

Aspectos geométricos y topológicos en Optimización Lineal

por

Fco. Javier Toledo Melero, Universidad Miguel Hernández

El objetivo de esta conferencia es proporcionar los conceptos y herramientas básicas necesarias para poder estudiar un problema de Programación Lineal Semi-Infinita y presentar algunos resultados relativos a este tipo de problemas. Puesto que el público al que está dirigida esta conferencia está esencialmente formado por estudiantes en Ciencias, se han ilustrado gráficamente una buena parte de las ideas geométricas que subyacen detrás de los conceptos y de los resultados, para facilitar una mejor comprensión de los mismos y, a su vez, para que se perciba la belleza de las matemáticas que se utilizan.

1. Introducción: tipos de problemas de optimización

Un *problema de optimización o de programación matemática*, o simplemente *programa matemático*, consiste en optimizar (maximizar o minimizar) una función f , llamada *función objetivo*, en cierto conjunto F , llamado *conjunto factible*, y se escribe

$$\begin{array}{ll} \text{Optimizar} & f(x) \\ \text{sujeto a} & x \in F \end{array} \quad (8.1)$$

En otras palabras, se trata de encontrar aquellos valores x de F que optimizan la función f . Un problema de este tipo se engloba dentro de los llamados *problemas de programación estática* en los que el conjunto F es un conjunto dado fijo. Existen otros tipos de problemas en los que, por ejemplo, las condiciones varían con el tiempo, *problemas de programación dinámica*, o aquellos en los que aparecen variables aleatorias, *problemas de programación estocástica*.

Es bastante habitual que el conjunto factible venga dado como el conjunto solución de ciertas ecuaciones $g(x) = 0$ y/o inecuaciones $h(x) \geq 0$, llamadas *restricciones*, esto es

$$F = \{x \mid g(x) = 0, h(x) \geq 0\}$$

en cuyo caso el problema (8.1) se escribe

$$\begin{array}{ll} \text{Optimizar} & f(x) \\ \text{sujeto a} & g(x) = 0; \\ & h(x) \geq 0. \end{array}$$

Si f es una función real definida en \mathbb{R}^n se tiene un sólo objetivo, $f(x)$, y si f toma valores en \mathbb{R}^p se tienen p objetivos, $(f_1(x), f_2(x), \dots, f_p(x))$. En este último caso se dice que el problema es de *programación multiobjetivo* o *programa vectorial*, y la primera cuestión que se plantea en tal caso es definir qué se entiende por óptimo.

Los programas con un solo objetivo se suelen clasificar según la naturaleza de la función objetivo y de las restricciones en:

- *Programas lineales (ordinarios)*: son aquellos en los que la función objetivo es lineal y las funciones que definen las restricciones son afines. Entenderemos, como es habitual, que un programa lineal ordinario¹ tiene una cantidad finita de restricciones.

- *Programas lineales semi-infinitos*: son programas lineales que tienen una cantidad arbitraria, posiblemente infinita, de restricciones. De esta forma, un programa lineal es un caso particular de un programa lineal semi-infinito.

- *Programas diferenciables*: tanto la función objetivo como las funciones que definen las restricciones son diferenciables.

- *Programas convexos*: la función objetivo es convexa (minimizar) o cóncava (maximizar) y el conjunto factible es convexo.

- *Programas enteros*: son aquellos en los que alguna o todas las variables están restringidas a tomar valores enteros. Si sólo alguna, pero no todas, de las variables toma valores enteros, se dice que el programa es *entero mixto*.

- *Programas no diferenciables*: la función objetivo o alguna de las funciones que definen las restricciones no son diferenciables.

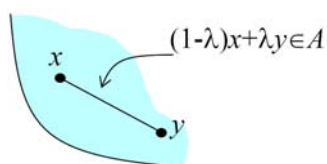
2. Herramientas básicas del análisis convexo

2.1 Conjuntos convexos, conos convexos, variedades afines

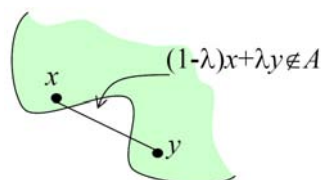
¹El término ordinario se suele poner cuando se quiere enfatizar el hecho de que estamos ante un programa lineal con una cantidad finita de restricciones.

Definición 2.1. Un conjunto $A \subset \mathbb{R}^n$ se dice *convexo* si para cualesquiera $x, y \in A$ y cualquier $0 \leq \lambda \leq 1$ se tiene que $(1 - \lambda)x + \lambda y \in A$.

Entenderemos que el conjunto vacío es convexo. Evidentemente \mathbb{R}^n es un conjunto convexo.



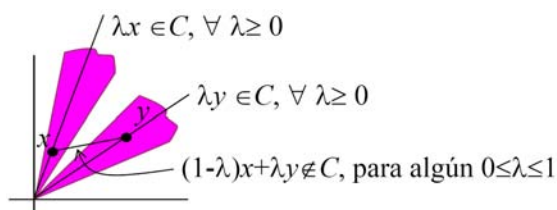
Conjunto convexo



Conjunto no convexo

Definición 2.2. Un conjunto $C \subset \mathbb{R}^n$ se dice que es un *cono* si para cualquier $x \in C$ y cualquier $\lambda \geq 0$ se tiene que $\lambda x \in C$.

Definición 2.3. Un conjunto $C \subset \mathbb{R}^n$ se dice que es un *cono convexo* si es un cono y además es un conjunto convexo.



Cono (no convexo)

Proposición 2.1. Se tiene que:

(i) Si $\{A_i\}_{i \in I} \subset \mathbb{R}^n$ es una colección arbitraria de conjuntos (resp. conos) convexos, entonces su intersección, $\bigcap_{i \in I} A_i$, también es un conjunto (resp. cono) convexo.

(ii) Si A es un conjunto (resp. cono) convexo, entonces su clausura, $cl(A)$, también es un conjunto (resp. cono) convexo.

