



Institución  
**Universitaria**  
Reacreditada en Alta Calidad

# Optimización

## Formulación matemática

**Docente: Cristian Guarnizo Lemus**

Somos Innovación Tecnológica con *Sentido Humano*



Alcaldía de Medellín



# Contenido

1. Definición formal de la optimización
2. Conocimientos matemáticos.
3. Condiciones de optimalidad en problemas suaves sin restricciones.

# Problema de optimización simple

$$\min_x f(x)$$

Función objetivo

$$\text{s.t. } c_i(x) = 0, \forall i \in E \quad \text{Restricciones de igualdad}$$

$$c_i(x) \leq 0, \forall i \in I \quad \text{Restricciones de desigualdad}$$

Ejemplo:

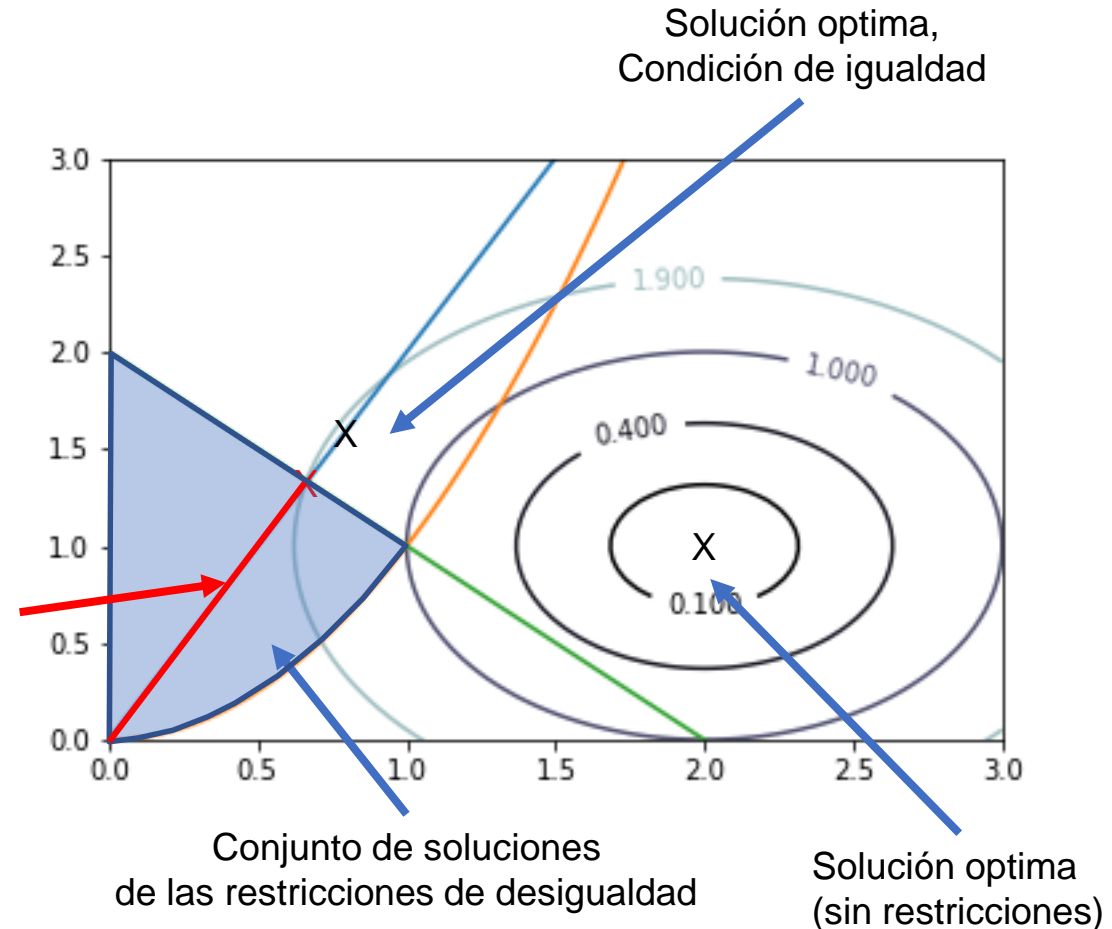
$$\min_x (x_1 - 2)^2 + (x_2 - 1)^2$$

$$\text{s.t. } x_2 - 2x_1 = 0$$

$$x_1^2 - x_2 \leq 0$$

$$x_1 + x_2 \leq 2$$

Conjunto de soluciones  
Factibles RI y RD



# Problema de optimización no-lineal

Formulación general:

$$\min_{x \in D} f(x)$$

$$\text{s.t. } c_i(x) = 0, i \in E$$

$$c_i(x) \leq 0, i \in I$$

$x = [x_1, x_2, \dots, x_n]^T \in D \subseteq R^n$  un vector (punto)

