EECS 16B Designing Information Devices and Systems II
Spring 2021 Discussion Worksheet **Discussion 10B**

1. Towards upper-triangulation by an orthonormal basis

This problem is a continuation of problem 1 from Discussion 10A.

Recall that in the last discussion we set out to show that any matrix M can be upper triangularized. In particular we want to find the coordinate transformation U such that M becomes upper triangular when represented in this coordinate system:

$$T = U^{-1}MU. \tag{1}$$

In the previous discussion we began with the example of a 3x3 matrix M. We first constructed U by extending the first eigenvector of M, $\vec{v}_1$, into an orthonormal basis using Gram-Schmidt:

$$U = \begin{bmatrix} \vec{v}_1 & R \end{bmatrix}. \tag{2}$$

This gave us the transformed matrix:

$$T = \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix}, \tag{3}$$

which is upper triangular if Q is upper triangular. Then we realized that we could do a similar transformation on Q to get

$$Q = \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix} \begin{bmatrix} \lambda_2 & \vec{b}^\top \\ \vec{0} & P \end{bmatrix} \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}^\top \tag{4}$$

where $\vec{v}_2$ is the first eigenvector of matrix Q and corresponds to λ_2. At this point, since Q was a 2x2 matrix, we can see that the transformed Q is upper triangular. We then plugged this Q in for M, and simplified to get the final result:

$$M = \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix} \begin{bmatrix} \lambda_1 & \breve{a}_1 & \breve{\vec{a}}_{\text{rest}} \\ 0 & \lambda_2 & \vec{b} \\ 0 & 0 & P \end{bmatrix} \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}^\top. \tag{5}$$

(a) Show that the matrix $\begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}$ is still orthonormal.

Answer: We originally constructed the columns of R to be orthogonal to $\vec{v}_1$, as they were produced by the Gram-Schmidt algorithm. Thus $\vec{v}_1^\top R\vec{v}_2 = 0$ and $\vec{v}_1^\top RY = \vec{0}$. As for the orthogonality of $R\vec{v}_2$ and RY, we can see that

$$(R\vec{v}_2)^\top RY = \vec{v}_2^\top R^\top RY = \vec{v}_2^\top Y = \vec{0} \tag{6}$$

for the reason that $\vec{v}_2$ and the columns of Y were constructed to be orthogonal.

To check for normality, we can consider the inner products of each element with itself:

$$\vec{v}_1^\top \vec{v}_1 = 1 \tag{7}$$
$$(R\vec{v}_2)^\top R\vec{v}_2 = \vec{v}_2^\top R^\top R\vec{v}_2 \tag{8}$$
$$= \vec{v}_2^\top \vec{v}_2 = 1 \tag{9}$$
$$(RY)^\top RY = Y^\top R^\top RY \tag{10}$$
$$= Y^\top Y = I. \tag{11}$$

Note that the final calculation also assures us that the columns of RY has orthonormal columns.

(b) We have shown how to upper triangularize a 3×3 and a 2×2 matrix. How can we generalize this process to any $n \times n$ matrix M?

Answer: For any $n \times n$ matrix $M = M_n$, we can decompose it into:

$$M_n = \begin{bmatrix} \vec{v_1} & R_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{a} \\ \vec{0} & M_{n-1} \end{bmatrix} \begin{bmatrix} \vec{v_1} & R_n \end{bmatrix}^\top \tag{12}$$
$$= U_n A_n U_n^\top, \tag{13}$$

where M_{n-1} is an $(n-1) \times (n-1)$ matrix.

We can recursively repeat this process on the submatrices M_i finding corresponding the U_i's until we've reached the M_2, the 2×2 case. Then we can combine these transformations from the bottom up, just like we did for the 3×3 case, until we construct our final basis $U_{n,\text{final}}$:

$$U_{i,\text{final}} = \begin{bmatrix} \vec{v}_{n-i+1} & R_i U_{i-1,\text{final}} \end{bmatrix} \tag{14}$$

Once we have our final basis $U = U_{n,\text{final}}$, we can transform into M into this basis to get our upper-triangular matrix T:

$$M = UTU^\top \tag{15}$$
$$T = U^\top MU. \tag{16}$$

(c) Show that the characteristic polynomial of square matrix M is the same as that of the square matrix UMU^{-1} for any invertible U.

Answer: The characteristic polynomial of the matrix M is given by $\det(M - \lambda I)$. Similarly the characteristic polynomial of UMU^{-1} is given by $\det\left(UMU^{-1} - \lambda I\right)$. Thus

$$\det\left(UMU^{-1} - \lambda I\right) = \det\left(UMU^{-1} - \lambda UU^{-1}\right) \tag{17}$$
$$= \det\left(U(M - \lambda I)U^{-1}\right) \tag{18}$$
$$= \det(U)\det(M - \lambda I)\det\left(U^{-1}\right). \tag{19}$$

Recognizing that $\det(U) \cdot \det\left(U^{-1}\right) = 1$ we can simplify eq. (19) to:

$$\implies \det\left(UMU^{-1} - \lambda I\right) = \det(M - \lambda I). \tag{20}$$

Thus the characteristic polynomials of M and UMU^{-1} are the same for square matrices M and U where U is invertible.

