Handout 12

Spectral Analysis

Analysis of time series has often two aspects
a) model fitting and forecasting, b) understanding hidden periodicities (spectral analysis).
We have discussed the first aspect. Now we are about to begin the second. Connection of the ARMA models to this second aspect will be clear later on.

The basis of spectral analysis is an important mathematical result which states that any stationary time series (causal or noncausal) can be written as (or approximated by) linear combinations of sines and cosines with random coefficients. Let us define an important quantity: frequency. A frequency is a real number between 0 and 1/2. For the energy data in Handouts 1 and, we see that, once the trend is subtracted, the detrended series has a similar pattern of behavior every 12 months. Energy consumptions in Januaries have similarities, energy consumptions in Marches have similarities an so on. In such a case we can say that energy consumption has an "important" frequency at $w = 1/12$. If we look at the sunspot data given in Handout 8, you can see that the peaks are occurring between 8 and 12 years. However, unlike in the energy data (which is seasonal), the peaks are not fixed in the sunspot series. There are 28.5 cycles in the series. So the series tends to behave similarly every 313/28.5=10.98 years on the average. We will see later that there is a peak at frequency at $w = 1/10.98 = 0.091$.

We will first begin with an important mathematical result.

Fact 1: Any stationary sequence (causal or non-causal) $\{X_t\}$ with zero mean can be approximated by a linear combinations of a large number of sines and cosines with random coefficients, i.e.,

$$X_t \approx \sum_{j=1}^{m} \{A_j \cos(2\pi w_j t) + B_j \sin(2\pi w_j t)\}, \tag{1}$$

where m is large, w_j's are distinct frequencies, $\{A_j\}$ and $\{B_j\}$ are random with zero means and $Var(A_j) = Var(B_j) = \sigma_j^2$. Moreover, $\{A_j\}$ and $\{B_j\}$ are all uncorrelated with each other. We will often take m to be equal to $n/2$ where n is the length of the observed series. Here is a fact.

Fact 2: The approximate variance and the autocovariances of a stationary sequence $\{X_t\}$ are

$$Var(X_t) \approx \sigma_1^2 + \cdots + \sigma_m^2, \ \gamma(h) \approx \sum_{j=1}^{m} \sigma_j^2 \cos(2\pi w_j h)$$

Remarks:

1. If X_t has mean μ, then the representation is valid with $X_t - \mu$ on the left hand side of the equation (1). From now on we will assume the mean to be equal to zero since we can always carry out spectral analysis by subtracting the mean.

2. In our analysis of energy data before, we modeled the data $Y_t = m_t + S_t + X_t$, where X_t is stationary. There are many approached to the analysis of this data set. In Handout 5 we estimated m_t and S_t from Y_t in order to get the stationary part. However, there is another way of modeling this. We can subtract the trend only, the the rest can be considered stationary (which contains a seasonal components). If a sequence has no trend, but has seasonality whose pattern does not change over time, then the sequence itself can be considered stationary.

3. In some trivial cases, m may be small. But in general m is large.

4. Goal of the spectral analysis is to find σ_j's. Note that

$$Var(X_t) \approx \sigma_1^2 + \cdots + \sigma_m^2.$$

Contribution of variability at frequency w_j is σ_j^2. It is of interest to find out which frequencies contribute more than others to this variability.

5. Since m is large and a substantial fraction of σ_j's are not equal to zero, then these σ_j's need to be small since $Var(X_t) \approx \sigma_1^2 + \cdots + \sigma_m^2$. It can be shown mathematically (under certain technical conditions), that there is a function f on $[-.5, .5]$ which is symmetric about zero, such that $\sigma_j^2 \approx 2f(j/n)/n$ when m is taken to be equal to $n/2$, where n is the length of the observed series. This function f is called the spectral density function. Note that $f(j/n)$ is the value of the spectral density function at frequency $w = j/n$ and it is a rescaled version σ_j^2.

How are the autocovariances and spectral density connected to one another? Here is an important theoretical result.

Fact 3: Let $\{X_t\}$ be a stationary sequence and let f be the spectral density function of this process.

a) For any positive integer h

$$\gamma(h) = \int_{-1/2}^{1/2} \exp(2\pi i w h) f(w) dw = 2 \int_0^{1/2} \cos(2\pi i w h) f(w).$$

b) For any $-1/2 \leq w \leq 1/2$

$$f(w) = \sum_{h=-\infty}^{\infty} \gamma(h) \exp(-2\pi i w h) = \gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h) \cos(2\pi w h)$$

Periodogram

From now on, we will take the approximation above in (1) with $m = n/2$ and $w_j = j/n, j = 1, \ldots, m$, with the understanding m equals $(n-1)/2$ if n is odd. We can estimate A_j and B_j from the data $\{X_1, ..., X_n\}$ using the method of least squares

$$\begin{aligned} \hat{A}_j &= \frac{\sum_{1\le t\le n} X_t \cos(2\pi w_j t)}{\sum_{1\le t\le n} \cos^2(2\pi w_j t)}, \hat{B}_j = \frac{\sum_{1\le t\le n} X_t \sin(2\pi w_j t)}{\sum_{1\le t\le n} \sin^2(2\pi w_j t)}, or \\ \hat{A}_j &= (2/n) \sum_{1\le t\le n} X_t \cos(2\pi w_j t), \hat{B}_j = (2/n) \sum_{1\le t\le n} X_t \sin(2\pi w_j t), \end{aligned}$$

the last equalities hold since it can be shown that

$$\sum_{1\le t\le n} \cos^2(2\pi w_j t) = n/2, \sum_{1\le t\le n} \sin^2(2\pi w_j t) = n/2.$$

