

INTRODUCCIÓN A LA OPTIMIZACIÓN

CESÁREO RAIMÚNDEZ

1. INTRODUCCIÓN

1.1. Generalidades. Dado un conjunto X y una función $f : X \rightarrow \mathbb{R}$ (la *función-objetiva*), se desea determinar $x^* \in X$ tal que, para todo $x \in X$ valga $f(x) \geq f(x^*)$. La variable x se denomina variable de *decisión* o *control*. Consideraremos apenas casos en que $X \subset \mathbb{R}^n$, definido a través de *restricciones*. o sea, dado un número $m_D + m_I$ de funciones $c_j : \mathbb{R}^n \rightarrow \mathbb{R}$ para $j = 1, \dots, m_D + m_I$ el problema es:

$$\begin{aligned} \text{mín } f(x) & \quad x \in \mathbb{R}^n \\ c_j(x) & \leq 0 \quad j \in D \\ c_j(x) & = 0 \quad j \in I \end{aligned}$$

donde D e I son conjuntos disjuntos de índices, de cardinalidad m_D, m_I respectivamente. Así tendremos m_D restricciones de desigualdad y m_I restricciones de igualdad.

1.2. Clasificación. Entre todas posibles clasificaciones propondremos una que está de acuerdo con la dificultad creciente de su resolución.

1. Problemas sin restricciones ($m_D = m_I = 0, D = I = \emptyset$)
 - a) Problemas cuadráticos: $f(x) = \frac{1}{2}x'Mx - b \cdot x$ (M simétrica $x \times x$)
 - b) Problemas no lineales.
2. Problemas con restricciones lineales.
 - a) Problemas con apenas restricciones de igualdad. ($m_D = 0$)
 - 1) Problemas lineal-cuadráticos: f cuadrática.
 - 2) Problemas no lineales
 - b) Problemas con restricciones de desigualdad
 - 1) Programación Lineal: f lineal y $m_D \geq n - m_I$
 - 2) Programación Cuadrática Lineal: f cuadrática.
 - 3) Problemas no lineales con restricciones lineales.
3. Programación No Lineal
 - a) Con restricciones lineales
 - b) Programación No Lineal *Lato Senso*.
4. Técnicas Modernas de Optimización.
 - a) Programación Dinámica.
 - b) Técnicas Estocásticas Inspiradas en Paradigmas Naturales.
 - 1) Recocido Simulado.
 - 2) Algoritmos Genéticos.
 - 3) Estrategias Evolutivas.
 - 4) Enjambres.

2. APLICACIONES

La Optimización en su sentido más amplio, puede ser aplicada en la resolución de cualquier problema de ingeniería. Algunas aplicaciones en diversos sectores de la ingeniería son.

1. Proyecto de estructuras aero-espaciales para mínimo peso.
2. Determinación de trayectorias óptimas en vehículos.
3. Proyecto de estructuras civiles tales como puentes, torres, chimeneas, fundaciones, represas etc, para coste mínimo.
4. Proyecto de peso mínimo para estructuras sometidas a terremotos, vendavales, y otros tipos de perturbaciones aleatorias de gran impacto.
5. Proyecto de sistemas de distribución de recursos acuíferos para proporcionar beneficios máximos.
6. Proyecto de componentes mecánicos de desempeño óptimo en algún sentido.
7. Selección de condiciones de fabricación para coste mínimo de producción.
8. etc.

3. PRELIMINARES

3.1. Regularidad de una función. La regularidad representa el buen comportamiento de la función relativamente a su suavidad. El hecho de afirmar que una función f es regular debe entenderse como: *La función f es continua y posee todas sus derivadas continuas en grado suficiente*

3.2. Gradiente y Hessiana. Para una función regular

$$f(x) : x \in \mathbb{R}^n \rightarrow \mathbb{R}$$

existe asociado un campo vectorial llamado campo gradiente, designado por ∇f que se calcula como

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

El campo gradiente evaluado en un punto genérico, apunta a la dirección en que la función crece.

La matriz hessiana se calcula como sigue

$$H(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial^2 x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial^2 x_2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial^2 x_n} \end{pmatrix}$$

La matriz hessiana calculada en un punto, confirma la concavidad de la función en ese punto.

