LECTURE 9

A third technique for estimating the significance of EOF is North's rule of thumb.

(North et. al 1982: show that the 95% conf. limits on the eigenvalues is approximately:

$$\Delta\lambda = \lambda\sqrt{\frac{2}{N^*}}$$

where λ is the eigenvalue and N^* is the number of d.o.f. So d.o.f. still needs to be estimated (like in the Monte Carlo example) but the error estimate is straightforward.

With errorbar estimates on the eigenvalues, a plot of eigenvalues will show those which are the same to within statistical certainty.

Gaps

The data matrix must be full, i.e. no gaps, for SVD

The gaps can be filled:

(1) interpolation

e.g. loess

optimal interpolation

done right, this will be an unbiased estimate

(2) zero

(3) random numbers with $\langle d \rangle$ and σ^2 estimated for each time series
- should really include effect of autocorrelation
- run several EOF analyses for different random samples & compare results.

For EIG solution, make estimate of elements $c_{i,j}$ from available data.

Rotation of EOFs

Sometimes the orthogonality constraint will cause structures to have significant amplitude all over the domain.

This is because orthogonality, combined with the requirement that the EOF explain as much variance as possible over the whole domain, can yield artificial global-looking structures.

If we expect the structures to be more local, we might be able to get a more compact & more physically interpretable set of structures by rotation of the EOFs.

This will sacrifice one or both of the orthogonalities.

Orthogonal rotations can be applied to retain the orthogonality property of the loadings at the cost of creating time series correlated in time.

There are many rotation criteria developed for different objectives. The most common method used in oceanography & meterology is the varimax rotation:

The objective of varimax is to produce "simplicity" in the loading pattern by driving loading pattern values toward either 0 or 1.

This says variability within a certain location is either related, or not.

Varimax maximizes the variance of the squared loadings, which tends to produce loadings which are half 0s and half 1s (as much as possible) thereby creating a set of loadings with high contrast.

The rotation procedure can be thought of a transformation of the data factorization:

$$D = EA$$

to $$D = E\underbrace{RR^{-1}}_{=I}A$$

The rotation $E^* = ER$ achieves the varimax criterion while retaining orthogonality of E^*:

... but the orthogonality (uncorrelatedness) of the rows of A^* $R^{-1}A$ is lost

The variance explained by each mode alters, with rotated mode 1 always having less variance than unrotated.

Loading patterns become more local, with possibly easier physical interpretation

Lagged correlations between time series may assist interpretation