Módulo 4: Introducción a programación No Lineal y Análisis Convexo

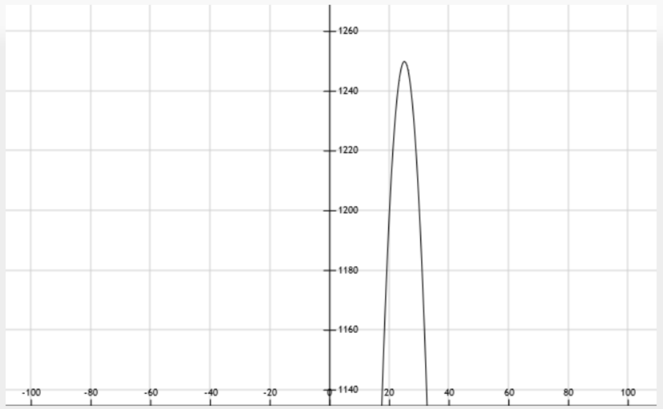
Departamento MACC
Matemáticas Aplicadas y Ciencias de la Computación
Universidad del Rosario

Primer Semestre de 2021

Agenda

Introducción (cont.)

- $f(p) = 100p - 2p^2$



Introducción (cont.)

- $f(p) = 100p - 2p^2$
- Máximo beneficio (sin restricciones)?
- Primera derivada: $\frac{df}{dp} = 100 - 4p = 0$
- $\Rightarrow p^* = \frac{100}{4}$
- Máximo? Mínimo?
- $\Rightarrow$ Segunda derivada: $\frac{d^2 f}{dp^2} = -4 < 0$
- Beneficio máximo: $B_t^* = 100 * 25 - 2 * 25^2 = 1250$

Introducción (cont.)

- Sin restricciones, una variable
- Máximo/mínimo: primera derivada (condición de primer orden)
- Máximo o mínimo?: segunda derivada (condición de segundo orden)
- Múltiples variables?
- Restricciones?

Agenda

1. Introducción
2. Definiciones
3. Condiciones de optimalidad local
   - Condición necesaria de primer orden
4. Introducción
   - Condición necesaria de segundo orden
   - Condición suficiente de segundo orden

Optimización restringida por conjuntos y no restringida

$$\min\ f(x)$$
$$\text{s.a. } x \in \Omega$$

- $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$
- $\min\ f(x) = \max\ -f(x)$
- Restricción de conjunto: $x \in \Omega$
- Ejemplos:
  - $\Omega = \{x : Ax = b, x \geq 0\}$
  - $\Omega = \{x : h(x) = 0, g(x) \leq 0\}$ $(h, g : \mathbb{R}^n \to \mathbb{R})$
  - $\Omega = \mathbb{R}^n$: no restringida

Mínimo local y global

Definición (Mínimo local)

$x^* \in \Omega$ *es un mínimo local de* f *en* Ω *si* $\exists \epsilon > 0 : f(x) \geq f(x^*)$, $\forall x \in \Omega \backslash \{x^*\} : ||x - x^*|| < \epsilon$

Definición (Mínimo global)

$x^* \in \Omega$ *es un mínimo global de* f *en* Ω *si* $f(x) \geq f(x^*)$, $\forall x \in \Omega \backslash \{x^*\}$

x^* mínimo global:

- $f(x^*) = \text{mín}_{x \in \Omega}\{f(x)\}$
- $x^* = \text{argmin}_{x \in \Omega}\{f(x)\}$

Derivadas de $f : \mathbb{R}^n \to \mathbb{R}$ (cont.)

- $Df(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$
- Gradiente de f: $\nabla f(x) = (Df(x))'$

Ejemplo: $f(x) = 5x_1 + 8x_2 + x_1 x_2 - x_1^2 - 2x_2^2$

- $Df(x) = \begin{bmatrix} 5 + x_2 - 2x_1 & 8 + x_1 - 4x_2 \end{bmatrix}$
- $\nabla f(x) = \begin{bmatrix} 5 + x_2 - 2x_1 \\ 8 + x_1 - 4x_2 \end{bmatrix}$

Direcciones factibles

Definición (Dirección factible)

$d \in \mathbb{R}^n$, $d \neq 0$, es una dirección factible en $x \in \Omega$ si $\exists \alpha_0 > 0$:
$x + \alpha d \in \Omega, \forall \alpha \in [0, \alpha_0]$

Definición (Derivada direccional)

Dada una dirección factible d en $x \in \Omega$, la derivada direccional de $f(x)$ en la dirección d es

$$\frac{\partial f(x)}{\partial d} = \lim_{\alpha \to 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha}$$

Direcciones factibles (cont.)

