Assignment 3: System identification and multi-output ARMAX models

Instructions: The assignment report is to be handed in via DTU Learn "FeedbackFruits" latest at April 12th at 23:59. You are allowed to hand in in groups of 1 to 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 report! Shortened software output is ok this time, to save some time formatting in tables. Arrange the report in sections and subsections according to the questions in this document. Please indicate your student numbers on the report.

NOTE, that this time there is no peer-review. The teachers evaluate the reports. The assignment is a little bit more "free" in the sense of the questions and thus also your answer can be a little more free, for example we don't expect that you all select the same models. The argumentation is thus more important than to select the same model as we do.

NOTE, that the report should not be too long! Include only one (max two figures) (i.e. one figure can have multiple plots) per question and make the text concise! Long and unprecise reports are not good!

All material needed is provided in the `assignment3.zip` file. There are included some R helping functions in files, see how they are loaded in the `assigment3_intro.R` script, which was presented during a lecture, see the recording if you missed it.

It will be difficult not to use R in this assignment, especially for the parts where `marima` is used. Matlab system identification toolbox can perhaps be used (we are not sure about how multiple-output models are handled). We don't have a particular alternative in Python, perhaps https://rviews.rstudio.com/2022/05/25/calling-r-from-python-with-rpy2/ can by used, if anyone has success with that we would like to know :). In any case, try the R helping functions!

Introduction

First a description of the experiment setup and data is given.

The system consists of a pot and a drinking glass, both filled with water. The drinking glass is placed in the pot, such that the water of the pot covers up to half the height of the glass. In both the pot and the glass is placed a 300 W water heater.

In Figure 1 two pictures of the system are shown. As it can be seen a small "model-ship outboard motor" was used to mix up the water, such that the water is mixed and the temperature in the each compartment can be assumed homogenous with the applied sampling frequency. In Figure 2 a system diagram depicts the setup.

The following variables have been measured:

- $T_{\mathrm{i},t}$ (`Tinner`) (°C): The temperature of the water in the glass
- $T_{\mathrm{o},t}$ (`Touter`) (°C): The temperature of the water in the pot
- $T_{\mathrm{a},t}$ (`Ta`) (°C): The temperature of the ambient air surrounding the system

Figure 1: In the left image the setup is seen from the side and in the right image from above (the right image don't have the mixing moter, but it was in the experiment used for the assignment).

- $P_{\mathrm{i},t}$ (`Pinner`) (W): The Power input of the heater in the glass
- $P_{\mathrm{o},t}$ (`Pouter`) (W): The power input of the heater in the pot

The variables are sampled with a sampling period of 10 second. In the data the column `t` is Unix time UTC.

Data from one experiment is used. In the experiment both heaters were switched on and off following to a designed test sequence, such that the on-off periods were independent.

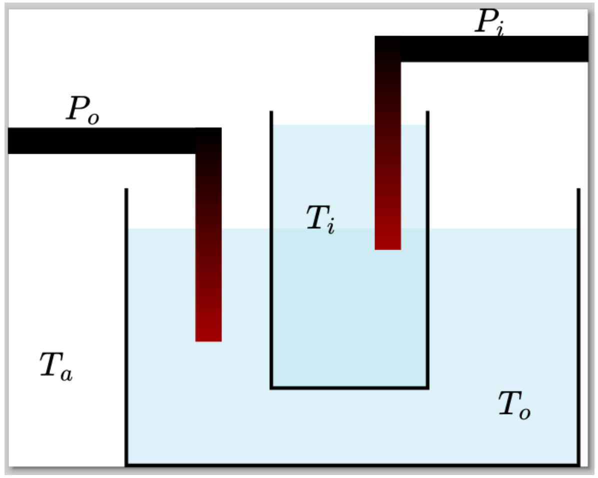

Figure 2: Diagram of the experiment. Each compartment can be assumed homogeneous meaning that the temperature is the same in all of the material (water in this case) in the compartment.

1 Explorative data analysis

1.1. Make plots of the data in `data/experiment1.csv`. Illustrate the time series to and make one or plots of a short period to zoom in on an interesting part of the data.

1.2. From the plot describe the experiment. What are the causalities, i.e. which variables are "dependent" and which are "independent"? Is the system a linear time-invariant (LTI) system or what could cause it not to be?

1.3. Try to investigate the relations between the variables using the cross-correlation function between the variables. Do the CCFs provide more insights than the time series plots?

2 ARMAX model of the glass water temperature

First, we want to build a model with the drinking glass water temperature as output and ambient temperature as input.

