Auxiliar 3: Geometría!

Best auxiliar para comprender geometría, no verán nada mejor :D

Profesores: Gonzalo Muñoz y Daniel Rossi
Auxiliares: Felipe Fierro, Felipe Hueitra, Anais Muñoz, Leonardo Navarro, Jimmy Pirul

Pregunta 1 - SBF y optimalidad

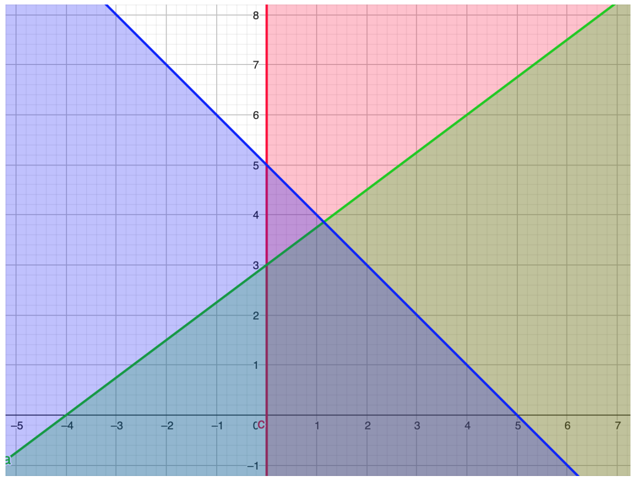
Considere el siguiente problema de programación lineal:

$$\begin{aligned} \max \quad & x + 2y \\ \text{s.a.} \quad & -3x + 4y \leq 12 \\ & 3x + 3y \leq 15 \\ & x \geq 0 \end{aligned}$$

1. Bosqueje el poliedro asociado al PL.
2. Determine las soluciones básicas. ¿Cuáles de estas son factibles? ¿Cuál es el óptimo?
3. Proponga un vector de costos para cada caso:
   a) La solución óptima es única y distinta al problema original.
   b) Existen múltiples soluciones óptimas y el conjunto de estas es acotado.
   c) Existen múltiples soluciones óptimas y el conjunto de estas es no acotado.
   d) El costo óptimo es ∞ y ninguna solución factible es óptima.
4. Proponga una nueva restricción linealmente independiente a las originales tal que alguna de las soluciones básicas sea degenerada.
5. Considere que se agrega la nueva restricción $y \geq 0$ ¿el punto $(0,0)$ es un vértice? ¿el punto $(2,0)$ es vértice? ¿el punto $(0,5)$ es vértice?
   Definición: "Sea P un poliedro. Un vector $x \in P$ es un vértice de P si existe algún c tal que $c'x < c'y$ para todo y que satisface $y \in P$ y $y \neq x$."

Solución:

1. El bosquejo en cuestión será:



2. Determine las soluciones básicas. ¿Cuáles de estas son factibles? ¿Cuál es el óptimo?

   Solución:

   Las soluciones básicas son aquellos puntos en donde hay n restricciones activas *(que una restricción esté activa quiere decir que se está cumpliendo con igualdad)*, con n la cantidad de variables. En este caso, $n = 2$, por lo que buscamos puntos en donde 2 restricciones se estén cumpliendo con igualdad. En este caso son las intersecciones entre rectas: $(0, 5)$, $(0, 3)$ y $(\frac{8}{7}, \frac{27}{7})$. Para que una SB sea factible, debe pertenecer al poliedro, por lo tanto $(0, 5)$ no es SBF. Finalmente, el óptimo corresponde al punto $(\frac{8}{7}, \frac{27}{7})$. La explicación detallada está en la siguiente parte.

3. Proponga un vector de costos para cada caso:

   Para esta parte, nos vamos a apoyar en la teoría gráfica que hay detrás de todos estos problemas. Sabemos que cambiar el vector costo implica cambiar la función objetivo. Además, sabemos que existirán las rectas de isocostos que se caracterizarán por formar un ángulo de 90º con el gradiente de la función.

