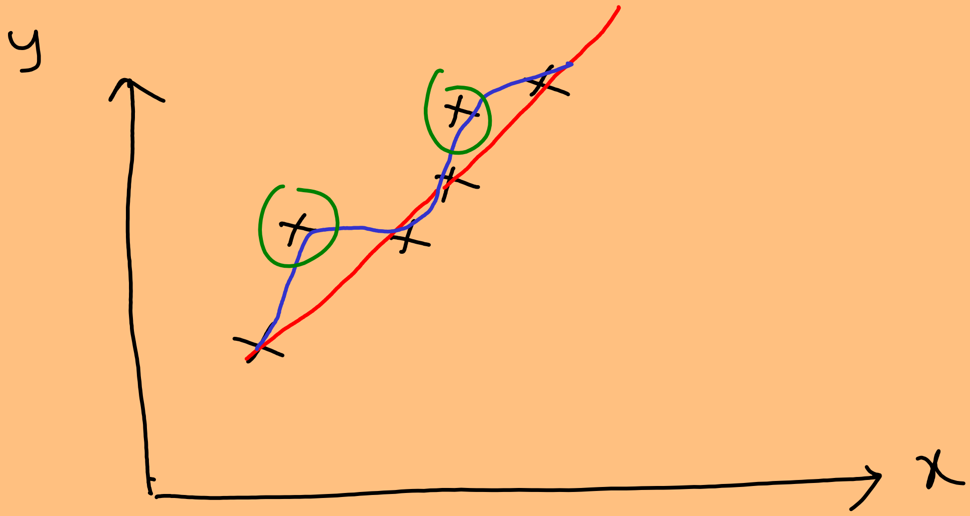
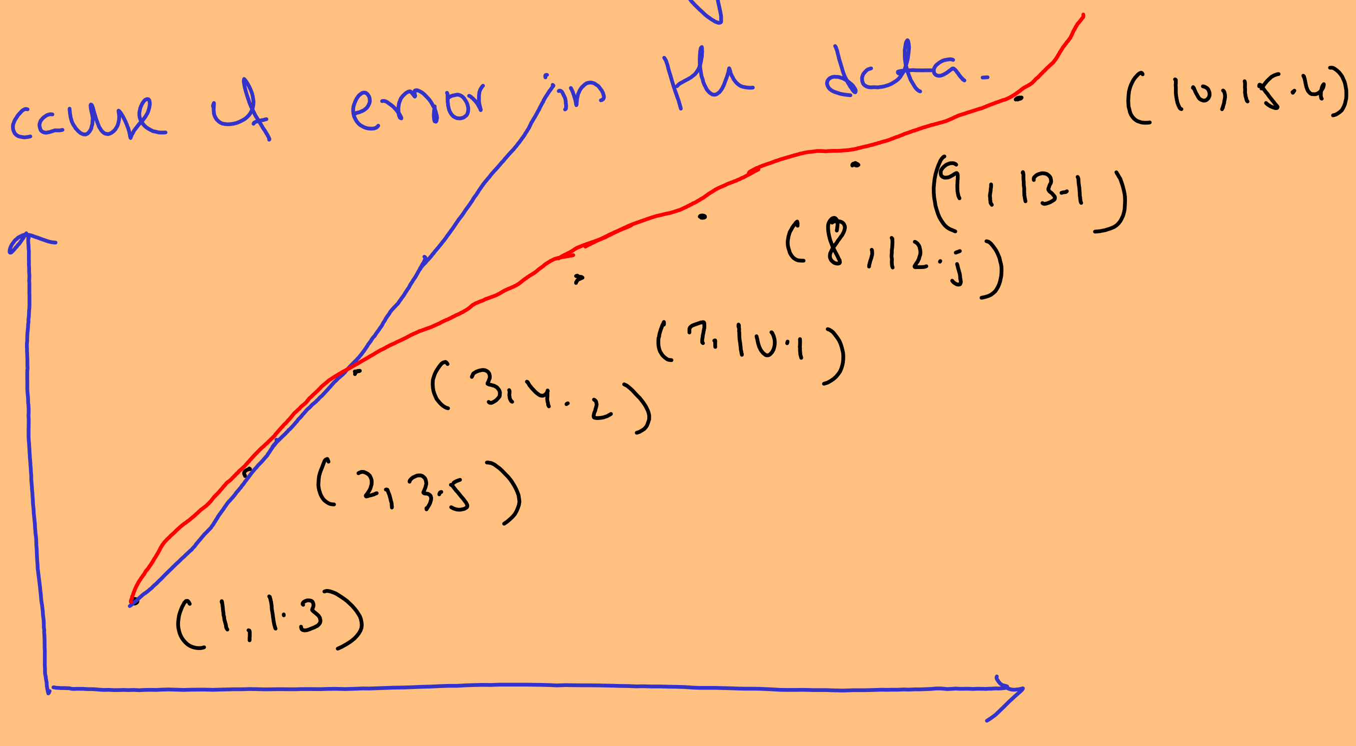
Curve Fitting:

Method of Least Squares Approximation:



* From the above figure, the actual relationship b/w x and y is linear.

