ECO374: Forecasting and Time Series Econometrics
Department of Economics, University of Toronto
Problem Set 2
Winter 2024

This Problem Set is based on the software file *Case Study 2.*

1. Load data on GDP, unemployment, target rate, CPI, from the file *Macro_data_us.csv* posted on Quercus. The start of the data series should be adjusted to 2012-01-01.
2. Using the Johansen test, obtain the number *r* of cointegration links.
3. Perform Time-series validation for the following models, each with lags 1 and 2:
    a. VAR model
    b. TVAR model with one threshold
    c. VEC model
4. Write a brief comment (1 sentence) stating the model that yields the smallest time series validation MSE for each variable.

Submission details:

Submit one pdf file with your code, Markdown text, and output. The file should be uploaded to Quercus by the due date and time. There is 10% grade penalty for each new day of late submission.

ECO374 Problem Set 2 Submission (2)

by Filip Zaleski

2024-03-22

Load necessary packages

```
options(repos = c(CRAN="http://cran.rstudio.com"))
if (!require("quantmod")) install.packages("quantmod")
if (!require("xts")) install.packages("xts")
if (!require("ggplot2")) install.packages("ggplot2")
if (!require("ggpubr")) install.packages("ggpubr")
if (!require("tseries")) install.packages("tseries")
if (!require("urca")) install.packages("urca")
if (!require("tsDyn")) install.packages("tsDyn")
if (!require("timetk")) install.packages("timetk")
if (!require("dplyr")) install.packages("dplyr")
library(quantmod); library(xts); library(ggplot2); library(ggpubr); library(tseries)
library(urca); library(tsDyn); library(timetk); library(dplyr)
```

Question 1: Loading Macro Data for the US

The start of the Data Series is the 1st of January 2021. Variables: index (date), GDP, unemployment, Fed funds rate, and CPI.

```
table <- utils::read.csv(file="C:/Users/filip/OneDrive - University of Toronto/Courses/Third Year/Winter/EC
ind <- as.Date(table$Index, format="%Y-%m-%d")
table <- subset(table, select=-c(Index))
Macro_data_us <- xts(x=table, order.by=ind)
Macro_data_us <- na.omit(Macro_data_us)
Macro_data_us <- stats::window(Macro_data_us, start="2012-01-01")
vnames <- names(Macro_data_us)
dat <- as.matrix(na.omit(diff(Macro_data_us)))
```

Question 2: Evaluating Cointegration via the Johansen Test

```
cointegration <- urca::ca.jo(Macro_data_us, type="trace", ecdet="trend", spec="transitory")
sc <- summary(cointegration)
sc@test.name
```

```
## [1] "Johansen-Procedure"
```

```
cbind(sc@teststat, sc@cval)
```

```
##                      10pct  5pct  1pct
## r <= 3 |   7.004467 10.49 12.25 16.26
## r <= 2 |  31.463007 22.76 25.32 30.45
## r <= 1 |  77.722336 39.06 42.44 48.45
## r = 0  | 145.419343 59.14 62.99 70.05
```

Question 3: TS Validation for Various Models

Performed with lags 1 through 2.

```
for (lag in 1:2) {
  print("lag:")
  print(lag)

  for (m in 1:3) {

    TT <- nrow(dat)
    T1 <- floor(0.5*TT)
    step <- 12
    tseq <- seq(from=T1, to=TT, by=step)
    tseq <- tseq[-length(tseq)]
    MSE.t <- matrix(0,nrow=tseq[length(tseq)]+step-T1,ncol=length(vnames))
    colnames(MSE.t) <- vnames

    for (j in tseq) {

      # The VAR model
      if (m==1) {model <- tsDyn::lineVar(data=dat[1:j-1,], lag=lag, model="VAR", I="diff")
                 fcst <- stats::predict(model, n.ahead=step)}

      # The TVAR model with 1 threshold
      if (m==2) {model <- tsDyn::TVAR(data=dat[1:j-1,], lag=lag, model="TAR", nthresh=1, trace=F)
                 fcst <- stats::predict(model, n.ahead=step)}

      # The VEC model
      if (m==3) {model <- tsDyn::lineVar(data=dat[1:j-1,], lag=lag, r=3, model="VEC")
                 fcst <- stats::predict(model, n.ahead=step)}

      js <- j+step-1
      MSE.t[(j-T1+1):(js-T1+1),] <- (dat[j:js,]-fcst)^2
    }

    if (m==1) print("VAR")
```

```
    if (m==2) print("TVAR")
    if (m==3) print("VEC")

    MSE <- matrix(colMeans(MSE.t), nrow=1)
    colnames(MSE) <- vnames
    print(MSE)
    print(" ")
  }
}
```

```
## [1] "lag:"
## [1] 1
## [1] "VAR"
##          CPI.US   GDP.US Unemployment.US Fed.Funds.Rate
## [1,] 0.07780628 114348.7        1.774793      0.1144626
## [1] " "
## [1] "TVAR"
##          CPI.US   GDP.US Unemployment.US Fed.Funds.Rate
## [1,] 0.07655317 101062.3        1.720554     0.05340935
## [1] " "
## [1] "VEC"
##         CPI.US GDP.US Unemployment.US Fed.Funds.Rate
## [1,] 0.0695728 102464        1.761289      0.1129596
## [1] " "
## [1] "lag:"
## [1] 2
## [1] "VAR"
##          CPI.US   GDP.US Unemployment.US Fed.Funds.Rate
## [1,] 0.07901373 107724.9        1.770911     0.08981221
## [1] " "
## [1] "TVAR"
##          CPI.US   GDP.US Unemployment.US Fed.Funds.Rate
## [1,] 0.07811312 102145.4        1.727105     0.06459959
## [1] " "
## [1] "VEC"
##          CPI.US   GDP.US Unemployment.US Fed.Funds.Rate
## [1,] 0.06887072 103558.9        1.800245      0.1023302
## [1] " "
```

Question 4: Evaluating TS Validation

Consider the yellow highlights in the results from question 3 above to view the minimized MSEs for each variable. Among 1-lag models, the TVAR model minimizes the validation MSE for all but the CPI variable. The situation is similar in the 2-lag model version, where the TVAR model minimizes the validation MSE for three variables while the VEC model does so for only one. The TVAR thus seems the best fit for both specifications. However, while the TVAR is optimal in both lag specifications, the lag 1 TVAR model returns smaller time series validation MSEs than the lag 2 alternative. The lag 1 TVAR is thus seemingly our best model.