Time-Series-Assessment-2.R

Grace Laryea

2022-04-03

```
#Code is in deep blue

#Loading of library packages to enable the functionality of my functions
library(tseries)

## Warning: package 'tseries' was built under R version 4.0.5

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

library(forecast)

## Warning: package 'forecast' was built under R version 4.0.5
```

Q1

```
#.........................................................................#

#(a)
#Setting random number generator at value 639
set.seed(639)
#Running ARIMA Simulation
#n= number of values, 350 values in this case
#model list defines parameters of AR and MA series
#alpha 1=0.4, alpha 2=-0.1 for AR series
#beta1 =0.57
x=arima.sim(n=350,model=list(ar=c(0.4,-0.1),ma=0.57))
```

....

Explanation:

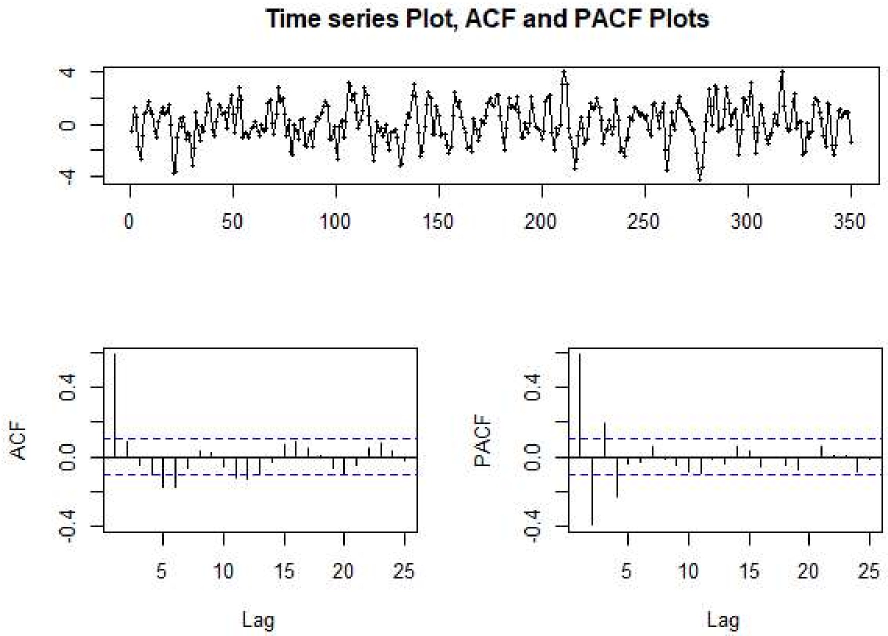
We are simulating a time series of length 350, and ARMA(2,1) with parameters $\alpha_1 = 0.4, \alpha_2 = -0.1 \ and \ \beta_1 = 0.57 \ respectively$

Also, the set.seed, ensure we are able to replicate our time series each time , without randomly generating different values each time.

....

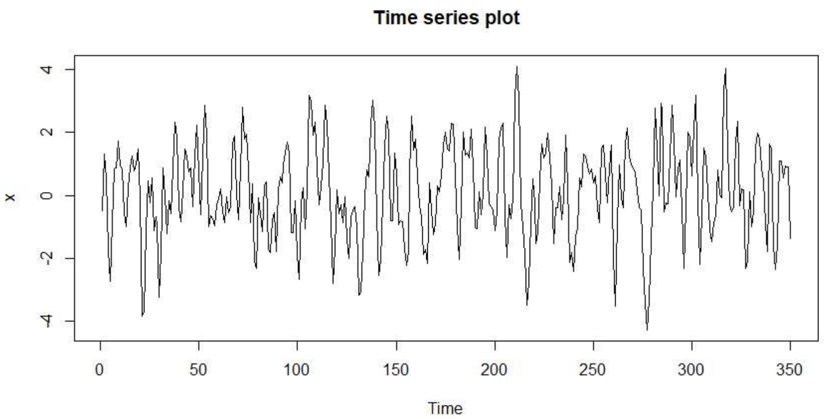
```
Q1(b)
#We use the tsdisplay function to plot the ARIMA time series simulation of x
#This produces a graph of the Time series, its respective ACF and PACF graphs

#code
tsdisplay(x,main="Time series Plot, ACF and PACF Plots ")
```



