EOF Analysis

- Interpretation of EOFS
  - localized vs. global / REOF
  - REOF calculation (?) Varimax
  - separation EIG spectra

- Applications
  - data compression – e.g. one or two words
  - noise filtering – wave pattern / matrix inversion
  - statistical prediction – homework example
  - multivariate dynamics – Pacific analysis
  - low-order dynamical systems

  AR1 – or

- EEOF / Complex EOFS

$$\underline{x} = \begin{bmatrix} \underline{x}_{t+1} \\ \underline{x}_t \end{bmatrix} \quad \langle\ \rangle \quad \underline{x}\,\underline{x}^T = \cancel{C_{t+1}, C_{t+1}} \begin{bmatrix} C_{t+1,t+1} & C_{t+1}, C_t \\ C_{t,t+1} & C_{t,t} \end{bmatrix}$$

4/

Signal Decomposition

- Empirical Orthogonal Function
- Principal Component Analysis
- Factor Analysis

$\left.\right\}$ Eigenmodes Methods

Supose you have

$f(s,t)$ → space (s) time (t) array
→ parameter time array

$$\underline{\underline{F}}_{s,t} = \begin{bmatrix} s & \xrightarrow{\quad t \quad} & T \\ \Big\downarrow & & \\ S & & \end{bmatrix}$$

T ← number of time element

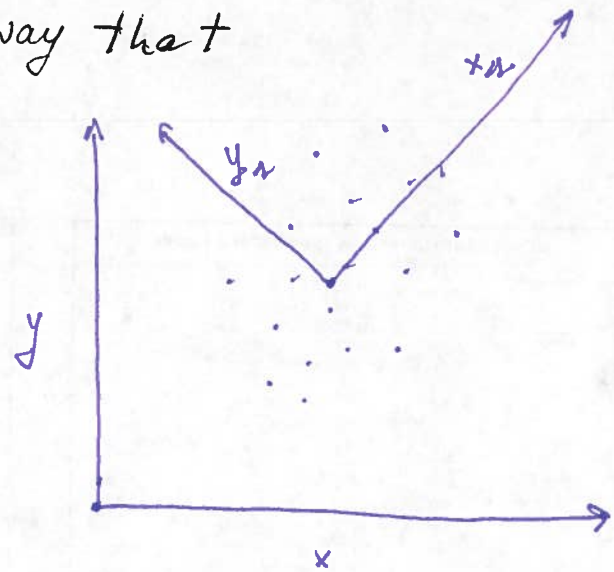
S ← number of space elements

Assume you want to decompose space and time variability

$$f(s,t) = \sum_{k}^{K} \underbrace{E(s,k)}_{\text{space patterns or mode}} \overbrace{P(k,t)}^{\text{time-series of each pattern}}$$

2

Choose a new coordinate system $E(s,k)$ to represent $f(s,t)$ in a way that most of the leading fall on fewer axis.



Example: by rotating $x, y \rightarrow x_1, y_1$ I can express most of my data on x_1. This also allows to reduce the dimension (1 axis rather than 2) where I can interpret the data

$$\underline{\underline{F}}(s,t) = \underline{\underline{E}}(s,k)\,\underline{P}(k,t)$$

choose the new coordinate so that maximum variance patterns

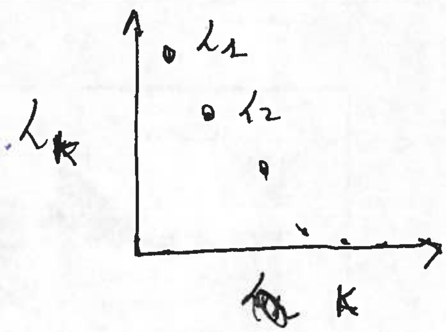
$$\underline{\underline{C}} = \cancel{E/N/t,k}\;\underline{\underline{F}}(s,t)\,\underline{\underline{F}}^T(t,s) = \underline{\underline{C}}(s,s)$$

$$\underline{\underline{C}}(s,s) = \underline{\underline{V}}(s,k)\,\underline{\underline{\Lambda}}(k,k)\,\underline{\underline{V}}^T(k,s)$$

my new basis $= \underline{\underline{E}}(s,k)$

(3)

$$\underline{\underline{\Lambda}} = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & \emptyset & \\ & & \lambda_3 & & \\ & \emptyset & & \ddots & \\ & & & & \lambda_K \end{bmatrix}$$



each λ_k, $\underline{\underline{E}}(s,k)$ will explain a certain fraction of variance. $v_k = \dfrac{\lambda_k}{\text{sum}(\lambda_k)}$

The new basis / orthogonal coordinate is $\underline{\underline{E}}(s,k) \equiv \underline{\underline{V}}(s,k)$

$$\underline{\underline{F}}(s,t) = \underline{\underline{E}}(s,k)\,\underline{\underline{P}}(k,t)$$

I know that $\underline{\underline{V}}^T \equiv \underline{\underline{V}}^{-1}$ property of eigenvectors.