$D$  conjunto anfetrón

$f: D \rightarrow R$  función objetivo

$c_i: D \rightarrow R$  funciones de restricción  $\forall i \in E \cup I$

$E$  el conjunto índice de las restricciones de igualdad

$I$  el conjunto índice de las restricciones de desigualdad

Las restricciones y el conjunto anfetrón define el conjunto factible, esto es, el conjunto de todas las soluciones factibles:

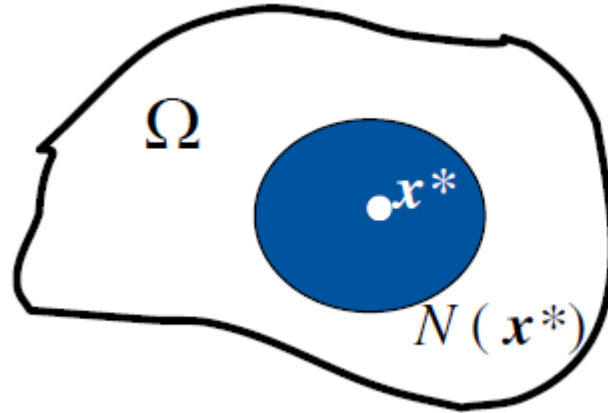
$$\Omega = \{x \in D | c_i(x) \leq 0 \forall i \in I, c_i(x) = 0 \forall i \in E\}$$

Formulación equivalente:

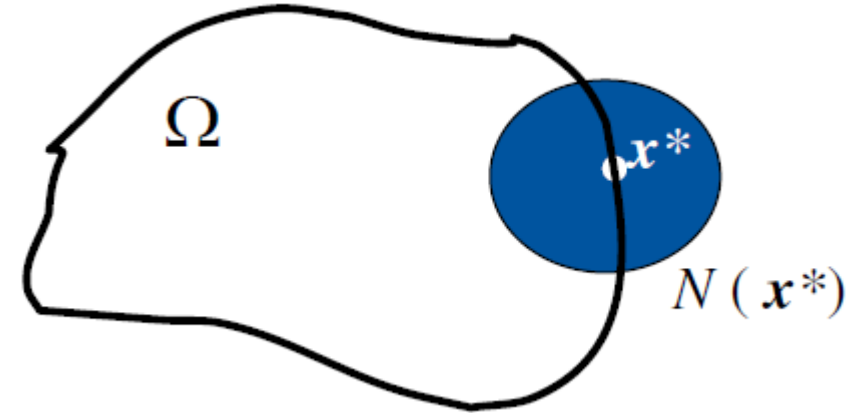
$$\min_{x \in \Omega} f(x)$$

# Cual es una solución optima?

$$\min_{x \in \Omega} f(x)$$



Solución en el interior  
de un conjunto factible



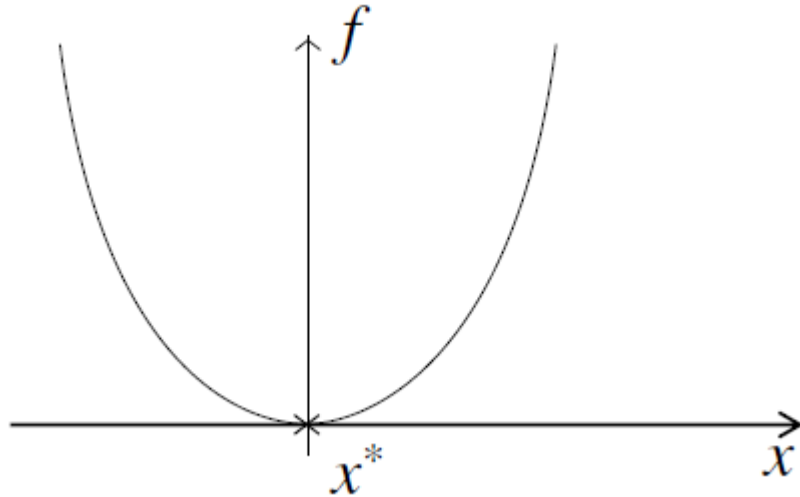
Solución en el perímetro  
del conjunto factible

- a)  $x^*$  es una **solución local** sí  $x^* \in \Omega$  y una vecindad  $N(x^*)$  de  $x^*$  existe:  $f(x^*) \leq f(x) \forall x \in N(x^*) \cap \Omega$
- b)  $x^*$  es una **solución local estricta** sí  $x^* \in \Omega$  y una vecindad  $N(x^*)$  de  $x^*$  existe:  $f(x^*) < f(x) \forall x \in N(x^*) \cap \Omega, x \neq x^*$
- c)  $x^*$  es una **solución global** sí  $x^* \in \Omega$  y  $f(x^*) \leq f(x) \forall x \in \Omega$

