System Identification – Practical Assignment 4
Correlation Analysis

Logistics

Please reread the logistics part of lab 2, the same rules will apply to this lab. The only thing that changes is the DropBox link, which will be communicated separately.

Assignment description

We apply the linear regression method to obtain finite impulse response (FIR) models from input-output data – see the course material, *Correlation Analysis*. This data is more general than just the step or impulse responses we have been seeing so far. [1]

Each student will obtain a data set using the DC motor and will identify the system, as detailed in the following instructions. **Important note**: To minimize system wear, separate the code that generates the data from the code that performs the rest of the steps below (easiest using different script sections, see *Code Sections* in the Matlab documentation), and **regenerate the data only when necessary**.

1. To keep things simple, we will create a single, longer sequence of data containing both the identification and validation data. We will use a sampling rate of 0.01 s (10 ms). After a short range of zero inputs, apply a uniformly random input signal with amplitudes in the interval $[-0.7, 0.7]$ and a length of 400 to 600 samples (our identification data), followed by another range of zero inputs, and then a step signal of magnitude around 0.2 and around 70 samples in length (this last step will be used for validation). Isolate the data range corresponding to the random inputs and copy it to new input and output vectors; this will be our identification data.

2. Plot and examine the data. Determine whether the identification 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`.

3. 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 and illustrations from the lecture. Hint: By checking the structure of the linear system, it will become clear that you only need to compute the covariances for positive τ.

4. Implement and solve the system of linear equations to obtain the FIR model for any value of M supported by the data. Again, use the formulas and illustrations from the lecture. Hint: Make sure you understand the structure of the regressor matrix in the linear system, with the values of r_u "reflected" around $\tau = 0$ on the main diagonal.

5. Use convolution to simulate the model response to the whole set of (identification and validation) inputs, and compare the model response to the whole set of true system outputs. Keep in mind you should qualitatively investigate the fit on the identification data, but the true test of the model is if it performs well on the validation data. Note: the time vector used for plotting the convolution result

[1] Presumably you already did this for the previous labs, but if you did not, prior to starting the lab familiarize yourself with the DC motor system and the ways in which input signals can be applied and output signals can be read, using the guide linked from the website.

should be equidistant, unlike the real time vector obtained from the system, in which the sampling instants are not fully equidistant. You can use the Matlab function `conv` for this step (although it is not required).

6. Study the influence of the length M of the FIR model on the accuracy of the model, and try to find a good value for M. A rule of thumb for selecting such a good value 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 trade them off against each other to get a reasonable fit given the amount of data you have.