Haven's Light is Our Guide

RUET

Rajshahi University of Engineering & Technology

Department of Computer Science & Engineering

Lab Report

Course Code:	CSE 2204
Course Title:	Numerical Methods Sessional
Experiment No:	04
Experiment Name:	Find the straight line that best fits some given data using Least Square Method.

Date:

December 15, 2023

Submitted By:	Submitted To:
Name: Md. Abdullah Al Mamun	Shyla Afroge
Section: A	Assistant Professor
Roll No: 2003028	Computer Science & Engineering
Year: 2nd Year Odd Semester	Rajshahi University of Engineering & Technology

Experiment No: 04

Experiment Name: **Find the straight line that best fits some given data using Least Square Method.**

Theory:

Least Square Method:

The least squares method is a statistical technique used to determine the optimal parameters of a linear regression model by minimizing the sum of the squared differences between observed and predicted values. It aims to find the coefficients (intercept a_0 and slope a_1) that minimize the overall squared residuals in the data. The optimization involves solving the normal equations, resulting in formulas for a_0 and a_1 that provide the best-fitting line. This method is widely applied in various fields for modeling and predicting relationships between variables, as it provides a systematic way to estimate parameters that yield the most accurate linear approximation to the given data.

Algorithm:

Given a set of data points (x_i, y_i) for $i = 1, 2, \ldots, n$:

1. Formulate the linear regression model: $Y = a_0 + a_1x + \varepsilon$, where Y is the dependent variable, x is the independent variable, and ε is the error term.

2. Define the cost function: $J(a_0, a_1) = \sum_{i=1}^{n}(y_i - (a_0 + a_1x_i))^2$, representing the sum of squared differences between observed y_i and predicted values.

3. Find the partial derivatives of J with respect to a_0 and a_1:

$$\frac{\partial J}{\partial a_0} = -2\sum_{i=1}^{n}(y_i - (a_0 + a_1x_i)), \quad \frac{\partial J}{\partial a_1} = -2\sum_{i=1}^{n}x_i(y_i - (a_0 + a_1x_i))$$

4. Set the derivatives to zero and solve for a_0 and a_1:

$$a_1 = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}, \quad a_0 = \frac{\sum y - a_1\sum x}{n}$$

5. The resulting a_0 and a_1 values represent the optimal coefficients for the best-fitting line through the given data points.

Program:

Listing 1: Least Square Method Model

```cpp
double leastSquare(double x[], double y[], int n, double *a1)
{
    double sumX = 0, sumY = 0, sumXY = 0, sumX2 = 0;
    for (int i = 0; i < n; i++)
    {
        sumX += x[i];
        sumY += y[i];
        sumXY += x[i] * y[i];
        sumX2 += x[i] * x[i];
    }
    *a1 = (n * sumXY - sumX * sumY) / (n * sumX2 - sumX * sumX);
    return (sumY - *a1 * sumX) / n;
}
```

Listing 2: Main Program

```cpp
#include <iostream>
using namespace std;

int main()
{
    int n;
    cout << "Enter the number of values: ";
    cin >> n;
    double x[n], y[n];

    cout << "Enter the values of X: ";
    for (int i = 0; i < n; i++)
        cin >> x[i];

    cout << "Enter the values of Y: ";
    for (int i = 0; i < n; i++)
        cin >> y[i];

    double a0, a1;
    a0 = leastSquare(x, y, n, &a1);

    cout << "\nThe value of a0 is " << a0 << " and the value of a1 is "
        << a1 << endl;
    cout << "The equation is: y = " << a0 << " + " << a1 << "x" << endl;
    return 0;
}
```

Result:

The resulting coefficients a_0 and a_1 represent the best-fitting line that minimizes the sum of squared differences between the observed and predicted values. This line provides a linear relationship that can be used for prediction or inference. In summary, the least squares method is a powerful and widely used approach for fitting linear models to data, providing a systematic way to estimate the parameters that best describe the relationship between variables.

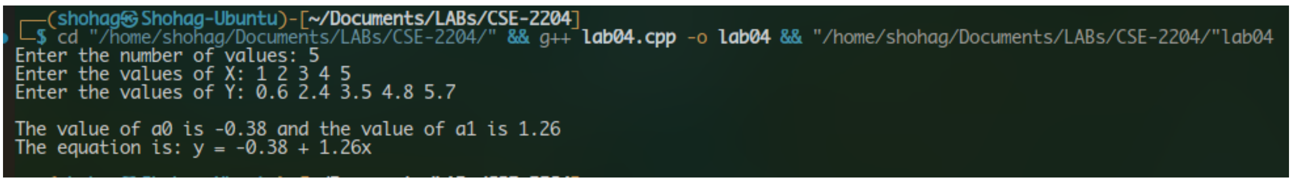

Figure 1: Output of the Program