KudrinKnobel.R

Albina

2023-05-06

# 1 - Kudrin-Knobel the simpliest mode  
  
library(seasonal)  
library(readxl)  
library(ggplot2)  
library(zoo)

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

library(reshape2)  
library(vars)

## Loading required package: MASS

## Loading required package: strucchange

## Loading required package: sandwich

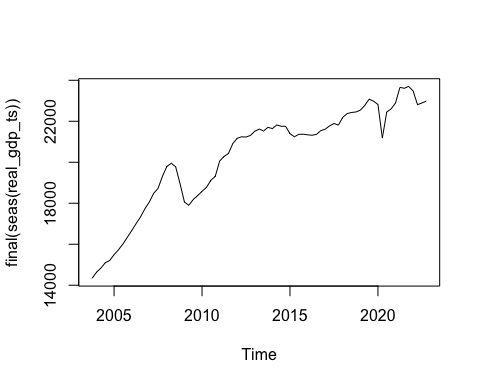
## Loading required package: urca

## Loading required package: lmtest

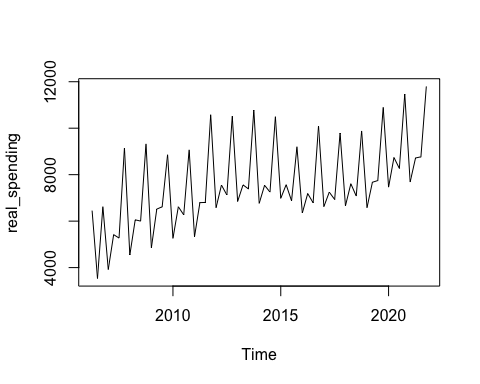
library(tseries)

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

library(dynlm)  
  
rm(list=ls())  
setwd('/Users/Albina/Documents/Studying/НИР/preparing\_to\_seminar\_01.03')  
  
# ВВП, приведенный в реальное выражение (2016 год), млрд руб  
real\_gdp <- na.omit(read\_excel('VVP\_kvartal\_s1995(3).xls', sheet = 'Real GDP'))  
real\_gdp <- na.omit(real\_gdp[8:84, 2])  
# урезаем данные, т.к. G на сайте казнчаейства только с 4 квартала 2003 по 4 квартал 2022  
real\_gdp\_ts <- ts(real\_gdp, start=c(2003, 4), end=c(2022, 4), frequency = 4)  
  
plot(final(seas(real\_gdp\_ts)))



# Расходы и доходы, составленные вручную с сайта казначейства  
# Приведены в реальное выражение с помощью ИПЦ  
  
real\_g\_new <- na.omit(read\_excel('NCGGRNSAXDCRUQ.xls', sheet = 'result'))  
real\_g\_new <- real\_g\_new["real\_spending"]  
real\_g\_new\_ts <- ts(real\_g\_new[11:73, 1], start=c(2006, 2), end=c(2021, 4), frequency = 4)  
  
real\_g\_old <- na.omit(read\_excel('VVP\_kvartal\_s1995(3).xls', sheet = 'G'))  
real\_g\_old\_ts <- ts(real\_g\_old["G"][8:17, 1], start=c(2003, 4), end=c(2006, 1), frequency = 4)  
  
real\_g\_merged <- ts(c(real\_g\_old\_ts, real\_g\_new\_ts), start = start(real\_g\_old\_ts), frequency = frequency(real\_g\_old\_ts))  
plot(real\_g\_new\_ts)



q1 <- read\_excel('/Users/Albina/Documents/Studying/НИР/Отчеты/2022.04.01.xls')

## New names:  
## • `` -> `...2`  
## • `` -> `...3`

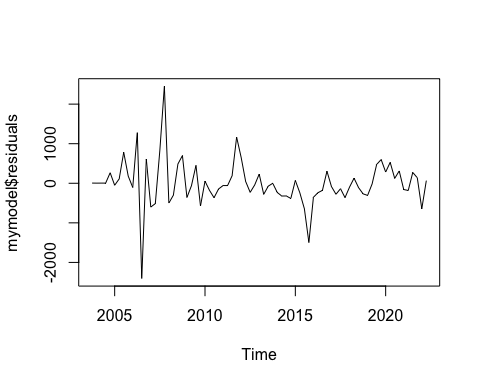
q1 <- q1[9:18, c(1, 3)]  
  
q2 <- read\_excel('/Users/Albina/Documents/Studying/НИР/Отчеты/2022.07.01.xls')

## New names:  
## • `` -> `...2`  
## • `` -> `...3`

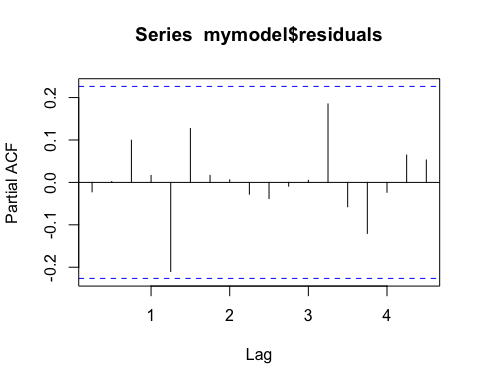
q2 <- q2[9:18, c(1, 3)]  
q2[,2] <- q2[,2] - q1[,2]  
  
q1[,2] <- q1[,2]/1.336050007/1.099468405  
q2[,2] <- q2[,2]/1.336050007/1.099468405/1.013259854  
  
  
g\_merged <- ts(c(real\_g\_merged, ts(q1[10, 2], start = c(2022, 1)), ts(q2[10, 2], start = c(2022, 2))), start=start(real\_g\_merged), frequency = 4)  
  
