Tema 6. Optimización numérica.

Computación Numérica

Antonio Palacio

Departamento de Matemáticas
Universidad de Oviedo

palacioantonio@uniovi.es

Curso 2021-2022

Contenidos I

Contenidos

Preliminares

Motivación

Dada una función

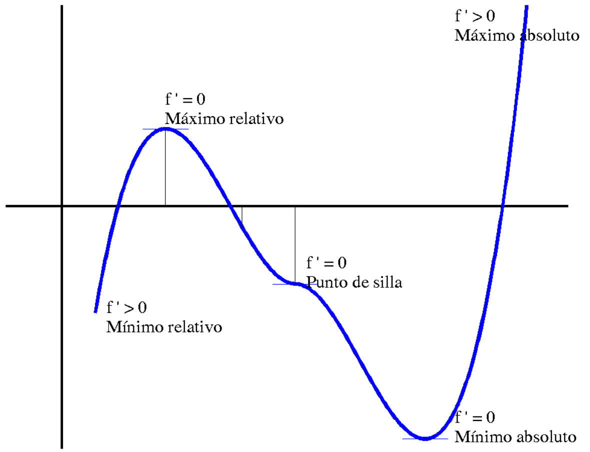
$$f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$$

nos planteamos el problema de encontrar los máximos y/o los mínimos de f en U. Cuando U es un conjunto abierto, se suele utilizar el término **Optimización sin restricciones** para este tipo de problemas.

Nota 6.1

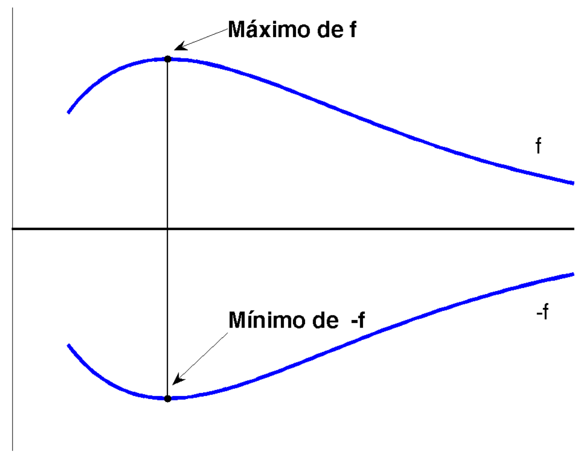
*Un punto $x^* \in U$ se dice que es un **mínimo relativo** de f si existe un entorno V de x^* tal que $f(x) \geq f(x^*)$ para todo $x \in V \cap U$.*
*Un punto se denomina **máximo relativo** si $f(x) \leq f(x^*)$ para todo $x \in V \cap U$.*
*El máximo y el mínimo se dice que son **absolutos** si $V = U$. Utilizaremos el término extremo para referirnos a los máximos y mínimos.*

Preliminares



Preliminares

Teniendo en cuenta que todo máximo de f es un mínimo para $-f$, limitaremos nuestro estudio al cálculo de mínimos.



Preliminares

Teorema 6.1 (Teorema de Weierstrass)

Sea $K \subset U$. Si K es compacto (cerrado y acotado) y f es continua en K entonces f alcanza un mínimo absoluto y un máximo absoluto en K.

Teorema 6.2

Sea $\mathbf{x}^*$ punto interior de U y supongamos que $f \in C^1(U)$.

- ***Condición necesaria de extremo**: Si x^* es un extremo relativo de f entonces:*

$$\nabla f(\mathbf{x}^*) = 0$$

- ***Condición suficiente de mínimo**: Supongamos ahora que $f \in C^2(U)$ y que $\nabla f(\mathbf{x}^*) = 0$. Si se verifica que*

$$\mathbf{y}^t \; \nabla^2 f(\mathbf{x}^*) \; \mathbf{y} > 0 \; para \; todo \; y \neq 0$$

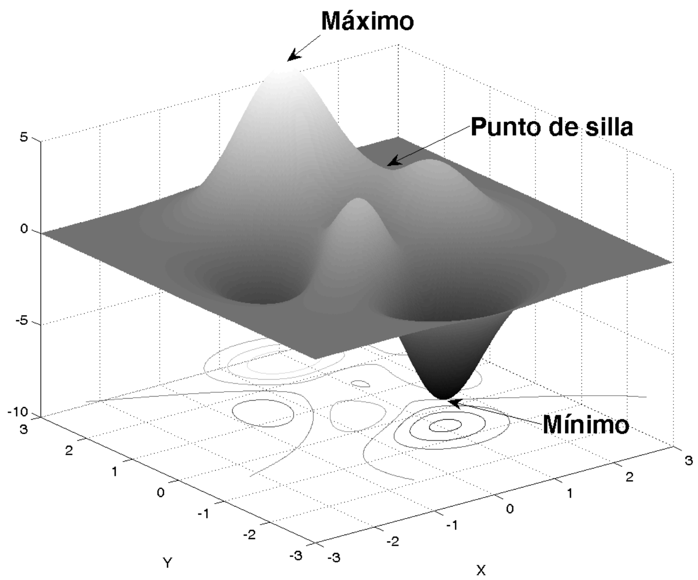
  entonces x^* es un mínimo de f.

Preliminares

Observación 6.1

- *Los puntos que verifican la condición $\nabla f(\mathbf{x}) = 0$ reciben el nombre de* **puntos estacionarios o críticos**. *Pueden existir puntos estacionarios que no son extremos y suelen denominarse puntos de silla.*
- *La matriz $H(\mathbf{x}^*) = \nabla^2 f(\mathbf{x}^*) = \left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}^*) \right)$ se denomina hessiano. La condición suficiente se puede expresar diciendo que el Hessiano es definido positivo.*
- *La condición suficiente de máximo es que el hessiano sea definido negativo.*
- *Observe que en dimensión 1 el hessiano está definido por $f''(x^*)$ y ser definido positivo significa que $f''(x^*) > 0$.*

Preliminares



Contenidos

Optimización en una variable

Introducción

Dada una función

$$g\ :\ [a,b] \longrightarrow \mathbb{R}$$

nos planteamos el problema de hallar sus mínimos en $[a,b]$.
Cuando g es derivable, la condición necesaria de extremo nos proporciona un procedimiento para obtener los mínimos en (a,b) ya que estos deben encontrarse entre las soluciones de la ecuación:

$$g'(x) = 0$$

La resolución de esta ecuación puede abordarse mediante los métodos numéricos vistos en el *Tema 2* (*bisección, Newton-Raphson, secante,...*). Sin embargo, en algunas ocasiones el uso de g' resulta muy complejo lo que hace necesario disponer de métodos que no requieran su cálculo. A continuación describiremos uno de estos procedimientos alternativos que consiste en reducir iterativamente la longitud del intervalo que contiene al mínimo mediante el análisis de los valores de g en ciertos puntos.