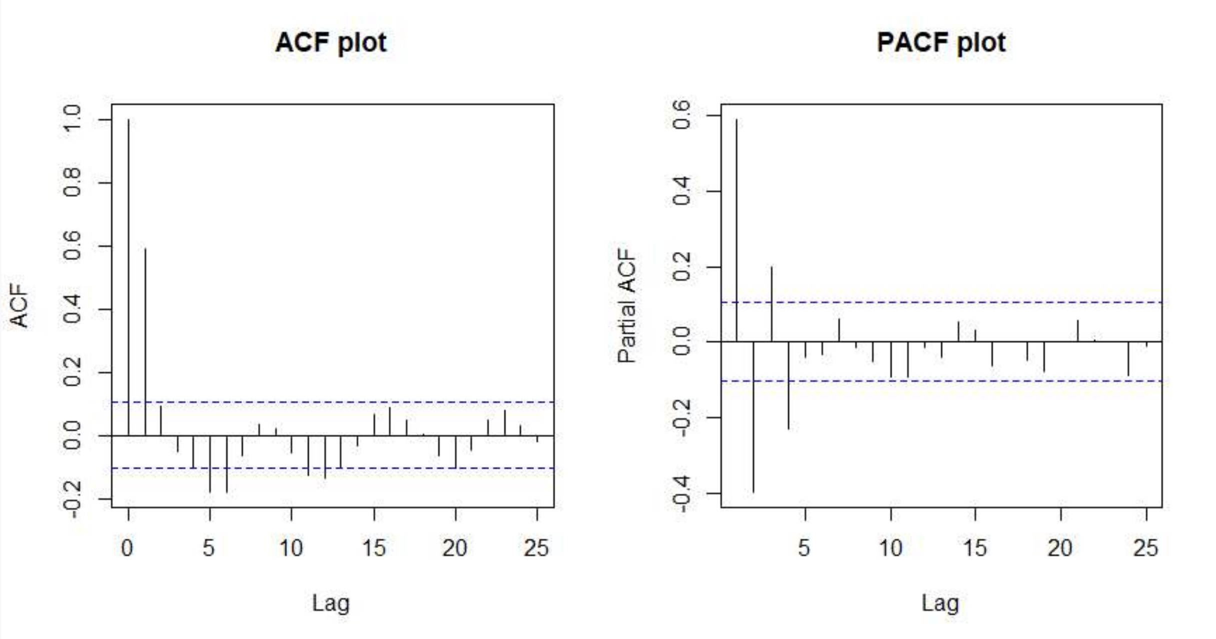

```
Q1b)a #Can also be produced using ts.plot(), acf() and pacf()
ts.plot(x)

Output
```



Q1b)b

```
#code:
par(mfrow=c(1,2))
acf(x)
pacf(x)
```



```
Q1(c)
Our simulation of Time series is stationary, but I am checking for clarity
Running Dicky-Fuller Test to prove stationarity

#code:
adf.test(x)
```

Output

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

##
##  Augmented Dickey-Fuller Test
##
## data:  x
## Dickey-Fuller = -6.695, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary

Q2c)a
#Fitting the ARIMA time series to get ARMA(2,2)
#ARIMA(2,0,2) or ARMA(2,2)
#I have assigned ARMA(2,2) as fit 1
fit1=arima(x,order=c(2,0,2))
#Calling fit1
```

Output

```
fit1

##
## Call:
## arima(x = x, order = c(2, 0, 2))
##
## Coefficients:
##           ar1      ar2     ma1     ma2  intercept
##       -0.0655  -0.0207  1.0440  0.3108     0.0982
## s.e.   0.6073   0.2067  0.6063  0.4021     0.1229
##
## sigma^2 estimated as 1.129:  log likelihood = -518.38,  aic = 1048.77

#(b)
#ARIMA(2,0,1) or ARMA(2,1)
#Assigning ARMA(2,1) as fit2
fit2=arima(x,order = c(2,0,1))
#Calling fit2
```

Output

```
fit2

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

```
## Coefficients:
##          ar1      ar2     ma1  intercept
##       0.3775  -0.1402  0.5969     0.0986
## s.e.  0.0832   0.0726  0.0694     0.1188
##
## sigma^2 estimated as 1.131:  log likelihood = -518.64,  aic = 1047.28
```

Q1(c)

I choose model 2, ARMA(2,1) as it has the lowest AIC, Maximum likelihood value. Hence AR MA(2,1) is the best model.

