→ Linear Regression (LR)

→ Statistical learning Algo, not a ML Algo

→ The output of a LR model is always a continuous variable

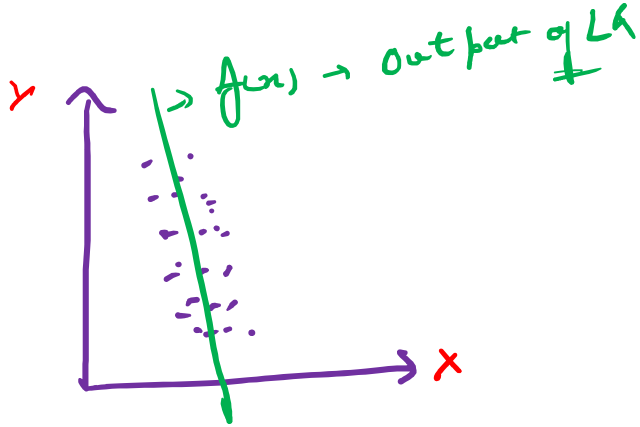
→ The target will be numerical → continuous | e.g. salary, stock price etc.

→ It tries to find a straight line, that fits best to your data

straight → linear

fits → passes

$$\boxed{y = mx + c}$$ → Mathematical Rep of a straight line.



LR

mathematical equation

Decision Tree

Classified function

O/P of any ML Algo

ML Model

Math eq

classification f^n

→ Std. dev (σ) → Variance (σ^2) → Co-variance (cov(x,y))

Avg dist from μ

(1-var)

Avg sq. dist from μ

(1-var)

dxn of relⁿ

+ve, -ve, 0

(2-var)

Co-relation (r, ρ) → Linear Regression

dxn + strength of the relⁿ

(2-var)

dxn + strength + Relⁿ

(2-var)

eg:

