CSE2202: Numerical Methods Lab

Online 6 A2

Instructions:
2 sets of problems are given for online lab test.
Set 1: Curve Fitting Using Linear Regression
Set 2: Curve Fitting Using Polynomial Regression.
If you choose Set 1 you will get 20% penalty and for choosing between Set 2 there is no penalty.
After completing your code you must upload you code and output in the following google form link.
Time for Set 1: 30 Minutes
Time for Set 2: 40 Minutes

Problem Description:

There is a relationship between water salinity & water temperature. The following are collected at 400 meter depth from the ocean surface.

T_degCWater = temperature in degree Celsius	Salnty = Salinity in g of salt per kg of water (g/kg)
5.88	34.131
5.87	34.078
7.44	34.223
7.24	34.219
6.68	34.229
6.89	34.263
6.57	34.208
6.24	34.145
6.04	34.123
5.96	34.089
7.39	34.264
7.5	34.261
7.26	34.271
6.93	34.214
6.64	34.194
6.72	34.237
6.91	34.223
6.68	34.278
6.15	34.138
6.29	34.094
7.29	34.241
7.19	34.224
7.39	34.268

1. Write a program that fit a straight line to fit above data and use the equation to predict the relationship between temperatures and salnty using linear regression. Print a table of saltines for the temperature 5 to 6 degree Celsius in step size 0.1 and also calculate Correlation Coefficient R. If the Correlation Coefficient R is *positive* then print strong positive relation, if the Correlation Coefficient R is $negative$ then print strong negative relation and if the Correlation Coefficient R is zero then print no relationship at all. [8]

[Note: Correlation coefficient formulas are used to find the relationship between data. The formula return a value between -1 and 1.]

Equation for calculating Pearson Correlation Coefficient R:

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

$$n = size\ of\ the\ data$$
$$x_i = independent\ varialble$$
$$y_i = dependent\ vaiable$$

2. Write a program that fit a 3rd order polynomial to above data and use the equation to approximate the relationship between temperatures and salnty using polynomial regression. Print a table of saltines for the temperature 5 to 6 degree Celsius in step size 0.1 and also calculate Correlation Coefficient R. If the Correlation Coefficient R is *positive* then print strong positive relation, if the Correlation Coefficient R is $negative$ then print strong negative relation and if the Correlation Coefficient R is zero then print no relationship at all. [10]

[Note: Correlation coefficient formulas are used to find the relationship between data. The formula return a value between -1 and 1.]

Equation for calculating Pearson Correlation Coefficient R:

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

$$n = size\ of\ the\ data$$
$$x_i = independent\ varialble$$
$$y_i = dependent\ vaiable$$

The set of m equations can be represented in matrix notation as follows: $CA = B$

The element of C matrix is: $C(j,k) = \sum_{i=1}^{n} x_i^{j+k-2}, j = 1,2,3,\ldots.m\ and\ k = 1,2,3,\ldots m$

$$B(j) = \sum_{i=1}^{n} y_i x_i^{j-1}, j = 1,2,3,\ldots.m$$