# на сайте приведены опубликованы неполные данные: не хватает 3 и 4 квартала 2022  
# попробуем оценить с помощью ARIMA  
  
library(forecast)  
mymodel <- auto.arima(g\_merged)  
mymodel

## Series: g\_merged   
## ARIMA(1,0,1)(0,1,1)[4] with drift   
##   
## Coefficients:  
## ar1 ma1 sma1 drift  
## 0.8394 -0.6030 -0.4543 69.3828  
## s.e. 0.1385 0.1836 0.1379 23.9434  
##   
## sigma^2 = 377277: log likelihood = -554.93  
## AIC=1119.86 AICc=1120.78 BIC=1131.18

plot.ts(mymodel$residuals)



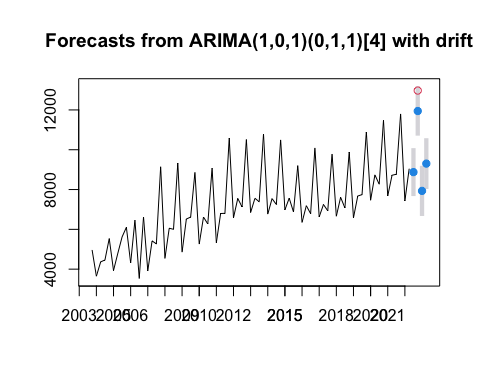
pacf(mymodel$residuals)



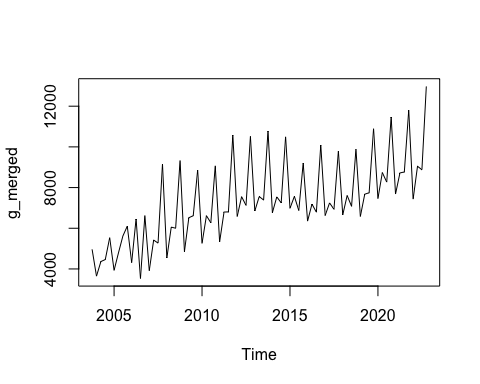
myforecast <- forecast(mymodel, level=c(95), h=1\*4)  
plot(myforecast)  
axis(1, at = seq(2003, 2022, by = 1))  
  
q4 <- read\_excel('/Users/Albina/Documents/Studying/НИР/Отчеты/2023.01.01.xls')

## New names:  
## • `` -> `...2`  
## • `` -> `...3`

q4 <- q4[9:18, c(1, 3)]  
  
  
q4[,2] <- q4[,2]/1.336050007/1.099468405/1.013259854/0.99141574/1.013349612  
q4\_ts <- ts(q4[10, 2] - myforecast$mean[1] - q1[10, 2] - q2[10, 2], start = c(2022, 4), end = c(2022, 4), frequency = 4)  
points(q4\_ts, col=2)



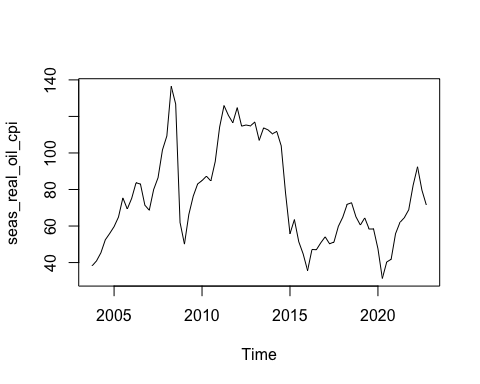
g\_merged <- ts(c(g\_merged, myforecast$mean[1], q4\_ts), start = start(g\_merged), frequency = 4)  
plot(g\_merged)



# цена на нефть марки Brent в долларах, приведенная в реальное выражение с помощью   
# индекса цен производителей  
real\_oil\_price <- na.omit(read\_excel('POILBREUSDQ.xls', sheet='FRED Graph'))

## New names:  
## • `` -> `...3`  
## • `` -> `...5`

#real\_oil\_price <- real\_oil\_price["real\_oil\_price"]  
#real\_oil\_price\_ts <- ts(real\_oil\_price, start=c(2003, 4), end=c(2022, 4), frequency = 4)  
  
#plot(log(real\_oil\_price\_ts))  
  
# приведено в реальное выражение с помощью Producer Price Index by Industry: Oil and Gas Extraction  
# PCU21112111  
  
#real\_oil\_price <- na.omit(read\_excel('POILBREUSDQ.xls', sheet='FRED Graph'))  
#real\_oil\_price\_ppi <- real\_oil\_price["real\_oil\_price\_ppi"]  
#real\_oil\_price\_ppi\_ts <- ts(real\_oil\_price\_ppi, start=c(2003, 4), end=c(2022, 4), frequency = 4)  
  
#plot(log(real\_oil\_price\_ppi\_ts), col=2)  
  
