Homework 4

YueLi

15620161152260

Q1

- #20171017 JSON input from CRIX , trial done in XMNinstall.packages("rjson",repos="http://cran.us.r-project.org")
- library("rjson")
- json_file = "http://crix.hu-berlin.de/data/crix.json"
- json_data = fromJSON(file=json_file)
- crix_data_frame = as.data.frame(json_data)
- x = crix_data_frame
- n = dim(x)
- a = seq(1,n[2],2)
- b = seq(2,n[2],2)

Q1

- #figure 3 : The daily value of CRIX
- date = t(x[1,a])
- price = t(x[1,b])
- plot(price)

- #figure 4 : The log returns of CRIX index
- dim(price)
- ts.plot(price)
- ret = diff(log(price))
- ts.plot(ret)

Q1

- #figure 5 : Histogram
- hist(ret, col = "grey", breaks = 20, freq = FALSE, ylim = c(0, 25), xlab = NA)
- lines(density(ret), lwd = 2)
- mu = mean(ret)l,
- sigma = sd(ret)
- x = seq(-4, 4, length = 100)
- curve(dnorm(x, mean = mean(ret), sd = sd(ret)), add = TRUE, col = "darkblue", lwd = 2)

- #figure 6 : QQ plot
- qqnorm(ret)
- qqline(ret, col = "blue", lwd = 3)

Q2

- rm(list = ls(all = TRUE))
- graphics.off()

- # install and load packages
- libraries = c("zoo", "tseries")
- lapply(libraries, function(x) if (!(x %in% installed.packages())) {
- install.packages(x)
- })
- lapply(libraries, library, quietly = TRUE, character.only = TRUE)

Q2

- #RET
- ("rjson",repos="http://cran.us.r-project.org")
- library("rjson")
- json_file = "http://crix.hu-berlin.de/data/crix.json"
- json_data = fromJSON(file=json_file)
- crix_data_frame = as.data.frame(json_data)
- x = crix_data_frame
- n = dim(x)
- a = seq(1,n[2],2)
- b = seq(2,n[2],2)
- date = t(x[1,a])
- price = t(x[1,b])
- dim(price)
- ret = diff(log(price))

Q2

- # d order
- Box.test(ret, type = "Ljung-Box", lag = 20)

- # stationary test
- adf.test(ret, alternative = "stationary")
- kpss.test(ret, null = "Trend")

- par(mfrow = c(1, 2))
- # acf plot
- autocorr = acf(ret, lag.max = 20, ylab = "Sample Autocorrelation", main = NA,
- lwd = 2, ylim = c(-0.3, 1))

- # LB test of linear dependence
- print(cbind(autocorr$lag, autocorr$acf))
- Box.test(ret, type = "Ljung-Box", lag = 1, fitdf = 0)
- Box.test(autocorr$acf, type = "Ljung-Box")

Q2

- # plot of pacf
- autopcorr = pacf(ret, lag.max = 20, ylab = "Sample Partial Autocorrelation",
- main = NA, ylim = c(-0.3, 0.3), lwd = 2)
- print(cbind(autopcorr$lag, autopcorr$acf))

- # arima model
- par(mfrow = c(1, 1))
- auto.arima(ret)
- fit1 = arima(ret, order = c(1, 0, 1))
- tsdiag(fit1)
- Box.test(fit1$residuals, lag = 1)

Q2

```
# aic
aic = matrix(NA, 6, 6)
for (p in 0:4) {
  for (q in 0:3) {
    a.p.q = arima(ret, order = c(p, 0, q))
    aic.p.q = a.p.q$aic
    aic[p + 1, q + 1] = aic.p.q
  }
}
aic
```

Q2

```
# bic
bic = matrix(NA, 6, 6)
for (p in 0:4) {
  for (q in 0:3) {
    b.p.q = arima(ret, order = c(p, 0, q))
    bic.p.q = AIC(b.p.q, k = log(length(ret)))
    bic[p + 1, q + 1] = bic.p.q
  }
}
bic
```

Q2

- # select p and q order of ARIMA model
- fit4 = arima(ret, order = c(2, 0, 3))
- tsdiag(fit4)
- Box.test(fit4$residuals, lag = 1)

- fitr4 = arima(ret, order = c(2, 1, 3))
- tsdiag(fitr4)
- Box.test(fitr4$residuals, lag = 1)

Q2

- # to conclude, 202 is better than 213
- fit202 = arima(ret, order = c(2, 0, 2))
- tsdiag(fit202)
- tsdiag(fit4)
- tsdiag(fitr4)

- AIC(fit202, k = log(length(ret)))
- AIC(fit4, k = log(length(ret)))
- AIC(fitr4, k = log(length(ret)))
- fit202$aic
- fit4$aic
- fitr4$aic

Q2

- # arima202 predict
- fit202 = arima(ret, order = c(2, 0, 2))
- crpre = predict(fit202, n.ahead = 30)

- dates = seq(as.Date("02/08/2014", format = "%d/%m/%Y"), by = "days", length = length(ret))

- plot(ret, type = "l", xlim = c(0, 644), ylab = "log return", xlab = "days",
- lwd = 1.5)
- lines(crpre$pred, col = "red", lwd = 3)
- lines(crpre$pred + 2 * crpre$se, col = "red", lty = 3, lwd = 3)
- lines(crpre$pred - 2 * crpre$se, col = "red", lty = 3, lwd = 3)