39/40

# 12.805 Homework 6

Nicely done

Bryan Kaiser

May 4, 2016

## Empirical Orthogonal Functions

For a data set with spatial and temporal dimensions $y = y(x_m, t_n)$ for $m = [1, ..., M]$ spatial points and $n = [1, ..., N]$ observations, so the matrix $\mathbf{y} = y_{mn}$ has dimensions $[M \times N]$. Note that two and three dimensional spatial data can be vectorized. Remove the mean and obtain the temporal autocovariance matrix at zero lag

$$\mathbf{C}_{yy} = \frac{1}{N}\mathbf{y}\mathbf{y}^\dagger = \frac{1}{N}\sum_{n=1}^{N}(y_{mn} \cdot y_{nm}) \tag{1}$$

so $\mathbf{C}_{yy}$ has dimensions of $[M \times M]$. To find the singular value decomposition of $\mathbf{y}$, $\mathbf{y}$ is multiplied by $\mathbf{U}$ (a self-orthogonal (unitary) matrix of size $[M \times M]$) and by $\mathbf{V}$ (a unitary matrix of size $[N \times N]$) such that $\mathbf{U}$ and $\mathbf{V}$ are orthogonal and the elements of another matrix $\mathbf{\Lambda}$ (dimensions $[M \times N]$) are the square roots of the eigenvalues of $\mathbf{y}^\dagger\mathbf{y}$

$$\mathbf{U}^\dagger\mathbf{y}\mathbf{V} = \mathbf{\Lambda} \quad \Rightarrow \quad \mathbf{y} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^\dagger. \tag{2}$$

Substituting equation 2 into equation 1 yields

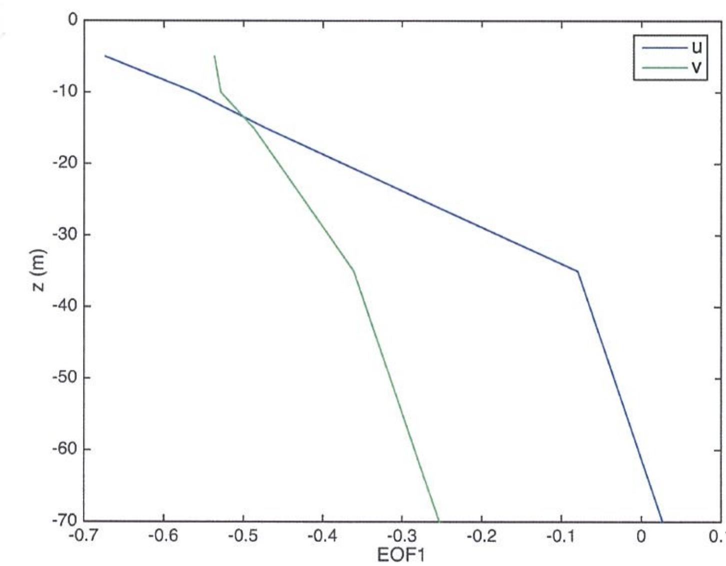
+1 Bonus, Nice

$$\mathbf{C}_{yy} = \frac{1}{N}(\mathbf{U}\mathbf{\Lambda}\mathbf{V}^\dagger)(\mathbf{U}\mathbf{\Lambda}\mathbf{V}^\dagger)^\dagger = \frac{1}{N}\mathbf{U}\mathbf{\Lambda}\mathbf{\Lambda}^\dagger\mathbf{U}^\dagger,$$

where $\mathbf{\Lambda}\mathbf{\Lambda}^\dagger$ is the squared singular values of $\mathbf{y}$. Note that this can be rearranged to form an eigenvalue problem

$$\underbrace{\mathbf{C}_{yy}\mathbf{U}}_{\text{covariance}} = \underbrace{\mathbf{U}\mathbf{\Gamma}}_{\text{SVD}}, \qquad \mathbf{\Gamma} = \frac{1}{N}\mathbf{\Lambda}\mathbf{\Lambda}^\dagger. \tag{3}$$