# приведено в реальное выражение с помощью Consumer Price Index for All Urban Consumers  
# CPIAUCSL  
  
seas\_real\_oil\_price\_cpi <- real\_oil\_price["seas\_real\_oil\_cpi"]  
seas\_real\_oil\_price\_cpi\_ts <- ts(seas\_real\_oil\_price\_cpi, start=c(2003, 4), end=c(2022, 4), frequency = 4)  
  
plot(seas\_real\_oil\_price\_cpi\_ts)



#legend(x = "topright", legend = c("PPI", "PPI\_oil", "CPI\_urban"),   
# lty = c(1, 1, 1),   
# col = c(1, 2, 3),   
# lwd = 1)   
#axis(1, at = seq(2003, 2022, by = 1))  
  
# исключаем сезонность и берем логарифм  
  
gdp <- log(final(seas(real\_gdp\_ts)))  
g <- log(final(seas(g\_merged)))  
oil <- log(seas\_real\_oil\_price\_cpi\_ts)  
  
# оценим модель в уровнях  
  
library(vars)  
library(tseries)  
  
# задаем вид матрицы при ошибках, NA кладем на место коэффициента, который хотим оценить  
  
bmat <- diag(2)  
bmat[2, 1] <- NA  
  
sv <- cbind(g, gdp)  
colnames(sv) <- cbind("Government spending", "GDP")  
  
# ur.df проводит тестирование Dickey-Fuller  
summary(ur.df(gdp, type = "none", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.077779 -0.005411 -0.000252 0.007900 0.070286   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## z.lag.1 0.0004618 0.0002181 2.118 0.0376 \*  
## z.diff.lag 0.2203546 0.1136862 1.938 0.0565 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.0177 on 73 degrees of freedom  
## Multiple R-squared: 0.1433, Adjusted R-squared: 0.1198   
## F-statistic: 6.105 on 2 and 73 DF, p-value: 0.003535  
##   
##   
## Value of test-statistic is: 2.1177   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.6 -1.95 -1.61

summary(ur.df(g, type = "drift", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression drift   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.58362 -0.01939 0.00147 0.02452 0.29871   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.90837 0.45529 1.995 0.0498 \*   
## z.lag.1 -0.10050 0.05135 -1.957 0.0542 .   
## z.diff.lag -0.53507 0.09617 -5.564 4.25e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09392 on 72 degrees of freedom  
## Multiple R-squared: 0.3654, Adjusted R-squared: 0.3477   
## F-statistic: 20.73 on 2 and 72 DF, p-value: 7.773e-08  
##   
##   
## Value of test-statistic is: -1.9573 3.2002   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau2 -3.51 -2.89 -2.58  
## phi1 6.70 4.71 3.86

#Не отвергаем H0  
  
#Проверка наличия коинтеграции. Включать константу "-1" или нет - вопрос дискуссионный.  
cointcy <- dynlm(gdp ~ g )  
summary(cointcy)

##   
## Time series regression with "ts" data:  
## Start = 2003(4), End = 2022(4)  
##   
## Call:  
## dynlm(formula = gdp ~ g)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.124424 -0.039879 -0.000942 0.031838 0.241380   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.22367 0.24782 21.08 <2e-16 \*\*\*  
## g 0.52845 0.02794 18.91 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.05524 on 75 degrees of freedom  
## Multiple R-squared: 0.8267, Adjusted R-squared: 0.8244   
## F-statistic: 357.7 on 1 and 75 DF, p-value: < 2.2e-16

#Сохраняем ошибку из регрессии (gdp ~ g) и проверяем стационарна ли она.  
ehat <- resid(cointcy)  
summary(ur.df(ehat, type = "none", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.131869 -0.017194 0.003078 0.017456 0.283757   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## z.lag.1 -0.4855 0.1290 -3.762 0.000337 \*\*\*  
## z.diff.lag -0.2821 0.1115 -2.529 0.013597 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.04996 on 73 degrees of freedom  
## Multiple R-squared: 0.3911, Adjusted R-squared: 0.3744   
## F-statistic: 23.45 on 2 and 73 DF, p-value: 1.365e-08  
##   
##   
## Value of test-statistic is: -3.7625   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.6 -1.95 -1.61

summary(adf.test(sv[,2]))

## Length Class Mode   
## statistic 1 -none- numeric   
## parameter 1 -none- numeric   
## alternative 1 -none- character  
## p.value 1 -none- numeric   
## method 1 -none- character  
## data.name 1 -none- character

lagselect <- VARselect(sv, lag.max = 4, "const")  
lagselect$selection

## AIC(n) HQ(n) SC(n) FPE(n)   
## 2 2 1 2

Model1 <- VAR(sv, p = 3, season = NULL, exogen = oil, type = "const")  
SVARMod1 <- SVAR(Model1, Bmat = bmat, hessian = TRUE, esmethod = "scoring")

## Warning in SVAR(Model1, Bmat = bmat, hessian = TRUE, esmethod = "scoring"):  
## Convergence not achieved after 100 iterations. Convergence value:  
## 0.000337158122369882 .