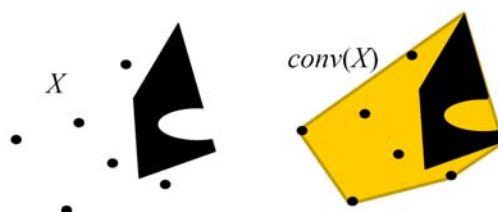
Definición 2.4. Un conjunto $A \subset \mathbb{R}^n$ se dice que es *afín* (también se dice que A es una *variedad afín* o *una variedad lineal*) si para cualesquiera $x, y \in A$ y cualquier $\lambda \in \mathbb{R}$ se tiene que $(1 - \lambda)x + \lambda y \in A$. Entenderemos que el conjunto vacío es afín.

Proposición 2.2. Un subconjunto de \mathbb{R}^n es afín si y sólo si es de la forma $a + L := \{a + l \mid l \in L\}$ donde $a \in \mathbb{R}^n$ y L es un subespacio vectorial de \mathbb{R}^n .

Es evidente que cualquier conjunto afín es convexo aunque el recíproco no es cierto en general.

Definición 2.5. La *envoltura convexa* de $X \subset \mathbb{R}^n$ es el menor subconjunto convexo de \mathbb{R}^n que contiene a X

$$\text{conv}(X) := \bigcap \{A \subset \mathbb{R}^n : A \text{ es convexo y } X \subset A\}.$$



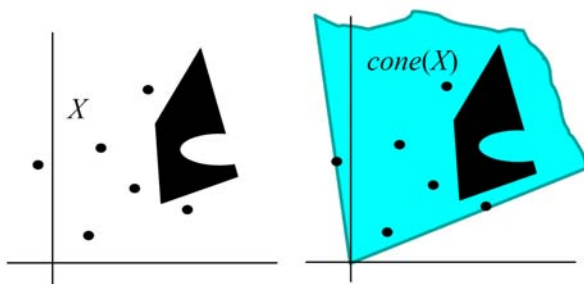
Proposición 2.3. Dado $X \subset \mathbb{R}^n$ se tiene que:

$$\text{conv}(X) = \left\{ \sum_{i=1}^k \lambda_i x^i : k \in \mathbb{N}, x^i \in X, \lambda_i \geq 0, i = 1, 2, \dots, k; \sum_{i=1}^k \lambda_i = 1 \right\}$$

Definición 2.6. La *envoltura cónica* de $X \subset \mathbb{R}^n$ es el menor cono convexo que contiene a X :

$$\text{cone}(X) := \bigcap \{C \subset \mathbb{R}^n : C \text{ es un cono convexo y } X \subset C\}.$$

Tomaremos como convenio que $\text{cone}(\emptyset) = \{0_n\}$.



Proposición 2.4. Dado $X \subset \mathbb{R}^n$, se tiene que:

$$\text{cone}(X) = \left\{ \sum_{i=1}^k \lambda_i x^i : k \in \mathbb{N}, x^i \in X, \lambda_i \geq 0, i = 1, 2, \dots, k \right\}.$$

Definición 2.7. Definimos por $\mathbb{R}_+^{(I)}$ al conjunto de funciones

$$\mathbb{R}_+^{(I)} := \{\lambda : I \rightarrow \mathbb{R}_+ : |\text{supp}(\lambda)| < +\infty\}$$

donde $\mathbb{R}_+ := [0, +\infty[$, y $supp(\lambda) := \{i \in I \mid \lambda(i) \neq 0\}$, es el *soporte de λ* . Denotamos $\lambda_i := \lambda(i)$.

Observación 2.1. Si $X := \{x_i, i \in I\} \subset \mathbb{R}^n$ para cierto conjunto de índices I , entonces un elemento de $conv(X)$ se puede escribir como $\sum_{i \in I} \lambda_i x_i$, para $\lambda \in \mathbb{R}_+^{(I)}$ con $\sum_{i \in I} \lambda_i = 1$. Análogamente, un elemento de $cone(X)$ se puede poner como $\sum_{i \in I} \lambda_i x_i$, para cierto $\lambda \in \mathbb{R}_+^{(I)}$.

Definición 2.8. Llamamos *envoltura afín* de $X \subset \mathbb{R}^n$ al menor subconjunto afín que contiene a X , es decir,

$$aff(X) := \bigcap \{A \subset \mathbb{R}^n : A \text{ es afín y } X \subset A\}.$$

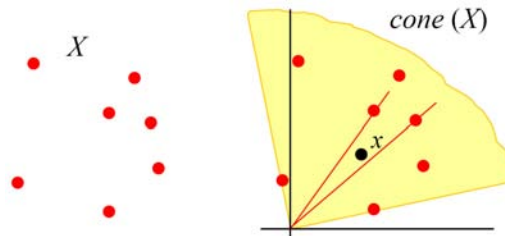
Análogamente a los resultados anteriores tenemos que:

Proposición 2.5. Dado $X \subset \mathbb{R}^n$, se tiene que:

$$aff(X) = \left\{ \sum_{i=1}^k \lambda_i x^i : k \in \mathbb{N}, x^i \in X, \lambda_i \in \mathbb{R}, i = 1, 2, \dots, k; \sum_{i=1}^k \lambda_i = 1 \right\}$$

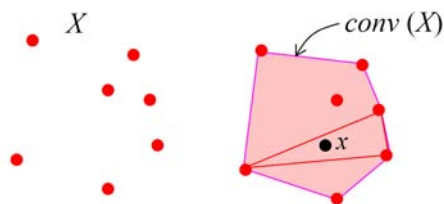
Teorema de Carathéodory para conos

Teorema 2.1. Si $X \subset \mathbb{R}^n$ y $x \in cone(X)$, existen $x^i \in X$ y $\lambda_i \geq 0, i = 1, 2, \dots, n$, tales que $x = \sum_{i=1}^n \lambda_i x^i$. Es decir, cualquier elemento de la envoltura cónica de X es combinación cónica de, a lo sumo, n elementos de X .



Teorema de Carathéodory para convexos

Teorema 2.2. Si $X \subset \mathbb{R}^n$ y $x \in conv(X)$, existen $x^i \in X$ y $\lambda_i \geq 0, i = 1, \dots, n + 1$, con $\sum_{i=1}^{n+1} \lambda_i = 1$, tales que $x = \sum_{i=1}^{n+1} \lambda_i x^i$. Es decir, cualquier elemento de la envoltura convexa de X es combinación convexa de, a lo sumo, $n + 1$ puntos de X .



