1. Towards upper-triangulation by an orthonormal basis

In lecture, we have been motivated by the goal of getting to a coordinate system in which the eigenvalues of a matrix representing a linear operation are on the diagonal. When this is done to the $A$ matrix representing a dynamical system (whether in continuous-time as a system of differential equations or in discrete-time as a relationship between the next state and the previous one), we can view the system as a cascade of scalar systems — with each one potentially being an input to the ones that come “after” it. We saw this in lecture, but it is good to spend more time to really understand this argument.

Note that in the next homework, you will be asked to derive this in a more formal way using induction. Here we will just provide some key steps along the way to a recursive understanding. Here, as in lecture, we will restrict attention to matrices that have all real eigenvalues.

In order for you to better understand the steps, you can consider a concrete case

$$S_{3\times 3} = \begin{bmatrix}
5 & 5 & 1 \\
12 & 12 & 12 \\
1 & 6 & 2
\end{bmatrix}$$

and figure out the general case by abstracting variables. This particular matrix has an additional special property of symmetry, but we won’t be invoking that here.

(a) Consider a non-zero vector $\vec{u}_0 \in \mathbb{R}^n$. **Can you think of a way to extend it to a set of basis vectors for $\mathbb{R}^n$?** In other words, find $\vec{u}_1, \cdots, \vec{u}_{n-1}$, such that span$(\vec{u}_0, \vec{u}_1, \cdots, \vec{u}_{n-1}) = \mathbb{R}^n$. To begin with, consider

$$\begin{bmatrix}
1 \\
-1 \\
0
\end{bmatrix}.$$ **Can you get an orthonormal basis from what you just constructed?**

**Answer:** Starting with the provided vector, add new vectors from the standard basis so that every row has a non-zero value along the diagonal. By doing this, we guarantee that the matrix spans $\mathbb{R}^n$. For $[1, -1, 0]^T$, we can do

$$\begin{bmatrix}
1 & 0 & 0 & 1 \\
-1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}$$

Then using this matrix, we can run Gram-Schmidt to convert it to an orthonormal basis. If you get a zero along the way, then discard the vector and move on. You are guaranteed to span the whole space by the end because the standard basis spans the whole space.

Using the Gram-Schmidt process for the basis obtained above,

$$\begin{bmatrix}
1 & 0 & 0 & 1 \\
-1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix} \Rightarrow \begin{bmatrix}
\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\
-\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\
0 & 0 & 1
\end{bmatrix}$$
(b) Now consider a real eigenvalue \( \lambda_0 \), and the corresponding eigenvector \( \vec{g}_0 \in \mathbb{R}^n \) of a square matrix \( M \in \mathbb{R}^{n \times n} \). From the previous part, we can extend \( \vec{g}_0 \) to an orthonormal basis of \( \mathbb{R}^n \), denoted by

\[
V = [\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_{n-1}]
\]

where \( \vec{v}_0 = \frac{\vec{g}_0}{\|\vec{g}_0\|} \).

Our goal is to look at what the matrix \( M \) looks like in the coordinate system defined by the basis \( V \).

**Compute** \( V^T M V \) by writing \( V = [\vec{v}_0, R] \), where \( R \triangleq [\vec{v}_1, \cdots, \vec{v}_{n-1}] \). If you prefer, you can do this and the next question with the concrete \( S_{[3 \times 3]} \) first.

**Answer:**

\[
V^T M V = \begin{bmatrix} \vec{v}_0^T \\ R^T \end{bmatrix} M [\vec{v}_0, R] = \begin{bmatrix} \vec{v}_0^T \\ R^T \end{bmatrix} [\lambda_0 \vec{v}_0, MR] = \begin{bmatrix} \lambda_0 \vec{v}_0^T \vec{v}_0 & \vec{v}_0^T M R \\ \lambda_0 R^T \vec{v}_0 & R^T M R \end{bmatrix}
\]

Concrete case: \( S_{[3 \times 3]} \) has zero as eigenvalue since it contains a repeated column vector, let the corresponding eigenvector be just \( [1, -1, 0]^T \), then we have

\[
V = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad R = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & 0 \\ 0 & 1 \end{bmatrix}
\]

Simple calculation yields

\[
V^T S_{[3 \times 3]} V = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{5}{6} & \frac{\sqrt{2}}{6} \\ 0 & \frac{\sqrt{2}}{6} & \frac{2}{3} \end{bmatrix}
\]

obviously \( Q = R^T S_{[3 \times 3]} R = \begin{bmatrix} \frac{5}{6} & \frac{\sqrt{2}}{6} \\ \frac{\sqrt{2}}{6} & \frac{2}{3} \end{bmatrix} \)