$\mathbf{C}_{yy}$ is empirical and eigenvectors (columns of $\mathbf{U}$) represent the EOFs or principal components of spatial system. The underbrackets indicate the separate methods that may be used to obtain the EOFs, and when both were computed for the current meter data the difference was $\mathbf{C}_{yy}\mathbf{U} - \mathbf{U}\mathbf{\Gamma} = \mathcal{O}(10^{-13})$, just a few orders of magnitude greater than machine precision (**Ans. 1**).




(just checked)

Figure 1: The spatial pattern of EOF 1 for both $u$ and $v$ suggest southwestward flow with greater variability near the surface. (**Ans. 2**).

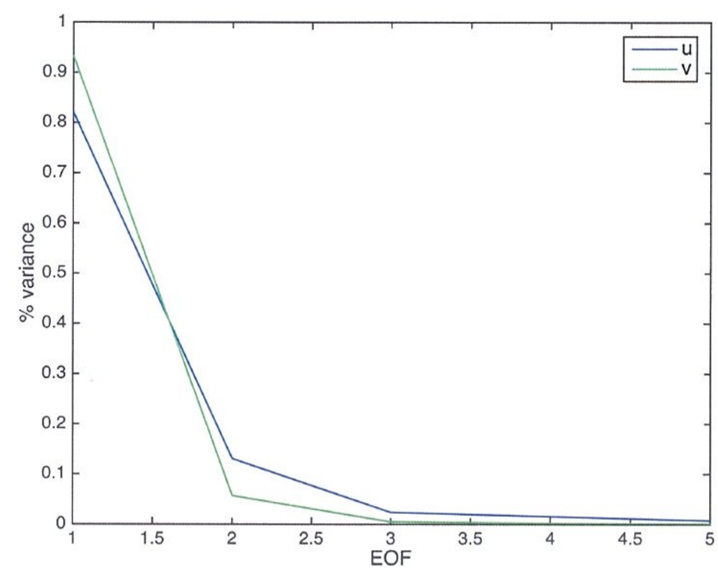



Figure 2: The % variance of all EOFs both $u$ and $v$. The first EOF dominates the variance of both; for $u$ the first EOF contains 83% of the variance of the signal and for $v$ the first EOF contains 94% of the signal. (**Ans. 2**).

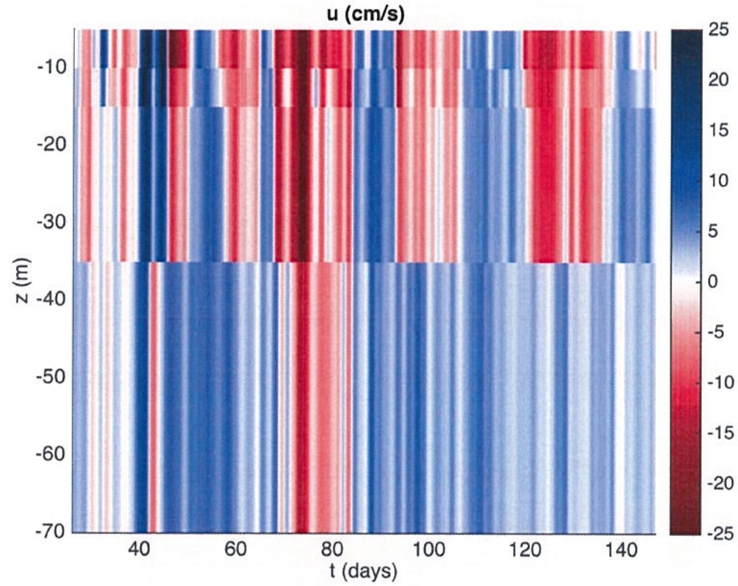



Figure 3: A Hovmöller plot of $u$. The plot shows increased variability and a predominantly westward direction near the surface and less variability and shoreward motion at depth, in agreement with first EOF of $u$ **Ans. 2**).

