# Simple Linear Regression (Supervised Machine Learning Algorithm)

* Regression :- To establish a relationship b/w the two Variables/more than two variables.

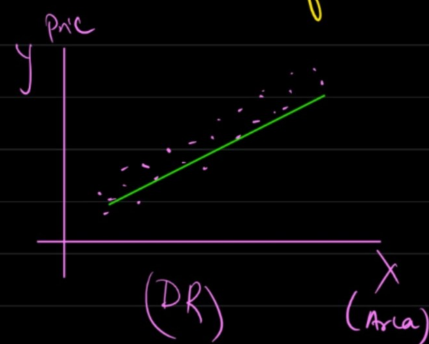
Supervised ML
- Regression
- Classification

* Linear → It establish a linear relationship.

* SLR attempts to determine the strength and characteristics of the relationship between one dependent Variable (y) and another variable independent Variable (X)

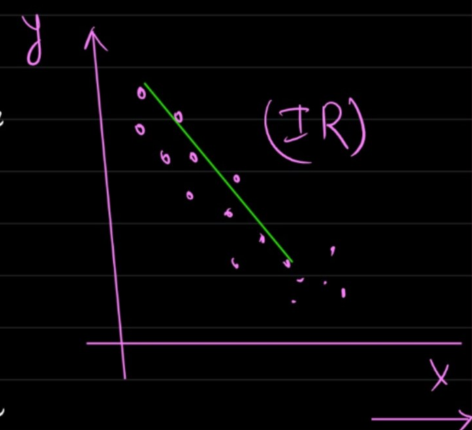
Predict Price of house based on Area of house.

| Area of house | y |
|---|---|
| 1100 | 50 |
| 1200 | 60 |
| 1300 | 70 |
| — | — |
| — | |

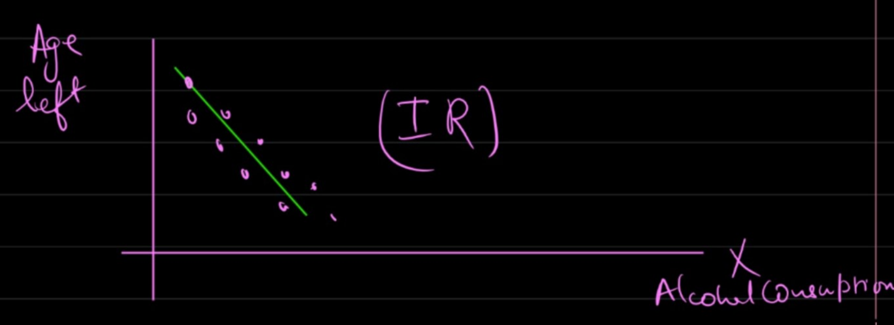



eg Selling Price of Car wrt age of car

| X age of car | Y Selling Price |
|---|---|
| 10 | 3.1 |
| 9 | 4.1 |
| 7 | 5.3 |
| — | — |
| — | |




eg Age left Vs Alcohol consumption



Simple → Only one X and y

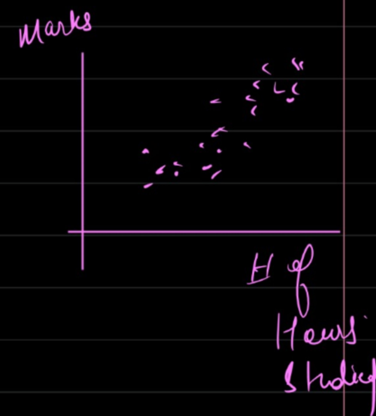
Linear → Captures the linear relationship

Regression → Understand relationship.

* if multiple IV (multiple X's) → multiple linear regression.

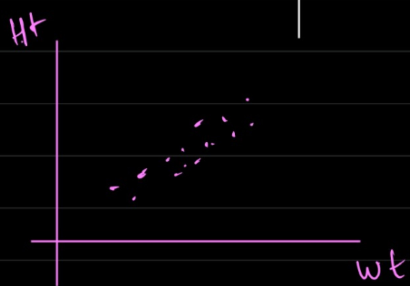
* Example

| No of Hours studied | marks obtained |
|---|---|
| 8 | 72 |
| 9 | 85.5 |
| – | – |
| – | – |
| – | – |




example

| Weight | Height |
|---|---|
| 65 | 168 |
| 70 | 172 |
| – | – |
| – | – |




* To predict price of a house based on number of room

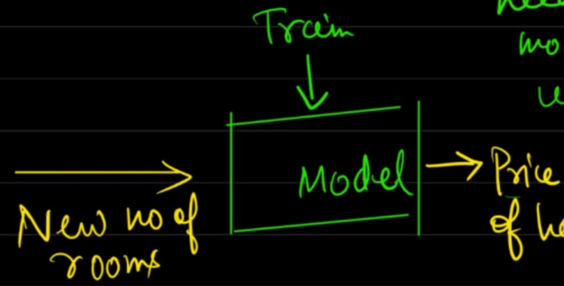
$y$ = price of house

$X$ → Number of rooms.

