Stat 5170: Assignment 10

due April 29, 5pm

1 R Tutorial

Let's look at the Australian labour data, "labour.dat", which was seen in previous assignments. We would like to estimate the spectral density of the data. The data contain the number of people in the civilian labor force in Australia from February 1978 to August 1995. We will only consider the first 144 observations since there was a recession after that observation.

```
labour<-scan("labour.dat")
labour<-labour[1:144]
tlabour<-diff(diff(log(labour)),12)
```

1. We'll start by creating the raw periodogram for the data, and computing confidence intervals for the power spectrum at certain frequencies.

   (a) Raw periodogram. We produce the raw periodogram using the following code:

   ```
   lab.per<-spec.pgram(tlabour,taper=0, log="no")
   ```

   (b) Confidence interval using the raw periodogram. Recall the approximate $100(1-\alpha)\%$ confidence interval for the power spectrum can be computed using

   $$\frac{2I(\omega_j)}{\chi^2_2(1-\alpha/2)} \leq f(\omega) \leq \frac{2I(\omega_j)}{\chi^2_2(\alpha/2)}. \quad (1)$$

   For this example, we compute this confidence interval for the power spectrum at ω_j^*, which we define as the fundamental frequency where the periodogram is maximised. Type the following to store the index of this fundamental frequency: `i<-which.max(lab.per$spec)`. To compute the confidence interval for the power spectrum at ω_j^*, type

   ```
   lower<-qchisq(0.975,2)
   upper<-qchisq(0.025,2)
   CI<-c(2*lab.per$spec[i]/lower,2*lab.per$spec[i]/upper)
   ```

(c) Overlaying confidence interval for the power using the raw periodogram. To overlay a computed confidence interval for the power spectrum at ω_j^*, use `abline()` as usual, e.g.:
`abline(h=CI[1], lty=2, col="blue")`. What does this say about the power spectrum at ω_j^*?

2. One of the things we noted in lecture is that the raw periodogram does not estimate the power spectrum well, especially when the estimated power is large. This is due to the high variability of the estimates. To reduce the variance, we consider smoothing the periodogram. First, we look at the Daniell kernel (symmetric moving average), which is defined as

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

where $L = 2m + 1$. To apply Daniell kernel smoothing, use:

```
k<-kernel("daniell",3)
lab.smooth<-spec.pgram(tlabour, k, taper=0, log="no")
```

The value `3` in `kernel()` is the value of of m as defined in (2), so $L = 7$. To confirm, type `lab.smooth$kernel` to check the weights. How many peaks we do have now?

3. A generalization of smoothing can be expressed as

$$\widehat{f}(\omega) = \sum_{k=-m}^{m} h_k I(\omega_j + k/n) \tag{3}$$

where $h_{-k} = h_k > 0$ and $\sum_{k=-m}^{m} h_k = 1$. Let $L_h = \left(\sum_{k=-m}^{m} h_k^2\right)^{-1}$. The approximate $(1-\alpha) \times 100\%$ confidence interval for the power spectrum is

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

(a) Modified Daniell kernel. The modified Daniell kernel can be used by typing the following:

```
k.mod<-kernel("modified.daniell",c(2,2))
lab.mod<-spec.pgram(tlabour, k.mod, taper=0, log="no")
```

The value `c(2,2)` in `kernel()` means we apply $m = 2$ for the modified Daniell kernel twice.

(b) The confidence intervals can be found by:

```
i<-which.max(lab.mod$spec)
Lh<-1/sum(k.mod[-k.mod$m:k.mod$m]^2)
lower.mod<-qchisq(0.95,2*Lh)
CI.mod<-2*Lh*lab.mod$spec[i]/lower.mod
abline(h=CI.mod, lty=2, col="blue")
```

What is the value of L_h for this particular modified Daniell kernel? The weights can always be recalled using `lab.mod$kernel`.