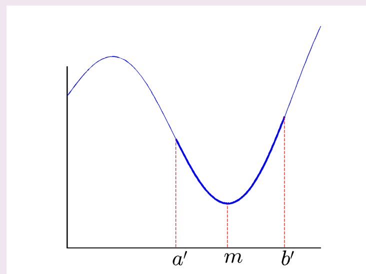
Optimización en una variable

Procedimiento

Supondremos que g es una función continua que posee un único mínimo en $x = m$, punto interior de su dominio. Por la continuidad, existe un intervalo $[a', b']$ con $m \in [a', b']$ y verificando que:

g **estrictamente decreciente en** $[a', m]$

g **estrictamente creciente en** $[m, b']$



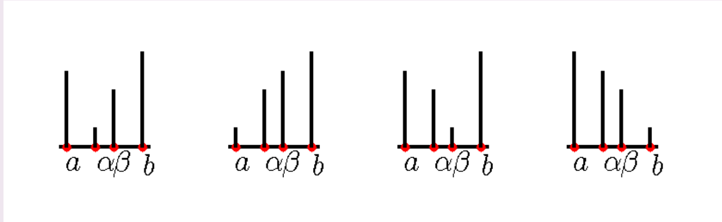
Metodología

A partir de estas propiedades se puede diseñar un método similar al método de bisección consistente en reducir iterativamente la longitud del intervalo inicial $[a', b']$ (que denotaremos en lo que sigue $[a,b]$ para simplificar las notaciones) que contiene a m.

- Se eligen dos puntos α y β tal que $a < \alpha < \beta < b$.
- Se reduce la longitud del intervalo que contiene al mínimo evaluando g en α y β:
  1. Si $g(\alpha) < g(\beta) \ \Rightarrow \ m \in [a, \beta]$
  2. Si $g(\alpha) \geq g(\beta) \ \Rightarrow \ m \in [\alpha, b]$
- Se utiliza como aproximación de m el punto medio del intervalo resultante.

Observación 6.1

- *La reducción del intervalo puede ser mejorada en algunas ocasiones. Por ejemplo, si en el caso 1 además se cumple que $g(\alpha) > g(a)$ podemos concluir que $m \in [a, \alpha]$. En el caso 2 se pueden plantear razonamientos análogos. Desde el punto de vista de la programación, esta mejora del intervalo no representa una ventaja significativa en la eficiencia del método.*
- *Eligiendo adecuadamente los puntos α y β se garantiza que la sucesión de puntos medios de los intervalos converge al mínimo. Por ejemplo, una elección muy sencilla sería $\alpha = (b-a)/3$ y $\beta = 2(b-a)/3$. Un elección no tan simple pero mas eficiente desde el punto de vista numérico es el método de la sección de oro.*



Método de la sección de oro

Se reduce el tamaño del intervalo en cada iteración en una razón fija dada por $\Phi - 1$, siendo Φ el **número de oro (o razón áurea)**:

$$\Phi = \frac{\sqrt{5}+1}{2} \approx 1{,}6180339887 \qquad \text{y} \qquad \frac{1}{\Phi} = \Phi - 1 \approx 0{,}6180339887$$

- Datos iniciales: $[a_1, b_1] = [a, b]$, $\qquad \Delta = (\Phi - 1)(b_1 - a_1)$.
- Se eligen $\alpha = b_1 - \Delta$, $\qquad \beta = a_1 + \Delta$.
- Se aplica el método para obtener un nuevo intervalo $[a_2, b_2]$ cuya longitud es Δ y que debe coincidir con $[a_1, a_1 + \Delta]$ o con $[b_1 - \Delta, b_1]$.

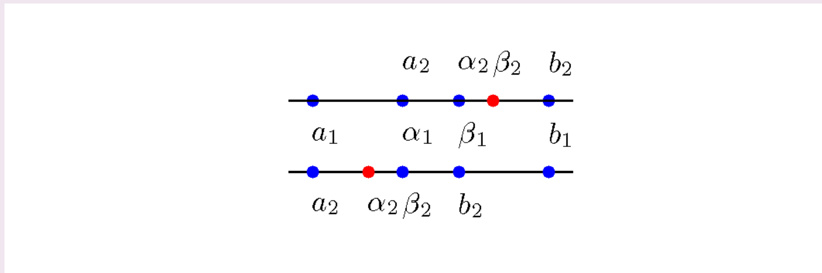
Del proceso anterior se deduce que $(b_n - a_n) = (\Phi - 1)^{n-1}(b-a)$.
Puesto que $0 < \Phi - 1 < 1$ se tiene que $\lim_{n\to\infty}(b_n - a_n) = 0$, lo que garantiza la convergencia del método.

Observación 6.1

La obtención de $[a_n, b_n]$ requiere realizar $n-1$ iteraciones.

Nota 6.2

El uso de la razón áurea hace que $\beta_2 = \alpha_1$ (si $[a_2, b_2] = [a_1, \beta_1]$) o bien $\alpha_2 = \beta_1$ (si $[a_2, b_2] = [\alpha_1, b_1]$).



Problema 6.1

Calcule una aproximación del mínimo de $g(x) = \dfrac{\operatorname{sen} x}{x}$, $x \in [\pi, 2\pi]$ aplicando dos iteraciones del método de la sección de oro. ¿Cuántas iteraciones del método garantizan que el valor que hace mínima la función está en un intervalo de longitud menor que una centésima?

