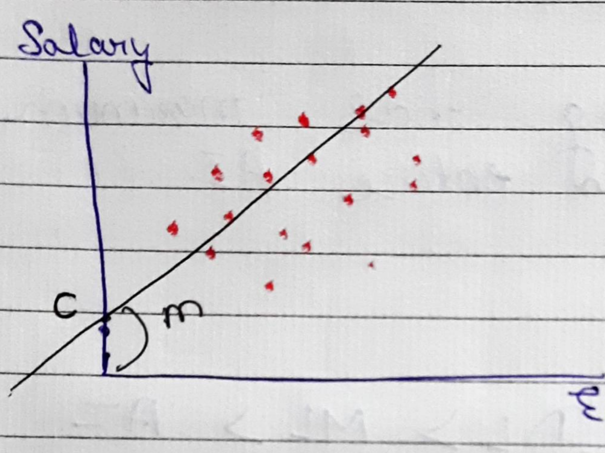
Linear Regression -

Hypothesis - $\hat{y} = mx + c$



exp^n (x)	Salary (y)
–	–
–	–
–	–
–	–

At the time of training : known parameters y & x

unknown parameters

m (slope)

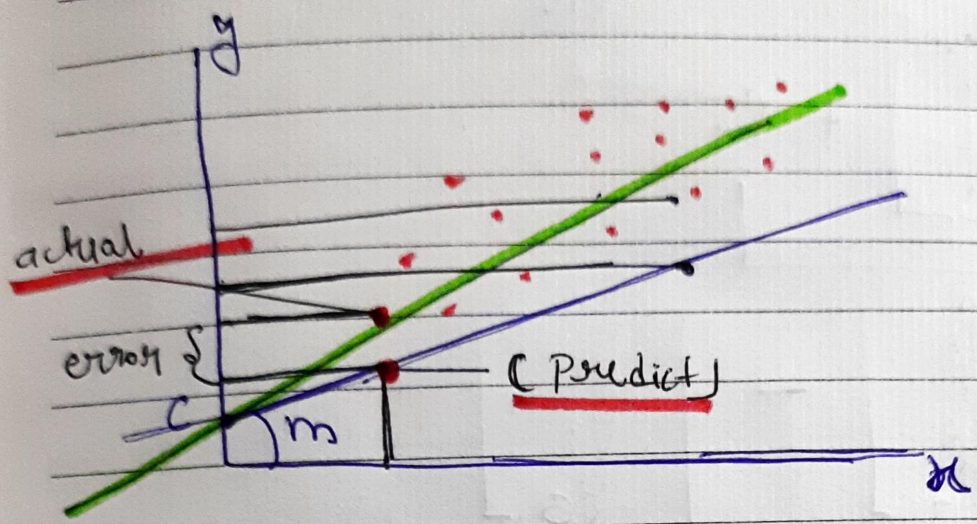
c (y-intercept)

Target :- To have best fit line will predict output with least error

Mo Tu We Th Fr Sa Su

Date / /

Putting random value of m & c -



Forward Propagation and Cost function -

forward propagation -

Step1 - $\hat{y} = mx + c$

Step2 - Calculate the error

$$\text{cost } f^n = \frac{\sum_{i=1}^{n} (\hat{y} - y)^2}{2n}$$

n - no. of instances

$\hat{y}$ — predicted value

y — actual value

$$\text{cost fun} = \frac{\sum_{i=1}^{n} (mx+c-y)^2}{2n}$$

$$m \times \begin{array}{c} x \\ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{array} + c$$

$$\underset{\text{Predicted}}{\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix}} - \underset{\text{actual}}{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}$$

\# Cost function — difference between predicted and actual results and helps improve the model

Step3 - Gradient Desent Algorith - (BackTward Propagation) -

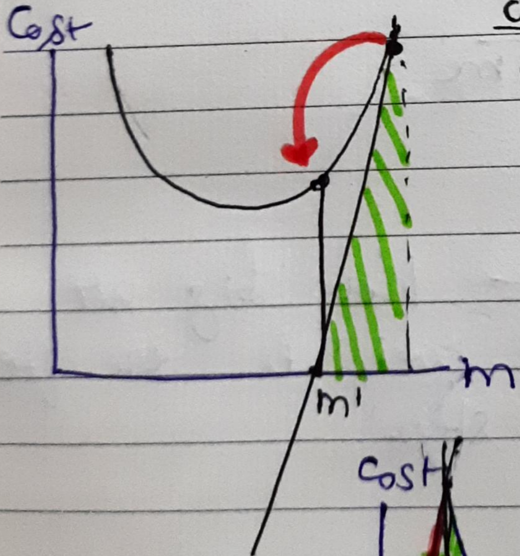
$$m' = m - \underset{\text{learning rate}}{\alpha} \frac{\partial \, Cost}{\partial m}$$

$$c' = c - \alpha \frac{\partial cost}{\partial c}$$

└ Partial differentiation

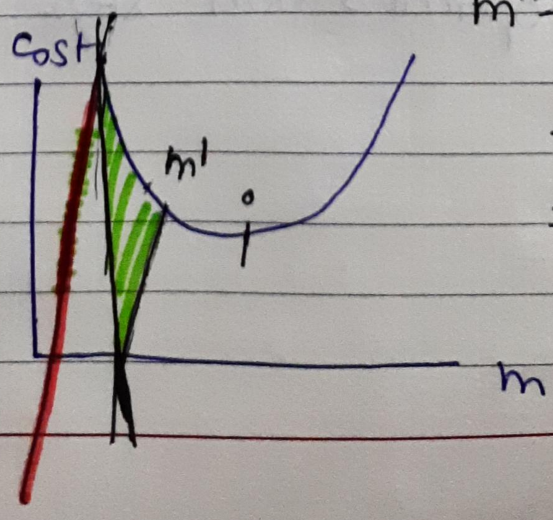
Why we use partial differentiation -

- there are two variable (m & c)
- if there ~~will be~~ is only one variable then it will be fully differentiation



