ISYE 6501 HW4

February 2021

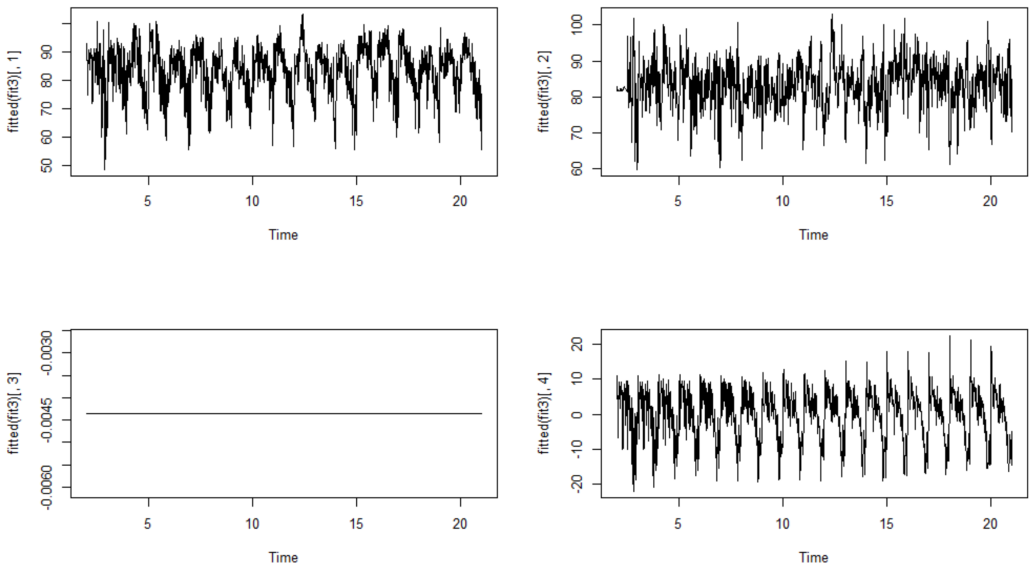
1 Question 7.1

Exponential Smoothing has been used a lot in business, economy, and financial field. It could also be used to forecast weather. It needs some data, like the set provided for the next question - Atlanta temperature. It also needs frequency which represent cycle. For weather, I would expect alpha to be closer to 1 since it's not so random (next day weather is related to current weather and the randomness is somewhat under control, despite sudden change of weather which occasionally would happen).

2 Question 7.2

First, I use the HoltWinters function in R to get Exponential Smoothed data. Second, I used Excel to perform CUSUM method to find unofficial end of summer dates for each year then compare the dates.I choose to use "additive" for seasonal factor in HoltWinters function because the temperatures are absolute terms. The frequency for this series is 123 (there are 123 days in July - October). When run the algorithm, it takes the first cycle - the 123 days in 1996 to "learn" the pattern and start predicting on 1997 temperatures. Thus, the total number of data predicted is 123 less than the original data.

When reviewing the predicted data, I plotted the predicted value xhat (top left), level (top right), trend (bottom left), and seasonality (bottom right) separately (see figure below). The predicted value xhat is combination of the other three elements. The seasonality shows a clear pattern, and since we are analyzing the end of summer day (temperature dropping), seasonality is giving valuable information here. I will use the seasonality value to do CUSUM to see end of summer days.



Import the seasonality data into Excel then apply CUSUM method to determine a changing in season pattern, this is equivalent to the changing in temperature but without much noise. Again, like what we did in previous homework, use only July data for mean and standard deviation since July is for sure summer. Then use sd to find C and T value, I have used the standared C is half sd and T is 5 times sd. I came up with information showing that since 2006 the end of summer day has gotten later each year, but just for a few days. From my analysis, it does show a trend of summer getting longer and end of summer getting later, but I would need more data to make a firm judgement.

DAY	St 2006	St 2007	St 2008	St 2009	St 2010	St 2011	St 2012	St 2013	St 2014	St 2015
29-Jul	0	0	0	0	0	0	0	0	0	1.3748
30-Jul	0	0	0	0	0	0	0	0	0	0
31-Jul	0	0	0	0	0	0	0	0.214845	1.847583	1.01641
1-Aug	2.61313	1.506609	0.501672	0.100909	0	0	0	0	0.097416	0.624077
2-Aug	6.874332	5.175767	3.421187	1.796964	2.273825	4.16107	2.366067	0.340227	0	0
3-Aug	10.95109	8.799386	6.42576	4.640335	3.62687	4.748414	2.92805	1.583024	1.058885	0
4-Aug	13.27392	11.2603	8.336706	6.787958	5.032374	5.341784	6.090797	5.834695	4.271511	1.927359
5-Aug	14.23331	13.11079	10.39077	8.739875	6.813415	7.285392	6.327833	6.351617	5.734739	3.567769
6-Aug	15.81641	16.47824	13.13136	10.93663	9.574932	9.988186	8.964739	9.348787	8.611428	5.600433
7-Aug	19.31202	18.61146	15.36242	13.99808	12.26035	13.18717	12.07431	12.83045	12.45722	8.350175
8-Aug	21.97478	21.33262	17.61762	17.43641	15.31886	16.01102	14.77351	13.97504	14.93949	11.83913
9-Aug	25.79592	25.27202	20.55473	20.90161	18.47707	18.95062	18.31966	17.58854	17.29693	14.80655
10-Aug	33.38939	32.02695	26.42263	26.38756	23.47025	23.57492	21.98368	21.38874	20.50412	18.70641
11-Aug	38.75528	39.40654	34.05353	32.94005	29.24553	28.94159	27.11743	26.60748	25.30549	24.07082
12-Aug	43.95172	46.30048	43.30215	43.02269	40.26231	38.8037	35.77481	33.83868	32.05307	29.5688
13-Aug	48.30093	52.17178	48.11933	49.12071	46.7129	45.71196	42.93827	40.12636	38.67243	36.66943

2.1 My R code is:

```
library(forcats)
df <- read.table("temps.txt", sep = '\t', stringsAsFactors = FALSE, header = TRUE)
dt <- as.vector(unlist(df[,2:21]))
tsdf <- ts(dt, start=1996, frequency = 123)
plot(decompose(tsdf))

# Triple exponential with seasonal "add" and trend
# reason for additive is temperature is absolute terms
fit3 <- HoltWinters(tsdf)
par(mfrow = c(2,2))
plot(fitted(fit3)[,1])
plot(fitted(fit3)[,2])
plot(fitted(fit3)[,3])
plot(fitted(fit3)[,4])
head(fitted(fit3))
# write data into csv, 123 data points are missingfilled with original data
table3 <- matrix(fitted(fit3)[,4], nrow=123)
head(table3)
write.csv(data.frame(table3), "xhat.csv")
```