Author:kkitonga and ewayagi

Date:15/08/2022

Vector Autoregression model

**Useful resources:https://www.youtube.com/c/JustinEloriaga (https://www.youtube.com/c/JustinEloriaga)

1.load libraries

```
library(AER)
```

```
## Warning: package 'AER' was built under R version 4.1.3
```

```
## Loading required package: car
```

```
## Loading required package: carData
```

```
## Loading required package: lmtest
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

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

```
## Loading required package: sandwich
```

```
## Warning: package 'sandwich' was built under R version 4.1.3
```

```
## Loading required package: survival
```

```
library(urca)
```

```
## Warning: package 'urca' was built under R version 4.1.3
```

```
library(DataExplorer)
```

```
## Warning: package 'DataExplorer' was built under R version 4.1.3
```

```
library(vars)
```

```
## Warning: package 'vars' was built under R version 4.1.3
```

```
## Loading required package: MASS
```

```
## Loading required package: strucchange
```

```
## Warning: package 'strucchange' was built under R version 4.1.3
```

```
library(tsDyn)
```

```
## Warning: package 'tsDyn' was built under R version 4.1.3
```

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

```
library(tseries)
```

```
## Warning: package 'tseries' was built under R version 4.1.3
```

```
library(lmtest)
```

2.Load data

```
#i.load data
data(USMacroG)
#ii.duplicating dataframe in new object
USmacroG1 <- na.omit(data.frame(USMacroG))
```

3.Descriptives

```
#i.first five observations
head(USMacroG)
```

```
##            gdp consumption invest government    dpi  cpi     m1 tbill unemp
## 1950 Q1 1610.5      1058.9  198.1      361.0 1186.1 70.6 110.20  1.12   6.4
## 1950 Q2 1658.8      1075.9  220.4      366.4 1178.1 71.4 111.75  1.17   5.6
## 1950 Q3 1723.0      1131.0  239.7      359.6 1196.5 73.2 112.95  1.23   4.6
## 1950 Q4 1753.9      1097.6  271.8      382.5 1210.0 74.9 113.93  1.35   4.2
## 1951 Q1 1773.5      1122.8  242.9      421.9 1207.9 77.3 115.08  1.40   3.5
## 1951 Q2 1803.7      1091.4  249.2      480.1 1225.8 77.6 116.19  1.53   3.1
##         population inflation interest
## 1950 Q1    149.461        NA       NA
## 1950 Q2    150.260    4.5071  -3.3404
## 1950 Q3    151.064    9.9590  -8.7290
## 1950 Q4    151.871    9.1834  -7.8301
## 1951 Q1    152.393   12.6160 -11.2160
## 1951 Q2    152.917    1.5494  -0.0161
```

```
#ii.last five observations
tail(USMacroG)
```

```
##            gdp consumption invest government    dpi   cpi     m1 tbill unemp
## 1999 Q3 8871.5      6000.0 1655.8     1533.2 6332.4 502.9 1093.4  4.70   4.2
## 1999 Q4 9049.9      6083.6 1725.4     1564.8 6379.2 504.1 1124.8  5.06   4.1
## 2000 Q1 9102.5      6171.7 1722.9     1560.4 6431.6 512.8 1113.7  5.54   4.0
## 2000 Q2 9229.4      6226.3 1801.6     1577.2 6523.7 516.5 1105.3  5.78   4.0
## 2000 Q3 9260.1      6292.1 1788.8     1570.0 6566.5 520.3 1096.0  6.03   4.1
## 2000 Q4 9303.9      6341.1 1778.3     1582.8 6634.9 521.1 1088.1  6.03   4.0
##         population inflation interest
## 1999 Q3    272.078    3.9968   0.6998
## 1999 Q4    272.691    0.9533   4.1067
## 2000 Q1    274.848    6.8445  -1.3012
## 2000 Q2    277.022    2.8758   2.9009
## 2000 Q3    279.213    2.9321   3.0979
## 2000 Q4    281.422    0.6146   5.4154
```

```
#iii.Variable type
str(USMacroG)
```

```
##  Time-Series [1:204, 1:12] from 1950 to 2001: 1610 1659 1723 1754 1774 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr [1:12] "gdp" "consumption" "invest" "government" ...
```

