PSTAT174_Lab06

Celeste Herrera

5/7/2021

1. (a) Summarize how you would carry out the following steps in time series analysis. You can describe briefly by words, or write down R commands you would use to implement these steps.

- **Step 1 Data processing (make the data stationary)**

So in order to make the data stationary the first thing I would do is find the differencing of the lags. So for this I would use " diff(dataset, lag = d) "

- **Step 2 Model identification**

For this following step (2) of Model identification I would be plotting ACF/PACF because it would help give the knowledge of what model is suppose to be used. So for the functions I would use acf() and pacf().

- **Step 3 Model estimation**

So after step (2) it will now be Model estimation and for this instance I would use and AR model to fit the sample PACF while also doing an instance of the estimation of coeffcents along side using Yule- Walker estimation So for the functions of this instance I would use ar(dataset.diff, method="yw")

- **Step 4 Model selection**

After completing the "Model estimation (Step 3)" we will now begin to fit the different models of ARMA by using the maximum likelihood estimation from step 3. After so we will be using the AICC for the arima() function then to find the model with the smallest AIC we will use AICC() function.

- **Step 5 Model diagnostics**

In step (5) we will be doing Model diagnostics to analyze and perform test as well as plots to acknowledge the test of the model. Stated below you will see different forms of code to analyze for different types of tests.

=> residual independence: Box.test(residuals(), type="Ljung")

=> residual analysis: arima(data, order=c(), method="ML")

=> normality: shapiro.test(residuals())

=> ts.plot(residuals()), acf(residuals())

=> more plots: pacf(residuals()), hist(residuals()), qqnorm(residuals()), qqline(residuals())

- **Step 6 Data forecast**

After Model Diagnostics (step(5)) we will now be in the Data forecast where in this instance we will use about 10 observations to use the model. The types of code we will be using regards the following below.

=> points()

=> predict()

=> ts.plot(data, xlim=c())

=> lines(,lty=)

