WK4_HW4

Anon Skywalker

9/15/2020

Quesion 7.2

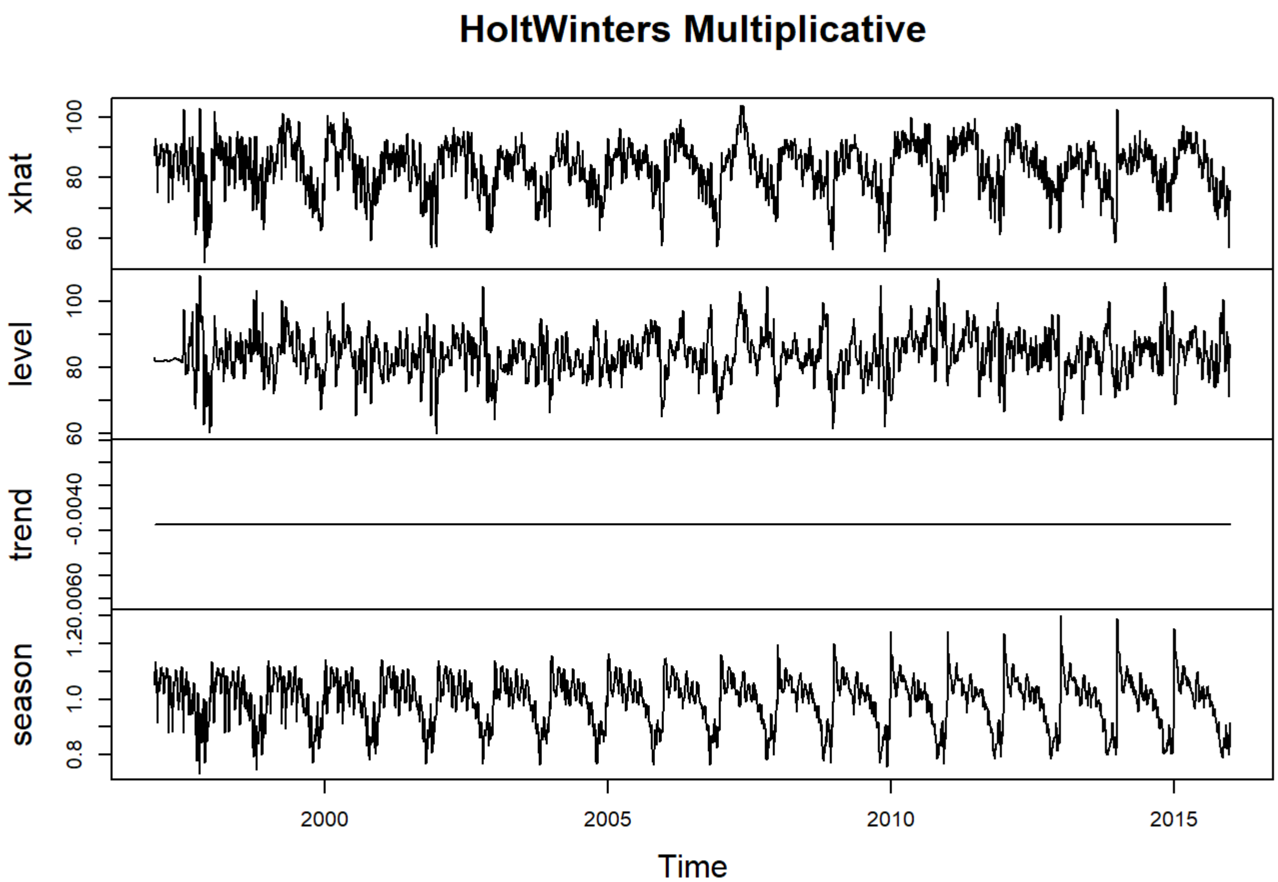
In the medical field, one area that could benefit from exponential smoothing is with ER room volumes. For certain days, aka weekends or holidays, the volume increases greatly and can throw off the average number of patients to expect during a given normal weekday. By applying HoltWinters to a set of data derived from several years worth of ER patient volumes from a given hospital, we could better understand what volumes we should expect during a given day and how many personnel we should have on site to handle it. I would expect the value of a to be closer to 0 due to how random ER volumes can be.