Teorema de Mazur

Teorema 2.3. Si $X \subset \mathbb{R}^n$ es un conjunto compacto, entonces $\text{conv}(X)$ es también un conjunto compacto. En otras palabras, la envoltura convexa de un compacto es compacta.

La envoltura convexa de un conjunto cerrado puede no ser un cerrado.



Proposición 2.6. Todo cono convexo finitamente generado es cerrado. Es decir, si $x^1, x^2, \dots, x^k \in \mathbb{R}^n$, $k \in \mathbb{N}$, entonces $\text{cone}(\{x^1, x^2, \dots, x^k\})$ es cerrado.

Proposición 2.7. El cono convexo generado por un conjunto convexo compacto que no contiene al origen es cerrado.

La hipótesis de no contener al origen es esencial como muestra el siguiente ejemplo en el que X es convexo y compacto pero $\text{cone}(X)$ no es cerrado.



Cono polar (positivo) de un conjunto

Definición 2.9. El cono polar (positivo) de un conjunto no vacío $X \subset \mathbb{R}^n$ viene dado por

$$X^o := \{u \in \mathbb{R}^n \mid u'x \geq 0, \forall x \in X\},$$

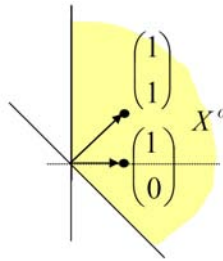
donde u' denota al traspuesto de u (entendido como vector-columna) y, por tanto, $u'x$ no es más que el producto escalar usual de u por x .

Proposición 2.8. Dado $X \subset \mathbb{R}^n$ se tiene que:

- (i) X^o es un cono convexo, cerrado y no vacío;
- (ii) $(\text{cone}(X))^o = X^o$;
- (iii) $X^{oo} := (X^o)^o = \text{cl}(\text{cone}(X))$.

Ejemplo 2.1. Sea $X = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$. Su cono polar viene dado por

$$X^o = \{(x_1, x_2)' \in \mathbb{R}^2 \mid x_1 + x_2 \geq 0, x_1 \geq 0\}.$$

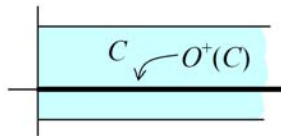


Cono de recesión de un conjunto

Definición 2.10. El *cono de recesión* de un conjunto $C \subset \mathbb{R}^n$ convexo y no vacío viene dado por:

$$O^+(C) := \{u \in \mathbb{R}^n \mid x + \lambda u \in C, \forall x \in C, \forall \lambda \geq 0\}$$

Si se supone además que C es cerrado, se puede reemplazar en la definición “ $\forall x \in C$ ” por “para algún $x \in C$ ”.



Proposición 2.9. Si $C \subset \mathbb{R}^n$ es un conjunto convexo cerrado y no vacío entonces $O^+(C)$ es un cono convexo y cerrado.

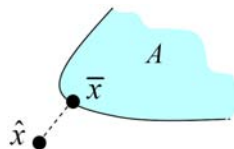
Proposición 2.10. Si $C \subset \mathbb{R}^n$ es un cono convexo entonces $O^+(C) = C$.

2.2 Mejor aproximación de un punto en un cerrado

Consideremos en \mathbb{R}^n una norma arbitraria, $\|\cdot\|$.

Definición 2.11. Sea $A \subset \mathbb{R}^n$ un conjunto cerrado no vacío y sea $\hat{x} \notin A$. Se dice que $\bar{x} \in A$ es una mejor aproximación de \hat{x} en A si

$$\|\bar{x} - \hat{x}\| = d(\hat{x}, A) := \inf \{\|x - \hat{x}\| : x \in A\}$$



Proposición 2.11. Sea $A \subset \mathbb{R}^n$ un conjunto cerrado y no vacío y sea $\hat{x} \notin A$. Entonces existe al menos una mejor aproximación de \hat{x} en A .

Observación 2.2. Si en la proposición anterior se supone además que A es convexo, y la norma considerada en \mathbb{R}^n es la norma Euclídea, puede probarse que la mejor aproximación de \hat{x} en A es única.

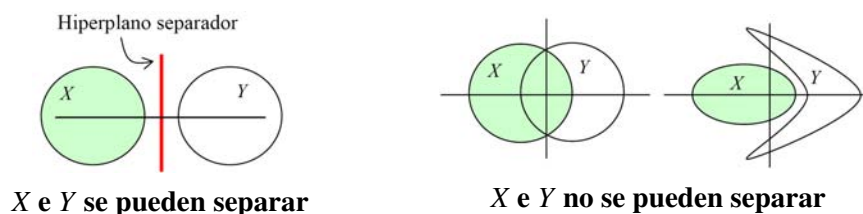
2.3 Separación en \mathbb{R}^n (mediante hiperplanos)

Nociones de separación

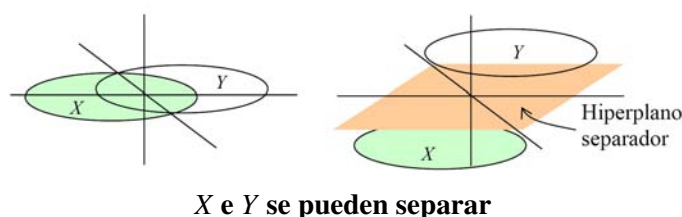
Definición 2.12. Sean X e Y dos subconjuntos no vacíos de \mathbb{R}^n . Se dice que X e Y se pueden *separar* si existe $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^{n+1}$, con $a \neq 0_n$, tal que

$$a'x \leq b \leq a'y, \quad \forall x \in X, \forall y \in Y.$$

En tal caso, se dice que X e Y están separados por el hiperplano de ecuación $a'z = b$, que se llama *hiperplano separador* de X e Y .

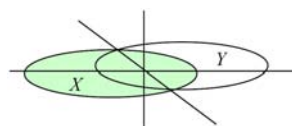


Observación 2.3. Informalmente hablando, el hiperplano separador divide el espacio \mathbb{R}^n en dos semiespacios, en uno de estos semiespacios estarían los puntos de X y en el otro los puntos de Y . Podría darse el caso particular en el que tanto los puntos de X como los de Y estuvieran todos contenidos en el hiperplano separador, es decir, $a'x = b = a'y$, $\forall x \in X, \forall y \in Y$, en cuyo caso, la noción de separación anterior no se ajustaría a nuestra intuición geométrica de separación.