(b) Review this week's lab material, Dow Jones Index question part 3) ('Make the data stationary'). Is differencing once at lag 1 sufficient to make the data stationary? If yes, justify it. If no, try to make it stationary. Please write related R codes.
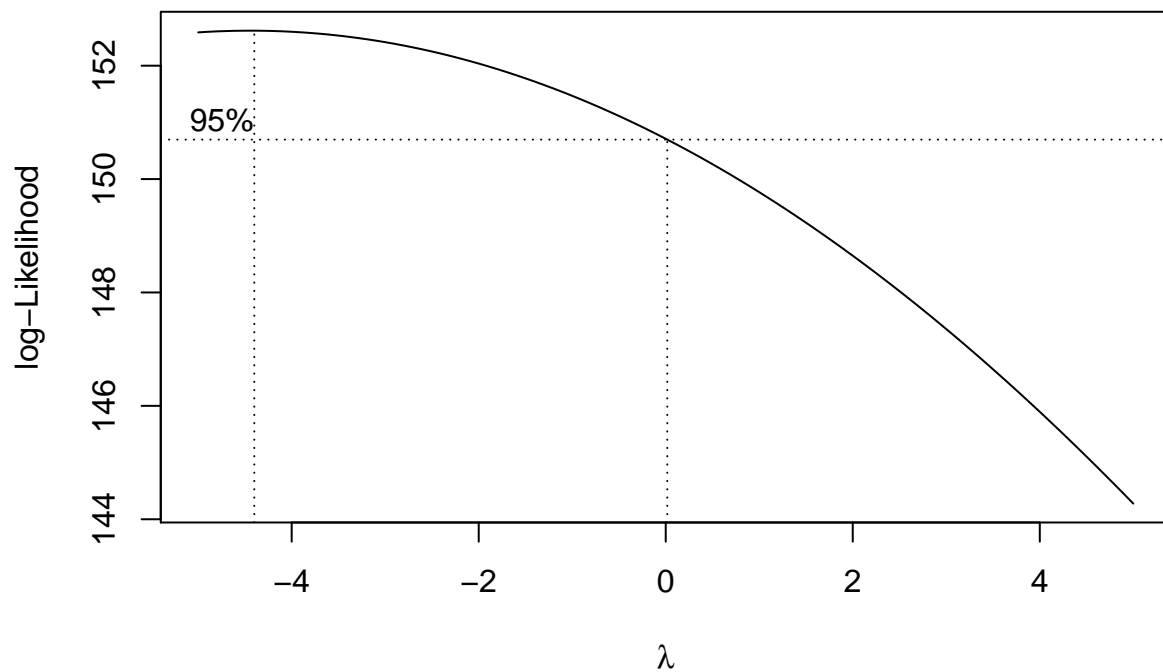