Quesion 7.2 Setup and Multiplicative Choice

To get useful data for determining whether or not summer is ending later in Atlanta, my approach involves the use of HoltWinters from the forcast R package and a manual CUSUM approach that is presented at the end of the document as function Atlanta.Warming. To do this, I converted my data into a time series and then used this data with both the multiplicative and additive seasonal definitions for the HoltWinters function. I then created a function that would take these data sets and return a data frame excluding the warm-up year that holt winters reserves for predicting future data points. I decided to opt for the multiplicative due to the nature of the data dealing with temperatures. As this is not a constant type of increase and is random due to nature influencing fluctuations between days, multiplicative is a good method for this data set. At the end of this code segment, we can see 3 new data frames, one for each relevant component returned by the holt winters model (xhat, seasonality, trend) that should be considered when choosing the best data set for CUSUM to work with.

Q 7.2 Continued: I’ve been plotting this all along…

As trend is displayed in the decomposition to be constant, we will not want this to represent any fluctuations that may be happening over the summers. This leaves us with level and season. By plotting these out as well and using different year ranges to see what the general trend of the data has, I discovered that seasonality has a pretty consistant downwards decline when plotting a line through it while level seems to have a slight incline. As both provide some level of consistency, I opted for the one that had a more noticeable trend over the years which was seasonality. Therefore, we will use the seasonality data from our holtwinters model for the CUSUM approach to derive whether or not summers are getting later over the years.

