Unit 21: Nonparametric Spectral Density Estimation

Jeffrey Woo

Department of Statistics, University of Virginia

Spring 2016

Readings for Unit 21

Textbook chapter 4.5.

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 ________ 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 never shrinks, regardless of the size of the sample.

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 an interval of frequencies that we will call a ________________ 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 ________________________ in the frequency band.

Smoothing the Periodogram

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

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

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 $n \to \infty$,

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

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

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

For large L, these intervals will be ________ 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 shrinking just like the regular sample variance. Because of the spacing of the fundamental frequencies, we call

$$B_w = \frac{L}{n} \tag{5}$$

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

$$2L = 2B_w n.$$

Bias-Variance Tradeoff

For bandwidth $B_w = \frac{L}{n}$, the variance of $\overline{f}(\omega) \approx \frac{c}{B_w n}$ for some constant c. So to ________________, 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 ___________.

Log Transformation

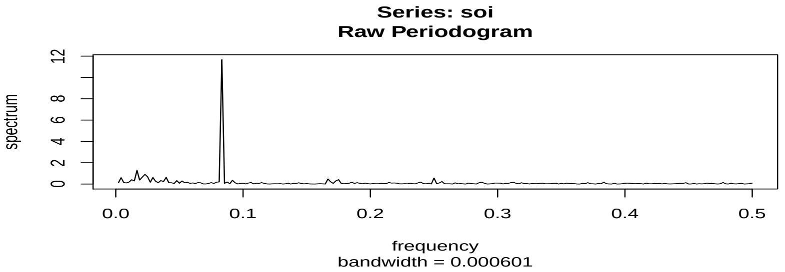
Many times, a log transformation of the spectrum can aid in the visual display of the spectral density plot. This can happen when regions of the spectrum exist with peaks of interest much smaller than some of the main components. For the log spectrum, the approximate $(1-\alpha) \times 100\%$ confidence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

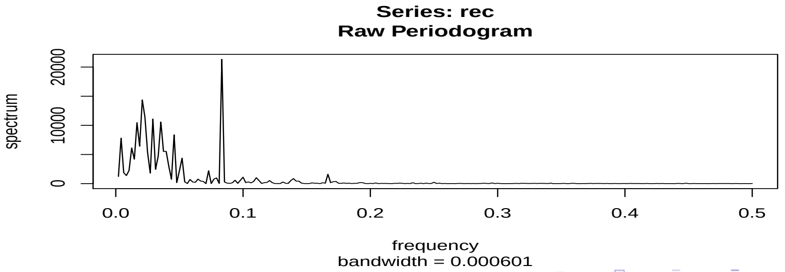
$$[\log \overline{f}(\omega)+\log 2L-\log \chi^2_{2L}(1-\alpha/2), \log \overline{f}(\omega)+\log 2L-\log \chi^2_{2L}(\alpha/2)] \tag{6}$$

SOI and Recruitment

We will continue with the example from Unit 20, with the Southern Oscillation Index ("soi.dat") and recruitment ("recruit.dat") 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 Nino.

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

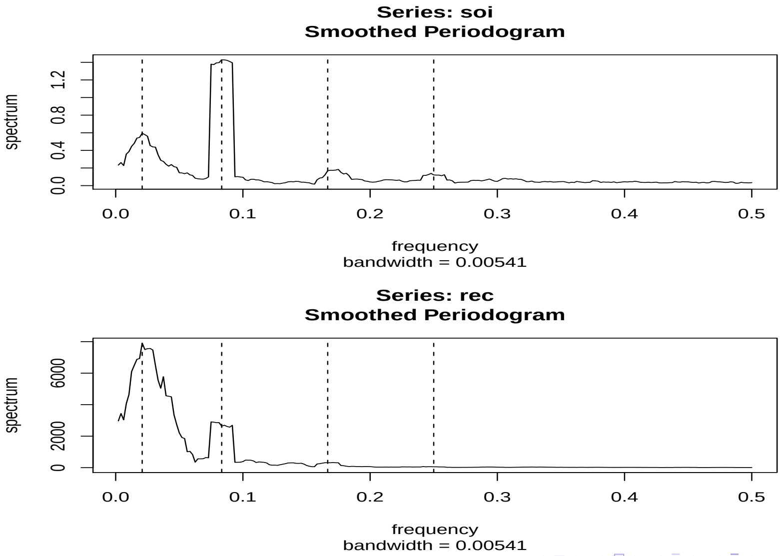
The estimates of the power spectra and the confidence intervals using the raw periodogram are listed below.

Series	ω	Estimated Power	Lower	Upper
SOI	$\frac{1}{48}$	0.6448	0.1748	25.4665
	$\frac{1}{12}$	11.6668	3.1627	460.8130
Recruit ($\times 10^3$)	$\frac{1}{48}$	14.368	3.895	567.52
	$\frac{1}{12}$	21.333	5.783	842.61

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 smoothed periodogram with $L = 9$ are listed below.

