Homework 3 - FE 570

Siddharth Iyer

3/30/2022

Packages

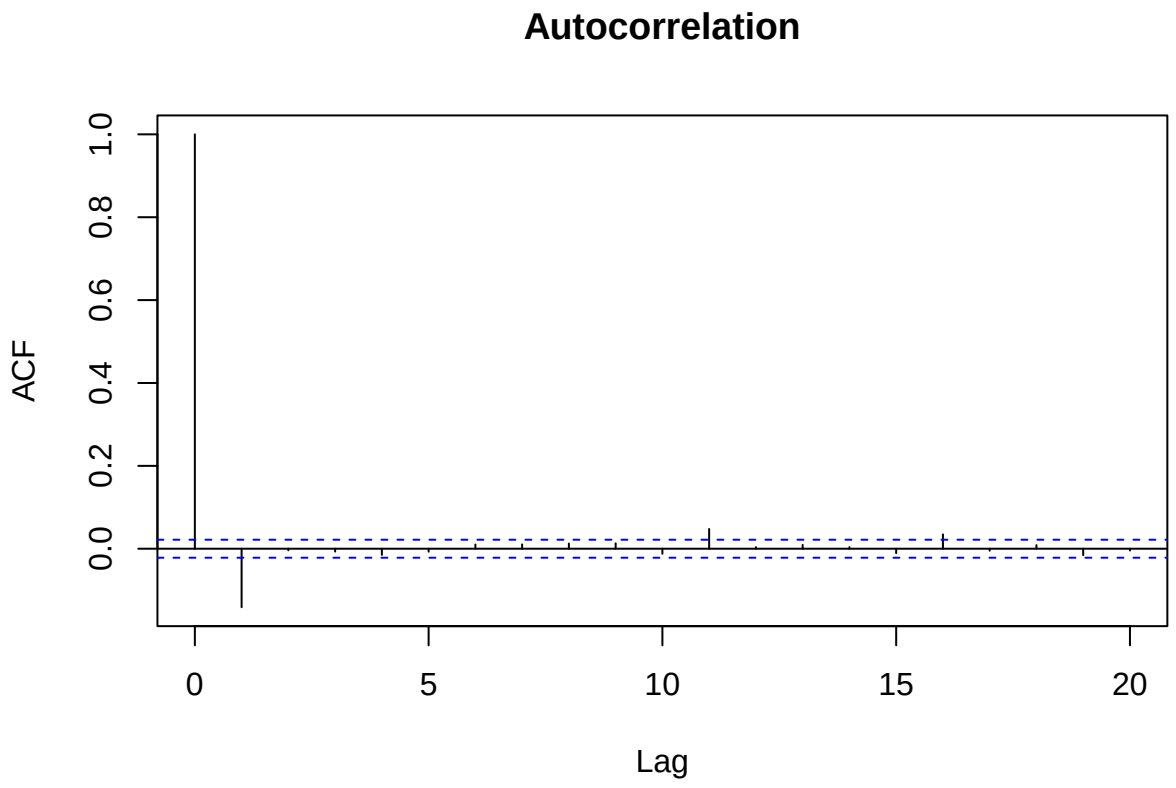
```
library(xts)
library(highfrequency)
```

Load Dataset

```
# load the TAQ dataset
Sys.setenv(TZ = "GMT") # work in East Coast Time Zone
options(digits.secs=3)
load("sampleTQdata.RData")
```