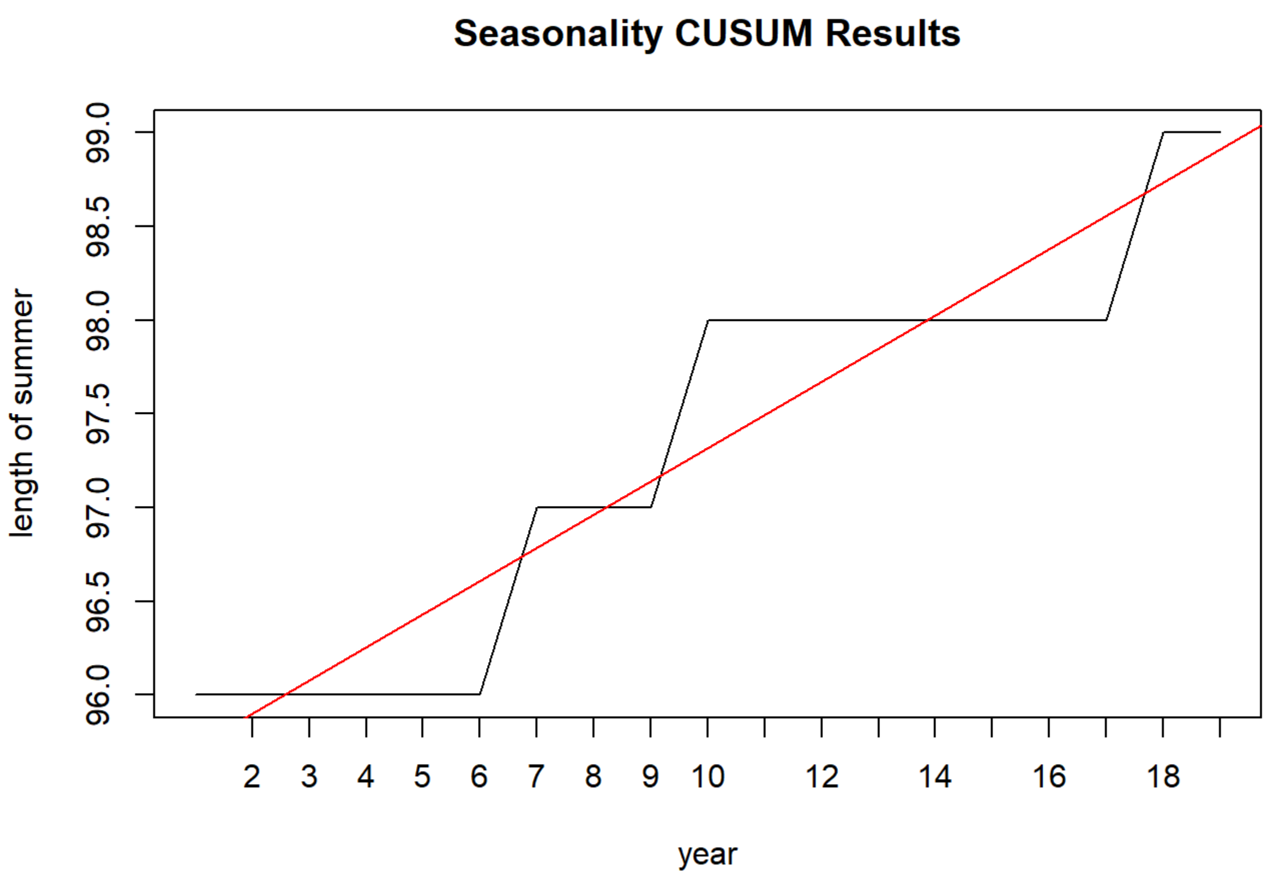


for the conclusion to this question, I will refer to the plots for the chosen data returned by the CUSUM function Atlanta.Warming to base my judgement. The data returned by the Atlanta.Warming function involves the day that summer ends, the temp that summer ended on, and most importantly, the length of each summer in the data set. I have plotted the summer lengths for each component of Holt Winters to determine my answer. Thanks to CUSUM, we can get a result that essentially tells us when the unofficial end of summer