STAT 5170: Applied Time Series

Course notes for part B of learning unit 6

Section 6B.1: Problem setup: inference on the spectral density.

Part A of the learning unit ended with a discussion of inference on the spectral density of a time series using the periodogram as a sample estimate. Recall the setup to that discussion:

The spectral density of the stationary time series (x_t) is $f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h}$. Given the first n measurements of the time series, $x_1, \ldots, x_n$, the discrete Fourier transform (DFT) is $d(\omega_k) = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} x_{t+1} e^{-2\pi i \omega_k t}$, and the periodogram is $I(\omega_k) = |d(\omega_k)|^2$, the squared magnitude of the DFT.

Both the DFT and the periodogram are typically defined at the fundamental Fourier frequencies $\omega_k = k/n$ for $k = 0, \ldots, (n-1)/2$ if n is odd and $k = 0, \ldots, n/2 - 1$ if n is even. Note that in either the odd or even case the upper index value may be written $\lfloor (n-1)/2 \rfloor$, where $\lfloor u \rfloor$ is the largest integer smaller than u (a.k.a., the "floor" of u). The DFT and the periodogram may also be calculated at other frequencies, but the restrictions above are most useful because of the following: (*i.*) working only with integer values admits use of the fast Fourier transform (FFT), a computationally algorithm for *calculating* the DFT; (*ii.*) when working with discrete data, consideration of frequencies ω larger than $1/2$ would be redundant because of aliasing. It is worthwhile to also point out that the frequency $\omega_0 = 0$, defined at $k = 0$, is sometimes excluded from the fundamental Fourier frequencies because the resulting DFT is simply

$$d(\omega_0) = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} x_{t+1} e^{-2\pi i \omega_0 t} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t = \sqrt{n}\bar{x}.$$

Key results for inference on the spectral density are provided by large-sample asymptotic study of the periodogram's sampling properties. When n is large,

(*i.*) The periodogram values $I(\omega_j)$ and $I(\omega_k)$ at distinct $j \neq k$ are approximately independent;

(*ii.*) For a given sequence $k_1, k_2, \ldots$ such that $\omega_{k_n} \to \omega$, the ratio $2I(\omega_k)/f(\omega)$ is approximately χ^2_2, a chi-square random variable with just two degrees of freedom.

Property (*ii*) in the list is especially motivating to the discussion that we take on here. Recall the following well-known formulas that provide that the mean and variance of a chi-square random variable random variable: if $V \sim \chi^2_\nu$, then $E[V] = \nu$ and $Var[V] = 2\nu$. By setting $V = 2I(\omega_k)/f(\omega)$, $\nu = 2$, and rearranging terms, the periodogram is $I(\omega_k) = f(\omega)V/2$, and, by property (*ii*), its approximate mean and variance when n is large are

$$\begin{aligned} E[I(\omega_k)] &\approx f(\omega)E[V]/2 = f(\omega) \\ Var[I(\omega_k)] &\approx f(\omega)^2 Var[V]/4 = f(\omega)^2 \end{aligned}$$

A troubling aspect of this expression is that the variance of the periodogram does not tend to zero as n increases. This means, *e.g.*, that a confidence interval for $f(\omega)$ would *not* tend to become more narrow, and would *not* eventually achieve an arbitrarily small width as n more time series data are measured.

Our present goal for this part of the learning unit is to identify an estimate, $\hat{f}(\omega_k)$, of the spectral density, $f(\omega)$, such that the ratio $\nu_n \hat{f}(\omega_k)/f(\omega)$ is approximately $\chi^2_{\nu_n}$, with a degrees of freedom parameter, ν_n, that diverges to infinity, $\nu_n \to \infty$. Supposing that such an estimate can be found, setting $V = \nu_n \hat{f}(\omega_k)/f(\omega)$ and rearranging terms will establish that the approximate mean and variance of $\hat{f}(\omega_k)$ are

$$\begin{aligned} E[\hat{f}(\omega_k)] &\approx f(\omega)E[V]/\nu_n = f(\omega) \\ Var[\hat{f}(\omega_k)] &\approx f(\omega)^2 Var[V]/\nu_n^2 = 2f(\omega)^2/\nu_n \to 0. \end{aligned}$$

That is, the variance of the estimate becomes arbitrarily close to zero as the time-series length, n, increases. Correspondingly, a confidence interval for $f(\omega)$ derived from this result would become more narrow as more time series data are measured. To be explicit, a formula for the approximate $(1-\alpha)100\%$ confidence interval for $f(\omega)$ associated with this result has endpoints

$$\frac{\nu_n \hat{f}(\omega_k)}{\chi^2_{\nu_n,1-\alpha/2}} \quad \text{and} \quad \frac{\nu_n \hat{f}(\omega_k)}{\chi^2_{\nu_n,\alpha/2}}$$

As $\nu_n \to \infty$, normal approximation to chi-square random variables provides that two chi-square quantiles in the above formula are approximately $\chi^2_{\nu_n,\alpha/2} \approx \nu_n + z_{\alpha/2}\sqrt{2\nu_n}$ and $\chi^2_{\nu_n,1-\alpha/2} \approx \nu_n + z_{1-\alpha/2}\sqrt{2\nu_n}$, having written z_p for the p'th quantile of a standard-normal distribution. The width of the confidence interval is therefore approximately

$$\begin{aligned} \frac{\nu_n \hat{f}(\omega_k)}{\chi^2_{\nu_n,\alpha/2}} - \frac{\nu_n \hat{f}(\omega_k)}{\chi^2_{\nu_n,1-\alpha/2}} &= \hat{f}(\omega_k)\left(\frac{\nu_n}{\chi^2_{\nu_n,\alpha/2}} - \frac{\nu_n}{\chi^2_{\nu_n,1-\alpha/2}}\right) \\ &\approx \hat{f}(\omega_k)\left(\frac{\nu_n}{\nu_n + z_{\alpha/2}\sqrt{2\nu_n}} - \frac{\nu_n}{\nu_n + z_{1-\alpha/2}\sqrt{2\nu_n}}\right) \\ &= \hat{f}(\omega_k)\left(\frac{1}{1 + z_{\alpha/2}\sqrt{2/\nu_n}} - \frac{1}{1 + z_{1-\alpha/2}\sqrt{2/\nu_n}}\right) \to 0. \end{aligned}$$

That is, the confidence interval width converges to zero.

