

Basic R Functions of Handling Financial Data

Zheng Tian

Contents

1	Linear Time Series Models with R	1
2	Unit-Root Non-Stationarity	5

1 Linear Time Series Models with R

We consider the example of the monthly value-weighted return of IBM stock, which uses the data file `m-ibm3dx2608.txt`.

```
ibm_rtn <- read.table('m-ibm3dx2608.txt', header = TRUE)
head(ibm_rtn)
```

```
##      date    ibmrtn    vwrtn    ewrtn    sprtn
## 1 19260130 -0.010381  0.000724  0.023174  0.022472
## 2 19260227 -0.024476 -0.033374 -0.053510 -0.043956
## 3 19260331 -0.115591 -0.064341 -0.096824 -0.059113
## 4 19260430  0.089783  0.038358  0.032946  0.022688
## 5 19260528  0.036932  0.012172  0.001035  0.007679
## 6 19260630  0.068493  0.056888  0.050487  0.043184
```

```
# tail(ibm_rtn)
```

We can convert the first column in `ibm_rtn` to a `Date` object, and convert the third column, which is the monthly value-weighted return, to a `ts` object.

```
ibm_date <- as.Date(as.character(ibm_rtn$date), "%Y%m%d")
class(ibm_date)
```

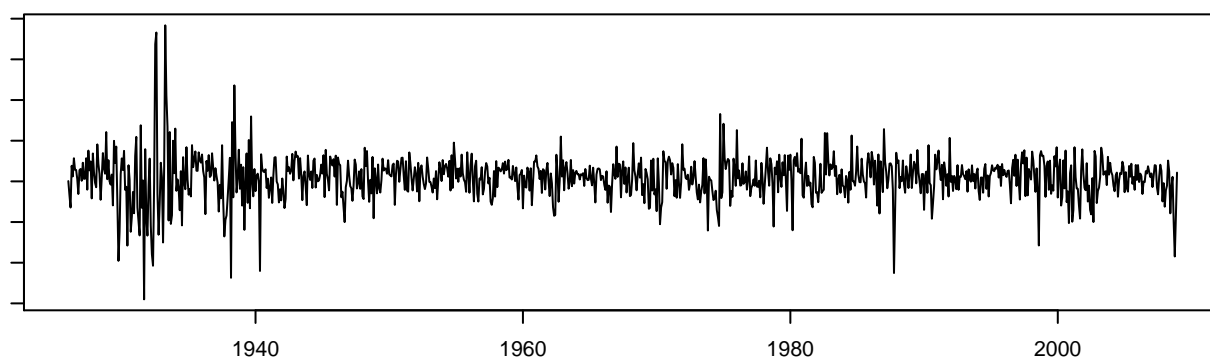
```
## [1] "Date"
```

```
vw_rtn <- ts(ibm_rtn[, 3], start=c(1926, 1), frequency = 12)
```

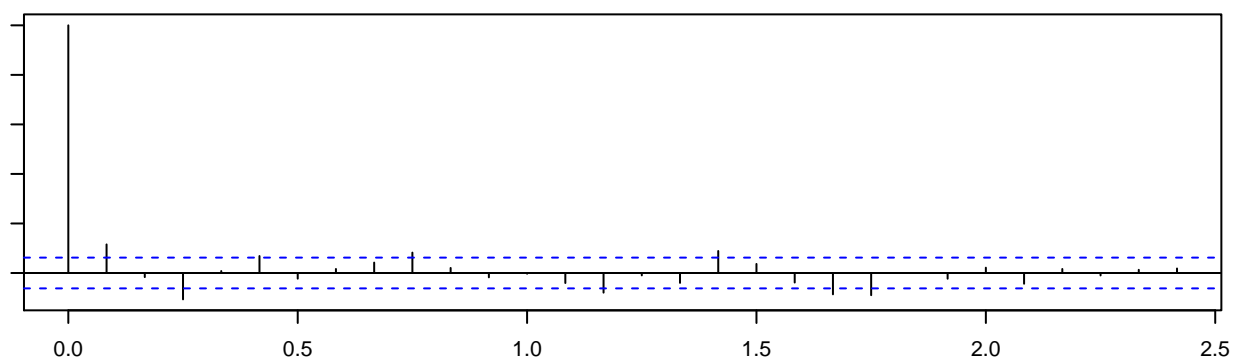
We plot the series and its ACF and PACF.

```
op <- par(mfrow = c(3, 1), mar=c(3, 1, 3, 1), pty = "m")
plot(vw_rtn, main="The monthly value-weighted return of IBM stock")
acf_vw_rtn <- acf(vw_rtn, main="The ACF of the series")
pacf_vw_rtn <- pacf(vw_rtn, main="The PACF of the series")
```

