Iteration 1 - MITAU Béla

1) Initial Parameters :

$m = -1$, $b = 1$
learning rate : $\alpha = 0.1$
data points $= (1,3)$ and $(3,6)$

2. Calculate predicted values $(\hat{y}_i)$:

for $(x_1, y_1) = (1,3) : \hat{y}_1 = (m)(x_1) + b$
$= (-1)(1) + 1 = 0$

for $(x_2, y_2) = (3,6) : \hat{y}_2 = (m)(x_2) + b$
$= (-1)(3) + 1 = -2$

3. Calculate Errors : $(y_i - \hat{y}_i)$:

$(y_1 - \hat{y}_1) = 3 - 0 = 3$
$(y_2 - \hat{y}_2) = 6 - (-2) = 8$

4. Calculate Gradients $\left(\frac{\partial J}{\partial m} \text{ and } \frac{\partial J}{\partial b}\right)$:

$$\frac{\partial J}{\partial m} = -\frac{2}{n}\sum(y_i - \hat{y}_i)x_i = -\frac{2}{2}[(3)(1) + (8)(3)]$$
$$= -1[3 + 24] = -27$$
$$\frac{\partial J}{\partial b} = -\frac{2}{n}\sum(y_i - \hat{y}_i) = -\frac{2}{2}[3+8] = -1[11] = -11$$

5. Update Parameters !

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

$$= -1 + 2.7 = 1.7$$

$$b_{new} = b - \alpha \frac{\partial J}{\partial b} = 1 - (0.1)(-11) = 1 + 1.1$$

$$= 2.1$$

Iteration 1 !

$m = 1.7$

$b = 2.1$

Truck Capacity

We are going to optimize Delivery Cost with Respect to the Truck

① Cost to deliver goods depends on the Truck capacity.

② Cost (c) or $f(c) = (c-4)^2+3$

* Minimum cost at Capacity $c = 4$ tons.
* Cost rises if capacity is too small or too large.

C_i : 0, learning rate (α) = 0.1

$$f(c) = (c-4)^2+3$$

$$\frac{df}{dc} = 2(c-4)$$

Iteration 1: on $C_i = 0$

$$\frac{df}{dc} = 2(0-4) = -8$$

$$C = 0 - 0.1 \cdot (-8) = 0 + 0.8 = 0.8$$

Iteration 2:

Current $C = 0.8$

$$\frac{df}{dc} = 2(0.8 - 4) = 2(-3.2)$$

$$\boxed{= -6.4}$$

$$C_{updated} = 0.8 - 0.1 \cdot (-6.4)$$

$$= 0.8 + 0.64$$

$$\boxed{= 1.44}$$

Iteration 3:

Current $C = 1.44$

$$\frac{df}{dc} = 2(1.44 - 4)$$

$$= 2(-2.56) = -5.12$$

$$C_{updated} = 1.44 + 0.512$$

$$\boxed{= 1.952}$$

Iteration 4:

Current $C = 1.952$

$$\frac{df}{dc} = 2(1.952 - 4)$$

$$= 2(-2.048)$$

$$= -4.096$$

$$C = 1.952 + 0.4096 = \boxed{2.3616}$$

Iteration Table

Iteration	Capacity(C)	Gradient	Cost f(c)
1	0.8	−8	$(0.8-4)^2+3 = 13.24$
2	1.44	−6.4	9.55
3	1.952	−5.12	7.52
4	2.3616	−4.096	5.68

Iteration 3

1. Initial Parametres

- $m = 1.26$
- $b = 1.9$
- Learning rate $a = 0.1$
- Data points: $(x_1, y_1) = (1, 3)$ and $(x_2, y_2) = 3, 6)$

2. Calculate Predicted Values $(y^\wedge i)$:

- For $(x1, y1) = (1, 3)$: $y^\wedge 1 = (m)(x1) + b = (1.26)(1) + 1.9 = 1.26 + 1.9 = 3.16$
- for $(x2, y2) = (3, 6)$: $y^\wedge 2 = (m)(x2) + b = (1.26)(3) + 1.9 = 3.78 + 1.9 = 5.68$

