EECS 16B Designing Information Devices and Systems II
Fall 2021 Discussion Worksheet Discussion 9A

The following notes are useful for this discussion: Note 12, Note 14.

1. Towards Upper-Triangularization By An Orthonormal Basis

Solution: 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, and there are only zeros below the diagonal. When this is done to the A matrix representing a time-evolving 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" or "above" 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. 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.

Previously in this course, we have seen the value of changing our coordinates to be eigenbasis-aligned, because we can then view the system as a set of parallel scalar systems. If we have a diagonalization, then these scalar equations are fully uncoupled, and can therefore be treated completely separately. But even when we cannot diagonalize, we can upper-triangularize in a way that allows us to solve the equations one at a time, from the "bottom up".

In this problem, to better understand the steps involved, we will use the following concrete example:

$$M = S_{[3\times 3]} = \begin{bmatrix} \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix} \tag{1}$$

and figure out the general case by abstracting variables. Note that there is a datahub link to a jupyter notebook on the website, which will allow you to perform the numerical calculations quickly to connect the symbolic analysis to an example, but without being time-consuming in the process.[1]

(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 $\text{span}(\vec{u}_0, \vec{u}_1, \cdots, \vec{u}_{n-1}) = \mathbb{R}^n$. **To make things concrete, consider** $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$**. Can you get an orthonormal basis where the first vector is a multiple of this vector?**

Hint: what was the last discussion all about? Also, the given vector isn't normalized yet!

Solution: Starting with the provided vector, we can include all the vectors from the standard basis (here, since we're in $\mathbb{R}^3$, we will add the $\mathbb{R}^3$ basis vectors) — namely, the columns of the identity matrix. By doing this, we guarantee that the matrix spans $\mathbb{R}^n$ (since the 3 vectors alone that we just

[1] This particular matrix has an additional special property of symmetry, but we won't be invoking that here.

added span $\mathbb{R}^3$, and the initial vector can be treated as "extra" for now.) Of course, to have a valid basis, we ultimately need a *minimal* set of spanning vectors, so we will only have 3 vectors in the end.

For $\begin{bmatrix} 1 & -1 & 0 \end{bmatrix}^\top$, we can form:

$$\begin{bmatrix} | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Then using this matrix (considering the constituent column vectors one at a time in order), we can run Gram-Schmidt (which was covered in great detail in the previous discussion) to convert this matrix to an orthonormal basis. Note that since we are starting with 4 vectors but only need 3, we will end up having to throw one out. But fortunately, Gram-Schmidt will tell us when to do this!

If we ever see a zero vector residual along the way while executing Gram-Schmidt orthonormalization, then we discard that vector and move on. Recall that if we see a zero residual, this indicates that the current vector under consideration is already in the span of the previous vectors in our set of basis vectors (so adding the current vector doesn't add any new degrees of freedom to our span).

The key is that we are guaranteed to span the whole space by the end because the standard basis spans the whole space and Gram-Schmidt guarantees that the final span of our constructed vectors must agree with the span of the set of vectors that we started with. We can use the Gram-Schmidt process for the basis obtained above, starting with $\vec{v}_1$ (using the same notation as dis08B):

$$\vec{q}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} \tag{2}$$

$$\implies \vec{z}_2 = \vec{v}_2 - \left(\vec{q}_1^\top \vec{v}_2\right) \vec{q}_1 \tag{3}$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} \right) \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} \tag{4}$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{\sqrt{2}}{2} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} \tag{5}$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} \tag{6}$$

$$\implies \vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} \tag{7}$$

$$\vec{z}_3 = \vec{v}_3 - \left(\vec{q}_1^\top \vec{v}_3\right) \vec{q}_1 - \left(\vec{q}_2^\top \vec{v}_3\right) \vec{q}_2 = \vec{0} \qquad \text{(unused in final basis)} \tag{8}$$

$$\vec{z}_4 = \vec{v}_4 - \left(\vec{q}_1^\top \vec{v}_4\right) \vec{q}_1 - \left(\vec{q}_2^\top \vec{v}_4\right) \vec{q}_2 \tag{9}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 0 \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} - 0 \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{10}$$