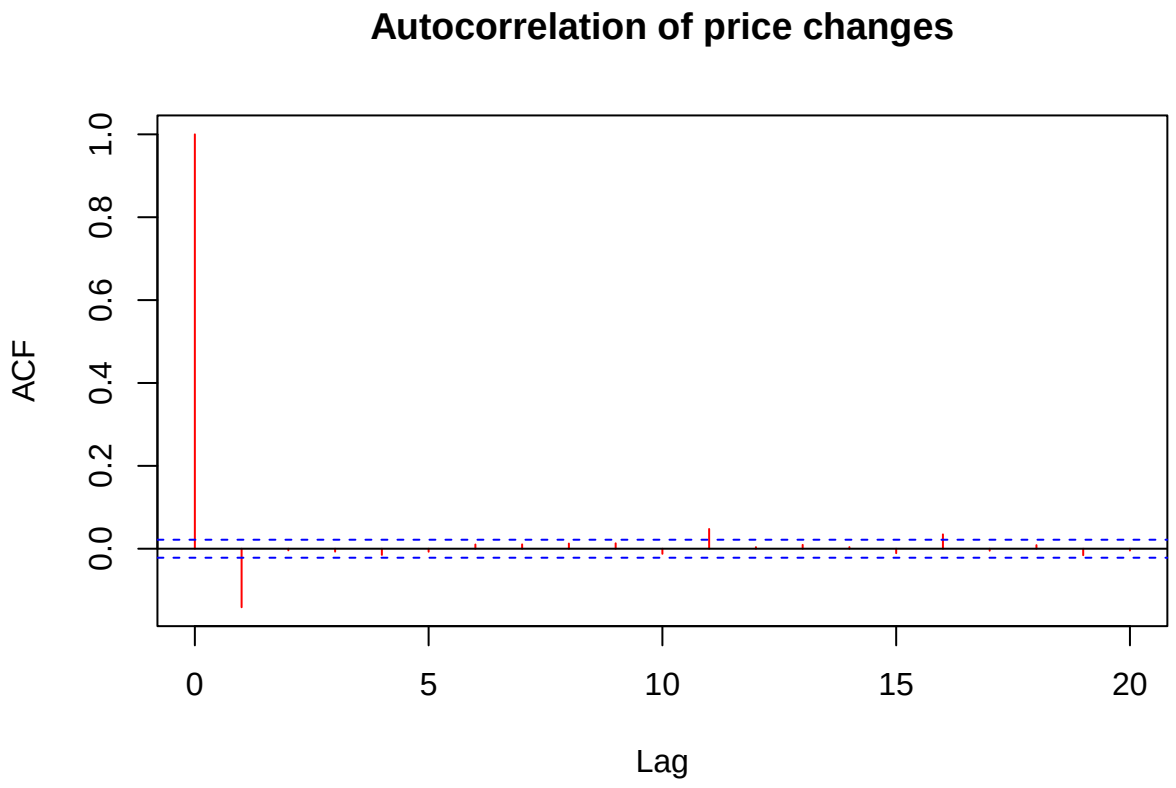
Problem 3.1.1: Calibrate Roll Model

```
# Calibrate the Roll model on the tqdata
pr <- as.numeric(tqdata$PRICE)
dpr <- diff(pr)  # Delta price
acpr <- acf(dpr, lag.max=20, type="correlation",
            plot=TRUE, main="Autocorrelation")
```




```
# acf: both autocorrelation and autocovariance

plot(acpr, col="red", lag.max=20,
     main="Autocorrelation of price changes")
```




```
# Roll estimate of bid-ask spread

covpr <- acf(dpr, lag.max=20, type="covariance", plot=FALSE)

gamma0 <- sd(dpr)^2
#print(gamma0)

gamma0alt <- covpr$acf[1]
#print(gamma0alt)

gamma1 <- covpr$acf[2]
#print(gamma1)

cparam.1 <- sqrt(-covpr$acf[2])
cat("bid-ask spread (2*c): ", 2*cparam.1, "\n")
```

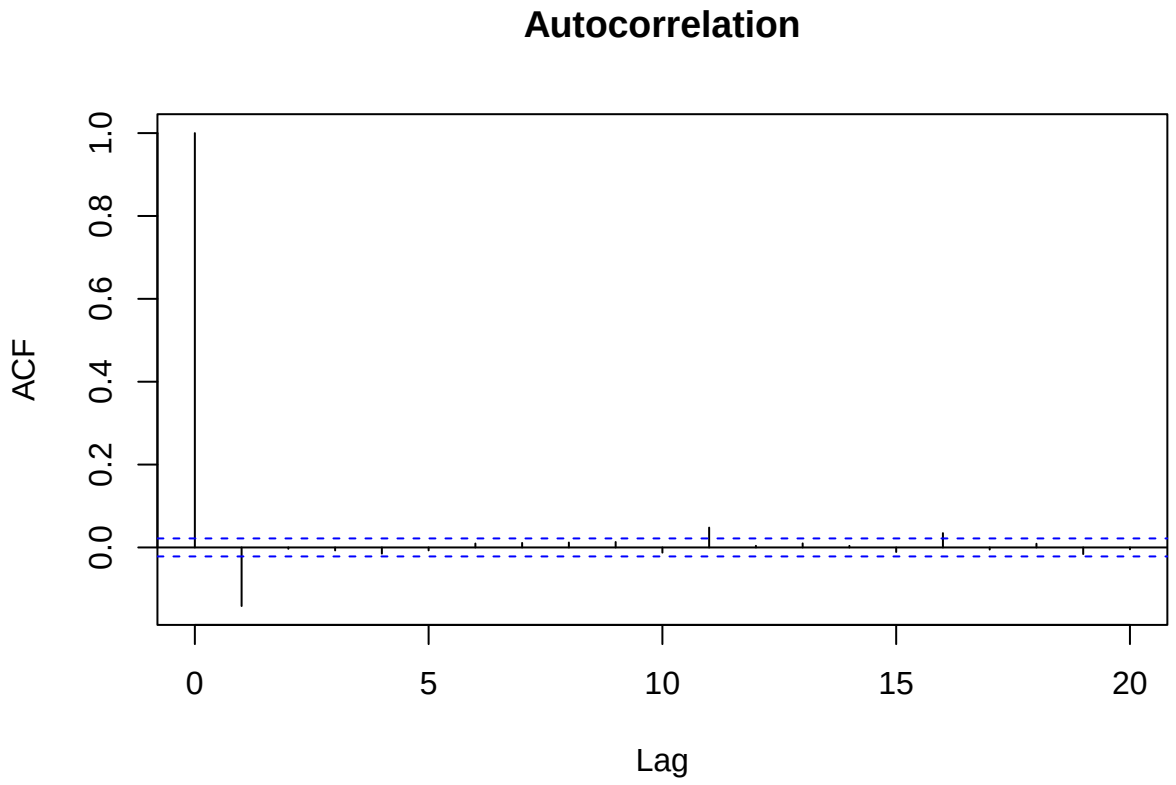
```
## bid-ask spread (2*c):  0.04036036
```

```
sig2u <- gamma0 + 2* gamma1
sigu.1 <- sqrt(sig2u)
cat("sigma: ", sigu.1, "\n")
```

```
## sigma:  0.04549637
```