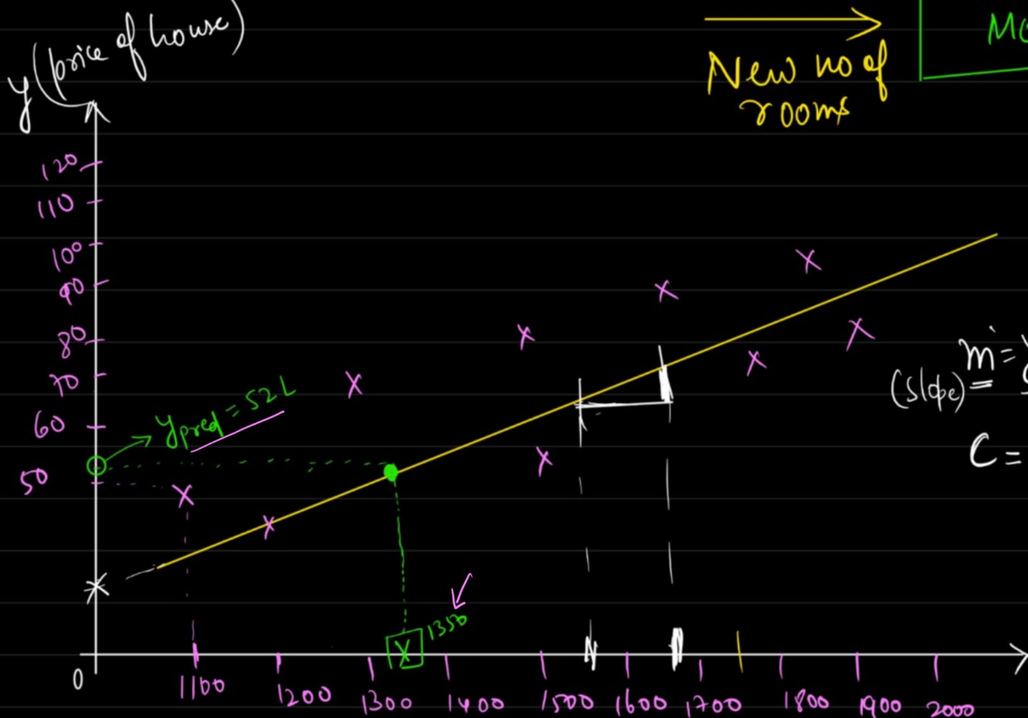
| # of rooms (x) | Price of house (y) |
|---|---|
| 1 | 50 |
| 1.5 | 60 |
| 2 | 65 |
| 3 | 75 |
| 4 | ? → 85L |

⟹ ML — To understand the mathematical relationship.

→ This was 4 dp's that's why you understood the trend. If many dp's you




need to train model to understand the mathematical relationship




Area of house (X) (sq ft)

$\hat{y} = mx + c$

$\hat{y} \rightarrow$ estimator

(you have estimated the price of house base on Area of house, m, c.)

$$(\text{slope}) = m = \frac{y_2 - y_1}{x_2 - x_1}$$

C = intercept (where it cuts Y axis; $x = 0$)

$\hat{y} = mx + c$

$\Downarrow$

$\hat{y} = \beta_0 + \beta_1 X$ $\quad (\beta_1 = m,\ \beta_0 = c) \rightarrow$ Coefficients)

$\boxed{h_\theta(x) = \theta_0 + \theta_1 x}$ $\quad (\theta_1, \theta_1 \rightarrow$ Coeffict)