One way to produce an estimate, $\hat{f}(\omega_k)$, with the desired properties described above is to apply a technique known as *smoothing* to the periodogram. We now turn our attention to this technique.

Section 6B.2: Smoothing the periodogram.

In the context of inference on the spectral density, the most basic version of smoothing is to group the periodogram and calculate what is sometimes called a "running mean." The technique is formulated as follow.

To start, let us associate with any frequency ω the *frequency band* of fundamental Fourier frequencies,

$$\mathcal{B}_m(\omega) = \{\omega_j : |\omega - \omega_j| \leq m/n, j = 0, \ldots, \lfloor (n-1)/2 \rfloor\}.$$

Denote by L the size of the frequency band, *i.e.*, the total number of fundamental Fourier frequencies it contains. An associated terminology defines the *bandwidth* of $\mathcal{B}_m(\omega_k)$ as $B = L/n$.

When $\omega = \omega_k$ is itself a fundamental Fourier frequency, the frequency band is

$$\mathcal{B}_m(\omega_k) = \{\omega_j = \omega_k + j/n : j = 0, \pm 1, \ldots, \pm m\},$$

its size is $L = 2m + 1$, and its bandwidth is approximately $B \approx 2m/n$, provided that ω_k is sufficiently far from 0 and $1/2$.

The *smoothed periodogram* at ω is defined by averaging values of the raw periodogram at frequencies within the frequency band

$$\hat{f}(\omega) \;=\; \frac{1}{L} \sum_{\omega_j \in \mathcal{B}_m(\omega)} I(\omega_j).$$

When $\omega = \omega_k$ is a fundamental Fourier frequency, this quantity is

$$\hat{f}(\omega_k) \;=\; \frac{1}{2m+1} \sum_{j=-m}^{m} I(\omega_k + j/n).$$

As it turns out, this statistic can achieve the type of asymptotic sampling property that we seek, provided that it is configured in the right way, and that the underlying spectral density does not take on a pathological form, such as having a discontinuity exactly at the frequency at which inference is directed. Typically, it is simply assumed the spectral density is well-behaved, unless worrisome patterns are revealed while carrying out the analysis. As for suitable configurations of the smoothed periodogram, the primary requirement is that the frequency band contains a fairly large number of frequencies, but its size, L, is still much smaller than n.

Few definitive, general rules are available for specifying the frequency band. One reason for this lack is that analysts often emphasize general criteria for how informative the smoothed periodogram looks to the eye. Another is that precise rules, though many have been derived, are typically specialized to the particular features of the spectral density, as approximated by the periodogram. It is often effective to simply experiment to find a bandwidth that avoids *oversmoothing* interesting peaks or *undersmoothing* noisy regions.

Mathematical study of the smoothed periodogram typically conceptualizes L as dependent on time-series length, n, as is indicated in the modified notation L_n. It is then supposed that

A. The size of the frequency band $L_n \to \infty$ as $n \to \infty$, which is interpreted to mean that the frequency band contains a fairly large number of frequencies; and

B. The bandwidth $B_n = L_n/n \to 0$ as $n \to \infty$, which is interpreted to mean that L_n is much smaller than n.

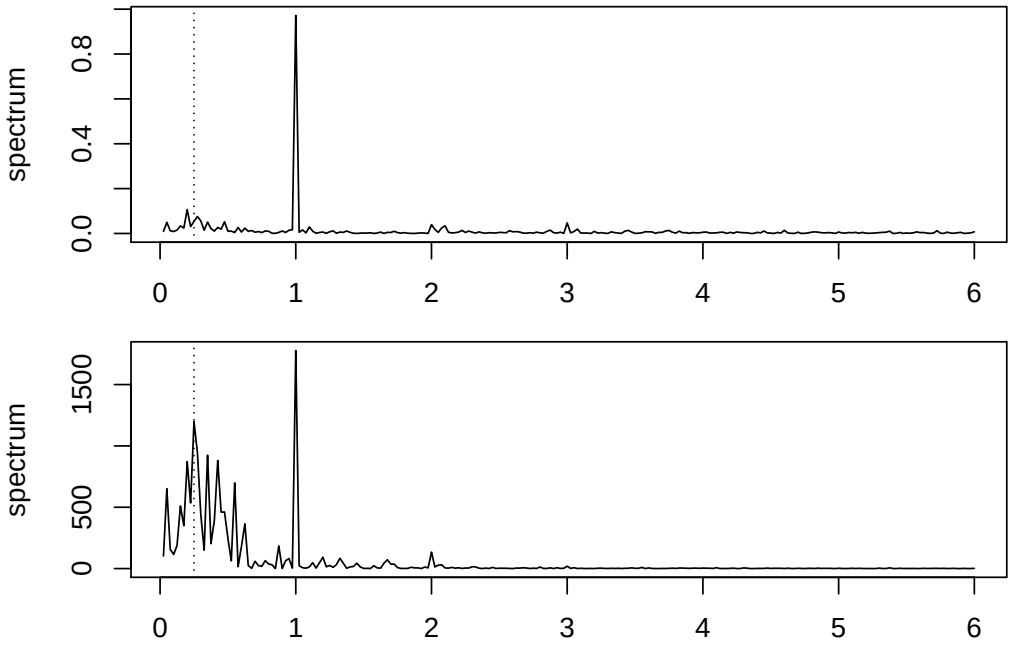
Under these conditions, it can be established that, when n is large...

(*iii.*) For a given sequence $k_1, k_2, \ldots$ such that $\omega_{k_n} \to \omega$, the ratio $\nu_n \hat{f}(\omega)/f(\omega)$ is approximately $\chi^2_{\nu_n}$, where the degrees of freedom parameter is $\nu_n = 2L_n$.

The following example illustrates the use of the running means smoother, as well as at least one of its deficiencies.

