ME 418/518 – Data-Based Control

Problem Set 2

Deadline – October 27th, 2025

Employment of LLMs is strictly prohibited

Problem 1. In this problem, we will first generate our synthetic data from a known plant. Then, using the input and output sequences we will estimate the state space model of this plant. Finally, we will validate our model using the synthetic data we produced and the data generated using the estimated model.

Consider a discrete time plant dynamics with a sampling interval Ts=0.02s, where

$$A = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.2 \end{bmatrix},\ B = \begin{bmatrix} 1 \\ 3 \end{bmatrix},\ C = [1 \quad -0.6],\ D = 0.$$

We provide a pseudo-random binary sequence (PRBS) input with a range [-1 1] and dwell time of 10 samples. The number of data points (input-output) is N=3000. The system picks a measurement noise $v_k \sim \mathcal{N}(0, \sigma^2)$, which means a zero mean Gaussian noise with a standard deviation of σ.

1.a) Create 3000 synthetic data points (input output pairs) by simulating the system above. Plot the first 15 seconds of input and the output. (provide the matlab code)

1.b) Is the choice of input given above suitable for the given problem?

1.c) Choose horizons (input-output depths) of i=j=30 and form the Hankel Matrices you will need for the implementation of the N4SID algorithm. (Explain what matrices you need, how you find the dimensions and then provide the matlab code)

1.d) Calculate the projector operator P using the matrix $W = \begin{bmatrix} U_p \\ U_f \end{bmatrix}$. (provide the matlab code)

1.e) Using SVD of the <u>projected</u> Y_f and find the number of significant singular values. What does this number mean? (provide the matlab code)

1.f) Find the estimated extended observability matrix and the state matrix. (show the matlab code).

1.g) Find $\hat{C}$ and $\hat{A}$. (show the matlab code)

1.h) For each column in Y_f, with boundary time $t_l = i + l$, propogate its boundary state x_{t_l} to the end of the window until $x_{t_{l+j-1}}$. This means, calculate $\hat{x}_{k|l} = \hat{A}^{k-t_l}\hat{x}_{t_l}, k = t_l, t_{l+1}, \dots, t_{l+j-1}$.

Then, for each estimated $\hat{x}_k$, average all estimates: $\frac{1}{|L(k)|} \sum_{l \in L(k)} \hat{x}_{k|l}$, where L(k) is the set of Hankel column indices whose time horizon (depth) contains k. $|L(k)|$ is the number of overlapping horizons that contain time k. Plot the first 15 seconds of the state vector elements for 15 seconds. We will use 15 seconds as our “estimation interval”. (show matlab code)

1.i) Estimate $\hat{B}$ and $\hat{D}$, using the data obtained above.

1.j) Calculate FIT% at the estimation interval. Calculate FIT% at the validation interval.

1.k) Use matlab function n4sid to create your state space model (using the same estimation segment) and report FIT%. Compare this with you manually crafted N4SID results.