Mo Tu We Th Fr Sa Su

Memo No.

Date

Initial parameters: $m = -1,\ b = 1$

Learning rate $\alpha = 0.1$

$(x_1, y_1) = (1, 3) \quad (x_2, y_2) = (3, 6),\ n = 2$

$\hat{y}_1 = m x_1 + b \qquad \hat{y}_2 = m x_2 + b$

$\hat{y}_1 = -1(1) + 1 \qquad \hat{y}_2 = -1(3) + 1$

$\hat{y}_1 = 0 \qquad \hat{y}_2 = -2$

$y_1 = 3 \qquad y_2 = 6$

Error for $y_1 = 3 - 0 = 3$

Error for $y_2 = 6 - (-2) = 8$

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

$$= \frac{1}{2} \left[(3-0)^2 + 6 - (-2)^2 \right]$$

$$= \frac{1}{2}(9 + 64) \rightarrow \frac{73}{2} = 36.5$$

Gradients

$$\frac{dJ}{dm} = \frac{-2}{n} \sum (y - \hat{y}) x$$

Mo Tu We Th Fr Sa Su

Memo No.

Date / /

$$\frac{dJ}{db} = -\frac{2}{n} \sum (y - \hat{y})$$

$$\frac{dJ}{dm} = -\frac{2}{n}\left[(y_1 - \hat{y}_1)x_1 + (y_2 - \hat{y}_2)x_2\right]$$

$$= \frac{-2}{2}\left[(3\times 1) + (8\times 3)\right] \rightarrow -1(27) = \cancel{-11}\ -27$$

$$\frac{dJ}{db} = \frac{-2}{n}\left[(y_1 - \hat{y}_1) + (y_2 - \hat{y}_2)\right]$$

$$= \frac{-2}{2}(3+8) \rightarrow -1(11) = -11$$

$$m_{new} = m_{old} - \alpha\left[\frac{dJ}{dm}\right]$$

$$b_{new} = b_{old} - \alpha\left[\frac{dJ}{db}\right]$$

$$m_{new} = -1 - \left[(0.1)(-27)\right]$$

$$= -1 - (-2.7) \rightarrow -1 + 2.7 = 1.7$$

$$b_{new} = 1 - \left[(0.1)(-11)\right]$$

$$= 1 - (-1.1) \rightarrow 1 + 1.1 = 2.1$$

$$m_{new} = 1.7, \quad b_{new} = 2.1$$

Gradient descent Iteration 2:

Using the updated values from our 1^{st} iteration:

- Data Point: (1,3) and (3,6)
- learning rate: 0.1
- m: 1.7
- b: 2.1

Next using the linear model we find the 2 errors:

So from $y = mx + b$, for $x = 1$: $\bar{y} = 1.7 \times 1 + 2.1 = 3.8$

$x = 3$: $\bar{y} = 1.7 \times 3 + 2.1 = 7.2$

So Error1 $= 3 - 3.8 = -0.8$

Error 2 $= 6 - 7.2 = -1.2$

Next we compute the gradients and so from The MSE where $n = 2$:

gradient of m: $\frac{\partial J}{\partial m} = -\frac{2}{n}\sum(y - \bar{y})x = \frac{-2}{2} \times [(-0.8)(1) + (-1.2)(3)]$

$= -1 \times (-0.8 - 3.6) = -1 \times (-4.4)$

$\frac{\partial J}{\partial m} = 4.4$

gradient of b: $\frac{\partial J}{\partial b} = -\frac{2}{n}\sum(y - \bar{y}) = -1 \times [-0.8 + (-1.2)]$

$\frac{\partial J}{\partial b} = 2.0$

next and final step we get to updating the values of both m and b

CS CamScanner

$m_{new} = m - \alpha \cdot \frac{\partial J}{\partial m} = 1.7 - 0.1 \times 4.4 = 1.26$

$b_{new} = b - \alpha \cdot \frac{\partial J}{\partial b} = 2.1 - 0.1 \times 2.0 = 2.1 - 0.2 = 1.9$

Final results:

new slope $m = 1.26$

new intercept $b = 1.90$

CamScanner

Data

$M_{new_1} = 1.7$ $\qquad (x_1, y_1) = (1,3)$

$B_{new_1} = 2.1$ $\qquad (x_2, y_2) = (3,6)$

$M_{new_2} = 1.26$

$B_{new_2} = 1.9$

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

$$\frac{\partial J}{\partial b} = \frac{-2}{n} \sum (y$$

* $\hat{y}_1 = 1.26(1) + 1.9 = 3.16$
* $\hat{y}_2 = 1.26(3) + 1.9 = 5.68$

* $y - \hat{y}$:

$y - \hat{y}_1 \Rightarrow 3 - 3.16 = -0.16$

$y - \hat{y}_2 \Rightarrow 6 - 5.68 = 0.32$

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

$= -\frac{1}{2}[-0.16 + 0.96] = -\frac{0.8}{2}$

$= -0.4$

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

$$= -\frac{0.16}{2}$$

$$= -0.08$$

So, $m_3 = m_2 - \alpha \cdot \frac{\partial J}{\partial m}$

$= 1.26 - 0.1(-0.4)$

$= 1.26 + 0.04$

$= 1.30$

$b_3 = b_2 - \alpha \cdot \frac{\partial J}{\partial b}$ ~~1.9 -~~

$= 1.9 - 0.1(-0.08)$

$= 1.9 + 0.008$

$= 1.908$

$m_{new\,3} = 1.30$

$B_{new\,4} = 1.908$

Scanned with CamScanner