Chapter 5: Inference in The Frequency Domain

STAT 443: Time Series and Forecasting

Prof. Bruce Dunham

Department of Statistics

Introduction

We introduce a key tool in the harmonic analysis of a discrete time series: *the periodogram.* This plot, which is in effect a histogram, attempts to replicate the spectral density function of a series.

We see the relationship between the periodogram and the variance of the data, and also show that the periodogram is the FT of the sample acvf.

Introducing the periodogram

We learned that if $x(1), \ldots, x(N)$ is a time series (with N even) then for each $t = 1, \ldots, N$,

$$\begin{aligned} x(t) &= \sum_{p=0}^{N/2} \left(a_p \cos(\omega_p t) + b_p \sin(\omega_p t)\right) \qquad (1) \\ &= a_0 + \sum_{p=1}^{N/2-1} \left(a_p \cos(\omega_p t) + b_p \sin(\omega_p t)\right) \\ &\quad + a_{N/2} \cos(\pi t) \end{aligned}$$

in which $\omega_p = 2\pi p/N$ and

$$
\begin{aligned}
a_0 &= \frac{1}{N}\sum_{t=0}^{N} x(t) = \bar{x}, \\
a_{N/2} &= \frac{1}{N}\sum_{t=1}^{N} (-1)^t x(t), \\
a_p &= \frac{2}{N}\sum_{t=1}^{N} x(t)\cos(\omega_p t), \\
b_p &= \frac{2}{N}\sum_{t=1}^{N} x(t)\sin(\omega_p t),
\end{aligned}
$$

for $p = 1, \ldots, N/2 - 1$.

Recall that

- There is no error term in (1) – it fits the data exactly.

- The Fourier coefficients, i.e., the a_p's and b_p's, are the least squares estimates found earlier.

This gives the so–called *harmonic analysis*, the variation in the data being split into components at each Fourier frequency, the component at frequency ω_p being the pth *harmonic.*

If we let

$$\tan\left(\phi_p\right) = -\frac{b_p}{a_p}$$

then

$$\begin{aligned}\cos\left(\phi_p\right) &= \frac{a_p}{\left(a_p^2 + b_p^2\right)^{\frac{1}{2}}}, \\ \sin\left(\phi_p\right) &= -\frac{b_p}{\left(a_p^2 + b_p^2\right)^{\frac{1}{2}}}\end{aligned}$$

and so

$$a_p \cos\left(\omega_p t\right) + b_p \sin\left(\omega_p t\right)$$

is

$$\left(a_p^2 + b_p^2\right)^{\frac{1}{2}} \left(\cos\left(\omega_p t\right)\cos\left(\phi_p\right) - \sin\left(\omega_p t\right)\sin\left(\phi_p\right)\right)$$

which can written $R_p \cos\left(\omega_p t + \phi_p\right)$ where

$$R_p = \left(a_p^2 + b_p^2\right)^{\frac{1}{2}}$$

is the amplitude of the pth harmonic, and

$$\phi_p = \tan^{-1}\left(-\frac{b_p}{a_p}\right)$$

is the phase of the harmonic. That is, the pth harmonic ($p \neq N/2$) is

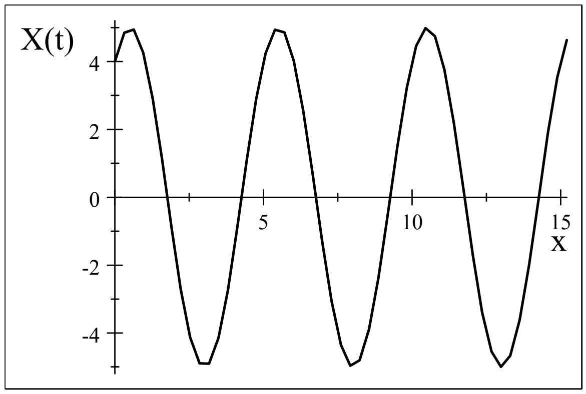
$$a_p \cos\left(\omega_p t\right) + b_p \sin\left(\omega_p t\right) = R_p \cos\left(\omega_p t + \phi_p\right)$$

Example 0.1 *The model*

$$X\left(t\right) = 4\cos\left(\frac{2\pi t}{5}\right) + 3\sin\left(\frac{2\pi t}{5}\right),$$

can be written in equivalent form as

$$X\left(t\right) = 5\cos\left(\frac{2\pi t}{5} + \tan^{-1}\left(-\frac{3}{4}\right)\right).$$



$X\left(t\right) = 4\cos\left(\omega t\right) + 3\sin\left(\omega t\right)$ *when* $\omega = 2\pi/5$.

Decomposing the sum of squares

Taking

$$x\left(t\right) = \sum_{p=0}^{N/2} \left(a_p \cos\left(\omega_p t\right) + b_p \sin\left(\omega_p t\right)\right)$$

squaring both sides and summing over $t = 1, \ldots, N$, we find (exercise)

$$\begin{aligned}\sum_{t=1}^{N} x(t)^2 - N\bar{x}^2 &= \sum_{t=1}^{N} (x(t) - \bar{x})^2 \\ &= \frac{N}{2} \sum_{p=1}^{N/2-1} R_p^2 + N a_{N/2}^2.\end{aligned}$$

Hence the variance of the data series can be written

$$\frac{1}{N} \sum_{t=1}^{N} (x(t) - \bar{x})^2 = \frac{1}{2} \sum_{p=1}^{N/2-1} R_p^2 + a_{N/2}^2.$$

This is known as *Parseval's theorem.* It shows that the contribution to the variance in the data from the pth harmonic is $R_p^2/2$ – and we met this in the simple sinusoidal model.

A plot of $R_p^2/2$ against ω_p would give a *line spectrum*, and would be appropriate for series with fixed components at Fourier frequencies.

However, most real series do not obey that property, and will most likely have **continuous** spectra. So we consider $R_p^2/2$ to be contribution to the variance from frequencies in the range

$$\omega_p \pm \frac{\pi}{N},$$

and plot a histogram, where

$$\begin{aligned}\frac{R_p^2}{2} &= \text{ area of } p\text{th histogram bar} \\ &= \text{height } \times \frac{2\pi}{N}.\end{aligned}$$

Hence the height of the histogram across $\omega_p \pm \frac{\pi}{N}$ is

$$I\left(\omega_p\right) := \frac{N R_p^2}{4\pi}.$$

This defines the *periodogram.* When $p = N/2$, the bar width is obviously halved, and

$$I\left(\pi\right) = \frac{N a_{N/2}^2}{\pi}.$$

Note that

- The total area under the periodogram equals the variance of the data.

- Some authors differ in notation, either by re-scaling or allowing negative frequencies.

- While apparently a natural estimator of the spectrum of the data–generating process, the periodogram needs some modification for this task.

Relationship between periodogram and acvf

Recall that the sample autocovariance function at lag k is

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} \left(x\left(t\right) - \bar{x}\right)\left(x\left(t+k\right) - \bar{x}\right),$$

and $c_{-k} = c_k$.

