2102531 System Identification

Term Project
Semester 1/2018

Jitkomut Songsiri

November 16, 2018

Abstract

This project aims to apply an identification technique to real-world problems. There are **three** topics selected from various applications, namely, (i) learning brain network differences from EEG data, (ii) photovaltaic system modeling, and (iii) time series models of stock prices. Students spend four weeks on this project where each week a progress report must be submitted.

Instruction to students

Each week students are supposed to submit the weekly progress report to CourseVille in Assessment section, where *everyone* must submit the group report in typesetting and .pdf format. Any proposed ideas, comments, or practical considerations should be stated in the report to keep track of your concerns or problem found as the work progresses. The use of LaTeX to type the document is highly encouraged. The progress reports should be named as `brain_network_diff_week1.pdf` `pvmodel_week1.pdf` `stockprice_week1.pdf` where the week number refers to the week of submitting the report. We have created a shared folder `ee531_project_2018` and all the established necessary files must be kept here except your codes in progress. Put the progress reports in the folder `report` each week as well.

Student lists

1. Parinthorn Manomaisaowapak

2. Kitinan Boonravee

3. Jeerapat Jitnuant

4. Maytus Piriyajitakonkij

Final report format

All groups should use the following formats:

- All vector graphics (such as MATLAB figures) must be in .eps or .pdf format.

- All MATLAB codes must be in the appendix with proper comments explaining the codes. You can use `listings` package for source code listing in LaTeX. Mark the language (MATLAB) to highlight the codes. Each MATLAB file should be referred to the section where it is used, or you can explain which function is used to generate each figure.

- Use BibTeX to create reference lists.

1 Learning Brain Network Differences from EEG Data

Student list:

1. Parinthorn Manomaisaowapak

1.1 Progress deadlines

Wed Oct 31

- Typesetting report of problem description.

- Explain the application and define the problem explicitly; the model, model characteristics, input, and output.

Wed Nov 7

- Revise the old contents.

  - Math notation.
  - Problem statement must be specific and clear.
  - Revise the explanation of GC and relate it with AR model.
  - Add references.

- New list:

  - Specify the description of data used in the experiment. Explain how to split data into train and test sets, and how to split the trials.
  - Explain the experiment setting and methodology used in MVGC toolbox.
  - Perform model estimation on both control and patient group. It is okay if we cannot finish estimation of all data trials.
  - Report model estimation and model selection results (preliminary is okay).

Wed Nov 14

- Revise the old contents.

  - Granger casaulity should be explained with VAR process (vector) and explain F_{ij} as an effect from y_j to y_i.
  - Revise the problem statement. Be specific about what to acheive. No need to explain the methods.
  - Add "Methodology" section. You should explain about all steps/procedures involved in order to acheive the goal. For example, the methodology consists of model estimation (by XXX method), statistical test (by XXX method with details on statistics used), and so on.
  - Revise the experiments where p should be selected from a range of 1 to XXX (as long as the model estimation has no problem). Make sure to the rank deficient problem is not observed. Finish the model selection experiment and report the results in an easier way to interpret. If you want to compare between OLS and LWR algorithms, you should display the RMSE in terms of bar or plots instead.

- New list:

  - Explore heuristic methods in the toolbox or literature when the rank deficient of the regressor matrix occurs.

- Add "Data description" section. Add every detail (and we will send this part to USM) for confirmation. Summarize all explanation from USM emails in to this section.

- Make "Experiment" as a section (not a subsection.) Explain about data splitting process (into trials, training and test sets). The numerical results can be under this section.

- Conclude the model selection results (from the data we have), and start Granger causality test.

Wed Nov 21

- Revise the old contents.

  - Introduction should contain more contents: problem motivation, literature, and a short description of this project.

  - Data spliting can be explained in data description. Give the list of channel number and the corresponding channel names. You can reorder them, so that the channel lables are sorted in spatial domain.

  - Correct the experimental results based on channel selections.

- New list:

  - Finish experimental results based on channel selection.

Mon Nov 26: Term project presentation

- XXX

2 Photovaltaic System Modeling

Student list:

1. Kitinan Boonravee

2. Jeerapat Jitnuant

2.1 Progress deadlines

Wed Oct 31

- Review about past research on system modeling (statistical models, neural network, electrical model). See example of papers in Web Resource of Courseville. See term project of EE531 in 2017 (by Jennarong) and technical report of Supachai (in Courseville - supplementary material folder).

