Unit 21: Nonparametric Spectral Density Estimation

Jeffrey Woo

Department of Statistics, University of Virginia

Spring 2020

Readings for Unit 21

Textbook chapter 4.4 (until page 198).

Last Unit

1. Discrete Fourier Transform
2. Periodogram

Motivation

We use the periodogram to estimate the spectral density (power spectrum). However, we noted that the confidence intervals for the spectral density based on the periodogram are too wide to be useful. We look at smoothing the periodogram to construct narrower confidence intervals.

Review: CIs for Periodogram

For the periodogram, the problem is that the number of parameters that we are fitting, $f(\omega_{j:n})$, is **growing** at the same rate as the data. Asymptotically,

$$\frac{2I(\omega_{j:n})}{f(\omega)} \to \chi_2^2.$$

Thus, the confidence interval for the power spectrum at ω is

$$\frac{2I(\omega_{j:n})}{\chi_2^2(1 - \alpha/2)} \leq f(\omega) \leq \frac{2I(\omega_{j:n})}{\chi_2^2(\alpha/2)}.$$

Note that the width of the confidence interval does not shrink as sample size increases.

1 Smoothing the Periodogram

2 Generalization of Smoothing

3 (Modified) Daniell Kernel

4 Worked Example

Estimating the Spectral Density

Another issue is this: we'd like to have a good representation of $f(\omega)$, which is presumably a continuous function in ω, and we want to obtain this continuous function from discrete observations. The periodogram $I(\omega_j)$ is a function across the discrete values ω_j We will use the fact that $f(\omega)$ and $f(\omega')$ should be close to one another if ω and ω' are. We can accomplish this with smoothing.

Smoothing the Periodogram

Define a **frequency band**, $\mathcal{B}$, of L contiguous fundamental frequencies centered around ω_j as

$$\mathcal{B} = [\omega_j - m/n, \omega_j + m/n] \tag{1}$$

where $L = 2m + 1$ is an odd number, and we also assume that ω_j is close to the frequency ω of interest. So, this will include $f(\omega_j + k/n)$ for $k = -m, ..., 0, ..., m$. Also assume that $f(\omega)$ is **approximately constant** in the frequency band.

Smoothing the Periodogram

One of the simplest types of smoothing is a symmetric moving average, so we have

$$\overline{f}(\omega) = \frac{1}{L} \sum_{k=-m}^{m} I(\omega_j + k/n) \tag{2}$$

for $\omega \in \mathcal{B}$. The elements of the periodogram are approximately independent; this means that the approximate distribution for this sum will be χ^2. So, as $n \to \infty$,

$$\frac{2L\overline{f}(\omega)}{f(\omega)} \to \chi^2_{2L} \tag{3}$$

CI for Smoothed Periodogram

Thus, when using the symmetric moving average smoother as defined in (2), the approximate $(1 - \alpha) \times 100\%$ confidence interval for the power spectrum at ω is

$$\frac{2L\overline{f}(\omega)}{\chi^2_{2L}(1 - \alpha/2)} \le f(\omega) \le \frac{2L\overline{f}(\omega)}{\chi^2_{2L}(\alpha/2)}. \tag{4}$$

For large L, these intervals will be **narrower** than for the unsmoothed (raw) periodogram.

Bandwidth

Note that as n gets very large, we can take L to be fairly large so that the variance is reduced. Because of the spacing of the fundamental frequencies, we call

$$B = \frac{L}{n} \qquad (5)$$

the **bandwidth**. Note that this means that the degrees of freedom in the chi-squared distribution will be

$$2L = 2Bn.$$

Bias-Variance Tradeoff

For bandwidth $B = \frac{L}{n}$, the variance of $\overline{f}(\omega) \approx \frac{c}{Bn}$ for some constant c. So to **reduce variance**, we want a bigger bandwidth.

However, the larger the bandwidth, the more questionable the assumption that $f(\omega)$ is approximately constant in the frequency band $\mathcal{B} = [\omega_j - m/n, \omega_j + m/n]$. The larger the bandwidth, the smoother the $\overline{f}(\omega)$, but we have introduced **more bias**.