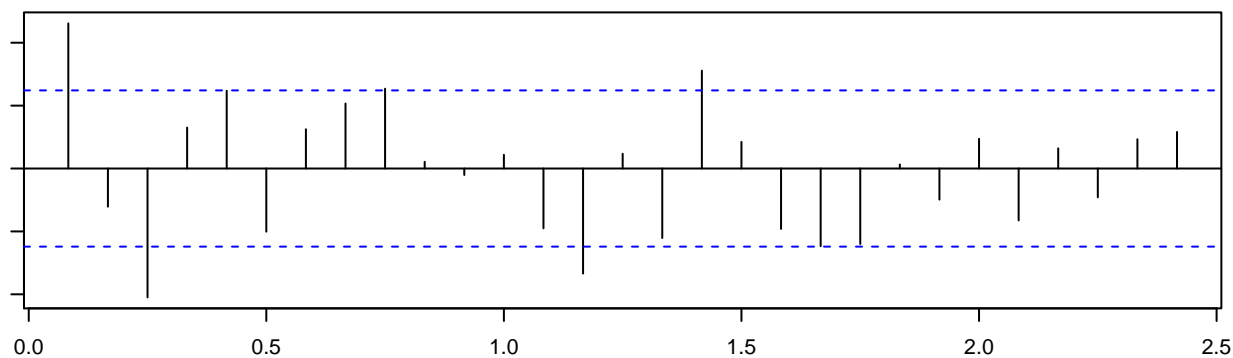
The monthly value-weighted return of IBM stock



The ACF of the series



The PACF of the series



```
par(op)
```

We use an AR(3) model to examine the return series, i.e.,

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \phi_3 r_{t-3} + a_t$$

```
vw_ar3 <- arima(vw_rtn, order=c(3,0,0))
vw_ar3
```

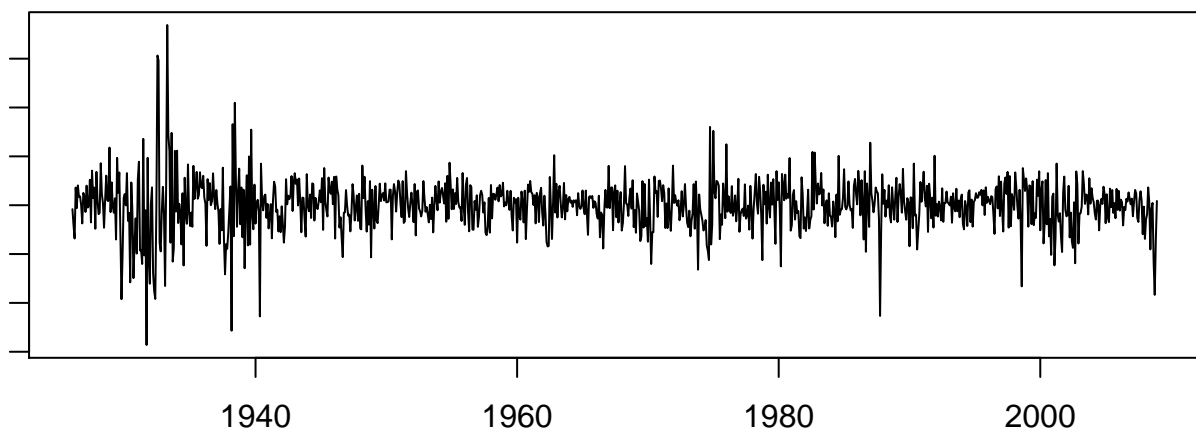
```
##
## Call:
## arima(x = vw_rtn, order = c(3, 0, 0))
##
```

```
## Coefficients:
##          ar1      ar2      ar3  intercept
##      0.1158 -0.0187 -0.1042    0.0089
## s.e.  0.0315  0.0317  0.0317    0.0017
##
## sigma^2 estimated as 0.002875:  log likelihood = 1500.86,  aic = -2991.73
```