The quantity $P(j/n) = \hat{A}_j^2 + \hat{B}_j^2$ is called the scaled periodogram and its rescaled version

$$I(j/n) = (n/4)P(j/n)$$

is called the periodogram. The main use of periodogram is that $I(j/n)$ is an estimator of $f(j/n)$, the spectral density function at frequency $w = j/n$. There is a related quantity called the discrete Fourier transform of the data

$$d(j/n) = (1/\sqrt{n}) \sum_{1\le t\le n} X_t \exp(-2\pi i w_j t),\ w_j = j/n,$$

where $i = \sqrt{-1}$ is the imaginary number and

$$\exp(-2\pi i w_j t) = \cos(2\pi w_j t) - i \sin(2\pi w_j t).$$

The connection between the discrete Fourier transform and the periodogram is

$$I(j/n) = |d(j/n)|^2 = (1/n)[\{ \sum_{1\le t\le n} X_t \cos(2\pi w_j t)\}^2 + \{ \sum_{1\le t\le n} X_t \sin(2\pi w_j t)\}^2]$$

Recall the sunspot series from Handout 8. We pointed out that the series needs to be transformed. You will find here both the sunspot series and the square rooted series. We will call the square rooted series $\{X_t\}$. The periodogram of $\{X_t\}$ is given here and plot looks quite unsmooth, whereas the periodogram is supposed to estimate the spectral density function which is usually smooth. The way out of this problem is to smooth the periodogram by local averaging. Before we get into more details of the smoothing, here is an important result.

Fact 4: a) When n is large, $I(j/n) \approx (\xi_j/2) f(j/n)$, where ξ_j has a chi-square distribution with 2 degrees of freedom. [It is known that $E(\xi_j) = 2$ and $Var(\xi_j) = 4$].
b) $I(\xi_j)$'s are independent.

Smoothing of periodogram

The easiest way to smooth a periodogram is to take a local (weighted) average. There are many methods for smoothing, but we will present only two. For notational convenience let us denote $f(j/n)$ by f_j. If we want to smooth over the periododogram values at four nearest frequencies (and at the frequency itself), then the smoothed periodogram is

$$\hat{f_j} = (1/5)[I_{j-2} + I_{j-1} + I_j + I_{j+1} + I_{j+2}],$$

where we write $I_j = I(j/n)$ for notational simplicity. This simple averaging method is called Daniell's method of smoothing.

A slightly improved version of Daniell's method (called modified Daniell) is to give half the weights for the farthest two neighboring frequencies. Hence modified Daniell's method for this case will be

$$\hat{f_j} = (1/4)[(1/2)I_{j-2} + I_{j-1} + I_j + I_{j+1} + (1/2)I_{j+2}].$$

The R code for this is - spec.pgram (x,spans=5,log="no").

This function will obtain and plot a smoothed estimate of the spectral density via modified Daniell's (default) method. Here 'spans' refer to the total number of periodogram values used in obtaining the smoothed estimate of f_j. Note that R automatically plots the logarithm of the estimated spectral density. That is why one need to tell R that it should not be plotted on the logarithmic scale.

Note that if you use k nearest neighbors for smoothing, then the modified Daniell's method is (for the R command, $spans = 2k + 1$)

$$\hat{f_j} = \frac{1}{2k}[(1/2)I_{j-k} + I_{j-k+1} + \cdots + I_j + \cdots + I_{j+k-1} + (1/2)I_{j+k}].$$

You will find below a plot with three modified Daniell estimates of the spectral density function: averaged periodogram over 2 neighbors, 6 neighbors and 12 neighbors. Note that smaller the number of neighbors used in smoothing, rougher is the plot. On the other hand, if the smoothing is done over too many neighbors, plot is smoother but flatter thereby loosing its resolution. In other words, one has to carefully select the number of neighbors used in smoothing in order to have a smooth plot without losing important details. There are AIC type criteria to decide how many neighbors should be used. Here we will use $k = 6$ by visual inspection and this is supported by a formal criterion given below.

A method for selecting k

We will write down a reasonable method for selecting k. If we use k neighbors for simple averaging and denote the estimates by $\{\hat{f_j}\}$, then consider the quantity

$$Q_A(k) = \sum (I_j - \hat{f_j})^2 + \frac{1}{2k+1} \sum I_j^2.$$

Then plot the criterion function $Q_A(k)$ against k and, if Q_A is the smallest at $k = k_A$, then k_A is the number of neighbors that should be used in smoothing. Note that $\{\hat{f}_j\}$ have to be calculated separately for every value of k.

If k neighbors are used in obtaining modified Daniell's estimator and the estimates are $\{\hat{f}_j\}$, then consider the criterion

$$Q_D(k) = \sum (I_j - \hat{f}_j)^2 + \frac{1}{2k} \sum I_j^2.$$

Then if Q_D attains its minimum at $k = k_D$, then k_D neighbors should be used in obtaining estimates of f_j if modified Daniell's method is to be employed for smoothing.

For the sunspot data, we have used modified Daniell's method to obtain a plot of $Q_D(k)$ against k. Note that the minimum is attained at $k = 6$, i.e., $k_D = 6$ (so *spans* = 13 in the R command 'spec.pgram') neighbors should be used in order to obtain the spectral density estimate.

R-commands

Suppose the series is stored as 'x'. Then here are two important commands

i) spec.pgram(x,log="no") [gives raw periodogram]

ii) spec.pgram(x,spans=2k+1,log="no") [gives smoothed spectral density estimates using modified Daniell's method with periodogram values at k nearest frequencies].

Suppose you want to obtain the frequencies and the estimates in the spctral density in (ii), then you can write

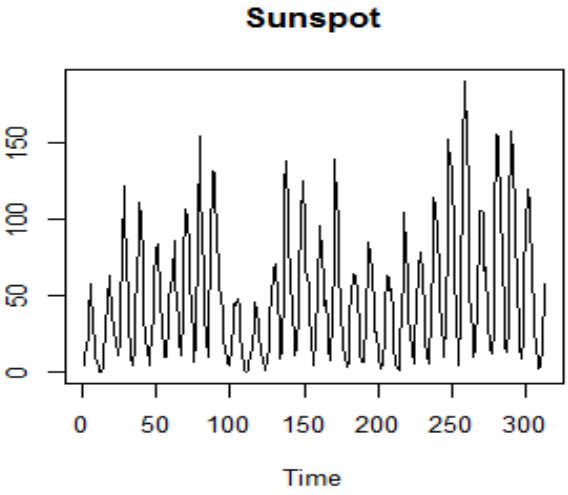
ss=spec.pgram(x,spans=2k+1,log="no").

