Part 3 : Gradient Descent Manual Calculations

X	Y
1	3
3	6

$m_1 = -1$

$b_1 = 1$

$\alpha = 0.1$

$\hat{y}_1 = \cancel{hw} \; mx + b$

$$MSE = \frac{\sum (y - \hat{y})^2}{n}$$

$$m_{new} = m_{old} - \alpha \frac{d(MSE)}{\partial m}$$

$$b_{new} = b_{old} - \alpha \frac{\partial(MSE)}{\partial b}$$

Step 1 : derive MSE with respect to m and b

* With m

$$MSE = \frac{1}{n} \sum (y - \hat{y})^2$$

$$= \frac{1}{n} \sum (y - (mx + b))^2$$

Let $u = y - (mx + b)$

$$\frac{\partial u}{\partial m} = -x$$

So $u^2 = 2u$

hence $\frac{\partial MSE}{\partial m} = \frac{1}{n} \sum 2u . (-x)$

$$= \frac{1}{n} \sum 2(y - (mx + b)).(-x)$$

$$= \frac{2}{n} \sum ((mx + b) - y_i)(x)$$

$$\boxed{\frac{\partial MSE}{\partial m} = \frac{2}{n} \sum ((mx+b) - y_i)(x)}$$

With b

$$MSE = \frac{1}{n} \sum (y - \hat{y})^2$$

$$= \frac{1}{n} \sum (y - (mx + b))^2$$

Let $u = y - (mx + b)$

$$\frac{\partial u}{\partial b} = -1$$

So $u^2 = 2u$

$$\frac{\partial MSE}{\partial b} = \frac{1}{n} \sum 2u . (-1)$$

$$= \frac{1}{n} \sum 2 (y_i - (mx + b)).(-1)$$

$$= \frac{1}{n} \sum 2 ((mx + b) - y_i)(1)$$

$$\boxed{\frac{\partial MSE}{\partial b} = \frac{2}{n} \sum ((mx + b) - y_i)}$$

$$m_{new} = m_{old} - \alpha \left(\frac{2}{n} \sum (mx + b) - y_i \right)(x)$$

$$b_{new} = b_{old} - \alpha \left(\frac{2}{n} \sum ((mx + b) - y_i) \right)$$

Mike iteration :

Iteration 1

$\hat{y}_1 = -1(1) + 1 = 0$

$\hat{y}_2 = -1(3) + 1 = -2$

$$MSE = \frac{((3-0)^2 + (6-(-2))^2)}{2}$$

$$= \frac{9+64}{2} = \frac{73}{2}$$

$$= 36.5$$

Update values of m and b

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

$$= -3 + (-24) = -27$$

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

$$= -3 + (-8) = -11$$

$m_{new} = -1 - (0.1)(-27) = 1.7$

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

Iteration 2 by Lorita

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

$\hat{y}_2 = 1.7(3) + 2.1 = 7.2$

$$MSE = \frac{((3-3.8)+(6-7.2))^2}{2}$$

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

$$= \frac{0.64 + 1.44}{2} = 1.04$$

Update values m and b

$$\frac{\partial MSE}{\partial m} = \frac{2}{2}((3.8-3)(1) + (7.2-6)(3))$$

$$= (0.8 \times 1) + (1.2 \times 3) = 4.4$$

$$\frac{\partial MSE}{\partial b} = \frac{2}{2}((3.8-3) + (7.2-6))$$

$$= 0.8 + 1.2 = 2$$

$$m_{new} = 1.7 - (0.1)(4.4) = 1.26$$

$$b_{new} = 2.1 - (0.1)(2) = 1.9$$

Miller Keza [illegible]

Iteration 3

Reference formulas

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

$$MSE = \frac{\sum (y - \hat{y})^2}{n}$$

$$m_{new} = m_{old} - \alpha \frac{dMSE}{dm}$$

$$b_{new} = b_{old} - \alpha \frac{dMSE}{db}$$

$$\frac{dMSE}{dm} = \frac{2}{n} \sum \left((mx + b) - y\right)(x)$$

$$\frac{dMSE}{db} = \frac{2}{n} \sum \left((mx + b) - y\right)$$

$$m_{new} = m_{old} - \alpha \left(\frac{2}{n} \sum (mx + b) - y\right)(x)$$

$$b_{new} = b_{old} - \alpha \left(\frac{2}{n} \sum (mx + b) - y\right)$$

Iteration 3

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

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

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

$$= \frac{0.0256 + 0.1024}{2}$$

$$= 0.064$$

Update values of m and b

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

$$= 0.16 + (-0.96) = -0.8$$

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

$$= 0.16 + (-0.32) = -0.16$$

$$m_{new} = 1.26 - (0.1)(-0.8)$$

$$= 1.26 + 0.08 = \underline{1.34}$$

$$b_{new} = 1.9 - (0.1)(-0.16)$$

$$= 1.9 + 0.016 = \underline{\underline{1.916}}$$

(By Nick - lemy)

iteration 4

$\hat{y}_1 = 1.34(1) + 1.916 = \cancel{3.10}\ 3.256$

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

$$MSE = \frac{(3-3.256)^{2} + (6-5.936)^2}{2}$$

$$= \frac{0.065536 + 0.004096}{2}$$

$= 0.034816$

iteration 5

$\hat{y}_1 = 1.3336(1) + 1.8968 = 3.2304$

$\hat{y}_2 = 1.3336(3) + 1.8968 = 5.8976$

$$MSE = \frac{(3-3.2304)^2 + (6-5.8976)^2}{2}$$

$$= \frac{0.053084 + 0.010449}{2}$$

$= 0.031785$

update values m and b

$$\frac{\partial MSE}{\partial m} = \frac{\cancel{2}}{\cancel{2}}(3.256-3)(1) + (5.936)(3)$$

$= 0.256 + (-0.192) = 0.064$

$$\frac{\partial MSE}{\partial m} = \frac{\cancel{2}}{\cancel{2}}(3.256-3) + (5.936-6)$$

$= 0.256 - 0.064$

$= 0.192$

$m_{new} = 1.34 - (0.1)(0.064)$

$= \underline{1.3336}$

$b_{new} = 1.916 - (0.1)(0.192) = \underline{1.8968}$

Conclusion

- MSE trend = After each Iteration, the values of MSE (the error decreases) which indicates that our m and b are being updated well to have a perfect fit.

- trend for m and b = Both m and b are being updated after each iteration to reduce the error. And that is shown by how the error is reduced after each iteration