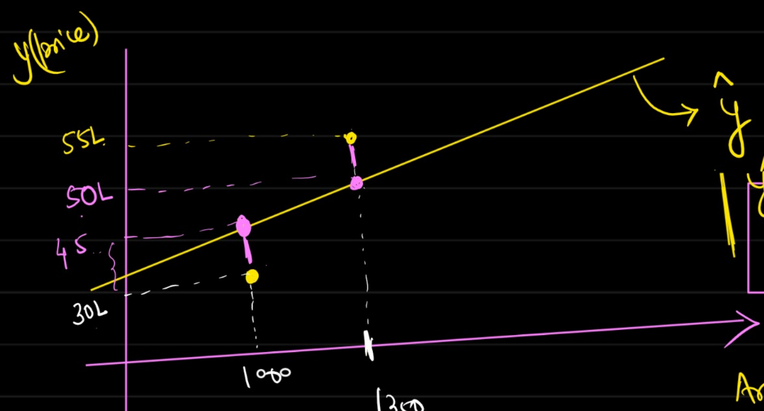
$\downarrow$

many book, you will find this format

$\rightarrow$ because we consider a S.L.R line to be hypothesis that it passes from these points

$X \rightarrow$ is dp

$\underbrace{(m, c), (\beta_0 \beta_1), (\theta_0, \theta_1)}_{\text{Coefficients.}}$




Area of house (X)

$\hat{y} = mx + c$

$\hat{y} = \beta_0 + \beta_1 X$

$\boxed{h_\theta(x) = \theta_0 + \theta_1 x}$

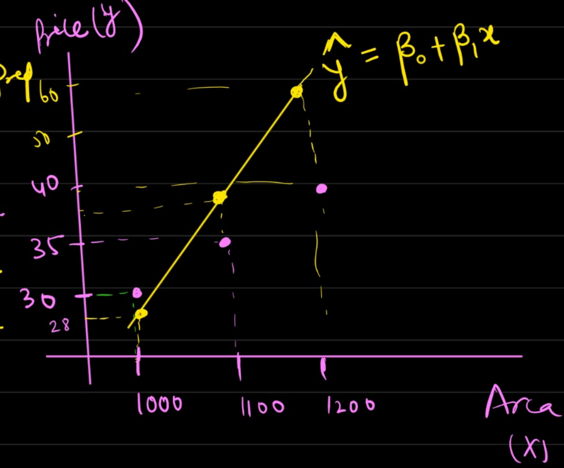
$y_{actual}$ for 1350 sq = 55 L

$y_{pred}$ for 1350 sq = 50 L.

Error = $y_{actual} - y_{pred}$

*

| X (Area of house) | y (Price of house) (y act) $y_i$ | $\hat{y}$ pred |
|---|---|---|
| 1000 | 30L | 28L |
| 1100 | 35L | 38L |
| 1200 | 40L | 60L |