$\underline{\underline{V}}^T \underline{\underline{V}} = \underline{\underline{I}}$

$$\underline{\underline{E}}^T \underline{\underline{F}} = \underline{\underline{E}}^T \underline{\underline{E}}\,\underline{\underline{P}} = \underline{\underline{P}}(k,t) \rightarrow \underline{\underline{P}}\,\underline{\underline{P}}^T = \underline{\underline{I}}$$

indep / not correl.

derive time series of projection on $\underline{\underline{E}}(s,k)$

Definitions:

$\underline{\underline{E}}(s,k)$ Empirical Orthogonal Functions (EOFs)

$\underline{\underline{P}}(k,t)$ Principal components

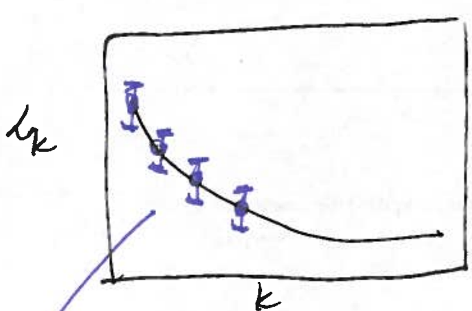
Normalization and Scaling of EOFs

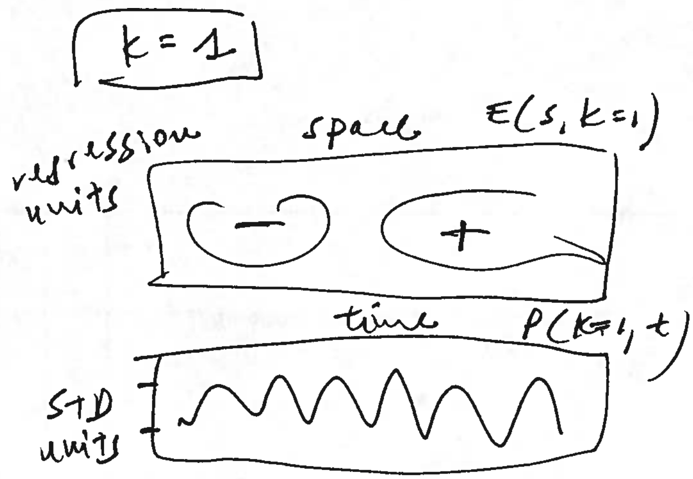
$$\underline{\underline{F}}(s,t) = \underline{\underline{E}}(s,k)\,\underline{\underline{N}}(k,k)\,\underline{\underline{N}}(k,k)^{-1}\,\underline{\underline{P}}(k,t)$$

choose $\underline{\underline{N}}$ so that $\underline{\underline{N}}^{-1}\underline{\underline{P}}$ → normalized time series of PCs by standard deviation.

→ $\underline{\underline{E}}(s,k)\,\underline{\underline{N}}(k,k)$ = Regression map between $\underline{\underline{F}}(s,t)$ and $\underline{\underline{P}}$

Plotting EOFs





~~Address~~

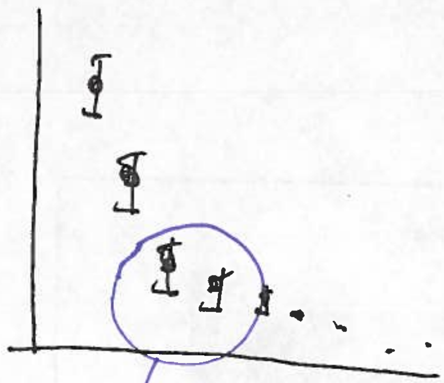
→ Error bars on λ_k

North et al (1982) 95% confidence interval

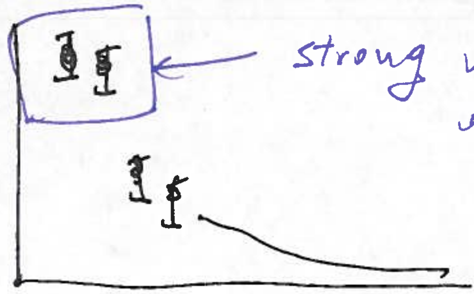
$$\Delta\lambda = \lambda\sqrt{2/N^*}$$

* ← effective degrees of freedom

If λ_i and λ_{i+1} are close within $\Delta\lambda$?



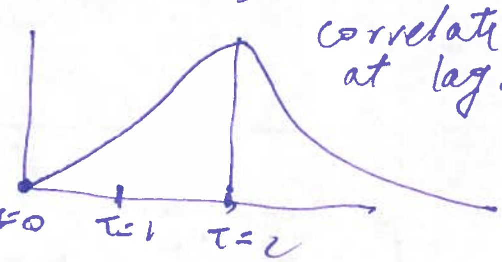
this structure is likely not significant, and any linear combination is equally significant



strong variance may indicate oscillatory pattern

$$\langle P(k=1, t+\tau)\, P(k=2, t) \rangle$$

=



correlation at lag.

