Inteligencia Artificial & Machine Learning

Applicaciones en movilidad

Dr. Iván S. Razo Zapata

Founded by the Royal Academy of Engineering
and Lloyd's Register Foundation

Aprendizaje Supervisado

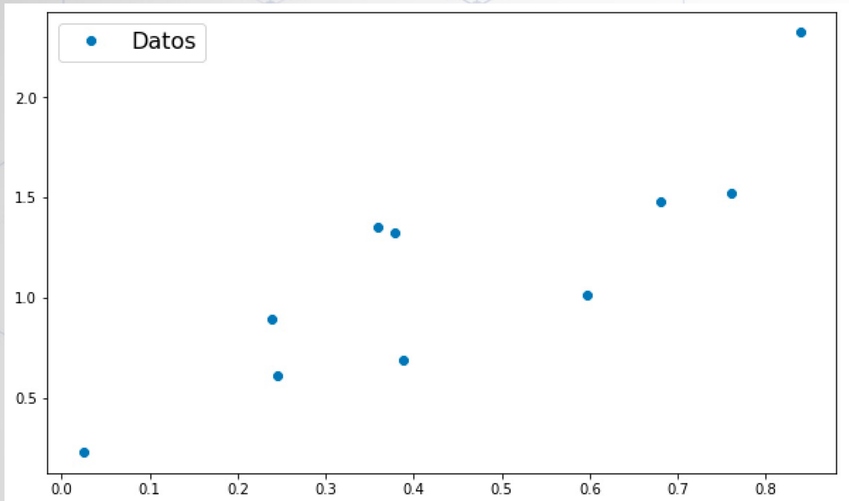
Tareas

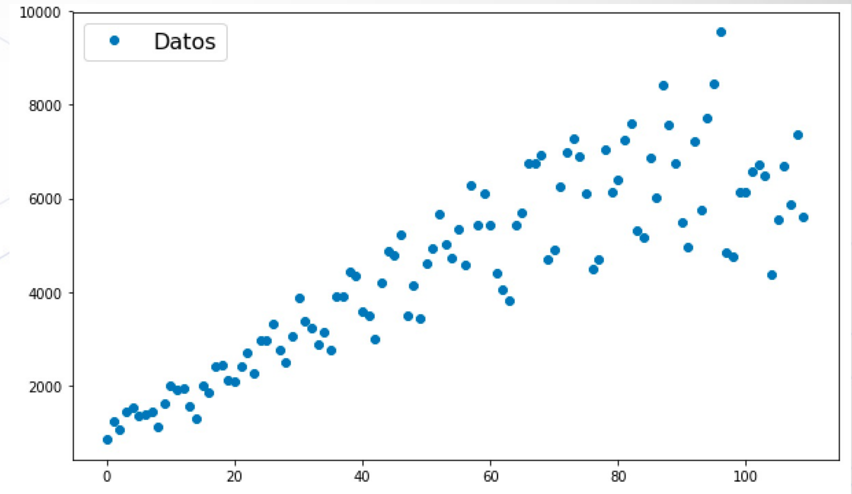
- Regresión

- Clasificación

Regresión Lineal

Idea general – Predecir nuevos valores en base a datos del pasado



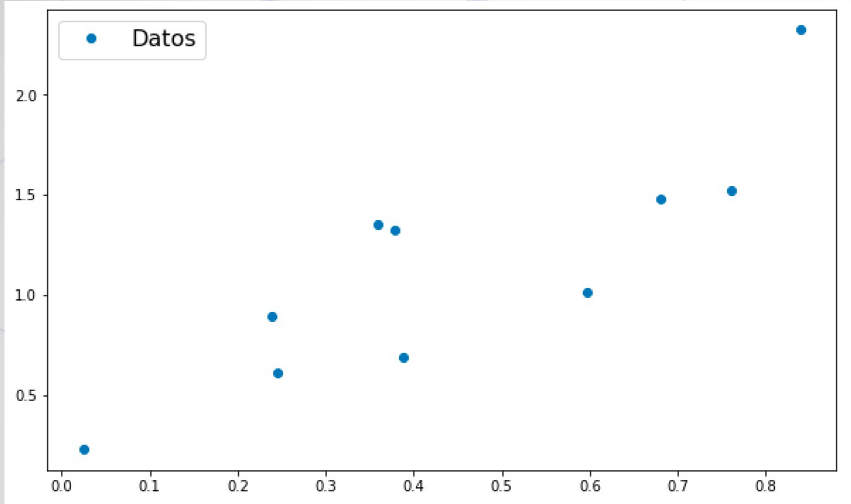


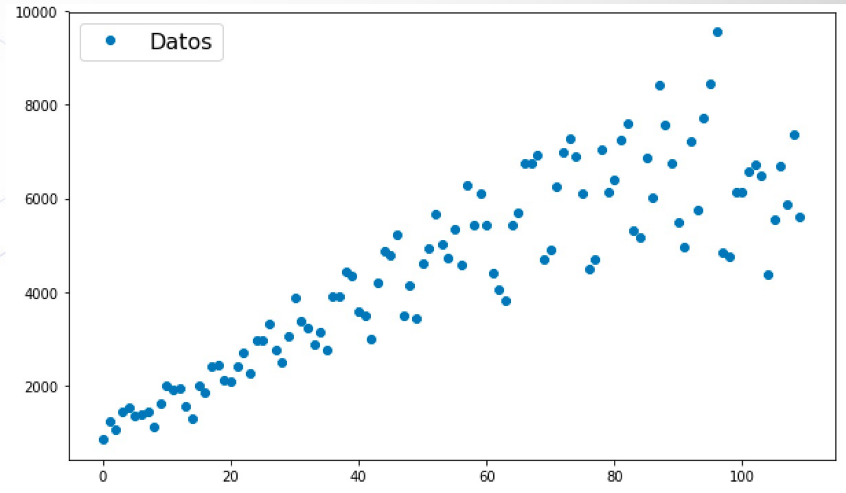
Peso del motor VS Consumo de combustible

m^2 VS Valor($)

¿Qué otros procesos/fenomenos pueden dar origen a estos datos?

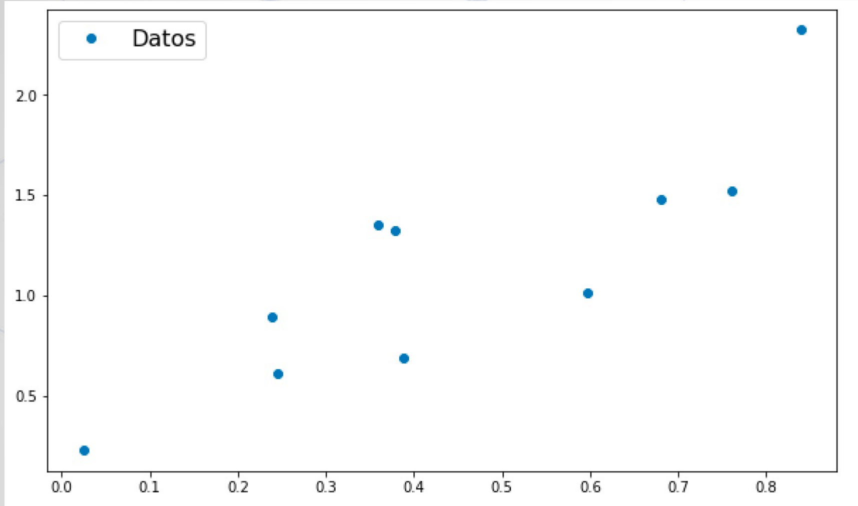
Idea general – Predecir nuevos valores en base a datos del pasado