Contenidos

Optimización en $\mathbb{R}^n$ sin restricciones

Introducción

Se considera ahora una función

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

que supondremos de clase dos. Un procedimiento para buscar los mínimos de f consiste en utilizar la condición necesaria de extremo y plantear la resolución del sistema:

$$\nabla f(\mathbf{x}) = \mathbf{0}$$

Este sistema, que posiblemente sea no lineal, puede resolverse mediante las técnicas desarrolladas en el tema 3.

El objetivo de esta sección es presentar varios métodos iterativos que permitan evitar la resolución del sistema. Estos métodos adoptan la forma:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda_k \mathbf{d}^k \tag{1}$$

siendo la elección del parámetro λ_k y del vector $\mathbf{d^k}$ característica de cada método.

Método de máxima pendiente

Método de máxima pendiente

Es conocido que el vector $\nabla f(\mathbf{x})$ determina la dirección y sentido en el que la función f crece más rápidamente en torno a $\mathbf{x}$. Si, partiendo de $\mathbf{x}^k$, queremos obtener un punto en el que la función tome un valor menor que $f(\mathbf{x}^k)$, en la fórmula (1) es natural elegir un vector $\mathbf{d}^k$ que verifique

$$\nabla f(\mathbf{x}^k) \cdot \mathbf{d}^k < 0$$

es decir, una *dirección de descenso*. Cuando se elige $\mathbf{d}^k = -\nabla f(\mathbf{x}^k)$, se obtiene el denominado **método de máxima pendiente**:

- Dato Inicial: $\mathbf{x}^0$
- Dado $\mathbf{x}^k$ se define $\mathbf{d}^k = -\nabla f(\mathbf{x}^k)$.
- Se calcula λ_k por $f(\mathbf{x}^k + \lambda_k \mathbf{d}^k) = \min\limits_{\lambda \geq 0} f(\mathbf{x}^k + \lambda \mathbf{d}^k)$
- Se define $\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda_k \mathbf{d}^k$
- Test de Parada: $\|\nabla f(\mathbf{x}^{k+1})\| \leq Tol$

Métodos de máxima pendiente

Observación 6.2

- *En el proceso anterior, el escalar λ_k no se puede calcular, en general, de forma exacta. Suele utilizarse algún método de optimización para funciones de una variable.*
- *El método de máxima pendiente tiene, en general, convergencia de orden 1.*

Problema 6.2

Realice una iteración del método de máxima pendiente para calcular

$$\min_{(x,y)\in\mathbb{R}^2} \frac{x^4 + y^4 + 2xy}{4}$$

partiendo del punto $\mathbf{x}^0 = (1,\ 0)$. Utilice un paso del método de la sección de oro para resolver $\min_{\lambda \geq 0} f(\mathbf{x}^k + \lambda \mathbf{d}^k)$ considerando que $\lambda \in [0,1]$.

Introducción

- Para conseguir que el método iterativo descrito en (1) tenga convergencia de orden dos es necesario hacer una elección mas eficiente de la dirección de descenso $\mathbf{d^k}$ que la realizada en el método de máxima pendiente.
- Una forma de mejorar esa elección es tener en cuenta que cada dirección de descenso debe cumplir ciertas *propiedades de ortogonalidad* con respecto a la dirección de la etapa anterior. Por ejemplo definiendo

$$\mathbf{d^k} = -\nabla f(\mathbf{x^k}) + \beta_k \mathbf{d^{k-1}}$$

y eligiendo $\beta_k \in \mathbb{R}$ de forma adecuada.

Método del gradiente conjugado

El **Método del gradiente conjugado** con la elección del parámetro β_k dada por **Fletcher-Reeves** se defini como:

- Dato inicial: $\mathbf{x}^0$
- Dado $\mathbf{x}^k$ se calcula $\mathbf{d^k} = -\nabla f(\mathbf{x^k}) + \beta_k \mathbf{d^{k-1}}$ siendo:

$$\text{Para } k=0, \qquad \beta_k = 0 \text{ y } \mathbf{d^{-1}} = \mathbf{0}$$

$$\text{Para } k \geq 1 \qquad \beta_k = \frac{\|\nabla f(\mathbf{x}^k)\|^2}{\|\nabla f(\mathbf{x}^{k-1})\|^2}$$

- Se calcula λ_k resolviendo $f(\mathbf{x}^k + \lambda_k \mathbf{d}^k) = \text{mín}_{\lambda \geq 0} f(\mathbf{x}^k + \lambda \mathbf{d}^k)$
- Se define $\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda_k \mathbf{d}^k$
- Test de Parada: $\|\nabla f(\mathbf{x}^{k+1})\| \leq Tol$

Problema 6.3

Realice dos iteraciones del método del gradiente conjugado (o método de Fletcher-Reeves) para calcular

$$\min_{(x,y)\in\mathbb{R}^2} \frac{1}{4}(x^4 + y^4 + 2xy)$$

partiendo del punto $\mathbf{x}^0 = (1,\ 0)$*. Utilice un paso del método de la sección de oro para*

$$\min_{\lambda} f(\mathbf{x}^k + \lambda \mathbf{d}^k)$$

suponiendo que $\lambda \in [0,1]$.

Observación 6.2

- *Si f es la función cuadrática* $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^t A\,\mathbf{x} + \mathbf{p}^t\mathbf{x} + q$ *con A matriz simétrica y definida positiva, $\mathbf{p} \in \mathbb{R}^n$ y $q \in \mathbb{R}$, el método del gradiente conjugado permite obtener en, a lo sumo, n iteraciones el mínimo de f (que es $\mathbf{x} = -A^{-1}\mathbf{p}$). Además, las direcciones de descenso son* ***conjugadas*** *respecto de la matriz A, es decir, verifican que $(\mathbf{d}^k)^t A\,\mathbf{d}^j = 0$ para $k \neq j$.*
- *Si f es una función en general y $\mathbf{x}$ es un mínimo en el que $\nabla f(\mathbf{x}) = 0$ y con hessiano* $H(\mathbf{x}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})\right)$ *simétrico y definido positivo entonces f se comporta, localmente, como una función cuadrática. En efecto, del desarrollo de Taylor en torno a $\mathbf{x}$ se deduce que*

$$f(\mathbf{x}+\Delta\mathbf{x}) \approx f(\mathbf{x}) + \nabla f(\mathbf{x})^t.\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^t\,\mathrm{H}(\mathbf{x})\,\Delta\mathbf{x} = f(\mathbf{x}) + \frac{1}{2}\Delta\mathbf{x}^t\,\mathrm{H}(\mathbf{x}).\Delta\mathbf{x}$$