* no line precisely fits the data is because of error in the data.



* To determine the best linear approximation involves finding the values of a_0 and a_1 to minimize

$$E = \sum_{i=1}^{10} \left(y_i - \left(a_1 x_i + a_0 \right) \right)$$

$$y = mx + c$$

* The least square method involves determining the best approximation line when the error involved is the sum of squares of the difference b/w y-values on the approximating line and the given values.

* Thus constants a_0 and a_1 must be found that minimize the least square error

$$E = \sum_{i=1}^{\overset{\text{no. of data}}{m}} \left[\underset{\text{actual value}}{y_i} - \underset{\text{approximate y value}}{\left(a_1 x_i + a_0 \right)} \right]^2 \quad \text{—} (1)$$

with respect to the parameter a_0 and a_1.

For a minimum to occur, we have

$$\frac{\partial E}{\partial a_0} = \frac{\partial E}{\partial a_1} = 0$$

$$\frac{\partial E}{\partial a_0} = 0 \Rightarrow 2\sum_{i=1}^{m}\left(y_i - (a_1 x_i + a_0)\right)(-1) = 0 \quad \text{--- (2)}$$

$$\frac{\partial E}{\partial a_1} = 0 \Rightarrow 2\sum_{i=1}^{m}\left(y_i - (a_1 x_i + a_0)\right)(-x_i) = 0 \quad \text{--- (3)}$$

From (2), (3)

$$\sum_{i=1}^{m} a_0 = m a_0$$

$$\boxed{\begin{aligned} a_0 m + a_1 \sum_{i=1}^{m} x_i &= \sum_{i=1}^{m} y_i \\ a_0 \sum_{i=1}^{m} x_i + a_1 \sum_{i=1}^{m} x_i^2 &= \sum_{i=1}^{m} x_i y_i \end{aligned}}$$

This is called Normal equation.

It means to fit the linear curve

$$y = a_0 + a_1 x$$

Normal Eqns:

$$a_0 m + a_1 \sum x_i = \sum y_i$$

$$a_0 \sum_{i=1}^{m} x_i + a_1 \sum x_i^2 = \sum x_i y_i$$

1) Find the least square line for the data

(1, 1.3) (2, 3.5) (3, 4.2) (4, 5) (5, 7)
(6, 8.8) (7, 10.1) (8, 12.5) (9, 13)
(10, 15.6)

x_i	y_i	x_i^2	$x_i y_i$
1	1.3	1	1.3
2	3.5	4	7

3	4.2	9	12.6
4	5	16	20
5	7	25	35.0
6	8.8	36	52.8
7	10.1	49	70.7
8	12.5	64	100
9	13	81	117
10	15.6	100	156
55	81	385	572.4

