ECON90033 – QUANTITATIVE ANALYSIS OF FINANCE I

TUTORIAL 11

Download the t11e1 Excel data file from the subject website and save it to your computer or USB flash drive. Read this handout and complete the tutorial exercises before your tutorial class so that you can ask for help during the tutorial if necessary.

Vector Autoregression (cont.)

Exercise 1

The *t11e1.xlsx* file contains the annual short-term and long-term interest rates (*SR*, *LR*, %) in Australia from 1970 to 2021.

a) Launch *RStudio*, create a new project and script, and name both *t11e1*. Import the data set from the *t11e1.xlsx* file to *RStudio*, save it as *t11e1.RData*, and attach it to your *R* project.

Plot these variables and briefly comment on the figures.

Execute the following commands:

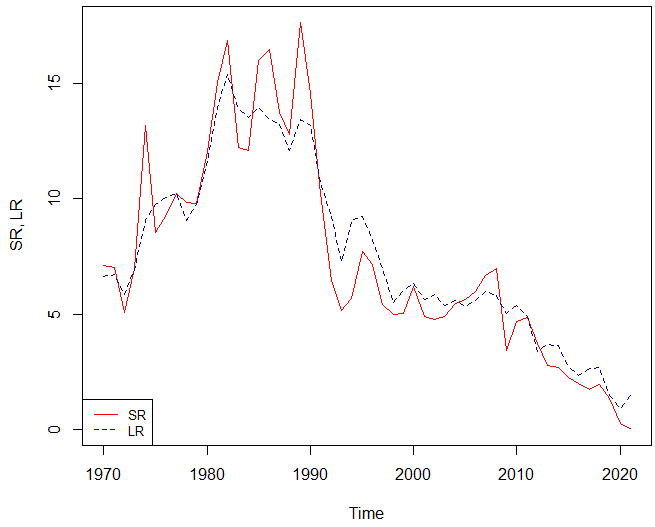
```
attach(t11e1)
SR = ts(SR, frequency = 1, start = 1970, end = 2021)
LR = ts(LR, frequency = 1, start = 1970, end = 2021)

plot.ts(SR, ylab = "SR, LR", col = "red")
lines(LR, col = "darkblue", lty = 2)
legend("bottomleft", legend = c("SR", "LR"), col = c("red", "blue"), lty = 1:2, cex = 0.8)
```

You should get the figure on the next page. Apparently, the two series are very closely aligned, and they fluctuate a lot but overall tend to decline during the sample period.

b) Perform the *ADF* and *KPSS* tests on the levels and first differences of *SR* and *LR*. What conclusions do you draw from these tests about the order of integration of these variables?

The time series plots suggest that for both variables the best is to use Model 3 in the tests on their levels and Model 2 on their first differences.



You can perform these tests like in Tutorials 8 and 9. Namely,

ADF tests:

```
library(urca)
adf.SR = ur.df(SR, type = "trend", selectlags = "BIC")
summary(adf.SR)
plot(adf.SR)

adf.LR = ur.df(LR, type = "trend", selectlags = "BIC")
summary(adf.LR)
plot(adf.LR)
```

The results are not shown here, but you can check on your screen that in both cases the unit root null hypothesis is maintained even at the 10% significance level.

On the correlograms, however, there are a few significant spikes. If you try to improve the test regressions by gradually increasing the lag length, you can see that for *SR* two lags are sufficient to eliminate serial correlation, while for *LR* even five lags are insufficient, but most importantly the unit root null hypotheses are still maintained.

If you repeat the tests on the first differences,

```
adf.DSR = ur.df(diff(SR), type = "drift", selectlags = "BIC")
summary(adf.DSR)
plot(adf.DSR)
```

```
adf.DLR = ur.df(diff(LR), type = "drift", selectlags = "BIC")
summary(adf.DLR)
plot(adf.DLR)
```

you find that the null hypotheses are now rejected, even at the 1% significance levels. This time one lag is sufficient for *SR*, but in the case of *LR* there is still sign of 5th order serial correlation.

