Lectures 6-7

Trend and Seasonality
Textbook Sections: 1.5

Departures from Stationarity

Before fitting a stationary model to the data, we need to make sure that the data does not have signs of non-stationarity. It is very common for the raw data to exhibit one or more of

- unequal variance,
- presence of a trend,
- presence of a seasonal component.

The **trend** is the slowly changing, smooth component of the data.

The **seasonal component**, or **seasonality** is the periodic component of the data. This is a pattern that repeaats at regular periods.

Transformations

If the data has unequal variance, then it is necessary to transform it. The hope is to find a transformation that makes the variance close to constant. If we find such a transformation, then we will fit a model to the transformed data. This model can be used to get forecasts and inferences for the transformed data. We can obtain forecasts and inferences of the original data by using the inverse transformation.

A common choice is power transformations. This is a family of transformations of the form $Y \to Y^\lambda$ for $\lambda \in \mathbb{R}$ with the convention that $Y \to ln(Y)$ for $\lambda = 0$.

Typically, several transformations are considered before one is selected. We could go through the full model fitting process for different transformations, and choose one that leads to the best results in the end. However, it is common to choose the transformation that leads to the most stable variance.

Model with Trend

Suppose the data has approximately constant variance, and exhibits a trend. We will use the model

$$X_t = m_t + Y_t,$$

where m_t is the trend, and $E(Y_t) = 0$ for $t = 1, \cdots, n$.

We will first compute an estimate of the trend, $\hat{m}_t$, then subtract it from the data to get an estimate of the detrended data $\hat{Y}_t = x_t - \hat{m}_t$.

Later on, we will analyze the sequence of $\hat{Y}_t$ using time series methods, and possibly fit another model to it. If we need to compute forecasts of X_{t+h} for $h > 0$, we can forecast m_{t+h} and Y_{t+h} separately, and combine them into forecasts of X_{t+h}.

Ultimately, estimating the trend comes down to *smoothing* the data. There are many techniques that could be employed here. Below are some common ones.

1. **Polynomial regression.**
   We can estimate the trend using polynomial regression. The degree of the polynomial can be chosen based on a model selection criterion such as AIC or BIC. A detailed description of polynomial regression is included in the supplementary materials for the course.

2. **LOESS.**
   LOESS, or LOWESS, stands for LOcally (Weighted) regrESSion. It is a popular method for smoothing the data. Rather than fitting one model based on the entire data set, this method employs local regressions based on subsets of the data set.

   At each point x_t, fit a local weighted regression model based on only the points $x_{t-q}, x_{t-q+1}, \cdots, x_t, x_{t+1}, \cdots,$ for some positive integer $q < n/2$. This model is usually linear or quadratic, and has weights giving the points closer to x_t more importance. The fitted value $\hat{x}_t$ is obtained using this model. We use these fitted values $\hat{x}_t$ as estimates of the trend.

   The size of the "window" of the data set on which the models are fit is called the "span," and is very important. The larger the span, the smoother the resulting fit. In practice, it is up to the user to decide on the value of the span. Different software packages use different formats to specify the span. It can be specified by

   a) the number of observaions on either side of x_t to include in the window (q in the above notation),

   b) the number of observations to include in the window ($2q + 1$) in above notation,

   c) the proportion of the data set to include in the window ($(2q + 1)/n$ in above notation). **This is the format of R's `loess()` function.**

   There are different ways of coping with points near the ends of the data sets, which do not have q neighboring points on one side.

   In this class, we will sometimes use this method for estimating trend, but we will not go deeply into the specifics.

3. **Moving Average.**
   Like with LOESS, the estimate of m_t at time t is computed based on the observations $x_{t-q}, x_{t-q+1}, \cdots, x_t, x_{t+1}$ for some positive integer $q < n/2$. However, instead of fitting a regression model, we simply take the mean of these values, so

$$\hat{m}_t = \sum_{j=-q}^{q} x_{t+j}/(2q+1).$$

   Note that in order to be able to obtain an estimate of m_t at time t, you will need to have q time points to the left and to the right of t. This is not possible when t is beween 1 and q or t is between $n-q+1$ and n. One can use the average of available observations to the left and to the right of t when $t = 1, .., q$ or when $t = n-q+1, ..., n$, but those estimates are not usually very good. For this reason, in this method, estimates of m_t are obtained for $t = q+1, ..., n-q$, and no estimates are usually provided for $m_1, , , , m_q$ and $m_{n-q+1}, ..., m_n$.

4. **Differencing.**
   Another technique important to time series analysis is **differencing**. This is a simple technique of taking first order differences $x_t - x_{t-1}$ for $t = 2, 3, \cdots, n$. Applying differencing to the sample, possibly several times, is an effective way to remove trend. We will revisit this technique in detail in the near future.

Model with Trend and Seasonality

Now suppose the data has approximately constant variance, and exhibits a trend and a seasonal component. We will use the model

$$X_t = m_t + s_t + Y_t,$$

where m_t is the trend, s_t is the seasonal component of period d, and $E(Y_t) = 0$ for $t = 1, \cdots, n$.

The seasonal component s_t has two conditions: $s_t = s_{t+d}$ and $\sum_{j=1}^{d} s_j = 0$.

This is called the **classical decomposition model**. It can be fit as follows. First, the trend is estimated using a two-sided moving average. Then the seasonal components are estimated by averaging the deviations from the estimated trend. Finally, the estimates are scaled to satisfy the condition that they add to 1. This is described in detail on pg.31 of the textbook.

We will use R's function `decompose()` to fit a classic decomposition model.

There are other techniques for estimating seasonal components.

1. Use harmonic function to detect and estimate periodic components.

2. Seasonalized differences can be used to remove the trend and seasonal components. We will briefly cover this later.

Plan of Action

1. Plot the data against time. Look for signs of trend, seasonality, unequal variance, outliers, or anything else noteworthy.
2. If there is unequal variance, consider transforming the data.
3. Estimate and remove trend and/or seasonal components.
4. Use the ACF plot or other methods to check if the data is i.i.d. If it is i.i.d., then there is no need to proceed further. Treat the trend and seasonality estimates as the fitted values, and this series as the residuals.
5. Fit a time series model to the detrended and deseasonalized data.
6. Check that the residuals are i.i.d., and hopefully normal.