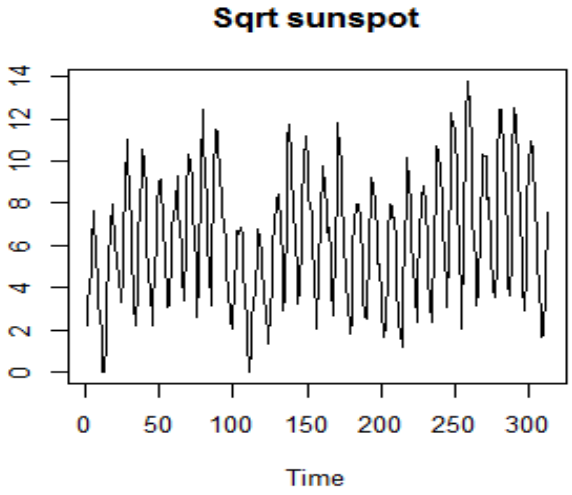
The frequencies and the spectral density estimates can be obtained from ss$freq and ss$spec, respectively.

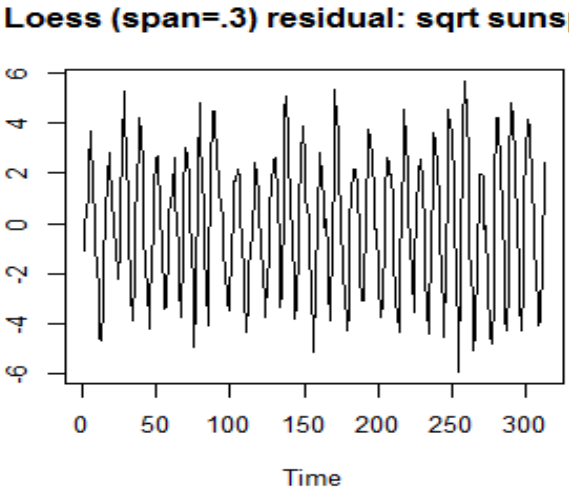
Sunspot data

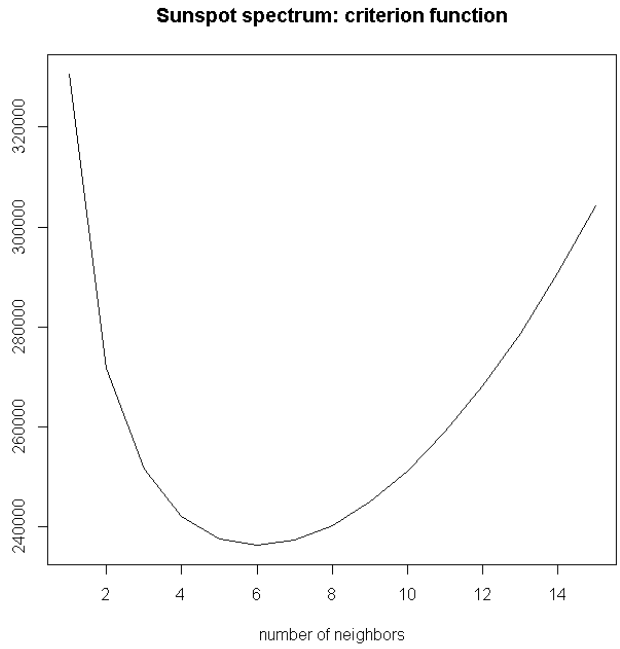
For the square rooted sunspot data, it is reasonable to choose six neighbors to obtain an estimate of the spectral density. Note that there is a sharp peak at frequency $w = 0.091$. This tells us we can expect a cycle of $1/0.091 = 11$ years, on the average. This is also confirmed by that the observation that in the plot of square rooted sunspot series, there are 28.5 cycles out of $n = 313$ years. This says that we can expect the cycles to appear every $313/28.5 = 10.98$ years, on the average. So we can expect a sharp peak of the spectral density at the frequency $28.5/313 = 0.091$ and this matches well with the peak of the estimated spectral density.

Also note that there is another peak at frequency zero. The peak at zero is not that prominent in comparison to the other peak. If the peak at zero is very sharp, this can be taken as an indication that a trend is present in the series. In this case the peak at zero is not that sharp, thus there is no clear indication that if there is a trend. Still, for the purpose of investigation, we estimate the trend by using the loess method with *span* = 0.3, subtract it from the data and then obtain a smoothed peridogram of the detrended data. Note that the peak at zero is gone.


Sunspot
0
50
100
150
0
50
100
150
200
250
300
Time


Sqrt sunspot
0
2
4
6
8
10
12
14
0
50
100
150
200
250
300
Time


Loess (span=.3) residual: sqrt sunspot
-6
-4
-2
0
2
4
6
0
50
100
150
200
250
300
Time



Filtering a series

If X_t is a series, we may decide to look at its first difference or a running weighted average of the series

