Time Series Analysis 02417
Spring 2025

Assignment 1

Friday 28th February, 2025 13:54

Instructions:
The assignment is to be handed in via DTU Learn "FeedbackFruits" latest at Monday 3rd at 23:59. You are allowed to hand in in groups of up to max 4 persons. You must hand in a single pdf file presenting the results using text, math, tables and plots (do not include code in the main report - your code must be uploaded as a separate file, it's not being evaluated directly). Arrange the report in sections and subsections according to the questions in this document. Please indicate your student numbers on the report.

1 Plot data

In this assignment we will be working with data from Statistics Denmark, describing the number of motor driven vehicles in Denmark. The data is provided to you, but if you are interested you can find it via `www.statistikbanken.dk` (search for the table: "BIL54"). Together with this document is a file with data `DST_BIL54.csv`, it holds timeseries of monthly data starting from 2018-Jan.

You can decide how to read the data – a script is available in the file `read_data.R`, where the data is read and divided into a training and a test set: The **training set** is from the beginning to 2023-Dec, the **test set** is the last 12 months (2024-Jan to 2024-Dec). **To begin with we will ONLY work with the training set**.

The variable of interest is `total`, which is the number vehicles in registrered in Denmark at a given time (in Danish "Drivmidler i alt"). We will ignore the other variables in the dataset.

Do the following:

1.1. Make a time variable, x, such that 2018-Jan has $x_1 = 2018$, 2018-Feb has $x_2 = 2018 + 1/12$, 2018-Mar has $x_3 = 2018 + 2/12$ etc. and plot the training data versus x.

1.2. Describe the time series in your own words.

2 Linear trend model

We will now make a linear trend model, which is a general linear model (GLM) of the form:

$$Y_t = \theta_1 + \theta_2 \cdot x_t + \epsilon_t \tag{1}$$

where $\epsilon_t \sim N(0, \sigma^2)$ is assumed i.i.d. The time is $t = 1, \dots, N$.

2.1. Write up the model on matrix form for the first 3 time points: First on matrix form (as vectors and matrices), then insert the elements in the matrices and vectors and finally, insert the actual values of the output vector $\boldsymbol{y}$ and the design matrix $\boldsymbol{X}$ (keep max 3 digits). All group participants do it – include picture for each in the report.

3 OLS - global linear trend model

Parameters of the model as a global linear trend model:

3.1. Estimate the parameters θ_1 and θ_2 using the training set (call it the Ordinary Least Squares (OLS) estimates). Describe how you calculated the estimates.

3.2. Present the values of the parameter estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ and their estimated standard errors $\hat{\sigma}_{\hat{\theta}_1}$ and $\hat{\sigma}_{\hat{\theta}_2}$. Plot the estimated mean as a line with the observations as points.

3.3. Make a forecast for the test set, hence the following 12 months - i.e., compute predicted values with corresponding prediction intervals for 2024-Jan to 2024-Dec. Present these values in a table.

3.4. Plot the fitted model together with the training data and the forecasted values (also plot the prediction intervals of the forecasted values).

3.5. Comment on your forecast – is it good?

3.6. Investigate the residuals of the model. Are the model assumptions fulfilled?

4 WLS - local linear trend model

We will now use WLS to fit the linear trend model in Eq. (1) as a local trend model, i.e., the observation at the latest timepoint (N) has weight $\lambda^0 = 1$, the observation at the second latest timepoint ($N-1$) has weight λ^1, the third latest observation ($N-2$) has weight λ^2 etc.

We start by setting $\lambda = 0.9$.

4.1. Describe the variance-covariance matrix (the $N \times N$ matrix Σ (i.e. 72×72 matrix, so present only relevant parts of it)) for the local model and compare it to the variance-covariance matrix of the corresponding global model.

4.2. Plot the "λ-weights" vs. time in order to visualise how the training data is weighted. Which time-point has the highest weight?

4.3. Also calculate the sum of all the λ-weights. What would be the corresponding sum of weights in an OLS model?

4.4. Estimate and present $\hat{\theta}_1$ and $\hat{\theta}_2$ corresponding to the WLS model with $\lambda = 0.9$.

4.5. Make a forecast for the next 12 months - i.e., compute predicted values corresponding to the WLS model with $\lambda = 0.9$.

   Plot the observations for the training set and the OLS and WLS the predictions for the test set (you are welcome to calculate the std. error also for the WLSand add prediction intervals to the plots).

   Comment on the plot, which predictions would you choose?

4.6. **Optional:** Repeat (estimate parameters and make forecast for the next 12 months) for $\lambda = 0.99, \lambda = 0.8$, $\lambda = 0.7$ and $\lambda = 0.6$. How does the λ affect the predictions?
   Comment on the forecasts - do the slopes of each model correspond to what you would (roughly) expect for the different λ's?