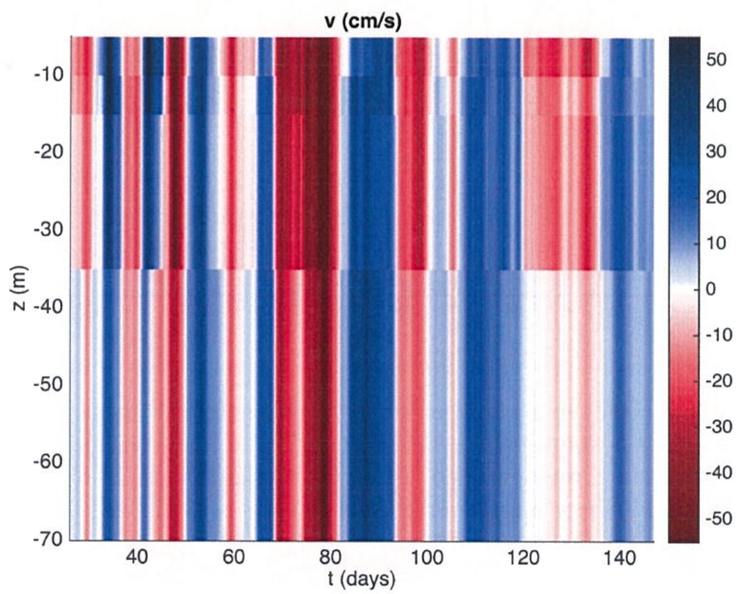



Figure 4: A Hovmöller plot of $v$. The plot shows that the variance of $v$ is weakly dependent on depth, in agreement with the first $EOF$ of $v$, and (arguably) a southward mean flow (**Ans. 2**).

✓

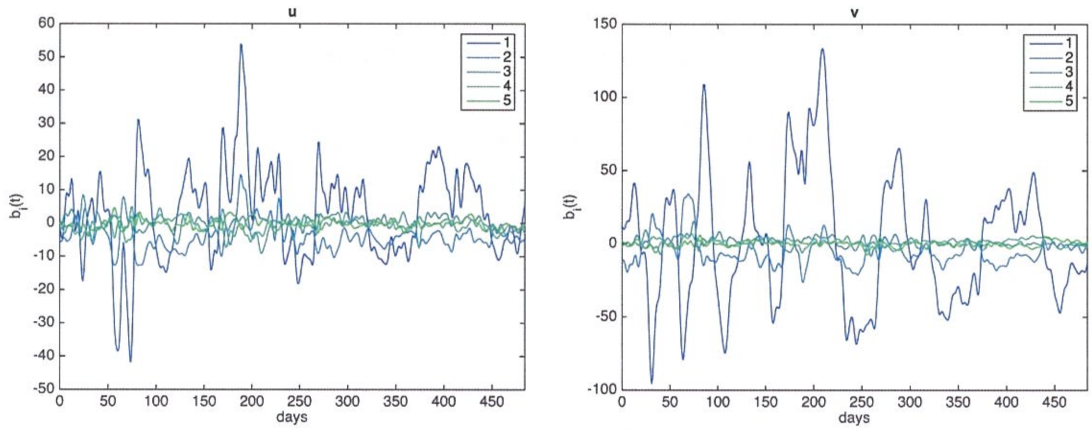



Figure 5: Expansion coefficients $b_i(t)$ for $u$ and $v$ which satisfy $\mathbf{b} = \mathbf{U}^{\dagger}\mathbf{y}$ (**Ans. 3**) .