In terms of the data, for $p \neq N/2$,

$$\begin{aligned}
I(\omega_p) &= \frac{1}{N\pi}\left(\left(\sum_{t=1}^{N} x(t)\cos(\omega_p t)\right)^2 \right.\\
&\qquad \left. + \left(\sum_{t=1}^{N} x(t)\sin(\omega_p t)\right)^2\right)\\
&= \frac{1}{N\pi}\left(\left(\sum_{t=1}^{N} (x(t)-\bar{x})\cos(\omega_p t)\right)^2 \right.\\
&\qquad \left. + \left(\sum_{t=1}^{N} (x(t)-\bar{x})\sin(\omega_p t)\right)^2\right)\\
&= \frac{1}{N\pi}\sum_{t=1}^{N}(x(t)-\bar{x})^2\left(\cos^2(\omega_p t)+\sin^2(\omega_p t)\right)\\
&\quad +\frac{1}{N\pi}\sum\sum_{s\neq t}(x(t)-\bar{x})(x(s)-\bar{x})\\
&\quad \times(\cos(\omega_p t)\cos(\omega_p s)+\sin(\omega_p t)\sin(\omega_p s)).
\end{aligned}$$

Now

$$\cos^2(\omega_p t)+\sin^2(\omega_p t)=1$$

and

$$\cos\left(\omega_p t\right)\cos\left(\omega_p\left(t-k\right)\right)+\sin\left(\omega_p t\right)\sin\left(\omega_p\left(t-k\right)\right)$$

is

$$\cos\left(\omega_p\left(t-k-t\right)\right)=\cos\left(\omega_p k\right).$$

Consequently, $I\left(\omega_p\right)$ is

$$\frac{1}{N\pi}\sum_{t=1}^{N}\left(x\left(t\right)-\bar{x}\right)^2+\frac{1}{N\pi}\sum\sum_{s\neq t}\left(x\left(t\right)-\bar{x}\right)\left(x\left(s\right)-\bar{x}\right)$$
$$\times\left(\cos\left(\omega_p t\right)\cos\left(\omega_p s\right)+\sin\left(\omega_p t\right)\sin\left(\omega_p s\right)\right)$$

and this is

$$\frac{1}{N\pi}\sum_{t=1}^{N}\left(x\left(t\right)-\bar{x}\right)^2$$
$$+\frac{2}{N\pi}\sum_{k=1}^{N-1}\sum_{t=k+1}^{N}\left(x\left(t\right)-\bar{x}\right)\left(x\left(t-k\right)-\bar{x}\right)$$
$$\times\left(\cos\left(\omega_p t\right)\cos\left(\omega_p\left(t-k\right)\right)+\sin\left(\omega_p t\right)\sin\left(\omega_p\left(t-k\right)\right)\right)$$

which equals

$$\frac{1}{\pi}c_0+\frac{2}{N\pi}\sum_{k=1}^{N-1}\sum_{t=1}^{N-k}\left(x\left(t\right)-\bar{x}\right)\left(x\left(t+k\right)-\bar{x}\right)\cos\left(\omega_p k\right).$$

Recalling the definition of c_k, we have shown that

$$I\left(\omega_p\right) = \frac{1}{\pi}\left(c_0 + 2\sum_{k=1}^{N-1} c_k \cos\left(\omega_p k\right)\right).$$

So the periodogram is the (discrete) FT of the acvf, assuming $c_k = 0$ for $|k| \geq N$.

The periodogram and the spectrum

Having defined the periodogram of a time series, we focus now of some properties that it possesses. We have already shown that the periodogram is the FT of the sample autocovariance function of the series.

Here it is shown that the periodogram is not a consistent estimator of the spectrum of the underlying process.

Recall that the periodogram of the series $x(1), \ldots, x(N)$ across $\omega_p \pm \frac{\pi}{N}$ is

$$I(\omega_p) = \frac{NR_p^2}{4\pi}.$$

for $p = 1, \ldots, N/2 - 1$, where $R_p = \left(a_p^2 + b_p^2\right)^{\frac{1}{2}}$, and when $p = N/2$,

$$I(\pi) = \frac{Na_{N/2}^2}{\pi}.$$

Moreover we showed that

$$I(\omega_p) = \frac{1}{\pi}\left(c_0 + 2\sum_{k=1}^{N-1} c_k \cos(\omega_p k)\right),$$

so the periodogram is the (discrete) FT of the acvf c_k.

Hence $I(\omega)$ seems a natural estimator of the spectrum of the series,

$$f(\omega) = \frac{1}{\pi}\left(\gamma(0) + 2\sum_{k=1}^{\infty} \gamma(k)\cos(\omega_p k)\right),$$

as c_k is a reasonable estimator of $\gamma(k)$. Indeed, for each $w \in (0, \pi)$,

$$E(I(\omega)) \to f(\omega)$$

as $N \to \infty$, so $I(\omega)$ is asymptotically unbiased. However, we will see that $I(\omega)$ is not a *consistent* estimator of $f(\omega)$, in that its variance does not decrease as N increases.

Consider the case where $x(t) \sim N\left(0, \sigma^2\right)$, independently for each t. As both a_p and b_p are linear combinations of Normals for each p, these coefficients must themselves be Normally distributed. Now, assuming N is even and $p \neq N/2$,

$$\begin{aligned} E(a_p) &= E\left(\frac{2}{N}\sum_{t=1}^{N} x(t)\cos(\omega_p t)\right) \\ &= \frac{2}{N}\sum_{t=1}^{N} E(x(t))\cos(\omega_p t) \\ &= 0 \end{aligned}$$

and similarly $E(b_p) = 0$. Further,

$$\begin{aligned}
\mathrm{Var}(a_p) &= \frac{4}{N^2}\mathrm{Var}\left(\sum_{t=1}^{N} x(t)\cos(\omega_p t)\right) \\
&= \frac{4\sigma^2}{N^2}\sum_{t=1}^{N}\cos^2(\omega_p t) \\
&= \frac{2\sigma^2}{N}.
\end{aligned}$$

Similarly

$$\mathrm{Var}(b_p) = \frac{2\sigma^2}{N}.$$

Further, since the $x(t)$ are independent,

$$\begin{aligned}
\mathrm{Cov}(a_p, b_p) &= \frac{4}{N^2}\mathrm{Cov}\left(\sum_{t=1}^{N} x(t)\cos(\omega_p t),\right. \\
&\qquad \left.\sum_{t=1}^{N} x(t)\sin(\omega_p t)\right) \\
&= \frac{4\sigma^2}{N^2}\sum_{t=1}^{N}\cos(\omega_p t)\sin(\omega_p t) \\
&= 0.
\end{aligned}$$

Recalling that sums of squares of n independent standard Normals follow the χ_n^2 distribution, we see that

$$\frac{N\left(a_p^2+b_p^2\right)}{2\sigma^2}=\frac{2\pi I\left(\omega_p\right)}{\sigma^2}\sim\chi_2^2.$$

Since the expectation and variance of χ_n^2 are n and $2n$ respectively, we have

$$\begin{aligned}\mathrm{Var}\left(I\left(\omega_p\right)\right) &= \frac{\sigma^4}{4\pi^2}\mathrm{Var}\left(\chi_2^2\right)\\ &= \frac{\sigma^4}{\pi^2},\end{aligned}$$

a constant. So in this special case, and in general, $\mathrm{Var}(I\left(\omega_p\right)) \nrightarrow 0$ as $N\rightarrow\infty$, so the periodogram is not a consistent estimator of the spectrum.

Why is this?

- As N increases, the data are being asked to find proportionately more *independent* values of $I\left(\omega\right)$, so that the precision for each is unchanged.

- Recall that $I(\omega_p)$ is a regression SS on two d.o.f., regardless of N.

The following relates the full story.

