STATIONARY TIMESERIES MODEL FOR QUARTERLY AUSTRALIAN GDP DATA

**INTRODUCTION**

***Problem Description:*** Here in this problem we are interested in,

Choosing a non seasonal time series data set and answer the following questions

1. Fit a suitable ARMA model for describing the patterns in the model. Draw your conclusions.
2. Which model is suitable among the class of ARMA models, while examining the ACF and PACF plot of the stationary data?
3. Fit the suitable model based on the answer you obtained from question no.2. Can you say it as the best fitted model for the given data?

***Objective:*** Here our main objective is to take a non seasonal data and fit a suitable ARMA model to describe the pattern in the model. We are also interested in examininig the acf and pacf plot of stationary data and fit a suitable model based on the examination of acf and pacf plot.

The three stationary time series model that we want to study are:

1. Auto regressive process
2. Moving average process
3. Auto regressive moving average process

***Auto regressive process:*** An autoregressive (AR) model predicts future behavior based on past behavior. It’s used for forecasting when there is some correlation between values in a time series and the values that precede and succeed them. You only use past data to model the behavior, hence the name autoregressive (the Greek prefix auto– means “self.” ). The process is basically a linear regression of the data in the current series against one or more past values in the same series.

AR models are also called conditional models, Markov models, or transition models.

An AR(p) model is an autoregressive model where specific lagged values of yt are used as predictor variables. Lags are where results from one time period affect following periods.

The value for “p” is called the order. For example, an AR(1) would be a “first order autoregressive process.” The outcome variable in a first order AR process at some point in time t is related only to time periods that are one period apart (i.e. the value of the variable at t – 1). A second or third order AR process would be related to data two or three periods apart.

***Moving average process:*** In time series analysis, the moving-average model, also known as moving-average process, is a common approach for modeling univariate time series. The moving-average model specifies that the output variable depends linearly on the current and various past values of a stochastic term.

***Auto regressive moving average process:*** Moving average models capture the fact that returns depend not only on current information, but also on signals that have arrived over a previous stretch of time. This could happen if new information is only gradually absorbed or reaches market participants at different points in time. As a consequence, any new signal has not only an immediate, but also a delayed, effect.

Autoregressive models assume that there is a linear relationship between current returns and their own history. This type of model can be used when (some) investors base their decisions on recent price movements: in a bull market, profits attract more buyers who will drive up the price even further; and falling prices are seen as a sell signal that will prolong the downward movement.

These two concepts can be combined; not surprisingly, the resulting model is then called autoregressive moving average model, ARMA(p,q)

*#Loading the package 'expsmooth'.*  
**library**(expsmooth)

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

## Loading required package: forecast

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

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

***Data Description:*** Below is the Quarterly Australian GDP per capita data recorded in the period 1971 - 1998. The data set consist of quarterly record of australian per capita gdp from the year 1971 - 1998.

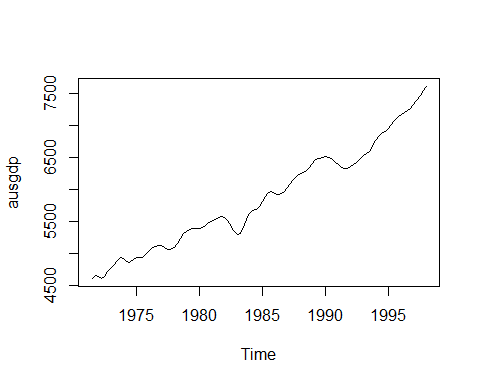
*#Loading the Australian gdp dataset.*  
**data**("ausgdp")  
  
*#Obtaining the dataset.*  
ausgdp

## Qtr1 Qtr2 Qtr3 Qtr4  
## 1971 4612 4651  
## 1972 4645 4615 4645 4722  
## 1973 4780 4830 4887 4933  
## 1974 4921 4875 4867 4905  
## 1975 4938 4934 4942 4979  
## 1976 5028 5079 5112 5127  
## 1977 5130 5101 5072 5069  
## 1978 5100 5166 5244 5312  
## 1979 5349 5370 5388 5396  
## 1980 5388 5403 5442 5482  
## 1981 5506 5531 5560 5583  
## 1982 5568 5524 5452 5358  
## 1983 5303 5320 5408 5531  
## 1984 5624 5669 5697 5736  
## 1985 5811 5894 5952 5965  
## 1986 5943 5924 5935 5979  
## 1987 6035 6097 6167 6227  
## 1988 6256 6272 6295 6345  
## 1989 6413 6468 6497 6511  
## 1990 6514 6512 6490 6442  
## 1991 6390 6346 6328 6340  
## 1992 6362 6389 6433 6491  
## 1993 6541 6566 6602 6671  
## 1994 6765 6847 6890 6918  
## 1995 6962 7018 7083 7134  
## 1996 7173 7212 7242 7276  
## 1997 7332 7400 7478 7550  
## 1998 7618

**ANALYSIS**

*#Obtaining timeseries plot for australian gdp data.*   
**ts.plot**(ausgdp)

**TIME SERIES PLOT OF GDP PER CAPITA [FIGURE 1]**



***Interpretation:*** From the above timeseries plot it is observed that the there exist only a trend component in the dataset, thus we proceed with the analysis.