Error = $y_{actual} - y_{pred}$

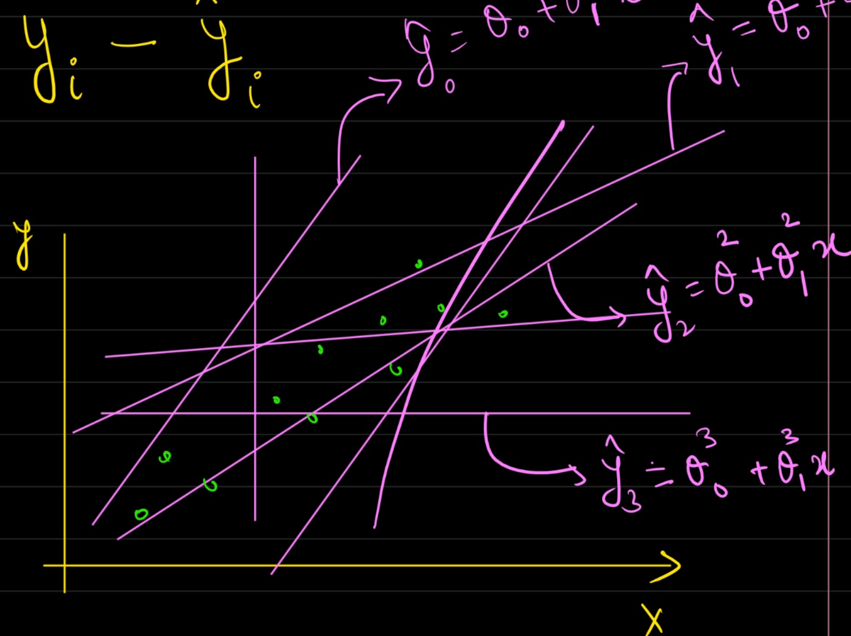
Error = $y_i - \hat{y}_i$

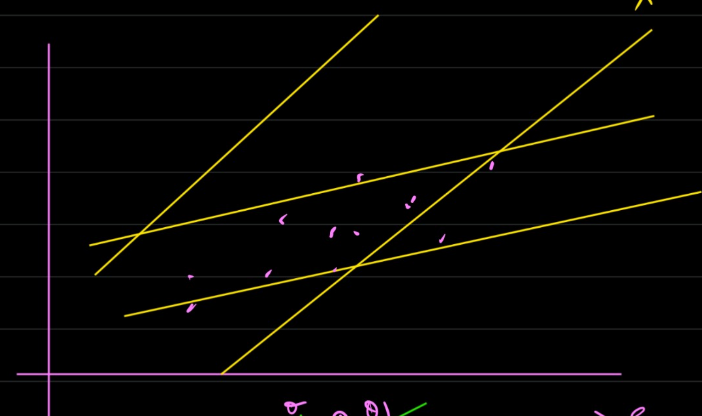
$\hat{y}_1$

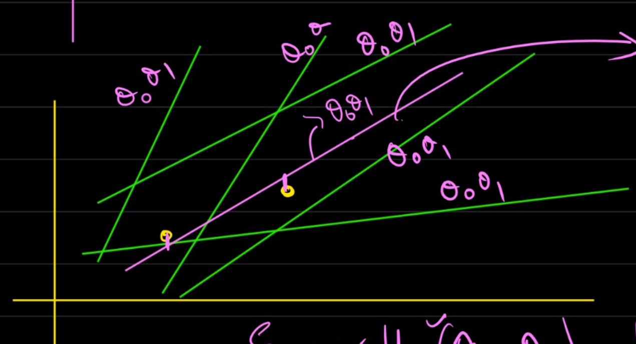
$\hat{y}_2$

$\hat{y}_3$

$\hat{y}_4$






→ Error will be very less.

Error ⇓ $(\theta_0, \theta_1)$, $(\beta_0 \beta_1)$

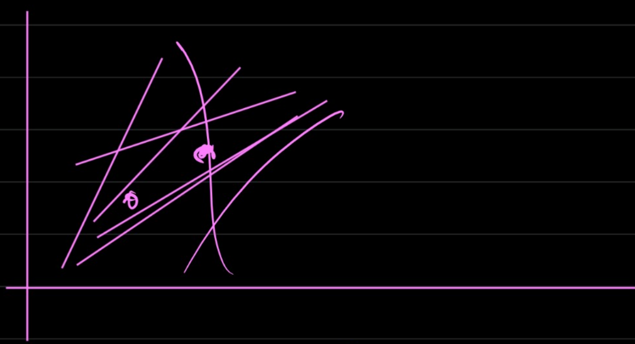
optimal.

for all lines, we want

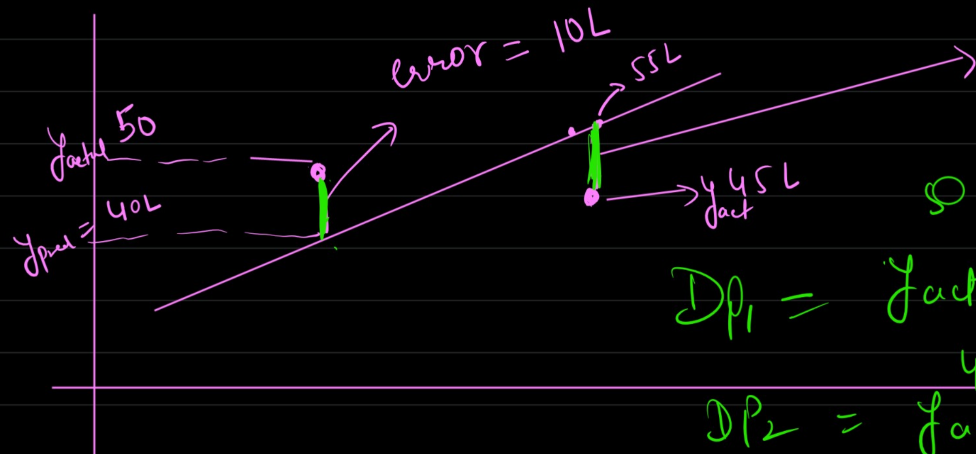
Error = $y_{act} - y_{pred}$ to be least.

