

## Problem Set 6: SIO 221B, Data Analysis

due Wednesday, November 27, 2002

1. Demonstrate that the relative sizes of the singular values of a matrix provide a measure of the variance represented by each singular mode. To do this, first generate an  $N$  by  $M$  matrix  $\mathbf{G}$  consisting of random numbers. Then compute the singular value decomposition of  $\mathbf{G} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$ . (Matlab will do this with the “svd” command.)

a. Plot the singular values.

b. Define a matrix  $\mathbf{H}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k$ , which is the reconstruction of  $\mathbf{G}$  using only the  $k$ th singular value and singular vectors. Compute the variance of all values in  $\mathbf{H}_k$  and  $\mathbf{G}$ . What is the relationship between the variance of  $\mathbf{H}_k$  and the  $k$ th singular value?

c. Now show (using equations) that the L2 norm of  $\mathbf{H}_k$  is the  $k$ th singular value,  $\lambda_k$ .

2. The temperature along a transect in the ocean is believed to satisfy a linear rule  $\theta = ar + b$ , where  $r$  is latitude measured in kilometers from a reference point, and  $a, b$  are constants. Measurements at sea yield the following table of values:

$$\begin{aligned} r &= 0, & \theta &= 10 \\ r &= 1, & \theta &= 9.5 \\ r &= 2, & \theta &= 11.1 \\ r &= 3, & \theta &= 12 \end{aligned}$$

a. Using ordinary least-squares, find an estimate of  $a, b$ , and estimate the noise in each measurement, and the standard errors of the estimated values of  $a, b$ .

b. Using the singular value decomposition (you can use an SVD program if you wish), also find an estimate of  $a, b$ , and its standard error. Which of the equations proved most important to the solution? Are the elements of the solution fully resolved?

3. From the course web site:

<http://www-mae.ucsd.edu/~sgille/sio221b/homework.html>

download the matlab data set “hw6.mat” or the individual ASCII files “ueq.dat” and “loneq.dat”.

The variable “ueq” is a 480 by 45 matrix containing monthly zonal velocities on the equator at 22 m depth from a 40 year run of the NCAR CSM Ocean model. The variable “loneq” identifies the longitudes of the velocities.

a. Represent “ueq” in empirical orthogonal modes. An easy way to do this is to use the singular value decomposition. Explain your methodology.

b. Plot the first 3 temporal modes and the first 3 spatial modes of the EOF.

c. What fraction of the variance in “ueq” is captured by the first 3 EOF modes?