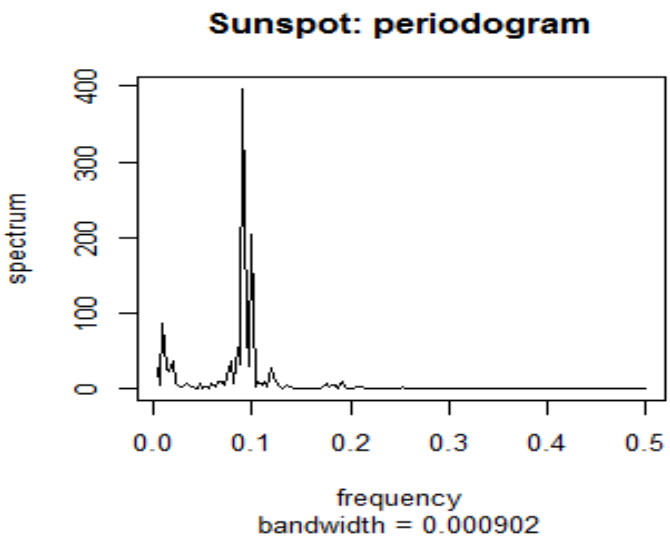
$$a)\ Z_t = X_t - X_{t-1},\ b)\ Z_t = (1/2)X_t + (1/2)X_{t-1},\ c)\ Z_t = (1/3)X_t + (1/3)X_{t-1} + (1/3)X_{t-2}.$$

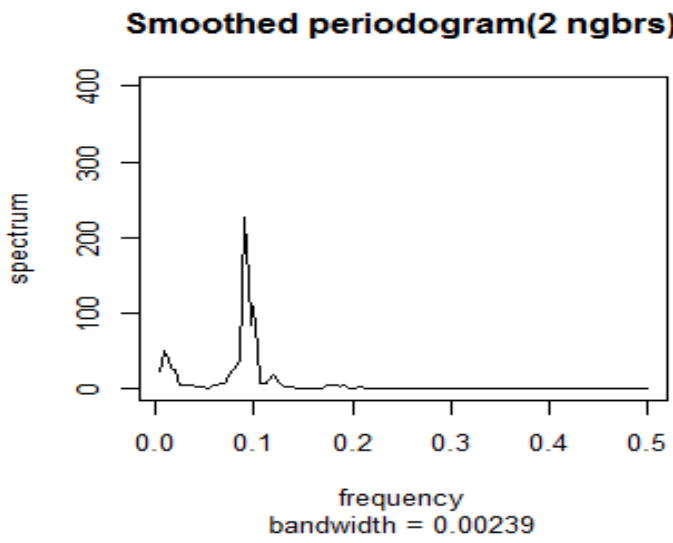
In each case above, Z_t is a linear combination of the X series. A linear combination of the X-series is called a filtered series of $\{X_t\}$. It turns out there is a nice formula connecting the spectral density of X series to that of the Z-series when $\{X_t\}$ is stationary. Denoting the spectral density functions of $\{X_t\}$ and $\{Z_t\}$ by f_X and f_Z, respectively, for the cases (a) and (b) above we have

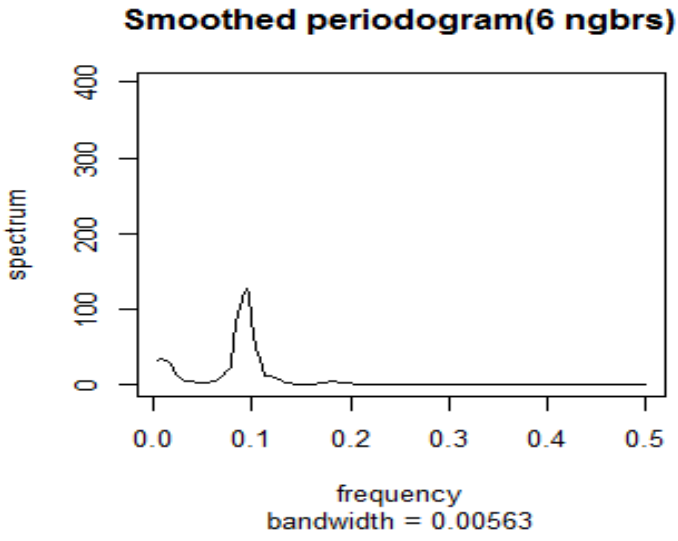
$$a)\ f_Z(w) = (2 - 2\cos(2\pi w))f_X(w),\ b)\ f_Z(w) = (1/2)(1 + \cos(2\pi w))f_X(w).$$

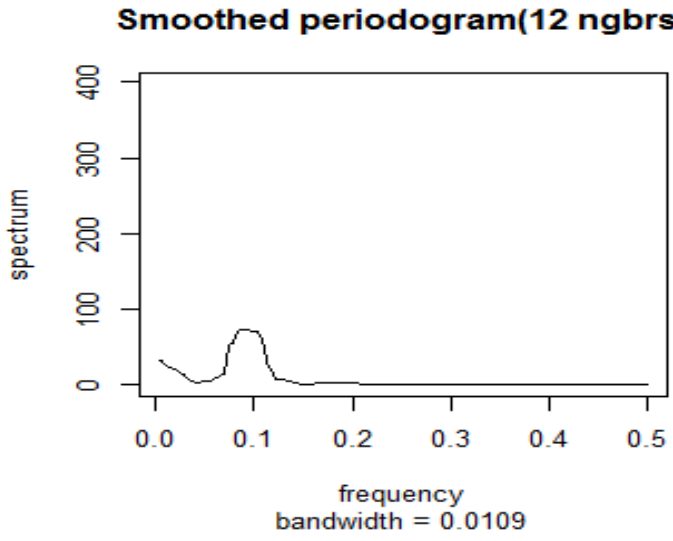
In other words, in each case the spectral density of Z_t is equal to the spectral density of X_t times a weight function. Note that in (a), the weight function is zero at zero and then monotonically increases to the value of 4 at $w = 1/2$. In other words, higher the frequency higher is the weight indicating that the first difference of $\{X_t\}$ is a rougher series than $\{X_t\}$. Now let us note (b). The weight function $(1/2)(1+\cos(2\pi w))$ equals 1 at $w = 0$ and then it decreases to zero $w = 1/2$. So here we have the scenario, higher the frequency lower is the weight. This indicates that the running average Z_t of the series $\{X_t\}$ is smoother than $\{X_t\}$. Here is a general result.

