Macroeconometrics - 2168 Joao B. Duarte
Nova SBE, Spring 2020

Problem Set 3

Cointegration and State Space Models

Problem 1 (40 pts) This exercise is on cointegration. There are two main packages to deal with cointegration in R: *vars* and *urca*. For an helpful reference for this problem check here. First open the *ps3_data_Q1_2020.xlsx* dataset. This data contains yearly time series of total revenue, total spending, and spending on health as share of GDP by the US government from 1970 to 2018.

a) (5 pts) Construct a data frame in R where you just have revenue and spending on health, where revenue is ordered last. Now put this data frame into the time series format using the ts() function.

b) (5 pts) Test for stationarity of all series. What do you conclude?

c) (10 pts) Test for cointegration following the Johansen methodology using urca library. Use both the trace and the eigen tests. What do you conclude? What is the cointegrating vector? Interpret it.

d) (10 pts) What are the coefficients of adjustment for health spending and revenue (1 for each variable, two in total)? Interpret them.

e) (10 pts) Transform your VECM into the VAR representation using the vec2var() function. Get the IRFs and plot them for the next 30 steps (30 years). Interpret them. Then estimate a VAR using the same dataset. In the same figure, plot the response of revenue to a shock in health spending of the VECM and of the VAR. What do you conclude? Given the current COVID-19 crisis, what are your predictions in for future tax revenue if you extrapolate your IRFs findings?

Problem 2 (60 pts) In this exercise you will estimate the unobservable level of inflation from 4 monthly observable price indexes. Use quantmod library to import the 4 time series: “CPIAUCSL”, “CPILFESL”, “PCEPI”, “PCEPILFE”.

a) (5 pts) Create a data frame in time series format with all variables starting in the first month of 1959. Transform the data to the first differences of the log. Plot all he resulting inflation time series.

b) (15 pts) Load the library dlm. Apply the Kalman Filter to CPI only. You can use $V = 0.1$, $W = 0.2$, $C_0 = 0.5$, $m0 = 0$. Assume true inflation to follow a random walk. Set the appropriate FF and GG matrices. Plot CPI vs. the filtered level of inflation. Explain intuitively how the posterior is being updated.

c) (15 pts) Apply the Kalman Filter to all series. Assume again inflation follows a random walk, the same prior, the same variance of the error in the transition equation and that the variance of the measurement error is the same in all series. Appropriately set FF and the V matrix and get the filtered inflation rate. How does in compare to b)? How do you results depend on your prior? Choose different values for both m_0 and C_0 and show how different your filtered inflation would be. (tip: you can use a for loop)

d) (15 pts) Apply principal components analysis to your data with the 4 series (don't forget to scale and center the data). Can you explain all 4 series reasonably well with just one factor? Explain. Estimate 1 principal components and show its loadings. Plot all the predicted time series of all variables with the raw data in the same figure (tip: you need to scale the predicted value back to the original series, i.e. multiply by the standard deviation and add the mean).

e) (10 pts) Compare the true level of inflation estimate of PCA with that of Kalman Filter. That is, plot the Kalman Filter estimate of true inflation with the predicted value of the principal component of CPI (need to scale back to original data units again). Analyze your findings.