Log Transformation

Many times, a log transformation of the spectrum can aid in the visual display of the spectral density plot. This can happen when regions of the spectrum exist with peaks of interest much smaller than some of the main components. For the log spectrum, the approximate $(1 - \alpha) \times 100\%$ confidence intervals are

$$[\log \overline{f}(\omega) + \log 2L - \log \chi^2_{2L}(1 - \alpha/2), \log \overline{f}(\omega) + \log 2L - \log \chi^2_{2L}(\alpha/2)] \tag{6}$$

Fast Fourier Transform

We noted in the previous unit that the Fast Fourier Transform (FFT) is used to compute the Discrete Fourier Transform (DFT) to make the computations more efficient. The FFT is used in the `mvspec()` and `spec.pgram()` functions in R when computing the periodogram.

Zero Padding and Adjusted DF

- For FFT, zero padding is done by appending 0s in the data, so that the sample size in the DFT, n', is highly composite. A highly composite integer has many factors of 2, 3, or 5.

- The degrees of freedom used to compute (4) needs to be adjusted since we do not gain more information with zero padding. The **adjusted degrees of freedom** is

$$df = \frac{2Ln}{n'} \tag{7}$$

Zero Padding and Adjusted DF

So the confidence interval specified in (4) becomes

$$\frac{df\,\overline{f}(\omega)}{\chi^2_{df}(1-\alpha/2)} \leq f(\omega) \leq \frac{df\,\overline{f}(\omega)}{\chi^2_{df}(\alpha/2)} \tag{8}$$

when the FFT is used.

SOI and Recruitment

We will continue with the example from Unit 20, with the
Southern Oscillation Index and recruitment datasets, which contain
monthly data on the changes in air pressure and estimated number
of new fish in the central Pacific Ocean from 1950 to 1987. The
central Pacific Ocean warms approximately every three to seven
years due to El Niño.

Raw Periodograms
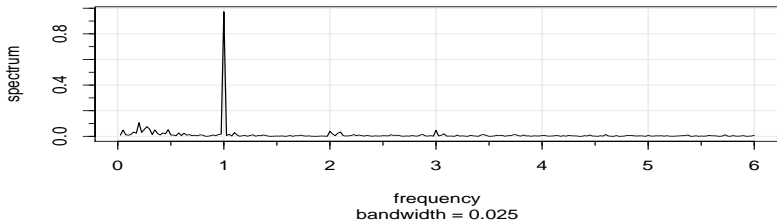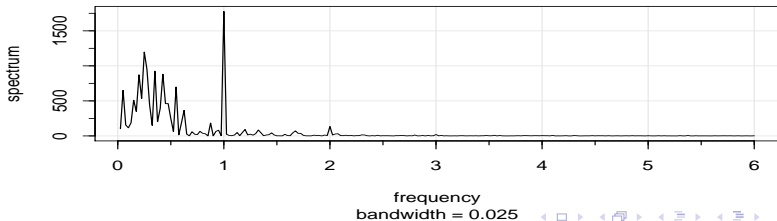
Observations from Raw Periodograms

From the raw periodograms:

- obvious peak at $\omega = 1/12$ for yearly cycle.
- some peaks at around $\omega = 1/48$ for El Nino cycle. The wide band of activity suggests that this cycle is not very regular.

Estimated Power and CIs from Raw Periodograms

The estimates of the power spectra and the confidence intervals
using the raw periodogram are listed below.

| Series	ω	Estimated Power	Lower	Upper
SOI	$\frac{1}{48}$	0.0537	0.0146	2.1222
	$\frac{1}{12}$	0.9722	0.2636	38.4011
Recruit ($\times 10^3$)	$\frac{1}{48}$	1.1974	0.3245	47.2935
	$\frac{1}{12}$	1.7777	0.4819	70.2172

Apply Smoothing

The periodogram might need smoothing to determine a predominant overall trend for El Nino. Applying $L = 9$ seems to work well. The resulting bandwidth is $9/480 = 0.01875$. This means that with $L = 9$, we are assuming an approximately constant spectrum over $0.01875/0.5 = 0.0375$ of the entire frequency interval $[0, 1/2]$.