Para distinguir este tipo de casos se da la siguiente definición.

Definición 2.13. Se dice que X e Y se pueden *separar propiamente* si se pueden separar de forma que no estén contenidos ambos en el hiperplano separador.

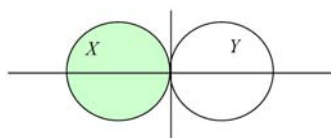


X e Y no se pueden separar propiamente

Definición 2.14. Se dice que X e Y se pueden *separar estrictamente* si existe $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^{n+1}$, con $a \neq 0_n$, tal que

$$a'x < b < a'y, \quad \forall x \in X, \forall y \in Y.$$

Observación 2.4. Evidentemente la separación estricta implica la separación propia. El recíproco no es cierto en general.

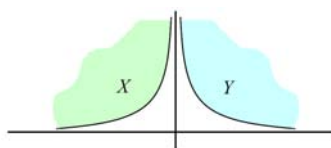


X e Y se pueden separar propiamente pero no estrictamente

Definición 2.15. X e Y se pueden *separar fuertemente* si existe $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^{n+1}$ tal que

$$\sup \{a'x : x \in X\} < \inf \{a'y : y \in Y\}.$$

Observación 2.5. Evidentemente la separación fuerte implica la separación estricta. El recíproco no es cierto en general.



X e Y se pueden separar estrictamente pero no fuertemente

Un caso particular aparece cuando el conjunto X es un conjunto compacto de \mathbb{R}^n , por ejemplo, cuando X se reduce a un punto. En tal caso, los conceptos de separación estricta y fuerte coinciden como muestra el siguiente resultado.

Proposición 2.12. Sea $X \subset \mathbb{R}^n$ compacto no vacío e $Y \subset \mathbb{R}^n$ no vacío. Son equivalentes:

- (i) X e Y se pueden separar fuertemente;
- (ii) X e Y se pueden separar estrictamente;
- (iii) Existe $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^{n+1}$, con $a \neq 0_n$, tal que

$$a'x < b \leq a'y, \quad \forall x \in X, \forall y \in Y. \quad (8.2)$$

Teoremas de separación

- Separación de un punto y un convexo cerrado

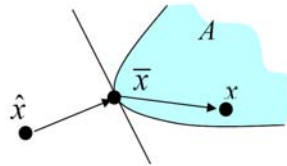
Teorema 2.4. Sea $A \subset \mathbb{R}^n$ un conjunto convexo cerrado y no vacío, y sea $\hat{x} \notin A$. Entonces existe $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^{n+1}$, con $a \neq 0_n$, tal que

$$a' \hat{x} < b \leq a' x, \forall x \in A, \quad (8.3)$$

es decir, $\{\hat{x}\}$ y A se pueden separar fuertemente. Es más, si $\bar{x} \in A$ es una mejor aproximación con la norma Euclídea de \hat{x} en A , se tiene que $a = \bar{x} - \hat{x}$ y $b = (\bar{x} - \hat{x})' \bar{x}$ satisfacen (8.3), es decir,

$$(\bar{x} - \hat{x})' \hat{x} < (\bar{x} - \hat{x})' \bar{x} \leq (\bar{x} - \hat{x})' x, \forall x \in A.$$

Observación 2.6. El hiperplano separador $(\bar{x} - \hat{x})' z = (\bar{x} - \hat{x})' \bar{x}$ es de hecho un hiperplano soporte de A en \bar{x} (esto es, es un hiperplano que contiene a \bar{x} y deja a A en uno de los semiespacios que define)



- Separación de un punto y un cono convexo cerrado

Teorema 2.15. Sea $C \subset \mathbb{R}^n$ un cono convexo cerrado y $\hat{x} \notin C$. Entonces existe $a \in \mathbb{R}^n$, $a \neq 0_n$, tal que

$$a' \hat{x} < 0 \leq a' x, \forall x \in C. \quad (8.4)$$

Es más, si $\bar{x} \in C$ es la mejor aproximación con la norma Euclídea de \hat{x} en C , se tiene que $a = \bar{x} - \hat{x}$ satisface (8.4), es decir,

$$(\bar{x} - \hat{x})' \hat{x} < 0 \leq (\bar{x} - \hat{x})' x, \forall x \in C.$$

3. Introducción a la programación lineal semi-infinita

3.1 El contexto de la programación lineal semi-infinita

Consideremos el problema de optimización

$$\begin{aligned} \pi : \text{Inf } & c'x \\ \text{s.a. } & a'_t x \geq b_t, \quad t \in T \end{aligned} \quad (8.5)$$

donde $a_t, c, x \in \mathbb{R}^n, b_t \in \mathbb{R}$. El conjunto de índices T es arbitrario, posiblemente infinito, luego el problema π es un problema de *programación lineal semi-infinita* (PLSI). Si T es finito, π no es más que un *problema de programación lineal* (PL) ordinaria.

El espacio paramétrico de todos los problemas en la forma (8.5) lo denotaremos por Π .

Denotamos por σ al sistema de restricciones del problema π , i.e.

$$\sigma = \{a'_t x \geq b_t, t \in T\}, \quad (8.6)$$

y lo identificamos con sus vectores de coeficientes, $\sigma \equiv \left\{ \begin{pmatrix} a_t \\ b_t \end{pmatrix} \right\}_{t \in T}$. De esta manera, podemos representar un sistema como un conjunto de puntos en $\mathbb{R}^n \times \mathbb{R}$. En otras palabras, si denotamos por Θ el espacio paramétrico de todos los sistemas en la forma (8.6), estamos identificando Θ con $(\mathbb{R}^n \times \mathbb{R})^T$, esto es, el espacio de todas las funciones de T en $\mathbb{R}^n \times \mathbb{R}$.

Asímismo, identificamos $\pi \equiv (c, \sigma)$ de manera que $\Pi \equiv \mathbb{R}^n \times (\mathbb{R}^n \times \mathbb{R})^T$.

