ATOC 7500: The Art of Climate and Environmental Modeling
Week 14 Lab Assignment
Spatial Analysis

Identify climate patterns using Empirical Orthogonal Functions (EOF), and determine how they are related to Colorado precipitation.

Decomposition of data (X) into orthogonal basis functions to reduce the dimensions of data - i.e. the data can be described by a linear combination of basis functions (**v**) –

$$X(\boldsymbol{x},t) = \sum a(t)\, \mathbf{v}(\boldsymbol{x})$$

Basis functions are linearly uncorrelated with each other, and are called 'EOFs'. The amplitudes $a(t)$ are time-series called 'principal components'.

Data reconstructed from the first few basis functions and amplitudes will approximate the original data, but have simplified the number of degrees of freedom. The more basis functions that are included, the closer the reconstruction will be to the original data (c.f. spherical harmonics and truncation). We will choose the basis functions which maximize the amount of variance along the primary axis.

The eigenvector of a square matrix **A** is a non-zero vector that, when multiplied by **A** yields the original vector, multiplied by a single number called the eigenvalue.

Therefore, the eigenvalue problem -

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$$

where $\mathbf{v}_i$ is the eigenvector (of a square matrix **A**), λ_i is the scalar eigenvalues.

$$\mathbf{A}\mathbf{v} = \boldsymbol{\lambda}\mathbf{v}$$

where **v** is the matrix with the eigenvectors $\mathbf{v}_i$ as the columns, and $\boldsymbol{\lambda}$ is the matrix with the eigenvalues λ_i along the diagonal, and zeros elsewhere.

The set of eigenvectors and associated eigenvalues represent a coordinate transformation, such that matrix **A** becomes diagonal. The eigenvalues indicate how much variance is explained by each eigenvector – the eigenvectors define directions in the original coordinate space along which the maximum possible variance can be explained. The new directions are uncorrelated with each other. Convention is to order the eigenvalues and eigenvectors in decreasing order, therefore you can explain most of the variance with the first few values.

Note that λ_j is equal to the variance explained by the corresponding eigenvector $\mathbf{v}_j$.

Solve for the eigenvalues -

$$\det(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) = 0$$

where **I** is the identity matrix.

For the EOF analysis **A** is the correlation or covariance matrix of **X**, the original data;

$$\mathbf{A} = \mathbf{X}^T\mathbf{X}$$

Solve for the corresponding eigenvectors **v** and eigenvalues $\boldsymbol{\lambda}$ as above.

We can project a single eigenvector onto the data to get the amplitude of the eigenvector at each time;

$$\mathbf{v}^T\mathbf{X} = \mathbf{a}$$

where $\mathbf{v}^T$ is the eigenvectors in the columns of a square matrix, and $\mathbf{a}$ is a vector of the principal component time-series for each EOF.

Tasks -

- Read in monthly mean data over specific region for 1980-2010
- Subtract mean from data
- Calculate the correlation matrix
- Solve for the eigenvectors
- Plot maps of the first two EOF patterns
- Plot the percentage variance explained by first 10 EOFs
- Project patterns onto data to calculate the time-series associated with each pattern
- Plot the amplitude vs time for the first two PCs
- Correlate each of the time-series with precipitation over North America

IDL routines

- NetCDF files – open, read variable, close
```
fileid = ncdf_open(filename)
varid = ncdf_varid(fileid, 'Z')
ncdf_getvar(fileid, varid, zg)
ncdf_close, fileid
```
- pcanal.pro to calculate PC patterns and time series
```
/home/atoc/sala7581/pcanal.pro
```
- Plot maps
```
map_set,30,300,,/isotropic
contour, something
map_continents
```

Discussion questions

- How much variance do the first three EOFs account for?
- How many EOF are required to reconstruct 70% of the data?
- What does the first EOF pattern look like?
- Can we tell what next winter will be like?