



UNIVERSIDAD  
PANAMERICANA

# Gradiente Descendente

# Gradiente descendente

Todo método de aprendizaje supervisado consta de:

Datos

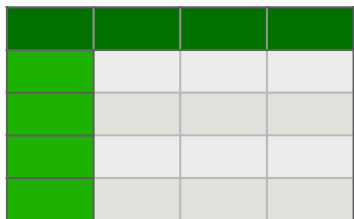
Modelo

Función de  
pérdida

Algoritmo de  
optimización

(error)

(entrenamiento)




$\{x, y\}$

$$\hat{y} = f(x, \theta)$$

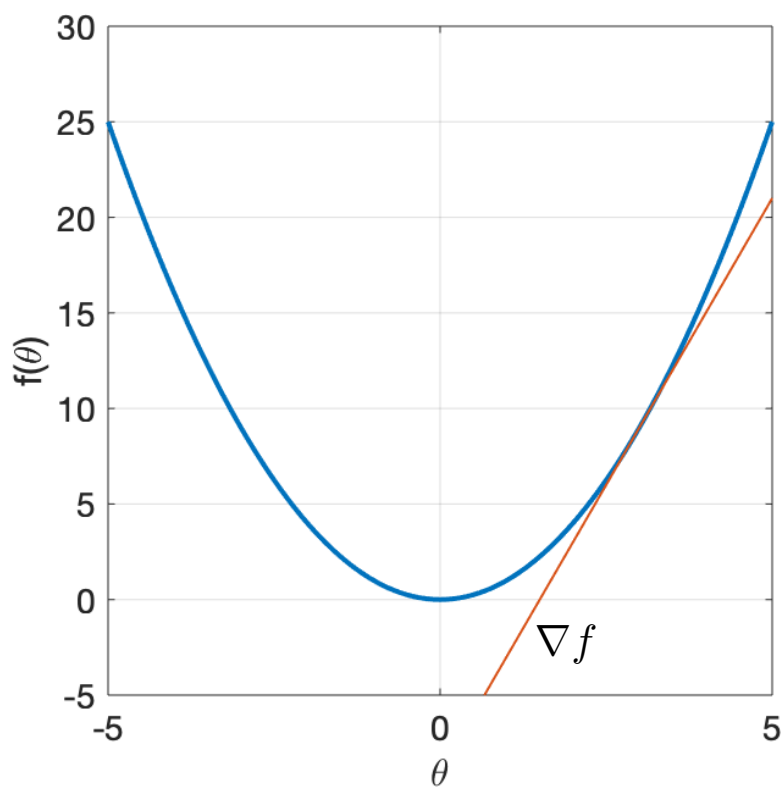
$$E = \sum_{i=1}^n (y_i - f(x_i))^2$$

$$\theta = ?$$

# Gradiente descendente

El **gradiente descendente** es un algoritmo iterativo de optimización que permite aproximar los valores de los parámetros que minimizan la función de error de un modelo dado un conjunto de datos.

$$\min_{\theta} f(\theta)$$



De manera iterativa, obtener el gradiente e ir en la dirección “contraria”.

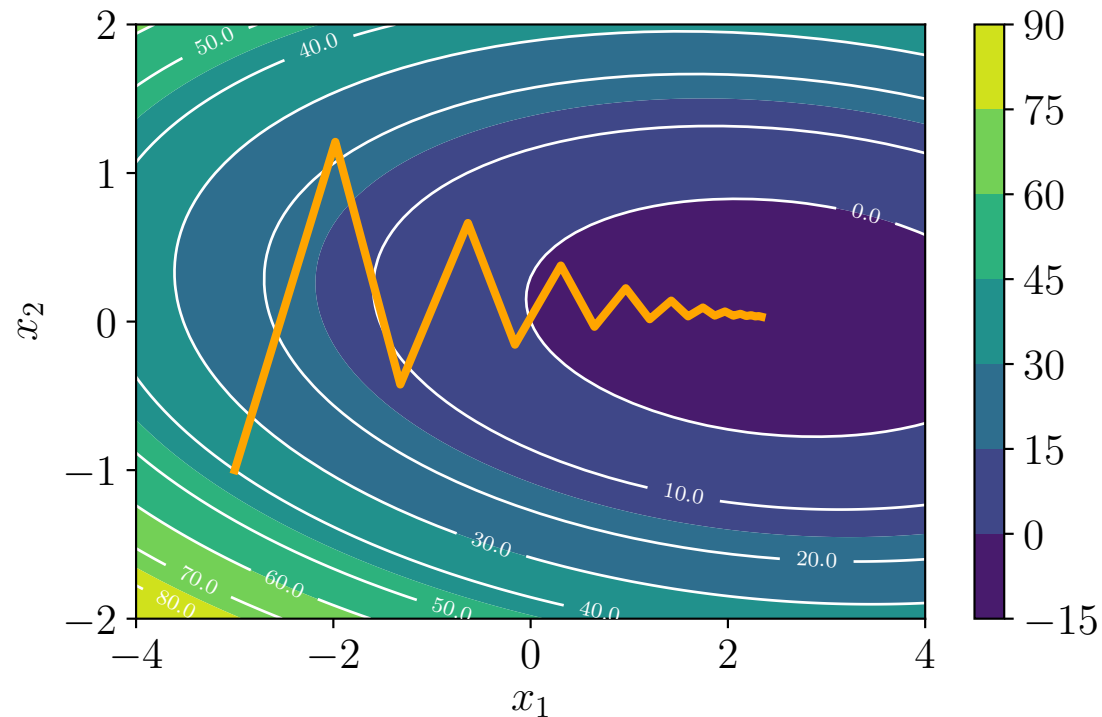
Regla de actualización de parámetros:

$$\theta_{i+1} = \theta_i - \eta \nabla f(\theta_i)$$

$$\eta > 0$$

# Gradiente descendente

El **gradiente descendente** es un algoritmo de optimización que permite aproximar los valores de los parámetros que minimizan la función de error de un modelo dado un conjunto de datos.



Deisenroth, M., Faisal, A., Ong, C., Mathematics for Machine Learning, Cambridge University Press, 2020.

## **Gradiente descendente** – *algoritmo simple*

1. Inicializar los parámetros del modelo.
2. Evaluar el modelo en la función de error.
3. Mientras (función de error  $>$  tolerancia):
  - i. Calcular el gradiente.
  - ii. Actualizar los parámetros del modelo.
  - iii. Evaluar el modelo en la función de error.
4. Regresar los valores de los parámetros

El **gradiente descendente con momento** es un algoritmo iterativo de optimización que permite aproximar los valores de los parámetros que minimizan la función de error de un modelo dado un conjunto de datos, tomando en cuenta un elemento de memoria sobre lo ocurrido en la iteración anterior.

Regla de actualización de parámetros:

$$\theta_{i+1} = \theta_i - \eta \nabla f(\theta_i) + \alpha \Delta \theta_i$$

$$\Delta \theta_i = \theta_i - \theta_{i-1} = \alpha \Delta \theta_{i-1} - \eta \nabla f(\theta_{i-1})$$

$$\alpha \in [0, 1]$$

## Tarea:

- \* Investigar el método del **gradiente descendente estocástico**
- \* Programar el método del gradiente descendente estocástico como función