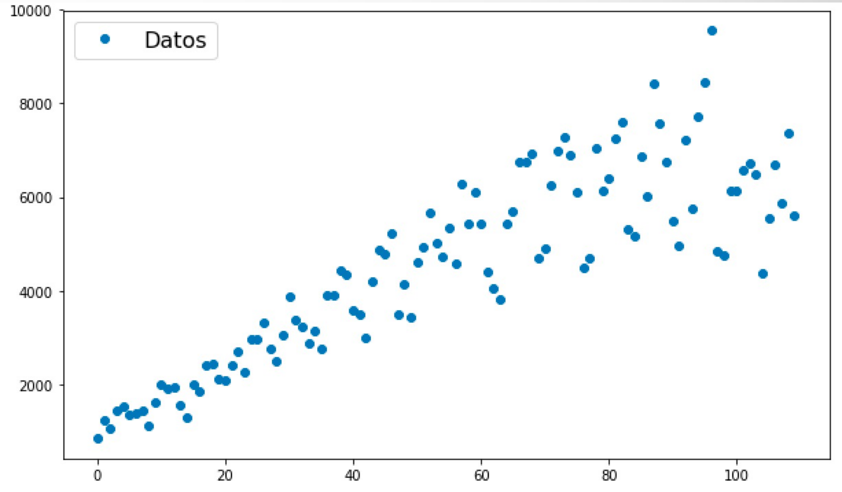




$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

Idea general – Predecir nuevos valores en base a datos del pasado

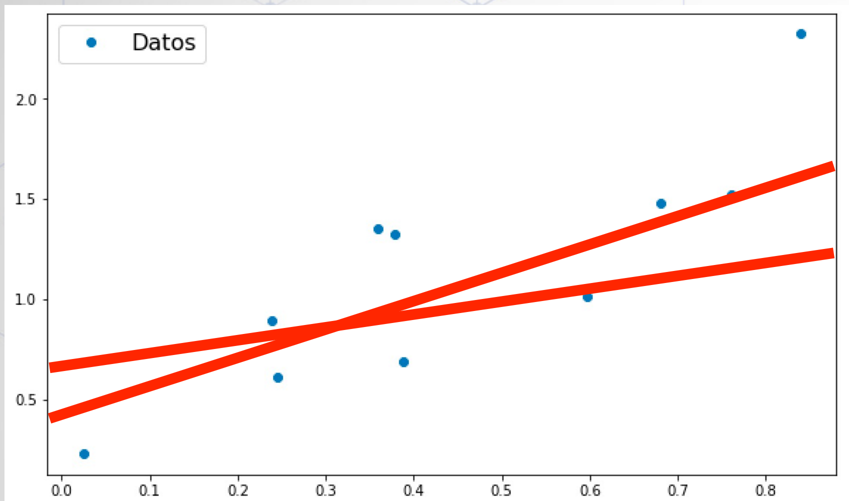


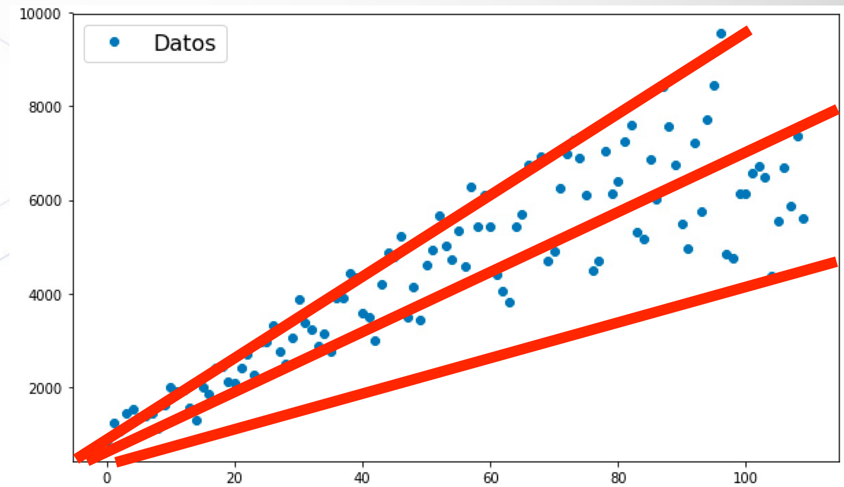


$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

$$y_{new} = f(x_{new})$$

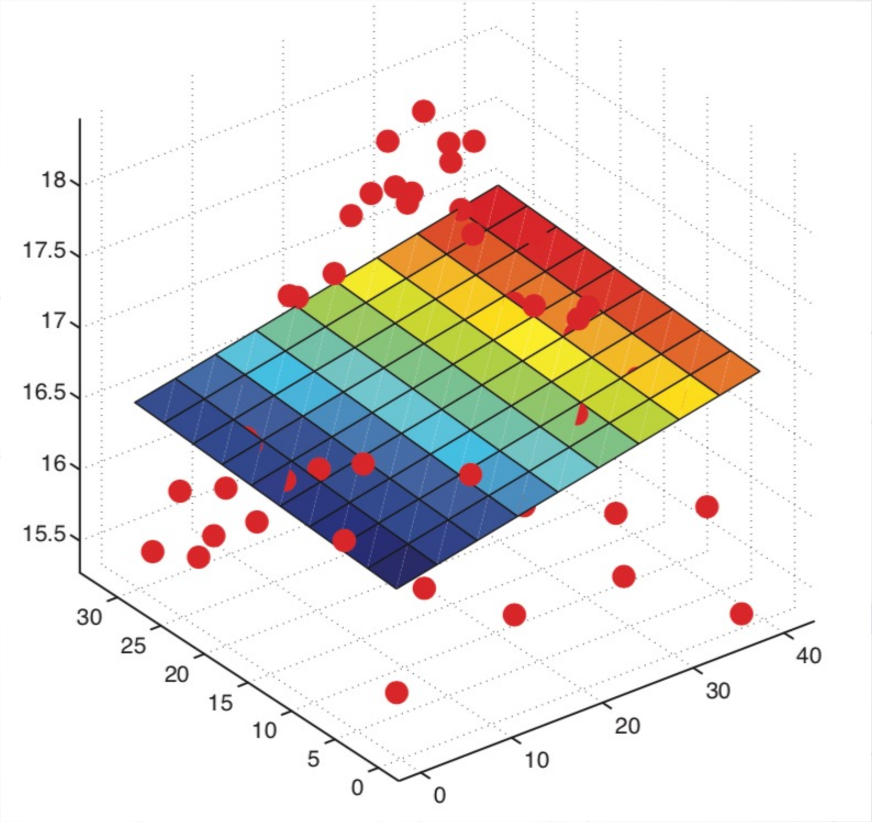
Idea general





$$y_{new} = f(x_{new})$$

Idea general



(Peso del motor, Km recorridos) VS Consumo de combustible

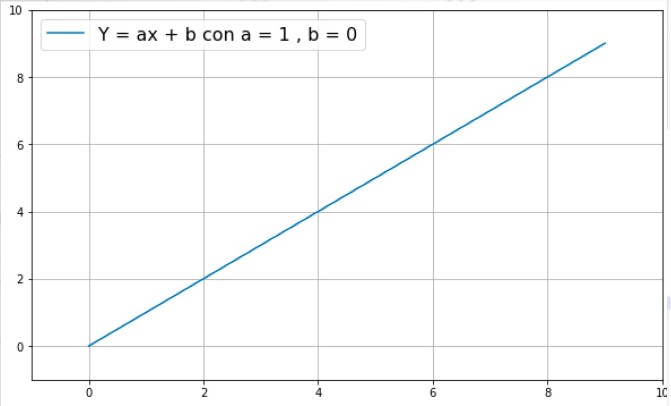
$$y = ax + b$$

Primer modelo

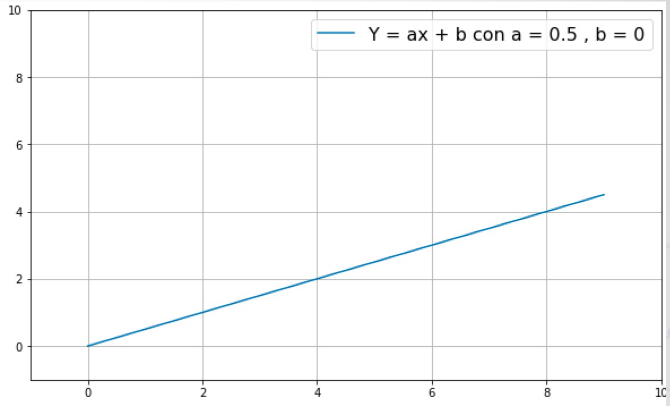
$$y = ax + b$$

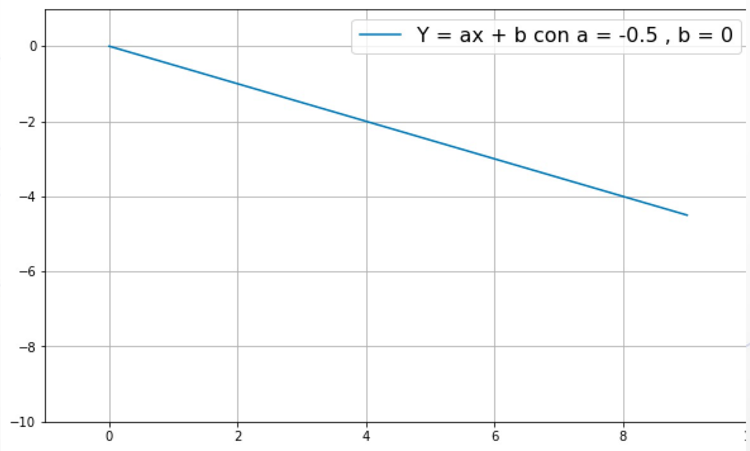
Parámetros

Ideas básicas - parámetro a (pendiente)

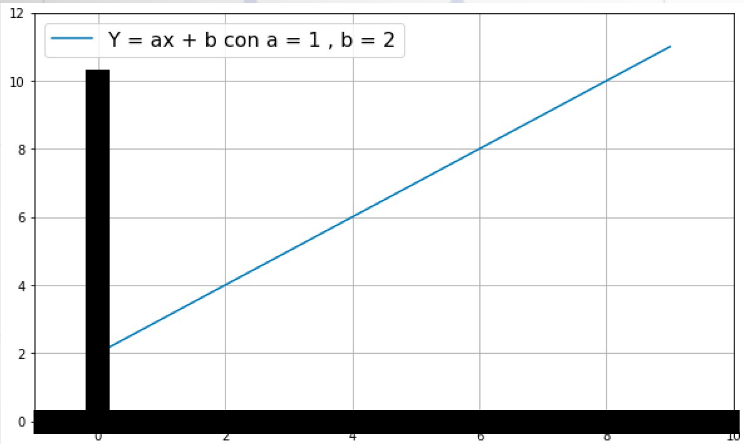


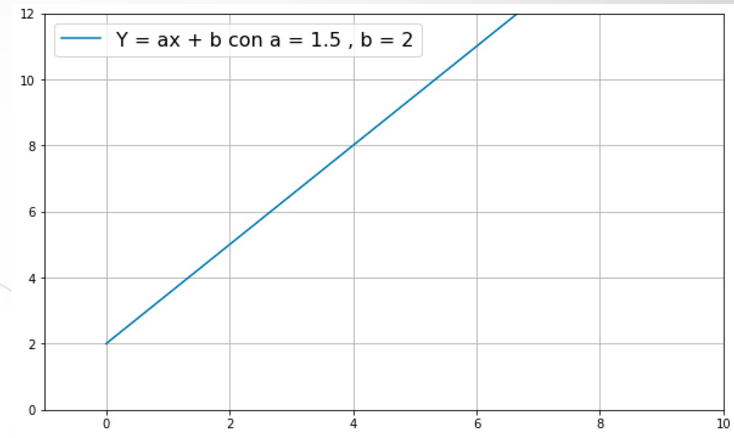
$$y = ax + b$$

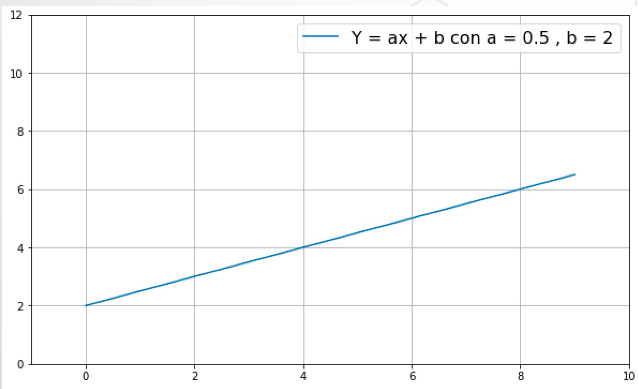




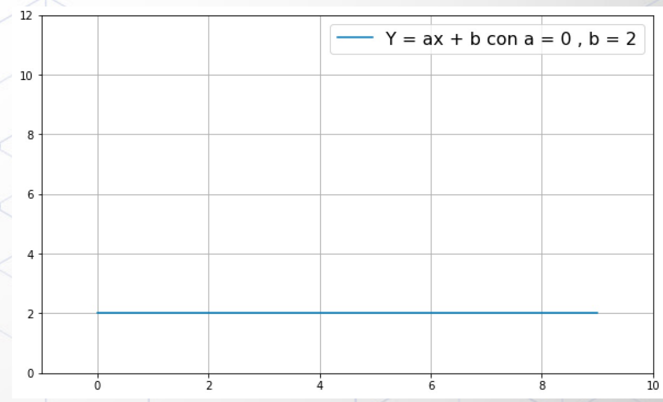
Ideas básicas - parámetro b (intercepto - cruce con el eje y)





