Denis Mitali

# Part 3: Gradient Descent Manual Calculation

Given linear equation: $y = mx + b$

where

Initial $m = -1$ Learning rate $\alpha = 0.1$

Initial $b = 1$ Given points = $(1,3)$ and $(3,6)$

$\hat{y}_i = mx_i + b$

for point $(1,3)$,

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

$y_i - \hat{y}_i = 3 - 0 = 3$

for point $(3,6)$

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

$y_i - \hat{y}_i = 6 - (-2) = 8$

Using the mean squared error where $n = 2$:

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

$$= -\frac{2}{2} \left[ (3 \times 1) + (8 \times 3) \right]$$

$$= -1 (3 + 24) = \underline{\underline{-27}}$$

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

$$= \frac{-2}{2} (3 + 8)$$

$$= -1 \times 11 = \underline{\underline{-11}}$$

Updating m and b using gradient descent:

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

where $\frac{\partial J}{\partial m} = -27$ and $\alpha = 0.1$.

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

$$= -1 + 2.7 = \underline{\underline{1.7}}$$

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

where $\frac{\partial J}{\partial b} = -11$

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

$$= 1 + 1.1 = \underline{\underline{2.1}}$$

After Iteration 1,

$m = 1.7$

$b = 2.1$



Makuochi Prince

Iteration 2

Step I: Calculate predicted values ($\hat{y}$)

For point (1,3):

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

For point (3,6):

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

Step 2: Calculate errors ($y - \hat{y}$)

for point (1,3):

$y_1 - \hat{y}_1 = 3 - 3.8 = -0.8$

For point (3,6)

$y_2 - \hat{y}_2 = 6 - 7.2 = -1.2$

Step 3: Compute gradients

For m:

$\partial J / \partial m = -2/2 \, [(-0.8 \times 1) + (-1.2 \times 3)] = -1(-0.8 - 3.6) = 4.4$

for b

$\partial J / \partial b = -2/2 \, [-0.8 + (-1.2)] = -1(-2) = 2$

update the parameters

New M:

$M\_New = 1.7 - 0.1 \times 4.4 = 1.7 - 0.44$

$1.26$ //

New b:

$b\_New = 2.1 - 0.1 \times 2 = 2.1 - 0.2 = 1.9$

After iteration 2

$M = 1.26, \; b = 1.9$

Hortance Irakoze

# Iteration 3

Step 1: Computing $\hat{y}$

→ for (1, 3)

$\hat{y} = 1.26 + 1.9 = 3.16$

$y - \hat{y} = 3 - 3.16$
$= -0.16$

→ for (3, 6)

$\hat{y} = 1.26(3) + 1.9$
$= 3.78 + 1.9$
$= 5.68$

$y - \hat{y} = 6 - 5.68$
$= 0.32$

Step 2: finding gradients

$\frac{\partial J}{\partial m} = -\frac{2}{2}\left[(-0.16)(1) + (0.32)(3)\right]$
$= -1(-0.16 + 0.96)$
$= -0.8$

$\frac{\partial J}{\partial b} = -\frac{2}{2}\left[-0.16 + 0.32\right] = -1(0.16)$
$= -0.16$

Step 3: set new values

$m = 1.26 - 0.1(-0.8)$
$= 1.26 + 0.08$
$= 1.34$

$b = 1.9 - 0.1(-0.16) = 1.9 + 0.016$
$= 1.916$

The values of m and b are moving towards reducing the error. Translating to our model learning and adjusting to the observations to better fit our data on each iteration.

Davy Ngamije

**Iteration IV**

step 1: Compute $\hat{y}$

→ for (1,3)

$$\hat{y} = 1.34 + 1.916$$
$$= 3.256$$
$$y - \hat{y} = -0.256$$

→ for (3,6)

$$\hat{y} = 1.34(3) + 1.916$$
$$= 4.02 + 1.016$$
$$= 5.936$$
$$y - \hat{y} = 0.064$$

step 2: find gradients

$$\frac{\partial J}{\partial m} = -1\left[(-0.256)(1) + (0.064)(3)\right]$$
$$= -1\left[-0.256 + 0.192\right]$$
$$= 0.064$$



$$\frac{\partial J}{\partial b} = -1(-0.256 + 0.064) = 0.192$$

Step 3: Update Values

$$m = 1.34 - 0.1(0.064)$$
$$= 1.34 - 0.0064$$
$$= 1.3336$$

$$b = 1.916 - 0.1(0.192) = 1.916 - 0.019$$
$$= 1.8068$$

After 4 iterations

$$m \simeq 1.3336$$
$$b \simeq 1.8068$$

Trend Observation:

m: -1 → 1.7 → 1.26 → 1.34 → 1.3336 (stabilizing ~1.33)

b: 1 → 2.1 → 1.9 → 1.916 → 1.8968 (stabilizing ~1.90)

Key Insight:

Parameters do not move toward (0,0).

Instead, they converge to values that minimize error for the given data.

By the 4th iteration, changes become very small (Δm ≈ 0.006, Δb ≈ 0.02), indicating near-convergence.

Final Line Equation:
y≈1.3336x+1.8968
y≈1.3336x+1.8968

This fits the data points (1,3) and (3,6) with minimal error (MSE ≈ 0.032).