*#Detrending the dataset to extract the stationary component from the dataset.*  
data=**diff**(ausgdp)

*#loading the package 'tseries'*  
**library**(tseries)

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

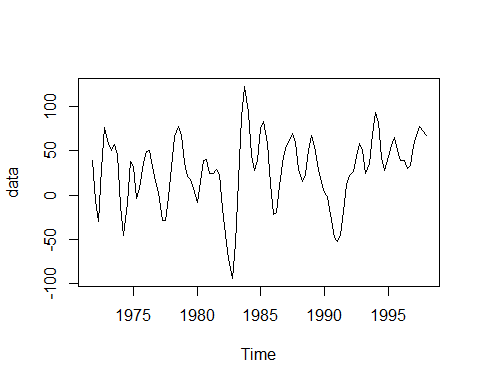
*#Checking for the stationary component of the dataset.*  
**adf.test**(data)

##   
## Augmented Dickey-Fuller Test  
##   
## data: data  
## Dickey-Fuller = -4.0368, Lag order = 4, p-value = 0.01007  
## alternative hypothesis: stationary

***Interpretation:*** We observe that p value = 0.01 which is less than 0.05, thus we conclude that the timeseries dataset has now become stationary thus we proceed to fit a suitable stationary time series model for the dataset.

*#Obtaining the time series plot for the stationary dataset.*  
**ts.plot**(data)

**TIMESERIES PLOT OF STATIONARY DATA [FIGURE 2]**



***Interpretation:*** Thus, now we observe that the timeseries dataset is stationary.

1. **Fit a suitable ARMA model for describing the patterns in the model. Draw your conclusions.**

*#Fitting the suitable ARMA model using stationary data.*  
fit=**auto.arima**(data, seasonal=FALSE)  
**summary**(fit)

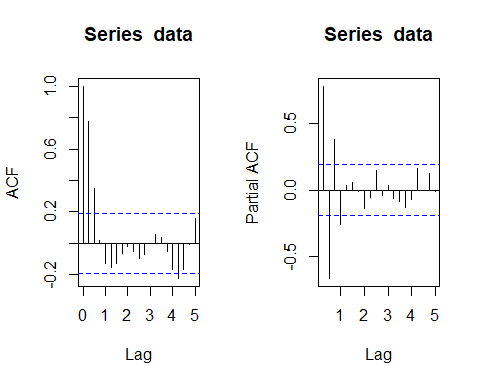
## Series: data   
## ARIMA(2,0,1) with non-zero mean   
##   
## Coefficients:  
## ar1 ar2 ma1 mean  
## 1.0696 -0.4949 0.6919 29.1407  
## s.e. 0.0930 0.0920 0.0726 5.7253  
##   
## sigma^2 estimated as 230.4: log likelihood=-438.34  
## AIC=886.67 AICc=887.27 BIC=899.99  
##   
## Training set error measures:  
## ME RMSE MAE MPE MAPE MASE  
## Training set -0.07136166 14.89125 11.47762 21.95504 58.15256 0.2472477  
## ACF1  
## Training set 0.03689456

***Interpretation:*** Thus, from the above summary it is observed that the model obtained using auto.arima command is ARMA(2,1) with AIC = 886.67.

1. **Which model is suitable among the class of ARMA models, while examining the ACF and PACF plot of the stationary data?**

*#Obtaining the acf and pacf plot for the stationary data.*  
**par**(mfrow=**c**(1,2))  
**acf**(data)  
**pacf**(data)

**ACF AND PACF PLOT FOR STATIONARY DATA OF AUS GDP PER CAPITA [FIGURE 3]**



***Interpretation:*** From, the acf and pacf plot we observe that acf plot is exponentially decreasing and we also observe that in the pacf plot has significant lags till lag 1 thus we can conclude from the acf and pacf plot that AR(1) model is the suitable model to fit.

1. **Fit the suitable model based on the answer you obtained from question no.2. Can you say it as the best fitted model for the given data?**

*#Fitting AR(1) model using 'arima' command.*  
**arima**(data, order=**c**(1,0,0))

##   
## Call:  
## arima(x = data, order = c(1, 0, 0))  
##   
## Coefficients:  
## ar1 intercept  
## 0.7807 29.9405  
## s.e. 0.0595 10.3182  
##   
## sigma^2 estimated as 578: log likelihood = -487.94, aic = 981.87

***Interpretation:*** Thus, above we obtained AR(1) model with aic value = 981.87. No it is not a best fitted model for this data because we have already obtained a ARMA(2,1) model which has less aic value than AR(1) process. Thus, ARMA(2,1) is the best fitted model.

**CONCLUSION:**

From the above analysis we observed that the ARMA(2,1) model is the best model among AR(2,1) and AR(1) model because it has both least variance and least AIC. Thus auto.arima gives the best fitted model.