Homework 5

Ma641 Time Series I

due by class time 6:15pm, Monday July 7, 2008

Please try to submit a hardcopy of the report in class. If you chose the elearn submission option elearn please convert the report to pdf format before submitting.

This assignment has an almost two weeks deadline. Please recall that the first Midterm is scheduled in class on July 14. It covers all the classical time series models from chapter 2 and all the testing procedures and methods we learned thus far.

1. Solve problem 2.2 on page 93.

2. For the next parts of this assignment we will refer to the attached 3 data files: *"SP500close1yearJune24.csv"*, *"SP500close2yearJune24.csv"*, and *"SP500close3yearJune24.csv"*. These files contain daily data on S&P500 for the periods Jan 2008-July 2008, Jan 2007 - July 2008 and Jan 2006 - July 2008. A great concern for time series estimation as you have discovered in the previous assignments is that one needs a large number of observations from the time series for a reliable estimation and forecast. On the other hand the data has to be stationary for a reliable estimation and in general it is believed that data stays stationary only for short periods of time.

   Here we will study the performance of various models and the effect of the data length on the estimation.

   Repeat each of the following for all three data sets.

(a) Work with continuously compounded returns. Pay attention to the date order in your data.

(b) First test the specific type of non-stationarity provided by the unit root models (i.e. perform a unit root test). Compare for all 3 data periods.

(c) Build an AR, an MA and a straight ARMA model after you remove non-stationarity (if needed) for each time period.

(d) Build a seasonal model for your data using whatever lag you feel approximates the data best.

(e) Using each of the models that you constructed forecast the next three observations. 24 June is a Tuesday so you are forecasting the daily adjusted closing data for the rest of the week. Use the three data sets separately when building the models. All in all you should have three categories of models. The models in each category should correspond to the respective time period length.

(f) At the time when you write this report look back at the adjusted closing price for 25, 26 and 27 July and compare with the forecasted values by each of the models that you constructed. Calculate the squared error of your forecast for each model and the absolute value of your forecast. What was the better model in retrospective?

(g) Was there a difference in the three time periods (i.e., is there a type of model that consistently outperformed the others for all the datasets under consideration)?

(h) Are the perceived differences between models due to chance or there is a model that performed significantly better? (Hint: Use the standard error of the forecast as given by the R output.)