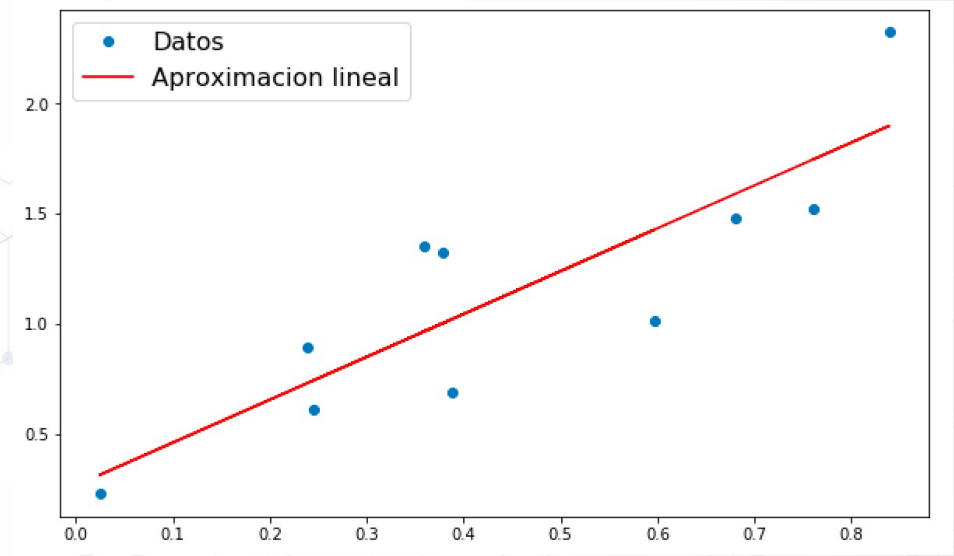


a = 0
b = 2



Ideas básicas

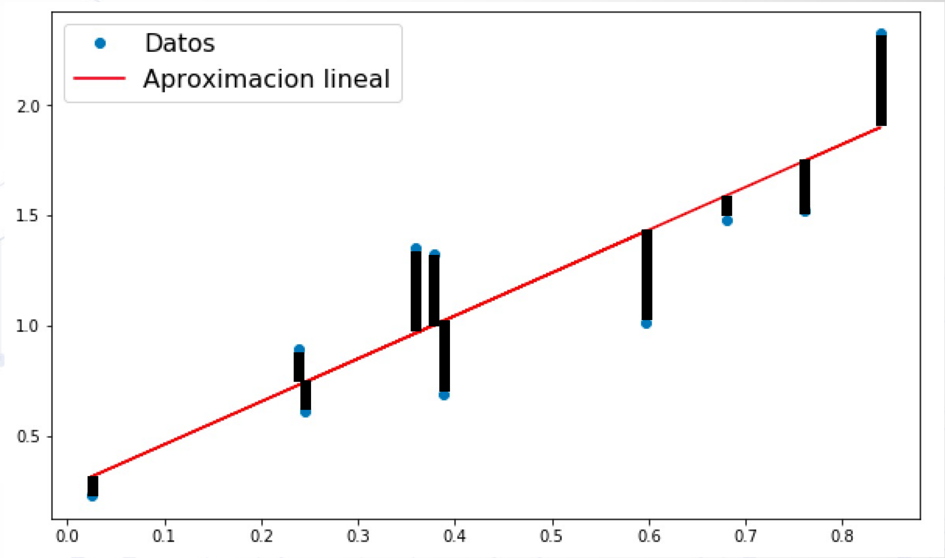
- **a**
  - Pendiente de la línea (slope)
- **b**
  - Intercepto
- ¿Cómo seleccionar los “mejores” valores de **a** y **b**?



$$y = ax + b$$

Ideas básicas

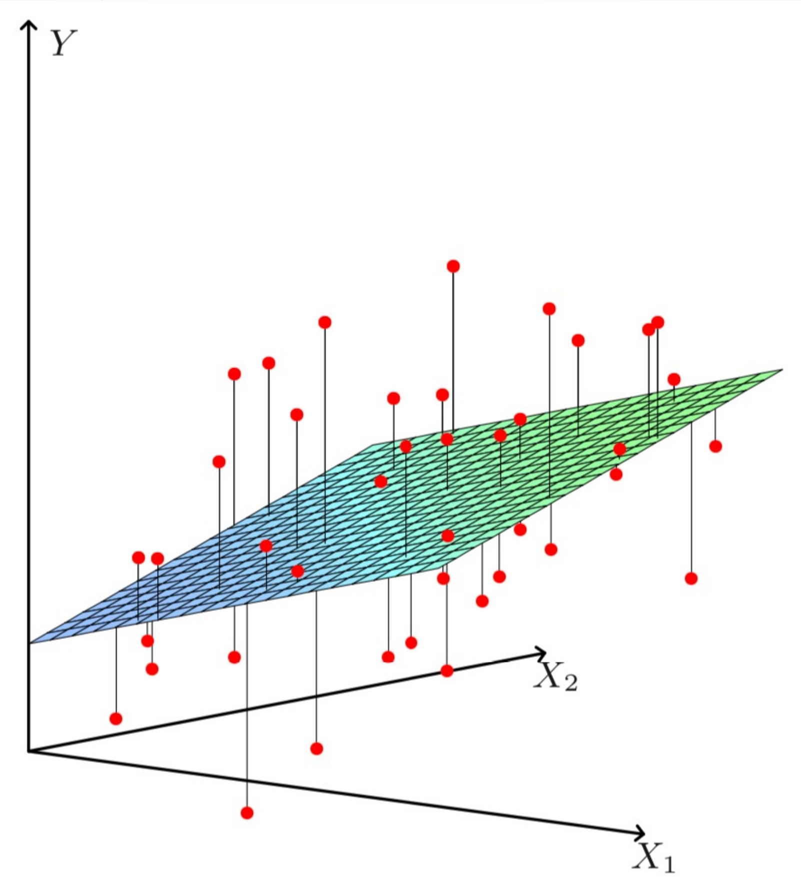
- **a**
  - Pendiente de la línea (slope)
- **b**
  - Intercepto
- ¿Cómo seleccionar los "mejores" valores de **a** y **b**?



Aprender en base a los datos observados

De los errores que se cometen durante la predicción

Ideas básicas



HUB IQ

EngineeringX
Founded by the Royal Academy of Engineering and Lloyd's Register Foundation

Ideas básicas

$$y = ax + b$$

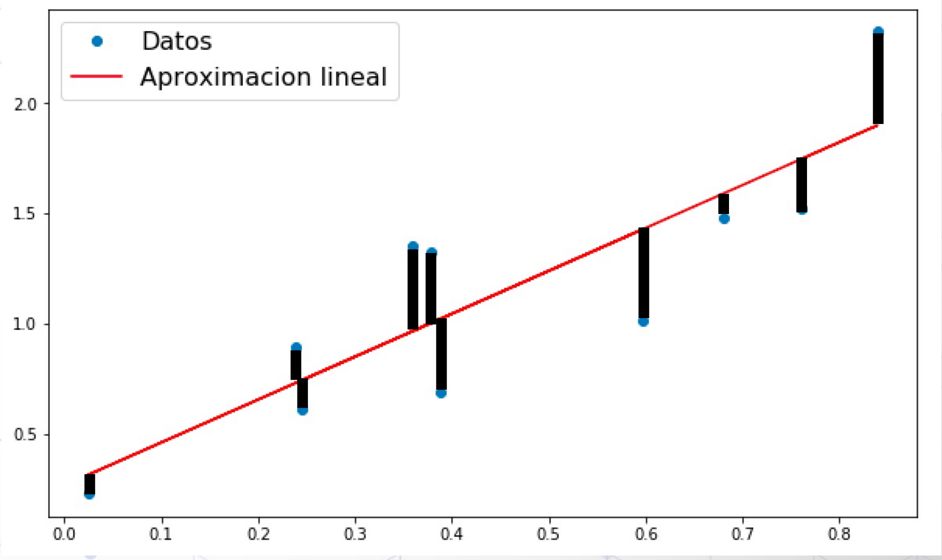
- Encontrar los valores para **a** y **b**
- Dos formas:
    - Solución iterativa con descenso de gradiente
        - El proceso comienza con valores aleatorios para a y b
    - Solución analítica

Error

- Minimizar el **error cuadrático medio** entre nuestra estimación y los datos

$$\hat{y}_{a,b}(x^{(i)}) = ax^{(i)} + b$$

$$E(a,b) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$



- Minimizar E(a,b) para **a** y **b**

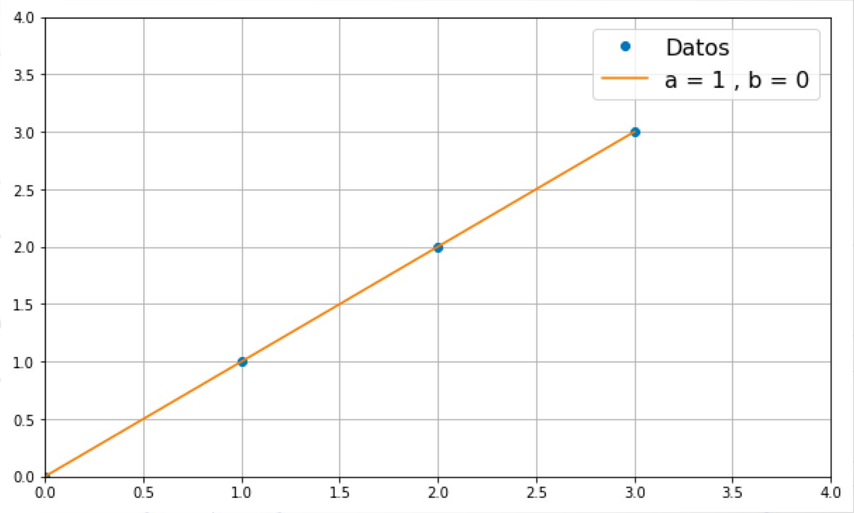
HUB IQ

EngineeringX
Founded by the Royal Academy of Engineering and Lloyd's Register Foundation