(c) Show that \( V^{-1} = V^T \)

**Answer:** Remember that \( V \) is an orthonormal basis. \( V^T V \) performs a dot product between all of the basis vectors. Since the basis vectors are orthogonal to each other, all non-diagonal elements are 0. Since the basis vectors are normalized, the dot product with itself is 1. As a result, \( V^T V = I \), and \( V^{-1} = V^T \)

(d) Define \( Q = R^T M R \). Look at the first column and the first row of \( V^T M V \) and **show that**:

\[
M = V \begin{bmatrix} \lambda_0 & \vec{a}^T \\ \vec{0} & Q \end{bmatrix} V^T
\]

Here the \( \vec{a} \) is just something arbitrary.

**Answer:** We observe that \( \vec{v}_0^T \vec{v}_0 = 1 \), \( R^T \vec{v}_0 = \vec{0} \) and \( \vec{v}_0^T M R = (M \vec{v}_0)^T R = \lambda_0 \vec{v}_0^T R = \vec{0}^T \), due to the orthonormal construction. Hence we have,

\[
V^T M V = \begin{bmatrix} \lambda_0 & \vec{a}^T \\ \vec{0} & Q \end{bmatrix}
\]

and we get desired form by using \( VV^T = I \). Note that in the homework, this will be proved via a different method. In the numerical example, we have \( Q = R^T S_{[3 \times 3]} R = \begin{bmatrix} \frac{5}{6} & \frac{\sqrt{2}}{6} \\ \frac{\sqrt{2}}{6} & \frac{2}{3} \end{bmatrix} \)
(e) Now, we can recurse on $Q$ to get:

$$Q = [\vec{u}_0, Y] \begin{bmatrix} \lambda_1 & \vec{b}^T \\ 0 & P \end{bmatrix} [\vec{u}_0, Y]^T$$

where we have taken $\vec{u}_0 \in \mathbb{R}^{n-1}$, a eigenvector of $Q$, associated with eigenvalue $\lambda_1$. Again $\vec{u}_0$ is extended into an orthonormal basis $[\vec{u}_0, \vec{u}_1, \cdots, \vec{u}_{n-2}]$ of $\mathbb{R}^{n-1}$. We denote $Y \triangleq [\vec{u}_1, \cdots, \vec{u}_{n-2}]$.

**Plug this into $M$ to show that:**

$$M = [\vec{v}_0, R\vec{u}_0, RY] \begin{bmatrix} \lambda_0 & a_1 & \vec{a}^T \\ 0 & \lambda_1 & \vec{b}^T \\ 0 & 0 & P \end{bmatrix} [\vec{v}_0, R\vec{u}_0, RY]^T$$

Again, using the concrete case may help you first.

**Answer:**
From part (d), we know that

$$M = V \begin{bmatrix} \lambda_0 & \vec{a}^T \\ 0 & Q \end{bmatrix} V^T$$

and that

$$V = [\vec{v}_0, R]$$

So we get that,

$$M = V \begin{bmatrix} \lambda_0 & \vec{a}^T \\ 0 & Q \end{bmatrix} V^T = [\vec{v}_0, R] \begin{bmatrix} \lambda_0 & \vec{a}^T \\ 0 & Q \end{bmatrix} \begin{bmatrix} \vec{v}_0^T \\ R^T \end{bmatrix}$$

Since,

$$Q = [\vec{u}_0, Y] \begin{bmatrix} \lambda_1 & \vec{b}^T \\ 0 & P \end{bmatrix} [\vec{u}_0, Y]^T$$

$$\therefore M = [\vec{v}_0, R] \begin{bmatrix} \lambda_0 & \vec{a}^T \\ 0 & Q \end{bmatrix} \begin{bmatrix} \vec{v}_0^T \\ R^T \end{bmatrix} = [\vec{v}_0, R\vec{u}_0, RY] \begin{bmatrix} \lambda_0 & a_1 & \vec{a}^T_{n-1} \\ 0 & \lambda_1 & \vec{b}^T \\ 0 & 0 & P \end{bmatrix} [\vec{v}_0, R\vec{u}_0, RY]^T$$

The numerical one has

$$Q = \begin{bmatrix} \sqrt{\frac{3}{4}} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\frac{3}{4}} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{2} \end{bmatrix}^T$$

