High-resolution frequency wavenumber analysis
Capon's f-k estimator

Matthias Ohrnberger

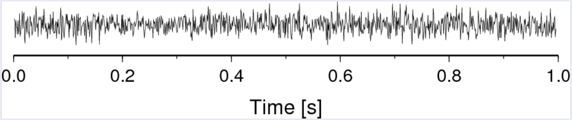
Institute of Earth and Environmental Sciences
University of Potsdam

21. Februar 2010

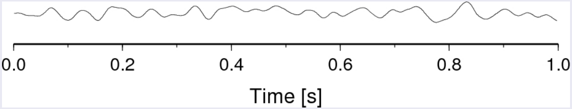
Section overview

process observation → from continuous infinite to finite discrete . . .

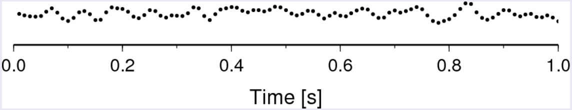
Continuous infinite process → here: analog time series



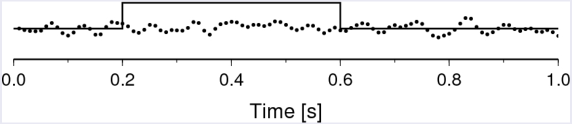
Lowpass filtering to remove high frequencies (analog AA-filter)



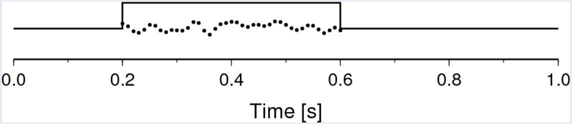
Sampling → from continuous to discrete (aliasing theorem!!)



Finite length observation (eternity is for God only)

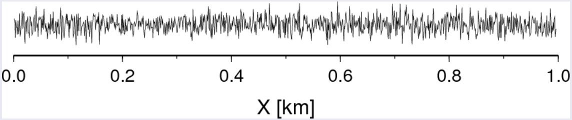


Equivalent to boxcar tapering

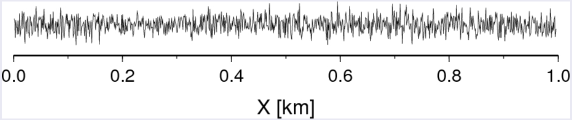


Time series sampling and analogy to spatial sampling

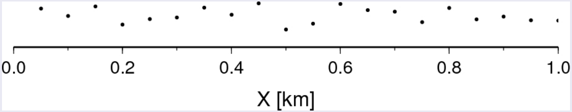
Continuous infinite process → here: wavefield snapshot



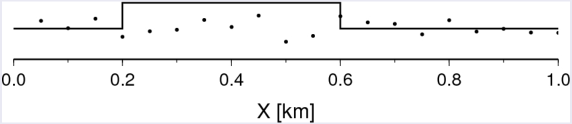
Lowpass filtering to remove high frequencies? NOT possible!



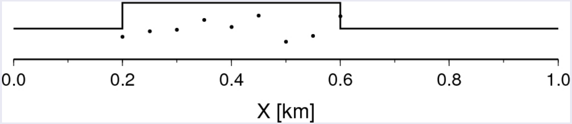
Spatial sampling → from continuous to discrete



Finite aperture observation (omnipresence is for God only)



Again: Equivalent to boxcar tapering



Consequences of spatio-temporal wave field sampling

aliasing (violation of the sampling theorem)

>2 samples are required for sampling a period (wavelength)

- time domain: $\Delta T < T_{min}/2$
- spatial domain: $\Delta x < \lambda^{a}_{min}/2$

[a] apparent

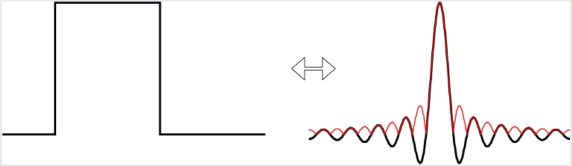
Spectral analysis of discrete series $\longrightarrow$ spectral resolution limit

- time domain: $\Delta\omega = 2\pi/((N-1)\Delta T)$
- spatial domain: $\Delta k = 2\pi/((N-1)d_{min}) = 2\pi/D_{max}$

A remedy against spectral leakage - tapering

Boxcar taper function

Spectrum of boxcar is sinc function $sinc(x) = \sin(x)/x$



For spectral analysis it is well known, that some taper functions perform better than others with respect to sidelobe height, roughness, leakage ... Does it make sense to introduce a tapering function for the spatial sampling process?