- Dados x, d: $f(x + \alpha d)$ en función de α
- $\frac{\partial f(x)}{\partial d} = \left.\frac{df(x+\alpha d)}{d\alpha}\right|_{\alpha=0}$
- $\frac{\partial f(x)}{\partial d} = Df(x)d = d'\nabla f(x)$
- Si $||d|| = 1$, $Df(x)d$ es la tasa de cambio de f en el punto x en la dirección d

Direcciones factibles - Ejemplo 1

$f(x) = 2x + x^2/2 \qquad \Omega = [-3, 0]$

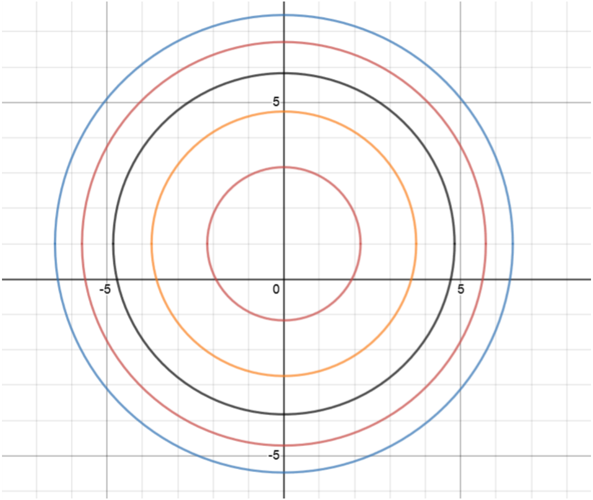
- $\hat{x} = -1,\ d = -1$
- $f(\hat{x} + \alpha d) = -3/2 - \alpha + \alpha^2/2$
- $\left.\frac{\partial f(\hat{x}+\alpha d)}{\partial \alpha}\right|_{\alpha=0} = -1$
- $\nabla f(x) = 2 + x,\ \nabla f(\hat{x}) = 1$
- $d' \nabla f(\hat{x}) = -1$

Direcciones factibles - Ejemplo 2

$f(x_1, x_2) = x_1^2 + (x_2 - 1)^2$ (círculo centrado en $(0, 1)$) $\quad \Omega = \mathbb{R}^2$

- $\hat{x} = (0, 0)$, $d' = [0\ 1]$
- $\nabla f(x) = \begin{bmatrix} 2x_1 \\ 2(x_2 - 1) \end{bmatrix}$
- $\nabla f(\hat{x}) = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$, $d' \nabla f(\hat{x}) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \end{bmatrix} = -2$

Direcciones factibles - Ejemplo 2



Agenda

Condición necesaria de primer orden

Teorema

$\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^1$ *(primera derivada continua). Si* x^* *es un mínimo local de* f *en* Ω*, entonces, para toda dirección factible* d *en* x^*,

$$d'\nabla f(x^*) \geq 0$$

- Equivalente a $\frac{\partial f(x^*)}{\partial d} \geq 0$ (incremento no negativo en toda dirección)

Condición necesaria de primer orden

Corolario

$\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^1$. *Si* x^* *es un mínimo local de* f *en* Ω *y* x^* *es un punto interior, entonces,*

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

- Si $\Omega = \mathbb{R}^n$, todo punto es interior

Condición necesaria de primer orden - Ejemplo 1

$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + 2x_2^2 \qquad \Omega = \mathbb{R}^2$

- ¿Es $\hat{x} = [2\ 3]$ un mínimo local?
- $\nabla f(x) = \begin{bmatrix} 6x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{bmatrix}$
- $\nabla f(\hat{x}) = \begin{bmatrix} 18 \\ 16 \end{bmatrix} \neq 0 \Rightarrow x^*$ no cumple CNPO
- $\Rightarrow x^*$ no es un mínimo local