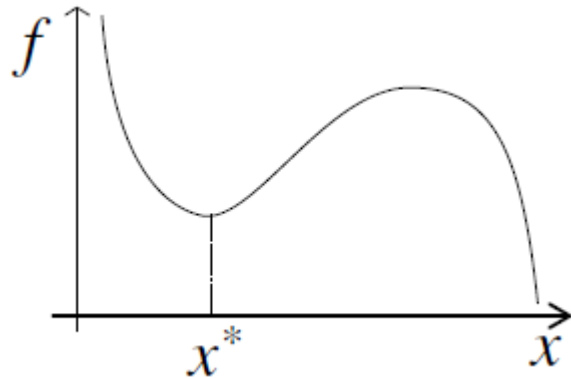


# Solución óptima: ejemplos

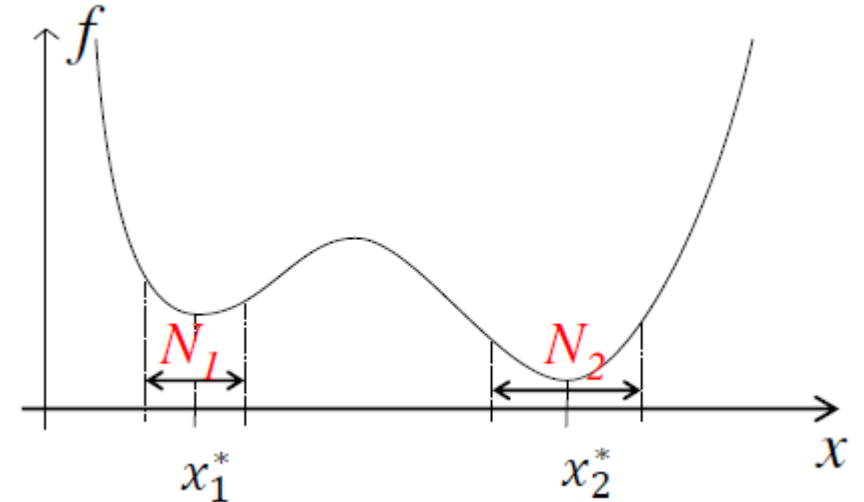
a) Mínimo global estricto



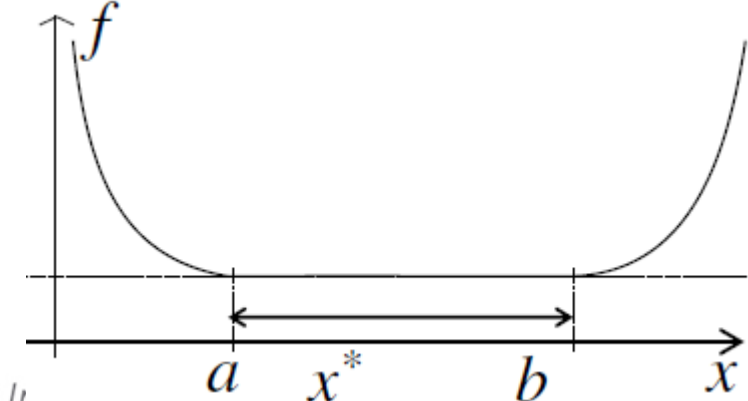
c) Un mínimo local estricto, no hay mínimo global.



b) Dos mínimos locales estrictos, donde uno es un mínimo global estricto.



d) Cada  $x \in [a, b]$  es un mínimo global y local sin mínimo estricto.



# Chequeo

- Escribir la definición general de un problema de optimización.
- Definición de una solución local y global de un problema de optimización?
- Son todas las soluciones locales también una solución global? Son todas las soluciones globales también una solución local?
- Que es el conjunto factible de un problema de optimización?
- Puede una solución ser el punto interior de un conjunto factible? Sobre su perímetro? Por fuera del conjunto factible? – Dibujar una figura explicativa.



# Contenido

1. Definición formal de la optimización
2. **Conocimientos matemáticos.**
3. Condiciones de optimalidad en problemas suaves sin restricciones.



# Problema de optimización no-lineal

Formulación general:

$$\min_{x \in D} f(x)$$

$$\text{s.t. } c_i(x) = 0, i \in E$$

$$c_i(x) \leq 0, i \in I$$

$x = [x_1, x_2, \dots, x_n]^T \in D \subseteq R^n$  un vector (punto)

$D$  conjunto anfiteón

$f: D \rightarrow R$  función objetivo

$c_i: D \rightarrow R$  funciones de restricción  $\forall i \in E \cup I$

$E$  el conjunto índice de las restricciones de igualdad

$I$  el conjunto índice de las restricciones de desigualdad

Las restricciones y el conjunto anfiteón define el conjunto factible, esto es, el conjunto de todas las soluciones factibles:

$$\Omega = \{x \in D | c_i(x) \leq 0 \forall i \in I, c_i(x) = 0 \forall i \in E\}$$