is based on the temperature changes. As the temperature gets colder, CUSUM will eventaully break through its set threshold.

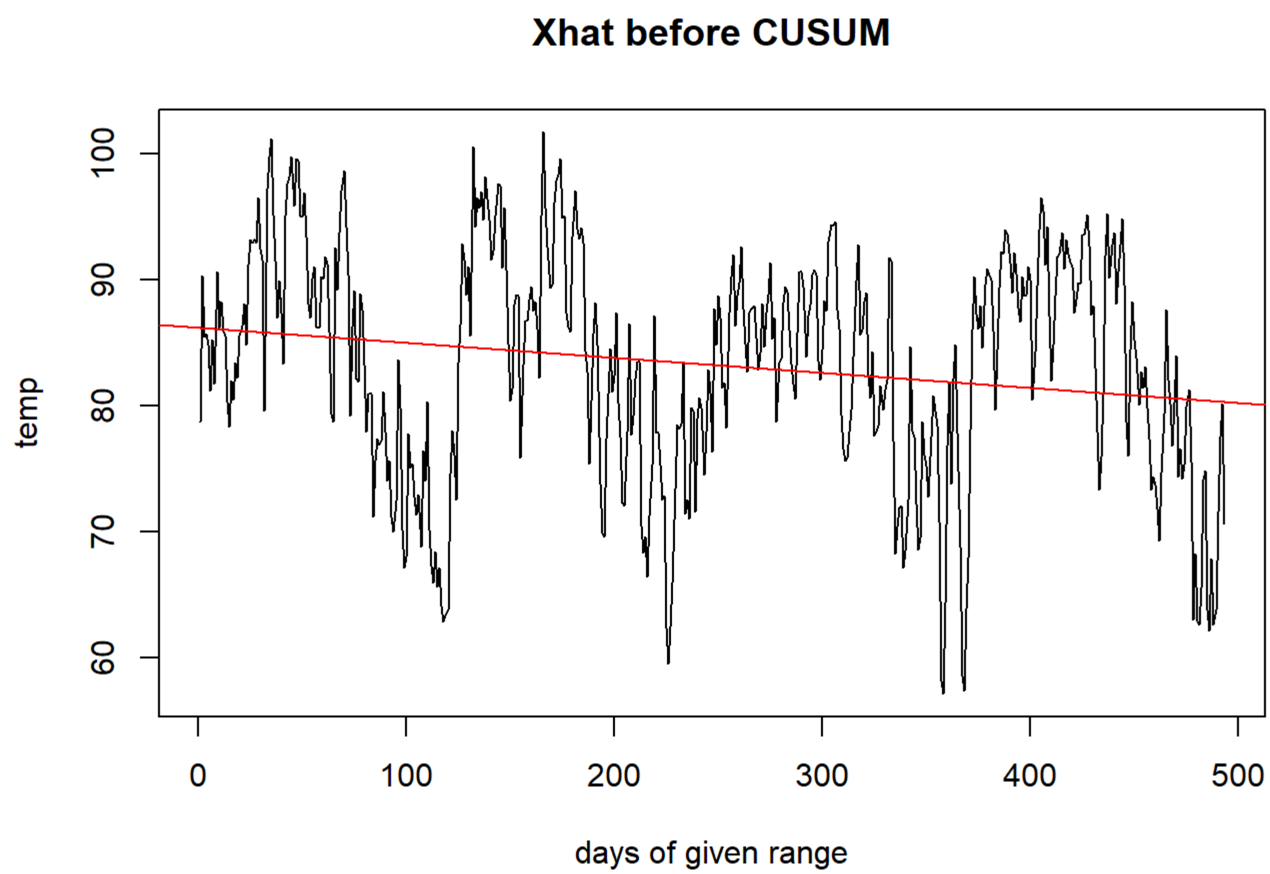
1. Xhat Plot: This plot shows a slight increase for the length of summer, but is diluted due to the noise that comes from being the raw fitted data created via the sum of level, trend and season value. This makes it a less reputable set of data we could be using to answer our question. By pulling out the factors causing noise, we can get a better representation of what is going on with the data.