Smoothed Periodograms
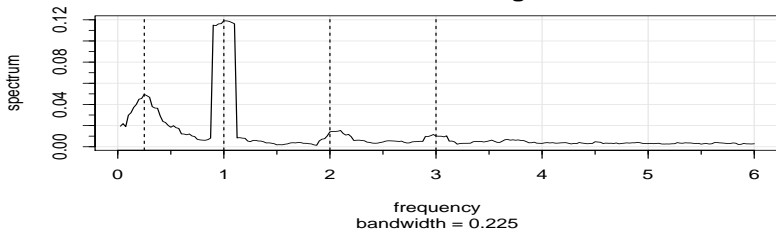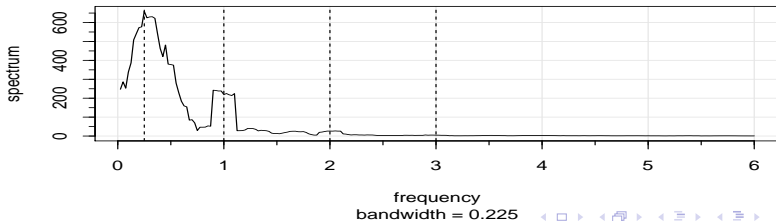
Observations from Smoothed Periodograms

From the smoothed periodograms:

- compromise with the noisy, unsmoothed version. Smoothing loses some of the peaks.

- sharp peaks at $\omega = 1/12$ now flattened to include nearby frequencies.

- smaller flat peaks at multiples of $\omega = 1/12$ appear.

Estimated Power and CIs from Smoothed Periodograms

The estimates of the power spectra and the confidence intervals using the symmetric average smoother with $L = 9$ are listed below.

Series	ω	Estimated Power	Lower	Upper
SOI	$\frac{1}{48}$	0.0495	0.0279	0.1113
	$\frac{1}{12}$	0.1191	0.0670	0.2677
Recruit ($\times 10^3$)	$\frac{1}{48}$	0.6590	0.3710	1.4815
	$\frac{1}{12}$	0.2194	0.1235	0.4932

Choice for Bandwidth

From (5), the bandwidth is $B = L/n$. Here are some
considerations as we decide the value of L in the bandwidth:

- If L is too large, our assumption that $f(\omega)$ is approximately
  constant within the frequency band may become more
  questionable. We may end up **smoothing out** valid peaks
  within a frequency band.

- If L is too small, the confidence intervals for $f(\omega)$ may
  become too wide, making it difficult to detect significance.

What is normally done: vary the value of L and examine the
resulting periodograms to see if common observations can be
made.

1 Smoothing the Periodogram

2 Generalization of Smoothing

3 (Modified) Daniell Kernel

4 Worked Example

Generalization of Smoothing

A generalization of smoothing for the periodogram is

$$\widehat{f}(\omega) = \sum_{k=-m}^{m} h_k I(\omega_j + k/n) \tag{9}$$

where $h_{-k} = h_k > 0$ and $\sum_{k=-m}^{m} h_k = 1$. The asymptotic distribution of chi-squared continues to hold as $m/n \to 0$ which will give that $\sum_{k=-m}^{m} h_k^2 \to 0$.

Generalization of Smoothing

Under these conditions $E(\widehat{f}(\omega)) \to f(\omega)$. This estimator also has the following asymptotic distribution

$$\frac{2L_h \widehat{f}(\omega)}{f(\omega)} \to \chi^2_{2L_h} \tag{10}$$

where $L_h = \left(\sum_{k=-m}^{m} h_k^2\right)^{-1}$.

Generalization of Smoothing

When we smooth a periodogram, we smooth across a frequency band. Recall that the periodogram is computed at the fundamental frequencies $\omega_j = j/n$ for $j = 1, 2, \cdots, n/2$. When smoothing is applied to a periodogram, $\widehat{f}(\omega)$ is a **weighted average** of the periodogram values for frequencies in the range $\frac{j-m}{n}$ to $\frac{j+m}{n}$.