KPSS tests:

```
kpss.SR = ur.kpss(SR, type = "tau", lags = "short")
summary(kpss.SR)
kpss_LR = ur.kpss(LR, type = "tau", lags = "short")
summary(kpss_LR)

kpss_DSR = ur.kpss(diff(SR), type = "mu", lags = "long")
summary(kpss_DSR)
kpss_DLR = ur.kpss(diff(LR), type = "mu")
summary(kpss_DLR)
```

Again, the printouts are not shown here, but the results are summarised in the table below:[1]

	Number of detected unit roots (α = 0.05)	
SR	***ADF***	***KPSS***
Level	0	0
Differenced	0	0
	I(0)	*I*(0)
LR	***ADF***	***KPSS***
Level	0	0
Differenced	0	0
	I(0)	*I*(0)

The results are unambiguous, both tests indicate that *SR* and *LR* are stationary, i.e., *I*(0).

[1] Take this opportunity to practice and try to verify this table.

c) Consider a *VAR* model with a constant of *p* and *m*, and determine the optimal lag length with the *VARselect()* function of the *vars* package.

VARselect(y, lag.max = , type =

where y is the list of endogenous variables and type is "const" or "trend" or "both" or "none", returns information criteria and final prediction error for VAR(p) models where the lag length (p) is increased gradually from 1 to k, and k is based on the same sample size.

Execute,

```
data = cbind(SR, LR)
library(vars)
VARselect(data, type = "const")
```

You should get this printout:

```
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     9      1      1      1

$criteria
                1         2         3        4         5         6
AIC(n) 0.5391064 0.5806811 0.6373563 0.809357 0.6990466 0.6786749
HQ(n)  0.6300956 0.7323298 0.8496645 1.082325 1.0326738 1.0729615
SC(n)  0.7873449 0.9944119 1.2165795 1.554073 1.6092545 1.7543751
FPE(n) 1.7153099 1.7913150 1.9033854 2.276981 2.0627437 2.0557779
                7         8         9        10
AIC(n) 0.6972732 0.5445278 0.4602149 0.6038779
HQ(n)  1.1522194 1.0601335 1.0364800 1.2408026
SC(n)  1.9384658 1.9512127 2.0323922 2.3415475
FPE(n) 2.1449807 1.9018050 1.8247635 2.2279959
```

HQ, *SC* and *FPE* take their smallest values for 1 lag, but *AIC* selects 10 lags. Estimate the more parsimonious *VAR*(1) model and test the residuals for first and second order autocorrelation with the Breusch-Godfrey *LM* test.

The

```
var1 = VAR(data, p = 1, type = "const")
serial.test(var1, lags.bg = 2, type = "BG")
```

commands return this printout:

```
        Breusch-Godfrey LM test

data:  Residuals of VAR object var1
Chi-squared = 9.4965, df = 8, p-value = 0.3022
```

Given the large *p*-value (0.3022), we maintain the null hypothesis that there is not first and second order residual serial correlation and accept the *VAR*(1) model.

The VAR(1) printout is displayed by the

summary(var1)

command:

```
VAR Estimation Results:
=========================
Endogenous variables: SR, LR
Deterministic variables: const
Sample size: 51
Log Likelihood: -155.585
Roots of the characteristic polynomial:
0.9876 0.3175
Call:
VAR(y = data, p = 1, type = "const")


Estimation results for equation SR:
===================================
SR = SR.l1 + LR.l1 + const

       Estimate Std. Error t value Pr(>|t|)
SR.l1   0.3676     0.2080   1.767  0.08356 .
LR.l1   0.6801     0.2464   2.760  0.00816 **
const  -0.5915     0.6791  -0.871  0.38808
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 1.989 on 48 degrees of freedom
Multiple R-Squared: 0.8219,     Adjusted R-squared: 0.8145
F-statistic: 110.7 on 2 and 48 DF,  p-value: < 2.2e-16
```

```
Estimation results for equation LR:
===================================
LR = SR.l1 + LR.l1 + const

       Estimate Std. Error t value Pr(>|t|)
SR.l1  0.04567    0.10966   0.416    0.679
LR.l1  0.93749    0.12988   7.218 3.41e-09 ***
const  0.03489    0.35797   0.097    0.923
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 1.049 on 48 degrees of freedom
Multiple R-Squared: 0.9303,     Adjusted R-squared: 0.9274
F-statistic: 320.5 on 2 and 48 DF,  p-value: < 2.2e-16



Covariance matrix of residuals:
      SR   LR
SR 3.958 1.62
LR 1.620 1.10

Correlation matrix of residuals:
       SR     LR
SR 1.0000 0.7765
LR 0.7765 1.0000
```

The first part of the printout shows, among others, the estimated characteristic roots. Since this is a bivariate system with 1 lag, there are 2 characteristic roots. The absolute values (lengths) of their point estimates are all smaller than one (i.e., inside the unit circle), indicating that this *VAR* is stable.

