Estimation of covariance function

Let $x_t, t = 0, 1, 2, \ldots n-1$ be a zero mean sequence of data. The usual biased covariance function estimate is given from

$$\widehat{r}_x(\tau) = \frac{1}{n} \sum_{t=0}^{n-1-|\tau|} x_t x_{t+|\tau|},$$

where

$$\begin{aligned} E[\widehat{r}_x(\tau)] &= \frac{1}{n} \sum_{t=0}^{n-1-|\tau|} E[x_t x_{t+|\tau|}], \\ &= \frac{(n-|\tau|)}{n} r_X(\tau) \rightarrow r_X(\tau), \quad n \rightarrow \infty \end{aligned}$$

if x_t is a realization of a weakly stationary stochastic process.

Estimation of spectral density

The *periodogram* is defined as

$$\widehat{R}_x(f) = \frac{1}{n}|\mathcal{X}(f)|^2,$$

where $\mathcal{X}(f) = \sum_{t=0}^{n-1} x_t e^{-i2\pi ft}$ is the Fourier transform of data. The periodogram is also the Fourier transform of the usual biased covariance estimate, as

$$\begin{aligned}
\widehat{R}_x(f) &= \frac{1}{n}\sum_{t=0}^{n-1}\sum_{s=0}^{n-1} x_t x_s e^{-i2\pi f(s-t)}, \\
&= \frac{1}{n}\sum_{\tau=-n+1}^{n-1}\sum_{t=0}^{n-1-|\tau|} x_t x_{t+|\tau|} e^{-i2\pi f\tau}, \\
&= \sum_{\tau=-n+1}^{n-1} \widehat{r}_x(\tau) e^{-i2\pi f\tau}.
\end{aligned}$$

This means that the periodogram is an estimate of the spectral density, as the spectral density relates to the covariance function through the Fourier transform.

Expected value of the periodogram for small n

We calculate the expected value of the periodogram,

$$E[\widehat{R}_x(f)] = \sum_{\tau=-n+1}^{n-1} \widehat{r}_x(\tau) e^{-i2\pi f \tau} = \sum_{\tau=-n}^{n} \frac{(n-|\tau|)}{n} r_X(\tau) e^{-i2\pi f \tau}.$$

using the usual expression for the biased covariance estimate. The expected value of the periodogram will accordingly also be biased. We define a lag-window as

$$k_n(\tau) = \frac{(n-|\tau|)}{n} \quad \text{for} -n \leq \tau \leq n,$$

and express the expected value of the periodogram as

$$\begin{aligned} E[\widehat{R}_x(f)] &= \sum_{\tau=-\infty}^{\infty} k_n(\tau) r_X(\tau) e^{-i2\pi f \tau} = \sum_{\tau=-\infty}^{\infty} k_n(\tau) \int_{-1/2}^{1/2} R_X(u) e^{i2\pi u \tau} du \; e^{-i2\pi f \tau}, \\ &= \int_{-1/2}^{1/2} R_X(u) \sum_{\tau=-\infty}^{\infty} k_n(\tau) e^{-i2\pi (f-u)\tau} du = \int_{-1/2}^{1/2} K_n(f-u) R_X(u) du. \end{aligned}$$

The expected value is the convolution between the true spectral density and the Fourier transform of the lag-window.

Variance of the periodogram

The variance of the periodogram is

$$V[\widehat{R}_x(f)] = \frac{1}{n^2}C[|\mathcal{X}(f)|^2, |\mathcal{X}(f)|^2] = \frac{1}{n^2}C[\mathcal{X}(f)\overline{\mathcal{X}}(f), \mathcal{X}(f)\overline{\mathcal{X}}(f)] =$$

$$= \frac{1}{n^2}C[\mathcal{X}(f), \mathcal{X}(f)]C[\overline{\mathcal{X}}(f), \overline{\mathcal{X}}(f)]) + \frac{1}{n^2}C[\mathcal{X}(f), \overline{\mathcal{X}}(f)]C[\overline{\mathcal{X}}(f), \mathcal{X}(f)],$$

using Isserlis theorem, see page 243-244. We study the different covariances (including $1/n$),

$$\frac{1}{n}C[\mathcal{X}(f), \overline{\mathcal{X}}(f)] = \frac{1}{n}E[\sum_{t=0}^{n-1}\sum_{s=0}^{n-1} x(t)x(s)e^{-i2\pi f(s-t)}] = R_x(f),$$

but it can be shown that

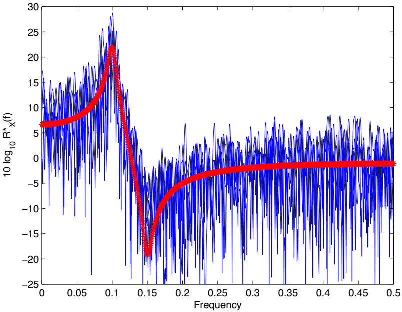
$$\frac{1}{n}C[\mathcal{X}(f), \mathcal{X}(f)] = \frac{1}{n}E[\sum_{t=0}^{n-1}\sum_{s=0}^{n-1} x(t)x(s)e^{-i2\pi f(s+t)}] \approx R_x(f),$$

only for $f = [0, 0.5]$, and ≈ 0 for all other values. We find

$$V[\widehat{R}_x(f)] \approx \begin{cases} R_X^2(f) & 0 < |f| < 0.5 \\ 2R_X^2(f) & f = 0 \text{ and } 0.5 \end{cases}$$

