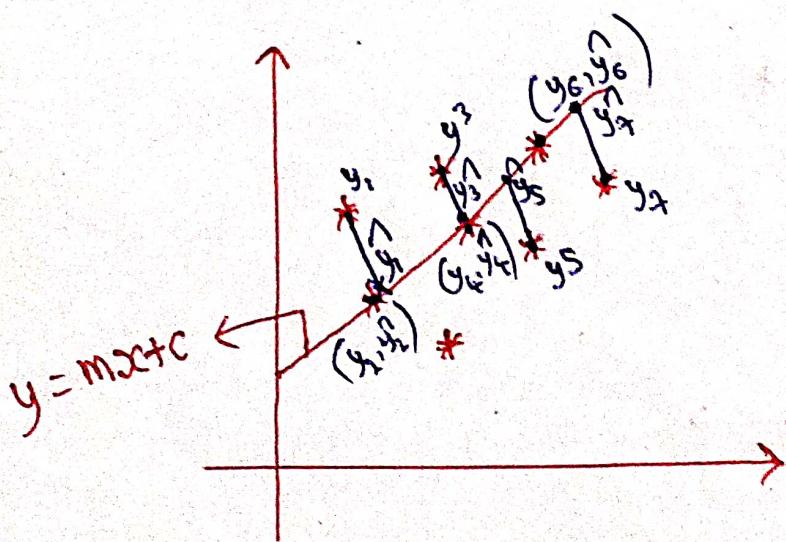


ML Algorithms

* Simple Linear Regression:

- * It is one of the Main Algorithm in Machine Learning, which is used for Supervised Learning.
- * It can be used, if we have one Independent and one Dependent feature.
- * It tries to find the best fit line which will have low cost of error from all the points.



where,
 $y_1, y_2, y_3, y_4, y_5, y_6$ and y_7
are original points
And
 $\hat{y}_1, \hat{y}_2, \hat{y}_3, \hat{y}_4, \hat{y}_5, \hat{y}_6$ and \hat{y}_7
are fitted points

* The best fit line is represented by the formulae, $y = mx + c$

where,

y is predicted point (Dependent point)

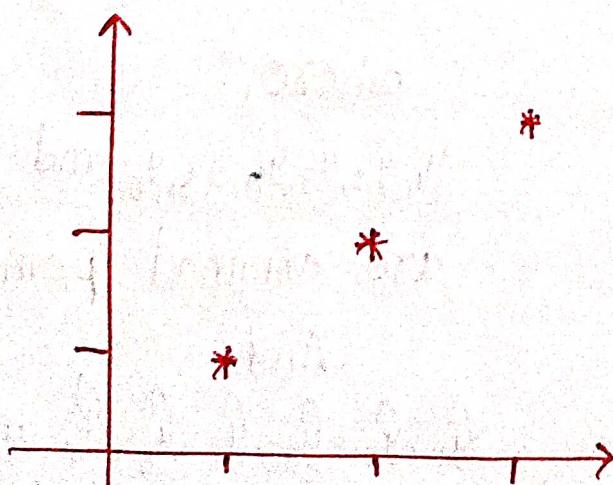
x is given point (Independent point)

m is slope

c is Intercept (Point of y at which Line meets x at zero)

Procedure of finding best fit line:

Let's take the points $(1, 1)$, $(2, 2)$ and $(3, 3)$



* Let's find y by using $mx+c$

case 1: Let's consider $m=1$ and $c=1$

we have x as 1, 2 and 3

$$y_1 = m(x) + c \\ = 1(1) + 1 \\ = 2$$

$$y_2 = m(x) + c \\ = 1(2) + 1 \\ = 3$$

$$y_3 = m(x) + c \\ = 1(3) + 1 \\ = 4$$

we got \hat{y} values as 2, 3 and 4.

our original values of y are 1, 2 and 3.

Let's find cost function using the formulae:

$$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \text{ where } n = \text{No of points}$$

y_i = original y value.