```
#iv.five point summary
summary(USMacroG)
```

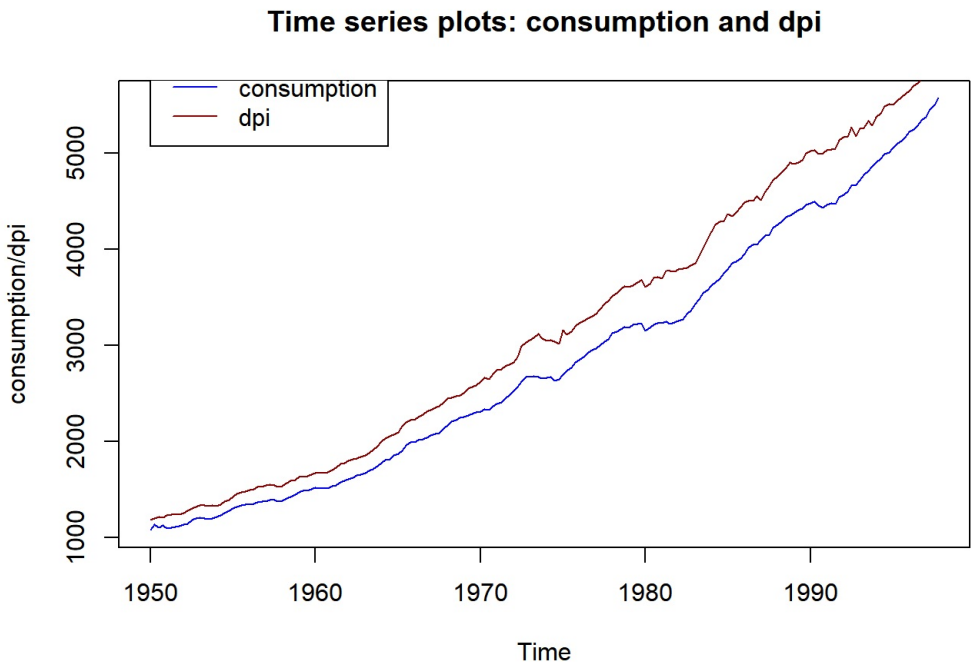
```
##       gdp        consumption       invest         government          dpi
##  Min.   :1610   Min.   :1059   Min.   : 197.7   Min.   : 359.6   Min.   :1178
##  1st Qu.:2602   1st Qu.:1640   1st Qu.: 309.3   1st Qu.: 740.6   1st Qu.:1822
##  Median :4142   Median :2715   Median : 568.5   Median : 952.0   Median :3133
##  Mean   :4563   Mean   :2999   Mean   : 652.3   Mean   : 997.0   Mean   :3341
##  3rd Qu.:6294   3rd Qu.:4235   3rd Qu.: 874.1   3rd Qu.:1300.8   3rd Qu.:4733
##  Max.   :9304   Max.   :6341   Max.   :1801.6   Max.   :1582.8   Max.   :6635
##
##       cpi               m1             tbill            unemp
##  Min.   : 70.60   Min.   : 110.2   Min.   : 0.810   Min.   : 2.600
##  1st Qu.: 91.15   1st Qu.: 147.5   1st Qu.: 3.087   1st Qu.: 4.400
##  Median :162.10   Median : 284.4   Median : 5.045   Median : 5.600
##  Mean   :225.82   Mean   : 453.9   Mean   : 5.229   Mean   : 5.675
##  3rd Qu.:350.12   3rd Qu.: 764.3   3rd Qu.: 6.645   3rd Qu.: 6.800
##  Max.   :521.10   Max.   :1152.1   Max.   :15.090   Max.   :10.700
##
##    population      inflation         interest
##  Min.   :149.5   Min.   :-2.530   Min.   :-11.2160
##  1st Qu.:185.6   1st Qu.: 1.762   1st Qu.: -0.1583
##  Median :214.7   Median : 3.138   Median :  1.5133
##  Mean   :214.0   Mean   : 3.939   Mean   :  1.3112
##  3rd Qu.:243.0   3rd Qu.: 5.591   3rd Qu.:  2.9155
##  Max.   :281.4   Max.   :16.864   Max.   : 10.6262
##                  NA's   :1        NA's   :1
```

4.Subsetting data

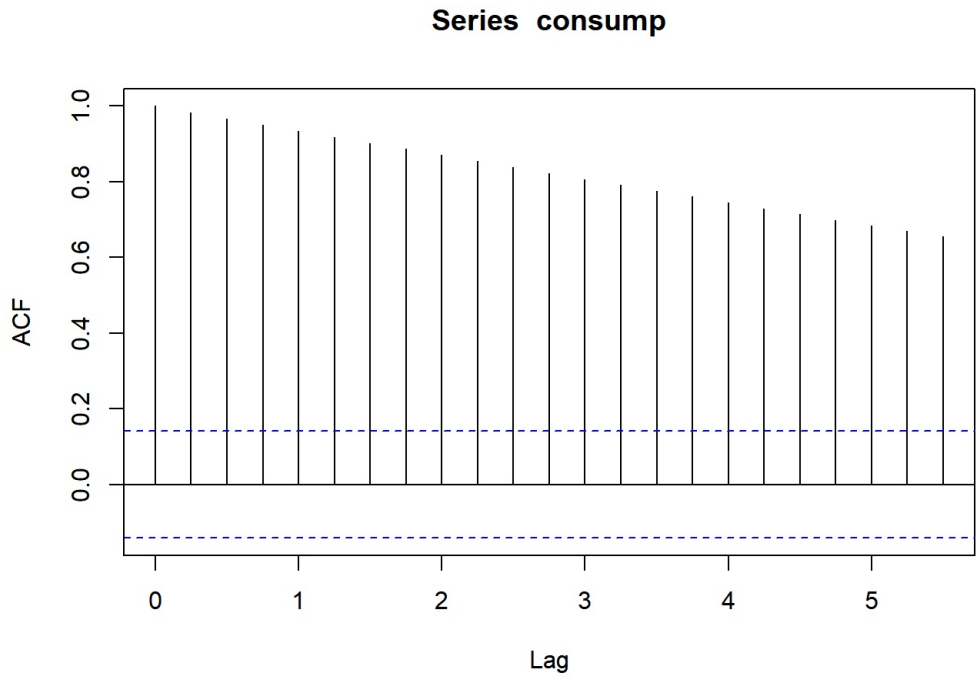
```
#1.creating a data frame of cpi and dpi
dpi <- ts(USmacroG1$dpi,start=c(1950,1),end=c(1997,4),frequency=4)