Recall beam computation

Signal model

$$X_i(\omega, \vec{k}_0) = S(\omega)\exp(j\vec{k}_0\vec{r}_i) + n_i(\omega)$$

shift and sum operation

$$B(\omega, \vec{k}, \vec{k}_0) = \sum_{i=1}^{N} X_i(\omega, \vec{k}_0)\exp(-j\vec{k}\vec{r}_i)$$

$$B(\omega, \vec{k}, \vec{k}_0) = \sum_{i=1}^{N} \left[S(\omega)\exp(j\vec{k}_0\vec{r}_i)\exp(-j\vec{k}\vec{r}_i) + \tilde{n}_i(\omega)\right]$$

$$B(\omega, \vec{k}, \vec{k}_0) = \sum_{i=1}^{N} \left[S(\omega)\exp(j(\vec{k}_0 - \vec{k})\vec{r}_i) + \tilde{n}_i(\omega)\right]$$

Weighted shift and sum

Generalization of array output

conventional beamforming

$$B(\omega, \vec{k}, \vec{k}_0) = \sum_{i=1}^{N} X_i(\omega, \vec{k}_0) \exp(-j\vec{k}\vec{r}_i)$$

and introducing (complex) sensor weights as spatial taper equivalent to windowing function in time series spectral analysis

$$B(\omega, \vec{k}, \vec{k}_0) = \sum_{i=1}^{N} W_i(\omega) X_i(\omega, \vec{k}_0) \exp(-j\vec{k}\vec{r}_i)$$

Introducing vector/matrix notation

For compact notation we use

$$\vec{X}(\omega,\vec{k}_0) = \begin{bmatrix} X_1(\omega,\vec{k}_0) \\ X_2(\omega,\vec{k}_0) \\ \vdots \\ X_N(\omega,\vec{k}_0) \end{bmatrix} \qquad \vec{A}(\omega,\vec{k}) = \begin{bmatrix} W_1(\omega)\exp(j\vec{k}\vec{r}_1) \\ W_2(\omega)\exp(j\vec{k}\vec{r}_2) \\ \vdots \\ W_N(\omega)\exp(j\vec{k}\vec{r}_N) \end{bmatrix}$$

and therefore

$$B(\omega,\vec{k},\vec{k}_0) = \sum_{i=1}^{N} W_i(\omega) X_i(\omega,\vec{k}_0)\exp(-j\vec{k}\vec{r}_i) = \vec{A}^H(\omega,\vec{k})\vec{X}(\omega,\vec{k}_0)$$

superscript ^{H} (also †) denotes conjugate transpose operation

$$\underline{Y}^H = \underline{Y}^\dagger = (\overline{\underline{Y}})^T = \overline{\underline{Y}^T}$$

From beam to beam power to the generalized beamformer

Compact notation of beam

$$B(\omega, \vec{k}, \vec{k}_0) = \vec{A}^H(\omega, \vec{k})\vec{X}(\omega, \vec{k}_0)$$

Compact notation of beam power

$$\left|B(\omega, \vec{k}, \vec{k}_0)\right|^2 = \left[\vec{A}^H(\omega, \vec{k})\vec{X}(\omega, \vec{k}_0)\right] \overline{\left[\vec{A}^H(\omega, \vec{k})\vec{X}(\omega, \vec{k}_0)\right]}$$

$$\left|B(\omega, \vec{k}, \vec{k}_0)\right|^2 = \vec{A}^H(\omega, \vec{k})\vec{X}(\omega, \vec{k}_0)\vec{X}^H(\omega, \vec{k}_0)\vec{A}(\omega, \vec{k})$$