3. Calculate Errors $(yi - y^\wedge i)$:

- $(y1 - y^\wedge 1) = 3 - 3.16 = -0.16$
- $(y2 - y^\wedge 2) = 6 - 5.68 = 0.32$

4. Calculate Gradients $(\partial m \, \partial J)$ and $(\partial b \, \partial J)$:

- $\partial m \, \partial b = -n 2 \sum (yi - y^\wedge i) xi = -22 [(-0.16)(1) + (0.32)(3)] = -1 [-0.16 + 0.96]$
$= -1 [0.8] = -0.8$

- $\partial b \, \partial J = n 2 \sum (yi - y^\wedge i) = -22 [-0.16 + 0.32] = 1 [0.16] = -0.16$

5. Update Parametres:

- $m_{new} = m - a \, \partial m \, \partial J = 1.26 - (0.1)(-0.8) = 1.26 + 0.08 = 1.34$
- $b_{new} = b - a \, \partial b \, \partial = 1.9 - (0.1)(-0.16) = 1.9 + 0.016 = 1.916$

Results fo interation 3

$m = 1.34$

$b = 1.916$

Date: 22/ 06 /2025

ITERATION 4

CHARLOTTE KARIZA

1st Initial Parameters for the iteration;

$m = 1.34$

$b = 1.916$

Learning rate $\alpha = 0.1$

Data points: $(x_1, y_1) = (1,3)$ and $(x_2, y_2) = (3,6)$

No of data points $n = 2$

2nd Calculate Predicted Values $(\hat{y}_1)$;

$\Rightarrow (x_1, y_1) = (1,3) : (\hat{y}_1) = (m)(x_1) + b = (1.34)$

$(1) + 1.916 = 1.34 + 1.916 = 3.256$

$\Rightarrow (x_2, y_2) = (3,6) : (\hat{y}_2) = (m)(x_2) + b = (1.34)$

$(3) + 1.916 = 4.02 + 1.916 = 5.936$

3rd Calculate Errors $(y_i - \hat{y}_i)$

$\Rightarrow (y_1 - \hat{y}_1) = 3 - 3.256 = -0.256$

$\Rightarrow (y_2 - \hat{y}_2) = 6 - 5.936 = 0.064$

Date: / /

4^{th} Calculate Gradients ($\partial m \partial J$ and $\partial b \partial J$):

$\Rightarrow \partial m \partial J = -n2\Sigma(yi - y^{\wedge}i)xi = -22[(-0.256)(1) + (0.064)(3)] = -1[-0.256 + 0.192] = -1[-0.064] = \underline{0.064}$

$\Rightarrow \partial b \partial J = -n2\Sigma(yi - y^{\wedge}i) = -22[-0.256 + 0.064] = -1[-0.192] = \underline{0.192}$

5^{th} Update Parameters;

$\Rightarrow mnew = m - \alpha \partial m \partial J = 1.34 - (0.1)(0.064) = 1.34 - 0.0064 = 1.3336$

$\Rightarrow bnew = b - \alpha \partial b \partial J = 1.916 - (0.1)(0.192) = 1.916 - 0.0192 = 1.8968$

Results for iteration 4;

$m = 1.3336$

$b = 1.8968$

Date: / /

Iteration	m Value	b Value	MSE (J)
0 (Initial	-1	1	36.5
1	1.7	2.1	1.04
2	1.26	1.9	0.064
3	1.34	1.916	0.034816
4	1.3336	1.8968	0.034816

TREND ANALYSIS - ALL GROUP MEMBERS

TREND OBSERVATIONS:

1. Slope (m) BEHAVIOR:
   - Started completely wrong (-1)
   - Made large correction (+1.7)
   - Fine-tuned toward optimal (~1.5)
   - Changes getting smaller each iteration

2. INTERCEPT (b) BEHAVIOR:
   - Started at 1
   - overshot 2.1, then settled around 1.9
   - Approaching optimal value (~1.5)

3. ERROR Reduction:

   - Massive improvement from iteration 1
   - Continued refinement in later iterations
   - Parameters stabilizing (Smaller changes)

YES! The parameters are clearly moving toward reducing the error. Both m and b are converging to optimal values, and the changes get smaller each iteration, showing successful gradient descent.