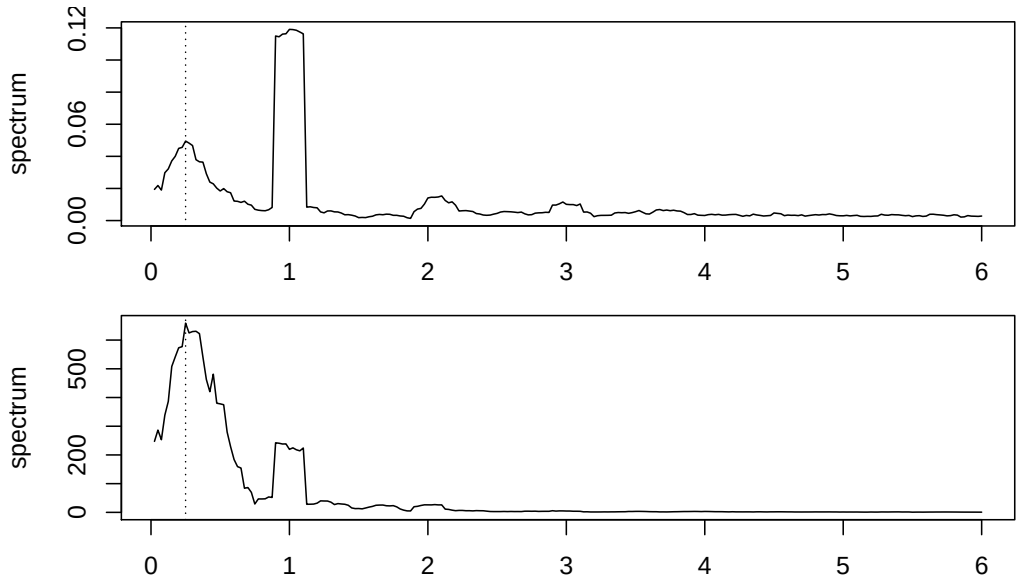
Example: Pacific Ocean time series

In previous examples, we explored the periodograms of two time series of $n = 453$ monthly measurements from the Pacific Ocean made during the years 1950-1987: Southern Oscillation Index (SOI) time series, which measures ocean surface temperature, and a times series of fish population counts. The raw periodograms of both series are displayed below, plotted with a horizontal axis labeled in units of 1/12, so that *e.g.*, the frequency $\omega = 1/12$ (*i.e.*, a yearly cycle) is one unit, and the frequency $\omega = 2/12$ (*i.e.*, a half-year cycle) is two units, *etc.*



As we can see, each series exhibits a strong yearly cycle, at $\omega = 1/12$, and indications of a four-year cycle, at $\omega = 1/48$.

The figure below displays smooth versions of the same periodogram, having set $m = 4$, so that there are $L = 2m + 1 = 9$ points in the typical frequency band $\mathcal{B}(\omega_k)$, whose bandwidth is $B = 9/453 = 0.0199$.



Observe there is some smearing around $\omega = 1/12$ at which the sharp peak in the spectral density may be sensible; however, overall, the smoothed periodogram depicts the nice general shape for a spectral density.

In a previous example, approximate 95% confidence intervals were calculated from the raw periodogram. A summary of these values is shown in the table below, on the original and logarithmic scale.

Series	ω	est. $f(\omega)$	lower CI	upper CI	log est. $\hat{f}(\omega)$	lower CI	upper CI
SOI	1/12	0.9722	0.2636	38.4011	-0.0282	-1.3335	3.6481
	1/48	0.0537	0.01457	2.1222	-2.9238	-4.2291	0.7525
Fish	1/12	1777.75	481.92	70217.18	7.4831	6.1778	11.1593
	1/48	1197.37	324.59	47293.54	7.0879	5.7826	10.7641

The table below shows corresponding estimates and confidence intervals based on the smoothed periodogram.

Series	ω	est. $f(\omega)$	lower CI	upper CI	log est. $\hat{f}(\omega)$	lower CI	upper CI
SOI	1/12	0.1191	0.0670	0.2677	-2.1280	-2.7025	-1.3178
	1/48	0.0495	0.0279	0.1113	-3.0054	-3.5799	-2.1952
Fish	1/12	219.3919	123.51	493.24	5.3909	4.8163	6.2010
	1/48	658.96	370.98	1481.50	6.4907	5.9162	7.3008

Besides being calculated around different estimates of the periodogram, recall that the CI formulas of the two tables also differ in the degrees of freedom of the chi-square quantile values: the intervals around raw periodogram have $\nu = 2$ degrees of freedom, while those around the smoothed periodogram would have $\nu = 2L = 18$ degrees of freedom.

Comparing values between the two tables, it is seen that the estimated values can change quite substantially due to smoothing; *e.g.*, estimates at $\omega = 1/12$ decrease, while those at

$\omega = 1/48$ tend to increase. The important characteristic to notice, however, is that the widths of the confidence intervals decrease with the smoothed periodogram, indicating a drastic improvement in precision. □

Substantial smearing in a graph of smooth periodogram around sharper details is sometimes undesirable because those details may be associated with an essential characteristic of the spectral density that is not adequately represented in the graph.

A modification of the straightforward averaging approach is a smoothed periodogram that is defined from weighted averaging. Here, the smoothed periodogram at the fundamental Fourier frequency ω_k is defined by

$$\hat{f}^*(\omega_k) \quad = \quad \sum_{j=-m}^{m} h_j I(\omega_k + j/n),$$

where h_j is a set of weights such that $h_{-m} + h_{-m+1} + \cdots + h(0) + \cdots + h_{m-1} + h_m = 1$. For example, when each $h_j = 1/L$, the weighted smoothing operation reduces to smoothing by straightforward averaging, discussed before. Weighted smoothing can also be defined for the general case where inference on the spectral density is at a frequency ω that is not among the fundamental Fourier frequencies, but the notation is cumbersome and so is not discussed.

For weighted smoothing to offer an effective approach to inference, a modification to previously stated conditions is required:

A*. The modified frequency-band size $L_n^* = 1/\sum_{k=-m}^{m} h_k^2 \to \infty$ as $n \to \infty$.