SVARMod1

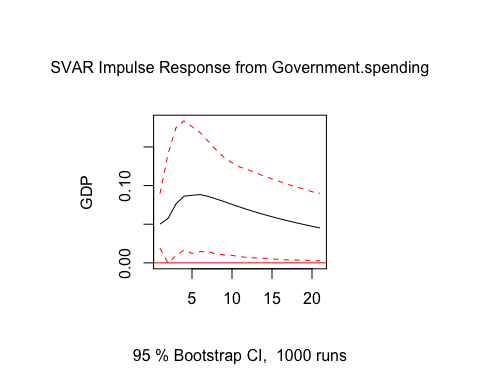
##   
## SVAR Estimation Results:  
## ========================   
##   
##   
## Estimated B matrix:  
## Government.spending GDP  
## Government.spending 1.00000 0  
## GDP 0.05023 1

#summary(SVARMod1)  
  
# считаем значение мультипликатора за год  
  
Phi(SVARMod1, nstep=4)[,,1][2,][1] + Phi(SVARMod1, nstep=4)[,,2][2,][1] + Phi(SVARMod1, nstep=4)[,,3][2,][1] + Phi(SVARMod1, nstep=4)[,,4][2,][1]

## [1] 0.2701571

Imp1 <- irf(SVARMod1, impulse = "Government.spending", response = "GDP", n.ahead = 20,  
 ortho = TRUE, runs = 1000)

plot(Imp1)



# 3 - Попробуем ту же модель в разностях  
  
gdp <- diff(log(final(seas(real\_gdp\_ts))))  
g <- diff(log(final(seas(g\_merged))))  
oil <- diff(oil)  
  
bmat <- diag(2)  
bmat[2, 1] <- NA  
  
sv3 <- cbind(g, gdp)  
colnames(sv3) <- cbind("Government spending", "GDP")  
  
# ur.df проводит тестирование Dickey-Fuller  
summary(ur.df(gdp, type = "none", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.072164 -0.002955 0.004355 0.010272 0.077102   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## z.lag.1 -0.5644 0.1357 -4.159 8.72e-05 \*\*\*  
## z.diff.lag -0.2045 0.1143 -1.789 0.0778 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.01795 on 72 degrees of freedom  
## Multiple R-squared: 0.3825, Adjusted R-squared: 0.3653   
## F-statistic: 22.3 on 2 and 72 DF, p-value: 2.909e-08  
##   
##   
## Value of test-statistic is: -4.1587   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.6 -1.95 -1.61

summary(ur.df(g, type = "drift", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression drift   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.51991 -0.02501 -0.00816 0.02558 0.33243   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.02292 0.01083 2.117 0.03779 \*   
## z.lag.1 -2.10972 0.19821 -10.644 2.35e-16 \*\*\*  
## z.diff.lag 0.33815 0.11162 3.029 0.00342 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09128 on 71 degrees of freedom  
## Multiple R-squared: 0.8124, Adjusted R-squared: 0.8071   
## F-statistic: 153.8 on 2 and 71 DF, p-value: < 2.2e-16  
##   
##   
## Value of test-statistic is: -10.6437 56.6439   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau2 -3.51 -2.89 -2.58  
## phi1 6.70 4.71 3.86

# отвергаем H0  
  
#Проверка наличия коинтеграции. Включать константу "-1" или нет - вопрос дискуссионный.  
cointcy <- dynlm(gdp ~ g )  
summary(cointcy)

##   
## Time series regression with "ts" data:  
## Start = 2004(1), End = 2022(4)  
##   
## Call:  
## dynlm(formula = gdp ~ g)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.080957 -0.006199 0.001417 0.009919 0.052329   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.006165 0.002084 2.958 0.00415 \*\*  
## g 0.004330 0.018063 0.240 0.81122   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.01808 on 74 degrees of freedom  
## Multiple R-squared: 0.0007759, Adjusted R-squared: -0.01273   
## F-statistic: 0.05746 on 1 and 74 DF, p-value: 0.8112

#Сохраняем ошибку из регрессии (gdp ~ g) и проверяем стационарна ли она.  
ehat <- resid(cointcy)  
summary(ur.df(ehat, type = "none", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.076941 -0.005546 0.001277 0.007186 0.069072   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## z.lag.1 -0.6668 0.1453 -4.590 1.83e-05 \*\*\*  
## z.diff.lag -0.1510 0.1160 -1.302 0.197   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.01759 on 72 degrees of freedom  
## Multiple R-squared: 0.4068, Adjusted R-squared: 0.3903   
## F-statistic: 24.68 on 2 and 72 DF, p-value: 6.858e-09  
##   
##   
## Value of test-statistic is: -4.5898   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.6 -1.95 -1.61

summary(adf.test(sv[,2]))

## Length Class Mode   
## statistic 1 -none- numeric   
## parameter 1 -none- numeric   
## alternative 1 -none- character  
## p.value 1 -none- numeric   
## method 1 -none- character  
## data.name 1 -none- character

lagselect <- VARselect(sv3, lag.max = 4, "const")  
lagselect$selection

## AIC(n) HQ(n) SC(n) FPE(n)   
## 2 1 1 2

Model3 <- VAR(sv3, p = 2, season = NULL, exogen = oil, type = "const")  
SVARMod3 <- SVAR(Model3, Bmat = bmat, hessian = TRUE, esmethod = "scoring")

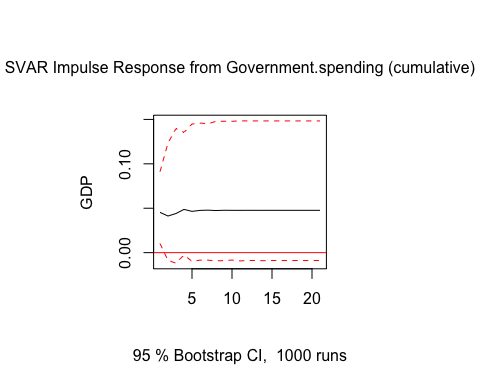
## Warning in SVAR(Model3, Bmat = bmat, hessian = TRUE, esmethod = "scoring"):  
## Convergence not achieved after 100 iterations. Convergence value:  
## 0.000348376077394595 .