Contenidos

Aplicación de la optimización a la resolución de sistemas

Aplicación de la optimización a la resolución de sistemas

Los métodos de optimización puede ser utilizados para la resolución de sistemas de ecuaciones. Por ejemplo:

- Para resolver el sistema no lineal $\left.\begin{array}{l} f_1(x,y)=0 \\ f_2(x,y)=0 \end{array}\right\}$ se pueden aplicar técnicas de optimización sobre la función

$$F(x,y)=(f_1(x,y))^2+(f_2(x,y))^2$$

  pues $F(x,y)\geq 0$ para todo (x,y) y el valor mínimo 0 se obtiene justamente en las soluciones del sistema.
- Para resolver el sistema lineal $A\mathbf{x}=\mathbf{b}$ con A simétrica y definida positiva, se puede aplicar el método del gradiente conjugado a la función cuadrática

$$f(\mathbf{x}) \;=\; \frac{1}{2}\mathbf{x}^t A\,\mathbf{x}-\mathbf{b}^t\mathbf{x}+q$$

Aplicación de la optimización a la resolución de sistemas

Problema 6.4

Calcule una aproximación de la solución del sistema no lineal

$$\left.\begin{array}{rcl} 2x+y & = & 2 \\ 2xy & = & 1 \end{array}\right\}$$

realizando una iteración del método de máxima pendiente sobre la función $f(x,y)=(2x+y-2)^2+(2xy-1)^2$. *Considere como estimación inicial* $\mathbf{x}^0=(0,0)$ *y aplique una iteración del método de la sección de oro para*

$$\min_{\lambda\geq 0} g(\lambda)=\min_{\lambda} f(\mathbf{x}^k+\lambda\mathbf{d}^k)$$

comenzando en el intervalo $[0,1]$.

Contenidos

Optimización con restricciones: Programación lineal

Se denomina problema de **optimización con restricciones** al problema del cálculo de mínimos de una función $f : \mathbb{R}^n \to \mathbb{R}$ sobre el conjunto de puntos que verifican las condiciones (restricciones):

$$g_1(\mathbf{x}^*) = 0, \dots, g_r(\mathbf{x}^*) = 0$$

$$g_{r+1}(\mathbf{x}^*) \leq 0, \dots, g_m(\mathbf{x}^*) \leq 0$$

Cuando además las funciones f y g_i sean lineales, es decir,

$$f(\mathbf{x}) = c_1 x_1 + \cdots + c_n x_n$$

$$g_i(\mathbf{x}) = a_{i1} x_1 + \cdots + a_{in} x_n$$

se hablará de problemas de **Programación Lineal**. En estos problemas f suele representar el coste o el beneficio en un proceso de fabricación.

Problema 6.5

Formule matemáticamente el siguiente problema: En una fábrica de cerveza que produce dos tipos distintos: negra (N) y rubia (R). Para su obtención son necesarios lúpulo, malta y levadura. La siguiente tabla proporciona la capacidad necesaria de cada uno de estos recursos escasos para producir un bidón de cada una de las respectivas cervezas, los kilos disponibles de cada recurso en un día y el beneficio por bidón de cerveza producida.

Recursos	*N*	*R*	
lúpulo	*1*	*2*	*14*
Malta	*3*	*1*	*15*
Levadura	*3*	*2*	*18*
Beneficio	*5*	*4*	

El problema del fabricante consiste en decidir cuanto debe fabricar diariamente de cada cerveza para que el beneficio sea máximo.

Solución del Ejercicio 5

Este problema puede plantearse como un problema matemático en el que se intenta maximizar una función lineal sujetas las variables (incógnitas) a una serie de restricciones lineales, es decir, puede plantearse como un problema de programación lineal:
$x_1 =$ Bidones de cerveza negra.
$x_2 =$ Bidones de cerveza rubia.

$$\begin{array}{rc} \text{Maximizar} & f(x_1, x_2) = 5x_1 + 4x_2 \\ \text{sujeto a} & x_1 + 2x_2 \leq 14 \\ & 3x_1 + x_2 \leq 15 \\ & 3x_1 + 2x_2 \leq 18 \\ & x_1, x_2 \geq 0. \end{array}$$

Optimización con restricciones: Programación lineal

Se dirá que un problema de programación lineal está en la **forma estándar** (por ejemplo el ejercicio 5) cuando se escribe de la forma:

$$\text{Maximizar} \quad f(x_1, \dots, x_n) = -c_1 x_1 - c_2 x_2 - \dots - c_n x_n \tag{2}$$

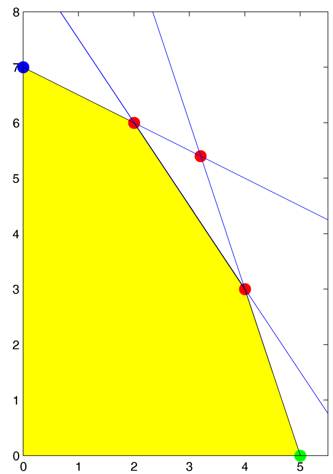
$$\begin{array}{c} \text{sujeto a las restricciones:} \\ a_{11}x_1 + a_{12}x_2 + \dots a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots a_{2n}x_n \leq b_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots a_{mn}x_n \leq b_m \\ x_1, x_2, \dots, x_n \geq 0. \end{array}$$

con $b_1, b_2, \dots, b_m \geq 0$.
Nota.- Todo problema de programación lineal se puede llegar a formular en la forma estándar, como se verá más adelante.

