

D1EAD – Análise Estatística para Ciência de Dados

2021.1



Regressão Linear Simples

Prof. Ricardo Sovat

sovat@ifsp.edu.br

Prof. Samuel Martins (Samuka)

samuel.martins@ifsp.edu.br



Régressão Linear

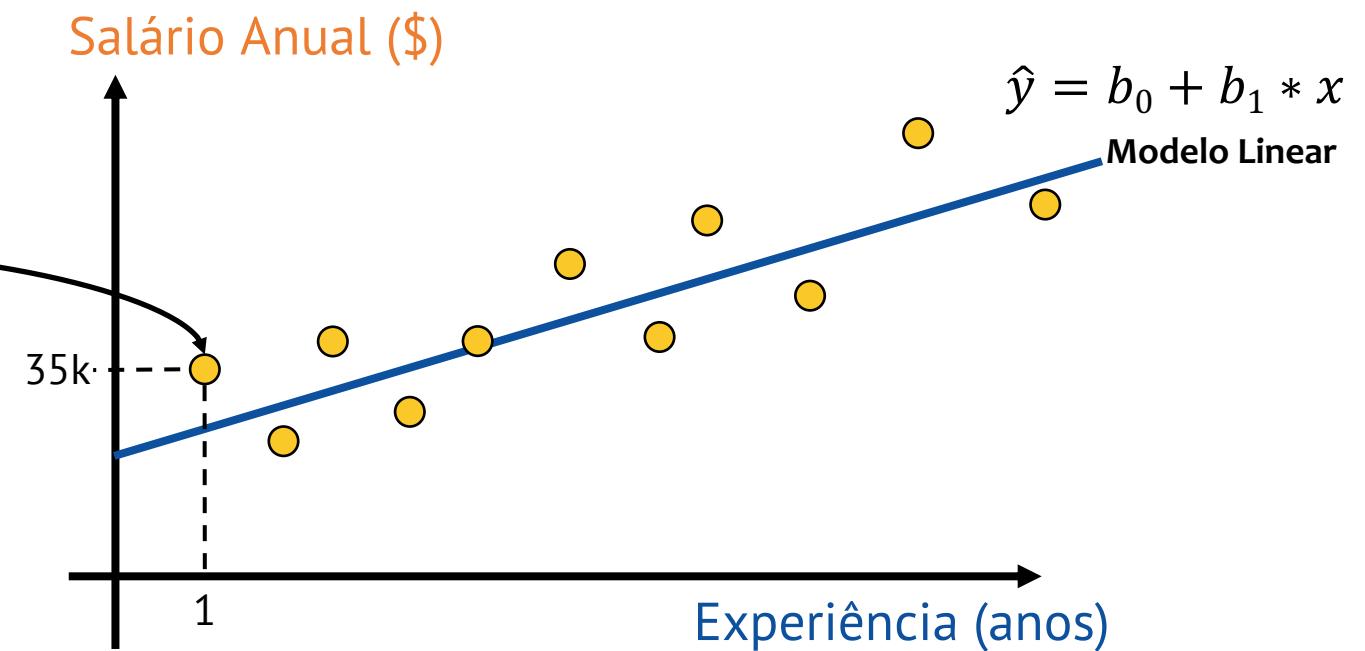
- É uma abordagem para **modelar** o **relacionamento** entre **variáveis independentes** (explicativas) e **dependentes** numéricas ajustando um **modelo linear** (p. ex., uma reta) para as observações de um conjunto treinamento.
- Tal modelo linear é usado para a **prever** variáveis dependentes numéricas para novas observações (ainda não vistas).

observações	variáveis independentes			variável dependente
	Idade	Titulação	Experiência (anos)	Salário Anual (\$)
	21	Graduação	1	35,000.00
	25	Especialização	5	80,000.00
	35	Doutorado	10	120,000.00

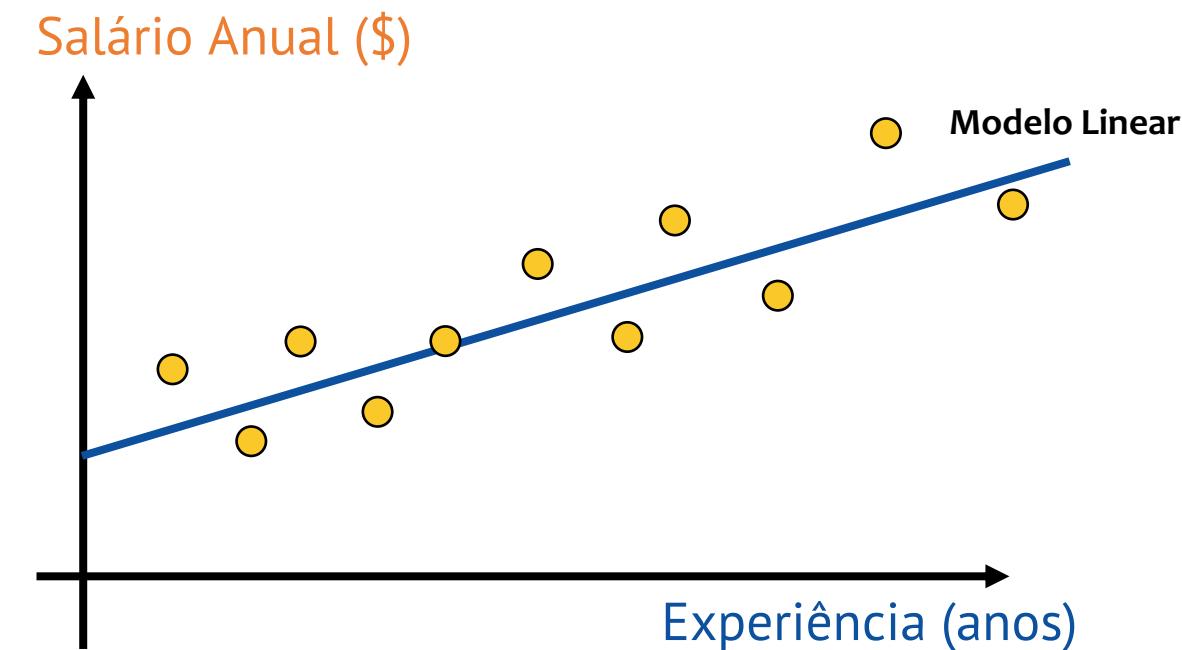
Conjunto de Treinamento				

Regressão Linear Simples: Uma variável independente

variáveis independentes	variável dependente
Experiência (anos)	Salário Anual (\$)
1	35,000.00
5	80,000.00
10	120,000.00
...	...
Conjunto de Treinamento	



Regressão Linear Simples: Uma variável independente



$$h(x_1) = \hat{y} = b_0 + b_1 * x$$

variável dependente

variável independente

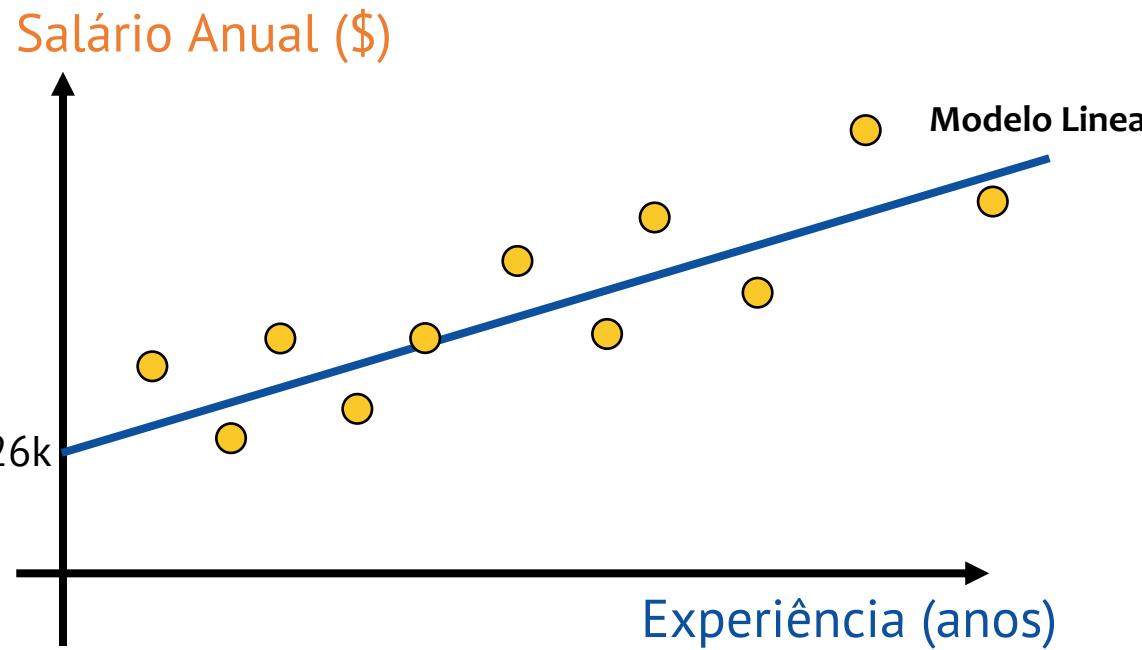
intercepto em y

modelo linear (hipótese)

inclinação da reta

$$\text{salário} = b_0 + b_1 * \text{experiência}$$

Regressão Linear Simples: Uma variável independente



$$\begin{aligned} b_0 &= 26000 \\ b_1 &= 9200 \end{aligned}$$

$\text{salário} = b_0 + b_1 * \text{experiência}$

↑
Salário Anual predito para
um trabalhador com **zero**
anos de experiência.

intercepto
em y

inclinação
da reta

Para cada **ano adicional de**
experiência, o salário
anual do trabalhador
cresce em \$ 9200.

Equação da Reta

$$(y - yo) = \mu * (x - xo)$$

simplificando para $x_o = 0$

$$\mu = b1 = \operatorname{tg}(\theta^\circ)$$

$$(y - yo) = \mu * x$$

$$\theta^\circ = \operatorname{arctg}(\mu)$$

$$y = \overline{y_o} + \underline{\mu} * x$$

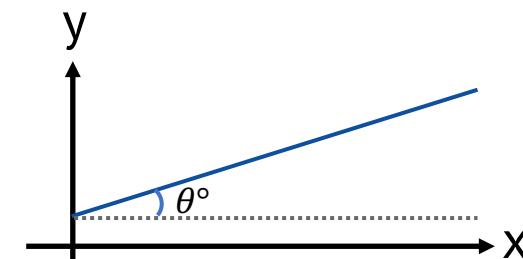
intercepto
em y

inclinação
da reta

$$h(x) = \hat{y} = \overline{b_0} + \underline{b_1} * x$$

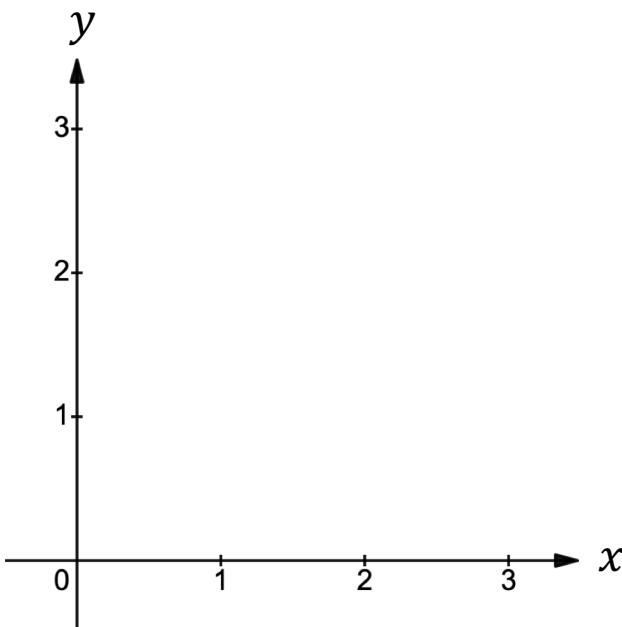
intercepto
em y

inclinação
da reta

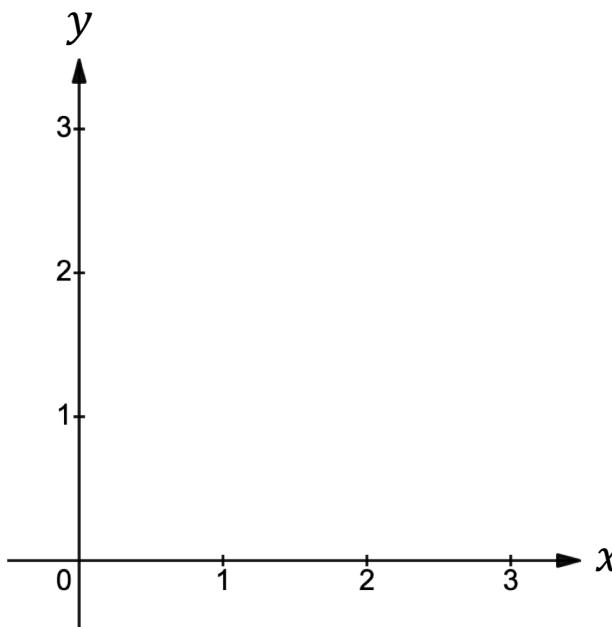


Equação da Reta: Exemplos

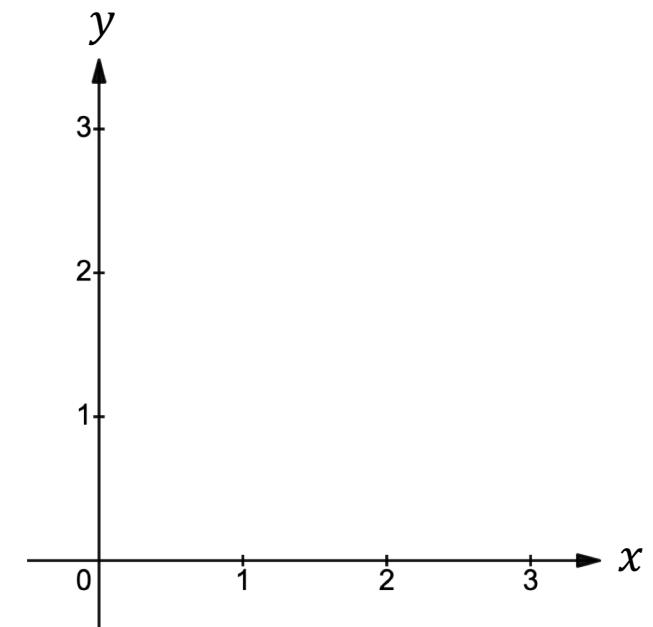
$$h(x) = \hat{y} = \underbrace{b_0}_{\substack{\text{intercepto} \\ \text{em } y}} + \underbrace{b_1 * x}_{\substack{\text{inclinação} \\ \text{da reta}}}$$