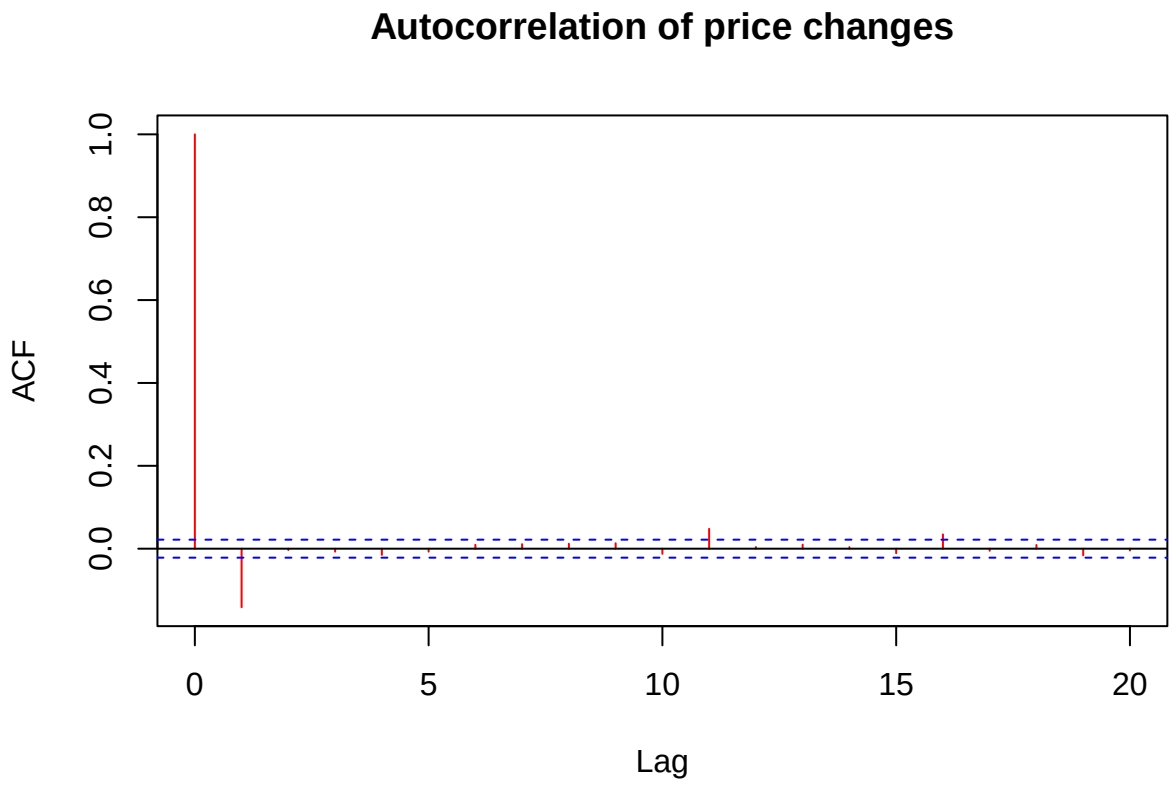
Problem 3.1.2: Calibrate Roll Model with log(pt)

```
# Calibrate the Roll model on the tqdata
pr <- log(as.numeric(tqdata$PRICE))
dpr <- diff(pr)  # Delta price
acpr <- acf(dpr, lag.max=20, type="correlation",
            plot=TRUE, main="Autocorrelation")
```




```
# acf: both autocorrelation and autocovariance

plot(acpr, col="red", lag.max=20,
     main="Autocorrelation of price changes")
```




```
# Roll estimate of bid-ask spread

covpr <- acf(dpr, lag.max=20, type="covariance", plot=FALSE)

gamma0 <- sd(dpr)^2
#print(gamma0)

gamma0alt <- covpr$acf[1]
#print(gamma0alt)

gamma1 <- covpr$acf[2]
#print(gamma1)

cparam.2 <- sqrt(-covpr$acf[2])
cat("bid-ask spread (2*c): ", 2*cparam.2, "\n")
```

```
## bid-ask spread (2*c):  0.0002112998
```

```
sig2u <- gamma0 + 2* gamma1
sigu.2 <- sqrt(sig2u)
cat("sigma: ", sigu.2, "\n")
```

```
## sigma:  0.0002380327
```