2 Assignment

1. (no R required) Recall from Unit 19 the autocovariance generating function for a causal and invertible ARMA process is

$$\gamma_B = \sigma_w^2 \psi(B^{-1})\psi(B)$$

where $\psi(B) = \frac{\theta(B)}{\phi(B)}$. Thus, the spectral density is

$$f(\omega) = \sigma_w^2 \left|\psi(e^{-2\pi i\omega})\right|^2.$$

   (a) For the following seasonal ARMA$(1,0)_{12}$ process

$$x_t = \Phi x_{t-12} + w_t,$$

   show that the spectral density is $f(\omega) = \frac{\sigma_w^2}{1-2\Phi\cos(24\pi\omega)+\Phi^2}$.

   (b) For the following multiplicative seasonal ARMA$(1,0)\times(1,0)_{12}$ process

$$(1-\Phi B^{12})(1-\phi B)x_t = w_t,$$

   show that the spectral density is $f(\omega) = \frac{\sigma_w^2}{(1-2\Phi\cos(24\pi\omega)+\Phi^2)(1-2\phi\cos(2\pi\omega)+\phi^2)}$.

2. (a) In this question, we will simulate a time series that follows an AR(1) model with $\phi_1 = 0.8$ (with $\sigma_w^2 = 1$). Use the `armasim()` function in `armasim.R` from Collab. We can use the following command to simulate 200 observations from this model:

```
source("armasim.R")
x<-armasim(c(0.8),0,1,200)
##arguments are phi vector, theta vector, sigmasq, no. of observations
```

   We know that there is a closed form representation for the spectral density of ARMA models. We want to see how much smoothing is necessary to obtain a good representation for the spectral density. Load the functions found in `A10function.R`. Use the following command to obtain a plot of the periodogram along with the TRUE spectral density:

```
source("A10function.R")
specplot(spec.pgram(x,taper=0, log="no"),phi=c(0.8))
```

Note that the red curve represents the spectral density, while the black lines represent the periodogram, which is an estimate of the spectral density based on our simulated sample. Next, introduce smoothing into the periodogram. For example,

```
k.mod<-kernel("modified.daniell",c(1,1))
specplot(spec.pgram(x, k.mod,taper=0, log="no"), phi=c(0.8))
```

produces the spectral density with the smoothed periodogram. The second argument in the `kernel()` function specifies the amount of smoothing you want to use. Adjust this value to see how much smoothing is required to obtain a fairly good estimate of the spectral density without over smoothing? Are there certain features of the true spectral density that are difficult to obtain? Be sure to turn in a plot of the unsmoothed periodogram as well as the smoothed periodogram you are satisfied with.

(b) Simulate a new time series with the same AR(1) model but with 400 data points. Repeat attempts to smooth the periodogram. Are there differences here with more data? Be sure to turn in a plot of the unsmoothed periodogram as well as the smoothed periodogram you are satisfied with.

3. Estimate the spectral density for the airline data, `airline.dat`. The data consist of monthly international airline passenger traffic from 1949 to 1956. Read the data using the following transformations:

```
airline<-scan("airline.dat")
tairline<-diff(diff(log(airline),12))
```

(a) Produce smoothed periodograms using the modified Daniell kernel, with $m = 1, 2, \cdots 8$. For what value of m do you think produces the "best" periodogram? R Hint: you might realize this can be produced quickly using a `for` loop.

(b) Describe how the shape/pattern of the smoothed periodogram changes as m goes from 1 to 8. What does this tell you about the effect of oversmoothing, and what is a practical implication of oversmoothing? What is a practical implication of undersmoothing?

4. Let $\{y_t\}$ be defined by

$$y_t = \frac{w_{t+1} - 2w_t + w_{t-1}}{3},$$

where w_t is white noise with variance σ_w^2.

(a) Show the frequency response function for this linear filter is $A(\omega) = \frac{2}{3}\left[\cos(2\pi\omega) - 1\right]$.

(b) Derive the spectral density $f_y(\omega)$. You may assume $f_w(\omega) = \sigma_w^2$.

(c) Using R, create a plot of the power transfer function and describe the effect of using this filter (i.e. what frequencies are retained/enhanced and what frequencies are dampened). Hint: This is very similar to assignment 9 where you created plots of the spectral density $f(\omega)$ against ω. Now you create a plot of the power transfer function, $|A(\omega)|^2$, against ω.