacio euclídeo.

Se cumplen las siguientes propiedades:

1. $\boldsymbol{X} \bullet \boldsymbol{Y} = \boldsymbol{Y} \bullet \boldsymbol{X}$,
2. $\boldsymbol{X} \bullet \boldsymbol{I} = \boldsymbol{I} \bullet \boldsymbol{X} = \operatorname{Tr}(\boldsymbol{X})$.

Problema primal de programación semidefinida

Con la notación introducida, escribimos un problema de programación semidefinida (en forma primal) como

$$\begin{aligned} \text{minimizar:} \quad & \boldsymbol{C} \bullet \boldsymbol{X} \\ \text{sujeto a:} \quad & \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i, \quad i = 1, \ldots, m, \\ & \boldsymbol{X} \succeq 0, \end{aligned}$$

donde $\boldsymbol{C}, \boldsymbol{A}_i \in \mathcal{S}^n$ (el conjunto de matrices simétricas).

La matrix $\boldsymbol{X} \in \mathcal{S}^n$ es la variable sobre la cual realizamos la minimización.

Tanto $\boldsymbol{C} \bullet \boldsymbol{X}$ como $\boldsymbol{A}_i \bullet \boldsymbol{X}, 1, \ldots, m$, representan combinaciones lineales de las coordenadas de $\boldsymbol{X}$.

Podemos ver en seguida la similitud con un problema de programación lineal.

Conjunto factible

La función a minimizar es una funcional lineal exactamente igual que en programación lineal.

Para describir el conjunto factible

$$\{\boldsymbol{X} : \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i, i = 1, \ldots, m, \quad \boldsymbol{X} \succeq 0\}$$

mediante una desigualdad matricial lineal, podemos despejar las coordenadas de $\boldsymbol{X}$ como expresiones lineales en función de un conjunto de variables libres.

Ejemplo

Podemos describir conjunto de matrices

$$S = \left\{ \boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \mid x_{11} + x_{22} = 1, \boldsymbol{X} \succeq 0 \right\}$$

mediante la condición

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{12} & 1 - x_{11} \end{pmatrix} \succeq 0$$

que a la vez es equivalente a la desigualdad matricial

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + x_{11} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + x_{12} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \succeq 0.$$

Obtenemos $S \cong \tilde{S} = \left\{ (y_1, y_2) \in \mathbb{R}^2 : \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + y_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \succeq 0 \right\}.$

Conjunto factible y espectrahedro

En nuestra definición, un espectrahedro S es un conjunto cerrado y convexo del espacio afín $\mathbb{R}^m$.

Siguiendo la terminología usual, llamaremos también espectrahedro al conjunto

$$\left\{ \boldsymbol{A}_0 + \sum_{i=1}^{m} \boldsymbol{A}_i y_i \succeq 0 \mid \boldsymbol{y} \in \mathbb{R}^m \right\}.$$

Si bien este es un conjunto de matrices y no un subconjunto de $\mathbb{R}^m$, si las matrices $\boldsymbol{A}_i$, $1 \leq i \leq m$, son linealmente independientes, ambos conjuntos son afínmente equivalentes.

El conjunto factible de un problema de programación semi-definida es un espectrahedro.

Ejercicio

Resolver el siguiente problema de programación semidefinida para $\boldsymbol{X} \in \mathbb{R}^{3\times 3}$:

$$
\begin{aligned}
\text{minimizar:} \quad & 2x_{22} + x_{11} \\
\text{sujeto a:} \quad & x_{31} = 0, \\
& x_{32} = 0, \\
& x_{22} = 2 - x_{11}, \\
& x_{33} = x_{11} + x_{21} - 1, \\
& \boldsymbol{X} \succeq 0.
\end{aligned}
$$

Formulación dual

Utilizando desigualdades matriciales lineales tenemos otra forma de plantear problemas SDP. Llamamos formulación dual a un problema planteado en la forma:

$$\begin{aligned} &\text{maximizar: } \boldsymbol{b} \cdot \boldsymbol{y} \\ &\text{sujeto a:} \quad \boldsymbol{A}_0 + y_1 \boldsymbol{A}_1 + \cdots + y_m \boldsymbol{A}_1 \succeq 0. \end{aligned}$$

Veremos en la próxima clase que cada problema primal tiene asociado un problema dual tal que las soluciones de uno de los problemas pueden usarse para obtener cotas sobre las soluciones del otro.

Ejemplo de formulación dual

Resolver el siguiente problema:

$$
\begin{aligned}
\text{maximizar:} \quad & y \\
\text{sujeto a:} \quad & \begin{pmatrix} 2-y & 1 \\ 1 & -y \end{pmatrix} \succeq 0
\end{aligned}
$$

