

RTP Exercise Sheet

Series 8

Exercise 8.1

As in exercise 7.1, we would like to use this exercise to simulate several time series by means of an ARMA model. Please perform the same steps as in exercise 5.1 for the following models:

- Why is it not possible to simulate the ARIMA(2,1,2) model with coefficients $\alpha_1 = 0.5, \alpha_2 = 0.5, \beta_1 = -0.4$ and $\beta_2 = 0.3$ with $d = 1$ with `arima.sim`?
- What is the equivalent ARMA model to the ARIMA(2,1,2) model in task c)?
Hint: Determine the equivalent ARMA process.

The innovation E_t shall follow a standard normal distribution $\mathcal{N}(0; 1)$ in every model.

Exercise 8.2

There is a study on the development of beluga whales that focusses on the nursing behaviour of mother and calf. During a total of 160 time periods (each lasting 6 hours) subsequent to birth, the following variables were observed for *Hudson*, a beluga calf. Zoologists use this data to ascertain the health of this young whale. A short description of the data is given in the following table.

PERIOD	Index of time period
BOUTS	Square root of the number of nursing bouts
LOCKONS	Square root of the number of lock-ons (docking attempts)
DAYNIGHT	Day (1, 8am - 8pm) or night (0, 8pm - 8am) indicator
NURSING	Square root of the number of seconds spent successfully nursing during the period

A nursing bout is defined as a successful nursing episode where milk was obtained. We would like to model the nursing time by means of the other variables. Count variables have already undergone a square root transformation to stabilize their variance (*first-aid-transformation*). You will find the data in the file *beluga.dat*. Load the data in the usual way and create a time series matrix:

```
d.beluga <- read.table("http://stat.ethz.ch/Teaching/Datasets/WBL/beluga.dat",
  header = TRUE)
d.beluga <- ts(d.beluga)
```

- a) Fit the model

$$\text{NURSING} = \beta_0 + \beta_1 \text{PERIOD} + \beta_2 \text{BOUTS} + \beta_3 \text{LOCKONS} + \beta_4 \text{DAYNIGHT}$$

using ordinary linear regression. Check the independence of the residuals. What conclusions can zoologists draw from this analysis?

- b) Due to the correlations involved, an AR(p) model should be assumed for the residuals. Determine the order p of this model, and estimate the parameters $\alpha_1, \dots, \alpha_p$. R-Hint:

```
r.burg <- ar(...)
```

- c) Estimate the regression coefficients and the AR parameters using Generalized Least Squares with Maximum Likelihood estimation. R-Hint:

```
library(nlme) # Load the package containing the procedure gls()
r.bel.gls <- gls(NURSING ~ BOUTS + LOCKONS + DAYNIGHT + PERIOD,
+ data = d.beluga, correlation = corARMA(form= ~ PERIOD,
+ p = r.burg$order, q = 0, fixed = FALSE), method = "ML")
summary(r.bel.gls)
d.resid <- ts(resid(r.bel.gls))
```

To ensure convergence of the algorithm, known estimates of the AR parameters can be passed to `corARMA()` as starting values using the optional argument `value`. In this particular case, this does not change the outcome. (`correlation = corARMA(..., value = r.burg$ar, ...)`)

- d) Optional: Simplify the model if possible.
- e) Optional: What transformation should you apply to obtain a linear model with independent errors? State it as a formula.

Hint: Cochrane-Orcutt Method.

- f) Optional: How would you perform this transformation (or these transformations) in R? Use the transformed time series to carry out another regression, and look at the correlation structure for the errors!

R-Hint: `lag()`.

Exercise 8.3

With the new material in the course we would like to return to Exercise 7.3.

- a) Choose a suitable model that fits the data. Does your model fit? Analyze the residuals and comment on your decision.

R-Hint:

```
r.varve <- arima(log(d.varve), order = c(..., ..., ...))
```

- b) Write down the model you chose in a) with its estimated coefficients.

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