SVARMod3

##   
## SVAR Estimation Results:  
## ========================   
##   
##   
## Estimated B matrix:  
## Government.spending GDP  
## Government.spending 1.00000 0  
## GDP 0.04535 1

#summary(SVARMod3)  
  
#Phi(SVARMod3, nstep=4)[,,1][2,][1] + Phi(SVARMod3, nstep=4)[,,2][2,][1] + Phi(SVARMod3, nstep=4)[,,3][2,][1] + Phi(SVARMod3, nstep=4)[,,4][2,][1]  
  
  
Imp3 <- irf(SVARMod3, impulse = "Government.spending", response = "GDP", n.ahead = 20, ortho = TRUE, runs = 1000, cumulative = TRUE)

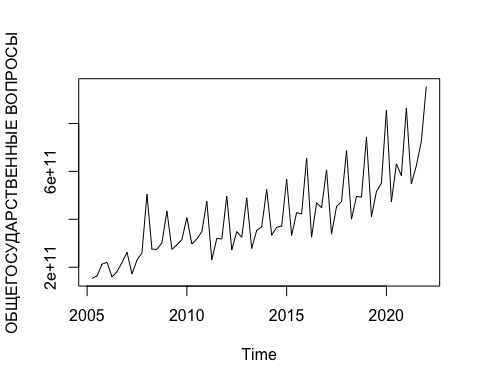
plot(Imp3)



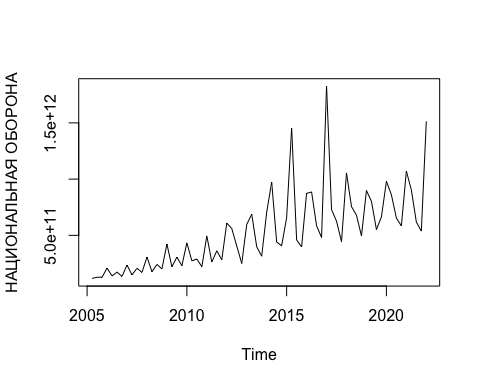
# считаем накопленный мульипликатор  
Imp3$irf$Government.spending[1] + Imp3$irf$Government.spending[2] + Imp3$irf$Government.spending[3] + Imp3$irf$Government.spending[4]

## [1] 0.1793357

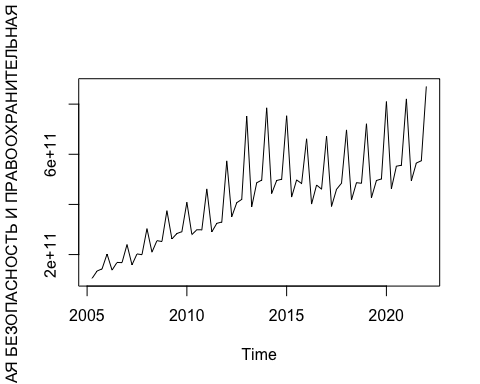
################################################################################  
  
# Расходы по отдельным функциональным разделам  
  
g\_by\_categories <- read\_excel('/Users/Albina/Documents/Studying/НИР/calculations/g\_by\_categories\_subtracted.xlsx', sheet='Sheet1')  
  
g\_by\_categories\_ts <- ts(g\_by\_categories, start=c(2005, 2), end=c(2022, 1), frequency = 4)  
g\_by\_categories\_1 <- ts(g\_by\_categories['ОБЩЕГОСУДАРСТВЕННЫЕ ВОПРОСЫ'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_1)



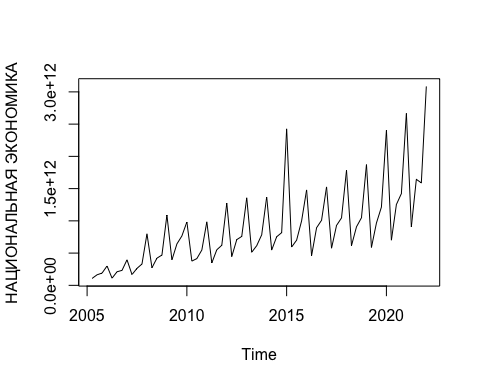
g\_by\_categories\_2 <- ts(g\_by\_categories['НАЦИОНАЛЬНАЯ ОБОРОНА'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_2)



