Topics of the lab: ARIMA models for non-periodic non-stationary series, unit root tests.

If the time series of interest is not stationary, then it happens quite often that the difference of consecutive values (or difference of differences and so on) is stationary and we may fit models to the differenced series. If we have a model for differences (or changes), then we can easily predict also future values of the original series. If an ARMA(p,q) model is fitted after taking d differences, the resulting model for the original series is an ARIMA(p,d,q) model.

Our procedure for selecting a suitable model is as follows:

1. Inspect visually the time series, it's autocorrelations and partial autocorrelations. If the mean level is clearly changing or the variance around the overall mean is changing in time, the series is not stationary. If the existence of a fixed mean is not clear (there may be an overall mean value, but it is crossed only a few times with long time intervals between crossings, autocorrelations are decaying towards 0 but quite slowly), we may use unit root tests (like Phillips-Perron or Augmented Dickey-Fuller test) in order to decide if a series should be considered stationary or not. NB! If you base your decisions on stationarity tests, you should read carefully the description of the test and also make sure you understand how the tests work.

2. If the series is not stationary in mean, we should think about the possibility of existence of a (linear, quadratic, exponential or some other) global trend. If the existence of a global trend is possible, remove the trend by fitting a trend curve to data and check, if the residuals may correspond to a stationary series.

3. If the existence of a global trend is not feasible or after removing a possible trend the series is still not stationary, we should consider a differenced series. If a differenced series is not stationary, we'll consider twice differenced series and so on. Important! Since each difference loses some information (the number of observations is reduced) it is important to avoid unnecessary differencing.

4. If after taking d differences the stationariness assumptions are not violated, we'll fit an ARMA(p,q) series for the differenced series (equivalently, we'll fitt ARIMA(p,d,q) model to the original series). It is important to understand that a non-zero mean of a d times differenced series corresponds to a polynomial global trend of order d for the original series, which is usually not reasonable. This is why in R (if one uses the command arima() from the `stats` library, which is loaded by default) it is assumed that the mean value of a differenced series is 0 and a global trend has to be included separately, if needed.

Exercises of the lab:

1. Let us look at time series data which is generated according to non-stationary ARIMA models. Look at the graphs of the series, first and second differences and autocorrelations/partial autocorrelations and at the result of unit root tests. The differences can be found by **diff**() command and tests can be applied by PP.test() and (after installing the additional package `#tseries#`) with the command adf.test().

   - Use the command `arima.sim(n=500,model=list(order=(0,1,1),ma=0))` to generate several versions of a random walk. Note that such series has no global/local trends that can be used in forecasting. Try to understand what is the difference between random changes of the mean value and changes corresponding to a global trend (see, for example the series generated by

     ```
     n <- 500
     trend <- 0.04*(1:n)
     z <- trend+arima.sim(n=n,model=list(ar=0.3))
     ```

     Look at the results of unit root tests. Do the tests give correct advice?

   - Generate several series corresponding to ARIMA(p,2,q) models. Should one use a stationarity test for the original series? Is the first difference clearly not stationary? What are the results of stationarity tests for the first difference?

2. Find a suitable ARIMA(p,d,q) model for time series in `lab7.txt`. For each series, predict next 4 values and write out the model.