Topics of the lab: Autocorrelation function, suitability of a model. Generating time series data according to Autoregressive (AR) and Moving Average (MA) models.

Autocorrelations of a time series measure how accurately it is possible to predict the next value of a time series by a linear function of an earlier value. Autocorrelations give useful information about what kind of model to use and about the presence of seasonal dependencies. Significantly non-zero autocorrelations of residuals (one step prediction errors) are clear indicators of unsuitability of a model - if the autocorrelations are non-zero, the assumptions of the model are not satisfied (hence information about prediction errors can not be trusted) and it is quite easy to improve the model by using the information in the residuals. So we should know how autocorrelations of different models behave and we always check the autocorrelations of residuals after fitting a model.

1. Let us look at autocorrelations of a series and of a model fitted to a series. Use the time series of Apple Inc closing prices

   Ex 1 Load the data saved in Lab1. Look visually at the relation betweeen a current value of a series and the value of the same seris k time moments ago. This can be done with a command of the form `plot(lag(series,-k),series)`.

   Ex 2 It is clear, that the last known value is important for predicting the next value. Look also at the graph of autocorrelations

   Ex 3 Fit Holt-Winters method for predicting the series. Look at the graph of autocorrelations. Can we consider the model to be suitable for predictions?

Autoregressive models describe situations when a finite number of last known time series values influence the expected value of the next observation. If the dependence is linear, we can write a corresponding model of the form

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + A_t,$$

which corresponds to AR model if the disturbances (innovations) A_t are uncorrelated and with constant variance.

Moving Average models describe a situation when a finite number of last disturbances influence the expected value of the next observation. Again, assuming the influence is linear and that the innovations are uncorrelated and with constant variance, we get models of the form

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

In this lab we consider generating data according to such models by using values of iid random variables instead of innovations A_t.

In addition to learning to generate time series data according to AR and MA models, we'll study the behaviour of the resulting series both directly (by looking at the graph of the series) and through autocorrelation function.

3. Let us start by AR models. We'll consider two ways to generate data according to AR models: the commands `filter` and `arima.sim`. In the examples we'll assume that, the innovations A_t are normally distributed with mean 0.

   (a) For generating values according to an AR model, we should use the following form of the `filter` command:

   ```
   filter(A, filter=c(phi_1,...,phi_p),method="recursive", init=c(z1,...,zp))
   ```

   where the vector A contains the values of the innovations A_t and $z1, \ldots, zp$ correspond to the first p values of the resulting series. We can generate A from the distribution $N(0, \sigma)$ by the command `A <-sigma*rnorm(n)`, where n is the length of the resulting vector.

   (b) The main form of the command `arima.sim` for generating data according to AR model is as follows:

   ```
   arima.sim(n=100,model=list(ar=c(phi_1,...,phi_p)),sd=sigma)
   ```

Here sigma denotes the standard deviation of the innovation series. By default, the command generates the values of the series A_t, but it is possible to use a given series by providing an option of the form innov=A, where A should contain n values.

Exercises:

Ex 4 Let $p = 1$, consider the AR(1) models with $\phi_1 = -1.5, -1, -0.8, 0, 0.8, 1, 1.5$. Assume that A_t are normally distributed with standard deviation $\sigma = 0.2$. Generate a vector A with 100 values of A_t and use them to generate an AR(1) time series for each value of ϕ_1. For each series, look at the graph and try to decide, if the series may have a constant mean and a constant standard deviation. Also look at the graphs of autocorrelations.

Ex 5 Consider the case $p = 2$ with $\phi_1 = 0.7$, $\phi_2 = 0.1$, let the standard deviation of the innovations A_t be $\sigma = 0.2$. Generate 200 values of of the innovation process and use them with the command filter to generate a time series Z. Look at the graph of the series (can it have a constant overall mean value?), find autocorrelations and also check, how much the last value of the series depends on the initial values (if you change one of the initial values by 1 unit, how much the final value changes).

Ex 6 Consider the AR(2) process with $\phi_1 = -0.5$, $\phi_2 = 0.66$. Repeat the steps of the previous exercise by using arima.sim.

4. In order to generate values according to MA(q) model $Z_t = A_t - \sum_{i=1}^{q} \theta_i A_{t-i}$ we can use the following commands:

   (a) filter with the options:

   ```
   filter(A, filter=c(1,-theta_1,...,-theta_q), sides=1).
   ```

   (b) arima.sim with the following parameters:

   ```
   arima.sim(n=100,model=list(ma=c(-theta_1,...,-theta_p)),sd=sigma).
   ```

Exercises:

Ex 7 Generate 200 values of a MA(1) process in the cases $\theta_1 = -1.25, -1, -0.8, 0.8, 1, 1.25$, check the existence of a fixed mean and variance, also look at the autocorrelations. How different are autocorrelations for $\theta_1 = -1.25$ and $\theta_1 = -0.8$?

Ex 8 Generate 500 values of a MA(2) process for various values of θ_1 and θ_2. Look at the graphs of autocorrelations. Can you spot anything common?