Homework 4 – Due Tuesday, March 7
STAT-GB.2302, STAT-UB.0018: Forecasting Time Series Data

Problem 1

Consider the AR(2) process $x_t = x_{t-1} - 5x_{t-2} + \varepsilon_t$. Determine whether the process is stationary.

Problem 2

Use the ACF and PACF to identify ARIMA(p, d, q) models for the Housing Starts series, the log of the GDP series, the first differences of the log of the GDP series, and the first differences of the log of the CPI series (commonly known as "inflation"). Give reasons for your choices of p, d, q, for each series. Do *not* try to estimate the parameters. Just select p, d, q. For the inflation series (first log of the CPI) determine if additional differencing results in over-differencing.

Problem 3

For the first difference of the log GDP series, use the method described in the handout for Chapter 3, Part IV, page 6 to estimate b in the invertible MA(1) model $x_t = \varepsilon_t + b\varepsilon_{t-1}$.

Problem 4

For the first difference of the log GDP series, use the Yule-Walker equation $r_1 = \hat{a}_1 r_0$ to estimate a_1 in the AR(1) model $x_t = a_1 x_{t-1} + \varepsilon_t$. Is your fitted model stationary?

Problem 5

(a) For the first difference of the log GDP series, use the two Yule-Walker equations

$$r_2 = \hat{a}_1 r_1 + \hat{a}_2 r_0,$$
$$r_1 = \hat{a}_1 r_0 + \hat{a}_2 r_1,$$

to estimate a_1 and a_2 in the AR(2) model $x_t = a_1 x_{t-1} + a_2 x_{t-2} + \varepsilon_t$.

(b) Prove that your fitted AR(2) model is stationary. (It must be stationary, since it can be proved in general that AR models estimated by solving the Yule-Walker equations are *always* stationary.)

(c) Use your fitted model to forecast the log GDP (*not* just the first difference of the log GDP, but the log GDP itself) for the first quarter of 2017. (This is a one-step-ahead forecast for log GDP, based on ARIMA$(2, 1, 0)$ model.)

New R commands used in this assignment

- `Acf`. Produce an ACF plot or compute autocorrelations. This is part of the `forecast` package. Note: this command is slightly different than the `acf` command, which includes the lag-0 autocorrelation in the plot. Examples:

```
Acf(x)              # produce an ACF plot for the time series x
Acf(x, lag.max=10)  # produce an ACF plot with lags up to 10
Acf(x, plot=FALSE)  # compute sample autocorrelations, don't plot
```

- `diff`. Difference a time series. If you want the resulting series to have the same length as the original series, you will need to pad the beginning of the result with `NA` values. This is important, for example, if you want to make a plot of the result using the dates from the original series. Examples:

```
y <- diff(x)                          # difference the x series
diff.log.gdp <- c(NA, diff(log.gdp))  # difference and pad with NA
```

- `library`. Load an R add-on package. You must first install the package using the "Tools ⇒ Install packages ..." command. Example:

```
# load the "forecast" package, which provides the Acf and Pacf commands:
library("forecast")
```

- `Pacf`. Produce a PACF plot or compute partial autocorrelations. Like the `Acf` command, this is part of the `forecast` package. Examples:

```
Pacf(x)              # produce a PACF plot for the time series x
Pacf(x, lag.max=10)  # produce a PACF plot with lags up to 10
Pacf(x, plot=FALSE)  # compute sample partial autocorrelations, don't plot
```