Homework 7

This homework is due on Friday, March 10, 2023, at 11:59PM. Self-grades and HW re-submissions are due on the following Friday, March 17, 2023, at 11:59PM.

1. **System Identification**

You are given a discrete-time system as a black box. You don't know the specifics of the system but you know that it takes one scalar input and has two states that you can observe. You assume that the system is linear and of the form

$$\vec{x}[i+1] = A\vec{x}[i] + Bu[i] + \vec{w}[i], \tag{1}$$

where $\vec{w}[i]$ is an external small unknown disturbance, $u[i]$ is a scalar input, and

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad x[i] = \begin{bmatrix} x_1[i] \\ x_2[i] \end{bmatrix}. \tag{2}$$

You want to identify the system parameters (a_1, a_2, a_3, a_4, b_1 and b_2) from measured data. However, you can only interact with the system via a black box model, i.e., you can see the states $\vec{x}[t]$ and set the inputs $u[i]$ that allow the system to move to the next state.

(a) You observe that the system has state $\vec{x}[i] = \begin{bmatrix} x_1[i] & x_2[i] \end{bmatrix}^\top$ at time i. You pass input $u[i]$ into the black box and observe the next state of the system: $\vec{x}[i+1] = \begin{bmatrix} x_1[i+1] & x_2[i+1] \end{bmatrix}^\top$.
Write scalar equations for the new states, $x_1[i+1]$ **and** $x_2[i+1]$. Write these equations in terms of the a_i, b_i, the states $x_1[i]$, $x_2[i]$ and the input $u[i]$. Here, assume that $\vec{w}[i] = \vec{0}$ (i.e., the model is perfect).

(b) Now we want to identify the system parameters. We observe the system at the start state $\vec{x}[0] = \begin{bmatrix} x_1[0] \\ x_2[0] \end{bmatrix}$. We can then input $u[0]$ and observe the next state $\vec{x}[1] = \begin{bmatrix} x_1[1] \\ x_2[1] \end{bmatrix}$. We can continue this for a sequence of ℓ inputs.
Let us define an ℓ-length trajectory to be an initial condition $\vec{x}[0]$, an input sequence $u[0], \ldots, u[\ell-1]$, and the corresponding states that are produced by the system $x[1], \ldots, x[\ell]$. **Assuming that the model is perfect ($\vec{w}[i] = \vec{0}$), what is the minimum value of ℓ you need to identify the system parameters?**

(c) We now remove our assumption that $\vec{w} = 0$. We assume it is small, so the model is approximately correct and we have

$$\vec{x}[i+1] \approx A\vec{x}[i] + Bu[i]. \tag{3}$$

Say we feed in a total of 4 inputs $u[0], \ldots, u[3]$, and observe the states $\vec{x}[0], \ldots, \vec{x}[4]$. To identify the system we need to set up an approximate (because of potential, small, disturbances) matrix equation

$$DP \approx S \tag{4}$$

using the observed values above and the unknown parameters we want to find. Let our parameter vector be

$$P := \begin{bmatrix} \vec{p}_1 & \vec{p}_2 \end{bmatrix} = \begin{bmatrix} a_1 & a_3 \\ a_2 & a_4 \\ b_1 & b_2 \end{bmatrix} \tag{5}$$

Find the corresponding D and S to do system identification. Write both out explicitly.

(d) Now that we have set up $DP \approx S$, we can estimate a_0, a_1, a_2, a_3, b_0, and b_1. **Give an expression for the estimates of $\vec{p}_1$ and $\vec{p}_2$ (which are denoted $\hat{\vec{p}}_1$ and $\hat{\vec{p}}_2$ respectively) in terms of D and S.** Denote the columns of S as $\vec{s}_1$ and $\vec{s}_2$, so we have $S = [\vec{s}_1\ \vec{s}_2]$. Assume that the columns of D are linearly independent. *(HINT: Don't forget that D is not a square matrix. It is taller than it is wide.) (HINT: Can we split $DP = S$ into separate equations for p_1 and p_2?)*

2. Identifying systems from their responses to known inputs

In many problems, we have an unknown system, and would like to characterize it. One of the ways of doing so is to observe the system response with different initial conditions (or inputs). This problem is also called system identification. It is a prototypical example of a problem that today is called machine learning — inferring an underlying pattern from data, and doing so well enough to be able to exploit that pattern in some practical setting. Go through the attached Jupyter notebook `demo_system_id.ipynb` and answer the following questions.

(a) In Example 2, we assume that instead of measuring the state $\vec{x}$, we are instead measuring a transformation of the state $\vec{y} = T\vec{x}$ where T is a full rank matrix. Assume that we perform system ID on our observations $\vec{y}[i]$ to recover A_y, B_y such that $\vec{y}[i+1] = A_y\vec{y}[i] + B_y u[i]$. **How do the identified A_y and B_y matrices relate to the original A and B matrices in the dynamics of $\vec{x}$?** Remember that our original state dynamics are $\vec{x}[i+1] = A\vec{x}[i] + B\vec{u}[i]$.

Hint: The answer is given in the Jupyter notebook but remember to show your work.

(b) **Please share your observations on Example 2. Comment on what impact a linear transformation of the state trace has on our ability to perform system identification.**

(c) **Prove that for any full rank transformation matrix T, the eigenvalues of A_y and A from part (a) are the same.**

(d) **Please share your observations on Example 3. Comment on the impact that changing the noise magnitude, number of samples and number of states has on the system identification performance.**

(e) **Please share your observations on Example 4. Comment on the sample efficiency of this method, i.e. do you need more or less samples for accurate system identification when given scalar observations rather than the entire state vector?**

(f) **Please share your observations on Example 5. Comment on how important the model size is for this setting.**

Contributors:

- Nikhil Shinde.
- Ashwin Vangipuram.
- Sally Hui.
- Alex Devonport.
- Tanmay Gautam.