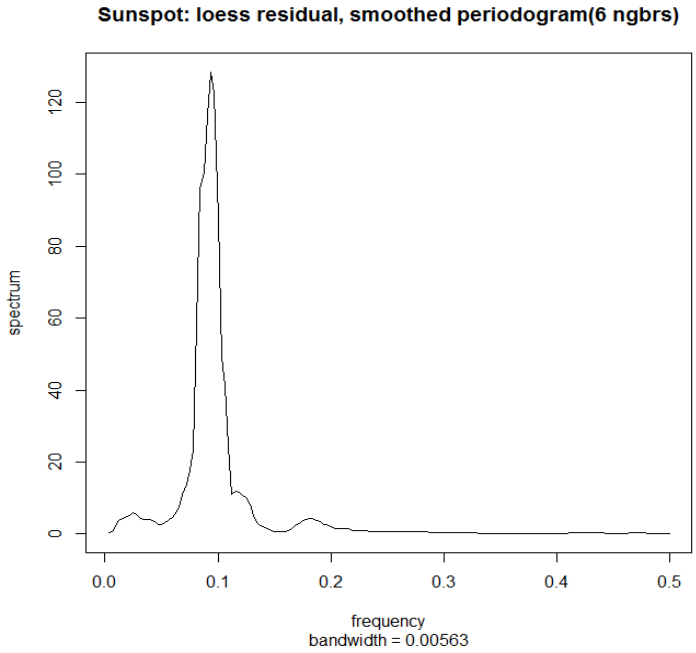
Fact 5. Let $Z_t = \sum_{-\infty<s<\infty} \psi_s X_{t-s}$, where $\{X_t\}$ is stationary and $\{\psi_s\}$ are constants satisfying


Sunspot: periodogram
spectrum
0
100
200
300
400
0.0
0.1
0.2
0.3
0.4
0.5
frequency
bandwidth = 0.000902


Smoothed periodogram(2 ngbrs)
spectrum
0
100
200
300
400
0.0
0.1
0.2
0.3
0.4
0.5
frequency
bandwidth = 0.00239


Smoothed periodogram(6 ngbrs)
spectrum
0
100
200
300
400
0.0
0.1
0.2
0.3
0.4
0.5
frequency
bandwidth = 0.00563


Smoothed periodogram(12 ngbrs
spectrum
0
100
200
300
400
0.0
0.1
0.2
0.3
0.4
0.5
frequency
bandwidth = 0.0109



the condition $\sum_{-\infty<s<\infty}|\psi_s| < \infty$. The **frequency response function** $\Psi(w)$ is defined as $\Psi(w) = \sum_{-\infty<s<\infty}\psi_s \exp(-2\pi i w s)$. Then

$$f_Z(w) = |\Psi(w)|^2 f_X(w), i = \sqrt{-1},$$

Here is a consequence of this fact.

Fact 6. Let $Z_t = (X_t + \cdots X_{t-L+1})/L$, where $\{X_t\}$ is stationary. Then

$$f_Z(w) = \frac{1}{L^2}\frac{1-\cos(2\pi w L)}{1-\cos(2\pi w)} f_X(w).$$

Spectral density for ARMA

Stationary ARMA sequences have formulas for their spectral densities and R can be used to obtain and plot them. Before we proceed any further. Let us look at the square rooted sunspot series again. We have used the function auto.arima in R to obtain the "best" ARMA model for the loess residual series using the AICC criterion. The fitted model turns out to be $ARMA(2,2)$ with the estimated parameters given below

$$\begin{aligned}\hat{\phi}_1 &= 1.5771, \hat{\phi}_2 = -0.8827, \hat{\theta}_1 = -0.4762, \hat{\theta}_2 = -0.0863, \hat{\sigma}^2 = 1.349,\\ s(\hat{\phi}_1) &= 0.0299, s(\hat{\phi}_2) = 0.0278, s(\hat{\theta}_1) = 0.0635, s(\hat{\theta}_2) = 0.0562.\end{aligned}$$

We use the function **arma.spec** in R to obtain the spectral density estimator along with the modified Daniell estimator (averaging over 6 neighbors) and they are plotted here. These two different estimators

are quite close as is evident from the plot. The estimator with the higher peak is the smoothed periodogram estimate.

We also used the function auto.arima to select the best ARMA model for the square rooted series and have plotted the spectral density of the fitted series. The best model according to the AICC criterion is now an $ARMA(2,1)$ model

$$\begin{aligned} \hat{\phi}_1 &= 1.4773, \hat{\phi}_2 = -0.7552, \hat{\theta}_1 = -0.1318, \hat{\mu} = 6.3960, \hat{\sigma}^2 = 1.05, \\ s(\hat{\phi}_1) &= 0.0509, s(\hat{\phi}_2) = 0.0462, s(\hat{\theta}_1) = 0.0759, s(\hat{\mu}) = 0.2053. \end{aligned}$$

Before we write down the general result, let us consider a few simpler cases.

White noise. The spectral density of white noise with variance σ^2 is a constant and it given by

$$f(w) = \sigma^2, \text{ for all } w.$$

MA(1). The spectral density of an $MA(1)$ sequence is given by

$$f(w) = \sigma^2(1 + \theta^2 + 2\theta\cos(2\pi w)).$$

AR(1). The spectral density of an $AR(1)$ sequence is given by

$$f(w) = \sigma^2/[1 + \phi^2 - 2\phi\cos(2\pi w)].$$

$ARMA(1,1)$**.** The spectral density of an $ARMA(1,1)$ sequence is given by

$$f(w) = \sigma^2 \frac{1 + \theta^2 + 2\theta\cos(2\pi w)}{1 + \phi^2 - 2\phi\cos(2\pi w)}.$$

Note the structure of the spectral density for an $ARMA(1,1)$ sequence. The numerator comes from the MA part and the denominator comes from the AR part. The general result for ARMA sequences is exactly the same. Before we state the general result let us once again remind ourselves of some notations used a few weeks ago. Let us recall the polynomials used in the discussion on issues of stationarity, inverrtibility etc.