Theorem 0.1 *If $X(t)$ is a stationary stochastic process where each finite realization is from a multivariate Normal distribution, let $x(1), \ldots, x(N)$ be a realization from $X(t)$ with periodogram $I(\omega)$. Then as $N \to \infty$, for $0 < p < N/2$,*

(a)

$$\frac{2I(\omega_p)}{f(\omega_p)} \sim \chi_2^2$$

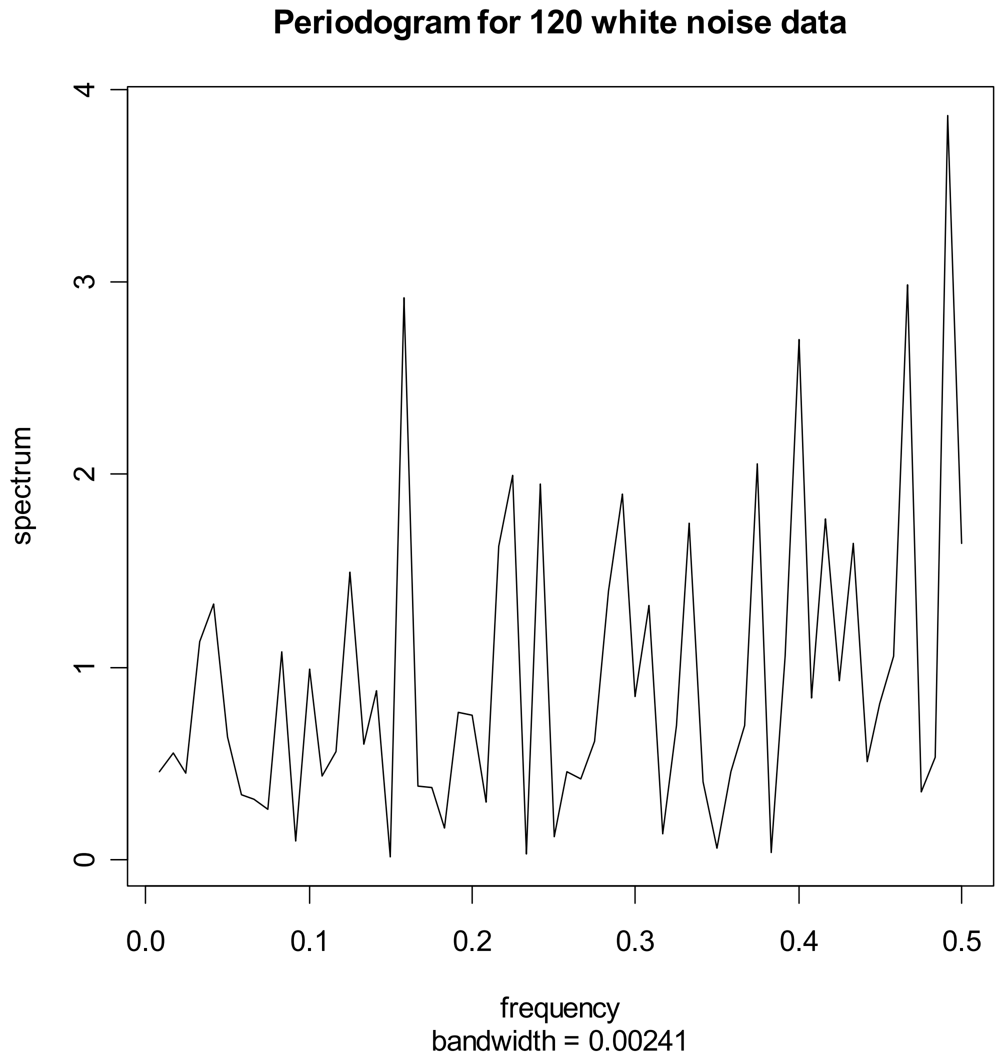
(b) *$I(\omega_p)$ and $I(\omega_q)$ are independent for all $p \neq q$.*

When $\omega = \pi$, (a) becomes

$$\frac{I(\pi)}{f(\pi)} \sim \chi_1^2,$$

recalling $b_{N/2} = 0$.

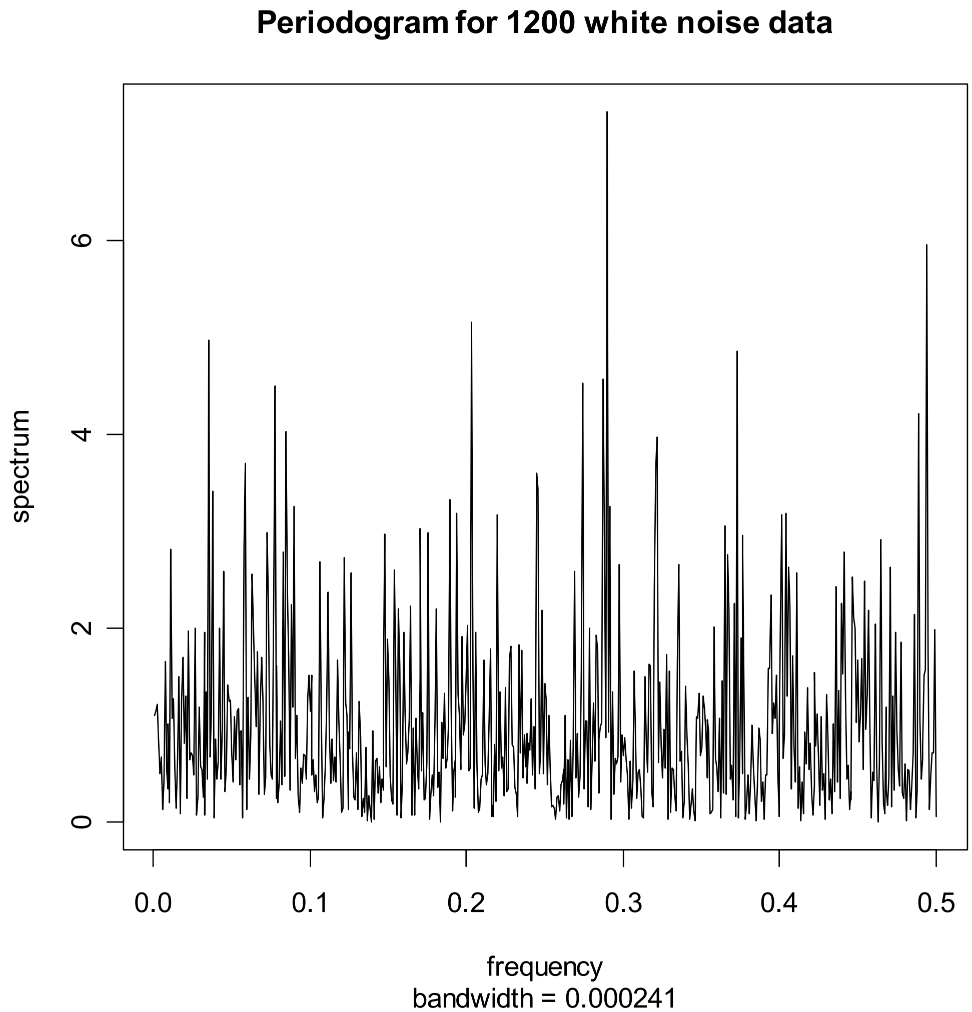
Example 0.2 *The vector* x *contains 120 observations simulated from* $N(0,1)$. *Its periodogram is plotted* *below:*



(R defines frequency as cycles per unit time, $f = \omega/2\pi$, *and defines the periodogram as* $\pi I(\omega)$*.) Of course, the spectrum here is actually flat. Note also that* $\chi^2_2(0.95) = 6.0$, *a "significant" value of* $2\pi I(\omega)$

could be taken as one over 6 since here the true spectrum is $f(\omega) = 1/\pi$. In this case since 5% of 60 is 3, we would expect three values of $\pi I(\omega)$ to exceed 3. In fact in the above two values of $\pi I(\omega)$ are above 3.

