4.R

harshitaggarwal

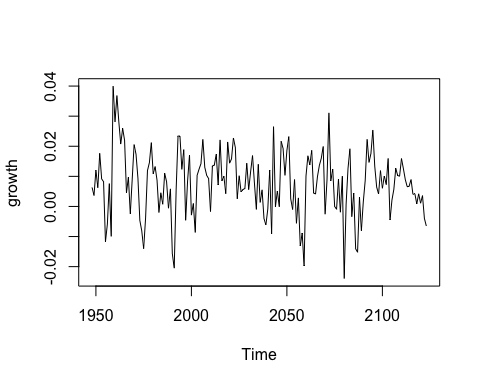
2022-04-21

setwd("~/Documents/GWU/Forecasting/Assignment 4")  
library(forecast)

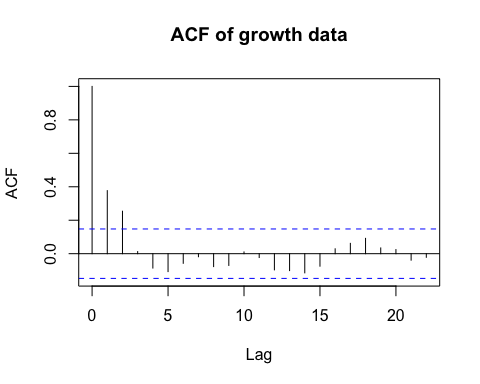
## Warning: package 'forecast' was built under R version 4.1.2

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

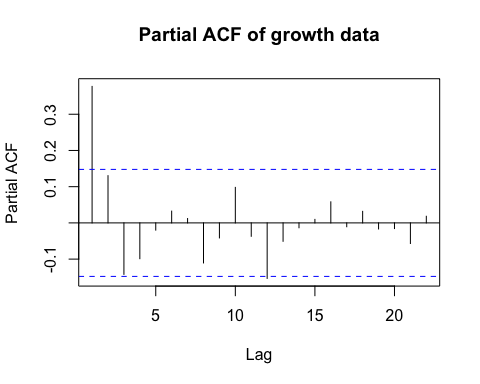
data <- read.table(file="GNP\_Quarterly.txt", header=TRUE)  
  
attach(data)  
  
growth <- ts(data=GROWTH, start=c(2, 1947))  
  
detach(data)  
  
ts.plot(growth)



acf(growth, plot=TRUE, main="ACF of growth data")



# From the time series plot we can see that the mean is constant in the series with no trend being observed  
# From the ACF plot we can see a quick decay in values with only lag 1 and lag 2 being significant  
  
pacf(growth, plot=TRUE, main="Partial ACF of growth data")



# From the Partial ACF plot we can see that only values at lag 1 are outside the 2 standard error bounds  
# Hence we will use a autoregressive function of order 1  
  
fit <- Arima(growth, order=c(1,0,0))  
fit

## Series: growth   
## ARIMA(1,0,0) with non-zero mean   
##   
## Coefficients:  
## ar1 mean  
## 0.3787 0.0077  
## s.e. 0.0698 0.0012  
##   
## sigma^2 = 9.913e-05: log likelihood = 562.47  
## AIC=-1118.94 AICc=-1118.8 BIC=-1109.43

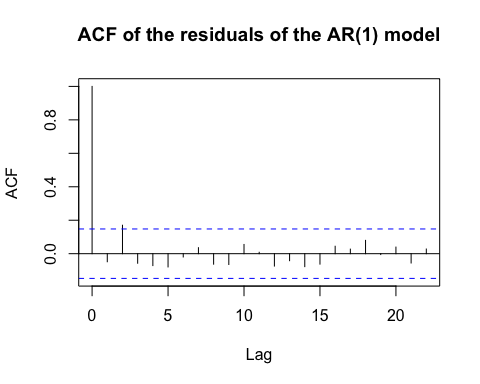
phi=fit$coef  
mu = phi[2]  
C = mu\*(1-phi[1])  
# Estimated Intercept:  
C

## intercept   
## 0.004776108

# Estimated Model:  
# S(t) = 0.004776108 + 0.3786541\*S(t-1)  
  
se\_phi = rep(0,2)  
  
se\_phi[1] = (fit$var.coef[1,1])^0.5  
se\_phi[2] = (fit$var.coef[2,2])^0.5  
  
t\_stat = phi/se\_phi  
t\_stat

## ar1 intercept   
## 5.422598 6.409186

# As we can see the T-stat for both AR1 coefficient and the intercept is over 1.96 so we can reject the null hypothesis that they are 0.  
  
acf(fit$resid, main="ACF of the residuals of the AR(1) model")



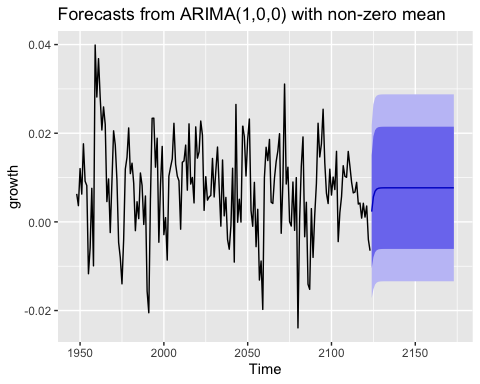
Box.test(fit$resid, lag=20)

##   
## Box-Pierce test  
##   
## data: fit$resid  
## X-squared = 15.216, df = 20, p-value = 0.7639

# As we can see from the residuals plot none of the ACF values are significant except at lag 0  
# We can see the p-value of the Box-Pierce test at 0.7639, having such a high values means we cannot reject the that all acf's are 0  
# Based on the above factors we can conclude that the residuals are white noise  
  
forecast(fit)

## Point Forecast Lo 80 Hi 80 Lo 95 Hi 95  
## 2124 0.002314857 -0.010444987 0.01507470 -0.01719964 0.02182935  
## 2125 0.005652639 -0.007991322 0.01929660 -0.01521400 0.02651928  
## 2126 0.006916503 -0.006849568 0.02068257 -0.01413689 0.02796989  
## 2127 0.007395070 -0.006388420 0.02117856 -0.01368496 0.02847510  
## 2128 0.007576282 -0.006209704 0.02136227 -0.01350756 0.02866013  
## 2129 0.007644898 -0.006141445 0.02143124 -0.01343949 0.02872929  
## 2130 0.007670880 -0.006115515 0.02145728 -0.01341359 0.02875535  
## 2131 0.007680718 -0.006105684 0.02146712 -0.01340376 0.02876520  
## 2132 0.007684444 -0.006101960 0.02147085 -0.01340004 0.02876893  
## 2133 0.007685854 -0.006100549 0.02147226 -0.01339863 0.02877034

autoplot(forecast(fit,h=50))



# As we can see from the forecast plot as k gets large the k-step ahead forecast will approach its mean.  
# This is due to the fact that lags are not correlated and as k gets large we will use the mean as the prediction