consump <- ts(USmacroG1$consump,start=c(1950,1),end=c(1997,4),frequency=4)

#2.Binding the two objects
df <- cbind(dpi,consump)
```

5.Visualization

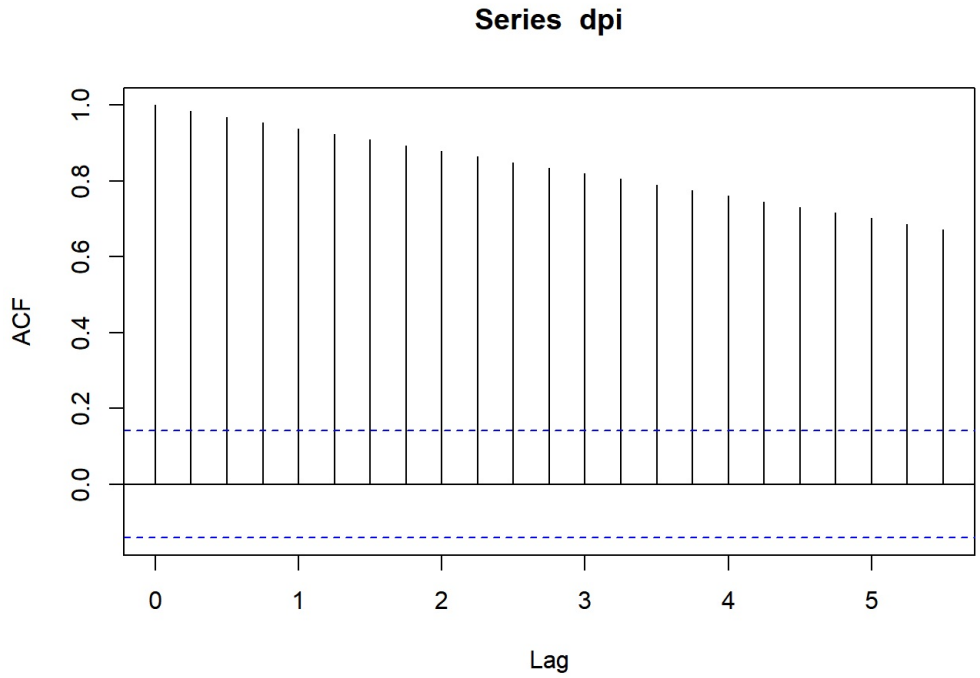
```
#1.Line Plots:Consumption and disposable personal income
plot(consump,
     ylab="consumption/dpi",
     col="blue",
     lty=1,
     main="Time series plots: consumption and dpi")
lines(dpi,col='dark red',lty=1)
legend(1950,6000,legend=c('consumption','dpi'),
       col=c('blue','dark red'),lty=c(1,1))
```




```
#2.Acf plots
acf(consump)
```




```
acf(dpi)
```



6.Stationarity testing

```
#1.Augmented Dickey Fuller test (consumption)
adf.test(consump,k=12)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  consump
## Dickey-Fuller = -0.4041, Lag order = 12, p-value = 0.9852
## alternative hypothesis: stationary
```

```
#2.Augmented Dickey Fuller test (dpi)
adf.test(dpi,k=12)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  dpi
## Dickey-Fuller = -1.5687, Lag order = 12, p-value = 0.7566
## alternative hypothesis: stationary
```

7.Differencing:Checking stationarity

```
#1.Differencing consumption then checking stationarity
adf.test(diff(consump,lag=1))
```

```
## Warning in adf.test(diff(consump, lag = 1)): p-value smaller than printed p-
## value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(consump, lag = 1)
## Dickey-Fuller = -4.6886, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