Example 0.3 *In the plot below y is 1200 observations* *from* $N(0,1)$.



Of course there are more ordinates at which $I(\omega)$ *has been found, but the result is no "smoother" or flatter than when* $N = 120$.

Modifying the periodogram

Although the periodogram appears to be a natural estimator of the spectrum of a process, we have seen that it is not consistent. We describe methods by which the periodogram can be modified to provide a consistent estimator of $f(\omega)$.This in turn leads to approximate confidence intervals for the spectrum.

Truncating and transforming the periodogram

With

$$I(\omega_p) = \frac{1}{\pi}\left(c_0 + 2\sum_{k=1}^{N-1} c_k \cos(\omega_p k)\right),$$

for a stationary series the autocovariances $\{c_k\}$ are likely to decrease with k. So it seems sensible to attach weights $\{\lambda_k\}$ to the c_k in an estimate of $f(\omega)$ which decrease with k, i.e.,

$$\hat{f}(\omega) = \frac{1}{\pi}\left(\lambda_0 c_0 + 2\sum_{k=1}^{M} \lambda_k c_k \cos(\omega k)\right).$$

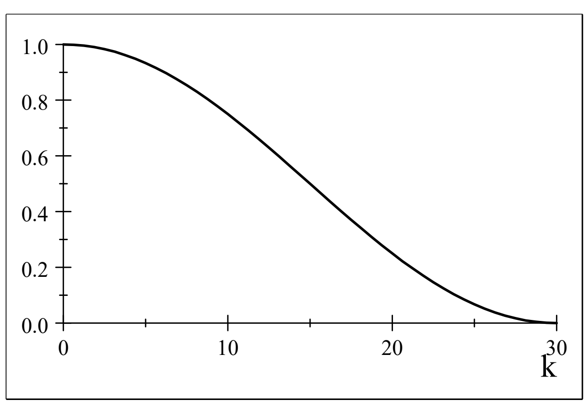
The weights $\{\lambda_k\}$ are called the *lag window*, and $M < N$ is the *truncation point.*

Examples include

Tukey window:

$$\lambda_k = \frac{1}{2}\left(1 + \cos\left(\frac{\pi k}{M}\right)\right)$$

for $k = 0, 1 \ldots, M$.



Tukey window when $M = 30$.

Parzen window:

$$\lambda_k = \begin{cases} 1 - 6\left(\frac{k}{M}\right)^2 + 6\left(\frac{k}{M}\right)^3 & 0 < k \leq M/2 \\ \\ 2\left(1 - \frac{k}{M}\right)^3 & \frac{M}{2} \leq k \leq M \end{cases}$$

This has a similar shape to the Tukey window.

Bartlett window:

$$\lambda_k = 1 - \frac{k}{M},$$

for $k = 0, \ldots, M$.

The Parzen and Tukey windows give similar results, and are superior to Bartlett's window.

The choice of M must balance "resolution" against "variance" – the lower M, the lower the variance of $\hat{f}$, but the larger the bias.

Several values of M can be tried: too small and $\hat{f}$ is over–smoothed, losing essential features; too large and there is insufficient smoothing. One suggestion is to take $M = 2\,(N)^{\frac{1}{2}}$.

Hanning and Hamming

Hanning is actually equivalent to a Tukey window, other than computationally, with lag window values taken as unity, $\lambda_k = 1$, for $k = 0, \ldots, M$. The result $\hat{f}_1$ is smoothed using the weights $\left\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right\}$. This gives

$$\hat{f} = \frac{1}{4}\hat{f}_1\left(\omega - \frac{\pi}{M}\right) + \frac{1}{2}\hat{f}_1(\omega) + \frac{1}{4}\hat{f}_1\left(\omega + \frac{\pi}{M}\right)$$

for $\omega = \pi j/M$, for $j = 1, \ldots, M-1$. Weights $\left\{\frac{1}{2}, \frac{1}{2}\right\}$ are used at the boundaries.

Hamming is similar in name and nature to Hanning, but uses weights $\{0.23, 0.54, 0.23\}$ instead of $\left\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right\}$, and weights $\{0.54, 0.46\}$ at $\omega = 0, \pi$.

Smoothing

An alternative to transforming the truncated periodogram involves smoothing $I(\omega)$, averaging sets of consecutive values. This is computationally harder, but viable now with modern computers. The idea is due to P.J. Daniell.

Let

$$\hat{f}(\omega) = \frac{1}{m} \sum_{j} I\left(\omega_j\right) \tag{2}$$

where $\omega_j = 2\pi j/N$ and j ranges over m consecutive integers so that $\left\{\omega_j\right\}$ are symmetric about ω. At the boundaries $\omega = 0, \pi$, we let

$$\hat{f}(0) = \frac{2}{m} \sum_{j=1}^{(m-1)/2} I\left(\omega_j\right)$$

and

$$\hat{f}(\pi) = \frac{1}{m} \left(I(\pi) + 2 \sum_{j=1}^{(m-1)/2} I\left(\pi - \omega_j\right) \right)$$

assuming m is odd and $I(0) = 0$.

Note that:

- Since consecutive values in the periodogram are (asymptotically) uncorrelated, the variance of (2) is $O\left(\frac{1}{m}\right)$. In fact (2) is (typically slightly) biased, since
$$E\left(\hat{f}(\omega)\right) \approx \frac{1}{m}\sum_{j} f\left(\omega_j\right)$$
but equality only holds when the spectrum is linear over the values of ω in the above sum.

- The choice of m is similar to the choice of M in windowing, only the effect is the opposite way round (so m too large implies over–smoothed, too small gives insufficient smoothing). Chatfield suggests $2(N)^{\frac{1}{2}}$.

- Smoothing can be applied more than once if required – typically smoothing twice with two different choices for m works best.

Confidence intervals

We have described how to produce a consistent point estimate of $f(\omega)$ by either truncating or smoothing the "raw" periodogram.

If using a lag window, i.e., an estimate

$$\hat{f}(\omega) = \frac{1}{\pi} \sum_{k=-M}^{M} \lambda_k c_k \cos(\omega k)$$

then it can be shown that asymptotically

$$\frac{n\hat{f}(\omega)}{f(\omega)} \sim \chi_n^2,$$

approximately, where

$$n := \frac{2N}{\sum_{k=-M}^{M} \lambda_k^2}$$

gives the degrees of freedom.

