Class summary/homework 8 – Due Friday April 25th, 2014 9:00am

1. Download from the web page the new script class21a.R and download two data files **class21a_tavg_Jun_1950-2012.csv**, **class21a_tavg_Jan_1950-2012.csv**. You can work with one of the two data files. These files contain monthly mean temperature anomalies from 20 (24) stations in NY from 1950-2012.
   (http://www.atmos.albany.edu/facstaff/timm/ATM315spring14/R)
2. Observe how the Principal Component Analysis (eigenvectors and eigenvalues of the Covariance matrix) change the representation of the stations' temperature variability. What happens to the total variability, and what does it mean when we see the variance of the first (few) eigenvectors 'piling up'. Take note of what happens to the total variance. Explain the observed behavior.
3. Eigenvectors are orthogonal (uncorrelated). What can be said about the first eigenvector structure (i.e. the values in the eigenvector) compared with the second.
4. Try to give a meaning to the 'Principal Component' (PC) time series. (Observe the correlation between the PC 1 time series and the station-averaged time series).
5. Conduct some systematic experiments: Control the number of stations used in the PCA (see line 63 : m<-2), and control the level of random error (noise line 36-39) that is added to the temperature time series. How does it affect the results of the PCA?
   Does it change the concentration of variance in the first (second) eigenvalues? What effect has a large random error on the correlation between PC time series and the station average time series.

Suggested reading:

http://www.atmos.albany.edu/facstaff/timm/ATM315spring14/wilks-pca1.pdf (11.1.1)