Brute Force Moving Average (MA) Smoother

Australian Wine Edition <3

In workshop, we will cover the actual high-level ITSMR functions used to implement this kind of smoother. However, these functions are *black boxes* in the sense that without diving into further documentation, it's not clear what they are doing *mathematically*. We'll recreate what's inside the black box now.

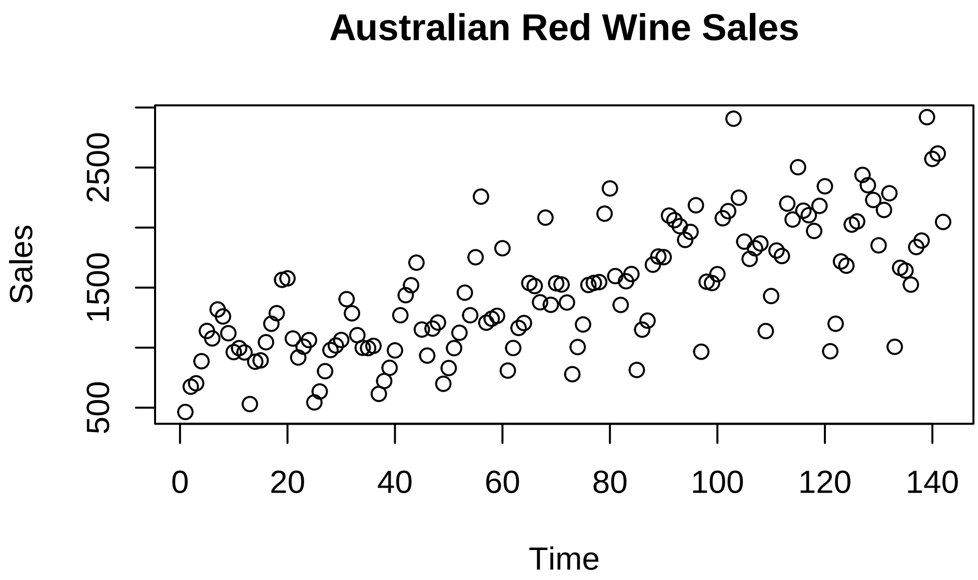
Agenda

1. Load the data
2. Review the theoretical definition of an MA smoother
3. Create our own MA smoothing code
4. Calculate and examine the output at a few select times.

1) Loading the Data

We'll use the Australian Wine Sales data. Recall that these data are dollar amounts, observed monthly from January 1980 to October 1991.

```r
# load the ITSMR package which contains the wine dataset
library(itsmr)

# store the data as a spreadsheet with 2 columns: time (mo) and sales ($)
wine.data <- data.frame(Time = 1:142, Sales = wine)

# plot sales vs. time - Make sure time is the first column in the dataframe
plot(wine.data, main = "Australian Red Wine Sales")
```



2) What's MA smoothing again?

At time t, we want to find $\hat{m}_t$: an estimate of the trend component m_t. The formula for an MA smoother with bandwidth $q \geq 0$ is

$$\hat{m}_t = \frac{1}{2q+1} \sum_{j=-q}^{q} x_{t-j}$$

3) Making our own MA smoother for a fixed t

```
q <- 5                  # choose bandwidth
t <- 10                 # choose timepoint to estimate (the t in m_t)
x <- wine.data$Sales    # this is our time series :)
```

Recall that the colon operator `a:b` concatenates (connects at the seams) integers ranging from `a` to `b`, inclusive. So to create the vector of $t - j$ values for $j \in \{-q, \ldots, q\}$, considering we've chosen $t = 10$ and $q = 5$,

$$
\begin{bmatrix} t-(-q) \\ t-(-q+1) \\ \vdots \\ t-(q-q)=t \\ \vdots \\ t-(q-1) \\ t-q \end{bmatrix} = \begin{bmatrix} 10-(-5) \\ 10-(-4) \\ \vdots \\ 10-0 \\ \vdots \\ 10-(4) \\ 10-(5) \end{bmatrix} = \begin{bmatrix} 15 \\ 14 \\ \vdots \\ 10 \\ \vdots \\ 6 \\ 5 \end{bmatrix}
$$

```
(t-(-q)):(t-q)
```

```
 [1] 15 14 13 12 11 10  9  8  7  6  5
```

So our indices are correct. The corresponding data is

```
x[ (t-(-q)):(t-q) ]
```

```
 [1]  894  883  530  960  996  963 1120 1260 1318 1077 1139
```

And the code representing the full RHS of the smoothing formula (part 2) is:

`(1/(2*q+1)) * sum(x[ (t-(-q)):(t-q) ])`

4) Calculating $\hat{m}_t$

For $t = 10$:

```
# Calculate and Store
hat.m10 <- (1/(2*q+1)) * sum(x[ (t-(-q)):(t-q) ])

# Print (you must print it if you want it to show up in the Quarto output)
hat.m10
```

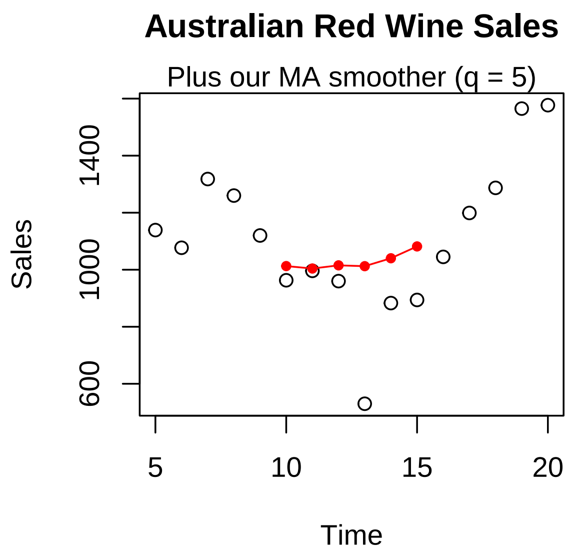
```
[1] 1012.727
```

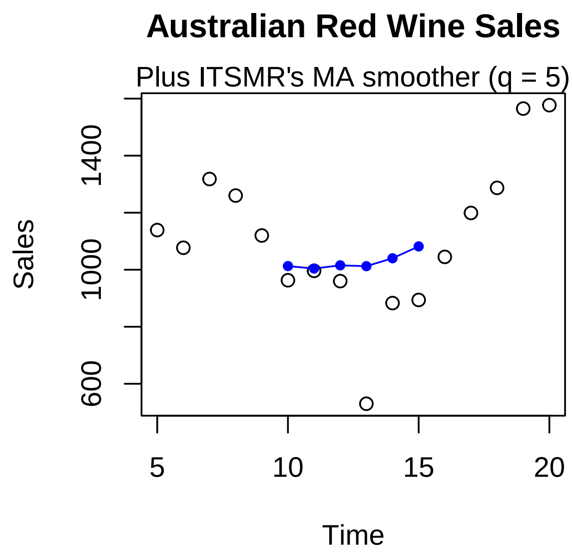
For $t = 11$ through to $t = 15$, we can compute $\hat{m}_t$ manually (obviously, in practice, we would write a function returning a vector, but we're doing it manually now for the sake of demonstration.)

```r
hat.m11 <- (1/(2*q+1)) * sum(x[ (11-(-q)):(11-q) ])
hat.m12 <- (1/(2*q+1)) * sum(x[ (12-(-q)):(12-q) ])
hat.m13 <- (1/(2*q+1)) * sum(x[ (13-(-q)):(13-q) ])
hat.m14 <- (1/(2*q+1)) * sum(x[ (14-(-q)):(14-q) ])
hat.m15 <- (1/(2*q+1)) * sum(x[ (15-(-q)):(15-q) ])

my.MA <- c(hat.m10, hat.m11, hat.m12, hat.m13, hat.m14, hat.m15)

check <- smooth.ma(wine,5)[10:15] # ITSMR function: should be the same as ours
```

Let's plot our results over the original data in that time-region, and compare it to the results from the ITSMR package's built-in function.

```r
par(mfrow = c(1,2), mar = c(4,4,4,1)) # puts 2 plots side-by-side

## --- PLOT 1
# Only shows region surrounding timepoints we've estimated
plot(wine.data[5:20,],
     ylim = range(x[5:20],my.MA),           # Makes it all fit in plot window
     main = "Australian Red Wine Sales")
mtext("Plus our MA smoother (q = 5)")       # Adds a subtitle
lines(10:15, my.MA,
      col = "red", type = "o", pch = 20)    # Adds our home-made MA


## --- PLOT 2
# Same plot as before, but with the ITSMR output highlighted
plot(wine.data[5:20,],
     ylim = range(x[5:20],my.MA),           # Makes it all fit in plot window
     main = "Australian Red Wine Sales")
mtext("Plus ITSMR's MA smoother (q = 5)")   # Adds a subtitle
lines(10:15, my.MA,
      col = "blue", type = "o", pch = 20)   # Adds ITSMR function's output
```





My questions for you

Does this look right?

Why do you think $\widehat{m}_t$ has this shape, in this region?

Are the results "successful?" What would that mean?