Hence with $\chi_n^2(\alpha)$ the $100\alpha\%$ of χ_n^2, we have

$$P\left(\chi_n^2(\alpha/2) < \frac{n\hat{f}(\omega)}{f(\omega)} < \chi_n^2(1-\alpha/2)\right) = 1-\alpha,$$

consequently a $100(1-\alpha)\,\%$ confidence interval for $f(\omega)$ is

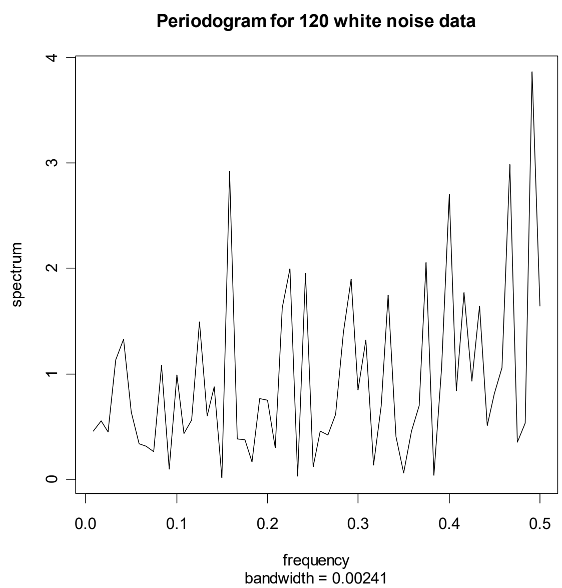
$$\left(\frac{n\hat{f}(\omega)}{\chi_n^2(1-\alpha/2)}, \frac{n\hat{f}(\omega)}{\chi_n^2(\alpha/2)}\right). \qquad (3)$$

Alternatively, if using an estimate of the form

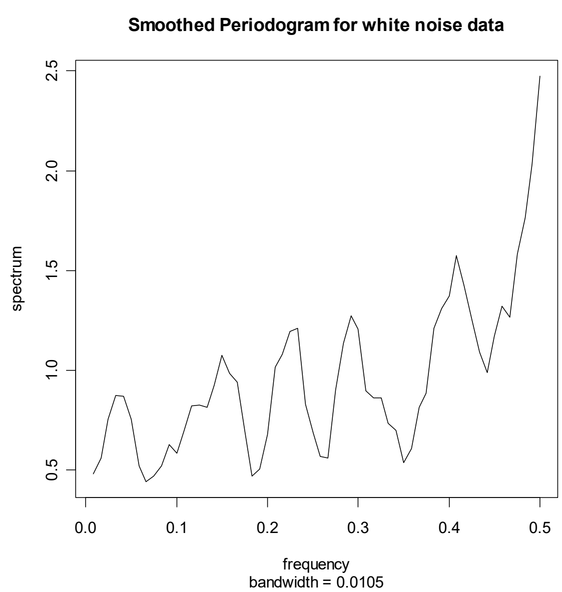
$$\hat{f}(\omega) = \frac{1}{m}\sum_j I\left(\omega_j\right)$$

this is simply averaging m values of the periodogram, i.e., an average of m (asymptotically) independent χ_2^2 variables. So $\hat{f}(\omega)$ will have $n = 2m$ degrees of freedom in (3).

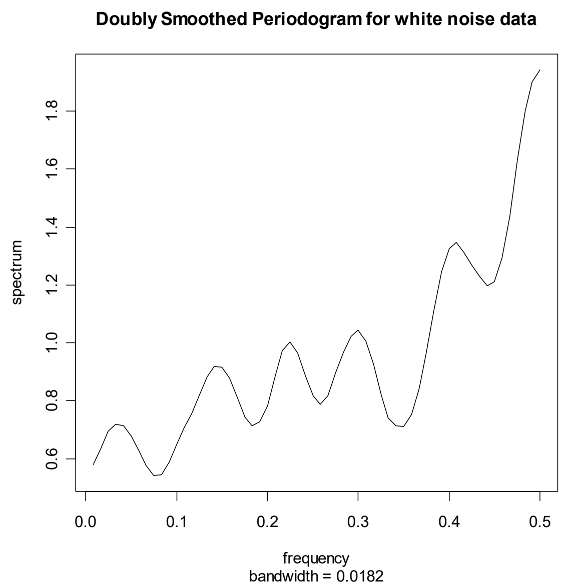
Example 0.4 *The vector* $\mathbf{x}$ *from above contains 120 observations simulated from* $N(0,1)$ *with periodogram* *below:*



A smoothed version of this, with $m = 5$, *is below:*



Smoothing again, taking $m = 7$ *this time, gives the* *plot below:*



Doubly Smoothed Periodogram for white noise data
spectrum
0.6
0.8
1.0
1.2
1.4
1.6
1.8
0.0
0.1
0.2
0.3
0.4
0.5
frequency
bandwidth = 0.0182

Comparing Estimates of the Spectrum

Since the raw periodogram is inconsistent in estimating the spectrum of a process, we met two methods of modifying the periodogram:

1. Truncating and transforming using a lag window, and

2. Smoothing.

Comparing the methods

The periodogram

$$I(\omega) = \frac{1}{\pi}\left(c_0 + 2\sum_{k=1}^{N-1} c_k \cos(\omega k)\right)$$

is an inconsistent estimator of the spectrum $f(\omega)$.

Recall that truncating and transforming provides an estimator of the form

$$\hat{f}_T(\omega) := \frac{1}{\pi}\left(\lambda_0 c_0 + 2\sum_{k=1}^{M} \lambda_k c_k \cos(\omega k)\right).$$

with $\{\lambda_k : k = 0, \ldots, M\}$ the lag window, and $M < N$ is the truncation point, whereas smoothing gives an estimator

$$\hat{f}_S(\omega) := \frac{1}{m}\sum_j I\left(\omega_j\right)$$

for some (usually odd) m.

Now it can be shown that $\hat{f}_S$ is a special case of an estimator of the form of $\hat{f}_T$, so both can be defined by some lag window $\{\lambda_k\}$. Letting $\lambda_{-k} = \lambda_k$ for all $k \in \{-M, \ldots, -1, 0, 1, \ldots, M\}$ we define the *spectral window* to be the FT of $\{\lambda_k\}$,

$$K(\omega) := \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \lambda_k e^{-\mathbf{i}k\omega}$$

for $\omega \in (-\pi, \pi)$.

Remark 0.1 *Negative frequencies cannot sensibly be avoided here, though $K(-\omega) = K(\omega)$ given the lag window is symmetric about 0.*

The inverse FT is

$$\lambda_k = \int_{-\pi}^{\pi} K(\omega)\, e^{\mathbf{i}\omega k} d\omega.$$

Now as both $\hat{f}_T$ and $\hat{f}_S$ are based on a lag window (see Exercises 4), we can write either at ω' as

$$\begin{aligned}
\hat{f}\left(\omega'\right) &= \frac{1}{\pi} \sum_{k=-(N-1)}^{N-1} \lambda_k c_k e^{-\mathbf{i}\omega' k} \\
&= \frac{1}{\pi} \sum_{k=-(N-1)}^{N-1} \left(\int_{-\pi}^{\pi} K(\omega)\, e^{\mathbf{i}\omega k} d\omega \right) c_k e^{-\mathbf{i}\omega' k} \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} K(\omega) \sum_{k=-(N-1)}^{N-1} c_k \exp\left(\mathbf{i}\left(\omega - \omega'\right) k\right) d\omega \\
&= \int_{-\pi}^{\pi} K(\omega)\, I\left(\omega' - \omega\right) d\omega.
\end{aligned}$$

Hence all estimates smooth $I(\omega)$ using some *smoothing kernel* $K(\omega)$.

