Time Series A3

Andrew Martin

September 7, 2018

1. Let $\{U_t\}$ be a white noise sequence, and let the series $\{Y_t\}$ be defined by

$$Y_t = \sum_{j=0}^{m} U_{t-j}/(m+1)$$

Show that the autocorrelation function for $\{Y_t\}$ is:

$$\rho_k = \begin{cases} \frac{(m+1-k)}{(m+1)} & k = 0, 1, \ldots, m \\ 0 & \text{otherwise} \end{cases}$$

Solution This is a moving average model. From lectures we know the moving average model

$$Y_t = \theta(B)Z_t = \sum_{j=0}^{q} B_j Z_{t-j}$$

Has autocorrelation

$$\rho_k = \frac{\sum_{j=0}^{q-k} \beta_j \beta_{j+k}}{\sum_{j=0}^{q} \beta_j^2}$$

for $k = 0, 1, \ldots, m$ and 0 otherwise. So in our case $\beta_j = \frac{1}{m+1} \forall j$, which gives

$$\begin{aligned}
\rho_k &= \frac{\sum_{j=0}^{q-k} \beta_j \beta_{j+k}}{\sum_{j=0}^{q} \beta_j^2} \\
&= \frac{\sum_{j=0}^{m-k} \frac{1}{(m+1)^2}}{\sum_{j=0}^{m} \frac{1}{(m+1)^2}} \\
&= \frac{\sum_{j=0}^{m-k} 1}{\sum_{j=0}^{m} 1} \\
&= \frac{m-k+1}{m+1} \text{ for } k = 0, 1, \ldots, m
\end{aligned}$$

For $k > m$ the top sum will instead be 0.

As required.

2. The file `electricity.csv` contains the Australian monthly electricity production from January 1956 to August 1995

   (a) Read the data into `R` and construct a time series object
   Solution This is shown in the code.
   As required.

   (b) Plot the log transformed series and also its periodogram. Describe and interpret the main features of the periodogram.
   Solution This can be seen in plots (a) and (b) respectively. The periodogram shows the effect of periodic elements in the data. There are clearly strong periodic elements with frequency 1,2,3, and 5. The one with greatest effect is 1. The highest peak still corresponds to frequency 1, meaning there is a strong 1 monthly periodic component.
   As required.

(c) Detrend the log transformed series by calculating residuals from a cubic regression on time. Plot the detrended data and its periodogram. Describe and interpret the main features of the periodogram.

Solution This is done in the code, and the plots are plot (c) and (d) respectively. The periodogram produced here is very similar to the original periodogram, with the amplitudes of the peaks decreased, and some of the shape changed. The highest peak still corresponds to frequency 1, meaning there is a strong 1 monthly periodic component.
As required.

(d) Estimate the periodic component corresponding to the dominant frequency in the residual series and plot it over a cycle of 12 months

Solution This frequency was estimated to be 1. This is found in plot (e).
As required.

3. This question is concerned with the sunspot data available as the build-in dataset `sunspots` in `R`.

(a) Obtain the residual series from a cubic regression on time.

Solution This is done in the code.
As required.

(b) Obtain and interpret the cumulative periodogram for the residual series.

Solution This is shown in plot (f). If the data lies within the blue lines, the assumption that the data is white noise can be made. Clearly this data is not in the blue bands, and is not white noise.
As required.

(c) Obtain and interpret the periodogram from the residual series

Solution This is plot (g). The plot shows a distinctive peak right near the start, with a negative trend as the frequency increases. There are many troughs in the plot corresponding to frequencies around 1,3,4,5,6, indicating that the periodic component is not 1,3,4,5, or 6 monthly.
As required.

(d) Identify the dominant frequency and express its period in years

Solution This is done in the code, and the dominant frequency was found to be 0.091666... And its period was 10.9. This corresponds to a sunspot every 10.9 years.
As required.

(e) Estimate the periodic component corresponding to the dominant frequency of the residual series and add it to a time series plot of that series.

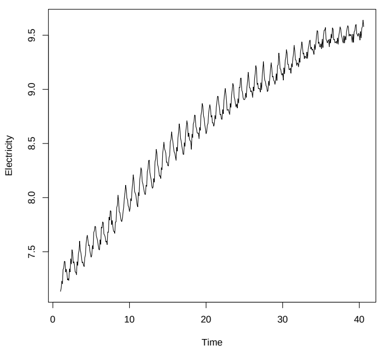
Solution Using the dominant frequency from the previous question, this is done in the code. The plot is plot (g), where the smaller, magenta wave is the periodic component generated, and the larger black plot is the sunspot data.
As required.

