Fitting a Straight Line to a Set of Data

In some of the experiments, a straight line is to be fitted to the experimental data. This can be done using the **Least Square Fitting Method**. The method can be used to fit any polynomial, but you would require it only for fitting a straight line. In the following the method for this purpose is described.

Let $(x_i, y_i\ ;\quad i = 1,2,\dots N)$ be the given set of data from a measurement. You are required to fit a straight line to this data. The best straight line, according to the least square method, can be fitted as follows.

Let the equation of the straight line to be fitted be written as

$$y = mx + c$$

where the slope m and the intercept c are to be determined from the given data. This is done using the least square method which gives the following working formulae:

$$m = \frac{AC - DN}{A^2 - NB} \quad \text{and} \quad c = \frac{C - mA}{N}$$

where $A = \sum_{i=1}^{N} x_i \qquad B = \sum_{i=1}^{N} x_i^2 \qquad C = \sum_{i=1}^{N} y_i \qquad D = \sum_{i=1}^{N} x_i y_i$

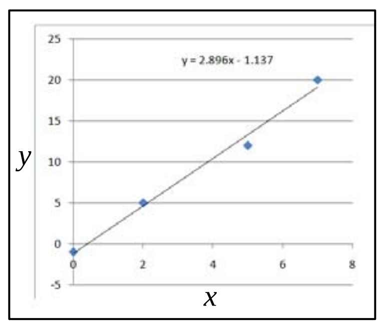
Example: Fit a straight line to the following data:

x	0	2	5	7
y	-1	5	12	20

x	x^2	y	xy
0	0	-1	0
2	4	5	10
5	25	12	60
7	49	20	140
Sum= 14	78	36	210
A	B	C	D

$$m = \frac{AC - DN}{A^2 - NB} = \frac{14 X 36 - 210 X 4}{14 X 14 - 4 X 78} = \frac{-336}{-116} = \frac{84}{29} = 2.8966$$

$$c = \frac{C - mA}{N} = \frac{36 - 14 X 336/116}{4} = \frac{36 X 116 - 14 X 336}{4 X 116} = \frac{-33}{29} = -1.1379$$



Thus the equation of the straight line is: $y = (84x - 33)/29$

A comparison of the given data and the fitted line:

x	0	2	5	7
y	-1	5	12	20
y_{fit}	-1.138	4.655	13.35	19.18

The straight line actually does not pass through any of the given points yet this the best fit line (please see the figure). In this case, the sum of the square of errors is minimum, i.e. $\sum (y - y_{fit})^2$, is minimum. Thus, it is not required that the straight line passes through any of the points for it to be the best fit line.