Data Science Winter School

Time Series Analysis and Forecasting

Assignment

We consider the following data sets:

- **Data 1**: retail sales of beer, wine and liquor in the U.S. Monthly data from January 1992 till November 2017 (Source: Federal Reserve Bank of St. Louis, file: `BeerWineUS.csv`).
- **Data 2**: real (inflation corrected) interest rate for bank deposits with investment durations between one and two years (Source: Deutsche Bundesbank / German Central Bank, file: `interestrate.csv`).
- **Data 3**: interest rates for private mortgages with duration between 5 and 10 years (Source: ECB, file: `credit.csv`)

1. Regression-type modelling and forecasting with `BeerWineUS.csv`
   (a) Import the 1st data set to R. Create a `ts`-object. Keep the last 5 years of data as test data set.
   (b) Plot the data and choose an appropriate order of time polynomial for modelling. Run a simple regression on the time polynomial and check the autocorrelation of the residuals (also using Durbin-Watson). What can be concluded?
   (c) Since the time series is rather short to define a dummy for every month, it is reasonable to define dummies for longer periods. Use `ggmonthplot` or `ggseasonplot` to decide if it is possible to assign dummies to quarters/specific months, etc. Run a corresponding regression and check the residuals as above.
   (d) With the last model compute the (interval) forecasts for the test data set and visualize the results.
2. Time series decomposition with `BeerWineUS.csv`
   (a) Try several moving average techniques to extract the trend (`ma`). Which orders/form of moving averages do provide the best results?
   (b) Apply the loess regression for trend extraction.

(c) For one of the above computed trends compute the seasonal and the irregular components. Visualize the decomposition.

(d) Check the ACF of the irregular component. Is there anything left in the time series?

3. Simple forecasting and forecasting with exponential smoothing with `BeerWineUS.csv`

   (a) Compute simple one-step-ahead forecasts for the test data set using naïve, absolute trend and relative trend methods. Compute the corresponding losses.

   (b) Compute forecasts using simple EWMA, Holt and Holt-Winters forecasts with the smoothing parameters estimated from the training data set. Compute the corresponding losses. Check the ACF of the forecast errors.

   (c) Compare the performance of the models using the three tests discussed in the lectures.

4. ARMA modelling with `interestrate.csv`. Leave the last year for forecasting.

   (a) Import the data and create a `zoo`-object.

   (b) Check the ACF and decide about the strength of the memory in the time series using Box-Ljung/Pierce tests.

   (c) Try MA(1), AR(1) and ARMA(1,1) processes and check the fit (ACF, AIC, etc.)

   (d) Try differencing and subsequent application of MA(1), AR(1) and ARMA(1,1). Check again the processes and check the fit (significance, ACF, AIC, etc.) Decide which model is the best one.

   (e) Try `autoarima` and compare the final model with the one you found in the previous step.

   (f) Compute the forecasts and forecast intervals using the final model.

5. Structural breaks with `credit.csv` (or with `interestrate.csv`).

   (a) Use `breakspoints` to find ex-post the structural breaks in the time series. Visualise the results and check the optimal number of breaks.

   (b) Compute moving (or expanding) window estimates for OLS parameters of the mortgage rate on the simple time variable. Plot the results and decide about the break.

   (c) Compute one-step-ahead forecast errors, the scaled recursive residuals (scaled forecast errors)) and plot them with the confidence intervals. Compare the results with the conclusion in the previous step.

   (d) Use the `sctrucchange` package to implement the CUSUM test.