Project 1, Math 324
Matrix Approximation[1]

Work in groups of two or three people. **All members of the group must contribute to all stages of the project process, from coding to writing.** Each group should turn in one typed report that has the standard format (Intro, Results, Conclusion, Bibliography, as a minimum). The report should be a single **narrative** that includes all relevant information (i.e. it should *not* read 'Question 1,..., Question 2,...'). All equations and figures should be labeled and explained and should be part of a sentence. **Matlab** code should be included when appropriate and should be interlaced with text.

Project Description

1. Load the Saturn image into a matrix and display it. Convert it to a double format gray-scale matrix and display it.

2. Compute the optimal rank 10 approximation to the Saturn image using the Singular Value Decomposition. Display the resulting matrix as an image. Note the corners that appear in the image in the upper left and lower right. Display the unused portion of the decomposition as an image.

3. Display the first 10 rank one matrices given by the Singular Value Decomposition as images. Explain the presence of the corners in the rank 10 approximation.

4. Display the first 10 left singular vectors as images. Compare these images with a column vector from the original image where a corner occurs. Which left singular vector does it appear will represent it best?

5. Compute the norm of the projection of the column vector you used in Question 4 onto each of the first 10 singular vectors. Was your guess in Question 4 correct?

6. Compute the total energy of the matrix.

7. Compute the normalized mean squared error for a specific truncation. When does it become less than 0.1?

8. When are both conditions below satisfied?

    - The normalized mean squared error is less than 0.1, and
    - The ratio of eigenvalues is less than $\delta = 0.01$.

[1]Some problems and terminology for this project are from *Geometric Data Analysis* by Michael Kirby.

9. At what value D is the normalized mean squared error less than 0.01? Do you think most of the important features of the image are captured by the rank D approximation to the image?

10. The normalized mean squared error and the ratio of eigenvalues are used as a means of estimating the dimension of the data. For this data set, what do you think are good values of γ and δ?

11. Identify which are the four most important left-singular vectors for representing the rings in the upper left and lower right corner of the image.

12. Save the image of each group member in the Kenyon directory as an image and load it as a matrix in Matlab. Display a low rank approximation to it and estimate its dimension using your recommended choices of γ and δ above.

Helpful Information

- You can store the Saturn image in a matrix with the command
M=imread('saturn.png');.
Display the image using
figure; imshow(M).
It is a color image, and for the purposes of this project, we will focus on one gray-scale image. Matlab will convert it to a gray-scale image if you use the command
Mgray=rgb2gray(M).
Mgray will be a single matrix in unsigned integer format. To get it to be a matrix in the usual format (on which, for instance, floating point number calculations are allowed), type
Mgray=double(Mgray).

- When you compute the svd, it will save you time and memory space to use the thin svd. That is, use
[U,S,V]=svd(A,0).

- The total energy of a data matrix is defined to be the sum of the eigenvalues of A^TA,

$$E_N = \sum_{i=1}^{N} \lambda_i.$$

- The normalized energy captured by a D-term expansion is given by

$$\frac{E_D}{E_N},$$

where

$$E_D = \sum_{i=1}^{D} \lambda_i.$$

The normalized mean squared error is then

$$\epsilon_{nmse} = 1 - \frac{E_D}{E_N} = \sum_{i=D+1}^{N} \frac{\lambda_i}{E_N}.$$

It is often useful to use a normalized energy criterion,

$$\epsilon_{nmse} < 1 - \gamma,$$

for some γ, typically taken to be in $[0.90, 0.99]$. We also often add a restriction that the following ratio of eigenvalues be less than some constant δ,

$$\frac{\lambda_{D+1}}{\lambda_1} < \delta.$$