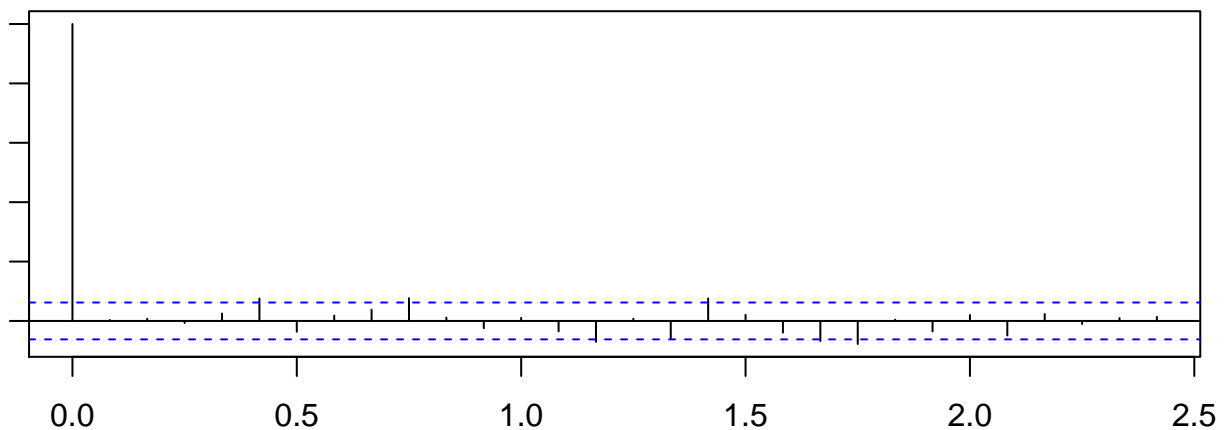
We can check the model adequacy by testing whether the residual series follows a white noise series. First, we plot the ACF of the residuals and then do a Ljung-Box test.

```
vw_ar3_resid <- resid(vw_ar3)
op <- par(mfrow = c(2, 1), mar=c(3, 1, 3, 1))
plot(vw_ar3_resid, main="The residuals from the AR(3) model")
acf(vw_ar3_resid, main="The ACF of residuals")
```

The residuals from the AR(3) model



The ACF of residuals



```
par(op)
```

The textbook suggest using the chi-squared distribution with the degree-of-freedom adjustment for the Ljung-Box statistics.

```
m <- 12
g <- length(coef(vw_ar3)) - 1
```

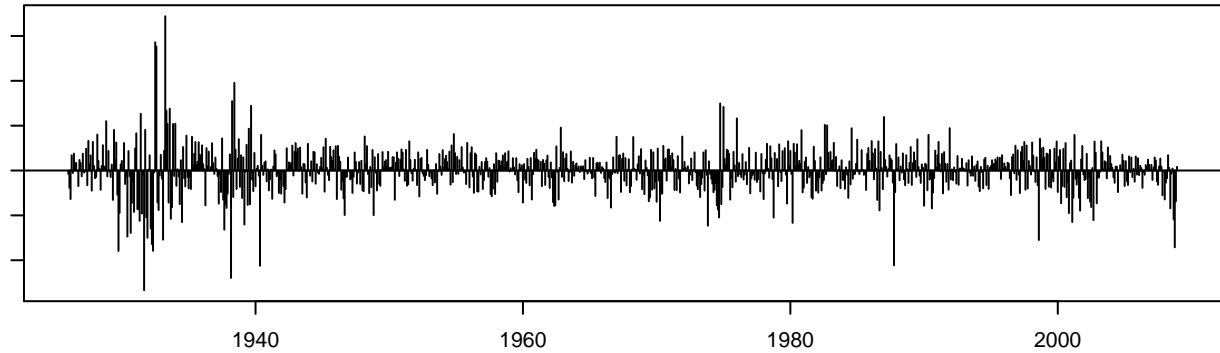
```
vw_ar3_lbtest <- Box.test(vw_ar3_resid, lag = m, type = "Ljung")
pv_ar3 <- 1 - pchisq(vw_ar3_lbtest$statistic, m-g)
```

The p-value is 0.0599 greater than the 5% significant level so that we cannot reject the null hypothesis that the residuals are white noises, implying that the AR(3) model is adequate.

