

Stat SVD

Homework: SVD with Matlab

1.) Use Matlab or some other canned routine set on a workstation to calculate all of the major subspaces of the following two 3 x 3 matrices using SVD. Discuss.

a) $A = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}$

b) $B = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 6 & 2 \\ 2 & 2 & 6 \end{bmatrix}$

2.) Two data sets for you to analyze. Get the text files x552a.dat and x552b.dat and put them in your working directory.

Now you are ready to start matlab. Once matlab is up you can ingest the data files by typing 'load x552a.dat' and this will give you the data as a matrix called x552a. Now you can SVD this matrix. This blows a bunch of stuff onto the screen. Be sure you use the method '[u,s,v]=svd(x552a,0)' so you don't get the null spaces printed. The matrices are 512x32, think of the long dimension as time and the short dimension as space. I arranged the matrices so that the longer column vectors print out first, and the spatial vectors will be at the end of the screen dump. You will want to plot the eof's, pc's and the singular values. You can do this using the command 'plot(v(:,1:4))', which plots the first 4 columns of v (or substitute u), and you can plot the singular values using 'plot(s)'. Matlab will put the spatial(dimension 32) structures into V and the temporal structures into U. The shorter dimension is structure and the longer dimension is sampling dimension. To print a copy of one of these graphs type 'print' while it is in the plot window. I think that's all you really need to know, unless you want to label the plots, which is probably a good idea. After you type a command Matlab prints the answer on the screen by default. To avoid this, end your commands with a semicolon ;.

Note: You can find a description of all the Matlab functions at the following web site. <http://www.mathworks.com/access/helpdesk/help/techdoc/matlab.shtml>

a) Evaluate the statistical significance of the singular values using the North method and discuss. OK, to do this you need to estimate the true number of degrees of freedom in the sample. Here is a little help file for doing this, with an example. You will need the auto.m function . [Homework Helper](#) autocorrelation function supplied by Matlab.

You can also now get the software at [Matlab Example Software for Fun and Homework!](#)

b) Interpret the dominant structures that you derive. Be sure to look at both the spatial structure (eof) and the corresponding temporal structure (pc time series). Discuss.

c) Compare and contrast the two data sets. Hint: One is red noise and the other is red noise plus some other stuff, which you are supposed to identify. Apply any other logical or statistical tests you can

think of to make sense of these data sets. Describe in words what you think I did to make data sets a and b.

d) Put your results in a little report, and include relevant figures and statistical analysis.

END