\hat{y}_i = fitted y value.

$$\Rightarrow \frac{1}{3} \sum_{i=1}^3 (y_i - \hat{y}_i)^2 \Rightarrow \frac{1}{3} \times ((2-1)^2 + (3-2)^2 + (4-3)^2)$$

$$\Rightarrow \frac{1}{3} \times (1+1+1) \Rightarrow \frac{3}{3} \Rightarrow ①$$

i.e., at $m=1$ and $c=1$,

we got cost function as 1 i.e., the error is,

case ② :

Let's take $m=0$ and $c=1$

$$\begin{aligned}y_1 &= m(x) + c & y_2 &= m(x) + c & y_3 &= m(x) + c \\&= 0(1) + 1 & &= 0(2) + 1 & &= 0(3) + c \\&= 1 & &= 1 & &= 1\end{aligned}$$

$$\begin{aligned}\text{cost Function} &= \frac{1}{n} \left((1-1)^2 + (2-1)^2 + (3-1)^2 \right) \\&= \frac{1}{3} (0 + 1 + 4) \\&= 5/3\end{aligned}$$

case ③ :

Lets take $m=1$ and $c=0$

$$\begin{aligned}y_1 &= m(x) + c & y_2 &= m(x) + c & y_3 &= m(x) + c \\&= 1(1) + 0 & &= 1(2) + 0 & &= 1(3) + 0 \\&= 1 & &= 2 & &= 3\end{aligned}$$

$$\text{cost function} = \frac{1}{3} ((1-1)^2 + (2-2)^2 + (3-3)^2)$$

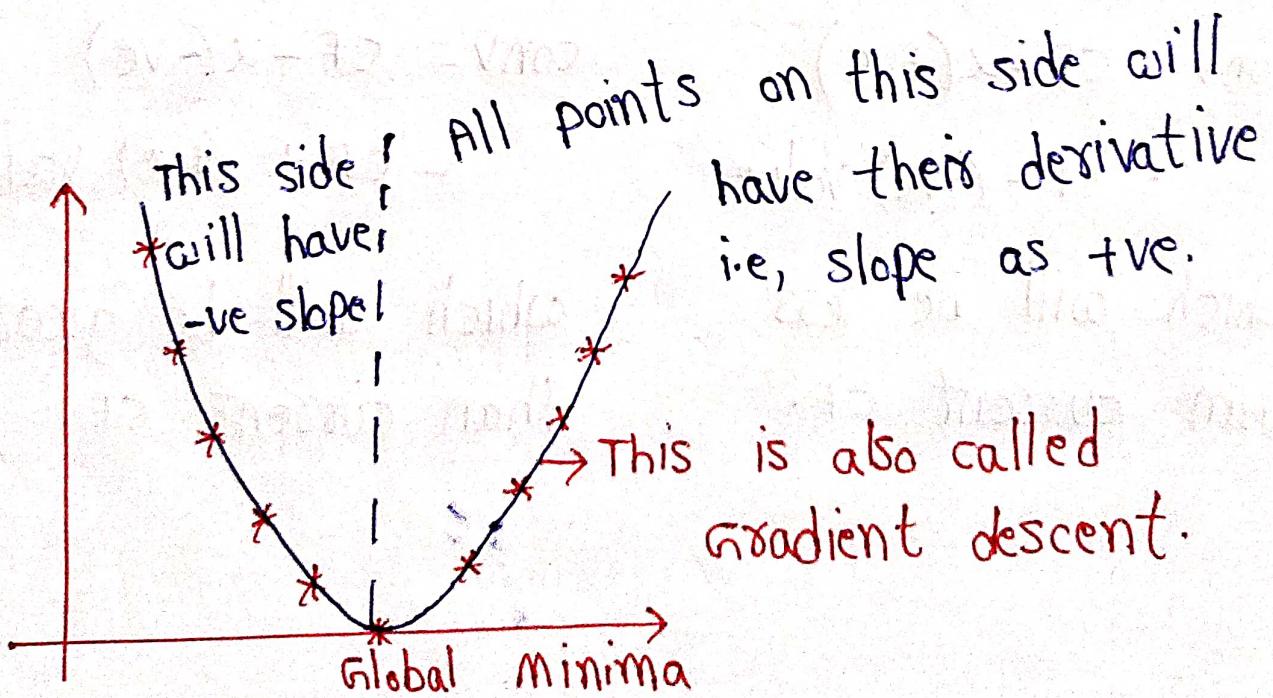
$$= 0/3$$

$$= 0$$

→ out of case 1, case 2 and case 3, we got global minima cost function at $m=1$ and $c=0$ where we got 'zero'

* So while predicting $y = mx + c$, we can take $m=1$ and $c=0$ for better predictions.