(f) **Show that the matrix $[\vec{v}_0, R\vec{u}_0, RY]$ is still orthonormal.**

**Answer:** By construction, we have $\vec{v}_0^T R\vec{u}_0 = 0$, $\vec{v}_0^T RY = 0$ because $\vec{v}_0$ is orthogonal to columns of $R$, and $(R\vec{u}_0)^T RY = \vec{u}_0^T R^T RY = \vec{u}_0^T Y = \vec{0}$ because we constructed $[\vec{u}_0, Y]$ as an orthonormal basis of $\mathbb{R}^{n-1}$. Check for the numerical one, that

$$RU = \begin{bmatrix} \sqrt{\frac{2}{3}} & 0 \\ \frac{\sqrt{6}}{3} & \sqrt{\frac{2}{3}} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\frac{2}{3}} & \frac{\sqrt{6}}{3} \\ -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}$$
which is orthogonal to $\vec{v}_0 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}$ and we finally get

$$S = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}^T$$

(g) Perform the above process recursively - what will you get in the end?

**Answer:** The final matrix would be a scalar, i.e., having dimensions $[1 \times 1]$ ($n = 1$) having an eigenvalue and eigenvector of 1. Although the above recursive process is intuitive and essentially rigorous, it is still not a very formal proof. Consult the homework for how this can be cast as a formal induction proof.

(h) Show that the characteristic polynomial of square matrix $A$ is the same as that of the square matrix $T^{-1}AT$ for any invertible $T$.

**Answer:** Evaluating the determinant, $\det(A - \lambda I)$ yields the characteristic polynomial of $A$. So we must show that the determinants of $A$ and $TAT^{-1}$ evaluate to the same expression.

$$\det(A - \lambda I) = \det(TT^{-1}(A - \lambda I)) = \det(T(A - \lambda I)T^{-1})$$

This is true since determinants are oriented volumes and hence the determinant of a product is the product of determinants. So, one can reorder terms inside the determinant even if the matrices themselves do not commute.

$$\Rightarrow \det(A - \lambda I) = \det(TAT^{-1} - \lambda I)$$

(i) What can you say about the characteristic polynomial $\det(\lambda I - Q)$ of $Q$ in relationship to the characteristic polynomial of the original $M$? Recall that $Q$ is an $(n - 1) \times (n - 1)$ matrix.

**Answer:** The idea here is to see that the characteristic polynomial of $Q$ must be the characteristic polynomial of the original $M$ divided by the factor $(\lambda - \lambda_0)$.

This is because of the structure of the block matrix. The determinant must be the first corner times the determinant of the bottom corner block. The oriented volume interpretation gives that to you nearly instantly, especially via the connection between determinants and the operations of Gaussian Elimination.

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:

$$x(t + 1) = ax(t) + bu(t) \quad (1)$$

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 $x(T)$. **Write a matrix**
equation for how a choice \( u(t) \) will determine the output at time \( T \).

(hint: write out all the inputs as a vector \([u(0) \ u(1) \ \cdots \ u(T-2) \ u(T-1)]^T\) and figure out the combination of \( a \) and \( b \) that gives you the state at time \( T \).)

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

\[ x(T) = [b \ \ a \ \ a^2b \ \ a^3b \ \ \cdots \ \ a^{T-2}b \ \ a^{T-1}b] \cdot \begin{bmatrix} u(T-1) \\ u(T-2) \\ \vdots \\ u(1) \\ u(0) \end{bmatrix} \]

(b) Consider the scalar system:

\[ x(t + 1) = 1.0x(t) + 0.7u(t) \]  

(2)

where \( x(0) = 0 \) is the initial condition and \( u(t) \) 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 \( x(T) = 14 \). **With our dynamics** \( a = 1 \), **solve for the best way to get to a specific state** \( x(T) = 14 \), **when** \( T = 10 \). When we say **best way** to control a system, we want the sum squared of the inputs to be minimized

\[
\arg\min_{u(t)} \sum_{t=0}^{T} u(t)^2.
\]

**Answer:** Starting from the previous part, we have

\[ x(T) = 0.7u(T-1) + 0.7u(T-2) + \cdots + 0.7u(1) + 0.7u(0). \]

Then, we can plug in what we know about \( x(T) \) to get

\[ 14 = 0.7u(T-1) + 0.7u(T-2) + \cdots + 0.7u(1) + 0.7u(0). \]

Then, we can see the minimum norm solution will be a constant input, so we have

\[ 20 = u(T-1) + u(T-2) + \cdots + u(1) + u(0). \]

Set \( u(t) = \bar{u} \ \forall \ t < T \) to get

\[ 20 = T \cdot \bar{u}, \quad \bar{u} = \frac{20}{T}. \]

This gives us a solution of \( u(t) = 2 \ \forall t \). This can be interpreted as pushing by two each time.

