Part 3: Gradient Descent Manual Calculation (4 updates)

Given information: - linear Model: $y = mx + b$

- Initial parameters: - Slope $m = -1$
  - Intercept $b = 1$
- learning rate: $\alpha = 0.1$
- Data points: • $(x_1, y_1) = (1, 3)$
  • $(x_2, y_2) = (3, 6)$
- N° of data points $n = 2$

Update 1: - Step 1: calculate predicted values $\hat{y}$

$$\hat{y}_1 = mx_1 + b = (-1)(1) + 1 = 0$$

$$\hat{y}_2 = mx_2 + b = (-1)(3) + 1 = -2$$

- Step 2: Calculate Gradients of cost function $J(m,b)$ using The Mean Squared Error (MSE) gradient formulas:

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

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

- so m = $(+3 - 0) \times 1 = 3$

$(-6 - (-2)) \times 3 = 8 \times 3 = 24$

$$\Rightarrow \frac{\partial J}{\partial m} = -\frac{2}{2}(3 + 24) = \boxed{-27}$$

- b = $(+3 - 0) = 3$

$(6 - (-2)) = 8$

$$\Rightarrow \frac{\partial J}{\partial b} = \frac{-2}{2}(3 + 8) = \boxed{-11}$$

- Step 3: update parameters

$$m_{new} = m_{old} - \alpha \frac{\partial J}{\partial m} = -1 - 0.1 \times (-27) = \boxed{1.7}$$

$$b_{new} = b_{old} - \alpha \frac{\partial J}{\partial b} = 1 - 0.1 \times (-11) = \boxed{2.1}$$

Update 2: Predicted values

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

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

Gradients

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

$$= -[-0.8 - 3.6] = 4.4$$

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

$$= -(-0.8 - 1.2) = 2.0$$

Update Parameters

$$m_{new} = 1.7 - 0.1 \times 4.4 = \boxed{1.26}$$

$$b_{new} = 2.1 - 0.1 \times 2.0 = \boxed{1.9}$$

Update 3

Predicted values $\hat{y}_1 = 1.26 \times 1 + 1.9 = 3.16$

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

Gradients $\frac{\partial J}{\partial m} = -\frac{2}{2}[(3 - 3.16) \times 1 + (6 - 5.68) \times 3] = -[-0.16 + 0.96] = -0.8$

$\frac{\partial J}{\partial b} = -\frac{2}{2}[(3 - 3.16) + (6 - 5.68)]$

$= -[-0.16 + 0.32] = -0.16$

Update Parameters

$m_{new} = 1.26 - 0.1 \times (-0.8) = 1.34$

$b_{new} = 1.9 - 0.1 \times (-0.16) = 1.916$

~~M new~~ $m_{new} = \boxed{1.34}$

$b_{new} = \boxed{1.916}$

BIHIBINDI AUBERT GLOIRE

AFRICA CEO FORUM
ANNUAL SUMMIT

Iteration 4 :

Current Values : $m = 1.84$

$b = 1.916$

Predicted Values

$$y_1 = 1.84 \times 1 + 1.916 = 3.256$$

$$y_2 = 1.84 \times 3 + 1.916 = 5.936$$

Gradients

$$\frac{\partial J}{\partial m} = -\frac{2}{2}\left[(3 - 3.256) \times 1 + (6 - 5.936) \times 3\right]$$

$$= -\left[-0.256 + 0.192\right] = 0.064$$

$$\frac{\partial J}{\partial b} = -\frac{2}{2}\left[(3 - 3.256) + (6 - 5.936)\right] = -(-0.256 + 0.064)$$

$$= 0.192$$

$$m_{new} = 1.84 - 0.1 \times 0.064 = \boxed{1.8336}$$

$$b_{new} = 1.916 - 0.1 \times 0.192 = \boxed{1.8968}$$

## Summary Table of parameter Updates

| Iteration | Slope m | Intercept b |
|---|---|---|
| 0 (initial) | -1.0000 | 1.0000 |
| 1 | 1.7000 | 2.1000 |
| 2 | 1.2600 | 1.9000 |
| 3 | 1.3400 | 1.9160 |
| 4 | 1.3336 | 1.8968 |

### Trend Analysis

- Both slope m and intercept b are Converging to values that better fit the data points.
- Gradients are reducing in magnitude with each update, showing the algorithm is moving towards Convergence
- The error between Predicted and actual values is decreasing
- The model parameters are moving closer to the optimal linear relationship $y = 1.5x + 1.5$
- This confirms that gradient descent is successful, minimizing the cost function and learning the data pattern