3x+4

5x-2

1.5x + 2.78

$$y = f(x)$$

Target features

ML objective

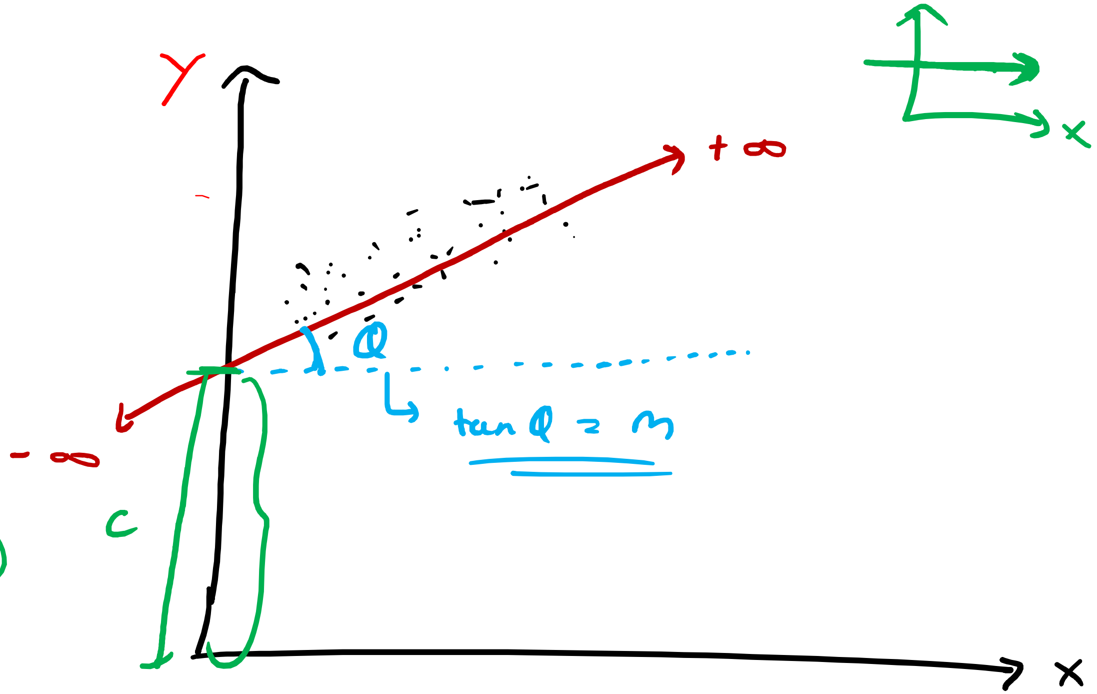
$\sin(\theta) \Rightarrow (-1, +1)$

$\cos(\theta) \rightarrow (-1, +1)$

$\tan(\theta) \Rightarrow [-\infty, +\infty]$

$\tan 0^\circ = 0$, $\tan(90^\circ) \rightarrow +\infty$

$\tan(270^\circ) = \tan(-90^\circ) = -\infty$



the range of output will always be $[-\infty, +\infty]$

$$y = mx + c$$

Target

features

Intercept

The pt. where the line cuts the y-axis

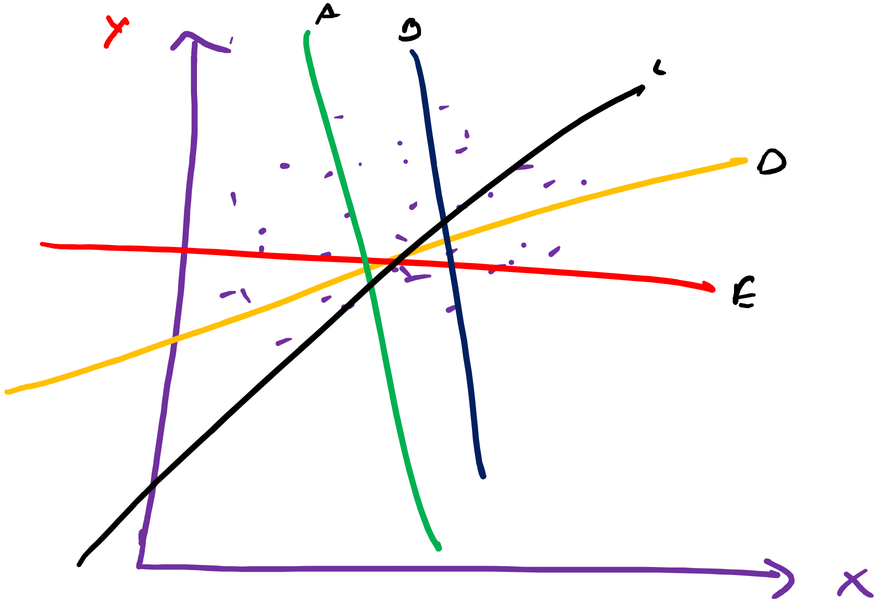
Slope = $\tan\theta$

the angle the line is making wrt x-axis

If $m = 0 \Rightarrow$ the line is parallel to x-axis

