Stat 510, Lab 4

September 18, 2009

1 In Class

In this lab, we will learn how to use several R commands to smooth a time series and remove trends.

1. Start by creating a new folder "lab4" and download the file "gas.dat". This data file contains monthly heating oil prices from July 1973 to December 1987. Load the data into R using the command

   ```
   gas=scan("gas.dat")
   ```

   Plot the data.

2. We can take a moving average of the data by using the `filter()` command in R. Take the moving average over twelve observations using the command

   ```
   magas=filter(gas,rep(1,12)/12)
   ```

   Plot this moving average. Plot the residuals (`gas-magas`). Do these residuals appear stationary? Why or why not?

3. From now on we will be working with the logarithm of the data. Use `lgas=log(gas)` to do this. Let us now recalculate the moving average using the following commands. We also plot the residuals, the ACF of the residuals, and the periodogram of the residuals.

   ```
   malgas= filter(lgas,rep(1,12)/12)
   plot.ts(lgas)
   ```

```
lines(malgas)
resmalgas=lgas-malgas
resmalgas=resmalgas[!is.na(resmalgas)]
plot.ts(resmalgas)
acf(resmalgas)
spec.pgram(resmalgas, taper=0, log="no")
```

Notice the fourth command, the filter command will register a "NA" when the window will not fit inside the data. Using the fifth command above deletes the "NA" entries. Do the residuals appear more reasonable?

4. Another method for smoothing the data is kernel smoothing. Use the following commands to perform kernel smoothing and to analyze the residuals.

```
ksmlgas= ksmooth(1:length(lgas),lgas,"normal",bandwidth=12)$y
plot.ts(lgas)
lines(ksmlgas)
resksmlgas=lgas-ksmlgas
plot.ts(resksmlgas)
acf(resksmlgas)
spec.pgram(resksmlgas, taper=0, log="no")
```

Adjust the bandwidth in the first command above to see how the curve changes. How does the ACF of the residuals change when using a bandwidth of 3 versus a bandwidth of 12? How does the fitted line change?

2 Homework

1. Verify that the following model is non-stationary:

$$x_t = \beta_0 + \beta_1 t + \beta_2 t^2 + w_t$$

where w_t is white noise. Now, verify that $\nabla^2 x_t$ is stationary.

2. Load the monthly temperature data for England from 1723 to 1970 using the following command

```
engtemp=scan("tpmon.dat",skip=1)
```

(a) Plot the data and create an ACF. Try doing this only on the first 300 observations.

(b) Fit the following model using `lm()`:

$$x_t = \beta_1 cos\left(2\pi\frac{1}{12}t\right) + \beta_2 sin\left(2\pi\frac{1}{12}t\right) + w_t$$

(You will need to "create" the variables for the covariates. It may be useful to know that there are R functions `sin()` and `cos`.)

(c) Plot the residuals of the above fit. Comment on these residuals. You may want to look at only a few hundred at a time. Are the residuals dependent? Are they stationary?

(d) Compare the periodograms of the original data and the residuals from the fit model.

3. Use the smoothing techniques introduced in class and above to estimate the trend in the global temperature data. (Example 2.1 from the textbook. The data can be found in `globtemp.dat`).

4. (No R required.) Show that the $MA(3)$ model is weakly stationary. You need to show that the mean is zero and the covariance function depends only on distance. This will be very similar to what was done in class for the $MA(2)$ model.