$\therefore$ normal equations are

$$10a_0 + 55a_1 = 81$$
$$55a_0 + 385a_1 = 572.4$$

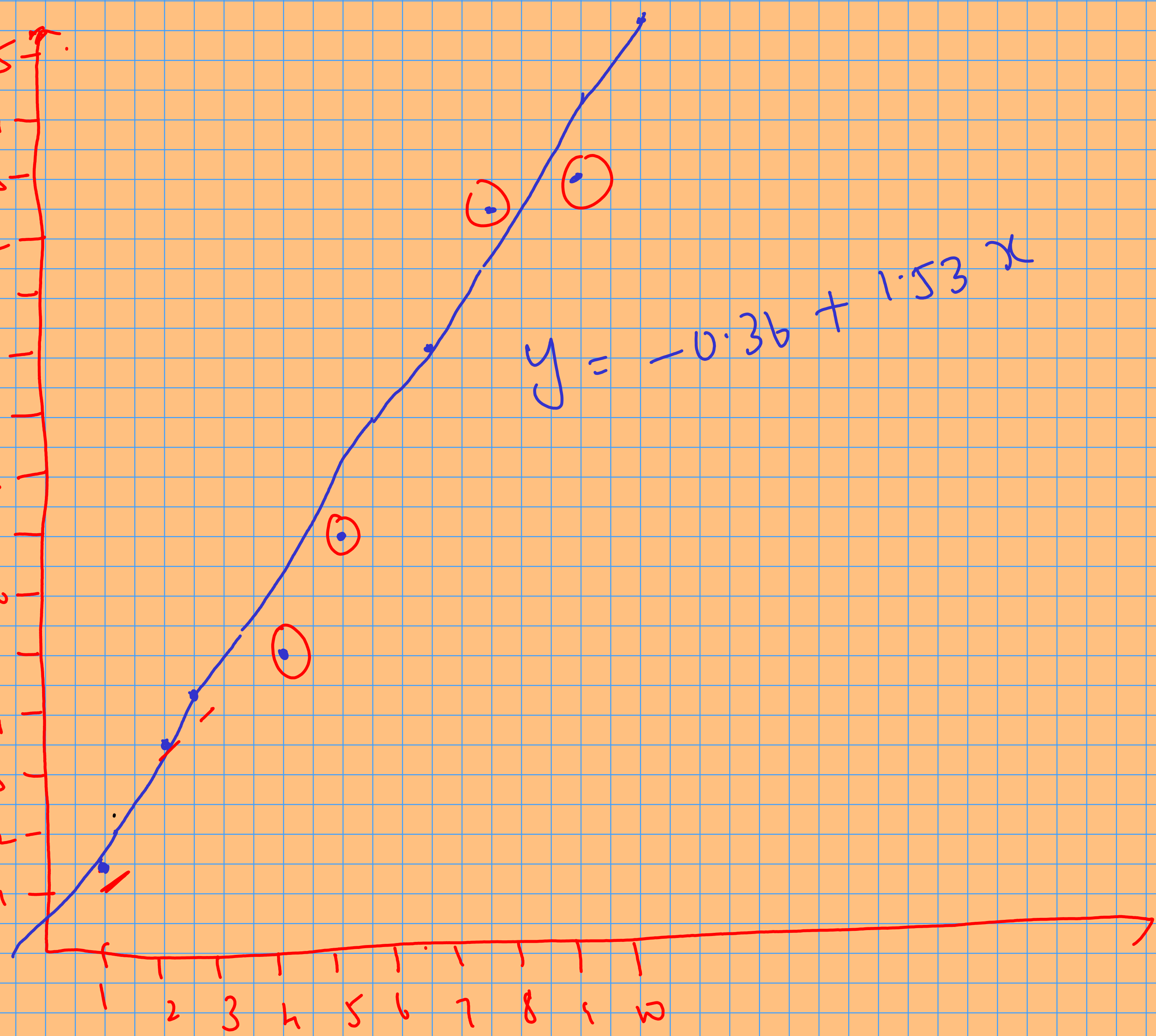
Use any iterative method to solve the above system of eqn, we get

$$a_0 = -0.36$$

$$a_1 = 1.53818$$

$\therefore$ approximating line

$$y = -0.36 + 1.53818\,x$$



Polynomial Least Squares:

The problem of approximating a set of data (x_i, y_i) $i = 1, 2 \ldots m$ with an polynomial

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

Choose $a_0, a_1, \ldots a_n$ to minimize the least square error

$$E = \sum_{i=1}^{m} \Big(\overset{\text{actual value}}{y_i} - \overset{\text{app. value}}{P_n(x_i)} \Big)^2$$

$$= \sum_{i=1}^{m} y_i^2 - 2 \sum_{i=1}^{m} y_i \, P_n(x_i) + \sum_{i=1}^{m} P_n(x_i)^2$$

$$= \sum_{i=1}^{m} y_i^2 - 2 \sum_{i=1}^{m} y_i \Big(\sum_{j=0}^{n} a_j x_i^j \Big) + \sum_{i=1}^{m} \Big(\sum_{j=0}^{n} a_j x_i^j \Big)^2$$

$$\boxed{\sum_{j=0}^{n} a_j x_i^j = a_n x_i^n + a_{n-1} x_i^{n-1} + \dots + a_1 x_i + a_0}$$

$$= \sum_{i=1}^{m} y_i^2 - 2 \sum_{j=0}^{n} a_j \left(\sum_{i=1}^{m} y_i x_i^j \right)$$

$$+ \sum_{j=0}^{n} \sum_{k=0}^{n} a_j a_k \sum_{i=1}^{m} x_i^{j+k}$$

We are looking to find

$$\frac{\partial E}{\partial a_j} = 0, \quad j = 0, 1, \dots n$$

For each j, we set

$$-2 \sum_{i=1}^{m} y_i x_i^j + 2 \sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = 0$$

$$\Rightarrow \boxed{\sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^j}$$

$j = 0, 1, \dots n$

This is called normal equations.

