1. (25 pts) Consider the following ARMA(2,1) time series model where $\varepsilon_t$ follows a white noise process:

$$y_t = 2 - 0.8y_{t-1} - 0.1y_{t-2} + \varepsilon_t + 0.8\varepsilon_{t-1} \qquad (1)$$

(a) (10 pts) Find the set of homogeneous solutions. Explain the shape of this solution. Is it stable?

*The homogeneous equation is given by*

$$y_t + 0.8y_{t-1} + 0.1y_{t-2} = 0$$

.

*The full set of homogeneous solutions is given by* $y_t^h = A_1\alpha_1^t + A_2\alpha_2^t$. *The characteristic equation is given by:*

$$\alpha^t + 0.8\alpha + 0.1 = 0$$

*The two roots are:*

$$\alpha_1 = \frac{-0.8 + \sqrt{0.8^2 - 4 * 0.1}}{2} = -0.155051$$

$$\alpha_2 = \frac{-0.8 - \sqrt{0.8^2 - 4 * 0.1}}{2} = -0.644949$$

*The final set of homogenous solution is then:*

$$\boxed{y_t^h = A_1(-0.155051)^t + A_2(-0.644949)^t}$$

*Since both roots are inside the unit circle, the difference equation is stable. Also, given that both roots are negative, the solution will oscillate between negative numbers and positive numbers and the size of these oscillations will decreases as t increases. When t is even, the solution features positive values, and when t is odd it features negative values.*

(b) (5 pts) Is (1) stationary and invertible? Explain.

*It is stationary because it is stable. If the roots of the characteristic equation are inside the unit circle, then the unconditional mean, variance and autocovariance will be time independent (weak stationarity). Looking at the moving average component, we have a MA(1) which means that if the root of the characteristic equation of this component is inside the unit circle, the process is invertible. In this case, the root is -0.8 which means its inside the unit circle and the process is thus invertible.*

(c) (5 pts) Given $y_0 = 1$ and $y_1 = 0.2$, use the iteration method to find the solution to $y_3$. Is the solution of the form of a constant plus past shocks? Does $y_3$ depend on $\varepsilon_2$? Explain why or why not.

$$y_2 = 2 - 0.8 * 0.2 - 0.1 * 1 + \varepsilon_2 + 0.8\varepsilon_1$$
$$y_2 = 1.74 + \varepsilon_2 + 0.8\varepsilon_1$$
$$y_3 = 2 - 0.8 * y_2 - 0.1 * 0.2 + \varepsilon_3 + 0.8\varepsilon_2$$
$$y_3 = 2 - 0.8 * (1.74 + \varepsilon_2 + 0.8\varepsilon_1) - 0.1 * 0.2 + \varepsilon_3 + 0.8\varepsilon_2$$
$$y_3 = 0.588 - \cancel{0.8\varepsilon_2} - 0.64\varepsilon_1 + \varepsilon_3 + \cancel{0.8\varepsilon_2}$$
$$y_3 = 0.588 - 0.64\varepsilon_1 + \varepsilon_3$$

*The reason why* $\varepsilon_2$ *does not affect* $y_3$ *is the interaction between the autoregressive component and the moving average component. For a variable* $y_t$*, the autoregressive coefficient of -0.8 will always cancel out the +0.8 coefficient of the moving average component for shocks at* $t-1$.

(d) (5 pts) Compute the ACF for the first and second lag. How does the moving average component affect the ACF for the first two lags?

*THe demeaned equation is given by*

$$y_t - \mu = -0.8(y_{t-1} - 1) - 0.1(y_{t-1} - \mu) + \varepsilon_t + 0.8\varepsilon_{t-1}$$

*Next, multiply both sides by* $(y_t - \mu)$*, take expectations, and use the definition of covariance to find:*

$$\gamma_0 = -0.8\gamma_1 - 0.1\gamma_2 + \sigma^2 + \cancel{0.8 * 0.8\sigma^2} - \cancel{0.8 * 0.8\sigma^2}$$
$$\gamma_0 = -0.8\gamma_1 - 0.1\gamma_2 + \sigma^2$$

*then, multiply both sides by* $(y_{t-1} - \mu)$*, take expectations, and use the definition of covariance to find*

$$\gamma_1 = -0.8\gamma_0 - 0.1\gamma_1 + 0.8\sigma^2$$

*finally, multiply both sides by* $(y_{t-2} - \mu)$*, take expectations, and use the definition of covariance to find*

$$\gamma_2 = -0.8\gamma_1 - 0.1\gamma_0$$

*Solve the system of 3 equations with three unknowns* $(\gamma_0, \gamma_1, \gamma_2)$*: Replace third equation in first equation:*

$$\gamma_0 = -0.8\gamma_1 - 0.1(-0.8\gamma_1 - 0.1\gamma_0) + \sigma^2$$
$$\gamma_0 = -0.72\gamma_1 + 0.01\gamma_0 + \sigma^2$$

*Next, solve second equation with respect to* $\gamma_1$

$$\gamma_1 = -\frac{0.8}{1.1}\gamma_0 + \frac{0.8}{1.1}\sigma^2$$

*Now replace for* $\gamma_1$*:*

$$\gamma_0 = -0.72(-\frac{0.8}{1.1}\gamma_0 + \frac{0.8}{1.1}\sigma^2) + 0.01\gamma_0 + \sigma^2$$
$$\gamma_0 = \frac{0.524}{0.513}\sigma^2$$

*Now, we know the autocorrelation is* $\rho_j = \frac{\gamma_j}{\gamma_0}$*. Use second and third equations of the system together with the ACF definition to find that:*

$$\rho_1 = -0.8 - 0.1\rho_1 + \frac{0.8\sigma^2}{\gamma_0}$$

$$\rho_1 = -\frac{0.8}{1.1} + \frac{0.8\cancel{\sigma^2}0.513}{1.1\cancel{\sigma^2}0.524} \approx 0.0153$$
$$\rho_2 = -0.8\rho_1 - 0.1 \approx -0.0878$$

*The the moving average component affects directly the ACF at the first lag by the following amount* $\frac{0.8*0.513}{1.1*0.524} \approx 0.712$ *And affects ACF at the second lag only indirectly by the following amount* $-0.8 * 0.712 \approx -0.570$

*The intuition is that the direct positive effect of the MAC actually offsets the negative effect of the AR component for the first autocorrelation. And indirectly by interacting with the AR components it contributes negatively to the ACF at the second lag.*

2. (25 pts) Suppose we have the following VAR with two variables, $y_t$ and $z_t$:

$$y_t = 0.1 + 0.1y_{t-1} + 0.4z_{t-1} + e_{1t}$$
$$z_t = 0.8 - 0.3y_{t-1} + e_{2t}$$

and that the residual variance-covariance matrix is given by:

$$\Sigma = \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 2 \end{bmatrix}$$

(a) (2 pts) Find the two steps-ahead conditional forecast of $y_t$, $E_t[y_{t+2}]$. Discuss how $z_t$ affects $E_t[y_{t+2}]$.

$$\begin{aligned} E_t[y_{t+1}] &= 0.1 + 0.1y_t + 0.4z_t \\ E_t[z_{t+1}] &= 0.8 - 0.3y_t \end{aligned}$$

*and*

$$\begin{aligned} E_t[y_{t+2}] &= 0.1 + 0.1E_t[y_{t+1}] + 0.4E[z_{t+1}] \\ E_t[z_{t+2}] &= 0.8 - 0.3E[y_{t+1}] \end{aligned}$$

*hence,*

$$E_t[y_{t+2}] = 0.1 + 0.1(0.1 + 0.1y_t + 0.4z_t) + 0.4(0.8 - 0.3y_t)$$
$$E_t[y_{t+2}] = 0.43 - 0.11y_t + 0.04z_t$$

*The effect of $z_t$ on $E_t[y_{t+1}]$ is only indirect and works it operates through the one-step ahead of $E_t[y_{t+1}]$. Note that $z_t$ does NOT affect $E_t[z_{t+1}]$.*

(b) (13 pts) Assume $z_t$ does not affect contemporaneously $y_t$ (recursive identification approach). Show your matrix $B$ after imposing this restriction and find all coefficients of the structural VAR (including $\sigma_y$ and $\sigma_z$).