B*. The modified bandwidth $B_n^* = L_n^*/n \to 0$ as $n \to \infty$.

Under the stated conditions, it can be established that, when n is large...

(*iii**.) For a given sequence $k_1, k_2, \ldots$ such that $\omega_{k_n} \to \omega$, the ratio $\nu_n \hat{f}^*(\omega)/f(\omega)$ is approximately $\chi^2_{\nu_n}$, where the degrees of freedom parameter is $\nu_n = 2L_n^*$.

Note that the modified frequency-band size in the unweighted case, where $h_j = 1/L$, reduces to

$$L_n^* = \frac{1}{\sum_{j=-m}^{m} h_j^2} = \frac{1}{(2m+1)(1/L)^2} = \frac{1}{(2m+1)/(2m+1)^2} = 2m + 1.$$

The next question is whether there is a scheme for selecting weights that would improve the appearance of the smoothed periodogram. One interesting and potentially helpful scheme is called the *Daniell kernel*. It is described from a certain perspective in the following way: from a sequence $\ldots, u_{-2}, u_{-1}, u_0, u_1, u_2, \ldots$, define a new sequence $\ldots, \hat{u}_{-2}, \hat{u}_{-1}, \hat{u}_0, \hat{u}_1, \hat{u}_2, \ldots$ according to

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

Repeating this defines another sequence $\ldots, \hat{\hat{u}}_{-2}, \hat{\hat{u}}_{-1}, \hat{\hat{u}}_0, \hat{\hat{u}}_1, \hat{\hat{u}}_2, \ldots$ according to

$$\begin{aligned}
\hat{\hat{u}}_t &= \frac{1}{3}\hat{u}_{t-1} + \frac{1}{3}\hat{u}_t + \frac{1}{3}\hat{u}_{t+1} \\
&= \left(\frac{1}{9}u_{t-2} + \frac{1}{9}u_{t-1} + \frac{1}{9}u_t\right) + \left(\frac{1}{9}u_{t-1} + \frac{1}{9}u_t + \frac{1}{9}u_{t+1}\right) + \left(\frac{1}{9}u_t + \frac{1}{9}u_{t+1} + \frac{1}{9}u_{t+2}\right) \\
&= \frac{1}{9}u_{t-2} + \frac{2}{9}u_{t-1} + \frac{3}{9}u_t + \frac{2}{9}u_{t+1} + \frac{1}{9}u_{t+2}
\end{aligned}$$

The constant appearing before the value u_{t-j} defines the weight h_j. Thus, the constants appearing in the example defined the weights $h_{-2}, h_{-1}, h_0, h_1, h_2 = 1/9, 2/9, 3/9, 2/9, 1/9$, which is suitable for a frequency band defined with $m = 2$.

The *modified Daniell kernel* applies the same scheme, but starts with an uneven set of initial weights. For example, when the starting set of weights are $h_{-1}, h_0, h_1 = 1/4, 1/2, 1/4$, the calculation works out as

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

and

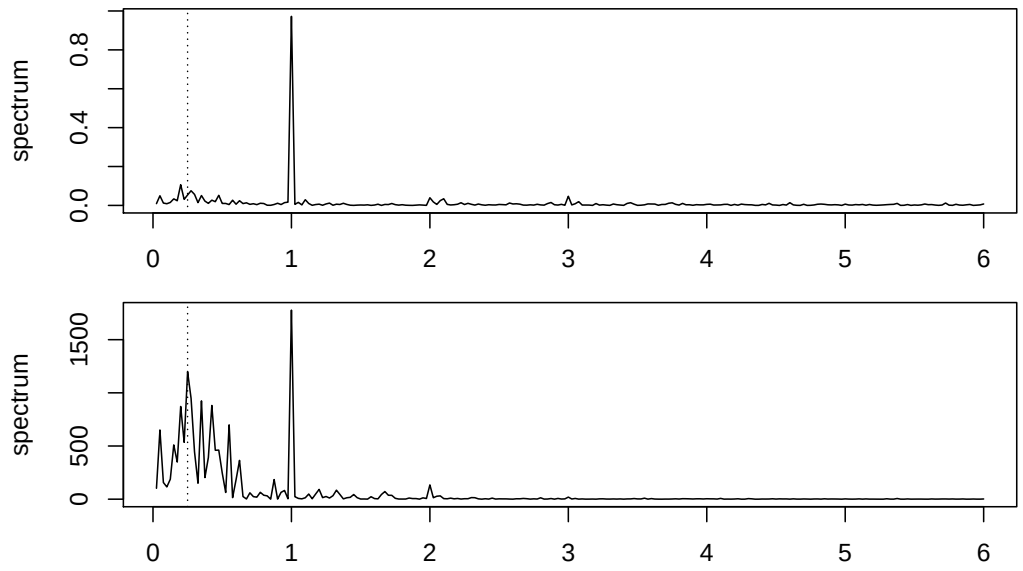
$$\begin{aligned}
\hat{\hat{u}}_t &= \frac{1}{4}\hat{u}_{t-1} + \frac{1}{2}\hat{u}_t + \frac{1}{4}\hat{u}_{t+1} \\
&= \left(\frac{1}{16}u_{t-2} + \frac{1}{8}u_{t-1} + \frac{1}{16}u_t\right) + \left(\frac{1}{8}u_{t-1} + \frac{1}{4}u_t + \frac{1}{8}u_{t+1}\right) + \left(\frac{1}{16}u_t + \frac{1}{8}u_{t+1} + \frac{1}{16}u_{t+2}\right) \\
&= \frac{1}{16}u_{t-2} + \frac{1}{4}u_{t-1} + \frac{3}{8}u_t + \frac{1}{4}u_{t+1} + \frac{1}{16}u_{t+2},
\end{aligned}$$

which produces the final set of weights $h_{-2}, h_{-1}, h_0, h_1, h_2 = 1/16, 1/4, 3/8, 1/4, 1/16$.

