

# System Identification – Practical Assignment 5

## Correlation Analysis

### Logistics

This practical assignment should very preferably be carried out by each student separately. Otherwise, if there are more students than computers, *with the explicit agreement of the lab teacher for each group*, students may team up in groups of 2.

The assignment solution consists of Matlab code. Develop this code in a single Matlab script. If you need to create functions, they can be local to the script, see local functions in scripts.

The overall rules for labs are described on the website. For each particular lab, your attendance will only be registered if you have a working, original solution. The teacher will check that your code **works** during the lab class. Only after this has been done, for the **originality** check, upload your solution here:

<https://www.dropbox.com/request/5viGSrmbGaCwSlG6Q1oz>

Upload *only once*, a single *m-file*, named *exactly according to the following pattern*:

L5\_ENgh\_LastnameFirstname.m

where *g* is your group, *h* your halfgroup, and your last and first names follow. E.g., L5\_EN32\_PopAlex.m.

*If you worked in a group per the above, only upload one file with both student names included.* Any files that are duplicate, nonstandard, inappropriately named, or corresponding to unchecked solutions will not be considered. Files will be automatically tested for plagiarism, and any solution that fails this test will be marked copied; only solutions that pass both the working and the originality test are definitively validated. Therefore, while you are encouraged to discuss ideas and algorithms amongst colleagues, sharing and borrowing pieces of code is forbidden.

### Assignment description

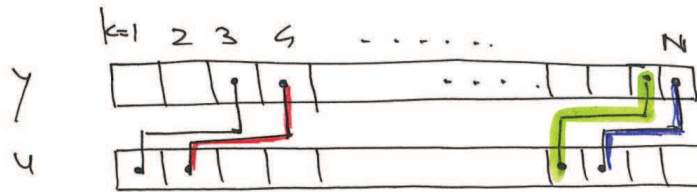
We apply the linear regression method to obtain finite impulse response (FIR) models from input-output data – see the course material, Part IV: *Correlation Analysis*. This data is more general than just the step or impulse responses we have been seeing so far.

Each student is assigned an index number by the lecturer. Then, the student downloads the data file that forms the basis of the assignment from the course webpage. The file contains the identification data in variable `id`, and separately the validation data in variable `val`. Both these variables are objects of type `iddata` from the system identification toolbox of Matlab, see `doc iddata`. The corresponding time vectors are `tid`, `tval`.

Requirements:

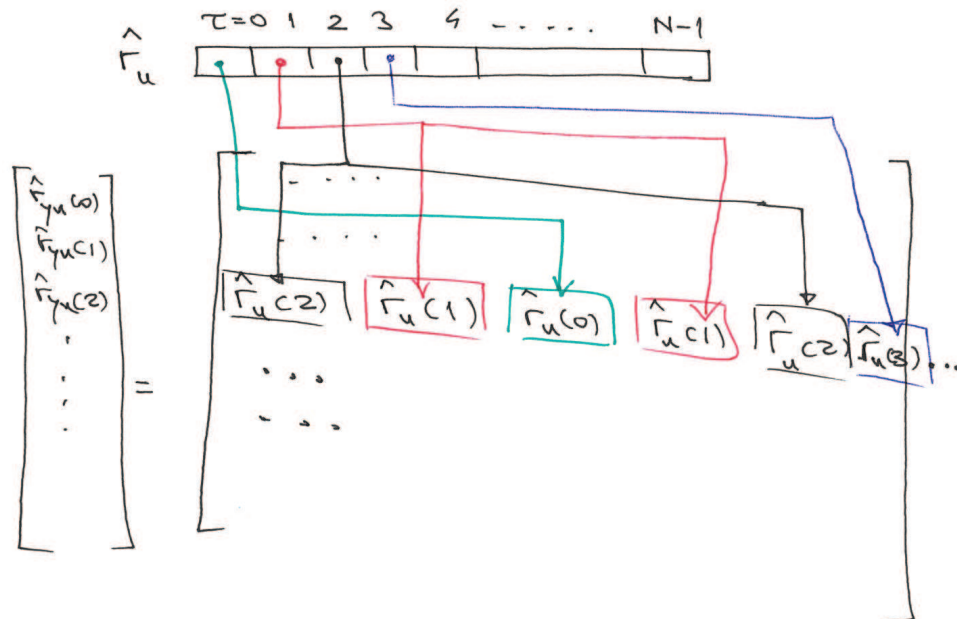
- Plot and examine the data supplied. Determine whether the input and output are zero-mean or not. If the signals are not zero-mean, remove the means, either manually or using e.g. `detrend`.
- After making sure the signals are zero-mean, compute the correlation functions  $r_u(\tau)$  and  $r_{yu}(\tau)$  from the identification data, using the formulas from the lecture. An illustration is provided below. Check whether the input signal is white noise. Hint: By checking the structure of the linear system, it will become clear that you only need to compute the covariances for positive  $\tau$ .

$\tau=2$ , for example



$$\frac{1}{N} \sum_{k=1}^{N-\tau} y[k+\tau] u[k] = \frac{1}{N} (y[c3] u[c1] + y[c4] u[c2] + \dots + y[cN-1] u[cN-3] + y[cN] u[cN-2])$$

- Implement the system of linear equations to obtain the FIR model for an arbitrary value of  $M$  supported by the data. Hint: Make sure you understand the structure of the regressor matrix, with the values of  $r_u$  “reflected” around 0, for example:



- Compute the convolutions to simulate the model response to the identification and validation inputs, and compare with the identification and validation outputs.
- Study the influence of the length  $M$  of the FIR model on the accuracy of the model. A rule of thumb for selecting a good  $M$  is as follows: preferably the entire transient regime of the impulse response must be modeled (until it reaches steady state), but without estimating too many parameters because this would introduce overfitting. These requirements may be conflicting so you may need to deviate one way or the other to get a good fit.

To allow for more insight, the true impulse response of the system is provided as a vector `imp` in the

datafile (note that this true response would not be available in a real identification experiment). Solve each part of the assignment without using the true response, but once you have solved it, you may optionally compare the FIR models obtained with the true impulse response.

This is the first lab where we start using in earnest the System Identification toolbox (`ident`). Relevant functions from this toolbox: `detrend`, `plot`, `compare`, `cra`, `impulseest`; and generic Matlab function `conv`. See also `doc ident` for the full documentation of the toolbox.