$$\begin{aligned} b_0 &= 2.0 \\ b_1 &= 0.0 \end{aligned}$$



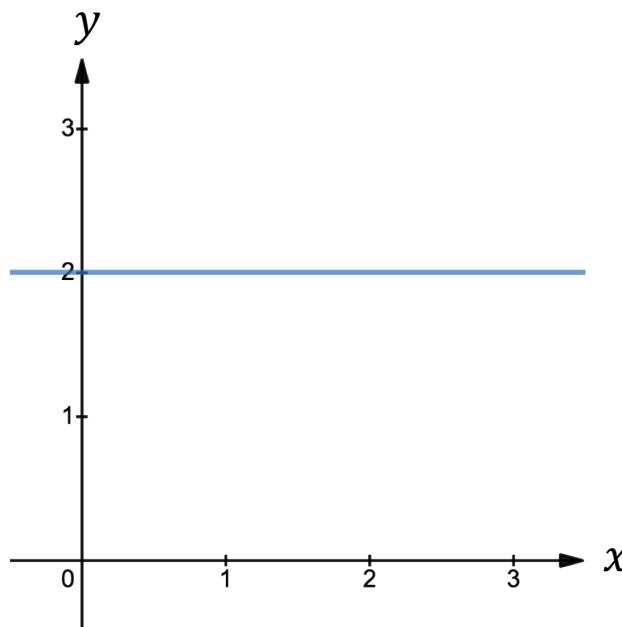
$$\begin{aligned} b_0 &= 0.0 \\ b_1 &= 1.0 \\ (\theta^\circ &= 45^\circ) \end{aligned}$$



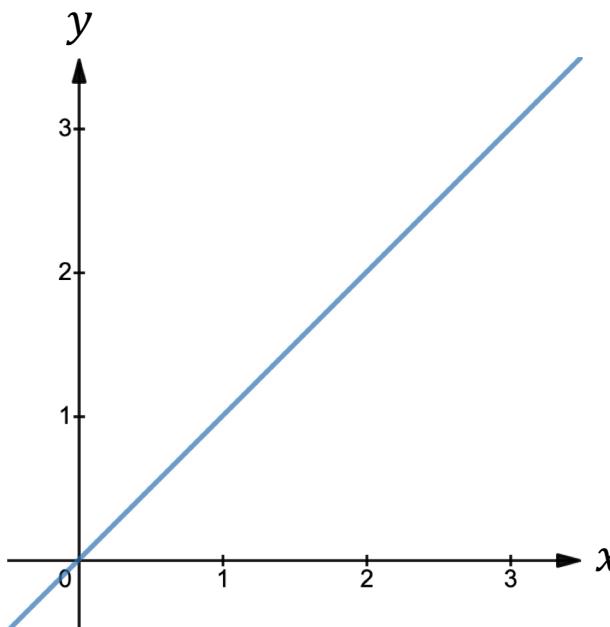
$$\begin{aligned} b_0 &= 1.0 \\ b_1 &= 0.5 \\ (\theta^\circ &\approx 27^\circ) \end{aligned}$$

Equação da Reta: Exemplos

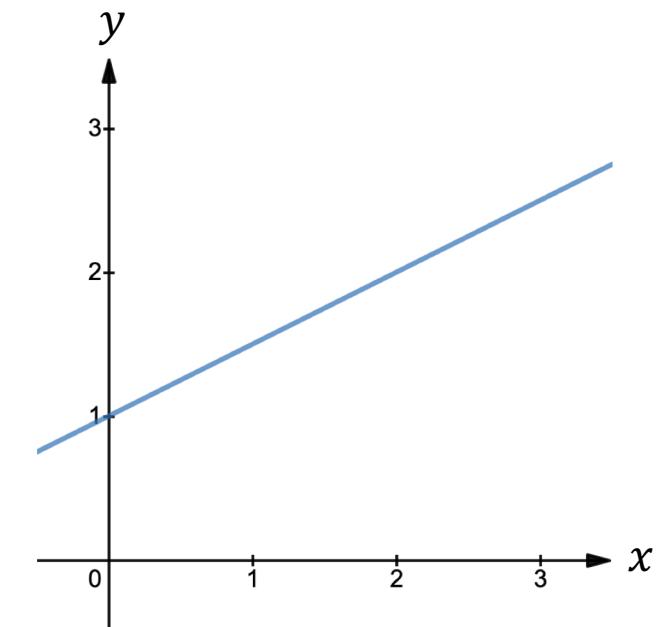
$$h(x) = \hat{y} = \underbrace{b_0}_{\substack{\text{intercepto} \\ \text{em } y}} + \underbrace{b_1 * x}_{\substack{\text{inclinação} \\ \text{da reta}}}$$



$$\begin{aligned} b_0 &= 2.0 \\ b_1 &= 0.0 \end{aligned}$$

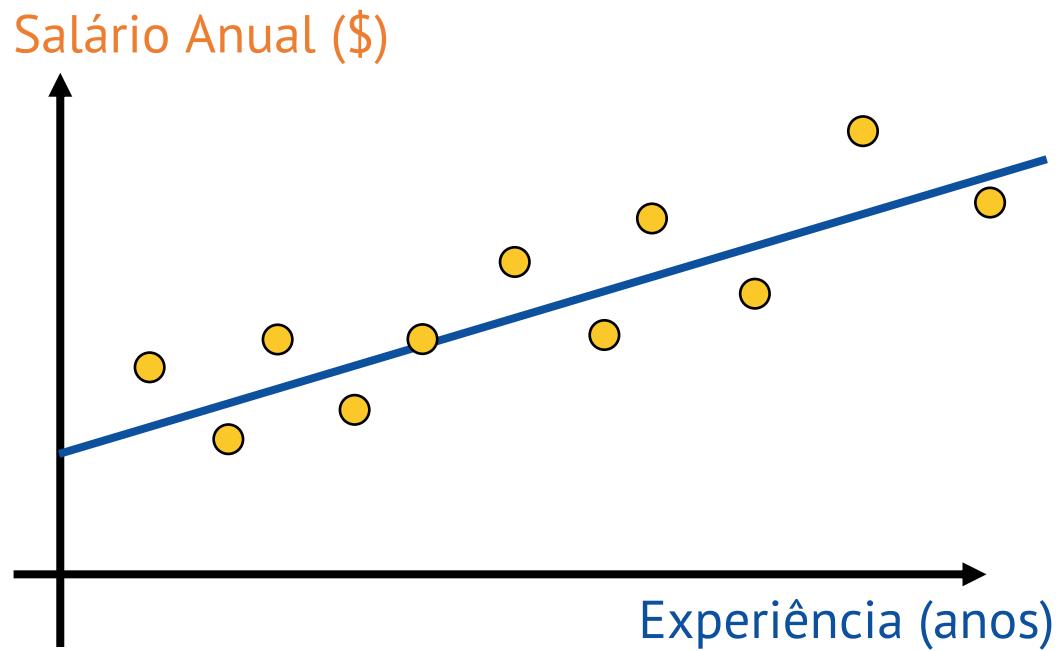


$$\begin{aligned} b_0 &= 0.0 \\ b_1 &= 1.0 \\ (\theta^\circ) &= 45^\circ \end{aligned}$$

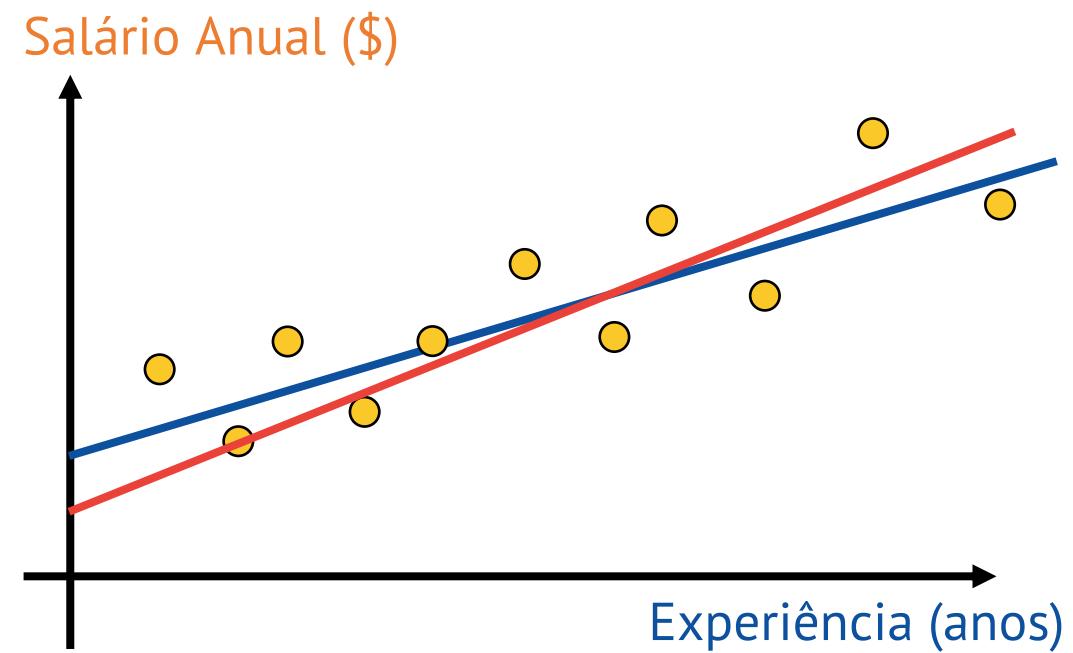


$$\begin{aligned} b_0 &= 1.0 \\ b_1 &= 0.5 \\ (\theta^\circ) &\approx 27^\circ \end{aligned}$$

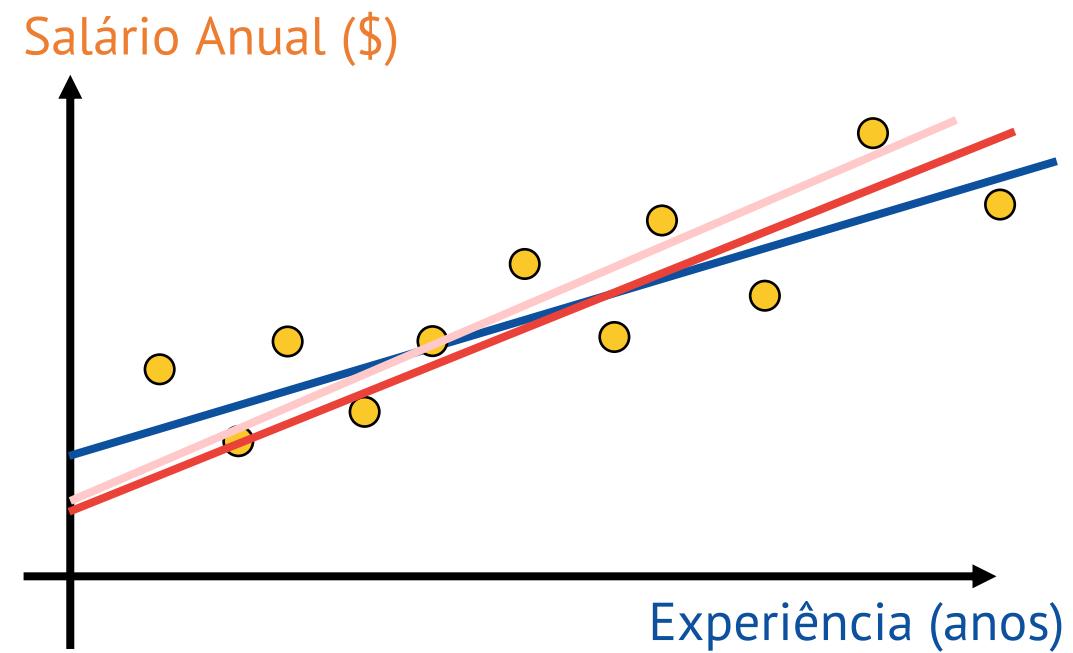
Qual reta é a melhor?