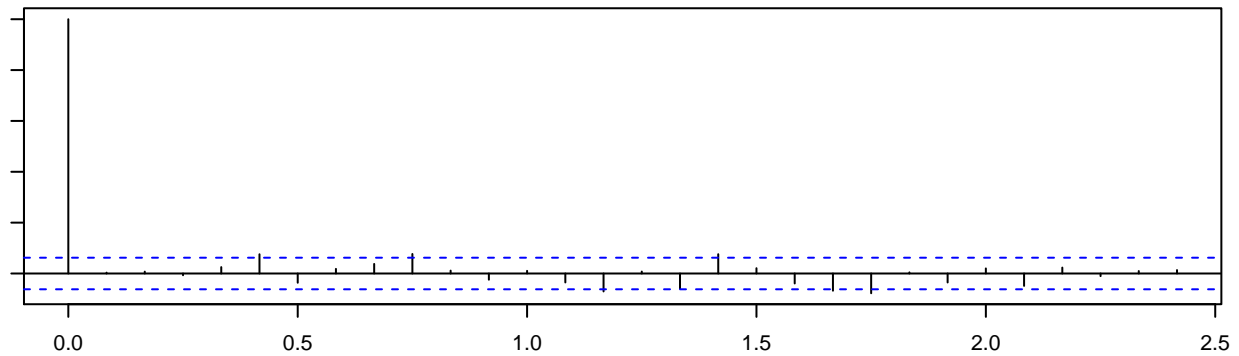
Also, there is a function called `tsdiag` to check the residuals from a model estimation.

```
op <- par(mfrow = c(3, 1), mar=c(3, 1, 3, 1), pty = "m")
tsdiag(vw_ar3, 12)
```