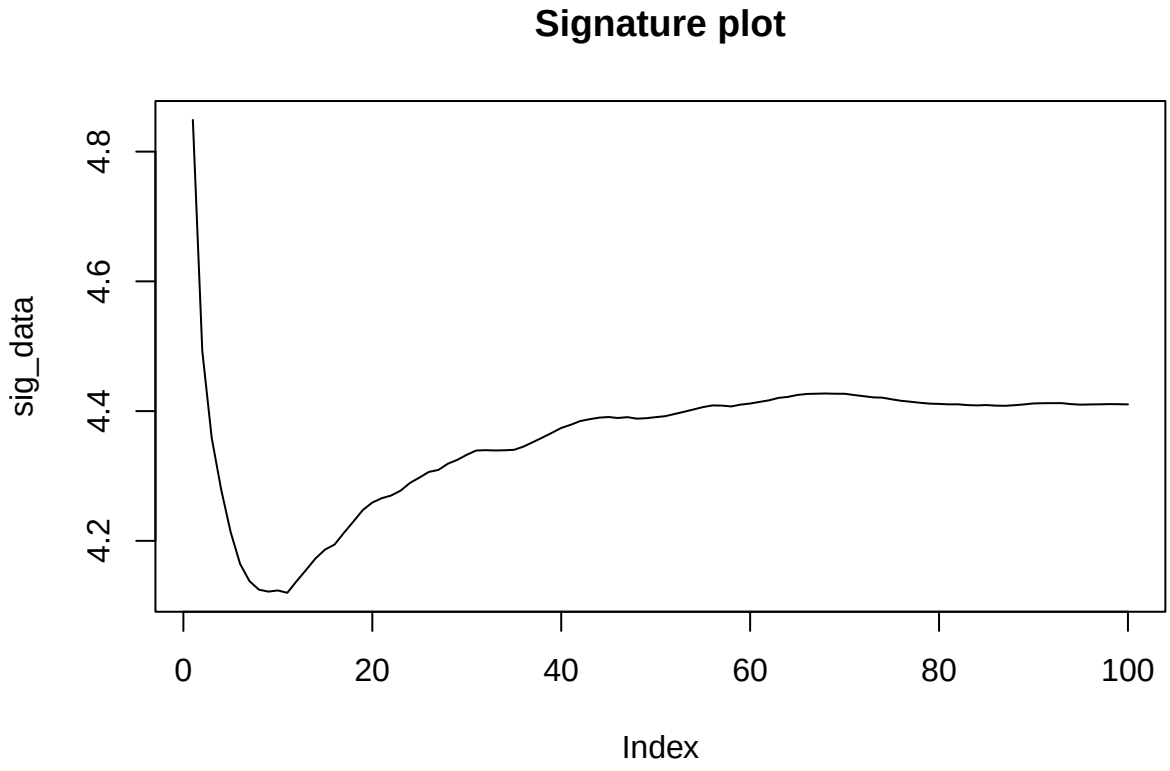
When we are using log(price) and applying diff on the datasets, we are effectively taking the spread on returns data. I don't this this makes much sense.

Problem 3.1.3

```
realizedVar <- function(q){
  pr <- as.numeric(tqdata$PRICE)
  rCov(diff(pr, lag=q, differences=1))/q
}
# compute the signature plot sigma.day(q) = sqrt(RV(q))
sig_data <- NULL

for(q in 1:100){
  sig_data <- c(sig_data, sqrt(realizedVar(q)))
}

plot(sig_data, type ="l", main="Signature plot")
```



We see that rougly 10 day lag is the minimum realized volatility, after which the graph grows at a decreasing rate and plateaus.

Problem 3.1.4

```
n_trades = length(tqdata)
sig_roll = sqrt(n_trades * (sigu.1)^2)
cat("Roll: ", sig_roll, "\n")
```

```
## Roll:  12.32414
```

```r
q_5min = n_trades*5/390
cat("Realized Var: ", realizedVar(q_5min))
```

```
## Realized Var:  12.79712
```

Incredible! The Roll model and the Realized volatility is approximately the same at a log of 5 minutes.