*If we impose the recursive identification restrictions we get the following $B$ matrix:*

$$B = \begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix}$$

*The structural shocks are related to the reduced form shocks in the following way:*

$$B\varepsilon = e$$

*Take variance on both sides:*

$$B\Sigma_\varepsilon B' = \Sigma$$
$$\begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_z^2 \end{bmatrix} \begin{bmatrix} 1 & b_{21} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 2 \end{bmatrix}$$
$$\begin{bmatrix} \sigma_y^2 & b_{21}\sigma_y^2 \\ b_{21}\sigma_y^2 & (b_{21})^2\sigma_y^2 + \sigma_z^2 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 2 \end{bmatrix}$$

$$\begin{aligned} \sigma_y^2 &= 0.5 \\ b_{21}\sigma_y^2 &= 0.4 \\ (b_{21})^2\sigma_y^2 + \sigma_z^2 &= 2 \end{aligned}$$

*Solving this system one finds $\sigma_y^2 = 0.5$, $b_{21} = 0.8$, $\sigma_z^2 = 1.68$*

*So we have recovered three structural parameters (i.e. the structural variances and the b coefficient). Now we just need to recover the* $\Gamma_0$ *and* $\Gamma_1$ *matrices.*

$$\Gamma_0 = BA_0$$

*and*

$$\Gamma_1 = BA_1$$

*so*

$$\Gamma_0 = \begin{bmatrix} 1 & 0 \\ 0.8 & 1 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.8 \end{bmatrix} = \begin{bmatrix} 0.1 \\ 0.88 \end{bmatrix}$$

$$\Gamma_1 = \begin{bmatrix} 1 & 0 \\ 0.8 & 1 \end{bmatrix} \begin{bmatrix} 0.1 & 0.4 \\ -0.3 & 0 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.4 \\ -0.22 & 0.32 \end{bmatrix}$$

*The structural model is then given by*

$$\begin{aligned} y_t &= 0.1 + 0.1 y_{t-1} + 0.4 z_{t-1} + \varepsilon_{yt} \\ z_t &= 0.88 - 0.8 y_t - 0.22 y_{t-1} + 0.32 z_{t-1} + \varepsilon_{zt} \end{aligned}$$

(c) (5 pts) Under the restriction imposed in b), compute the IRF for both variables to a one standard deviation shock in $z_t$ at impact and step 1. Interpret them. What if the size of the shock doubles. What are the responses at impact and step 1?

*Let* $x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}$. *Casting the model in its infinite lag MA representation we have:*

$$x_t = \mu + e_t + A_1 e_{t-1} + \ldots$$

*Which in terms of structural shocks is:*

$$x_t = \mu + B^{-1}\varepsilon_t + A_1 B^{-1}\varepsilon_{t-1} + \ldots$$

*So the IRF matrix at step j is given by* $\phi_j = A_1^j B^{-1}$, *from which we are interested in computing* $\phi_0$ *and* $\phi_1$.

$$B^{-1} = \begin{bmatrix} 1 & 0 \\ 0.8 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -0.8 & 1 \end{bmatrix}$$

*So,*

$$\begin{aligned} \phi_0 &= B^{-1} = \begin{bmatrix} 1 & 0 \\ -0.8 & 1 \end{bmatrix} \\ \phi_1 &= A_1 B^{-1} = \begin{bmatrix} 0.1 & 0.4 \\ -0.3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -0.8 & 1 \end{bmatrix} = \begin{bmatrix} -0.22 & 0.4 \\ -0.3 & 0 \end{bmatrix} \end{aligned}$$

*Now, the IRF has the responses of variables in rows to 1 unit shocks in columns. But what we want to know what is what is the responses of all variables to a 1 unit standard deviation shock in* $z_t$. *The variance of* $z_t$ *is* $\sigma_z^2 = 1.68$. *So um standard deviation is* $\sigma_z = \sqrt{1.68} = 1.296148$. *So we need to scale the second column of our IRFs by 1.296148. The responses of y and z at impact to a one standard deviation shock in z are then given:*

$$1.296148 \frac{\partial y_t}{\partial \varepsilon_{zt}} = 1.296148 \phi_0(1,2) = 0$$

$$1.296148 \frac{\partial z_t}{\partial \varepsilon_{zt}} = 1.296148 \phi_0(2,2) = 1.296148$$

Cont.

*And at one-step:*

$$1.296148\frac{\partial y_{t+1}}{\partial \varepsilon_{zt}} = 1.296148\phi_1(1,2) = 0.5184592$$
$$1.296148\frac{\partial z_{t+1}}{\partial \varepsilon_{zt}} = 1.296148\phi_1(2,2) = 0$$

*Since the IRFs are linear, we can scale them by any constant $C$. If the size of the shock doubles $1.296148 * 2 = 2.592296$, we just need to double the previously computed IRFS. Hence,*

$$2.592296\frac{\partial y_t}{\partial \varepsilon_{zt}} = 2.592296\phi_0(1,2) = 0$$
$$2.592296\frac{\partial z_t}{\partial \varepsilon_{zt}} = 2.592296\phi_0(2,2) = 2.592296$$

*And at one-step:*

$$2.592296\frac{\partial y_{t+1}}{\partial \varepsilon_{zt}} = 2.592296\phi_1(1,2) = 1.036918$$
$$2.592296\frac{\partial z_{t+1}}{\partial \varepsilon_{zt}} = 2.592296\phi_1(2,2) = 0$$

(d) (5 pts) Compute the FEVD of $y_t$ at impact and step 1. How do your findings in terms of FEVD complement your conclusions in c)?