   Por lo tanto, al momento de elegir dicho vector de costo (que es lo mismo que escoger el gradiente de la función, pues recuerden que, como la función objetivo es lineal, c $= \nabla f$, con $\nabla f =$ gradiente de la función), tenemos que tener en consideración que **la solución o soluciones que queremos que sean óptimas** (es decir, que al evaluarlas en la función, dicho valor $c^T x^*$ cumpla $c^T x^* \geq c^T x, \forall x \in P$, puesto que estamos maximizando) **pertenezca o pertenezcan a la recta de isocosto de mayor nivel tal que choque con el conjunto**. Si eventualmente se quisiese minimizar, el análisis sería al revés. Veámos un ejemplo ultra detallado para la parte (a).

a) La solución óptima es única y distinta al problema original.
Solución: Supongamos que ahora quiero que mi óptimo se encuentre en el punto $(0,3)$ en vez del anterior, pues la solución estaría siendo única y sería diferente al problema original. Entonces, tenemos que escoger un vector costo tal que cumpla con lo detallado anteriormente. Siendo aún más especificos, como dicha solución quiero que sea única, la recta de isocostos de mayor nivel debe tocar de forma tangencial a mi punto (0,3). Es así como tendré todo un conjunto de costos que cumplen con lo pedido:

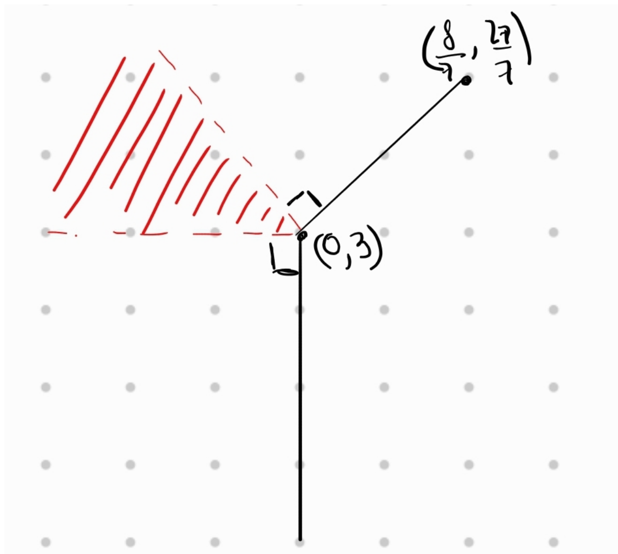


Figura 1: conjunto rojo representa todas las posibles inclinaciones que podría tener mi vector costo para que el punto (0,3) sea óptimo

Entre todas las posibles inclinaciones, podríamos encontrar algunos ejemplos más concretos:

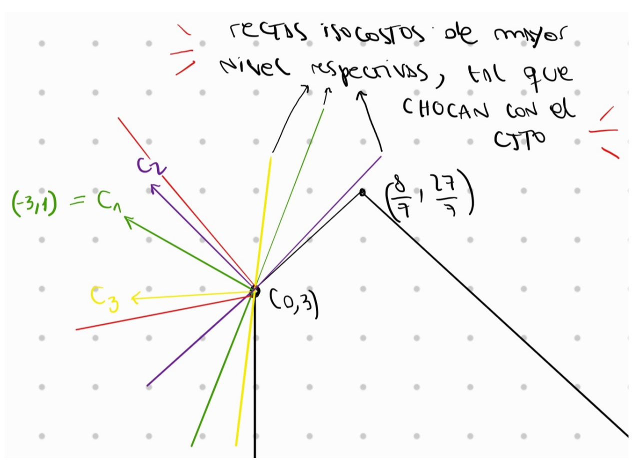


Figura 2: Tres diferentes costos, tal que la inclinación de los vectores en cuestión pertenecen al conjunto rojo definido en la imagen anterior.

Donde en particular podemos elegir el costo c = (-3, 1).