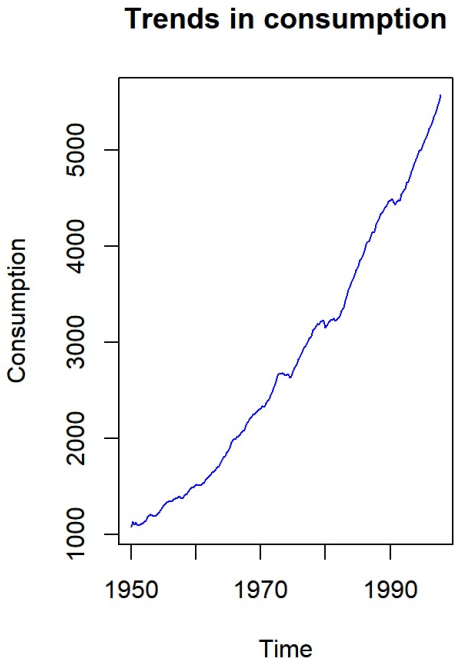
```
#2.Differencing dpi then checking stationarity
adf.test(diff(dpi,lag=1))
```

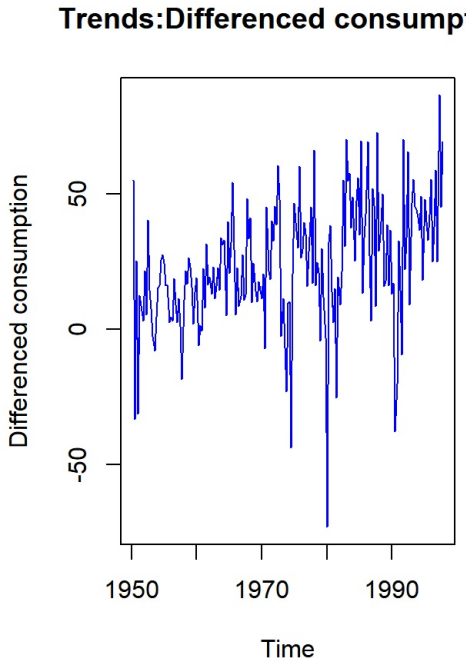
```
## Warning in adf.test(diff(dpi, lag = 1)): p-value smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(dpi, lag = 1)
## Dickey-Fuller = -5.7376, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

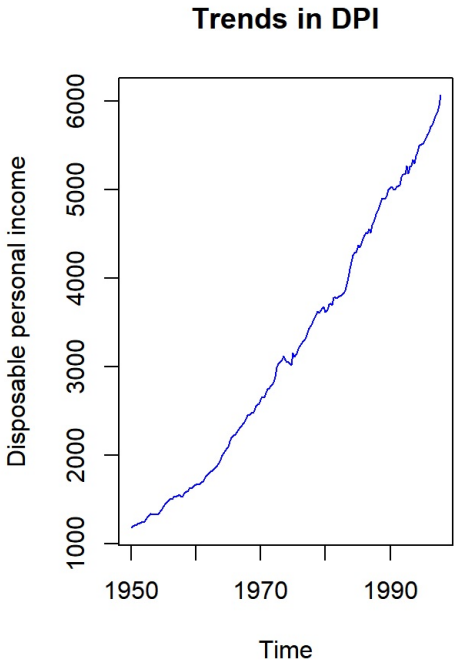
8.Plotting:level versus differenced data

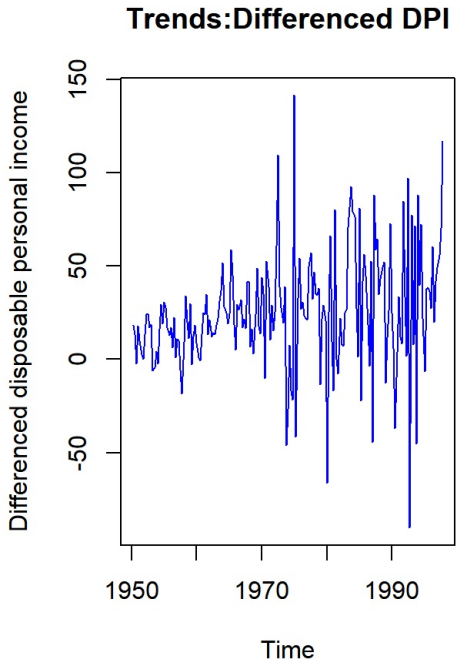
```
#i.Taking first differences and creating objects
consump_diff <-diff(consump,1)
dpi_diff <-diff(dpi,1)

#ii.Creating dataframe of differenced data
consump_dpi_diff <-cbind(consump_diff,dpi_diff)

#iii.Subplots:consumption versus differenced

par(mfrow=c(1,2))
plot(consump,
          ylab="Consumption",
          col="blue",
          lty=1,
          main="Trends in consumption")
plot(consump_diff,
     ylab="Differenced consumption",
     col="blue",
     lty=1,
     main="Trends:Differenced consumption")
```






```
par(mfrow=c(1,1))

#iv.Subplots:dpi versus differenced dpi
par(mfrow=c(1,2))
plot(dpi,
     ylab="Disposable personal income",
     col="blue",
     lty=1,
     main="Trends in DPI")
plot(dpi_diff,
     ylab="Differenced disposable personal income",
     col="blue",
     lty=1,
     main="Trends:Differenced DPI")
```






