

Multivariate time series

R examples

JEN-WEN LIN, PhD, CFA

November 22, 2017

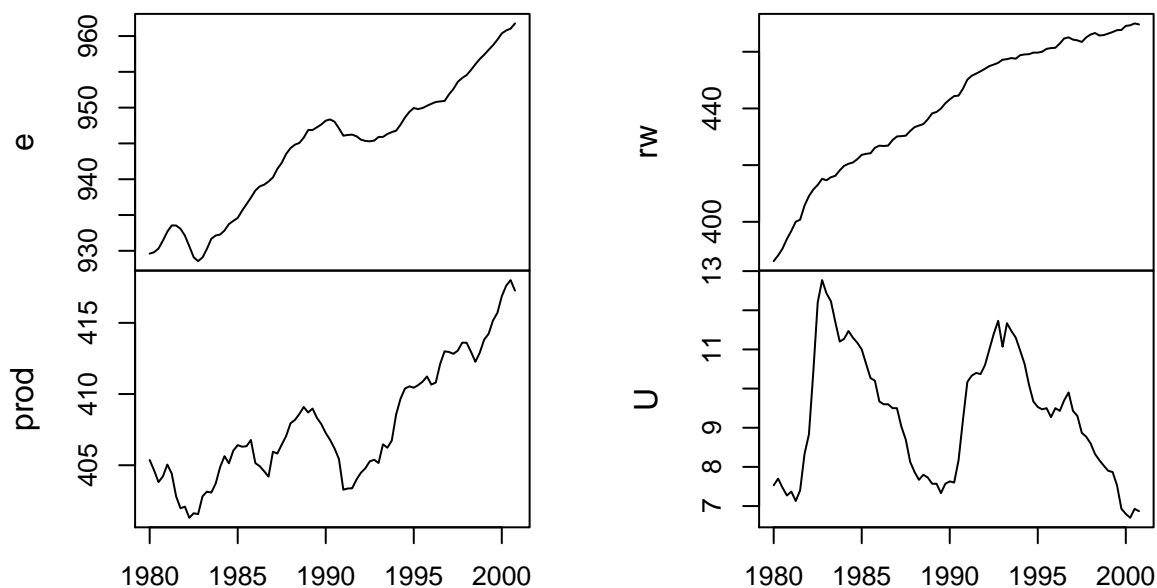
1 Data sets

1.1 Canadian macroeconomic time series

The original time series are published by the OECD. The sample range is from the 1980:Q1 until 2000:Q4. Four economic series are included in this data set, including **prod** is used as a measure of labour productivity; **e** is used for employment; **U** is the unemployment rate and **rw** assigns the real wage.

```
data("Canada")
par(mar=c(1,1,1,1),cex=0.6)
plot(Canada, nc=2, xlab="", main="Canada")
```

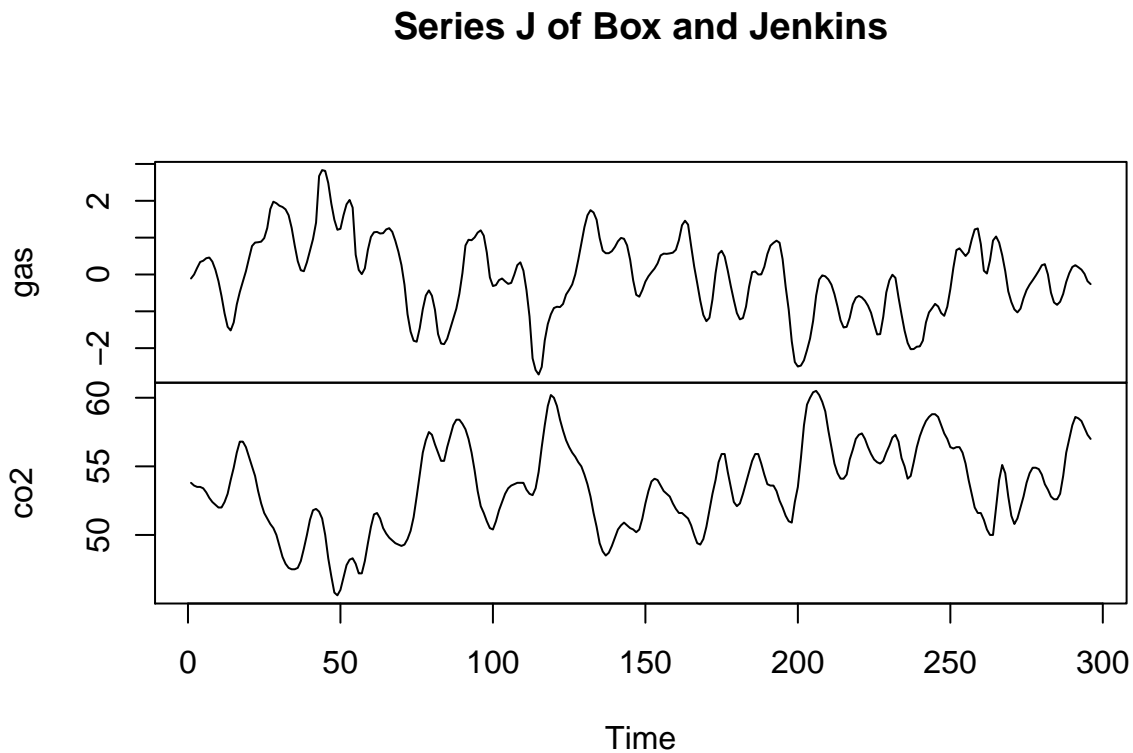
Canada



1.2 Gas furnace data set

The gas furnace data set contains the gas rate and the percentage CO2 in the gas. The source of the data is from Box and Jenkins' book on Time Series Analysis (series J) and available at <http://openmv.net/info/gas-furnace>.

```
dat<-read.csv("gasfur.csv",header = TRUE)
#Import data from working directory
gas<-ts(dat[,1],freq=1,start=1)
co2<-ts(dat[,2],freq=1,start=1)
SeriesJ<-cbind(gas,co2)
par(cex=0.6)
plot(SeriesJ, main="Series J of Box and Jenkins")
```

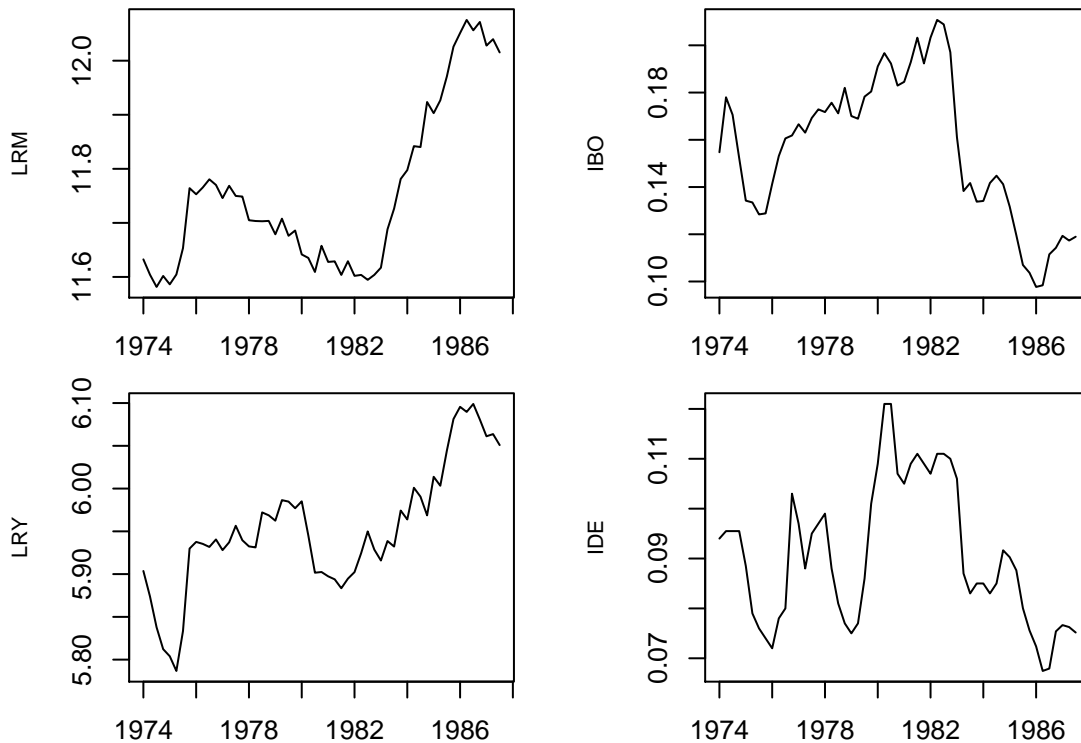


1.3 Estimating money demand function for Denmark

A data frame with 55 observations (source: <http://www.math.ku.dk/~sjo/data/data.html>) on the following 6 variables.

1. period: Time index from 1974:Q1 until 1987:Q3;
2. LRM: Logarithm of real money, M2;
3. LRY: Logarithm of real income;
4. LPY: Logarithm of price deflator;
5. IBO: Bond rate;
6. IDE: Bank deposit rate.