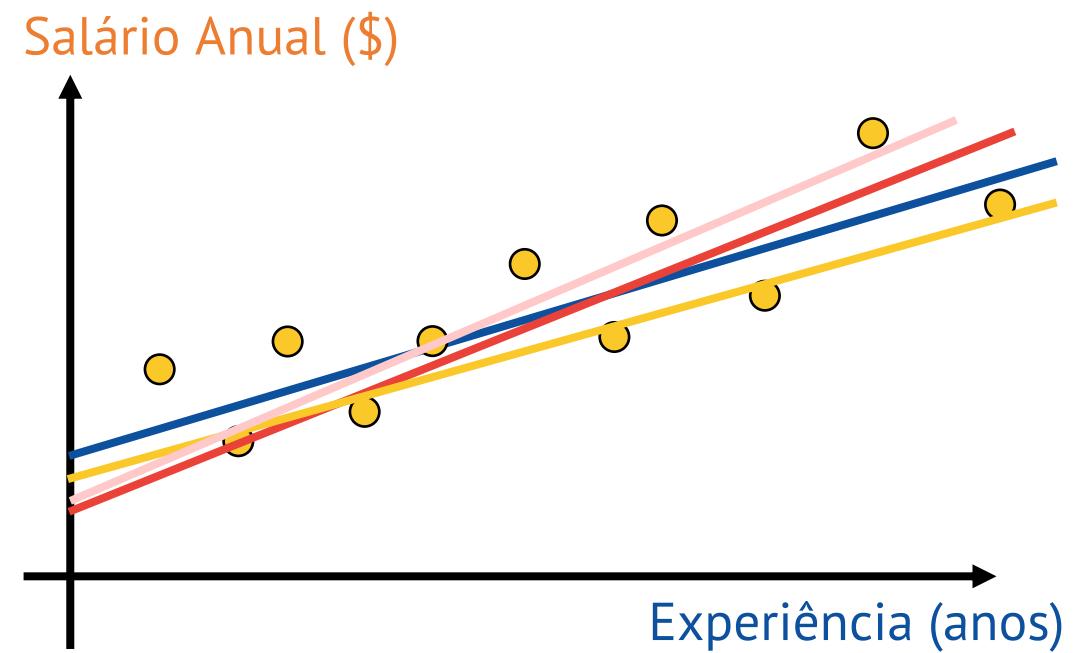
Qual reta é a melhor?



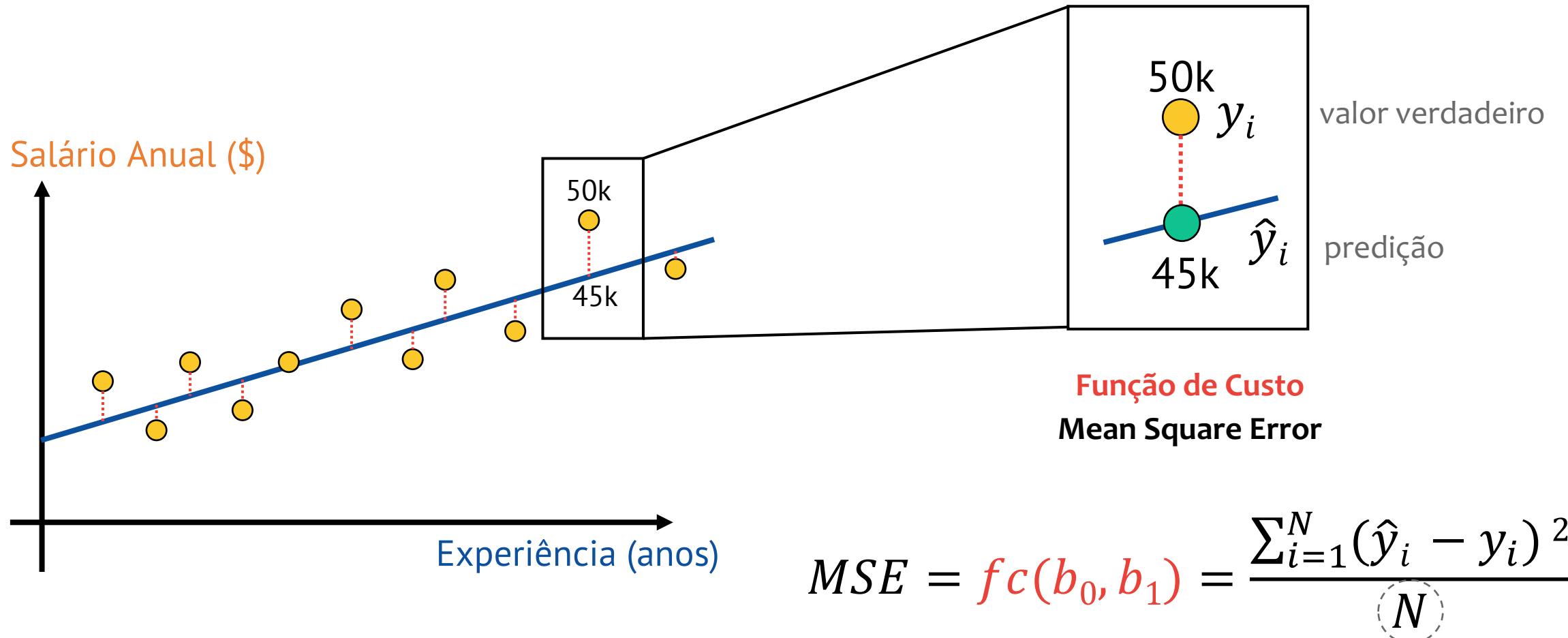
Qual reta é a melhor?



Qual reta é a melhor?



Qual reta é a melhor? – Erro do Modelo Linear



Objetivo: Achar os parâmetros b_0 e b_1 que **minimizam a função de custo** para o conjunto de treinamento.

Modelo linear (hipótese)

$$h(x) = \hat{y} = b_0 + b_1 * x$$

Modelo linear Simplificado

$$h(x) = \hat{y} = b_1 * x$$

simplificando: $b_0 = 0$

Função de Custo

$$fc(b_0, b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$

Função de Custo

$$fc(b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$

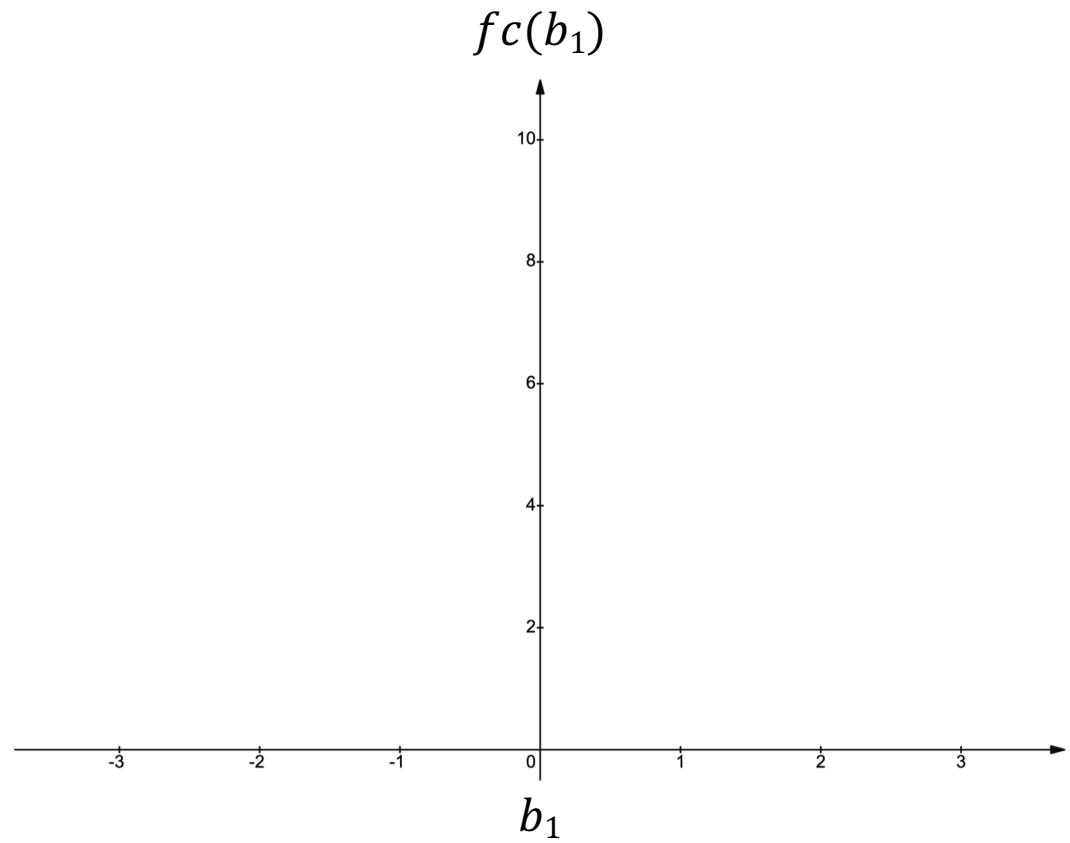
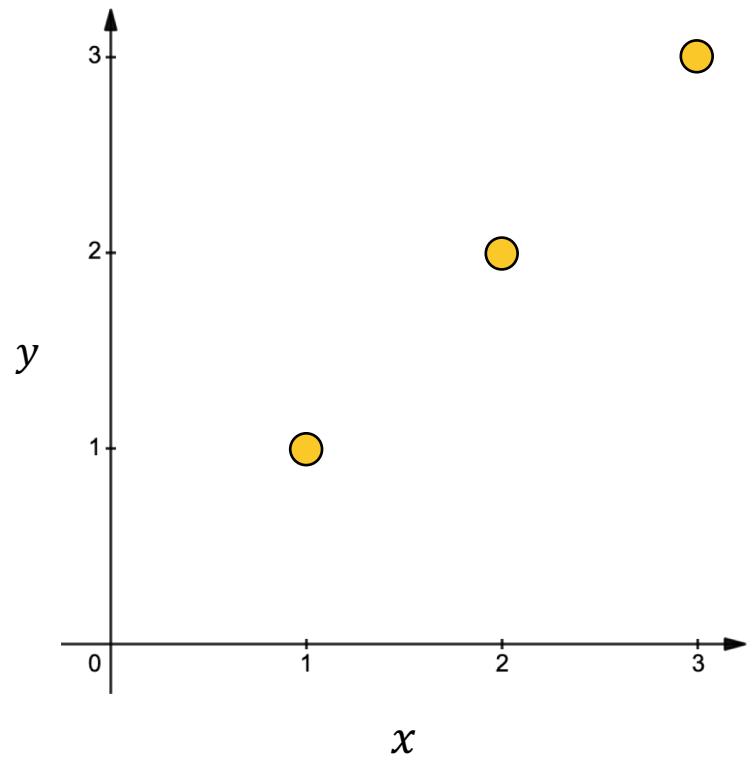
Objetivo: $\min_{b_0, b_1} fc(b_0, b_1)$

