GTx: ISYE6501x - Homework 3

Muh Alif Ahsanul Islam

5/31/2019

Question 7.1

Describe a situation or problem from your job, everyday life, current events, etc., for which exponential smoothing would be appropriate. What data would you need? Would you expect the value of (the first smoothing parameter) to be closer to 0 or 1, and why?

Answer:

When analyzing the trend of stock price of a company, it is often hard to determine the clear trend (up, down, or flat) because the price move up and down erratically. To solve this issue we can use exponential smoothing. The data needed is stock price.
Exponential smoothing formula:

$$S_t = \alpha x_t + (1 - \alpha) S_{t-1}$$

From formula above we can see by increasing alpha we give more weight to the current real data.
For big capitalization company I will choose alpha closer to one because I trust the current observation. Big company tend to have lower volatility and higher liquidity and thus make it a good reflection of underlying true value. Therefore, we don't need much smoothing to determine the trend (giving too many weight to past observation will obscure the most updated trend).
For small cap company I will use smaller alpha, giving less weight to current observation and more weight to previous values. Smaller company is very volatile so we can't really trust the current observation to determine the trend.

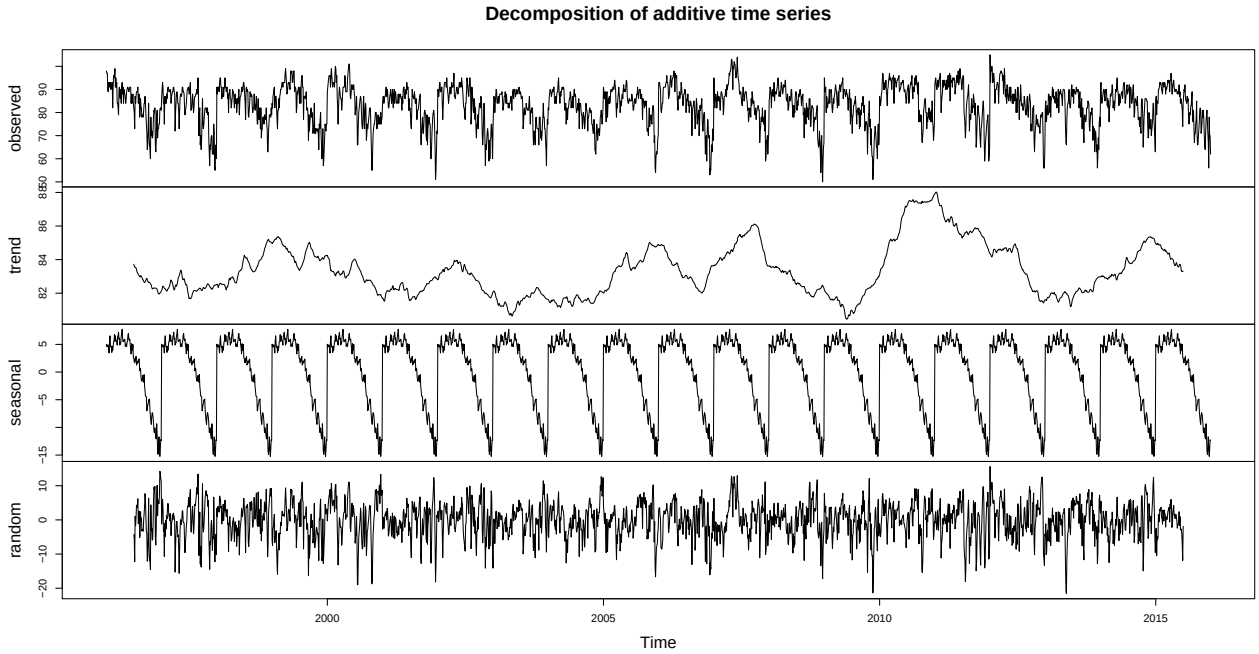
Question 7.2

Using the 20 years of daily high temperature data for Atlanta (July through October) from Question 6.2 (file temps.txt), build and use an exponential smoothing model to help make a judgment of whether the unofficial end of summer has gotten later over the 20 years. (Part of the point of this assignment is for you to think about how you might use exponential smoothing to answer this question. Feel free to combine it with other models if you'd like to. There's certainly more than one reasonable approach.).
Note: in R, you can use either HoltWinters (simpler to use) or the smooth package's es function (harder to use, but more general). If you use es, the Holt-Winters model uses model="AAM" in the function call (the first and second constants are used "A"dditively, and the third (seasonality) is used "M"ultiplicatively; the documentation doesn't make that clear).