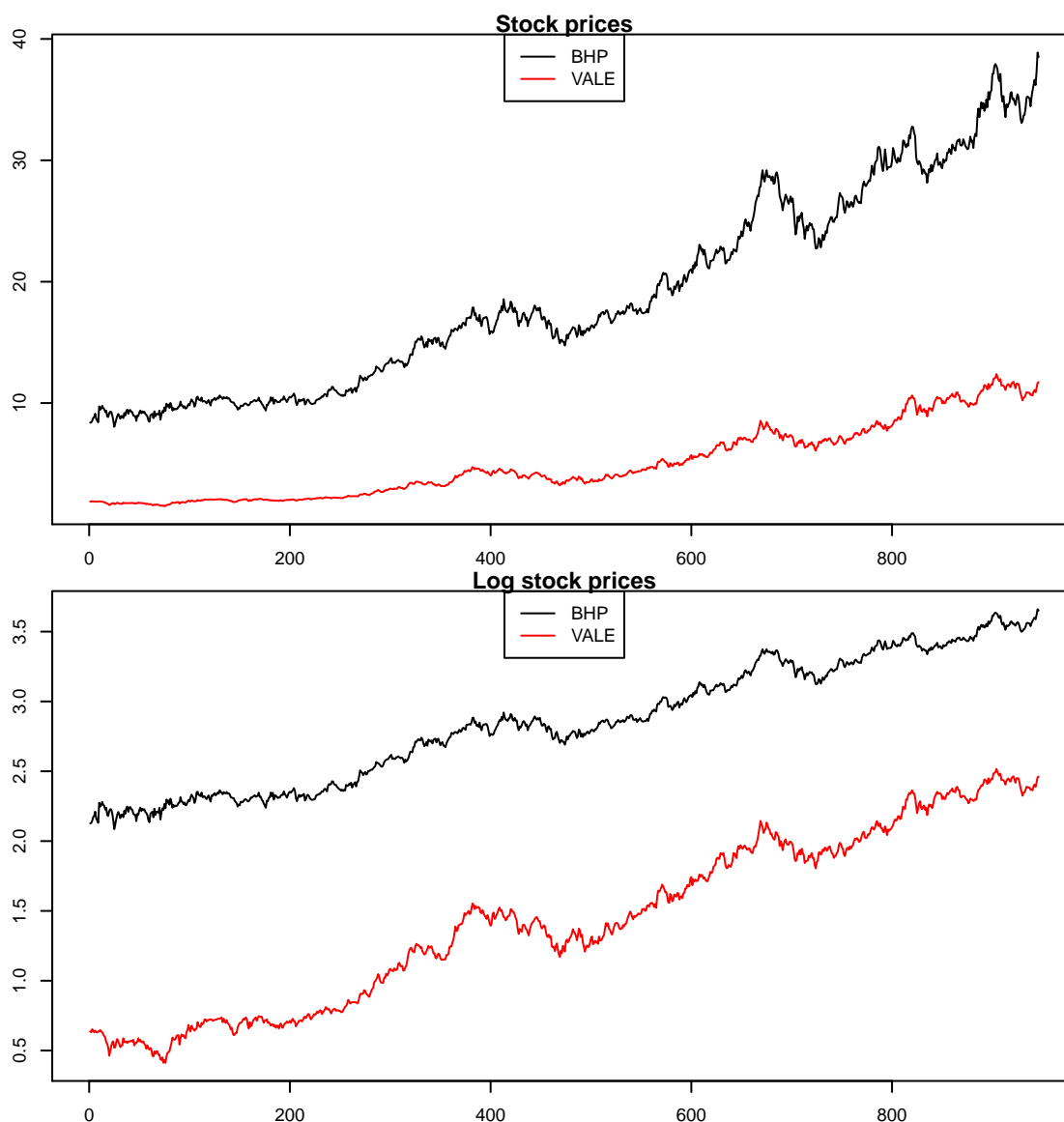
```
data("denmark")
sjd<-ts(denmark[,c("LRM","LRY","IBO","IDE")],freq=4,start=c(1974,1))
par(mar=c(2,4,1,2),mfrow=c(2,2),cex.lab=0.85)
ts.plot(sjd[, "LRM"],ylab="LRM")
ts.plot(sjd[, "IBO"],ylab="IBO")
ts.plot(sjd[, "LRY"],ylab="LRY")
ts.plot(sjd[, "IDE"],ylab="IDE")
```



1.4 Stock prices

The daily closing prices of the two stocks, the Billion Ltd. of Australia BHP and the Vale S.A. of Brazil VALE, were obtained from Yahoo Finance from July 1, 2002, to March 31, 2006. The time series plots on the original scale and the log-scale are plotted below.

```
dat.pairs<-read.csv("pairsTrading.csv", header = TRUE)[,4:5]
par(mfrow=c(2,1), mar=c(2,2,1,2), cex=0.6)
ts.plot(dat.pairs, main="Stock prices", col=c(1,2))
legend("top",c("BHP","VALE"), lty=1, col=c(1,2))
ts.plot(log(dat.pairs), main="Log stock prices", col=c(1,2))
legend("top",c("BHP","VALE"), lty=1, col=c(1,2))
```

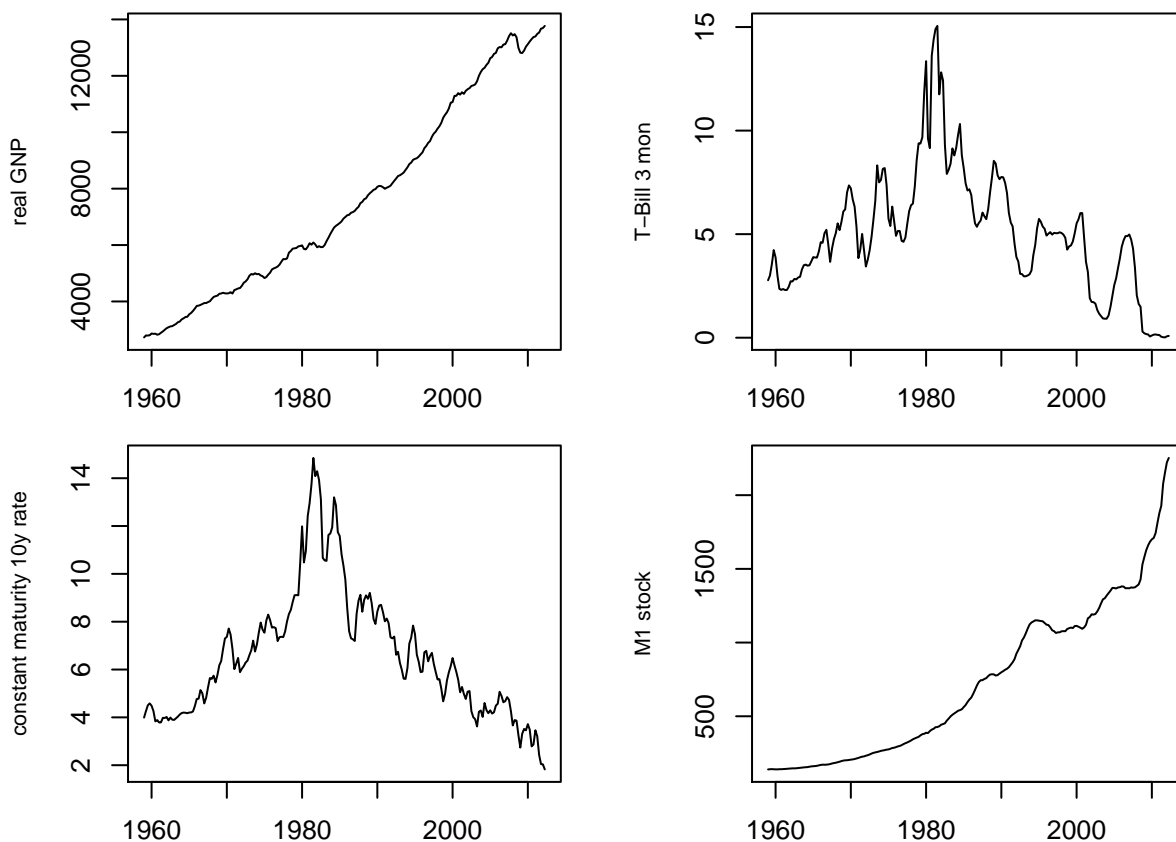


1.5 US macroeconomic data

Four quarterly macroeconomic time series from 1959:Q1 until 2012:Q2 are considered. The definitions are given below.

1. `rgnp`: US GNP measured in billions (2005 as deflator) and seasonally adjusted;
2. `tb3m`: US 3-month treasury bills (simple average of the monthly rates within a quarter);
3. `gs10`: US M1 money stock as the average of monthly data (measured in billions and seasonally adjusted);
4. `m1sk`: Rate of US 10-year constant maturity interest rate.

```
USmacro<-ts(read.csv("q-4macro.csv", header = TRUE)[-c(1:3)],
             freq=4,start=c(1959,1))
par(mar=c(2,4,1,2),mfrow=c(2,2), cex.lab=0.85)
ts.plot(USmacro[, "rgnp"], main="", ylab="real GNP")
ts.plot(USmacro[, "tb3m"], main="", ylab="T-Bill 3 mon")
ts.plot(USmacro[, "gs10"], main="", ylab="constant maturity 10y rate")
ts.plot(USmacro[, "m1sk"], main="", ylab="M1 stock")
```



2 Vector autoregression modeling

A vector autoregression of order n is given by

$$\mathbf{z}_t = \mathbf{c} + Trend + \sum_{i=1}^n \Phi_i \mathbf{z}_{t-i} + \mathbf{u}_t,$$

where \mathbf{c} denotes the $(k \times 1)$ intercept vector, *Trend* denotes linear time trends, \mathbf{z}_τ denotes the $(k \times 1)$ random vector at time τ , Φ_i denotes the $(k \times k)$ autoregressive coefficient matrix of order i , and \mathbf{u}_τ denotes the $(k \times 1)$ error vector.