```
library(MASS)
dowj_data <-scan("dowj.txt")
dowj <- ts(dowj_data)

# Shown below will be the use of the box-cox transformation.

par(mfrow=c(1,1))
ts.plot(dowj, main = "Dow Jones Index")
```
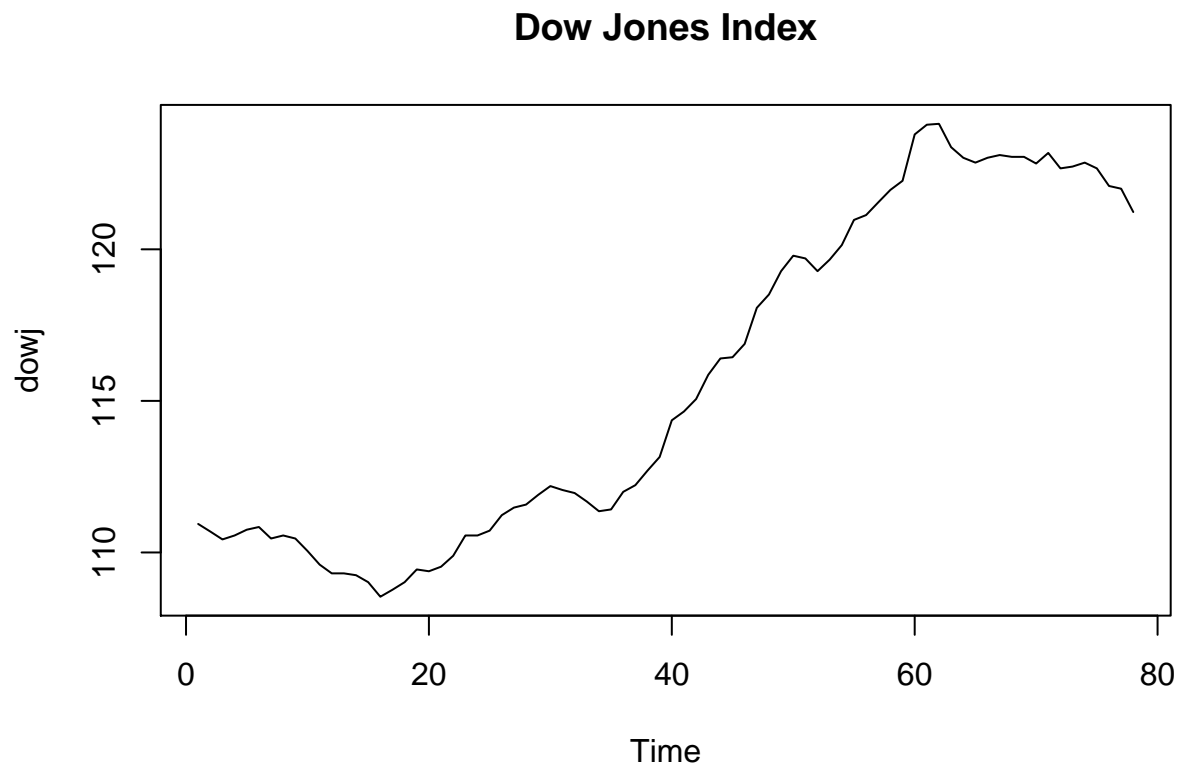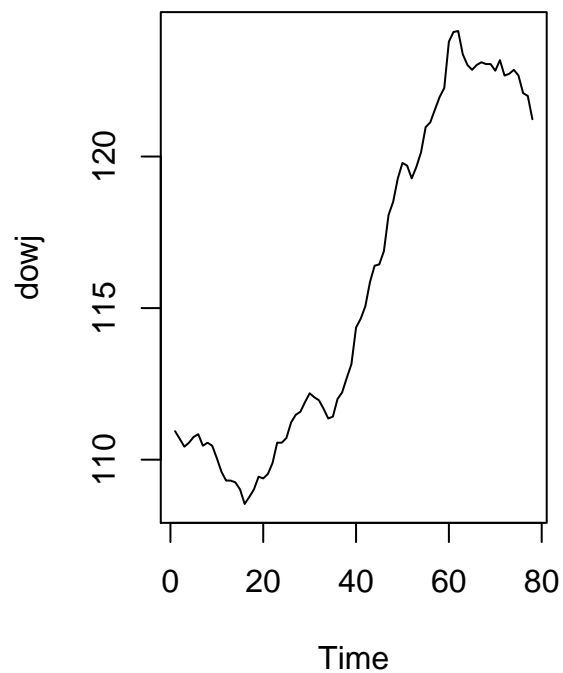

