Homework 9

This homework is due on Friday, October 28, 2022, at 11:59PM. Self-grades and HW Resubmissions are due on the following Friday, November 4, 2022, at 11:59PM.

1. **(OPTIONAL) Mid-Semester Survey**

   Please fill out this mid-semester survey to let us know how the class has been going so far! This survey is optional and anonymous, but you can submit a screenshot of the final page of the survey to Gradescope to receive 2 global extra credit points! We will be accepting submissions on Gradescope until Sunday, October 30 at 11:59pm.

2. Change of Basis

(a) For any given vector, we have to choose a basis to write this vector in. Typically, we choose the standard basis $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ where $\vec{e}_i$ is a vector with a 1 in the ith position and zeros everywhere else. **Given a vector** $\vec{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$**, write** $\vec{x}$ **as a linear combination of standard basis vectors.**

(b) We can also represent the same vector $\vec{x}$ in a different basis. Let us write this new basis as $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. **Find a way to write** $\vec{x}$ **from the previous subpart as a linear combination of** $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$**.** Simplify your answer as an equation with matrix-vector multiplication, and assume that $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are linearly independent.

(HINT: One representation of $\vec{x}$ is the one you determined in the previous subpart. Another representation of $\vec{x}$ is $\widetilde{\alpha}_1 \vec{v}_1 + \widetilde{\alpha}_2 \vec{v}_2 + \cdots + \widetilde{\alpha}_n \vec{v}_n$. We need these two representations to be algebraically equal to indicate that they both represent the same vector. For your convenience, you may define $\vec{\widetilde{\alpha}} = \begin{bmatrix} \widetilde{\alpha}_1 \\ \widetilde{\alpha}_2 \\ \vdots \\ \widetilde{\alpha}_n \end{bmatrix}$.)

(c) Suppose that we truncated our basis so that we now only have $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ where $m < n$ linearly independent vectors, but we could still represent $\vec{x}$ as a linear combination of these vectors. **How do you modify your method from the previous part?** You may not assume that you know $\vec{v}_{m+1}, \ldots, \vec{v}_n$.

(HINT: Think about using projections. Specifically, consider projecting onto the column space of a matrix that you define.)

(d) Suppose that all the vectors $\vec{v}_i$ from the previous part were orthonormal. **Simplify your answer from the previous subpart under this assumption.**

(HINT: Let $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \end{bmatrix} \in \mathbb{R}^{n \times m}$ where $n > m$. If $S = U^\top U$, then $S_{ij} = \vec{u}_i^\top \vec{u}_j$.)

3. Cayley-Hamilton and Controllability Matrix

(a) We can define the *characteristic polynomial* of a matrix $A \in \mathbb{R}^{n\times n}$ as

$$p_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0\lambda^0 \tag{1}$$

where each $c_i \in \mathbb{R}$ is a constant. The characteristic polynomial has roots that are the eigenvalues of A. That is, we can equivalently define

$$p_A(\lambda) = \det\{\lambda I - A\} \tag{2}$$

We say that any of the eigenvalues of A "satisfy" the characteristic polynomial in that

$$p_A(\lambda_i) = 0 \tag{3}$$

where λ_i is the ith eigenvalue of A. Now, let A be a diagonalizable matrix, where we may write $A = V\Lambda V^{-1}$. **Prove that A satisfies its own characteristic polynomial.** In other words, prove that $p_A(A) = 0_{n\times n}$, where $0_{n\times n}$ is a $n \times n$ matrix of zeros.

(HINT: It is not correct to simply plug in $\lambda = A$ into $\det\{\lambda I - A\}$.)

