732A62 Time Series Analysis
Division of Statistics and Machine Learning
Department of Computer and Information Science

Computer lab 2

Instructions

- The lab is assumed to be done in groups.
- Create a report to the lab solutions in PDF.
- Be concise and do not include unnecessary printouts and figures produced by the software and not required in the assignments.
- **Include all your codes as an appendix into your report.**
- A typical lab report should 2-4 pages of text plus some amount of figures plus appendix with codes.
- The group lab report should be submitted via LISAM before the deadline.
- Use 12345 as a random seed everywhere where the result of the simulation differs with the run unless stated otherwise.

Assignment 1. Computations with simulated data

a) Generate 1000 observations from AR(3) process with $\phi_1 = 0.8, \phi_2 = -0.2, \phi_3 = 0.1$. Use these data and the definition of PACF to compute ϕ_{33} from the sample, i.e. write your own code that performs linear regressions on necessarily lagged variables and then computes an appropriate correlation. Compare the result with the output of function pacf() and with the theoretical value of ϕ_{33}

b) Simulate an AR(2) series with $\phi_1 = 0.8, \phi_2 = 0.1$ and $n = 100$. Compute the estimated parameters and their standand errors by using three methods: method of moments (Yule-Walker equations), conditional least squares and maximum likelihood (ML) and compare their results to the true values. Which method does seem to give the best result? Does theoretical value for ϕ_2 fall within confidence interval for ML estimate?

c) Generate 200 observations of a seasonal $ARIMA(0,0,1)\times(0,0,1)_{12}$ model with coefficients $\Theta = 0.6$ and $\theta = 0.3$ by using arima.sim(). Plot sample ACF and PACF and also theoretical ACF and PACF. Which patterns can you see at the theoretical ACF and PACF? Are they repeated at the sample ACF and PACF?

d) Generate 200 observations of a seasonal $ARIMA(0,0,1)\times(0,0,1)_{12}$ model with coefficients $\Theta = 0.6$ and $\theta = 0.3$ by using arima.sim(). Fit $ARIMA(0,0,1)\times(0,0,1)_{12}$ model to the data, compute forecasts and a prediction band 30 points ahead and plot the original data and the forecast with the prediction band. Fit the same data with function *gausspr* from package **kernlab** (use default settings). Plot the original data and predicted data from $t = 1$ to $t = 230$. Compare the two plots and make conclusions.

e) Generate 50 observations from ARMA(1,1) process with $\phi = 0.7, \theta = 0.5$. Use first 40 values to fit an ARMA(1,1) model with $\mu = 0$. Plot the data, the 95%

prediction band and plot also the true 10 values that you initially dropped. How many of them are outside the prediction band? How can this be interpreted?

Assignment 2. ACF and PACF diagnostics.

1. For data series *chicken* in package **astsa** (denote it by x_t), plot 4 following graphs up to 40 lags: ACF(x_t), PACF(x_t), ACF(∇x_t), PACF(∇x_t) (group them in one graph). Which $ARIMA(p,d,q)$ or $ARIMA(p,d,q)\times(P,D,Q)_s$ models can be suggested based on this information only? Motivate your choice.
2. Repeat step 1 for the following datasets: *so2, EQcount, HCT* in package **astsa.**

Assignment 3. ARIMA modeling cycle.

In this assignment, you are assumed to apply a complete ARIMA modeling cycle starting from visualization and detrending and ending up with a forecasting.

1. Find a suitable $ARIMA(p,d,q)$ model for the data set *oil* present in the library **astsa.** Your modeling should include the following steps in an appropriate order: visualization, unit root test, detrending by differencing (if necessary), transformations (if necessary), ACF and PACF plots when needed, EACF analysis, Q-Q plots, Box-Ljung test, ARIMA fit analysis, control of the parameter redundancy in the fitted model. When performing these steps, always have 2 tentative models at hand and select one of them in the end. Validate your choice by AIC and BIC and write down the equation of the selected model. Finally, perform forecasting of the model 20 observations ahead and provide a suitable plot showing the forecast and its uncertainty.
2. Find a suitable $ARIMA(p,d,q)\times(P,D,Q)_s$ model for the data set *unemp* present in the library **astsa.** Your modeling should include the following steps in an appropriate order: visualization, detrending by differencing (if necessary), transformations (if necessary), ACF and PACF plots when needed, EACF analysis, Q-Q plots, Box-Ljung test, ARIMA fit analysis, control of the parameter redundancy in the fitted model. When performing these steps, always have 2 tentative models at hand and select one of them in the end. Validate your choice by AIC and BIC and write down the equation of the selected model (write in the backshift operator notation without expanding the brackets). Finally, perform forecasting of the model 20 observations ahead and provide a suitable plot showing the forecast and its uncertainty.