3.3. Aproximación de Taylor de una función regular. Desarrollando en series la función $f(x)$ en una vecindad regular de x_0 se obtiene

$$f(x) = f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)'H(x_0)(x - x_0) + O^3(x - x_0)$$

3.4. Máximos y mínimos locales. Una función regular f posee un máximo local en x_0 si cumple las condiciones:

$$\begin{aligned}\nabla f(x_0) &= 0 \\ H(x_0) &\prec 0\end{aligned}$$

El símbolo $\prec 0$ indica que $H(x_0)$ es definida negativa.

Análogamente para que ocurra un mínimo deben cumplirse las condiciones

$$\begin{aligned}\nabla f(x_0) &= 0 \\ H(x_0) &\succ 0\end{aligned}$$

El símbolo $\succ 0$ indica que $H(x_0)$ es definida positiva.

3.5. Ejemplo 1. La función $f(x) = x_1^2 - x_2^2$ en el punto $x_0 = (0, 0)$ no posee máximo o mínimo local a pesar de que $\nabla f(x_0) = 0$.

3.6. Ejemplo 2. Sea el cálculo de los puntos estacionarios de la función

$$z = 3e^{-x^2-(y+1)^2}(1-x)^2 - \frac{1}{3}e^{-(x+1)^2-y^2} - 10e^{-x^2-y^2}\left(-y^5 - x^3 + \frac{x}{5}\right)$$

Resolviendo $\nabla z = 0$ o sea las ecuaciones

$$\begin{aligned}-\frac{2}{3}e^{-x^2-2x-(y+1)^2}(-e^{2y}(x+1) + 9e^{2x}(x^3 - 2x^2 + 1) + e^{2x+2y+1}(30xy^5 + 30x^4 - 51x^2 + 3)) &= 0 \\ -\frac{2}{3}e^{-x^2-2x-(y+1)^2}(9e^{2x}(y+1)(x-1)^2 - e^{2y}y + 3e^{2x+2y+1}y(10x^3 - 2x + 5y^3(2y^2 - 5))) &= 0\end{aligned}$$

tendremos como soluciones

k	$\{x_k, y_k\}$	autovalores	tipo
1	$\{1.09827, 0.854461\}$	$\{11.1484, -10.014\}$	inflexión
2	$\{-1.3474, 0.204519\}$	$\{15.5714, 8.5517\}$	mínimo
3	$\{-0.265938, 0.46674\}$	$\{15.2026, -13.875\}$	inflexión
4	$\{1.28568, -0.00484756\}$	$\{-15.7511, -7.11291\}$	máximo
5	$\{-0.00931758, 1.58137\}$	$\{-32.4281, -16.2793\}$	máximo
6	$\{0.416323, -0.394145\}$	$\{20.8892, -8.17109\}$	inflexión
7	$\{0.228279, -1.62553\}$	$\{30.1822, 17.9197\}$	mínimo
8	$\{1.09827, 0.854461\}$	$\{11.1484, -10.014\}$	inflexión
9	$\{-0.460025, -0.629197\}$	$\{-30.413, -14.4383\}$	máximo

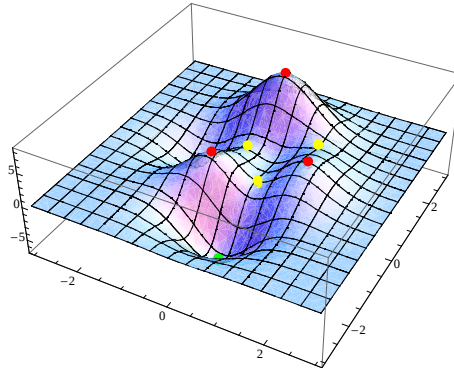


FIGURA 1. Ubicación de los puntos estacionarios

3.7. Procedimientos iterativos para alcanzar puntos estacionarios. Si $f(x)$ es lo suficiente regular tendremos

$$f(x) \approx f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)'H(x_0)(x - x_0)$$

