Institución Universitaria
Reacreditada en Alta Calidad

Optimización

Optimización sin restricciones

Docente: Cristian Guarnizo Lemus

Contenido

1. Optimización sin restricciones.
2. Convexidad en optimización.

Optimización sin restricciones suave

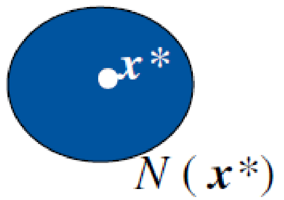
Problema de optimización sin restricciones:

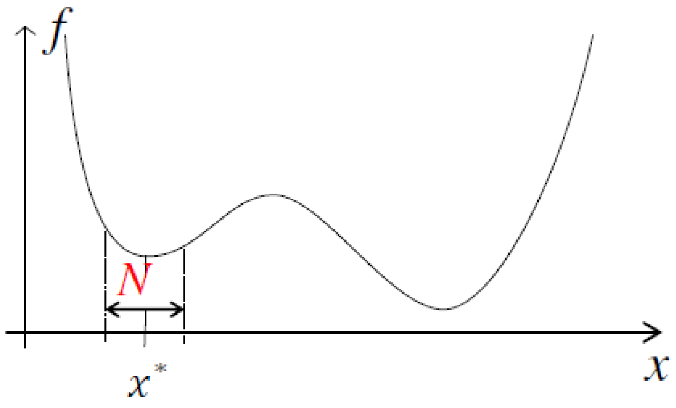
$$\min f(\boldsymbol{x})$$
$$\boldsymbol{x} \in R^n$$

Caso especial cuando el conjunto factible $\Omega = R^n$

Definición:

$\boldsymbol{x}^*$es una solución local si $\exists N(\boldsymbol{x}^*): f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ $\forall \boldsymbol{x} \in N(x^*)$.





Condiciones de optimalidad:

1er orden Necesaria: sí $\boldsymbol{x}^*$ es un mínimo local entonces $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$.

2do orden Necesaria: sí $\boldsymbol{x}^*$ es un mínimo local entonces $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$ y $H(\boldsymbol{x}^*)$ es semidefinida positiva.

2do orden Suficiente: sí $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$ y $H(\boldsymbol{x}^*)$ es definida positiva entonces $\boldsymbol{x}^*$ es un mínimo local.

Ejemplo de aplicación de Condición necesaria

Problema

Encontrar todos los puntos estacionarios de la función

$$f(x) = x_1^4 + x_1^2(1 - 2x_2) + 2x_2^2 - 2x_1x_2 + 4.5x_1 - 4x_2 + 4$$

Y úselos para determinar los mínimos.

$$\boldsymbol{\nabla} f(\boldsymbol{x}) = \begin{bmatrix} \left.\dfrac{\partial f}{\partial x_1}\right|_{\boldsymbol{x}} \\ \left.\dfrac{\partial f}{\partial x_2}\right|_{\boldsymbol{x}} \end{bmatrix} = \begin{bmatrix} 4x_1^3 + 2x_1(1 - 2x_2) - 2x_2 + 4.5 \\ -2x_1^2 + 4x_2 - 2x_1 - 4 \end{bmatrix} = \mathbf{0}$$

$$\Rightarrow \begin{cases} 4x_1^3 + 2x_1(1 - 2x_2) - 2x_2 + 4.5 = 0 \\ -2x_1^2 + 4x_2 - 2x_1 - 4 = 0 \end{cases}$$