Cosas importantísimas a notar: en la figura 1, fíjense que el conjunto en cuestión presenta un ángulo de 90º entre sus límites y algunas restricciones del poliedro; y fíjense en que las líneas límites están punteadas. Esto sucede debido a que, si tenemos un vector costo que no pertenece a mi conjunto rojo, entonces el punto (0,3) que quiero que sea el óptimo, **no pertenecerá a la recta de isocostos de mayor nivel tal que choque con el conjunto.** Basta ver el ejemplo inicial con el que partimos. Ese vector no pertenecía al conjunto rojo, haciendo así que el óptimo fuese otro punto, el $(\frac{8}{7}, \frac{27}{7})$

Y si eventualmente eligo el vector costo que justo tenga la inclinación de las lineas punteadas, entonces mi punto (0,3) será óptimo, pero no será único, lo que se evidencia en la siguiente imagen:

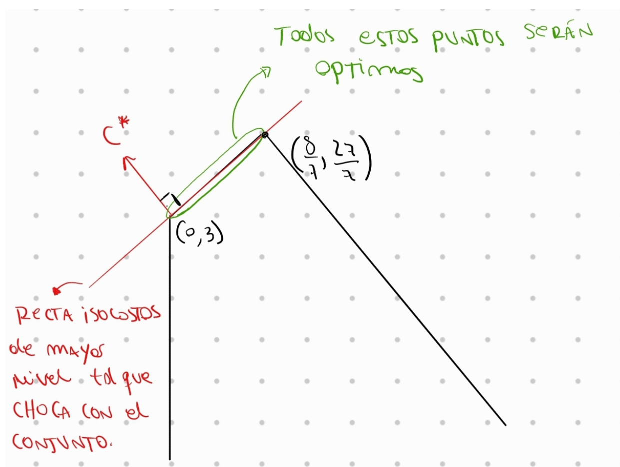


Figura 3: Vector costo c^* justo tiene la inclinación de una de las lineas punteadas y por tanto forma 90º con una de las restricciones del conjunto

Se puede notar como cada punto del conjunto verde, que de hecho se puede expresar como un punto de la linea que une (0,3) y $(\frac{8}{7}, \frac{27}{7})$, será óptimo.

b) Existen múltiples soluciones óptimas y el conjunto de estas es acotado.
Solución: $c = (-3, 4)$

c) Existen múltiples soluciones óptimas y el conjunto de estas es no acotado.
Solución: $c = (-1, 0)$

d) El costo óptimo es ∞ y ninguna solución factible es óptima.
Solución: $c = (-3, -1)$

4. Proponga una nueva restricción linealmente independiente a las originales tal que alguna de las soluciones básicas sea degenerada.

Solución:

Recordemos que una solución es degenerada cuando hay más de n restricciones activas. Si se agrega la restricción $y \leq 3$, el punto $(0,3)$ pasa a ser degenerado.

5. Considere que se agrega la nueva restricción $y \geq 0$ ¿el punto $(0,0)$ es un vértice? ¿el punto $(2,0)$ es vértice? ¿el punto $(0,5)$ es vértice?
   Definición: "Sea P un poliedro. Un vector $x \in P$ es un vértice de P si existe algún c tal que $c'x < c'y$ para todo y que satisface $y \in P$ y $y \neq x$."

   Solución: El punto $(0,0)$ es vértice, ya que por la definición, si tomamos $c = (1,1)$ entonces $c'x < c'y$ para todo y que satisface $y \in P$ y $y \neq x$. (Ojo: este c no es el vector de costos, sino que es el que utilizamos para la definición de vértice). El punto $(2,0)$ no es vértice ya que no cumple la desigualdad de la definición de forma estricta para ningún c. El punto $(0,5)$ no es vértice porque no pertenece al poliedro.

   Noten que el conjunto es poliedro no vacío, por lo que se podría utilizar el teorema de equivalencia para demostrar lo pedido, demostrando primero si los puntos son SBF y luego concluir (teorema equivalencia: descrito en el teorema 18 de la página 28 del apunte IN3171 M.Osorio y M.Muñoz).

Pregunta 2: Convexidad

1. **Ejercicio 1.2 del libro *Introduction to Linear Optimization***
   Supongamos que $f_1, \ldots, f_m$ son funciones convexas de $\mathbb{R}^n$ en $\mathbb{R}$ y sea
   $$f(x) = \sum_{i=1}^{m} f_i(x).$$
   Demuestre que si cada f_i es convexa, entonces f también es convexa.

2. **Ejercicio 2.2 del libro *Introduction to Linear Optimization***
   Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa y c una constante. Demuestre que el conjunto
   $$S = \{x \in \mathbb{R}^n \mid f(x) \leq c\}$$
   es convexo.

