

Optimización Semidefinida

Introducción

Segundo Cuatrimestre 2021

Facultad de Ciencias Exactas y Naturales, UBA

Organización de la materia

- Régimen de aprobación
- Inscripción en el SIU Guaraní
- Puntaje para licenciatura y doctorado
- Licencias de Mosek

Introducción: programación lineal

Programación lineal es el problema de minimizar o maximizar una función lineal sujeta a restricciones lineales.

Ejemplo. Queremos maximizar la función

$$f(x_1, x_2) = 4x_1 + 6x_2$$

sujeta a las restricciones

$$-x_1 + x_2 \leq 11,$$

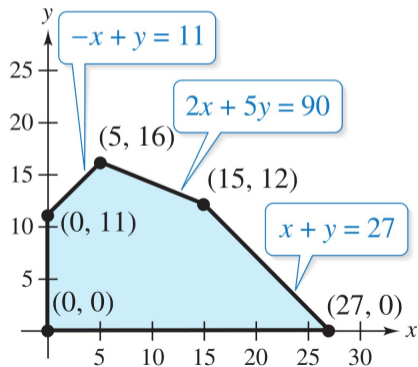
$$x_1 + x_2 \leq 27,$$

$$2x_1 + 5x_2 \leq 90,$$

$$x_1, x_2 \geq 0.$$

Introducción: programación lineal

Cada una de las desigualdades en las restricciones define un semiplano en \mathbb{R}^2 . Podemos graficar el conjunto de todos los puntos de \mathbb{R}^2 que cumplen todas las restricciones intersecando los semiplanos correspondientes.

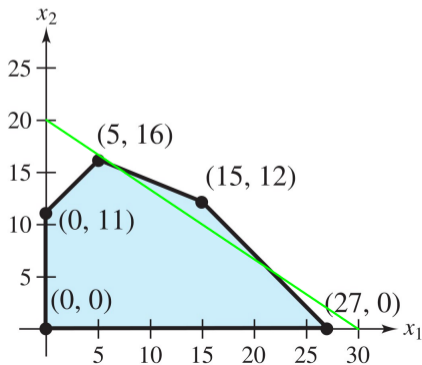


Introducción: programación lineal

El conjunto de puntos del plano para los cuales la función toma un valor fijo z_0 es una recta

$$l(z_0) = \{(x_1, x_2) : 4x_1 + 6x_2 = z_0\}.$$

Por ejemplo, si tomamos $z_0 = 120$ obtenemos la recta verde.



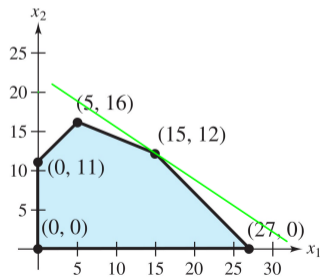
Introducción: programación lineal

Geoméricamente, si variamos el valor de z_0 , estamos desplazando la recta obteniendo siempre rectas paralelas.

En nuestro ejemplo, si desplazamos la recta hacia arriba, el valor de z_0 aumenta, mientras que si desplazamos la recta hacia abajo, el valor de z_0 disminuye.

Podemos resolver el problema gráficamente, desplazando la recta hacia arriba todo lo que podamos mientras que la intersección de la recta con la figura sea no vacía.

El máximo es $4 \times 15 + 6 \times 12 = 132$.



Introducción: programación lineal

Mediante esta resolución gráfica, podemos observar algunas propiedades del problema de programación lineal.

- Si la región de puntos que satisfacen las restricciones es acotada, forma un polígono, y el óptimo de la función a optimizar se alcanza en el borde del polígono.
- Más precisamente en un vértice del polígono (puede suceder que el óptimo se alcance también sobre todo un lado del polígono).
- En particular, dado que un polígono tiene una cantidad finita de vértices, el problema de programación lineal puede resolverse algorítmicamente evaluando la función objetivo sobre todos los vértices del polígono.

Introducción: programación semidefinida

Ejemplo. Dado el polinomio $f(x) = 10x^4 + 2x^3 + 27x^2 - 24x + 5$, queremos determinar si $f(x) \geq 0$ para todo $x \in \mathbb{R}$.

Propiedad. Un polinomio $f \in \mathbb{R}[x]$ de grado $2d$ es no-negativo para todo $x \in \mathbb{R}$ si y solo si f se puede escribir como una suma de cuadrados

$$f = p_1^2 + p_2^2 + \cdots + p_s^2,$$

con polinomios $p_i \in \mathbb{R}[x]$ de grado $\leq d$, $1 \leq i \leq s$.

