Arima

Akhil Isaac John

11/10/2022

ARMIA MODELLING

Auto Regressive Integrated Moving Average. In the following paper we try to run this model on the dataset under consideration using the Box-Jenkins methodology which is in these following steps 1. Model selection 2. Parameter Estimation 3. Testing for Significance 4. Prediction We use the stock data for Trent Ltd from 1st September 2021 - 1st September 2022

Load data set

```
df = read.csv("C:\\Users\\John\\Desktop\\Sem 3 notes\\Econ\\TRENT.NS.csv")
head(df)
```

```
##         Date    Open    High    Low   Close Adj.Close  Volume
## 1 2021-09-01 1004.45 1036.70 988.20  994.40  992.7692  861143
## 2 2021-09-02  999.40 1013.65 985.55 1000.50  998.8592  600794
## 3 2021-09-03 1008.90 1017.65 984.15  995.70  994.0671  454344
## 4 2021-09-06 1000.10 1023.90 995.70 1009.90 1008.2438  646921
## 5 2021-09-07 1009.80 1021.95 993.25  999.00  997.3616  597743
## 6 2021-09-08 1001.00 1050.00 995.00 1033.85 1032.1544 2226426
```

```
summary(df)
```

```
##      Date                Open           High            Low
##  Length:248         Min.   : 971   Min.   : 991.1   Min.   : 953
##  Class :character   1st Qu.:1044   1st Qu.:1065.0   1st Qu.:1025
##  Mode  :character   Median :1094   Median :1116.8   Median :1073
##                     Mean   :1129   Mean   :1150.0   Mean   :1108
##                     3rd Qu.:1204   3rd Qu.:1219.4   3rd Qu.:1185
##                     Max.   :1481   Max.   :1522.5   Max.   :1448
##      Close          Adj.Close          Volume
##  Min.   : 979.1   Min.   : 977.5   Min.   : 141116
##  1st Qu.:1041.6   1st Qu.:1039.9   1st Qu.: 462463
##  Median :1095.2   Median :1094.1   Median : 620163
##  Mean   :1129.0   Mean   :1127.9   Mean   : 772909
##  3rd Qu.:1207.7   3rd Qu.:1207.7   3rd Qu.: 877675
##  Max.   :1474.7   Max.   :1474.7   Max.   :5869336
```

```
attach(df)
```

Check ARIMA parameters

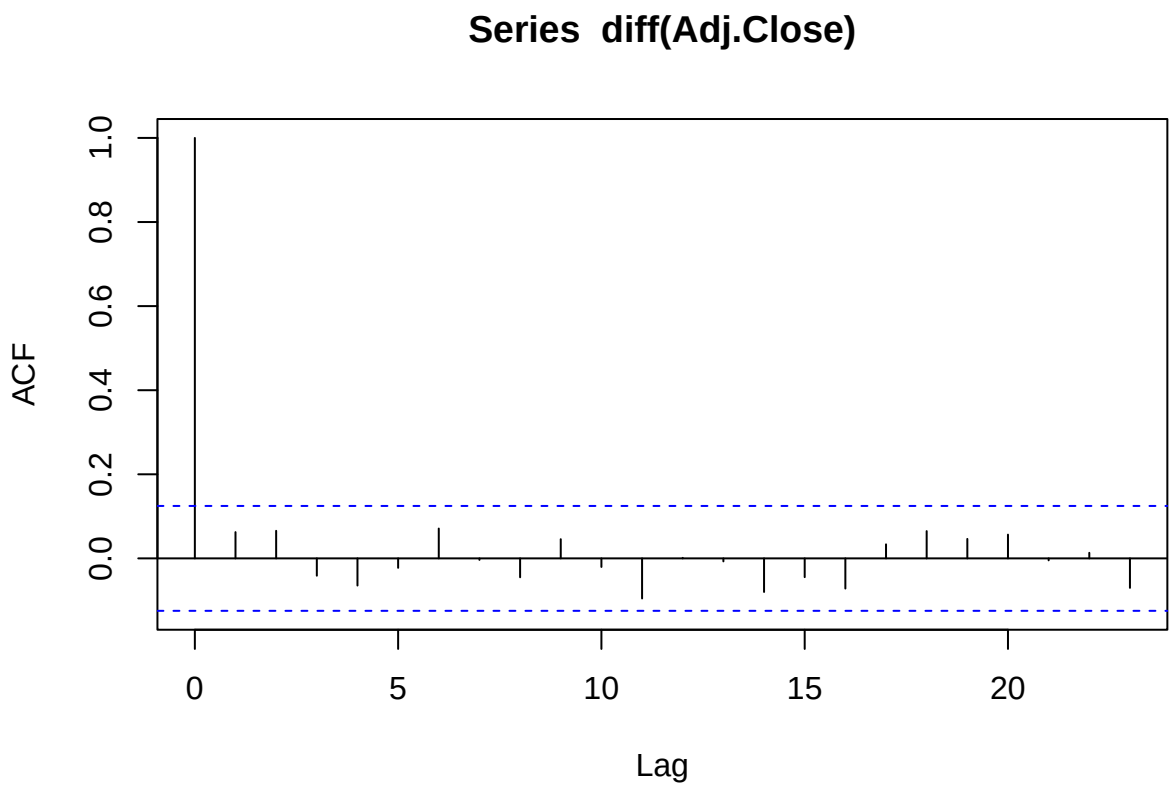
```
auto.arima(Adj.Close)
```

```
## Series: Adj.Close
## ARIMA(0,1,0)
##
## sigma^2 estimated as 644.1:  log likelihood=-1149.26
## AIC=2300.51   AICc=2300.53   BIC=2304.02
```