Objetivo: $\min_{b_1} fc(b_1)$

Exemplo: Função de Custo

$$h(x) = \hat{y} = b_1 * x$$

$$fc(b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$

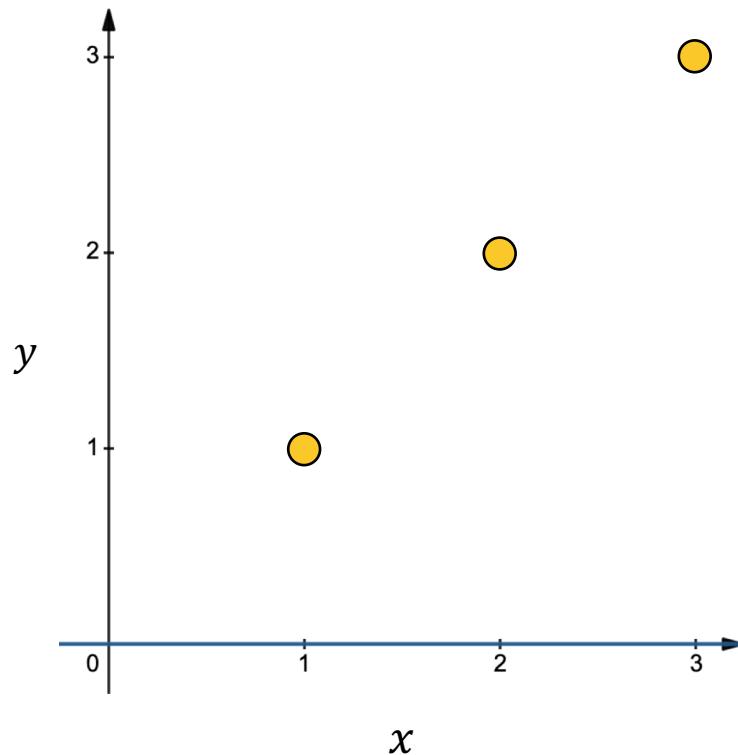


Exemplo: Função de Custo

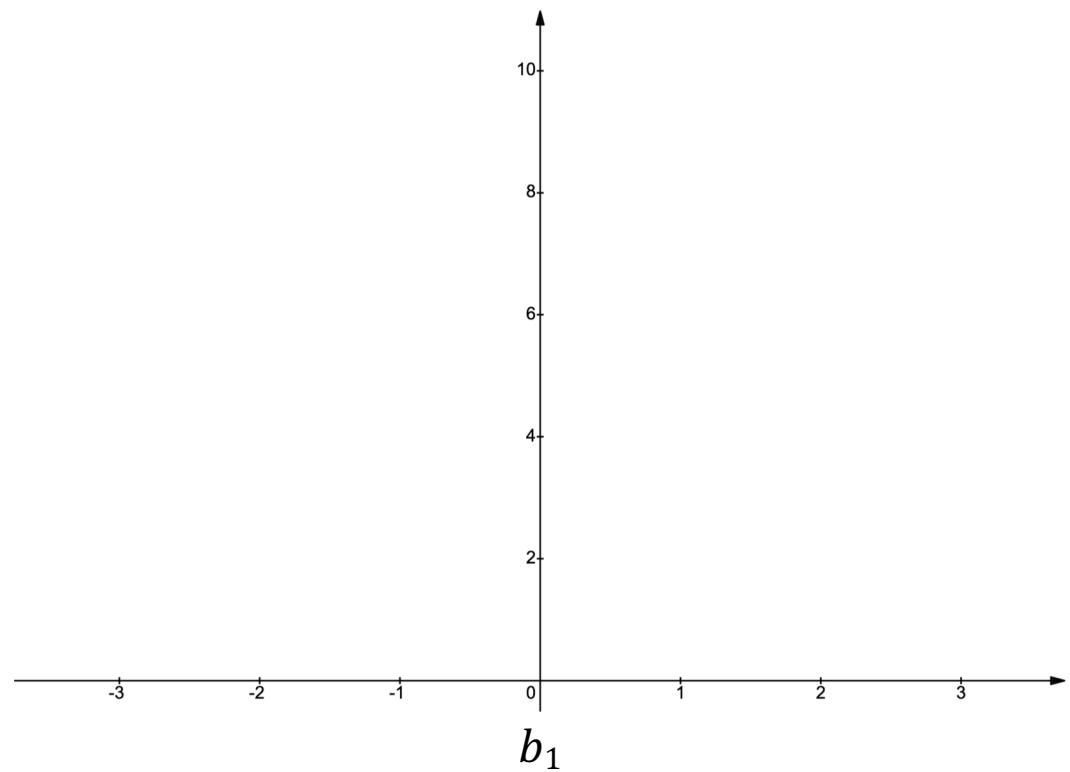
$$h(x) = \hat{y} = b_1 * x$$

$$fc(b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$

$$b_1 = 0$$



$$fc(b_1)$$

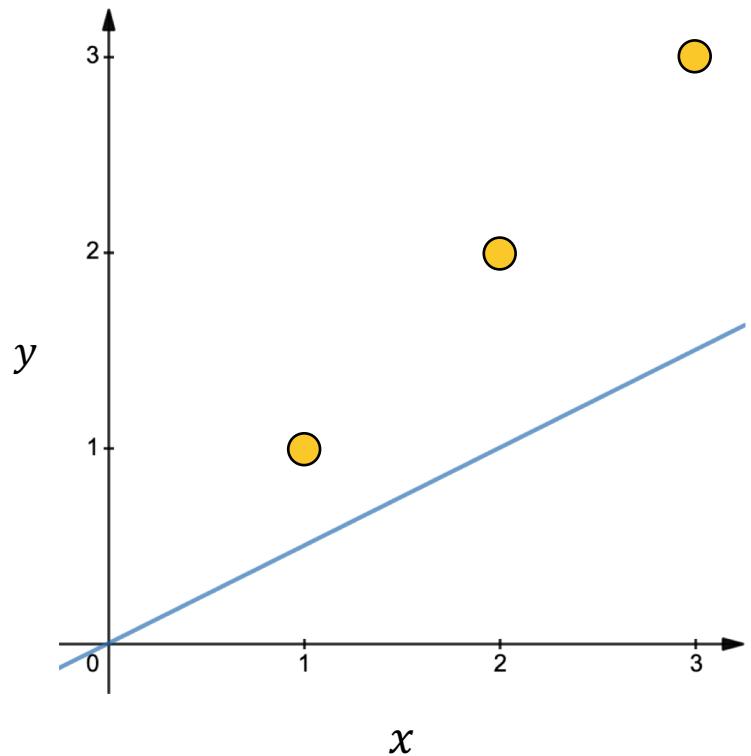


Exemplo: Função de Custo

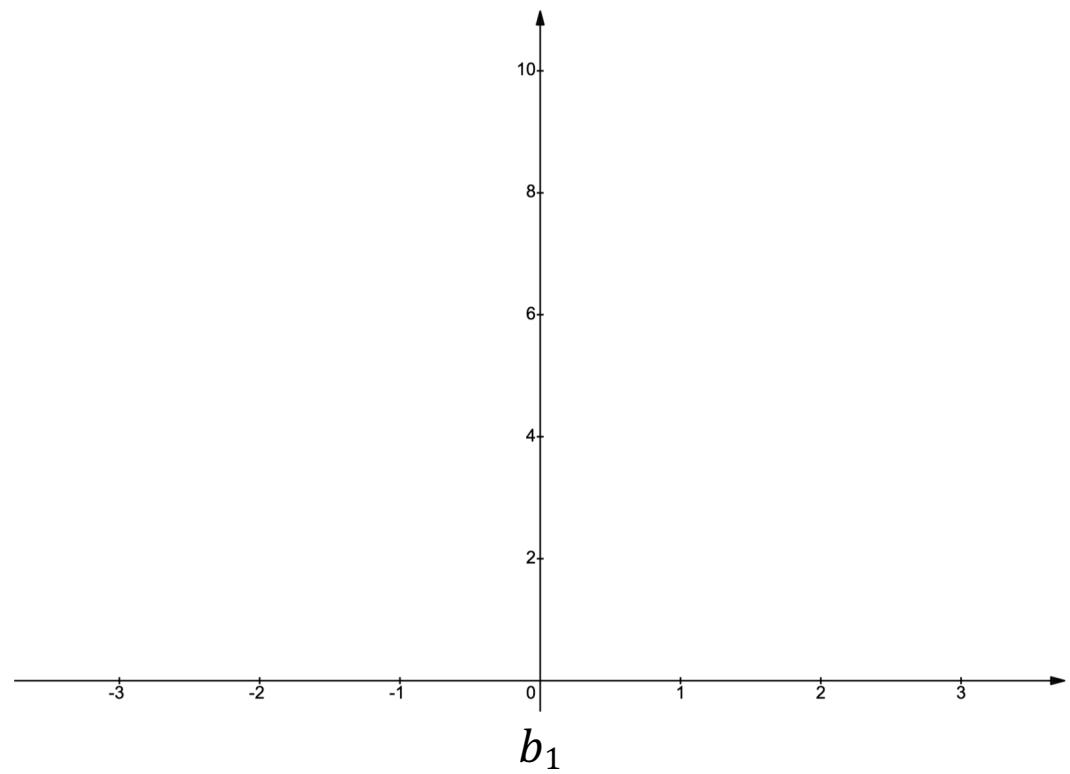
$$h(x) = \hat{y} = b_1 * x$$

$$fc(b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$

$$b_1 = 0.5$$



$$fc(b_1)$$

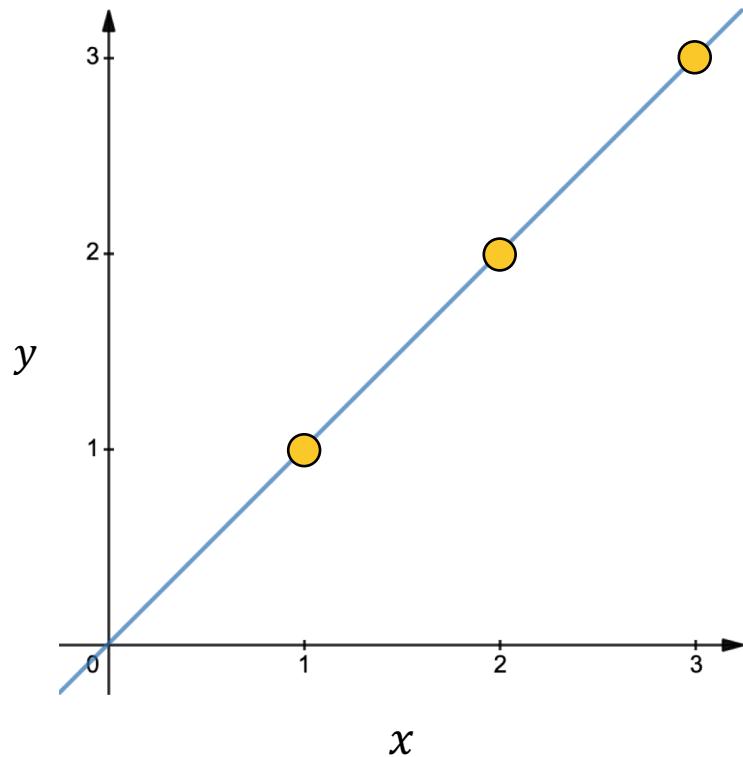


Exemplo: Função de Custo

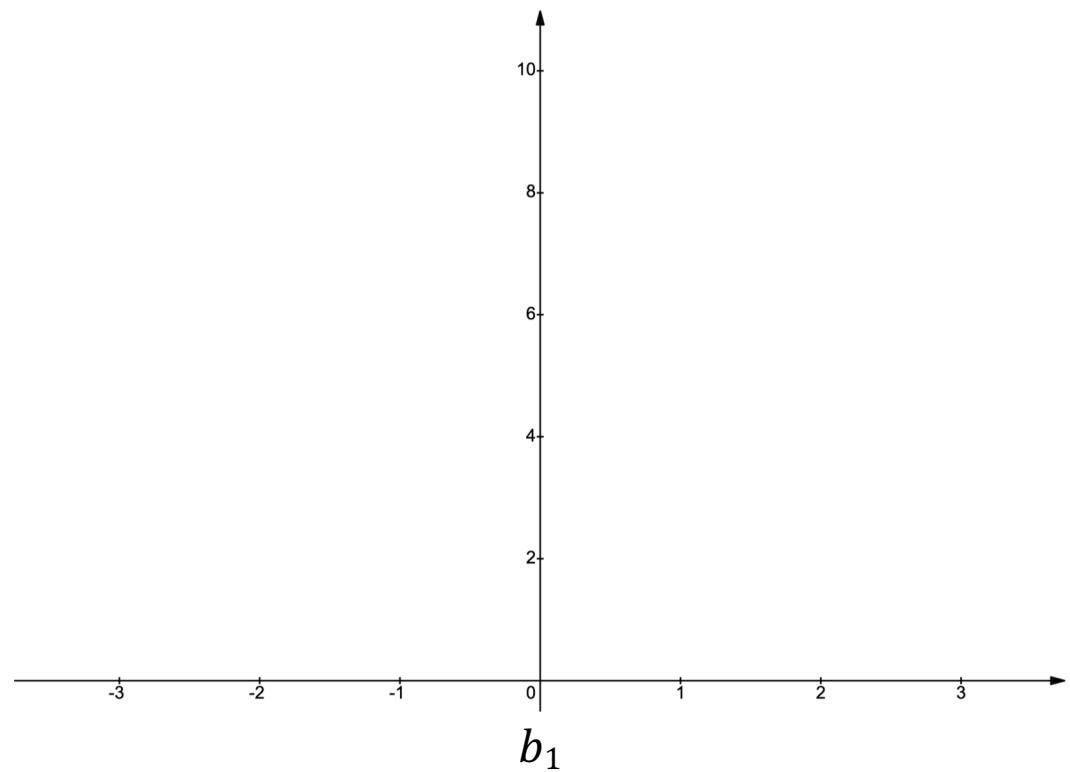
$$h(x) = \hat{y} = b_1 * x$$

$$fc(b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$

$$b_1 = 1$$



$$fc(b_1)$$

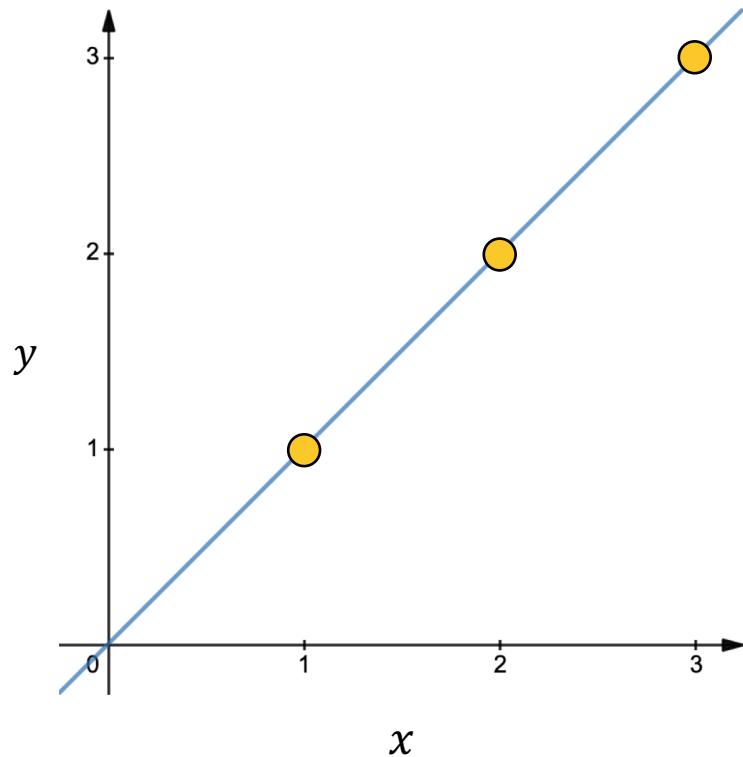


Exemplo: Função de Custo

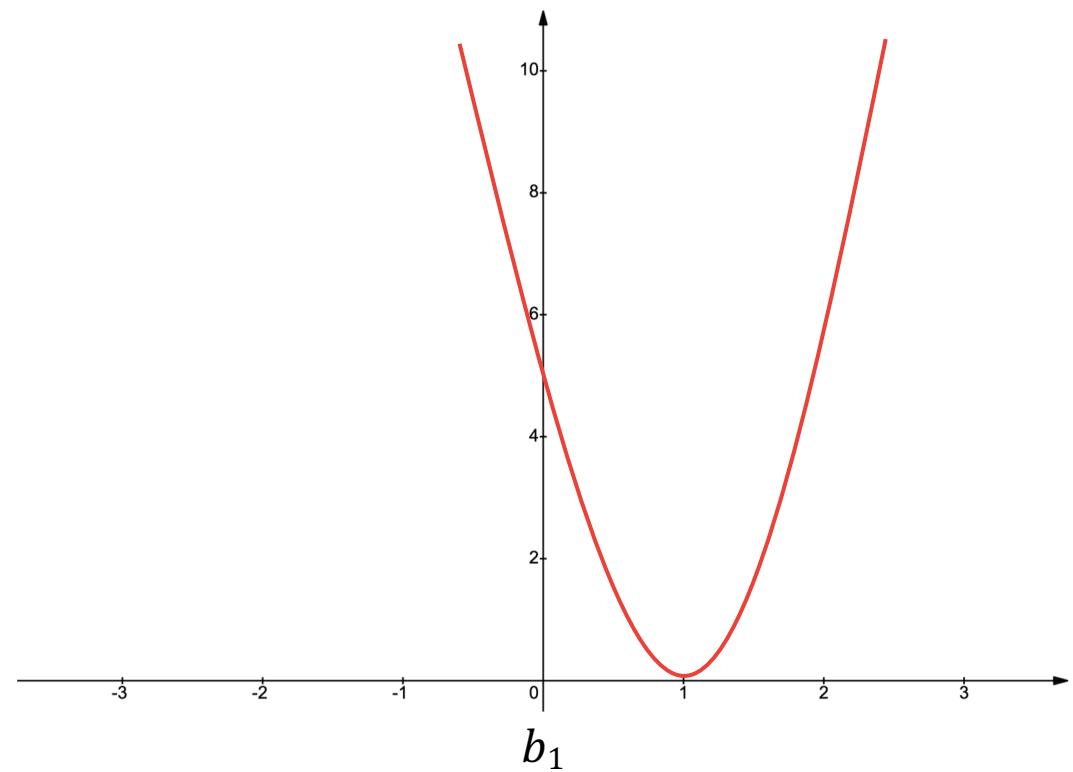
$$h(x) = \hat{y} = b_1 * x$$

$$fc(b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$

$$b_1 = 1$$



$$fc(b_1)$$



Modelo linear (hipótese)