2.1 Model/order selection criteria

The order of a vector autoregression may be selected using the information criteria. Some commonly used criteria for a VAR(n) model include

$$\begin{aligned} AIC(n) &= \ln \det(\hat{\Sigma}_u(n)) + \frac{2}{T}nk^2, \\ HQ(n) &= \ln \det(\hat{\Sigma}_u(n)) + \frac{2 \ln(\ln(T))}{T}nk^2, \\ SC(n) &= \ln \det(\hat{\Sigma}_u(n)) + \frac{\ln(T)}{T}nk^2, \\ FPE(n) &= \left(\frac{T + n^*}{T - n^*} \right)^k \det(\hat{\Sigma}_u(n)), \end{aligned}$$

with $\hat{\Sigma}_u(n) = T^{-1} \sum_{t=1}^T \hat{\mathbf{u}}_t \hat{\mathbf{u}}_t'$ and n^* is the total number of the parameters in each equation and n assigns the lag order.

```
vars::VARselect(Canada, lag.max=8, type="both")
```

```
## $selection
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      3      2      1      3
##
## $criteria
##              1              2              3              4              5
## AIC(n) -6.272579064 -6.636669705 -6.771176872 -6.634609210 -6.398132246
## HQ(n)  -5.978429449 -6.146420347 -6.084827770 -5.752160366 -5.319583658
## SC(n)  -5.536558009 -5.409967947 -5.053794411 -4.426546046 -3.699388378
## FPE(n)  0.001889842  0.001319462  0.001166019  0.001363175  0.001782055
##              6              7              8
## AIC(n) -6.307704843 -6.070727259 -6.06159685
## HQ(n)  -5.033056512 -4.599979185 -4.39474903
## SC(n)  -3.118280272 -2.390621985 -1.89081087
## FPE(n)  0.002044202  0.002768551  0.00306012
```

2.2 Model estimation

```
mod1<-vars::VAR(Canada,p=1,type="both")
summary(mod1)

##
## VAR Estimation Results:
## =====
## Endogenous variables: e, prod, rw, U
## Deterministic variables: both
## Sample size: 83
## Log Likelihood: -207.525
## Roots of the characteristic polynomial:
## 0.9504 0.9504 0.9045 0.7513
## Call:
## vars::VAR(y = Canada, p = 1, type = "both")
##
##
## Estimation results for equation e:
## =====
## e = e.l1 + prod.l1 + rw.l1 + U.l1 + const + trend
##
##           Estimate Std. Error t value Pr(>|t|)
## e.l1         1.23892    0.08632  14.353 < 2e-16 ***
## prod.l1       0.19465    0.03612   5.389 7.49e-07 ***
## rw.l1        -0.06776    0.02828  -2.396 0.018991 *
## U.l1          0.62301    0.16927   3.681 0.000430 ***
## const       -278.76121   75.18295  -3.708 0.000392 ***
## trend        -0.04066    0.01970  -2.064 0.042378 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.4701 on 77 degrees of freedom
## Multiple R-Squared: 0.9975, Adjusted R-squared: 0.9973
## F-statistic: 6088 on 5 and 77 DF, p-value: < 2.2e-16
##
##
## Estimation results for equation prod:
## =====
## prod = e.l1 + prod.l1 + rw.l1 + U.l1 + const + trend
##
##           Estimate Std. Error t value Pr(>|t|)
```

```

## e.l1      0.01291    0.12583    0.103    0.919
## prod.l1   0.96314    0.05266   18.291   <2e-16 ***
## rw.l1     -0.03909    0.04122   -0.948    0.346
## U.l1      0.21109    0.24674    0.855    0.395
## const    16.24341   109.59380    0.148    0.883
## trend     0.04613    0.02872    1.607    0.112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.6852 on 77 degrees of freedom
## Multiple R-Squared: 0.9754, Adjusted R-squared: 0.9738
## F-statistic: 610.6 on 5 and 77 DF, p-value: < 2.2e-16
##
##
## Estimation results for equation rw:
## =====
## rw = e.l1 + prod.l1 + rw.l1 + U.l1 + const + trend
##
##           Estimate Std. Error t value Pr(>|t|)
## e.l1      -0.05104    0.14143  -0.361 0.719153
## prod.l1   -0.22309    0.05918  -3.769 0.000319 ***
## rw.l1      0.94891    0.04633  20.480 < 2e-16 ***
## U.l1     -0.36864    0.27734  -1.329 0.187709
## const    163.02453   123.18318    1.323 0.189608
## trend      0.07142    0.03228    2.213 0.029868 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.7702 on 77 degrees of freedom
## Multiple R-Squared: 0.9989, Adjusted R-squared: 0.9988
## F-statistic: 1.412e+04 on 5 and 77 DF, p-value: < 2.2e-16
##
##
## Estimation results for equation U:
## =====
## U = e.l1 + prod.l1 + rw.l1 + U.l1 + const + trend
##
##           Estimate Std. Error t value Pr(>|t|)
## e.l1      -0.24844    0.06404  -3.880 0.000219 ***
## prod.l1   -0.12319    0.02680  -4.597 1.65e-05 ***
## rw.l1      0.06581    0.02098    3.137 0.002420 **
## U.l1      0.39158    0.12558    3.118 0.002559 **

```



```

## const    259.98201    55.77671    4.661 1.30e-05 ***
## trend      0.03452     0.01461    2.362 0.020712 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.3487 on 77 degrees of freedom
## Multiple R-Squared: 0.9557, Adjusted R-squared: 0.9528
## F-statistic: 331.9 on 5 and 77 DF, p-value: < 2.2e-16
##
##
##
## Covariance matrix of residuals:
##           e      prod      rw      U
## e      0.22096  0.067667 -0.082793 -0.13200
## prod   0.06767  0.469517  0.002141 -0.04128
## rw    -0.08279  0.002141  0.593174  0.06374
## U     -0.13200 -0.041280  0.063738  0.12161
##
## Correlation matrix of residuals:
##           e      prod      rw      U
## e      1.0000  0.210085 -0.228688 -0.8052
## prod   0.2101  1.000000  0.004057 -0.1728
## rw    -0.2287  0.004057  1.000000  0.2373
## U     -0.8052 -0.172753  0.237307  1.0000

```

2.3 Stationarity of vector autoregression

Consider a $k \times 1$ VAR(1) model (WLOG ignore the time trend and intercept)

$$\mathbf{z}_t = \Phi_1 \mathbf{z}_{t-1} + \mathbf{u}_t.$$

Using recursive substitution, we can express \mathbf{z}_t as

$$\mathbf{z}_t = \Phi_1^t \mathbf{z}_0 + \sum_{i=0}^{t-1} \Phi_1^i \mathbf{u}_{t-i}.$$

According to the above equation, we need $\lim_{n \rightarrow 0} \Phi_1^t = \mathbf{0}$ for \mathbf{z}_t to be independent of \mathbf{z}_0 . Mathematically, this is equivalent to the statement that all eigenvalues of Φ_1 are less than

one.¹ Since the eigenvalues of Φ_1 are solutions of

$$\lambda^k |\mathbf{I}_k - \Phi_1 \frac{1}{\lambda}| = 0.$$

It is equivalent to solve

$$|\mathbf{I}_k - \Phi_1 B| = |\Phi(B)| = 0,$$

where $B = 1/\lambda$ and $\Phi(B) = \mathbf{I}_k - \Phi_1 B$. Consequently, the conditions for the stationarity/stability of a VAR(1) model require that the solutions of the determinant equation $|\Phi(B)| = 0$ be outside the unit circle.

```
Phi<-rbind(mod1$varresult$e$coef[1:4],
           mod1$varresult$prod$coef[1:4],
           mod1$varresult$U$coef[1:4],
           mod1$varresult$rw$coef[1:4])
colnames(Phi)<-c("e", "prod", "U", "rw")
row.names(Phi)<-c("e", "prod", "U", "rw")
# Phi<-Acoef(mod1)
kable(Phi, digits = 3)
```

	e	prod	U	rw
e	1.239	0.195	-0.068	0.623
prod	0.013	0.963	-0.039	0.211
U	-0.248	-0.123	0.066	0.392
rw	-0.051	-0.223	0.949	-0.369

```
eigen(Phi, only.values = TRUE)$values
```

```
## [1] 1.1025869+0.00000i -0.8342208+0.00000i 0.8154310+0.04999i
## [4] 0.8154310-0.04999i
```

```
ev<-eigen(Phi, only.values = TRUE)$values
all(sqrt(Re(ev)^2+Im(ev)^2)<1)
```

```
## [1] FALSE
```

Exercise Discuss how to test the stationarity of a vector autoregression of order n using the idea of the companion matrix.

¹Recall that if $\{\lambda_1, \dots, \lambda_k\}$ are the eigenvalues of Φ_1 , then $\{\lambda_1^n, \dots, \lambda_k^n\}$ are the eigenvalues of Φ_1^n . Also, if all eigenvalues of a matrix are 0, then the matrix must be 0. Consequently, the condition for $\Phi_1^t \rightarrow 0$ as $t \rightarrow \infty$ is that all eigenvalues λ_j of Φ_1 must satisfy $\lambda_j^t \rightarrow 0$ as $t \rightarrow \infty$. This implies the absolute values of all eigenvalues λ_j of Φ_1 must be less than 1.

2.4 Model diagnostics

2.4.1 Portmanteau test

The Portmanteau statistic for testing the absence of up to the order h serially correlated disturbances in a stationary VAR(n) is defined as:

$$Q_h = T \sum_{j=1}^h \text{tr}(\hat{C}_j' \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}),$$

where $\hat{C}_i = \frac{1}{T} \sum_{t=i+1}^T \hat{\mathbf{u}}_t \hat{\mathbf{u}}_{t-i}'$. The test statistic is approximately distributed as $\chi^2(k^2(h-n))$. In R **vars** package, this test statistic is chosen by setting `type = PT.asymptotic`. For smaller sample sizes and/or values of h that are not sufficiently large, a corrected test statistic is computed as:

$$Q_h^* = T^2 \sum_{j=1}^h \frac{1}{T-j} \text{tr}(\hat{C}_j' \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}),$$

Similar, in R **vars** package, this test statistic can be accessed, if `type = PT.adjusted` is set.

```
serial.test(mod1, lags.pt = 16, type = "PT.asymptotic")
```

```
##
## Portmanteau Test (asymptotic)
##
## data: Residuals of VAR object mod1
## Chi-squared = 233.5, df = 240, p-value = 0.606
```

```
serial.test(mod1, lags.pt = 16, type = "PT.adjusted")
```

```
##
## Portmanteau Test (adjusted)
##
## data: Residuals of VAR object mod1
## Chi-squared = 256.88, df = 240, p-value = 0.2167
```