Formulación equivalente:

$$\min_{x \in \Omega} f(x)$$

# Derivada direccional

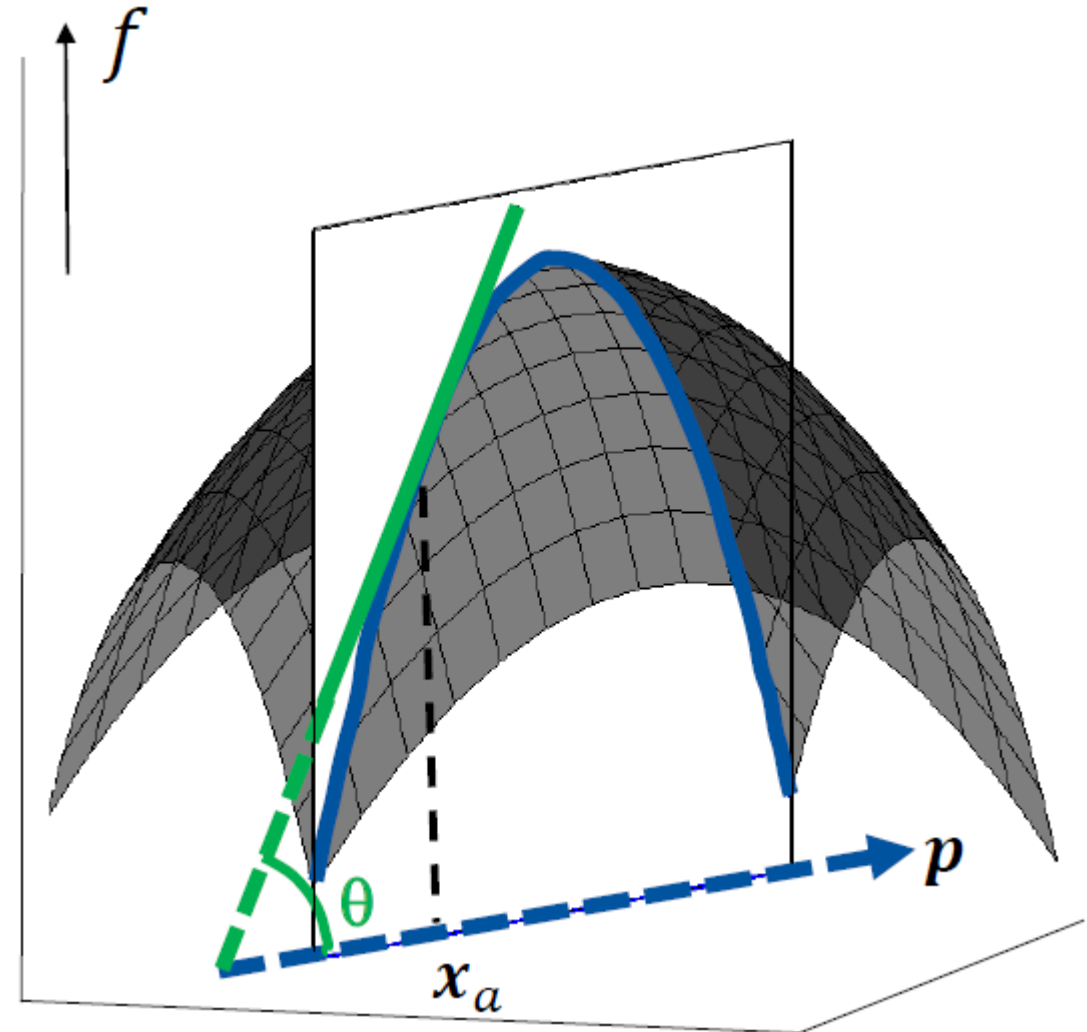
## Definición:

Sea  $f: D \rightarrow R$ ,  $D \subseteq R^n$ ,  $x \in D$  y  $p \in R^n$  con  $\|p\| = 1$ .

$f$  es diferenciable en el punto  $x = x_a$  en la dirección  $p$  si el limite,

$$D(f, p)|_{x=x_a} = \lim_{\varepsilon \rightarrow 0} \frac{f(x_a + \varepsilon p) - f(x_a)}{\varepsilon} =: \nabla_p f(x_a)$$

$D(f, p)$  es llamada la derivada direccional de  $f$  en la dirección  $p$ .



# Gradiente

## Definición:

- La primera derivada de una función continua escalar es llamada gradiente de  $f$  en el punto  $x$ :

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \Big|_x \\ \vdots \\ \frac{\partial f}{\partial x_n} \Big|_x \end{bmatrix}.$$

El vector gradiente también indica la dirección en la cual crece la función (pendiente en el punto  $x$ )

## Observaciones:

- Si  $x$  es función de  $t$ , se aplica la regla de la cadena:

$$\frac{df}{dt} \Big|_{x(t)} = \nabla f(x)^\top \frac{dx}{dt} \Big|_t = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{x(t)} \frac{\partial x_i}{\partial t} \Big|_t.$$

- La derivada direccional esta relacionada con el gradiente:

$$D(f(x), p) = \nabla_p f(x) = \lim_{\varepsilon \rightarrow 0} \frac{f(x + \varepsilon p) - f(x)}{\varepsilon} = \nabla f(x)^\top p$$

en Sentido Humano

# Hessiana (matriz)

Definición:

- La segunda derivada de una función continua escalar diferenciable dos veces es una matriz simétrica llamada Hessiana,  $H(\mathbf{x})$ :

$$H(\mathbf{x}) = \nabla^2 f(\mathbf{x}) = \begin{bmatrix} \left. \frac{\partial^2 f}{\partial x_1^2} \right|_{\mathbf{x}} & \cdots & \left. \frac{\partial^2 f}{\partial x_1 \partial x_n} \right|_{\mathbf{x}} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial^2 f}{\partial x_1 \partial x_n} \right|_{\mathbf{x}} & \cdots & \left. \frac{\partial^2 f}{\partial x_n^2} \right|_{\mathbf{x}} \end{bmatrix}.$$

T

# Condiciones suficientes y necesarias

- **Condición necesaria:** Declaración  $A$  es una condición necesaria para la declaración  $B$  si (y solo si) la **falsedad** de  $A$  garantiza la **falsedad** de  $B$ . En notación matemática:  $\text{not } A \Rightarrow \text{not } B$ .
- **Condición suficiente:** Declaración  $A$  es una condición suficiente para la declaración  $B$  si (y solo si) la verdad de  $A$  garantiza la verdad de  $B$ . En notación matemática:  $A \Rightarrow B$ .

T

# Condiciones suficientes y necesarias

- Si la declaración  $A$  es una condición necesaria para la declaración  $B$ , entonces  $B$  es una condición suficiente para la declaración  $A$ .
  - $\text{not } A \Rightarrow \text{not } B$  implica  $B \Rightarrow A$
- Si la declaración  $A$  es una condición suficiente para la declaración  $B$ , entonces  $B$  es una condición necesaria para la declaración  $A$ .
  - $A \Rightarrow B$  implica  $\text{not } B \Rightarrow \text{not } A$

T



# Condiciones suficientes y necesarias

- En optimización queremos verificar fácilmente las condiciones de un punto candidato
  - es un optimo local (condición suficiente para optimalidad es suficiente)
  - no es una condición optima (condición necesaria es violada)

Idealmente queremos condiciones que son necesarias y suficientes para un optimo local (o mejor aun para global)

T

# Condiciones suficientes y necesarias: ejemplos

- **Ejemplo simple:** sea  $x \in R$  y  $y = x^2$ .

Declaración  $A$  “ $x$  es positivo” y declaración  $B$  “ $y$  es positivo”

-  $A$  es suficiente para  $B$ .

Prueba:  $A$  verdadero  $\Leftrightarrow x > 0 \Rightarrow x^2 > 0 \Rightarrow y > 0 \Leftrightarrow B$  verdadero.

-  $A$  no es necesario para  $B$ .

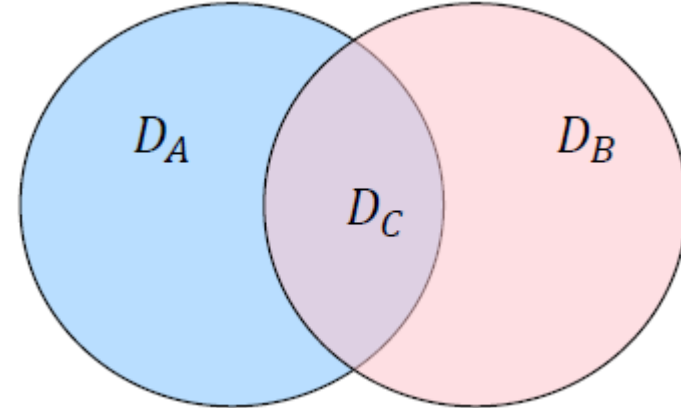
Prueba: contra ejemplo

$x = -1 \Rightarrow y = x^2 = 1$ , entonces  $B$  es verdadero y  $A$  es falso.

T

# Condiciones suficientes y necesarias: ejemplos

- **Ejemplo con conjuntos:** sea  $D_A, D_B \subset R^n$  y  $D_C = D_A \cap D_B$ 
  - Declaración A:  $x \in D_A$
  - Declaración B:  $x \in D_B$
  - Declaración C:  $x \in D_C$
  - A es necesario para C,  
B es necesario para C.
  - C es suficiente para A,  
C es suficiente para B.
  - (A y B) son ambos necesarios  
y suficientes para C.



T

# Chequeo

- Que funciones son continuas, diferenciables, continuas y diferenciables?
- Como se define la derivada direccional de una función? Como esta la derivada parcial relacionada con la derivada direccional?
- Cual es la definición del gradiente y la Hessiana de una función?



# Contenido

1. Definición formal de la optimización
2. Conocimientos matemáticos.
3. Condiciones de optimalidad en problemas suaves sin restricciones.