Example

Examples of 5 realizations of an ARMA(2,2)-process of sequence length $n = 128$ (blue). The true spectral density (red). The estimates show too large variance.



We need better spectral estimation methods!

Short history on variance reduction of spectral estimates

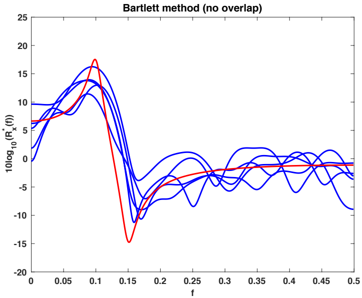
- Averaging of spectral estimates from non-overlapping sequences, the Bartlett method, (around 1950, Bartlett M. S).
- Welch showed the advantages of averaging of spectral estimates from overlapping windowed sequences. (Welch, P.D. ,1967, "The use of Fast Fourier Transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms", IEEE Transactions on Audio and Electroacoustics.)
- Thomson suggested totally overlapping windows with different shapes, *multiple windows, multitapers*. (Thomson, D.J. ,1982, "Spectrum estimation and harmonic analysis", Proceedings of the IEEE.)
- More tailored multitapers, e.g. for spectra with peaks: The Peak Matched Multiple Windows (PM MW). (Hansson M. and Salomonsson G., 1997, "A Multiple Window Method for Estimation of Peaked Spectra", IEEE Trans. on Signal Processing.)

Averaging of periodograms

The Bartlett method: We calculate the average of K estimates, $\widehat{R}_{x,j}(f)$, $j = 1 \ldots K$, where the length of each sequence is n/K,

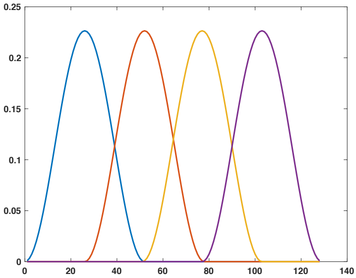
$$\widehat{R}_{mv}(f) = \frac{1}{K} \sum_{j=1}^{K} \widehat{R}_{x,j}(f).$$

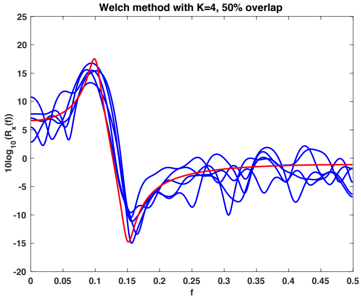
For large values of n, $V[\widehat{R}_{mv}(f)] \approx \frac{1}{K} R_X^2(f)$, if the properties of the data sequences are such that the different spectral estimates in the average are uncorrelated. Example of 5 estimates.



The Welch method

A standard estimator is the Welch method or WOSA (Weighted Overlap Segment Averaging) using Hanning windows and 50% overlap.





Multitaper spectral estimates

Definition:

$$\widehat{R}_{mv}(f) = \frac{1}{K}\sum_{j=1}^{K}\widehat{R}_{j,x}(f),$$

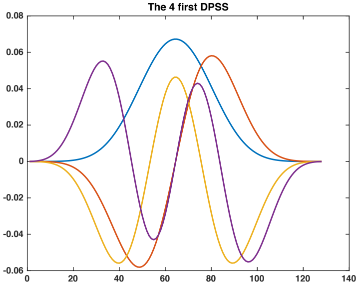
where

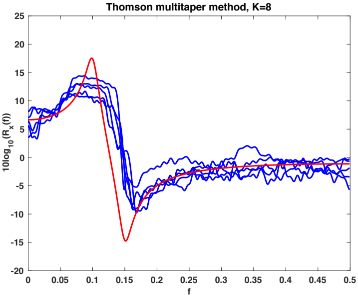
$$\widehat{R}_{j,x}(f) = \frac{1}{n} \mid \sum_{t=0}^{n-1} x(t)w_j(t)e^{-i2\pi ft} \mid^2 .$$

The windows $w_j(t)$, $j = 1 \ \ldots \ K$, should be chosen such that we get uncorrelated spectra from the same data sequence.

The Thomson method

The Thomson method uses the Discrete Prolate Spheroidal Sequences (DPSS) as multitapers.





The Peak Matched Multiple Windows

Multitapers tailored for spectra with peaks. (My own contribution to the field)