# Week 3 Exercise, due Wed wk 4

- Create a matrix M0 of 100 time series (columns), each 1000 points long, and fill with Gaussian random numbers with standard deviation 1.
- Perform an EOF analysis on M0. Plot log10 of the percent variance explained by the 100 EOFs and note the percent variance explained by the first EOF.
- Create a single column matrix y and fill with Gaussian random numbers with standard deviation 0.2
- Add y to each column of M0 to give M1.
- Perform an EOF analysis on M1. Plot log10 of the percent variance explained by the 100 EOFs and note the percent variance explained by the first EOF.
- The percent variance of M1 explained by y (not by the EOF) should be $100*0.2^2/(1^2+0.2^2) = 3.85\%$, compare this to the percent variance explained by the first EOF and explain the difference.
- Calculate an approximation M2 to M1 using only the first EOF. Work out the percent variance of y explained by any column m2 of M2, i.e. $100 \times \left[1 - \left(\frac{\sigma(y-m_2)}{\sigma(y)}\right)^2\right]$ where σ(x) means standard deviation of (x). Repeat this calculation using the first column of M1 (NB lots of noise often results in negative % variance explained).
- Give the values you have calculated, and explain why the EOF produces such a good estimate of y.
- NB all percentages should be given to 2 decimal places. For standard deviations, use std(x,1), because std(x) gives the unbiased estimate, where you divide by n-1, but for these 'variance explained' calculations we want the sample standard deviation, i.e. dividing by n, which you get from std(x,1).