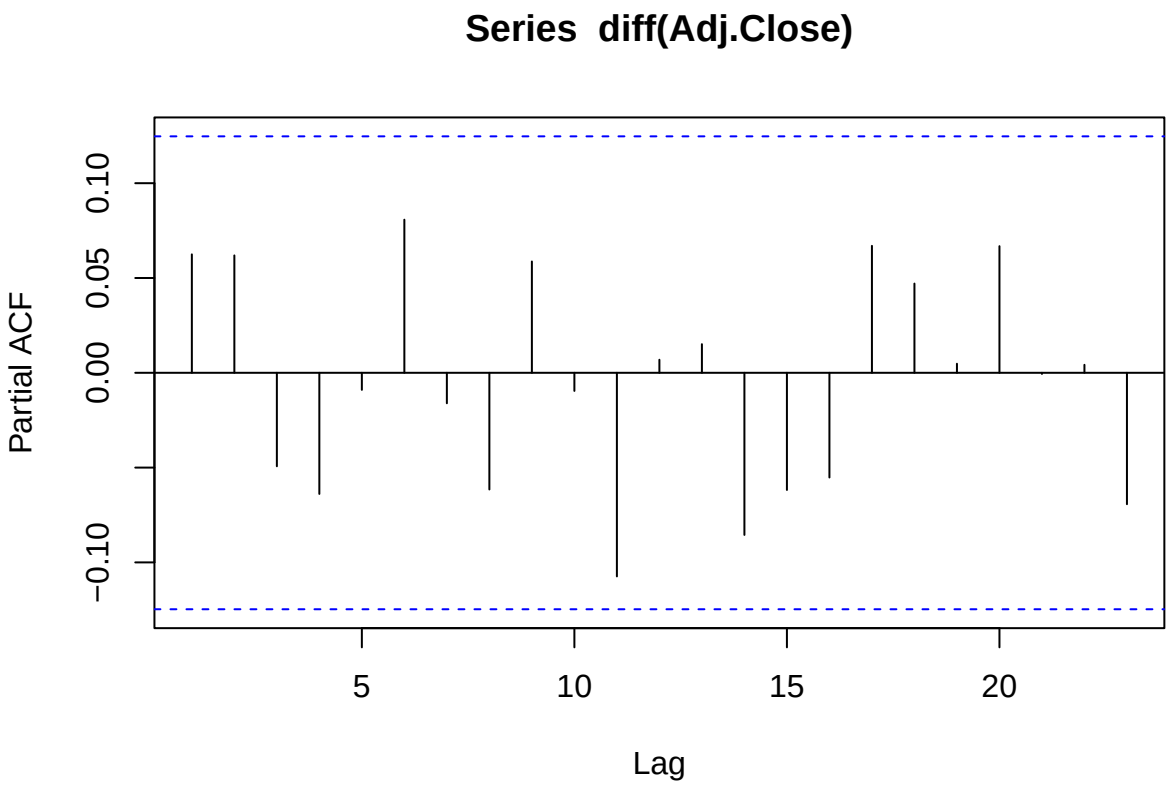
(0,1,0) shows thats lags are irrelevant

we plot acf and pacf to confirm this

```
acf(diff(Adj.Close))
```




```
pacf(diff(Adj.Close))
```



As both plot are withing the boundaries the lags are irrelavant

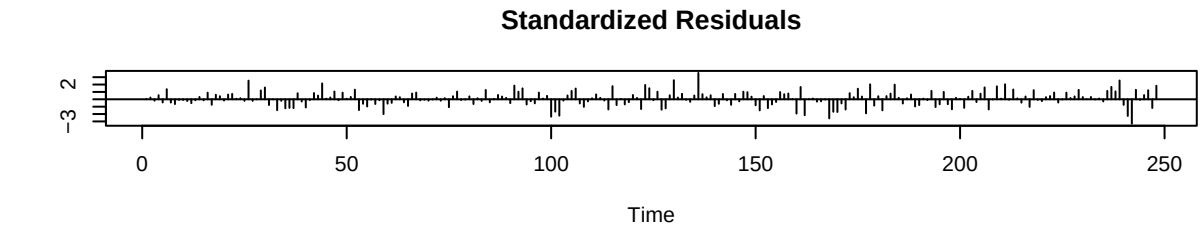