If $m > 0 \rightarrow$ +ve dxn

$m < 0 \Rightarrow$ -ve dxn



Multiple lines can pass thru the data

Inf lines → passing →

↳ 1 Best line

⇓

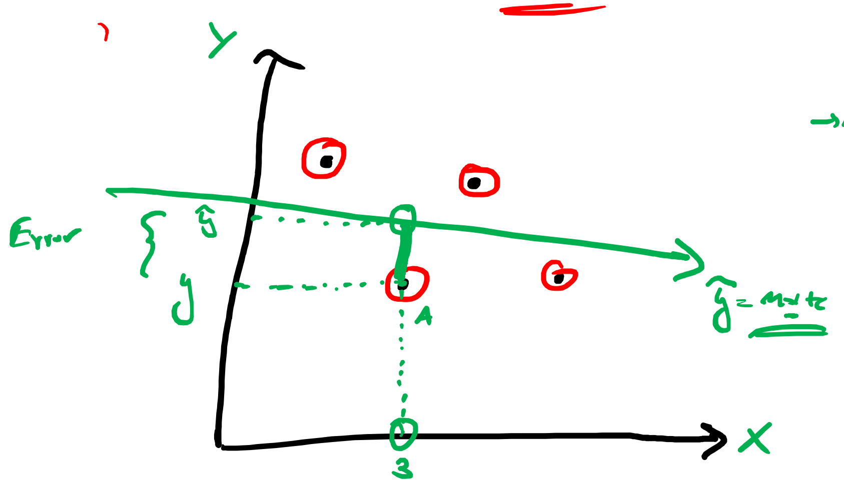
→ Best fit line

↓

→ $\boxed{\text{Least Error}}$

minimum error

$$\boxed{\text{Error} = \underline{y} - \underline{\hat{y}}}$$



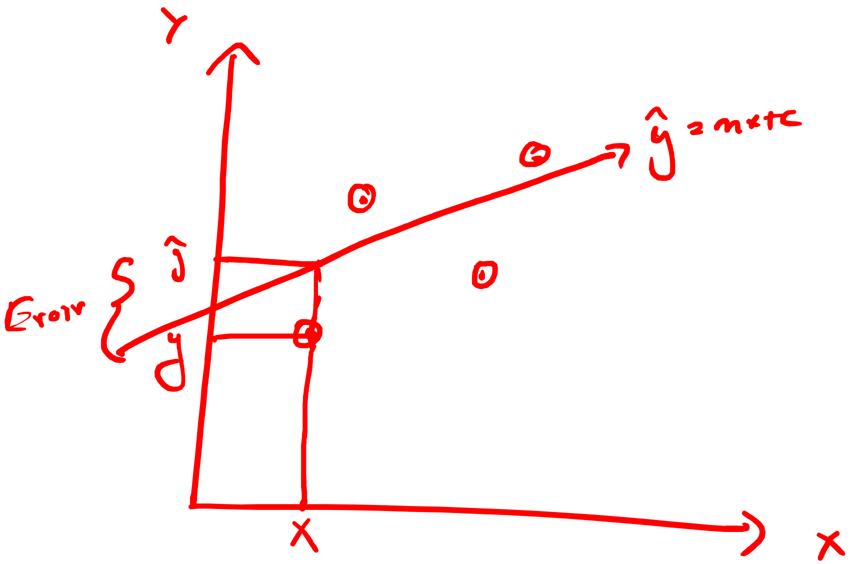
	X	X
→A	3	2

$(x, y) \rightarrow \text{LR} \rightarrow \hat{y} = mx + c$

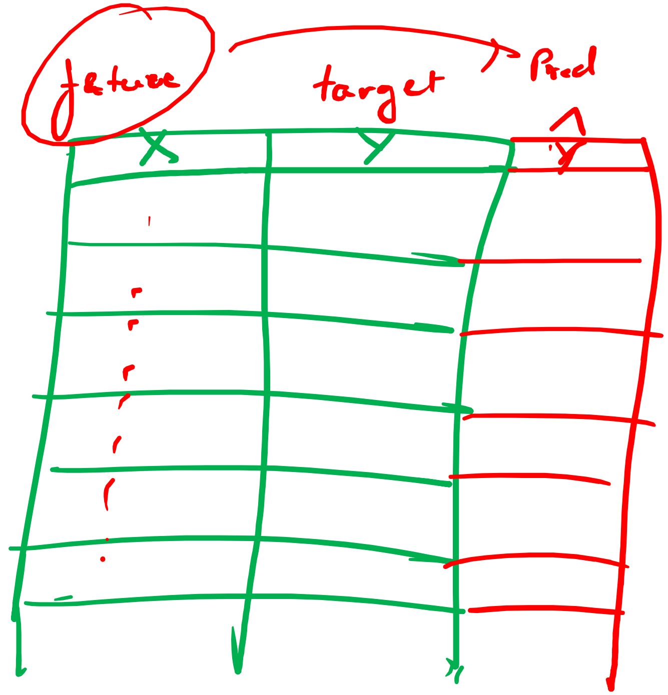
model

$\hat{y} = 3x + 3$

$\hat{y} = -2x + 1$



$$\boxed{\text{Error} = y - \hat{y}}$$

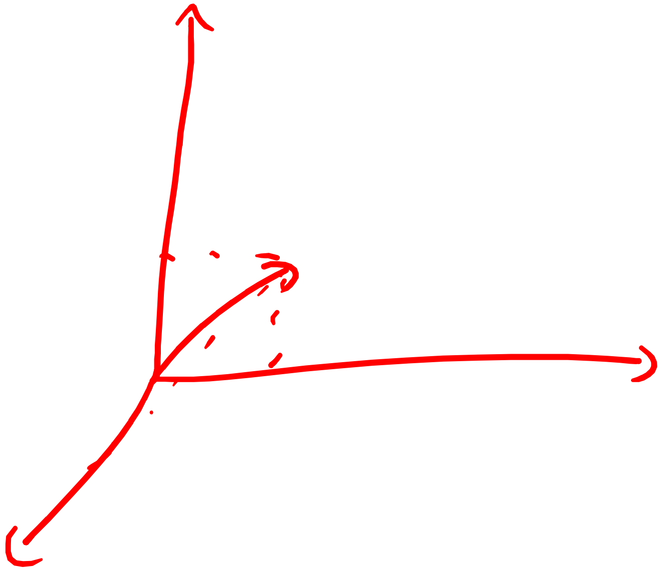


