HW 1: Introduction

Due Date: 2/23/2022

Lecturer: Dr. Abd-AlRahman AlMomani *Data-Driven Science & ML*

Part A

Problem 1:
Make a drawing of a circle of radius $r = 1$ by creating a number of points (say 1000) that lies on the circle.

Then, on the same figure, plot the image of a unit circle in R^2 when each point is multiplied by:

$$A = \begin{pmatrix} 3 & -2 \\ -1 & 5 \end{pmatrix}$$

Also overlay the scaled left singular vectors $\sigma_1 u_1$ and $\sigma_2 \mathrm{u}_2$, that you obtain from the SVD of A, on your plot and verify that they line up with the axes of the ellipse. Does the placement of the (center) of the circle matter? Discuss.

Note: No code submission is required for this question, just show your results and discussion.

Problem 2:
Duffing oscillator is an example of a periodically forced oscillator with a nonlinear elasticity, written as

$$\ddot{x} + \delta\dot{x} + \beta x + \alpha x^3 = \gamma\cos(\omega t) \tag{1.1}$$

Review the Duffing oscillator documentation provided on Moodle.

1. Write the system as two, first-order ODEs.
2. Assume we have a dataset (time series) for the 2-dimension system from the previous step.
3. Consider the white-box modeling, meaning that we know that the data comes from the duffing oscillator and we just wish to estimate the values of the parameters $\alpha, \beta, \delta, \gamma$, and most importantly, ω. Explain your strategy, steps, and data arrangement that you will follow to estimate the parameters using the Least Squares Solution.

Problem 3:

Starting from the SVD of a matrix :

$$A = U\Sigma V^* \tag{1.2}$$

explain how we can write the svd in terms of outer product, and explain why the outer product is a useful form.

Part B

Problem 1:
Using the GIF image sequence (PIA02829.gif) that you find in Jupiter dataset, apply the SVD similar to the Navir-Stokes example we discussed in the class.

In your work, show the following:

1. Explain, and show how you arrange the sequence of images in a matrix form, say the matrix A.
2. Show the power curve of the singular values, and discuss the optimal number of singular values that we can use to recover the system accurately.
3. Show the first three left singular vectors (the first three columns of U).
4. Discuss your results and what type of applications this method can be useful for.

Problem 2:
Load the attached dataset set in your matlab workspace, and you will find two matrices: cat, and dog. Each matrix has 80 columns for 80 images of cats and dogs. Each column represent 1 image (size 64x64) reshaped as a column vector.

Follow the Eigenfaces example we discussed in the class to create a classification model. Use the first 70 images for training, and the last 10 images for testing.

Show the results for all the test images, and report the accuracy rate of detection. Discuss your results.

Remember, the purpose is follow the mathematical principles correctly in creating your model. It doesn't matter how accurate your results is. The important is to tell why you got good or bad results.

If you are using Python, read the documentation and use the function (scipy.io.loadmat) to read the mat files.