ENV 797 - Time Series Analysis for Energy and Environment Applications | Spring 2024

Assignment 6 - Due date 02/28/24

Cara Kuuskvere

#Load/install required package here  
library(ggplot2)  
library(forecast)

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

library(tseries)  
#install.packages(sarima)  
#library(sarima)  
library(cowplot)

This assignment has general questions about ARIMA Models.

## Q1

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

* AR

Answer:For AR models the ACF will decay exponentially with time, and the PACF will identify the order of the AR model.

* MA

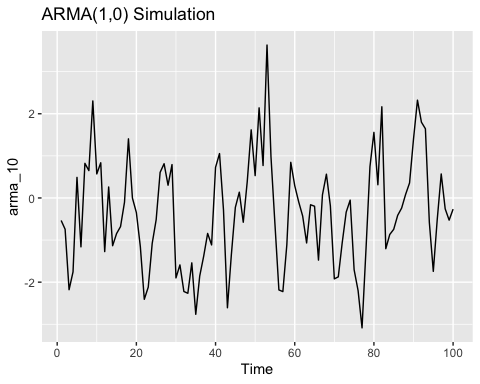
Answer: For MA models the ACF will identify the order of the MA model and the PACF will decay exponentially.

## Q2

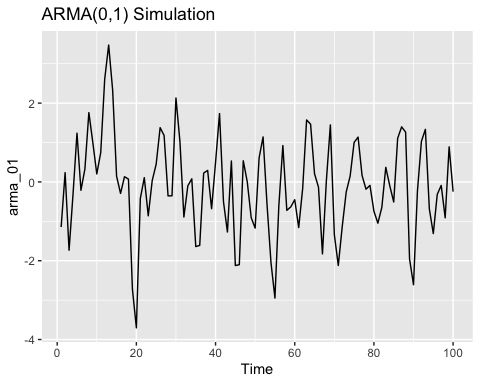
Recall that the non-seasonal ARIMA is described by three parameters ARIMA where p is the order of the autoregressive component, d is the number of times the series need to be differenced to obtain stationarity and q is the order of the moving average component. If we don’t need to difference the series, we don’t need to specify the “I” part and we can use the short version, i.e., the ARMA

1. Consider three models: ARMA(1,0), ARMA(0,1) and ARMA(1,1) with parameters phi=0.6 and theta= 0.9. The phi refers to the AR coefficient and the theta refers to the MA coefficient. Use the arima.sim() function in R to generate n=100 observations from each of these three models. Then, using autoplot() plot the generated series in three separate graphs.

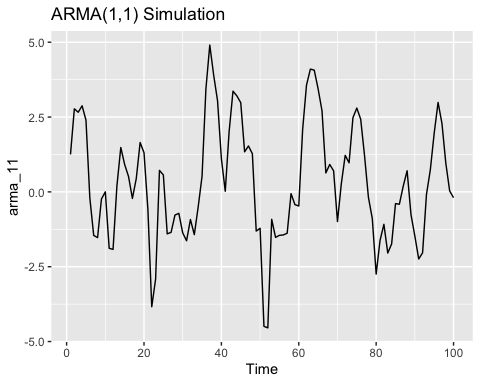
phi <- 0.6  
theta <- 0.9  
n <- 100  
  
# Generate data for ARMA(1,0)  
arma\_10 <- arima.sim(model = list(order = c(1, 0, 0), ar = phi), n = n)  
  
# Generate data for ARMA(0,1)  
arma\_01 <- arima.sim(model = list(order = c(0, 0, 1), ma = theta), n = n)  
  
# Generate data for ARMA(1,1)  
arma\_11 <- arima.sim(model = list(order = c(1, 0, 1), ar = phi, ma = theta), n = n)  
  
# Plot the generated series using autoplot()  
autoplot(arma\_10, main = "ARMA(1,0) Simulation")



autoplot(arma\_01, main = "ARMA(0,1) Simulation")