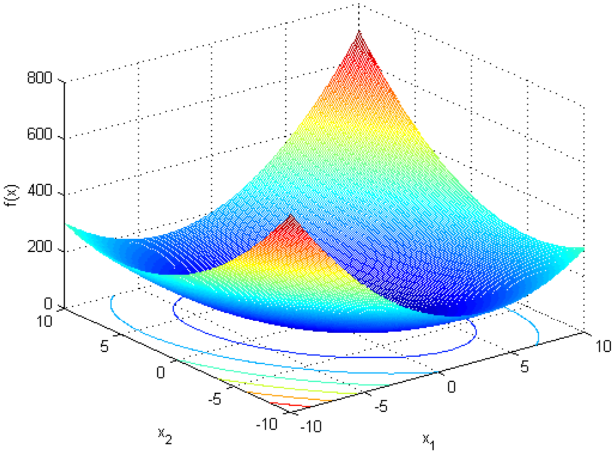
Condición necesaria de primer orden - Ejemplo 1 (cont.)

$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + 2x_2^2 \quad \Omega = \mathbb{R}^2$

- Condición de primer orden: $\nabla f(x) = 0$
- $\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- $\Rightarrow x = [0\ 0]$ cumple CNPO (único candidato a ser mínimo local de f en Ω)

Condición necesaria de primer orden - Ejemplo 1 (cont.)

$$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + 2x_2^2 \qquad \Omega = \mathbb{R}^2$$



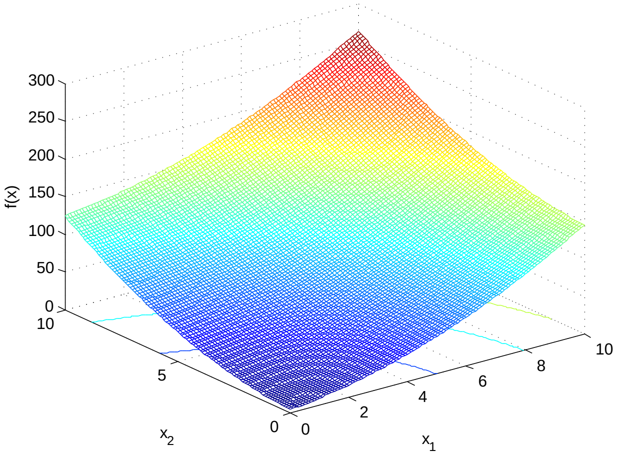
Condición necesaria de primer orden - Ejemplo 2

$$f(x_1, x_2) = (x_1 + 2)^2 + (x_2 + 1)^2 \qquad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$$

- ¿Es $\hat{x} = [0\ 0]$ un mínimo local?
- $\nabla f(x) = \begin{bmatrix} 2(x_1 + 2) \\ 2(x_2 + 1) \end{bmatrix}$
- $\nabla f(\hat{x}) = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$
- $d' \nabla f(\hat{x}) = 4d_1 + 2d_2 \geq 0$ dado que $d_1 \geq 0$ y $d_2 \geq 0$
- $\Rightarrow \hat{x}$ cumple CNPO (candidato a ser mínimo local de f en Ω)

Condición necesaria de primer orden - Ejemplo 2 (cont.)

$$f(x_1, x_2) = (x_1 + 2)^2 + (x_2 + 1)^2 \qquad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$$



Agenda

Derivadas de $f : \mathbb{R}^n \to \mathbb{R}$ (cont.)

Ejemplo: $f(x) = 5x_1 + 8x_2 + x_1x_2 - x_1^2 - 2x_2^2$

- $Df(x) = \begin{bmatrix} 5 + x_2 - 2x_1 & 8 + x_1 - 4x_2 \end{bmatrix}$
- $\nabla f(x) = \begin{bmatrix} 5 + x_2 - 2x_1 \\ 8 + x_1 - 4x_2 \end{bmatrix}$
- $H(x) = \begin{bmatrix} -2 & 1 \\ 1 & -4 \end{bmatrix}$

