S-mode and T-mode EOFs from a GCM modeller's perspective: Notes on the linear algebra

R.E. Benestad

DNMI, February 7, 2003

Contents

1 Definitions

The vectors are written as $\vec{x}$ and matrices are denoted by using the capital letters: $X = [\vec{x_1}, \vec{x_2}, .., \vec{x_T}]$. The vector quantities are used to represent several observations at a given time, i.e. they can be regarded as maps. Let the *number of observers* mean the number of grid points or stations where observations are made (number of observers = R), and the *number of observations* be the length of the time series at each location (number of observations = T). We use the notation $\overline{x}$ to mean the temporal mean of x and $\langle x \rangle$ the spatial (ensemble) mean of x.

2 S-mode

Let the matrix X_{rt} contain T observations from R different locations, where X can be expressed in the form $X = [\vec{x_1}, \vec{x_2}, .., \vec{x_T}]$ and $\vec{x_t} = [x_1(t), x_2(t), .., x_R(t)]$. Each column represents one set of observations, with each element holding the data from the r different locations:

$$X = \begin{pmatrix} .. & \rightarrow & T \\ \downarrow & .. & .. \\ R & .. & .. \end{pmatrix}. \quad (1)$$

Let anomalies in X be defined as:

$$X'_{rt} = X_{rt} - \Sigma_{t=1}^{T} X_{rt} = X_{rt} - \overline{X_r}. \quad (2)$$

The variance-covariance matrix is defined as

$$C_{rr} = X'X'^T = \begin{pmatrix} .. & \rightarrow & R \\ \downarrow & .. & .. \\ R & .. & .. \end{pmatrix}. \quad (3)$$

The S-mode Empirical Orthogonal Functions (EOFs) of Xrt are defined as:

$$C_{rr}\vec{e_s} = \lambda \vec{e_s}. \quad (4)$$

Let $E_s = [\vec{e_1}, \vec{e_2}, .., \vec{e_{R^*}}]$ be a matrix with the columns holding the eigenvectors (EOFs) and R^* be the rank of X. The data may be expressed in terms of the orthogonal set spanned by the EOFs:

$$F = EY. \quad (5)$$

where Y is the projection of F onto the EOF space.

We have used singular value decomposition (SVD) to compute the EOFs. Using SVD, we can express the matrix X' as:

$$X' = U\Sigma V^T. \tag{6}$$

Note that the SVD algorithm is written in such a way that the numbers of columns must be less than number of rows. In this example, the number of observers are assumed to be greater than the number of observations (which often is the case for gridded climate data). If the number of columns is greater than the number of rows, then the SVD must be applied to the transpose of the matrix (U and V will now by swapped). The columns of U and V are orthogonal respectively:

$$U^T U = V^T V = I. \tag{7}$$

The matrix Σ is a diagonal matrix, with R^* non-zero singular values and $R - R^*$ zero values in descending order along the diagonal. The inverse of Σ is a diagonal matrix with the reciprocal of the non-zero singular values along the diagonal. The reciprocal of the small singular values or zeros are taken to be zero.

The variance-covariance matrix can be expressed in terms of the SVD products:

$$C_{rr} = X'X'^T = U\Sigma V^T(U\Sigma V^T)^T = U\Sigma V^T(V\Sigma U^T) = U\Sigma^2 U^T. \tag{8}$$

A right operation of U gives:

$$C_{rr}U = U\Sigma^2. \tag{9}$$

$$U\Sigma^2 = \begin{pmatrix} .. & \rightarrow & R \\ \downarrow & .. & .. \\ R & .. & .. \end{pmatrix} \begin{pmatrix} .. & \rightarrow & T \\ \downarrow & .. & .. \\ R & .. & .. \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 & 0 & .. \\ 0 & \sigma_2^2 & 0 & .. \\ .. & 0 & \sigma_n^2 & .. \end{pmatrix}. \tag{10}$$

$$C_{rr}\vec{u} = \sigma^2 \vec{u}. \tag{11}$$

Hence, $U = E_s$ and $\sigma^2 = \Lambda$, and the SVD routine applied to X gives the S-mode EOFs of X.

The S-mode, described above, has been employed where spatial EOF maps have been discussed.