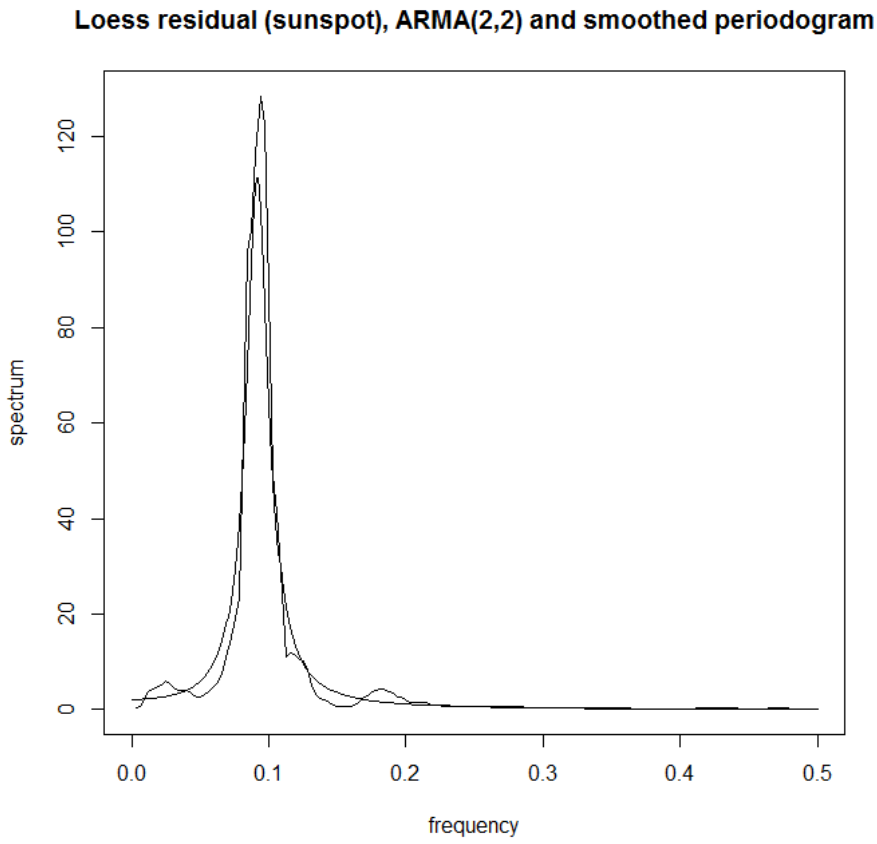
$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q, \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p.$$

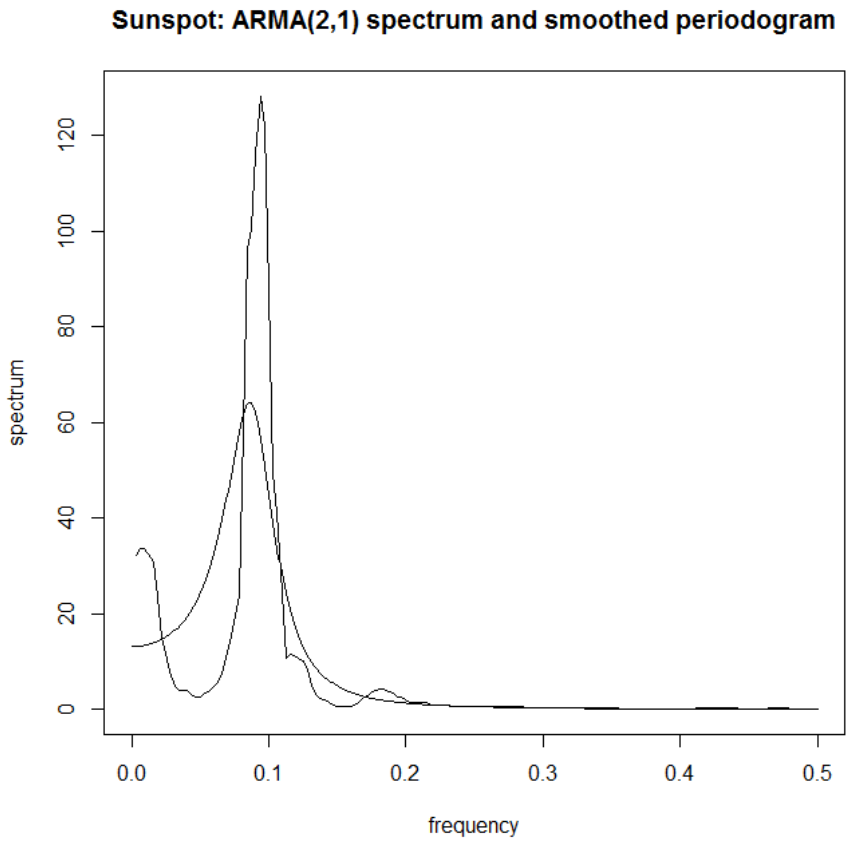
Fact 7. All the series considered here are stationary.
a) If $\{X_t\}$ is $MA(q)$, then its spectral density is given by

$$f(w) = \sigma^2|\theta(z)|^2, where\ z = \exp(-2\pi i w).$$

b) If $\{X_t\}$ is $AR(p)$, then its spectral density is given by

$$f(w) = \sigma^2/|\phi(z)|^2, where\ z = \exp(-2\pi i w).$$


Loess residual (sunspot), ARMA(2,2) and smoothed periodogram
spectrum
0
20
40
60
80
100
120
0.0
0.1
0.2
0.3
0.4
0.5
frequency