To produce weights for a larger frequency band, construction with $m > 2$, the scheme may be carried through further iterations, or it may begin with a larger set of initial weights. For instance, an example below starts with the initial weights $h_{-3}, \ldots, h_3 = 1/12, 1/6, 1/6, 1/6, 1/6, 1/6, 1/12$. Refer to the numerical examples accompanying this part of the learning unit for a general algorithm for calculating the weights of a modified Daniell kernel.

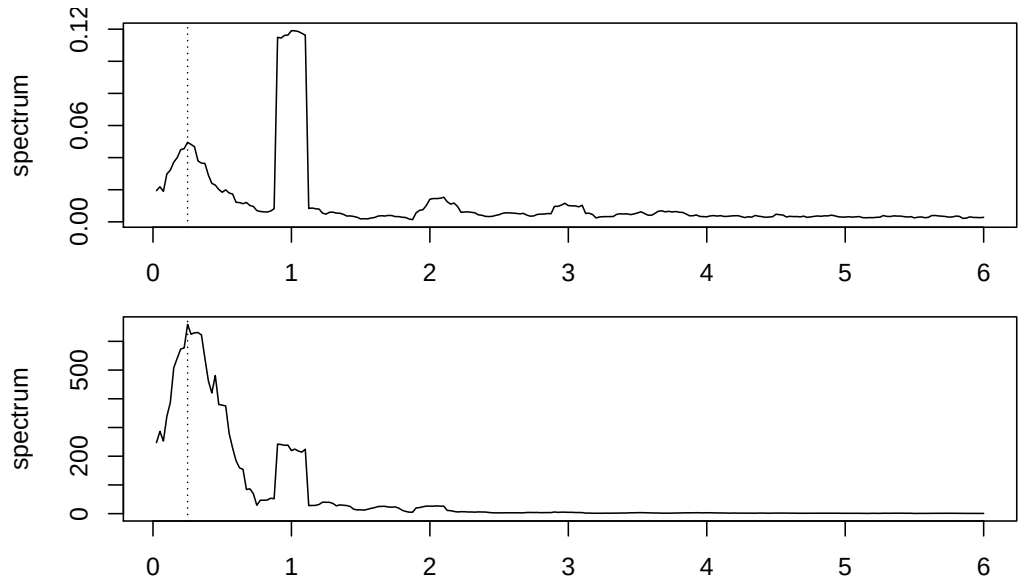
Illustration: Smoothed periodogram comparisons

Let us use the Pacific Ocean time series examined in previous example to compare the behavior of different smoothing approaches. To start, let us recall the raw periodograms of the SOI and fish-count time series, which are plotted as follows.

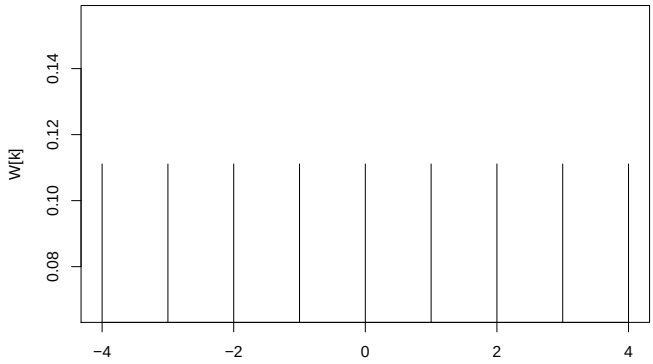


These periodgrams are too noisy to be useful, which is why we smooth them.

In the previous example, straightforward averaging was applied to produce the following smoothed periodograms.



This approach might be viewed weighted smoothing with even weights, $h_k = 1/L$. Here $L = 2m + 1 = 9$. A plot of the weights is as follows.



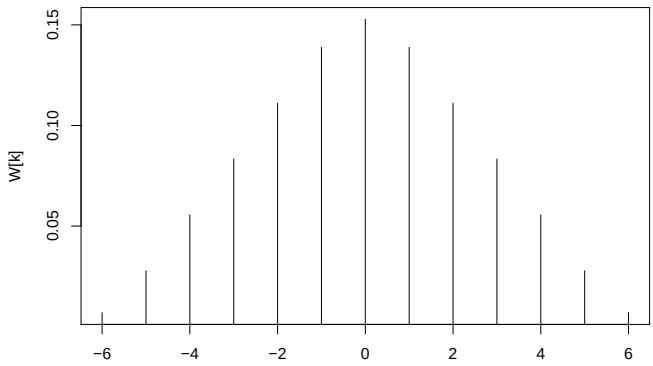
The plot below is of smoothed periodograms obtained using a modified Daniell kernel scheme that begins with the initial weight values

$$\tfrac{1}{12}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{12},$$

an carries them through one iteration of the calculation. The result is

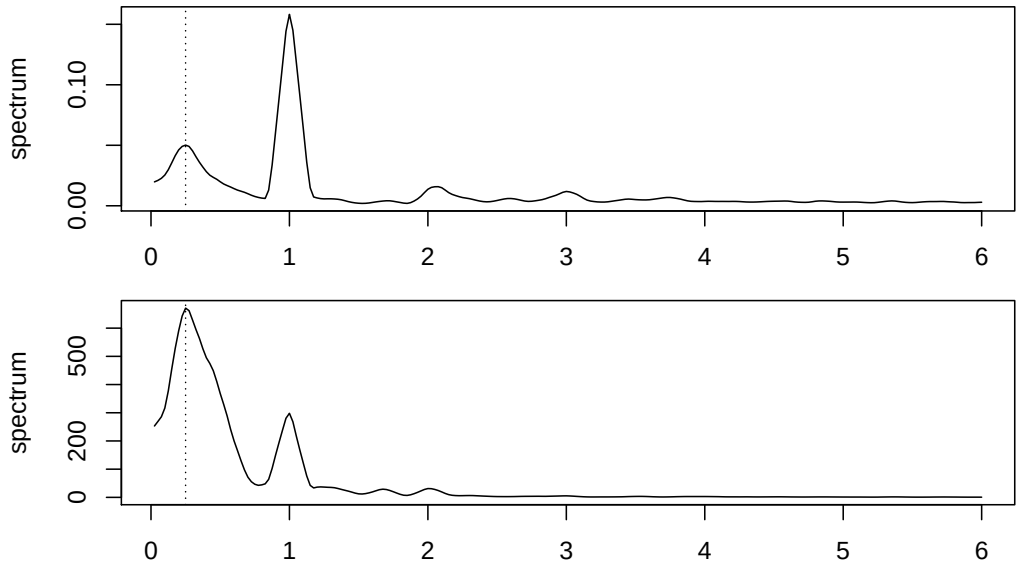
$$h_1, \ldots, h_{13} = \tfrac{1}{144}, \tfrac{4}{144}, \tfrac{8}{144}, \tfrac{12}{144}, \tfrac{16}{144}, \tfrac{20}{144}, \tfrac{22}{144}, \tfrac{20}{144}, \tfrac{16}{144}, \tfrac{12}{144}, \tfrac{8}{144}, \tfrac{4}{144}, \tfrac{1}{144},$$