Start with an ARX model

$$\phi(B)T_{\mathrm{i},t} = \omega_{\mathrm{a}}(B)T_{\mathrm{a},t} + \varepsilon_t \tag{1}$$

where ε_t is assumed i.i.d. Keep the same order of AR and the all the coefficient polynomials.

For example the ARX model above of order 1 written out as a regression model is

$$T_{\mathrm{i},t} = -\phi_1 T_{\mathrm{i},t-1} + \omega_{\mathrm{a},1}T_{\mathrm{a},t-1} + \varepsilon_t \tag{2}$$

and of order 2

$$T_{\mathrm{i},t} = -\phi_1 T_{\mathrm{i},t-1} - \phi_2 T_{\mathrm{i},t-2} + \omega_{\mathrm{a},1}T_{\mathrm{a},t-1} + \omega_{\mathrm{a},2}T_{\mathrm{a},t-2} + \varepsilon_t \tag{3}$$

and so on for higher orders.

2.1. Implement the ARX order 1 model in Equation (2) as a linear regression model and estimate the parameters with the least squares method. Provide the result and comment on it.

It's ok to copy the output from the software in the report (not code, only output), for example the relevant part of the summary table from `lm()`.

Tip., in R you can make the lags with the provided `lagdf()` function and use the `lm()` function and use the provided `ARX()` function to easily generate a formula for higher order models.

2.2. Make a model validation with the plot of residuals, acf and pacf, as well as CCFs from the inputs to the residuals. Is the assumption of white noise fulfilled? Comment on the result.

2.3. Identify a suitable ARX model for $T_{\mathrm{i},t}$. Add relevant inputs and increase the order (use same for all inputs and AR, i.e. always the same number of lags (from 1 to "order") for all variables).

Use the significance of parameters and validation plots. Furthermore, backward selection and AIC can be used. Argue for the selected inputs and order of the suitable model (i.e. neigher over- or under-fitted).

Tip: In general, as experience grows, one tends to become more "conservative" (as with other things in life ;) and choose simpler models. One argument, that usually many other effects come into play, e.g. some phenomena not in the data, etc.

Tip: In general, it's usually a good idea to first find out which inputs to use and then start increase the order of the model, but there is no final correct way, but try to be structured following some kind of model selection procedure.

2.4. Present the validation of the selected model.

2.5. Discuss the results: was it easy to select the inputs and the order? Did the different modelling techniques give the same structure? Are you still in doubt? What considerations did you have?

3 ARMAX model of the glass water temperature

Install the marima package from the provided file:

```r
# Install it (also update )
download.file("https://02417.compute.dtu.dk/material/marima2_0.1.tar.gz", "marima2_0.1.tar.gz")
install.packages("marima2_0.1.tar.gz", repos=NULL)
library(marima2)
```

3.1. Use the marima package to estimate the ARX order 1 model (from Equation (2)) and check that the residuals are equal (or very very close).

3.2. Fit an order 1 ARMAX model: Add the inputs (as for the ARX model in previous section) and compare the difference to the similar ARX (use marima also, just remove the MA part). Comment on the validation plots and discuss what the MA part does.

3.3. Identify the suitable order of the ARMAX model. Increase the order (same order for AR and all inputs), use same techniques as for the ARX from last section for model selection and validation.

3.4. Discuss the results, did you get the same results or is there some difference between the results for ARX and ARMAX?

3.5. Make a multi-step prediction with the selected model. Can it predict the temperature through out the exeriment?

3.6. Make some simulated "step responses" of the temperature with a step from 0 to 100 W on the heat input, does it seem realistic? Make also an extreme one and discuss the limitations of a linear model.

4 ARMAX model of the pot water temperature

4.1. Identify an ARMAX model for the water temperature of the pot ($T_{\mathrm{o},t}$). Use the same selection techniques. Provide a few considerations from the selection results.

4.2. Present the selected model and shortly comment on the results.

4.3. Compare the two identified ARMAX models (i.e. for $T_{\mathrm{i},t}$ and $T_{\mathrm{o},t}$). Discuss for example: Is the structure the same? parameter estimates? Any interesting differences and similarities?

5 ARMAX multi output model

5.1. Use marima to estimate an ARMAX multi-output model of the coupled system. Use the same model inputs and order as identified in the individual models.

5.2. Identify the order. Increase or decrease the order for all inputs together, use the same techniques as before.

5.3. Make a multi-step prediction with the selected model. Can it predict both temperatures through out the exeriment?

5.4. Simulation of step response from one input to both temperatures simulatiously. E.g. step 100 W on $P_{\mathrm{i},t}$ and see the effect on both temperatures, play around, and present three interesting simulations.

5.5. Discuss the pros and cons of the coupled model over the two independent models.