*Again using the MA representation, the structural shocks in to reduced form shocks on impact imply $x_{t+1} - E_t[x_{t+1}] = \phi_0\varepsilon_{t+1}$.*

$$\sigma_y(1)^2 = \phi_0(1,1)^2\sigma_y^2 + \phi_0(1,2)^2\sigma_z^2$$
$$\sigma_y(1)^2 = \sigma_y^2$$

*So the share of variation of $y$ that comes from $y$ is 100% and from $z$ is 0% which is expected given our recursive assumption that $y$ is only affected by itself contemporaneously.*

*Now at the one-step (two-steps in forecast), $x_{t+2} - E_t[x_{t+2}] = \phi_0\varepsilon_{t+2} + \phi_1\varepsilon_{t+1}$.*

$$\sigma_y(2)^2 = (\phi_0(1,1)^2 + \phi_1(1,1)^2)\sigma_y^2 + (\phi_0(1,2)^2 + \phi_1(1,2)^2)\sigma_z^2$$
$$\sigma_y(2)^2 = (1 + (-0.22)^2)0.5 + ((0.4)^2)1.68 = 0.793$$

*So the share of variation in $y$ at step-one that comes from itself is approximately 66.1% and from $z$ is 33.9%. While the IRFs tells us how variables move after the system is hit by one-time shock to a particular variable, the forecast error variance decomposition tells us the proportion of the movements in a sequence due to its own shocks versus shocks to the other variable. Hence, it complements the IRF analysis by telling us how relevant the IRFs are once we allow all shocks to hit the system at the same time.*

3. (25 pts) ARMA in practice. Use the library quantmod to get Amazon's stock price data. Use getSymbols("AMZN", src="yahoo", from="2007-12-01", to="2020-03-13").

(a) (2 pts) Plot the daily close price data. Does it look stationary? Plot its ACF, does your conclusion change?

(b) (5 pts) Compute the Amazon daily stock returns. Take log of the close date and take the first difference. Plot the daily returns, do they look stationary? Plot its ACF, does it change your conclusion? Now formally test for the unit root of the close price in levels and for the daily returns. Are they stationary?

(c) (5 pts) Fit an ARMA(2,1) with an intercept to the daily Amazon stock returns. What is the unconditional mean of the Amazon daily stock return in percentage?

(d) (3 pts) Check the model adequacy with the ACF plot of the residuals and the Ljung-Box test with 10 lags.

(e) (5 pts) Apply the as.ts() to your xts return series. Re-estimate the ARMA(2,1) to this new formatted series. Load the Forecast library and forecast the stock return for the next 5 days. Would you recommend buying the Amazon stock? Why or why not.

(f) (5 pts) How do the AR and MA components estimated coefficients relate to your forecasts in e)?

4. (25 pts) VAR in practice. Use the library readxl to load the dataset data_midterm_2020.xlsx. This dataset consists of quarterly Portuguese data on 5 variables in volume: GDP, household private consumption, government consumption, investment and taxes on products.

   (a) (5 pts) Get the data ready for analysis. 1) Remove the first column of the data; 2) apply the ts() function to the data in order to get the data in time series format; 3) put the data in logs and multiply by 100; 4) Take the first difference to all variables; 5) put the variables of the data in the following order: G, TAX, C, I, GDP.

   (b) (5 pts) Fit a VAR model to the transformed data. Choose the lags based on the Schwarz information criterion (SC). Check if the VAR model is adequate (use an ADF test on all 5 residuals series).

   (c) (10 pts) Compute the IRFs (20 steps) and plot them. What do you learn from them in terms of how a government expenditure shock affects the economy?

   (d) (5 pts) Compute the FEVD (20 steps). How relevant, in relative terms, are G shocks in explaining variation in GDP ?

   (e) (5 pts, Bonus points, challenge question for the bold!) Create a new dataset in memory that removes the last observation of the dataset you were working on (removes observation 99, note that you may need to apply the ts() function to the new dataset again). Re-estimate a VAR model to this subsample and forecast one step-ahead GDP. Now fit an ARMA model with the auto.arima() function to GDP series alone and make also a one step-ahead forecast. Compare both forecasts (VAR and ARMA) to the original observation you removed. Which forecast is better? Explain.

# Midterm_2020_Q3-4_solutions

March 29, 2020

## 0.1 Question 3

[36]:
```
library(quantmod)
library(tseries)
library(forecast)
library(xts)
```

[38]:
```
Amazon = getSymbols("AMZN", src="yahoo",from="2007-12-01", to="2020-03-13")
```

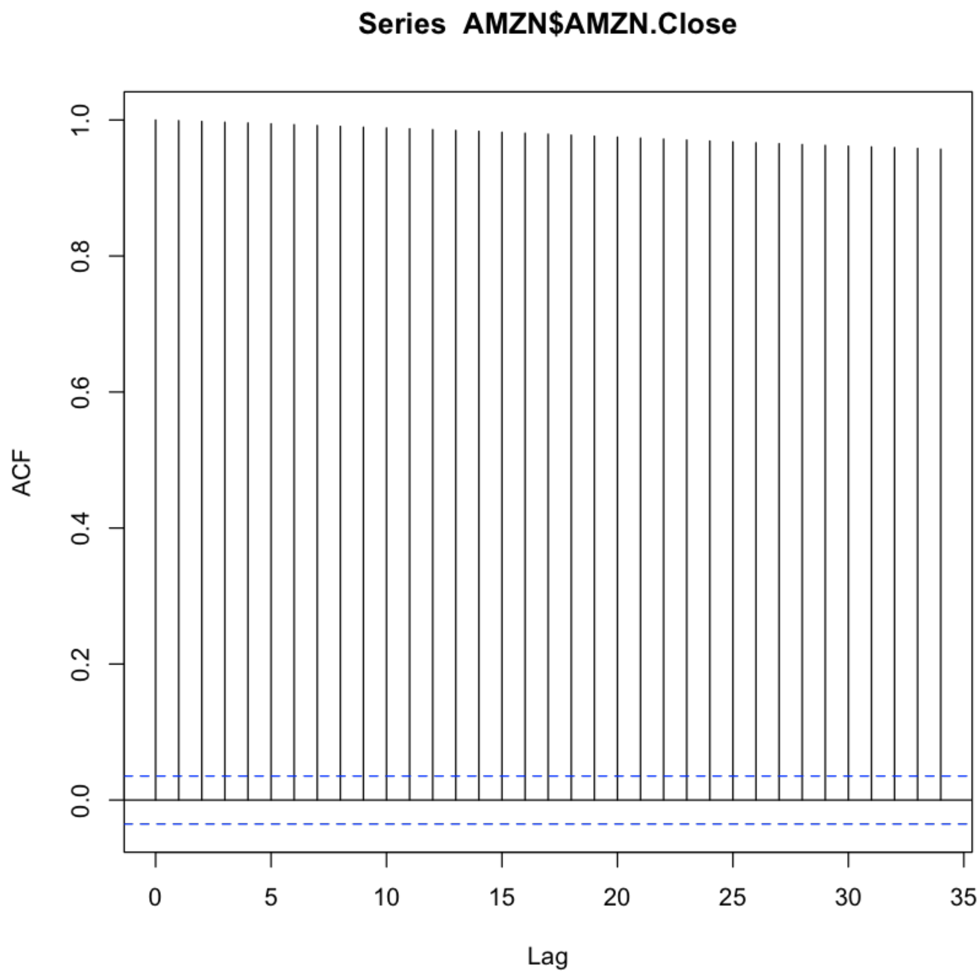
```
Warning message:
"'indexClass<-' is deprecated.
Use 'tclass<-' instead.
See help("Deprecated") and help("xts-deprecated")."
```