3. Muestre gráficamente que es posible que dos conjuntos distintos, S y T, tengan la misma envoltura convexa.

4. Construya un ejemplo que muestre que la unión de conjuntos convexos no es necesariamente convexa, y demuestre que la intersección de conjuntos convexos sí es convexa.

Solución:

1. Para probar esto, recordemos la definición de una función convexa. Una función $f : \mathbb{R}^n \to \mathbb{R}$ es convexa si, para todo $x, y \in \mathbb{R}^n$ y todo $\lambda \in [0, 1]$, se cumple que:
   $$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$
   Dado que cada f_i, con $i = 1, \ldots, m$, es convexa, sabemos que todas satisfacen esta propiedad.

   Ahora, consideremos la función $f(x) = \sum_{i=1}^{m} f_i(x)$. Queremos demostrar que f es convexa, es decir, que:
   $$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$
   Por la definición de f, tenemos que:
   $$f(\lambda x + (1-\lambda)y) = \sum_{i=1}^{m} f_i(\lambda x + (1-\lambda)y)$$
   Como cada f_i es convexa, podemos aplicar la desigualdad de convexidad a cada término de la suma:
   $$f_i(\lambda x + (1-\lambda)y) \leq \lambda f_i(x) + (1-\lambda)f_i(y)$$

Obteniendo así:

$$\sum_{i=1}^{m} f_i(\lambda x + (1-\lambda)y) \leq \sum_{i=1}^{m} (\lambda f_i(x) + (1-\lambda) f_i(y))$$

$$= \lambda \sum_{i=1}^{m} f_i(x) + (1-\lambda) \sum_{i=1}^{m} f_i(y)$$

Dado que $f(x) = \sum_{i=1}^{m} f_i(x)$ y $f(y) = \sum_{i=1}^{m} f_i(y)$, esto se convierte en:

$$\lambda f(x) + (1-\lambda) f(y)$$

Por lo tanto, como esta desigualdad se cumple para todo $x, y \in \mathbb{R}^n$ y todo $\lambda \in [0,1]$, concluimos que f es convexa.

2. Se utilizará la definición de convexidad de conjuntos. Sean $\mathbf{x}, \mathbf{y} \in S$ y sea $\lambda \in [0,1]$. Para demostrar lo que se pide, basta ver que $z \in \mathbb{R}^n$ definido como $\mathbf{z} = \lambda \mathbf{x} + (1-\lambda)\mathbf{y}$ pertenece al conjunto S. Para ello, se tiene que:

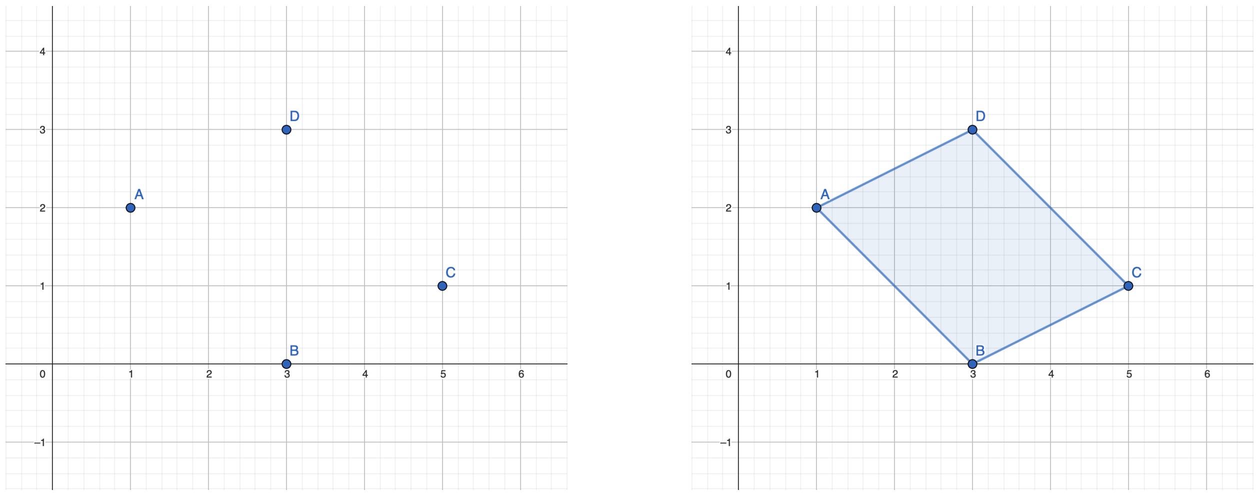
$$\begin{aligned} f(\mathbf{z}) &= f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \\ &\leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) \\ &\leq \lambda c + (1-\lambda) c \\ &= c \end{aligned}$$