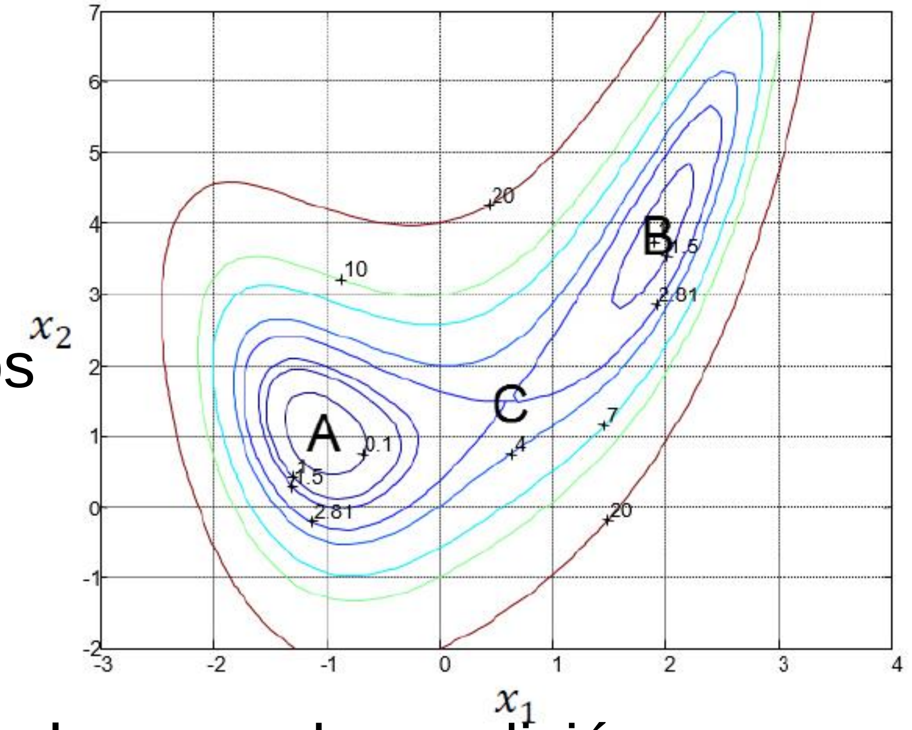
Ejemplo de aplicación de Condición necesaria

- Solucionando el sistema de ecuaciones resultan en los puntos estacionarios

  A (-1.053, 0.9855), B (1.941, 3.854), C (0.6117, 1.4929)

- Para clasificar los puntos estacionarios, investigamos la definición de la Hessiana $H(\boldsymbol{x})$

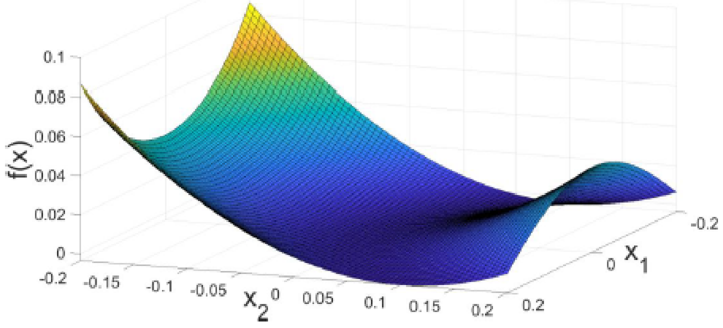
$$\boldsymbol{H}(\boldsymbol{x}) = \begin{bmatrix} 12x_1^2 + 2(1 - 2x_2) & -4x_1 - 2 \\ -4x_1 - 2 & 4 \end{bmatrix}$$



