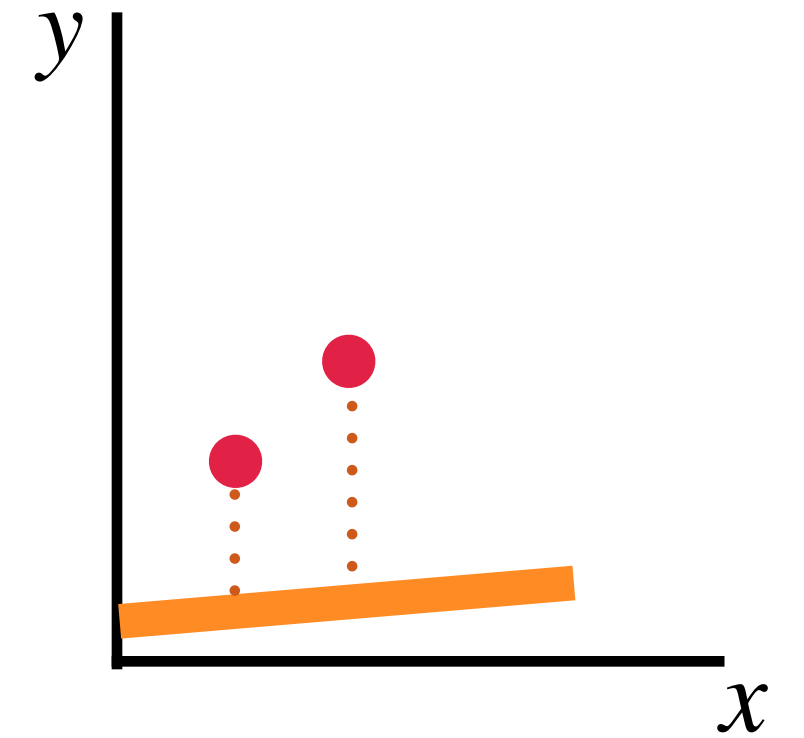


Example: Very Simple Linear Regression

Gradient:

$$\begin{bmatrix} \frac{\partial RSS}{\partial b_0} \\ \frac{\partial RSS}{\partial b_1} \end{bmatrix} = \begin{bmatrix} -2 \sum_i^N (y_i - (b_0 + b_1 x_i)) \\ -2 \sum_i^N x_i (y_i - (b_0 + b_1 x_i)) \end{bmatrix}$$



Apply the changes (learning rate 0.01):

$$\begin{bmatrix} b_{0_{new}} \\ b_{1_{new}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - 0.01 \begin{bmatrix} -10 \\ -16 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} b_{0_{new}} \\ b_{1_{new}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + 0.01 \begin{bmatrix} 10 \\ 16 \end{bmatrix} = \begin{bmatrix} 0.10 \\ 0.16 \end{bmatrix}$$

Example: Very Simple Linear Regression

Gradient:

$$\begin{bmatrix} \frac{\partial RSS}{\partial b_0} \\ \frac{\partial RSS}{\partial b_1} \end{bmatrix} = \begin{bmatrix} -2 \sum_i^N (y_i - (b_0 + b_1 x_i)) \\ -2 \sum_i^N x_i (y_i - (b_0 + b_1 x_i)) \end{bmatrix}$$

Apply the changes:

$$\begin{bmatrix} b_{0_{new}} \\ b_{1_{new}} \end{bmatrix} = \begin{bmatrix} 0.10 \\ 0.16 \end{bmatrix}$$

Compute loss function value:

$$RSS = \sum_i^N (y_i - b_0 - b_1 x_i)^2 = (2 - 0.10 - 0.16 \cdot 1)^2 + (3 - 0.10 - 0.16 \cdot 3)^2 = 8.88$$