donde se usó que f es una función convexa y que los vectores $\mathbf{x}$ e $\mathbf{y}$ pertenecen al conjunto S. Como se llegó a que $f(\mathbf{z}) \leq c$, se concluye que $\mathbf{z} \in S$ y por tanto que S es un conjunto convexo.

3. Considere un conjunto S con k puntos, y un conjunto T con los mismos k puntos junto con un punto arbitrario formado por alguna combinación convexa de los k puntos. Podemos ver que $S \neq T$; sin embargo, la envoltura convexa de ambos conjuntos es la misma, ya que ambos contienen los mismos puntos salvo uno que es combinación convexa de los otros.

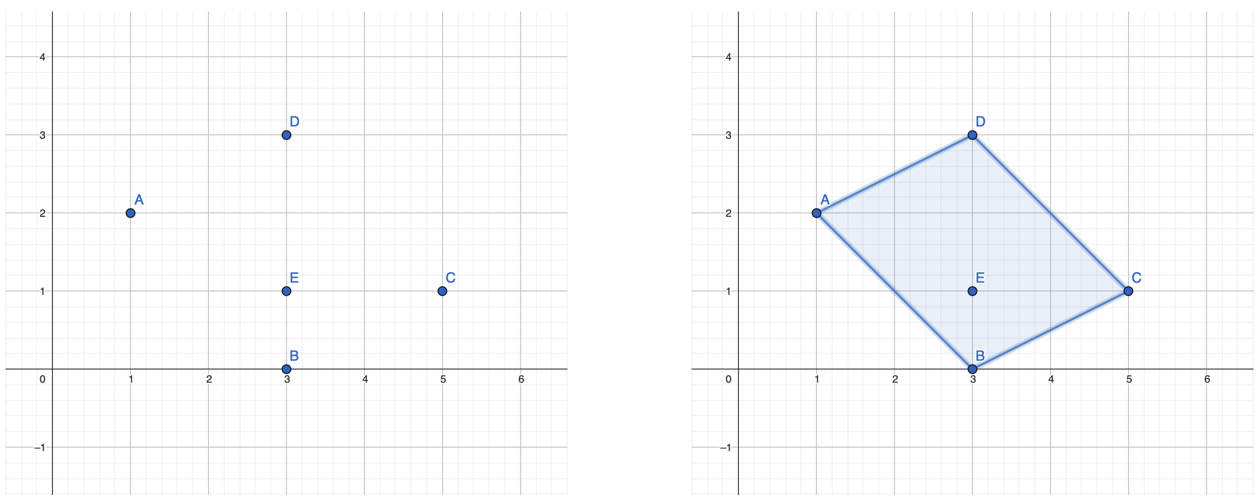
   Como ejemplo, considere los siguientes puntos: $A = (1,2)$, $B = (3,0)$, $C = (5,0)$, $D = (3,3)$ y $E = (3,1)$. Definimos los conjuntos:

$$S = \{A, B, C, D\} \quad \text{y} \quad T = \{A, B, C, D, E\}$$



(a) Conjunto S (b) conv(S)

Figura 4: Conjunto S y su envoltura convexa



(a) Conjunto T (b) conv(T)

Figura 5: Conjunto T y su envoltura convexa

La envoltura convexa de S está representada por todo lo que se encuentra dentro de la figura geométrica en la Figura 1(b), mientras que la envoltura convexa de T está representada por todo lo que se encuentra dentro de la figura geométrica en la Figura 2(b).

Vemos gráficamente que los conjuntos S y T son distintos, sin embargo su envoltura convexa es la misma.