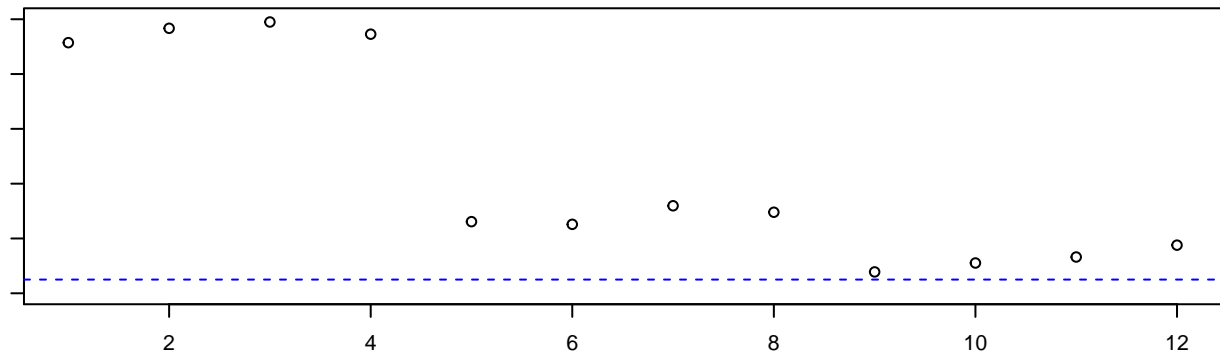
Standardized Residuals



ACF of Residuals



p values for Ljung-Box statistic



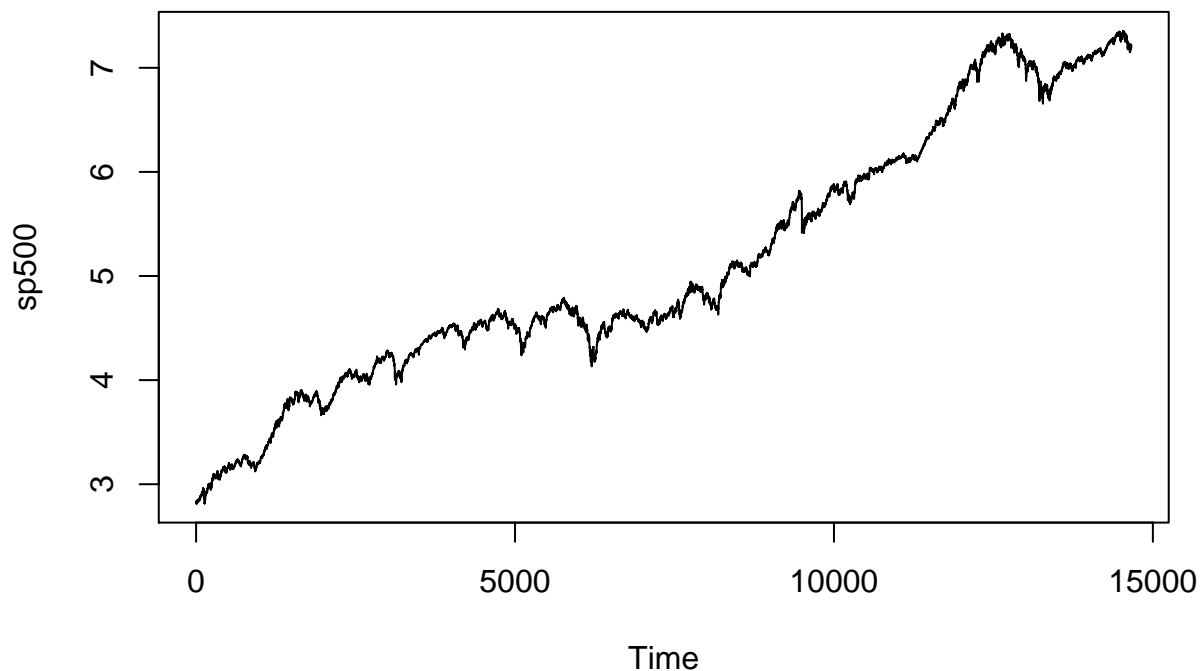
```
par(op)
```

2 Unit-Root Non-Stationarity

Finally, we demonstrate how to use R to examine a unit-root test with the example of the S&P 500 index.

```
sp_data <- read.table("d-sp55008.txt", header=TRUE)
sp500 <- ts(log(sp_data[, 7]))
plot(sp500, main="The time series of S&P 500 index")
```

