ALGONQUIN COLLEGE

Computer Engineering Technology – Computing Science

Course: Numerical Computing – CST8233
Term: Fall 2021

Lab #7

• Objectives

The main objective of this lab is to use R program to perform linear regression.

• Earning

This lab worth 1.5% of your final course mark. Each student should complete this lab and demo the codes of the exercises to the lab professor during the lab session.

• Steps

Step 1. Linear Regression

Linear regression is widely used to predict the response to n data point (x_1, y_1) (x_2, y_2) (x_3, y_3) …. (x_n, y_n) by a model given by:

$$y = a_0 + a_1 x$$

where a_0 and a_1 are constants of the regression model. The values of these two constants are calculated using the following formulas:

$$a_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

and

$$a_0 = \frac{\sum_{i=1}^{n} y_i}{n} - a_1 \frac{\sum_{i=1}^{n} x_i}{n}$$

Step 2. Exercise

A. Write an R program that implements the linear regression algorithm. The program asks the user to input the data and then prints the best fit equation. Your output should look like the following test case data.

```
Please enter the number of data pairs:
5

Enter the x-axis values:
20.5
32.7
51.0
73.2
95.7

Enter the y-axis values:
765
826
873
942
1032

The best linear fit is of the form:

y = 3.39487 x + 702.172

No.   x           y(observed)       y(fitted)
-----------------------------------------------
1     20.5        765               771.767
2     32.7        826               813.184
3     51          873               875.311
4     73.2        942               950.677
5     95.7        1032              1027.06
```

B. Calculate the sum of the square of the residuals.

```
The sum of the square of the residuals Sr = 315.077
```

<u>You need to demo this to your lab professor.</u>