### 0.1.1 a)

[39]:
```
plot(AMZN$AMZN.Close)

acf(AMZN$AMZN.Close)
```


AMZN$AMZN.Close
2007-12-03 / 2020-03-12
2000
1500
1000
500
Dec 03 2007
Jun 01 2009
Dec 01 2010
Jun 01 2012
Dec 02 2013
Jun 01 2015
Dec 01 2016
Jun 01 2018
Nov 29 2019

2

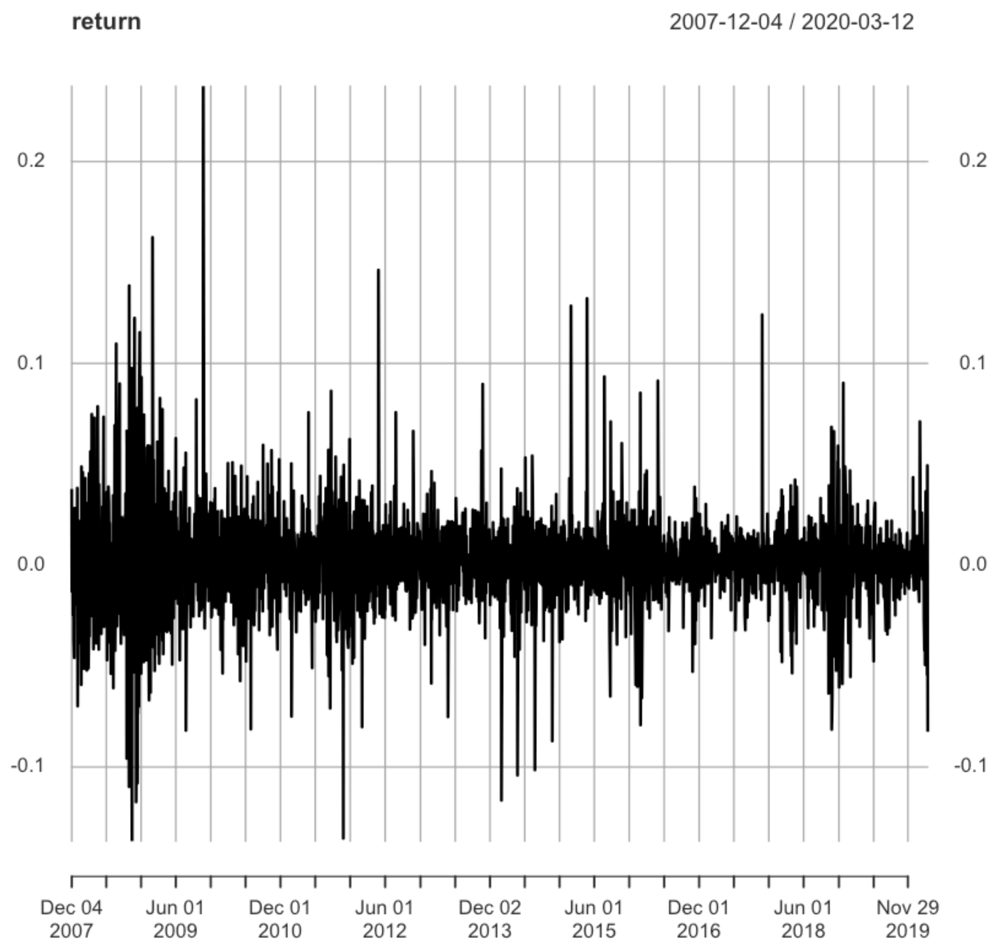



The plot shows a trend whichs makes the series nonstationary. Moreover, the ACF shows an extremly small decay which confirms our suspicion of nonstationarity of the series.

#### 0.1.2 b)

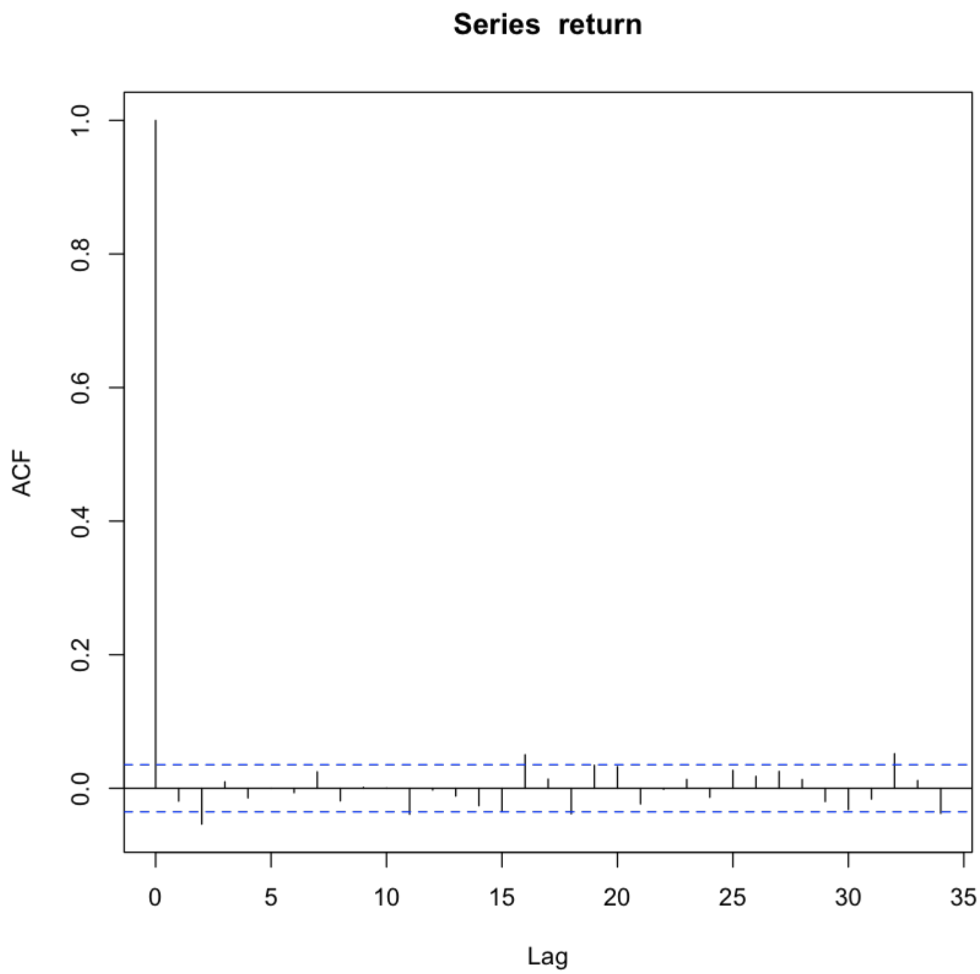
[40]:
```
return = diff(log(AMZN$AMZN.Close))
return = return[return$AMZN.Close != 'NA', ]

plot(return)
acf(return)

adf.test(AMZN$AMZN.Close)
```

```r
adf.test(return)
```




```
Augmented Dickey-Fuller Test

data:  AMZN$AMZN.Close
Dickey-Fuller = -1.7044, Lag order = 14, p-value = 0.7035
alternative hypothesis: stationary


Warning message in adf.test(return):
"p-value smaller than printed p-value"

Augmented Dickey-Fuller Test
```

```
data:  return
Dickey-Fuller = -16.017, Lag order = 14, p-value = 0.01
alternative hypothesis: stationary
```




The return is stationary. The plot suggests the series first and second moments are constant overtime. The ACF shows a quick decay to zero and the ADF rejects the null hypothesis of nonstationarity.