Parameter estimation

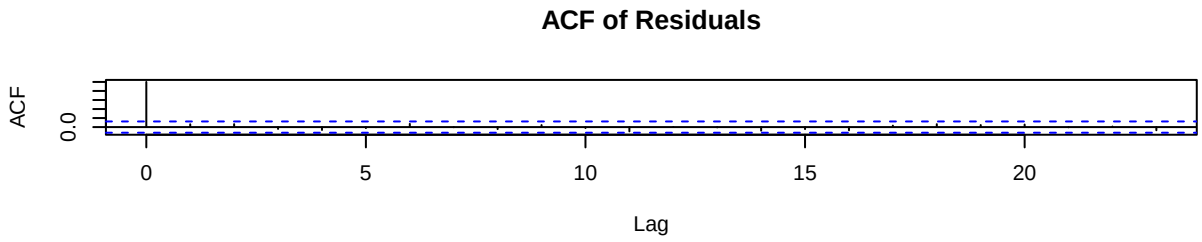
```
aa<- arima(Adj.Close,order = c(0,1,0))
aa
```

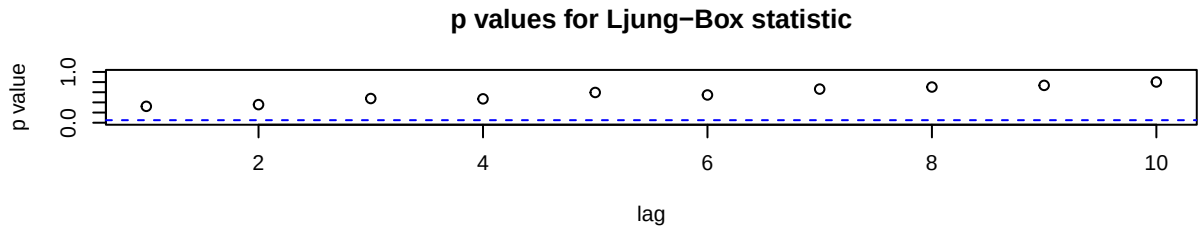
```
##
## Call:
## arima(x = Adj.Close, order = c(0, 1, 0))
##
##
## sigma^2 estimated as 644.1:  log likelihood = -1149.26,  aic = 2300.51
```