Introducción a Mosek

Como vemos en la página

`https://docs.mosek.com/latest/pythonapi/tutorial-sdo-shared.html`,

Mosek puede resolver problemas de optimización semidefinida de la forma:

$$
\begin{aligned}
\text{minimizar:} \quad & \sum_{j=0}^{n-1} c_j x_j + \sum_{j=0}^{p-1} \bar{\boldsymbol{C}}_j \bullet \bar{\boldsymbol{X}}_j \\
\text{sujeto a:} \quad & l_i^c \leq \sum_{j=0}^{n-1} a_{ij} x_j + \sum_{j=0}^{p-1} \bar{\boldsymbol{A}}_{ij} \bullet \bar{\boldsymbol{X}}_j \leq u_i^c, \quad i = 0, \ldots, m-1, \\
& l_i^x \leq x_j \leq u_i^x, \quad j = 0, \ldots, n-1, \\
& \boldsymbol{x} \in \mathbb{R}^n, \bar{\boldsymbol{X}}_j \succeq 0, j = 0, \ldots, p-1
\end{aligned}
$$

Problema primal en Mosek

Para llevar a esa forma un problema primal de la forma

$$\begin{aligned} \text{minimizar:} \quad & \boldsymbol{C} \bullet \boldsymbol{X} \\ \text{sujeto a:} \quad & \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i, \quad i = 0, \ldots, m-1, \\ & \boldsymbol{X} \succeq 0, \end{aligned}$$

observamos:

- p es la cantidad de matrices semidefinidas positivas $\boldsymbol{X}$. Por lo tanto, tenemos $p = 1$ y una única matrix $\bar{\boldsymbol{X}}_0$.
- El vector $\boldsymbol{x}$ es un vector de variables auxiliares que no aparecen en el problema primal, por lo tanto $n = 0$.

Problema primal en Mosek

Por lo tanto, el problema primal planteado en el formato de Mosek es

$$\begin{array}{ll} \text{minimizar:} & \bar{\boldsymbol{C}}_0 \bullet \bar{\boldsymbol{X}}_0 \\ \text{sujeto a:} & l_i^c \leq \bar{\boldsymbol{A}}_{i0} \bullet \bar{\boldsymbol{X}}_0 \leq u_i^c, \quad i = 0, \ldots, m-1, \\ & \bar{\boldsymbol{X}}_0 \succeq 0 \end{array}$$

donde $l_i^c = u_i^c = b_i$ para todo $i = 0, \ldots, m-1$.

Debemos ingresar en Mosek la matrix $\bar{\boldsymbol{C}}_0$, las matrices $\bar{\boldsymbol{A}}_{i0}, i = 0, \ldots, m-1$ y las cotas $l_i^c = u_i^c = b_i$, $i = 0, \ldots, m-1$.

Mínimo autovalor en Mosek

Para calcular el mínimo autovalor de una matriz $\boldsymbol{A} \in \mathcal{S}^n$, planteamos el problema

$$\begin{aligned} \text{maximizar: } & \alpha \\ \text{sujeto a: } \quad & \boldsymbol{A} - \alpha \boldsymbol{I} \succeq 0. \end{aligned}$$

Para llevarlo a la forma de Mosek, definimos $\boldsymbol{X} = \boldsymbol{A} - \alpha \boldsymbol{I}$. Por lo tanto $\boldsymbol{X} + \alpha \boldsymbol{I} = \boldsymbol{A}$ y planteamos el problema equivalente:

$$\begin{aligned} \text{maximizar: } & \alpha \\ \text{sujeto a: } \quad & \boldsymbol{X} + \alpha \boldsymbol{I} = \boldsymbol{A} \\ & \boldsymbol{X} \succeq 0. \end{aligned}$$

Observamos que las restricciones son restricciones lineales en las coordenadas de $\boldsymbol{X}$ y una variable adicional α y la función a maximizar depende de esta variable α.

Ejemplo: mínimo autovalor en Mosek

Veamos como plantear en Mosel el problema de calcular el menor autovalor de

$$\boldsymbol{A} = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}.$$

La ecuación $\boldsymbol{X} + \alpha \boldsymbol{I} = \boldsymbol{A}$ es

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} + \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}$$

y nos da las ecuaciones:

$$\begin{aligned} x_{11} + \alpha &= 3 \\ x_{12} &= 2 \\ x_{22} + \alpha &= 1 \end{aligned}$$

Ejemplo: mínimo autovalor en Mosek

Observamos

- Tenemos $p = 0$ y una única matrix $\boldsymbol{X}_0 = \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix}$.
- Tenemos una variable extra α, por lo tanto $n = 1$ y x_0 representa a α.
- Tenemos tres restricciones, $m = 3$.
- Las restricciones dependen linealmente de la variable x_0 y de las coordenadas de la matrix $\boldsymbol{X}_0$, por lo tanto podemos plantearlas en la forma

$$l \leq a_{i0}x_0 + \boldsymbol{A}_{i0} \bullet \boldsymbol{X}_0 \leq u.$$

Ejemplo: mínimo autovalor en Mosek

Obtenemos el problema

$$
\begin{aligned}
\text{maximizar: } & x_0 \\
\text{sujeto a: } \quad & 3 \leq x_0 + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \bullet \boldsymbol{X}_0 \leq 3 \\
& 2 \leq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \bullet \boldsymbol{X}_0 \leq 2 \\
& 1 \leq x_0 + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \bullet \boldsymbol{X}_0 \leq 1 \\
& \boldsymbol{X}_0 \succeq 0.
\end{aligned}
$$

Referencias

Fawzi, H. (2018).
Topics in convex optimisation (lecture notes).