### 0.1.3 c)

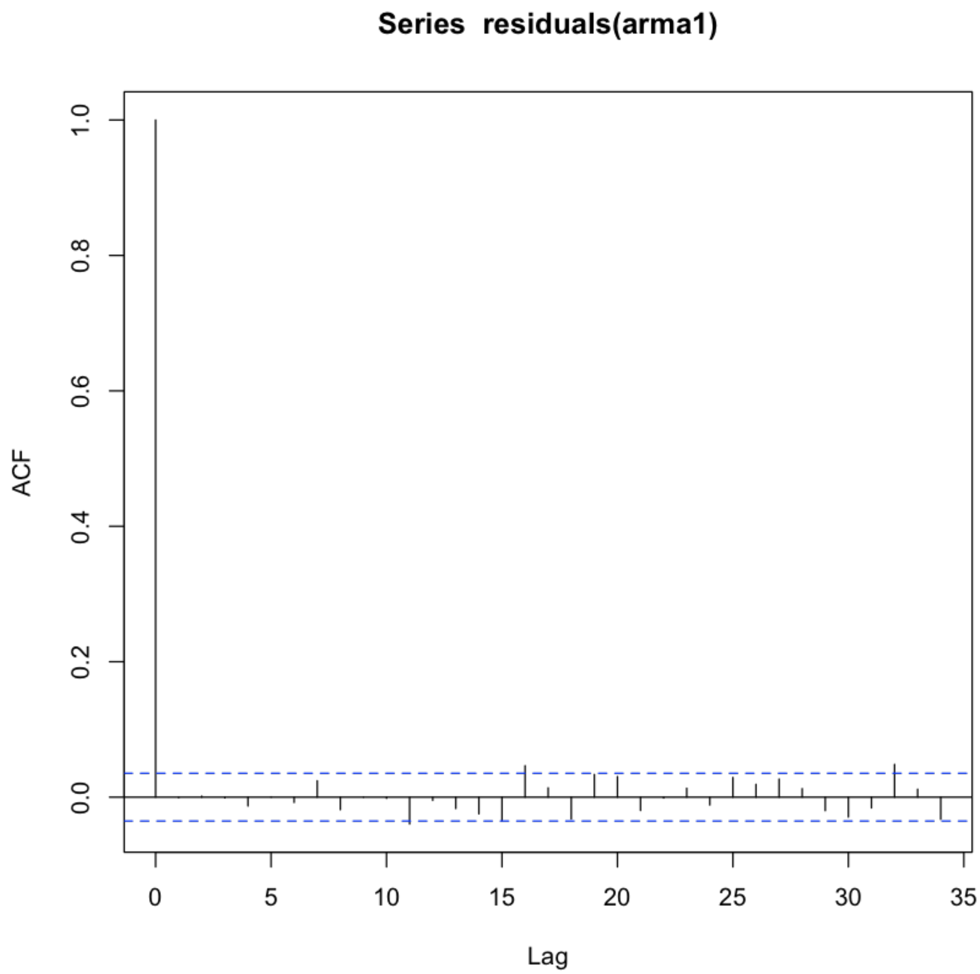
[41]:
```
arma1 = arima(return, order = c(1,0,2))
uncon_mean = arma1$coef[4]/(1-arma1$coef[1])
print(uncon_mean)
```

```
   intercept
0.0008004175
```

### 0.1.4 d)

[42]:
```
acf(residuals(arma1))

Box.test(arma1$residuals, lag = 10, type = 'Ljung-Box')
```

```
Box-Ljung test

data:  arma1$residuals
X-squared = 3.6265, df = 10, p-value = 0.9626
```




The model is adequate. The Ljung-Box test do not reject that the autocorrelation of the residuals up to 10 lags is zero.

### 0.1.5 e)

[43]:
```
pred1 = predict(arma1,
        n.ahead=5,
        se.fit=TRUE)
pred1
plot(pred1$pred)
```

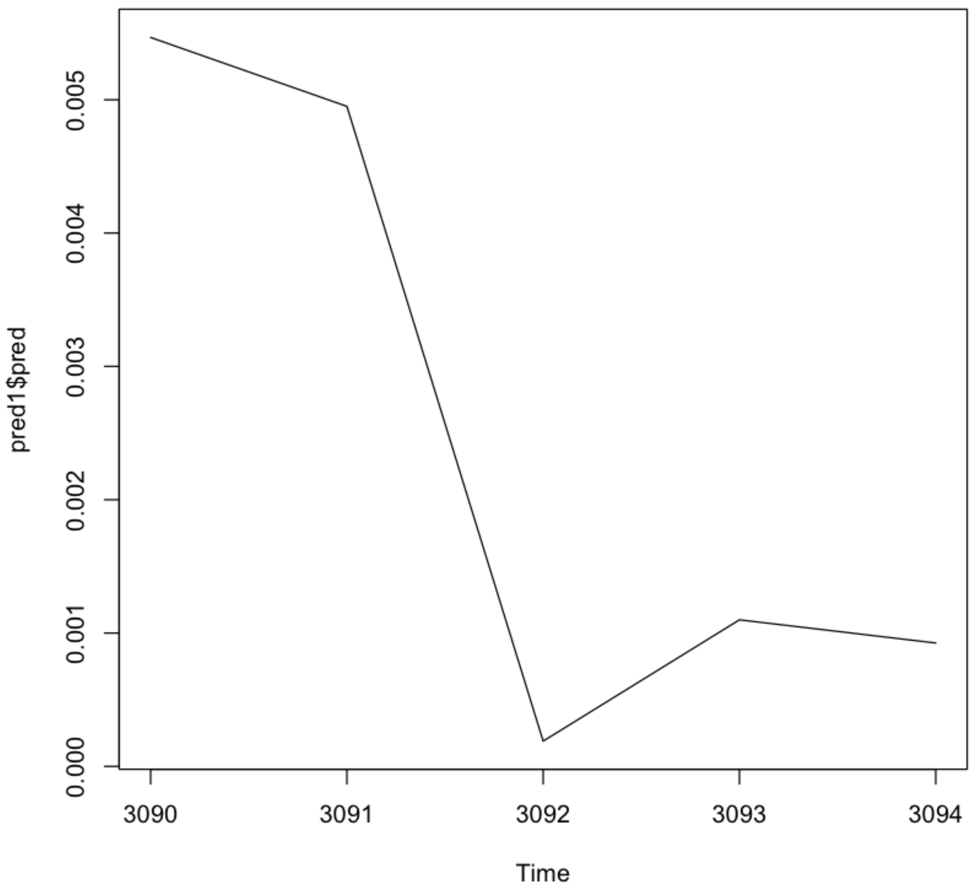
**$pred** A Time Series:



Cont.

