<http://home.ubalt.edu/ntsbarsh/Business-stat/stat-data/Forecast.htm>

## Smoothing

<https://en.wikipedia.org/wiki/Smoothing>

### Moving Average Smoothing

reference:

<http://machinelearningmastery.com/moving-average-smoothing-for-time-series-forecasting-python/>

more:

<https://en.wikipedia.org/wiki/Moving_average>

moving average-Rob J Hyndman.pdf

**moving average** (**rolling average** or **running average**)

It is also called a **moving mean (MM)**[[1]](https://en.wikipedia.org/wiki/Moving_average#cite_note-1) or **rolling mean** and is a type of [finite impulse response](https://en.wikipedia.org/wiki/Finite_impulse_response) filter.

A moving average is commonly used with [time series](https://en.wikipedia.org/wiki/Time_series) data to smooth out short-term fluctuations and highlight longer-term trends or cycles.

Moving average smoothing is a naive and effective technique, can be used for data preparation, feature engineering, and even directly for making predictions.

**(In feature engineering, this is so called aggregate features, it** **summarize the historical activity of each asset.)**

#### Assumption

Calculating a moving average of a time series makes some assumptions about your data.

It is assumed that both trend and seasonal components have been removed from your time series.

This means that your time series is stationary, or does not show obvious trends (long-term increasing or decreasing movement) or seasonality (consistent periodic structure).

There are many methods to remove trends and seasonality from a time series dataset when forecasting. Two good methods for each are to use the differencing method and to model the behavior and explicitly subtract it from the series.

#### Window size

A moving average requires that you specify a window size called the window width. This defines the number of raw observations used to calculate the moving average value.

The “moving” part in the moving average refers to the fact that the window defined by the window width is slid along the time series to calculate the average values in the new series.

* There is no definitive answer, but there is a trade-off to consider.
* Suppose the mean of the underlying process remains stable:   
  *If we include very few data points, then the moving average exhibits more variability than if we include a larger number of data points. In that sense, we get more stability from including more points.*
* Suppose there is an unanticipated change in the mean of the underlying process:  
  *If we include very few data points, our moving average will tend to track the changed process more closely than if we include a larger number of data points. In that case, we get more responsiveness from including fewer points.*

#### Central moving average

<https://www.otexts.org/fpp/6/2/>

##### odd order

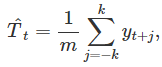
The value at time (t) is calculated as the average of raw observations at, before, and after time (t).

For example, a center moving average with a window of 3 would be calculated as:

center\_ma(t) = mean(obs(t-1), obs(t), obs(t+1))

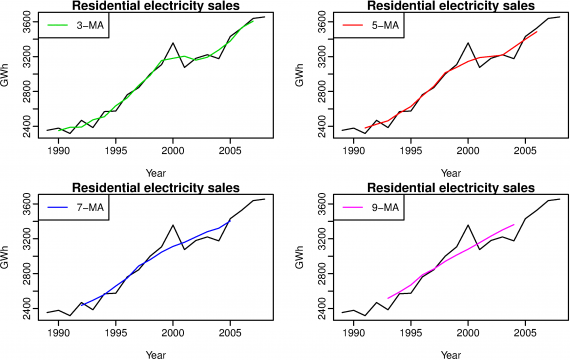
A center moving average can be used as a general method to remove trend and seasonal components from a time series, a method that we often cannot use when forecasting.

A moving average of order m can be written as



where m=2k+1. That is, the estimate of the trend-cycle at time t is obtained by averaging values of the time series within k periods of t. Observations that are nearby in time are also likely to be close in value, and the average eliminates some of the randomness in the data, leaving a smooth trend-cycle component. We call this an m-MA meaning a moving average of order m (通常为奇数).

The order of the moving average determines the smoothness of the trend-cycle estimate. In general, a larger order means a smoother curve. The following graph shows the effect of changing the order of the moving average for the residential electricity sales data.

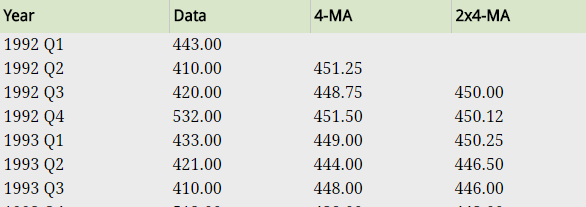


*Figure 6.8: Different moving averages applied to the residential electricity sales data.*

##### even order

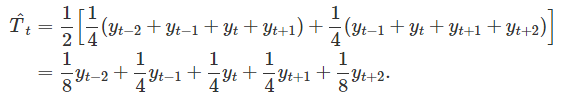
In order to make an even-order moving average symmetric, we can apply a moving average to a moving average.

For example, we might take a moving average of order 4, and then apply another moving average of order 2 to the results.



The notation “2×4-MA” in the last column means a 4-MA followed by a 2-MA. The values in the last column are obtained by taking a moving average of order 2 of the values in the previous column. For example, the first two values in the 4-MA column are 451.2=(443+410+420+532)/4 and 448.8=(410+420+532+433)/4. The first value in the 2×4-MA column is the average of these two: 450.0=(451.2+448.8)/2. When a 2-MA follows a moving average of even order (such as 4), it is called a "centered moving average of order 4".

we can write the 2×4-MA as follows:



()

It is now a weighted average of observations.