$$\implies \vec{q}_3 = \frac{\vec{z}_4}{\|\vec{z}_4\|} \tag{11}$$

Once we carry out this procedure, we find that our orthonormal matrix consisting of the new basis vectors is:

$$\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12}$$

(b) Now consider a real eigenvalue λ_1, and the corresponding (normalized) eigenvector $\vec{v}_1 \in \mathbb{R}^n$ of a square matrix $M \in \mathbb{R}^{n\times n}$. From the previous part, we know that we can extend $\vec{v}_1$ to an orthonormal basis of $\mathbb{R}^n$. We will denote the basis by

$$U = \begin{bmatrix} | & | & \cdots & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \\ | & | & \cdots & | \end{bmatrix}$$

where $\vec{u}_1 = \vec{v}_1$ (note that this eigenvector is already normalized).

Our goal is to look at what the matrix M looks like in the coordinate system defined by the basis U.

Compute $U^\top MU$ by writing $U = \begin{bmatrix} \vec{v}_1 & R \end{bmatrix}$, where $R \triangleq \begin{bmatrix} | & | & \cdots & | \\ \vec{r}_1 & \vec{r}_2 & \cdots & \vec{r}_{n-1} \\ | & | & \cdots & | \end{bmatrix}$. *(Note : $\vec{r}_i = \vec{u}_{i+1}$)*

Solution: Symbolic analysis:

$$U^\top MU = \begin{bmatrix} \vec{v}_1^\top \\ R^\top \end{bmatrix} M \begin{bmatrix} \vec{v}_1 & R \end{bmatrix} \tag{13}$$

$$= \begin{bmatrix} \vec{v}_1^\top \\ R^\top \end{bmatrix} \begin{bmatrix} M\vec{v}_1 & MR \end{bmatrix} \tag{14}$$

$$= \begin{bmatrix} \vec{v}_1^\top \\ R^\top \end{bmatrix} \begin{bmatrix} \lambda_1\vec{v}_1 & MR \end{bmatrix} \tag{15}$$

$$= \begin{bmatrix} \lambda_1\vec{v}_1^\top\vec{v}_1 & \vec{v}_1^\top MR \\ \lambda_1 R^\top\vec{v}_1 & R^\top MR \end{bmatrix}. \tag{16}$$

$$= \begin{bmatrix} \lambda_1 & \vec{v}_1^\top MR \\ \lambda_1 R^\top\vec{v}_1 & R^\top MR \end{bmatrix}. \tag{17}$$

Concrete case: $S_{[3\times3]}$ has zero as eigenvalue since it contains a repeated column vector. So, we let the corresponding eigenvector (which can be anything) be just $\begin{bmatrix} 1 & -1 & 0 \end{bmatrix}^\top$, the starting vector from the previous subppart. Then, we have:

$$U = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad R = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & 0 \\ 0 & 1 \end{bmatrix}$$

Performing the matrix multiplication yields:

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

From here, we can form a connection to the result of a couple subparts later, seeing that:

$$Q = R^\top S_{[3\times 3]} R = \begin{bmatrix} \frac{5}{6} & \frac{\sqrt{2}}{6} \\ \frac{\sqrt{2}}{6} & \frac{2}{3} \end{bmatrix}$$

(c) **Verify that $U^{-1} = U^\top$, where U is the matrix we get from Gram-Schmidt process.**

Solution: One way to reason through this proof is with definitions and properties. U is an orthonormal basis by our construction. $U^\top U$ performs an inner product between each of the basis vectors. Since these basis vectors are orthogonal to each other (we performed Gram-Schmidt to make them this way!), all the non-diagonal elements have to be 0. Since the basis vectors are normalized, the inner product with itself is 1. As a result, $U^\top U = I$, and $U^{-1} = U^\top$.