# Optimización sin restricciones

Problema de optimización sin restricciones:

Caso especial para el cual el conjunto factible  $\Omega = \mathbb{R}^n$

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

- $x^*$  es solución local si  $x^* \in \mathbb{R}^n$  y un vecindario  $N(x^*)$  de  $x^*$  existe:  
 $f(x^*) \leq f(x), \forall x \in N(x^*)$

Queremos condiciones fáciles de evaluar

Necesaria: si  $x^*$  es óptimo entonces las condiciones se satisfacen.

Suficiente: si la condición está satisfecha entonces  $x^*$  es óptimo.

Idealmente ambos son necesarios y suficientes!



# Condiciones necesarias de primer orden

**Teorema (Condiciones necesarias de primer orden):**

Sea  $f$  continuamente diferenciable y sea  $\mathbf{x}^* \in \mathbb{R}^n$  un minimizador local de  $f$ , entonces

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

**Prueba:**

Como  $\mathbf{x}^*$  es un minimizador local de  $f$ , para cada  $\mathbf{p} \in \mathbb{R}^n$ , existe un  $\tau > 0$ , tal que  $f(\mathbf{x}^* + \varepsilon \mathbf{p}) \geq f(\mathbf{x}^*) \forall \varepsilon \in [0, \tau]$

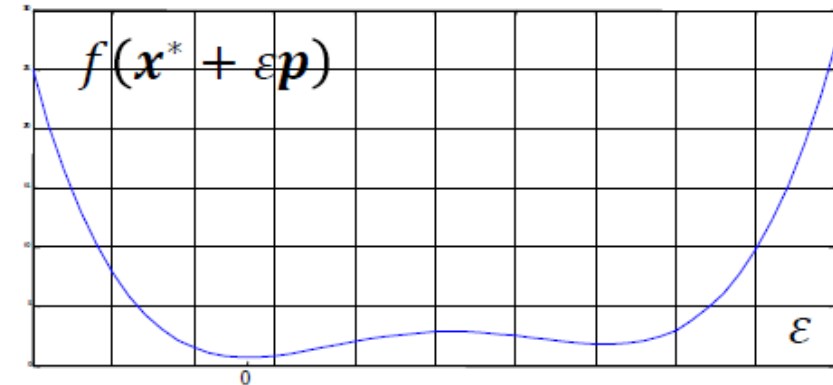
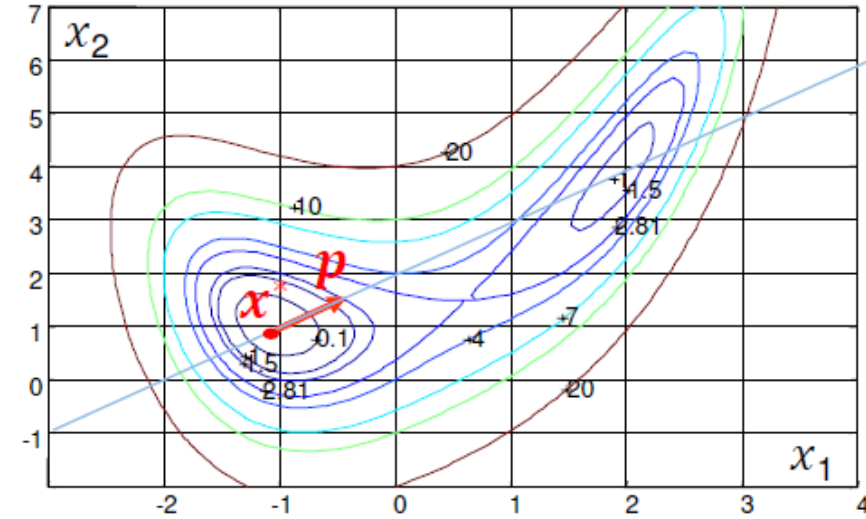
By the definition of the directional derivative:

$$\nabla_{\mathbf{p}} f(\mathbf{x}^*) = \lim_{\varepsilon \rightarrow 0} \frac{f(\mathbf{x}^* + \varepsilon \mathbf{p}) - f(\mathbf{x}^*)}{\varepsilon} = \nabla f(\mathbf{x}^*)^T \mathbf{p} \geq 0 \quad (1)$$

The special choice,  $\mathbf{p} = -\nabla f(\mathbf{x}^*)$ , leads to

$$\nabla f(\mathbf{x}^*)^T \mathbf{p} = -\nabla f(\mathbf{x}^*)^T \nabla f(\mathbf{x}^*) = -\|\nabla f(\mathbf{x}^*)\|^2 \leq 0 \quad (\text{norm property}) \quad (2)$$

(1) and (2)  $\Rightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$ .

