Gradient descent

initial values

$m = -1$ $b = 1$, $\alpha = 0.1$

x	1	3
y	3	6

let's find $\hat{y}$ using m and b

$\hat{y}_1 = mx + b$

$= (-1)(1) + 1 = 0$

$\hat{y}_2 = (-1)(3) + 1 = -2$

expected y	predicted y
3	0
6	-2

let's find the new value of m and b

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

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

Suppose MSE is J

then the ~~define~~ differentiation of MSE with respect to m is

$$\frac{\partial J}{\partial m} = \frac{\partial}{\partial m}\left(\frac{1}{n}\sum(y_i - \hat{y}_i)^2\right)$$

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

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

Using the chain rule where

$$u = y_i - mx_i + b$$

now we have

$$\frac{\partial J}{\partial m} = \frac{\partial}{\partial m}\left(\frac{1}{n}\sum u^2\right)$$

$$= \frac{1}{n}\sum \frac{\partial}{\partial m}(u^2)$$

$$= \frac{1}{n}\sum \frac{\partial J}{\partial u} \cdot \frac{\partial u}{\partial m}$$

$$\frac{\partial J}{\partial u} = \frac{\partial}{\partial u}(u^2) = 2u$$

$$\frac{\partial u}{\partial m} = \frac{\partial}{\partial m}(y_i - mx_i - b)$$

$= -x_i$ bcz m is the changing value others are costants bringing everything together we get

$$\frac{\partial J}{\partial m} = \frac{1}{n} \sum 2(y_i - mx_i - b)(-x_i)$$

$$= \boxed{-\frac{2}{n}(y_i - \hat{y}_i)x_i}$$

now that we have the formular of $\frac{\partial J}{\partial m}$

let's find m_{new}

$$\frac{\partial J}{\partial m} = -\frac{2}{2}\left((3-0)(1) + (6-(-2))(3)\right)$$

$$= -1(3+24)$$

$$= -1\ (27) = \underline{\underline{-27}}$$

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

$$= \underline{\underline{1.7}}$$

let's differentiate past formular again with respect to b

$$\frac{\partial J}{\partial b} = \frac{\partial}{\partial b}\left(\frac{1}{n}\sum(y_i - \hat{y}_i)^2\right)$$

$$= \frac{\partial}{\partial b}\left(\frac{1}{n}\sum\left(y_i - (mx_i + b)\right)^2\right)$$

using chain ~~rule~~ ~~to~~ we have

$$\frac{\partial J}{\partial b} = \frac{1}{n}\left(\frac{\partial J}{\partial u} \cdot \frac{\partial u}{\partial b}\right) \text{ where } u = y_i - mx_i - b$$

$$\frac{\partial J}{\partial u} = \frac{\partial}{\partial u}(u^2) = 2u$$

$$\frac{\partial u}{\partial b} = \frac{\partial}{\partial b}(y_i - mx_i - \underline{b})$$

$= -1$ because b is the ~~ot~~ only changing value the rest are just costonts

$$\frac{\partial J}{\partial u} = \frac{1}{n}\sum 2(y_i - mx_i - b)(-1)$$

$$= -\frac{2}{n}\sum(y_i - mx_i - b)$$

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

replacing with ~~the~~ our values of $y_i, \hat{y}_i$ and n we get

$$\frac{\partial J}{\partial u} = -\frac{2}{2}\sum\left((3-0) + (6-(-2))\right)$$

$$= -1(11) = \underline{\underline{-11}}$$

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

$$= 1 - (0.1)(11) = \underline{\underline{2.1}}$$

Now let's find the new value of $\hat{y}$ using the new m and b

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

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

expected y	predicted y
3	3.8
6	7.2

No

Date

Iteration 2:

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

$$= -\frac{2}{2} \sum \left[(3-3.8)(1) + (6-7.2)(3) \right]$$

$$= -1(-0.8-3.6)$$

$$= 4.4$$

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

$$= +1.7 - 0.1(4.4) = 1.26$$

$$\frac{\partial J}{\partial b} = -\frac{2}{2}\left[(3-3.8) + (6-7.2)\right] = 2$$

$$b_{new} = b_{old} - \alpha \frac{\partial J}{\partial b} = 2.1 - 0.1(2) = 1.9$$

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

Expected y	Predicted y
3	3.16
6	5.68

Date: / /

iteration 3 :

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

$$= -\frac{2}{2} \left((3 - 3.16)(1) + (6 - 5.68)(3) \right)$$

$$= -1 \left(-0.16 + 0.96 \right) = \underline{\underline{-0.8}}$$

$$m_{new} = 1.26 - 0.1(-0.8) = \underline{\underline{1.34}}$$

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

$$= -1 \left((3 - 3.16) + (6 - 5.68) \right)$$

$$= -1 \left(-0.16 + 0.32 \right) = \underline{\underline{-0.16}}$$

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

$$= 1.9 - 0.1(-0.16)$$

$$= 1.916$$

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

$$\hat{y}_2 = (1.34)(3) + 1.916 = 5.936$$

expected y	predicted y
3	3.256
6	5.936

Iteration 4

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

$$= -1\left((3-3.256)(1)+(6-5.936)(8)\right)$$

$$= -1(-0.256+0.196)$$

$$= 0.064$$

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

$$= 1.39 - 0.1(0.064) = 1.3336$$

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

$$= -1\left((3-3.256)+(6-5.936)\right)$$

$$= -1(-0.256+0.0064)$$

$$= 0.192$$

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

$$= 1.916 - 0.1(0.192)$$

$$= 1.8968$$

Find $\hat{y}$ using new m and b

$\hat{y}_1 = (1.3336)(1) + 1.8968$

$= 3.2304$

$\hat{y}_2 = (1.3336)(3) + 1.8968$

$= 5.8976$

expected	predicted y
3	3.2307
6	5.8976

~~Conclusion~~

Conclusion

Iteration	m	b
1	1.7	2.1
2	1.26	1.9
3	1.34	1.916
4	1.3336	1.8968

In iteration one, there is a huge jump since error was high but the rest is low because because prediction improved. So with more iterations we can get the optimal value for b and m.