How all the cost functions looks on graph:



* while getting cost functions, we will use something called as convergence to get the global minima point.

convergence \Rightarrow $CF - \alpha(\text{slope})$

where CF = current cost function.

α = Learning Rate.

slope = Derivative.

If we have +ve derivative
If we have -ve derivative

$$\text{conv} = CF - \alpha(+\text{ve})$$

$$= CF - (-\text{ve}) \text{ value}$$

which will be less than current CF.

$$\text{conv} = CF - \alpha(-\text{ve})$$

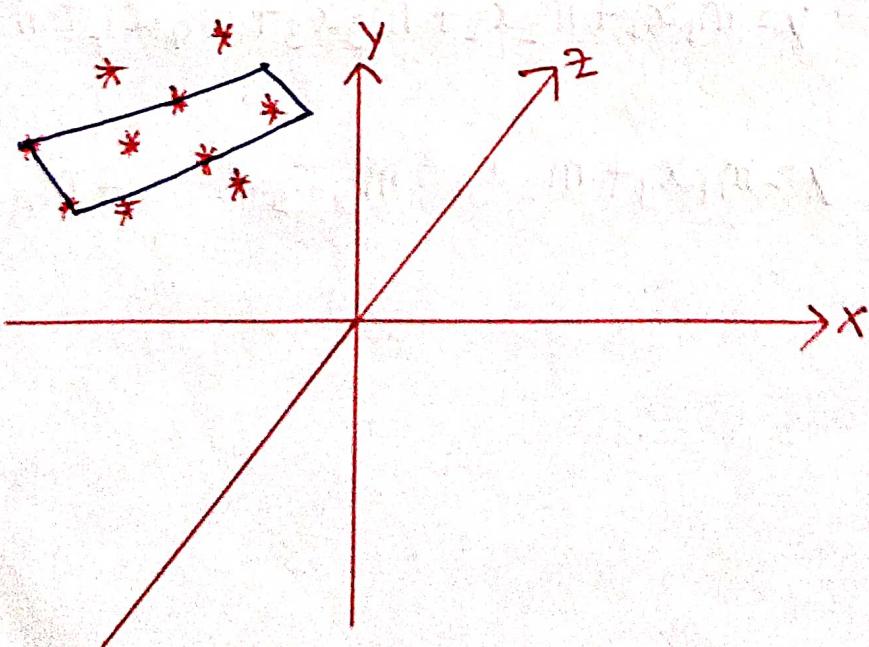
$$= CF + (+\text{ve}) \text{ value}$$

which will be greater than current CF.

- * while taking Learning rate (α) if it is very high, there may be times it can't find better Global minima.
- * But if we take it less, although it moves slowly, But can find better Global minima.

Multi Linear Regression:

- * It is same as simple linear regression, but will be used if we have more than 1 independent feature.



- * Instead of 2D Line, we'll use 3D Plane box here to fit the points.
- * Formulae for plane depends on number of Independent Features.

Independent Features Formulae

$$2 \quad y = m_1x_1 + m_2x_2 + c$$

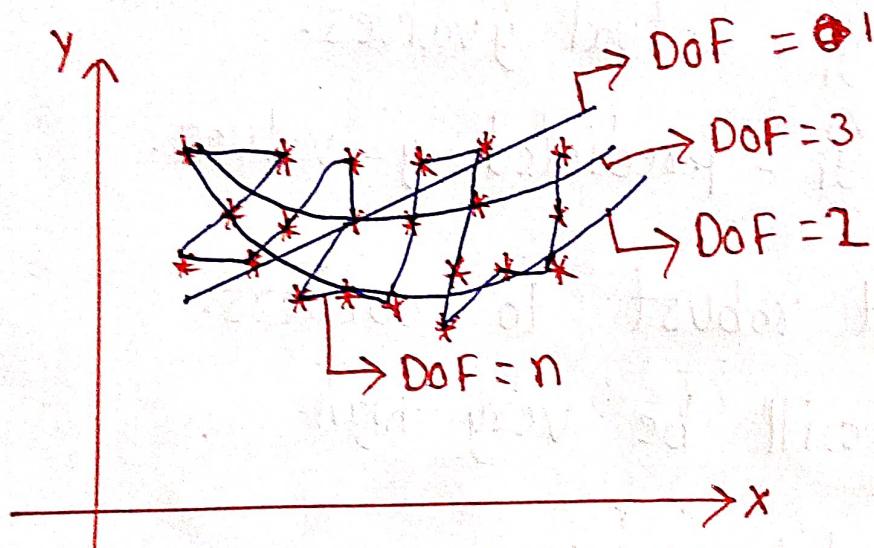
$$3 \quad y = m_1x_1 + m_2x_2 + m_3x_3 + c$$

