Generating and studying time series  
based on AR & MA processes

Pranav Gopalkrishna, 1940223

2022-02-15

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

QUESTION 1

# AIM

Generate an AR(2) process of size 1000. Choose the parameters from the stationarity region.

# NOTES

## AR(2) process

An AR(2) process is an autoregressive process of order 2, i.e. the value of an observation linearly depends on the previous two observations, i.e. i.e. z-bar\_t depends on z-bar\_(t-1) and z-bar\_(t-2). AR(2) models are given by:

where

* is the error term for time t i.e. difference between the actual observation at time t and the estimated observation at time t based on the other deterministic components
* is the *i*th constant coefficient i.e. the constant coefficient for the observation at lag *i*

## Stationarity region

If all the coefficients of an AR model (in our case, these are  and ) lie between -1 and 1, the AR process will be stationary.

# GENERATING A RANDOM AR(2) PROCESS

To generate an AR process in R, we will use the function to generate autoregressive moving average processes (ARIMA), since ARIMA generalizes autoregressive, integrated and moving average models. With the right parameters, we can create AR(2) models as required.

To generate an AR process in R, we will use the ‘**arima.sim’** function. The arguments are ‘**model**’ (for taking model parameters) and ‘**n**’ (for determining the output time series size). In ‘**model**’, we will give a list object containing the element  
**ar = c(phi1, phi2)**  
where ‘**ar**’ denotes that the parameters belong to an AR process, phi1 and phi2 are the AR(2) model parameters. Since no other parameters will be given, the output will be an AR(2) process specifically.

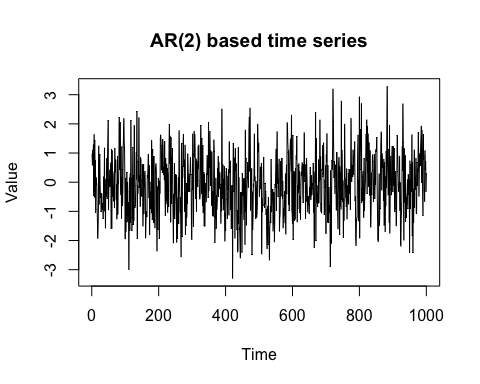
## Time series generation

AR2\_ts = arima.sim(model = list(ar = c(0.1, 0.2)), n = 1000)  
summary(AR2\_ts)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Minimum | 1st Quadrant | Median | Mean | 3rd Quadrant | Maximum |
| -3.29952 | -0.80839 | -0.07965 | -0.07145 | 0.62223 | 3.28477 |

## **Time plot**

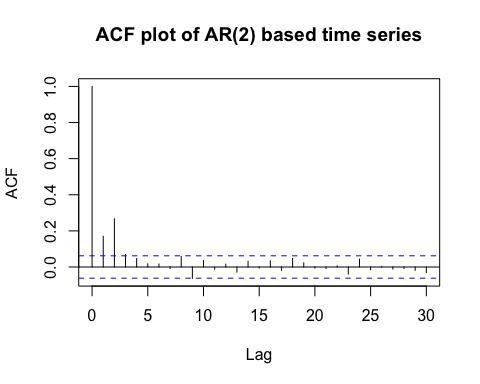
ts.plot(  
 AR2\_ts,  
 main = "AR(2) based time series",  
 xlab = "Time",  
 ylab = "Value")



From the above plot, we observe that the time series seems to be centered around a constant mean of zero. The range of variations in the data also seem to be constant across time. There may be some long-term fluctuations (cyclic fluctuations), but there seems to be no seasonal fluctuations. These factors can lead one to conclude that the time series is stationary.

## ACF plot (ACF => Autocorrelation function)

acf(  
 AR2\_ts,  
 main = "ACF plot of AR(2) based time series")



Here, we can observe that only the autocorrelation between the current observation and the previous two lagged observations is significant. Furthermore, both autocorrelation coefficients are positive, which matches the fact that the weights given the the previous two lagged observations in the defined AR(2) model are both positive.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

QUESTION 2

# AIM

Generate an MA(3) process of size 500. Is the process stationary? Comment on its ACF and PACF plots.