2. *Seasonal Plot: This plot gives us a definite increase in the lengths of summer for Atlanta and is the answer to our question! This plot shoes that we have a clear increase in the length of summer as the years go on.The seasonality factor is the amount we need to increase/decrease by based on where in the cycle we are. Since we know that the higher baseline values we have, the higher the seasonality (Ct) in order to adjust to that higher baseline number. As summer gets warmer and winter colder, the plot of seasonality before CUSUM would indicate that we are getting higher seasonality values when summer is warm and smaller values as we approach fall and get into colder temperatures relative to the baseline. Therefore, after sending this data through my CUSUM function Atlanta.Warming, what we are looking for is the size of these peaks in the plot to indicate that the threshold of CUSUM is not being met until later in the summer. Since my CUSUM function records the day that summer ended for each year and calculates the length of each summer, we can easily plot this information and get our answer!*

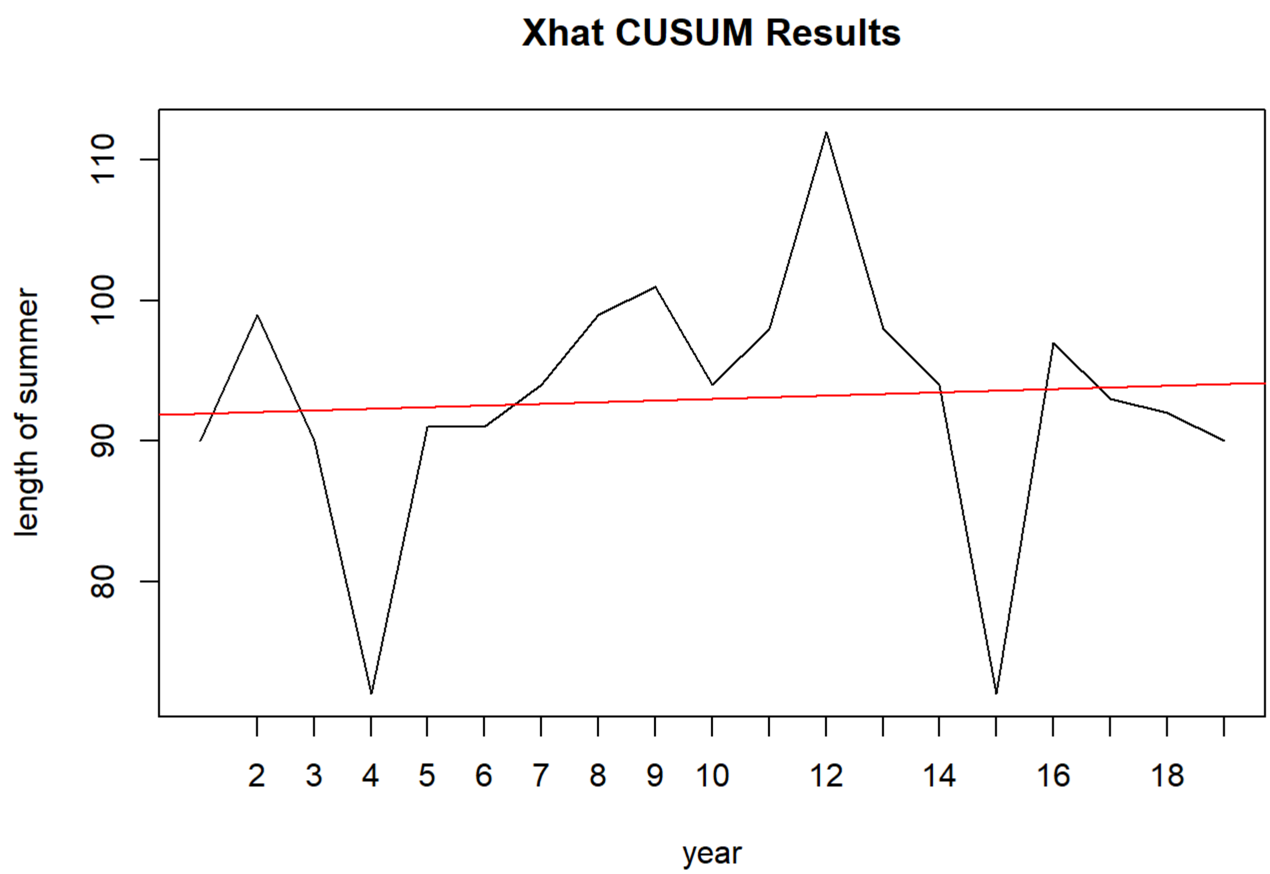
```
yrs1997_2001 <- 246:738
yrs2001_2006 <- 738:1353
yrs2006_2011 <- 1353:1968
yrs2011_2015 <- 1968:2337
#provide df, thresh, and c value. Runs through CUSUM!
season.result.df <- Atlanta.Warming(season.df, 1, 0)
level.result.df <- Atlanta.Warming(level.df, 40, 0)
xhat.result.df <- Atlanta.Warming(xhat.df, 30, 3)

#plot for seasonal CUSUM results (Length of summer)
x.seq <- c(2:20)
plot(season.result.df$summer.length, main = "Seasonality CUSUM Results", xlab = "year",
                                     ylab = "length of summer", xaxt = "n", type = "l")
axis(side = 1, at = x.seq)
abline(reg = lm(season.result.df$summer.length~time(season.result.df$summer.length)), col = "re
d")
```