Ejemplo 3.1. Aplicación a la aproximación uniforme de funciones

Tratamos de encontrar una función polinómica, p , de grado menor que n , que esté “lo más próxima posible” a una función dada, f , en $[a, b]$.

Si consideramos en el espacio, X , de todas las funciones $g : [a, b] \rightarrow \mathbb{R}$ la distancia extendida $d : X \times X \rightarrow [0, +\infty]$ dada por

$$d(g_1, g_2) := \sup_{t \in [a, b]} |g_1(t) - g_2(t)|,$$

el problema que nos planteamos se puede formular como un problema de optimización en la forma

$$\begin{aligned} \pi : \text{Inf} \quad & d(p, f) \\ \text{s.a} \quad & p(t) = x_1 + x_2 t + \cdots + x_n t^{n-1}, \\ & x_1, x_2, x_3, \dots, x_n \in \mathbb{R}, \end{aligned}$$

o equivalentemente

$$\begin{aligned} \pi : \text{Inf} \quad & x_{n+1} \\ \text{s.a} \quad & \sup_{t \in [a, b]} |x_1 + x_2 t + \cdots + x_n t^{n-1} - f(t)| \leq x_{n+1}, \end{aligned}$$

el cual se puede escribir como el problema de PLSI

$$\begin{aligned} \pi : \text{Inf} \quad & x_{n+1} \\ \text{s.a} \quad & x_1 + x_2 t + \cdots + x_n t^{n-1} + x_{n+1} \geq f(t), t \in [a, b], \\ & -x_1 - x_2 t - \cdots - x_n t^{n-1} + x_{n+1} \geq -f(t), t \in [a, b]. \end{aligned}$$

3.2 Conjunto factible, conjunto de nivel inferior, valor óptimo y conjunto óptimo

Definición 3.1. Dado el problema $\pi \equiv (c, \sigma)$ definimos:

- El *conjunto factible* de π (y de σ) como

$$F := \{x \in \mathbb{R}^n : a'_t x \geq b_t \ \forall t \in T\}$$

- El *conjunto de nivel inferior* α ($\alpha \in \mathbb{R}$) de π como

$$\begin{aligned} L(\alpha) &:= \{x \in F : c'x \leq \alpha\} \\ &= \{x \in \mathbb{R}^n : a'_t x \geq b_t \ \forall t \in T, \ -c'x \geq -\alpha\} \end{aligned}$$

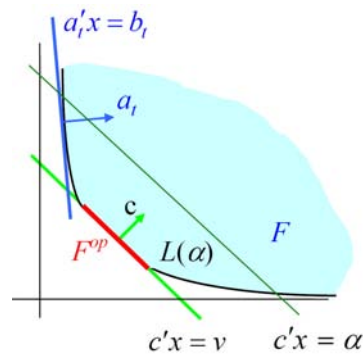
- El *valor óptimo* de π como

$$\begin{aligned} v &:= \inf \{c'x : x \in F\} \in [-\infty, +\infty] \\ v &:= +\infty \text{ si } F = \emptyset; \\ v &:= -\infty \text{ si } x \mapsto c'x \text{ no está acotada inferiormente en } F \end{aligned}$$

- El *conjunto óptimo* de π como

$$F^{op} := \{x \in F : c'x = v\} = L(v)$$

En la siguiente figura se ilustran los elementos anteriores.



Observación 3.1. Se tiene que F y F^{op} son conjuntos convexos y cerrados (posiblemente vacíos) de \mathbb{R}^n .

3.3 Representación de un conjunto convexo cerrado como conjunto factible de un sistema semi-infinito

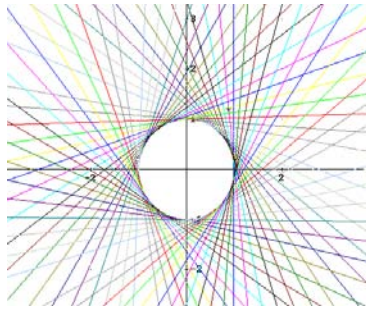
Teorema 3.1. Sea $C \subset \mathbb{R}^n$ un conjunto convexo, cerrado y no vacío. Entonces existe un sistema $\sigma = \left\{ a'_t x \geq b_t, \begin{pmatrix} a \\ b \end{pmatrix} \in S \right\}$, para algún $S \subset \mathbb{R}^{n+1}$, tal que C es el conjunto factible de σ .

Observación 3.2. La representación de un convexo cerrado como conjunto de soluciones de un sistema de desigualdades lineales (semi-infinito) no es única.

Ejemplo 3.2. Consideremos el círculo $C := \{(x_1, x_2)' \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$. Tenemos, entre otras, las siguientes representaciones del mismo:

$$C = \{(x_1, x_2)' \in \mathbb{R}^2 \mid -\cos(t)x_1 - \sin(t)x_2 \geq -1, t \in [0, 2\pi]\}$$

$$C = \left\{ (x_1, x_2)' \in \mathbb{R}^2 \mid \begin{array}{l} -tx_1 - \sqrt{1-t^2}x_2 \geq -1, t \in [-1, 1] \\ -(t-2)x_1 + \sqrt{1-(t-2)^2}x_2 \geq -1, t \in [1, 3] \end{array} \right\}$$



3.4 Definiciones de consistencia, acotación y resolubilidad

Definición 3.2 Distinguimos en Θ los siguientes subconjuntos:

$\Theta_c := \{\sigma \in \Theta \mid F \neq \emptyset\}$ subconjunto de los *sistemas consistentes*

$\Theta_i := \Theta \setminus \Theta_c$ subconjunto de los *sistemas inconsistentes*

El subconjunto de los sistemas inconsistentes se divide a su vez en:

$\Theta_{wi} := \{\sigma \in \Theta_i \mid \text{todo subsistema finito de } \sigma \text{ es consistente}\}$ subconjunto de los *sistemas débilmente inconsistentes* (en PL ordinaria no hay sistemas débilmente inconsistentes).

$\Theta_{si} := \{\sigma \in \Theta_i \mid \text{algún subsistema finito de } \sigma \text{ es inconsistente}\}$ subconjunto de los *sistemas fuertemente inconsistentes* (en PL ordinaria $\Theta_{si} = \Theta_i$)

Ejemplo 3.3. El sistema $\sigma = \left\{ \frac{1}{r}x \geq 1, r \in \mathbb{N} \right\}$ es débilmente inconsistente. El sistema $\sigma = \{x \geq 1, -x \geq 1\}$ es fuertemente inconsistente.