5 Recursive estimation and optimization of λ

Now we will fit the local trend model using Recursive Least Squares (RLS). The smart thing about recusive estimation is that we can update the parameter estimates with a minimum of calculations, hence it's very fast and we don't have to have keep the old data.

5.1. Write on paper the update equations of $\boldsymbol{R}_t$ and $\hat{\boldsymbol{\theta}}_t$.

For $\boldsymbol{R}_t$ insert values and calculate the first 2 iterations, i.e. until you have the value of $\boldsymbol{R}_2$.

Initialize with

$$R_0 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$$

and

$$\boldsymbol{\theta}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Note, that now the parameters are noted as a vector and thus the subscript is time, so for the linear trend model

$$\hat{\boldsymbol{\theta}}_t = \begin{bmatrix} \theta_{1,t} \\ \theta_{2,t} \end{bmatrix}$$

Everyone in the group must do this on paper and put a picture with the result for each in the report.

5.2. Implement the update equations in a for-loop in a computer. Calculate $\hat{\boldsymbol{\theta}}_t$ up to time $t = 3$. Present the values and comment: Do you think it is intuitive to understand the details in the matrix calculations? If yes, give a short explanaition.

5.3. Calculate the RLS estimates at time $t = N$ (i.e. $\hat{\boldsymbol{\theta}}_N$) and compare them to the OLS estimates, are they close? Can you find a way to decrease the difference by modifying some of the RLS initial values and explain why initial values are important to get right?

5.4. Now implement RLS with forgetting (you just have to multiply with λ at one position in the R_t update).

Calculate the parameter estimates: $\hat{\theta}_{1,t}$ and $\hat{\theta}_{2,t}$, for $t = 1, \ldots, N$ first with $\lambda = 0.7$ and then with $\lambda = 0.99$. Provide a plot for each parameter. In each plot include the estimates with both λ values (a line for each). Comment on the plots.

You might want to remove the first few time points in the plot, they are what is called a "burn-in" period for a recursive estimation.

Tip: It can be advantageous to put the loop in a function, such that you don't repeat the code too much (it's generally always a good idea to use functions, as soon as you need to run the same code more than once).

You might want to compare the estimates for $t = N$ with the WLS estimates for the same λ values. Are they equal?

5.5. Make one-step predictions

$$\hat{y}_{t+1|t} = \boldsymbol{x}_{t+1|t}\hat{\boldsymbol{\theta}}_t$$

The notation $t+1|t$ means the variable one-step ahead, i.e. at time $t+1$, given information available at time t. So this notation is used to denote predictions. For $\boldsymbol{x}_{t+1|t}$ we do have the values ahead in time for a trend model – in most other situations we must use forecasts of the model inputs.

Now calculate the one-step ahead residuals

$$\hat{\varepsilon}_{t+1|t} = \hat{y}_{t+1|t} - y_{t+1}$$

Note, they could also be written

$$\hat{\varepsilon}_{t|t-1} = \hat{y}_{t|t-1} - y_t$$

Applying a shift from "$t+1|t$" to "$t|t-1$" makes no difference.

Plot them for $t = 5, \ldots, N$ first with $\lambda = 0.7$ and then $\lambda = 0.99$ (note, we remove a burn-in period ($t = 1, \ldots, 4$ or more, might not be necessary, but usually a good idea when doing recursive estimation – depends on the initialization values).

Comment on the residuals, e.g. how do they vary over time?

5.6. Optimize the forgetting for the horizons $k = 1, \ldots, 12$. First calculate the k-step residuals

$$\hat{\varepsilon}_{t+k|t} = \hat{y}_{t+k|t} - y_{t+k}$$

then calculate the k-step Root Mean Square Error (RMSEk)

$$RMSE_k = \sqrt{\frac{1}{N-k}\sum_{t=k}^{N} \hat{\varepsilon}^2_{t|t-k}}$$

Do this for a sequence of λ values (e.g. 0.5,0.51,...,0.99) and make a plot.

Comment on: Is there a pattern and how would you choose an optimal value of λ? Would you let λ depend on the horizon?

5.7. Make predictions of the test set using RLS. You can use a single λ value for all horizons, or choose some way to have different values, and run the RLS, for each horizon.

Make a plot and compare to the predictions from the other models (OLS and WLS).

You can play around a bit, for example make a plot of the 1 to 12 steps forecasts at each time step to see how they behave.

5.8. Reflexions on time adaptive models - are there pitfalls!?

- Consider overfitting vs. underfitting.
- Are there challenges in creating test sets when data depends on time (in contrast to data not dependend on time)?
- Can recursive estimation and prediction aliviate challenges with test sets for time dependent data?
- Can you come up with other techniques for time adaptive estimation?
- Additional thoughts and comments?