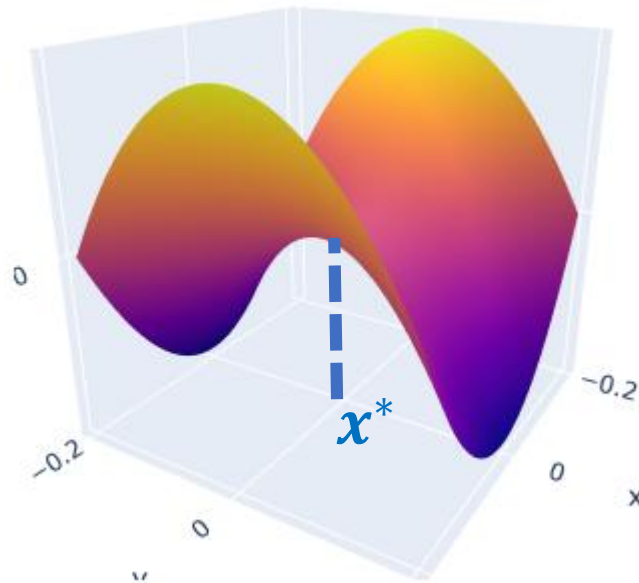


# Puntos estacionarios

- Sea  $f$  continuamente diferenciable y  $x \in R^n$ . Si  $\nabla f(x^*) = \mathbf{0}$  se mantiene, entonces  $x^*$  es llamado un **punto estacionario** de  $f$ .
- Esta condición es **necesaria**, pero no una condición **suficiente** para un mínimo local.
- Ejemplo:  $f(x) = -x^2$  posee un solo punto estacionario en  $x^* = 0$ , debido que  $\nabla f(x^*) = -2x^* = 0$ . Este punto no es un mínimo sino un máximo global.

# Punto de ensilladura

- Un punto estacionario no tiene que ser un mínimo o un máximo. Este punto estacionario es llamado **un punto de ensilladura** (saddle point).
- Ejemplo: el gradiente de  $f(x) = x_1^2 - x_2^2$  es  $\nabla f(x) = [2x_1, -2x_2]^T$ . Entonces,  $x^* = 0$  es el único punto estacionario. Como  $f$  es curvada positiva en la dirección  $x_1$  y negativamente curvada en la dirección  $x_2$ ,  $x^*$  es un punto de ensilladura.



# Condiciones necesarias de segundo orden

Teorema (condiciones necesarias de segundo orden):

Sea  $f$  dos veces diferenciable y sea  $\mathbf{x}^* \in \mathbb{R}^n$  un minimizador local de  $f$ , entonces

1.  $\nabla f(\mathbf{x}^*) = \mathbf{0}$ ,
2.  $\nabla^2 f(\mathbf{x}^*)$  es semidefinida positiva.

Estas condiciones son **solamente necesarias pero no suficientes**

- El único punto estacionario de  $f(x) = x^3$  es  $x^* = 0$ , con  $\nabla f(0) = 0$ ,  $\nabla^2 f(0) = 0$ . Se cumplen ambas condiciones.  $x^* = 0$  no es un mínimo local pero si es un **punto de ensilladura**.
- El único punto estacionario de  $f(x) = -x^4$  es  $x^* = 0$ , con  $\nabla f(0) = 0$ ,  $\nabla^2 f(0) = 0$ . Se satisfacen las dos condiciones. Pero  $x^* = 0$  no es un mínimo local sino un **máximo local**.

# Condiciones necesarias de segundo orden: demostración

Sea  $f$  dos veces diferenciable y sea  $\mathbf{x}^* \in R^n$  un minimizador local de  $f$ .

Asumir  $\nabla^2 f(\mathbf{x}^*)$  no es semidefinida positiva.

Entonces,  $\exists \mathbf{p} \in R^n: \mathbf{p}^T \nabla^2 f(\mathbf{x}^*) \mathbf{p} < 0$

La expansión de Taylor en  $\mathbf{x}^*$  nos da

$$f(\mathbf{x}^* + \epsilon \mathbf{p}) = f(\mathbf{x}^*) + \epsilon \nabla f(\mathbf{x}^*)^T \mathbf{p} + \frac{1}{2} \epsilon^2 \mathbf{p}^T \nabla^2 f(\mathbf{x}^*) \mathbf{p} + O(\epsilon^3).$$

$\mathbf{x}^*$  es un mínimo local y entonces la condición necesaria de primer orden

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

Para  $\epsilon$  lo suficientemente pequeño,  $O(\epsilon^2)$  domina sobre  $O(\epsilon^3)$ . Debido que

$$\mathbf{p}^T \nabla^2 f(\mathbf{x}^*) \mathbf{p} < 0 \rightarrow f(\mathbf{x}^* + \epsilon \mathbf{p}) < f(\mathbf{x}^*)$$

$\mathbf{x}^*$  no es un mínimo local.

# Condiciones suficiente para optimalidad

Teorema (condiciones suficientes para optimalidad):

Sea  $f$  dos veces diferenciable y sea  $\mathbf{x}^* \in R^n$ , sí

1.  $\nabla f(\mathbf{x}^*) = \mathbf{0}$ ,
2.  $\nabla^2 f(\mathbf{x}^*)$  es definida positiva.