Ejemplo 1

$$y = ax + b$$

- Caso "trivial"
- Datos = (1,1), (2,2), (3,3)
- **a** = 1, **b** = 0



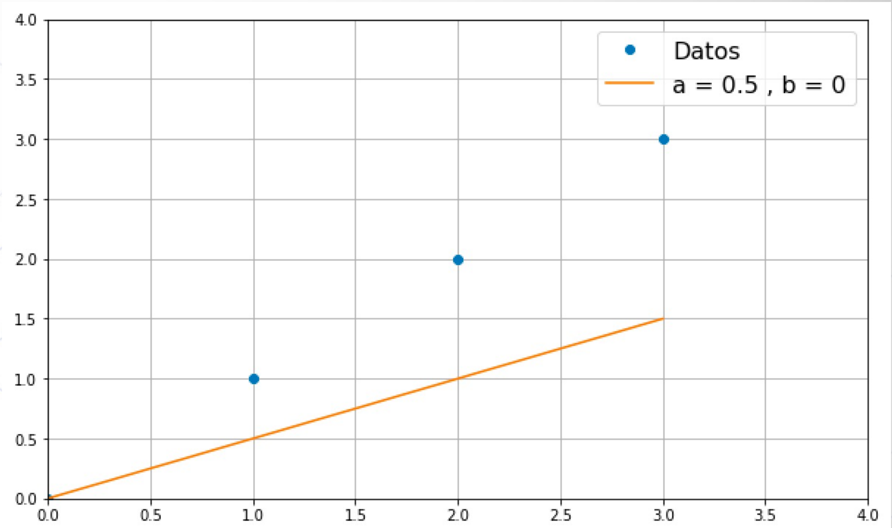
$$E(a,b) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$E(a,b) = \frac{1}{2\times 3}[((1-1)^2 + (2-2)^2 + (3-3)^2] = \frac{1}{6}[0^2 + 0^2 + 0^2] = 0$$

Ejemplo 2

$$y = ax + b$$

- Datos = (1,1), (2,2), (3,3)
- **a** = 0.5, **b** = 0



$$E(a,b) = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$E(a,b) = \frac{1}{2\times3} [((0.5-1)^2 + (1-2)^2 + (1.5-3)^2] = \frac{1}{6} [-0.5^2 + -1^2 + -1.5^2]$$

$$E(a,b) = \frac{1}{6} [0.25 + 1 + 2.25] = \frac{1}{6} [3.5] = \frac{3.5}{6} = 0.58$$

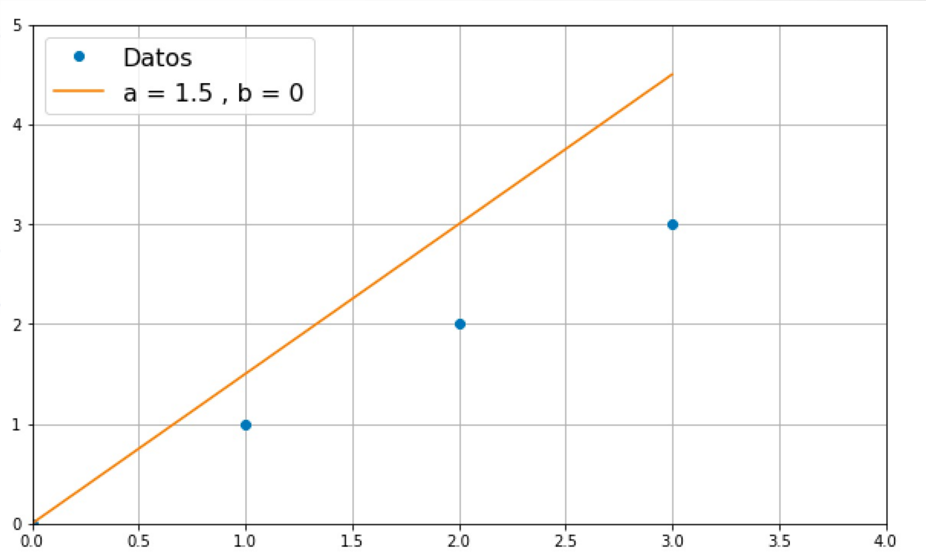
HUB IQ

EngineeringX
Founded by the Royal Academy of Engineering and Lloyd's Register Foundation

Ejemplo 3

$$y = ax + b$$

- Datos = (1,1), (2,2), (3,3)
- **a** = 1.5, **b** = 0



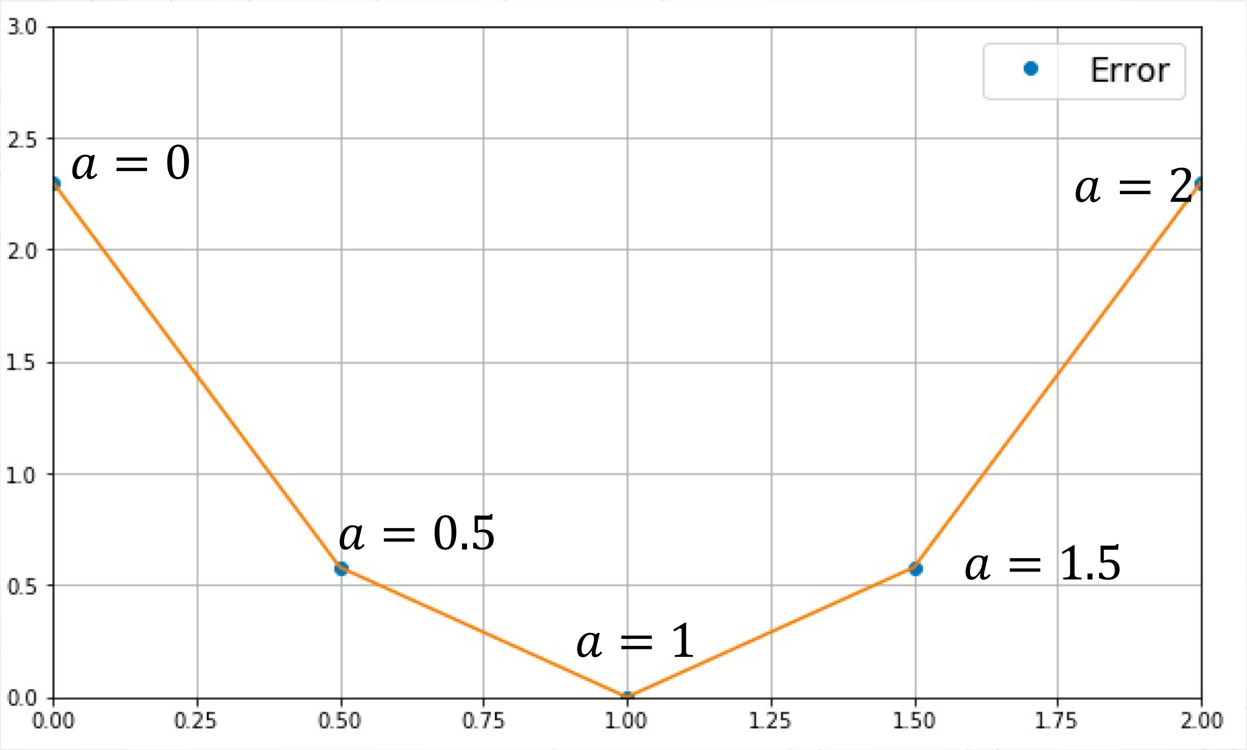
$$E(a,b) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$E(a,b) = \frac{1}{2\times3}[((1.5-1)^2 + (3-2)^2 + (4.5-3)^2] = \frac{1}{6}[0.5^2 + 1^2 + 1.5^2]$$

$$E(a,b) = \frac{1}{6}[0.25 + 1 + 2.25] = \frac{1}{6}[3.5] = \frac{3.5}{6} = 0.58$$

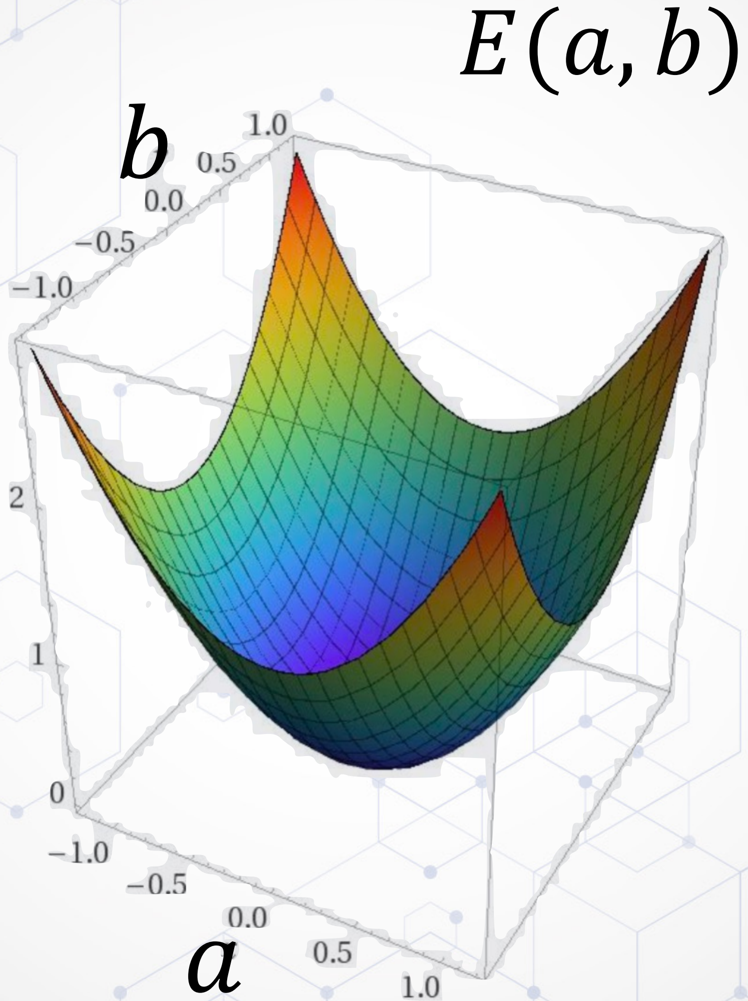
Curva de error – b fija

$$E(a, b = 0)$$



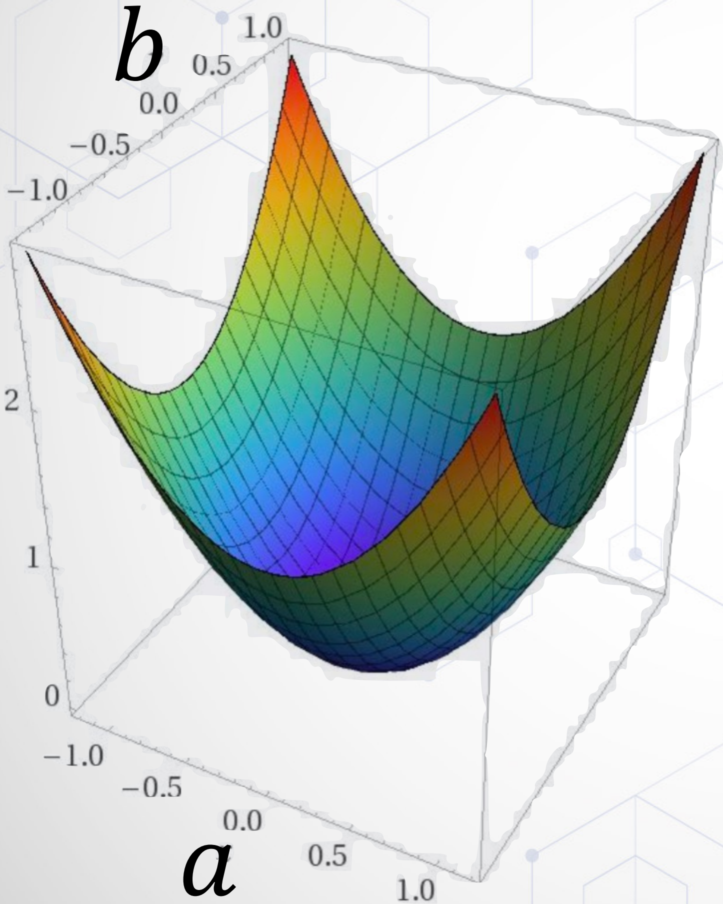
Al alejarnos del valor óptimo el error se incrementa