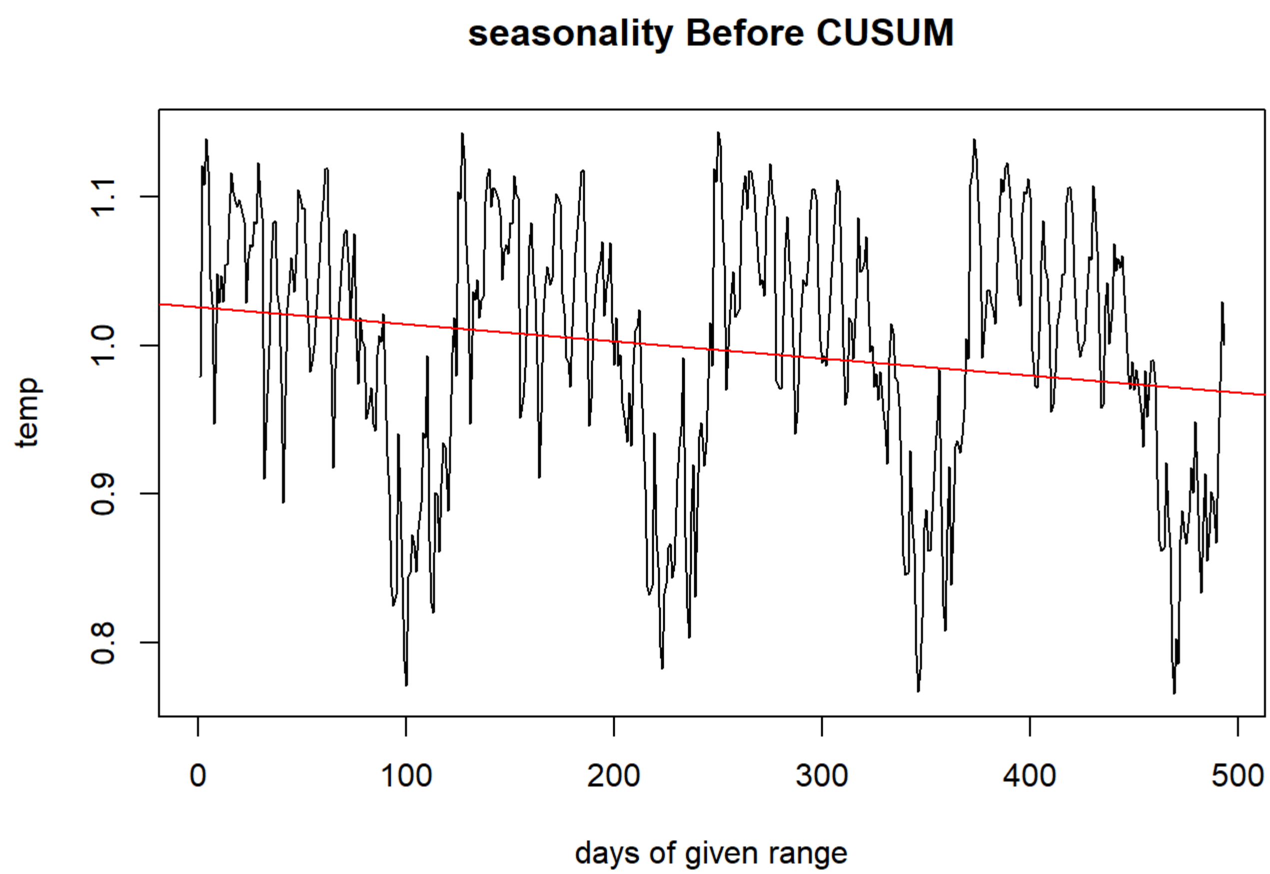

```
#plot for xhat before cusum
plot(fitted.temps_mult$fitted[yrs1997_2001,1], main = "Xhat before CUSUM", ylab = "temp", type =
"l", xlab = "days of given range")
abline(reg = lm(fitted.temps_mult$fitted[yrs1997_2001,1]~time(fitted.temps_mult$fitted[yrs1997_2
001,1])), col = "red")
```




```
#plot for xhat CUSUM results (Length of summer)
x.seq <- c(2:20)
plot(xhat.result.df$summer.length, main = "Xhat CUSUM Results" ,xlab = "year",
                              ylab = "length of summer", xaxt = "n", type = "l")
axis(side = 1, at = x.seq)
abline(reg = lm(xhat.result.df$summer.length~time(xhat.result.df$summer.length)), col = "red")
```




```
#plot for seasonality before CUSUM
#works for 1997-2001 and 2001-2006
plot(fitted.temps_mult$fitted[yrs1997_2001,4], main = "seasonality Before CUSUM", type = "l",  y
lab = "temp", xlab = "days of given range")
  abline(reg = lm(fitted.temps_mult$fitted[yrs1997_2001,4]~time(fitted.temps_mult$fitted[yrs1997
_2001,4])), col = "red")
```


seasonality Before CUSUM
temp
1.1
1.0
0.9
0.8
0
100
200
300
400
500
days of given range