Time Series Econometrics
Jacek Suda

Homework 1

Due date: Wednesday, 17 April, 13:45

1. **Question 1**
   Consider an ARMA(1,2) model:
   $$y_t = \phi y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \qquad \varepsilon_t \sim iidN(0, \sigma^2).$$
   (a) What are the stationarity conditions for this process?
   (b) What are the invertibility conditions for this process?
   (c) Express the process in the Wold form and specify Wold form coefficients in terms of ϕ and θ_is.

2. **Question 2**
   For an AR(1) with $\phi = 0.5$, calculate and plot the true ACF and PACF and the IRF for $j = 0, 1, \ldots, 5$. Do the same for an AR(2) with $\phi_1 = 1.25$ and $\phi_2 = -0.75$. See Hamilton p. 111 on solving PACF given ACF.

3. **Question 3**
   Stata Exercise
   We will replicate what we did in class using the longer series of US GDP data. The updated `do` file is available on the class website.
   Go to the FRED Economic Data website operated by the St. Louis Fed. Download (from the https://fred.stlouisfed.org/series/GDPC1) the data for U.S. quarterly real GDP. (If you run into insurmountable problem use the gdpq.prn from the class website). While downloading the data you can choose whichever format is more convenient for you but commands below work for the CSV delimited format.
   (a) Load the data using the command `import delimited` *filename* where default name given by Fred is *GDPC1.csv*. Two imported series are named `date` and `gdpc1`. Note that `date` is a series containing strings rather than proper date format and we cannot use `tsset date` right away. To create the time/date indicator create new series similarly to how we did it in class.
   (b) Rename the series in levels *gdpc1* to *gdpq*. Plot the series in levels. Transform the series to logs (calling it *lgdpq*). Examine *lgdpq*. The resulting series should look more linear and, to some extent, more like it has a stable variance. Paste graph and comment.

(c) Create the First Differences Series and call it *dlgdpq*. The linearity of *lgdpq* should imply that *dlgdpq* has a fairly stable mean. Meanwhile, the log transformation should keep the variance of the series reasonably stable. Having said that, note that the variance of output has been lower since 1984. Paste and comment.

(d) Do the correlogram for *dlgdpq* and plot AC and PAC functions. Include the output. What ARMA model(s) would you consider for *dlgdpq* given this correlogram?

(e) Since it is not as clear as before what the appropriate model is, estimate all the possible ARMA(p,q) models for max(p)=2 and max(q)=2. (I.e., run ARMA models in Stata such as AR(1), MA(1), ARMA(1,1), ARMA(1,2), ARMA(2,2), etc...) For this exercise consider conditional ML estimator. For each estimation adjusted sample period to 1950:1-2018:1. Don't worry about pasting the output for each model. But, record the AIC and BIC numbers for each case and report. What model fits best according to the two criteria? Paste output for best model(s). Look at correlogram and plot AC and PAC functions. Comment on the output.

(f) Take two most preferred models and compare their forecasting performance using Diebold-Mariano statistics. Report which model does better job forecasting.

(g) Re-estimate your preferred model and collect residuals by typing the command `predict model_residuals, residuals` after the estimation command. Check the correlogram for *model_residuals*, plot AC and PAC functions, do the Q-test for white noise. Comment on the output.