Sensibly we would have

$$\lambda_0 = 1 = \int_{-\pi}^{\pi} K(\omega)\, d\omega.$$

Further, taking expectations of $\hat{f}(\omega')$ we have

$$\lim_{N\to\infty} E\left(\hat{f}\left(\omega'\right)\right) = \int_{-\pi}^{\pi} K(\omega)\, f\left(\omega' - \omega\right) d\omega$$

so $K(\omega)$ expresses the contribution to the spectrum f at each frequency to the mean of $\hat{f}$.

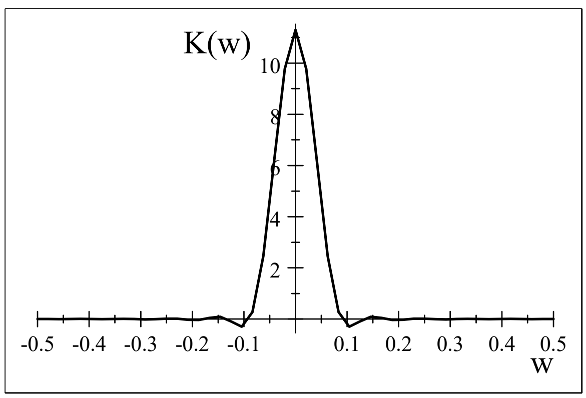
We consider three cases, each with $N = 800$.

Example 0.5 *For the Tukey window with* $M = 71$,

$$\lambda_k = \frac{1}{2}\left(1 + \cos\left(\frac{\pi k}{71}\right)\right)$$

and

$$K(\omega) = \frac{1}{4\pi} \sum_{k=-71}^{71} \left(1 + \cos\left(\frac{\pi k}{71}\right)\right) e^{-\mathbf{i}k\omega}$$
$$= \frac{1}{4\pi} \sum_{k=-71}^{71} \left(1 + \cos\left(\frac{\pi k}{71}\right)\right) \cos(k\omega).$$



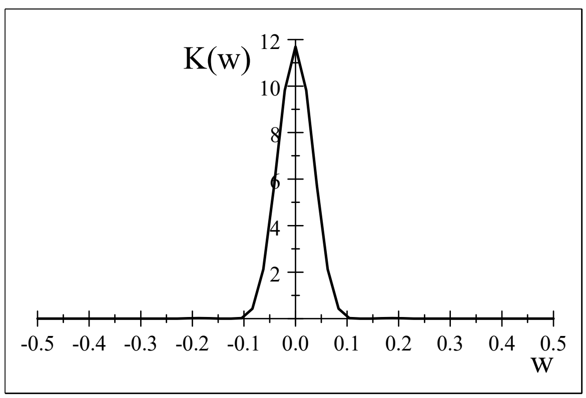
Spectral window for Tukey window with $M = 71$.

Example 0.6 *The Parzen window with $M = 98$ has*

$$\lambda_k = \begin{cases} 1 - 6\left(\frac{k}{98}\right)^2 + 6\left(\frac{k}{98}\right)^3 & 0 < k \leq 49 \\ 2\left(1 - \frac{k}{98}\right)^3 & 49 \leq k \leq 98 \end{cases}$$

and with $\lambda_{-k} = \lambda_k$ *we have*

$$\begin{aligned} K(\omega) &= \frac{1}{2\pi}\sum_{k=-98}^{98} \lambda_k e^{-\mathbf{i}k\omega} \\ &= \frac{1}{2\pi} + \frac{2}{\pi}\sum_{k=1}^{49}\left(1 - 6\left(\frac{k}{98}\right)^2 + 6\left(\frac{k}{98}\right)^3\right)\cos(\omega k) \\ &\quad + \frac{4}{\pi}\sum_{k=50}^{98}\left(1-\frac{k}{98}\right)^3\cos(\omega k)\,. \end{aligned}$$



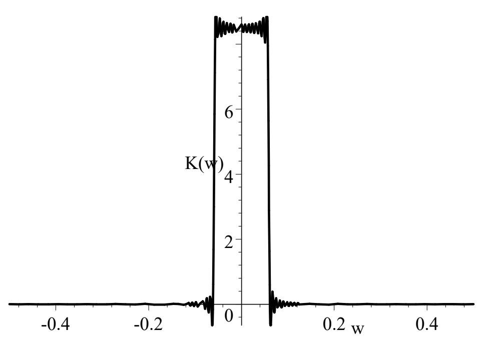
Spectral window for Parzen window with $M = 98$.

Example 0.7 *When smoothing the periodogram here with* $m = 15$, *it can be shown that this takes* $\lambda_0 = 1$ *and*

$$\lambda_k = \frac{\sin(15\pi k/800)}{15\pi k/800}$$

for $k = 1, 2, \ldots, 799$. *So*

$$\begin{aligned} K(\omega) &= \frac{1}{2\pi} \sum_{k=-799}^{799} \lambda_k e^{-\mathbf{i}k\omega} \\ &= \frac{1}{2\pi} + \frac{2}{\pi} \sum_{k=1}^{799} \frac{\sin\left(\frac{3\pi k}{160}\right) \cos(\omega k)}{\frac{3\pi k}{160}}. \end{aligned}$$



Spectral window for smoothed periodogram with $m = 15$.

This is effectively flat over the range $\pm 15\pi/800$.

Variances of the estimates

In the examples earlier, M and m were taken to make the variances of the estimates approximately equal. We have said that for the smoothed estimate $\hat{f}_S$ with $2m = n$,

$$\frac{n\hat{f}_S(\omega)}{f(\omega)} \sim \chi^2_n$$

approximately, and for the "truncated" estimates

$$\frac{n\hat{f}_T(\omega)}{f(\omega)} \sim \chi^2_n,$$

approximately, where

$$n := \frac{2N}{\sum_{k=-M}^{M} \lambda_k^2}.$$

This is approximately $3.71N/M$ for the Parzen window, and $8N/3M$ for the Tukey.

As $\mathrm{Var}\left(\chi^2_n\right) = 2n$, we find

$$\mathrm{Var}\left(\frac{n\hat{f}(\omega)}{f(\omega)}\right) = n^2 \mathrm{Var}\left(\frac{\hat{f}(\omega)}{f(\omega)}\right)$$

and approximate variances are

	$\hat{f}_S$	Parzen	Tukey
Var $\left(\frac{\hat{f}(\omega)}{f(\omega)}\right)$	$\frac{1}{m}$	$\frac{2M}{3.17N}$	$\frac{3M}{4N}$

Bandwidth

Bandwidth is the "width" of the spectral window. For the smoothed periodogram, this is approximately $2\pi m/N$. For the Bartlett, Parzen, and Tukey windows

$$\text{Bandwidth} \propto \frac{1}{M}$$

and so as M grows the bias diminishes but the variance increases, whereas for smoothing

$$\text{Bandwidth} \propto m.$$

Discussion and overview of harmonic analysis

The aim of analysis in the frequency domain is to detect cyclical variation in a series, typically at non-deterministic frequencies. The main tool is the periodogram, which once modified estimates the spectral density function of the process.

While probably less intuitive than the time domain approach, the methods are *non-parametric*, so no models have been assumed. The approach may be the natural one where underlying cyclical patterns govern the variation in the series, as might be expected in data from biostatistics, geological and electrical engineering applications.

Remark 0.2 *Parametric harmonic analysis exists, whereby the spectrum of a fitted ARMA model is estimated via the periodogram.*

Some disadvantages of the approach include:

- Long series are often needed, typically $N > 100$.