Check if parameters are iid

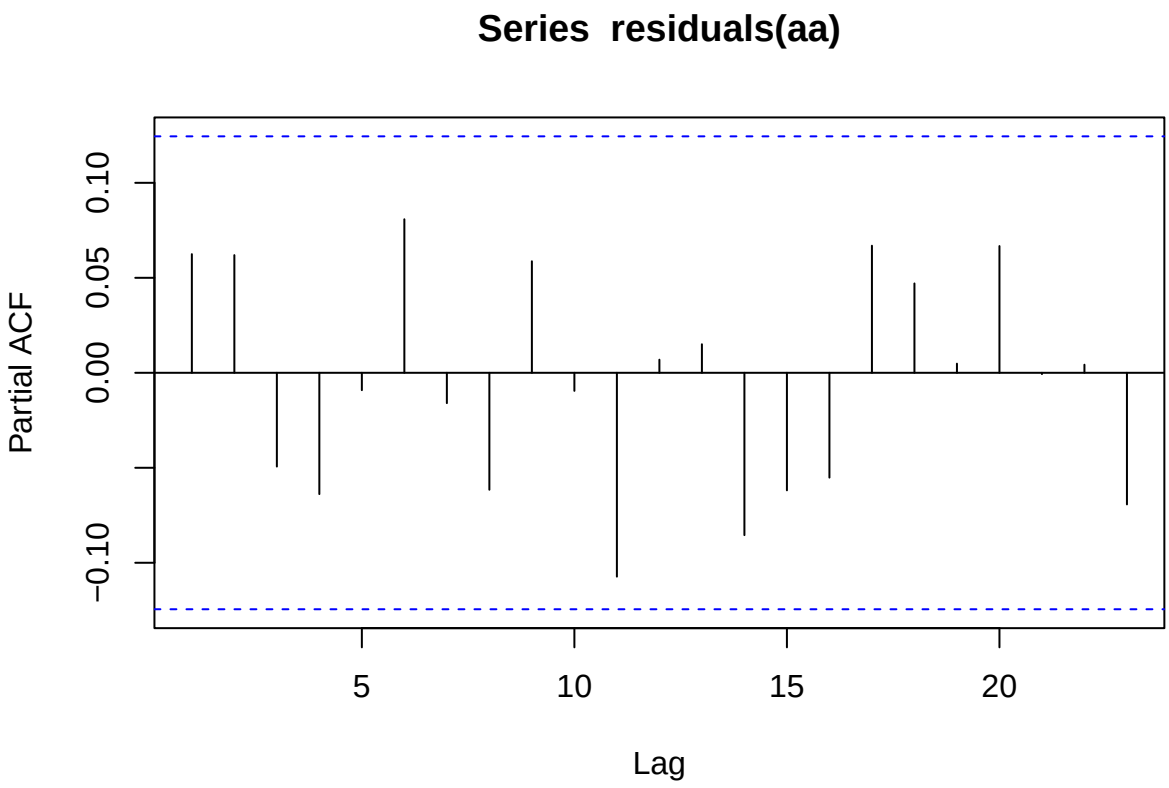
```
tsdiag(aa)
```








```
pacf(residuals(aa))
```



The parameters are iid.

Test Significance

```
# box test looks at residuals are iid, h0: they are iid
Box.test(residuals(aa),lag=10)
```

```
##
##  Box-Pierce test
##
## data:  residuals(aa)
## X-squared = 5.9957, df = 10, p-value = 0.8156
```

As p-value is more than 0.05 we fail to reject null hypothesis The residuals are iid.

