Fall 2016 MSAN 604

Assignment 3

Due Date:
Submit a hard copy in person to me by **5:00pm on Tuesday, December 6th.**

The Problem:
For this assignment you will be modeling and forecasting monthly Chinese export figures (in 100 million USD). Available to you is an "export" time series beginning in January 1984 and ending in December 2008. Also available to you is an "import" time series over the same period that provides monthly Chinese import figures (in 100 million USD). The data of interest can be found in the `china.csv` file.

Your task, by a variety of methods, is to model and forecast the export time series. To evaluate the forecasts made by your models it is helpful to split the available data into "training" and "test" sets according to the 80–20 rule. In particular, partition the data in the following way:
- Training: January 1984 – December 2002
- Test: January 2003 – December 2008

Using this partition the predictive accuracy of each of your models may be determined. The modeling approaches you must consider are:
(a) Holt-Winters
(b) SARIMA
(c) SARIMA + exogenous variable (i.e., the "import" time series)
(d) VAR + seasonal indicator (where both "export" and "import" time series are treated as endogenous variables)

For each modeling approach, choose the model that you deem to perform the best, and use it to predict the export time series for January 2003 – December 2008. Please display these forecasts and their associated 95% prediction intervals in both a tabular and graphical format. Using the held out "test" set, compare your predictions to the observed data and compute the predictive root mean squared error (RMSE).

The Deliverable:
The deliverable for this assignment is a written report including a brief introduction, four main sections corresponding to each modeling approach, and a brief conclusion. Within each of the four main sections describe your model selection approach and indicate which model you choose as "optimal", with a justification for this choice. It is here that you will display, in both tabular and graphical format, your optimal model's forecasts, as well as the predictive RMSE. In the conclusion section be sure to make a recommendation for which model is to be preferred (on the basis of predictive RMSE).

Note that these reports should be clear, concise, and generally well written. You may write as if the reader is an expert in time series analysis.