Definición 3.3. Distinguimos en Π los siguientes subconjuntos:

$\Pi_c := \{\pi = (c, \sigma) \in \Pi \mid \sigma \in \Theta_c\}$ subconjunto de los *problemas consistentes*

$\Pi_i := \Pi \setminus \Pi_c$ subconjunto de los *problemas inconsistentes*

$\Pi_b := \{\pi \in \Pi \mid -\infty < v < +\infty\}$ subconjunto de los *problemas acotados*

$\Pi_s := \{\pi \in \Pi \mid F^{op} \neq \emptyset\}$ subconjunto de los *problemas resolubles*

Observación 3.3. En el caso particular de la programación lineal ordinaria se tiene que $\Pi_s = \Pi_b$.

En el contexto más general de la programación lineal semi-infinita se tiene que $\Pi_s \subset \Pi_b \subset \Pi_c$. Sin embargo $\Pi_b \not\subset \Pi_s$ como se puede ver en el siguiente ejemplo.

Ejemplo 3.4. Consideremos el problema de PLSI

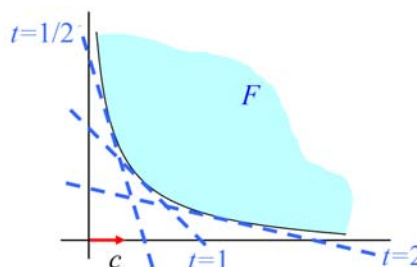
$$\begin{aligned} \pi : \text{Inf } x_1 \\ \text{s.a. } x_1 + t^2 x_2 \geq 2t, t \in]0, +\infty[\end{aligned}$$

Se tiene que:

$$F = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_2 \geq \frac{1}{x_1}, x_1 > 0 \right\}$$

$$v = 0 \Rightarrow \pi \in \Pi_b$$

$$F^{op} = \emptyset \Rightarrow \pi \notin \Pi_s$$



3.5 Conjuntos convexos asociados a un problema de PLSI

Definición 3.4 Asociados a un problema $\pi \equiv (c, \sigma)$, se definen los siguientes conjuntos:

$$A := \text{conv}(\{a_t, t \in T\}); \quad M := \text{cone}(\{a_t, t \in T\}) = \mathbb{R}_+ A;$$

$$Z^- := \text{conv}(\{a_t, t \in T; -c\}); \quad Z^+ := \text{conv}(\{a_t, t \in T; c\});$$

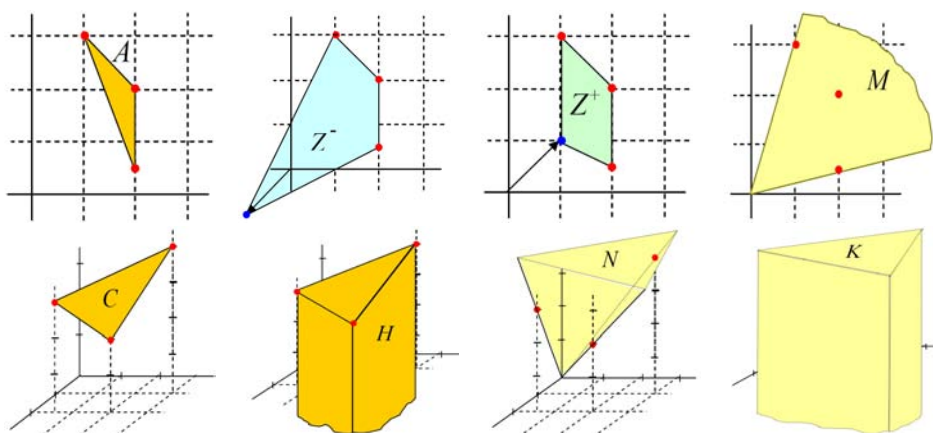
$$C := \text{conv}\left(\left\{\begin{pmatrix} a_t \\ b_t \end{pmatrix}, t \in T\right\}\right); \quad N := \text{cone}\left(\left\{\begin{pmatrix} a_t \\ b_t \end{pmatrix}, t \in T\right\}\right);$$

$$H := C + \mathbb{R}_+ \left\{\begin{pmatrix} 0_n \\ -1 \end{pmatrix}\right\}; \quad K := N + \mathbb{R}_+ \left\{\begin{pmatrix} 0_n \\ -1 \end{pmatrix}\right\}.$$

M y N se llaman *primer* y *segundo cono de momentos*, respectivamente. K se llama *cono característico* y H *conjunto hipográfico*.

Ejemplo 3.4. En este ejemplo ilustramos los conjuntos anteriores asociados al problema:

$$\begin{aligned} \pi : \text{Inf } x_1 + x_2 \\ \text{s.a. } 2x_1 + \frac{1}{2}x_2 \geq 3, \quad 2x_1 + 2x_2 \geq 2, \quad x_1 + 3x_2 \geq 4. \end{aligned}$$

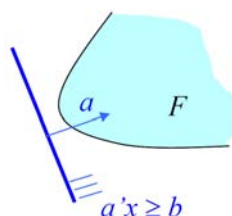


3.6 Relaciones consecuentes de un sistema consistente. Lema de Farkas

Definición 3.5. Se dice que $a'x \geq b$ es una *relación consecuyente* o una *consecuencia* de un sistema consistente $\sigma = \{a'_t x \geq b_t, t \in T\}$ si $a'y \geq b$ para todo $y \in F$.

Teorema 3.2. (Lema de Farkas) $a'x \geq b$ es una consecuencia de un sistema consistente $\sigma = \{a'_t x \geq b_t, t \in T\}$ si y sólo si $\begin{pmatrix} a \\ b \end{pmatrix} \in cl(K)$.

Por esto, $cl(K)$ también se llama *cono de las relaciones consecuentes* de σ .



Proposición 3.1. Sean $\sigma, \sigma_1 \in \Theta_c$, son equivalentes:

- (i) $F = F_1$;
- (ii) $cl(K) = cl(K_1)$;
- (iii) $cl(M) = cl(M_1)$.