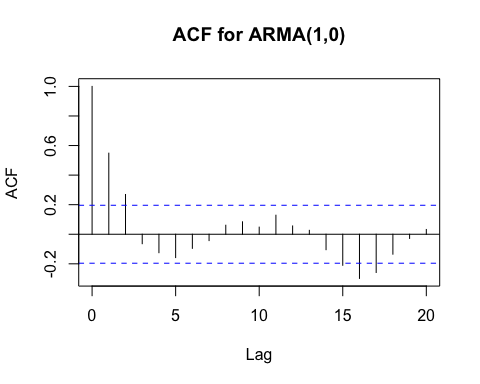


autoplot(arma\_11, main = "ARMA(1,1) Simulation")

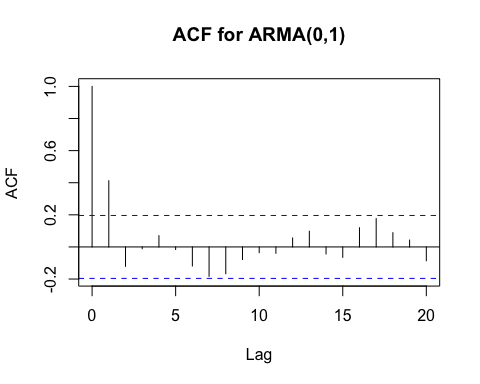


1. Plot the sample ACF for each of these models in one window to facilitate comparison (Hint: use cowplot::plot\_grid()).

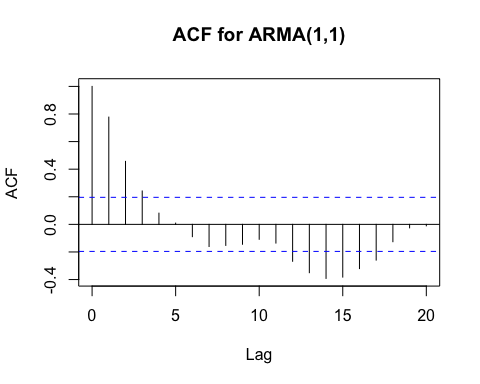
# Calculate and plot ACF for ARMA(1,0)  
acf\_arma\_10 <- acf(arma\_10, main = "ACF for ARMA(1,0)")



# Calculate and plot ACF for ARMA(0,1)  
acf\_arma\_01 <- acf(arma\_01, main = "ACF for ARMA(0,1)")



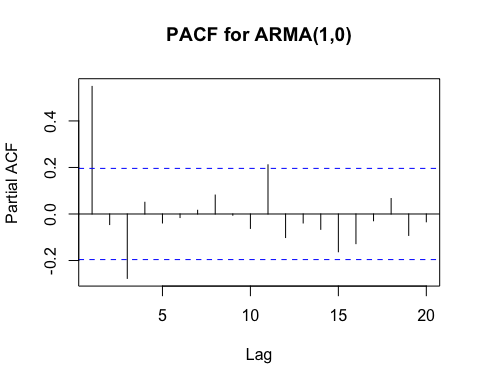
# Calculate and plot ACF for ARMA(1,1)  
acf\_arma\_11 <- acf(arma\_11, main = "ACF for ARMA(1,1)")



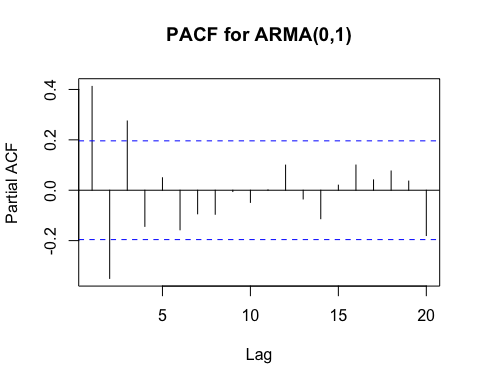
# Create a grid of ACF plots using cowplot::plot\_grid()  
# cowplot::plot\_grid(acf\_arma\_10, acf\_arma\_01, acf\_arma\_11, ncol = 3)  
  