which are applied to the periodogram for smoothing. A plot of the weights is



With such h_k, one would calculate $L_h = 1/\sum_{k=-m}^{m} h_k^2 = 9.2324$, which is comparable to the value $L = 9$ of the smoothed periodogram obtained by straightforward averaging.

The smoothed periodograms obtained by weighted averaging is as follows



Observe this approach to smoothing handles the peaks at $\omega = 1/12$ in a more pleasing way than smoothing by straightforward averaging, but produces the same general smoothed shape at lower frequencies, such as around $\omega = 1/48$. □

Section 6B.3: Linear filters.

Linear filtering refers to a general framework for emphasizing patterns within a time series. It includes the smoothing techniques discussed in the previous sections of these notes, and goes beyond to give perspective to other ideas that have been discussed in previous learning units. The definition of its central concepts is as follows.

Definition: A *linear filter* is the operation translating the time series (x_t) into (y_t) according to

$$y_t = \sum_{j=-\infty}^{\infty} a_j x_{t-j}$$

for coefficients a_j such that

$$\sum_{j=-\infty}^{\infty} |a_j| < \infty.$$

□

An intriguing aspect of linear filters is how compatible they are with ideas connected to spectral densities. An important property is the following:

Property: In the definition above, the spectral densities of (x_t) and (y_t) are related according to

$$f_y(\omega) = |A_{yx}(\omega)|^2 f_x(\omega)$$

where

$$A_{yx}(\omega) = \sum_{j=-\infty}^{\infty} a_j e^{-2\pi i \omega j}.$$

□

This property suggests that the patterns emphasized by a linear filter can be understood through the spectral density, as a later example will demonstrate. In the terminology associated with linear filters, the coefficients the a_j are called, as a collection, the *impulse response function*, the quantity $A_{yx}(\omega)$ is called the *frequency response function*, and its squared magnitude, $|A_{yx}(\omega)|^2$, is called the *power transfer function*. The formula $f_y(\omega) = |A_{yx}(\omega)|^2 f_x(\omega)$ is sometimes regarded as an analogue to the elementary relationship $Var[Y] = a^2 Var[X]$ for random variables X and Y related by $Y = aX$.

It is clear from the definition of a linear filter that it captures the underlying intuition of a moving average. Consider, for example, the $MA(\infty)$ representation of an $ARMA(p, q)$ time series,

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j},$$

which we can now characterize as a linear filter of white noise. Taking this further let us observe that the frequency response function at play here is

$$A_{xw}(\omega) = \sum_{j=0}^{\infty} \psi_j e^{-2\pi i \omega j} = \psi(e^{-2\pi i \omega}) = \frac{\theta(e^{-2\pi i \omega})}{\phi(e^{-2\pi i \omega})}.$$

If we furthermore recall that the spectral density of the white noise time series (w_t) is $f_w(\omega) = \sigma_w^2$, we have that the spectral density of (x_t) is

$$f_x(\omega) = |A_{xw}(\omega)|^2 f_w(\omega) = \sigma^2 \frac{|\theta(e^{-2\pi i \omega})|^2}{|\phi(e^{-2\pi i \omega})|^2}$$

This reestablishes a property that we have seen before. The point of these comments is to illustrate that linear filters extend the previous theory of spectral densities.

The following example describes other ways in which linear filters extend other ideas we have seen before, including smoothing. Along the way we come across at least one new tool that is potentially useful for applied analysis.

Example: High- and low-pass filters

Suppose the impulse response function is defined according to

$$a_j = \begin{cases} 1 & \text{if } j = 0 \\ -1 & \text{if } j = 1 \\ 0 & \text{otherwise} \end{cases}$$

The corresponding linear filter is

$$y_t = \sum_{j=-\infty}^{\infty} a_j x_{t-j} = x_t - x_{t-1} = \nabla x_t.$$

This shows that the differencing operator may be understood from the viewpoint of linear filters theory.

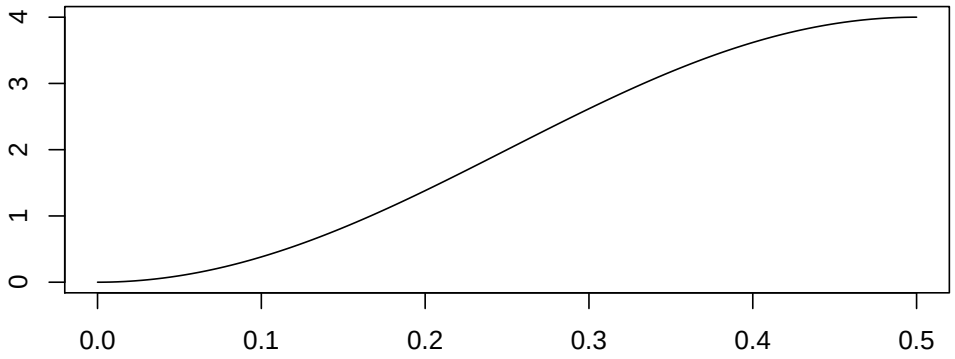
The frequency response function is

$$A_{yx}(\omega) = \sum_{j=-\infty}^{\infty} a_j e^{-2\pi i\omega j} = 1 - e^{-2\pi i\omega},$$

whose squared magnitude (*i.e.*, the power transfer function) is

$$|A_{yx}(\omega)|^2 = (1 - e^{-2\pi i\omega})(1 - e^{-2\pi i\omega}) = 2 - (e^{2\pi i\omega} + e^{-2\pi i\omega}) = 2\{1 - \cos(2\pi\omega)\}.$$