1. 0.00546744383322206 2. 0.00495041914103155 3. 0.000189540524276657 4. 0.00109934765719857 5. 0.000925482896983406

**$se** A Time Series:
1. 0.0232756250334324 2. 0.0232799120734797 3. 0.0233167805473954 4. 0.0233181258612074 5. 0.0233181749898848




According to the point estimate which shows positive return you would want to buy the stock now and sell it tomorrow if this return is higher than the transaction costs. However, note that the standard errors are too large which means that you cannot rule out that the return will be negative tomorrow at the 95% confidence level. Hence, you probaly should not reccomend buying the stock.

### 0.1.6 f)

[45]:
```
arma1
```

```
Call:
arima(x = return, order = c(1, 0, 2))

Coefficients:
          ar1     ma1      ma2  intercept
      -0.1911  0.1719  -0.0600      1e-03
s.e.   0.4392  0.4387   0.0193      4e-04

sigma^2 estimated as 0.0005418:  log likelihood = 7232.61,  aic = -14455.22
```

The estimated ARMA(1,2) is then:

$$y_t = 0.001 - 0.1911 y_{t-1} + \varepsilon_t + 0.1719\varepsilon_{t-1} - 0.06\varepsilon_{t-2}$$

so

$$E_t[y_{t+1}] = 0.001 - 0.1911 y_t + 0.1719\varepsilon_t - 0.06\varepsilon_{t-1}$$

we need to know what is the last observation $y_t$ and the last two residuals $\varepsilon_t$ and $\varepsilon_{t-1}$

[48]:
```
tail(return,1)
```

```
             AMZN.Close
2020-03-12 -0.08253502
```

[23]:
```
tail(residuals(arma1), 2)
```

A Time Series:
1. -0.0414716226318796 2. -0.0810212650812879

The AR component implies in an oscillatory behaviour. Since the last observation was negative this contributes for a positive value in our one-step-ahead forecast. The first lag of the MA contributes negatively to the one-step-ahead forecast because the last residual is negative and the second lag contributes positively because the second to last residual is negative and the coefficient on the second lag MA is negative. Going forward in forecasts horizon, the AR component will imply in oscillatory contributions that fade away overtime and the MA components will only affect directly the one-step and second-step forecasts.

## 0.2 Question 4

[25]:
```
library(readxl)
library(vars)
```

### 0.2.1 a)

[26]:
```
data = read_xlsx("/Users/joaobduarte/Dropbox/Lecture_Macroeconometrics/Data/
↪data_midterm_2020.xlsx")
data  = data[, c(4,6,3,5,2)]
data = log(data)*100
data_ts = ts(data)
data_diff = diff(data_ts, lag = 1)
```

### 0.2.2 b)

[28]:
```
var2 = VAR(data_diff, lag.max = 12, ic = "SC")
res = residuals(var2)
adf.test(res[,1])
adf.test(res[,2])
adf.test(res[,3])
adf.test(res[,4])
adf.test(res[,5])
```

```
Warning message in adf.test(res[, 1]):
"p-value smaller than printed p-value"
```

```
Augmented Dickey-Fuller Test

data:  res[, 1]
Dickey-Fuller = -4.5667, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```

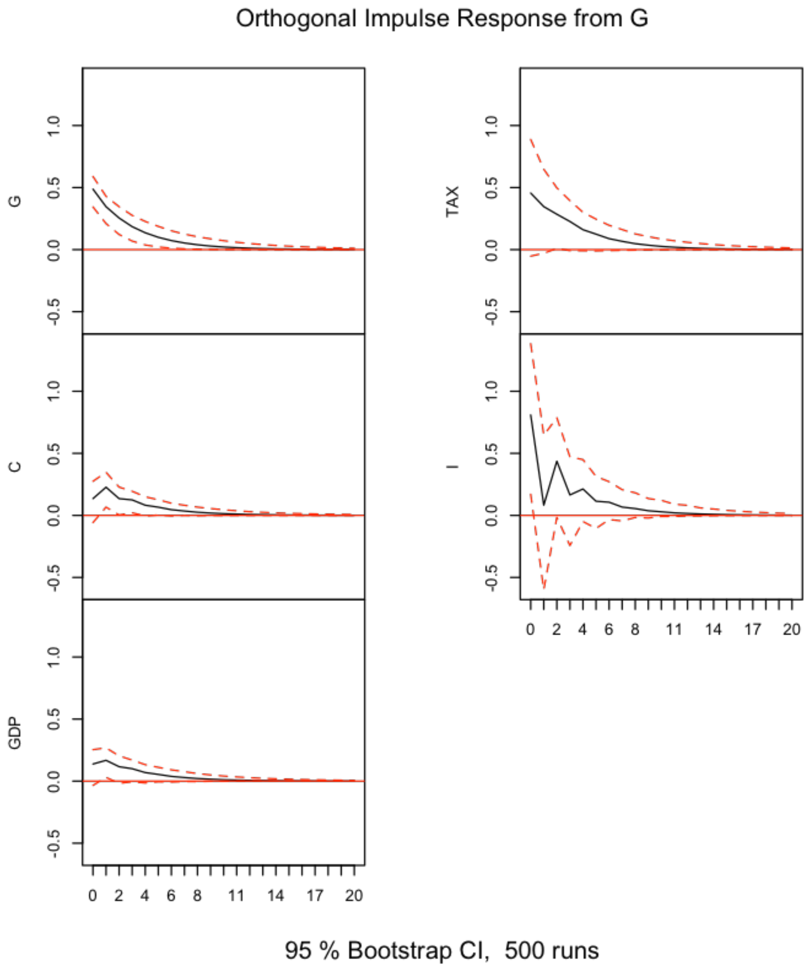
```
Warning message in adf.test(res[, 2]):
"p-value smaller than printed p-value"
```

```
Augmented Dickey-Fuller Test

data:  res[, 2]
Dickey-Fuller = -4.2466, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```

```
Augmented Dickey-Fuller Test

data:  res[, 3]
Dickey-Fuller = -3.523, Lag order = 4, p-value = 0.04377
alternative hypothesis: stationary


Augmented Dickey-Fuller Test

data:  res[, 4]
Dickey-Fuller = -3.7437, Lag order = 4, p-value = 0.02454
alternative hypothesis: stationary


Augmented Dickey-Fuller Test

data:  res[, 5]
Dickey-Fuller = -3.7606, Lag order = 4, p-value = 0.02373
alternative hypothesis: stationary
```