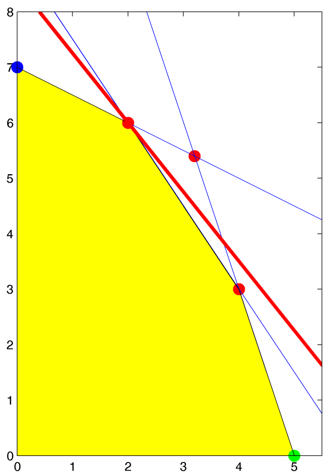
La formulación del problema en la forma standard requiere localizar las variables de decisión, la función objetivo y las restricciones:

- Variables de decisión: representan las n decisiones cuantificables para las que se deben determinar los valores respectivos. (En el ejemplo: x_i.- bidones de cerveza a producir al día.)
- Función objetivo: es la función matemática de las variables de decisión que queremos optimizar. (En el ejemplo: el beneficio, es decir, $z = 5x_1 + 4x_2$.)
- Restricciones: expresan matemáticamente todas las limitaciones que se puedan imponer sobre los valores de las variables de decisión, casi siempre en forma de ecuaciones o inecuaciones. En este contexto, el conjunto de puntos que cumplen las restricciones suele denominarse **Región factible**.

Por ejemplo, en el ejercicio 5 la región factible está determinado por: $\{(x_1, x_2) \in \mathbb{R}^2 / x_1 + 2x_2 \leq 14, 3x_1 + x_2 \leq 15, 3x_1 + 2x_2 \leq 18, x_1, x_2 \geq 0\}$.



Cuando el número de variables de decisión es menor o igual que tres, la solución óptima puede obtenerse gráficamente a través de las curvas de nivel de la función objetivo:



La formulación standard (2) se puede reescribir de forma equivalente en un formato que recibe el nombre de **forma aumentada**. Consiste en añadir a cada desigualdad una **variable de holgura** positiva y considerarlas como nuevas incógnitas del problema. Se asigna también a cada ecuación un índice y se modifican los índices de las variables del segundo miembro:

$$\begin{array}{cc} \text{Índices} & \\ 0 & f(x_1,\dots,x_n,x_{n+1},\dots,x_{n+m}) + c_1x_1 + c_2x_2 + \dots + c_nx_n = 0 \\ n+1 & a_{n+11}x_1 + a_{n+12}x_2 + \dots a_{n+1n}x_n + x_{n+1} = b_{n+1} \\ n+2 & a_{n+21}x_1 + a_{n+22}x_2 + \dots a_{n+2n}x_n + x_{n+2} = b_{n+2} \\ \vdots & \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ n+m & a_{n+m1}x_1 + a_{n+m2}x_2 + \dots a_{n+mn}x_n + x_{n+m} = b_{n+m} \\ & x_1, x_2, \dots, x_n, x_{n+1}, \dots x_{n+m} \geq 0. \end{array} \tag{3}$$

con $b_{n+1}, b_{n+2}, \dots, b_{n+m} \geq 0$.

La forma aumentada para el ejercicio 5 es:

$$\begin{array}{cc}
\text{Índices} & \\
0 & f(x_1,x_2,\ldots,x_5)-5x_1-4x_2=0 \\
3 & x_1+2x_2+x_3=14=b_3 \\
4 & 3x_1+x_2+x_4=15=b_4 \\
5 & 3x_1+2x_2+x_5=18=b_5 \\
 & x_1,x_2,x_3,x_4,x_5\geq 0.
\end{array} \quad (4)$$

A la vista de la solución gráfica, sabemos la solución es $x_1=2$ y $x_2=6$ y que las variables de holgura toman los valores $x_3=x_5=0$ y $x_4=3$.
La obtención analítica de la solución (sin usar la resolución gráfica) es un problema muy complejo en general. Una idea para resolver el problema es transformar la forma aumentada inicial en una equivalente en la que se pueda deducir fácilmente la solución.

Por ejemplo, veremos en el ejercicio 8 que la forma aumentada original (4) es equivalente a:

$$\begin{array}{cc}
0 & f(x_1,x_2,\ldots,x_5)+(1/2)x_3+(3/2)x_5=34 \\
4 & (3/4)x_3+x_4-(5/4)x_5=3 \\
1 & x_1-(1/3)x_3+(1/2)x_5=2 \\
2 & x_2+(3/4)x_3-(1/4)x_5=6 \\
 & x_1,x_2,x_3,x_4,x_5\geq 0.
\end{array} \quad (5)$$

y teniendo en cuenta que $x_3,\ x_5\geq 0$, se obtiene que el máximo de la función se alcanza si $x_3=x_5=0$, dando el valor 34. Además, $x_3=x_5=0$ generan una solución factible ya que al sustituir en el sistema se deduce que $x_4=3$, $x_1=2$ y $x=6$.
La cuestión ahora es cómo pasar de la formulación (4) a la (5). Un procedimiento para hacerlo es el algoritmo del símplex que se introduce a continuación.

El algoritmo del símplex

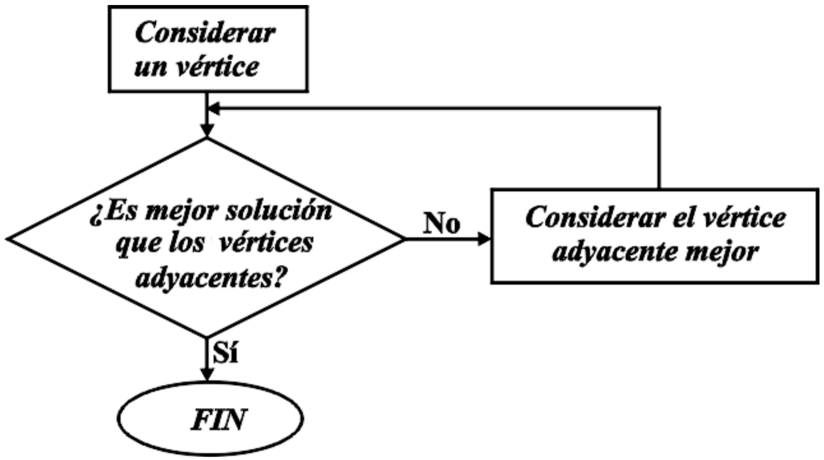
El algoritmo del símplex permitirá simplificar la formulación aumentada de un problema en base a ciertas propiedades geométricas:

1. Si hay una sola solución óptima, es un vértice de la región factible.
2. Si hay varias soluciones óptimas, al menos dos son vértices.
3. Sólo hay un número finito de vértices.
4. Si el valor en un vértice es mejor que en todos los adyacentes, entonces es óptimo (igual o mejor que todos los demás vértices).