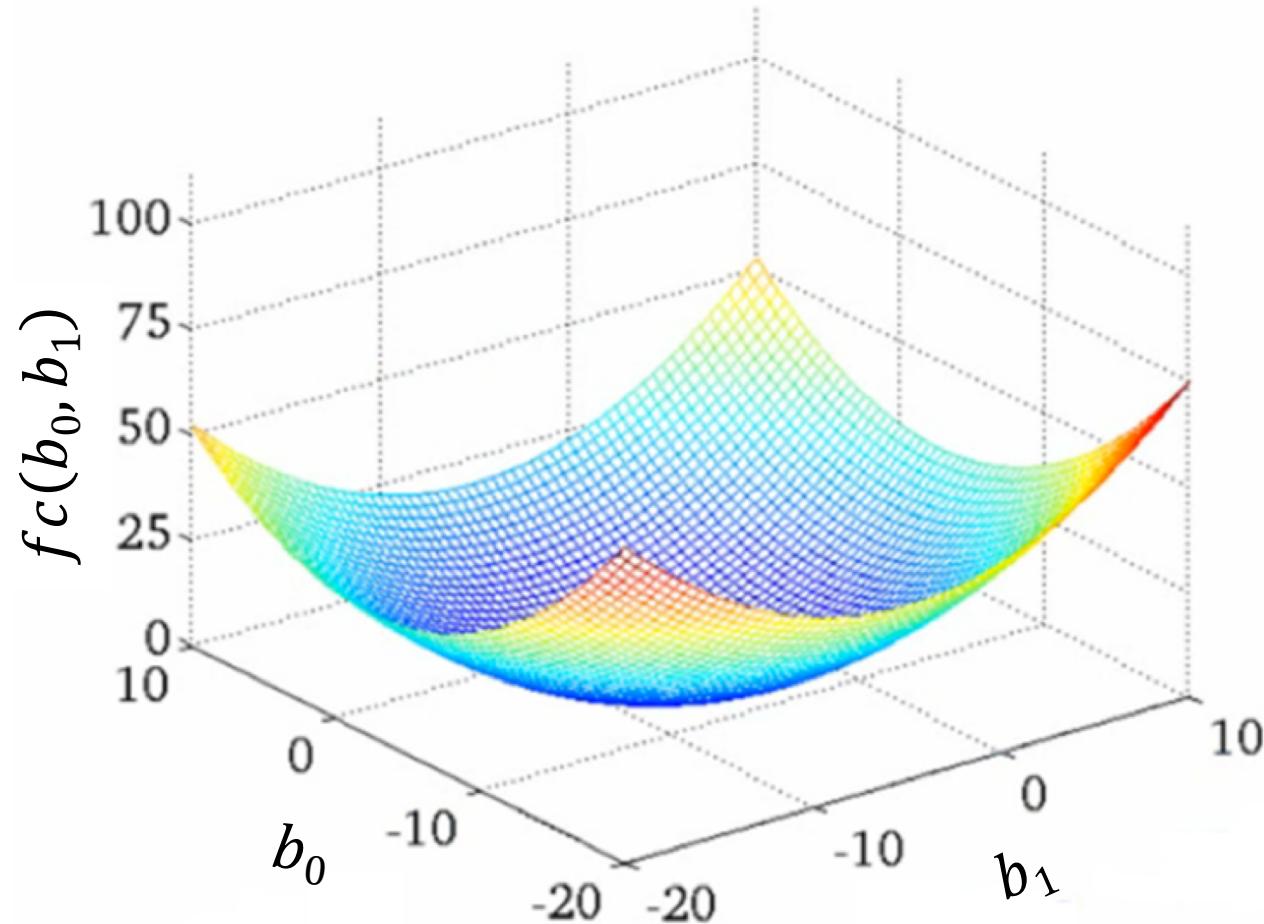
$$h(x) = \hat{y} = b_0 + b_1 * x$$

Função de Custo

$$fc(b_0, b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$

Objetivo: $\min_{b_0, b_1} fc(b_0, b_1)$

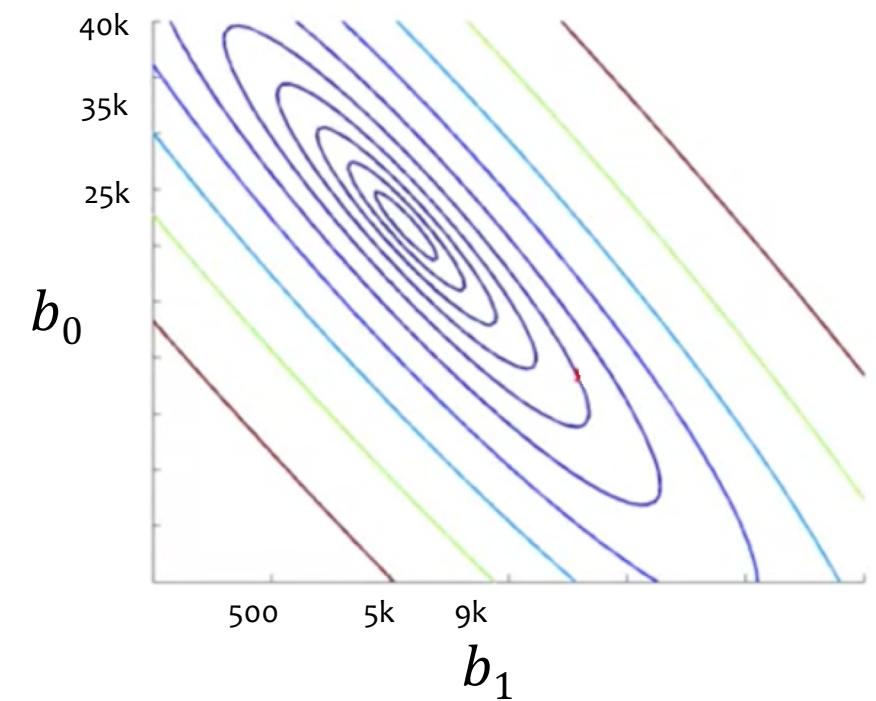
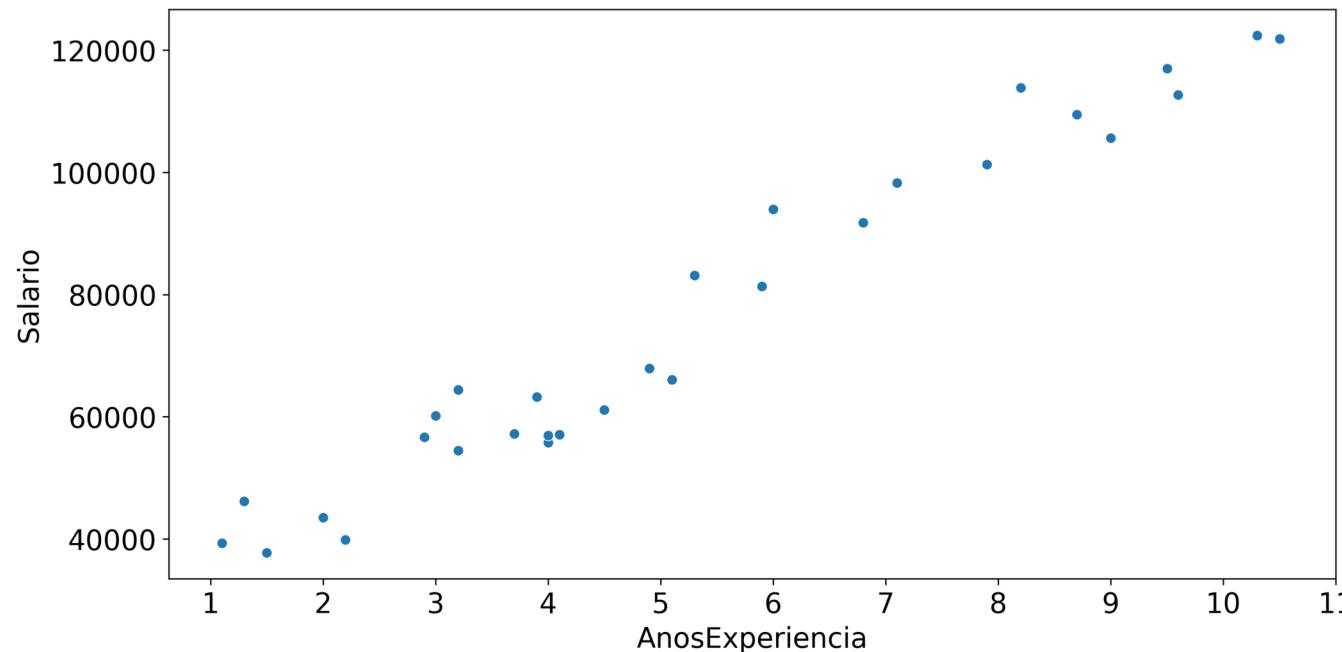
$$fc(b_0, b_1) = MSE = \frac{\sum_{i=1}^N (\widehat{y}_i - y_i)^2}{N}$$



Exemplo: Função de Custo

$$h(x) = \hat{y} = b_0 + b_1 * x$$

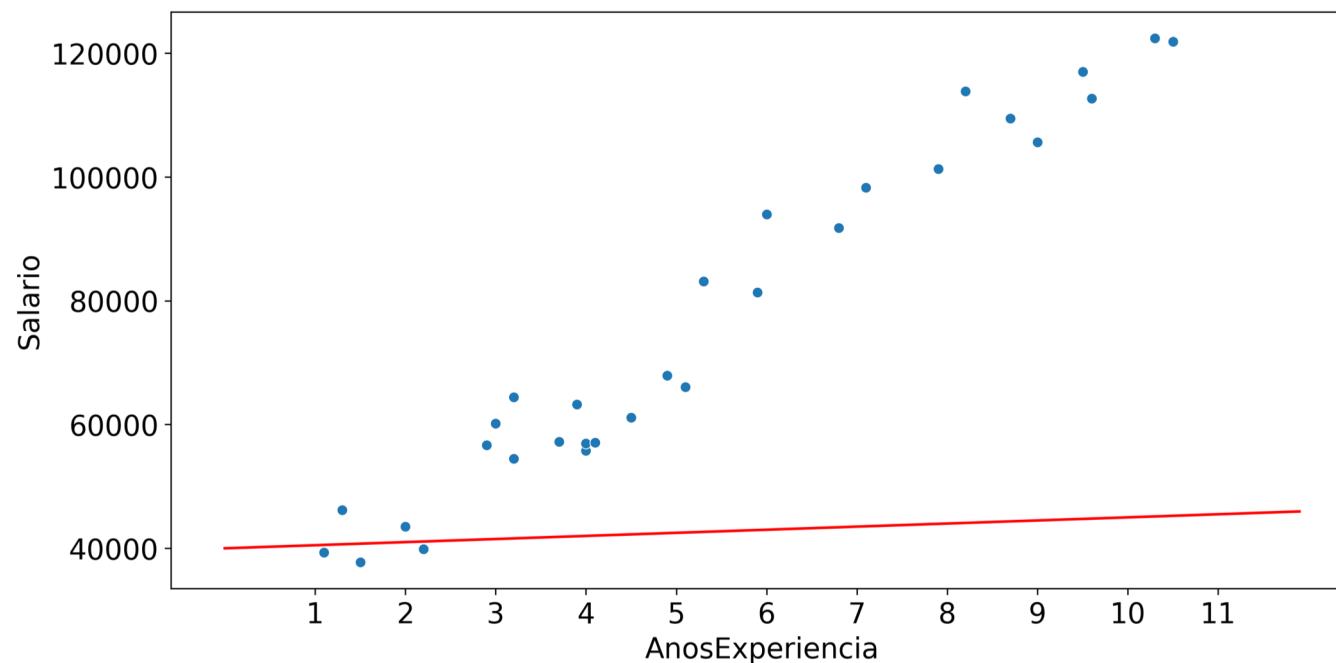
$$fc(b_0, b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$