Generalization of Smoothing

Note that when we set $h_k = L^{-1}$ for all k, where $L = 2m + 1$, we obtain the symmetric moving average smoother we defined earlier in (2),

$$\overline{f}(\omega) = \frac{1}{L} \sum_{k=-m}^{m} I(\omega_j + k/n).$$

Bandwidth

We earlier defined the bandwidth as $B = \frac{L}{n}$. The bandwidth is a measure of the width of the frequency intervals used for smoothing the periodogram. When **unequal weights** are used, this definition is modified to become

$$B = \frac{L_h}{n} = \frac{1/\sum_{k=-m}^{m} h_k^2}{n}. \tag{11}$$

Equation (11) holds for equal weights and is a generalization.

CI for Smoothed Periodogram

When using the smoothed periodogram as defined in (9), the approximate $(1 - \alpha) \times 100\%$ confidence interval for the power spectrum at ω is

$$\frac{2L_h \widehat{f}(\omega)}{\chi^2_{2L_h}(1 - \alpha/2)} \leq f(\omega) \leq \frac{2L_h \widehat{f}(\omega)}{\chi^2_{2L_h}(\alpha/2)}. \tag{12}$$

If FFT and zero padding are used, then replace $2L_h$ in (12) with $df = 2L_h n/n'$.

Simultaneous CIs

- In some situations, there is more than one frequency of interest, so we want to to make simultaneous inferences about the collection of frequencies we are interested in.
- To control the overall type I error rate, we work with the Bonferroni inequality $1 - K\alpha$ where K denotes the number of frequencies we are interested in.
- Thus, to maintain $1 - \alpha$ confidence, the significance level for each frequency should be α/K.

Simultaneous CIs

The confidence interval (12) then becomes

$$\frac{2L_h\widehat{f}(\omega)}{\chi^2_{2L_h}(1-\alpha/2K)} \leq f(\omega) \leq \frac{2L_h\widehat{f}(\omega)}{\chi^2_{2L_h}(\alpha/2K)}. \tag{13}$$

Going back to the worked example with the SOI and recruitment datasets, we constructed confidence intervals for the power spectrum at $\omega = 1/48$ and $1/12$. So $K = 4$ if we want to have at least 95% confidence that we do not make at least one type I error.

Daniell Kernel

Next, we explore some of the ways to generate the weights h_k, in R. A common choice for smoothing is to use the Daniell kernel. The Daniell kernel (initially) puts same weights on neighbors.

Question: Does this sound familiar?

Daniell Kernel

For example, consider $m = 1$ and $L = 2m + 1 = 3$, the Daniell kernel has weights $\{h_k\} = \{1/3, 1/3, 1/3\}$. Applying this kernel to a sequence $\{u_t\}$ produces

$$\hat{u}_t = \frac{1}{3} u_{t-1} + \frac{1}{3} u_t + \frac{1}{3} u_{t+1}.$$

Daniell Kernel

We can apply the same kernel again to the $\hat{x}_t$ to obtain

$$\hat{\hat{u}}_t = $$
$$= $$

Daniell Kernel

Applying the Daniell kernel to obtain a smoothed periodogram as defined in (9), we obtain

$$\widehat{f}(\omega) \;\; = \;\; \frac{1}{3}I(\omega_j - 1/n) + \frac{1}{3}I(\omega_j) + \frac{1}{3}I(\omega_j + 1/n).$$

Daniell Kernel

If we use the Daniell kernel twice, we obtain

$$\begin{aligned}
\widehat{f}(\omega) &= \frac{1}{9}I(\omega_j - 2/n) + \frac{2}{9}I(\omega_j - 1/n) + \frac{3}{9}I(\omega_j) \\
&+ \frac{2}{9}I(\omega_j + 1/n) + \frac{1}{9}I(\omega_j + 2/n).
\end{aligned}$$

A higher degree of smoothing is achieved by repeating the kernel.