Series	ω	Estimated Power	Lower	Upper
SOI	$\frac{1}{48}$	0.5942	0.3345	1.336
	$\frac{1}{12}$	1.4290	0.8045	3.2126
Recruit ($\times 10^3$)	$\frac{1}{48}$	7.908	4.452	17.778
	$\frac{1}{12}$	2.633	1.482	5.919

Generalization of Smoothing

A generalization of smoothing for the periodogram is

$$\widehat{f}(\omega) = \sum_{k=-m}^{m} h_k I(\omega_j + k/n) \tag{7}$$

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

$$\frac{2L_h \widehat{f}(\omega)}{f(\omega)} \to \chi^2_{2L_h} \tag{8}$$

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 interval. 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 ________ ________ 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 explored earlier,

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

Bandwidth

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

$$B_w = \frac{L_h}{n} = \frac{1/\sum_{k=-m}^{m} h_k^2}{n}. \tag{9}$$

Equation (9) holds for equal weights and is a generalization.

CI for Smoothed Periodogram

When using the smoothed periodogram as defined in (7), the approximate $(1-\alpha)\times 100\%$ confidence interval for the power spectrum at ω is

$$\frac{2L_h\widehat{f}(\omega)}{\chi^2_{2L_h}(1-\alpha/2)} \leq f(\omega) \leq \frac{2L_h\widehat{f}(\omega)}{\chi^2_{2L_h}(\alpha/2)}. \tag{10}$$

Bandwidth

- A bandwidth too great will oversmooth the periodogram and we may miss seeing important peaks.
- The bandwidth is predominantly controlled by the value m or L, and whether we repeat the kernel (modified Daniell).
- It takes some experimentation to find a bandwidth that gives suitable smoothing.
- Note that the bandwidth reported in the periodogram in R is not calculated using formula (9).

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

$$
\begin{aligned}
\hat{\hat{u}}_t &= \\
&=
\end{aligned}
$$

Daniell Kernel

Applying the Daniell kernel to obtain a smoothed periodogram as defined in (7), 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

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

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

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

More about Bandwidth

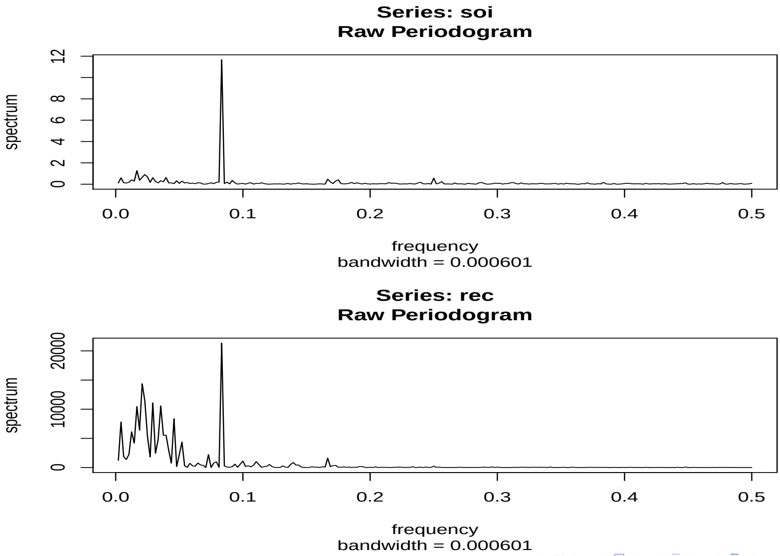
The authors of our textbook define the bandwidth per equation (9). This definition is not used by R (and some other authors). An alternate definition of the bandwidth is that it is the standard deviation of the weighting distribution. Let's look at the Daniell kernel.

More about Bandwidth: Daniell Kernel

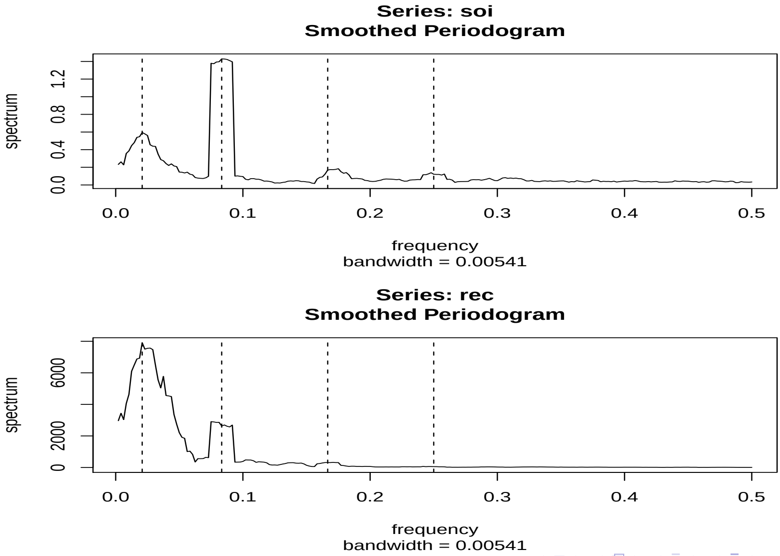
Worked Example

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 $B_w = 9.232/480 = 0.019$. Compare these to $L = 9$ and $B_w = 0.01875$ from the earlier worked example, when we used the Daniell kernel.

Raw Periodograms



Smoothed Periodograms with Daniell Kernel



Smoothed Periodograms with Modified Daniell Kernel