FEB23001(X) Assignment 1

Annika Camehl

Time Series Analysis

Instructions

- This assignment should be made with groups of three students. It is not allowed to cooperate with other groups, or to copy their results.
- The results of this assignment should be written down in a solution report consisting of max. 5 pages (not including a possible appendix which can contain part of the tables and graphs). Answer each question in a separate section (clearly indicating the number of the question).
- Do not include your code!
- Reports need to be submitted via Canvas Assignments as a PDF file with the name "`A1-XXX-YYY-ZZZ-.pdf`", where `XXX`, `YYY` and `ZZZ` are the student numbers of the three group members.
- Motivate your answers! For example, answers like Yes or No will not be rewarded any points.
- Some questions may not be specified up to the last detail and require you to make some choices yourself. Explain and motivate your choices in your report.
- It is recommended to use RStudio (Desktop) for programming.

Questions

The CSV file `CPILFESL` contains monthly observations on US Consumer Price Index (known as core CPI, aggregate of prices paid by urban consumers for a typical basket of goods, excluding food and energy) over the period January 1957 until February 2023 (obtained from `https://fred.stlouisfed.org/series/CPILFESL`). The index is seasonally adjusted. We denote the series by y_t.

Before you start, install the following packages with

```
install.packages("readr")   # for data import
install.packages("ggplot2") # for plotting
install.packages("tseries") # for Jarque-Bera test
install.packages("forecast")# for prediction
```

and load them

```
library(readr)
library(ggplot2)
library(forecast)
library(tseries)
```

Use the function `Arima` of the package `forecast` to estimate AR, MA or ARMA models.

1. (4 points) Import the time series and plot the data. Do you observe any of the key features discussed in Chapter 2 by analyzing the graphs? Discuss all key features briefly.

2. (1 point) Check whether a linear or quadratic time trend is present in the data by estimating

$$y_t = \alpha + \beta t + u_t, \quad t = 1, \ldots, T$$
$$y_t = \alpha + \beta t^2 + u_t$$

with ordinary least squares and discussing your results.

3. (2 points) Fit an AR(1) model with a linear trend to the data. Discuss what conclusions you can draw from the estimated coefficients.

4. (2 points) Fit a random walk

$$y_t = y_{t-1} + u_t, \quad t = 1, \ldots, T$$

and a random walk with linear trend

$$y_t = y_{t-1} + \beta t + u_t, \quad t = 1, \ldots, T$$

to the data. Which of the two models fits the data better based on the Akaike Information Criterion (AIC)? What is the difference between the random walk with linear trend and the AR(1) model with linear trend estimated before?

5. (2 points) Is the time series normally distributed? Perform a Jarque-Bera test. Clearly state the null hypothesis of the test, the value of the test statistic, p-value, whether you reject or not and what conclusions you can draw.

6. (1 point) Create a time series of monthly growth rates of CPI by taking the first differences of the log series and multiplying by 100 (to put the growth rates in terms of percentage points). Are these growth rates normally distributed?

7. (3 points) Compute the first 48 empirical autocorrelations and partial autocorrelations for the monthly growth rates of CPI. What conclusions can you draw based on the ACF and PACF regarding selecting an MA, AR or ARMA model?

8. (2 points) Estimate AR(p) and MA(q) models for $p = 1, \ldots, 24$ and $q = 1, \ldots, 24$ using the calculated monthly growth rates of CPI. Include a constant in the models. Select which model you would chose based on the values of the Akaike Information Criterion (AIC).

9. (1 point) Fit an ARMA(p, q) model to the growth rates of CPI. Chose the best lag length based on AIC with maximum lag length of 24. Include a constant in the model. Compare your findings to the before selected AR and MA models.

10. (3 points) Estimate an AR(9) model (including a constant). Calculate fitted values based on your estimated coefficients of the AR(9) process and compare these to the true data by plotting them in one graph. What do you conclude about the fit of your model?

11. (3 points) Inspect the properties of the time series residuals of the AR(9) model. Are the residuals following a White Noise process? Are they autocorrelated? Are they homoskedastic? You can address these questions by calculating mean and standard deviation, calculating the autocorrelation function based on the residuals, and plotting the residuals. What conclusions do you draw with respect to the adequacy of the estimated AR(9) model?

12. (6 points) Evaluate the forecast performance of the AR(9) model. Split the sample into an estimation period and a forecasting period from January 2015 to February 2023. Produce one-step, two-step, three-step and four-step ahead forecasts for the period January 2015 to February 2023 using a moving window. Thus, to obtain a one-step ahead forecast for January 2015 use the observations from January 1957 to December 2014, to obtain a one-step ahead forecast for February 2015 use the observations from

February 1957 to January 2015 and so on. Compare you results to forecasts based on an AR(1). Include a constant in all your AR models. Compare the forecasting accuracy of the AR(9) model based on mean squared forecast errors for each forecasting horizon relative to forecasts based on an AR(1). What do you conclude about the prediction accuracy of the AR(9) model?