(c) Consider the scalar system:

\[ x(t + 1) = 0.5x(t) + 0.7u(t) \]  

(3)

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 \( x(T) = 14 \), when \( T = 10 \). **Explain in words the trend of the control input that will be used to solve this problem**
**Answer:** The equation to solve for the inputs should be set up as follows:

$$x(T) = \begin{bmatrix} 0.7 & 0.5 \cdot 0.7 & \cdots & 0.5^{T-2} \cdot 0.7 & 0.5^{T-1} \cdot 0.7 \end{bmatrix} \cdot \begin{bmatrix} u(T-1) \\ u(T-2) \\ \vdots \\ u(1) \\ u(0) \end{bmatrix}$$

Expanding this relation out, we see that

$$x(T) = 0.7u(T-1) + 0.5 \cdot 0.7u(T-2) + \cdots + 0.5^{T-2} \cdot 0.7u(1) + 0.5^{T-1} \cdot 0.7u(0).$$

This relation shows us that the minimum norm solution will be an approach that more heavily weights the last few actions. Consider the SVD of the matrix

$$\vec{m} = \begin{bmatrix} 0.7 & 0.5 \cdot 0.7 & \cdots & 0.5^{T-2} \cdot 0.7 & 0.5^{T-1} \cdot 0.7 \end{bmatrix}.$$ 

That means, set up $M = U\Sigma V^T$ where $U \in \mathbb{R}^{1 \times 1}$, $\Sigma \in \mathbb{R}^{1 \times 10}$, $V^T \in \mathbb{R}^{10 \times 10}$. In order for $U$ to be orthonormal (definition of SVD), it must be $I$. Next, we construct $V^T$ to be

$$V^T = \begin{bmatrix} \vec{m} \\ \|\vec{m}\| \\ \vec{a}_1 \\ \vdots \\ \vec{a}_9 \end{bmatrix},$$

where $a_i$ is any vector of length 10. Now, in order for the SVD $M = U\Sigma V^T$ to be true, we have

$$\Sigma = \begin{bmatrix} \|\vec{m}\| & 0 & 0 & \cdots & 0 \end{bmatrix}.$$ 

The critical and last step here is how to read off this SVD as the solution to the minimum norm control problem. What has the SVD given us here? It created two orthonormal bases and a stretching matrix as an alternate representation of $\vec{m}$. The SVD gave us the equation

$$x(T) = U\Sigma V^T \vec{u},$$

So in order to solve for the vector $\vec{u}$, we need to use the pseudo inverse

$$\vec{u} = V\Sigma^+ U^T x(T)$$

From here, we know that our solution is going to be the vector that aligns perfectly with the solution in SVD form (See note below on nullspace). We know our solution will have most of the control effort later on, and this mirrors the basis we created ($\|\vec{m}\| = .643$).

$$\vec{u} = 14 \cdot \begin{bmatrix} 0.7 \\ 0.5 \cdot 0.7 \\ \vdots \\ 0.5^{T-2} \cdot 0.7 \\ 0.5^{T-1} \cdot 0.7 \end{bmatrix} = 14 \cdot \begin{bmatrix} 0.7 \\ 0.432 \\ 0.35 \\ \vdots \\ 0.0027 \\ 0.0014 \\ 0.0012 \end{bmatrix}.$$
Here is a brief comment on why our solution aligns perfectly with our one basis vector. Effectively, if our input in the 10 dimension space covered any of the other basis vectors, this control effort would fall into the nullspace. Maybe it would help to view the problem as

\[
x(T) = U \cdot \begin{bmatrix} \|\vec{m}\| & 0 & 0 & \cdots & 0 \\ \vec{a}_1^T & \vdots & \vdots & \ddots & \vdots \\ \vec{a}_9^T & u(T-1) & u(T-2) & \cdots & u(0) \\ \vec{a}_0^T & u(1) & u(0) & \cdots & u(0) \end{bmatrix}
\]

Now, consider any vector \( \vec{v}_1 \), another vector \( \vec{c} \). We can define \( \ell = \vec{v}_1^T \vec{c} \). We want to choose the minimum norm \( \vec{c} \) that solves this equality. Suppose

\[
\vec{c} = \alpha \vec{v}_1 + \beta \vec{v}_\perp.
\]

Then we get the \( \|\vec{c}\| = \sqrt{\alpha^2 + \beta^2} \). Then we want to choose \( \vec{c} \) so that

\[
\alpha = \vec{v}_1^T \vec{c}
\]

This is because we can isolate any part of the solution in the \( \vec{v}_1 \) portion, anything in \( \vec{v}_\perp \) will increase \( \|\vec{c}\| \).

