5261-Project

Bo Peng

April 17, 2017

Read data

```
library(zoo)
```

```
## Warning: package 'zoo' was built under R version 3.3.3
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
LIBOR <- read.csv("../data/1-Month_LIBOR_daily.csv", header = TRUE, as.is = T)
LIBOR[,1] <- as.Date(LIBOR[,1], "%m/%d/%Y")
SHIBOR <- read.csv("../data/SHIBOR.csv", header = TRUE, as.is = T)
SHIBOR[,1] <- as.Date(SHIBOR[,1], "%m/%d/%Y")
CSI300 <- read.csv("../data/CSI300_daily.csv", header = TRUE, as.is = T)
CSI300[,1] <- as.Date(CSI300[,1], "%m/%d/%Y")
GSPC <- read.csv("../data/S&P500(^GSPC)_daily.csv", header = TRUE, as.is = T)
GSPC[,1] <- as.Date(GSPC[,1], "%m/%d/%Y")
EX <- read.csv("../data/ExRate.csv", header = TRUE, as.is = T)
EX[,1] <- as.Date(EX[,1], "%m/%d/%Y")
TRADE <- read.csv("../data/Trade.csv", header = TRUE, as.is = T)
TRADE[,1] <- as.Date(as.yearmon(TRADE[,1], "%Y/%m"))
TRADE <- TRADE[order(TRADE[,1]),]
M_US <- read.csv("../data/m1m2_US.csv", header = TRUE, as.is = T)
M_US[,1] <- as.Date(as.yearmon(M_US[,1], "%Y-%m"))
M_CN <- read.csv("../data/m1m2_China.csv", header = TRUE, as.is = T)
M_CN[,1] <- as.Date(as.yearmon(M_CN[,1], "%Y-%m"))
M_CN <- M_CN[order(M_CN[,1]),]
LIBOR <- LIBOR[-nrow(LIBOR),]
TRADE[1,1] <- TRADE[1,1]+1
M_CN[1,1] <- M_CN[1,1]+1
M_CN <- M_CN[-c(nrow(M_CN), nrow(M_CN)-1), ]
M_US[1,1] <- M_US[1,1]+1
```

Combine into main dataframe

```
DATE <- seq(min(LIBOR[,1]), max(LIBOR[,1]), by="days")
DATE <- data.frame(DATE)
main <- merge(CSI300, DATE, by.x='Date', by.y='DATE', all.x=T, all.y=T)
main <- merge(main, EX, by.x='Date', by.y='Date', all.x=T, all.y=T)
main <- merge(main, GSPC, by.x='Date', by.y='Date', all.x=T, all.y=T)
main <- merge(main, LIBOR, by.x='Date', by.y='DATE', all.x=T, all.y=T)
main <- merge(main, M_CN, by.x='Date', by.y='Time.Period', all.x=T, all.y=T)
main <- merge(main, M_US, by.x='Date', by.y='Time.Period', all.x=T, all.y=T)
main <- merge(main, SHIBOR, by.x='Date', by.y='Date', all.x=T, all.y=T)
colnames(main) <- c("Date", "CSI300", "RMBperUSD", "GSPC", "LIBOR", "m2_CN", "m1_CN", "m1_US", "m2_US",
main <- main[, c("Date", "RMBperUSD", "SHIBOR", "LIBOR", "CSI300", "GSPC", "m1_CN", "m2_CN", "m1_US", "
```

Use interpolation to fill missing spaces

```
for (i in 4:10) {
  main[,i][is.na(main[,i])] <- "."
  main[,i] <- as.numeric(paste(main[,i]))
}

for (i in 2:6) {
  main[3:nrow(main),i] <- na.approx(main[3:nrow(main),i])
}

for (j in 7:10) {
  for (i in 1:nrow(main)) {
    if (is.na(main[i,j])) {
      k <- max(which(is.na(main[1:i,j]) == FALSE))
      main[i,j] <- main[k,j]
    }
  }
}

main <- main[-c(1,2),]
save(main, file = "main.RData")
```

Scale regressors

```
main[,3:8] <- scale(main[,3:8])
```

Divide dataset into training and test sets

```
train <- main[1:3500,]
test <- main[3501:3709,]
```