#plot\_grid(plotlist= autoplot(acf(arma\_10, main = "ACF for ARMA(1,0)"), autoplot(acf(arma\_01, main = "ACF for ARMA(0,1)"),autoplot(acf(arma\_11, main = "ACF for ARMA(1,1)")

1. Plot the sample PACF for each of these models in one window to facilitate comparison.

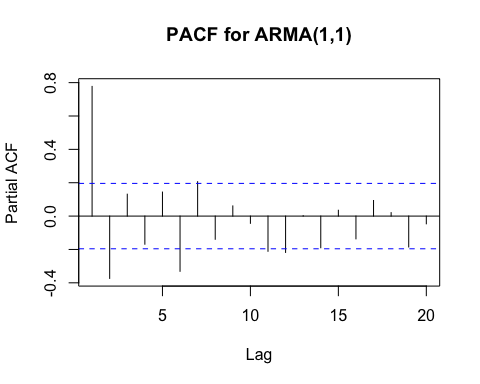
# Calculate and plot PACF for ARMA(1,0)  
pacf\_arma\_10 <- pacf(arma\_10, main = "PACF for ARMA(1,0)")



# Calculate and plot PACF for ARMA(0,1)  
pacf\_arma\_01 <- pacf(arma\_01, main = "PACF for ARMA(0,1)")



# Calculate and plot PACF for ARMA(1,1)  
pacf\_arma\_11 <- pacf(arma\_11, main = "PACF for ARMA(1,1)")



# Create a grid of PACF plots using cowplot::plot\_grid()  
#pacf\_grid <- plot\_grid(pacf\_arma\_10, pacf\_arma\_01, pacf\_arma\_11, ncol = 3)  
  
# Display the grid  
#print(pacf\_grid)

1. Look at the ACFs and PACFs. Imagine you had these plots for a data set and you were asked to identify the model, i.e., is it AR, MA or ARMA and the order of each component. Would you be able identify them correctly? Explain your answer.

Answer: The ARMA(1,0) model appears to have exponential decay in the ACF and a significant spike at lag 1. This would Indicate an autoregressive process with an order of 1. The ARMA 0,1 acf has a significant spike at lag 1 and a pacf with exponential decay. This could be identified as a moving average process with an order of 1. For the ARMA 1,1 model, it has exponential decay in both its acf and pacf, with would make it challenging to identify them correctly because it could indicate a a AR or MA process with orders 1. However, oftentines interpreting these graphs can be a challenge so I may not necessarily be able to identify the model order correctly as if I did not know the information about the models I expect to see.

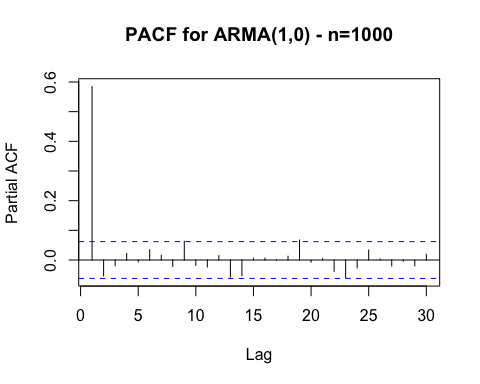
1. Compare the PACF values R computed with the values you provided for the lag 1 correlation coefficient, i.e., does phi=0.6 match what you see on PACF for ARMA(1,0), and ARMA(1,1)? Should they match?

Answer: For ARMA(1,0), representing only an autoregressive (AR) component, the PACF at lag 1 is expected to closely match phi = 0.6, indicating the influence of the AR component.

