Problem Solving Related to Gradient Descent Algorithm for Finding the Regression Coefficients:

Question:

Consider the following sample dataset

X	Y
2	3
3	6
4	7

Estimate the regression coefficients b_0 and b_1 by using Gradient Descent Algorithm where learning rate $\alpha = 0.05$. Perform three iteration and compute Loss function (Mean Square Error (MSE)) at each iteration as well. Initialize regression coefficients with $(0,0)$

Solution:

$$Estimated\ Regression\ Line\ Equation:\ \ \hat{y} = b_0 + b_1 x$$

$$Loss\ Function = L = MSE = \frac{1}{2n}\sum_{1}^{n}(\hat{y} - y)^2 = \frac{1}{2n}\sum_{1}^{n}(b_0 + b_1 x - y)^2$$

$$L = \frac{1}{6}((b_0 + 2b_1 - 3)^2 + (b_0 + 3b_1 - 6)^2 + (b_0 + 4b_1 - 7)^2)$$

$$b_0^{new} = b_0^{old} - \alpha\frac{\partial L}{\partial b_0} \qquad ------equation1$$

$$b_1^{new} = b_1^{old} - \alpha\frac{\partial L}{\partial b_1} \qquad ------equation2$$

$$\frac{\partial L}{\partial b_0} = \frac{1}{n}\sum_{1}^{n}(b_0 + b_1 x - y) = \frac{1}{3}(b_0 + 2b_1 - 3 + b_0 + 3b_1 - 6 + b_0 + 4b_1 - 7)$$

$$\frac{\partial L}{\partial b_0} = \frac{1}{3}(3b_0 + 9b_1 - 16) \quad ------equation3$$

$$\frac{\partial L}{\partial b_1} = \frac{1}{n}\sum_{1}^{n}(b_0 + b_1 x - y)x = \frac{1}{3}(2(b_0 + 2b_1 - 3) + 3(b_0 + 3b_1 - 6) + 4(b_0 + 4b_1 - 7))$$

$$\frac{\partial L}{\partial b_1} = \frac{1}{3}(2(b_0 + 2b_1 - 3) + 3(b_0 + 3b_1 - 6) + 4(b_0 + 4b_1 - 7))$$

$$\frac{\partial L}{\partial b_1} = \frac{1}{3}(9b_0 + 29b_1 - 52) \quad ------equation4$$

Put values of equation 3 and 4 in equation 1 and 2,

$$b_0^{new} = b_0^{old} - 0.05\frac{1}{3}\left(3b_0^{old} + 9b_1^{old} - 16\right)$$

$$b_1^{new} = b_1^{old} - 0.05\ \tfrac{1}{3}\left(9b_0^{old} + 29b_1^{old} - 52\right)$$

Iteration 1:

$$b_0^{(1)} = b_0^{(0)} - 0.05\frac{1}{3}\left(3b_0^{(0)} + 9b_1^{(0)} - 16\right) = 0 - 0.05\frac{1}{3}(3(0) + 9(0) - 16) = 0.26667$$

$$b_1^{(1)} = b_1^{(0)} - 0.05\ \tfrac{1}{3}\left(9b_0^{(0)} + 29b_1^{(0)} - 52\right) = 0 - 0.05\ \tfrac{1}{3}(9(0) + 29(0) - 52) = 0.86667$$

$$L = \frac{1}{6}((0.26667 + 2 * 0.86667 - 3)^2 + (0.26667 + 3 * 0.86667 - 6)^2 + (0.26667 + 4 * 0.86667 - 7)^2)$$

$$L = 3.58145$$

Iteration 2:

$$b_0^{(2)} = b_0^{(1)} - 0.05\frac{1}{3}\left(3b_0^{(1)} + 9b_1^{(1)} - 16\right) = 0.26667 - 0.05\frac{1}{3}(3(0.26667) + 9(0.86667) - 16) = 0.39$$

$$b_1^{(2)} = b_1^{(1)} - 0.05\ \tfrac{1}{3}\left(9b_0^{(1)} + 29b_1^{(1)} - 52\right)$$

$$= 0.86667 - 0.05\ \frac{1}{3}(9(0.26667) + 29(0.86667) - 52) = 1.27445$$

$$L = \frac{1}{6}((0.39 + 2 * 1.27445 - 3)^2 + (0.39 + 3 * 1.27445 - 6)^2 + (0.39 + 4 * 1.27445 - 7)^2)$$

$$L = 0.91377$$

Iteration 3:

$$b_0^{(3)} = b_0^{(2)} - 0.05\frac{1}{3}\left(3b_0^{(2)} + 9b_1^{(2)} - 16\right) = 0.39 - 0.05\frac{1}{3}(3(0.39) + 9(1.27445) - 16) = 0.446$$

$$b_1^{(3)} = b_1^{(2)} - 0.05\ \tfrac{1}{3}\left(9b_0^{(2)} + 29b_1^{(2)} - 52\right)$$

$$= 1.27445 - 0.05\ \frac{1}{3}(9(0.39) + 29(1.27445) - 52) = 1.46663$$

$$L = \frac{1}{6}((0.446 + 2 * 1.46663 - 3)^2 + (0.446 + 3 * 1.46663 - 6)^2 + (0.446 + 4 * 1.46663 - 7)^2)$$

$$L = 0.32474$$