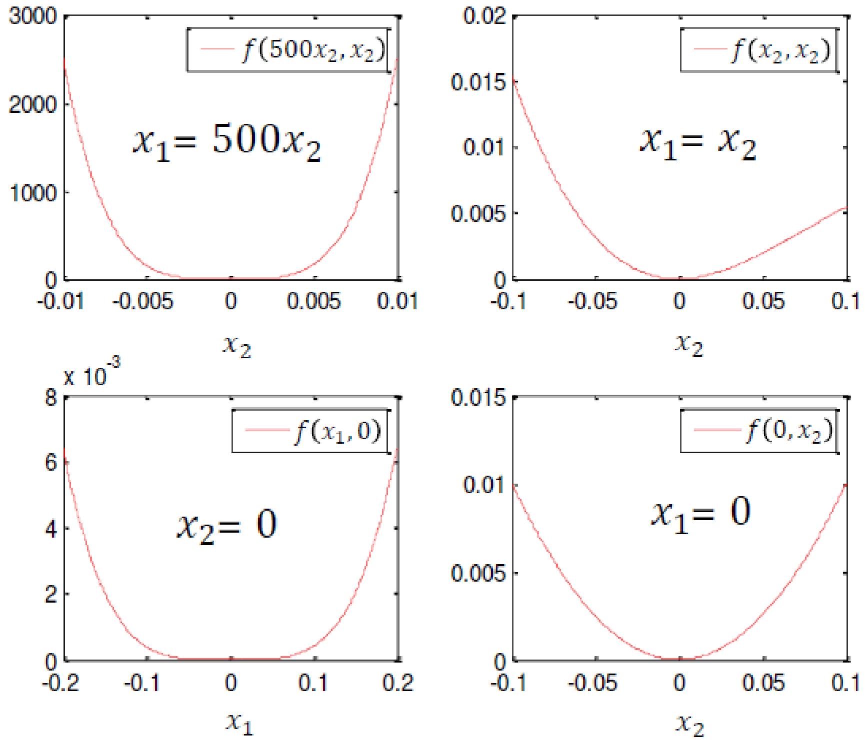
- En A y B todos las valores propios son positivos. Por la segundo condición suficiente A y B son mínimos locales. (A es de hecho un mínimo global)
- En C, la Hessiana tiene un valor propio positivo y uno negativo. Es un punto de ensilladura.

Una función funky

$$f(\boldsymbol{x}) = (x_2 - x_1^2)(x_2 - 4x_1^2)$$



- $$\nabla f(\boldsymbol{x}) = \begin{bmatrix} -10x_1x_2 + 16x_1^3 \\ 2x_2 - 5x_1^2 \end{bmatrix}, \quad \nabla f(\mathbf{0}) = \mathbf{0}$$
- $$\boldsymbol{H}(\boldsymbol{x}) = \begin{bmatrix} -10x_2 + 48x_1^2 & -10x_1 \\ -10x_1 & 2 \end{bmatrix}, \quad \boldsymbol{H}(\mathbf{0}) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

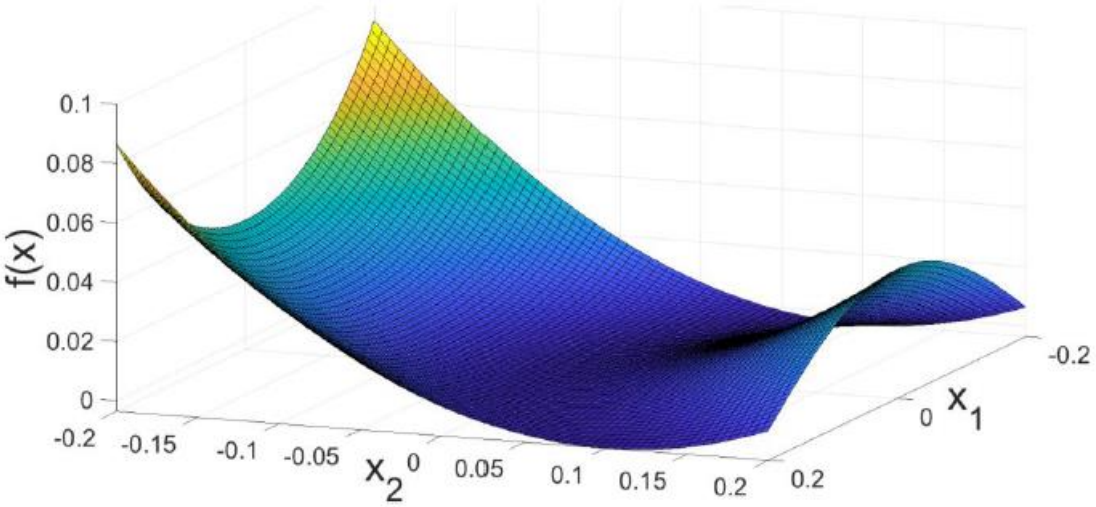


- Se satisfacen las condiciones necesarias (1ro y 2do).
- No se satisfacen las condiciones suficientes.