X_{train} Y_{train}

X_{test} Y_{test}

y_{pred}

100, 1m, 2/3



theory

$\hat{y} = m x_1 + c$ → 2D

$\hat{y} = m_1 x_1 + m_2 x_2 + c$ → 3D

$\hat{y} = m_1 x_1 + m_2 x_2 + m_3 x_3 \ldots$

$m_n x_n + c$

N → n

Sales = 2.TV + 3 N + 1.5 Radio + 100

AWS, 1m → 690x + 1400

10

6900 + 1400 = $\hat{y}$

$y > \hat{y}$ = +ve

$y < \hat{y}$ = -ve

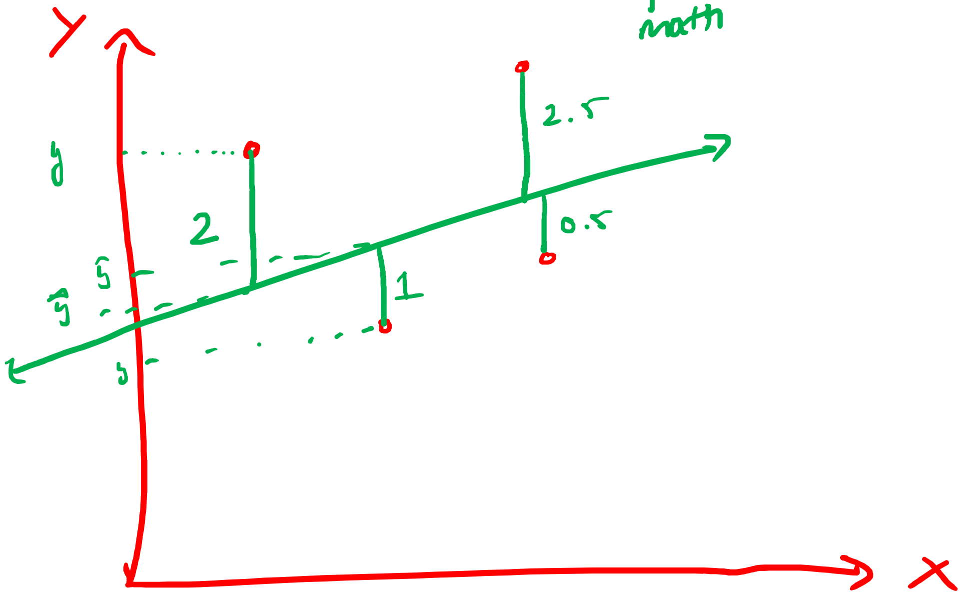
$\boxed{\text{Error} = y - \hat{y}}$

↑ math

Total Error = $\sum (y - \hat{y})$

$\Rightarrow 2 + (-1) + 2.5 + (-0.5)$

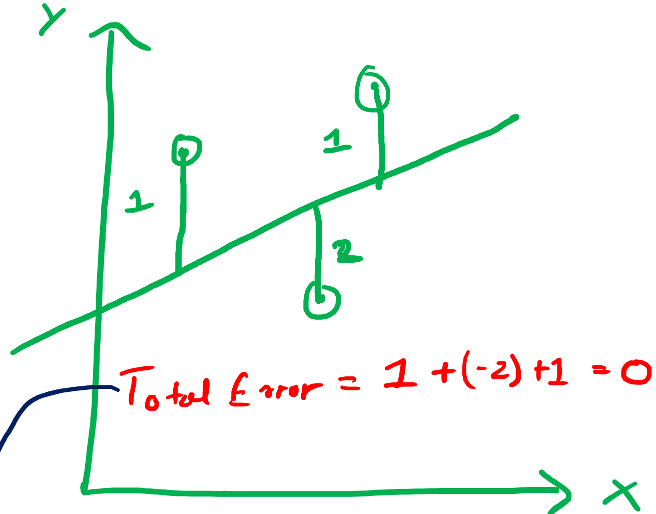
$\boxed{\Rightarrow 3}$



Error fns

How to get rid of (-ve) sign?