3 Granger causality

This section considers a general Granger causality test based on a vector autoregression of order n

$$\mathbf{z}_t = \sum_{i=1}^n \Phi_i \mathbf{z}_{t-i} + \mathbf{u}_t.$$

Let's first split the vector of endogenous variables \mathbf{z}_t into two subvectors $\mathbf{z}_{1,t}$ and $\mathbf{z}_{2,t}$ with dimensions $(k_1 \times 1)$ and $(k_2 \times 1)$ with $k = k_1 + k_2$, and rewrite the VAR(n) model as

$$\begin{bmatrix} \mathbf{z}_{1,t} \\ \mathbf{z}_{2,t} \end{bmatrix} = \sum_{i=1}^n \begin{bmatrix} \phi_{11}^{(i)}, \phi_{12}^{(i)} \\ \phi_{21}^{(i)}, \phi_{22}^{(i)} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{1,t-i} \\ \mathbf{z}_{2,t-i} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{1,t} \\ \mathbf{u}_{2,t} \end{bmatrix},$$

where

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{z}_{1,t} \\ \mathbf{z}_{2,t} \end{bmatrix}, \quad \Phi_i = \begin{bmatrix} \phi_{11}^{(i)}, \phi_{12}^{(i)} \\ \phi_{21}^{(i)}, \phi_{22}^{(i)} \end{bmatrix}, \quad \mathbf{u}_t = \begin{bmatrix} \mathbf{u}_{1,t} \\ \mathbf{u}_{2,t} \end{bmatrix}.$$

Using the above notation, we can test whether the subvector $\mathbf{z}_{1,t}$ does not Granger-cause $\mathbf{z}_{2,t}$. The null and alternative hypotheses are respectively given by

1. The null hypothesis: $\phi_{21}^{(i)} = \mathbf{0}$, $i = 1, 2, \dots, n$.
2. The alternative hypothesis: for any $\phi_{21}^{(i)} \neq \mathbf{0}$, $i = 1, 2, \dots, n$.

```
vars::causality(mod1, cause="e")$"Granger"

##
## Granger causality H0: e do not Granger-cause prod rw U
##
## data: VAR object mod1
## F-Test = 5.2493, df1 = 3, df2 = 308, p-value = 0.001517

vars::causality(vars::VAR(Canada, p=2, type="both"), cause="e")$"Granger"

##
## Granger causality H0: e do not Granger-cause prod rw U
##
## data: VAR object vars::VAR(Canada, p = 2, type = "both")
## F-Test = 5.7683, df1 = 6, df2 = 288, p-value = 1.092e-05
```

Exercise As discussed in class, we can also test Granger causality using the portmanteau test of Pierce and Haugh (1977)—a univariate approach. Describe and implement the univariate approach to test whether \mathbf{e} does not Granger-cause prod , rw , and U .

4 Transfer function noise model

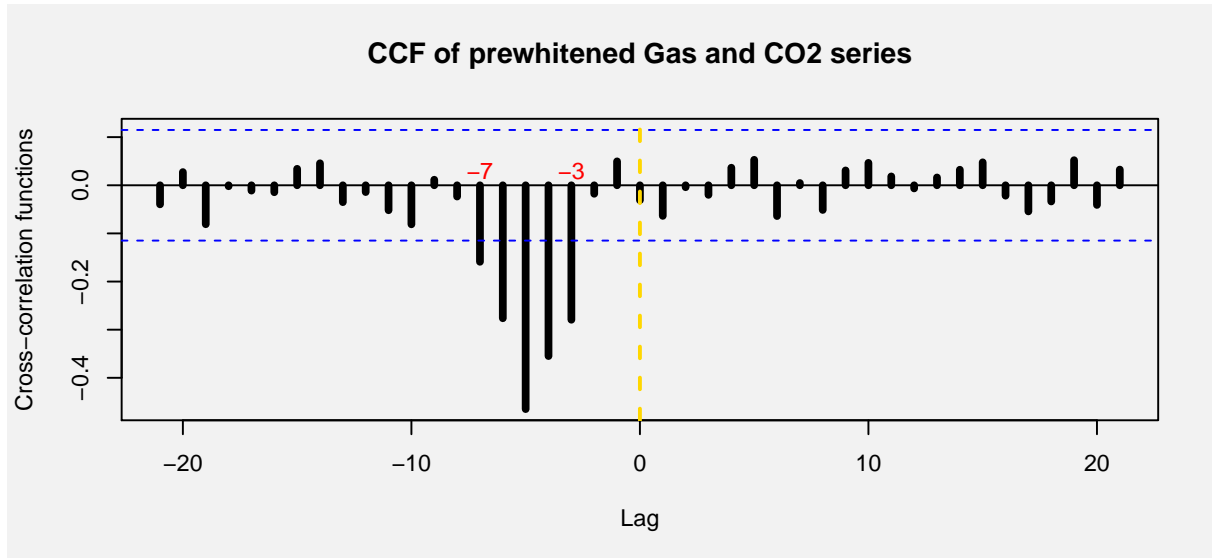
In this section, we shall use the gas and furnace data to demonstrate the techniques for implementing TFN models. Consider

$$co2_t = \alpha + \sum_{i=0}^s \nu_i \cdot gas_{t-i} + e_t, \quad (1)$$

where e_t is serially correlated.

4.1 Model identification using prewhitening

```
# Assume that the gas series follows an AR(5) model
mod.ar<-arima(gas,c(5,0,0)) # fit an AR(5) model to {gas} series
f1<-c(1,-mod.ar$coef[1:5]) # construct filter 1
co2f<-filter(co2,f1,method=c("convo"),sides=1) # construct filter 2
yf<-co2f[6:length(co2)] # loss degrees of freedom
xf<-mod.ar$res[6:length(co2)]
par(cex=0.75,bg="gray95")
ccf(xf, yf, lwd=4, ylab="Cross-correlation functions",
    main="CCF of prewhitened Gas and CO2 series")
abline(v=0, col="gold", lwd=2, lty="dashed")
text(-3, 0.03, "-3", col=2)
text(-7, 0.03, "-7", col=2)
```



According to the above CCF plot, we select x_{t-3}, \dots, x_{t-7} to be included in the TFN model. Mathematically, we have

$$co2_t = a + v_3 gas_{t-3} + v_4 gas_{t-4} + v_5 gas_{t-5} + v_6 gas_{t-6} + v_7 gas_{t-7} + e_t.$$

Note that the disturbance terms e_t are serially correlated so we have to fit a model for them. For simplicity, I choose an MA(6) model for the error term and fit the above TFN model using `r arima` function.²

4.2 Model estimation

```
# Arrange lagged explanatory variables
head(cbind(co2, lag(gas, 3), lag(gas, 4), lag(gas, 5), lag(gas, 6), lag(gas, 7)), 10)
```

```
## Time Series:
## Start = -6
## End = 3
## Frequency = 1
##      co2 lag(gas, 3) lag(gas, 4) lag(gas, 5) lag(gas, 6) lag(gas, 7)
## -6    NA          NA          NA          NA          NA      -0.109
## -5    NA          NA          NA          NA      -0.109      0.000
## -4    NA          NA          NA      -0.109      0.000      0.178
## -3    NA          NA      -0.109      0.000      0.178      0.339
## -2    NA      -0.109      0.000      0.178      0.339      0.373
## -1    NA      0.000      0.178      0.339      0.373      0.441
##  0    NA      0.178      0.339      0.373      0.441      0.461
##  1 53.8      0.339      0.373      0.441      0.461      0.348
##  2 53.6      0.373      0.441      0.461      0.348      0.127
##  3 53.5      0.441      0.461      0.348      0.127     -0.180
```

```
X<-ts.intersect(co2, lag(gas, 3), lag(gas, 4), lag(gas, 5), lag(gas, 6), lag(gas, 7))
head(X, 10)
```

```
## Time Series:
## Start = 1
## End = 10
## Frequency = 1
##      co2 lag(gas, 3) lag(gas, 4) lag(gas, 5) lag(gas, 6) lag(gas, 7)
##  1 53.8      0.339      0.373      0.441      0.461      0.348
##  2 53.6      0.373      0.441      0.461      0.348      0.127
##  3 53.5      0.441      0.461      0.348      0.127     -0.180
##  4 53.5      0.461      0.348      0.127     -0.180     -0.588
```

²Alternatively, we may use the `tfm` function in the *MTS* package.

```
## 5 53.4      0.348      0.127      -0.180      -0.588      -1.055
## 6 53.1      0.127      -0.180      -0.588      -1.055      -1.421
## 7 52.7     -0.180     -0.588     -1.055     -1.421     -1.520
## 8 52.4     -0.588     -1.055     -1.421     -1.520     -1.302
## 9 52.2     -1.055     -1.421     -1.520     -1.302     -0.814
## 10 52.0    -1.421    -1.520    -1.302    -0.814    -0.475

# fit transfer function noise model using arima function in r
m1 <- arima(X[,1], order = c(0, 0, 6), xreg = X[,-1])
m1

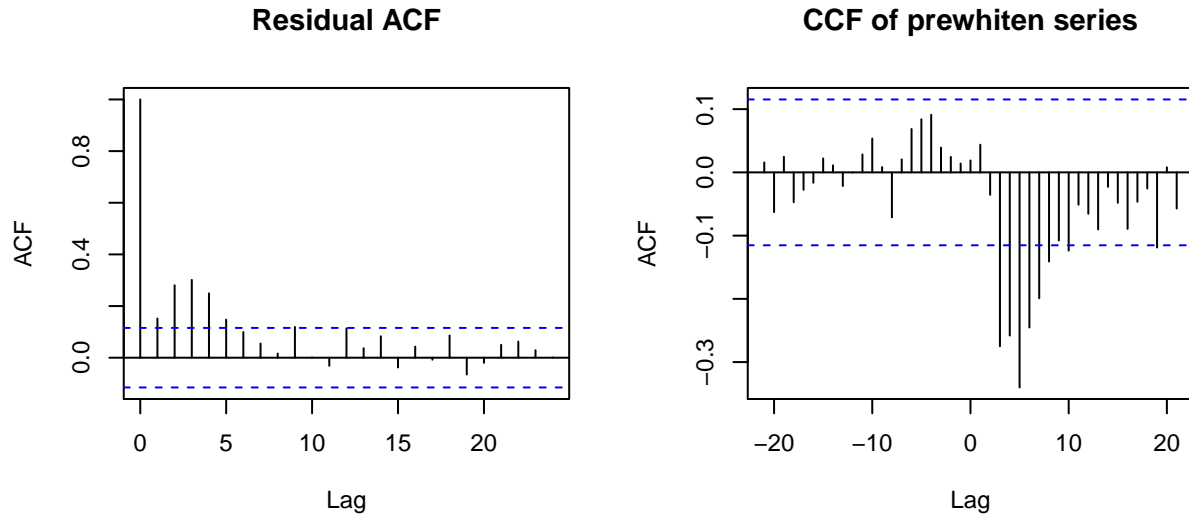
##
## Call:
## arima(x = X[, 1], order = c(0, 0, 6), xreg = X[, -1])
##
## Coefficients:
##          ma1      ma2      ma3      ma4      ma5      ma6  intercept
##          2.1622  3.0509  3.1622  2.4714  1.3609  0.4379   53.4004
## s.e.      0.0681  0.1455  0.1744  0.1356  0.0886  0.0493    0.2956
##          lag(gas, 3) lag(gas, 4) lag(gas, 5) lag(gas, 6) lag(gas, 7)
##          -0.1742    -0.2080    -0.0960    -0.1946     0.0312
## s.e.         0.1198     0.1255     0.1116     0.1270     0.1194
##
## sigma^2 estimated as 0.1378:  log likelihood = -127.05,  aic = 280.1
```