Dow Jones Index

```
t <- 1:length(dowj)
box_trans <- boxcox(dowj~t, lambda = seq(-5,5,0.1), plotit = TRUE)
```
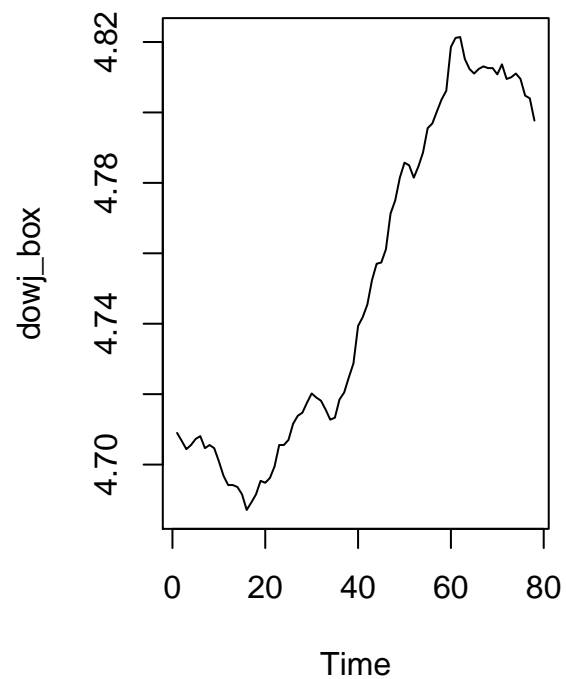
```
dowj_box <- log(dowj)

par(mfrow=c(1,2))
ts.plot(dowj, main = " Dow Jones Index")
ts.plot(dowj_box, main = " Dow Jones Box Cox Transformed")
```
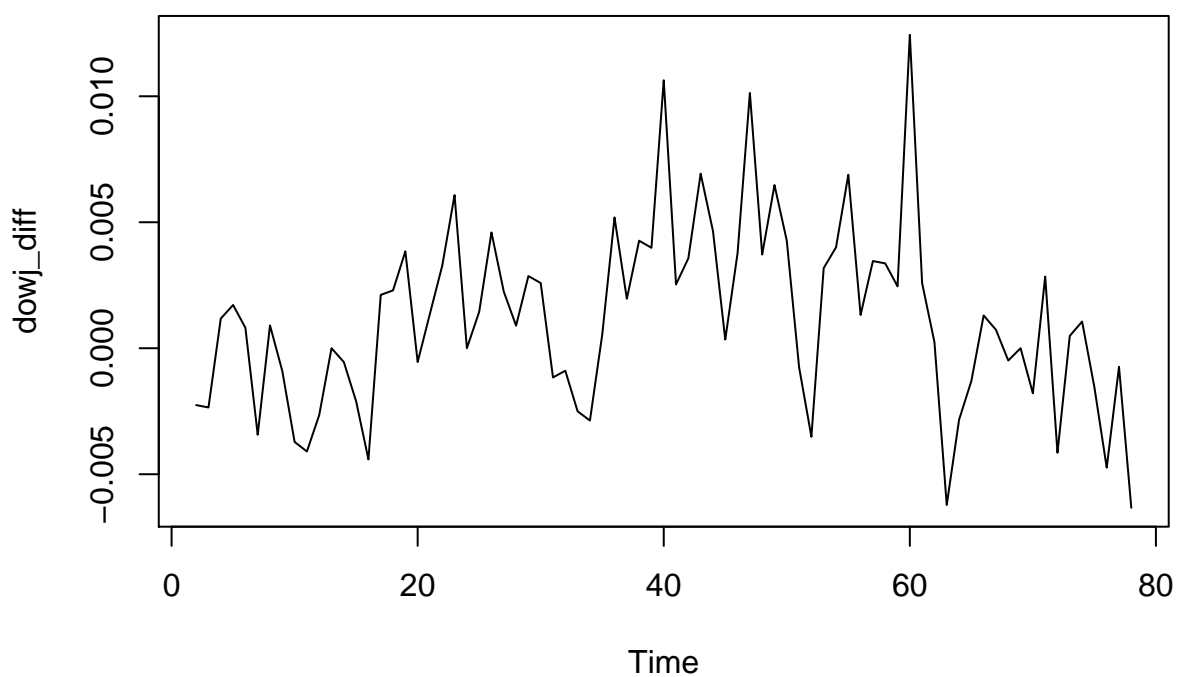
Dow Jones Index   **Dow Jones Box Cox Transforme**


```
#Differencing the data
dowj_diff <- diff(dowj_box, 1)
par(mfrow=c(1,1))
```
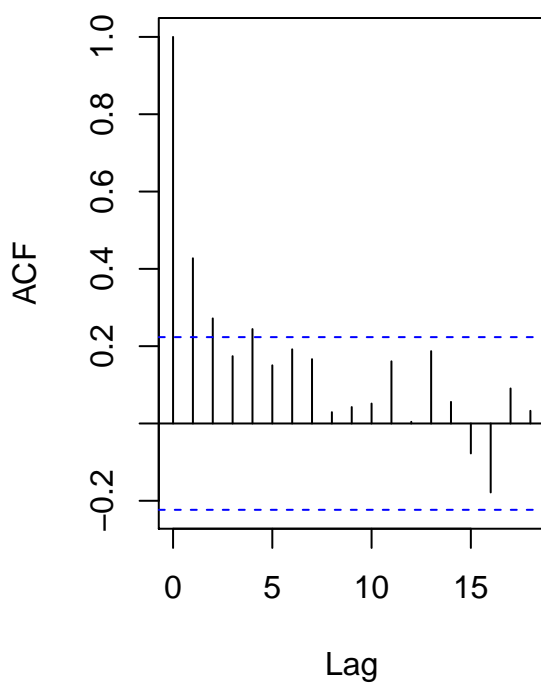
```
ts.plot(dowj_diff, main = "Detrended data")
```
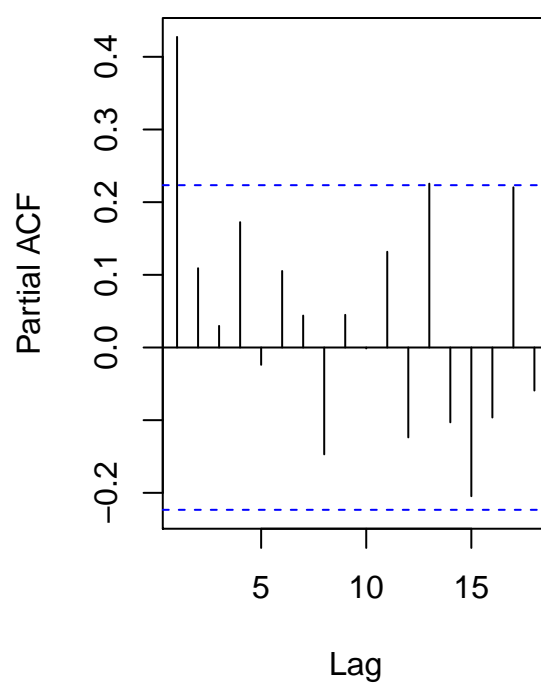
Detrended data


```
par(mfrow=c(1,2))
acf(dowj_diff); pacf(dowj_diff)
```

