Topics of the lab: Inputting a time series data into R from a text file, visual inspection, stationarity of a series, using autocorrelations and partial autocorrelations for preliminary model selection.

In this lab, our aim is to decide if a given time series may correspond to a MA(q) model of the form

$$Z_t - \mu = A_t - \sum_{i=1}^{q} \theta_i A_{t-1},$$

where A_t is a white noise series or to an AR(P) model of the form

$$\tilde{Z}_t = \sum_{i=1}^{p} \phi_i \tilde{Z}_{t-i} + A_t,$$

where $\tilde{Z}_t = Z_t - \mu$ and A_t is again a white noise series.

It is quite easy to check that theoretical autocorrelations of a MA(q) model are zero for lags larger than q, so we can use this fact for deciding if an MA(q) model (and of which order) may describe a given stationary series.

It turns out that a similar property holds for AR(p) models if we consider partial autocorrelations instead of autocorrelations. Intuitively, the k-th order partial autocorrelation describes, how well we can predict by Z_{t-k} the part of Z_t, which can not be predicted by the intermediate values $Z_{t-i},\ i = 1, \ldots, k-1$. If $k > p$, then the only part of Z_t, we can use the prediction $\mu + \sum_{i=1}^{p} \phi_i \tilde{Z}_{t-i}$, which leaves only A_t as unpredicted part. Since A_t is uncorrelated with all earlier values of Z_{t-i}, the partial autocorrelations between Z_{t-k} and Z_t is 0. So we can use partial autocorrelations (which can be computed by pacf()) to identify an order of a suitable AR(p) model.

If we choose a model for a time series, we should first decide if the series is (weakly second order) stationary an if it is, choose a suitable model with the least number of parameters. If the mean value of the series is clearly increasing or decreasing over all available time period, the series is clearly not stationary. If the variance around mean is clearly increasing (or decreasing) in time, it is also not stationary. In other cases it is not so easy to decide about stationarity by looking at a graph of the series and later in the course we are going to learn about different tests which help us to decide in such cases.

Exercises of the lab:

Ex.1 Read the data of 7 different time series into R from the file `lab5.csv`.

Ex.2 For each series, decide if it can correspond to a stationary time series model. For all stationary series, look at both autocorrelations and partial autocorrelations to identify the best MA(q) and AR(p) type models which can be used to describe the series. Which one would you use for a starting point for model fitting?