Curve Fitting

The general problem of finding equations of approximating curves which fit given data is called curve fitting.

Principle of least Squares

Let a set of points (x_i, y_i) where $i = 1, 2, 3, \ldots, n$ represents the given data and x, y are variates.

Let Y_i be the expected value of y corresponding to $x = x_i$ then

$E_i = y_i - Y_i$ is called Error of Estimate or Residual of y

$$E = \sum_{i=1}^{n} E_i^2$$

Algorithm

To fit the straight line $y = a + bx$ — (1)

Step 1: First substitute the given set of n values in the eqn (1)

Step 2: write the normal eqn

$$\Sigma y = na + b\Sigma x$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2$$

Step 3: Solve the normal eqn for values of a & b

Step 4: Sub. a and b in (1) to get the reqd eqn.

Q. By the method of least square, find the straight line that best fits the following data

x	1	2	3	4	5
y	14	27	40	55	68

9. Fit a straight line to the following data regarding x as independent variable

x	1	2	3	4	5	6
y	1200	900	600	200	110	50

Soln: Let the eqn of straight line be

$$y = a + bx$$

x	y	xy	x^2
1	1200	1200	1
2	900	1800	4
3	600	1800	9
4	200	800	16
5	110	550	25
6	50	300	36

$\Sigma x = 21$ $\Sigma y = 3060$ $\Sigma xy = 6450$ $\Sigma x^2 = 91$

Here, $n = 6$

from normal eqn,

$$\Sigma y = na + b\Sigma x$$

and $$\Sigma xy = a\Sigma x + b\Sigma x^2$$

$$\Rightarrow 3060 = 6a + b(21)$$

and $$6450 = a(21) + b(91)$$

$$\Rightarrow a = 1361.97$$

and $$b = -243.42$$

$$\therefore y = 1361.97 - 243.42x$$

Q. Fit a straight line to the following data

x	71	68	73	69	67	65	66	67
y	69	72	70	70	68	67	68	64