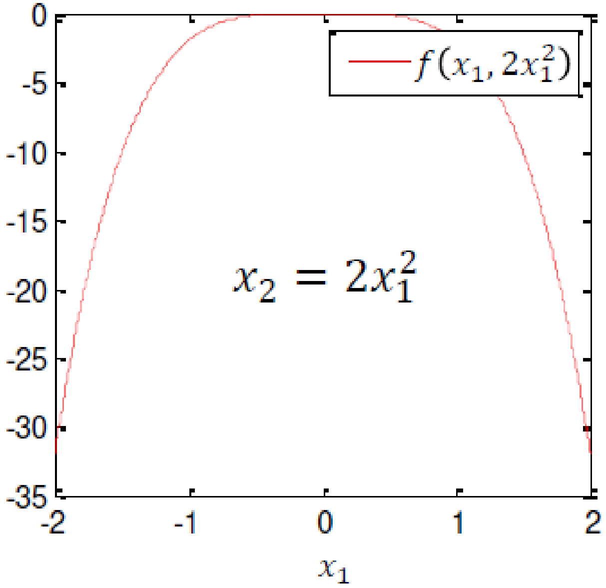
- **0** es mínimo local con respecto a toda línea que lo atraviese.
- **0** no es mínimo local de f.
- Como puede ser?

Una función funky

$$f(\boldsymbol{x}) = (x_2 - x_1^2)(x_2 - 4x_1^2)$$



- Tomar $x_2 = 2x_1^2$. Tenemos $f(x) = f(x_1, 2x_1^2) = -2x_1^4$ y claramente (0,0) no es un mínimo a lo largo de la curva.



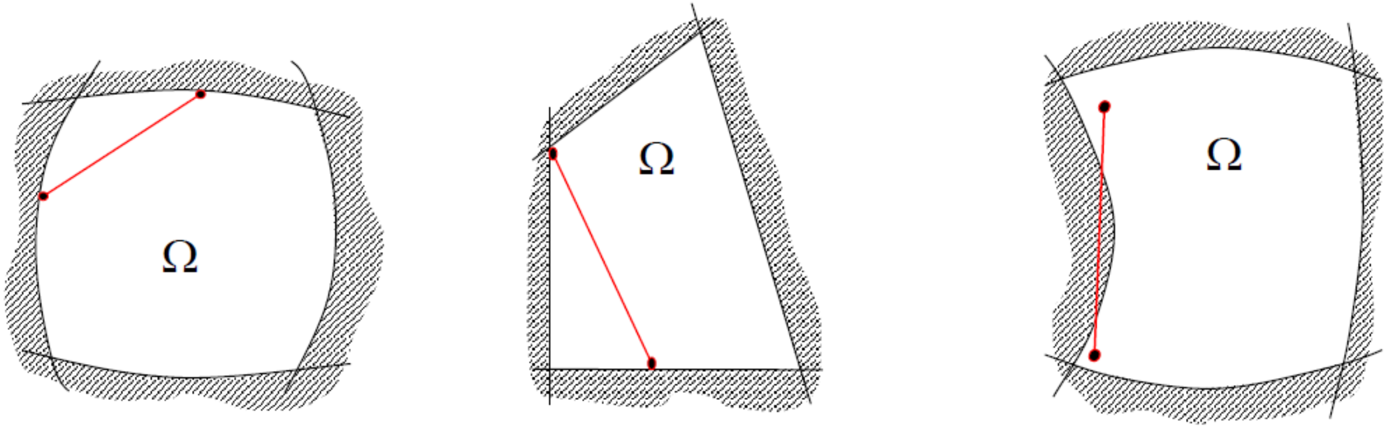
Contenido

1. Optimización sin restricciones.
2. Convexidad en optimización.

Convexidad de un conjunto

Definicion conjunto convexo:

- Un conjunto $\Omega \subseteq R^n$ es convexo, si $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \Omega$ y $\forall \alpha \in [0,1], \alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \Omega$



convex nonconvex

- Las restricciones definen el conjunto factible $\Omega = \{\boldsymbol{x} \in D | c_i(\boldsymbol{x}) \leq 0 \; \forall i \in I, c_i(\boldsymbol{x}) = 0 \forall i \in E\}$.
- La convexidad de Ω hace una gran diferencia en las propiedades teóricas y en una solución numérica.
- Un conjunto es convexo o no convexo.

