Chicago_Crimes

Shicheng Huang

October 22, 2016

Loading Necessary Packages

```r
sshhh <- function(a.package){
  suppressWarnings(suppressPackageStartupMessages(
    library(a.package, character.only=TRUE)))
}
pkgs <- c("knitr", "readr", "forecast", "dplyr")
for (pkg in pkgs) {
  sshhh(pkg)
}
```

Setup

```r
opts_chunk$set(cache=TRUE)
opts_chunk$set(root.dir = "./chicagoCrime")
```

Loading Datasets

```r
#loading crimes from 2001 to 2006 as training set
crimes_month <- read_csv("crimes_month_2001to2006.csv")
```

```
## Parsed with column specification:
## cols(
##   X1 = col_integer(),
##   monthYear = col_character(),
##   Count = col_integer()
## )
```

```r
crimes_month <- select(crimes_month, monthYear, Count)
#only need the crime report data by months vector
crimes <- crimes_month$Count
#using log transformation
crimes_log <- log(crimes)

#doing similar things for the datasets from 2007to2014
crimes_future <- read_csv("./chicagoCrime/crime_months.csv")
```

```
## Parsed with column specification:
## cols(
##   X1 = col_integer(),
##   monthYear = col_character(),
##   Count = col_integer()
## )
```

```r
crimes_future <- select(crimes_future, monthYear, Count)
crimes_f <- crimes_future$Count
```

```r
#trying to forecase 2007 to 2015 by updating each year with given information
crimes_log_temp <- crimes_log
next_year_temp <- numeric()
for (i in 1:8) {
  month_indices <- seq((i-1)*12 + 1, i*12)
  crimes_log_temp <- c(crimes_log_temp, log(crimes_f[month_indices]))
  auto_mod_log <- auto.arima(crimes_log_temp,
                             lambda=0,
                             d=0, D=12,
                             max.order=9,
                             stepwise=FALSE,
                             approximation=FALSE)
  forecast_temp <- exp(forecast(auto_mod_log, 12)$mean)
  next_year_temp <- c(next_year_temp, forecast_temp)
}
```

```r
plot(crimes_f, type="l")
par(new=TRUE)
lines(next_year_temp, col = "red")
```
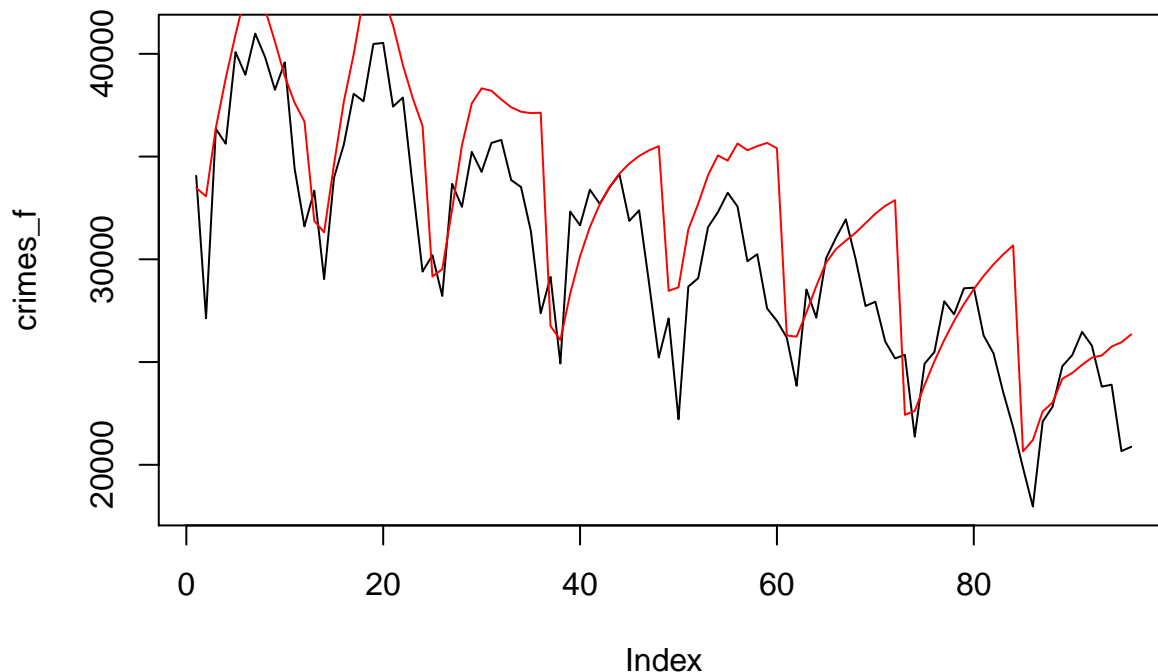


```r
## Forecasting without log transformation
crimes_temp <- crimes
next_year_temp <- numeric()
for (i in 1:8) {
  month_indices <- seq((i-1)*12 + 1, i*12)
  crimes_temp <- c(crimes_temp, crimes_f[month_indices])
  auto_mod <- auto.arima(crimes_temp,
                         lambda=0,
                         d=0, D=12,
                         max.order=9,
```

2

```r
                    stepwise=FALSE,
                    approximation=FALSE)
  forecast_temp <- forecast(auto_mod, 12)$mean
  next_year_temp <- c(next_year_temp, forecast_temp)
}
```

```r
plot(crimes_f, type="l")
par(new=TRUE)
lines(next_year_temp, col = "red")
```