```
par(mfrow=c(1,1))
```

9.Granger causality

```
#i.consumption causing dpi
grangertest(consump,dpi,order=4)
```

```
## Granger causality test
##
## Model 1: dpi ~ Lags(dpi, 1:4) + Lags(consump, 1:4)
## Model 2: dpi ~ Lags(dpi, 1:4)
##   Res.Df Df      F    Pr(>F)
## 1    179
## 2    183 -4 6.7529 4.386e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#ii.dpi causing consumption
grangertest(dpi,consump,order=4)
```

```
## Granger causality test
##
## Model 1: consump ~ Lags(consump, 1:4) + Lags(dpi, 1:4)
## Model 2: consump ~ Lags(consump, 1:4)
##   Res.Df Df      F Pr(>F)
## 1    179
## 2    183 -4 0.8103 0.5201
```

10.Training versus testing dataset

```
#I.Create training dataset
con_diff_tr <- diff(USmacroG1$consumption,lag=1)
dpi_diff_tr <- diff(USmacroG1$dpi,lag=1)
train_df <- cbind(con_diff_tr,dpi_diff_tr)

#ii.Create test dataset
con_diff <- diff(USmacroG1$consumption,lag=1)
dpi_diff <- diff(USmacroG1$dpi,lag=1)

#iii.training dataset
test_df <- data.frame(cbind(con_diff[191:202],dpi_diff[192:201]))
```

```
## Warning in cbind(con_diff[191:202], dpi_diff[192:201]): number of rows of result
## is not a multiple of vector length (arg 2)
```

11.Lag optimality

```
#i.Information on VAR select syntax
??VARselect
```

```
## starting httpd help server ... done
```

```
#ii.Establishing optimal lag
optimal_lag <-VARselect(consump_dpi_diff, lag.max = 10)
#iii.Getting information on selection criteria
optimal_lag$selection
```

```
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      3      3      1      3
```

12.VAR model

```
 #i.Information on var model
??VAR

#ii.running var model
var_model <- VAR(consump_dpi_diff, p = 3, type = c("const",                  "trend", "both", "none"),season = NU
LL,                    exogen = NULL, lag.max = NULL,
                ic = c("AIC", "HQ", "SC", "FPE"))

#iii.Output of var model
summary(var_model)
```

```
##
## VAR Estimation Results:
## ========================
## Endogenous variables: consump_diff, dpi_diff
## Deterministic variables: const
## Sample size: 188
## Log Likelihood: -1701.03
## Roots of the characteristic polynomial:
## 0.7937 0.4556 0.4556 0.3957 0.3957 0.1818
## Call:
## VAR(y = consump_dpi_diff, p = 3, type = c("const", "trend", "both",
##     "none"), exogen = NULL, lag.max = NULL, ic = c("AIC", "HQ",
##     "SC", "FPE"))
##
##
## Estimation results for equation consump_diff:
## =============================================
## consump_diff = consump_diff.l1 + dpi_diff.l1 + consump_diff.l2 + dpi_diff.l2 + consump_diff.l3 + dpi_diff.l3 +
const
##
##                 Estimate Std. Error t value Pr(>|t|)
## consump_diff.l1  0.14489    0.08266   1.753 0.081347 .
## dpi_diff.l1      0.03766    0.06029   0.625 0.533001
## consump_diff.l2  0.27509    0.08401   3.274 0.001269 **
## dpi_diff.l2     -0.05487    0.06294  -0.872 0.384552
## consump_diff.l3  0.29460    0.08395   3.509 0.000567 ***
## dpi_diff.l3     -0.08739    0.05936  -1.472 0.142734
## const            9.79053    2.67816   3.656 0.000336 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 20.32 on 181 degrees of freedom
## Multiple R-Squared: 0.2233,  Adjusted R-squared: 0.1976
## F-statistic: 8.674 on 6 and 181 DF,  p-value: 2.653e-08
##
##
## Estimation results for equation dpi_diff:
## =========================================
## dpi_diff = consump_diff.l1 + dpi_diff.l1 + consump_diff.l2 + dpi_diff.l2 + consump_diff.l3 + dpi_diff.l3 + con
st
##
##                 Estimate Std. Error t value Pr(>|t|)
## consump_diff.l1  0.49112    0.11715   4.192 4.32e-05 ***
## dpi_diff.l1     -0.26231    0.08544  -3.070 0.002470 **
## consump_diff.l2  0.18123    0.11906   1.522 0.129706
## dpi_diff.l2     -0.05415    0.08920  -0.607 0.544563
## consump_diff.l3  0.30864    0.11897   2.594 0.010253 *
## dpi_diff.l3     -0.07563    0.08413  -0.899 0.369839
## const           12.90588    3.79540   3.400 0.000828 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 28.8 on 181 degrees of freedom
## Multiple R-Squared: 0.1837,  Adjusted R-squared: 0.1567
## F-statistic:  6.79 on 6 and 181 DF,  p-value: 1.652e-06
##
##
##
## Covariance matrix of residuals:
##              consump_diff dpi_diff
## consump_diff        413.0    274.3
## dpi_diff            274.3    829.4
##
## Correlation matrix of residuals:
##              consump_diff dpi_diff
## consump_diff       1.0000   0.4687
## dpi_diff           0.4687   1.0000
```