Arima model with regressors

```
library(zoo)
library(forecast)
```

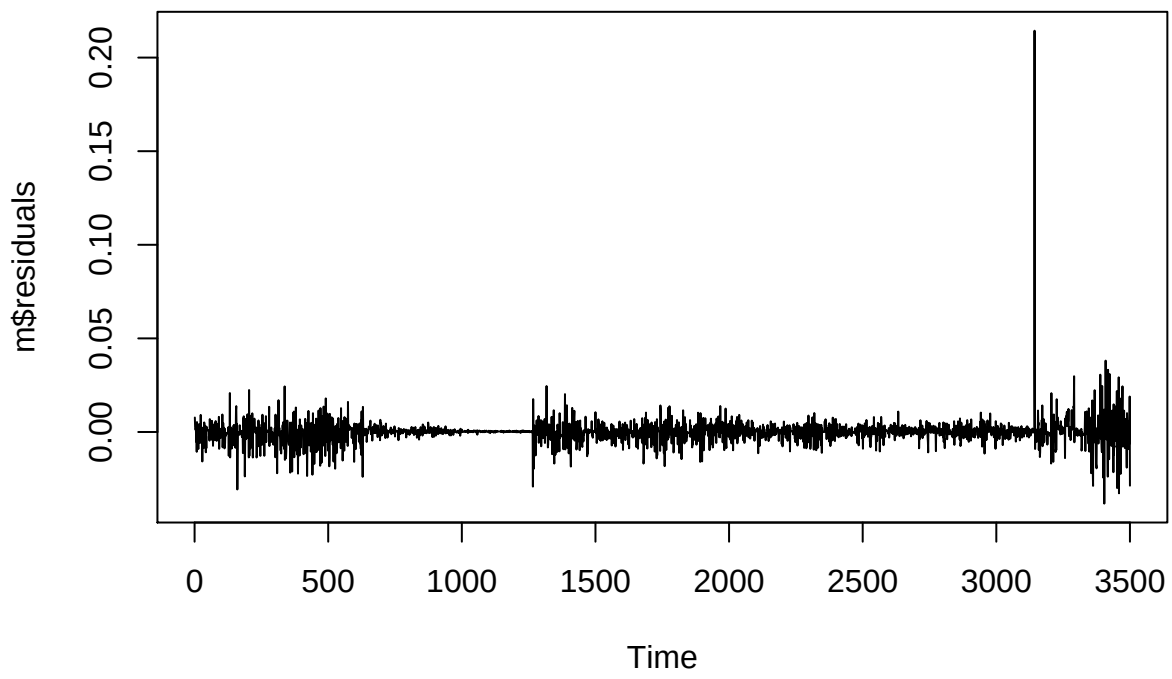
```
## Warning: package 'forecast' was built under R version 3.3.3
```

```
m <- auto.arima(train$RMBperUSD, xreg = train[,3:10])
m
```

```
## Series: train$RMBperUSD
## Regression with ARIMA(0,1,1) errors
##
## Coefficients:
##          ma1   drift  SHIBOR   LIBOR   CSI300    GSPC    m1_CN   m2_CN
##       0.1777  -3e-04  -1e-04  0.0236  -0.0051  0.0009  -0.0045  0.0090
## s.e.  0.0167   1e-04   9e-04  0.0109   0.0017  0.0035   0.0096  0.0233
##       m1_US  m2_US
##       0e+00      0
## s.e.  1e-04      0
##
## sigma^2 estimated as 3.776e-05:  log likelihood=12858.25
## AIC=-25694.5   AICc=-25694.43   BIC=-25626.74
```

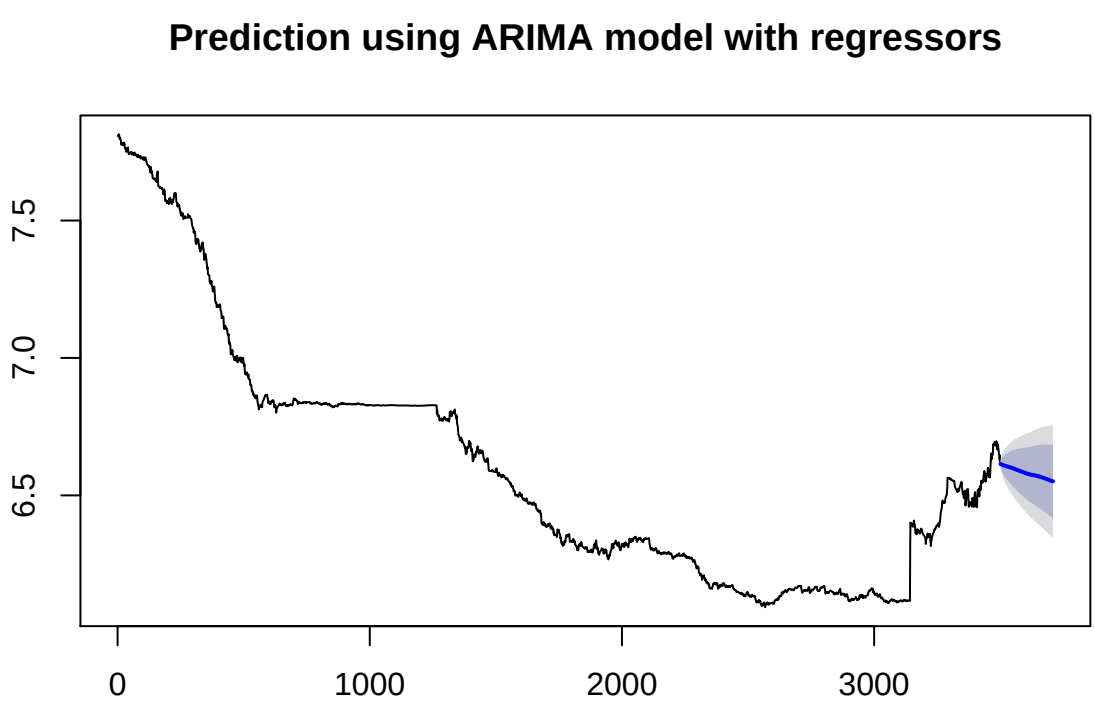
Plot residuals

```r
plot(m$residuals)
```



Use model to forecast over test set

```r
pred <- forecast(m, xreg = test[,3:10], h = 209)
plot(pred, main = "Prediction using ARIMA model with regressors")
```



Calculate mean test error

```
forecasted <- as.numeric(pred$mean)
actual <- test$RMBperUSD
errors <- (forecasted - actual)^2
mse <- mean(errors)
mse
```

```
## [1] 0.06113678
```