3.7 Caracterización de la consistencia de sistemas de desigualdades lineales

Teorema 3.3. Dado $\sigma = \{a'_t x \geq b_t, t \in T\}$, se tiene que:

- (i) $\sigma \in \Theta_c \Leftrightarrow \begin{pmatrix} 0_n \\ 1 \end{pmatrix} \notin cl(K)$;
- (ii) $\sigma \in \Theta_{si} \Leftrightarrow \begin{pmatrix} 0_n \\ 1 \end{pmatrix} \in K$;
- (iii) $\sigma \in \Theta_{wi} \Leftrightarrow \begin{pmatrix} 0_n \\ 1 \end{pmatrix} \in cl(K) \setminus K$.

Para ilustrar como se utilizan las herramientas proporcionadas anteriormente, damos a continuación la demostración del teorema anterior.

Demostración (i) “ \Rightarrow ” Sea σ consistente y supongamos, razonando por reducción al absurdo, que $\begin{pmatrix} 0_n \\ 1 \end{pmatrix} \in cl(K)$; esto es, que existen $\{\lambda^r\} \subset \mathbb{R}_+^{(T)}$ y $\{\mu^r\} \subset \mathbb{R}_+$ tales que $\begin{pmatrix} 0_n \\ 1 \end{pmatrix} = \lim_r \left\{ \sum_{t \in T} \lambda_t^r \begin{pmatrix} a_t \\ b_t \end{pmatrix} + \mu^r \begin{pmatrix} 0_n \\ -1 \end{pmatrix} \right\}$. Sea $\bar{x} \in F$. Si multiplicamos en la última expresión por $\begin{pmatrix} \bar{x} \\ -1 \end{pmatrix}$ se obtiene que

$$-1 = \lim_r \left\{ \sum_{t \in T} \lambda_t^r \begin{pmatrix} a_t \\ b_t \end{pmatrix}' \begin{pmatrix} \bar{x} \\ -1 \end{pmatrix} + \mu^r \right\} = \lim_r \left\{ \sum_{t \in T} \lambda_t^r (a_t' \bar{x} - b_t) + \mu^r \right\} \geq 0,$$

que es absurdo.

(i) “ \Leftarrow ” Supongamos ahora que $\begin{pmatrix} 0_n \\ 1 \end{pmatrix} \notin cl(K)$ y veamos que σ es consistente. En efecto, por el teorema de separación de un punto y un cono convexo cerrado se tiene que existe $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^{n+1}$ tal que

$$\begin{pmatrix} u \\ v \end{pmatrix}' \begin{pmatrix} 0_n \\ 1 \end{pmatrix} < 0 \quad \text{y} \quad \begin{pmatrix} u \\ v \end{pmatrix}' \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \geq 0, \quad \forall \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in cl(K),$$

de donde obtenemos que $v < 0$ y, como $\begin{pmatrix} a_t \\ b_t \end{pmatrix} \in cl(K) \forall t \in T$, $\begin{pmatrix} u \\ v \end{pmatrix}' \begin{pmatrix} a_t \\ b_t \end{pmatrix} \geq 0 \forall t \in T$, es decir, $a_t' u + b_t v \geq 0 \forall t \in T$. Dividiendo por $-v > 0$ en esta última expresión obtenemos que $\frac{-1}{v} u \in F$, y por tanto σ es consistente.

(ii) “ \Rightarrow ” Supongamos que σ es fuertemente inconsistente, entonces el sistema $\{a_i' x \geq b_i, i = 1, 2, \dots, m\}$ es inconsistente para algún $\{t_1, t_2, \dots, t_m\} \subset T$, $m \in \mathbb{N}$. Aplicando el apartado (i) se tiene entonces, teniendo en cuenta que todo cono finitamente generado es cerrado, que

$$\begin{aligned} \begin{pmatrix} 0_n \\ 1 \end{pmatrix} &\in cl \left(\text{cone} \left\{ \begin{pmatrix} a_{t_i} \\ b_{t_i} \end{pmatrix}, i = 1, 2, \dots, m; \begin{pmatrix} 0_n \\ -1 \end{pmatrix} \right\} \right) = \\ &= \text{cone} \left\{ \begin{pmatrix} a_{t_i} \\ b_{t_i} \end{pmatrix}; i = 1, 2, \dots, m; \begin{pmatrix} 0_n \\ -1 \end{pmatrix} \right\} \subset K \end{aligned}$$

(ii) “ \Leftarrow ” Si $\begin{pmatrix} 0_n \\ 1 \end{pmatrix} \in K$, entonces existen $\lambda \in \mathbb{R}_+^{(T)}$, $\mu \geq 0$ tales que

$$\begin{pmatrix} 0_n \\ 1 \end{pmatrix} = \sum_{t \in T} \lambda_t \begin{pmatrix} a_t \\ b_t \end{pmatrix} + \mu \begin{pmatrix} 0_n \\ -1 \end{pmatrix}.$$

Sea $\{t_1, t_2, \dots, t_m\} = \text{supp}(\lambda)$ (recordemos que $|\text{supp}(\lambda)| < +\infty$), entonces

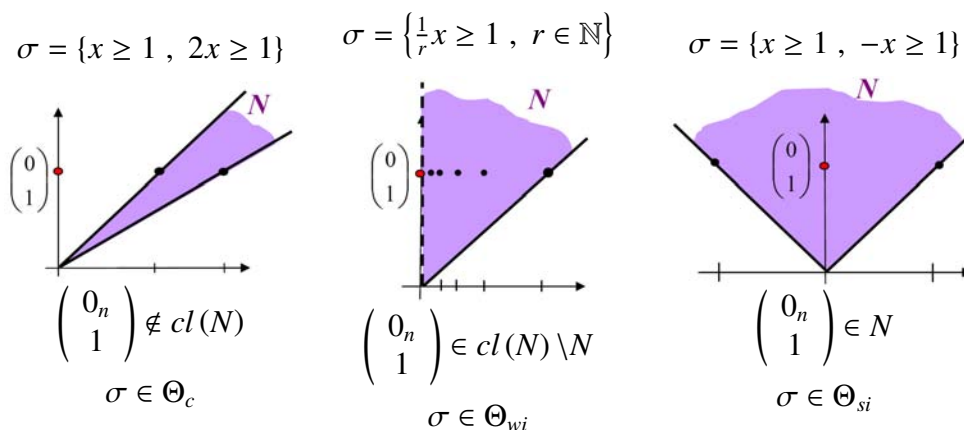
$$\begin{pmatrix} 0_n \\ 1 \end{pmatrix} = \sum_{i=1}^m \lambda_{t_i} \begin{pmatrix} a_{t_i} \\ b_{t_i} \end{pmatrix} + \mu \begin{pmatrix} 0_n \\ -1 \end{pmatrix} \in \text{cone} \left\{ \begin{pmatrix} a_{t_i} \\ b_{t_i} \end{pmatrix}, i = 1, 2, \dots, m; \begin{pmatrix} 0_n \\ -1 \end{pmatrix} \right\} \subset K$$