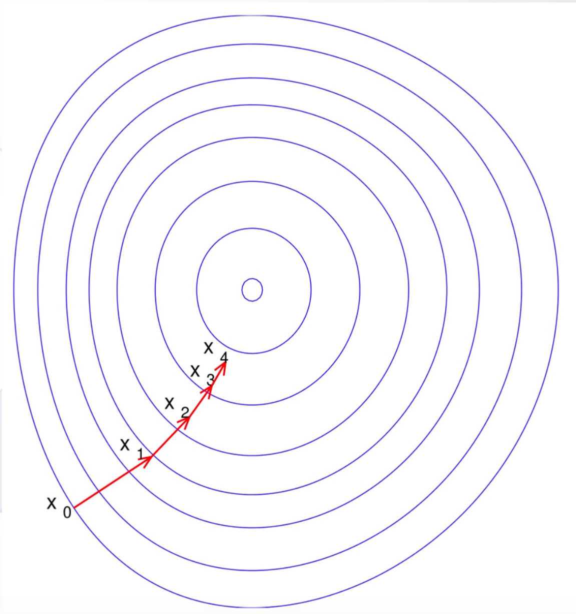
Superficie de error



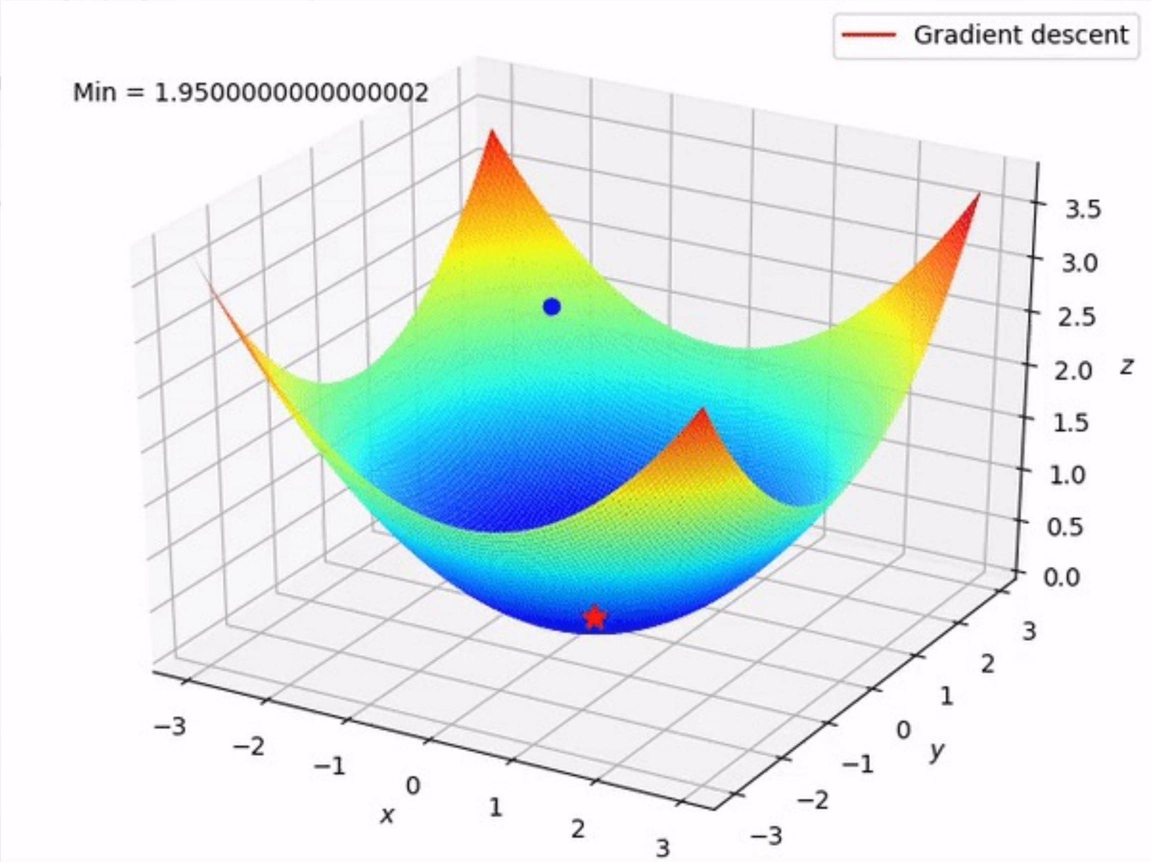
Solución iterativa con descenso de gradiente

$$E(a,b)$$





Solución iterativa con descenso de gradiente



¿Cómo?

Solución iterativa con descenso de gradiente

- Iniciando con valores aleatorios **a** y **b**
- Actualizar el valor de **a** y **b** de manera iterativa hasta llegar al punto mínimo

$$E(a,b) = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$\nabla E(a,b)$$

$$a \leftarrow a - \alpha \frac{\partial E(a,b)}{\partial a}$$

$$b \leftarrow b - \alpha \frac{\partial E(a,b)}{\partial b}$$

Solución iterativa con descenso de gradiente

- Iniciando con valores aleatorios **a** y **b**
- Actualizar el valor de **a** y **b** de manera iterativa hasta llegar al punto mínimo

$$E(a,b) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$\nabla E(a,b)$$

$$a \leftarrow a - \alpha\frac{\partial E(a,b)}{\partial a} \qquad b \leftarrow b - \alpha\frac{\partial E(a,b)}{\partial b}$$

$$\alpha$$

Tasa de aprendizaje (learning rate): Hyperparámetro

Solución iterativa con descenso de gradiente

$$\hat{y}_{a,b}(x^{(i)}) = ax^{(i)} + b$$

$$E(a,b) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$a \leftarrow a - \alpha \frac{\partial}{\partial a} E(a,b)$$

$$a \leftarrow a - \alpha \frac{\partial}{\partial a}\frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$a \leftarrow a - \alpha \frac{2}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})\frac{\partial}{\partial a}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})$$

$$a \leftarrow a - \alpha \frac{2}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})\frac{\partial}{\partial a}(ax^{(i)} + b - y^{(i)})$$

$$a \leftarrow a - \alpha \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})\; x^{(i)}$$

Solución iterativa con descenso de gradiente

$$\hat{y}_{a,b}(x^{(i)}) = ax^{(i)} + b$$

$$E(a,b) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$b \leftarrow b - \alpha\frac{\partial}{\partial b}E(a,b)$$

$$b \leftarrow b - \alpha\frac{\partial}{\partial b}\frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})^2$$

$$b \leftarrow b - \alpha\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})\frac{\partial}{\partial b}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})$$

$$b \leftarrow b - \alpha\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})\frac{\partial}{\partial b}(ax^{(i)} + b - y^{(i)})$$

$$b \leftarrow b - \alpha\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_{a,b}(x^{(i)}) - y^{(i)})$$

Solución iterativa con descenso de gradiente

$$\hat{y}_{a,b}(x^{(i)}) = ax^{(i)} + b$$

$a \leftarrow$ *valor aleatorio, p.ej.* $\in [0,1]$

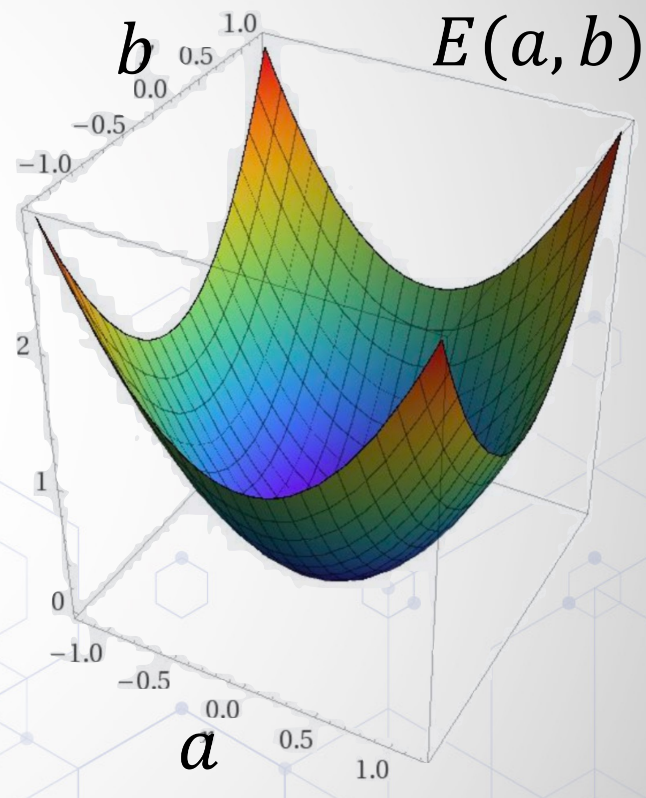
$b \leftarrow$ *valor aleatorio, p.ej.* $\in [0,1]$

