Time Series Midterm Fall 2020

TakeHome Portion (20 points)

Consider the data on Github and also distributed during the midterm, ***MidtermFall2020TakeHome.csv***.

This should be completed and submitted back to the Midterm assignment by 11:59pm Saturday, October 17 at 11:59pm.

1. (2 pts) First let’s assume the data come from a stationary process. What is the most favored correlation structure (identify p and q) with respect to the AIC? Assume you only want to consider ARMA(p,q) models with p < 15 and q < 5. All you have to do here is identify p and q.
2. Next we would like to compare the usefulness of a stationary and non-stationary model in terms of ASE rather than the AIC. Be sure and include any code you use to answer these questions.
   1. (*8 pts) Find* and report the ASE (and show your code) for a horizon of 24 from the following two models:
   2. (1 pt) Given your answer in part a, which model would be most favorable for forecasting the next 24 observations with respect to the ASE?
   3. (8 pts) Next evaluate each model using a rolling window ASE of at least 10 windows. Report each of your 10 or more individual ASEs, mean and median of the ASEs and a histogram of the individual ASEs. Did you change your preferred model? Why or why not?
   4. (1 pts) Use your most preferred model to forecast the next 24 observations from the end of the dataset. Provide the numerical values of the forecasts as well as a plot with probability intervals.

BONUS (up to 3 points): Plotting the entire dataset with the forecasts makes the forecasts tough to see and assess. Come up with a good visualization that you would show your client that would clearly show them the forecasts of the last 24 observations superimposed on the last 24 observations (visualizing the ASE.) Note these are the forecasts used to find the ASE **not** the forecasts from the end of the realization.