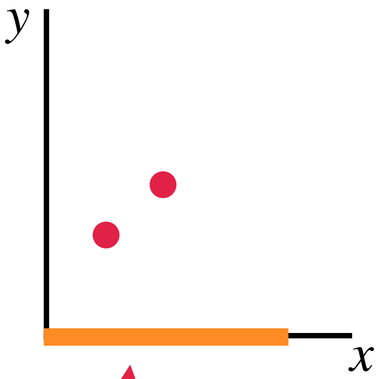
Example: Very Simple Linear Regression

Gradient:

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

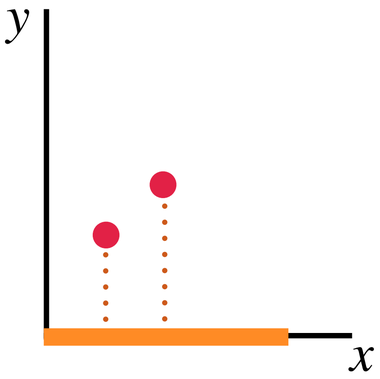
Initialize the gradient algorithm at **(0,0)**

$$\implies \begin{bmatrix} -2\sum_i^N (y_i - (0 + 0x_i)) \\ -2\sum_i^N x_i(y_i - (0 + 0x_i)) \end{bmatrix} = \begin{bmatrix} -2\sum_i^N (y_i) \\ -2\sum_i^N x_i(y_i) \end{bmatrix}$$



plug in our actual values

Example: Very Simple Linear Regression



Gradient:

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

Initialize the gradient algorithm at (0,0)

$$\Longrightarrow \begin{bmatrix} -2\sum_i^N (y_i) \\ -2\sum_i^N x_i(y_i) \end{bmatrix} = \begin{bmatrix} -2(2+3) \\ -2(1\cdot 2 + 2\cdot 3 \end{bmatrix} = \begin{bmatrix} -10 \\ -16 \end{bmatrix}$$

These are the changes we need to make to intercept and slope in order to reduce our loss function.