The second part of the printout shows the two estimated equations of *VAR*(1). Both equations are acceptable as they are strongly significant and have reasonable adjusted R^2 statistics. This is not unusual in *VAR* models, and it is not an issue because in *VAR* analyses the individual coefficients are of little importance.[2]

d) Use the estimated *VAR*(1) model to forecast *SR* and *LR* 1- 4 years ahead.

Execute

library(forecast)
var1_ea = forecast(var1, h = 4)

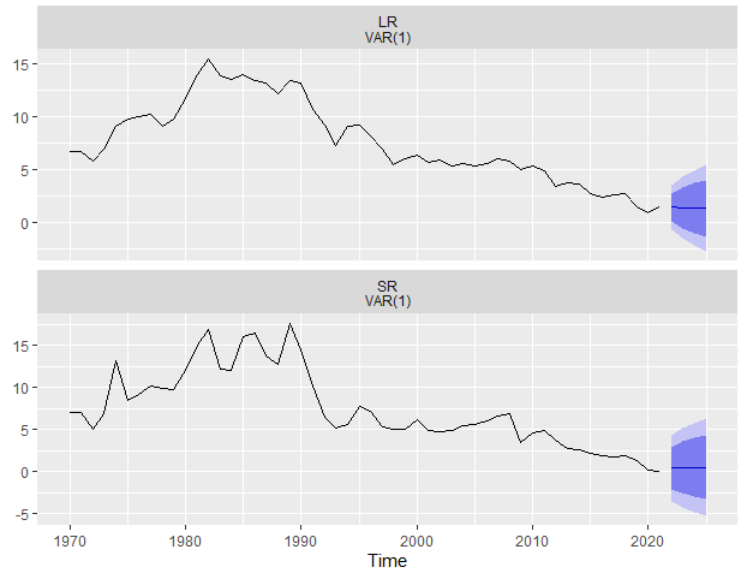
[2] Still, to save degrees of freedom, you might wish to eliminate the individually insignificant variables. Remember, however, that if you do so and as a result various equations have different sets of right-hand-side variables, the system becomes a near *VAR* and it should be estimated using seemingly unrelated regression (*SUR*).

print(var1_ea)
autoplot(var1_ea)

to get

```
SR
     Point Forecast     Lo 80    Hi 80     Lo 95    Hi 95
2022      0.4253596 -2.124218 2.974937 -3.473882 4.324602
2023      0.5330142 -2.556413 3.622441 -4.191856 5.257885
2024      0.5490090 -2.905246 4.003264 -4.733818 5.831836
2025      0.5361235 -3.221987 4.294234 -5.211411 6.283658

LR
     Point Forecast       Lo 80    Hi 80      Lo 95    Hi 95
2022       1.423667  0.07980634 2.767528 -0.6315906 3.478925
2023       1.388988 -0.51746862 3.295445 -1.5266858 4.304662
2024       1.361394 -0.97133100 3.694118 -2.2062008 4.928988
2025       1.336255 -1.34920087 4.021710 -2.7707951 5.443304
```



The *forecast()* function provides the point predictions and the 80% and 95% prediction bands. On the plots, the narrower dark blue areas illustrate the 80% prediction bands and the light blue areas the 95% prediction bands.

Granger Causality

In econometrics causality is a synonym for predictability. Considering two stationary variables, Y and Z, Z is said to be Granger causal to Y (denoted as $Z \rightarrow Y$) if and only if y_{t+1} can be predicted better when the information set includes z_t, z_{t-1}, …. , and Y is said to be Granger causal to Z (denoted as $Y \rightarrow Z$) if and only if z_{t+1} can be predicted better when the information set includes y_t, y_{t-1}, …. If Z is Granger causal to Y and Y is Granger causal to Z, there is a two-way (or feedback) Granger causal relationship (denoted as $Z \leftrightarrow Y$) between the two variables. The last possibility is that there is no Granger causal relationship at all between Y and Z (denoted as $Z \nrightarrow Y$ and $Y \nrightarrow Z$).