Series dowj_diff

Series dowj_diff

```
var(dowj_box); var(dowj_diff)
```

```
## [1] 0.002243933
```
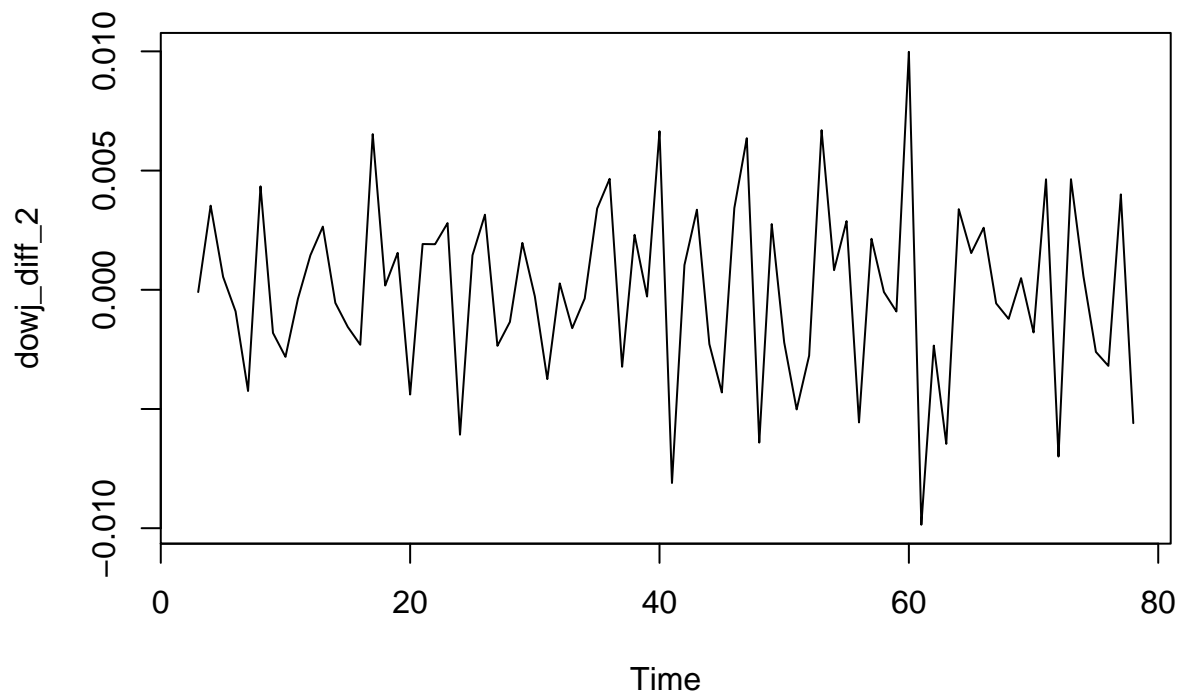
```
## [1] 1.325112e-05
```

```
# differencing again
dowj_diff_2 <- diff(dowj_diff, 1)
par(mfrow = c(1,1))
ts.plot(dowj_diff_2, main = "Detrended data")
```
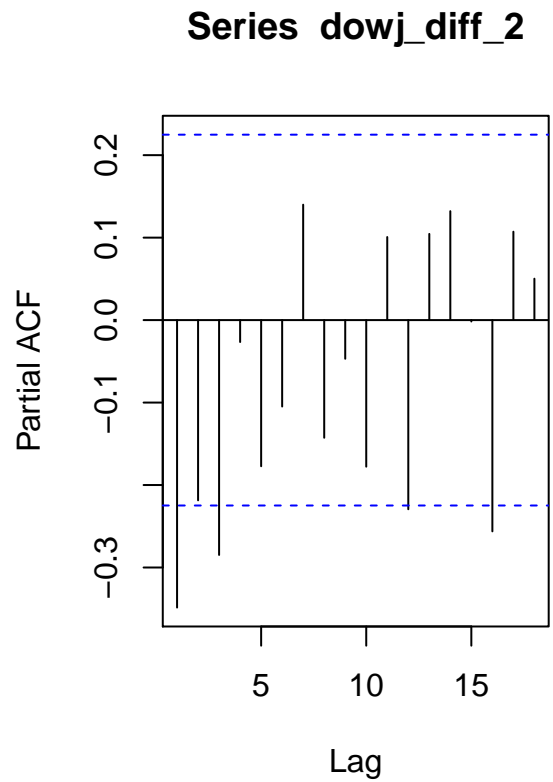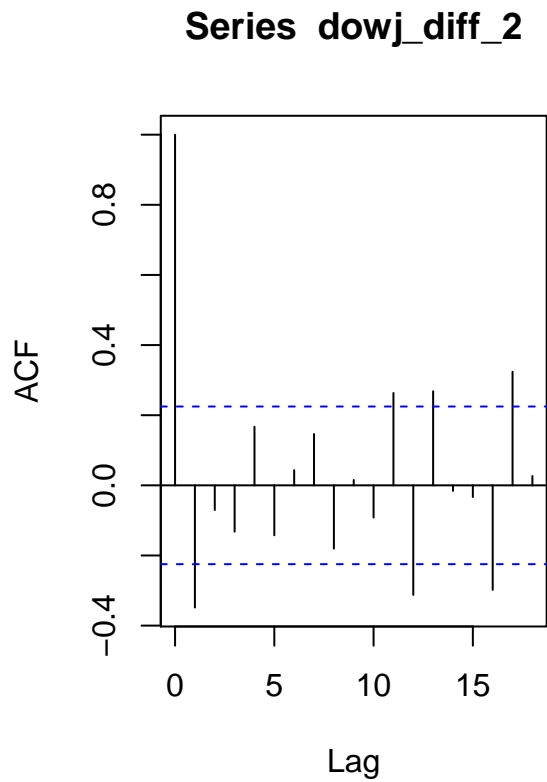
Detrended data


```
par(mfrow = c(1,2))
acf(dowj_diff_2); pacf(dowj_diff_2)
```

Series dowj_diff_2

Series dowj_diff_2

```r
var(dowj_diff); var(dowj_diff_2)
```

```
## [1] 1.325112e-05
```

```
## [1] 1.447361e-05
```

After differencing again during the second time as shown above the data looks much more stationary.