Las propiedades 1 y 2 significan que la búsqueda de la solución óptima se puede reducir a la consideración de los vértices, por 3 sólo existe un número finito (el proceso se acaba) y 4 proporciona una prueba de optimalidad (criterio de parada).
Para asegurarnos de que el problema va a tener solución óptima basta suponer que la región factible es no vacía y acotada.
El método símplex se basa en las propiedades anteriores: consiste en trasladarse de un vértice a otro adyacente mejor hasta que no se encuentre uno mejor.

El algoritmo del símplex

El procedimiento se resume como sigue:



El algoritmo del símplex

- Una vez que tenemos la forma ampliada, un vértice de la región factible se obtiene fijando unas variables con el valor 0 y el resto con los valores que se obtienen resolviendo el sistema correspondiente (siempre y cuando estos sean valores positivos).
- Si tenemos un vértice, donde hemos fijado unas variables con el valor cero, cambiar a otro consecutivo, significa seguir fijando las mismas variables menos una con el valor cero y fijar una **nueva variable** con el valor cero.
- Si todos los coeficientes c_i de la función objetivo son positivos, el máximo se alcanza cuando las variables correspondientes a los $c_i > 0$ son cero (el resto de variables lo que se obtenga en el sistema).
- El algoritmo descrito a continuación está basado en estas propiedades.

El algoritmo de simplex

- Se parte del vértice inicial $x^{(0)} = (0, \dots, 0, b_{n+1}, \dots, b_{n+m})$
- Mientras algún coeficiente c_i sea negativo, se realiza lo siguiente:
  1. Se toma el índice j correspondiente a la variable de la función objetivo cuyo coeficiente c_j sea más pequeño (negativo).
  2. Se toma el índice i tal que
  $$\frac{b_i}{a_{ij}} = \text{mín}\left\{ \frac{b_k}{a_{kj}} \text{ con } a_{kj} > 0 \text{ y } k > 0 \text{ índice de ecuación } \right\}$$
  3. Se multiplica la ecuación i por $\frac{1}{a_{ij}}$, se cambia el índice de esta ecuación por j y con esta ecuación, se elimina la variable x_j del resto de ecuaciones y de la función objetivo.
  4. El nuevo vértice se obtiene haciendo cero las variables cuyo índice no coincide con los índices de las ecuaciones y dando el valor actual de b_j a las variables cuyos índices coinciden con los índice de las ecuaciones.

Problema 6.6

La WYNDOR GLASS CO. produce artículos de vidrio de alta calidad, incluyendo ventanas y puertas de vidrio. Tiene tres plantas. Los marcos y molduras de aluminio se hacen en la primera, los marcos de madera en la segunda y en la tercera se produce el vidrio y se ensamblan los productos.
Debido a que las ganancias se han reducido, la gerencia ha decidido reorganizar la línea de producción: Se dejarán de fabricar varios productos no rentables y se dejará libre una parte de la capacidad de producción para fabricar uno o dos productos nuevos que han tenido demanda. Uno de los productos propuestos (Producto 1) es una puerta de vidrio de 8 pies con marco de aluminio. El otro (Producto 2) es una ventana grande (4× 6 pies) para vidrio doble con marco de madera. El departamento de mercadotecnia ha llegado a la conclusión de que puede vender todo lo que pueda producir de ambos productos. Sin embargo, como ambos compiten por la misma capacidad de producción en la planta tercera, no es obvio qué mezcla de los dos sería más rentable. Por todo esto, la gerencia pidió al departamento de I.O. que estudiara el asunto.

Problema 6.6

Después de hacer algunas investigaciones el departamento determinó:

1. *La capacidad de producción de cada planta disponible para estos productos.*
2. *La capacidad que requiere cada unidad producida.*
3. *La ganancia unitaria de cada producto.*

obteniéndose la siguiente tabla:

Planta	Producto 1	Producto 2	Capacidad disponible
Primera	1	0	4
Segunda	0	2	12
Tercera	3	2	18
Ganancia	3	5	

- *Formula el problema como un problema de programación lineal y resuélvelo gráficamente.*
- *Resuelve el problema por el método del símplex.*

El algoritmo del símplex

La aplicación del método del símplex puede simplificarse utilizando una tabla en la que se registra sólo la información esencial, es decir, coeficientes de las variables, constantes del lado derecho de las ecuaciones e índices de las ecuaciones.

Problema 6.7

Aplica el algoritmo del símplex en forma tabular al Ejercicio 6

Solución del Ejercicio 6

Forma tabular para el método del símplex:

Índices	x_1	x_2	x_3	x_4	x_5	$=$	mín
0							
0							
0							

El algoritmo del símplex

Problema 6.8

Aplica el algoritmo del símplex en forma tabular al Ejercicio 5

Partimos de la forma aumentada de dicho algoritmo que expresada en forma tabular será:

Índices	x_1	x_2	x_3	x_4	x_5	$=$	mín
0	-5	-4	0	0	0	0	
3	1	2	1	0	0	14	$\frac{14}{1}=14$
4	3	1	0	1	0	15	$\frac{15}{3}=5$
5	3	2	0	0	1	18	$\frac{18}{3}=6$

Solución del Ejercicio 8

Índices	x_1	x_2	x_3	x_4	x_5	$=$	mín
0							
0							
0							

Contenidos

Apéndice

Aplicando el método de símplex se puede detectar si el problema inicial tiene solución única (como en los ejemplos vistos en la sección anterior), solución múltiple (si el método permite seguir iterando indefinidamente) o, incluso, si no tiene solución (por ejemplo si los coeficientes de la variable a eliminar son todos negativos). Se presenta a continuación, mediante dos ejemplos, los casos de infinitas soluciones e inexistencia de solución:

Ejemplo con infinitas soluciones

Problema 6.9

Aplique El algoritmo del símplex al siguiente problema de programación lineal y explique los resultados obtenidos.