Repetir {

$$a \leftarrow a - \alpha \frac{1}{m} \sum_{i=1}^{m} (ax^{(i)} + b - y^{(i)}) x^{(i)}$$

$$b \leftarrow b - \alpha \frac{1}{m} \sum_{i=1}^{m} (ax^{(i)} + b - y^{(i)})$$

}



Solución iterativa con descenso de gradiente

$$\hat{y}_{a,b}(x^{(i)}) = ax^{(i)} + b$$

$a \leftarrow$ *valor aleatorio, p.ej.* $\in [0,1]$

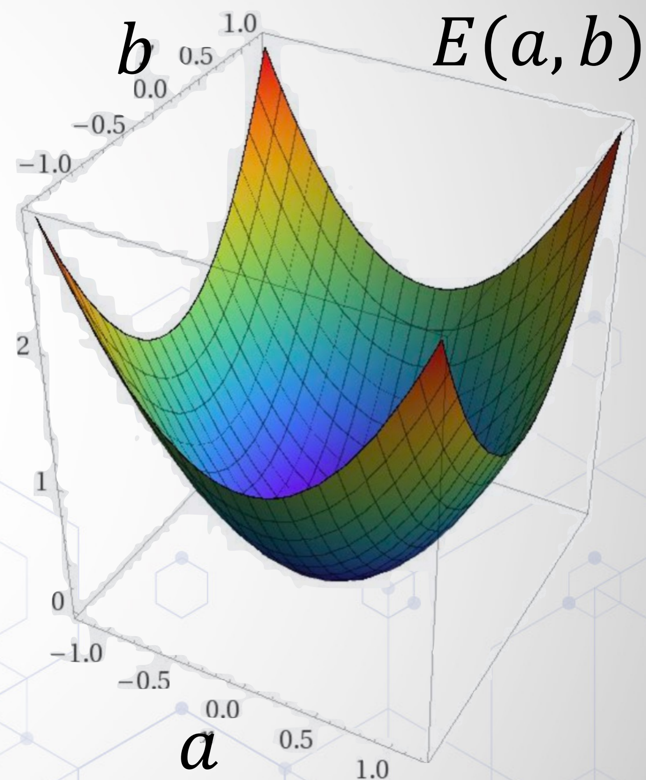
$b \leftarrow$ *valor aleatorio, p.ej.* $\in [0,1]$

Repetir {

$$a \leftarrow a - \alpha \frac{1}{m} \sum_{i=1}^{m} (\hat{y} - y^{(i)}) x^{(i)}$$

$$b \leftarrow b - \alpha \frac{1}{m} \sum_{i=1}^{m} (\hat{y} - y^{(i)})$$

}



HUB IQ

EngineeringX

Founded by the Royal Academy of Engineering and Lloyd's Register Foundation

Solución iterativa con descenso de gradiente

$$\hat{y}_{a,b}(x^{(i)}) = ax^{(i)} + b$$

$a \leftarrow$ *valor aleatorio, p.ej.* $\in [0,1]$

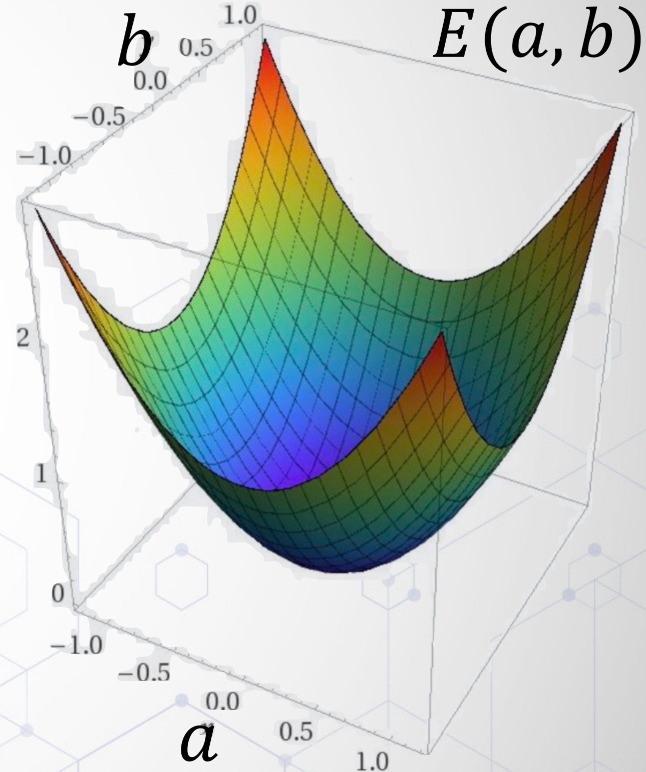
$b \leftarrow$ *valor aleatorio, p.ej.* $\in [0,1]$

Repetir {

$$a \leftarrow a - \alpha \frac{1}{m} \sum_{i=1}^{m} (Error) x^{(i)}$$

$$b \leftarrow b - \alpha \frac{1}{m} \sum_{i=1}^{m} (Error)$$

}



Regresión Lineal

Implementación en Python

Solución iterativa con descenso de gradiente

$$\hat{y}_{a,b}(x^{(i)}) = ax^{(i)} + b$$

$a \leftarrow$ *valor aleatorio, p.ej.* $\in [0,1]$

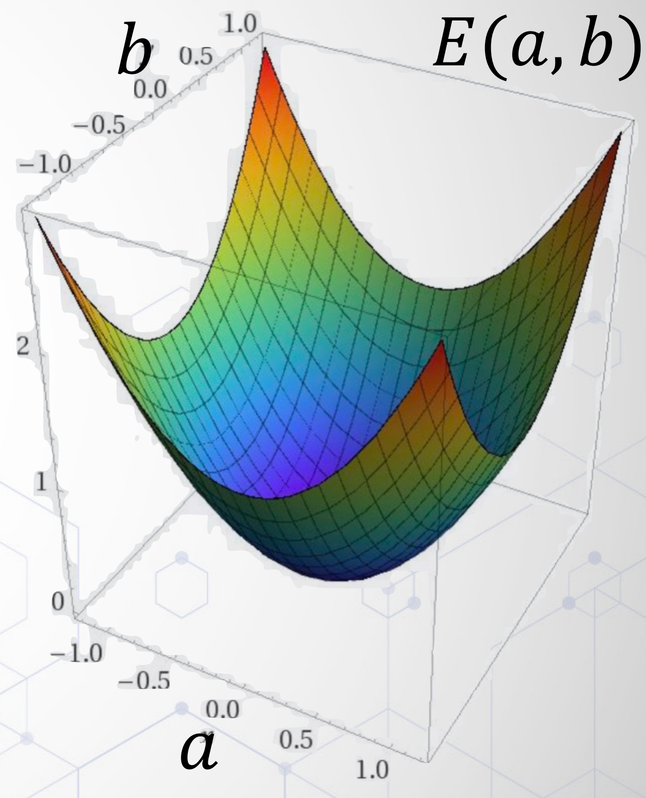
$b \leftarrow$ *valor aleatorio, p.ej.* $\in [0,1]$

Repetir {

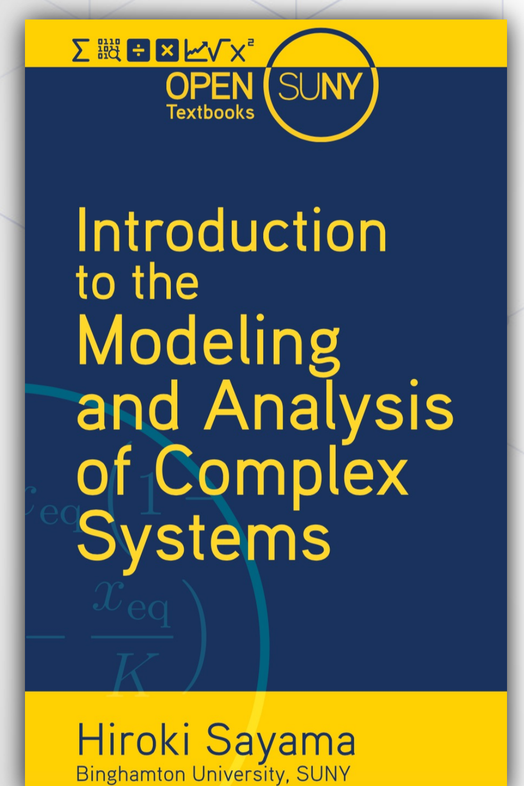
$$a \leftarrow a - \alpha \frac{1}{m} \sum_{i=1}^{m} (Error) x^{(i)}$$

$$b \leftarrow b - \alpha \frac{1}{m} \sum_{i=1}^{m} (Error)$$

}



Idea general



Three essential components of computer simulation

Initialize. You will need to set up the initial values for all the state variables of the system.

Observe. You will need to define how you are going to monitor the state of the system. This could be done by just printing out some variables, collecting measurements in a list structure, or visualizing the state of the system.

Update. You will need to define how you are going to update the values of those state variables in every time step. This part will be defined as a function, and it will be executed repeatedly.

Idea general

Regresión lineal (descenso por gradiente)

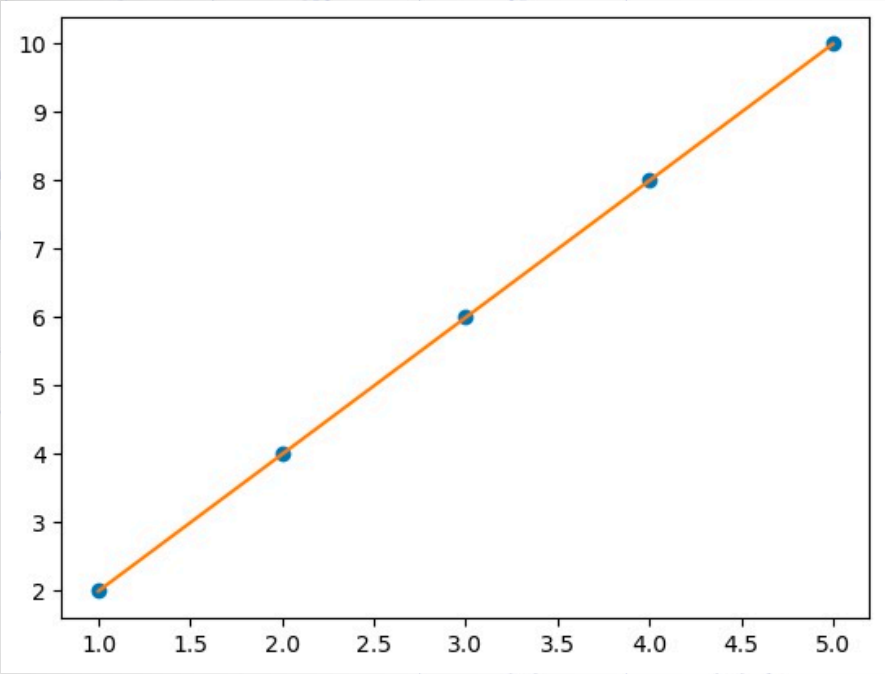
Three essential components of computer simulation

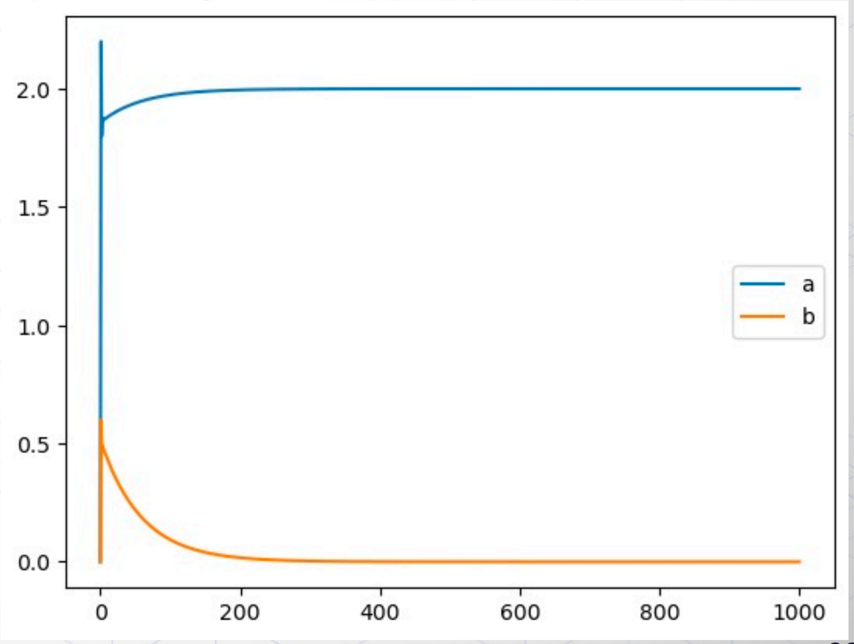
Initialize. You will need to set up the initial values for all the state variables of the system.