# NOTES

## MA(3) process

A MA(3) process is a moving average process of order 3, i.e. the value of an observation linearly depends on the current as well as the past three error terms (white noise sequence elements). i.e. z-bar\_t depends on at, at-1, at-2 and at-3. MA(3) models are given by:

where

* is the error term for time t i.e. difference between the actual observation at time t and the estimated observation at time t based on the other deterministic components
* is the *i*th constant coefficient i.e. the constant coefficient for the observation at lag *i*

# GENERATING A RANDOM MA(3) PROCESS

The method and function used here is the same that was used to generate the AR(2) process time series in the previous question. The difference is in the arguments, ‘model’, we will give a list object containing the element ma = c(theta\_1, theta\_2, theta\_3), where ‘ma’ denotes that the parameters belong to an MA process, and theta\_2 are the MA(3) parameters. Since no other parameters will be given, the output will be an MA(3) process specifically.

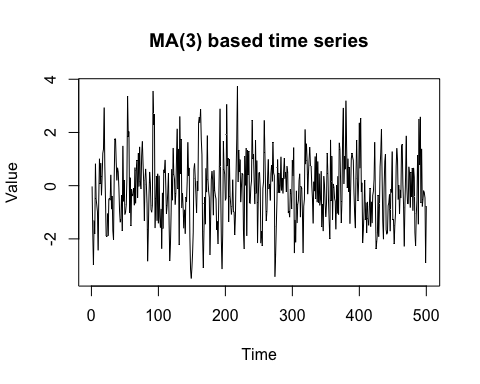
## Time series generation

MA3\_ts = arima.sim(model = list(ma = c(0.5, 0.6, -0.3)), n = 500)  
summary(MA3\_ts)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Minimum | 1st Quadrant | Median | Mean | 3rd Quadrant | Maximum |
| -3.4854 | -0.9822 | -0.1912 | -0.1517 | 0.6477 | 3.7311 |

## **Time plot**

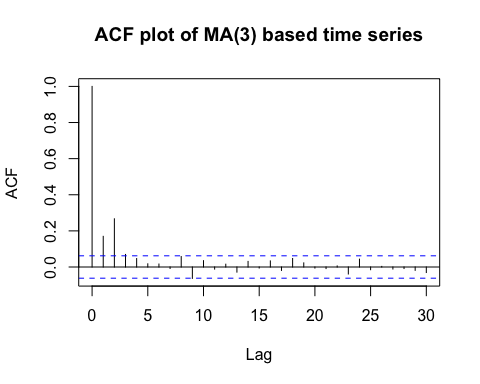
ts.plot(  
 MA3\_ts,  
 main = "MA(3) based time series",  
 xlab = "Time",  
 ylab = "Value")



From the above plot, we observe that the time series seems to be centered around a constant mean of zero. The range of variations in the data also seem to be constant across time. There may be some long-term fluctuations (cyclic fluctuations), but there seems to be no seasonal fluctuations. These factors can lead one to conclude that the time series is stationary.

## ACF plot (ACF => Autocorrelation function)

acf(  
 AR2\_ts,  
 main = "ACF plot of MA(3) based time series")



Here, we can observe that only the autocorrelation between the current observation and the previous two lagged observations is significant. Furthermore, both autocorrelation coefficients are positive, which matches the fact that the weights given the the previous three lagged error terms in the defined MA(3) model are all positive.

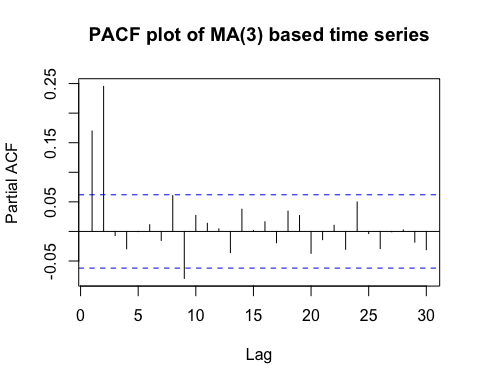
## PACF plot (PACF => Partial autocorrelation function)

### Note on partial autocorrelation

Partial autocorrelation has the same logic as partial correlation, where to compare the correlation between two variables when there are more variables present, we control for the effect of the other variables to give a more accurate association between the two variables of focus. Similarly, partial autocorrelation aims to find the correlation between the current observation and a lagged observation, while controlling for the effect of other lagged observations. This gives a more accurate view about the autocorrelation between the current observation and a past observation at a certain lag.

### Creating the graph

pacf(  
 AR2\_ts,  
 main = "PACF plot of MA(3) based time series")



Here, we can see that the autocorrelation between the previous and current observations is still significant, and even after correcting for this effect, the autocorrelation between the twice lagged and current observations is also still significant. However, we see that the autocorrelation between of the nine-times lagged and current observations is also significant, but negative, which can suggest that the effect of certain lagged observations nullifies or masks to effect of the nine-times lagged observation on the current observation.