- Typesetting report of problem description.

- Explain the application and define the problem explicitly; the model, model characteristics, input, and output.

Wed Nov 7

- Revise the old contents.

  - Contents in the progress is very minimal. Add all information that you read from the literature and conclude it in the report. Finalize the list of PV model candidates.

  - Add references.

   – Explain scope and how to evaluate the performance of all models; how to compare them.

- New list:

  – Explain the methods used to estimate in each model.

  – Perform PV model estimation from simple models first (the ones can be done by regression). See if models should be estimated independently from different times.

  – Perfrom one nonlinear PV model estimation.

  – Report the results and discussion.

Wed Nov 14

- Revise the old contents.

  – The problem statement is still not clear.

  – The description of models used in reference [1,2,4,6] should be explained more explicitly in the literature.

  – Background should contain all the contents on the PV electrical circuit, ANN, and SVR. For SVR, you should really understand the concept of this problem and how to apply to the PV modeling.

  – Change "Proposed Model" section to "Methodology" and describe the list of models (5 models) explicitly. Write down the input list of each model and the output (predicted power). Describe the assumption of each model why such form is proposed. In this section, also explain what performance index you will use to compare all the models.

  – In "Experiment" section, explain i) data description, ii) experiment setting (how to split data into training and test sets, etc.)

- New list:

  – Check the data of irradiance and power to see if some pair of the two values do not agree. Make a criterion to filter out these outliers.

  – Finish the experiment of estimating polynomial models and ANN models.

Wed Nov 21

- Revise the old contents.

  – Finish and respond to the old comments on Nov 14.

  – Revise the introduction: the problem motivation, literature, and what you're going to do (in short).

  – Explain the background in the way that the reader understand basic concepts of the tools you will use.

  – Methodology should contain all steps involved in order to solve the problem statement.

  – The description of outlier detection is not clear to understand. This is part of data pre-processing.

  – Think before presenting the results. Select a proper way to visualize the results.

- New list:

  – Finish all the experiments and conclude the results.

Mon Nov 26: Term project presentation

- XXX

3 Time Series Models of Stock Price

Student list:

1. Maytus Piriyajitakonkij

3.1 Progress deadlines

Wed Oct 31

- Survey about typical time series models used in stock price modeling. Basic methods can be found from time series textbooks.

- Typesetting report of problem description.

- Explain the application and define the problem explicitly; the model, model characteristics, input, and output.

Wed Nov 7

- Revise the old contents.

  - Specify the scope of this work. Specify the stock variable used in this experiment.
  - Add references on the chosen models and literature detail (how to choose inputs, etc.)
  - Problem statement must be clear and specific. What forecast value will you provide ? $\hat{y}(t+1)$ only ?
  - Conclude the list of model candidates and how to compare them.

- New list:

  - Explain data description.
  - Explain experiment settings and perform one model estimation.
  - Report results and discussion.

Wed Nov 14

- Revise the old contents.

  - In "Model" section, change to "Background" and explain about existing tools (time series, ANN, RNN) with literature detail.
  - In "Methodology" sectio, explain the list of models you will examine in this project. Specify the model structure and the input list clearly (especially, for ANN, RNN). Explain all the processs involved, and the performance indices you will use to compare the model. Check the correctness of R^2 and AIC, BIC.
  - The selection of p, q, d should follow the guideline in time series textbook. This will be examined from ACF and PACF plots (decaying behaviour).
  - In "Experiment" section, explain the data description more clearly. How to split the data into training and test sets. Then explain the experiment setting and report the results.

- New list:

  - Finish the estimation experiments using time series modela and one ANN model. Conclude the results.

Wed Nov 21

- Revise the old contents.

  - Estimation results of time series model are strange. Check the correctness.
  - Give more details about ANN in terms of estimation method (algorithm, library, etc.) Varying parameter of ANN in a more suitable way. Running neuron from 2 to 20 is meaningless.
  - Re-organize the document and think about how to group the same contents together. If the performance index is RMSE, present the results in terms of this index.
  - Compare the results from all models. Think about how to visualize the results.

- New list:

  - Finish the experiment and conclude the results.

Mon Nov 26: Term project presentation

- XXX