The generalized beamformer

$$\left|B(\omega, \vec{k}, \vec{k}_0)\right|^2 = \vec{A}^H(\omega, \vec{k})\underline{R}(\omega, \vec{k}_0)\vec{A}(\omega, \vec{k})$$

where $\underline{R}(\omega, \vec{k}_0) = \vec{X}(\omega, \vec{k}_0)\vec{X}^H(\omega, \vec{k}_0)$

The Cross spectral matrix

We introduced the abbreviation $\underline{R}(\omega)$

$$\underline{R}(\omega) = \vec{X}(\omega)\vec{X}^H(\omega)$$

as outer product of the vector of observations at each station with itself. This quantitiy is the primary source of information for all f-k estimates! (In the following we drop the dependency of $\vec{k}_0$)

$\underline{R}(\omega)$ has several names

- cross spectral matrix
- spatio-spectral matrix
- spatial correlation matrix
- sensor covariance martix
- ...

Cross spectral matrix continued

Elements of the cross spectral matrix

$$\underline{R}(\omega) = \begin{bmatrix} X_1(\omega)\overline{X}_1(\omega) & X_1(\omega)\overline{X}_2(\omega) & \dots & X_1(\omega)\overline{X}_N(\omega) \\ X_2(\omega)\overline{X}_1(\omega) & X_2(\omega)\overline{X}_2(\omega) & \dots & X_2(\omega)\overline{X}_N(\omega) \\ X_3(\omega)\overline{X}_1(\omega) & X_3(\omega)\overline{X}_2(\omega) & \dots & X_3(\omega)\overline{X}_N(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ X_N(\omega)\overline{X}_1(\omega) & X_N(\omega)\overline{X}_2(\omega) & \dots & X_N(\omega)\overline{X}_N(\omega) \end{bmatrix}$$

Matrix elements contain phase difference between sensors , e.g.

$$X_1(\omega)\overline{X}_2(\omega) = |X_1(\omega)|\,|X_2(\omega)|\exp(j(\phi_1 - \phi_2))$$

the phase differences are equivalent to time delays.

Cross spectral matrix estimation

Block averaging procedure

$$(\underline{\hat{R}})_{ij} = \frac{1}{M}\sum_{m=1}^{M} X_i(\omega)\overline{X}_j(\omega)$$

$$(\underline{\hat{R}})_{ij} = \frac{1}{M}\sum_{m=1}^{M} \frac{X_i(\omega)\overline{X}_j(\omega)}{\sqrt{X_i(\omega)\overline{X}_i(\omega)}\sqrt{X_j(\omega)\overline{X}_j(\omega)}}$$

- In order to obtain good estimates of the phase differences, an averaging procedure is required! Thus, the assumption of stationarity of the wavefield has to be introduced.
- Variance reduction of cross spectral matrix estimates are usually obtained by the block averaging procedure

Block averaging procedure depicted

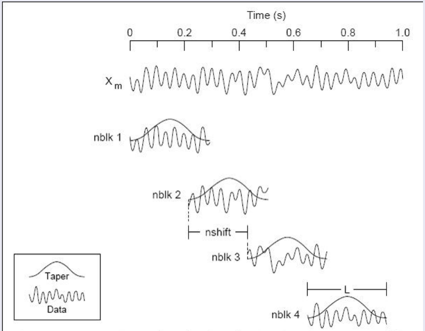


Figure 3. Segmentation and overlapping of a time-domain component of the signal vector. Each tapered block is transformed into the frequency domain via an FFT.

Note: block-averaging in time domain (as depicted here) is equivalent to smoothing over some bandwidth in frequency domain! Matter of taste which implementation is to be preferred - in both cases long time windows are needed to gather sufficient number of samples for averaging/smoothing.

Capon's method (1969)

Based on the formulation of the generalized beamformer

$$\left|B(\omega,\vec{k})\right|^2 = \vec{A}^H(\omega,\vec{k})\underline{R}(\omega)\vec{A}(\omega,\vec{k})$$

- Idea followed by Capon: Find optimum weights for the generalized beamformer which provide an f-k estimate that has unity gain at the true wavenumber and is minimized elsewhere!
- Additionally a particular constraint should be met: Array output equals signal observation at each station for true wavenumber. This constraint is called the distortionless constraint.