Convexidad de una función

Definición función convexa: asumir que D es convexo

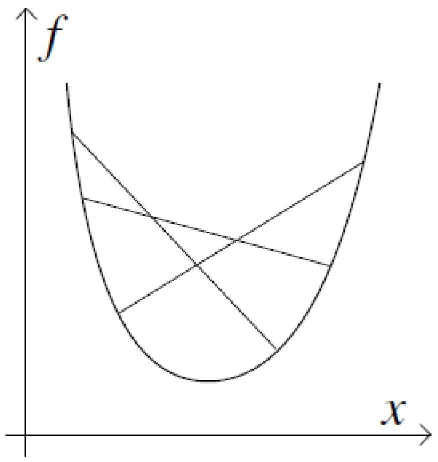
- Una función f es convexa en D, si $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in D$, $\forall \alpha \in [0,1]$:

$$f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2).$$

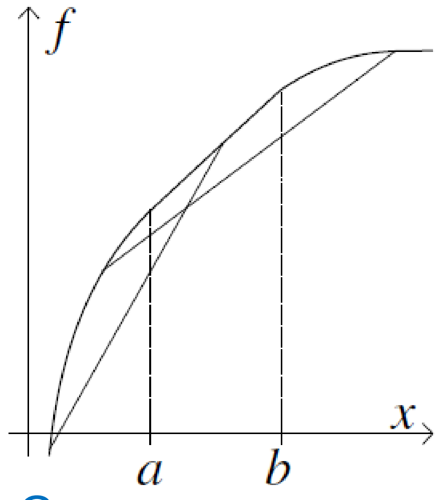
- f es estrictamente convexa en D, si $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in D$, $\forall \alpha \in [0,1]$:

$$f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) < \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2).$$

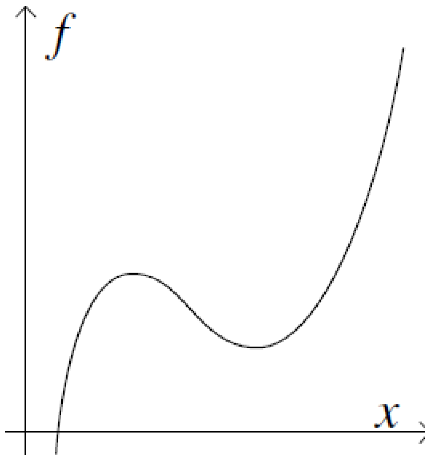
- f es estrictamente concava en D, si $-f$ es estrictamente convexa.



Estrictamente convexa



Concava, pero no estrictamente



Ni convexa, ni concava

Criterio de convexidad

Definición (semidefinido positivo):

- Una matriz $\boldsymbol{A}$ (nxn) simétrica es llamada definida positiva, si $\boldsymbol{p}^\top \boldsymbol{A}\boldsymbol{p} > 0\ \forall \boldsymbol{p} \in R^n, \boldsymbol{p} \neq \boldsymbol{0}$.
- Una matriz $\boldsymbol{A}$ (nxn) simétrica es llamada semi-definida positiva, si $\boldsymbol{p}^\top \boldsymbol{A}\boldsymbol{p} \geq 0\ \forall \boldsymbol{p} \in R^n, \boldsymbol{p} \neq \boldsymbol{0}$.
- Si $(-\boldsymbol{A})$ es positiva (semi-)definida, entonces $\boldsymbol{A}$ es llamada negativa (semi-)definida.

