Gradient descent

Initial values

$m = -1 \quad , \quad b = 1 \quad , \quad \alpha = 0.1$

x	1	3
y	3	6

let's find $\hat{y}$ using m and b

$\hat{y}_1 = mx + b$

$= (-1)(3) + 1 = -2$

expected y	predicted y
3	0
6	-2

let's find the new m and b

$$M_{new} = M_{old} - \alpha \frac{\partial J}{\partial m}$$

$$\frac{\partial J}{\partial m} = \frac{-2}{n} \sum_{i=1}^{2} (y_i - \hat{y}_i) x_i$$

$$\frac{\partial J}{\partial m} = \frac{-2}{2} \left((3-0)(1) + (6-(-2))(3) \right)$$

$$= -1 \ (3 + 24)$$

$$= -1 \ (27) = \underline{\underline{-27}}$$

$$M_{new} = -1 - 0.1\,(-27)$$

$$= \underline{\underline{1.7}}$$

$$b_{new} = b_{old} - \alpha \frac{GJ}{Gb}$$

$$\frac{GJ}{Gb} = \frac{-2}{n} \sum_{i=1}^{2} (y_i - \hat{y}_i)$$

$$= \frac{-2}{2} \left((3-0) + 6 - (-2) \right)$$

$$= -1\,(11) = \underline{\underline{-11}}$$

$$b_{new} = 1 - 0.1\,(-11) = \underline{\underline{2.1}}$$

Date

let's find $\hat{y}$ using the new b and m

$\hat{y}_1 = (1.7)(1) + 2.1 = 3.8$

$y_2 = (1.7)(3) + 2.1 = 7.2$

Expected y	Predicted y
3	3.8
6	7.2

Iteration 2:

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

$$= \frac{-2}{2} \left((3-3.8)(1) + (6-7.2)(3) \right)$$

$$= -1 \left(-0.8 - 3.6 \right) = (-1)(-4.4)$$

$$= \underline{\underline{4.4}}$$

$$M_{new} = M_{old} - \alpha \frac{\partial J}{\partial m}$$

$$= 1.7 - 0.1(4.4) = \underline{\underline{1.26}}$$

Date

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

$$= \frac{-2}{2} \left((3 - 3.8) + (6 - 7.2) \right)$$

$$= -1 (-0.8 - 1.2) = \underline{2}$$

$$b_{new} = b_{old} - \alpha \frac{\partial J}{\partial b}$$

$$= 2.1 - 0.1 (2.0) = \underline{1.9}$$

let's find $\hat{y}$ using the new b and m

$$\hat{y}_1 = (1.26)(1) + 1.9 = 3.16$$

$$\hat{y}_2 = (1.26)(3) + 1.9 = 5.68$$

expected y	predicted y
3	3.16
6	5.68

iteration 3 :

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

$$= \frac{-2}{2} \left((3 - 3.16)(1) + (6 - 5.68)(3) \right)$$

$$= -1 (-0.16 + 0.96) = \underline{\underline{-0.8}}$$

$$M_{new} = 1.26 - 0.1 (-0.8) = \underline{\underline{1.34}}$$

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

$$= -1 \left((3 - 3.16) + (6 - 5.68) \right)$$

$$= -1 (-0.16 + 0.32) = \underline{\underline{-0.16}}$$

$$b_{new} = b_{old} - \alpha \frac{\partial J}{\partial b}$$

$$= 1.9 - 0.1 (-0.16)$$

$$= 1.916$$

$\hat{y}_1 = 1.34(1) + 1.916 = 3.256$

$\hat{y}_2 = (1.34)(3) + 1.916 = 5.936$

expected y	predicted y
3	3.256
6	5.936

<u>iteration 4 :</u>

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

$$= -1 \left((3 - 3.256)(1) + (6 - 5.936)(3) \right)$$

$$= -1(-0.256 + 0.192)$$

$$= \underline{\underline{0.064}}$$

$$M_{new} = M_{old} - \alpha \frac{\partial J}{\partial m}$$

$$= 1.34 - 0.1(0.064) = \underline{\underline{1.3336}}$$

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

$$= -1\left((3 - 3.256) + (6 - 5.936)\right)$$

$$= -1(-0.256 + 0.0064)$$

$$= \underline{\underline{0.192}}$$

$$b_{new} = b_{old} - \alpha \frac{\partial J}{\partial b}$$

$$= 1.916 - 0.1(0.192)$$

$$= \underline{\underline{1.8968}}$$

Let's find $\hat{y}$ using the new m and b

$$\hat{y}_1 = (1.336)(1) + 1.8968$$

$$= 3.2304$$

$$\hat{y}_2 = (1.3336)(3) + 1.8968$$

$$= 5.8976$$

Scanned with CamScanner

expected y	predicted y
3	3.2304
6	5.8976

Conclusion

iteration	m	b
1	1.7	2.1
2	1.26	1.9
3	1.34	1.916
4	1.3336	1.8968

in iteration 1, there is a large jump in both m and b because the initial error was high. In the rest iteration the change in m and b are very small because the prediction improved and the error decreased.
So yes m and b are moving towards value that reduce MSE.