$$5 \quad y = m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 + m_5x_5$$

$$n \quad y = m_1x_1 + m_2x_2 + m_3x_3 + \dots + m_nx_n$$

Polynomial Regression:

- * It is used while data is non-linear.
- * we will have a new term in it called Degree of Freedom.



- * Higher the DOF, Higher (or) best the best fit line will be.

Polynomial Degree

0

Formulae

$$m_0 x_i^0$$

1

$$m_0 x_i^0 + m_1 x_i^1$$

2

$$m_0 x_i^0 + m_1 x_i^1 + m_2 x_i^2$$

3

$$m_0 x_i^0 + m_1 x_i^1 + m_2 x_i^2 + m_3 x_i^3$$

Types of cost functions:

1) Mean Squared Error

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

where, n = no of data points.

y_i = actual y values.

\hat{y}_i = predicted y values.

- * It's not robust to outliers.
- * values will be very high.

2) Mean Absolute Error:

$$\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

- * It's robust to outliers
- * values will be less.

Root Mean Squared error:

- * It's square root of Mean Squared Error.

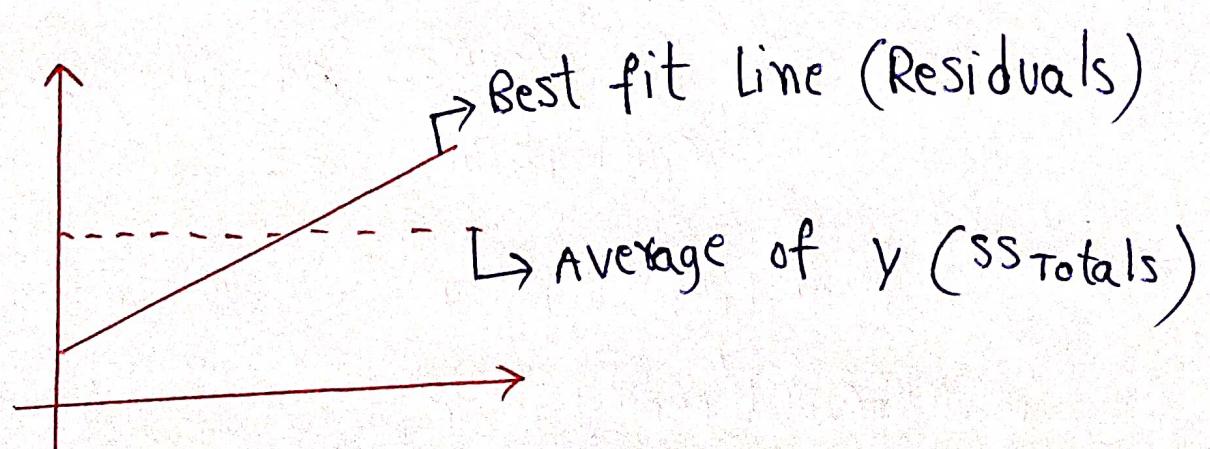
$$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

- * It's not robust to outliers.
- * values will be less.

Accuracy checks:

→ In Regression (Linear), we'll check accuracy of our models by r^2 and adjusted r^2 's.

$$R^2 = 1 - \frac{SS_{\text{Residuals}}}{SS_{\text{Total}}}$$



Adjusted R^2 :

$$1 - \frac{(1-R^2)(N-1)}{N-P-1} * P = \text{No. of Independent features}$$