$\Downarrow$

Optimation process.



$$\text{Squared error} = \left(y_{act} - y_{pred}\right)^2$$



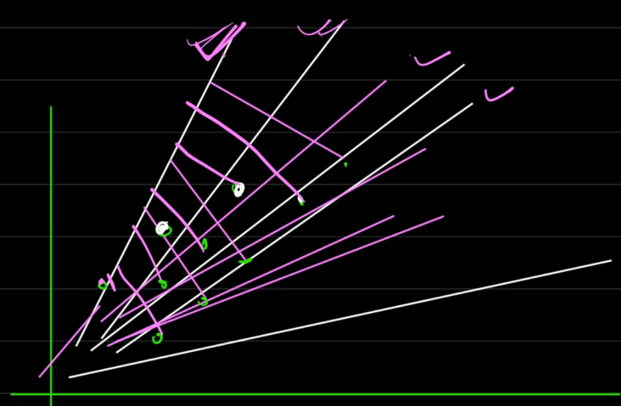

50 − 40

$DP_1 = y_{act} - y_{pred} = 10L$

45L − 55

$DP_2 = y_{act} - y_{pred} = -10L$

for all the dp's = $\underline{10L} + \underline{(-10L)} = 0$



$10^2 + (-10)^2$

$= 100 + 100 = 200$

$$= \min\left(y_{act} - y_{pred}\right)^2$$

$$= \min \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2$$

$y_{actual}$ $\quad y_{predict} = \min \sum_{i=1}^{n} \left(y_i - h_\theta(x)\right)^2$

$$\boxed{\min \sum_{i=1}^{n} \left(y_i - h_\theta(x)\right)^2}$$

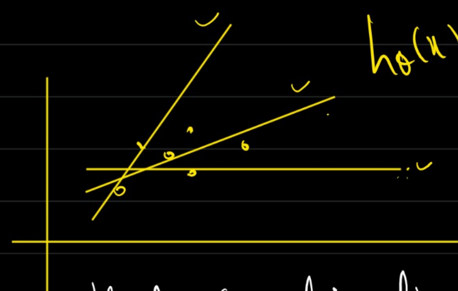
$n$ is no. of datapoint

$y_i$ - $y$ actual

$h_\theta(x) \rightarrow y$ predicted

least square method
or
Ordinary least square method.

$h_\theta(x) = \theta_0 + \theta_1 x$



* We want to have that combination of $m, c$ $(\theta_0, \theta_1)$, $(\beta_0, \beta_1)$ where the overall error is least and the line will be best fit line.

$$\min \sum_{i=1}^{n} \left(y_i - h_\theta(x)\right)^2$$

$$\Rightarrow \min \sum_{i=1}^{n} \left(y_i - (\theta_0 + \theta_1 x)\right)^2$$

$$\begin{Bmatrix} \theta_0^1 & \theta_1^1 \\ \theta_0^2 & \theta_1^2 \\ \theta_0^3 & \theta_1^3 \\ - & - \\ - & - \\ - & - \\ \theta_0^\infty & \theta_1^\infty \end{Bmatrix} \rightarrow \text{optimal } (\theta_0, \theta_1)$$

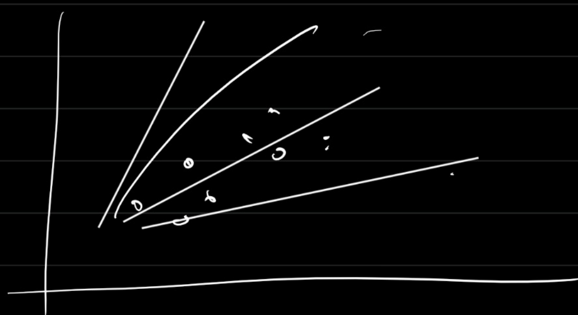
IQ
Technique to find best fit line in SLR
$\rightarrow$ OLS.