Modified Daniell Kernel

The modified Daniell kernel is such that the two endpoints in the averaging receive half the weight that the interior points do. For this example, the modified Daniell kernel puts one half on the center and one quarter on each. So, the sequence $\{u_t\}$ appears as follows:

$$\hat{u}_t = \frac{1}{4}u_{t-1} + \frac{1}{2}u_t + \frac{1}{4}u_{t+1}.$$

Modified Daniell Kernel

Applying the modified Daniell kernel again produces

$$
\begin{aligned}
\hat{\hat{u}}_t &= \frac{1}{4}\hat{u}_{t-1} + \frac{1}{2}\hat{u}_t + \frac{1}{4}\hat{u}_{t+1} \\
&= \frac{1}{16}u_{t-2} + \frac{4}{16}u_{t-1} + \frac{6}{16}u_t + \frac{4}{16}u_{t+1} + \frac{1}{16}u_{t+2}.
\end{aligned}
$$

1 Smoothing the Periodogram

2 Generalization of Smoothing

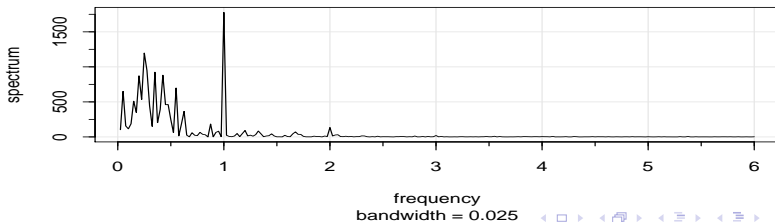3 (Modified) Daniell Kernel

4 Worked Example

Worked Example

We return to the SOI and recruit datasets. We use the modified Daniell kernel twice, with $m = 3$ both times. This results in $L_h = 1/\sum_{k=-m}^{m} h_k^2 = 9.232$. The bandwidth now is $B = 9.232/480 = 0.019$. Compare these to $L = 9$ and $B = 0.01875$ from the earlier worked example, when we used the Daniell kernel.

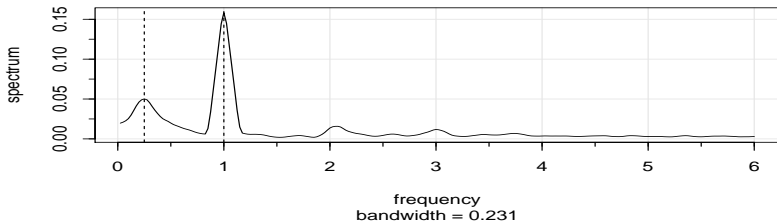# Raw Periodograms

Series: soi
Raw Periodogram

frequency
bandwidth = 0.025
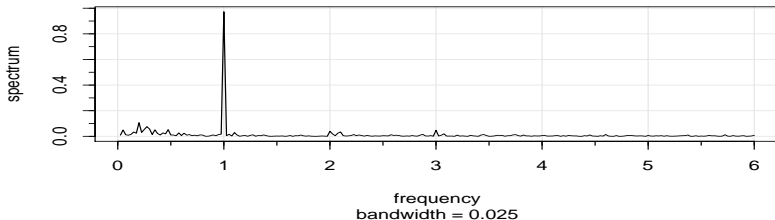
Series: rec
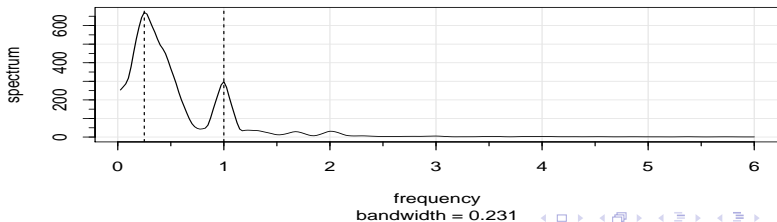**Raw Periodogram**

frequency
bandwidth = 0.025

Smoothed Periodograms with Modified Daniell Kernel

Series: soi
Smoothed Periodogram

Series: rec
Smoothed Periodogram

Estimated Power and CIs from Smoothed Periodograms

The estimates of the power spectra and the confidence intervals using the modified Daniell kernel twice are listed below.

Series	ω	Estimated Power	Lower	Upper
SOI	$\frac{1}{48}$	0.0501	0.0284	0.1113
	$\frac{1}{12}$	0.1582	0.0896	0.3513
Recruit $(\times 10^3)$	$\frac{1}{48}$	0.6709	0.3801	1.4898
	$\frac{1}{12}$	0.2981	0.1688	0.6619