3 T-mode

The spatial variance-covariance matrix is defined as

$$C_{tt} = X'^T X' = \begin{pmatrix} .. & \rightarrow & T \\ \downarrow & .. & .. \\ T & .. & .. \end{pmatrix}. \tag{12}$$

The T-mode Empirical Orthogonal Functions (EOFs) of Xrt are defined as:

$$C_{tt}\vec{e_t} = \lambda \vec{e_t}. \tag{13}$$

The spatial variance-covariance matrix can be expressed in therms of the SVD products:

$$C_{tt} = X'^T X' = (U\Sigma V^T)^T U\Sigma V^T = (V\Sigma U^T)U\Sigma V^T = V\Sigma^2 V^T. \tag{14}$$

A right operation of V gives:

$$C_{tt}V = V\Sigma^2. \tag{15}$$

Hence, $V = E_t$ and $\sigma^2 = \lambda$, and the SVD routine applied to X also gives the T-mode EOFs of X.

The T-mode has been employed where temporal evolution of coherent spatial structures have been discussed. The T-mode forms the basis for both canonical correlation analysis (CCA) and regression.

4 Autocorrelation considerations

4.1 Spatial coherence

4.1.1 Degrees of Freedom

If there is spatial coherence, i.e. that observers at some locations r, $x_r(t)$, are correlated (spatial autocorrelation), then the actual number of independent spatial observations is smaller than the number of observers. We assume that the data are uncorrelated in time, i.e. X consists of independent temporal realisations. Hence, the principal component analysis (PCA) may represent the data in terms of a small number of EOFs describing the coherent spatial structures with similar "behaviour". Each of these structures, or maps, is associated with a time evolution: $Y_t = E^T X$.

The number of independent realisations in R is often smaller than the (effective) time dimension. Therefore, the estimation of the spatial variance-covariance matrix tends to be associated with large sampling errors. In this case, the S-mode is preferred method.

4.1.2 Geographical weighting

It is important to apply a geographical weighting factor if the data is represented on grids that cover large latitudinal ranges, as the boxes (on a regular lon-lat grid) near the poles tend to represent a much smaller area than those near the equator. Unweighted data will therefore give too much weight to polar regions. Similarily, for a net work of unevenly distributed observers, a weighting function must be applied in order to ensure equal contribution from each observer.

4.2 Temporal coherence

If there is serial temporal correlation, but no correlation between each observer, then the actual number of independent observations is smaller than R. The EOFs hence yield a smaller set of temporal structures, or "trajectories". Each of these trajectory is associated with a spatial structure given by: $Y_r = E^T X^T$.

The number of independent realisations in T is often smaller than the (effective) spatial dimension. Therefore, the estimation of the variance-covariance matrix tends to be associated with large sampling errors. In this case, the T-mode is preferred method.

5 Spatial anomalies

We have so far only considered anomalies where the temporal mean value at each location is subtracted from the respective time series. It is also possible to perform EOF analysis on "spatial anomalies" where the mean observation at time t, $\langle \vec{x}(t) \rangle$, is subtracted from all observations at this time:

$$X^+_{rt} = X_{rt} - \Sigma_{r=1}^R X_{rt} = X_{rt} - \langle X_t \rangle. \quad (16)$$

Whereas the temporal (the usual definition of) anomalies captures trends in time (such as a global warming) and oscillations, EOF analysis based on spatial anomalies will be insensitive to the evolution of global mean values. The PCA on spatial anomalies, on the other hand, will be sensitive to large spatial gradients, although oscillating structures that have sufficiently small scales to produce large spatial variance (heterogeneous structures) will also be captured by the spatial anomaly EOFs.

6 Further reading

EOF analysis is commonly used among geophysicists, and there is a large number of references giving further details about EOF analysis and related mathematical considerations. *Press* et al. (1989) and *Strang* (1995) discuss the SVD algorithm in terms of numerical solutions and linear algebra respectively. *Anderson* (1958) gives an account of principal component analysis from a statistical point of view on an advanced level, whereas *Wilks* (1995) gives a simpler introduction to EOF analysis. *Preisendorfer* (1988) is a commonly used text, giving detailed recipes on how to do the calculations, and *Peixoto & Oort* (1992) gives a brief overview of EOF analysis in one appendix.

References

Anderson, T.W. 1958. *An Introduction to Multivariate Statistical Analysis.* 1 edn. New York: John Wiley & Sons, Inc.

Peixoto, J.P., & Oort, A.H. 1992. *Physics of Climate.* AIP.

Preisendorfer, R.W. 1988. *Principal Component Analysis in Meteorology and Oceanology.* Amsterdam: Elsevier Science Press.

Press, W.H., Flannery, B.P., Teukolsky, S.A., & Vetterling, W.T. 1989. *Numerical Recipes in Pascal.* Cambridge University Press.

Strang, G. 1995. *Linear Algebra and its Application.* San Diego, California, USA: Harcourt Brace & Company.

Wilks, D.S. 1995. *Statistical Methods in the Atmospheric Sciences.* Orlando, Florida, USA: Academic Press.