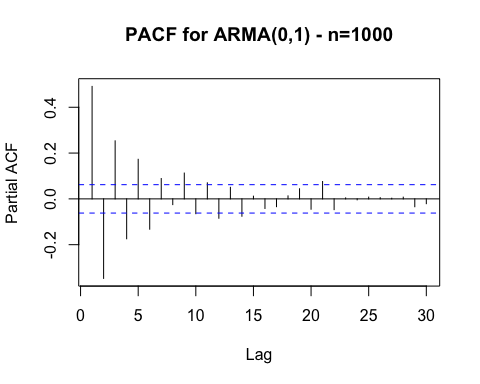
In ARMA(1,1), combining autoregressive (AR) and moving average (MA) components, the PACF at lag 1 should also reflect phi = 0.6. However, at lag 2, it reveals the combined effect of AR and MA coefficients. For ARMA(0,1), with only a moving average (MA) component, the PACF at lag 1 is anticipated to align with the specified MA coefficient (theta = 0.9). Overall, observed PACF values aligning with the specified AR and MA coefficients validate that the simulated data adheres to the anticipated autoregressive and moving average behaviors in the respective ARMA models.

1. Increase number of observations to n=1000 and repeat parts (b)-(e).

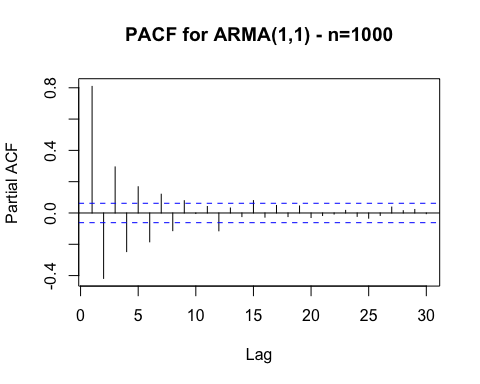
# Increase number of observations  
n <- 1000  
  
# (b) Generate data for ARMA(1,0)  
arma\_10\_large <- arima.sim(model = list(order = c(1, 0, 0), ar = phi), n = n)  
  
# (c) Generate data for ARMA(0,1)  
arma\_01\_large <- arima.sim(model = list(order = c(0, 0, 1), ma = theta), n = n)  
  
# (d) Generate data for ARMA(1,1)  
arma\_11\_large <- arima.sim(model = list(order = c(1, 0, 1), ar = phi, ma = theta), n = n)  
  
# (e) PACF analysis  
pacf\_arma\_10\_large <- pacf(arma\_10\_large, main = "PACF for ARMA(1,0) - n=1000")



pacf\_arma\_01\_large <- pacf(arma\_01\_large, main = "PACF for ARMA(0,1) - n=1000")



pacf\_arma\_11\_large <- pacf(arma\_11\_large, main = "PACF for ARMA(1,1) - n=1000")



# Display the plots  
print(pacf\_arma\_10\_large)

##   
## Partial autocorrelations of series 'arma\_10\_large', by lag  
##   
## 1 2 3 4 5 6 7 8 9 10 11   
## 0.585 -0.054 -0.020 0.022 -0.006 0.035 0.016 -0.022 0.063 -0.018 -0.024   
## 12 13 14 15 16 17 18 19 20 21 22   
## 0.015 -0.058 -0.053 0.006 0.007 0.000 0.013 0.067 -0.007 0.005 -0.038   
## 23 24 25 26 27 28 29 30   
## -0.060 -0.027 0.034 0.003 -0.020 -0.004 -0.020 0.019

print(pacf\_arma\_01\_large)

##   
## Partial autocorrelations of series 'arma\_01\_large', by lag  
##   
## 1 2 3 4 5 6 7 8 9 10 11   
## 0.491 -0.347 0.253 -0.174 0.173 -0.132 0.089 -0.025 0.113 -0.065 0.070   
## 12 13 14 15 16 17 18 19 20 21 22   
## -0.085 0.051 -0.076 0.012 -0.042 -0.034 0.013 0.044 -0.045 0.076 -0.047   
## 23 24 25 26 27 28 29 30   
## 0.004 -0.004 0.008 0.005 0.003 0.008 -0.034 -0.021

print(pacf\_arma\_11\_large)