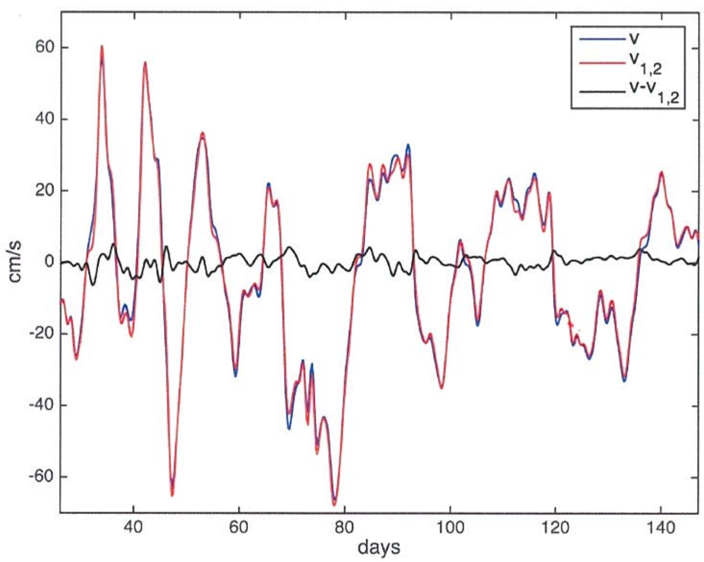



Figure 6: The expansion coefficients and the % variance of each EOF suggest that approximately 10% of the velocity signal is contained in the second EOF. Therefore, EOFs 3 through 5 were assumed to be random error removed and the velocity signals are reconstructed. It is also assumed that all systematic error has been removed from the signal. The plot above shows $v$ at 5m depth, where $v_{1,2}$ indicates the reconstructed signal from EOFs 1 and 2. The raw $v$, filtered $v$, and their difference (i.e. the sum of the last 3 EOFs or the assumed noise) are plotted (**Ans. 4**).

## Objective Mapping

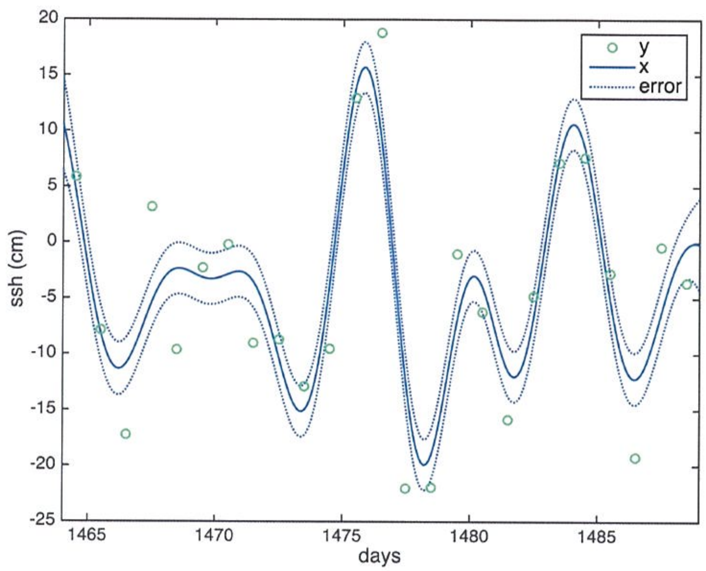



Figure 7: A 1D signal $\mathbf{y}$ with a resolution of $\Delta t = 1$ days and length $M = 25$ is objectively mapped to a signal $\mathbf{x} = \mathbf{R}_{xy}\mathbf{R}_{xy}^{-1}\mathbf{y}$ with a resolution of $\Delta t = 1/10$ days and length $N = 251$. The covariance matrix for $\mathbf{x}$ is constructed $\mathbf{R}_{xx} = 100\exp(-(\tau/4)^2)\cos(\tau\pi/4)$ and $\mathbf{R}_{xn} = 0$ assumed. The error is shown as the region between the dotted lines (**Ans.**).

– Also wanted errorbars on observations

– They help check whether estimates are consistent

−2