Condición necesaria de segundo orden

Teorema

$\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^2$. Si x^* es un mínimo local de f en Ω, y d una dirección factible en x^* tal que $d'\nabla f(x^*) = 0$, entonces

$$d'H(x^*)d \geq 0$$

Condición necesaria de segundo orden

Corolario

$\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^2$. *Si* x^* *es un mínimo local de* f *en* Ω *y* x^* *es un punto interior, entonces*

$$\nabla f(x^*) = 0 \quad y \quad H(x^*) \text{ es semidefinida positiva}$$

Matrices definidas positivas

Una matriz simétrica $A \in \mathbb{R}^{n \times n}$ es *definida positiva* ($A \succ 0$) si

$$\forall x \in \mathbb{R}^n, \; x'Ax > 0.$$

Otros criterios:

- Determinante de todas las menores principales es mayor que cero
- Todos los eigenvalores son positivos

Equivalente matricial de número positivo

Matrices semidefinidas positivas

Una matriz simétrica $A \in \mathbb{R}^{n \times n}$ es *semidefinida positiva* si

$$\forall x \in \mathbb{R}^n, \; x' A x \geq 0.$$

Otros criterios:

- Todos los eigenvalores son no-negativos

Para una matriz simétrica semidefinida positiva, los determinantes de todas las menores principales deben ser no-negativos. Condición necesaria pero no suficiente.

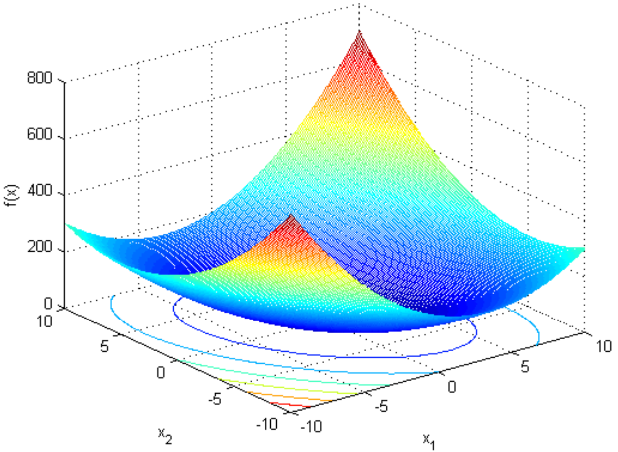
Condición necesaria de segundo orden - Ejemplo 1

$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + 2x_2^2 \quad \Omega = \mathbb{R}^2$

- Condición de primer orden: $\nabla f(x) = \begin{bmatrix} 6x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{bmatrix} = 0$
- $\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- $\Rightarrow \hat{x} = [0\ 0]$ único candidato a ser mínimo local
- $H(x) = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \succ 0 \Rightarrow H(\hat{x})$ es definida positiva
- $\Rightarrow \hat{x} = [0\ 0]$ cumple CNSO para ser un mínimo local

Condición necesaria de segundo orden - Ejemplo 1

$$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + 2x_2^2 \ - \ \Omega = \mathbb{R}^2$$



Condición necesaria de segundo orden - Ejemplo 2

$f(x_1, x_2) = (x_1 + 2)^2 + (x_2 + 1)^2 \quad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$

- ¿Es $\hat{x} = [0\ 0]$ un mínimo local?
- $\nabla f(x) = \begin{bmatrix} 2(x_1 + 2) \\ 2(x_2 + 1) \end{bmatrix}$, $H(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$
- $\nabla f(\hat{x}) = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$, $H(\hat{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$
- $d' \nabla f(\hat{x}) = 4d_1 + 2d_2 \geq 0$ dado que $d_1 \geq 0$ y $d_2 \geq 0$
- $d' H(\hat{x}) d = 2d_1^2 + 2d_2^2 \geq 0$, pero no existe $d \neq 0$ tal que $d' \nabla f(\hat{x}) = 0$
- $\Rightarrow$ no puede evaluarse la CNSO para $\hat{x}$

