Polynomial Linear Regression :-

Before going to polynomial eqn, we just discuss the Simple, multiple linear regression use case, the eqns are look like —

S.L.R — $\hat{y} = mx + c$

M.L.R — $\hat{y} = m_1x_1 + m_2x_2 + m_3x_3 + \dots + c$

So in this first we have to check the linearity Assumption.

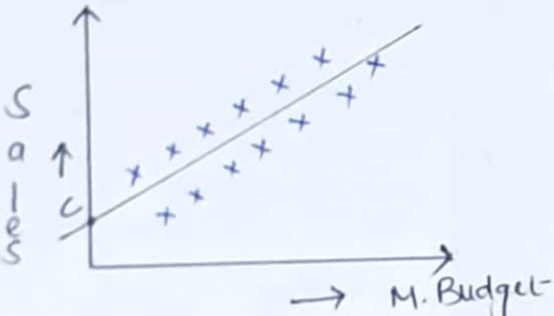
Task : So here Our task is to fit a line the training data. (m, c)

Use case 1 :— A data which Contain market budget and sales. on it for diff market budget gets diffrent sales in this market budget is Independent variable and sales is dependent variable

if i draw Scatterplot on this table then that shows some relationship b/w them. may be +ve or −ve

M. Budget	Sales
1 lack	10 lacks
2 lack	25 lacks
50 lack	75 lacks
⋮	⋮
Ind_V	dep_V

pred — 10L → ?



So if we make a line of liniarity ie show the relation b/w dependent & Indepent variable ⇒ $y = mx + c$

Sales = m (some slope) * Marketing budget + c

for Suppose if i put c = 3 from graph ⇒ Sales = m * M-B + 3

if i put M-Budget as zero ⇒ Sales = 0 + 3 ⇒ 3 lacks Sales

Scanned with CamScanner

ie if i put '0' Marketing budget then also i get 3 lack Revenue on Sales from data we Observe that when marketing budget increases Sales also increases ie there is exist some positive Co-relation.

Use case 2 : for Suppose if we have a multiple features in dataset like marketing budget Contain Social media, TV advertisement as independent variable and Sales dependent variable as shown in figure.

M.B.

TV	Social	Sales
50L	50L	10 L
1.3L	70L	25 L
⋮	⋮	⋮

Ind-var (TV, Social) ; dep-var (Sales)

from dataset if we make aline eqn —

$$y = m_1 x_1 + m_2 x_2 + \cdots + c$$

Put the features in above eqn - we get -

$$Sales = m_1 * TV + m_2 * Social\_media + c$$

from the data we can ~~if its~~ for Say that if we put Investments on budget either TV advertisement or Social media we get Increment of Sales. this depends on the angle making Value w.r.t to plane ie Slope.

for Suppose if we Consider two Cases where Slopes m_1 & m_2 let Consider $c=3, m_1 = 2$ and $m_2 = 10$. put it in above eqn.

$Sales\_1 = 2*50 + 0*10 + 3$ if we didn't Consider Social Media budget

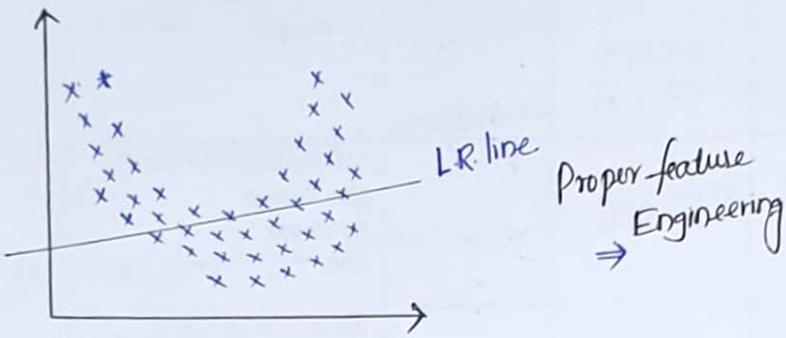
$Sales\_1 = 1$ cr 3lack rupees.

$Sales\_2 = 0*2 + 50*10 + 3$ when Investment On Social-Media.

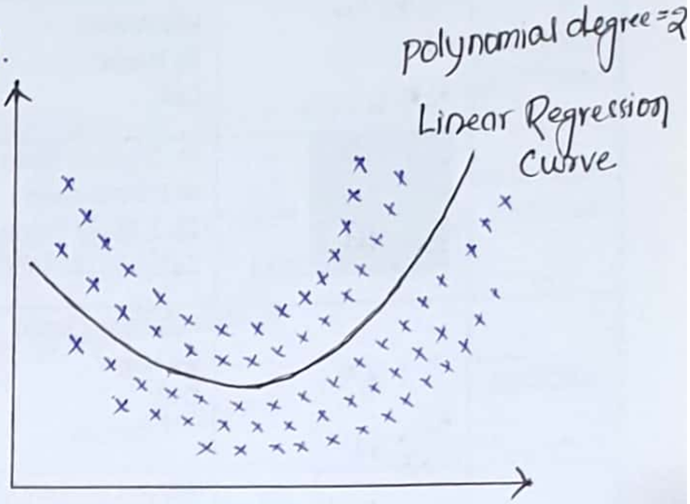
$= 5$ cr 3lack rupees

from the above we infer that when we invest budget on Social media we get more sales as compare to TV advertisement

Polynomial Regression :- if we plot a graph in between dep and Ind-variable and the graph is looklike below mentioned.