Sus puntos estacionarios cumplen

$$\nabla f(x) = 0$$

pero

$$\nabla f(x) \approx \nabla f(x_0) + H(x_0)(x - x_0) = 0$$

lo que puede utilizarse como procedimiento iterativo para su determinación. Así

$$x_{k+1} = x_k - H^{-1}(x_k)\nabla f(x_k)$$

3.8. Con la ayuda del gradiente. Sea el problema sin restricciones

$$\min_{x \in R^n} f(x)$$

Adoptaremos una estrategia apoyada en el vector gradiente y sus propiedades

1. Escoger un punto de partida.
2. Escoger una dirección $\hat{v}$ que apunte a valores decrecientes de f .
3. Caminar a lo largo de esta dirección mientras se reduzca el valor de f .
4. Verificar si todavía se puede reducir el valor de f a lo largo de una dirección nueva.

Un algoritmo posible para este procedimiento puede ser:

- 1 x_0 punto inicial
- 2 $\hat{v}_0$ dirección inicial
- 3 $\lambda_1 = \arg \min_{\lambda} f(x_0 + \lambda \hat{v}_0)$ escoger el próximo punto
- 4 $x_1 = x_0 + \lambda_1 \hat{v}_0$
- 5 si $f(x_1) < f(x_0)$ entonces $x_0 \leftarrow x_1, \hat{v}_0 \leftarrow \nabla f(x_1) / \|\nabla f(x_1)\|$ sigue para 3
- 6 para

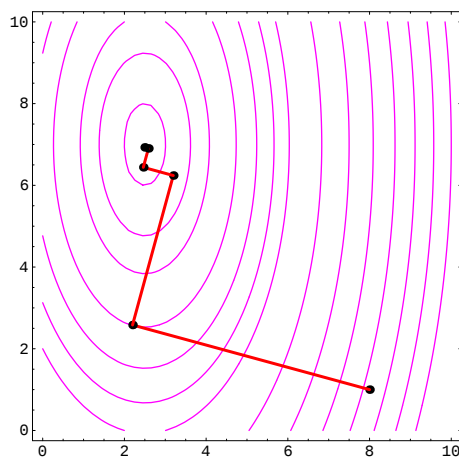


FIGURA 2. Evolución del Algoritmo del gradiente

3.9. Con la ayuda del gradiente y la hessiana. Sea el problema sin restricciones que se sabe poseer un mínimo local

$$\min_{x \in \mathbb{R}^n} f(x)$$

Adoptaremos una estrategia apoyada en el vector gradiente, la hessiana y sus propiedades.

- Escoger un punto de partida x_0 en la vecindad del mínimo.
- Calcular el gradiente $\nabla f(x_0)$ y la hessiana $H(x_0)$ haciendo $k = 0$.
- Calcular $x_{k+1} = x_k - H^{-1}(x_k)\nabla f(x_k)'$.
- Si $|x_{k+1} - x_k| < \epsilon_1$ y $|\nabla f(x_k)| < \epsilon_2$ para caso contrario, $k \leftarrow k + 1$ y salta para el paso anterior.

Si el punto obtenido obedece $H(x_k) \succ 0$ tendremos localizado un punto de mínimo local.

4. CONVEXIDAD

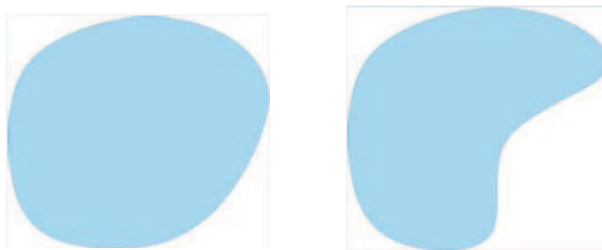


FIGURA 3. Conjuntos convexo y cóncavo

4.1. Propiedad fundamental de convexidad. Para cualesquiera puntos $x_1, x_2, \dots, x_n$ pertenecientes a un conjunto convexo C y cualesquiera parámetros $\lambda_1, \lambda_2, \dots, \lambda_n$ todos no negativos, tales que $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$ vale que $\sum_{k=1}^n \lambda_k x_k \in C$.