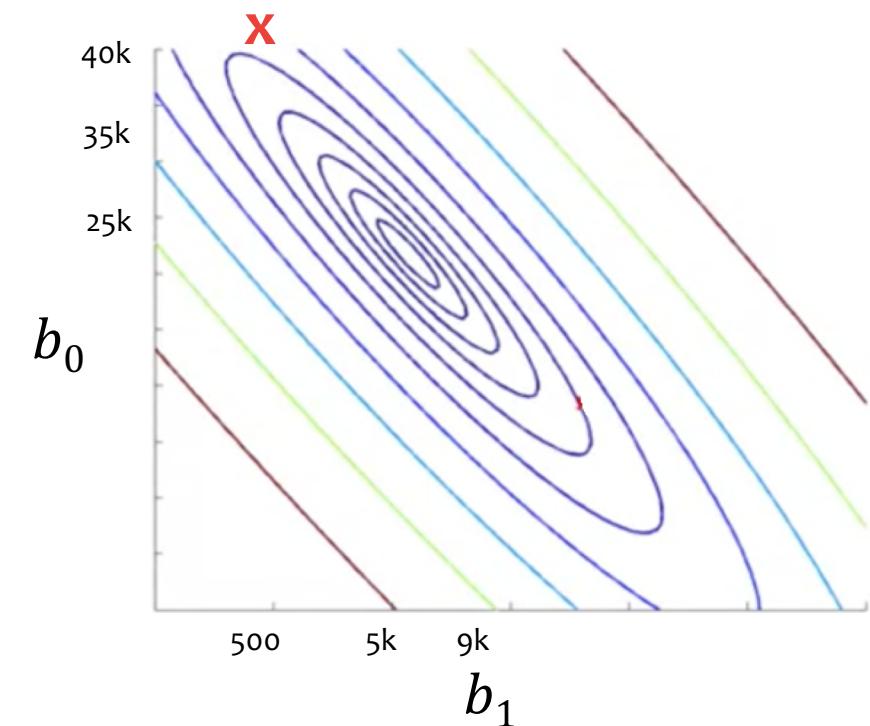
Exemplo: Função de Custo

$$h(x) = \hat{y} = b_0 + b_1 * x$$

$$fc(b_0, b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$



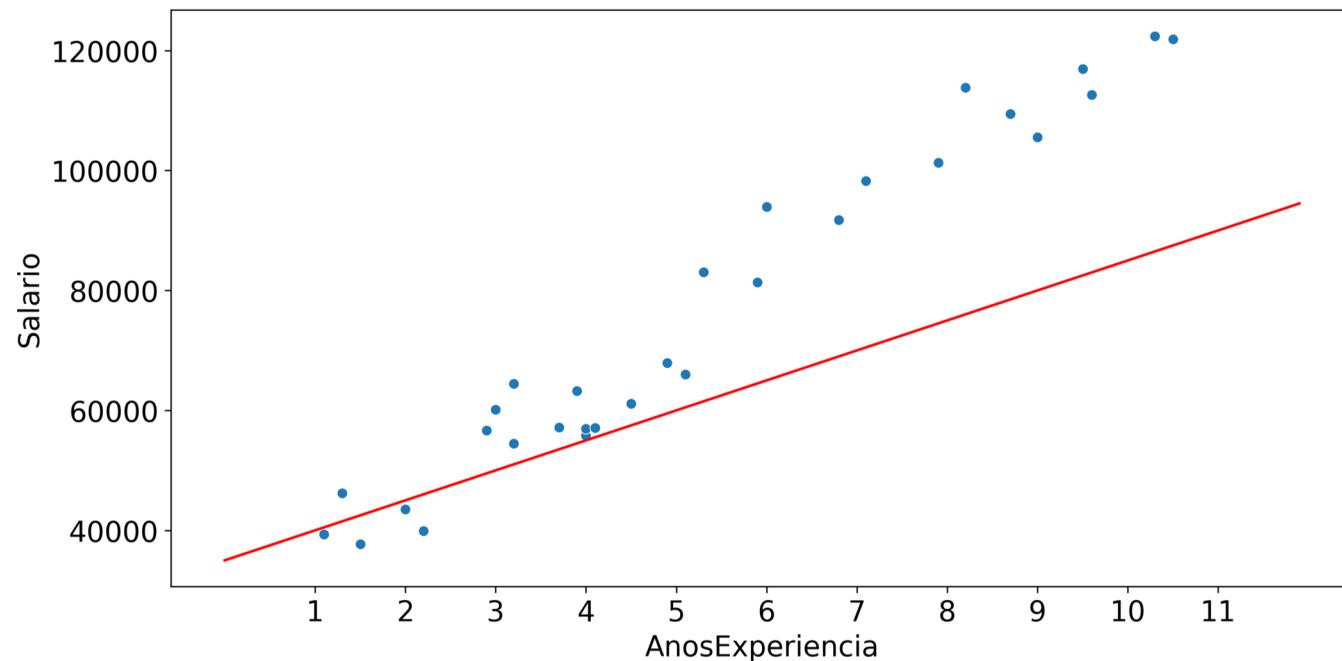
$$\begin{aligned}b_0 &= 40000 \\b_1 &= 500\end{aligned}$$



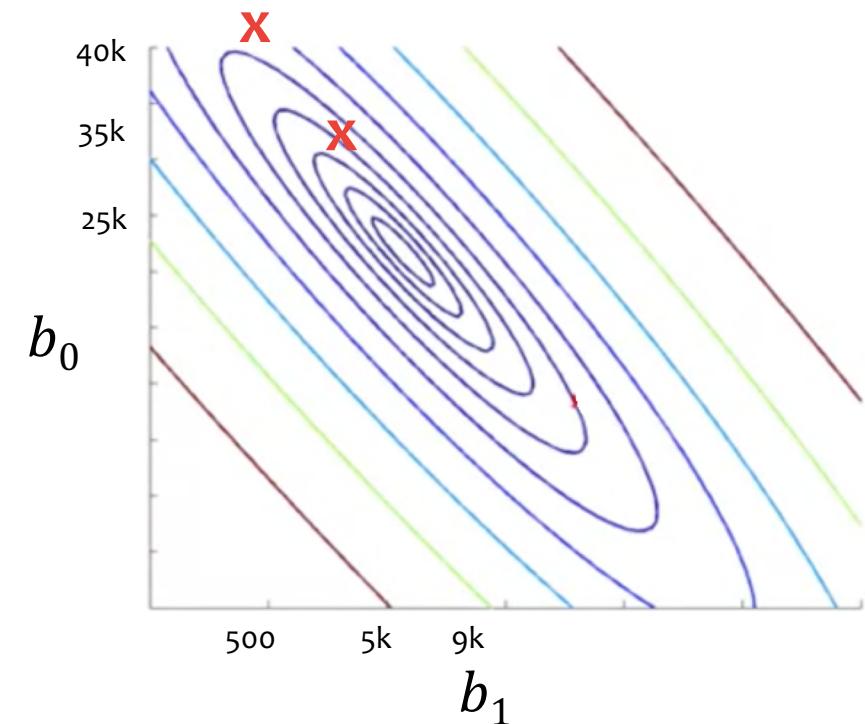
Exemplo: Função de Custo

$$h(x) = \hat{y} = b_0 + b_1 * x$$

$$fc(b_0, b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$



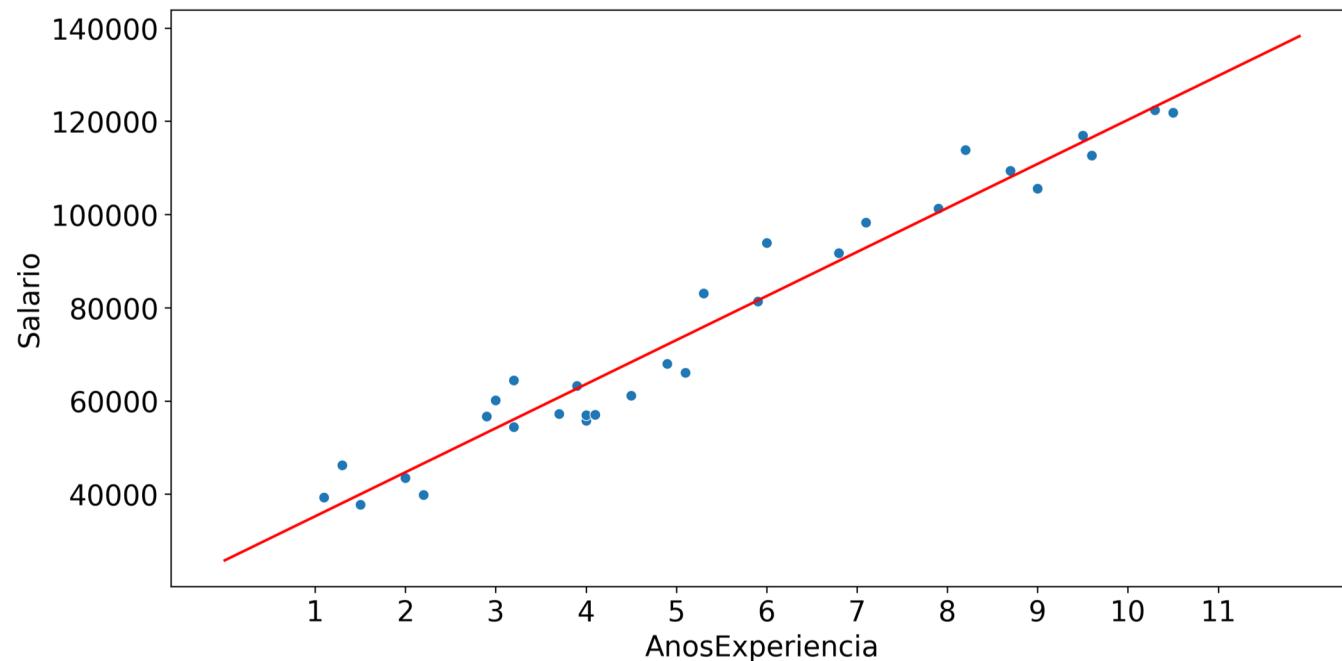
$$\begin{aligned}b_0 &= 35000 \\b_1 &= 5000\end{aligned}$$



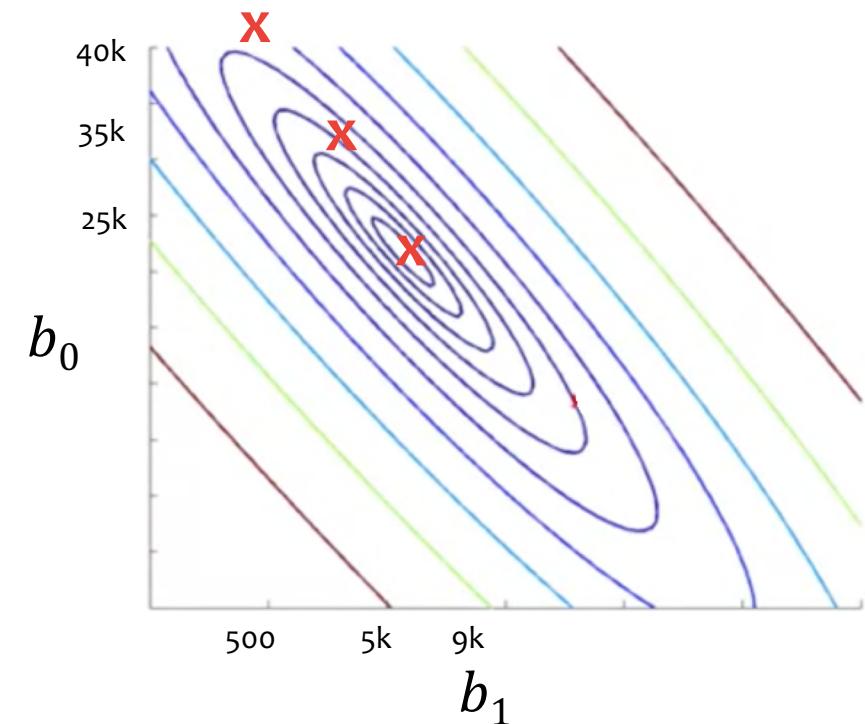
Exemplo: Função de Custo

$$h(x) = \hat{y} = b_0 + b_1 * x$$

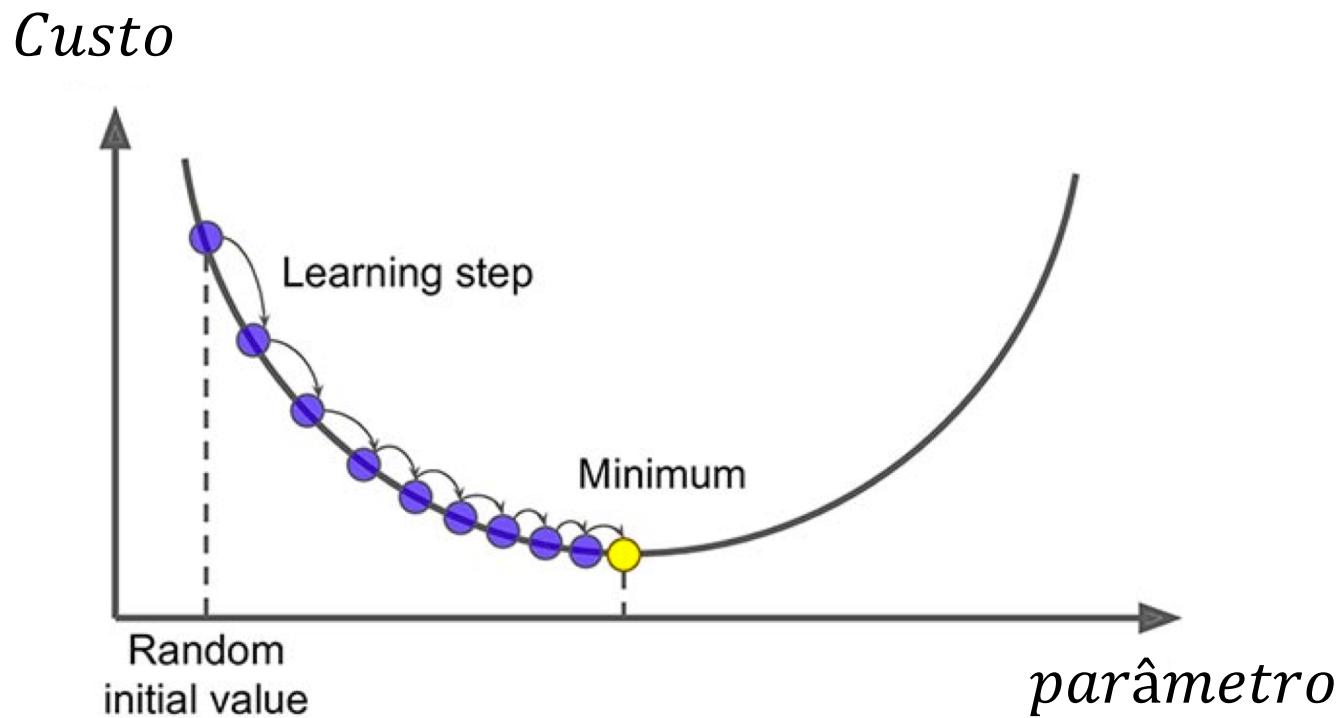
$$fc(b_0, b_1) = MSE = \frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}$$