```
#iv.generating forecasts
forecast <-predict(var_model,n.ahead = 12, ci = 0.95,                          dumvar = NULL)
```

13.Diagnostics

```
              #i.serial autocorellation
autocorr <- serial.test(var_model,lags.pt = 16,
                        type = c("PT.asymptotic"))
#autocorrelation result
autocorr
```

```
##
##  Portmanteau Test (asymptotic)
##
## data:  Residuals of VAR object var_model
## Chi-squared = 54.155, df = 52, p-value = 0.3922
```

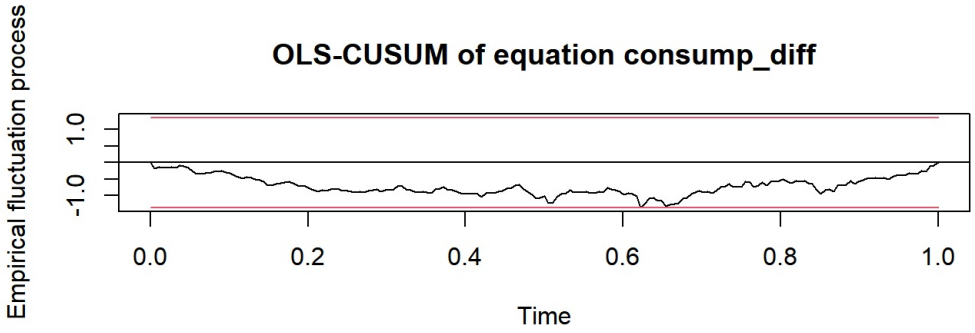
```
            #ii.Heteroskedasticity
hetesk <- arch.test(var_model,lags.multi=5, multivariate.only=TRUE)
hetesk
```

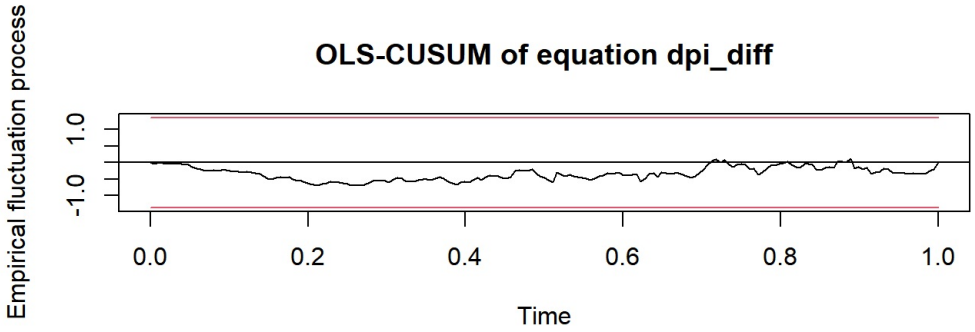
```
##
##  ARCH (multivariate)
##
## data:  Residuals of VAR object var_model
## Chi-squared = 93.496, df = 45, p-value = 2.987e-05
```

```
             #iii.Normality :Residual distribution
normality <- normality.test(var_model,
                            multivariate.only=TRUE)
normality
```

