Time Series Analysis

Computer Aided Lab 1

This is an ungraded lab sheet. Graded work will appear under Assignments.

For this lab we will introduce modelling Time Series in `R`.
`R` is a statistical computing language available for free on Windows, Mac, and linux and can be downloaded from https://cran.r-project.org/. Once installed, either open `R` or `Rstudio` on your desktop / laptop. We will be making use of the TSA package associated with the Cryer and Chan book. You can install it by running

```
install.packages("TSA")
```

This is only done once. However you still need to load the package each time you run `R`, using

```
require(TSA)
```

1 Setting working directory and downloading data

Download the **lab1.csv** file from Brightspace and place it in your working directory. To view the current working directory run

```
getwd()
```

in the `R` console and to change it use

```
setwd()
```

2 Reading data into `R`

Read the CSV file for this lab into a dataframe:

```
lab1data = read.csv("lab1.csv", header=TRUE)
```

This will create a dataframe of 6 columns which can be accessed e.g. using

```
lab1data[,2]
```

for the 2nd column. Run

```
names(lab1data)
```

to check the column names.

3 Functions for Time Series

1. To **plot** a series:

   ```
   plot(lab1data[,1], type="l")
   ```

   the `"l"` option draws a line between all observations which is traditional in time series plots.

2. To **difference** a series and store it in a variable:

   ```
   series1diff = diff(lab1data[,1], lag=1)
   ```

   Run

   ```
   ?diff
   ```

   to understand what `diff` does.

3. To view the **autocorellation** and **partial autocorellation** plots, use the functions

   ```
   acf()
   ```

   and

   ```
   pacf()
   ```

   To view both plots together run:

   ```
   par(mfrow=c(2,1)) # par sets plotting options; here we set 2 plots
   acf(lab1data[,1])
   pacf(lab1data[,1])
   ```

Practice Exercises

Do try these before looking at the solutions on the next page!

Exercise 1. For Series 1 comment on the plot of the series, does it appear weakly stationary? Why?

Exercise 2. For Series 2 and 3, plot the series and comment on the shape of the `ACFs` and `PACFs` of each series.

Exercise 3. For Series 4, 5, and 6, plot the series and identify which Series appear to have a linear trend.

Exercise 4. Now individually difference each series once and check if it was difference stationary.

Exercise 5. Look again at Series 4: it could also be made stationary be removing a linear trend. Use the function lm() for this regression

```
tt=1:nrow(lab1data)
TC=lm(lab1data[,4] ~ tt)
```

to linearly regress the series on time and then remove this trend. What is the independent variable in this regression and what is the dependent variable? Is the series now stationary after removing the linear trend?

Exercise 6. Try running the `arima_game.R` script, then type

```
play_arima_game()
```

to see if you can correctly identify the ARIMA model used to generate each time series based on the plots.

Practice Exercises: Solutions

Exercise 1. For Series 1 comment on the plot of the series, does it appear weakly stationary? Why?

Yes, the ACF and PACF go quickly to zero. In fact it looks like a white noise.

Exercise 2. For Series 2 and 3, plot the series and comment on the shape of the `ACFs` and `PACFs` of each series.

Series 2 looks like it is also a white noise. Series 3 looks like AR(1).

Exercise 3. For Series 4, 5, and 6, plot the series and identify which Series appear to have a linear trend.

Series 4 looks like a linear trend. Series 5 looks integrated maybe with a trend also, Series 6 looks integrated.

Exercise 4. Now individually difference each series once and check if it was difference stationary.

diff(Series 4) looks like it is MA(1). diff(Series 5) also looks like it is MA(1). diff(Series 6) looks like it is MA(1) or AR(1) or perhaps ARMA(1,1). They are all difference stationary.

Exercise 5. Look again at Series 4: it could also be made stationary be removing a linear trend. Use the function lm() for this regression

```
tt=1:nrow(lab1data)
TC=lm(lab1data[,4] ~ tt)
```

to linearly regress the series on time and then remove this trend. What is the independent variable in this regression and what is the dependent variable? Is the series now stationary after removing the linear trend?

Yes, removing an estimated trend makes it stationary.

```
y=lab1data[,4]-fitted(TC)
plot(y,t='l')
acf(y)
pacf(y)
```