Methodology to solve this problem:
We set the null and alternative hypothesis: H_0 : Unofficial end of summer has NOT gotten later H_a : Unofficial end of summer has gotten later

We check

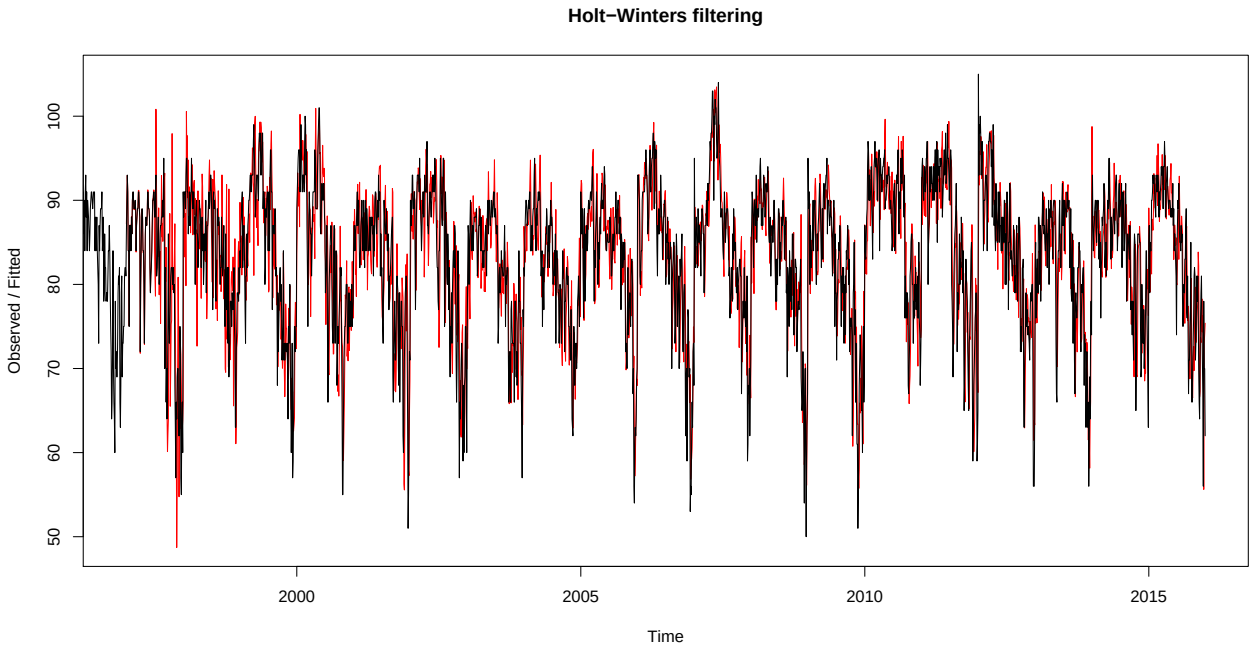
```
rm(list=ls())
temps = read.table('temps.txt', header=TRUE)
temp_ts = ts(as.vector(unlist(temps[,-1])),
             start = 1996, frequency = nrow(temps))
# plot(temp_ts)
plot(decompose(temp_ts))
```




```
temp_hw = HoltWinters(temp_ts)
summary(temp_hw)
```

```
##              Length Class  Mode
## fitted       9348   mts    numeric
## x            2460   ts     numeric
## alpha           1   -none- numeric
## beta            1   -none- numeric
## gamma           1   -none- numeric
## coefficients  125   -none- numeric
## seasonal        1   -none- character
## SSE             1   -none- numeric
## call            2   -none- call
```

```
plot(temp_hw)
```


Holt−Winters filtering
Observed / Fitted
50
60
70
80
90
100
2000
2005
2010
2015
Time

Question 8.1

Describe a situation or problem from your job, everyday life, current events, etc., for which a linear regression model would be appropriate. List some (up to 5) predictors that you might use.

Answer

Analytics can be used to help increase productivity of your banana farm. Using linear regresion model, we can know what type of stimulation we need to give to the plant to make it produce more bananas. Target variable (y): number of bananas produced Predictor: 1. Temperature 2. Humidity 3. Fertilizer 4. Water

Question 8.2

Using crime data from http://www.statsci.org/data/general/uscrime.txt (file uscrime.txt, description at http://www.statsci.org/data/general/uscrime.html), use regression (a useful R function is lm or glm) to predict the observed crime rate in a city with the following data: M = 14.0 So = 0 Ed = 10.0 Po1 = 12.0 Po2 = 15.5 LF = 0.640 M.F = 94.0 Pop = 150 NW = 1.1 U1 = 0.120 U2 = 3.6 Wealth = 3200 Ineq = 20.1 Prob = 0.04 Time = 39.0 Show your model (factors used and their coefficients), the software output, and the quality of fit.