(show example EOF_SVD)

Computation of EOF using SVD

$$\underline{\underline{F}} = \underline{\underline{E}}\,\underline{\underline{P}} \qquad \underline{\underline{E}}\,\underline{\underline{\Lambda}}\,\underline{\underline{E}}^T = \underline{\underline{C}} = \underline{\underline{F}}\,\underline{\underline{F}}^T$$

we can compute $\underline{\underline{E}}$ and $\underline{\underline{P}}$ from

$$\mathrm{svd}(\underline{\underline{F}}) = \underbrace{\underline{\underline{U}}(s,k)}_{\underline{\underline{E}}}\ \underbrace{S(k,k)\, V^T(k,t)}_{\underline{\underline{P}}}$$

6

$$\underline{\underline{E}}\,\underline{\underline{E}}^T = \underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^T \underline{\underline{V}}\,\underline{\underline{S}}^T \underline{\underline{U}}^T = \underline{\underline{C}}$$

$$= \underline{\underline{U}}\,\underline{\underline{S}}^2\,\underline{\underline{U}}^T$$

$$\underline{\underline{\Lambda}}$$

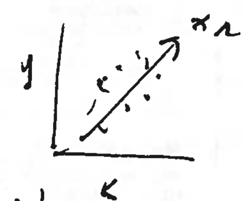
$$\underline{\underline{S}}^2 = \begin{bmatrix} s_1^2 = \lambda_1 & & & \\ & s_2^2 = \lambda_2 & & \\ & & \ddots & \\ & & & s_k^2 = \lambda_k \end{bmatrix}$$

some times SVD may be more computationaly efficient.

Applications of EOFs

DATA COMPRESSION:

- reducing dimensionality



$$\hat{\underline{F}}(s,t) = E(s, k=1\ldots K)\, P(k=1\ldots K, t)$$

$$\frac{dSST(x,y,t)}{dt} = F(x,y,t) - \cancel{\lambda SST(x,y)}$$

$P(\cancel{t,t}) =$

number of dimensions to retain / ~~filter~~

- noise filtering

STATISTICAL PREDICTION

• Objective Mapping

$$f(x,t) = \underline{\underline{\alpha}}\, f(y,t)$$

x and y indicate different location

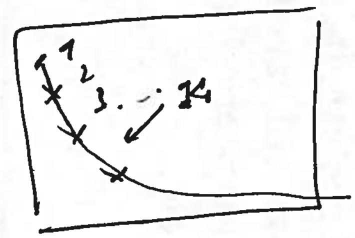
$\underline{\underline{\alpha}} \sim \underline{\underline{C}}_{xy}$

If I have previos info of C_{xx}

$$f(x,t) = E(x,k)\, P(k,t)$$

$k = 1 \ldots K$ modes of interest

$$P(k,t) = \underline{\underline{\alpha}}\, f(y,t)$$



from data

$$P(k,t)\, f^T(t,y) = \underline{\underline{\alpha}}\, C_{yy}$$

$\underline{\underline{E}}^T(k,y)$

$$\underline{\underline{\alpha}} = \underline{\underline{E}}^T(k,y)\, \underline{\underline{C}}_{yy}^{-1}$$

$\underline{\underline{\alpha}}\, f(y,t)$ projecting observations onto a limited set of EOF.

• same thing for a time model

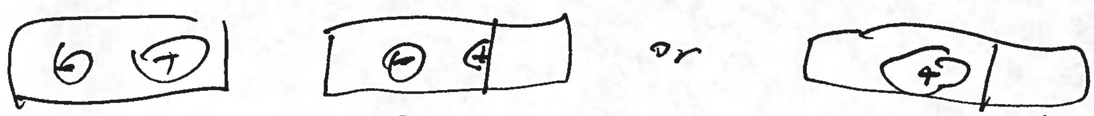
Exploratory Analysis and EOF interpretation

- Is the variance explained by EOF more than you would expect if the data had no structure?

  e.g. Monte carlo Test

- Do the structure fit with a priory-hypothesis

- Are the structure sensitive to domain?

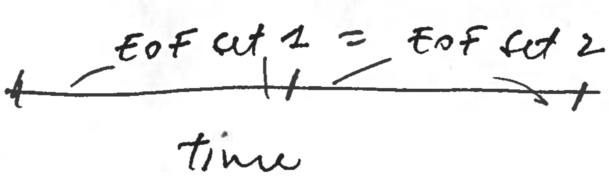


same

distinct

(spatial sensitivity)

- Are the EOF sensitive to samples?



Rotation of EOFs

- EOF orthogonality in space → global structures all over the domain
  v.s.
  local in space phenomena

rotate factors $v(s,k)$

goal is to maximize the simplicity of the factors.

VARIMAX method

simplicity criteria = variance of squared loading

e.g. $\underline{\underline{v}}(s,k) \longrightarrow \underline{\underline{B}}(s,k)$

$$s_k^2 = \frac{1}{S}\sum_1^S \left[b_{s,k}^2 - \overline{b_{s,k}^2} \right]^2$$

$$= \frac{1}{S}\sum_1^S \left[b_{s,k}^2 - \frac{1}{S}\sum b_{s,k}^2 \right]^2$$

$$= \frac{1}{S}\sum_1^S b_{s,k}^4 - \frac{1}{S^2}\left(\sum_1^S b_{s,k}^2\right)^2$$

→ factors will tend to be 0 or 1 in sp.