When we set up a *VAR* model, the left-hand side variables are considered endogenous variables. Whether they are indeed endogenous, i.e., are determined within the system, as opposed to exogenous variables that are fully determined outside the system, can be checked with joint Granger causality tests. Namely, a variable is an endogenous variable in the given system if the other variables jointly Granger cause it, and it is exogenous otherwise.

Granger causality can be tested with the general *F*-test or the Wald chi-square test on all lags of a variable (or several variables) jointly. Under the null hypothesis all these lags have zero coefficients, while under the alternative hypothesis some lag(s) has (have) non-zero coefficient(s).

Based on vector autoregression, Granger causality can be tested with the following function of the *bruceR* package:

granger_causality(model, test =)

where model is an estimated VAR object and test is "F" or "Chisq" (the default is both), performs Granger (predictive) causality tests between multivariate time series in VAR framework.

We return to Exercise 1 to illustrate Granger causality testing.

e) Use the *VAR*(1) model to test for Granger causality between *SR* and *LR* at the 5% significance level.

The

```
library(bruceR)
granger_causality(var1)
```

commands return the following printout:

```
Granger Causality Test (Multivariate)

F test and Wald χ² test based on VAR(1) model:
-------------------------------------------------------------------
                F df1 df2     p      Chisq df     p
-------------------------------------------------------------------
 ---------
SR <= LR     7.62   1  48  .008 **   7.62  1  .006 **
SR <= ALL    7.62   1  48  .008 **   7.62  1  .006 **
 ---------
LR <= SR     0.17   1  48  .679      0.17  1  .677
LR <= ALL    0.17   1  48  .679      0.17  1  .677
-------------------------------------------------------------------
```

The top part of the printout shows the Granger causality test results for the null hypothesis that *LR* does not cause *SR*, and the bottom part shows the Granger causality test results for the null hypothesis that *SR* does not cause *LR*. In both panels a pair of tests are reported. The first test in each pair is for causality from one variable to the other, and the second for causality from all other variables to the one on the left. In this case, however, there are only two variables, so in the top and bottom parts alike the two tests are the same.

This time the *F* and *Chisq* tests lead to the same conclusions at the conventional significance levels. Namely, at the 1% level both tests indicate that *LR* is causing *SR* (i.e., $LR \rightarrow SR$) but *SR* is not causing *LR* ($SR \nrightarrow LR$).

These results imply that in this bivariate *VAR* system of the short-term and long-term interest rates, *SR* proves to be an endogenous variables and *LR* appears to be exogenous.

In this example there is agreement between the *F* and *Chisq* versions of the Granger causality test. Occasionally, however, they might lead to contradicting, or at least ambiguous, conclusions and it might be important to decide which test to rely on. To make this call, remember that the *F* test assumes normality while the *Chisq* does not.

After having estimated a *VAR* model, multivariate Jarque-Bera tests and multivariate skewness and kurtosis tests for the residuals can be performed with the *normality.test* function. The Jarque-Bera test, in general, is based on the comparison of the skewness and kurtosis statistics to the skewness and kurtosis parameters of a multivariate normal distribution whose expected values and standard deviations are equal to the corresponding sample means and sample standard deviations.

This time the

normality.test(var1)

command returns the following printout:

```
$JB

        JB-Test (multivariate)

data:  Residuals of VAR object var1
Chi-squared = 22.256, df = 4, p-value = 0.0001782


$Skewness

        Skewness only (multivariate)

data:  Residuals of VAR object var1
Chi-squared = 9.662, df = 2, p-value = 0.007978


$Kurtosis

        Kurtosis only (multivariate)

data:  Residuals of VAR object var1
Chi-squared = 12.594, df = 2, p-value = 0.001842
```

The *p*-value of the multivariate *JB* test is0.0001782, so the null hypothesis of normality can be safely rejected.

