Gradient Descent

Feature (x)	Output (y)
1	1
3	4
2	3

YouTube @ Study Mart

We know, $y = mx + c$ (Liear Equation)

Hypothesis for Linear Regression,

$$h_w(x) = w_0 + w_1 x$$

And, Cost function for Linear Regression,

$$Cost(w_0, w_1) = \frac{1}{2m} \sum_{i=1}^{m} \left\{ h_w(x^i) - y^i \right\}^2$$

Now, We have to update our weight

$$w_{new} = w_{old} - \alpha \frac{\partial \text{ Lose}}{\partial \text{ Old Weight}}$$

α = Learnig Rate

Let's Assume starting weight,

$w_0 = 0.00$ (zero)

$w_1 = 0.00$

So, Hypothesis will be,

$$h_w(x) = 0 + 0 \times x$$
$$= 0$$

Now, Draw a table for predicted output and Error:

x	$y_{out.}$	Predicted Outp.	Error $\{y_p^i - y_o^i\}$	~~cost~~ Error^2
1	1	0	-1	1
3	4	0	-4	16
2	3	0	-3	9

Error = ~~original~~ - ~~predicted~~
or, Predicted - original

Total Error = -8
Total Error^2 = 26

So, cost for w_0 and w_1 weight is,

$$\text{Cost}(w_0, w_1) = \frac{1}{2m}\sum_{i=1}^{m}\left\{h_w\{x^i - y^i\}\right\}^2$$

$$\text{or, } = \frac{1}{2m}\sum_{i=1}^{m}\left(y_{Predicted} - y_{original}\right)^2$$

$$= \left(\frac{1}{2*3}\right) * 26$$

$$= \frac{26}{5}$$

$$= 5.2$$

This is the cost value for first data point, Now, we have need to update our weight.

Weight update Formula

$$w_{new} = w_{old} - \eta \frac{\partial \text{lose}}{\partial w_{old}}$$

Derivatives of lose of old w

$$w_j = w_i - \eta \frac{\partial}{\partial w_j} \text{Lose}(w_0, w_1)$$

when, lose = 0 then,

$$w_{new} = w_{old} - \alpha \frac{1}{m} \sum_{i=1}^{m} \{ h_w(x^i) - y^{(i)} \}$$

when lose > 0 :

$$w_{new} = w_{old} - \alpha \frac{1}{m} \sum_{i=1}^{m} \{ h_w(x^i) - (y^i) \} \cdot x^{(i)}$$

Now, Let's Assume Learning Rate $\alpha = 0.01$

So, We can update our weight :

$$w_0 = 0 - 0.01 \times \frac{1}{3} \times (-8)$$
$$= 0.027$$
$$w_1 = 0 - 0.01 \times \frac{1}{3}(-8) \times 1$$
$$= 0.027$$

Now, $w_0 = 0.027$ and $w_1 = 0.027$

Now we calcuate for 2nd data point :-

$y = 0.027 + 0.027 * x$

x	y	Predicted y	Error	Error^2
1	1	.054	-.946	.894
3	4	.108	-3.892	15.147
2	3	.081	-2.919	8.52

Total = 24.56

$$Cost(w_0, w_1) = \frac{1}{2m} \sum_{i=1}^{m} \{h_w(x^i) - (y^i)\}^2$$

$$= \frac{1}{2 \times 3} \times 24.56$$
$$= 4.09$$

Again update weight and calculate for 3rd data point.