4.2. Funciones convexas. Una función $f(x)$ es convexa cuando el conjunto definido por $C = \{(x, y) | f(x) \leq y\}$ es un conjunto convexo.

4.3. Propiedad de las funciones convexas.

- El conjunto de los puntos de mínimo de una función convexa es convexo.
- La hessiana de una función convexa es siempre $H(x) \succeq 0$.

Ejercicio Muestre que la función

$$f(x) = x_1^4 + 2x_2^4 + x_1^2 - x_1x_2 + x_2^2 + 2x_2x_3 + 2x_3^2 - 9x_2 - 16x_3$$

es una función convexa.

5. RESTRICCIONES DE IGUALDAD - MULTIPLICADORES DE LAGRANGE

Sea inicialmente el problema

$$\min_{x \in R^n} f(x) \text{ con condiciones } g(x) = b$$

Se forma inicialmente la forma Lagrangiana del punto de silla (*ver apéndice*)

$$L(x, \lambda) = f(x) + \lambda(g(x) - b)$$

Se determinan las soluciones de la condición de punto de silla

$$\left. \begin{array}{l} \frac{\partial L}{\partial x} = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0 \\ \frac{\partial g}{\partial x} - b = 0 \end{array} \right.$$

resolviendo estas ecuaciones se obtendrán los puntos de silla que verifican la solución.

5.1. Ejemplo. Calcular

$$\min_{x \in R^n} x_1^2 + x_1 x_2 + x_2^2 \text{ con condiciones } x_1 + x_2 = -1$$

Solución

$$L(x, \lambda) = x_1^2 + x_1 x_2 + x_2^2 + \lambda(x_1 + x_2 + 1)$$

en seguida

$$\begin{array}{rcl} 2x_1 + x_2 + \lambda & = & 0 \\ x_1 + 2x_2 + \lambda & = & 0 \\ x_1 + x_2 + 1 & = & 0 \end{array}$$

Resolviendo el sistema se obtiene

$$x_1^* = -\frac{1}{2}, \quad x_2^* = -\frac{1}{2}, \quad \lambda^* = \frac{3}{2}$$

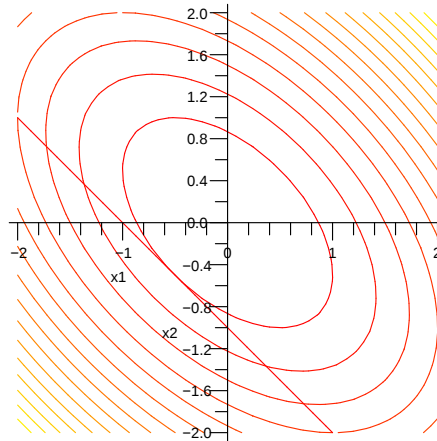


FIGURA 4. ilustración del problema

El valor mínimo de f es: $f_{min} = f(x^*) = f(-1/2, -1/2) = 3/4$. El valor de λ^* es tal que

$$\nabla f(x^*) = -\lambda^* \nabla g(x^*)$$

También se pueden resolver problemas sencillos utilizando técnicas de procesamiento simbólico como por ejemplo el `toolbox` simbólico de **MATLAB**. Veamos a guisa de ejemplo, el problema de determinar las dimensiones de la caja que contiene máximo volumen con la superficie de sus caras limitada al valor S . La lagrangiana para este problema es

$$L(x, y, \lambda) = x^2 y + \lambda(x^2 + 4xy - S)$$

y el programa que lo resuelve es:

```
syms x y p S
L = x^2*y + p*(x^2+4*x*y-S)
equ1 = diff(L,x)
equ2 = diff(L,y)
equ3 = diff(L,p)

sol = solve(equ1,equ2,equ3,x,y,p);

valx = simple(sol.x)
valy = simple(sol.y)
valp = simple(sol.p)
```

5.2. Multiplicadores de Lagrange - Refinamiento de Soluciones. Los resultados obtenidos con el criterio de Lagrange para problemas con restricciones de igualdad pueden ser refinados efectuando un procedimiento iterativo basado en la expansión de Taylor en un entorno del punto de silla solución.