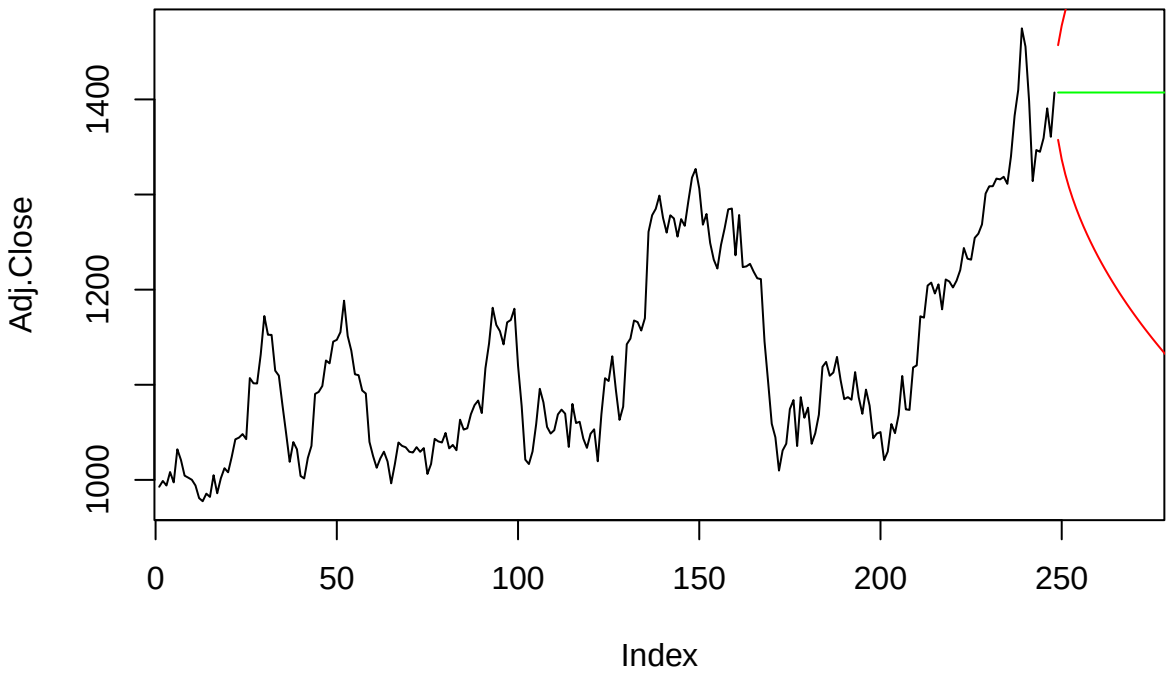
Prediction

```
p1<- predict(aa,n.ahead = 100)
p1
```

```
## $pred
## Time Series:
## Start = 249
## End = 348
## Frequency = 1
##   [1] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [10] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [19] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [28] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [37] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [46] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [55] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [64] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [73] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [82] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
##  [91] 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25 1407.25
## [100] 1407.25
##
## $se
## Time Series:
## Start = 249
## End = 348
## Frequency = 1
##   [1]  25.37895  35.89125  43.95762  50.75789  56.74905  62.16547  67.14638
##   [8]  71.78250  76.13684  80.25528  84.17244  87.91525  91.50509  94.95932
##  [15]  98.29224 101.51579 104.64008 107.67375 110.62426 113.49810 116.30094
##  [22] 119.03781 121.71315 124.33094 126.89473 129.40774 131.87287 134.29276
##  [29] 136.66981 139.00621 141.30399 143.56500 145.79095 147.98342 150.14387
##  [36] 152.27368 154.37410 156.44633 158.49147 160.51055 162.50455 164.47437
##  [43] 166.42088 168.34489 170.24715 172.12839 173.98929 175.83050 177.65263
##  [50] 179.45625 181.24193 183.01019 184.76152 186.49641 188.21530 189.91864
##  [57] 191.60684 193.28030 194.93939 196.58447 198.21591 199.83402 201.43914
##  [64] 203.03157 204.61161 206.17954 207.73563 209.28015 210.81336 212.33550
##  [71] 213.84680 215.34750 216.83781 218.31795 219.78812 221.24853 222.69935
##  [78] 224.14079 225.57301 226.99620 228.41052 229.81613 231.21321 232.60189
##  [85] 233.98232 235.35467 236.71905 238.07562 239.42450 240.76583 242.09972
##  [92] 243.42630 244.74570 246.05801 247.36337 248.66188 249.95363 251.23875
##  [99] 252.51733 253.78946
```

plots

```
plot(Adj.Close,type = 'l',xlim=c(10,268))
lines(p1$pred,col='green')
#range of prediction i.e pred +1.96*se
lines(p1$pred+(1.96*p1$se),col='red')
lines(p1$pred-(1.96*p1$se),col='red')
```


Adj.Close
1000
1200
1400
0
50
100
150
200
250
Index