Code

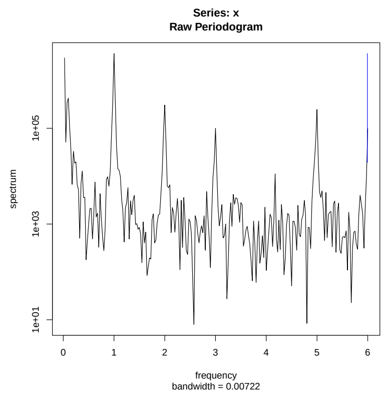
```
library(MASS)
setwd("~/Uni/2018/Sem2/Time Series")
pdf(file="A3Plots.pdf")

##Q2----
#a) Read in electricity as a ts object

elec = read.csv("electricity-1.csv")
elec = ts(elec,deltat=1/12)

#b) plot log and periodogram
plot(log(elec))
spectrum(elec,log="yes")

#c) detrend by using cubic regression on time.
```

```
#Plot detrended data & periodogram
#describe and interpret features of periodogram
et = time(elec)
eres = residuals(lm(log(elec)~et + I(et^2) + I(et^3)))
eres =ts(eres,deltat=1/12)
plot(ts(eres))
spectrum(eres)

#d) estimate the periodic component
#corresponding to the dominant frequency of residuals
#plot it over a cycle of 12 months
#the dominant freq is 1
t = time(eres)
c12 = cos(2*pi*t)
s12 = cos(2*pi*t)
plot(fitted(lm(eres~c12+s12))[1:12],type = "l")


##Q3----
sun = sunspots
st = time(sun)

#a) obtain residuals from cubic regression
sres = residuals(lm(sun~st + I(st^2) + I(st^3)))
sres = ts(sres,start=1749,deltat=1/12)

#b) cumulative periodogram for the residuals
cpgram(sres)
#c) periodogram from the residuals
sspec = spectrum(sres)
#d) dominant frequency
#estimate the periodic component of the dominant frequency

sMaxIndex = which.max(sspec$spec)
sMaxFreq = sspec$freq[22]
#Corresponds to a 10.9 yearly cycle
speriod = 1/sMaxFreq

#e)
t = time(sres)
cosPart = cos(2 * pi * t * sMaxFreq)
sinPart = sin(2 * pi * t * sMaxFreq)
plot(sres)
cosSints =  ts(fitted(lm(sres~cosPart+sinPart)),start=1749,deltat=1/12)
lines(cosSints,col="magenta")


dev.off()
```

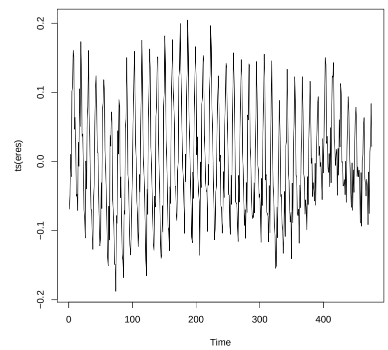
Plots



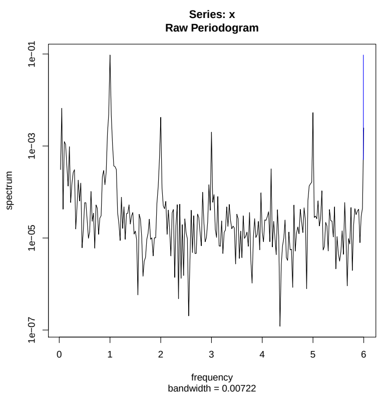
(a) Plot of Log, detrended electricity



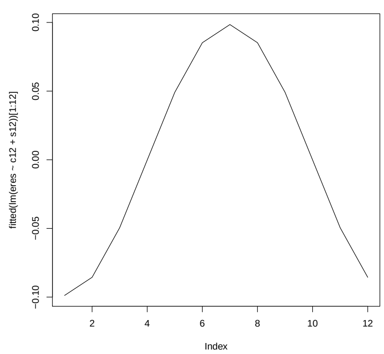
(b) Periodogram of electricity data



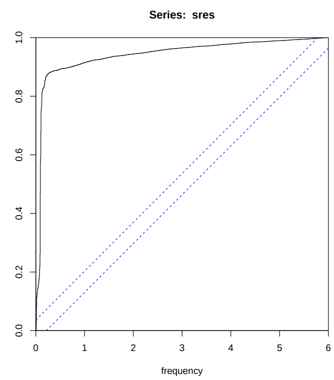
(c) Detrended electricity using regression



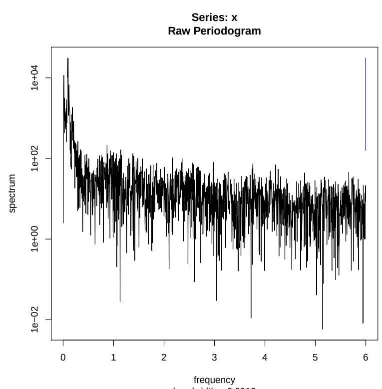
(d) Periodogram of detrended electricity



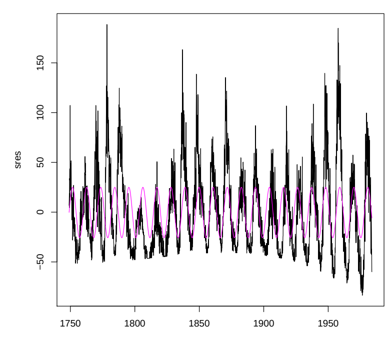
(e) Periodic Component of electricity



(f) Cumulative Periodigram of detrended sunspots



(g) Periodogram of detrended sunspots



(h) Sunspots with overlayed periodic component