```
## $JB
##
##  JB-Test (multivariate)
##
## data:  Residuals of VAR object var_model
## Chi-squared = 111.42, df = 4, p-value < 2.2e-16
##
##
## $Skewness
##
##  Skewness only (multivariate)
##
## data:  Residuals of VAR object var_model
## Chi-squared = 11.743, df = 2, p-value = 0.002819
##
##
## $Kurtosis
##
##  Kurtosis only (multivariate)
##
## data:  Residuals of VAR object var_model
## Chi-squared = 99.679, df = 2, p-value < 2.2e-16
```

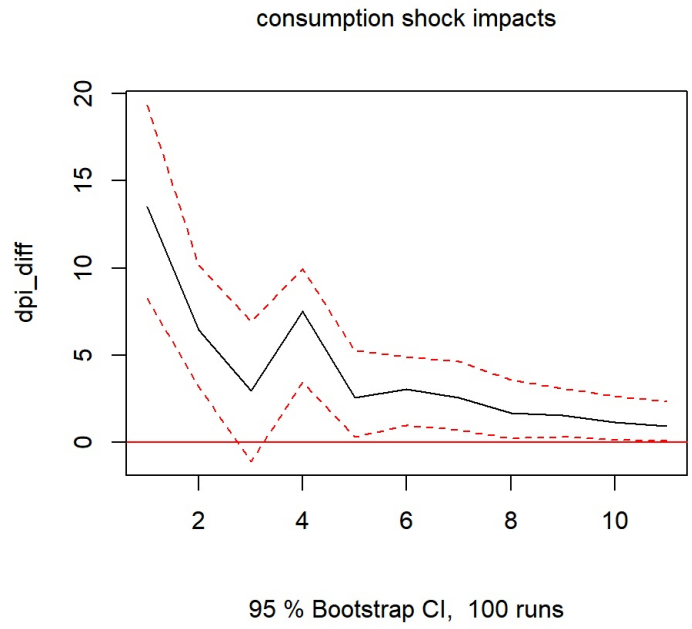
```
var_stab <-stability(var_model, type = c("OLS-CUSUM"))
plot(var_stab)
```



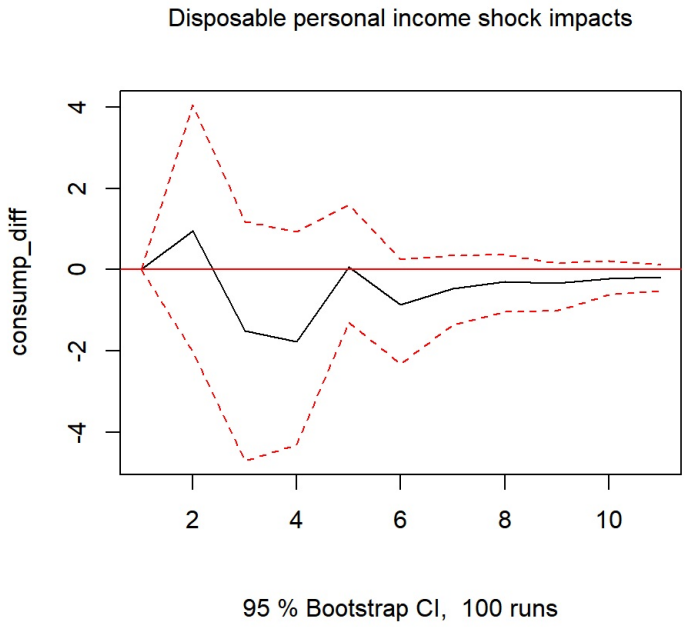


14.Impulse response functions

```
#i.Consumption shock

con_dpi_irf <- irf(var_model,impulse='consump_diff',response='dpi_diff',n.ahead=10)
plot(con_dpi_irf,main="consumption shock impacts")
```

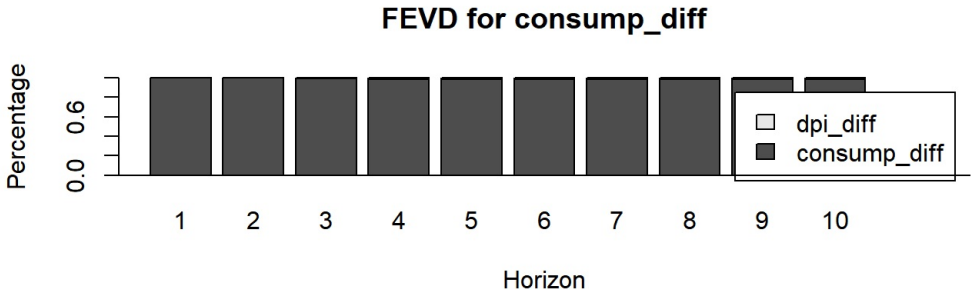


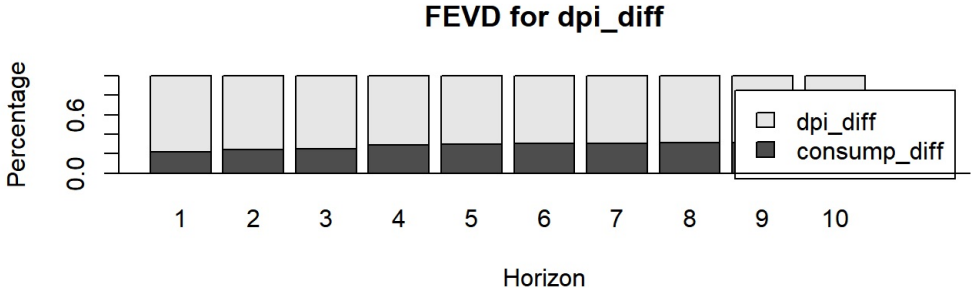

```
#ii.Dpi shock

dpi_con_irf <- irf(var_model,impulse='dpi_diff',response='consump_diff',n.ahead=10,boot=TRUE)
plot(dpi_con_irf,main= "Disposable personal income shock impacts")
```



15.Variance decomposition

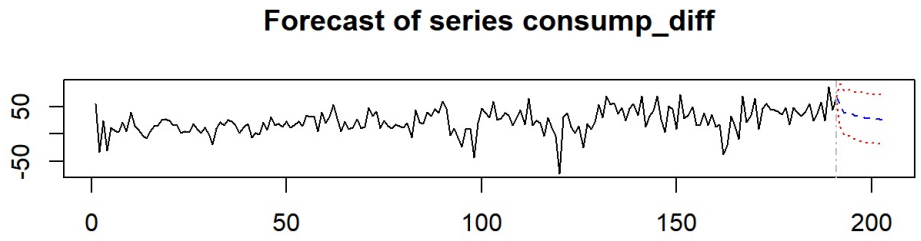
```
#I.About fevd
??fevd