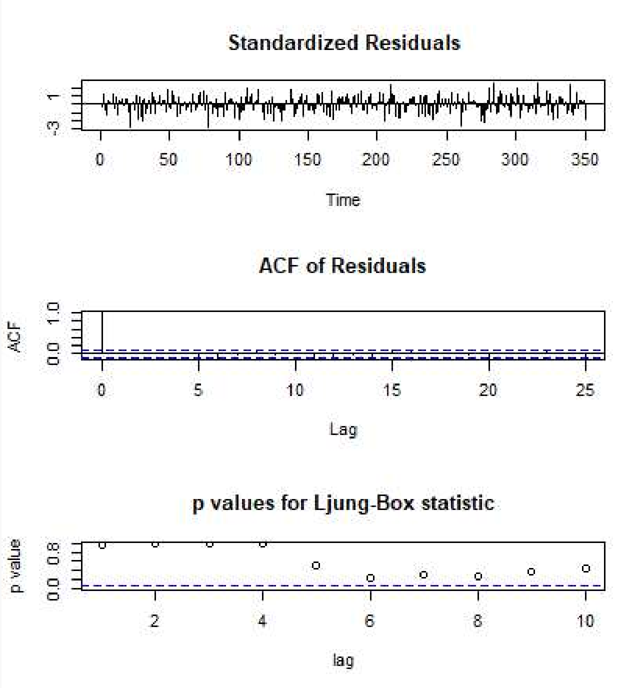
ARMA(2,2)--→ log likelihood = -518.38, aic = 1048.77

ARMA(2,1)--→log likelihood = -518.64, aic = 1047.28

Q1(d)

```
#Running a diagnostic test of fit2
tsdiag(fit2)
```

Output



Equation of model chosen:

ARMA(2,1)

$y_t = 0.4y_1 - 0.1y_2 - 0.57\varepsilon_1$

The output suggests some forecasting can be done, also, the ACF plots suggest white noise a s well as the p-values of Ljung-Box statistic being greater than the significance value of 5%.

Also, since our p-values are greater than 5%, we can suggest, our null hypothesis is not true , hence no evidence of serial correlation amongst the fitted model (fit2).

Looking at the 2nd plot, we can observe the lags of the ACF residuals do not exceed the 95% level, hence model is quite a good fit!

We will look into this further in part(e)

```r
Q1(e)
# I assumed my x value to be my chosen fitted ARIMA series
#fitdf= number of degrees of freedom to be subtracted if x is a series of
#residuals.
#In our case its 4, as our p=2 from AR terms and q=1 from MA terms and -1
#We normally don't tend to use it, but as I am using my fitted time series to
#check for serial correlation
#The question demanded a lag of length 10
#Type of Box.test is Ljung-Box
Box.test(residuals(fit2),fitdf =4,lag = 10,type = "Ljung-Box")
```

Output

```
##
##  Box-Ljung test
##
## data:  residuals(fit2)
## X-squared = 9.8832, df = 6, p-value = 0.1297
```

Explanation:

$$Q = T(T+2)\sum_{k=1}^{s} r_k^2/(T-k) \rightarrow \chi^2_{s-p-q} \qquad (4.11)$$

X-squared represents the Q value for Ljung-Box test(Portmanteau statistic),so our Q value is 9.8832, total degree of freedom= 10-3-1=6,p-value= is probability of our Q value occurring.

So in this context:

Our p-value > 5%, hence adequacy of fitted ARMA(p,q) should be re-analyzed.

$H_0 = Residuals\ of\ white\ noise\ observed$

$H_1 = No\ residuals\ of\ white\ noise\ observed$

T= 350

```r
#using the checkresiduals() to see if my answer is right
checkresiduals(fit2)
```

Output

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,0,1) with non-zero mean
```

```
## Q* = 9.8832, df = 6, p-value = 0.1297
##
## Model df: 4.   Total lags used: 10

#Since p-value is greater than 5%, we need to use the GARCH model

#.......................................................................#
```