4. Para ello, consideremos dos conjuntos convexos como esferas de radio k y r, respectivamente.
Sea $A \subset \mathbb{R}^2$ el conjunto de puntos que forman un disco de radio $k > 0$ con centro en $(0, 0)$:

$$A = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; x^2 + y^2 \leq k^2 \right\}$$

Sea $B \subset \mathbb{R}^2$ el conjunto de puntos que forman un disco de radio $r > 0$ con centro en $(1, 0)$:

$$B = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; (x-1)^2 + y^2 \leq r^2 \right\}$$

Notemos que gráficamente se puede ver que ambos son convexos, pues no importa el par de puntos que yo eliga, la línea que los une pertenecerá al conjunto. **propuesto hacerlo formalmente. Puede ser algo desafiante, pero definiendose un x_λ y un y_λ de forma estratégica, más ocupando desigualdades conocidas, se llega al resultado.**
Luego, el conjunto $C = A \cup B$ no será convexo. Esto se puede notar graficamente al elegir los puntos a, b. Vemos cómo la recta que los une no pertenece completamente al conjunto, y por lo tanto C no es convexo.

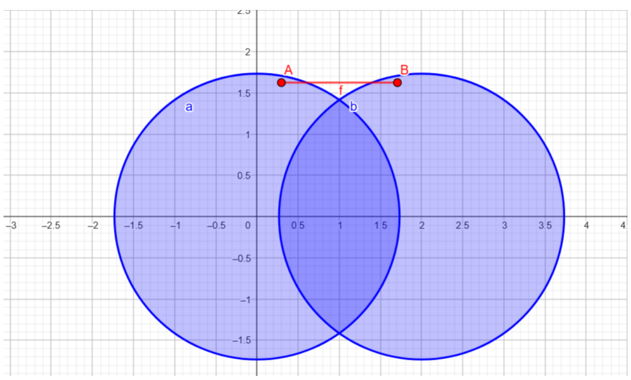


Figura 6: Recta entre dos puntos no pertenece completamente al conjunto C. Recuerden que C es la unión y no la intersección.

Demostración intersección es convexa:
Sea A convexo y sea B convexo. Entonces, por definición de convexidad:

$$\forall x, y \in A,\ \forall \lambda \in [0,1], \quad \lambda x + (1-\lambda) y \in A$$
$$\forall x, y \in B,\ \forall \lambda \in [0,1], \quad \lambda x + (1-\lambda) y \in B$$

Sea $D = A \cap B$, y tomemos $x^*, y^* \in D$. Entonces, por pertenecer a la intersección, se cumple que:

$$x^*, y^* \in A \quad \text{y} \quad x^*, y^* \in B$$

Luego, para todo $\lambda \in [0, 1]$:

$$\lambda x^* + (1-\lambda)y^* \in A \quad \text{y} \quad \lambda x^* + (1-\lambda)y^* \in B$$

Pues son conjuntos convexos. Por lo tanto como esta expresión pertenece tanto a A y a B a la vez:

$$\lambda x^* + (1-\lambda)y^* \in A \cap B = D$$

Así concluimos que D es convexo.

Pregunta 3 - Geometría

Considere los polítopos

$$P_1 = \left\{ x \in \mathbb{R}^2 \,\middle|\, 2x_1 - x_2 \leq 0,\ x_1 + x_2 \leq 0,\ -\frac{1}{2}x_1 + x_2 \leq 3,\ x_2 \geq -2 \right\}$$

$$P_2 = \text{conv}\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right)$$

1. Dibuje los conjuntos P_1 y P_2.

2. Defina $A = \text{conv}(P_1 \cup P_2)$. Dibuje A y deduzca una descripción de desigualdades de éste.

3. Muestre para este ejemplo que si x es punto extremo de A, entonces x es punto extremo o de P_1, o de P_2. Muestre también que la inversa no es cierta, es decir, encuentre x punto extremo de P_1 o P_2 que no es punto extremo de A.

4. Generalice y formalice el resultado anterior. Es decir, considere $P_1, \ldots, P_K$ politopos en $\mathbb{R}^n$ y A la envoltura convexa de $P_1 \cup P_2 \cup \cdots \cup P_K$. Equivalentemente,

$$A = \left\{ \sum_{k=1}^{K} \lambda_k x^k \,\middle|\, x^k \in P_k,\ \sum_{k=1}^{K} \lambda_k = 1,\ \lambda \geq 0 \right\}.$$

   Muestre que si x es punto extremo de A, entonces es punto extremo de P_k para algún $k = 1, \ldots, K$.