Podemos escribir esta igualdad matricialmente:

$$f = \begin{pmatrix} p_1 & p_2 & \cdots & p_s \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_s \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & \cdots & p_s \end{pmatrix} \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & & 1 \\ 0 & & & \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_s \end{pmatrix}.$$

Introducción: programación semidefinida

Como cada $p_i \in \mathbb{R}[x]_d$,

$$p_i = \sum_{j=0}^d c_{ij} x^j$$

y obtenemos

$$\begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_s \end{pmatrix} = \begin{pmatrix} c_{10} & c_{11} & \cdots & c_{1d} \\ c_{20} & c_{21} & \cdots & c_{2d} \\ & & \ddots & \\ c_{s0} & c_{s1} & & c_{sd} \end{pmatrix} \begin{pmatrix} 1 \\ x \\ \vdots \\ x^d \end{pmatrix}.$$

Introducción: programación semidefinida

Concluimos que podemos escribir a f como

$$f = \begin{pmatrix} 1 & x & x^2 & \dots & x^d \end{pmatrix} \mathbf{C}^T \mathbf{C} \begin{pmatrix} 1 \\ x \\ \vdots \\ x^d \end{pmatrix},$$

donde $\mathbf{C} \in \mathbb{R}^{s \times (d+1)}$ y $\mathbf{C}^T \mathbf{C} \in \mathbb{R}^{(d+1) \times (d+1)}$ es una matriz semidefinida positiva.

Recíprocamente, si $f(x) = \mathbf{v}^T \mathbf{Q} \mathbf{v}$, con \mathbf{v} el vector de monomios y $\mathbf{Q} \in \mathbb{R}^{(d+1) \times (d+1)}$ semidefinida positiva, existe \mathbf{C} tal que $\mathbf{Q} = \mathbf{C}^T \mathbf{C}$ y podemos obtener una descomposición como suma de cuadrados.

Introducción: programación semidefinida

Volviendo al ejemplo, planteamos la ecuación:

$$10x^4 + 2x^3 + 27x^2 - 24x + 5 = \begin{pmatrix} 1 & x & x^2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix}$$

Para determinar si $f \geq 0$ para todo $x \in \mathbb{R}$ tenemos que determinar si existe $\mathbf{A} \in \mathbb{R}^{3 \times 3}$ tal que

$$a_{11} = 5, \quad 2a_{12} = -24, \quad 2a_{13} + a_{22} = 27, \quad 2a_{23} = 2, \quad a_{33} = 10, \\ \mathbf{A} \succeq 0$$

Introducción: programación semidefinida

Observamos que las primeras restricciones son ecuaciones lineales en los coeficientes de la matriz incógnita. En general podemos plantear un problema de programación semidefinida en la forma

$$\begin{aligned} \text{minimizar:} \quad & \sum_{1 \leq i, j \leq n} c_{ij} x_{ij} \\ \text{sujeto a:} \quad & \sum_{1 \leq i, j \leq n} a_{ij}^{(1)} x_{ij} = b^{(1)}, \dots, \sum_{1 \leq i, j \leq n} a_{ij}^{(s)} x_{ij} = b^{(s)} \\ & \mathbf{X} \succeq 0, \end{aligned}$$

donde \mathbf{C} , \mathbf{A}_i ($1 \leq i \leq m$) son matrices simétricas y la matrix simétrica \mathbf{X} es la variable sobre la cual realizamos la minimización.

Hiperplanos y semiespacios

Dados vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, definimos $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i$ el producto interno usual de \mathbb{R}^n .

Definición

Dado un vector no-nulo $\mathbf{a} \in \mathbb{R}^n$ y un escalar b ,

- 1 el conjunto $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \cdot \mathbf{x} = b\}$ es un *hiperplano*,
- 2 el conjunto $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \cdot \mathbf{x} \leq b\}$ es un *semiespacio*.

Dada una matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$ y un vector $\mathbf{b} \in \mathbb{R}^m$, en la condición $\mathbf{A}\mathbf{x} = \mathbf{b}$, cada fila \mathbf{a}_i de \mathbf{A} impone la restricción $\mathbf{a}_i \cdot \mathbf{x} = b_i$.

El conjunto $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}\}$ corresponde por lo tanto a una intersección de hiperplanos, que llamamos *espacio afín*.

El conjunto $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ corresponde a una intersección de semiespacios.