Teorema:

Una matriz $\boldsymbol{A}$ (nxn) simétrica es llamada definida positiva, si $\lambda_k > 0, \forall k \in \{1, \ldots, n\}$ donde λ_k representa el k-ésimo valor propio (eigenvalue) de $\boldsymbol{A}$, o sea, las soluciones de $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$. De manera similar, semi-definida positiva si $\lambda_k \geq 0, \forall k \in \{1, \ldots, n\}$.

Teorema:

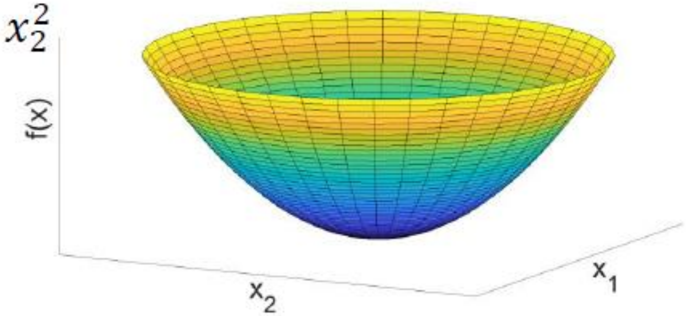
- f es convexa, si y solo si la Hessiana $\boldsymbol{H}(\boldsymbol{x})$ es semi-definida positiva $\forall \boldsymbol{x} \in D$.
- Si $\boldsymbol{H}(\boldsymbol{x})$ es definida positiva $\forall \boldsymbol{x} \in D$, entonces f es estrictamente convexa.

Criterio de convexidad

	if $\forall \boldsymbol{x} \in D$		
f is	$\boldsymbol{H}(\boldsymbol{x})$ is	all λ_k are	$\forall \boldsymbol{p} \in R^n$, $\boldsymbol{p}^T \boldsymbol{H}(\boldsymbol{x})\boldsymbol{p}$ is
strictly convex	positive definite	> 0	> 0
convex	positive semi-definite	≥ 0	≥ 0
strictly concave	negative definite	< 0	< 0
concave	negative semi-definite	≤ 0	≤ 0
neither convex, nor concave	-	≥ 0 or ≤ 0	≥ 0 or ≤ 0
	Definiteness of the Hessian	Sign of the eigenvalues	Sign of the quadratic form

