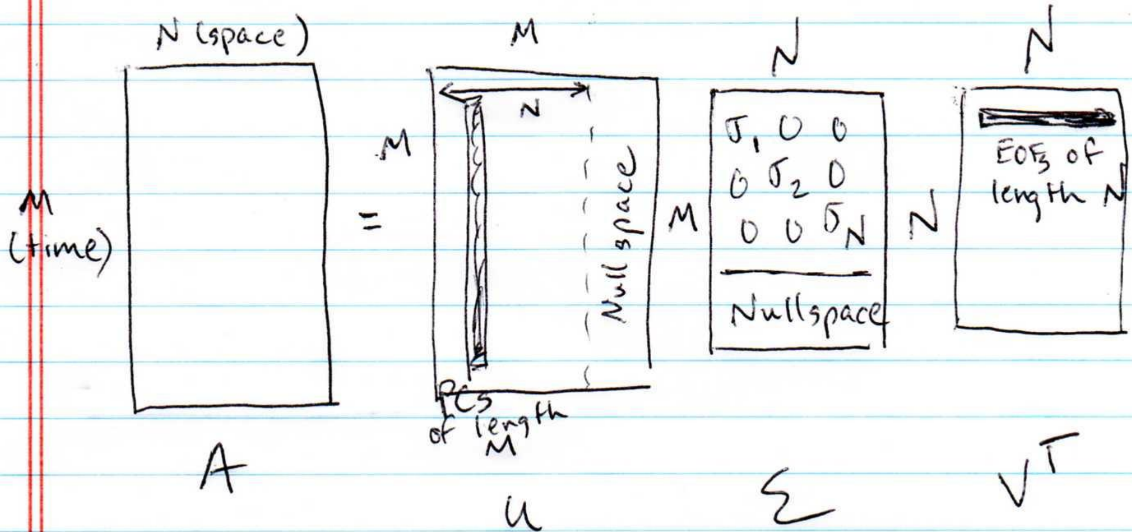
Example

If $M > N$ and A is full rank.

* Reverse the logic if $N < M$, which is typically the case ...

- U spans R^m ; V spans R^N
- N columns in V span R^N
- 1st N columns in U span the column space of A.
- $N+1 \rightarrow M$ columns in U span the null space of A.

Graphically



$u_1 \rightarrow u_N$ spanning A's column space (PCs)

$u_{N+1} \rightarrow u_M$ Null space not representable in terms of A's columns

$V_1 \rightarrow V_N$ spans A's row space (EOFs)

Physical interpretation

If the σ's are arranged such that σ_1 is the largest value:

- → 1^{st} column in U is the leading PC of A.
- → 1^{st} column in V is the leading EOF of A.

EOF modes

$$A = U \Sigma V^T$$

Can be written

$$A = \sum_{i=1}^{N} \sigma_i u_i v^T$$

Schematically:

$$\underset{(M \times N)}{A} = \sigma_i \underbrace{\begin{bmatrix} u_1 \\ \downarrow \end{bmatrix} \begin{bmatrix} v_1 \longrightarrow \end{bmatrix}}_{\text{EOF mode } 1} + \cdots$$

$$\sigma_N \underbrace{\begin{bmatrix} u_N \\ \downarrow \end{bmatrix} \begin{bmatrix} v_N \longrightarrow \end{bmatrix}}_{\text{EOF mode } N.} \cdots$$

Equivalence of SVD to Eigenanalysis (or diagonalization) of covariance matrix. (proof)

SVD : $A = U \Sigma V^T$

Square both sides of $A = U \Sigma V^T$

$$A^T A = V \Sigma^T U^T U \Sigma V^T$$

$$U^T U = I \qquad V^T V = I$$ (since they're symmetric)

$$A^T A = V \Sigma_N^2 V^T$$

or equivalently, since $A^T A =$ Cov. matrix

$$C V = V \Sigma_N^2$$

$$C V = V \lambda$$ → Original eigenvalue problem

Cov. matrix   Eigen- vectors   Eigenvalues.

