RTP Exercise Sheet

Series 5

Exercise 5.1

In this exercise, we look at the time series **prodn**, which is available in the package astsa. It contains monthly data about the Federal Reserve Board Production Index from 1948-1978, in total the time series contains data for n = 372 months.

a) Plot the time series. What kind of non-stationarity is evident?

b) How can the time series be made stationary?

c) Based on your considerations in b), what kind of model would you fit to the original time series **prodn**? Try different fits and choose your favourite.

   R-Hint:

```
arima(..., order = c(..., ..., ...), seasonal = c(..., ..., ...))
```

Exercise 5.2

In this exercise we again have a look at the sunspot data (Series 4). We figured out that AR(10) is a suitable model to describe the log transformed data.

```
library(fpp)
lsunspot100 <- window(log(sunspotarea), start = 1875, end = 1974)
fit.ar10 <- arima(lsunspot100, order = c(10, 0, 0))
```

a) For the AR(10) model, predict the next 100 observations of the log-transformed time series and plot them together with the log-transformed time series. Also add a line for the estimated global mean to the plot. What do you observe?

   R-Hint:

```
## prediction example with an AR(2) model
fit <- arima(data, order = c(2, 0, 0))
pred <- predict(fit, n.ahead = 100)
plot(data, xlim = c(1875, 2074))
abline(h = fit$coef["intercept"], lty = 2)
lines(pred$pred, col = "green")
```

b) Perform an out-of-sample evaluation, i.e. compare your prediction with the last 37 observed values of the time series. Plot the full log-transformed time series

(1875 - 2011) and add your prediction (1975 - 2011) as well as prediction intervals to the plot and comment on the plot. Also compute the mean squared forecasting error of your prediction:

$$\frac{1}{k}\sum_{i=1}^{k}(x_{n+i} - \hat{X}_{n+i;1:n})^2$$

Exercise 5.3

We want to compare methods for forecasting the airplane data: the forecasting using a SARIMA model and the forecasting using an STL-decomposition. For being able to compare the prediction methods with the real data, first read in the airplaine data and use only the observations from 1949 to 1956.

```
d.air <- AirPassengers
d.airshort <- window(d.air, end = c(1956, 12))
```

a) Fit an ARIMA/SARIMA Model for the shorter dataset d.airshort. Use transformations if suitable. Compute a prediction for the years 1957-1960 and plot it along with the prediction interval and the actual observations for this period.

   R-Hint:

   ```
   predict(..., n.ahead = ...)
   ```

b) Now do prediction for the years 1957-1960 as seen in the lecture with linear extraploation of the trend estimate, continuation of the seasonal effect and and ARMA(p,q) forecast for the stationary remainder. Plot the predicted timeseries (this time without prediction interval, why?) and compare them to the actual observations.

   R-Hint:

   ```
   ## example trend extraction
   fit <- stl(data, s.window = "periodic")
   trend <- fit$time.series[, 2]
   ```

c) Compare the different forecasts. Which of the methods seems to work best for the airplane data and why?

Disclaimer: Parts of the exercises are adopted from 'Applied Time Series Analysis' course at ETHZ by Marcel Dettling.