Tutorial - 26.

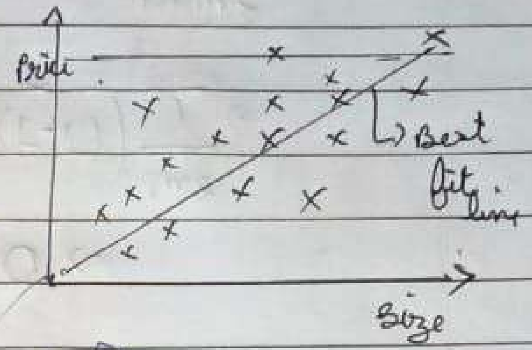
Linear Regression → Indepth Maths Intuition

$\{y = mx + c\}$

↳ Best fit line

m = slope

c ⇒ Intercept.



In the Best fit line for some limit say x to y. the error is summed up to give the least error or the minimal error. gives value of m or c

when → size $= 0$ i.e $x = 0$

$y = mx + c$

$y = m \cdot 0 + c \Rightarrow c$

$\boxed{y = c}$

Cost function is made.

↳ $\frac{1}{2m}$ ↳ number of pts.

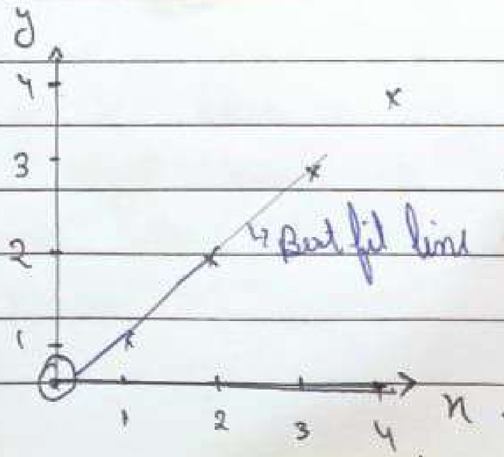
$$\boxed{\frac{1}{2m} \sum_{i=1}^{m} (\hat{y} - y)^2 .}$$

$\hat{y} \Rightarrow y = mx + c$

↳ the points on best fit line (Predicted)

y → Preexisting points. wala line

Example.



For BFL.

→ $y = mx + c = 0$

↳ must pass through origin

↳ $y = mx$

$x = 1$

if $m = 1$ ⇒ $\hat{y} = 1 \cdot 1$

$x = 2$, $m = 1$ → $\hat{y}$ → 1 ·

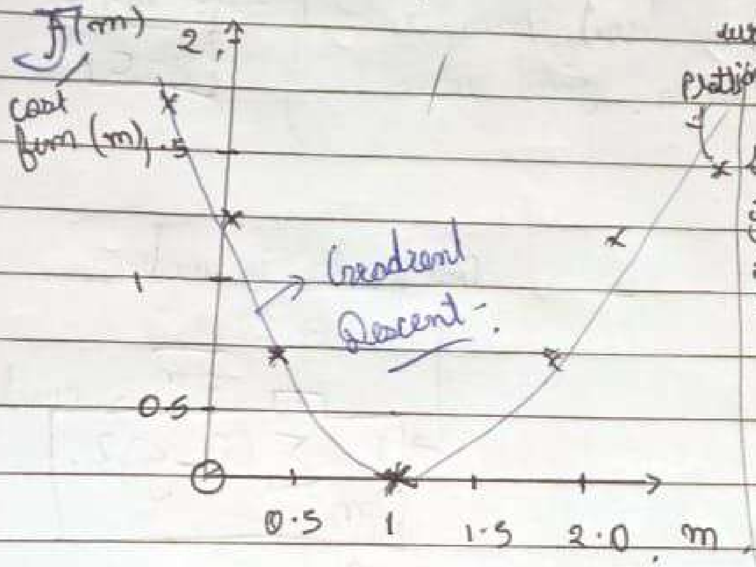
Cost function,

$$= \frac{1}{2m} \sum_{i=1}^{m} (\hat{y} - y)^2 .$$

$$\Rightarrow \frac{1}{2m} \left((1-1)^2 + (2-2)^2 + (3-3)^2 \right.$$

$$\Rightarrow 0 \text{ = cost.}$$

Eq2. Cost function slope.