$$\begin{array}{rc} \textit{Maximizar} & f(x_1,x_2) = x_1 + x_2 \\ \textit{sujeto a} & -x_1 + 2x_2 \leq 4 \\ & 2x_1 + 2x_2 \leq 14 \\ & x_1 - x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{array}$$

Ejemplo con infinitas soluciones

Aplicando el método símplex la tabla óptima se obtiene:

Índices	x_1	x_2	x_3	x_4	x_5	$=$	mín
0	-1	-1	0	0	0	0	
3	-1	2	1	0	0	4	-4
4	2	2	0	1	0	14	7
5	1	-1	0	0	1	3	3
0	0	-2	0	0	1	3	
3	0	1	1	0	1	7	7
4	0	4	0	1	-2	8	2
1	1	-1	0	0	1	3	-3
0	0	0	0	$1/2$	0	7	
3	0	0	1	$-1/4$	$3/2$	5	$10/3$
2	0	1	0	$1/4$	$-1/2$	2	-4
1	1	0	0	$1/4$	$1/2$	5	10

El vértice donde se obtiene el máximo es: $(5,2,5,0,0)$ Como se puede observar en el ejercicio, una de las variables que se fijan con el valor cero, x_5, tiene el coeficiente de la función objetivo nulo ($c_5 = 0$), por tanto, se podría continuar El algoritmo del símplex tomando como $j = 5$ y obtendríamos un nuevo vértice donde la función objetivo tendría el mismo valor.

Índices	x_1	x_2	x_3	x_4	x_5	$=$	mín
0	0	0	0	$1/2$	0	7	
5	0	0	$2/3$	$-1/6$	1	$10/3$	
2	0	1	$1/3$	$1/6$	0	$11/3$	
1	1	0	$-1/3$	$1/3$	0	$10/3$	

El nuevo vértice es $(10/3, 11/3, 0, 0, 10/3)$. Por tanto, tenemos soluciones múltiples. Si continuamos el algoritmo volviendo a fijar con valor cero la variable x_5, se tendría un ciclo infinito.
Son soluciones cualquier combinación convexa de las soluciones óptimas obtenidas con el símplex, que formarán los vértices de la región solución.

$$\text{Sol} = (5,2) + t((10/3, 11/3) - (5,2)) = (1-t)(5,2) + t(10/3, 11/3)$$

Cuando hay soluciones múltiples se reconoce porque hay variables fijadas con valor 0 cuyo coeficiente en la función objetivo es 0.

También se puede detectar con el método de símplex que no existe solución, por ejemplo, debido a que la región factible no es acotada:

Problema 6.10

Aplique El algoritmo del símplex al siguiente problema de programación lineal y explique los resultados obtenidos.

$$\begin{array}{rc} \textit{Maximizar} & f(x_1,x_2) = x_1 + 3x_2 \\ \textit{sujeto a} & x_1 - 2x_2 \leq 4 \\ & -x_1 + x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{array}$$

Aplicando el método símplex la tabla óptima es:

Índices	x_1	x_2	x_3	x_4	$=$	mín
0	-1	-3	0	0	0	
3	1	-2	1	0	4	
4	-1	1	0	1	3	3
0	-4	0	0	3	9	
3	-1	0	1	2	10	
2	-1	1	0	1	3	

Para continuar el algoritmo, tendríamos que usar la variable x_1 para eliminarla del resto de ecuaciones, pero no se puede usar porque sus coeficiente en las ecuaciones de las restricciones son negativos. Esto significa que la región factible no está acotada.
En general, cuando los coeficientes de las restricciones para la variable x_j con la que se realiza la eliminación gaussiana son todos menores o iguales que cero, la solución no está acotada.

En esta sección vamos a estudiar cómo adaptar otros modelos cuando no están escritos en la forma adecuada.
Hasta ahora se ha explicado un método para resolver un problema que se encuentra dado en la forma:

$$\begin{array}{ll} \text{Maximizar} & f(x) \\ \text{sujerto a} & Ax \leq b \\ & x \geq 0, b \geq 0. \end{array}$$

es decir, maximizar $f(x)$ sujeta a restricciones de la forma $\leq$, restricciones de no negatividad para las variables y $b_i \geq 0, \forall i$.
Vamos a ver como hacer ajustes para convertir cualquier problema en uno escrito en la forma adecuada. Dichos ajustes deben ser realizados siempre en el paso inicial.

La función objetivo es Min $f(x)$.
Puede resolverse de dos formas:

- Teniendo en cuenta que Min $f(x)$ es equivalente a Max $-f(x)$, y aplicar el método símplex con la nueva función objetivo.
- Modificando el criterio de selección de la variable x_j. En lugar de usar la variable con coeficiente más negativo y poder usar sólo las variables con coeficiente negativo, ahora se usaría la variable con coeficiente más positivo y serían candidatas las variables con coeficientes positivos.

Restricciones distintas de $\leq b_i, b_i \geq 0$. Método de penalización.
Mediante el algoritmo símplex que hemos descrito se pueden resolver problemas lineales en los que las restricciones vengan dadas de la forma $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \leq b_i$ con $b_i \geq 0$, al reescribir dichas restricciones mediante variables de holgura como $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n + x_3 = b_i, x_3 \geq 0$ donde $b_i \geq 0$.
El problema que se nos plantea es como reescribir estas restricciones como igualdades en el caso en el que no vengan expresadas de la forma $\leq b_i$ con $b_i \geq 0$. Vamos a ir comentándolo paso a paso:

- Si la restricción es de la forma $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \geq b_i$ con $b_i \leq 0$, sin más que multiplicar por -1 tenemos la ecuación en la forma estándar.

$$x_1 + x_2 \geq -2 \Longrightarrow -x_1 - x_2 \leq 2 \Longrightarrow -x_1 - x_2 + x_3 = 2, x_3 \geq 0.$$