Exercise Fit the above model using the Box and Tiao transformation. See the supplement document, *Box and Tiao Transformation.pdf*, posted on portal.

4.3 Model adequacy

After the model is estimated, we check model adequacy of the fitted model. A naive way to do so is to examine the residual autocorrelation functions (RACF) and cross-correlation functions (CCF) to see if there are signs indicating the violation of assumptions. Finally, according to the plots below, the proposed model pass our visual inspection (95% confidence level).

```
par(mfrow=c(1,2), cex=0.75)
acf(m1$resid, main="Residual ACF")
ccf(m1$res,mod.ar$res, main="CCF of prewhiten series")
```



For formal assessments of model adequacy, we use portmanteau tests. First, we examine the residuals of the fitted TFN model.

```
# p-values of portmanteau tests over different lags
c(m20=Box.test(m1$res, type="Ljung", lag=20, fitdf=6)$p.value,
  m30=Box.test(m1$res, type="Ljung", lag=30, fitdf=6)$p.value,
  m40=Box.test(m1$res, type="Ljung", lag=40, fitdf=6)$p.value
)
```

```
##           m20           m30           m40
## 3.219647e-15 1.980638e-13 1.521927e-11
```

It is seen clearly that the p-values of portmanteau tests are above the 5% conventional significance level so we could conclude that the model is fitted adequately.

Exercise Note that the error term $\{e_t\}$ is assumed to be independent of the input series gas_t . As a result, $\{e_t\}$ will be independent of the *prewhitened* input (gas) times series, and we can check this assumption by examining whether $\{\hat{e}_t\}$ (estimates of error terms obtained from **3.2**) are independent/uncorrelated of the residuals from fitting gas_t . See the supplement document, *Diagnostic checks of TFN model.pdf*, posted on portal.

5 Modeling cointegration

5.1 Review cointegration

5.1.1 Cointegration

Definition (Cointegration) Engle and Granger (1987) provided the following definition of cointegration:

The components of the vector $\mathbf{z}_t = (z_{1t}, \dots, z_{kt})'$ are said to be *cointegrated of order d* and b , denoted by $\mathbf{z}_t \sim CI(d, b)$ if

- 1) All components of x_t are integrated of order d ;
- 2) There exists a vector $\beta = (\beta_1, \dots, \beta_k)'$ such that the linear combination

$$\beta' \mathbf{z}_t = \sum_{i=1}^k \beta_i z_{i,t}$$

is integrated of order $(d - b)$ where $b > 0$, and the vector β is called the *cointegrated vector*.

Exercise Consider the following bivariate process

$$\begin{aligned}\Delta z_{1t} &= \alpha_1(z_{1,t-1} - z_{2,t-1}) + u_{1,t} \\ \Delta z_{2t} &= \alpha_2(z_{1,t-1} - z_{2,t-1}) + u_{2,t}.\end{aligned}$$

Show that the process $y_t = z_{1t} - z_{2t}$ is autoregressive and stationary if $|1 + \alpha_1 - \alpha_2| < 1$.

5.1.2 Some examples of cointegration in economics/Finace.

1. Money demand

The demand for nominal money holdings should be (positively) related to the price level, real income, and the associated number of transactions but negatively related to the interest rate. Mathematically, we can express the money demand as

$$m_t = \beta_0 + \beta_1 p_t + \beta_2 y_t + \beta_3 r_t + \epsilon_t,$$

where m_t denotes demand for money, p_t denotes price level, y_t denotes real income, ϵ_t is the stationary disturbance term, and $\beta_i, i = 0, 1, 2, 3$ are parameters to be estimated. Note that all variables but the interest rate are expressed in logarithms.

2. Purchasing power parity (PPP)

In the short run, prices of similar products in varied markets might differ but arbiters will prevent the various prices from moving too far apart even if the prices are nonstationary. That is, in the long-run PPP requires that the linear combination $e_t + p_{ft} - p_t$ be stationary, where e_t denotes the log of the price of foreign exchange, and p_t and p_{ft} denote, respectively, the logs of domestic and foreign price levels.

3. Unbiased forward rate hypothesis

Under the efficient market hypothesis, the forward (or futures) price of an asset should equal to the expected value of the asset's spot price in the future. Mathematically, we have

$$E_t s_{t+1} = f_t,$$

where f_t denotes the log of the forward price (on an asset of interest) at time t , and s_t denotes the log of the spot price at time t . Additionally, if agents' expectations are rational, the forecast error for the spot rate at $t + 1$ will have a conditional mean equal to zero, i.e. $s_{t+1} - E_t s_{t+1} = \varepsilon_t$, where $E_t \varepsilon_{t+1} = 0$. Combining the above two equations yields

$$s_{t+1} = f_t + \varepsilon_t.$$

Since $\{s_t\}$ and $\{f_t\}$ are $I(1)$, the *unbiased forward rate hypothesis* necessitates that there be a linear combination of nonstationary $\{s_t\}$ and $\{f_t\}$ that is stationary.

4. Pairs trading

Pairs trading is a market-neutral trading strategy. Pairs trading based on statistical arbitrage and cointegration is one of the most popular methods by practitioners. These arbitragers believe the following statements: 1) if two stocks have similar characteristics, their stock prices must be more or less the same, and 2) if the prices differ, then it is likely that one of the stocks is overpriced and the other under-priced. As a result, pairs trading involves selling a higher priced stock and buying a lower priced stock simultaneously with the hope that the mispricing will correct itself in the future. The value of the portfolio containing aforesaid two stocks is called the spread. For pairs trading, the greater the spread, the larger the magnitude of mispricing and the greater the profit potential. As a result, traders usually buy/sell the spread portfolio when the portfolio value is higher/lower than a predetermined threshold and then hold the portfolio until the mispricing in the spread disappears.

5.2 Statistical model for cointegration

Currently there are three ways of modeling cointegration in the literature: 1) regression formulation, 2) autoregressive formulation, and 3) unobserved components formulation.³ In this section, we discuss the first two approaches to modeling cointegration for $I(1)$ processes.

³Consider

$$\mathbf{z}_t = \xi \eta' \sum_{i=1}^t \mathbf{u}_{1,i} + \mathbf{u}_{2,t},$$

where \mathbf{u}_{2t} is typically independent of \mathbf{u}_{1t} . Models for different cointegration ranks are nested and the smallest, for $\xi = \eta = \mathbf{0}$, corresponds to stationary process. Estimation is performed by the Kalman filter, and asymptotic theory of the rank tests has been worked out by Nyblom and Harvey (2000), see also Durbin and Koopman (2012).

5.2.1 The regression formulation

5.2.1.1 Ideas behind the regression formulation

Let $\mathbf{z}_t = (z'_{1t}, \mathbf{z}'_{2t})'$ of dimension $k = 1 + (k - 1)$ satisfy

