Time Series Analysis and Forecasting 2019

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**: crude oil price per barrel; (Source: Federal Reserve Bank of St. Louis, file: `CrudeOil.csv`)

1. Trend extraction for non-seasonal time series of the crude oil price
   (a) Try several moving average techniques to extract the trend (`ma` in R). Which orders/form of moving averages do provide the best results? Use just figures of the trend and the difference $Y_t - T_t$.
   (b) Does it make sense to apply centered moving average ($2 \times k$ MA) smoothing? Explain and motivate your answer.
   (c) Since the time series exhibits a very irregular trend apply and visualize the local polynomial regression. In your particular implementation verify the functional form of the weights used for trend extraction.
   (d) Apply the B-spline approach to extract the trend. Explain precisely how many splines you use, the underlying time grid and the order of polynomials.
   (e) For every of the above approaches extract the irregular component (there is no seasonal component) and plot its autocorrelation. If there are significant lags is this good or bad for further analysis?
2. Simple forecasting and forecasting with exponential smoothing of the crude oil price. Begin the forecasting starting with the observation 201. Thus the out-of-sample performance of the below methods will be assessed using the forecasts for the periods 201 to 311.
   (a) By just looking at the time series identify the phases when the naive forecast can outperform the alternatives and intervals where exponential smoothing is potentially better. Motivate your decision.

(b) Split the time series into three phases using the observations [1;600]; [600;1400] and [1500;end]. For each phase do the following (use the first 200 observations to calibrate the model if needed):

   i. Compute simple one-step-ahead forecasts for the test data set using naïve, absolute trend and relative trend methods.

   ii. Compute simple one-step-ahead forecasts using EWMA and Holt (with optimal smoothing parameters).

   iii. Compute the MSE, MAE, U and Minzer-Zarnowitz R^2 for the time points $[201; end]$.

(c) Conclude which model is the best one for each phase and write down your ideas and conclusions about the applicability of the considered forecasting methods for different types of time series.

3. Simple forecasting and forecasting with exponential smoothing with `BeerWineUS.csv`. Begin the forecasting starting with the observation 201. Thus the out-of-sample performance of the below methods will be assessed using the forecasts for the periods 201 to 311.

   (a) Compute forecasts using simple EWMA, Holt and Holt-Winters forecasts with the smoothing parameters calibrated from the first 200 observations.

   (b) Compute the corresponding MSE 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`. Keep the last year for forecasting.

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

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

   (c) 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.

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

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

   (f) Explain why multi-step-ahead forecasts have wider forecast intervals than one-step-ahead-forecasts.

   (g) Imagine an ACF with only the first two correlations being significant. Which process is suitable to model this and why?

   (h) Imagine an ACF which consists only of positive values and quickly decays towards zero. Which process is suitable to model this and why?

   (i) Consider an AR(1) process with parameter α_1. Assume we have a shock to a time series (a large error term, unexpected event) at the time point $t = 10$. Which impact do you expect this shock to have on the observation at time point $t = 15$? Provide the formula and give formal motivation.

5. GARCH modelling of the crude oil prices. Compute the log-returns of the oil prices, i.e. $r_t = log(Y_t/Y_{t-1})$ and work with this time series.
   (a) Check the ACF of the returns and squared returns. Which specific features do the ACFs have?
   (b) Fit ARCH(1) and GARCH(1,1) processes to the data. Extract the conditional volatilities and plot them.
   (c) Check the properties of the residuals and choose the optimal model relying on BIC.
   (d) Explain in your own words the difference between conditional and unconditional volatility.
   (e) Explain in your own words idea behind the news curve.
6. Structural breaks with the crude oil price.
   (a) The objective is to determine the intervals where the time series shows homogenous behaviour. Apply either `breakspoints` or a regression-tree type approach to find the intervals where
      i. the time is series is approximated by constants;
      ii. the time is series is approximated by linear trends;
   (b) Compute moving window estimates for OLS parameters (use 100 observations) of the oil price on a simple time variable. Plot the results and decide about the breaks.
   (c) Explain the consequences of breaks in time series for modelling. You can pick up a particular model to motivate your ideas.
   (d) Explain the consequences of breaks in time series for forecasting. Think about the forecasting method, time point of the break, etc.