The middle and bottom parts of this printout focus on the two crucial components of the *JB* test, skewness and kurtosis. The *p*-value for skewness is 0.007978 and that of kurtosis is 0.001842. Both of these *p*-values are small, so the rejection of normality can be attributed both to skewness and kurtosis of the residuals.

Save your *R* code and quit *RStudio*.

Linear Combinations of Stationary Variables versus Linear Combinations of Random Walks

It can be shown that any linear combination of stationary variables is also stationary. On the contrary, linear combinations of random walks are likely nonstationary, but some of them might be stationary. This latter possibility is illustrated by the following exercise.

Exercise 2

Consider the following three processes:

$$x_t = \beta_1 y_t + \beta_2 z_t + \varepsilon_{1t}$$
$$y_t = \beta_3 z_t + \varepsilon_{2t}$$
$$z_t = z_{t-1} + \varepsilon_{3t}$$

where the error terms are uncorrelated white noise (*WN*) processes.

It is clear from the definitions that z_t is a random walk and thus it is I(1).

Using the third equation, the second equation can be manipulated as follows:

$$\begin{aligned} y_t &= \beta_3 z_t + \varepsilon_{2t} \\ &= \beta_3 (z_{t-1} + \varepsilon_{3t}) + \varepsilon_{2t} \\ &= (\beta_3 z_{t-1} + \varepsilon_{2,t-1}) + \varepsilon_{2t} - \varepsilon_{2,t-1} + \beta_3 \varepsilon_{3t} \\ &= y_{t-1} + \varepsilon_{2t} - \varepsilon_{2,t-1} + \beta_3 \varepsilon_{3t} \end{aligned}$$

Since the error terms are *WN*, their linear combination on the right-hand side is stationary, and thus y_t is also I(1).

Using the second and the third equations, the first equation can be manipulated as follows:

$$\begin{aligned} x_t &= \beta_1 y_t + \beta_2 z_t + \varepsilon_{1t} \\ &= \beta_1 (y_{t-1} + \varepsilon_{2t} - \varepsilon_{2,t-1} + \beta_3 \varepsilon_{3t}) + \beta_2 (z_{t-1} + \varepsilon_{3t}) + \varepsilon_{1t} \\ &= (\beta_1 y_{t-1} + \beta_2 z_{t-1} + \varepsilon_{1,t-1}) + \varepsilon_{1t} - \varepsilon_{1,t-1} + \beta_1 (\varepsilon_{2t} - \varepsilon_{2,t-1}) + (\beta_1 \beta_3 + \beta_2) \varepsilon_{3t} \\ &= x_{t-1} + \varepsilon_{1t} - \varepsilon_{1,t-1} + \beta_1 (\varepsilon_{2t} - \varepsilon_{2,t-1}) + (\beta_1 \beta_3 + \beta_2) \varepsilon_{3t} \end{aligned}$$

Hence, x_t is also I(1).

However, by rearranging the first and second equations we get,

$$x_t - \beta_1 y_t - \beta_2 z_t = \varepsilon_{1t}$$
$$y_t - \beta_3 z_t = \varepsilon_{2t}$$

which shows that although x_t, y_t, z_t are I(1) individually, their two linear combinations above are equal to the ε_{1t} and ε_{2t} white noise, respectively, so they are stationary. The second stationary linear combination implies that *yt* and *zt* are cointegrated, CI(1,1); while the first implies that *xt*, *yt*, *zt* are also CI(1,1).

Using vectors, these linear combinations can be written as

$$\begin{bmatrix} 1 & -\beta_1 & -\beta_2 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \varepsilon_{1t} \quad ; \quad \begin{bmatrix} 0 & 1 & -\beta_3 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \varepsilon_{2t}$$

These linear combinations are linearly independent if the two row vectors are linearly independent, i.e., if the only solution of

$$a\begin{bmatrix} 1 & -\beta_1 & -\beta_2 \end{bmatrix} + b\begin{bmatrix} 0 & 1 & -\beta_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

is the trivial solution, $a = b = 0$.

This equation is actually a system of three equations:

$$a = 0$$
$$-a\beta_1 + b = 0 \quad \rightarrow \quad b = 0$$
$$-a\beta_2 - b\beta_3 = 0$$

and its only solution is indeed $a = b = 0$.

Hence, the two stationary linear combinations of x_t, y_t, z_t above are linearly independent.