- Si la restricción es de la forma $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \geq b_i$ con $b_i \geq 0$, debemos restar una variable de holgura, con lo que queda $a_{n+i1}x_1 + a_{n+i2}x_2 + \ldots + a_{n+in}x_n - x_{n+i} = b_{n+i}$. Esto no permite hallar una solución básica factible ($x_{n+i} = -b_{n+1}$!!!), por lo que es necesario añadir una nueva variable que vamos a llamar **variable artificial**, de forma que la restricción queda $a_{n+i1}x_1 + a_{n+i2}x_2 + \ldots + a_{n+in}x_n - x_{n+i} + A_i = b_{n+i}, x_{n+i}, A_i \geq 0$.

$$x_1 + x_2 \geq 2 \Longrightarrow x_1 + x_2 - x_3 = 2, x_3 \geq 0 \Longrightarrow$$

$$x_1 + x_2 - x_3 + A_1 = 2, x_3, A_1 \geq 0.$$

Ahora debemos asegurar que $-x_3 + A_1 \leq 0$, para que la restricción inicial se cumpla. Para no imponer una nueva restricción que complique el problema, lo que se hace es procurar que A_1 valga 0. Para ello se transforma la función objetivo a $f(x_1,\ldots,x_n,x_{n+1},\ldots,x_{n+m},A_1,\ldots,A_p) = c_1x_1 + c_2x_2 + \ldots + c_nx_n - MA_i$, donde M es una constante positiva muy grande. Ahora se aplicaría el método símplex como siempre, salvo un paso previo que consiste en anular el coeficiente de las variables artificiales para así obtener la forma de eliminación de Gauss. A_i no se fija con el valor cero en la primera iteración pero con la nueva función objetivo f se asegura que en alguna etapa pasa a ser una de las variables que se fijan con valor cero.
Ilustramos lo anterior con un nuevo ejercicio.

Problema 6.11

Aplique el algoritmo del símplex al siguiente problema de programación lineal y explique los resultados obtenidos.

$$\begin{array}{ll} \textit{Minimizar} & f(x_1,x_2) = x_1 + 2x_2 \\ \textit{sujeto a} & -3x_1 + 2x_2 \leq 6 \\ & x_1 + x_2 \leq 10 \\ & -x_1 + 3x_2 \geq 6 \\ & x_1, x_2 \geq 0 \end{array} \Rightarrow \begin{array}{ll} \textit{Maximizar} & f(x_1,x_2) = -x_1 - 2x_2 \\ \textit{sujeto a} & -x_1 + 3x_2 \geq 6 \\ & -3x_1 + 2x_2 \leq 6 \\ & x_1 + x_2 \leq 10 \\ & x_1, x_2 \geq 0 \end{array}$$

$$\begin{array}{ll} \textit{Índices} & \\ 0 & f(x_1,x_2,x_3,x_4,x_5,A_1) + x_1 + 2x_2 + MA_1 = 0 \\ A_1 & -x_1 + 3x_2 - x_3 + A_1 = 6 \\ 3 & -3x_1 + 2x_2 + x_4 = 6 \\ 4 & x_1 + x_2 + x_5 = 10 \\ & x_1,x_2,x_3,x_4,x_5,A_1 \geq 0 \end{array}$$

Aplicando el método símplex la tabla óptima es:

Índices	x_1	x_2	x_3	x_4	x_5	A_1	$=$	mín
0	1	2	0	0	0	M	0	
A_1	-1	3	-1	0	0	1	6	
3	-3	2	0	1	0	0	6	
4	1	1	0	0	1	0	10	
0	$1+M$	$2-3M$	M	0	0	0	$-6M$	
A_1	-1	3	-1	0	0	1	6	2
3	-3	2	0	1	0	0	6	3
4	1	1	0	0	1	0	10	10
0	$5/3$	0	$2/3$	0	0	$-2/3+M$	-4	
2	$-1/3$	1	$-1/3$	0	0	$1/3$	2	
3	$-7/3$	0	$2/3$	1	0	$-2/3$	2	
4	$4/3$	0	$1/3$	0	1	$-1/3$	8	

Por tanto la solución es $(0,2)$. El valor de la f.objetivo es 4 (el valor opuesto del que nos da el método).

Debe tenerse en cuenta, ver por ejemplo el libro de Bazaraa, que si una vez obtenida una solución óptima alguna de las variables artificiales aparece entre la variables que no se fijan con valor cero y tiene valor positivo, el problema original no tiene soluciones factibles.

- Si la restricción es de la forma $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \leq b_i$ con $b_i \leq 0$, multiplicando por -1 tendríamos $-a_{i1}x_1 - a_{i2}x_2 - \ldots - a_{in}x_n \geq -b_i$ con $-b_i \geq 0$, con lo que ya estaríamos en el caso anterior.

$$x_1 + x_2 \leq -2 \Longrightarrow -x_1 - x_2 \geq 2.$$

- Si la restricción es de la forma $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i$ con $b_i \geq 0$, se descompone en dos desigualdades y se procede como en los casos estándar y a) respectivamente.

$$x_1 + x_2 = 2 \Longleftrightarrow \begin{cases} x_1 + x_2 \geq 2 \Longleftrightarrow x_1 + x_2 - x_4 + A_1 = 2 \\ x_1 + x_2 \leq 2 \Longleftrightarrow x_1 + x_2 + x_3 = 2 \end{cases}$$

- Si la restricción es de la forma $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i$ con $b_i \leq 0$, se multiplicaría por -1 y se resolvería aplicando el caso anterior.

$$x_1 + x_2 = -2 \Longrightarrow -x_1 - x_2 = 2 \Longleftrightarrow$$

$$\begin{cases} -x_1 - x_2 \leq 2 \Longleftrightarrow -x_1 - x_2 + x_3 = 2 \\ -x_1 - x_2 \geq 2 \Longleftrightarrow -x_1 - x_2 - x_4 + A_1 = 2 \end{cases}$$

Adaptación de otras formas de modelo

Las variables no son no negativas.

- Supongamos por ejemplo que $x_1 \in \mathbb{R}$, en este caso se redefine el problema tomando x_1' y x_1'' ambas positivas y $x_1 = x_1' - x_1''$ y a continuación aplicar el algoritmo del símplex.
- Supongamos por ejemplo que $x_1 \leq 0$, en este caso se trabaja con la variable $x_1' = -x_1 \geq 0$.

Además de estas posibilidades tenemos combinaciones de las mismas, que se resuelven análogamente.