(b) Now, consider some vector $\vec{b} \in \mathbb{R}^n$. **Using the result from the previous part, show that $A^n\vec{b}$ is linearly dependent on $A^{n-1}\vec{b}, A^{n-2}\vec{b}, \ldots, A\vec{b}, \vec{b}$.**

(c) Instead of setting $\vec{b}$ to be a vector, let it be a matrix $B \in \mathbb{R}^{n\times m}$. **Now, show that the columns of A^nB are linearly dependent on the columns of $A^{n-1}B, A^{n-2}B, \ldots, AB, B$.**

(HINT: If we were to write $B = \begin{bmatrix}\vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_m\end{bmatrix}$ where each column is n-dimensional, we can write $A^iB = \begin{bmatrix}A^i\vec{b}_1 & A^i\vec{b}_2 & \cdots & A^i\vec{b}_m\end{bmatrix}$. Make sure you convince yourself of this.)

(d) Consider a discrete time system of the form

$$\vec{x}[i+1] = A\vec{x}[i] + B\vec{u}[i] \tag{4}$$

where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. The controllability matrix for this discrete time system is given by

$$\mathcal{C} = \begin{bmatrix}A^{n-1}B & A^{n-2}B & \cdots & AB & B\end{bmatrix} \tag{5}$$

Conclude that the rank of your controllability matrix will not change if, instead, you made your controllability matrix $\begin{bmatrix}A^nB & A^{n-1}B & \cdots & AB & B\end{bmatrix}$ (i.e., you prepended A^nB to your original controllability matrix).

4. CCF Transformation and Controllability

(a) Consider the following discrete time system

$$\vec{x}[i+1] = A\vec{x}[i] + B\vec{u}[i] \tag{6}$$

Suppose we define a change of basis operation given by $M\vec{z}[i] = \vec{x}[i] \iff \vec{z}[i] = M^{-1}\vec{x}[i]$. This yields a new discrete time system of the form

$$\vec{z}[i+1] = \widetilde{A}\vec{z}[i] + \widetilde{B}\vec{u}[i] \tag{7}$$

for some $\widetilde{A}$ and $\widetilde{B}$ defined in terms of M, A, and B. **What is the controllability matrix for the system in eq. (7), in terms of M, A, and B?**

(b) Consider the change of basis given by $\vec{z}[i] = T^{-1}\vec{x}[i]$ where, under this change of basis transformation, we have the following discrete time system

$$\vec{z}[i+1] = A_{\text{CCF}}\vec{z}[i] + B_{\text{CCF}}\vec{u}[i] \tag{8}$$

Using the result from the previous part, determine an expression for T in terms of $\mathcal{C}$, the controllability matrix of the original system in eq. (6), and $\mathcal{C}_{\text{CCF}}$, the controllability matrix of the system in eq. (8).

(c) We know that the controllability matrix for a system in CCF will always be full rank. **Using this, prove that you can find a transformation matrix T as in the previous part if and only if your original system is controllable.** *(HINT: To prove this, you can first show that, if such a T exists, then your original system is controllable. Then, you can show that, if your original system is controllable, there will exist such a transformation matrix T.) (HINT: Recall that T must be invertible (equivalently, full rank) in order for it to be a valid transformation matrix. You may use without proof the fact that* $\text{rank}(AB) = \min(\text{rank}(A), \text{rank}(B))$*.)*

(d) Consider the following discrete-time dynamics model:

$$\vec{x}[i+1] = \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{A} \vec{x}[i] + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\vec{b}} \vec{u}[i] \tag{9}$$

Find the transformation matrix T such that the dynamics model for $\vec{z}[i] = T^{-1}\vec{x}[i]$ is in CCF. You may use a calculator/computer to perform any computations, if you wish.

(HINT: First, find the characteristic polynomial of A. Use this to determine what A_{CCF} and $\vec{b}_{\text{CCF}}$ should be, and then use this to determined $\mathcal{C}_{\text{CCF}}$.)

5. QR System ID

(a) Suppose we are given the following discrete time dynamical system:

$$x[i+1] = ax[i] + b_1 u_1[i] + b_2 u_2[i] + \cdots + b_{n-1} u_{n-1}[i] \tag{10}$$

We would like to estimate $a, b_1, b_2, \ldots, b_{n-1}$ using system ID. Suppose we have collected data up to $x[m]$, where $m < n$. **Set up a linear system of the form $D\vec{p} = \vec{s}$ to solve this system ID problem. Show that D has dimensions $m \times n$.**

(b) As we saw in the previous part, we have a wide matrix D. Assuming that D is rank m, we would technically have infinitely many solutions for $a, b_1, b_2, \ldots, b_{n-1}$. We can find the solution with the smallest norm using QR decomposition.