 Other combinations of moving averages are also possible. For example a 3×3-MA is often used, and consists of a moving average of order 3 followed by another moving average of order 3. **In general, an even order MA should be followed by an even order MA to make it symmetric. Similarly, an odd order MA should be followed by an odd order MA.**

##### Estimating the trend-cycle with seasonal data

The most common use of centered moving averages is in estimating the trend-cycle from seasonal data. Consider the 2×4-MA:



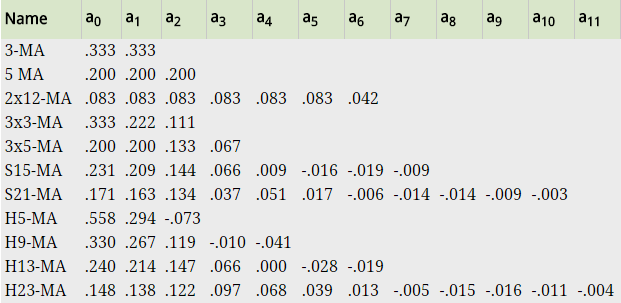
When applied to quarterly data, each quarter of the year is given equal weight as the first and last terms apply to the same quarter in consecutive years. Consequently, the seasonal variation will be averaged out and the resulting values of T^t will have little or no seasonal variation remaining. A similar effect would be obtained using a 2×8-MA or a 2×12-MA. **In general, a 2×m-MA is equivalent to a weighted moving average of order m+1 with all observations taking weight 1/m except for the first and last terms which take weights 1/(2m)**. So if the seasonal period is **even** and of order m, use a 2×m-MA to estimate the trend-cycle. If the seasonal period is **odd** and of order m, use a m-MA to estimate the trend cycle. **In particular, a 2×12-MA can be used to estimate the trend-cycle of monthly data and a 7-MA can be used to estimate the trend-cycle of daily data.** Other choices for the order of the MA will usually result in trend-cycle estimates being contaminated by the seasonality in the data.

#### [Weighted moving average](https://en.wikipedia.org/wiki/Moving_average#Weighted_moving_average)

For example, the 2x4-MA discussed above is equivalent to a weighted 5-MA with weights given by [18,14,14,14,18]. In general, a weighted m-MA can be written as



where k=(m−1)/2 and the weights are given by [a−k,…,ak]. It is important that the weights all sum to one and that they are symmetric so that aj=a−j. The simple m-MA is a special case where all the weights are equal to 1/m. A major advantage of weighted moving averages is that they yield a smoother estimate of the trend-cycle. Instead of observations entering and leaving the calculation at full weight, their weights are slowly increased and then slowly decreased resulting in a smoother curve. Some specific sets of weights are widely used. Some of these are given in table below.



*S = Spencer’s weighted mov­ing average  
H = Henderson’s weighted mov­ing average  
Com­monly used weights in weighted mov­ing averages*

These are all symmetric, so .

For example:

3-MA

, 



#### Trailing Moving Average

The value at time (t) is calculated as the average of the raw observations at and before the time (t).

For example, a trailing moving average with a window of 3 would be calculated as:

trail\_ma(t) = mean(obs(t-2), obs(t-1), obs(t))

Trailing moving average only uses historical observations and is used on time series forecasting.

#### [Simple moving average](https://en.wikipedia.org/wiki/Moving_average#Simple_moving_average)

#### [Cumulative moving average](https://en.wikipedia.org/wiki/Moving_average#Cumulative_moving_average)

#### [Exponential moving average](https://en.wikipedia.org/wiki/Moving_average#Exponential_moving_average)

#### Exponentially weighted moving average

### [Moving Median](https://en.wikipedia.org/wiki/Moving_average#Moving_median) Smoothing

<https://en.wikipedia.org/wiki/Moving_average#Moving_median>

### Exponential smoothing

<https://en.wikipedia.org/wiki/Exponential_smoothing>

<https://www.otexts.org/fpp/7>

## Filter

### Median\_filter

<https://en.wikipedia.org/wiki/Median_filter>

### Kalman Filter

<https://en.wikipedia.org/wiki/Kalman_filter>

<https://zhuanlan.zhihu.com/p/21294526>

## Decomposition

<https://www.otexts.org/fpp/6>

STL stands for “Seasonal Decomposition of Time Series by LOESS”

STL分解基于Loess，即局部加权回归散点平滑法，是1990年由密歇根大学的R. B. Cleveland教授以及AT&T Bell实验室的W. S. Cleveland等人提出来的一种对时间序列进行分解的方法。STL分解将时间序列分解成季节项、趋势项及残余项。

为了研究这种方法，我花了一天的时间仔细研读这篇论文，完成了17页的翻译稿（原文31页，不含讨论comment部分），基本对这种方法的原理有了大致的了解。本质上讲，这种方法是基于Loess，由内循环和外循环组成。其中，内循环包含了①去趋势、②周期子序列平滑、③对平滑后的周期子序列的低通滤波处理等6个步骤；而外循环主要作用则是引入了一个稳健性权重项，以控制数据中异常值产生的影响，这一项将会考虑到下一阶段内循环的临近权重中去。实际上，趋势分量和季节分量都是在内循环中得到的。循环完后，季节项将出现一定程度的毛刺现象，因为在内循环中平滑时是在每一个截口中进行的，因此，在按照时间序列重排后，就无法保证相邻时段的平滑了，为此，还需要进行季节项的后平滑，后平滑基于局部二次拟合，并且不再需要在loess中进行稳健性迭代。

R软件以及S-PLUS软件均提供了STL分解函数，但两个软件之间存在一些微小区别。具体可参考stl函数帮助。

下面直接给出对R软件中的co2数据进行STL分解的结果。图中最上端是原始的co2浓度随时间的变化，从1959年1月~1997年12月，共468个数据。下面依次为季节项、趋势项和残余项。从结果来看，分解的效果还是不错的。除了给出分解图，R还给出一些统计特征量，通过summary即可得到。

<http://blog.sina.com.cn/s/blog_5ffd41cf01012i5e.html>

more:

<https://anomaly.io/seasonal-trend-decomposition-in-r/>