Sunspot: ARMA(2,1) spectrum and smoothed periodogram
spectrum
0
20
40
60
80
100
120
0.0
0.1
0.2
0.3
0.4
0.5
frequency

c) If $\{X_t\}$ is $AR(p,q)$, then its spectral density is given by

$$f(w) = \sigma^2[|\theta(z)|^2/|\phi(z)|^2], where\ z = \exp(-2\pi i w).$$

Some proofs

Fact 2. Note that, since A_j and B_j are uncorrelated with each other, we have

$$\begin{aligned}
\gamma(h) &= Cov(X_t, X_{t+h}) \\
&\approx \sum Var(A_j)\cos(2\pi w_j t)\cos(2\pi w_j(t+h)) + \sum Var(B_j)\sin(2\pi w_j t)\sin(2\pi w_j(t+h)) \\
&= \sum \sigma_j^2 \cos(2\pi w_j t)\cos(2\pi w_j(t+h)) + \sum \sigma_j^2 \sin(2\pi w_j t)\sin(2\pi w_j(t+h)) \\
&= \sum \sigma_j^2 [\cos(2\pi w_j t)\cos(2\pi w_j(t+h)) + \sin(2\pi w_j t)\sin(2\pi w_j(t+h))] \\
&= \sum \sigma_j^2 \cos(2\pi w_j h).
\end{aligned}$$

The last is true since for any constants c and d

$$\cos(c)\cos(d) + \sin(c)\sin(d) = \cos(c-d).$$

The result for $\gamma(0)$ follows easily from the expression of $\gamma(h)$ when $h = 0$.

Spectral density of $Z_t = X_t - X_{t-1}$.
Note that we can write this as a special case of the general result stated in Fact 5 with $\psi_s = 0$ when $s < 0$, $\psi_0 = 1$, $\psi_1 = -1$, and $\psi_s = 0$ when $s > 1$. Then

$$\begin{aligned}
\sum \psi_s \exp(-2\pi i w s) &= 1 - \exp(-2\pi i w) = 1 - [\cos(2\pi w) - i\sin(2\pi w)] \\
&= (1 - \cos(2\pi w)) + i\sin(2\pi w).
\end{aligned}$$

Hence

$$\begin{aligned}
|\sum \psi_s \exp(-2\pi i w s)|^2 &= (1-\cos(2\pi w))^2 + \sin^2(2\pi w) \\
&= 1 + \cos^2(2\pi w) - 2\cos(2\pi w) + \sin^2(2\pi w) \\
&= 2 - 2\cos(2\pi w).
\end{aligned}$$

Spectral density of $Z_t = (1/2)X_t + (1/2)X_{t-1}$.
Note that here we have $\psi_s = 0$ when $s < 0$, $\psi_0 = 1/2$, $\psi_1 = 1/2$, and $\psi_s = 0$ when $s > 1$. Rest of the arguments are similar to the case for the first difference.

Spectral density of $Z_t = (X_t + \cdots X_{t-L+1})/L$.
Note that we can take $\psi_s = 0$ when $s < 0$, $\psi_0 = \cdots = \psi_{L-1} = 1/L$, and $\psi_s = 0$ when $s > L-1$. Hence the frequency response function is

$$\Psi(w) = \sum \psi_s \exp(-2\pi i w s) = (1/L)[1 + z + \cdots + z^{L-1}],\ \text{where } z = \exp(-2\pi i w).$$

Note that

$$1 + z + \cdots + z^{L-1} = \frac{1 - z^L}{1 - z}.$$

When $z = \exp(-2\pi i w)$ we have

$$\begin{aligned} 1 - z &= (1 - \cos(-2\pi w)) + i \sin(2\pi w), \\ 1 - z^L &= 1 - \exp(-2\pi i w L) = (1 - \cos(2\pi w L)) + i \sin(2\pi w L). \end{aligned}$$

From these we have, when $z = \exp(-2\pi i w)$,

$$\begin{aligned} |1 - z|^2 &= (1 - \cos(2\pi w))^2 + \sin^2(2\pi w) = 2 - 2\cos(2\pi w), \\ |1 - z^L|^2 &= (1 - \cos(2\pi w L))^2 + \sin^2(2\pi w L) = 2 - 2\cos(2\pi w L). \end{aligned}$$

Combining all the steps we conclude that

$$|\Psi(w)|^2 = \frac{|1 - z^L|^2}{|1 - z|^2} = \frac{1 - \cos(2\pi w L)}{1 - \cos(2\pi w)}.$$

The result now follows from Fact 5.

Spectral density of White noise.
For a white noise $\gamma(0) = \sigma^2$ and $\gamma(h) = 0$ whenever $h \neq 0$. Hence, for any w, the spectral density is

$$f(w) = \sum \gamma(h) \exp(-2\pi i w h) = \sigma^2.$$

Spectral density of $MA(1)$
Note that we can re-express X_t as $X_t = \sum \psi_s \varepsilon_{t-s}$, where $\psi_s = 0$ when $s < 0$, $\psi_0 = 1$, $\psi_1 = \theta$, and $\psi_s = 0$ whenever $s > 1$. Hence by Fact 5 we have

$$f_X(w) = |\Psi(w)|^2 f_\varepsilon(w),$$

where $\Psi(w)$ is the frequence response function. Since

$$\Psi(w) = \sum \psi_s \exp(-2\pi i w s) = 1 + \theta \exp(-2\pi i w) = (1 + \theta \cos(2\pi w)) - i\theta \sin(2\pi w),$$

we have

$$|\Psi(w)|^2 = 1 + \theta^2 + 2\theta \cos(2\pi w).$$

The result now follows once we note that $\{\varepsilon_t\}$ is white noise and $f_\varepsilon(w) = \sigma^2$ for all w.

Spectral density for $AR(1)$.
We can re-express the $AR(1)$ sequence as $\varepsilon_t = X_t - \phi X_{t-1}$. Take $\psi_s = 0$ when $s < 0$, $\psi_0 = 1$, $\psi_1 = -\phi$, and $\psi_s = 0$ whenever $s > 1$. Then

$$f_\varepsilon(w) = |\Psi(w)|^2 f_X(w).$$

Since the frequency response function is

$$\Psi(w) = \sum \psi_s \exp(-2\pi i w s) = 1 - \phi \exp(-2\pi i w) = (1 - \phi\cos(2\pi w)) + i\phi \sin(2\pi w),$$

we have

$$|\Psi(w)|^2 = 1 + \phi^2 - 2\phi\cos(2\pi w).$$

Hence we get

$$\begin{aligned} f_\varepsilon(w) &= (1 + \phi^2 - 2\phi\cos(2\pi w)) f_X(w), i.e., \\ f_X(w) &= f_\varepsilon(w)/(1 + \phi^2 - 2\phi\cos(2\pi w)). \end{aligned}$$

The result follows once we note that $f_\varepsilon(w) = \sigma^2$ for all w.