A plot of the power transfer function on $0 \leq \omega \leq 1/2$ is as follows



Interpreting this through the formula $f_y(\omega) = |A_{yx}(\omega)|^2 f_x(\omega)$ provides some understanding of the impact of differencing on a time series: noting that $|A_{yx}(\omega)|^2$ takes large values when ω is near $1/2$, the differencing operation is seen to strengthen patterns of high-frequency periodic oscillations in the data; similarly, as $|A_{yx}(\omega)|^2$ takes small values when ω is near 0, it is seen that differencing dampens patterns of low-frequency periodic oscillations.

Reflecting that the differencing operator strengthens high-frequency oscillations—or, in other words, it allows high frequencies to "pass through" the filtering formula—, the linear filter defined by differencing is said to be a "high-pass" filter.

Now, suppose the impulse response function is defined according to

$$a_j = \begin{cases} 1/12 & \text{if } j = -5, -4, \ldots, 4, 5 \\ 1/24 & \text{if } j = -6, 6 \\ 0 & \text{otherwise} \end{cases}$$

The corresponding linear filter is

$$y_t = \sum_{j=-\infty}^{\infty} a_j x_{t-j} = \frac{1}{24}(x_{t-6} + x_{t+6}) + \frac{1}{12}\sum_{j=-5}^{5} x_{t-j}.$$

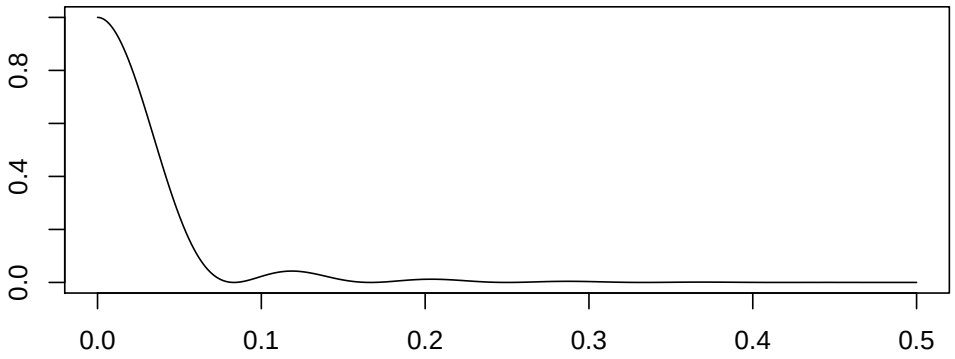
This is sometimes called a "moving-average filter," since it derives from a similar mathematical operation as moving average smoothing. This particular moving-average is derived from a modified Daniell kernel with $m = 6$.

The frequency response function is

$$\begin{aligned} A_{yx}(\omega) &= \sum_{j=-\infty}^{\infty} a_j e^{-2\pi i\omega j} \\ &= \frac{1}{12} + \frac{1}{24}(e^{12\pi i\omega} + e^{-12\pi i\omega}) + \frac{1}{12}\sum_{j=1}^{5}(e^{2\pi i\omega j} + e^{-2\pi i\omega j}) \\ &= \frac{1}{12}\left\{1 + \cos(12\pi\omega) + 2\sum_{j=1}^{5}\cos(2\pi\omega j)\right\}. \end{aligned}$$

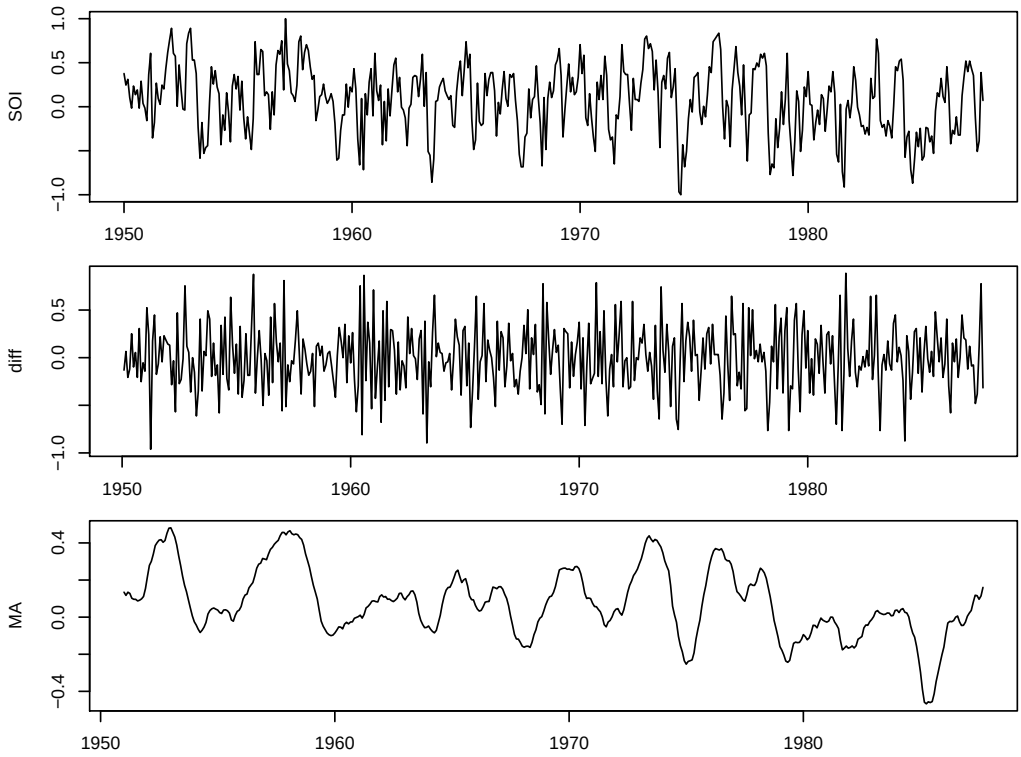
Since this evaluates to a real-valued function, the corresponding power transfer function is just its square, $|A_{yx}(\omega)|^2 = \{A_{yx}(\omega)\}^2$.

