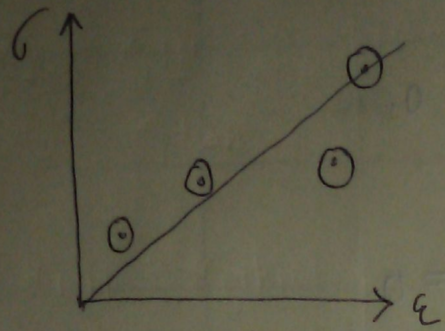
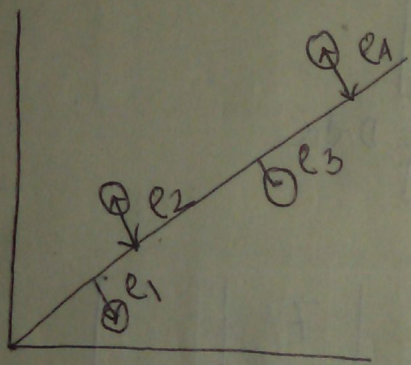
Motivation :-

Least Square Regression :-

best fit straight line
↓
curve fitting by linear regression.



Stress - strain diagram



e_1, e_2, e_3, e_4
L errors

$y = a_0 + \cancel{a_0 e}\; a_1 x$

$a_0 = ?$

$a_1 = ?$

Least Square Regression :-

sum of square of 'errors' are minimized
$e_1 e_2, e_3, e_4$
↓
gives best fitted straight line

$(e_1^2 + e_2^2 + e_3^2 + e_4^2)$ ← minimize this,

$S_r = \sum e_i^2$ ← minimize this → can be done by taking first derivative.

$y = a_0 + a_1 x + e$ - (errors)

$e = y - a_0 - a_1 x$

$\sum e_i^2 = \sum (y_i - a_0 - a_1 x_i)^2$ [a_0, a_1 variables]

$$\frac{\partial S_e}{\partial a_0} = 0 = -2 \sum (y_i - a_0 - a_1 x_i) = 0$$

$$\frac{\partial S_e}{\partial a_1} = 0 = -2 \sum x_i (y_i - a_0 - a_1 x_i) = 0$$

$$\boxed{a_0 = \bar{y} - a_1 \bar{x}}$$

$$\boxed{a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}}$$

Fit a straight line to the following Data

x	1	2	3	4	5	6	7	
y	.5	2.5	2	4	3.5	6	5.5	

Sol^n :- $y = a_0 + a_1 x$

x	y	x^2	xy
1	.5	1	.5
2	2.5	4	~~10~~ 5
3	2	9	~~18~~ 6
4	4	16	~~64~~ 16
5	3.5	25	17.5
6	6	36	36
7	5.5	49	38.5
$\Sigma x = 28$	$\Sigma y = 24$	$\Sigma x^2 = 140$	$\Sigma xy = 119.5$

$a_0 = \bar{y} - a_1 \bar{x}$

$$a_1 = \frac{n\Sigma x_i y_1 - \Sigma x_i \Sigma y_i}{n \Sigma x_i^2 - (\Sigma x_i)^2}$$

$\bar{y} = \frac{\Sigma y}{n} = \frac{24}{7} = 3.428$

$\bar{x} = \frac{\Sigma x}{n} = \frac{28}{7} = 4$

$a_0 = 3.428 - a_1 (4)$

$$a_1 = \frac{7 \times 119.5 - 28 \times 24}{7 \times 140 - (28)^2}$$

$\boxed{a_1 = .83928}$ Ans

$\boxed{a_0 = 0.07}$

$\boxed{y = .07 + .83928x}$

Co-efficient of Determination :- } goodness- of fit test

R^2 value

$$R^2 = \frac{S_t - S_r}{S_t}$$

$$S_t = \Sigma (y_i - \bar{y})^2$$

$$S_r = \Sigma (y_i - a_0 - a_1 x)^2$$

if $S_t = S_r \rightarrow R^2 = 0$, the horizontal line is the best fit straight line.

