Curve Fitting:- Sometimes a function may be defined in the form of an expression like

$$F(x)=Ax^{m}+Bx^{m-1}+Cx^{m-2}+- -$$

In this particular case function may be defined and plotted easily and may be extended to required limit, and function is continuous so any point on the curve may be observed and calculated. But in some problems function is not defined directly but given in the form of a table, in that case we have to approximate the function such that the curve passes through maximum number of points. Such method of approximating the curves is called curve fitting.

In brief we can define curve fitting as the process of finding equations of approximating curves which fit the given set of data.

Let F(x) is the approximated function

Let us suppose x_1, x_2, x_3, -------, x_n are the values of independent variable and y_1, y_2, y_3,------- , y_n are the values of dependent variables respectively that are the result of experiment.

Let y'=F(x)

$D_1=y_1-F(x_1) = y_1-y_1'$, $D_2=y_2-F(x_2) = y_2-y_2'$,----------------------------, $D_n=y_n-F(x_n)=y_n-y_n'$

The function is chosen in such a way that D_1, D_2, D_3,----------, D_n are small.

Different methods of minimization of errors may be

1. $\text{Min}\sum_{i=1}^{n} Di = \text{Min}\sum_{i=1}^{n}(yi - F(xi))$
2. $\text{Min}\sum_{i=1}^{n}|Di| = Min\sum_{i=1}^{n}|yi - F(xi)|$
3. $\text{Max}(Di) \leq \varepsilon$
4. $\text{Min}\sum_{i=1}^{n} Di^2 = Min\sum_{i=1}^{n}(yi - F(xi))^2$

Method of least square Fitting:-
To fit the observe sample points the error can be minimized by computing the sum of square of deviation between the proposed curves of data points i.e. in this we use approach of equation 4.

Advantages of least square fitting:-
1. Small calculation and better result
2. The small errors should be neglected and larger differences should be highlighted and method of least square fitting perform the same.

3. The opposite differences cannot be cancelled each other so as to minimize the errors. In the method of least square fitting the same task is performed.

Linear Regression or straight line fitting:-

Let F(x) is the approximated function. The experimental values are given in the Table:

x	y
X_1	Y_1
X_2	Y_2
|	|
|	|
X_n	Y_n

$$D_1=y_1-F(x_1)$$

$$D_2=y_2-F(x_2)$$

|

|

$$D_n=y_n-F(x_n)$$

$$S=D1^2+D2^2+D3^2+ - - - - - - - - - - - - - +Dn^2$$

Let $F(x) = a_0+a_1x$ be the approximated function

$$D_1=y_1-(a_0+a_1x_1)$$

$$D_2=y_2-(a_0+a_1x_2)$$

|

|

|

$$D_n=y_n-(a_0+a_1x_n)$$

$$\text{Min } S=\text{Min}\sum_{i=1}^{n}\left(yi-f(xi)\right)^2$$

Differentiating S partially w.r.t. a0 and equating it to zero

$$\frac{\partial S}{\partial ao} = \sum_{i=1}^{n} 2(yi - (ao + a1xi))(-1)=0\text{---------------------}①$$

Now differentiating S partially w.r.t a_1

$$\frac{\partial S}{\partial a1}= \sum_{i=1}^{n} 2(yi - (ao + a1xi))(-xi)=0\text{--------------------}②$$

On rearranging equations ① &② we finally get

$$nao + a_1\sum_{i=1}^{n} xi = \sum_{i=1}^{n} yi - - - - - - - - - - - - ③$$

$$ao\sum_{i=1}^{n} xi + a1\sum_{i=1}^{n} xi^2 = \sum_{i=1}^{n} xiyi - - - - - - - ④$$

On solving equations ③ &④ we have

$$a_o=\frac{\sum_{i=1}^{n} yi \sum_{i=1}^{n} xi^2 - \sum_{i=1}^{n} xi \sum_{i=1}^{n} xiyi}{n\sum_{i=1}^{n} xi^2 - \left(\sum_{i=1}^{n} xi\right)^2}$$

$$a_1=\frac{n\sum_{i=1}^{n} xiyi - \sum_{i=1}^{n} xi \sum_{i=1}^{n} yi}{n\sum_{i=1}^{n} xi^2 - \left(\sum_{i=1}^{n} xi\right)^2}$$

Polynomial Regression:-

Let us suppose $F(x) = a_0 + a_1xi + a_2xi^2$ be the polynomial

Let xi's and yi's are the data points where we want to fit this 2nd order polynomial.
By using method of least square fitting

$$S=\sum_{i=1}^{n}\left(yi - F(xi)\right)^2$$

$$\text{Min } S=\text{Min}\sum_{i=1}^{n}\left(yi - (ao + a1xi + a2xi^2)\right)^2$$

