Differencing and smoothing

Many time series models are built under the assumption of stationarity. However, time series data often present non-stationary features such as trends or seasonality. Practitioners may consider techniques for detrending, deseasonalizing and smoothing that can be applied to the observed data to obtain a new time series that is consistent with the stationarity assumption. We briefly discuss two methods that are commonly used in practice for detrending and smoothing.

Differencing

The first method is **differencing**, which is generally used to remove trends in time series data. The first difference of a time series is defined in terms of the so called difference operator denoted as D, that produces the transformation

$$Dy_t = y_t - y_{t-1}.$$

Higher order differences are obtained by successively applying the operator D. For example,

$$D^2 y_t = D(Dy_t) = D(y_t - y_{t-1}) = y_t - 2y_{t-1} + y_{t-2}.$$

Differencing can also be written in terms of the so called backshift operator B, with

$$By_t = y_{t-1},$$

so that $Dy_t = (1 - B)y_t$ and $D^d y_t = (1 - B)^d y_t$.

Smoothing

The second method we discuss is moving averages, which is commonly used to "smooth" a time series by removing certain features (e.g., seasonality) to highlight other features (e.g., trends). A moving average is a weighted average of the time series around a particular time t. In general, if we have data $y_{1:T}$, we could obtain a new time series such that

$$z_t = \sum_{j=-q}^{p} a_j y_{t+j},$$

for $t = (q+1) : (T-p)$, with $a_j \geq 0$ and $\sum_{j=-q}^{p} a_j = 1$. Often we work with moving averages in which $p = q$ (centered) and $a_j = a_{-j}$ (symmetric) for all j.

Assume we have periodic data with period d. Then, symmetric and centered moving averages can be used to remove such periodicity as follows:

- **If** $d = 2q$:

$$z_t = \frac{1}{d}\left(\frac{1}{2}y_{t-q} + y_{t-q+1} + \ldots + y_{t+q-1} + \frac{1}{2}y_{t+q}\right).$$

- **If** $d = 2q+1$:

$$z_t = \frac{1}{d}\left(y_{t-q} + y_{t-q+1} + \ldots + y_{t+q-1} + y_{t+q}\right).$$

Example: To remove seasonality in monthly data (i.e., seasonality with a period of $d = 12$ months), one can use a moving average with $p = q = 6$, $a_6 = a_{-6} = 1/24$, and $a_j = a_{-j} = 1/12$ for $j = 0, \ldots, 5$, resulting in

$$z_t = \frac{1}{24}y_{t-6} + \frac{1}{12}y_{t-5} + \ldots + \frac{1}{12}y_{t+5} + \frac{1}{24}y_{t+6}.$$