if $S_t > S_r \rightarrow R^2 \approx 1 \rightarrow$ Very good fit

$R^2 = -ve$ হলে Mean will best possible fit

x	1	2	3	4	5	6	7
y	.5	2.5	2	4	3.5	6	5.5

$$\bar{y} = \frac{\sum y}{n} = 3.428$$

x	y	xy	x^2	$(y-\bar{y})^2$	$(y-a_0-a_1x)^2$
1	.5	.5	1	8.57	.1687
2	2.5	5	4	.8611	.5634
3	2	6	9	2.039	.3462
4	4	16	16	.327	.3278
5	3.5	17.5	25	.00518	.5873
6	6	36	36	6.61	.8
7	5.5	38.5	49	4.29	.1974
$\sum x = 28$	$\sum y = 24$	$\sum xy = 119.5$	$\sum x^2 = 140$	$\sum (y-\bar{y})^2 = 22.7$	$2.99 = \sum (y-a_0-a_1x)^2$
$\bar{x} = \frac{28}{7} = 4$	$\bar{y} = \frac{24}{n} = \frac{24}{7} = 3.428$			S_t	S_r

$$a_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{7\times 119.5 - (28\times 24)}{7\times 140 - (28)^2} = .83928$$

$$a_0 = \bar{y} - a_1\bar{x} = 3.428 - .83928\times(4) = .0714$$

$$\boxed{y = .714 + .83928x}$$

$$r^2 = \frac{S_t - S_r}{S_t} = \frac{22.7 - 2.99}{22.71} = \boxed{.868 = r^2}$$

r^2 = coefficient of determination

r = co-relation co-efficient

Scanned by CamScanner

Polynomial Regression :-

$y = a_0 + a_1 x + a_2 x^2$ $\left\{ \begin{matrix} a_0 \\ a_1 \\ a_2 \end{matrix} = ?? \right\}$

$\frac{\partial S_p}{\partial a_0} = 0$ $\frac{\partial S_p}{\partial a_1} = 0$ $\frac{\partial S_p}{\partial a_2} = 0$

$n a_0 + \sum x a_1 + \sum x^2 a_2 = \sum y$

$\sum x a_0 + \sum x^2 a_1 + \sum x^3 a_2 = \sum xy$

$\sum x^2 a_0 + \sum x^3 a_1 + \sum x^4 a_2 = \sum x^2 y$

$$\begin{bmatrix} n & \sum x & \sum x^2 \\ \sum x & \sum x^2 & \sum x^3 \\ \sum x^2 & \sum x^3 & \sum x^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum xy \\ \sum x^2 y \end{bmatrix}$$

Fit a Parabola for the following data :-

x	0	1	2	3	4	5
y	2.1	7.7	13.6	27.2	40.9	61.1

x	y	x^2	x^3	x^4	xy	x^2y	$(y-\bar{y})^2$	$\{y-(a_0+a_1x+a_2x^2)\}^2$
0	2.1	0	0	0	0	0	544.42	
1	7.7	1	1	1	7.7	7.7	314.459	
2	13.6	4	~~27~~ 8	16	27.2	54.4	839.08	
3	27.2	9	27	81	81.6	244.8	3.122	
4	40.9	16	64	256	163.6	654.4	239.22	
5	61.1	25	125	625	305.5	1527.5	1272.134	
15	152.6	55	225	979	585.6	2488.8	~~161 71.44~~	
Σx	Σy	Σx^2	Σx^3	Σx^4	Σxy	Σx^2y	3212	

$\bar{x} = \frac{15}{6} = 2.5$ $\quad \bar{y} = 25.433$

$\Sigma(y-\bar{y}^2)$ St $\quad$ Sr

$$\begin{vmatrix} n & \Sigma x & \Sigma x^2 \\ \Sigma x & \Sigma x^2 & \Sigma x^3 \\ \Sigma x^2 & \Sigma x^3 & \Sigma x^4 \end{vmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma xy \\ \Sigma x^2 y \end{bmatrix}$$

$$r^2 = \frac{St - Sr}{St}$$

$=$

$$= \begin{vmatrix} 6 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{vmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 152.6 \\ 585.6 \\ 2488.8 \end{bmatrix}$$