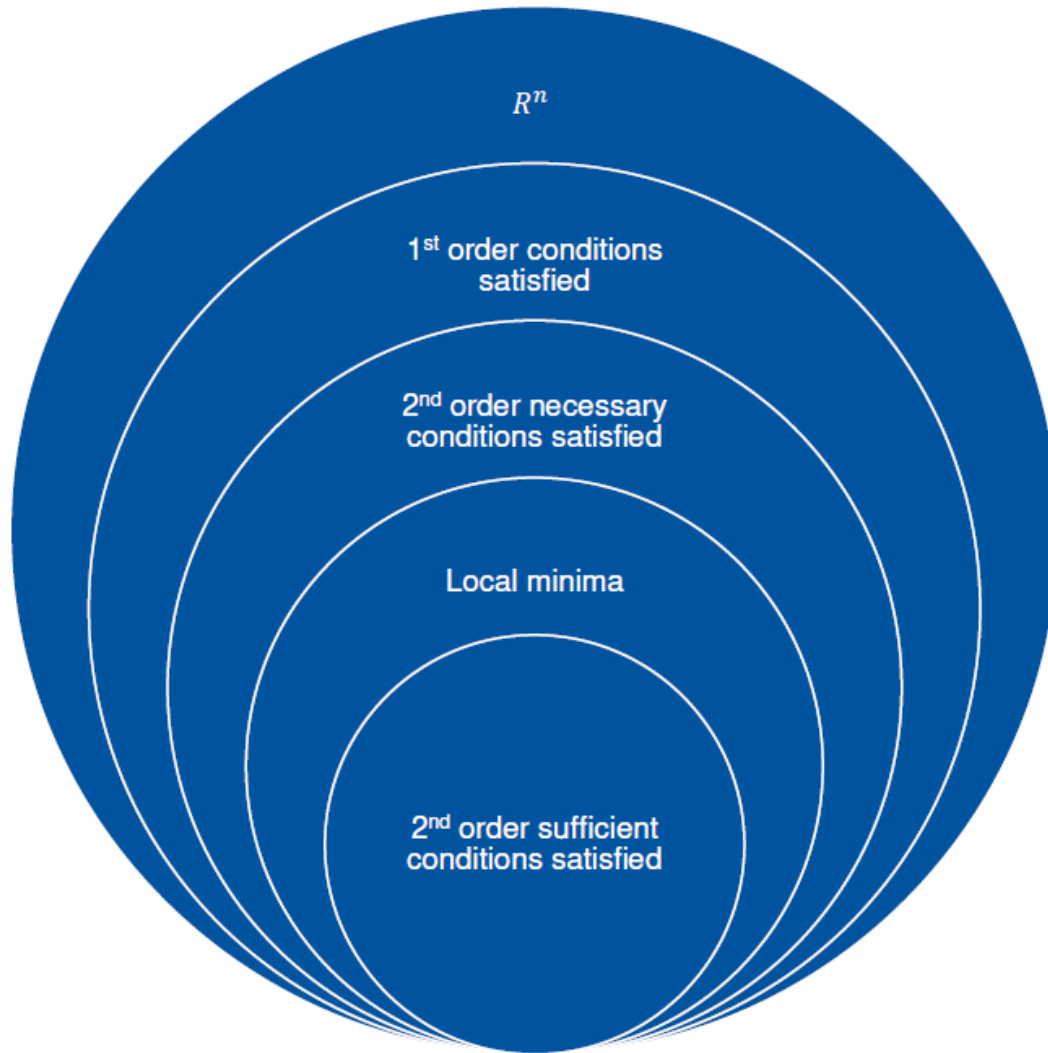
Entonces  $\mathbf{x}^*$  es un **minimizador local estricto** de  $f$ .

**Observación:**

- $f(x) = x^4$  tiene en  $x = 0$  su (único) mínimo global estricto. Además,  $\nabla f(0) = 0$  y  $\nabla^2 f(0) = 0$  se mantienen, entonces la segunda condición del teorema se viola.
- Por eso, las condiciones mencionadas en el teorema son **suficientes pero no necesarias**.



# Condiciones de optimalidad para problemas suaves



- Condiciones de optimalidad están en un punto, no para todo  $R^n$ .
- Todos los conjuntos mostrados son subconjuntos.
- Las condiciones necesarias de primer orden excluyen los puntos no estacionarios.
- Las condiciones necesarias de segundo orden excluyen algunos puntos y algunos máximos locales, pero no todos.

# Chequeo

- Que es un punto estacionario? Existen diferentes tipos de puntos estacionarios?
- Cuales son las condiciones necesarias de primer orden para problemas suaves sin restricciones?
- Cuales son las condiciones necesarias de segundo orden para problemas suaves sin restricciones?
- Cuales son las condiciones suficientes de segundo orden para problemas suaves sin restricciones?



Institución  
**Universitaria**  
Reacreditada en Alta Calidad

# *¡Gracias!*

Somos Innovación Tecnológica con *Sentido Humano*



Alcaldía de Medellín