Mathematical formulation of Capon's approach

We had introduced before:

$$\vec{A}(\omega) = \left[W_1(\omega)\exp(j\vec{k}\vec{r}_1), W_2(\omega)\exp(j\vec{k}\vec{r}_2), \ldots, W_N(\omega)\exp(j\vec{k}\vec{r}_3)\right]^T$$

which consists of weights and shifts: $\vec{A}(\omega, \vec{k}) = \underline{W}(\omega)\vec{e}(\vec{k})$:

$$\underline{W}(\omega) = \begin{bmatrix} W_1(\omega) & 0 & \ldots & 0 \\ 0 & W_2(\omega) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & W_N(\omega) \end{bmatrix} \qquad \vec{e}(\vec{k}) = \begin{bmatrix} \exp(j\vec{k}\vec{r}_1) \\ \exp(j\vec{k}\vec{r}_2) \\ \vdots \\ \exp(j\vec{k}\vec{r}_N) \end{bmatrix}$$

Split between the weights $\underline{W}(\omega)$ and the steering vector $\vec{e}(\vec{k})$.

Mathematical formulation of Capon's approach

Therefore the generalized beamformer reads now:

$$\left|B(\omega,\vec{k})\right|^2 = (\underline{W}(\omega)\ \vec{e}(\vec{k}))^H \underline{R}(\omega)(\underline{W}(\omega)\ \vec{e}(\vec{k}))$$

$$\left|B(\omega,\vec{k})\right|^2 = \vec{e}^H(\vec{k})\underline{W}^H(\omega)\underline{R}(\omega)(\underline{W}(\omega)\vec{e}(\vec{k}))$$

Now we can formulate Capon's ideas:

- f-k estimate that has unity gain at the true wavenumber and is minimized elsewhere → f-k estimate should approximate at best a 3D delta function (in k_x, k_y, ω), or cross spectral power is to be minimized:

$$\underline{W}^H\ \underline{R}\ \underline{W}$$

Distortionless constraint

As additional constraint for an optimal beamforming algorithm, one can formulate the following requirement:

- Array output equals observation for true wavenumber

or put into different words:

- the true signal (waveform / spectrum) should not be distorted when we steer the array beam to it.

Not sruprisingly, this constraint is termed "distortionless".

Formulating this contraint in math

$$W_i(\omega)X_i(\omega)\exp(-j\vec{k}\vec{r}_i) = X_i(\omega)$$

which results in

$$\underline{W}(\omega)\vec{e}(\vec{k}) = \vec{1}$$

Minimization with constraints

Problem to be solved

Minimize $\underline{W}^H \; \underline{R} \; \underline{W}$ obeying the constraint $\underline{W}(\omega)\vec{e}(\vec{k}) = \vec{1}$

Solution to problem using Lagrangian multipliers

Minimize $f(x)$ and K constraints $\Phi_k(x) = 0, \;\; k = 1, \ldots, K$

$$\text{Minimize Lagrangian} \quad \Lambda(x, \lambda) = f(x) + \sum_{k=1}^{K} \lambda_k \Phi_k(x)$$

$$\nabla \Lambda(x, \lambda) = 0$$

gives $K + 1$ equations for $K + 1$ variables

Optimum weights

Using Lagrangian multipliers, the solution to the problem is:

$$\underline{W}(\omega) = \frac{\underline{R}^{-1}(\omega)\vec{e}(\vec{k})}{\vec{e}^{H}(\vec{k})\underline{R}^{-1}(\omega)\vec{e}(\vec{k})}$$

What does this mean?

