ENV 790.30 - Time Series Analysis for Energy Data | Spring 2021

Assignment 6 - Due date 03/16/22

Yared S. Asfaw

## Directions

You should open the .rmd file corresponding to this assignment on RStudio. The file is available on our class repository on Github. And to do so you will need to fork our repository and link it to your RStudio.

Once you have the project open the first thing you will do is change “Student Name” on line 3 with your name. Then you will start working through the assignment by **creating code and output** that answer each question. Be sure to use this assignment document. Your report should contain the answer to each question and any plots/tables you obtained (when applicable).

When you have completed the assignment, **Knit** the text and code into a single PDF file. Rename the pdf file such that it includes your first and last name (e.g., “LuanaLima\_TSA\_A06\_Sp22.Rmd”). Submit this pdf using Sakai.

## Questions

This assignment has general questions about ARIMA Models.

Packages needed for this assignment: “forecast”,“tseries”. Do not forget to load them before running your script, since they are NOT default packages.\

#Loading/installing required package   
library(forecast)  
library(tseries)  
#install.packages("sarima")  
library(sarima)

## Warning: package 'sarima' was built under R version 4.1.3

## Q1

Describe the important characteristics of the sample autocorrelation function (ACF) plot and the partial sample autocorrelation function (PACF) plot for the following models:

## Q2

Recall that the non-seasonal ARIMA is described by three parameters ARIMA where is the order of the autoregressive component, is the number of times the series need to be differenced to obtain stationarity and is the order of the moving average component. If we don’t need to difference the series, we don’t need to specify the “I” part and we can use the short version, i.e., the ARMA. Consider three models: ARMA(1,0), ARMA(0,1) and ARMA(1,1) with parameters and . The refers to the AR coefficient and the refers to the MA coefficient. Use R to generate observations from each of these three models

# ARMA(1,0) $\phi=0.6$ and $\theta= 0.9$, $n=100$  
set.seed(1)  
Model\_10 <- arima.sim(list(order=c(1,0,0), ar=0.6), n=100)  
print(Model\_10)

## Time Series:  
## Start = 1   
## End = 100   
## Frequency = 1   
## [1] -2.07153341 -0.11798913 -0.11572708 -0.08562651 0.89246030 1.35669738  
## [7] 1.40791975 1.76372922 1.84037383 1.17878928 -1.28207813 -0.14942113  
## [13] -0.14578142 -0.24326436 -1.61671100 -1.44817665 -0.45096443 1.08810089  
## [19] 0.55007281 0.71771530 0.37682414 -1.15096507 -1.10557361 -1.05763412  
## [25] -0.69389387 0.68368905 1.17338918 0.53950991 0.07034427 0.73916994  
## [31] 1.00016516 -0.08865660 -0.76068912 -0.09183151 0.71343402 0.31571420  
## [37] 1.07053625 1.04042763 0.01223018 0.34845780 -0.92028842 0.88085065  
## [43] 2.50891029 1.13812470 -0.36125981 0.35296374 0.07672364 2.44765195  
## [49] 1.42935116 1.54735006 0.95641220 -0.16942589 0.08713676 -1.75267657  
## [55] 0.41394892 0.40162269 2.41358528 1.92366070 0.44424999 0.87727635  
## [61] -0.40773182 -1.49827249 -0.60751726 -0.80780223 -0.48357599 -0.21580427  
## [67] -0.71900351 -1.00007084 -0.73522112 0.73695433 -1.08139420 -0.05489034  
## [73] 0.30001617 1.24310954 0.44168180 0.63502789 0.64811552 -0.15365072  
## [79] 1.11567738 1.82980904 1.79809907 2.66569290 2.15790217 0.01814909  
## [85] -0.56237596 -1.56203819 -1.41062355 -1.46674081 -0.83792861 -1.41367882  
## [91] -0.69017852 -1.06869175 1.12607222 1.39235081 1.74558471 1.43153619  
## [97] 2.54109779 0.88892222 0.07170860 1.47530740

# ARMA(0,1) $\phi=0.6$ and $\theta= 0.9$, $n=100$  
set.seed(2)  
Model\_01 <- arima.sim(list(order=c(0,0,1), ma=0.9), n=100)  
print(Model\_01)

## Time Series:  
## Start = 1   
## End = 100   
## Frequency = 1   
## [1] -0.62237391 1.75420960 0.29868512 -1.09758986 0.06019370 0.82713299  
## [7] 0.39746123 1.76874571 1.64723953 0.29274244 1.35763845 0.49088214  
## [13] -1.39309480 0.84652688 -0.70706302 -1.20135760 0.82655084 1.04505474  
## [19] 1.34381098 2.47985784 0.68181147 0.50970496 3.38532602 1.76412425  
## [25] -2.44726239 -1.72929845 -0.16704460 0.25530092 1.00261965 0.99961164  
## [31] 0.98400514 1.36322871 0.68439020 -1.03241722 -1.29466825 -2.26207423  
## [37] -2.45596628 -1.37138795 -0.74966829 -0.60544754 -2.30433078 -2.60489792  
## [43] 1.14601291 2.33568665 2.55116497 1.48634467 -0.36577959 -0.26592126  
## [49] -1.36451307 -1.91717814 1.31184292 1.29742417 0.76969316 0.10057133  
## [55] -2.90869361 -2.09226151 0.64518854 1.98150608 2.69692559 -0.28378532  
## [61] 0.42182453 1.12497393 -0.47466514 0.64858308 -0.36438379 -2.73684259  
## [67] -2.27825489 -0.34718343 -0.81972470 -1.72721362 -0.49869860 0.15574374  
## [73] 0.30735303 0.33764036 -0.95546074 0.48711289 1.94495078 1.74713639  
## [79] -0.46276530 -0.27304992 -0.79937877 -2.05956056 -1.85230438 -3.44296228  
## [85] -0.16500528 0.98651686 -0.87273529 -0.64316270 0.03844033 1.94841633  
## [91] 3.12150672 0.32943307 -2.42370300 -2.73528232 -2.61450861 0.83156267  
## [97] 1.77106724 -0.83573391 -1.35951378 0.53341531

