

LAB EXPERIMENT # 8: Find the value of a function at a given point from given data set using linear regression of curve fitting

Name & Roll No.: Ashraf Al- Khalique, 1801171	Grade (20)
Registration Number: 351	
Lab Section: C-2	

8.1 Objectives

- To find the value of a function at a given point from given data set using linear regression of curve fitting
- To understand the linear regression method of curve fitting
- To understand MATLAB implementation of the linear regression method.

8.2 Theory

Regression analysis is a statistical procedure for determining the relationship between independent variables (explanatory variables) and dependent variables. It is termed as best fit or curve fitting if we obtain the optimal equation linking these variables.

Linear Regression is used to describe the discovery of a linear relationship between independent factors and dependent variables. When there is just one independent variable in linear regression, the procedure is known as Simple Linear Regression.

A linear regression line has an equation of the form $Y = a + bX$, where X is the explanatory variable and Y is the dependent variable. The slope of the line is b , and a is the intercept (the value of y when $x = 0$).

Where,

$$a = \frac{[(\sum y)(\sum x^2) - (\sum x)(\sum xy)]}{[n(\sum x^2) - (\sum x)^2]}$$

$$b = \frac{[n(\sum xy) - (\sum x)(\sum y)]}{[n(\sum x^2) - (\sum x)^2]}$$

8.3 Apparatus

- MATLAB

8.4 Algorithm

Step: 1 Start

Step: 2 Read Number of Data (n)

Step: 3 . For i=1 to n:

Read Xi and Yi

Next i

Step: 4 Initialize:

Sum1 = 0, Sum2 = 0, Sum3 = 0, Sum4 = 0

Step: 5 Calculate Required Constant a and b of $y = a + bx$:

$a0 = (n * \text{Sum4} - \text{Sum1} * \text{Sum3}) / (n * \text{sum2} - \text{sum1} * \text{sum1})$

$a1 = (\text{sum3} - b * \text{sum1}) / n$

$y1 = a0 + a1 * x$

Step: 6 Display result as a & b

Step: 7 Stop.

8.5 Pseudocode

Start

Read Number of Data (n)

For i=1 to n:

 Read Xi and Yi

Next i

Initialize:

 Sum1 = 0

 sum2 = 0

 sum3 = 0

 sumXY = 0

Calculate Required Sum

 For i=1 to n:

 Sum1 = sum1 + Xi

 sum2 = sum2 + Xi * Xi

 sum3 = sum3 + Yi

 sum4 = sum4 + Xi * Yi

 Next i

Calculate Required Constant a and b of $y = a + bx$:

$b = (n * \text{sum4} - \text{sum1} * \text{sum3}) / (n * \text{sum2} - \text{sum1} * \text{sum1})$

$a = (\text{sum4} - b * \text{sum1}) / n$

Display value of a and b

$y1 = a + b * x$

Stop

8.6 MATLAB Code

```
clc;
clear all;
x=[1 2 4 5 6 8 9];
y=[2 5 7 10 12 15 19];
n=length(x)
sum1=0;
sum2=0;
sum3=0;
sum4=0;
```

```

for i=1:n
    sum1=sum1+x(i);
    sum2=sum2+y(i);
    sum3=sum3+x(i)*y(i);
    sum4=sum4+x(i)*x(i);
end
den=n*sum4-sum1*sum1;
a0=(sum2*sum4-sum1*sum3)/den;
a1=(n*sum3-sum1*sum2)/den;
a0
a1
y1 = a0+a1*x
plot(x,y,x,y1)

```

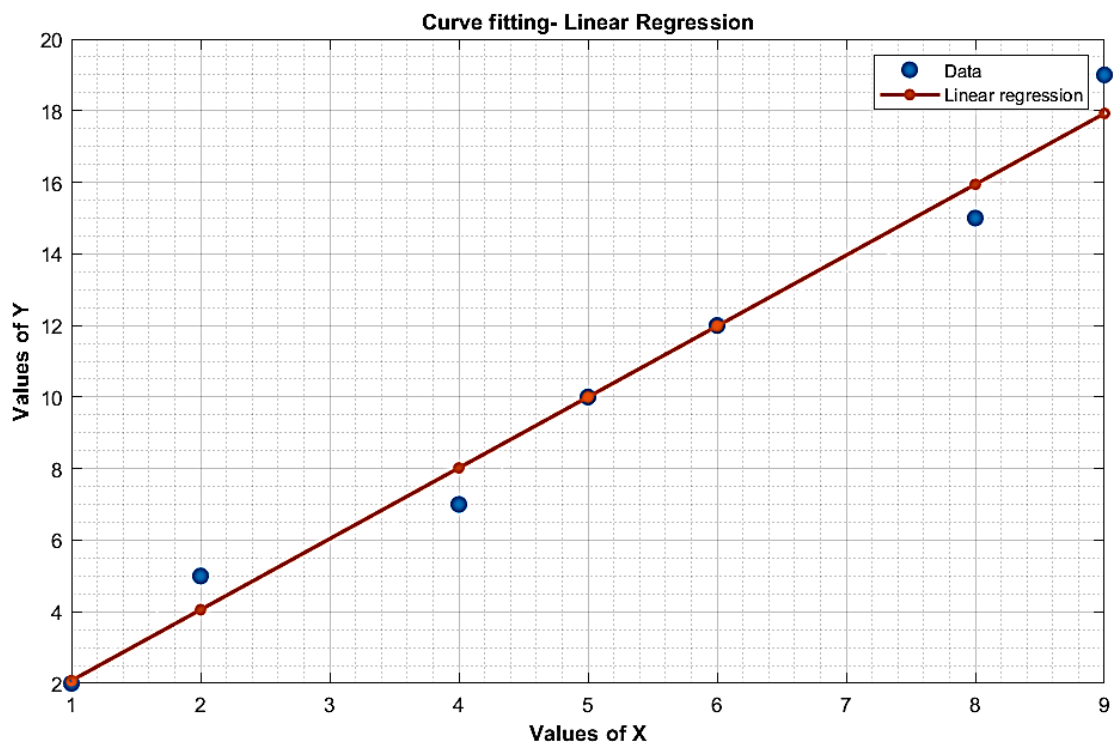
8.7 MATLAB Output

```

a0 = 0.0962
a1 = 1.9808
y1 = 1x7
      2.0769      4.0577      8.0192     10.0000     11.9808     15.9423     17.9231

```

8.8 Graph



8.9 Discussion & Analysis

In this experiment, the concept of linear regression consists of finding the best-fitting straight line through the given points. Here a comparison is observed between the regression line and the given data. This comparison is actually the error of predication. For the regression line where the regression parameters a_0 and a_1 are defined, the line reduces the sum of squared differences between observed values and predicted values.

Here, the regression coefficient (a_0) is the slope of the regression line which is equal to the average change in the dependent variable (Y) for a unit change in the independent variable (X).