M435 Projects in Applied Mathematics
Singular Value Decomposition Project

This sample project explores the application of the singular value decomposition to image analysis. If your final project is limited to this sample then please hand in your work individually. As a group you are expected to extend this project by solving the group computing project below.

The first application concerns the low-rank approximation of a digital image of the planet Saturn. This image will be loaded into a data matrix X.

1 Theory

1. Show that the non-zero singular values of the matrix X are the same as the positive square roots of the eigenvalues of XX^T and X^TX.

2. Use the SVD factorization to prove that the row rank and column rank of a matrix are the same.

3. Describe the meaning of a plot of the eigenvalues of XX^T in decreasing order, i.e., largest to smallest.

4. What can you infer about a singular vector if it is associated with a singular value that is zero? Answer this by considering the projection of the data onto this vector.

5. Show that the left (and right) singular vectors may be viewed as data-dependent by expressing them as sums of the columns (rows) of the data matrix X.

6. It is often stated that one can truncate eigenvectors associated with small eigenvalues. Argue with illustrative examples when this is a good idea and when it is not.

2 Computing

The Saturn image for this problem may be downloaded from the class web page. The format of the data is "uint8", which stands for unsigned integer, 8 bits. Before you apply the SVD, change it to "double" format:

```
A=imread('saturn.tif')
A=double(A);
```

To display the image use

```
imagesc(A);
colormap(gray);
```

Use the singular value decomposition to construct a low-rank approximation to Saturn.

```
[U,D,V] = svd(X);
```

1. Plot the squares of the singular values in decreasing order. Based on this graph can you comment on how many terms you might expect to need to accurately reconstruct Saturn?

2. Plot the rank one images for several different eigenvalues (large, medium, and small) and compare.

3. Recalculate the SVD of the data matrix of Saturn but first subtract the ensemble (row) average. Compare your singular values and first 4 rank one images and comment.

4. You may observe that the rings of Saturn are poorly approximated at the edges. Plot the 4 left singular vectors that are most important (based on maximum mean squared projection) for "fitting" the rings. Repeat for the right singular vectors and comment.

5. Can you modify the ordering of your basis to better approximate the rings? Show your results.

6. Write a subroutine to automatically identify whether a column of the picture contains only rings or both planet and rings.

3 Group Computing Problem

Select a data set to analyze using the singular value decomposition. You may generate this data set using a simulation, e.g., solutions to a system of differential equations, or obtain it from a data repository such as the one available from the Machine Learning Lab at UCI:

http://mlearn.ics.uci.edu/MLSummary.html)

Answer the following questions about the data:

- Based on examination of the singular values, is a reduced representation possible?
- Illustrate the errors incurred with your reduced representation by showing a reduced rank approximation.
- Can this reduced representation be used to assist in data classification? Show your results.