   Indicación. *Si lo desea, puede suponer conocido el siguiente resultado: un politopo es la envoltura convexa de sus puntos extremos.*

Solución:

1. Los conjuntos P_1 y P_2 se muestran en las siguientes figuras:

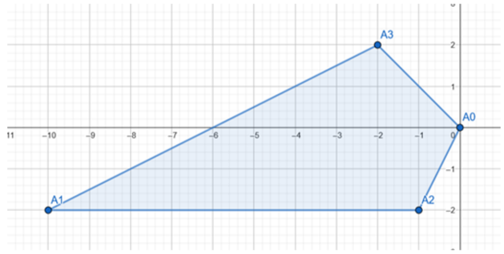


Figura 3: Conjunto P_1

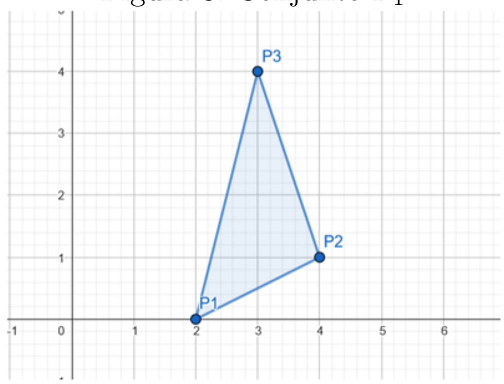


Figura 4: Conjunto P_2

2. La envoltura convexa de los conjuntos P_1 y P_2 se muestra en la siguiente figura:

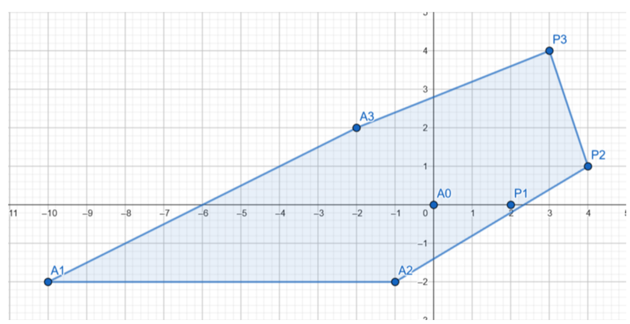


Figura 5: Conjunto $\text{conv}(P_1 \cup P_2)$

Donde el conjunto $A = \text{conv}(P_1 \cup P_2)$ está definido por las siguientes restricciones:

$$
\begin{aligned}
f &: \; x_2 \leq \frac{1}{2}x_1 + 3 \\
g &: \; x_2 \leq \frac{2}{5}x_1 + \frac{14}{5} \\
h &: \; x_2 \leq -3x_1 + 13 \\
q &: \; x_2 \geq \frac{2}{3}x_1 - \frac{4}{3} \\
r &: \; x_2 \geq -2 \\
p &: \; x_2 \geq \frac{1}{2}x_1 - 1
\end{aligned}
$$

Las restricciones (1) y (2) pertenecen a P_1, mientras que las restantes provienen de las rectas que unen puntos extremos de P_1 y P_2: (4,1) con (3,4), (4,1) con (–1,2), y (–2,2) con (3,4).

3. Los puntos extremos de A son: $(-2,2)$, $(-10,2)$, $(-1,2)$, $(4,1)$ y $(3,4)$.

    - $(-2,2)$, $(-10,2)$ y $(-1,2)$ son puntos extremos de P_1.
    - $(4,1)$ y $(3,4)$ son puntos extremos de P_2.

    De esta forma, si x es punto extremo de A, entonces x es punto extremo de P_1 o de P_2.

    Para ver que la recíproca no es cierta, consideremos el punto $(0,0)$, que es punto extremo de P_1, pero no de A, ya que:

    $$(0,0) = \frac{1}{2}(-2,-2) + \frac{1}{2}(2,2)$$

    Como $(-2,-2)$ y $(2,2)$ pertenecen a A, entonces $(0,0)$ es combinación convexa de elementos en A y por tanto no es punto extremo de A.