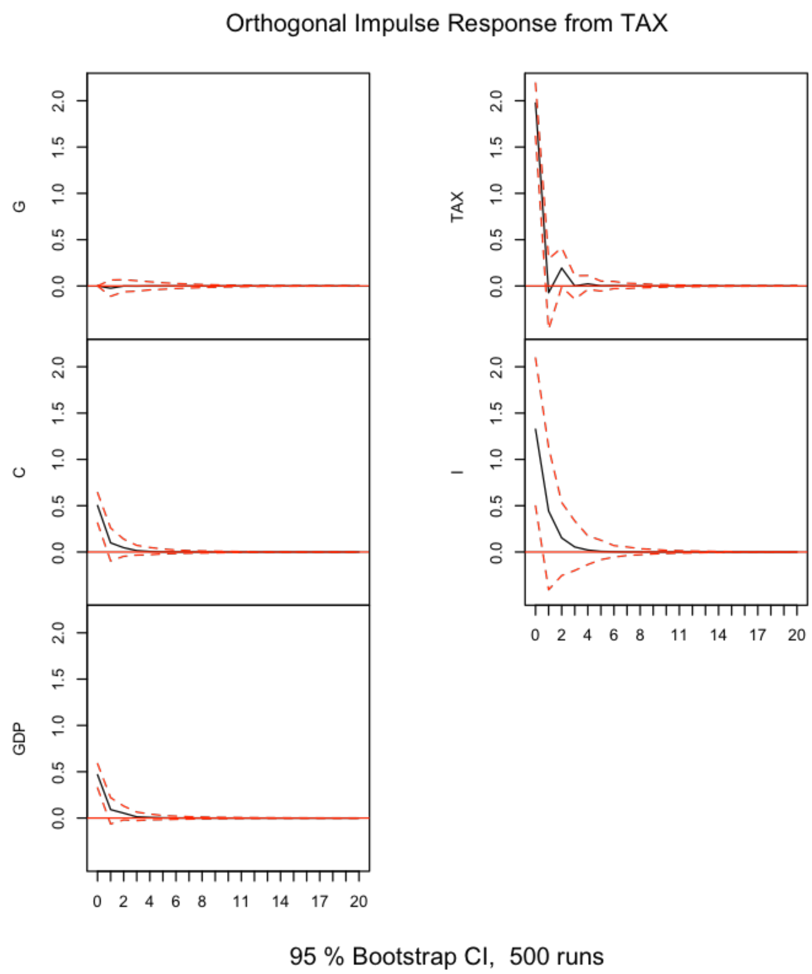
### 0.2.3 c)

[29]:
```
irf1 = irf(var2, n.ahead = 20, ci = 0.95, runs = 500)
plot(irf1)
```


Orthogonal Impulse Response from G
G
TAX
C
I
GDP
1.0
0.5
0.0
-0.5
0 2 4 6 8 11 14 17 20
95 % Bootstrap CI, 500 runs

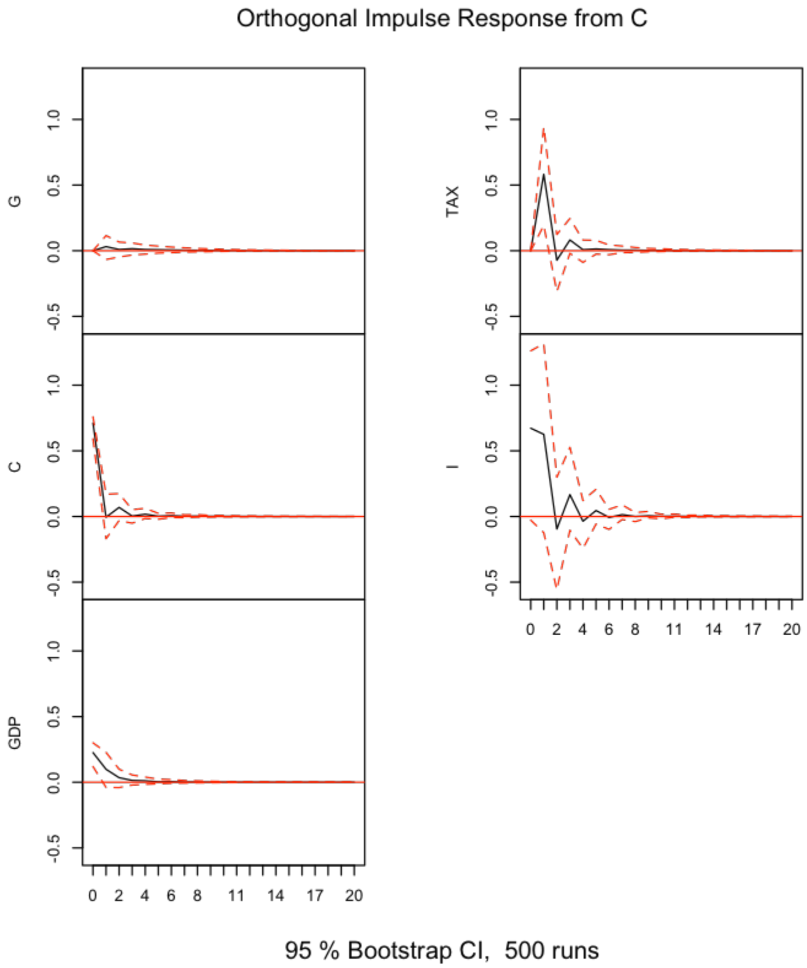


Cont.


Orthogonal Impulse Response from TAX
G
TAX
C
I
GDP
0.0
0.5
1.0
1.5
2.0
0 2 4 6 8 11 14 17 20
95 % Bootstrap CI, 500 runs