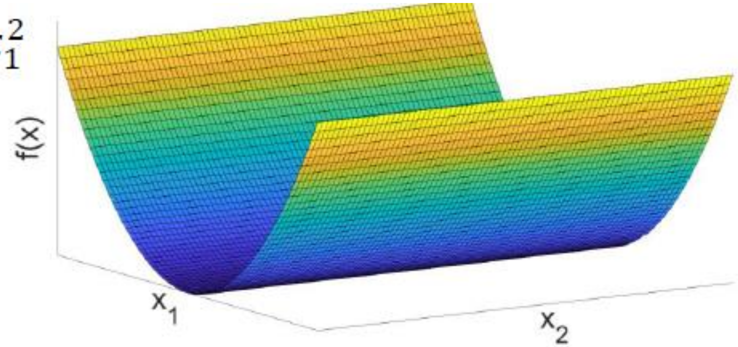
Ejemplos de convexidad

$f(\boldsymbol{x}) = x_1^2 + x_2^2$



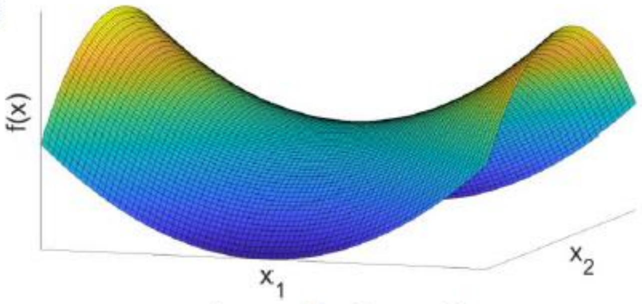
$\lambda_{1,2} > 0$

$f(\boldsymbol{x}) = x_1^2$



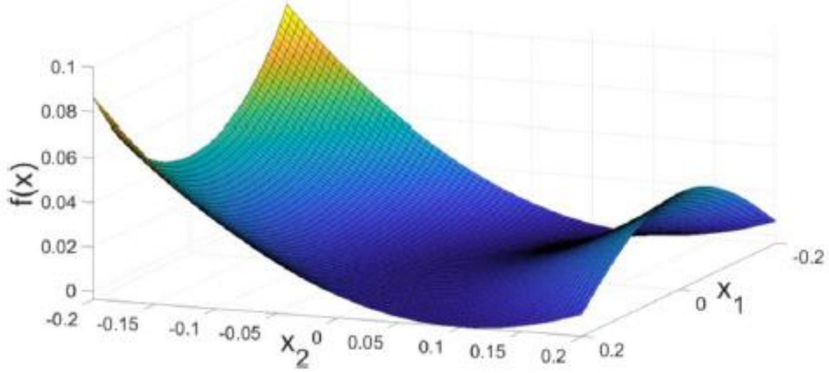
$\lambda_1 > 0,\ \lambda_2 = 0$

$f(\boldsymbol{x}) = x_1^2 - x_2^2$



$\lambda_1 > 0,\ \lambda_2 < 0$

$f(\boldsymbol{x}) = (x_2 - x_1^2)(x_2 - 4x_1^2)$



Problemas de Opt. Convexa

Definición - problema de optimización convexa:

- El problema de optimización $\min f(\boldsymbol{x})$ con $\boldsymbol{x} \in \Omega$ es convexo, si la función objetivo f es convexa y si el conjunto factible Ω es convexo.
- Si D es un conjunto convexo, $c_i \forall i \in I$ con convexas en D y $c_i \forall i \in E$ son lineales entonces
  $\Omega = \{x \in D | c_i(x) \leq 0 \ \forall i \in I, c_i(x) = 0 \ \forall i \in E\}$ es convexo.
- Aparte de algunas excepciones:
  - Ω es no convexo, si cualquier $c_i \forall i \in E$ es una función no-lineal.
  - Ω es no convexo, si cualquier $c_i \forall i \in I$ es no convexa en D.
- “de hecho el mayor punto de inflexión en la optimización no es entre lineal y no-lineal, sino entre convexo y no convexo.” R. Tyrrell Rockafellar in SIAM Review, 1993.

Condiciones de optimalidad para problemas convexos suaves

- Sea f dos veces diferenciables y convexa.
- Debido que f es suave y convexa, $\nabla^2 f(\boldsymbol{x})$ es positiva semi-definida para todo $\boldsymbol{x}$.
- Si $\boldsymbol{x}^* \in R^n$ es un punto de solución local, entonces también es punto de solución global. – Prueba: La convexidad implica que la primera derivada no decrece cuando nos alejamos de $\boldsymbol{x}^*$. Entonces no podemos encontrar otro mínimo local distinto.
- Un punto $\boldsymbol{x}^* \in R^n$ es un punto de solución global si y solo si $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$.

  - La convexidad implica $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*) + \left(\nabla f(\boldsymbol{x}^*)\right)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*)$
  - Con estacionariedad $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$

Condiciones de optimalidad para problemas convexos suaves

- Se puede decir:
  - La condición de primer orden es necesaria y suficiente.
  - Un punto estacionario es equivalente a una solución local y una solución global.
  - En problemas con restricciones la convexidad implica que las condiciones de primer orden son necesarias y suficientes.

Chequeo

- Cuando un problema de optimización es convexo?
- Como podemos verificar la convexidad de un problema de optimización sin restricciones suave?

Referencias

- Basado en el curso “Applied Numerical Optimization” por el profesor Alexander Mitsos.

Institución
Universitaria
Reacreditada en Alta Calidad

¡Gracias!

Somos Innovación Tecnológica con Sentido Humano

Alcaldía de Medellín