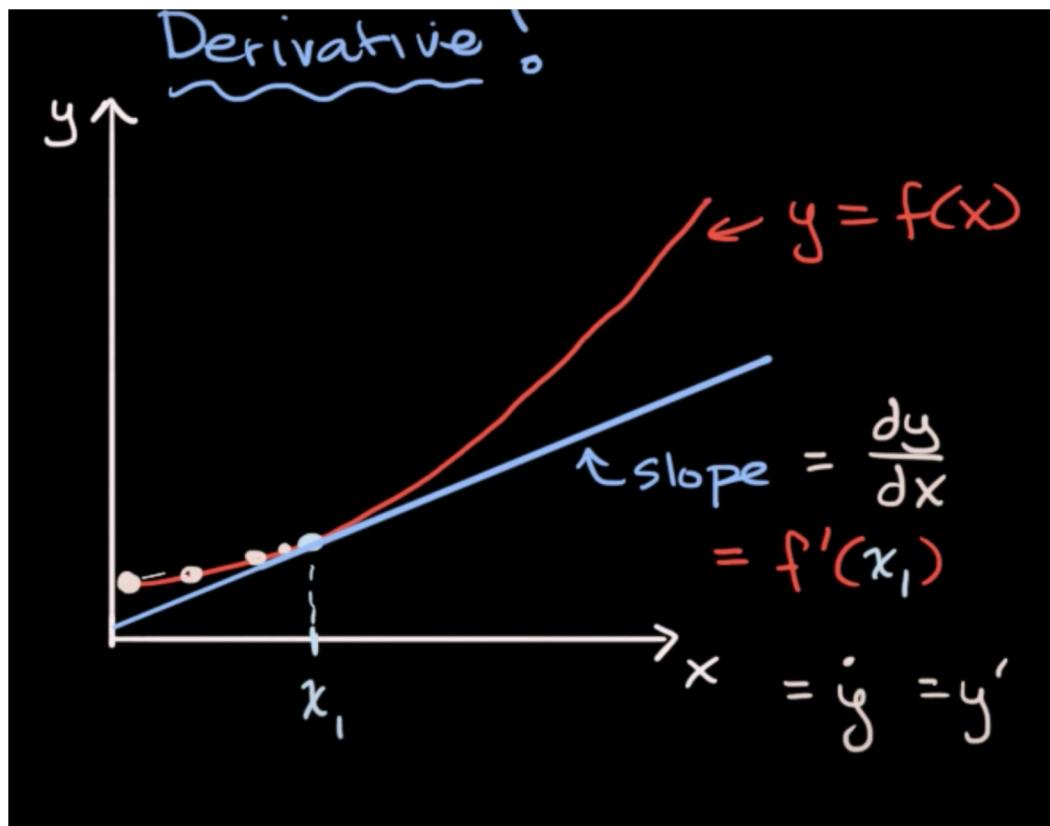
$$\begin{aligned}b_0 &= 25000 \\b_1 &= 9000\end{aligned}$$



Gradiente Descendente



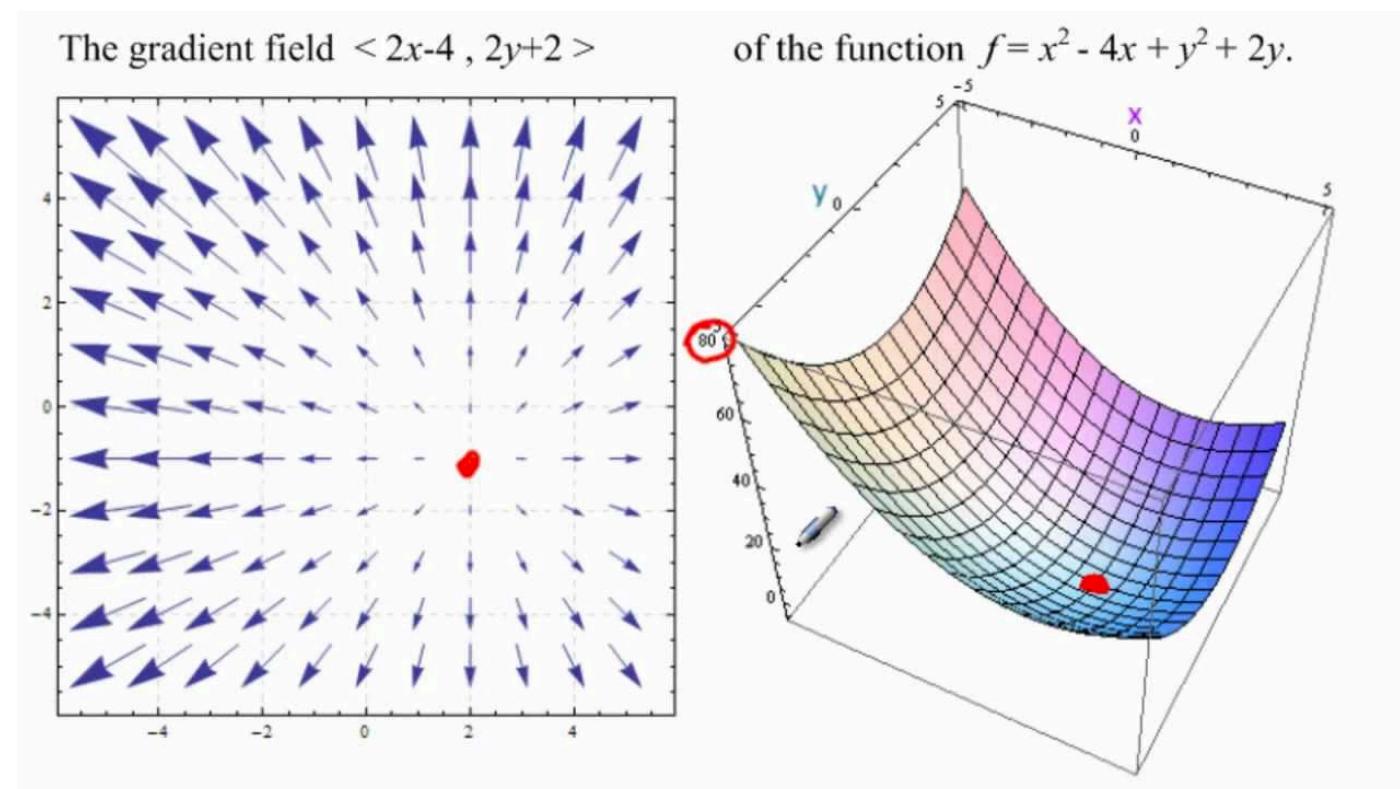
Derivada de uma função



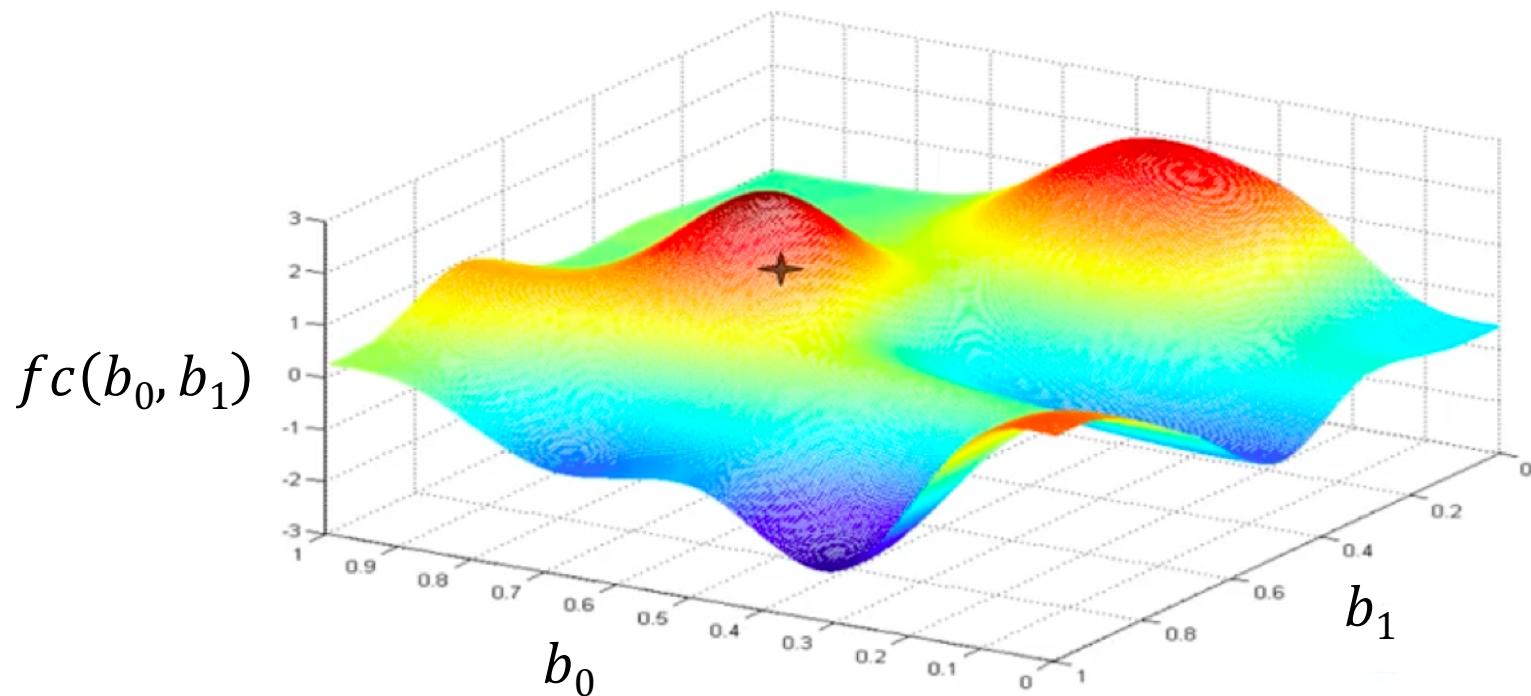
A **derivada** de uma **função/curva $y = f(x)$** em um dado ponto x_1 é a inclinação da **reta tangente à curva** no ponto x_1 ,
i.e., a tangente do ângulo da **reta tangente à curva** no ponto x_1 .

Vetor Gradiente

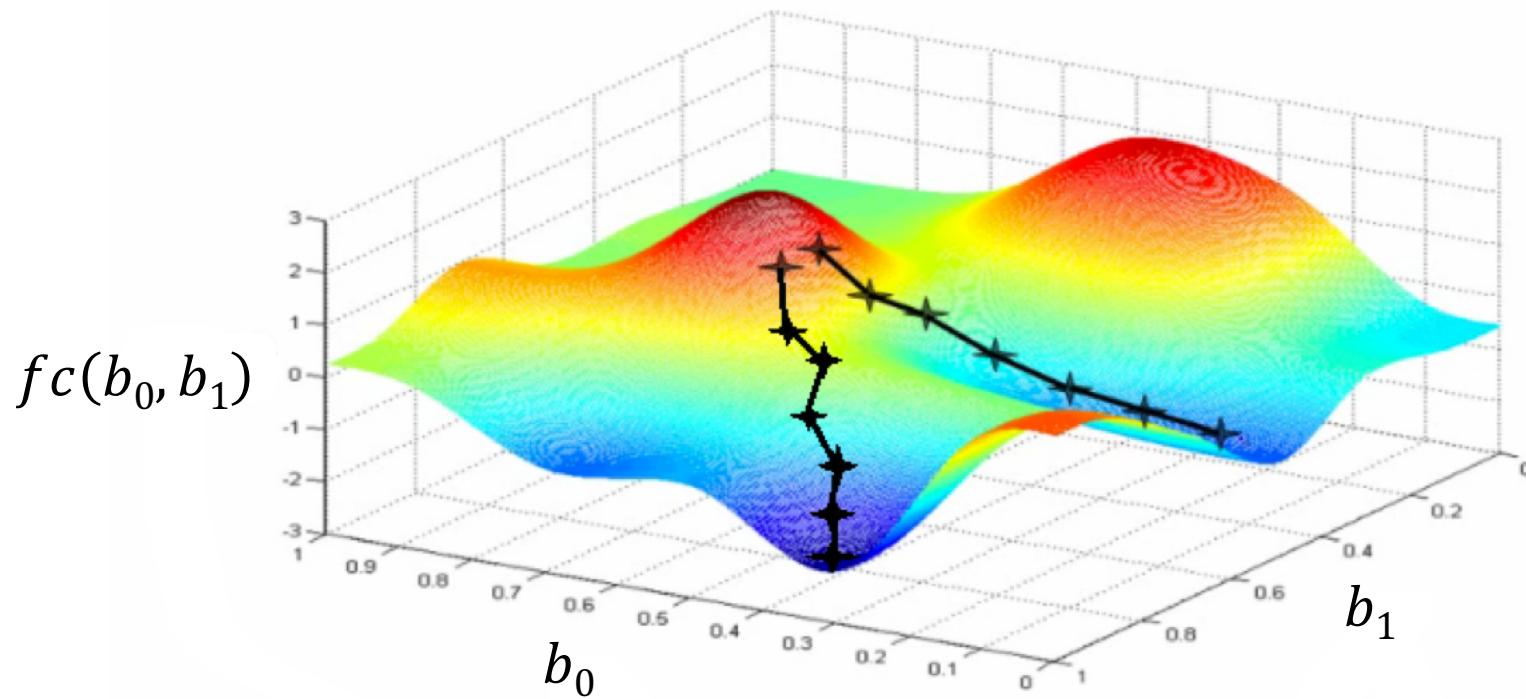
- O **vetor gradiente**, ou simplesmente **gradiente**, é um vetor que aponta para **a direção de maior crescimento da função**, para um dado ponto da função, com uma dada **taxa de crescimento**.
- É formado pela **derivada parcial** de cada variável da função.