(a)



(b)

from figure (a) if we fit a line across data its passes through data but its not a best fit line for that. if we do some proper feature engineering then we get a linear regression line that best fit the data as shown in fig (b) that line also called as "Linear Regression Curve"

S.L.R 2D line Eqn $\Rightarrow$ $w_1 x_1 + w_2 x_2 + w_0 = 0$

M.L.R $\Rightarrow$ plane Eqn $\Rightarrow$ $w_1 x_1 + w_2 x_2 + w_3 x_3 + w_0 = 0$ [$\because$ more than 2 dimentions]

from the above plane Eqns $\Rightarrow$ we can write as

$$S.L.R \Rightarrow \begin{bmatrix} w_1 & w_2 & w_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = 0$$

$$M.L.R \Rightarrow \underbrace{\begin{bmatrix} w_1 & w_2 & w_3 & w_0 \end{bmatrix}}_{\text{Coeffitionts}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}}_{\text{feature vectors}} = 0$$

So here linear Regression means if we have linear Coeffitent in eqn (w_1, w_2, w_0), it not depend on feature vectors in eqn $\begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$ ie the Slopes are Contain linearity, not features.

How to fit a polynomial Curve :-

from figure (a) we draw ~~a poly~~ a line that is linear eqn now we have to fit that line Curve becouse its not best-fit line, In Order to do it we have to transform Our train data in Such away as shown in below.

x_train	y_trn
x	y

(i)

Feature Eng

New feature

x^2	x	y_trn
x^2	x	y

(ii)

reg = LinearRegression()
reg . fit (x_trn, y_trn)
[x , y]

reg = Linear Regression ()
Polynomial degree (–)
reg.fit (x_trn, y_trn)
[x^2, x , y]

In graph (i) we do training of data and make a best fit line but thats not fitted well then we do Some feature-Enginering On Same train data then that form a new feature called as x^2 now in x_trn Containing (x^2, x) In erlier In x_trn (x) only by this diffrence of train data we get a Curve that best fitted to the points is Called as a Polynomial Curve.

So In this case linear regression means fitting the data Using a line then that line Can be 1° line, 2 degree line or etc.

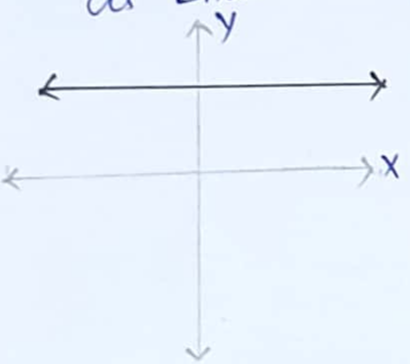
Simply we can say linear regression is nothing but Coefficents of

line is going to be linear

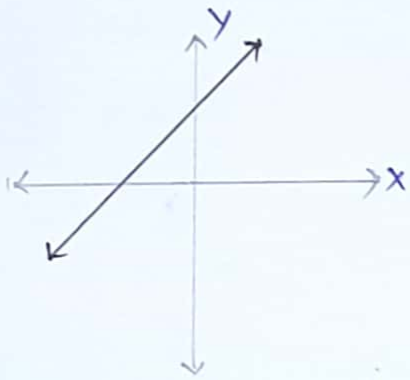
Note :- if Suppose we have data follow Some patterns like log normal, exponential eqn then we dont need to do polynomial Curve fitting. Simply do according transformation (ie data forming exponential eqn do exponential transformation, if follow lognormal then do Some log normal transformation) because that best fitted to that graph, not a polynomial Curve.

So here polynomial is nothing but fitt a best line (linear Regression) that may be a straight line or Curve, when we apply Some degree that will be change the line formatting as linear, Quadratic, Cubic eqn.

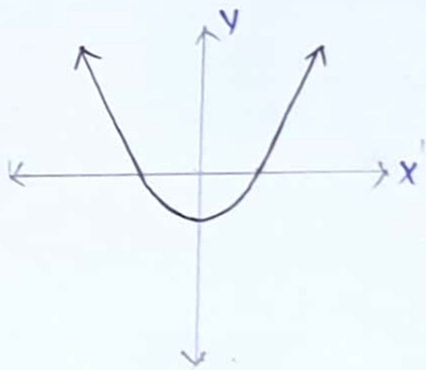
as Shown in figure -



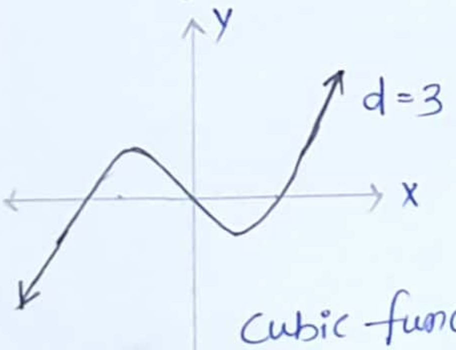
When we apply degree = 0 forming a Constant function



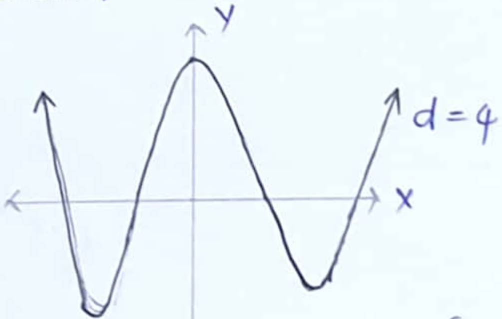
degree = 1
Linear function



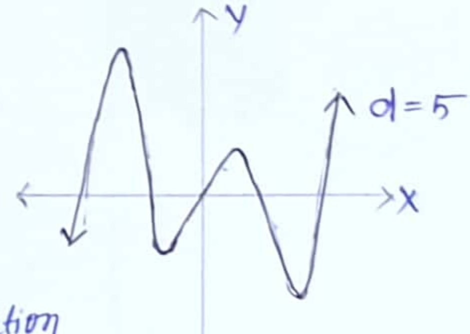
d = 2
Quadratic function



Cubic function



Quartic function



Quintic function