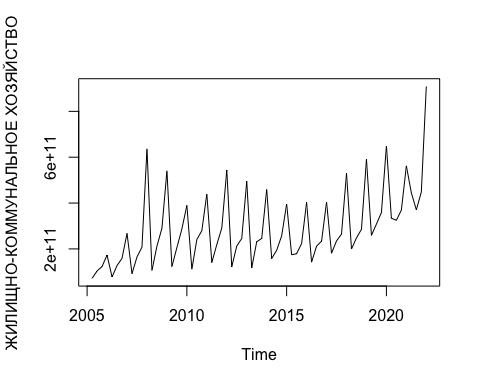
g\_by\_categories\_3 <- ts(g\_by\_categories['НАЦИОНАЛЬНАЯ БЕЗОПАСНОСТЬ И ПРАВООХРАНИТЕЛЬНАЯ ДЕЯТЕЛЬНОСТЬ'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_3)



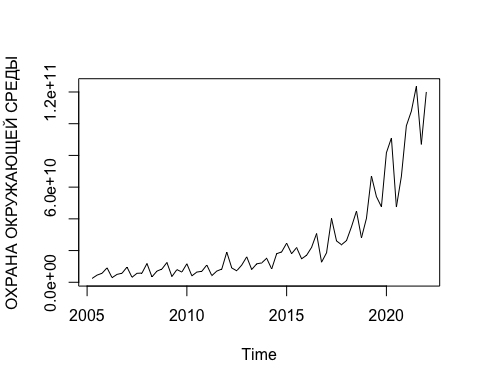
g\_by\_categories\_4 <- ts(g\_by\_categories['НАЦИОНАЛЬНАЯ ЭКОНОМИКА'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_4)



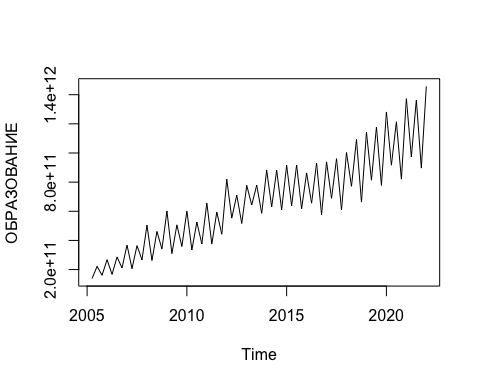
g\_by\_categories\_5 <- ts(g\_by\_categories['ЖИЛИЩНО-КОММУНАЛЬНОЕ ХОЗЯЙСТВО'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_5)



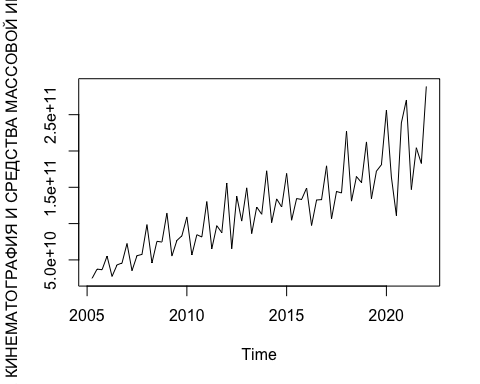
g\_by\_categories\_6 <- ts(g\_by\_categories['ОХРАНА ОКРУЖАЮЩЕЙ СРЕДЫ'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_6)



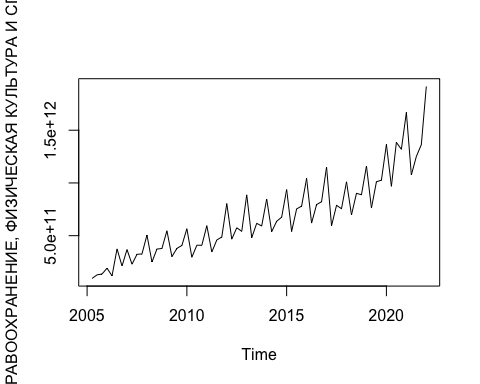
g\_by\_categories\_7 <- ts(g\_by\_categories['ОБРАЗОВАНИЕ'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_7)



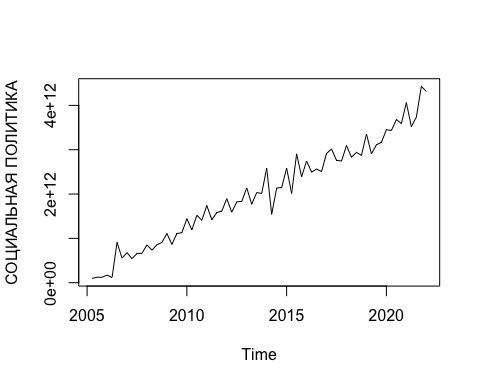
g\_by\_categories\_8 <- ts(g\_by\_categories['КУЛЬТУРА, КИНЕМАТОГРАФИЯ И СРЕДСТВА МАССОВОЙ ИНФОРМАЦИИ'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_8)



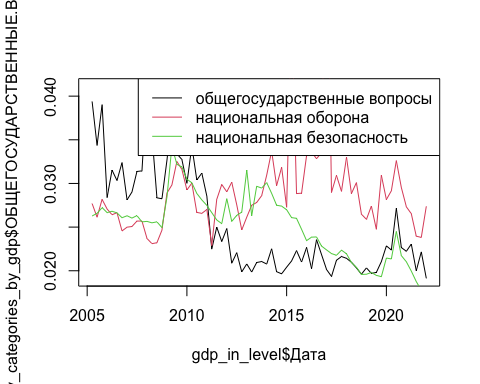
g\_by\_categories\_9 <- ts(g\_by\_categories['ЗДРАВООХРАНЕНИЕ, ФИЗИЧЕСКАЯ КУЛЬТУРА И СПОРТ'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_9)



