Economics 104: Project 2
Winter 2024, UCLA
Due Date: February 15, 2024 by 5PM (PST)

For this project, you will work any with dataset you like, however, it must contain at least 2 predictors and two response variables (y_{1t} and y_{2t}) which you will aim to predict. The response variables need to exhibit some shared dynamics since we will also consider a VAR model fit. Your task will be to find the best fit model by following the steps outlined below.

1. Provide a descriptive analysis of your variables. This should include histograms and fitted distributions, correlation plot, boxplots, scatterplots, and statistical summaries (e.g., the five-number summary). All figures must include comments.

2. Show the `tsdisplay` plot for each variable and comment on the stationarity, ACF, and PACF results.

3. Fit two AR(p) models to each variable, and evaluate the model performance as follows:

   - Plot and comment on the ACF and PACF of the residuals.
   - Evaluate the training/testing performance by splitting the data into 2/3 training and 1/3 testing, and computing the MSE and/or RMSE for each subset. Comment on which model is better. Make sure to also look at AIC and/or BIC.
   - Compute and plot a 10-step-ahead forecast for each model.

4. For this question, you need to identify an appropriate predictor(s) for your two series. Fit two ARDL(p, q) models to each variable, and evaluate the model performance as follows:

   - Plot and comment on the ACF and PACF of the residuals.
   - Evaluate the training/testing performance by splitting the data into 2/3 training and 1/3 testing, and computing the MSE and/or RMSE for each subset. Comment on which model is better. Make sure to also look at AIC and/or BIC.
   - Compute and plot a 10-step-ahead forecast for each model.

5. Fit a VAR(p) model to your data (y_{1t} and y_{2t}), and evaluate the model performance as follows:

   - Plot the CCF and comment on the results.
   - Perform a Granger-Causality test, and discuss whether it is possible to identify any causality between the variables.
   - Plot the IRFs and comment on the plots.
   - Show the plot that includes the data, fitted values, ACF, and PACF all in one figure. Comment on the results.

- Evaluate the training/testing performance by splitting the data into 2/3 training and 1/3 testing, and computing the MSE and/or RMSE for each subset. Comment on which model is better (e.g., the one for y_{1t} or for y_{2t}). Make sure to also look at AIC and/or BIC.
- Compute and plot an n-step-ahead forecast for each model. You can choose the number of steps-ahead.
- Plot and discuss the FEVD plot.
- Plot and discuss the CUSUM plot.

6. Provide a short summary of your overall conclusions/findings, and discuss which model is your preferred one for each response variable.