$$z_{1t} = \beta' \mathbf{z}_{2t} + u_{1t},$$

$$\Delta \mathbf{z}_{2t} = \mathbf{u}_{2t},$$

where u_t is a stationary and invertible process. The above model implies 1) \mathbf{z}_{2t} is a random walk, and 2) The cointegration rank of \mathbf{z}_t is 1.⁴ The above formulation uses the super-consistent least squares method for estimation and the residual based tests for cointegration.⁵ This formulation implies that there is at most one cointegrating vector. Finally, a simulation study of is used to demonstrate the property of *superconsistency* below.

```
set.seed(1234)
NSim<-1000; n100<-100; rho<-0.5;
beta<-numeric(NSim); betaZ<-numeric(NSim)

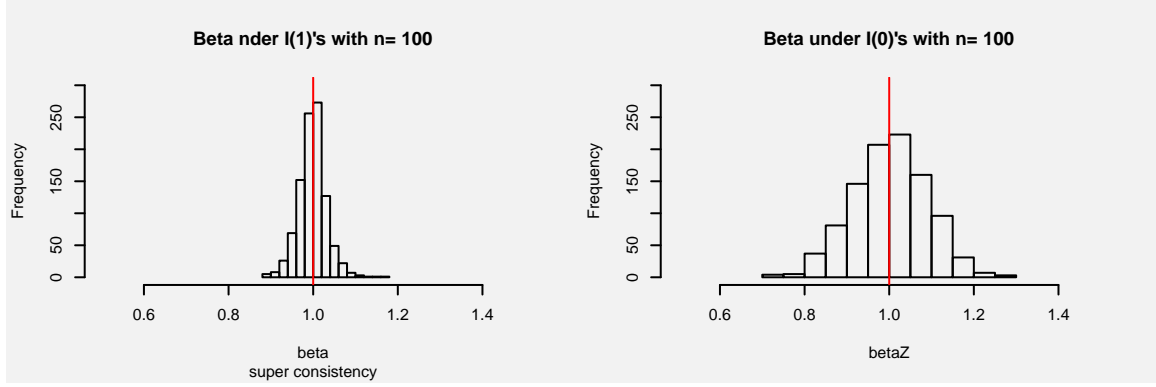
simAR<-function(phi=0.5,z){
  e<-numeric(length(z)); e[1]<-z[1]
  for(i in 2:length(z)) e[i]<-phi*e[i-1]+z[i]
  e
}

for(i in 1:NSim){
  z<-rnorm(n100)
  x1<-cumsum(z); x2<-x1+rnorm(n100)
  z1<-simAR(rho,z); z2<-z1+rnorm(n100)
  mod1<-lm(x2~x1); modZ<-lm(z2~z1)
  beta[i]<-mod1$coef[2]
  betaZ[i]<-modZ$coef[2]
}

par(mfrow=c(1,2), cex=0.5, bg="gray95")
hist(beta, ylim=c(0,300),xlim=c(0.5,1.5),
      main=paste("Beta nder I(1)'s with n=",n100))
abline(v=1,col=2)
title(sub = "super consistency")
hist(betaZ, ylim=c(0,300),xlim=c(0.5,1.5),
      main=paste("Beta under I(0)'s with n=", n100))
abline(v=1,col=2)
```

⁴That is, $z_{1t} - \beta' \mathbf{z}_{2t}$ gives one stationary linear combinations.

⁵Engle and Granger (1987) show the super-consistent estimator by Stock (1987) and the residual based tests for cointegration. Note that the least squares estimator $\hat{\beta}$ for $I(1)$ converges β at rate T instead of the usual rate $T^{\frac{1}{2}}$.



5.2.2 The autoregression formulation

5.2.2.1 Ideas behind the autoregression formulation

For simplicity, let's consider a naive case

$$\Delta \mathbf{z}_t = \alpha \beta' \mathbf{z}_{t-1} + \mathbf{u}_t, \quad (\star)$$

where \mathbf{u}_t are i.i.d. errors, and α and β are both $k \times r$ matrices. This formulation allows modeling 1) The long-run relations $\beta' \mathbf{z}$, and 2) The adjustment, or feedback coefficient α , toward the attractor set $\{\mathbf{z} : \beta \mathbf{z}\}$ defined by the long-run relations.

Note that this formulation implies that there are possible $0 \leq r \leq k$ cointegrating vectors. As a result, in this formulation, models for different cointegration ranks are nested and the smallest, for $\alpha = \beta = \mathbf{0}$, corresponds to k random walks. The rank can be tested using the likelihood ratio tests. Methods applied for this analysis are usually derived from the Gaussian process. See Johansen (1988, 1996) and Ahn and Reinsel (1990).

Exercise Use Equation (\star) to show that we need the absolute values of the eigenvalues of $\mathbf{I}_r + \beta' \alpha$ are less than 1 for $\beta' \mathbf{z}_t$ to be stationary.

5.2.2.2 Statistical model

Consider a vector autoregression of order n

$$\mathbf{z}_t = \Phi_1 \mathbf{z}_{t-1} + \dots + \Phi_n \mathbf{z}_{t-n} + \mathbf{u}_t.$$

The corresponding characteristic function is

$$\Phi(B) = \mathbf{I}_k - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_n B^n,$$

where

$$\Phi(1) = \mathbf{I}_k - \Phi_1 - \Phi_2 - \dots - \Phi_n.$$

Rearranging the above VAR(n) model, we have

$$\Delta \mathbf{z}_t = \Gamma_1 \Delta \mathbf{z}_{t-1} + \cdots + \Gamma_{n-1} \Delta \mathbf{z}_{t-n+1} + \Pi \mathbf{z}_{t-n} + \mathbf{u}_t,$$

where Γ_i matrices measure the cumulative long-run impacts and are given by

$$\Gamma_i = -(\mathbf{I}_k - \Phi_1 - \cdots - \Phi_i), \quad i = 1, \dots, n-1,$$

and finally

$$\Pi = -(\mathbf{I}_k - \Phi_1 - \cdots - \Phi_n) = -\Phi(1).$$

Note that the above model is a traditional first-order differenced VAR($n-1$) model except for the term $\Pi \mathbf{z}_{t-n}$ term.⁶ Johansen and Juselius (1990) shows how the coefficient matrix Π contains information about long-run relationships between the variables in the data vector. There are three possible cases:

1. $Rank(\Pi) = k$. The fact that Π has full rank implies that $\Phi(B)$ contains no unit root or $\Phi(B) \neq 0$, indicating that \mathbf{z}_t is stationary;
2. $Rank(\Pi) = 0$. In this case, the matrix Π is the null matrix and the above model corresponds to a traditional first-order differenced vector time series model;
3. $0 < Rank(\Pi) = r < k$. This case implies that there are $k \times r$ matrices α and β such that $\Pi = \alpha\beta'$,⁷ and the vector series $\mathbf{w}_t = \beta' \mathbf{z}_t$ is an $I(0)$ process, which is referred to as cointegrating series, and α denotes the impact of the cointegrating series on $\Delta \mathbf{z}_t$.

Remark: An important feature of “reduced rank” matrices like α and β is that they have orthogonal complements, denoted by α_\perp and β_\perp . α_\perp and β_\perp are $k \times (k-r)$ matrices orthogonal to α and β (so $\alpha'_\perp \alpha = \mathbf{0}$ and $\beta'_\perp \beta = \mathbf{0}$), where the $k \times k$ matrices $\alpha'_\perp \alpha$ and $\beta'_\perp \beta$ both have full rank r . These orthogonal matrices play a crucial role in understanding the relationship between cointegration and “common trends”. Specifically, $\beta'_\perp \mathbf{z}_t$ has $k-r$ unit roots and can be considered as the $k-r$ common trends of \mathbf{z} .

⁶There exists an alternative VECM specification as follows.

$$\Delta \mathbf{z}_t = \tilde{\Gamma}_1 \Delta \mathbf{z}_{t-1} + \cdots + \tilde{\Gamma}_{n-1} \Delta \mathbf{z}_{t-n+1} + \Pi \mathbf{z}_{t-1} + \mathbf{u}_t,$$

where

$$\tilde{\Gamma}_i = -(\Phi_{i+1} + \cdots + \Phi_n), \quad i = 1, \dots, n-1$$

and

$$\Pi = -(\mathbf{I}_k - \Phi_1 - \cdots - \Phi_n).$$

Note that the Π matrix is the same as in the first specification. However, the Γ_i matrices now differ, in the sense that they measure transitory effects. Please note that inferences drawn on Π will be the same, regardless which specification is chosen and that the explanatory power is the same, too.

⁷Note that this decomposition is not unique. Specifically, for any $r \times r$ orthogonal matrix \mathbf{P} such that $\mathbf{P}\mathbf{P}' = \mathbf{I}_r$, we have

$$\alpha\beta' = \alpha\mathbf{P}\mathbf{P}'\beta' = (\alpha\mathbf{P})(\beta\mathbf{P})'.$$

Thus, $\alpha\mathbf{P}$ and $\beta\mathbf{P}$ are also of rank r and may serve as another decomposition of Π . Hence, some additional normalization is needed to uniquely identify α and β .

5.3 Likelihood ratio tests for cointegration

5.3.1 Review of likelihood ratio tests

Consider a random vector \mathbf{z}_t that follows a multivariate normal distribution with mean zero and positive-definite covariance matrix Σ_z . Suppose that $\mathbf{z}' = (\mathbf{x}', \mathbf{y}')$, where the dimensions of \mathbf{x} and \mathbf{y} are p and q , respectively. Without loss of generality, assume that $q \leq p$ and

$$\Sigma_z = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}.$$

Consider the null hypothesis that \mathbf{x} and \mathbf{y} are uncorrelated, or $H_0 : \Sigma_{xy} = \mathbf{0}$ versus $H_a : \Sigma_{xy} \neq \mathbf{0}$. This is equivalent to testing the coefficient matrix Π being $\mathbf{0}$ in the multivariate regression

$$\mathbf{x}_i = \Pi \mathbf{y}_i + \mathbf{u}_i$$

where \mathbf{u}_i denotes the error term.

Under $H_0 : \Sigma_{xy} = \mathbf{0}$, the maximum likelihood estimate of Σ_z is

$$\hat{\Sigma}_0 = \begin{bmatrix} \hat{\Sigma}_{xx} & \mathbf{0} \\ \mathbf{0} & \hat{\Sigma}_{yy} \end{bmatrix},$$

where $\hat{\Sigma}_{xx} = T^{-1} \sum_{i=1}^T \mathbf{x}_i \mathbf{x}_i'$ and $\hat{\Sigma}_{yy} = T^{-1} \sum_{i=1}^T \mathbf{y}_i \mathbf{y}_i'$. The maximized likelihood function under H_0 is

$$l_0 \propto |\hat{\Sigma}_0|^{-\frac{T}{2}} = (|\hat{\Sigma}_{xx}| |\hat{\Sigma}_{yy}|)^{-\frac{T}{2}}.$$

On the other hand, under H_a , the maximum likelihood estimate of Σ_z is

$$\hat{\Sigma}_a = \frac{1}{T} \sum_{i=1}^T \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} \begin{bmatrix} \mathbf{x}_i' & \mathbf{y}_i' \end{bmatrix} = \begin{bmatrix} \hat{\Sigma}_{xx} & \hat{\Sigma}_{xy} \\ \hat{\Sigma}_{yx} & \hat{\Sigma}_{yy} \end{bmatrix}.$$

The maximized likelihood function under H_a is

$$l_a \propto |\hat{\Sigma}_a|^{-\frac{T}{2}} = \left(\hat{\Sigma}_{yy} - \hat{\Sigma}_{yx} \hat{\Sigma}_{xx}^{-1} \hat{\Sigma}_{xy} \right)^{-\frac{T}{2}}.$$

The likelihood ratio test statistic is therefore

$$L = \frac{l_0}{l_a} = \left(\frac{|\hat{\Sigma}_0|}{|\hat{\Sigma}_a|} \right)^{\frac{T}{2}} = \left(|\mathbf{I} - \hat{\Sigma}_{yy}^{-1} \hat{\Sigma}_{yx} \hat{\Sigma}_{xx}^{-1} \hat{\Sigma}_{xy}| \right)^{\frac{T}{2}}.$$

We will reject the null hypothesis if L is small.

