Assignment - 2

Ernin Debbarma

1) a)

x_1	x_2	y
2	1	5
4	3	7
6	5	8

Weights ($\theta_0 = 1.5, \theta_1 = 0.3, \theta_2 = 0.7$)

Learning rate (α) = 0.01

Cost function (MSE) $= \frac{1}{N} \sum_{i=1}^{N} (y_{pred,i} - y_i)^2$

∴ Batch size $e_i = y_{pred,i} - y_i$

$N = 3$

∴ Precision = 4 decimal places

∴ Gradient Descent (for each θ_j)

∴ Using MSE,

$$\frac{dJ}{d\theta_j} = \frac{2}{N} \sum_{i=1}^{N} (y_{pred,i} - y_i) \times (\text{feature})_{j,i}$$

∴ Weight update:

$$\theta_j \leftarrow \theta_j - \alpha \times \frac{dJ}{d\theta_j}$$

i) $(x_1, x_2, y) = 2, 1, 5$

$y_{pred} = 1.5 + (0.3)(2) + (0.7)(1) = 1.5 + 0.6 + 0.7$
$= 2.8$

$e_1 = y_{pred} - y = 2.8 - 5 = -2.2$

ii) $(x_1, x_2, y) = (4, 3, 7)$

$y_{pred} = 1.5 + (0.3)(4) + (0.7)(3) = 1.5 + 1.2 + 2.1$
$= 4.8$

$e_2 = 4.8 - 7 = -2.2$

$(x_1, x_2, y) = (6, 5, 8)$

$$y_{pred} = 1.5 + (0.3)(6) + (0.7)(5)$$
$$= 1.5 + 1.8 + 3.5 = 6.8$$
$$e_3 = 6.8 - 8 = -1.2$$

ii) Gradients $(N=3)$

; Gradient θ_0 (feature $=1$) :

$$\frac{dJ}{d\theta_0} = \frac{2}{3}\left[(-2.2) + (-2.2) + (-1.2)\right] = \frac{2}{3} \times (-5.6)$$
$$= -3.7333$$

; Gradient θ_1 (feature $= x_1$)

$$\frac{dJ}{d\theta_1} = \frac{2}{3}\left[(-2.2\times 2) + (2.2 \times 4) + (-1.2\times 6)\right]$$
$$= \frac{2}{3}\left[-4.4 - 8.8 - 7.2\right]$$
$$= \frac{2}{3} \left[-20.4\right]$$
$$= -13.6000$$

; Gradient θ_2 (feature $= x_2$)

$$\frac{dJ}{d\theta_2} = \frac{2}{3}\left[(-2.2\times 1) + (-2.2\times 3) + (-1.2\times 5)\right]$$
$$= \frac{2}{3}\left[-2.2 - 6.6 - 6.0\right]$$
$$= \frac{2}{3} \times (-14.8)$$
$$= -9.8667$$

iii) Weight update:

- $\theta_0 \leftarrow 1.5 - 0.01 \times (-3.7333)$
$= 1.5 + 0.037333$
≈ 1.5373

- $\theta_1 \leftarrow 0.3 - 0.01 \times (-13.6000)$
$= 0.3 + 0.1360$
≈ 0.4360

- $\theta_2 \leftarrow 0.7 - 0.01 \times (-9.8667)$
$= 0.7 + 0.098667$
≈ 0.7987

iv) Loss:

$(e_1)^2 = (-2.2)^2 = 4.84$
$(e_2)^2 = (-2.2)^2 = 4.84$
$(e_3)^2 = (-1.2)^2 = 1.44$

$\rightarrow \sum (e_i)^2 = 4.84 + 4.84 + 1.44$
$= 11.12$

$MSE = \frac{11.12}{3} \approx 3.7067$

$\rightarrow$ updated weights and loss:

$\theta_0 = 1.5373$
$\theta_1 = 0.4360$
$\theta_2 = 0.7987$
$MSE \approx 3.7067$

Iteration 2

$\theta_0 = 1.5373, \ \theta_1 = 0.4360, \ \theta_2 = 0.7987$

i) $(x_1, x_2, y) = (2, 1, 5)$

$y_{pred} = 1.5373 + (0.4360)(2) + (0.7987)(1)$

$= 1.5373 + 0.8720 + 0.7987$

$= 3.2080$

$e_1 = 3.2080 - 5$

$= -1.7920$

$(x_1, x_2, y) = (4, 3, 7)$

$y_{pred} = 1.5373 + (0.4360)(4) + (0.7987)(3)$

$= 1.5373 + 1.7440 + 2.3961$

$= 5.6774$

$e_2 = 5.6774 - 7 = -1.3226$

$(x_1, x_2, y) = (6, 5, 8)$

$y_{pred} = 1.5373 + (0.4360)(6) + (0.7987)(5)$

$= 1.5373 + 2.6160 + 3.9935$

$= 8.1468$

$e_3 = 8.1468 - 8$

$= 0.1468$

ii) Gradient: θ_0

$$\frac{dJ}{d\theta_0} = \frac{2}{3}\left[(-1.7920) + (-1.3226) + (0.1468)\right]$$

$$= \frac{2}{3} \times (-2.9678)$$

$$= -1.9785$$

· θ_1

$$\frac{dJ}{d\theta_1} = \frac{2}{3}\left[(-1.7920 \times 2) + (-1.3226 \times 4) + (0.1468 \times 6)\right]$$

$$= \frac{2}{3}\left[-3.5840 - 5.2904 + 0.8808\right]$$

$$= \frac{2}{3} \times (-7.9936)$$

$$= -5.3291$$

· θ_2

$$\frac{dJ}{d\theta_2} = \frac{2}{3}\left[(-1.7920 \times 1) + (-1.3226 \times 3) + (0.1468 \times 5)\right]$$

$$= \frac{2}{3}\left[-1.7920 - 3.9678 + 0.7340\right]$$

$$= \frac{2}{3}\left[-4.0258\right]$$

$$= -3.3505$$

iii) Weight updates:

· $\theta_0 \leftarrow 1.5373 - 0.01 \times (-1.9785) = 1.5373 + 0.019785$
≈ 1.5571

· $\theta_1 \leftarrow 0.4360 - 0.01 \times (-5.3291) = 0.4360 + 0.053291$
≈ 0.4893

· $\theta_2 \leftarrow 0.7987 - 0.01 \times (-3.3505) = 0.7987 + 0.033505$
≈ 0.8322

iv) Loss

$(e_1)^2 = (-1.7920)^2$
≈ 3.2113

$(e_2)^2 = (-1.3226)^2$
≈ 1.7503

$(e_3)^2 = (0.1468)^2$
≈ 0.0216

$\therefore \Sigma (e_i)^2 \approx 3.2113 + 1.7503 + 0.0216$
$= 4.9832$

$MSE = \dfrac{4.9832}{3}$
≈ 1.6611

$\therefore$ Updated weights and loss:

$\theta_0 = 1.3571$, $\theta_1 = 0.4893$, $\theta_2 = 0.8322$,

$MSE \approx 1.6611$

b) Prediction for $X_1 = 5, X_2 = 4$

$$y_{pred} = \theta_0 + \theta_1 \times 5 + \theta \times 4$$
$$= 1.5571 + (0.4893 \times 5) + (0.8322 \times 4)$$

$$\theta_1 \times 5 = 0.4893 \times 5 = 2.4465$$

$$\theta_2 \times 4 = 0.8322 \times 4 = 3.3288$$

$$\Rightarrow y_{pred} = 1.5571 + 2.4465 + 3.3288$$
$$= 7.3324.$$

2) Using Python’s numerical operations, we performed linear regression on the two-feature dataset. Gradient descent successfully updated the weights toward minimizing the MSE, and a 3D visualization illustrated that the final model fits the data as expected.