2. Minimum Energy Control

In this question, we build up an understanding for how to get the minimum energy control signal to go from one state to another

(a) Consider the scalar system:

$$x[k+1] = ax[k] + bu[k] \tag{21}$$

where $x[0] = 0$ is the initial condition and $u[k]$ is the control input we get to apply based on the current state. Consider if we want to reach a certain state, at a certain time, namely $x[K]$. Write a matrix equation for how a choice of values of $u[k]$ for $k \in \{0, 1, \dots K-1\}$ will determine the output at time K.

[Hint: write out all the inputs as a vector $\begin{bmatrix} u[0] \ u[1] \ \cdots \ u[K-2] \ u[K-1] \end{bmatrix}^T$ and figure out the combination of a and b that gives you the state at time K.]

Answer: The equation to solve for the inputs should be set up as follows:

$$x[K] = \begin{bmatrix} b \ ab \ \cdots \ a^{K-2}b \ a^{K-1}b \end{bmatrix} \cdot \begin{bmatrix} u[K-1] \\ u[K-2] \\ \vdots \\ u[1] \\ u[0] \end{bmatrix}$$

(b) Consider the scalar system:

$$x[k+1] = 1.0x[k] + 0.7u[k] \tag{22}$$

where $x[0] = 0$ is the initial condition and $u[k]$ is the control input we get to apply based on the current state. Suppose if we want to reach a certain state, at a certain time, namely $x[K] = 14$. With our dynamics $a = 1$, solve for the best way to get to a specific state $x[K] = 14$, when $K = 10$. When we say **best way** to control a system, we want the sum squared of the inputs to be minimized

$$\underset{u[k]}{\operatorname{argmin}} \sum_{k=0}^{K} u[k]^2.$$

Hint: recall the Cauchy-Schwarz inequality $\langle \vec{a}, \vec{b} \rangle \leq \|\vec{a}\| \|\vec{b}\|$ where equality holds if $\vec{a}$ and $\vec{b}$ are linearly dependent.

Answer: Starting from the previous part, we have

$$x[K] = 0.7u[K-1] + 0.7u[K-2] + \cdots + 0.7u[1] + 0.7u[0].$$

Then, we can plug in what we know about $x[K]$ to get

$$14 = 0.7u[K-1] + 0.7u[K-2] + \cdots + 0.7u[1] + 0.7u[0].$$

We can rewrite this as

$$14 = \langle \vec{u}, \vec{v} \rangle,$$

where $\vec{v} = \begin{bmatrix} 0.7 & \cdots & 0.7 \end{bmatrix}^\top$.

Cauchy-Schwarz tells us that $\langle \vec{u}, \vec{v} \rangle \leq \|\vec{u}\|\|\vec{v}\|$ with equality holding if $\vec{u}$ and $\vec{v}$ are linearly dependent. Thus, we can see the minimum norm solution will be a constant input, so we have

$$20 = u[K-1] + u[K-2] + \cdots + u[1] + u[0].$$

Set $u[t] = \bar{u} \quad \forall\, t < K$ to get

$$20 = K \cdot \bar{u}, \qquad \bar{u} = \frac{20}{K} = 2.$$

This gives us a solution of $u[t] = 2 \quad \forall t$. This can essentially be interpreted as pushing our state by two each timestep.

(c) Consider the scalar system:

$$x[t+1] = 0.5x[t] + 0.7u[t] \tag{23}$$

where $x[0] = 0$ is the initial condition and $u[t]$ is the control input we get to apply based on the current state. Consider if we want to reach a certain state, at a certain time, namely $x[K] = 14$, when $K = 10$. Explain in words the trend of the control input that will be used to solve this problem.

Answer: Similar to the previous part, we begin by writing $x[K]$ in terms of the inputs $u[0] \cdots u[K-1]$:

$$x[K] = \begin{bmatrix} 0.7 & 0.7 \cdot 0.5 & \cdots & 0.7 \cdot 0.5^{K-1} & 0.7 \cdot 0.5^{K-1} \end{bmatrix} \begin{bmatrix} u[K-1] \\ u[K-2] \\ \vdots \\ u[1] \\ u[0] \end{bmatrix} = \langle \vec{v}, \vec{u} \rangle. \tag{24}$$

By Cauchy-Schwarz, we know that the minimum norm solution $\vec{u}$ will be one that is linearly dependent to $\vec{v}$. Thus:

$$\begin{aligned} \vec{u} &= \alpha \vec{v} && (25) \\ x[K] = 14 &= \langle \vec{v}, \vec{u} \rangle && (26) \\ &= \alpha \langle \vec{v}, \vec{v} \rangle && (27) \\ &= \alpha \|\vec{v}\|^2 && (28) \\ &\approx \alpha \cdot 0.808 && (29) \\ \alpha &\approx 17.32 \vec{u} = 17.32 \cdot \vec{v}. && (30) \end{aligned}$$

This gives us a final solution of $x[t] = 12.12 \cdot 0.5^t \quad \forall t \in \{0 \ldots K-1\}$.

Contributors:

- Yuxun Zhou.
- Edward Wang.
- Anant Sahai.
- Sanjit Batra.
- Pavan Bhargava.
- Kareem Ahmad.
- Nathan Lambert.