- *Harmonics* may appear: an important frequency component at ω' may also appear as secondary (spurious) spikes in the periodogram at frequencies $2\omega', 3\omega', \ldots$, because the variation is not exactly sinusoidal.

- *Leakage* occurs when the underlying process has cyclic components at non-Fourier frequencies, the "power" being "leaked" in the periodogram to neighbouring Fourier frequencies. Only really an issue with short series.

- *Aliasing* is common – frequencies outside $(0, \pi)$ will be unidentifiable.

When inspecting a periodogram, what should we be looking for?

1. Are there noticeable "spikes" in the spectrum?

2. If so, at what frequencies?

3. Is there a physical interpretation for these frequencies?

4. Does the low–frequency behaviour indicate non–stationarity?

In the last situation above perhaps *pre–whitening* – removing the mean, trend and any obvious seasonal effects from the data – was not performed. The aim is to examine a series "close to" being white noise. Otherwise nonstationary effects can dominate the periodogram.

Variances of the estimates

In the examples earlier, M and m were taken to make the variances of the estimates approximately equal. We have said that for the smoothed estimate $\hat{f}_S$ with $2m = n$,

$$\frac{n\hat{f}_S(\omega)}{f(\omega)} \sim \chi_n^2$$

approximately, and for the "truncated" estimates

$$\frac{n\hat{f}_T(\omega)}{f(\omega)} \sim \chi_n^2,$$

approximately, where

$$n := \frac{2N}{\sum_{k=-M}^{M} \lambda_k^2}.$$

This is approximately $3.71N/M$ for the Parzen window, and $8N/3M$ for the Tukey.

As $\mathrm{Var}\left(\chi_n^2\right) = 2n$, we find

$$\mathrm{Var}\left(\frac{n\hat{f}(\omega)}{f(\omega)}\right) = n^2 \mathrm{Var}\left(\frac{\hat{f}(\omega)}{f(\omega)}\right)$$

and approximate variances are

	$\hat{f}_S$	Parzen	Tukey
Var $\left(\frac{\hat{f}(\omega)}{f(\omega)}\right)$	$\frac{1}{m}$	$\frac{2M}{3.17N}$	$\frac{3M}{4N}$

Bandwidth

Bandwidth is the "width" of the spectral window. For the smoothed periodogram, this is approximately $2\pi m/N$. For the Bartlett, Parzen and Tukey windows

$$\text{Bandwidth} \propto \frac{1}{M}$$

and so as M grows the bias diminishes but the variance increases, whereas for smoothing

$$\text{Bandwidth} \propto m.$$

Spectral analysis in R

Various commands exist for periodogram analysis in R. In some cases the output/default options are arguably rather non-standard.

A class of objects `spec` exists, which can be created via

```
> spec.pgram(x, spans = NULL, kernel,
taper = 0.1, pad = 0, fast = TRUE,
demean = FALSE, detrend = TRUE, plot = TRUE, ...)
```

in which

- `x` is a time series object,
- `spans` defines the smoothing – typically a single odd integer, or `c(m,n)` if smoothing twice,

- `taper` specifies a constant for a smoothing method – see below,

- `pad` is proportion of data to be padded for FFT (0 value is overridden if `fast=T`)

- `fast` is a logical operator for choice of whether to use FFT,

- `detrend` removes trend at start,

- `plot=T` opts to show plot,

The `plot` command acts on a `spec` object as follows:

```
> plot(x, ci = 0.95, log = c(''yes'', ''dB'',
''no''), xlab = ''frequency'',
ylab = NULL, type = ''l'', ci.col = ''blue'',
ci.lty = 3, main = NULL, sub = NULL)
```

in which

- `x` is a `spec` object,

- `ci` gives coverage probability of confidence interval,

- `log` – what scale is periodogram plotted on? Default is natural log, `dB` is "decibels" (see below),

- `type` defines type of plot, "line" being default.

Remark 0.3 *What is the tapering parameter? R applies a cosine bell taper where $x(t)$ is replaced by*

$$y(t) = \begin{cases} \left(1 - \cos\left(\frac{\pi\left(t-\frac{1}{2}\right)}{\alpha N}\right)\right) x(t) & t \leq \alpha N \\ x(t) & \alpha N < t < N - \alpha N \\ \left(1 - \cos\left(\frac{\pi\left(N-t+\frac{1}{2}\right)}{\alpha N}\right)\right) x(t) & t \geq (1-\alpha) N. \end{cases}$$

Remark 0.4 *Why plot the periodogram on the log scale? Recall that*

$$\frac{2I(\omega_p)}{f(\omega_p)} \sim \chi^2_2$$

so that

$$\log(I(\omega)) = const + \log(f(\omega)) + Z$$

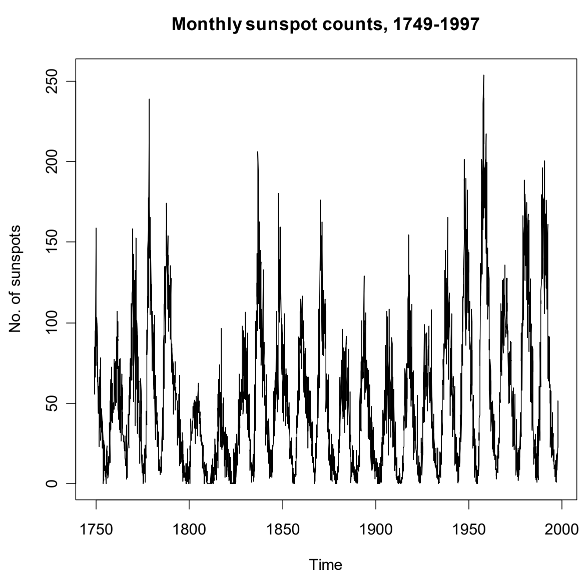
where Z *is a r.v. with a distribution related to* χ^2_2*. Hence log*$(I(\omega))$ *has a constant, additive variance – its variability about log*$(f(\omega))$ *is independent of* ω*.*

Remark 0.5 *The* decibel scale *is* $10\log_{10}$.

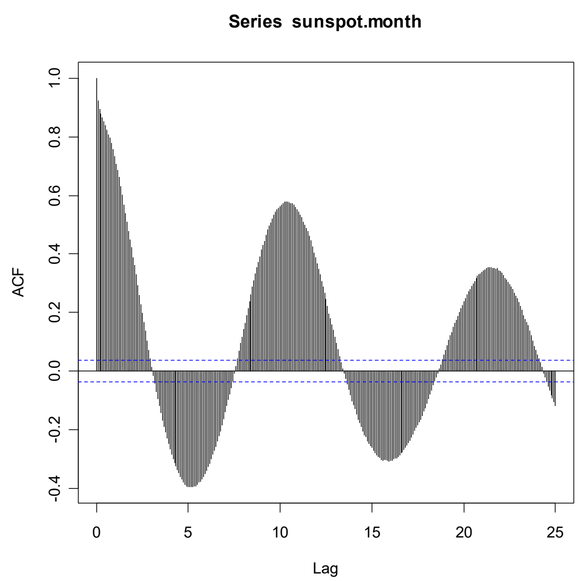
Remark 0.6 *R defines frequency as* **cycles per unit time***, though its choice of "unit time" may not always be obvious.*