# ARMA(1,1) $\phi=0.6$ and $\theta= 0.9$, $n=100$  
set.seed(3)  
Model\_11 <- arima.sim(list(order=c(1,0,1), ar=0.6, ma=0.9), n=100)  
print(Model\_11)

## Time Series:  
## Start = 1   
## End = 100   
## Frequency = 1   
## [1] -0.78733391 -0.64321564 -1.61583750 -2.47546091 -0.84438145 0.79506500  
## [7] 0.07838572 -1.41590465 -1.90134163 -2.99063518 -3.77866357 -3.44428040  
## [13] -1.57291786 1.11287061 1.50650431 -0.29775034 -0.30112954 1.48165498  
## [19] 2.38330156 2.82142674 2.00357836 1.59074587 2.88976947 2.94243589  
## [25] 0.82060458 0.40476858 1.74375313 1.44364493 2.28565499 2.10579564  
## [31] 0.89678077 -1.41373867 -3.04930945 -0.84470543 0.05540926 -1.27591214  
## [37] -0.84795807 -0.66406073 -0.85984813 -2.01279599 -4.18264007 -4.23341789  
## [43] -3.98373811 -1.96004212 0.96736007 0.62098494 0.23948119 1.57805128  
## [49] 2.02949463 1.80032188 2.57130038 2.11883748 0.41971103 0.82999267  
## [55] 0.06859051 -0.93275421 -0.42340034 0.18486416 1.55550375 2.12243861  
## [61] 1.74591187 2.00378308 3.09217158 3.33402295 1.35738346 -0.23167037  
## [67] 1.45536537 2.08290261 1.62131314 2.81697002 3.58644380 2.86033055  
## [73] 1.92958681 0.46852391 -0.07694274 -0.53677113 -1.97363191 -2.94261009  
## [79] -3.30743869 -1.51630651 -0.35210995 -0.63364259 -0.90576412 -0.41839593  
## [85] 1.16741485 1.32964604 1.22679076 0.88661199 1.22050119 2.28237835  
## [91] 2.92894401 3.28124878 2.72283585 -0.07796339 -2.43731543 -1.59537539  
## [97] -0.98705170 -1.28870065 -1.43867806 0.38323856

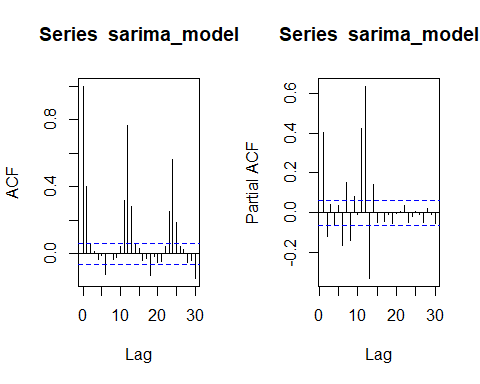
## Q3

Consider the ARIMA model

## Q4

Plot the ACF and PACF of a seasonal ARIMA model with and using R. The after the bracket tells you that , i.e., the seasonal lag is 12, suggesting monthly data whose behavior is repeated every 12 months. You can generate as many observations as you like. Note the Integrated part was omitted. It means the series do not need differencing, therefore . Plot ACF and PACF for the simulated data. Comment if the plots are well representing the model you simulated, i.e., would you be able to identify the order of both non-seasonal and seasonal components from the plots? Explain.

# Plotting ACF and PACF plots of a seasonal ARIMA(0,0,1)(1,0,0)12 with parameters phi = 0.8, theta = 0.5, s=12  
  
sarima\_model <- sim\_sarima(model=list(ma=0.5,sar=0.8, nseasons=12), n=1000)   
  
par(mfrow=c(1,2))  
acf(sarima\_model)  
pacf(sarima\_model)



### To identify the non-seasonal component of the model, in the PACF plot, there is a kind of tail off or a gradual decrease in significance of few spikes and in the ACF plot there is a significant spike at lag 1 that serve as a cut off. These show the process is an MA with an order of q=1.Thus, the non-seasonal part of the model is (0,0,1). For the seasonal component, in the ACF plot, there are multiple spikes at seasonal lag that indicate the process is a seasonal AR, and in the same manner in the PACF plot,there is a single most significant spike at lag 12 that show the process is a seasonal AR. These demonstrate the process for the seasonal component is an AR process with the order of 1 (from the single spike in the PACF plot). Thus, the seasonal part of the model ise (1,0,0). Therefore, we can say that the plots are well representing the model simulated.

.