Honourgod Levison

Formative 3 - Iteration 1

Part 3 - Gradient Descent Manual Calculation.

Given; $y_{predict} = mx + b$

Where; Initial $m = 1$, Initial $b = 1$

Learning rate $(\alpha) = 0.1$

Given Points: $(1,3)$ and $(3,6)$

Required to:

1 Compute the predicted value $y_{predict}$ for each data point

2 Derive the gradient of the cost function $J(m,b)$. Using MSE

- Iteratively update m and b

Solution.

1. Finding $(y_{predict} - y)$; given $y_{predict} = mx+b$

For Point 1 $(x=1, y=3)$:

- $y_{predict} = (-1)(1) + 1 = 0$
- Error$_1$ $= (\hat{y} - y) = 0 - 3 = -3$

For Point 2 $(x=3, y=6)$

- $y_{predict} = (-1)(3) + 1 = -2$
- Error 2 $= (y_{predict} - y) = -2 - 6 = -8$

2 Deriving the gradient of the cost function.

Using cost function $J(m,b) = \frac{1}{n}\sum_{i=1}^{n}(mx_i + b - y_i)^2$

* With respect to m (slope)

$$\frac{\partial J}{\partial m} = \frac{\partial}{\partial m}\left[\frac{1}{n}\sum_{i=1}^{n}(mx_i + b - y_i)^2\right]$$

Having a linear differentiation, we can move the derivative inside the summation

$$\frac{\partial J}{\partial m} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial m}\overbrace{(mx_i + b - y_i)^2}^{u^2}$$

Let $(mx_i + b - y_i)^2$ be u^2

Applying Chain rule:

The derivative of u^2 is $2u \cdot u'$

$$\frac{\partial J}{\partial m} = \frac{1}{n}\sum_{i=1}^{n} 2(mx_i + b - y_i) \cdot \frac{\partial}{\partial m}(mx_i)$$

derivative of u with respect to m is x_i (b and y_i are constants)

$$\therefore \frac{\partial J}{\partial m} = \frac{1}{n}\sum_{i=1}^{n} 2(mx_i + b - y_i) \cdot x_i$$

$$= \frac{\partial J}{\partial m} = \frac{2}{n}\sum_{i=1}^{n}(mx_i + b - y_i)x_i$$

* With respect to b (intercept)

$$\frac{\partial J}{\partial b} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial b}(mx_i + b - y_i)^2$$

Let $(mx_i + b - y_i)^2$ be u^2

Derivative of $u^2 = 2u \cdot u'$

Derivative of u with respect to b $= 1$ (mx_i and y_i are constants)

$$\frac{\partial J}{\partial b} = \frac{1}{n}\sum_{i=1}^{n} 2(mx_i + b - y_i) \cdot \frac{\partial}{\partial b}(b)$$

$$\therefore \frac{\partial J}{\partial b} = \frac{1}{n}\sum_{i=1}^{n} 2(mx_i + b - y_i) \cdot (1)$$

$$= \frac{\partial J}{\partial b} = \frac{2}{n}\sum_{i=1}^{n}(mx_i + b - y_i)$$

For Matitle 3 Continued

B. .

2. ~~Get~~ Getting gradient m and b

$$G_m = \frac{\partial J}{\partial m} = \frac{\cancel{2}}{\cancel{2}}[(-3.1)+(-8.3)]$$

Meaning sumation of all $(m x_i + b - y_i)$

$$-3 + (-24) = -27$$

$$G_b = \frac{\partial J}{\partial m} = \frac{\cancel{2}}{\cancel{2}}[-3+(-8)] = -11$$

3. Iteratively ~~updating~~ updating m and b

for Iteration 1

using;

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

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

New_m =

$$m_{new} = -1 - (0.1 \cdot (-27)) = 1.7$$

$\therefore m_{new} = 1.7$

New b =

$$b_{new} = 1 - (0.1 \cdot (-11)) = 2.1$$

$\therefore b_{new} = 2.1$

Summary:

Parameter	Initial Value	Gradient	New Value
Slope (m)	-1	-27	1.7
Intercept (b)	1	-11	2.1

Name : MUHIRE NGABO Jean

Starting Values : $m = 1.7$, $b = 2.1$

Step 1: Calculate Predictions ($\hat{y}$) & Errors

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

Error (e) $= y - \hat{y}$ or $\hat{y} - y$

And we have two points $(1,3)$ and $(3,6)$.