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) = f(x) + \lambda^T g(x)$$

efectuando la expansión en series de Taylor en un entorno del punto (x^k, λ^k) tendremos:

$$\begin{aligned} L(x, \lambda) = & L(x^k, \lambda^k) + \frac{\partial L}{\partial x}(x - x^k) + \frac{\partial L}{\partial \lambda}(\lambda - \lambda^k) \\ & + \frac{1}{2}(x - x^k)^T \frac{\partial^2 L}{\partial x^2}(x - x^k) + \frac{1}{2}(\lambda - \lambda^k)^T \frac{\partial^2 L}{\partial \lambda^2}(\lambda - \lambda^k) \\ & + (x - x^k)^T \frac{\partial^2 L}{\partial x \partial \lambda}(\lambda - \lambda^k) + O^3(x - x^k, \lambda - \lambda^k) \end{aligned}$$

o en forma más compacta

$$\begin{aligned} L(x, \lambda) = & L(x^k, \lambda^k) + \begin{pmatrix} \nabla f + \lambda^T \nabla g & g \end{pmatrix} \begin{pmatrix} x - x^k \\ \lambda - \lambda^k \end{pmatrix} \\ & + \frac{1}{2} \begin{pmatrix} x - x^k \\ \lambda - \lambda^k \end{pmatrix}^T \begin{pmatrix} \nabla^2 f + \lambda^T \nabla^2 g & \nabla g \\ \nabla g^T & 0 \end{pmatrix} \begin{pmatrix} x - x^k \\ \lambda - \lambda^k \end{pmatrix} + O^3(x - x^k, \lambda - \lambda^k) \end{aligned}$$

Calculando los puntos estacionarios de esta aproximación cuadrática de la Lagrangiana, llegamos a:

$$\begin{pmatrix} \nabla f + \lambda^T \nabla g & g \end{pmatrix} + \begin{pmatrix} x - x^k \\ \lambda - \lambda^k \end{pmatrix}^T \begin{pmatrix} \nabla^2 f + \lambda^T \nabla^2 g & \nabla g \\ \nabla g^T & 0 \end{pmatrix} = 0$$

finalmente obteniéndose el proceso iterativo

$$\begin{pmatrix} x \\ \lambda \end{pmatrix}_{k+1} = \begin{pmatrix} x \\ \lambda \end{pmatrix}_k - \begin{pmatrix} \nabla^2 f + \lambda^T \nabla^2 g & \nabla g \\ \nabla g^T & 0 \end{pmatrix}_k^{-1} \begin{pmatrix} \nabla f + \lambda^T \nabla g \\ g \end{pmatrix}_k$$

6. EJEMPLOS DE APLICACIÓN

6.1. Problema - 1. Determinar la configuración de equilibrio asumida por una catenaria formada de n barras uniformes de un material con masa m_i y largura L_i concatenadas consecutivamente entre dos puntos P_a y P_b . La conexión entre barras se efectúa sin rozamiento.

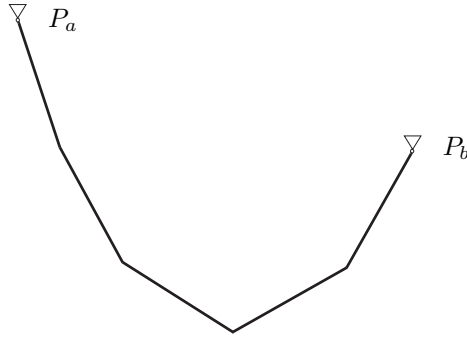


FIGURA 5. Catenaria 5 barras

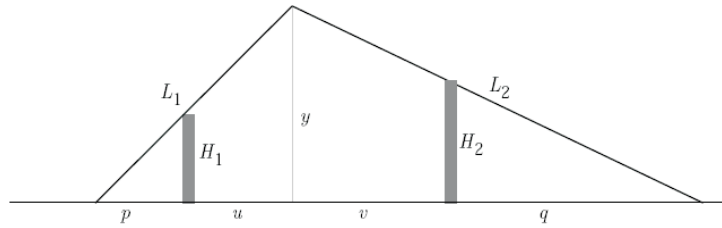


FIGURA 6. Las dos escaleras

6.2. Problema - 2. Un río plagado de cocodrilos debe ser atravesado por un aventurero que cuenta para tal propósito con dos escaleras de largura L_1 y L_2 respectivamente. Las escaleras deben ser apoyadas en los muros que circundan el río por sus márgenes, con alturas H_1 y H_2 respectivamente. Las escaleras deben apoyarse en sus extremidades y en los muros conforme la figura. Se pide dimensionar las escaleras más cortas que pueden realizar este propósito. Se conocen: $H_1 = 10$, $H_2 = 15$, $u + v = 50$.