Remark: In polynomial $P_n(x)$

(1) $n=1$, $P_1(x) = a_1 x + a_0$

$$\sum_{k=0}^{1} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^{j}$$

$j = 0, 1$

$k=0$ $\quad$ $k=1$

$$\boxed{\begin{array}{ll} j=0: & a_0 m + a_1 \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} y_i \\ j=1: & a_0 \sum_{i=1}^{m} x_i + a_1 \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i y_i \end{array}}$$

(2) $n=2$, $\boxed{P_2(x) = a_2 x^2 + a_1 x + a_0}$

$$\sum_{k=0}^{2} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^{j}$$

$j = 0, 1, 2$

$$\begin{array}{llll}
 & k=0 & k=1 & k=2 \\
j=0 & a_0 m & + \; a_1 \sum_{i=1}^{m} x_i^1 & + \; a_2 \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} y_i \\
j=1 & a_0 \sum_{i=1}^{m} x_i & + \; a_1 \sum_{i=1}^{m} x_i^2 & + \; a_2 \sum_{i=1}^{m} x_i^3 = \sum_{i=1}^{m} x_i y_i \\
j=2 & a_0 \sum_{i=1}^{m} x_i^2 & + \; a_1 \sum_{i=1}^{m} x_i^3 & + \; a_2 \sum_{i=1}^{m} x_i^4 = \sum_{i=1}^{m} x_i^2 y_i
\end{array}$$

2) Fit the data $(0,1)$ $(0.25, 1.2840)$ $(0.5, 1.6487)$ $(0.75, 2.1170)$ $(1, 2.7183)$ with least square polynomial of deg 2.

x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y_i$
0	1					
0.25	1.2840					

0.5	1.6487				
0.75	2.1170				
1	2.7183				
2.5	8.7680	1.875	1.5625	1.3828	5.4514
					4.4015

$\therefore$ normal equations are

$$5a_0 + 2.5a_1 + 1.875a_2 = 8.7680$$

$$2.5a_0 + 1.875a_1 + 1.5625a_2 = 5.4514$$

$$1.875a_0 + 1.5625a_1 + 1.3828a_2 = 4.4015$$

Solve the above eqns using any iterative method, we get

$$a_0 = 1.005075$$

$$a_1 = 0.864676$$

$$a_2 = 0.843164$$

$\therefore$ required polynomial is :

$$\boxed{y = 1.005075 + 0.864676x + 0.843164x^2}$$

Exponential Function:

Approximating function to be of the form

$$y = be^{ax}$$

$$\ln y = \ln b + ax$$

$$\Rightarrow \boxed{Y = B + ax}$$

where $Y = \ln y$

$B = \ln b$

Normal Eqn:

$$mB + a\sum_{i=1}^{m} x_i = \sum_{i=1}^{m} Y_i$$

$$B\sum_{i=1}^{m} x_i + a\sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} Y_i x_i$$

Eg: Fit a curve $y = be^{ax}$

x_i	1	1.25	1.5	1.75	2
y_i	5.10	5.79	6.53	7.45	8.46

x_i	y_i	$Y_i = \ln y_i$	x_i^2	$x_i Y_i = x_i \ln y_i$
1	5.1	1.629	1	1.629
1.25	5.79	1.756	1.5625	2.195
1.5	6.53	1.876	2.25	2.814
1.75	7.45	2.008	3.0625	3.514
2	8.46	2.175	4	4.270
7.5		9.404	11.875	14.422

Normal Eqn:

$$5B + 7.5a = 9.404$$

$$7.5B + 11.875a = 14.422$$

$a = 0.5056$

$B = \ln b = 1.1224$

$\Rightarrow b = e^{B} = e^{1.1224} = 3.072$

$\therefore y = b e^{ax}$

$= 3.072 e^{0.5056 x}$

(2) $y = b x^{a}$

$\ln y = \ln b + a \ln x$

$Y = B + a X$ where

$Y = \ln y$

$B = \ln b$

$X = \ln x$

Normal Eqn:

$$B m + a \sum_{i=1}^{m} X_i = \sum_{i=1}^{m} Y_i$$

$$B \sum_{i=1}^{m} X_i + a \sum_{i=1}^{m} X_i^2 = \sum_{i=1}^{m} X_i Y_i$$

(2) Ex Fit a curve $y = b x^a$

x_i	1	1.25	1.5	1.75	2
y_i	5.1	5.79	6.53	7.45	8.46