Solving

$a_0 = 2.47$

$a_1 = 2.35$

$a_2 = 1.86$

$$\boxed{y = 2.47 + 2.35x + 1.86x^2}$$

Non Liner regression :-

$\boxed{y = ax^b}$ Non linear equation

$a = ?$
$b = ?$

$$\underbrace{\log y}_{Y} = \underbrace{\log a}_{a_0} + \underbrace{b}_{a_1}\underbrace{\log x}_{x}$$

$Y = a_0 + a_1 X$ - equation of a straigwtline

Fit the model $y = ae^{bx}$ to the following data

$x = X$	y	$Y = \ln y$	X^2	XY
2	4.077	1.405	4	2.81
4	11.084	2.405	16	9.62
6	30.184	3.405	36	20.43
8	81.897	4.405	64	35.24
10	222.62	5.405	100	54.05

$\Sigma x = 30$ $\Sigma y = 349.862$ $\Sigma Y = 17.025$ $\Sigma X^2 = 220$ $\Sigma XY = 122.15$

$n = 5$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{30}{5} = 6$$

$$\bar{y} = \frac{\Sigma y}{n} = \cancel{\frac{349.86}{5}} = \cancel{69.97} = \frac{17.025}{5} = 3.405$$

$$\underbrace{\ln y}_{Y} = \underbrace{\ln a}_{a_0} + \underbrace{b}_{a_1}\underbrace{x}_{X}$$

$$a_1 = \frac{n\Sigma XY - \Sigma x \Sigma Y}{n\Sigma (X)^2 - (\Sigma x_i)^2}$$

$$= \frac{5 \times (122.15) - (30 \times 17.025)}{(5 \times 220 - 30^2)}$$

$$= .5$$

$a_0 = \bar{y} - a_1\bar{x}$

$= \cancel{69.97 - .5 \times 6}$

$= 3.405 - (.5 \times 6)$

$= .405$

$a_0 = \ln a = .405$

$a = 1.499$

$\boxed{y = 1.499e^{.5x}}$

Scanned by CamScanner

$y = ax^b$ থাকলে 10 base logarithm নিতে হবে।

$$y = \alpha_3 \frac{x}{\beta_3 + x}$$

$$\underbrace{\boxed{\frac{1}{y}}}_{Y} = \underbrace{\frac{1}{\alpha_3}}_{a_0} + \underbrace{\frac{\beta_3}{\alpha_3}}_{a_1} \underbrace{\boxed{\frac{1}{x}}}_{X}$$

Multiple linear Regression :-

$$y = a_0 + a_1 \overset{x_1}{x} + a_2 x^2$$

$$y = a_0 + a_1 x_1 + a_2 x_2$$

$$y = f(x_1, x_2)$$

$$\begin{bmatrix} n & \Sigma x_1 & \Sigma x_2 \\ \Sigma x_i & \Sigma x_1^2 & \Sigma x_1 x_2 \\ \Sigma x_2 & \Sigma x_1 x_2 & \Sigma x_2^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma x_1 y \\ \Sigma x_2 y \end{bmatrix}$$

Scanned by CamScanner

Fit the following Data with the equation $y = a_0 + a_1x_1 + a_2x_2$

x_1	x_2	y	x_1^2	x_2^2	x_1x_2	x_1y	x_2y
0	0	5	0	0	0	0	0
2	1	10	4	1	2	20	10
2.5	2	9	6.25	4	5	22.5	18
1	3	0	1	9	3	0	0
4	6	3	16	36	24	12	18
7	2	27	49	4	14	189	54
$\Sigma x_1 = 16.5$	$\Sigma x_2 = 14$	$\Sigma y = 54$	$\Sigma x_1^2 = 76.25$	54	48	243.5	100

$$\begin{bmatrix} n & \Sigma x_1 & \Sigma x_2 \\ \Sigma x_1 & \Sigma x_1^2 & \Sigma x_1x_2 \\ \Sigma x_2 & \Sigma x_1x_2 & \Sigma x_2^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma x_1y \\ \Sigma x_2y \end{bmatrix}$$

$$\begin{bmatrix} 6 & 16.5 & 14 \\ 16.5 & 76.25 & 48 \\ 14 & 48 & 54 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 54 \\ 243.5 \\ 100 \end{bmatrix}$$

$a_0 = 5$

$a_1 = 4$

$a_2 = -3$

$\boxed{y = 5 + 4x_1 - 3x_2}$ Ans