Time Series Analytics and Forecasting

Student Individual Project - Instructions

Your grade for this course will be decided through this individual project and its presentation to the class next week. The project consists of:

1. Identifying a suitable time series data set for analysis;

2. Applying the methods that we have looked at and will look at in the course;

3. Critically assessing the performance of each of the methods, and identifying the most appropriate for your data set;

4. Writing up your results in a professionally-looking report;

5. Presenting the work to the class as a 10–15 minute presentation.

Please follow the instructions below.

Identifying a Time Series

If you have any specific time series dataset in mind (e.g. financial data, economic or demographic data of Vietnam or any other country, sport data, sales data etc.), you are encouraged to use it, as long as it satisfies the points described in **task 1** (see below).

An excellent source of economic, demographic and environmental time series data has been made available by the Irish Central Statistics Office (CSO) at:

<center>http://www.cso.ie/px/pxeirestat/statire/SelectTable/Omrade0.asp?Planguage=0</center>

From this webpage, you can select a time series to download, and also specify the range of years that you wish to look at. The series can be downloaded as a `csv` file.

Task 1: *Investigate the data sets that are available from this webpage or any other sourse, and choose a time series that you want to analyse for your project. The series should have the following properties:*

- *It has at least 100 observations;*

- *It shows either a trend or seasonal behaviour (or both);*

- *No one else in the course is using the series.*

Analysis

In the second part, you will apply the methods that have been, or will be, described in the course to your data set.

Task 2:

1. *Plot the data as a time series. Also plot the ACF and PACF with limits for the test that the value is 0.*

2. *Fit the following models to your data set (**notice**: if for some reason we won't have time to cover a specific method during the course, then the use of this method won't be required for your project):*

   - *AR(1) and an AR model with the strongest seasonal lag (if your data show seasonal behaviour);*
   - *Exponential smoothing;*
   - *Holt-Winters;*
   - *ARMA;*
   - *ARIMA;*
   - *Kalman filter with $F_t = 1$;*
   - *GARCH model (or, alternatively, another model depending on what models we will be able to cover during the course).*

3. *You should transform your data where you think it is necessary.*

4. *For the AR model, exponential smoothing, Holt-Winters, ARMA and ARIMA you should find the optimal parameters of each model by both RSME and MAPE.*

5. *For the case of the Kalman Filter, you have 3 sets of free parameters to find: G_t, V_t and W_t (F_t is fixed to be 1). Assume that each set is constant with t e.g. $G_t = G$, $V_t = V$ and $W_t = W$ for all t, and so there are 3 parameters G, V and W to identify. Choose a set of combinations of values (G, V, W) and find the best combination by RMSE and MAPE. For extra credit, try to find an optimal set of values with* `R` *by using the* `optim` *function, or otherwise, but please note that the problem of optimising these parameters together can be very difficult!*

6. *You can fit each of the models in Excel and/or* `R`*, whichever you prefer. It is perfectly fine if you use another software you are more familiar with (e.g. Python or Matlab).*

Assessing Model Performance

Task 3: *Assess the performance of each of your models by:*

- *Producing 1-step ahead forecasts for each observation e.g. computing a predicted value for x_{t+1} given previous values $x_t, x_{t-1}, \ldots$. Plotting a time series plot of these predictions next to the observed values. Computing RMSE and MAPE values for these forecasts;*

- *Computing forecasts for 12 time points into the future, with 95% prediction intervals where possible. Plotting these forecasts for each model for comparison.*

- *Use a criterion such as AIC to compare as many of the models as possible.*

Report Write-Up

Task 4: *Write down the results of Tasks 1, 2 and 3 in a report. The report can be written in Word or LaTeX.*
More detailed instructions for the report:

- *The number of pages should be in the range 10-15 (including figures, tables and other material you find meaningful).*

- *A possible outline of the report could be:*

  1. *Data description*

  2. *Fitted models*

  3. *Model comparison and conclusions*

- *Make sure that the report, and in particular plots and tables, are presented in a clear and professional way (e.g. with axis labels, legend if needed, captions and so on).*

- *Recall that the report should describe the work you did for your project and it's not a review of the classes. So, if you decide to include some theory (e.g. description of the model that are fitted), make sure this introduction is concise.*

Presentation

Task 5: *Each student will present their results for Tasks 1, 2 and 3 to the rest of the class on **Thursday 29th of March 2018**. Prepare a presentation using Powerpoint, LaTeX(e.g. Beamer), Keynote, etc. The presentation should last between 10 and 15 minutes. The presentation should include a description of the data set, exploratory analysis, description of how each model was fitted, plots of k-step ahead predictions against the data, and conclusions on which model or models performed best.*
Make sure that:

- *The presentation follows a clear path: introduction, methods, results and conclusions.*

- *Plots and Tables are properly presented, including title, labels for the axes and legend if suitable.*

- *The material presented is suitable for a 10 - 15 minutes presentation (nor too little, nor too much).*

Your grade for this course is assessed as 50% from the report and 50% from your presentation. Extra marks are awarded if you apply methods beyond those that we have looked at in the course e.g. looking at confidence and prediction intervals, alternative ways of assessing model performance and comparing models. If you do this then be sure to explain clearly what you have done in both the report and presentation.

Your report must be e-mailed to me at `nipotib@tcd.ie` *not later than Friday 13th April 2018*