1. Modulus = $\sum |y - \hat{y}|$ = Absolute Error (SAE)
2. Square = $\sum (y - \hat{y})^2$ = SSE, Sum of Sq. Errors
3. Root of Square = $\sqrt{\sum (y - \hat{y})^2}$ = RSE. root of Sum of Sq. Error



Total Error = $1 + (-2) + 1 = 0$

Is this logically correct?

logically Incorrect

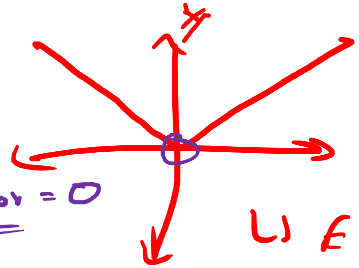
because of (-ve) sign

Error fn = Cost fn = $h(\theta)$ $\quad$ $r, \rho, \sigma, \alpha, \mu$

$h(\theta) = SAE = \sum |y - \hat{y}|$ → degree = 1 → linear → straight line

$\frac{1}{2} * 2$

$h(\theta) = RSE = \sqrt{\sum (y - \hat{y})^2}$ → degree = 1



$y = |x|$

min error = 0

L1 Error

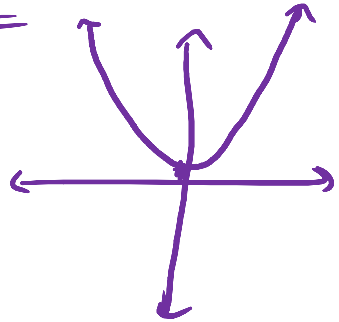
Best fitt line is minimum error

To calc the best line

$$\boxed{h(\theta) = SSE = \sum (y - \hat{y})^2}$$ ⇒ degree = 2 ⇒ parabola

$y = |2x|$

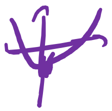
$4x$



$y = x^2$

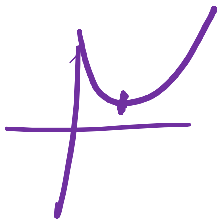
$y = 3x^2$

$(3x+2)^2$



(m, c)

The best line



→ degree = 1 → Abne Reeck

degree - 2 : $y = 2x^2 + 3$

$4x^2 - 9$

$y = ax^2 + b$

SSE

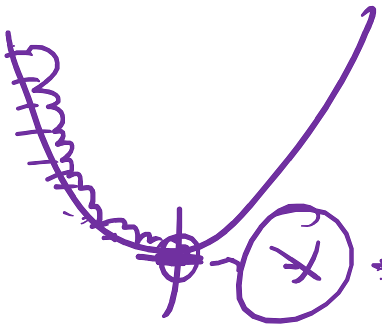
$\frac{dy}{dx} = 2ax$

$x = 100$

2000

$x = 4$

8



x → m, c

$y = 2x + 3$

$4x - 9$

$y = ax + b$

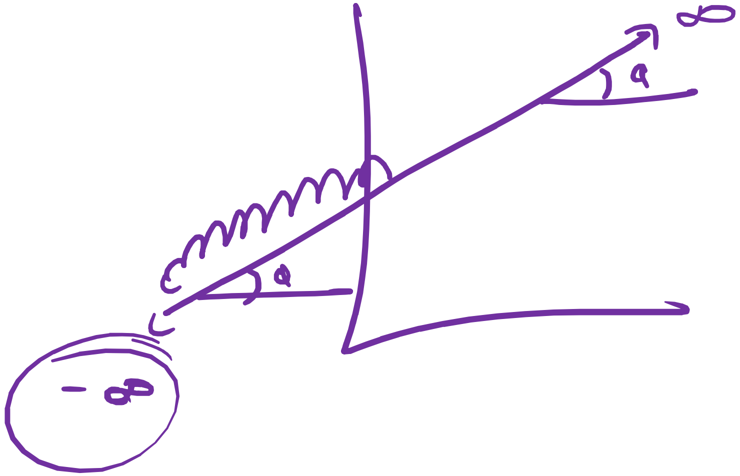
$\frac{dy}{dx} = a$

↳ Slope

$x = \infty$

$x = 100$

$x = -1.34$



→ Gradient Descent (Linear Regression)