#ii.Variance decomposition
var_fevd <- fevd(var_model,n.ahead=10)
plot(var_fevd)
```

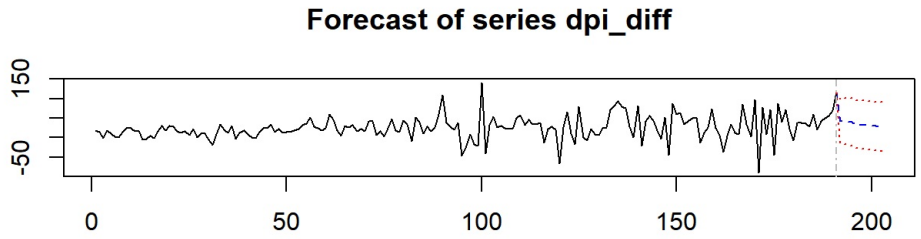




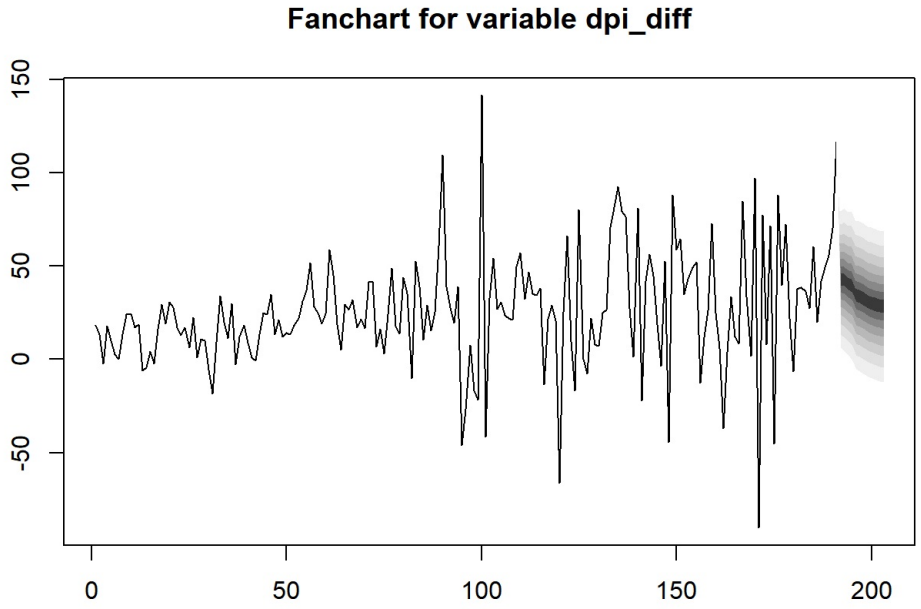
16.Forecasting

```
#i.About forecast synatx
??forecast

#ii.Forecast syntax and plot
forecast <- predict(var_model,n.ahead=12,ci=0.95)
plot(forecast)
```

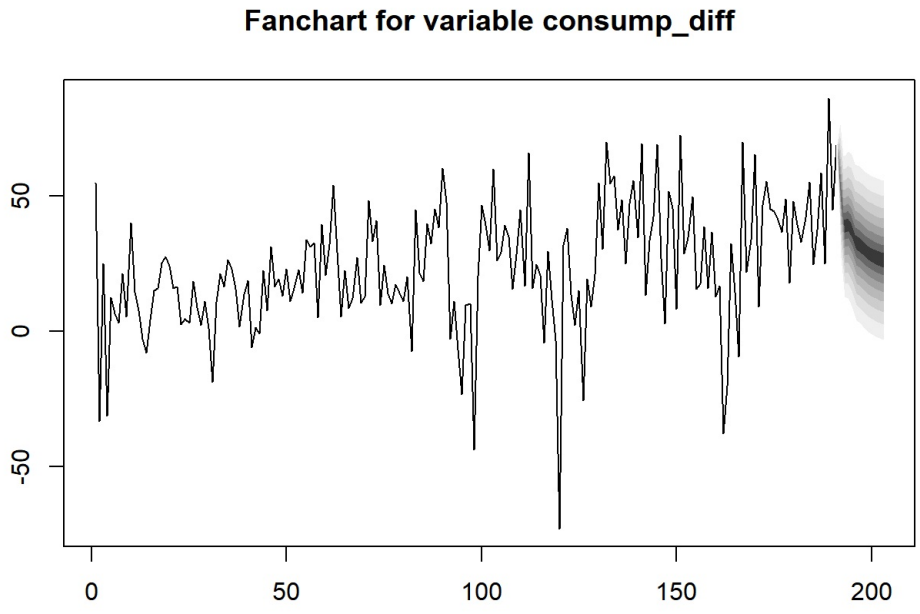





```
#iii.Plotting a fanchart
fanchart(forecast,names='dpi_diff')
```




```
fanchart(forecast,names='consump_diff')
```



END