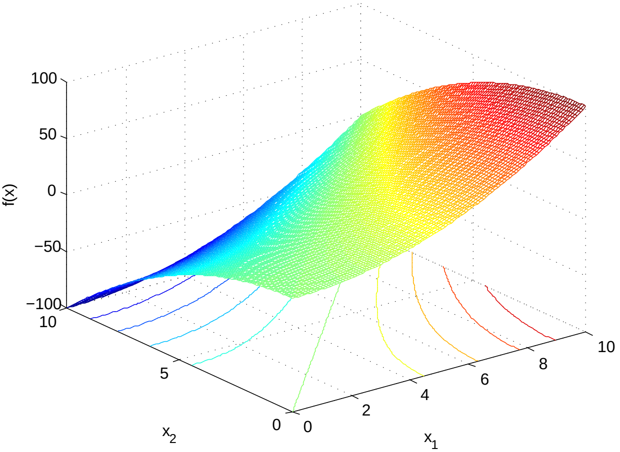
Condición necesaria de segundo orden - Ejemplo 3

$f(x_1, x_2) = x_1^2 - x_2^2 \quad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$

- ¿Es $\hat{x} = [0\ 0]$ un mínimo local?
- $\nabla f(x) = \begin{bmatrix} 2x_1 \\ -2x_2 \end{bmatrix}$, $H(x) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$
- $\nabla f(\hat{x}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $H(\hat{x}) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$
- $d'\nabla f(\hat{x}) = 0 \geq 0$ para toda dirección factible (cumple CNPO)
- $d'H(\hat{x})d = 2d_1^2 - 2d_2^2$, y $d'\nabla f(\hat{x}) = 0$ para toda dirección factible
- $\Rightarrow \hat{x}$ no cumple CNSO, ya que $d'H(\hat{x})d$ puede ser negativo $(d_1 < d_2)$

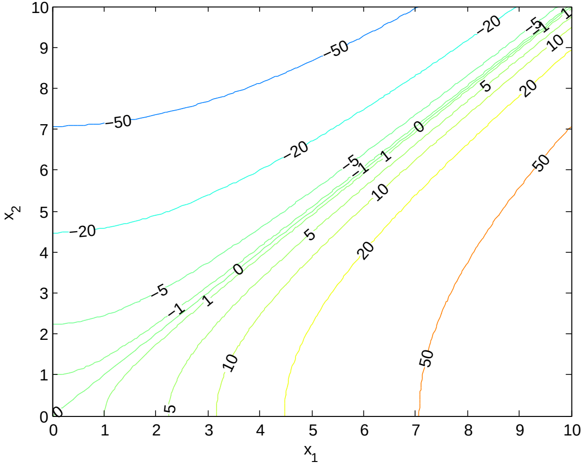
Condición necesaria de segundo orden - Ejemplo 3 (c.)

$$f(x_1, x_2) = x_1^2 - x_2^2 \qquad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$$



Condición necesaria de segundo orden - Ejemplo 3 (c.)

$f(x_1, x_2) = x_1^2 - x_2^2 \quad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$



Condición suficiente de segundo orden

Teorema

$\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^2$ y x^* un punto interior de Ω. Si

1. *$\nabla f(x^*) = 0$, y*
2. *$H(x^*)$ es definida positiva,*

entonces x^* es un mínimo local estricto de f en Ω.

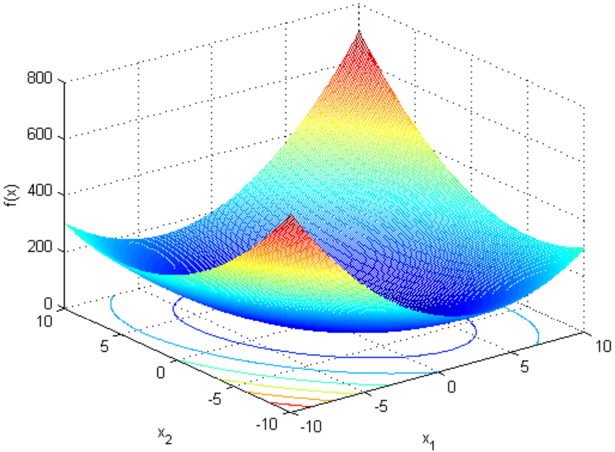
Condición suficiente de segundo orden - Ejemplo 1

$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + 2x_2^2 \quad \Omega = \mathbb{R}^2$

- Condición de primer orden: $\nabla f(x) = 0$
- $\nabla f(x) = \begin{bmatrix} 6x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{bmatrix}$
- $\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- $\Rightarrow x^* = [0\ 0]$ único candidato a ser mínimo local
- $H(x) = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \succ 0 \Rightarrow H(x^*)$ es definida positiva
- $\Rightarrow x^* = [0\ 0]$ cumple CSSO para un mínimo local (es un mínimo loca estricto)