with $m = 1$, $cf = 0$

putting ~~$\hat{y} = 0.5$~~

say $m = 0.5$

$\hat{y}$ for $x = 1$

$\hat{y} \Rightarrow 0.5$

for $x = 2$

$\hat{y} = 1$

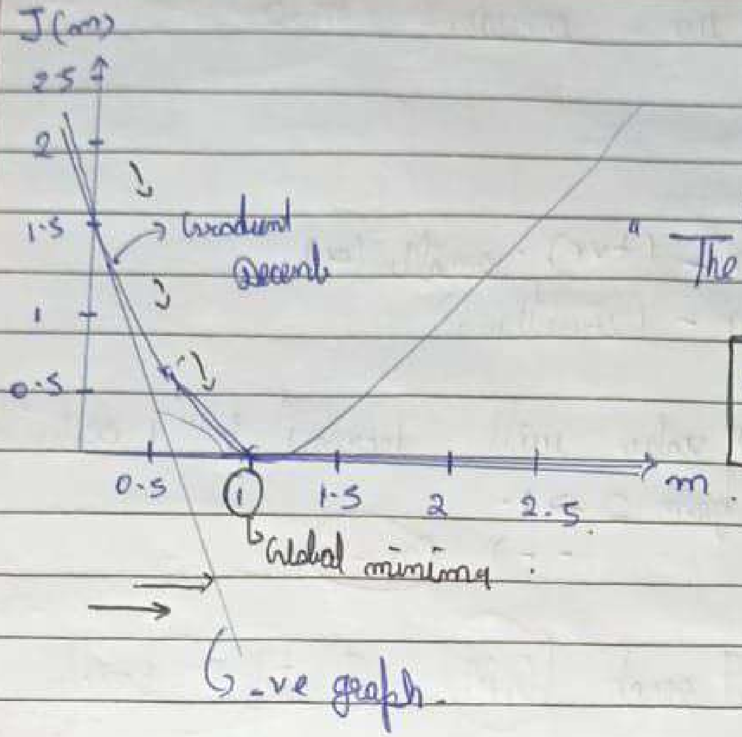
for $x = 3$

$\hat{y} = 1.5$

$$\frac{1}{2m} \sum_{i=1}^{m} \left((0.5-1)^2 + (1-2)^2 + (1.5-3)^2 \right)$$

$$\Rightarrow 0.58$$

PTO.



To reach Global minima we have to follow a Theorem to go down from LHL "The Convergence Theorem".

$$\left[m = m - \left(\frac{\partial J}{\partial m}\right) \times \alpha \right]$$

func: $\frac{\partial J}{\partial m}$ ↳ slope.

α ↳ learning rate.

applying Convergence theorem.

$$m = m - \left(\frac{\partial m}{\partial m}\right) * \alpha \qquad \alpha \approx 0.001$$

$$m = m - (-ve) * (\text{small})$$

$$m = m + (+ve) * \text{smaller}$$

↳ This suggests that value of m will be +ve & tend to come nearer to 1 or Global minima.

* If α is taken as larger value say, 1. it might jump up the result of m & never allow it to reach the Global minima.

Similarly, if we consider $m = 2$.

$m = +ve$

$m = m - (+ve) \cdot \text{smaller}(\alpha)$

$m = m - (\text{something})$

m value will decrease to 1 certainly from $2 \Rightarrow 1$.

(- -) (m - -)

The time I reach GM I stop searching or training.

Say we have more than one feature in model. our Gradient decent would be a 3-D figure. Thus we get a complex answer.