We can write $D^\top = \begin{bmatrix} \vec{d}_1 & \vec{d}_2 & \cdots & \vec{d}_m \end{bmatrix}$ where each $\vec{d}_i \in \mathbb{R}^n$. We can also define an orthonormal matrix $Q \in \mathbb{R}^{n \times n}$ which can be written as $Q = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_m & \vec{q}_{m+1} & \vec{q}_{m+2} & \cdots & \vec{q}_n \end{bmatrix}$, where $\text{Span}(\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_m) = \text{Span}\left(\vec{d}_1, \vec{d}_2, \ldots, \vec{d}_m\right)$. **In this case, what is $\vec{d}_j^\top \vec{q}_i$ for $j \in \{1, \ldots, m\}$ and $i \in \{m+1, \ldots, n\}$? Explain your answer.**

(HINT: If we say that $\text{Span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n) = \text{Span}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$, then we may say that $\vec{v}_i$ can be written as a linear combination of $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ (and equivalently, $\vec{u}_i$ can be written as a linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$).)

(c) Suppose that $D^\top$ can be written as

$$D^\top = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_m \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ 0 & r_{22} & \cdots & r_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{mm} \end{bmatrix} \tag{11}$$

Using this result and the result from the previous part, show that the QR decomposition of $D^\top$ can be written as

$$D^\top = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0_{(n-m)\times m} \end{bmatrix} \tag{12}$$

Using eq. (12), write an expression for $Q_1^\top \vec{p}$ where $D\vec{p} = \vec{s}$, and show that the value of $Q_2^\top \vec{p}$ does not matter. Here, $R_1 \in \mathbb{R}^{m \times m}$ is a square, upper triangular matrix, $\begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \in \mathbb{R}^{n \times n}$ is an orthonormal matrix, and $0_{(n-m)\times m} \in \mathbb{R}^{(n-m)\times m}$ is a matrix of all zeros. Q_1 is $n \times m$ and Q_2 is $n \times (n-m)$. Note that R_1 is invertible.

(HINT: Equation (12) uses block matrix form. When multiplying block matrices, they obey the same rules as regular matrix-vector multiplication. That is, $\begin{bmatrix} M & N \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = MA + NB$. When transposing block matrices, we may write $\begin{bmatrix} A \\ B \end{bmatrix}^\top = \begin{bmatrix} A^\top & B^\top \end{bmatrix}$.) (HINT: First, simplify eq. (12) using the previous hint. Then, use the previous problem to find a potential candidate for Q_2. Use the previous part again to confirm that this candidate would work by computing R_{ij} using the formula provided in lecture (for $j \in \{1, \ldots, m\}$ and $i \in \{m+1, \ldots, n\}$).)

(d) From the previous part, we determined that the value of $Q_2^\top \vec{p}$ did not matter. Hence, we can set $Q_2^\top \vec{p} = \vec{0}$ for the purposes of minimizing $\|\vec{p}\|$ (the reason why we do this will be covered a little bit later, but take this as a given for now). **Solve for $\vec{p}$ using the QR decomposition of $D^\top$, assuming $Q_2^\top \vec{p} = \vec{0}$.** *(HINT: The following identity holds true:* $\begin{bmatrix} A\vec{x} \\ B\vec{x} \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix} \vec{x}$*.) (HINT: Stack the two expressions for $Q_1^\top \vec{p}$ and $Q_2^\top \vec{p}$ to obtain an expression for* $\begin{bmatrix} Q_1^\top \vec{p} \\ Q_2^\top \vec{p} \end{bmatrix}$*. Use the previous hint to determine your final expression for $\vec{p}$.)*

Contributors:
- Anish Muthali.