CS 6210 Introduction to Scientific Computing
Assignment 3

<u>Note</u>: For both questions you should use the supplied Matlab codes for as a starting point. The code must compile on one of the lab machines with your instructions. Document your code thoroughly!

1. Filip" data set. This question is a somewhat modified form of question 5.10 of Cleve Moler's book. The data consist of a few dozen observations of a variable y at different values of x. The task is to model y by a polynomial of different degrees from degree 10 upwards in x.

This data set is controversial. A search of the Web for "filip strd" will find several dozen postings, including the original page at NIST, the National Institute for Standards and technology. Some mathematical and statistical packages are able to reproduce the polynomial coefficients that NIST has decreed to be the "certified values." Other packages give warning or error messages that the problem is too badly conditioned to solve. A few packages give different coefficients without warning. The Web offers several opinions about whether or not this is a reasonable problem. Let's see what Matlab does with it.

While the data set is available from the NIST Web site, this data file is also on the canvas page .There is one line for each data point. The data are given with the first number on the line a value of y, and the second number the corresponding x. The x-values are not monotonically ordered, but it is not necessary to sort them. Let n be the number of data points and p = 11 the number of polynomial coefficients.

(a) As your first experiment, edit the file to produce a data file that contains only the x and y data load the write code to read this data file and sort it and plot it . You should see something like the figure of sample output below. [10 Marks]

(b) Write a program that uses the Monomial Vandermode matrix approximation to produce a least squares approximation to work with this data set. In this case n=82 as there are 82 data points and m is the degree of the polynomial with m+1 terms. Use values of m equal to 9,11,13,15,17,19,21,23,25. Make use of the normal equations and matlab QR or the Gramschmidt QR method given . Plot the values of the polynomial by evaluating the Monomial Vandemonde polynomial at (say) 1001 points and use plots in showing how the different polynomials behave. [10 marks]

Note that you do not have to plot every single case.

(c)In each case calculate the least squares error

$$E = \sum_{i=1}^{82} (y_i - p(x_i))^2 \quad \text{where } y_i \text{ are the data values at points } x_i \text{ and } p(x_i)$$ are the

calculated values of the polynomial.. [10 marks]

(d) I found that the QR method did not work well for larger values of m and the normal equations approach did. Use a linear sc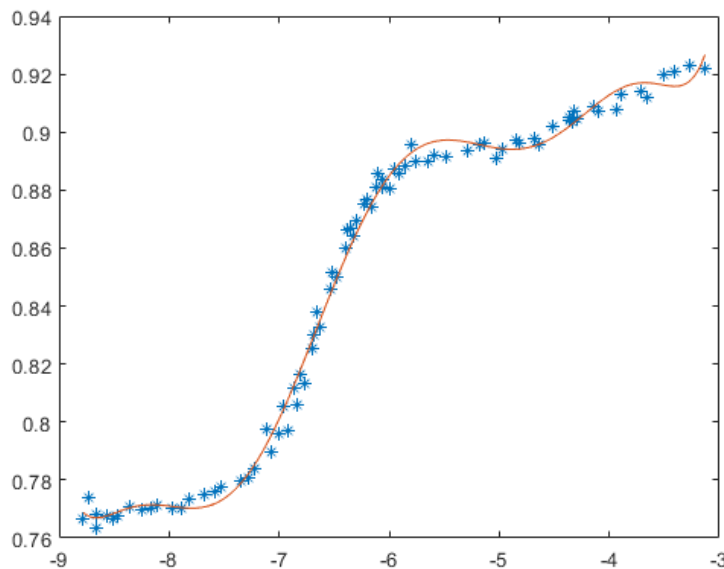aling to put both x and y values in the range [-1,1] such as $y_i = \dfrac{2}{(y_{max} - y_{min})} y_i - \dfrac{(y_{max} + y_{min})}{(y_{max} - y_{min})}$. Repeat the experiments and show that the QR results are better by producing before and after plots. [15 marks]

(e) On the basis of these experiments which value of m is best in terms of accuracy and which method provides the most reliable results? [5 marks]

2. SVD for image processing

    In this question you will use the SVD and randomized factorizations to modify and compress the gatlin and durer images that can be found in matlab.   You   can access the images using the commands below.

    load gatlin

    load durer

(a)

Use the provided code Gatlin_image.m to compute the SVD of both the gatlin I and durer images. Change the colormap from colormap(map) to colormap(gray) and modify the code as needed. Use a rank r = 2, 8, 32, and 128 to represent both images. Show your results.   [10 marks]

(b)

Comment on the performance of the truncated SVD for each image and compare the results. How much storage is required in term of the number of singular values r? [10 marks]

(c)

Use the Randomized SVD code supplied (rsvd.m) to contrast against the ordinary svd code. What are the differences in quality of image, accuracy of singular values and compute time for the values of r as in part (a)

[10 marks]

(d)

Modify the randomized SVD routine to perform q steps of the power method internally where $0 < q < 4$. How does this change image quality, accuracy of singular values and and compute time?

[10 marks]

3. Consider the matrix given in matlab form by

```
[611.0 196.0 -192.0 407.0 -8.0 -52.0 -49.0 29.0;

 196.0 899.0 113.0 -192.0 -71.0 -43.0 -8.0 -44.0;

 192.0 113.0 899.0 196.0 61.0 49.0 8.0 52.0;
 407.0 -192.0 196.0 611.0 8.0 44.0 59.0 -23.0;
  -8.0 -71.0 61.0 8.0 411.0 -599.0 208.0 208.0;
 -52.0 -43.0 49.0 44.0 -599.0 411.0 208.0 208.0;
  -49.0 -8.0 8.0 59.0 208.0 208.0 99.0 -911.0;
   29.0 -44.0 52.0 -23.0 208.0 208.0 -911.0 99.0]
```

Modify your program to show how well (or not) the SVD and rsvd codes perform with this matrix in terms of accuracy of the matrix and of the singular values. Explain your results using eigenvalue information.

[10 marks]

What to turn in

For this assignment, we expect both ***SOURCE CODE*** and a written ***REPORT*** be uploaded as a zip or tarball file to Canvas.

- Source code for all programs that you write, thoroughly documented.
  - Include a README file describing how to compile and run your code.
- Your report should be in PDF format and should stand on its own.
  - It should describe the methods used.
  - It should explain your results and contain figures.
  - It should also answer any questions asked above.
  - It should cite any sources used for information, including source code.
  - It should list all of your collaborators.

This homework is due on November 5th by 11:59 pm. If you don't understand these directions, please use Zoom office hours , send questions to me at mb@sci.utah.edu or contact the TA.