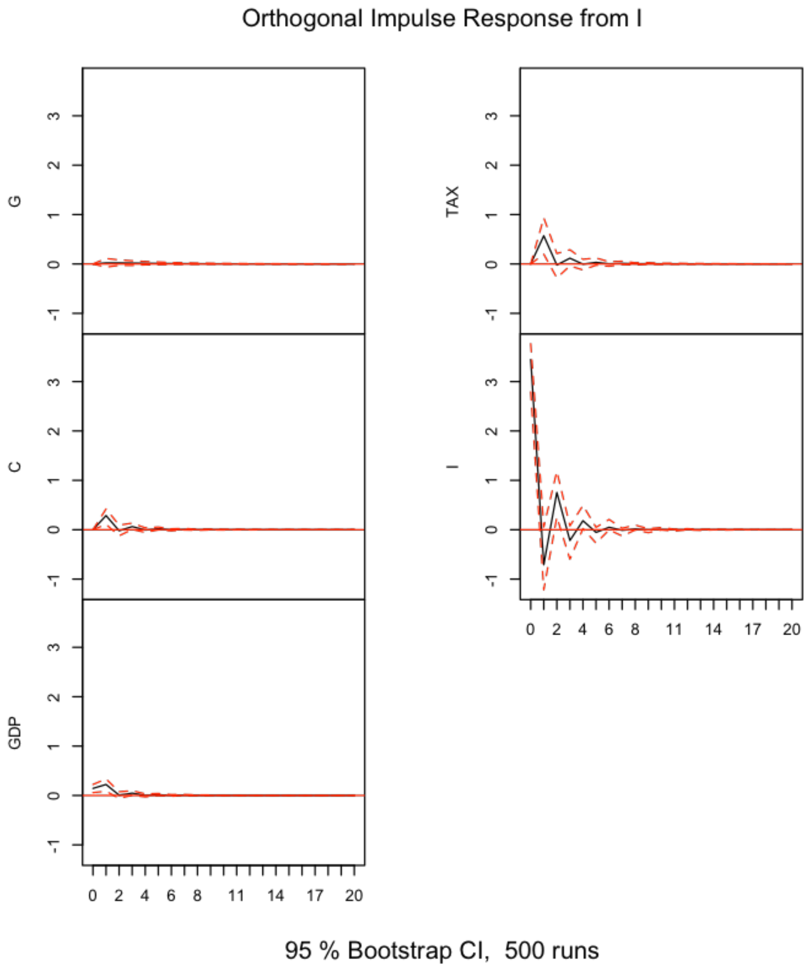


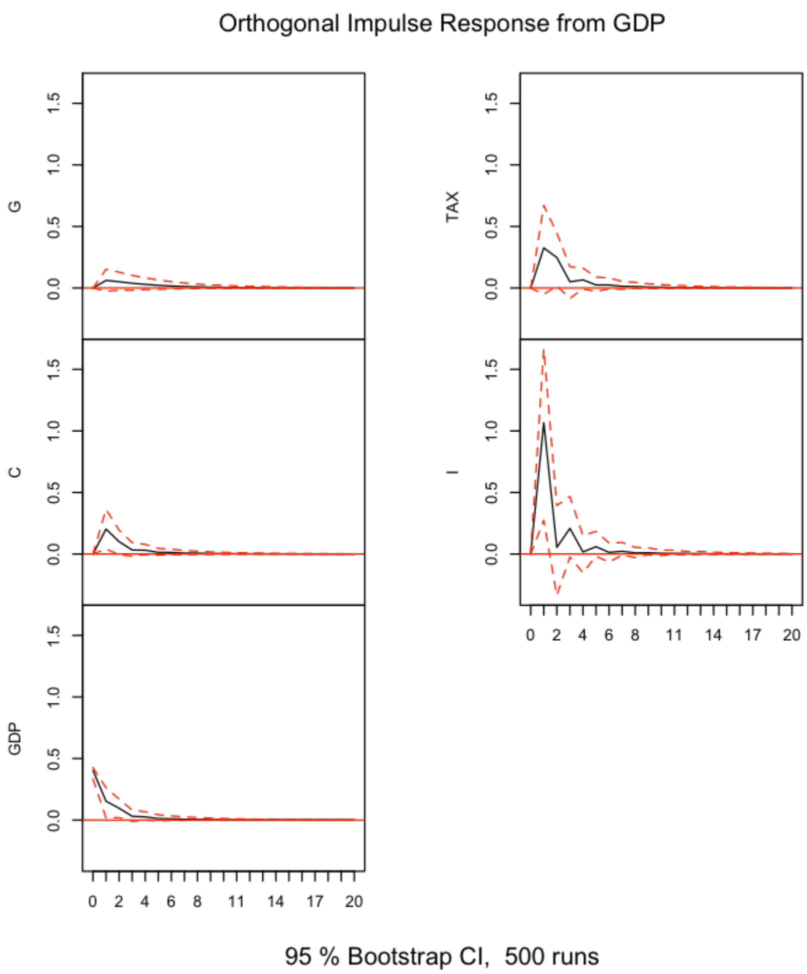
Cont.


Orthogonal Impulse Response from C
G
TAX
C
I
GDP
1.0
0.5
0.0
-0.5
0 2 4 6 8 11 14 17 20
95 % Bootstrap CI, 500 runs

14

Cont.


Orthogonal Impulse Response from I
G
TAX
C
I
GDP
-1
0
1
2
3
0 2 4 6 8 11 14 17 20
95 % Bootstrap CI, 500 runs




An increase in G affects positively GDP, C and I. GDP and C on the second lag (its the only significan coefficent) and I on impact only. You can also conclude that the fiscal multiplier is less than one as an increase in G of 0.5 has an impact on GDP much lower than 0.5.

### 0.2.4 d)

[32]:
```
fe = fevd(var2, n.ahead = 20)
fe$GDP
```

A matrix: 20 × 5 of type dbl

| G | TAX | C | I | GDP |
|---|---|---|---|---|
| 0.04034955 | 0.4640846 | 0.10820140 | 0.04338808 | 0.3439764 |
| 0.07986986 | 0.3843743 | 0.10289238 | 0.11852299 | 0.3143404 |
| 0.09834162 | 0.3721936 | 0.10038479 | 0.11344936 | 0.3156306 |
| 0.11214399 | 0.3645009 | 0.09854789 | 0.11441515 | 0.3103921 |
| 0.11866067 | 0.3613406 | 0.09786928 | 0.11340282 | 0.3087266 |
| 0.12265433 | 0.3595013 | 0.09738901 | 0.11307543 | 0.3073799 |
| 0.12464178 | 0.3586047 | 0.09717502 | 0.11279749 | 0.3067810 |
| 0.12577519 | 0.3581008 | 0.09704340 | 0.11267059 | 0.3064100 |
| 0.12636069 | 0.3578413 | 0.09697886 | 0.11259228 | 0.3062269 |
| 0.12668411 | 0.3576985 | 0.09694186 | 0.11255317 | 0.3061224 |
| 0.12685458 | 0.3576232 | 0.09692285 | 0.11253094 | 0.3060684 |
| 0.12694736 | 0.3575823 | 0.09691231 | 0.11251940 | 0.3060386 |
| 0.12699675 | 0.3575606 | 0.09690678 | 0.11251305 | 0.3060229 |
| 0.12702345 | 0.3575488 | 0.09690375 | 0.11250969 | 0.3060143 |
| 0.12703773 | 0.3575425 | 0.09690215 | 0.11250787 | 0.3060098 |
| 0.12704542 | 0.3575391 | 0.09690128 | 0.11250689 | 0.3060073 |
| 0.12704954 | 0.3575373 | 0.09690082 | 0.11250637 | 0.3060060 |
| 0.12705176 | 0.3575363 | 0.09690057 | 0.11250609 | 0.3060053 |
| 0.12705295 | 0.3575358 | 0.09690043 | 0.11250594 | 0.3060049 |
| 0.12705359 | 0.3575355 | 0.09690036 | 0.11250585 | 0.3060047 |

At impact the G shocks account for 4% of GDP variation but after 5 years (j=20) is accounts for 12.7% of variation.

### 0.2.5 e)

[35]:
```
data3 = ts(data_diff[1:98,])

var3 = VAR(data3, lag.max = 12, ic = "SC")
forecast(var3, h = 1)
```

```
G
   Point Forecast       Lo 80     Hi 80       Lo 95    Hi 95
99      0.1830022 -0.4449319 0.8109363 -0.7773402 1.143345

TAX
   Point Forecast      Lo 80    Hi 80      Lo 95   Hi 95
99      0.5668705 -2.033804 3.167545 -3.410519 4.54426

C
   Point Forecast      Lo 80   Hi 80      Lo 95    Hi 95
99     0.08641907 -1.048422 1.22126 -1.649171 1.822009

I
   Point Forecast      Lo 80    Hi 80      Lo 95    Hi 95
99      0.1489819 -4.734984 5.032948 -7.320399 7.618363
```

```
GDP
   Point Forecast      Lo 80    Hi 80     Lo 95    Hi 95
99      0.2894059 -0.5947295 1.173541 -1.062762 1.641574
```

[34]:
```
arma3 = auto.arima(data3[,5], max.order = 20  ,ic = "bic")
forecast(arma3, h = 1)
```

```
   Point Forecast      Lo 80    Hi 80     Lo 95    Hi 95
99        0.51917 -0.3754505 1.413791 -0.849034 1.887374
```

[55]:
```
tail(data_diff)[6,5]
```

**GDP:** 0.656354911400399

The ARMA model does a much better job at out-of-sample forecast. This is expected given the overfitting problem of VARs.



The End.