For point 1 $(1,3)$

- $\hat{y}_1 = (1.7)(1) + (2.1) = 3.8$
- Error (e) $= 3.8 - 3 = 0.8$

For point 2 $(3,6)$

- $\hat{y}_2 = (1.7)(3) + (2.1) = 5.1 + 2.1 = 7.2$
- $e_2 = 7.2 - 6 = 1.2$

Step 2: Calculate Gradients.

- $\frac{\partial J}{\partial m} = \frac{2}{n} \sum_{i=1}^{n} (\hat{y} - y)x$; Gradient for m (Sum of Error . x), $n = 2$

$$\frac{\partial J}{\partial m} = (0.8)(1) + (1.2)(3) = 0.8 + 3.6 = 4.4$$

- $\frac{\partial J}{\partial b} = \frac{2}{n} \sum_{i=1}^{n} (\hat{y} - y)$; Gradient for b (Sum of Errors) and $n = 2$

$$\frac{\partial J}{\partial b} = 0.8 + 1.2 = 2.0$$

Step 3 : Update Parameter

$m_{new} = m_{old} - \alpha \frac{\partial J}{\partial m}$ as $\alpha = 0.1$

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

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

$b_{new} = (1.7) - (0.1)(2.0) = 2.1 - 0.2 = 1.9$

Name: NKUSI ORRIE DAN

Starting Value: $m = 1.26$, $b = 1.9$

Step 1 . Calculate Y' & Error

Y' , Error

Point 1 (1,3)

$y'_1 = (1.26 \times 1) + 1.9 = 3.16$

Error = $3.16 - 3 = 0.16$

Point 2 (3,6)

$y'_2 = (1.26 \times 3) + 1.9 = 3.78 + 1.9 = 5.68$

Error = $5.68 - 6 = -0.32$

Step 2 : Calculate Gradients

Gradient for m (Sum of Error . x)

$\frac{dr}{dm} = (0.16 \times 1) + (-0.32 \times 3) = 0.16 - 0.96$
$= -0.8$

Gradient for b

$\frac{dl}{dm} = 0.16 + (-0.32) = -0.16$

Step 3 . Update m & b

$m_{new} = 1.26 - (0.1 \times (-0.8)) = 1.26 + 0.08$

$m_{new} = 1.34$

$b_{new} = 1.9 - (0.1 \times (-0.16))$
$= 1.9 + 0.016$

$b_{new} = 1.916$

iteration 4

David Akintayo

Starting values ! $m = 1.34$, $b = 1.916$

Step 1 : Calculate Predictions ($\hat{y}$) $ Errors

at Point 1 (1,3)

prediction: $\hat{y}_1 = (1.34 \times 1) + 1.916 = 3.256$

Error : $3.256 - 3 = 0.256$ //

at Point 2 (3,6)

prediction: $(\hat{y}_2) = (1.34 \times 3) + 1.916 = 4.02 + 1.916$

$= 5.936$ //

$\therefore$ Error : $= 5.936 - 6 = -0.064$ //

Step 2 : Calculate Gradients

Gradient for m (Sum of Error x)

$$\frac{\partial l}{\partial m} = (0.256 \cdot 1) + (-0.064 \cdot 3) = 0.256 - 0.192$$

$$= 0.064 \;//$$

Gradient for b (sum of Errors)

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

Step 3 : update parameters

$\therefore$ $m_{new} = 1.34 - (0.1 \times 0.064) = 1.34 - 0.0064$ -

$= 1.3336$ //

$$\therefore\ b_{new} = 1.916 - (0.1 \times 0.192) = 1.916 - 0.0192$$

$$= 1.8968$$

Final Group Summary Table

Iteration	Gradient m	Gradient b	New m	New B
1 (start)	-27	-11	1.7	2.1
2	4.4	2.0	1.26	1.9
3	0.8	-0.16	1.34	1.916
4	0.064	0.192	1.3556	1.8968