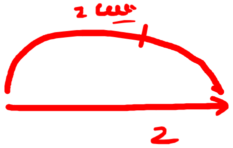
$$\hat{y} = mx + c$$

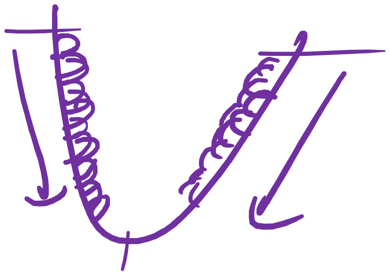
$$h(\theta) = \underset{\downarrow}{SSE} = \Sigma (y - \hat{y})^2$$

$$MSE = \frac{1}{n} \Sigma (y - \hat{y})^2$$

$$h(\theta) = \frac{1}{N} \Sigma (y - (mx + c))^2$$

$$\boxed{h(\theta) = \frac{1}{N} \Sigma (y - mx - c)^2}$$



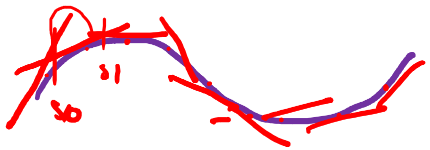


Gradient Descent
↓ ↓
Steepness down

$S_1 = S_0 + \text{step size}$

$S_2 = S_1 + \text{step size}$

moving on a straight path



$S_1 = S_0 + (\text{step size} * \text{slope})$

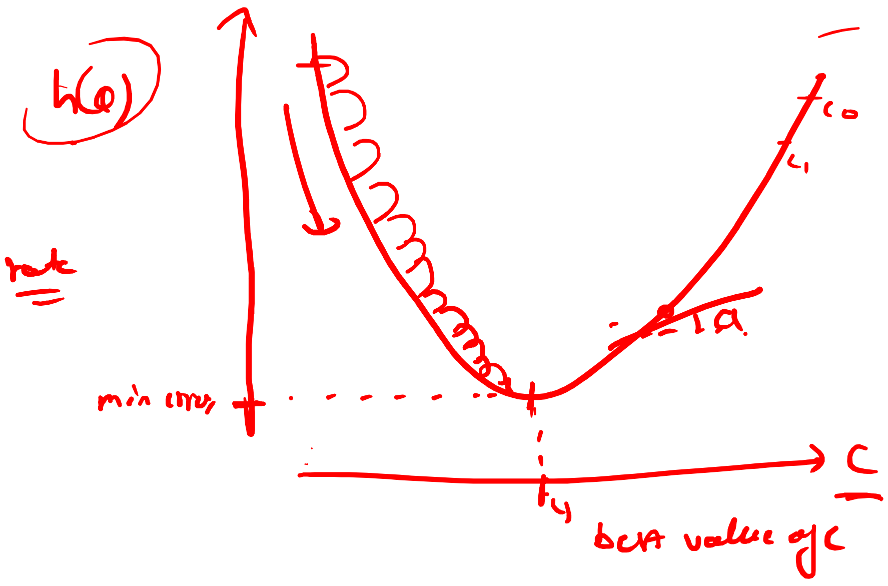
move on a curved path

$C_0 =$ Arbitrary value

$C_1 = C_0 - \text{stepsize} * \text{slope}$ ($\text{stepsize} \to \alpha$, $\text{slope} \to \dfrac{d\,h(\theta)}{dc}$)

