Initial Given Values

$m = -1$
$b = 1$
Learning rate $= 0.1$
data point $= (1,3)$ and $(3,6)$
$n = 2$
formula $\Rightarrow y = mx + b$

$\hat{y}_1 = m \cdot x_1 + b = (-1) \cdot 1 + 1 = 0$
$\hat{y}_2 = m x_2 + b = (-1) \cdot 3 + 1 = -2$

finding errors
Error $\Rightarrow y - \hat{y}$

Error 1 $= y_1 - \hat{y}_1 = 3 - 0 = 3$
Error 2 $= y_2 - \hat{y}_2 = 6 - (-2) = 8$

Updating m and b

Updating m

$$\frac{dJ}{dm} = -\frac{2}{n} \sum (y_i - \hat{y}_i) x_i = -\frac{2}{2} \sum$$

$$= -\frac{\cancel{2}}{\cancel{2}} \left[(3-0) \cdot 1 + 6 - (-2) \cdot 3 \right]$$

$$= -1 \left[3 + 24 \right]$$

$$= -27$$

$$\frac{dJ}{db} = -\frac{2}{n} \sum y_i - \hat{y}_i = -\frac{\cancel{2}}{\cancel{2}} \sum (3-0) + 6 - (-2)$$

$$\frac{dJ}{db} = -1 \left[3 + 8 \right] = -11$$

updated m & b

$$m_{new} = m_{old} - \alpha \frac{dJ}{dm} = -1 - 0.1 \cdot (-27) = -1 + 2.7$$
$$= 1.7$$

$$b_{new} = b_{old} - \alpha \frac{dJ}{db} = 1 - 0.1 \cdot (-11) = 1 + 1.1$$
$$= 2.1$$

$$MSE = \frac{1}{n} \sum (y_i - \hat{y})^2 = \frac{1}{2} \sum 3^2 + 8^2 = \frac{1}{2}(9 + 64) = \frac{73}{2} = 36.5$$

$y = mx + b$

updated $m = 1.7$

updated $b = 2.1$

Error $= y - \hat{y}$

Updating m and b

$$\frac{\partial J}{\partial m} = -\frac{2}{n} \sum (y_i - \hat{y}_i) x_i$$

$$\frac{\partial J}{\partial b} = -\frac{2}{n} \sum (y_i - \hat{y}_i)$$

Iteration 2

New predicted values:

$\hat{y}_1 = 1.7 \cdot 1 + 2.1 = 3.8$, $\hat{y}_2 = 1.7 \cdot 3 + 2.1 = 7.2$

Errors:

$e_1 = 3 - 3.8 = -0.8$, $e_2 = 6 - 7.2 = -1.2$

Gradients:

$$\frac{\partial J}{\partial m} = -\frac{2}{2}\left[1(-0.8) + 3(-1.2)\right] =$$

$$-1 \cdot (-0.8 - 3.6) = 4.4$$

$$\frac{\partial J}{\partial b} = -\frac{2}{2}(-0.8 - 1.2) = -1.(-2) = 2$$

update parameters :

$m = 1.7 - 0.1 . 4.4 = 1.7 - 0.44 = 1.26$

$b = 2.1 - 0.1 . 2 = 2.1 - 0.2 = 1.9$

$$MSE = \frac{1}{n} \sum y_i - \hat{y}$$

$$= \frac{1}{2} \sum ((-0.8)^2 + ((-1.2)^2)$$

$$= \frac{1}{2}(0.64 + 1.44) = \frac{2.08}{2} = 1.04$$

$y = mx + b$

$m = 1.26$

$b = 0.42$

learning rate $= 0.1$

$\text{Error} = y - \hat{y}$

Gradients (m & b) changed

$$\frac{dJ}{dm} = -\frac{2}{m} \sum (y_i - \hat{y}_i) x_i$$

$$\frac{dJ}{db} = -\frac{2}{m} \sum y_i - \hat{y}_i$$

$$\text{Cost or MSE} = \frac{1}{2} \sum y_i - \hat{y}$$

predicted values

$\hat{y}_1 = 1.26 + 1.9 = 3.16$

$\hat{y}_2 = 1.26 \cdot 3 + 1.9 = 5.58$

Errors

$\text{error}_1 = 3 - 3.16 = 0.16$

$\text{error}_2 = 6 - 5.58 = 0.42$

Gradients

$$\frac{dJ}{dm} = -1 \cdot [1(-0.16) + 3(0.42) = -1 \cdot (-0.16 + 1.26) =]$$

$$= -1.10$$

$$\frac{dJ}{db} = -1 \cdot (-0.16 + 0.42) = -1.026 = -0.26$$

Updated m & b

$m = 1.26 - 0.1 \cdot (-1.10) = 1.26 + 0.11 = 1.37$

$b = 1.9 - 0.1 \cdot (-0.26) = 1.9 + 0.026 = 1.926$

$$\text{MSE} = \frac{1}{2} \sum (-0.16)^2 + (0.42)^2$$

$$= \frac{1}{2} [0.0256 + 0.1764]$$

$$= 0.101$$

~~initial~~

iteration	m	b	$\hat{y}_1$	$\hat{y}_2$	Errors	Squared Errors	Cost (MSE)
0	-1	1	0	-2	8	73	36.5
1	1.7	2.1	3.8	7.2	-1.2	2.08	1.04
2	1.26	1.9	3.16	5.58	0.42	0.202	0.101

Our observation

We observed that ~~our~~ after updating our m & b are moving toward reducing the error because our Cost function are decreasing where we started with 36.5, then 1.04 and final 0.101.