Observe. You will need to define how you are going to monitor the state of the system. This could be done by just printing out some variables, collecting measurements in a list structure, or visualizing the state of the system.

Update. You will need to define how you are going to update the values of those state variables in every time step. This part will be defined as a function, and it will be executed repeatedly.

Idea general





Polinomios

$$y = ax + b$$

$$y = \theta_0 + \theta_1 x$$

Polinomios

$$y = ax + b$$

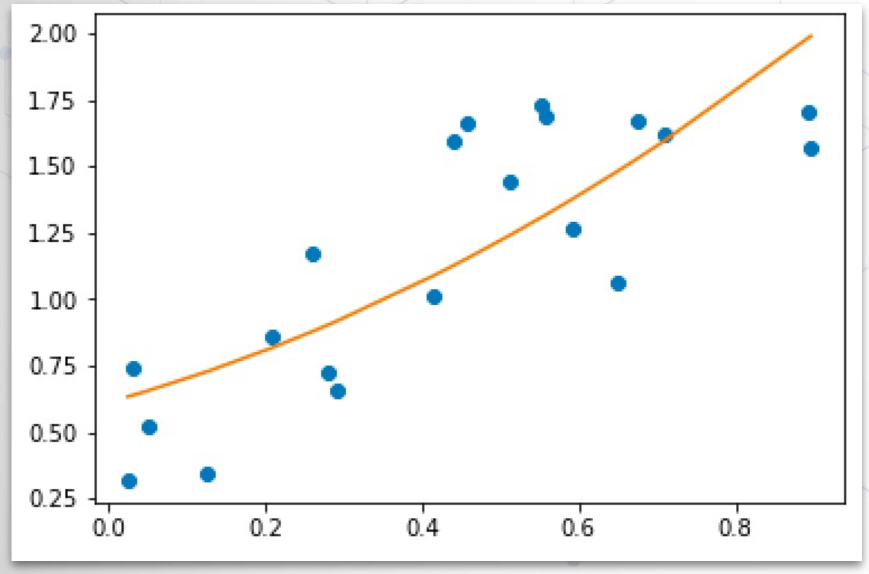
$$y = \theta_0 + \theta_1 x$$

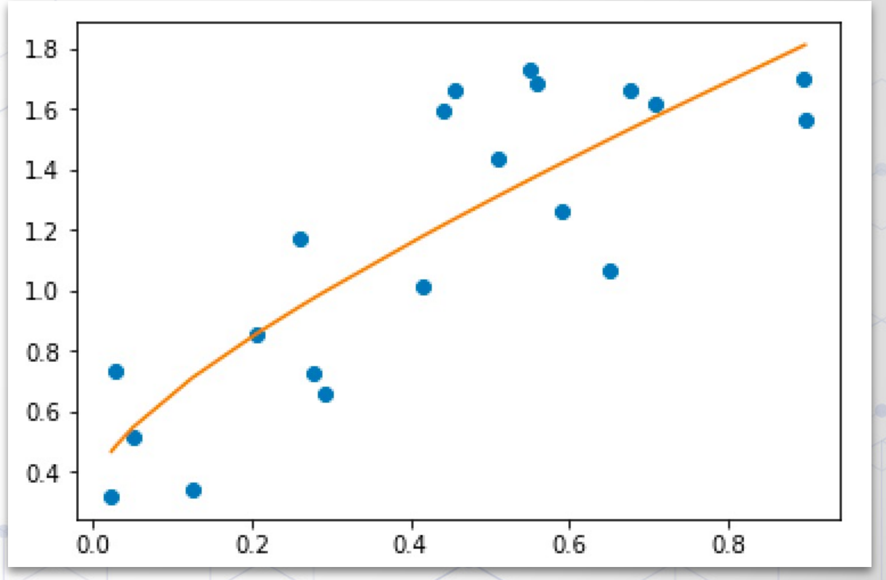
$$y = \theta_0 + \theta_1 x + \theta_2 x^2$$

Polinomios

$$y = \theta_0 + \theta_1 x + \theta_2 x^2$$

$$y = \theta_0 + \theta_1 x + \theta_2 \sqrt{x}$$





Polinomios

$$y = \theta_0 + \theta_1 x + \theta_2 x^2$$

$$E(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_\theta(x) - y)^2$$

$$\theta_0 = \theta_0 - \alpha\frac{\partial E(\theta)}{\partial \theta_0}$$

$$\theta_1 = \theta_1 - \alpha\frac{\partial E(\theta)}{\partial \theta_1}$$

$$\theta_2 = \theta_2 - \alpha\frac{\partial E(\theta)}{\partial \theta_2}$$

HUB IQ

EngineeringX
Founded by the Royal Academy of Engineering and Lloyd's Register Foundation

Polinomios

$$y = \theta_0 + \theta_1 x + \theta_2 x^2$$

$$E(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_\theta(x) - y)^2$$

$$\frac{\partial E(\theta)}{\partial \theta_0} = \frac{\partial}{\partial \theta_0}\frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_\theta(x) - y)^2$$

$$= \frac{\partial}{\partial \theta_0}\frac{1}{2m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)^2$$

$$= \frac{2}{2m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)\,\frac{\partial}{\partial \theta_0}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)$$

$$= \frac{1}{m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)\,1$$

$$\theta_0 = \theta_0 - \alpha\frac{1}{m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)$$

Polinomios

$$y = \theta_0 + \theta_1 x + \theta_2 x^2$$

$$E(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_\theta(x) - y)^2$$

$$\frac{\partial E(\theta)}{\partial \theta_1} = \frac{\partial}{\partial \theta_1}\frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_\theta(x) - y)^2$$

$$= \frac{\partial}{\partial \theta_1}\frac{1}{2m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)^2$$

$$= \frac{2}{2m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)\frac{\partial}{\partial \theta_1}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)$$

$$= \frac{1}{m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)\, x$$

$$\theta_1 = \theta_1 - \alpha\frac{1}{m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)x$$

Polinomios

$$y = \theta_0 + \theta_1 x + \theta_2 x^2$$

$$E(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_\theta(x) - y)^2$$

$$\frac{\partial E(\theta)}{\partial \theta_2} = \frac{\partial}{\partial \theta_2}\frac{1}{2m}\sum_{i=1}^{m}(\hat{y}_\theta(x) - y)^2$$

$$= \frac{\partial}{\partial \theta_2}\frac{1}{2m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)^2$$

$$= \frac{2}{2m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)\frac{\partial}{\partial \theta_2}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)$$

$$= \frac{1}{m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)x^2$$

$$\theta_2 = \theta_2 - \alpha\frac{1}{m}\sum_{i=1}^{m}(\theta_0 + \theta_1 x + \theta_2 x^2 - y)x^2$$

Polinomios

$$y = \theta_0 + \theta_1 x + \theta_2 x^2 + \cdots + \theta_n x^n$$

$\theta \leftarrow$ *valores aleatorios, p.ej.* $\in [0,1]$

Repetir {

$$\theta_j \leftarrow \theta_j - \alpha \times \sum_{i=1}^{m} (\theta x^i - y^i) x^j$$

}

Generalización a múltiples dimensiones

- Solución iterativa con descenso de gradiente

$$y = ax + b$$

$$y = bx_0 + ax_1 \qquad x_0 = 1$$

$$y = \theta_0 x_0 + \theta_1 x_1$$

$$y = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

Generalización a múltiples dimensiones

$\theta \leftarrow$ *valores aleatorios, p.ej.* $\in [0,1]$

$$\nabla E(\theta) = \left[\frac{\partial E}{\partial \theta_1}, \frac{\partial E}{\partial \theta_2}, \dots \frac{\partial E}{\partial \theta_n}\right]$$

Repetir {

$$\theta_j \leftarrow \theta_j - \alpha \times \sum_{i=1}^{m} \left(\theta x^i - y^i\right) x_j^i$$

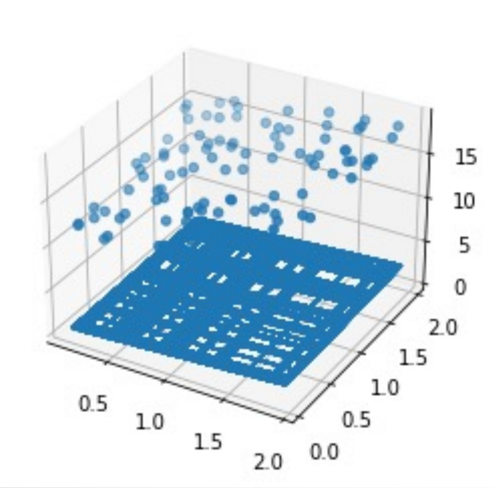
Para cada j (característica)