$$\boxed{C_2 = C_1 - \alpha * \frac{d\,h(\theta)}{dc}}$$

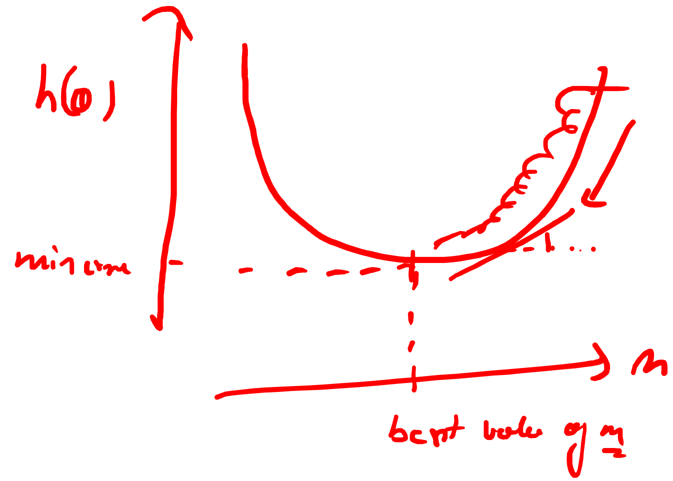
α = step size = learning rate



Any

$m_0 =$ Arbitrary value

$$\boxed{m_1 = m_0 - \alpha * \frac{d\,h(\theta)}{dm}}$$



$$h(\theta) = \frac{1}{N} \sum (y - mx - c)^2$$

$$\frac{d\,(h(\theta))}{dm} = \frac{1}{N} (-x).(2) \sum (y - mx - c) = \frac{-2x}{N} \sum (y - mx - c) = \frac{-2x}{N} \sum (y - \hat{y})$$

$$\frac{d(h(\theta))}{dc} = \frac{1}{N} (2).(-1) \sum (y - mx - c) = \frac{-2}{N} \sum (y - mx - c) = \frac{-2}{N} \sum (y - \hat{y})$$

α - Step size → Learning rate

$\alpha = 1$

$\alpha = 0.1$

m_0

m_1

m_1

$\vdots$

$m_n - \infty$

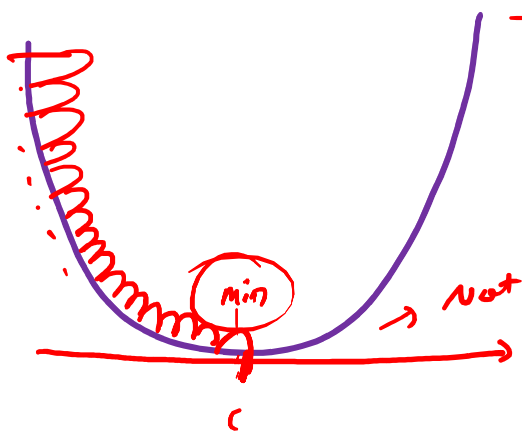
m_0

m_1

m_2

$\vdots$

m_{15} → min

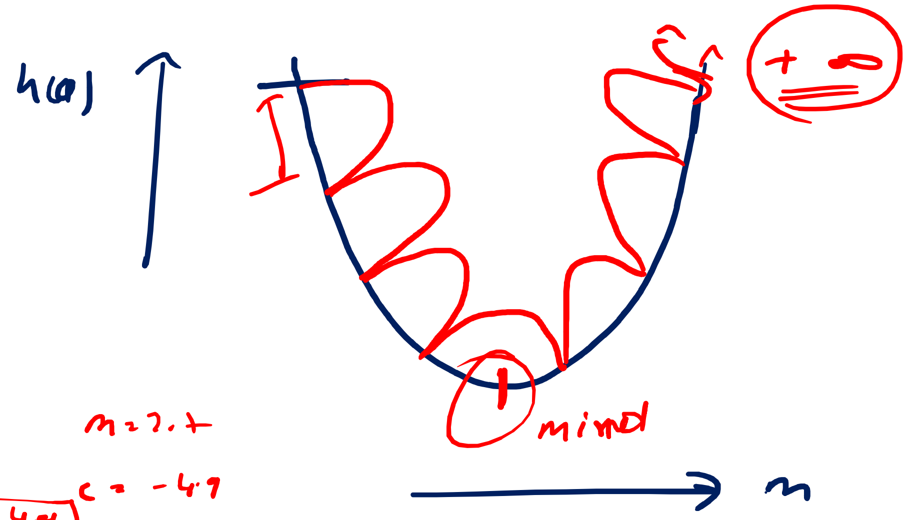


$\hat{y} = m x + c$

$m = 2.7$

$c = -4.9$

$\boxed{\hat{y} = 2.7x - 4.9}$



While moving If we reach ∞

it means the minima is missed

→ Not able to move forward → Convergence

Slope = 0

$m_{18} = m_{17} - \alpha * \text{slope}$

$m_{18} = m_{17}$

$m_{19} = m_{17}$

m_{21} $m_{100} = m_{17}$

$m_1 = m_0 - \alpha \cdot \text{slope}$