

Part 3: Gradient Descent Manual Calculations

X	Y	$m_1 = -1$	$\hat{y}_1 = mx + b$
1	3	$b_1 = 1$	$MSE = \frac{\sum (y - \hat{y})^2}{n}$
3	6	$d = 0.1$	$m_{new} = m_{old} - d \frac{\partial (MSE)}{\partial m}$
			$b_{new} = b_{old} - d \frac{\partial (MSE)}{\partial b}$

Step 1: derive MSE with respect to m and b

* With m

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

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

$$\text{Let } u = y - (mx + b)$$

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

$$\text{So } u^2 = 2u$$

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

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

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

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

With b

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

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

$$\text{Let } u = y - (mx + b)$$

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

$$\text{So } u^2 = 2u$$

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

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

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

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

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

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

Step 2: Iterations

Iteration 1

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

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

$$MSE = \frac{((3-0)^2 + (6-(-2))^2)}{2} = \frac{9 + 64}{2} = \frac{73}{2}$$

$$= 36.5$$

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

$$= 36.5$$

Iteration 2

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

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

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

$$= \frac{((-0.8)^2 + (-1.2)^2)}{2} = \frac{0.64 + 1.44}{2} = 1.04$$

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

Iteration 3

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

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

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

$$= \frac{((-0.16)^2 + (0.32)^2)}{2}$$

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

$$= 0.064$$

Update values of m and b

$$m = -1 - (0.1)(-)$$

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

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

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

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

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

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

Update values m and b

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

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

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

$$= 0.8 + 1.2 = 2$$

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

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

Update values m and b

$$\frac{\partial MSE}{\partial m} = \frac{2}{2} ((3.16-3) \times 1) + ((5.68-6) \times 3)$$

$$= 0.16 + (-0.86) = -0.8$$

$$\frac{\partial MSE}{\partial b} = \frac{2}{2} ((3.16-3) \times 1) + ((5.68-6) \times 3)$$

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

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

$$= 1.26 + 0.08 = 1.34$$

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

$$= 1.9 + 0.016 = 1.916$$

Iteration 4

$$\hat{y}_1 = 1.34(1) + 1.816 = 3.256$$

$$\hat{y}_2 = 1.34(3) + 1.816 = 5.836$$

$$\begin{aligned} \text{MSE} &= \frac{(3 - 3.256)^2}{2} + \frac{(6 - 5.836)^2}{2} \\ &= \frac{0.065536}{2} + \frac{0.004096}{2} \\ &= 0.034816 \end{aligned}$$

Iteration 5

$$\hat{y}_1 = 1.3336(1) + 1.8868 = 3.2304$$

$$\hat{y}_2 = 1.3336(3) + 1.8868 = 5.8976$$

$$\begin{aligned} \text{MSE} &= \frac{(3 - 3.2304)^2}{2} + \frac{(6 - 5.8976)^2}{2} \\ &= \frac{0.053084}{2} + \frac{0.010448}{2} \\ &= 0.031785 \end{aligned}$$

Update values m and b

$$\begin{aligned} \frac{\partial \text{MSE}}{\partial m} &= \frac{1}{2} ((3.256 - 3)(1) + (5.836 - 6)(3)) \\ &= 0.256 + (-0.182) = 0.064 \end{aligned}$$

$$\begin{aligned} \frac{\partial \text{MSE}}{\partial m} &= \left(\frac{1}{2} (3.256 - 3) + (5.836 - 6) \right) \\ &= 0.256 - 0.064 \\ &= 0.182 \end{aligned}$$

$$\begin{aligned} m_{\text{new}} &= 1.34 - (0.1)(0.064) \\ &= \underline{1.3336} \end{aligned}$$

$$\begin{aligned} b_{\text{new}} &= 1.816 - (0.1)(0.182) \\ &= \underline{1.8868} \end{aligned}$$

Conclusion :

- MSE Trend : After each iteration, the values of MSE decreases which indicates that our gradient descent is ~~update~~ updating m and b well.

- m and b : both m and b are reducing the MSE after each iteration when we update them.