Gradiente Descendente



Gradiente Descendente



Gradiente Descendente: Algoritmo

Dado a função de custo: $fc(b_0, b_1)$

Objetivo: $\min_{b_0, b_1} fc(b_0, b_1)$

Algoritmo:

- Inicialize os parâmetros b_0, b_1 (p. ex., $b_0 = 0, b_1 = 0$);
- Varie b_0, b_1 para reduzir $fc(b_0, b_1)$ até encontrar mínimo (local ou global) da função de custo

Gradiente Descendente: Algoritmo

Algoritmo:

Repita até convergir {

$$temp_0 = b_0 - \alpha \frac{\partial}{\partial b_0} fc(b_0, b_1)$$

$$temp_1 = b_1 - \alpha \frac{\partial}{\partial b_1} fc(b_0, b_1)$$

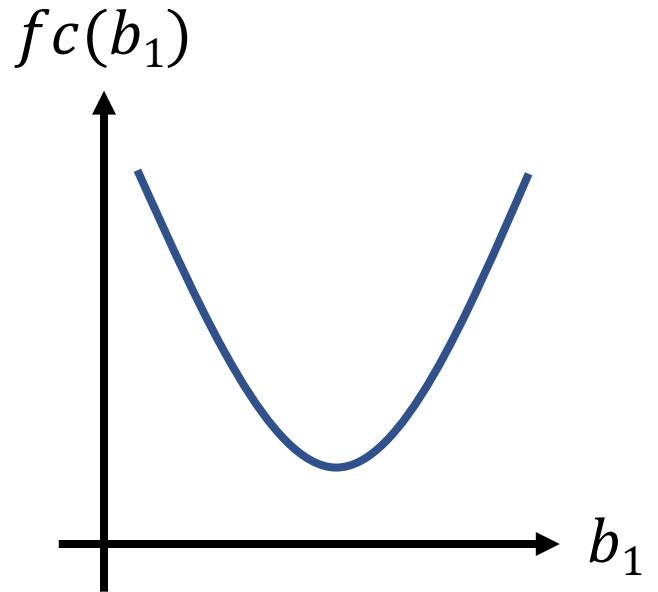
$$b_0 = temp_0$$

$$b_1 = temp_1$$

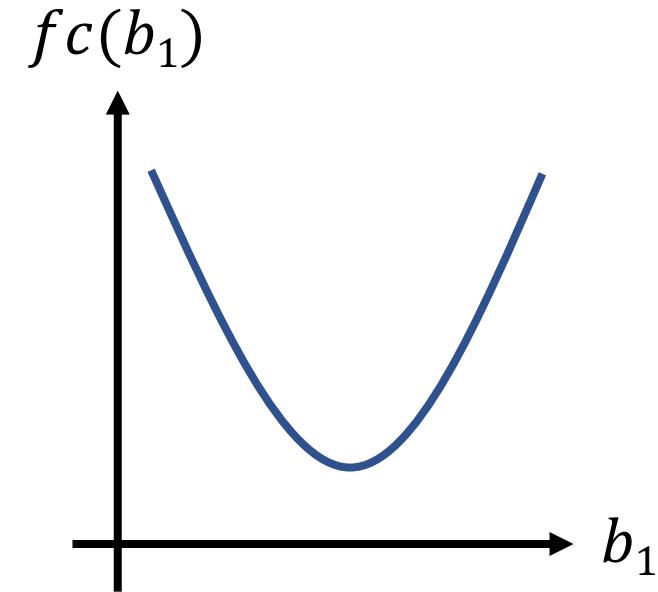
}

learning rate





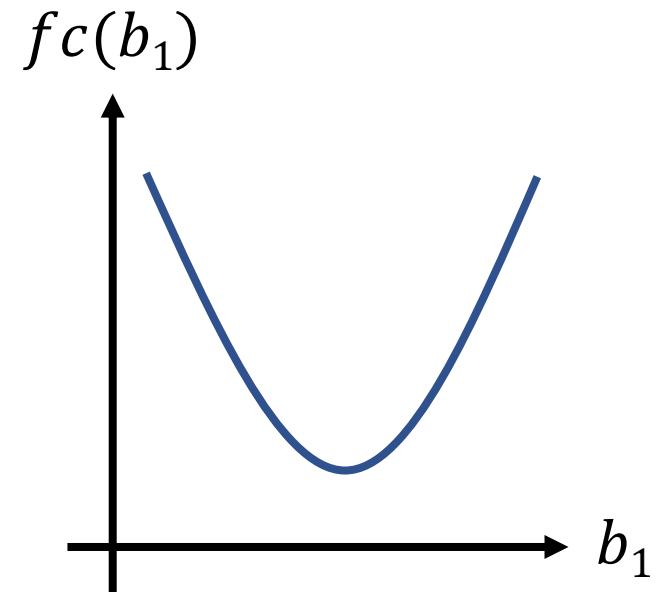
$$b_1 = b_1 - \alpha \boxed{\frac{\partial}{\partial b_1} f_c(b_1)} \geq 0$$



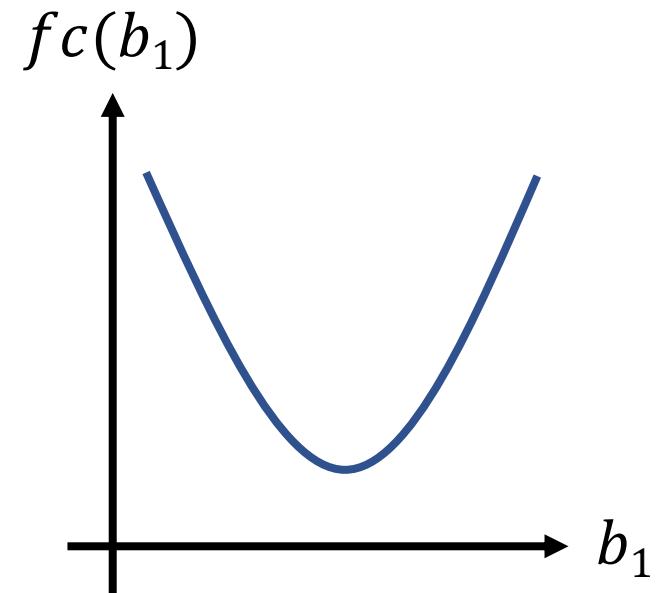
$$b_1 = b_1 - \alpha \boxed{\frac{\partial}{\partial b_1} f_c(b_1)} \leq 0$$

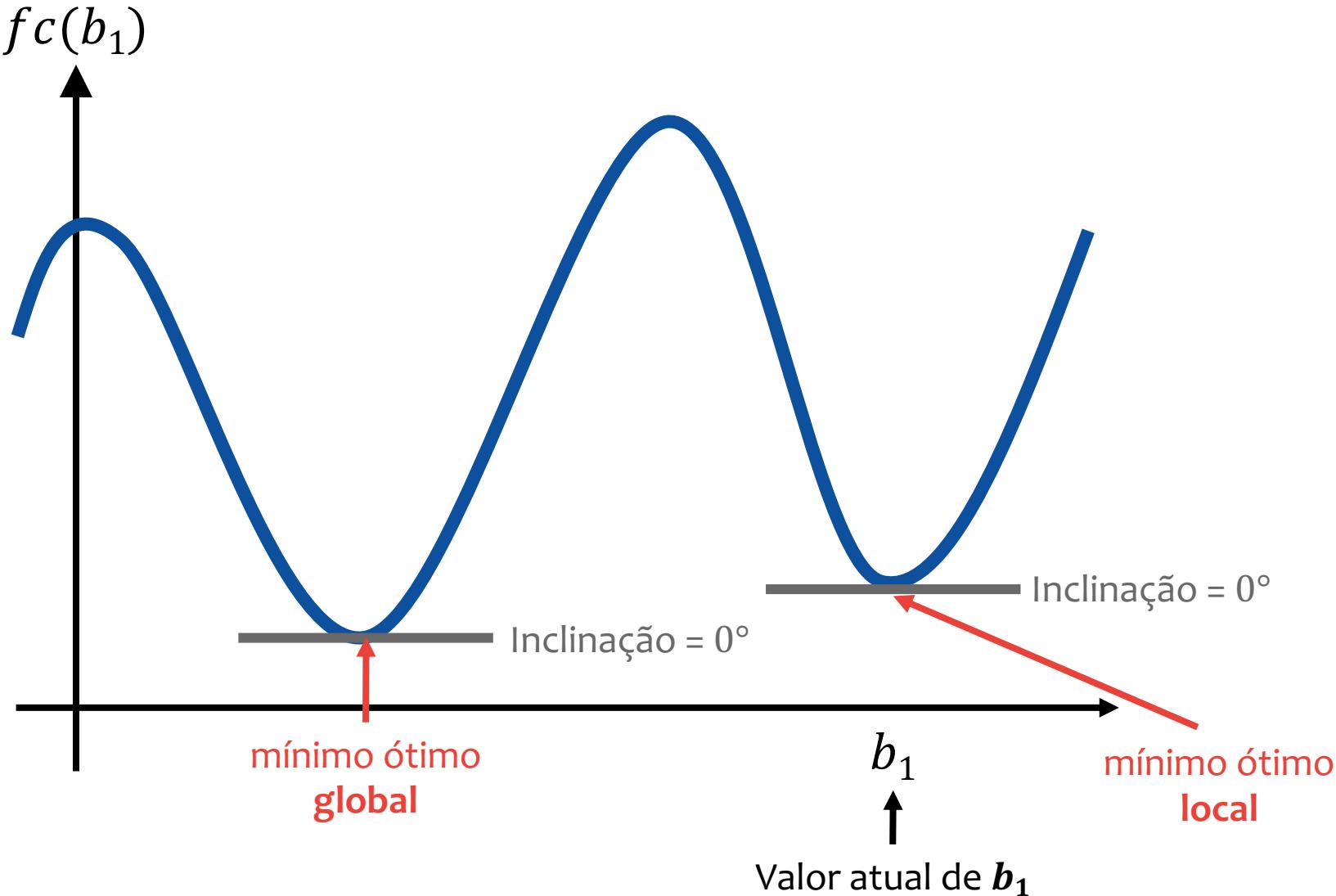
$$b_1 = b_1 - \alpha \frac{\partial}{\partial b_1} fc(b_1)$$

Se a é **muito pequeno**, o gradiente descendente pode ser lento!



Se a é **muito grande**, o gradiente descendente pode ser “chutado” do mínimo → a convergência pode falhar.



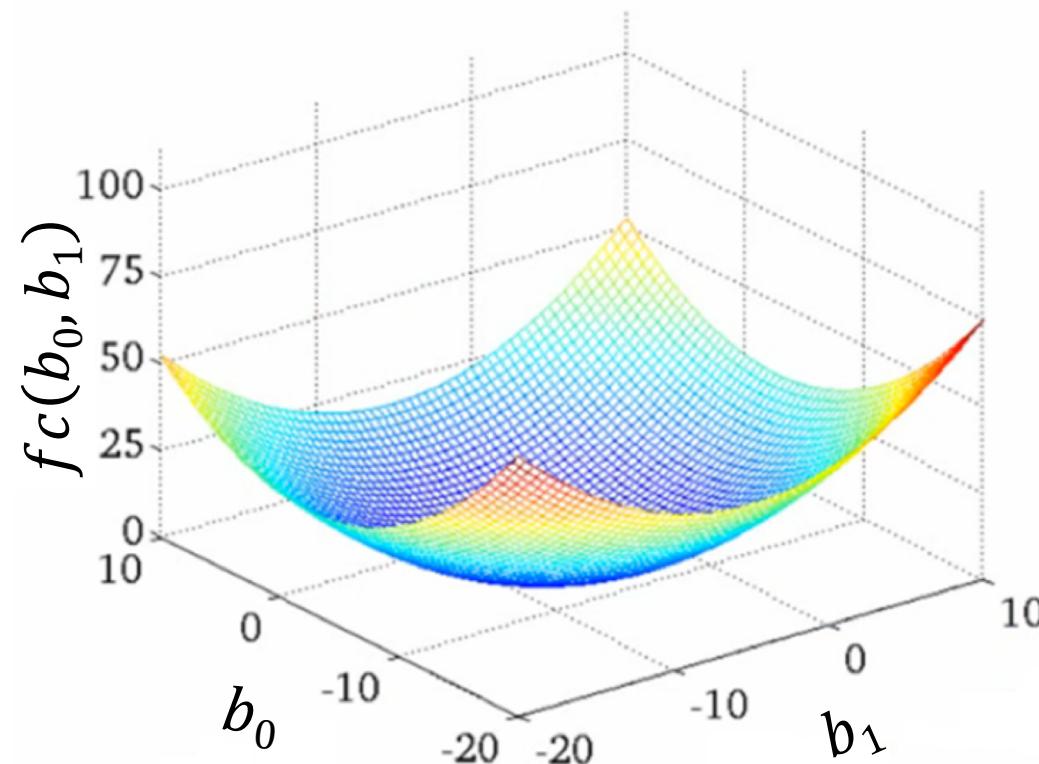


$$b_1 = b_1 - \alpha \frac{\partial}{\partial b_1} f_c(b_1)$$

$$\begin{aligned}\operatorname{tg}(0^\circ) &= 0 \\ \frac{\partial}{\partial b_1} f_c(b_1) &= 0\end{aligned}$$

Gradiente Descendente para Regressão Linear

- Embora o algoritmo do **Gradiente descendente** possa convergir para **um mínimo ótimo local**, isso **não** será problema para a **regressão linear**, uma vez que estamos utilizando uma **função convexa** como **função de custo**



D1EAD – Análise Estatística para Ciência de Dados

2021.1



Regressão Linear Simples

Prof. Ricardo Sovat

sovat@ifsp.edu.br

Prof. Samuel Martins (Samuka)

samuel.martins@ifsp.edu.br