$$\frac{\partial cost}{\partial m}$$

$$m' = m - \frac{\partial cost}{\partial m}$$

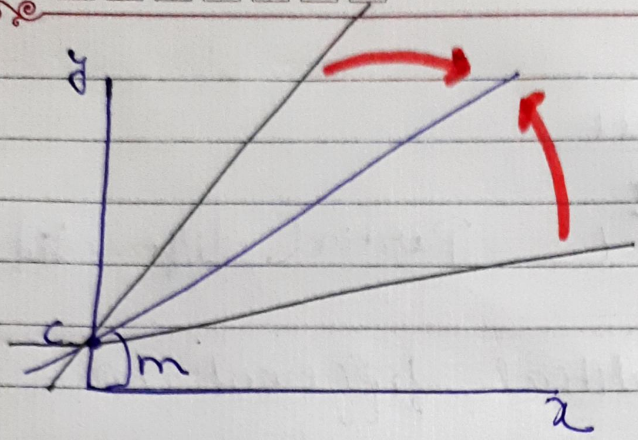


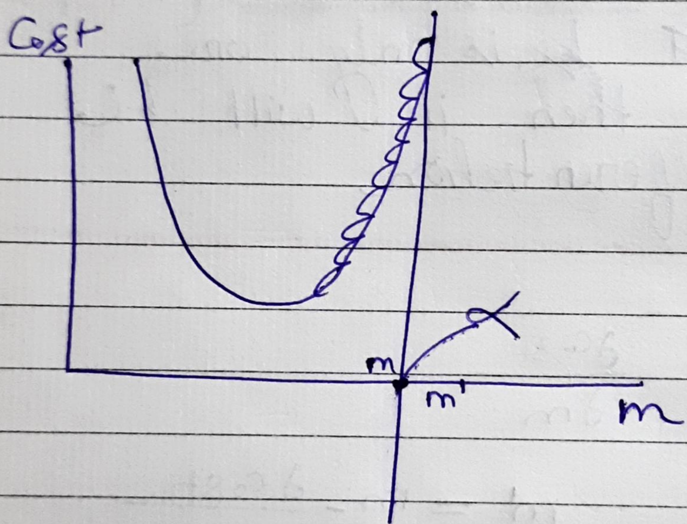
$$m' = m - \frac{\delta cost}{\delta m}$$

$$= m - (-ve)$$

$$= m + \text{value}$$

(reduce cost)





learning rate (α) – How big or small changes in weights are after each optimisation step.

Mo Tu We Th Fr Sa Su Date / /

Step 1 Forward propagation-

$$y = mx + c$$

Step 2 - Cost function -

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

Step 3 - Gradient Desent

$$m' = m - \alpha \frac{\partial Cost}{\delta m}$$

$$c' = c - \alpha \frac{\partial Cost}{\partial m}$$

Mo Tu We Th Fr Sa Su Date / /

$$\Rightarrow \quad m' = m - \alpha \frac{\partial Cost}{\partial m}$$

$$\frac{\partial Cost}{\partial m} = \frac{\partial Cost}{\partial f} \cdot \frac{\partial f}{\partial m}$$

where $\{ f = \hat{y} = mx + c \}$

$$df = \frac{\partial Cost}{\partial f} = \frac{\frac{1}{2n} (\hat{y} - y)^2}{\partial \hat{y}}$$

$$= \frac{2}{2n} (\hat{y} - y)' \cdot \frac{\partial (\hat{y} - y)}{\partial \hat{y}}$$

$$= \frac{\hat{y} - y}{n} \cdot \left(1 - \underbrace{\frac{\partial y}{\partial \hat{y}}}_{\to 0}\right)$$

$$\boxed{df = \frac{\hat{y} - y}{n}} \quad \text{——— (I)}$$

$$\frac{\partial f}{\partial m} = \frac{\partial (mx + c)}{\partial m}$$

$$= \frac{\partial mx}{\partial m} + \frac{\partial c}{\partial m}$$

$$= x + 0$$

$$\boxed{\frac{\partial f}{\partial m} = x} \quad \text{——— (II)}$$

$$\frac{\partial f}{\partial c} = \frac{\partial (mx + c)}{\partial c}$$

$$= 0 + 1$$

$$\boxed{\frac{\partial f}{\partial c} = 1} \text{ --- (III)}$$

m & c — randomly initialised

x & y : known

training input

training output

hypothesis ⇒ $\hat{y} = mx + c$

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

$$dm = \frac{d\,\text{Cost}}{dm} = \frac{\partial\,\text{Cost}}{\partial f} \cdot \frac{df}{dm}$$

$$\boxed{dm = \frac{\hat{y} - y}{n} \cdot x}$$

$$\boxed{dc = \frac{\partial\,\text{Cost}}{\partial c} = \frac{\hat{y} - y}{n}}$$

Mo Tu We Th Fr Sa Su

Date / /

$$\boxed{\begin{aligned} m' &= m - \alpha\, dm \\ c' &= c - \alpha\, dc \end{aligned}}$$

$$\boxed{\begin{aligned} m' &= m - \alpha\left(\frac{\hat{y} - y}{n} \cdot x\right) \\ c' &= c - \alpha\left(\frac{\hat{y} - y}{n}\right) \end{aligned}}$$