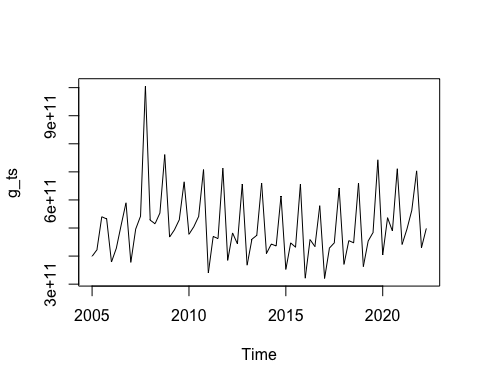
g\_by\_categories\_10 <- ts(g\_by\_categories['СОЦИАЛЬНАЯ ПОЛИТИКА'], start=c(2005, 2), end=c(2022, 1), frequency = 4)  
plot(g\_by\_categories\_10)



gdp\_in\_level <- na.omit(read\_excel('/Users/Albina/Documents/Studying/НИР/calculations/gdp.xls', sheet='in\_level'))  
df\_g\_by\_categories <- data.frame(g\_by\_categories\_1, g\_by\_categories\_2, g\_by\_categories\_3,   
 g\_by\_categories\_4, g\_by\_categories\_5, g\_by\_categories\_6,  
 g\_by\_categories\_7, g\_by\_categories\_8, g\_by\_categories\_9,  
 g\_by\_categories\_10)  
df\_g\_by\_categories\_by\_gdp <- df\_g\_by\_categories  
for(i in 1:ncol(df\_g\_by\_categories)) { # for-loop over columns  
 df\_g\_by\_categories\_by\_gdp[ , i] <- ts(df\_g\_by\_categories[ , i] / gdp\_in\_level$`GDP in level` / 10^9, start=c(2005, 2), end=c(2022, 1), frequency = 4)  
 df\_g\_by\_categories\_by\_gdp[ , i] <- final(seas(df\_g\_by\_categories\_by\_gdp[ , i]))  
}  
plot(x=gdp\_in\_level$Дата, y=df\_g\_by\_categories\_by\_gdp$ОБЩЕГОСУДАРСТВЕННЫЕ.ВОПРОСЫ, 'l')  
lines(x=gdp\_in\_level$Дата, y=df\_g\_by\_categories\_by\_gdp$НАЦИОНАЛЬНАЯ.ОБОРОНА, 'l', col=2)  
lines(x=gdp\_in\_level$Дата, y=df\_g\_by\_categories\_by\_gdp$НАЦИОНАЛЬНАЯ.БЕЗОПАСНОСТЬ.И.ПРАВООХРАНИТЕЛЬНАЯ.ДЕЯТЕЛЬНОСТЬ, 'l', col=3)  
legend(x = "topright", legend = c("общегосударственные вопросы", "национальная оборона", "национальная безопасность"),   
 lty = c(1, 1, 1),   
 col = c(1, 2, 3),   
 lwd = 1)



#for(i in 2:ncol(df\_g\_by\_categories\_by\_gdp)) { # for-loop over columns  
# plot(x=gdp\_in\_level$Дата, y=df\_g\_by\_categories\_by\_gdp[ , i], 'l', col=i)  
 #lines(x=gdp\_in\_level$Дата, y=df\_g\_by\_categories\_by\_gdp[ , i], col=i, 'l')  
#}  
  
################################################################################  
  
# оцениваем мультипликатр для разных статей гос расходов  
  
real\_g\_by\_categories <- read\_excel('/Users/Albina/Documents/Studying/НИР/calculations/g\_by\_categories\_subtracted.xlsx',sheet='real')  
  
g\_ts <- ts(real\_g\_by\_categories$`ОБЩЕГОСУДАРСТВЕННЫЕ ВОПРОСЫ`, start=c(2005, 1), end=c(2022, 2), frequency = 4)  
plot(g\_ts)



real\_gdp\_ts <- ts(real\_gdp[5:75, 1] \* 10^9, start=c(2005, 1), end=c(2022, 2), frequency = 4)  
  
real\_oil\_price\_ts <- ts(real\_oil\_price[5:75, 1], start=c(2005, 1), end=c(2022, 2), frequency = 4)  
  
gdp <- diff(log(final(seas(real\_gdp\_ts))))  
g <- diff(log(final(seas(g\_ts))))  
oil <- diff(log(real\_oil\_price\_ts))  
  
bmat <- diag(2)  
bmat[2, 1] <- NA  
  
sv <- cbind(g, gdp)  
colnames(sv) <- cbind("ОБЩЕГОСУДАРСТВЕННЫЕ ВОПРОСЫ", "ВВП")  
  