Relationship between EOFs / PCs.

$$A = U \Sigma V^T$$

Multiply by V

$$AV = U \Sigma$$

Hence, for an individual mode:

$$A v_i = \sigma_i u_i$$

Physically: PCs are found by projecting the data on to the EOFs and vice versa.

Summary of finding EOFs / PCs via SVD:

1) Decompose matrix A such that

$$A = U \Sigma V^T$$

2) for A $(M \times N)$
   if M is the sample space:
   - EOFs (spatial patterns) correspond to the columns in V
   - PC (time series) correspond to the columns in U.

3) The fraction of variance explained by the i^{th} EOF / PC pair is:

$$\frac{\sigma_i'^2}{\sum_{i=1}^{N} \sigma_i'^2} \quad \text{or} \quad \frac{\lambda_i'}{\sum_{i=1}^{N} \lambda_i'}$$

Presentation of EOFs

Whether the EOFs or PCs are shown depends on which domain yields interesting structure.

Problem with looking directly at EOFs is that they are normalized, so don't get a sense of physical units.

Common way to show the EOF spatial patterns is to regress the data (i.e A matrix) onto the standarized PC time series.

Steps

1) Get the PC $\begin{cases} \text{via SVD} \\ \text{projecting A onto EOFs via diag. method} \end{cases}$

2) Standardize the PC so it has a variance = 1 (unit variance)

$$PC = \frac{PC}{\sqrt{\overline{PC'^2}}} \rightarrow \text{std. dev. of PC.}$$

3) Regress the data onto the $\hat{PC}$ (normalized) yields the EOF scaled in actual units.

* Assumes a linear regression.

→ So don't show the "raw" EOF

In the case of SVD.

$$A = U \Sigma V^T$$

Multiply by U^T (i.e. project PCs)

$$U^T A = \underbrace{U^T U}_{\substack{\text{yields} \\ \text{identify matrix}}} \Sigma V^T$$

$$U^T A = \Sigma V^T$$

Multiplying the orthonormal EOFs by corresponding singular values equivalent to projecting data on the orthonormal PCs.

<u>Caveats</u>

→ "Orthonormal" is not the same as standardized

→ σ^2 (or λ), that is the eigenvalues, are useful for telling % variance explained, but <u>not</u> for scaling EOFs.

EOFs using the correlation matrix

Recall get the correlation matrix instead of the covariance matrix ($A^T A$) if the time series of the data are standardized before the analysis, typically assuming a normal distribution.

$$z = \frac{x - \mu}{\sigma}$$

μ = time avg. over M time intervals

σ = std. deviation over M time intervals.

Do this at each point.

Reasons to do:

1) If the state vector is comprised of different parameters
2) If the variance of the state vector varies greatly from one location to another.

$\Rightarrow$ Generally NOT the case when working with space-time data.

In general, suggest retaining amplitude in data because:

1) EOFs of unnormalized data explain more variance.
2) Centers of unnormalized EOFs are shifted from actual EOFs. (i.e. spatial pattern may change).

* The spatial patterns and gradients in variance matter, as they're linked to the spatial structure of the field.

Possible disadvantage: Wipe out ~~?~~ regions of small variance (e.g. mid-latitudes vs. tropics)

How many EOFs should be retained?

"White noise" = time series with equal power at all periods (i.e. no dominant time frequency in the data) and no spatial structure.

In such a case, all the eigenvalues (λ) would be equal to one. Therefore no one pattern would be any more significant than another pattern. ~~[illegible]~~ off diagonal ~~[illegible]~~ elements in the covariance matrix $= 0$. because no spatial structure.

The processes in the atmosphere, and most geophysical systems, are "red noise."

"Red noise": Because data are autocorrelated in time, the largest space scales and lowest frequencies tend to explain more of the variance than smaller scales and higher frequencies.