

TITLE: TO FIT A STRAIGHT LINE TO DATASET USING LEAST SQUARE METHOD

Definition:

The **least square method** is a statistical technique used to find the best-fitting line or curve for a set of data points by minimizing the sum of the squares of the differences (residuals) between the observed values and the values predicted by the model. It is widely used in regression analysis to determine the line of best fit that represents the relationship between variables.

Algorithm for least square method (linear case):

1. Collect Data:

Gather the data points $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

2. Assume the Model:

For a straight line, assume the equation $y = mx + c$, where m is the slope and c is the intercept.

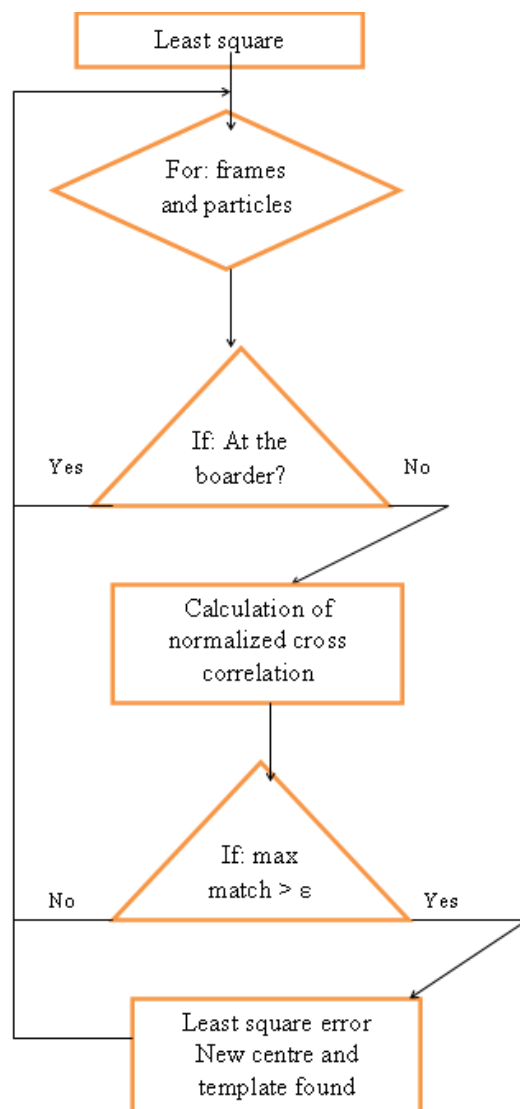
3. Calculate Slope (m) and Intercept (c):

- $$m = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$
- $$c = \frac{\sum y - m \sum x}{n}$$

4. Formulate the Line:

Substitute m and c into the equation $y = mx + c$ to get the best fit line.

Flowchart:



```

%Title:-
% To fit a straight line to the given daata set using least square method
%Developed by:-Arpan Adhikari
%Date: July 3, 2025
%-----Three critical statements-----
close all;
clear variables;
clc;
%----- User I/P section -----
X = input('Enter value for X[] =');
Y = input('Enter value for Y[] =');
while(length(X)~=length(Y))
    clc;
    disp('X and Y must have same dimension');
    X = input('Enter value for X[] =');
    Y = input('Enter value for Y[] =');
end
%----- Summation Calculation section -----
Sx = 0; Sy = 0; Sxy=0; Sxx=0;
n = length(X);
disp(' ');
disp('   X       Y       XY       X^2' );
disp(' ');
for i=1:n
    Sx = Sx + X(i);
    Sy = Sy + Y(i);
    Sxx = Sxx + X(i) * X(i);
    Sxy = Sxy + X(i) * Y(i);
    out = [ X(i), Y(i), X(i)*Y(i), X(i)*X(i)];
    disp(out);
end
grain = [Sx, Sy, Sxy, Sxx];
disp(' ');
disp(grain);
%----- Calculation section -----
b=(n*Sxy-Sx*Sy)/(n*Sxx -Sx*Sx);
a = (Sy-b*Sx)/n;
%----- Output section -----
result = strcat('The best curve to be fitted is Y =', num2str(a), '+',
num2str(b), '*X');
disp(' ');
disp(result);

```

Output:

```

[1,2,3,4,5]
Enter value for Y[] =
[1,2,1.3,3.75,2.25]

```

X	Y	XY	X^2
1	1	1	1
2	2	4	4
3.0000	1.3000	3.9000	9.0000
4.0000	3.7500	15.0000	16.0000
5.0000	2.2500	11.2500	25.0000
15.0000	10.3000	35.1500	55.0000

```

The best curve to be fitted is Y =0.785+0.425*X

```

Conclusion:

By applying the least square method, we successfully determined the equation of the best-fit straight line for the given dataset. This approach minimizes the error between the observed values and the predicted values, providing an optimal linear representation of the data.