S.S.E -

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

Mean squared Error = $\min \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

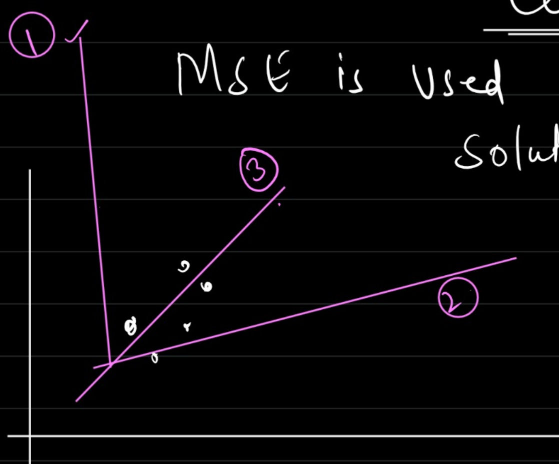
$\frac{1+2+3}{3}$ → mean



* For best fit line MSE should be least

Penalty

Cost function

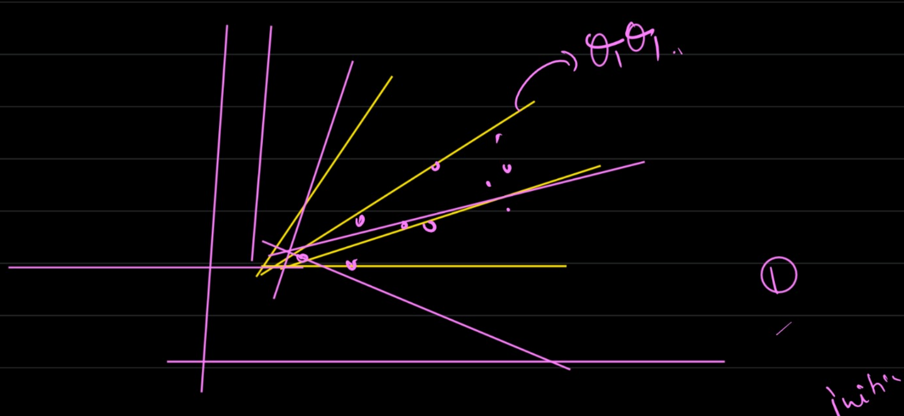


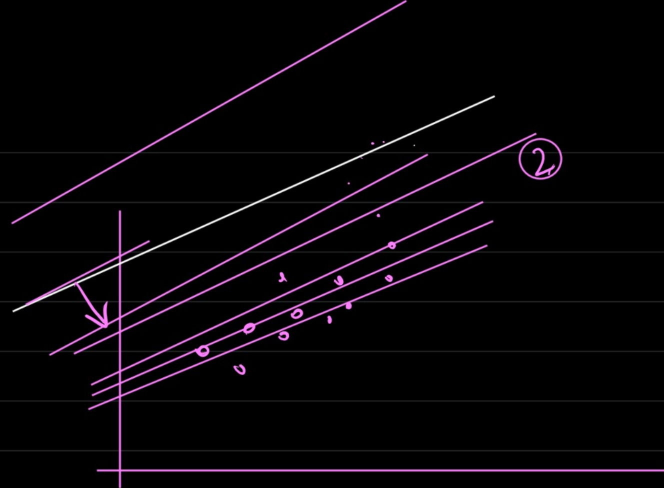

MSE is used as CF to get solution → $\beta_0, \beta_1$ $(\theta_0, \theta_1)$

Cost Function

$$J(\theta_0, \theta_1) = \frac{1}{n} \min \sum_{i=1}^{n} (y_i - h_\theta(x))^2$$

$$h_\theta(x) = (\theta_0 + \theta_1 x)^2$$

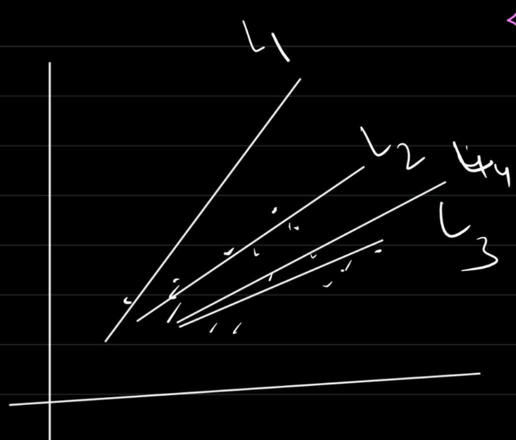






To get optimal line

1. choose righ combination of line
2. Find Mean Square Error
3. best fit line




$L_1(m_1 c_1)$ — $MSE_1$

$L_2(m_2 c_2)$ — $MSE_2$

$L_3(m_3 c_3)$ — $MSE_3$

$\vdots$

$L_\infty$ — $MSE_\infty$

$\Leftarrow$ MSE least

$\Downarrow$

best fit line

(optimal $\beta_0, \beta_1$)

Optimization process $\Rightarrow$ Gradient descent