A plot of this power transfer function on $0 \le \omega \le 1/2$ is



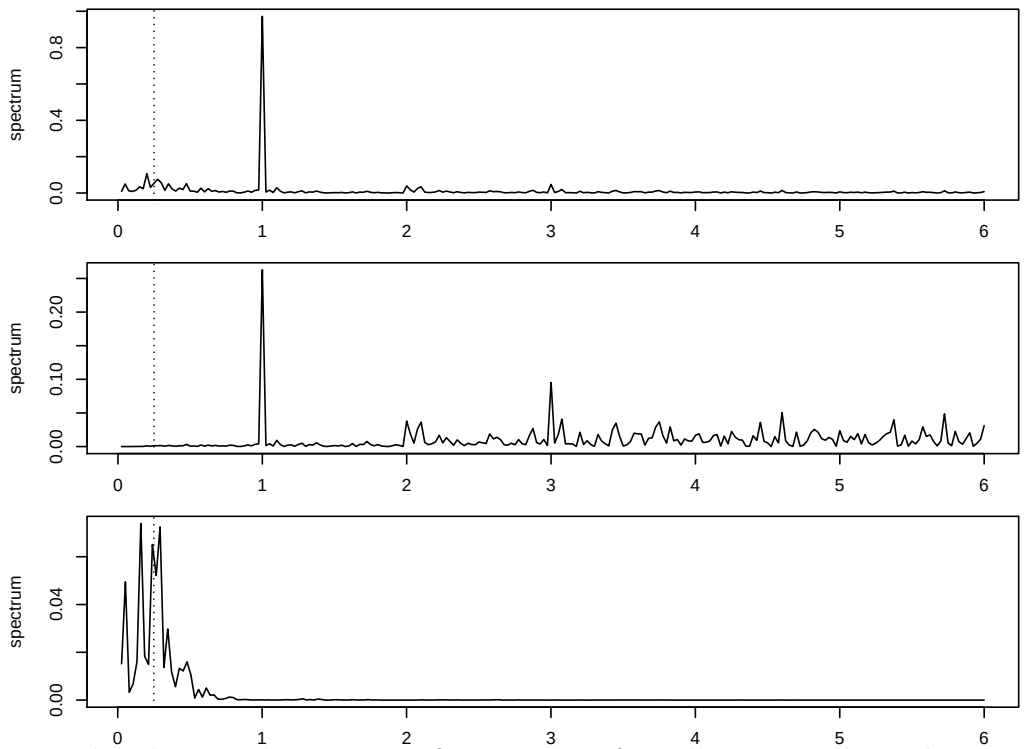
From this, the impact of this moving-average smoother is seen to strengthen periodic oscillations in the data at frequencies, ω, below about 0.08, and dampen patterns at frequencies above that value. Because this filter allows low frequencies to "pass through," it is said to be a "low-pass" filter.

The following figure illustrates how these filters act on the Southern Oscillation Index (SOI) time series, which we have examined before. Recall that this time series consists of $n = 453$ monthly measurements of ocean surface temperature in the Pacific Ocean made during the years 1950-1987. The top panel displays the raw SOI data; below that, the middle panel displays the output of high-pass filter (*i.e.*, the differencing operation); and, the bottom panel displays the output of the low-pass filter (*i.e.*, the moving-average filter).



The output of the high-pass filter appears very rough and noisy, which is as to be expected, since that filter strengthens high-frequency patterns and dampen low-frequency patterns. In contrast, the output of the low-pass filter emphasizes the smoother patterns in the SOI data, reflecting its emphasis on low-frequency patterns. Observe how that filtered time series follows the broader sweeps of the SOI data.

In fact, the characteristics of the particular moving-average filter suggest an interpretation of these patterns as clarifying the El Niño effect. To understand this, consider the figure below which plots the periodograms of the raw SOI data (top panel), the output of the high-pass filter (middle panel), and that of the low-pass filter (bottom panel). The horizontal axis of each panel is labeled in units of 1/12, so that a yearly cycle, having frequency $\omega = 1/12$, is one unit.



The top panel indicates the primary frequencies of oscillation in these data: there is a strong yearly cycle, at frequency $\omega = 1/12 = 0.0833$ (one unit on the horizontal axis), and indications of another cycle that repeats with a frequency that ranges between $\omega = 1/84 = 0.0119$ and $\omega = 1/36 = 0.0278$ (between $1/7 = 0.1429$ and $1/3 = 0.3333$ on the horizontal axis). In other words, the period of the latter cycle ranges between three and seven years. That cycle describes the El Niño effect.

From our previous examination of the power transfer function, $|A_{yx}(\omega)|^2$, of the high-pass filter, we know that it emphasizes the high-frequency oscillations in the data. Reflecting this, we see in the output time series none of the broader sweeping shapes of the yearly cycle or El Niño cycle. At the same time, the corresponding periodogram indicates that high-frequency are exhibited more strongly than in the original series.

Our examination of the power transfer function, $|A_{yx}(\omega)|^2$, of the low-pass filter indicates that the filter emphasizes oscillations at frequencies below about 0.08, and dampens patterns above that frequency. This cutoff value is important, as it is just below $\omega = 1/12 = 0.0833$, the frequency of the yearly cycle. Thus, an interpretation of this filter is that it dampens the yearly cycle, and other noisy patterns in the data, while emphasizing periodic oscillations that are associated with the El Niño cycle. This interpretation is reflected in the periodogram of the output time series to the low-pass filter, where the spike at $\omega = 1/12 = 0.0833$ (one unit on the horizontal axis) has disappeared, and the high-density values at lower-frequencies remain. The output time series would be interpreted as a clarified depiction at the El Niño cycle. □