Next, let $\{\lambda_i\}_{i=1}^q$ be the eigenvalues of the matrix $\hat{\Sigma}_{yy}^{-1} \hat{\Sigma}_{yx} \hat{\Sigma}_{xx}^{-1} \hat{\Sigma}_{xy}$. Then $\{1 - \lambda_i\}_{i=1}^q$ are the eigenvalues of $\mathbf{I} - \hat{\Sigma}_{yy}^{-1} \hat{\Sigma}_{yx} \hat{\Sigma}_{xx}^{-1} \hat{\Sigma}_{xy}$. As a result, the negative log likelihood ratio statistic is

$$LR = -\frac{T}{2} \cdot \ln \left(|\mathbf{I} - \hat{\Sigma}_{yy}^{-1} \hat{\Sigma}_{yx} \hat{\Sigma}_{xx}^{-1} \hat{\Sigma}_{xy}| \right)$$

$$\begin{aligned}
&= -\frac{T}{2} \cdot \ln \left(\prod_{i=1}^q (1 - \lambda_i) \right) \\
&= -\frac{T}{2} \sum_{i=1}^q \ln(1 - \lambda_i).
\end{aligned}$$

We will reject the null hypothesis for large test statistic LR .

5.3.2 Cointegration tests of VAR models

Since Π is related to the covariance matrix between \mathbf{z}_{t-1} and $\Delta \mathbf{z}_t$, we can estimate it using the following two regressions

$$\begin{aligned}
\Delta \mathbf{z}_t &= \sum_{i=1}^{n-1} \Psi_i \Delta \mathbf{z}_{t-i} + U_t, \\
\mathbf{z}_{t-1} &= \sum_{i=1}^{n-1} \Psi_i^* \Delta \mathbf{z}_{t-i} + V_t,
\end{aligned}$$

where U_t and V_t denote the error terms. Let \hat{U}_t and \hat{V}_t denotes the least-squares residuals of the above regressions. We have

$$\hat{U}_t = \Pi \hat{V}_t + E_t,$$

where E_t denotes the error term.

Let

$$H(0) \subset H(1) \subset \cdots \subset H(k)$$

be the nested model such that under $H(r)$ there are r cointegrating vectors in \mathbf{z}_t ($Rank(\Pi) = r$). Define

$$\begin{aligned}
\hat{\Sigma}_{00} &= T^{-1} \sum_{t=1}^T \hat{U}_t \hat{U}_t', \\
\hat{\Sigma}_{11} &= T^{-1} \sum_{t=1}^T \hat{V}_t \hat{V}_t', \\
\hat{\Sigma}_{01} &= T^{-1} \sum_{t=1}^T \hat{U}_t \hat{V}_t'.
\end{aligned}$$

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ be the ordered eigenvalues of the sample matrix $\hat{\Sigma}_{11}^{-1} \hat{\Sigma}_{10} \hat{\Sigma}_{00}^{-1} \hat{\Sigma}_{01}$, and \mathbf{g}_i be the eigenvector associated with eigenvalue λ_i .⁸ Using the above results, we can define two cointegration tests below.

⁸That is,

$$\hat{\Sigma}_{11}^{-1} \hat{\Sigma}_{10} \hat{\Sigma}_{00}^{-1} \hat{\Sigma}_{01} \mathbf{g}_i = \lambda_i \mathbf{g}_i.$$

1. Trace test:

Consider the nested hypotheses $H_0 : r = r_0$ versus $H_a : r > r_0$, where r_0 is a given integer between 0 and $k - 1$ with k being the dimension of \mathbf{z}_t . Johansen's trace statistic is then defined as

$$\lambda_{trace}(r_0) = -\tilde{T} \cdot \sum_{i=r_0+1}^k \ln(1 - \hat{\lambda}_i),$$

where $\hat{\lambda}_i$ denotes the estimated values of the characteristic roots (or eigenvalues) obtained from the estimated Π matrix, and \tilde{T} is the number of usable observations. Suppose that $\lambda_1 > \lambda_2 > \dots > \lambda_k$ for k characteristic roots. Under H_0 , all eigenvalues in $\{\lambda_i\}_{i=r_0+1}^k$ should be zero, and thus the test statistic is small⁹. On the other hand, Under H_a , some of the eigenvalues in $\{\lambda_i\}_{i=r_0+1}^k$ are nonzero and the test statistic should be large. Because of the presence of the unit roots, the limiting distribution of λ_{trace} is not chi-square.

2. Maximum eigenvalue test:

Consider the hypothesis $H_0 : r = r_0$ versus $H_a : r = r_0 + 1$. The test statistic is given by

$$\lambda_{max}(r_0) = -\tilde{T} \cdot \ln(1 - \hat{\lambda}_{r_0+1}).$$

5.4 Empirical studies

5.4.1 Estimation of Danish money demand function

```
sjd.trace<-ca.jo(sjd, ecdet = "const", type="trace", K=2, spec="longrun",
season=4)
sjd.eigen <- ca.jo(sjd, ecdet = "const", type="eigen", K=2, spec="longrun",
season=4)
seigen<-attributes(sjd.eigen)
strace<-attributes(sjd.trace)
names(strace)
```

```
## [1] "x"          "Z0"          "Z1"          "ZK"          "type"
## [6] "model"      "ecdet"       "lag"         "p"           "season"
## [11] "dumvar"     "cval"       "teststat"   "lambda"     "Vorg"
```

Equivalently, we have

$$\hat{\Sigma}_{10}\hat{\Sigma}_{00}^{-1}\hat{\Sigma}_{01}\mathbf{g}_i = \lambda_i\hat{\Sigma}_{11}\mathbf{g}_i, \quad i = 1, 2, \dots, k.$$

Note that the eigenvectors can be normalized such that

$$\mathbf{G}'\hat{\Sigma}_{11}^{-1}\mathbf{G} = \mathbf{I},$$

where $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_k]$ is the matrix of eigenvectors.

⁹The rank of a matrix is equal to the number of non-zero characteristic roots.


```
## [16] "V"          "W"          "PI"          "DELTA"       "GAMMA"
## [21] "R0"         "RK"         "bp"          "spec"        "call"
## [26] "test.name" "class"
```

```
lambda<-round(strace$lambda, digits = 2)
```

A model without a linear trend in the non-stationary part of the process was assumed and $n = 2$ is assumed.

Table 1. Trace test (λ_{trace})

```
kable(cbind("trace"=strace$teststat,strace$cval),digits = 2)
```

	trace	10pct	5pct	1pct
r <= 3	2.35	7.52	9.24	12.97
r <= 2	8.69	17.85	19.96	24.60
r <= 1	19.06	32.00	34.91	41.07
r = 0	49.14	49.65	53.12	60.16

First we consider the number of cointegrating vectors, beginning with the hypothesis $r \leq 1$. Using the trace test procedure gives us

$$-T \sum_{i=2}^4 \ln(1 - \lambda_i) = 19.06$$

The 95 percentile quantile, 35.07, is not significant. Hence there is no evidence in the Danish data for more than one cointegrating relation.

Table 2. Maximum eigenvalue test (λ_{max})

```
kable(cbind("Eigen.max"=seigen$teststat,seigen$cval), digits = 2)
```

	Eigen.max	10pct	5pct	1pct
r <= 3	2.35	7.52	9.24	12.97
r <= 2	6.34	13.75	15.67	20.20
r <= 1	10.36	19.77	22.00	26.81
r = 0	30.09	25.56	28.14	33.24

We consider the maximum eigenvalue test, and test $H_0 : r = 0$ versus $H_a : r = 1$. The test statistic is $-T \cdot \ln(1 - \lambda_1) = 30.09$ which is in the upper tail of the distribution of λ_{max} for $r = 0$ with a p -value of 2.5 percent.

Exercise Conflicting conclusions are found in the above two trace tests. How do you resolve the conflicting inference results? Explain your reasons.

Table 3. Estimates of eigenvalues (λ) and cointegrating vectors (β)

```
kable(rbind("Eigenvalues"=strace$lambda,strace$V), digits = 2)
```

	LRM.l2	LRY.l2	IBO.l2	IDE.l2	constant
Eigenvalues	0.43	0.18	0.11	0.04	0.00
LRM.l2	1.00	1.00	1.00	1.00	1.00
LRY.l2	-1.03	-1.37	-3.23	-1.88	-0.63
IBO.l2	5.21	0.24	0.54	24.40	1.70
IDE.l2	-4.22	6.84	-5.65	-14.30	-1.90
constant	-6.06	-4.27	7.90	-2.26	-8.03

Table 4. Estimates of speed of adjustment (α)

```
kable(strace$W, digits = 3)
```

	LRM.l2	LRY.l2	IBO.l2	IDE.l2	constant
LRM.d	-0.213	-0.005	0.035	0.002	0
LRY.d	0.115	0.020	0.050	0.001	0
IBO.d	0.023	-0.011	0.003	-0.002	0
IDE.d	0.029	-0.030	-0.003	0.000	0

The coefficient estimates of the cointegrating relation are found in the first column of **Table 3**. The cointegration vector can be seen as an error correction mechanism measuring the excess demand for money with the estimate of the equilibrium relation given by

$$m_2 = 1.03y - 5.21i_b + 4.22i_d + 6.06.$$

Moreover,

$$\hat{\alpha} = (-0.213, 0.115, 0.023, 0.029)'$$

can be found in the first column of **Table 4**. These coefficients are natural to give them an economic meaning in terms of the average speed of adjustment towards the estimated equilibrium state, such that a low coefficient indicates slow adjustment and a high coefficient indicates rapid adjustment.