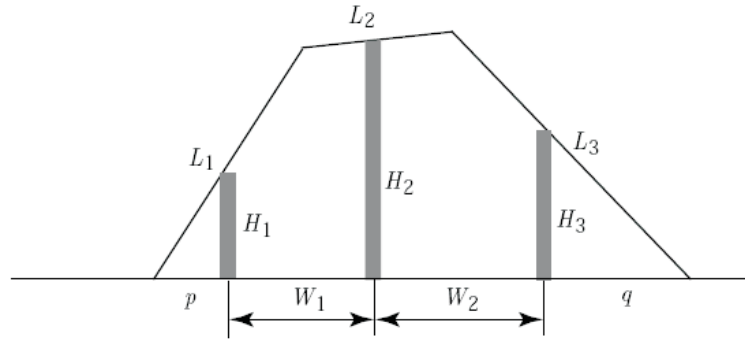


FIGURA 7. Las tres escaleras

6.3. Problema - 3. Resolver el problema de la misma familia del anterior, conforme presentado en la figura 2. Los valores conocidos son: $H_1 = 10$, $H_2 = 20$, $H_3 = 15$, $W_1 = 25$, $W_2 = 30$.

7. APÉNDICE

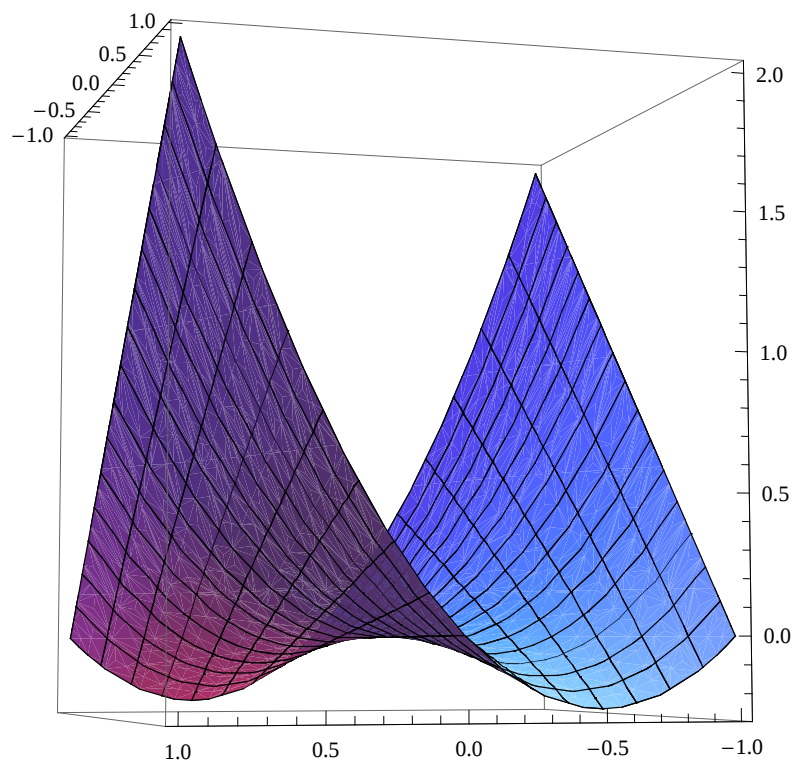
A título de ilustración sigue la gráfica de punto de silla asociada al problema

$$\min x^2 \text{ con la restricción } x = 0$$

El punto de silla de la lagrangiana

$$L(x, \lambda) = x^2 + \lambda x$$

puede observarse en la figura que sigue

FIGURA 8. Forma de silla de la gráfica de $L(x, \lambda)$