Condición suficiente de segundo orden - Ejemplo 1

$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + 2x_2^2 \qquad \Omega = \mathbb{R}^2$



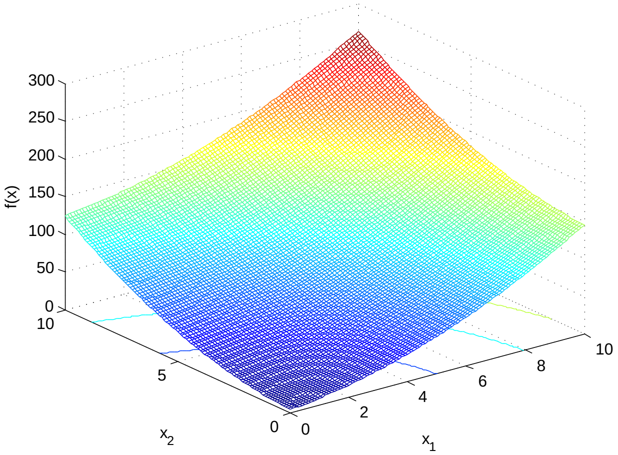
Condición suficiente de segundo orden - Ejemplo 2

$f(x_1, x_2) = (x_1 + 2)^2 + (x_2 + 1)^2 \quad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$

- ¿Es $\hat{x} = [0\ 0]$ un mínimo local?
- $\hat{x}$ no es punto interior $\Rightarrow$ no puede verificarse CSSO

Condición suficiente de segundo orden - Ejemplo 2

$$f(x_1, x_2) = (x_1 + 2)^2 + (x_2 + 1)^2 \qquad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$$



Condición suficiente de segundo orden - Ejemplo 3

$f(x_1, x_2) = x_1^2 - x_2^2 \quad \Omega = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$

- Es $\hat{x} = [0\ 0]$ un mínimo local?
- $\hat{x}$ no cumple CNSO $\Rightarrow$ $\hat{x}$ no es mínimo local
- No hace falta comprobar CSSO

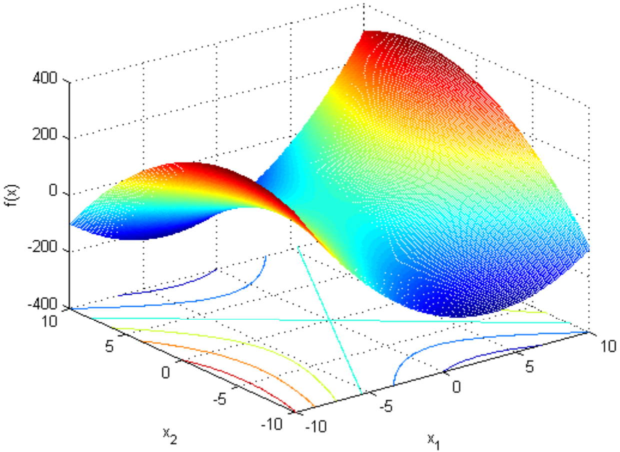
Condición suficiente de segundo orden - Ejemplo 4

$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 - 2x_2^2 \qquad \Omega = \mathbb{R}^2$

- Condición de primer orden: $\nabla f(x) = 0$
- $\nabla f(x) = \begin{bmatrix} 6x_1 + 2x_2 \\ 2x_1 - 4x_2 \end{bmatrix}$
- $\begin{bmatrix} 6 & 2 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- $\Rightarrow x^* = [0\ 0]$ único candidato a ser mínimo local
- $H(x) = \begin{bmatrix} 6 & 2 \\ 2 & -4 \end{bmatrix} \Rightarrow H(x^*)$ no es definida positiva
- $\Rightarrow x^* = [0\ 0]$ NO cumple CSSO para un mínimo local

Condición suficiente de segundo orden - Ejemplo 4

$$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 - 2x_2^2 \qquad \Omega = \mathbb{R}^2$$



Condición suficiente de segundo orden - Ejemplo 4

$$f(x_1, x_2) = 3x_1^2 + 2x_1x_2 - 2x_2^2 \qquad \Omega = \mathbb{R}^2$$