4. **Demostración general:** Supongamos por contradicción que x es punto extremo de $A = \text{conv}(P_1 \cup \cdots \cup P_K)$ pero no es punto extremo de ningún P_k con $k = 1, \ldots, K$.

    Como $x \notin \text{ext}(P_k)$ para todo k, entonces existe una representación

    $$x = \lambda y + (1-\lambda)z, \quad \text{con } 0 < \lambda < 1 \text{ y } y, z \in P_k$$

    Pero como cada P_k es convexo y es la envoltura convexa de sus puntos extremos, entonces $y, z \in A$.

    Esto contradice el hecho de que x es punto extremo de A, ya que entonces se escribiría como combinación convexa de dos puntos distintos en A.

    Por lo tanto, si x es punto extremo de A, entonces x debe ser punto extremo de algún P_k para $k = 1, \ldots, K$.

    Observación: Un politopo es la envoltura convexa de sus puntos extremos.

Resumen

- **Forma canónica de un PPL:** Todo problema de programación lineal puede escribirse como:
  $$\min c^T x \quad \text{s.a.} \quad Ax \geq b$$

- **Factibilidad y optimalidad:**
  - x es **factible** si cumple todas las restricciones.
  - x es **óptimo** si, además de ser factible, cumple $c^T x \leq c^T x'$ para todo x' factible.

- **Tipos de soluciones posibles en un PPL:**

- 1. Existe un único punto óptimo x^*.
- 2. Existen múltiples puntos óptimos x^*.
- 3. El problema es no acotado: para todo C, existe x tal que $c^T x \leq C$.

- **Definiciones claves:**

- **Poliedro:** $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$, intersección finita de semiespacios.
- **Acotado:** $S \subseteq \mathbb{R}^n$ es acotado si $|x_i| \leq K$ para todo $x \in S$.
- **Polítopo:** Un poliedro que es acotado.
- **Convexo:** S es convexo si $\lambda x + (1-\lambda)y \in S$ para $x, y \in S$, $\lambda \in [0,1]$.
- **Combinación convexa:** $y = \sum_{i=1}^K \lambda_i x_i$, con $\sum \lambda_i = 1$ y $\lambda_i \geq 0$.
- **Envoltura convexa:** $\text{Conv}\{x_1, ..., x_K\}$ es el conjunto de todas las combinaciones convexas de esos puntos.
- **Punto extremo:** $x \in P$ es extremo si no puede escribirse como combinación convexa de otros dos puntos distintos en P.
- **Vértice:** $x \in P$ es un vértice si existe c tal que x es la única solución óptima de $\min c^T x$.

- **Restricción activa:** $a_i^T x = b_i$ se dice activa en x.
- **Linealmente independientes:** Restricciones cuyos vectores a_i son l.i.
- **Solución básica:** Punto x^* con n restricciones activas l.i. (incluyendo igualdades).
- **Solución básica factible (SBF):** x^* es solución básica y pertenece a P.
- **Poliedro con línea:** Existe $x \in P$ y $d \neq 0$ tal que $x + \lambda d \in P$ para todo $\lambda \in \mathbb{R}$.

- **Lema: Caracterización de SBF:** Sea $x^* \in \mathbb{R}^n$ y $I = \{i \mid a_i^T x^* = b_i\}$:
  - 1. Hay n vectores a_i, $i \in I$, l.i.
  - 2. $\text{span}(S) = \mathbb{R}^n$ con $S = \{a_i \mid i \in I\}$.
  - 3. El sistema $a_i^T x = b_i$, $i \in I$, tiene solución única.
- **Teorema: Equivalencia vértice, punto extremo y SBF:** Si $P = \{x \mid Ax \geq b\}$ y $x^* \in P$, entonces:
  - x^* es un vértice $\Leftrightarrow$
  - x^* es un punto extremo $\Leftrightarrow$

  - x^* es una solución básica factible.

- **Teorema: Condición para que P tenga puntos extremos:** Sean equivalentes:
  - 1. P tiene al menos un punto extremo.
  - 2. P no contiene una línea.
  - 3. Existen n restricciones activas linealmente independientes.