- optimum weights are in general complex quantities
- weights depend explicitely on steering vector $\vec{e}(\vec{k})$!
- weights are adaptive and tune the shape of the spatial taper function in dependence of the wavenumber

So, we are able to compute the optimum weights given the wavenumber $\vec{k}$ for a particular beam realization. For computing the beam power, we insert the weights into the expression for the generalized beamformer:

$$\left|B(\omega,\vec{k})\right|^{2} = (\underline{W}(\omega)\ \vec{e}(\vec{k}))^{H}\underline{R}(\omega)(\underline{W}(\omega)\ \vec{e}(\vec{k}))$$

Capon's estimator

Weights introduced in generalized beampower notation:

$$\left|B(\omega,\vec{k})\right|^2 = \left(\frac{\underline{R}^{-1}(\omega)\vec{e}(\vec{k})}{\vec{e}^H(\vec{k})\underline{R}^{-1}(\omega)\vec{e}(\vec{k})}\vec{e}(\vec{k})\right)^H \underline{R}(\omega)\left(\frac{\underline{R}^{-1}(\omega)\vec{e}(\vec{k})}{\vec{e}^H(\vec{k})\underline{R}^{-1}(\omega)\vec{e}(\vec{k})}\vec{e}(\vec{k})\right)$$

and after some simplification we obtain:

$$\left|B(\omega,\vec{k})\right|^2 = P_{Capon}(\omega,\vec{k}) = \frac{1}{\vec{e}(\vec{k})^H\underline{R}^{-1}(\omega)\vec{e}(\vec{k})}$$

This is a surprising (or maybe better: a convenient) result. For beam power computation the weights don't have to be estimated explicitly. All the information is contained in the cross spectral matrix!

Capon's method put into practice

- Select time window of observation
- Divide data window into smaller subwindows.
- DFT data subwindows
- Estimate cross spectral matrix by averaging frequency dependent covariance over all subwindows and for each station pair
- invert cross spectral matrix
- sweep over wavenumber space using trial steering vectors and compute:

$$P_{Capon}(\omega, \vec{k}) = \frac{1}{\vec{e}(\vec{k})^H \underline{R}^{-1}(\omega)\vec{e}(\vec{k})}$$

- display results as f-k map

Does it work?

High-resolution or not? Yes, but not always . . .

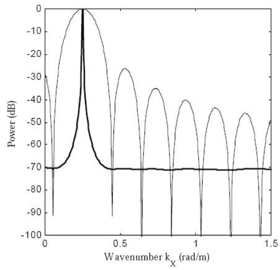


Figure 4.12 Power Output for the Minimum Variance Distortionless Look (MVDL) Method for a Single Wave at $k_x = 0.25$ rad/m Propagating Along the Main Axis of the 16 Sensor Uniform Linear Array. The MVDL output is shown with the dark line, and the FDBF output is shown with the light line for reference.

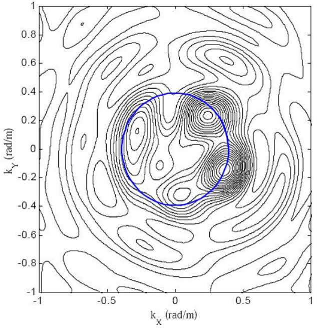


Figure 8.17 Example of Multiple Signals Arriving at a Frequency = 9.875 Hz.

from Ph.D. Thesis of Zywicki, 1999. (www.zywicki.com)
Result depends crucial on the stabilization of the CSM inversion and "quality" of CSM estimate from the data!

References

- D.H.Johnson & D.E.Dudgeon, Array Signal Processing - concepts and techniques, Prentice Hall Signal Processing Series, Alan v. Oppenheim, Series Editor, 1993.
- H.L. Van Trees, Optimum Array Processing, Part IV of 'Detection, Estimation, and Modulation Theory, John wiley and Sons, Inc., New York, 2002.
- Schweitzer, J., Fyen, J., Mykkeltveit, S. and Kvaerna, T. (2002). Chapter 9: Seismic Arrays, In: Bormann, P. (Ed.), IASPEI New Manual of Seismological Observatory Practice, GeoForschungsZentrum Potsdam, Vol. 1, 51pp.
- Mykkeltveit, S., Astebol, K., Doornbos, D.J., and Husebye, E.S. (1983). Seismic Array configuration Optimization, BSSA, 73(1), pp. 173-186.
- Rost, S. and Ch. Thomas, Array Seismology, Reviews of Geophysics, 2002.