(d) Now, consider the following linear discrete time system

\[
\vec{x}(t+1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t)
\]

Set up the system of equations to calculate the state at time \( T = 20 \).

i Write out the matrices in symbolic form:

**Answer:**

\[
\vec{x}(T) = \begin{bmatrix} \vec{b} & A\vec{b} & \cdots & A^{T-2}\vec{b} & A^{T-1}\vec{b} \end{bmatrix} \cdot \begin{bmatrix} u(T-1) \\ u(T-2) \\ \vdots \\ u(1) \\ u(0) \end{bmatrix}
\]

ii Write out the matrix with powers of \( A \) with numbers:

**Answer:**

\[
\vec{x}(T) = \begin{bmatrix} 0 & 1 & 0 & 1 & \cdots & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & \cdots & 1 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} u(T-1) \\ u(T-2) \\ \vdots \\ u(1) \\ u(0) \end{bmatrix}
\]
(e) What form does the minimum norm solution take in this problem?

**Answer:** If you expand the matrix of 1s and 0s with an SVD as done in part c), one will see that an elegant solution emerges where the odd controls effect $\vec{x}[2]$ and the even controls effect $\vec{x}[1]$. Set up $M = U \Sigma V^T$ where $U \in \mathbb{R}^{2 \times 2}, \Sigma \in \mathbb{R}^{2 \times 20}, V^T \in \mathbb{R}^{20 \times 20}$. Next, we construct $V^T$ to be

$$V^T = \begin{bmatrix} 0 & \frac{1}{\sqrt{10}} & 0 & \cdots & \frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{10}} & 0 \\
\frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{10}} & \cdots & \frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{10}} & 0 \\
\vec{a}_1 \\
\vdots \\
\vec{a}_{18} 
\end{bmatrix}$$

where $a_i$ is any vector of length 20. Now, in order for the SVD $M = U \Sigma V^T$ to be true, we have

$$\Sigma = \begin{bmatrix} \sqrt{10} & 0 & 0 & \cdots & 0 \\
0 & \sqrt{10} & 0 & \cdots & 0 
\end{bmatrix}$$

Finally, we construct $U$ to be orthonormal and satisfy the equality

$$U = \begin{bmatrix} 1 & 0 \\
0 & 1 
\end{bmatrix}$$

Recall that we showed that $\vec{u} = V^T \Sigma^+ U \vec{x}(T)$. Without specifying a target vector $\vec{x}$, the solution will be $\vec{u} = M \vec{x}(T)$, where $M = V^T \Sigma^+ U$. This gives us the following action choice for when $\vec{x}(20) = \begin{bmatrix} 1 \\
1 
\end{bmatrix}$.

$$\vec{u} = \begin{bmatrix} \frac{1}{\sqrt{10}} \\
\frac{1}{\sqrt{10}} \\
\vdots \\
\frac{1}{\sqrt{10}} \\
\frac{1}{\sqrt{10}} 
\end{bmatrix}$$

(f) Repeat part e) with a time horizon of $T = 21$.

**Answer:** Here we will find that the symmetry of the solution from the previous part no longer holds up. This breaks down because when $T = 21$, we get a problem of the form

$$\vec{x}(T) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\
1 & 0 & 1 & \cdots & 1 & 0 & 1 
\end{bmatrix} \begin{bmatrix} u(T-1) \\
\vdots \\
u(1) \\
u(0) 
\end{bmatrix}$$

The symmetry of even and odd inputs no longer exists. Therefore, when an SVD is performed, the ordering of the bases is flipped because the $\Sigma$ matrix is sorted by magnitude of the singular values. Set
up $M = U \Sigma V^T$ where $U \in \mathbb{R}^{2 \times 2}, \Sigma \in \mathbb{R}^{2 \times 21}, V^T \in \mathbb{R}^{21 \times 21}$. Next, we construct $V^T$ to be

$$V^T = \begin{bmatrix} \frac{1}{\sqrt{11}} & 0 & \frac{1}{\sqrt{11}} & \cdots & 0 & \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{11}} \\ 0 & \frac{1}{\sqrt{10}} & 0 & \cdots & \frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{10}} \end{bmatrix}$$

(16)

where $a_i$ is any vector of length 21. Now, in order for the SVD $M = U \Sigma V^T$ to be true, we have

$$\Sigma = \begin{bmatrix} \sqrt{11} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{10} & 0 & \cdots & 0 \end{bmatrix}$$

(17)

Finally, we construct $U$ to be orthonormal and satisfy the equality (See how we lost the identity structure here)

$$U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(18)

Contributors:

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