##   
## Partial autocorrelations of series 'arma\_11\_large', by lag  
##   
## 1 2 3 4 5 6 7 8 9 10 11   
## 0.809 -0.418 0.294 -0.247 0.168 -0.184 0.120 -0.113 0.079 -0.003 0.042   
## 12 13 14 15 16 17 18 19 20 21 22   
## -0.114 0.032 -0.023 0.079 -0.027 0.048 -0.023 0.045 -0.029 -0.016 -0.007   
## 23 24 25 26 27 28 29 30   
## 0.017 -0.022 -0.033 -0.016 0.038 0.015 0.022 -0.005

ARMA(1,0) - Autoregressive (AR) Model

ACF:

The ACF may show exponential decay, with decreasing correlations as the lag increases. A significant spike at lag 1 indicates the presence of the autoregressive term.

PACF:

The PACF at lag 1 should be close to the specified autoregressive coefficient (phi = 0.6) due to the absence of a moving average component.

ARMA(0,1) - Moving Average (MA) Model

ACF:

A significant spike at lag 1 indicates the presence of the moving average term. Exponential decay in the ACF may be observed.

PACF:

The PACF should show a sharp cutoff after lag 1, representing the effect of the moving average term. ARMA(1,1) - Autoregressive Moving Average (ARMA) Model

ACF:

Exponential decay may be observed, indicating the presence of both AR and MA components. A significant spike at lag 1 may indicate the autoregressive term.

PACF:

The PACF at lag 1 reflects the autoregressive coefficient. The PACF at lag 2 may indicate the combined effect of AR and MA coefficients.

In summary, the ACF and PACF plots for each model provide insights into the presence and behavior of autoregressive and moving average components, guiding the identification of the underlying ARMA structure in the time series data.

## Q3

1. Identify the model using the notation ARIMA notation i.e., identify the integers from the equation.

p: AR determined by lag of highest dependent variable = 1 d: differencing order to achieve stationarity– 0 since there is no differencing q: MA determined by highest lag of MA term, = 1

Seasonal: P: seasonal AR = 1, D: Seasonal differencing, D=0 since no seasonal differencing Q: Seasonal MA = 0 since no seasonal moving avg term

s: seasonal period, s=12

ARIMA(1, 0, 1)(1, 0, 0)12

1. Also from the equation what are the values of the parameters, i.e., model coefficients.

The coefficients in the equation represent the values of the parameters:

Autoregressive (AR) coefficients: 0.7

Moving Average (MA) coefficients: −0.1

Seasonal Autoregressive (SAR) coefficients: −0.25

Seasonal Moving Average (SMA) coefficients: none

Constant term: None.

Therefore, the values of the parameters are:

AR(1): 0.7 MA(1): -0.1 SAR(1): -0.25 In summary, the identified ARIMA model is ARIMA(1, 0, 1)(1, 0, 0)12 with the corresponding parameter values.

## Q4

Simulate a seasonal ARIMA(0, 1) model with phi =0 .8 and theta = 0.5 using the sim\_sarima() function from package sarima.

phi <- 0.8  
theta <- 0.5  
  
# Create SARIMA model  
sarima\_model <- list(order = c(0, 1, 0), seasonal = list(order = c(1, 0, 0), period = 12), phi = phi, theta = theta)  
#sarima\_model <- as.ts(sarima\_model)  
  
# Plot the generated series using autoplot()  
#autoplot(sarima\_model)

It would look seasonal if it had regular variations, but I cannot see the plot due to an error. ## Q5

Plot ACF and PACF of the simulated series in Q4. 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.

# Calculate and plot ACF and PACF for the simulated series with greater n  
#acf\_simulated\_large <- acf(sarima\_model, main = "ACF of Simulated Series (n=1000)")  
#pacf\_simulated\_large <- pacf(sarima\_model, main = "PACF of Simulated Series (n=1000)")  
  
# Display the plots  
#print(acf\_simulated\_large)  
#print(pacf\_simulated\_large)

Given the plots that should be outputted if I could get this to run properly, I should be able to understand waht is simulated given the decay of the lags and the impact at the seasonal intervals.