Poliedros y polítopos

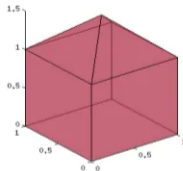
Llamamos *poliedro* a un conjunto definido por ecuaciones e inecuaciones lineales,

$$P = \{x \in \mathbb{R}^n \mid A_1x = b_1, A_2x \geq b_2\}.$$

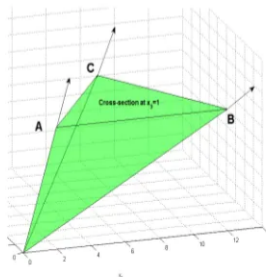
En el caso de que el conjunto resulte acotado, lo llamamos *polítopo*.

Según nuestra definición, un poliedro no necesariamente es un conjunto acotado.

Bounded- Polytope



Unbounded - polyhedron



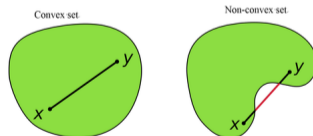
Conjuntos convexos

Definición

Un conjunto $S \subset \mathbb{R}^n$ es convexo si para todos $\mathbf{x}, \mathbf{y} \in S$, el segmento con vértices \mathbf{x} e \mathbf{y} está incluido en S . Es decir, para todo $\lambda \in [0, 1]$,

$$\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in S.$$

El punto $\mathbf{z} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{y}$, $\lambda \in [0, 1]$, es una *combinación convexa* de \mathbf{x} y \mathbf{y} .



Propiedad: La intersección de dos conjuntos convexos es un conjunto convexo.

El problema de programación lineal

Consideramos el problema de programación lineal

minimizar: $c \cdot x$

sujeto a: $Ax = b$

$x \geq 0$.

Podemos interpretar geoméricamente el conjunto factible (*feasible set* en inglés), es decir la región sobre la que queremos minimizar $c \cdot x$, como la intersección de un espacio afín (definido por la ecuación $Ax = b$) y el ortante positivo $x \geq 0$.

El conjunto factible es un poliedro (en particular es un conjunto convexo).

El problema de programación lineal

Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$f(\mathbf{x}) = \mathbf{c} \cdot \mathbf{x} = c_1x_1 + \cdots + c_nx_n,$$

$\mathbf{c} \in \mathbb{R}^n$, la llamamos *funcional lineal*.

El problema de programación lineal consiste en optimizar una funcional lineal sobre un poliedro.

Ejercicio

Dado el problema

$$\text{minimizar: } 3x_1 + 5x_2$$

$$\text{sujeto a: } x_1 + x_2 = 6$$

$$\mathbf{x} \geq 0,$$

graficar el conjunto factible, y resolver el problema.

Variables de holgura

Dado el problema de programación lineal

$$\text{minimizar: } \mathbf{c} \cdot \mathbf{x}$$

$$\text{sujeto a: } \mathbf{Ax} \leq \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0},$$

cada fila \mathbf{a}_i de \mathbf{A} impone una restricción $\mathbf{a}_i \cdot \mathbf{x} \leq b_i$.

Podemos convertir el problema en un problema con restricciones de igualdad reemplazando cada desigualdad $\mathbf{a}_i \cdot \mathbf{x} \leq b_i$ por las restricciones

$$\mathbf{a}_i \cdot \mathbf{x} + s_i = b_i, \quad s_i \geq 0,$$

donde s_i es una nueva variable del problema. Estas nuevas variables se llaman variables de holgura.

Variables de holgura

Obtenemos el problema equivalente con restricciones de igualdad

minimizar: $\mathbf{c} \cdot \mathbf{x}$

sujeto a: $\mathbf{a}_1 \cdot \mathbf{x} + s_1 = b_1$

...

$\mathbf{a}_m \cdot \mathbf{x} + s_m = b_m$

$\mathbf{x} \geq \mathbf{0}, \mathbf{s} \geq \mathbf{0},$

con $\mathbf{s} = (s_1, \dots, s_m)$.

Ejercicio

Dado el problema

$$\text{minimizar: } 3x_1 + 5x_2$$

$$\text{sujeto a: } x_1 \leq 1, x_2 \leq 1$$

$$\mathbf{x} \geq 0,$$

- 1 *graficar el conjunto factible,*
- 2 *convertirlo a un problema con igualdades agregando las variables de holgura necesarias,*
- 3 *calcular las coordenadas de los vértices del poliedro en el nuevo problema.*