luego el subsistema finito $\{a'_i x \geq b_{t_i}, i = 1, 2, \dots, m\}$ es inconsistente.

(iii) Es inmediato, pues σ es débilmente inconsistente si es inconsistente y no es fuertemente inconsistente, es decir, en virtud de los apartados (i) y (ii) si $\begin{pmatrix} 0_n \\ 1 \end{pmatrix} \in \text{cl}(K)$ y $\begin{pmatrix} 0_n \\ 1 \end{pmatrix} \notin K$, respectivamente. \diamond

Observación 3.4. En el Teorema 3.3 se puede reemplazar el cono K por el cono N .

Ejemplo 3.6. Los siguientes ejemplos ilustran el resultado anterior



3.8 Caracterización de la acotación del conjunto factible y del conjunto óptimo

Teorema 3.4. Si $\sigma = \{a'_t x \geq b_t, t \in T\} \in \Theta_c$, las siguientes condiciones son equivalentes:

- (i) F está acotado;
- (ii) $M^o = \{0_n\}$;
- (iii) $0_n \in \text{int}(M)$.

Teorema 3.5. Sea $\pi = (c, \sigma) \in \Pi_c$, las siguientes condiciones son equivalentes:

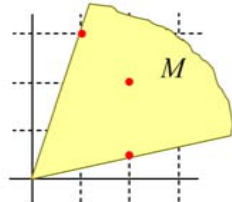
- (i) F^{op} está acotado y no vacío;
- (ii) $(Z^-)^o = \{0_n\}$;
- (iii) $0_n \in \text{int}(Z^-)$.

En el siguiente ejemplo se ilustran los dos teoremas anteriores.

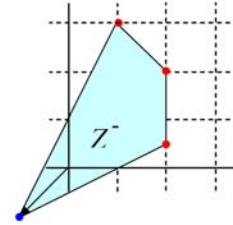
Ejemplo 3.7. Consideremos el mismo problema que en el Ejemplo 3.5

$$\pi : \text{Inf } x_1 + x_2$$

$$\text{s.a } 2x_1 + \frac{1}{2}x_2 \geq 3, \quad 2x_1 + 2x_2 \geq 2, \quad x_1 + 3x_2 \geq 4.$$

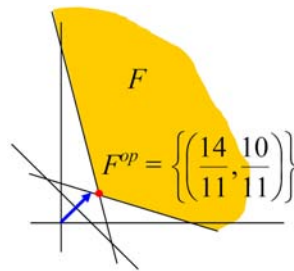


$$0_n \notin \text{int}(M)$$



$$0_n \in \text{int}(Z^-)$$

Por el Teorema 3.4 tenemos que F no está acotado, y por el Teorema 3.5 tenemos que F^{op} sí que lo está. El siguiente gráfico muestra lo anterior.



Observación final

Mi intención inicial en esta conferencia era contar algunas cosas más acerca de estabilidad, distancia al mal-planteamiento, medidas de la variación de F , v y F^{op} respecto de los parámetros, etc., en el contexto de la PLSI (véanse, entre otros, [1], [2] y [3], donde se utilizan todos los conjuntos descritos anteriormente). La verdad es que la mayoría de las veces, nos ocurre que queremos contar más cosas de las que el tiempo nos permite y se acaba desbordando al público con un exceso de información. Puesto que mi objetivo principal era el de presentar un problema de PLSI junto con las herramientas que se utilizan y algunos resultados acerca de este tipo de problemas, creo que ese objetivo se ha cumplido. No obstante, para aquellos que hayan quedado interesados en este tópico de las matemáticas (espero que muchos), les propongo a continuación una serie de referencias en las que, tomando como punto de partida lo expuesto en esta charla, podrán profundizar hondamente en el tema. Concretamente, véanse, entre otros, [6] y [7] para la teoría y las aplicaciones del análisis convexo, [5] para las técnicas y métodos de programación matemática más usuales y [4] como manual indispensable para abordar un problema de programación lineal semi-infinita.

Por último, quiero agradecer a Marta y a Raúl la oportunidad que me han brindado de contar, en este ciclo de conferencias, este tema con el que tanto disfruto.

Bibliografía

- [1] M.J. Cánovas, M.A. López, J. Parra and F.J. Toledo, *Distance to ill-posedness and the consistency value of linear semi-infinite inequality systems*, Math. Program. 103A, 95-126, 2005.
- [2] M.J. Cánovas, M.A. López, J. Parra and F.J. Toledo, *Distance to solvability/unsolvability in linear optimization*, SIAM J. Optim. 16(3), 629-649, 2006.
- [3] M.J. Cánovas, M.A. López, J. Parra and F.J. Toledo, *Ill-posedness with respect to the solvability in linear optimization*, Linear Algebra and Appl. 416, 520-540, 2006.
- [4] M.A. Goberna and M.A. López, *Linear Semi-Infinite Optimization*, John Wiley and Sons, 1998.
- [5] V. Novo, *Teoría de la optimización*, Aula Abierta, UNED, 2000.
- [6] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
- [7] J.v. Tiel, *Convex Analysis*, John Wiley and Sons, 1984.

Fco. Javier Toledo Melero
Universidad Miguel Hernández de Elche
Centro de Investigación Operativa
Avenida de la Universidad s/n,
03202 Elche (Alicante)
e-mail: javier.toledo@umh.es