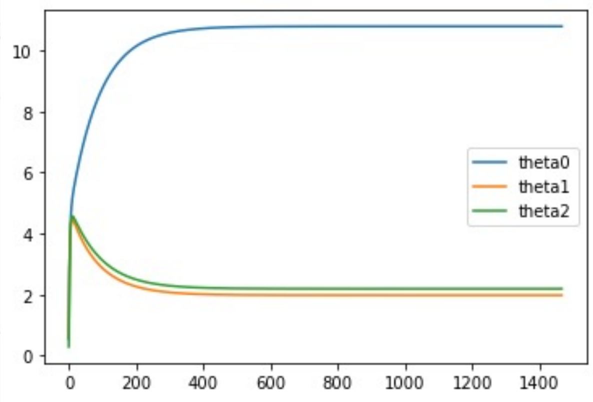
}

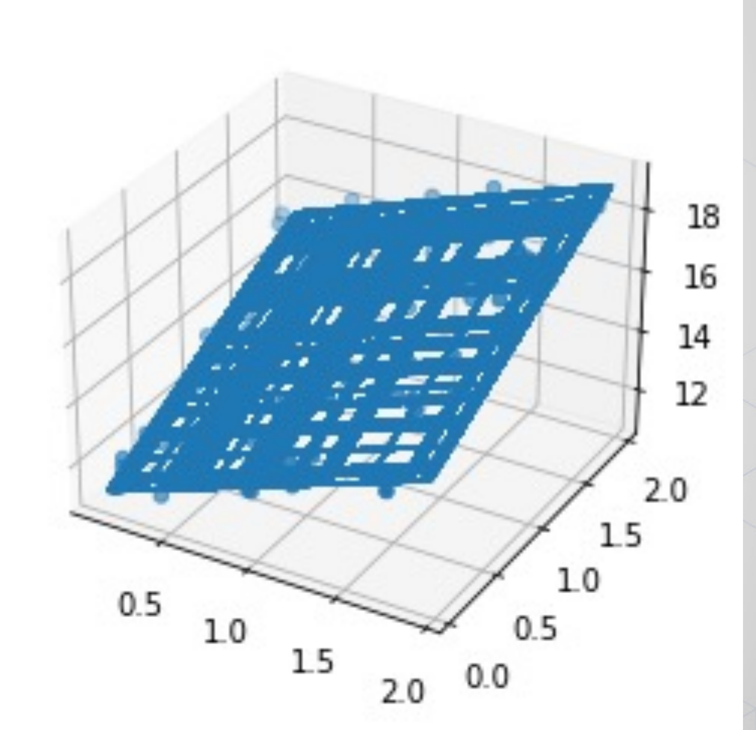
i dato/muestra

Múltiples dimensiones

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$







Datos reales

`sklearn.datasets`.load_diabetes

`sklearn.datasets.load_diabetes(*, return_X_y=False, as_frame=False)` [source]

Load and return the diabetes dataset (regression).

Samples total	442
Dimensionality	10
Features	real, -.2 < x < .2
Targets	integer 25 - 346

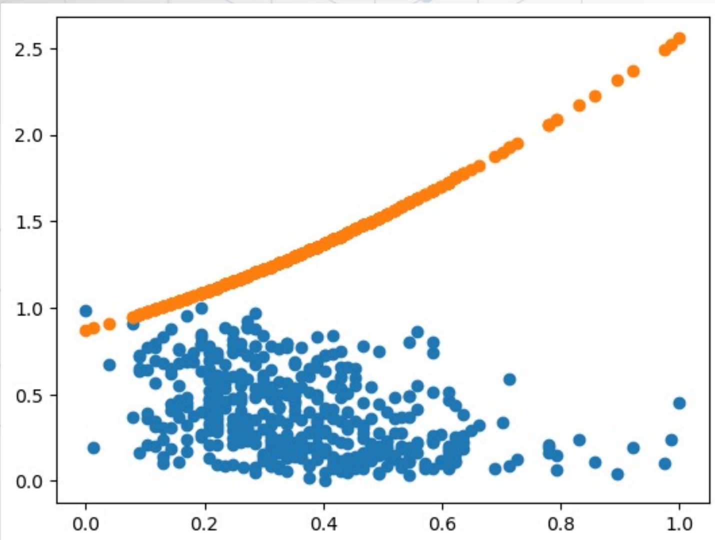
Note: The meaning of each feature (i.e. `feature_names`) might be unclear (especially for `ltg`) as the documentation of the original dataset is not explicit. We provide information that seems correct in regard with the scientific literature in this field of research.

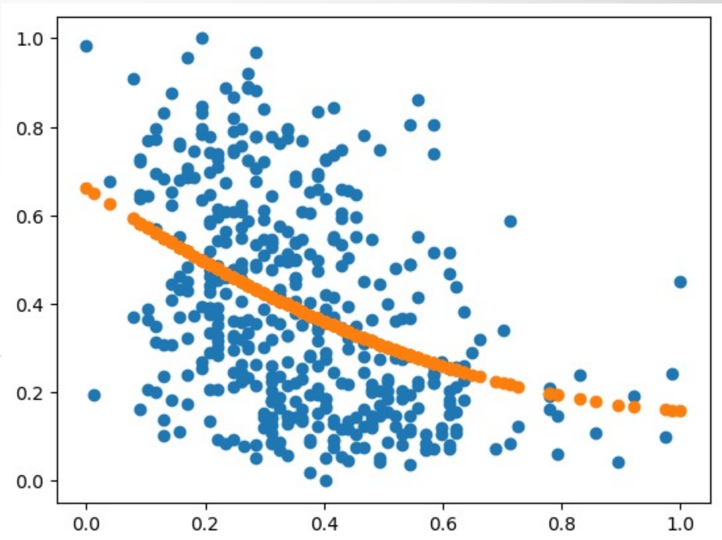
Read more in the User Guide.

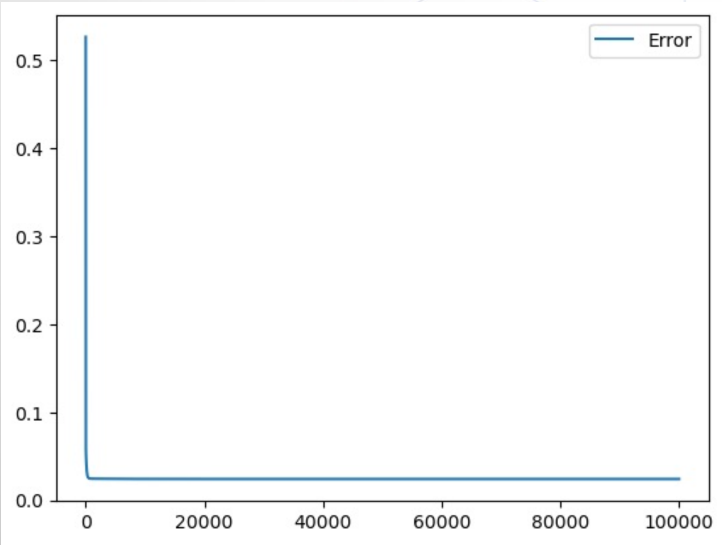
Datos reales

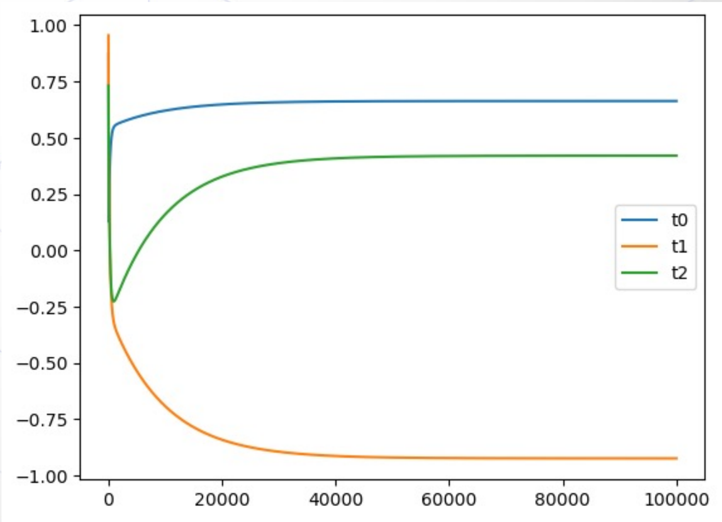
```
# Datos
from sklearn import datasets
xT, y = datasets.load_diabetes(return_X_y=True)
x = xT[:,6]
m = x.size
plt.plot(x,y,'o')
```

Datos reales









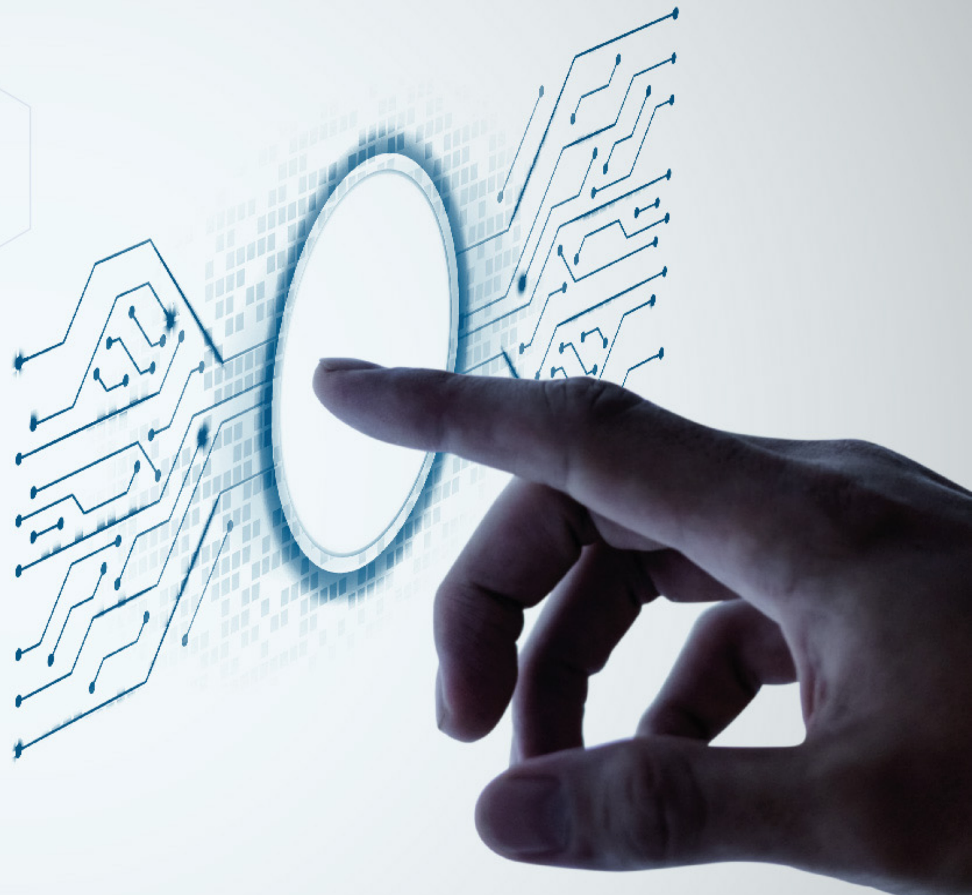
Tarea moral

- Condiciones de paro
- Descenso estocástico
- Batch learning

EngineeringX

Founded by the Royal Academy of Engineering
and Lloyd's Register Foundation

GRACIAS



https://hubiq.mx/

HUBIQRO HUBIQ HUBIQRO