5.4.2 Pairs trading

We demonstrate *Pairs Trading* in equity markets using the example in Tsay (2010), where two stocks, Billiton Ltd. of Australia, BHP, and Vale S.A. of Brazil, VALE are used in the analysis.

```
pairs.eigen<-attributes(ca.jo(log(dat.pairs), type="eigen",
                             ecdet = "const", spec="transitory"))
pairs.trace<-attributes(ca.jo(log(dat.pairs), type="trace",
                             ecdet = "const", spec="transitory"))
```

Table 5-1. Trace test (λ_{trace})

```
kable(cbind("trace"=pairs.trace$teststat, pairs.trace$cval), digits = 2)
```

	trace	10pct	5pct	1pct
r <= 1	7.78	7.52	9.24	12.97
r = 0	47.77	17.85	19.96	24.60

Table 5-2. Maximum eigenvalue test (λ_{max})

```
kable(cbind("Eigen.max"=pairs.eigen$teststat,pairs.eigen$cval), digits = 2)
```

	Eigen.max	10pct	5pct	1pct
r <= 1	7.78	7.52	9.24	12.97
r = 0	40.00	13.75	15.67	20.20

Exercise Do we see conflicting results in this example? Explain your reasons.

Table 6-1. Estimates of eigenvalues and cointegrating vectors

```
kable(rbind("Eigenvalues"=pairs.trace$lambda,pairs.trace$V), digits = 2)
```

	BHP.l1	VALE.l1	constant
Eigenvalues	0.04	0.01	0.00
BHP.l1	1.00	1.00	1.00
VALE.l1	-0.72	-0.73	2.05
constant	-1.83	-1.54	-5.71

Table 6-2. Estimates of speed of adjustment and vector autoregressive coefficient matrix

```
kable(pairs.trace$W, digits = 3)
```

	BHP.l1	VALE.l1	constant
BHP.d	-0.067	0.005	0
VALE.d	0.025	0.008	0

```
kable(pairs.trace$GAMMA, digits = 3)
```

	BHP.dl1	VALE.dl1
BHP.d	-0.115	0.069
VALE.d	0.053	0.045

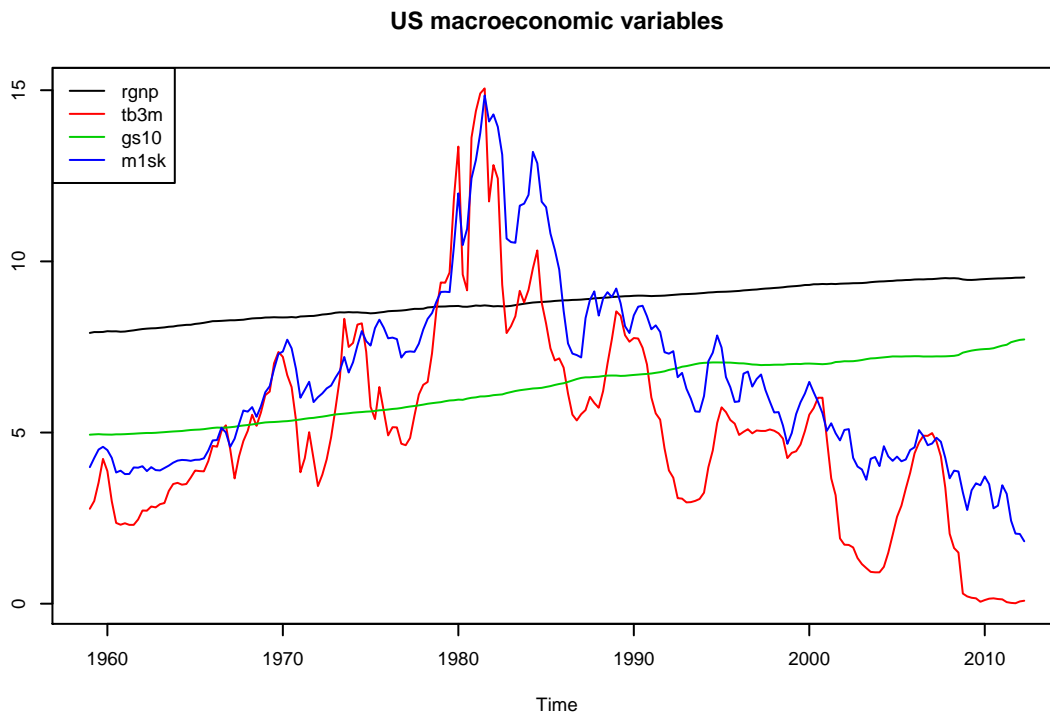
Exercise Write down the error correction model for BHP and VALE using the above results.

Exercise Redo the cointegration analysis on BHP and VALE using the Engle and Granger's approach.

5.4.3 US macroeconomic variables

Finally, let's use US macro data to demonstrate the Johansen modeling procedure and explore the case of multiple cointegrating vectors.

```
zt<-cbind(log(USmacro[, "rgnp"]), USmacro[, "tb3m"],
          log(USmacro[, "m1sk"]), USmacro[, "gs10"])
colnames(zt)<-c("rgnp", "tb3m", "gs10", "m1sk")
par(cex=0.6)
ts.plot(zt, main="US macroeconomic variables", col=1:4)
legend("topleft", c("rgnp", "tb3m", "gs10", "m1sk"), lty=1, col=1:4)
```



```
m1=VARorderI(zt, output=FALSE)
c("AIC order"=m1$aicor, "BIC order"=m1$bicor, "HQ order"=m1$hqor)
```

```
## AIC order BIC order HQ order
##          6          2          2
```

```
m2<-attributes(ca.jo(zt, K=5, ecdet = "const",
                     spec = "transitory", type = "eigen"))
m3<-attributes(ca.jo(zt, K=5, ecdet = "const",
                     spec = "transitory", type="trace"))
```

Table 7-1. Trace test (λ_{trace})

```
kable(cbind("trace"=m3$teststat, m3$cval), digits = 2)
```

	trace	10pct	5pct	1pct
r <= 3	4.25	7.52	9.24	12.97
r <= 2	12.79	17.85	19.96	24.60
r <= 1	43.26	32.00	34.91	41.07
r = 0	87.87	49.65	53.12	60.16

Table 7-2. Maximum eigenvalue test (λ_{max})

```
kable(cbind("Eigen.max"=m2$teststat, m2$cval), digits = 2)
```

	Eigen.max	10pct	5pct	1pct
r <= 3	4.25	7.52	9.24	12.97
r <= 2	8.53	13.75	15.67	20.20
r <= 1	30.47	19.77	22.00	26.81
r = 0	44.61	25.56	28.14	33.24

Table 8-1. Estimates of eigenvalues and cointegrating vectors

```
kable(rbind("Eigenvalues"=m3$lambda,m3$V), digits = 2)
```

	rgnp.l1	tb3m.l1	gs10.l1	m1sk.l1	constant
Eigenvalues	0.19	0.14	0.04	0.02	0.00
rgnp.l1	1.00	1.00	1.00	1.00	1.00
tb3m.l1	-0.28	-0.78	-0.01	-0.05	0.04
gs10.l1	-0.79	-0.67	-0.54	-0.76	-0.17
m1sk.l1	0.31	0.77	0.03	0.01	-0.19
constant	-3.73	-5.94	-5.61	-3.89	-6.46

Table 8-2. Estimates of speed of adjustment

```
kable(m3$W, digits = 3)
```

	rgnp.l1	tb3m.l1	gs10.l1	m1sk.l1	constant
rgnp.d	0.008	0.000	-0.014	0.002	0
tb3m.d	-0.211	0.166	-0.131	0.394	0
gs10.d	0.009	-0.001	0.017	0.002	0
m1sk.d	-0.274	-0.084	-0.192	0.271	0

```
ci1<-m3$V[-5,1]
ci2<-m3$V[-5,2]
w1t<-zt%*%ci1
w2t<-zt%*%ci2
```

```
library(tseries)
adf.test(w1t)
```

```
## Warning in adf.test(w1t): p-value smaller than printed p-value
```

```
##
```

```
## Augmented Dickey-Fuller Test
```

```
##
```

```
## data: w1t
```

```
## Dickey-Fuller = -4.9722, Lag order = 5, p-value = 0.01
```

```
## alternative hypothesis: stationary
```

```
adf.test(w2t)
```

```
## Warning in adf.test(w2t): p-value smaller than printed p-value
```

```
##
```

```
## Augmented Dickey-Fuller Test
```

```
##
## data: w2t
## Dickey-Fuller = -5.4555, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

```
wt<-cbind(w1t,w2t)
m4<-ECMvar1(zt, 6, wt, include.const = T, output = FALSE)
m5<-refECMvar1prime(m4, thres = 0.8);
```

```
## Equation: 1 npar = 18
## Equation: 2 npar = 17
## Equation: 3 npar = 12
## Equation: 4 npar = 17
```

```
# names(m4)
```

```
alpha<-m4$coef[1:2,]
alpha1<-m5$coef[1:2,]
colnames(alpha)<-c("rgnp","tb3m","gs10","m1sk")
row.names(alpha)<-c("w1t","w2t")
colnames(alpha1)<-c("rgnp","tb3m","gs10","m1sk")
row.names(alpha1)<-c("w1t","w2t")
kable(alpha, digits = 4)
```

	rgnp	tb3m	gs10	m1sk
w1t	0.0097	0.113	0.0081	0.0505
w2t	-0.0006	0.140	-0.0010	-0.1610

```
kable(alpha1, digits = 4)
```

	rgnp	tb3m	gs10	m1sk
w1t	0.0083	0.0000	0.0053	0.0000
w2t	0.0000	0.1718	0.0000	-0.1504