# ur.df проводит тестирование Dickey-Fuller  
summary(ur.df(gdp, type = "none", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.081030 -0.000568 0.003931 0.012193 0.082623   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## z.lag.1 -0.6842 0.1506 -4.542 2.48e-05 \*\*\*  
## z.diff.lag -0.1073 0.1224 -0.877 0.384   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.02031 on 65 degrees of freedom  
## Multiple R-squared: 0.3919, Adjusted R-squared: 0.3732   
## F-statistic: 20.95 on 2 and 65 DF, p-value: 9.53e-08  
##   
##   
## Value of test-statistic is: -4.5417   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.6 -1.95 -1.61

summary(ur.df(g, type = "drift", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression drift   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.25925 -0.02186 0.00018 0.02257 0.41455   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.002460 0.008954 0.275 0.784   
## z.lag.1 -1.395517 0.187217 -7.454 2.96e-10 \*\*\*  
## z.diff.lag 0.196419 0.122577 1.602 0.114   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07324 on 64 degrees of freedom  
## Multiple R-squared: 0.5993, Adjusted R-squared: 0.5868   
## F-statistic: 47.86 on 2 and 64 DF, p-value: 1.954e-13  
##   
##   
## Value of test-statistic is: -7.454 27.7811   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau2 -3.51 -2.89 -2.58  
## phi1 6.70 4.71 3.86

# отвергаем H0  
  
#Проверка наличия коинтеграции. Включать константу "-1" или нет - вопрос дискуссионный.  
cointcy <- dynlm(gdp ~ g )  
summary(cointcy)

##   
## Time series regression with "ts" data:  
## Start = 2005(2), End = 2022(2)  
##   
## Call:  
## dynlm(formula = gdp ~ g)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.087169 -0.004714 -0.000311 0.010370 0.057620   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.006488 0.002365 2.743 0.0078 \*\*  
## g 0.042174 0.032410 1.301 0.1976   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.01964 on 67 degrees of freedom  
## Multiple R-squared: 0.02465, Adjusted R-squared: 0.01009   
## F-statistic: 1.693 on 1 and 67 DF, p-value: 0.1976

#Сохраняем ошибку из регрессии (gdp ~ g) и проверяем стационарна ли она.  
ehat <- resid(cointcy)  
summary(ur.df(ehat, type = "none", selectlags = "AIC"))

##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.084759 -0.005256 -0.000520 0.008482 0.071039   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## z.lag.1 -0.80292 0.16020 -5.012 4.4e-06 \*\*\*  
## z.diff.lag -0.05076 0.12347 -0.411 0.682   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.01955 on 65 degrees of freedom  
## Multiple R-squared: 0.4244, Adjusted R-squared: 0.4067   
## F-statistic: 23.97 on 2 and 65 DF, p-value: 1.596e-08  
##   
##   
## Value of test-statistic is: -5.012   
##   
## Critical values for test statistics:   
## 1pct 5pct 10pct  
## tau1 -2.6 -1.95 -1.61

summary(adf.test(sv[,2]))

## Length Class Mode   
## statistic 1 -none- numeric   
## parameter 1 -none- numeric   
## alternative 1 -none- character  
## p.value 1 -none- numeric   
## method 1 -none- character  
## data.name 1 -none- character

lagselect <- VARselect(sv, lag.max = 4, "const")  
lagselect$selection

## AIC(n) HQ(n) SC(n) FPE(n)   
## 1 1 1 1

Model <- VAR(sv, p = 1, season = NULL, exogen = oil, type = "const")

SVARMod <- SVAR(Model, Bmat = bmat, hessian = TRUE, esmethod = "scoring")

SVARMod

##   
## SVAR Estimation Results:  
## ========================   
##   
##   
## Estimated B matrix:  
## ОБЩЕГОСУДАРСТВЕННЫЕ.ВОПРОСЫ ВВП  
## ОБЩЕГОСУДАРСТВЕННЫЕ.ВОПРОСЫ 1.00000 0  
## ВВП 0.07347 1

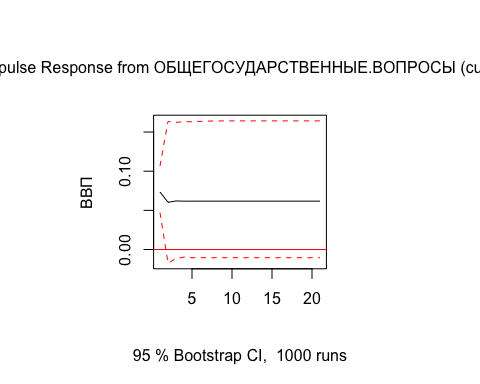
#summary(SVARMod)  
  
Imp3 <- irf(SVARMod, impulse = "ОБЩЕГОСУДАРСТВЕННЫЕ.ВОПРОСЫ", response = "ВВП", n.ahead = 20, ortho = TRUE, runs = 1000, cumulative = TRUE)

value:  
## 0.000250115624352551 .

## Warning in VAR(y = ysampled, p = 1, type = "const", exogen = oil): No column names supplied in exogen, using: exo1 , instead.

## Warning in SVAR(x = varboot, Bmat = bmat, hessian = TRUE, esmethod =  
## "scoring"): Convergence not achieved after 100 iterations. Convergence value:  
## 8.28674245158256e-05 .

plot(Imp3)



s = 0  
for(i in 1:4){  
 s = s + Imp3[1]$irf$ОБЩЕГОСУДАРСТВЕННЫЕ.ВОПРОСЫ[i]  
}  
s

## [1] 0.2574835

#################################################################################