Spectral density for $ARMA(1,1)$

We can re-express $ARMA(1,1)$ sequence as $X_t - \phi X_{t-1} = \varepsilon_t + \theta\varepsilon_{t-1}$. So the spectral density of $X_t - \phi X_{t-1}$ equals the spectral density of $\varepsilon_t + \theta\varepsilon_{t-1}$. Note that the spectral density of $X_t - \phi X_{t-1}$ is

$$(1 + \phi^2 - 2\phi\cos(2\pi w)) f_X(w).$$

The spectral density of $\varepsilon_t + \theta\varepsilon_{t-1}$ is

$$(1 + \theta^2 + 2\theta\cos(2\pi w)) f_\varepsilon(w).$$

Consequently we have

$$\begin{aligned} (1 + \phi^2 - 2\phi\cos(2\pi w)) f_X(w) &= (1 + \theta^2 + 2\theta\cos(2\pi w)) f_\varepsilon(w), i.e., \\ f_X(w) &= \frac{1 + \theta^2 + 2\theta\cos(2\pi w)}{1 + \phi^2 - 2\phi\cos(2\pi w)} f_\varepsilon(w). \end{aligned}$$

The result now follows since $f_\varepsilon(w) = \sigma^2$ for all w.

Appendix.

Fact 5. Note that the autocovariance function of the Z series is

$$\begin{aligned} \gamma_Z(h) &= Cov(Z_t, Z_{t+h}) = Cov(\sum \psi_{s_1} X_{t-s_1}, \sum \psi_{s_2} X_{t+h-s_2}) \\ &= \sum\sum \psi_{s_1}\psi_{s_2} Cov(X_{t-s_1}, X_{t+h-s_2}) = \sum\sum \psi_{s_1}\psi_{s_2}\gamma_X(h + s_1 - s_2). \end{aligned}$$

From part (b) of Fact 3 we have

$$\begin{aligned} f_Z(w) &= \sum_{h=-\infty}^{\infty} \gamma_Z(h)\exp(-2\pi i w h) \\ &= \sum_h [\sum_{s_1}\sum_{s_2} \psi_{s_1}\psi_{s_2}\gamma_X(h + s_1 - s_2)]\exp(-2\pi i w h) \\ &= \sum_h \sum_{s_1}\sum_{s_2} \psi_{s_1}\exp(2\pi i w s_1)\psi_{s_2}\exp(-2\pi i w s_2)\gamma_X(h + s_1 - s_2)\exp(-2\pi i w(h + s_1 - s_2)) \end{aligned}$$

Now denote $h + s_1 - s_2$ by j. Since s_1, s_2 and j vary between $-\infty$ and ∞, we have

$$\begin{aligned} f_Z(w) &= [\sum_{s_2} \psi_{s_1} \exp(2\pi i w s_1)][\sum_{s_2} \psi_{s_2} \exp(-2\pi i w s_2)][\sum_{j} \gamma_X(j) \exp(-2\pi i w j)] \\ &= |\sum \psi_{s_1} \exp(2\pi i w s_1)|^2 f_X(w). \end{aligned}$$

Fact 7. In all the proofs we note that the spectral density of $\{\varepsilon_t\}$ is a constant, i.e., $f_\varepsilon(w) = \sigma^2$ for all w. a) For the $MA(q)$ sequence take $\psi_s = 0$ when $s < 0$, $\psi_0 = 1$, $\psi_1 = \theta_1, ..., \psi_q = \theta_q$, and $\psi_s = 0$ whenever $s > q$. Hence we can write X_t as $X_t = \sum \psi_s \varepsilon_{t-s}$. Now

$$|\sum \psi_s \exp(-2\pi i w s)|^2 = |1 + \theta_1 z + \cdots + \theta_q z^q|^2 = |\theta(z)|^2, \text{ where } z = \exp(-2\pi i w).$$

From Fact 5 we get the result

$$f_X(w) = |\sum \psi_s \exp(-2\pi i w s)|^2 f_\varepsilon(w) = |\theta(z)|^2 \sigma^2, \text{ where } z = \exp(-2\pi i w).$$

b) We can re-express the $AR(p)$ sequence as

$$\varepsilon_t = X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}.$$

Take $\psi_s = 0$ when $s < 0$, $\psi_0 = 1$, $\psi_1 = -\phi_1, ..., \psi_p = -\phi_p$, and $\psi_s = 0$ whenever $s > p$. Then we have

$$\begin{aligned} f_\varepsilon(w) &= |\sum \psi_s \exp(-2\pi i w s)|^2 f_X(w), i.e., \\ f_X(w) &= f_\varepsilon(w) / |\sum \psi_s \exp(-2\pi i w s)|^2. \end{aligned}$$

The result follows once we note that

$$|\sum \psi_s \exp(-2\pi i w s)|^2 = |1 - \phi_1 z - \cdots - \phi_p z^p|^2 = |\phi(z)|^2, \text{ where } z = \exp(-2\pi i w).$$

c) We can re-express an $ARMA(p, q)$sequence as

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}.$$

The spectral density of the left hand side and the right hand side are $|\phi(z)|^2 f_X(w)$ and $|\theta(z)|^2 f_\varepsilon(w) = |\theta(z)|^2 \sigma^2$, respectively. Hence we get

$$|\phi(z)|^2 f_X(w) = |\theta(z)|^2 \sigma^2, i.e., f_X(w) = \frac{|\theta(z)|^2}{|\phi(z)|^2} \sigma^2.$$