The time series of S&P 500 index



The series appears to be non-stationary and have a time trend. We can use the augmented Dickey-Fuller test to check the existence of unit-roots, and model the series with a time trend.

```
library(fUnitRoots, quietly = TRUE)
```

```
##
## Rmetrics Package fBasics
## Analysing Markets and calculating Basic Statistics
## Copyright (C) 2005-2014 Rmetrics Association Zurich
## Educational Software for Financial Engineering and Computational Science
## Rmetrics is free software and comes with ABSOLUTELY NO WARRANTY.
## https://www.rmetrics.org --- Mail to: info@rmetrics.org
##
## Attaching package: 'fUnitRoots'
```

```

## The following objects are masked from 'package:urca':
##
##      punitroot, qunitroot, unitrootTable
# Carry out the Dickey-Fuller test with a constant and time trend
adfTest(sp500, lags=2, type="ct")

##
## Title:
## Augmented Dickey-Fuller Test
##
## Test Results:
##   PARAMETER:
##     Lag Order: 2
##   STATISTIC:
##     Dickey-Fuller: -2.0179
##   P VALUE:
##     0.5708
##
## Description:
## Tue Apr 18 15:45:09 2017 by user:
# Model the series with time trend
dsp500 <- diff(sp500)
tdx <- 1:length(dsp500)
m3 <- arima(dsp500, order=c(2,0,0), xreg=tdx)
m3

##
## Call:
## arima(x = dsp500, order = c(2, 0, 0), xreg = tdx)
##
## Coefficients:
##           ar1          ar2  intercept    tdx
##      0.0721  -0.0387      4e-04     0
## s.e.  0.0083   0.0083      2e-04     0
##
## sigma^2 estimated as 8.068e-05:  log likelihood = 48286.95,  aic = -96563.91

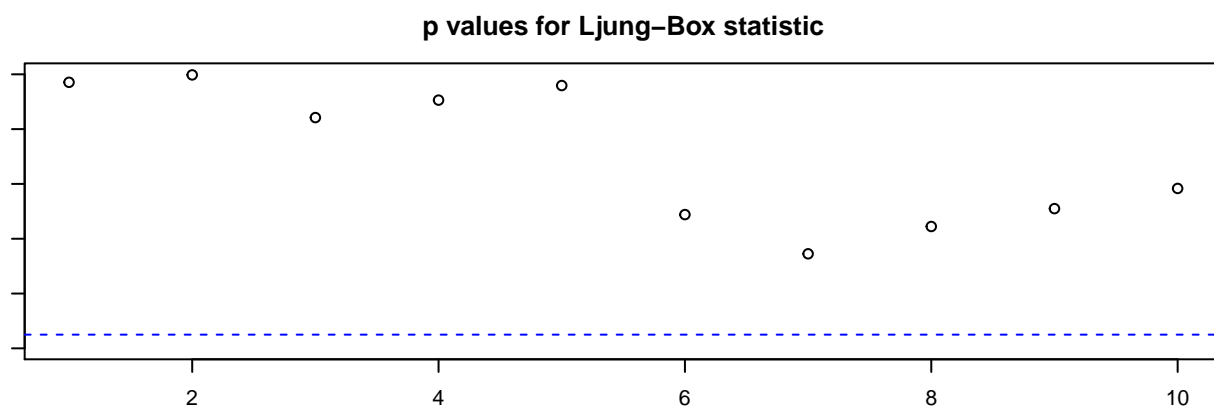
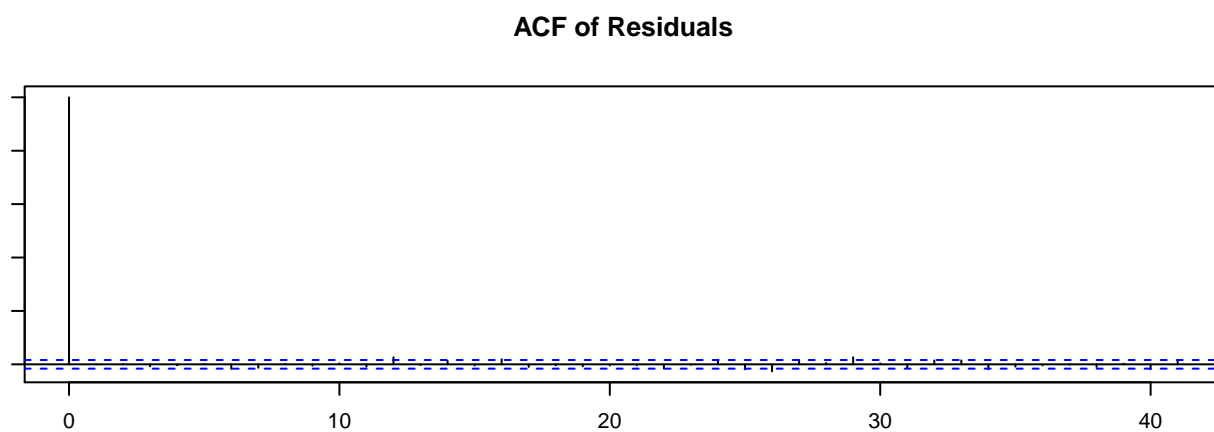
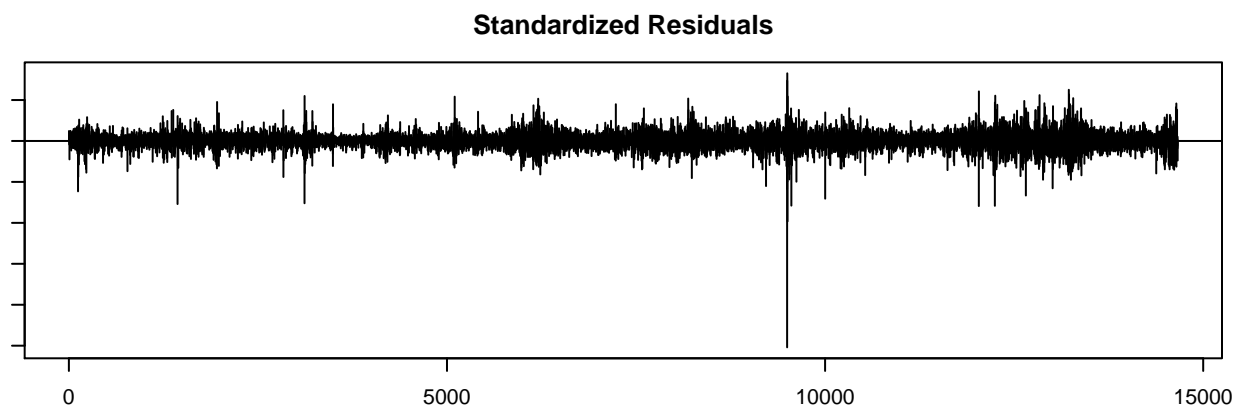
We then check the model adequacy with t statistics and the Ljung-Box test for the residuals.

# check the t statistics
tratio <- m3$coef / sqrt(diag(m3$var.coef)) # compute t-ratio
tratio

##           ar1          ar2  intercept          tdx
## 8.683863146 -4.669296791  2.285809057 -0.000858162

# check the residuals
op <- par(mfrow = c(3, 1), mar=c(3, 1, 3, 1), pty = "m")
tsdiag(m3)

```



```
par(op)
```