

Part 3: Gradient Descent Manual Calculation

Data given

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

Data points: (1,3) and (3,6)

Number of data points (n) = 2

Iteration 1

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

Predicted values of \hat{y}

$$\text{For (1,3): } \hat{y}_1 = (-1)(1) + 1 = 0$$

$$\text{For (3,6): } \hat{y}_2 = (-1)(3) + 1 = -2$$

Gradients:

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

$$= -\frac{2}{2} [(3-0)(1) + (6-(-2))(3)]$$
$$= -[3+24] = -27$$

$$\frac{\Delta J}{\Delta b} = -\frac{2}{n} \sum (y_i - \hat{y}_i) = -\frac{2}{2} [(3-0) + (6-(-2))]$$
$$= -[3+8] = -11$$

$$m_{\text{new}} = m_{\text{old}} - \alpha \frac{\Delta J}{\Delta m} = -1 - 0.1(-27) = -1 + 2.7 = 1.7$$

$$b_{\text{new}} = b_{\text{old}} - \alpha \frac{\Delta J}{\Delta b} = 1 - 0.1(-11) = 1 + 1.1 = 2.1$$

$$\underline{m_{\text{new}} = 1.7}, \underline{b_{\text{new}} = 2.1}$$

2nd Iteration
From the first iteration:

$$m = 1.7$$

$$b = 2.1$$

Values of \hat{y} using values of m and b from 1st iteration.

$$\text{For } (1, 3); \hat{y}_1 = (1.7) \times 1 + 2.1 \Rightarrow 1.7 + 2.1 = 3.8$$

$$\text{For } (3, 6); \hat{y}_2 = (1.7) \times 3 + 2.1 \Rightarrow 5.1 + 2.1 = 7.2$$

Expected values of y	Predicted values of y
3	3.8
6	7.2

New values of m and b

$$m_{\text{new}} = m_{\text{old}} - \alpha \frac{\Delta J}{\Delta m}$$

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

$$\frac{\Delta J}{\Delta m} = -\frac{2}{2} [(3 - 3.8)(1) + (6 - 7.2)(3)]$$

$$\frac{\Delta J}{\Delta m} = -[(-0.8) + (-1.2)(3)]$$

$$\frac{\Delta J}{\Delta m} = -(-0.8 - 3.6)$$

$$= -(-4.4)$$

$$\frac{\Delta J}{\Delta m} = 4.4$$

$$m_{\text{new}} = 1.7 - (0.1 \times 4.4) \Rightarrow 1.7 - 0.44 = \underline{\underline{1.26}}$$

$$m_{\text{new}} = 1.26$$

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

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

$$\frac{\Delta J}{\Delta b} = -\frac{2}{2} [(3 - 3.8) + (6 - 7.2)]$$

$$= -[(-0.8) + (-1.2)]$$

$$= -(-2)$$

$$b_{\text{new}} = 2.1 - (0.1 \times 2) \Rightarrow 2.1 - 0.2$$

$$b_{\text{new}} = \underline{\underline{1.9}}$$

2nd Iteration

$$b_{\text{new}} = 1.9$$

$$m_{\text{new}} = 1.26$$

3rd iteration

using m & b from iteration 2

$$m = 1.26$$

$$b = 1.9$$

y-Predictions:

$$\text{for } (1, 3): \hat{y}_1 = (1.26 \times 1) + 1.9 = \underline{3.16}$$

$$\text{for } (3, 6): \hat{y}_2 = (1.26 \times 3) + 1.9 = \underline{5.68}$$

expected y	Predicted y
3	3.16
6	5.68

Finding m_{new} & b_{new}

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

$$m_{\text{new}} = 1.26 - \left[0.1 \times \left(-\frac{2}{n} \right) \left[(3 - 3.16)(1) + (6 - 5.68)(3) \right] \right]$$

$$m_{\text{new}} = 1.26 - \left[0.1 \left(-\frac{2}{4} \right) \left[(-0.16) + (0.96) \right] \right]$$

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

$$\underline{\underline{m_{\text{new}} = 1.34}}$$

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

$$b_{\text{new}} = 1.9 - \left[0.1 \times \left(-\frac{2}{n} \right) \left[(3 - 3.16) + (6 - 5.68) \right] \right]$$

$$b_{\text{new}} = 1.9 - \left[0.1 \times \left(-\frac{2}{4} \right) \left[(-0.16) + (0.32) \right] \right]$$

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

$$b_{\text{new}} = 1.9 + 0.016$$

$$\underline{\underline{b_{\text{new}} = 1.916}}$$

Iteration 4

1. Predicted values

$$\text{point } (1, 3) \rightarrow \hat{y} = 1.34 \times 1 + 1.916 = 3.256$$

$$(3, 6) \rightarrow \hat{y} = 1.34 \times 3 + 1.916 = 5.936$$

2. Gradients

$$\begin{aligned} \frac{\partial J}{\partial m} &= -\frac{2}{2} [(3 - 3.256)(1) + (6 - 5.936)(3)] \\ &= -[-0.256 + 0.192] = 0.064 \end{aligned}$$

$$\begin{aligned} \frac{\partial J}{\partial b} &= -\frac{2}{2} [(3 - 3.256) + (6 - 5.936)] \\ &= -[-0.256 + 0.064] = 0.192 \end{aligned}$$

To update parameters

$$\begin{aligned} m &= 1.34 - 0.1 \times 0.064 \\ &= 1.34 - 0.0064 = 1.3336 \end{aligned}$$

$$\begin{aligned} b &= 1.916 - 0.1 \times 0.192 \\ &= 1.916 - 0.0192 = 1.8968 \end{aligned}$$

Summary

Iteration Results	m	b
Initially	-1	1
1	1.2	2.1
2	1.26	1.9
3	1.34	1.916
4	1.3336	1.8968

Our observation

The parameters show an oscillating but converging pattern

1. The first update made large changes

m from -1 to 1.7

b " " 1 to 2.1

2. Subsequent updates made smaller adjustments (3rd ~~and~~)

3. By the u th iteration changes became very smaller

$$\Delta m \approx 0.0064, \Delta b \approx 0.0192$$

4. The values are stabilizing around $m \approx 1.33$, $b \approx 1.9$
suggesting convergence

Each iteration successfully reduced the error with the magnitude of updates decreasing as we approach optimal values.