To apply the criteria of least square fitting we differentiate S w.r.t a0, a1 & a2 and equate them to zero

$$\frac{\partial S}{\partial ao} = \sum_{i=1}^{n} 2(yi - (ao + a1xi + a2xi^2) = 0 - - - - - - - - ①$$

$$\frac{\partial S}{\partial a1} = \sum_{i=1}^{n} 2\big(yi - (ao + a1xi + a2xi^2)\big)(-xi) = 0 ------ ②$$

$$\frac{\partial S}{\partial a2} = \sum_{i=1}^{n} 2(yi - (ao + a1xi + a2xi^2)(-xi^2) = 0 ---- ③$$

On rearranging equation①, ② &③

$$nao + a1\sum_{i=1}^{n} xi + a2\sum_{i=1}^{n} xi^2 = \sum_{i=1}^{n} yi -------- ④$$

$$ao\sum_{i=1}^{n} xi + a1\sum_{i=1}^{n} xi^2 + a2\sum_{i=1}^{n} xi^3 = \sum_{i=1}^{n} xiyi ------------ ⑤$$

$$ao\sum_{i=1}^{n} xi^2 + a1\sum_{i=1}^{n} xi^3 + a2\sum_{i=1}^{n} xi^4 = \sum_{i=1}^{n} xi^2yi --- ⑥$$

Equations ④, ⑤ &⑥ are three linear equations with three unknowns. These equations are called normal equations for quadratic regression. These may be solved by Gauss Elimination Method.

Exponential Regression:-

Fitting of an exponential curve to a given problem.

The basic assumption in picking up least square error criteria for curve fitting is that for linear functions F(x), this leads to optimum fit.

This assumption is no more valid for nonlinear curve, inspite of that we use least square method due to its case of easy implementation.

Let y=aoe^{a1x} be the function to be fitted by using method of least square

y=aoe^{a1x}------------------①

Taking Log of equation ①

$$\log y = \log(a0e^{a1x})$$

$\log y$ =$\log a0 + a1x$--------------②

Let $\log y = z$, $\log ao = b\,o$, a1=b1-----------③

By putting these assumption from equations ③ into equation ②, *we get*

$$Z = bo + b1x \text{------------------} ④$$

Equation ④ is linear and we can use the normal equations for linear regression and obtain

$$bo = \frac{\sum_{i=1}^{n} zi \sum_{i=1}^{n} xi^2 - \sum_{i=1}^{n} xizi}{n\sum_{i=1}^{n} xi^2 - \left(\sum_{i=1}^{n} xi\right)^2}$$

$$bo = \frac{\sum_{i=1}^{n} \log yi \sum_{i=1}^{n} xi2 - \sum_{i=1}^{n} xi \log yi}{n\sum_{i=1}^{n} xi^2 - \left(\sum_{i=1}^{n} xi\right)^2}$$

$$b1 = \frac{n\sum_{i=1}^{n} xizi - \sum_{i=1}^{n} xi \sum_{i=1}^{n} zi}{n\sum_{i=1}^{n} xi^2 - \left(\sum_{i=1}^{n} xi\right)^2}$$

$$b1 = \frac{n\sum_{i=1}^{n} xi \log yi - \sum_{i=1}^{n} xi \sum_{i=1}^{n} \log yi}{n\sum_{i=1}^{n} xi^2 - \left(\sum_{i=1}^{n} xi\right)^2}$$

$$ao = e^{bo}$$

$$a1 = b1$$

Hyperbolic Regression:-

Fitting of a hyperbolic curve

$$y = \frac{1}{a+bx}$$

$$z = a + bx$$

$$a = \frac{\sum_{i=1}^{n} xi \sum_{i=1}^{n} zi - \sum_{i=1}^{n} xi \sum_{i=1}^{n} xizi}{n\sum_{i=1}^{n} xi^2 - \left(\sum_{i=1}^{n} xi\right)^2}$$

$$a = \frac{\sum_{i=1}^{n} xi^2 \sum_{i=1}^{n} \frac{1}{yi} - \sum_{i=1}^{n} xi \sum_{i=1}^{n} \frac{xi}{yi}}{n\sum_{i=1}^{n} xi^2 - \left(\sum_{i=1}^{n} xi\right)^2}$$

$$b=\frac{n\sum_{i=1}^{n} xizi-\sum_{i=1}^{n} xi\sum_{i=1}^{n} zi}{n\sum_{i=1}^{n} xi^{2}-\left(\sum_{i=1}^{n} xi\right)^{2}}$$

$$b=\frac{n\sum_{i=1}^{n}\frac{xi}{yi}-\sum_{i=1}^{n} xi\sum_{i=1}^{n}\frac{1}{yi}}{n\sum_{i=1}^{n} xi^{2}-\left(\sum_{i=1}^{n} xi\right)^{2}}$$