Example 1

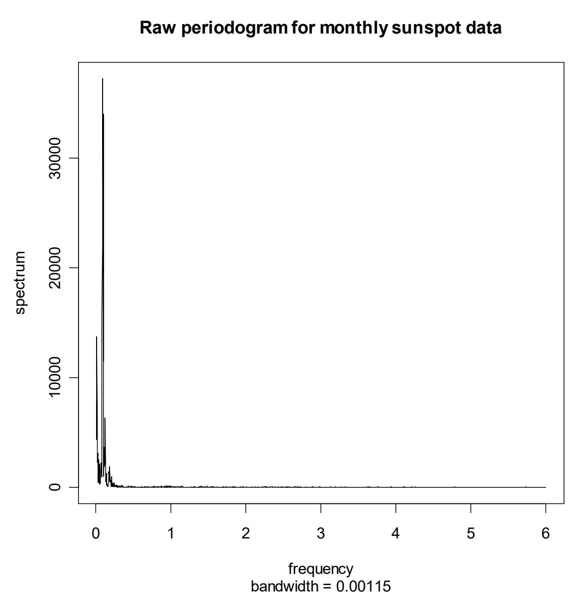
The data set `sunspot.month` in R is obtained via the Solar Influences and Data Analysis Centre (`www.sidc.oma.be`) and contains monthly sunspot numbers from 1749 to 1997, 2988 observations.



Examining the time domain first, observe the acf:



Note the lag unit is year, not month. The raw periodogram looks like:



We remark that:

- The data have been padded to make
$$3000 = 2^3 \times 3 \times 5^3$$
observations.

- The "un-logged" version of the periodogram is plotted above.

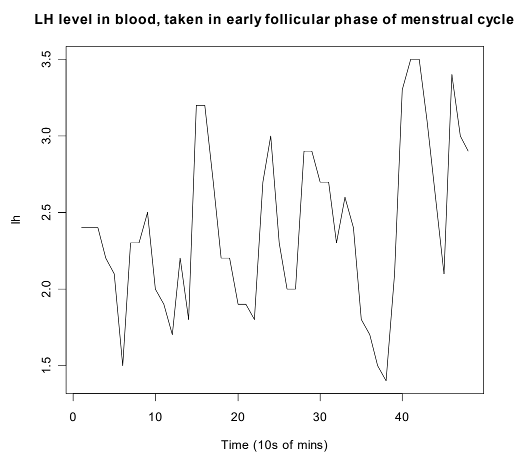
- The frequency scale is
$$\begin{aligned} f &= \frac{\omega_p}{2\pi\delta t} \\ &= \frac{\omega_p}{2\pi} \times 12 \end{aligned}$$
where $\delta t = 1/12$ and $\omega_p = 2\pi p/3000$, for $p = 1, 2, \ldots, 1500$.

The spectrum is dominated by low frequency values, with the maximum value, 37,273.16, at $p = 23$ (though 6 of the first 25 values are over 10,000). This corresponds to a wavelength of 10.87 years, an (approximate) 11 year "period" often claimed for sunspots. But the two frequencies either side this are of the same order of magnitude, and this cannot be totally explained by "leakage".

Example 2

Recall the lh data, which gives the luteinizing hormone in blood samples at ten–minute intervals from a woman over an eight–hour time period. Here $N = 48$

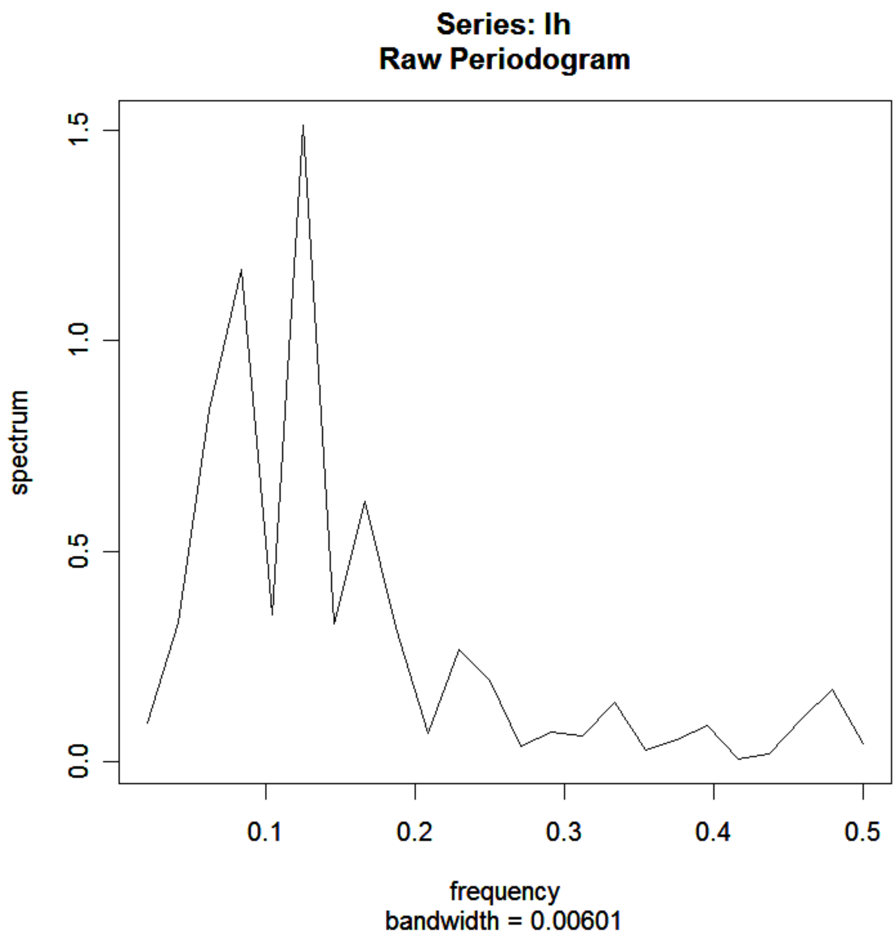
and the data are plotted again below:



There is a suggestion of a cyclical effect here, but of unknown period. We consider first the unsmoothed periodogram, plotted unlogged:

```
> speclh <- spec.pgram(lh)
```

```
> plot.spec(speclh, log="no")
```



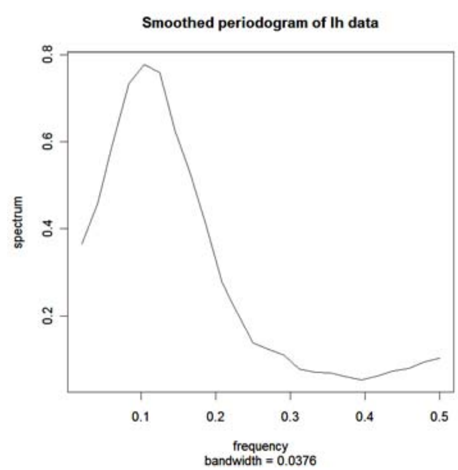
Inspecting the spectrum values via `speclh$spec` shows that the largest value occurs at the 6th harmonic, so this is at frequency

$$f_6 = \frac{6}{48} = \frac{1}{8}$$

cycles per unit time, where “unit time” here is ten minutes (see `speclh$freq`). This suggests a wavelength of around eighty minutes, completing six cycles over the course of the series.

We can smooth the spectrum of course: for example

```
> speclh <- spec.pgram(lh, spans=6)
> plot.spec(speclh, log="no")
```



We note though here that the cyclical component is not behaving like a simple sinusoidal model, since there are non-negligible periodogram values at other frequencies that contribute to the variation in the data.