Also, this result was shown rigorously in lec9A. We outline the same approach below in the 3×3 case. Suppose we have an orthonormal matrix P:

$$P = \begin{bmatrix} | & | & | \\ \vec{p}_1 & \vec{p}_2 & \vec{p}_3 \\ | & | & | \end{bmatrix}$$

We can compute $P^\top P$. We use the fact that for a set of mutually orthonormal vectors, the inner product of a vector with any *other* vector in the set is 0, but the inner product of a vector with itself is 1:

$$\begin{aligned} P^\top P &= \begin{bmatrix} - & \vec{p}_1^\top & - \\ - & \vec{p}_2^\top & - \\ - & \vec{p}_3^\top & - \end{bmatrix} \begin{bmatrix} | & | & | \\ \vec{p}_1 & \vec{p}_2 & \vec{p}_3 \\ | & | & | \end{bmatrix} \\ &= \begin{bmatrix} \vec{p}_1^\top \vec{p}_1 & \vec{p}_1^\top \vec{p}_2 & \vec{p}_1^\top \vec{p}_3 \\ \vec{p}_2^\top \vec{p}_1 & \vec{p}_2^\top \vec{p}_2 & \vec{p}_2^\top \vec{p}_3 \\ \vec{p}_3^\top \vec{p}_1 & \vec{p}_3^\top \vec{p}_2 & \vec{p}_3^\top \vec{p}_3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= I \end{aligned}$$

This shows that $P^\top = P^{-1}$.

(d) **Look at the first column and the first row of $U^\top MU$ and show that:**

$$M = U \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix} U^\top \tag{18}$$

where $Q = R^\top MR$. Here, $\vec{a}$ is a vector related to M, R, and $\vec{v}_1$ (we will show the relation!).

Solution: We found above that:

$$U^\top MU = \begin{bmatrix} \lambda_1 & \vec{v}_1^\top MR \\ \lambda_1 R^\top \vec{v}_1 & R^\top MR \end{bmatrix} \tag{19}$$

Now, we need to show why:

$$\begin{bmatrix} \lambda_1 & \vec{v}_1^\top MR \\ \lambda_1 R^\top \vec{v}_1 & R^\top MR \end{bmatrix} \stackrel{?}{=} \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix} \tag{20}$$

We start simplifying the left side. First, we note that $\lambda_1 R^\top \vec{v}_1 = \vec{0}$ because R consists of all of the other $\vec{u}_i$ vectors that compose our orthonormal basis; taking the inner product between any one of these and $\vec{v}_1$ yields zero (same logic as outlined in the previous part for vectors that are mutually orthogonal). $\vec{v}_1^\top MR$ currently takes the place of $\vec{a}^\top$, suggesting that $\vec{a} = \left(\vec{v}_1^\top MR\right)^\top = R^\top M^\top \vec{v}_1$. So, finally substituting that $Q = R^\top MR$, we have:

$$U^\top MU = \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix} \tag{21}$$

We want an expression for M and so we can use the fact that $U^\top = U^{-1}$ to see:

$$U^\top MU = \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix} \implies M = U \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix} U^\top \tag{22}$$

In the numerical example with $S_{[3\times 3]}$, we have:

$$Q = R^\top S_{[3\times 3]} R = \begin{bmatrix} \frac{5}{6} & \frac{\sqrt{2}}{6} \\ \frac{\sqrt{2}}{6} & \frac{2}{3} \end{bmatrix} \tag{23}$$

(e) Now, we can recurse on Q to get:

$$Q = \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix} \begin{bmatrix} \lambda_2 & \vec{b}^\top \\ \vec{0} & P \end{bmatrix} \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}^\top \tag{24}$$

where we have taken $\vec{v}_2 \in \mathbb{R}^{n-1}$, a normalized eigenvector of Q, associated with eigenvalue λ_2. Again $\vec{v}_2$ is extended into an orthonormal basis to form $\begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}$.

Plug this form of Q into M above, to show that:

$$M = \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix} \begin{bmatrix} \lambda_1 & \breve{a}_1 & \breve{\vec{a}}_{\textbf{rest}}^\top \\ 0 & \lambda_2 & \vec{b}^\top \\ \vec{0} & \vec{0} & P \end{bmatrix} \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}^\top \tag{25}$$

where we define $\breve{\vec{a}}$ to be the "adjusted" $\vec{a}$ to account for the subsitution of Q; $\breve{\vec{a}}^\top = \vec{a}^\top \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}$.

Solution: From part (d), we know that

$$M = U \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix} U^\top \tag{26}$$

with $U = \begin{bmatrix} \vec{v}_1 & R \end{bmatrix}$

In the given definition of Q, let's denote $\begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}$ as U_2, since this is the orthonormal basis that upper triangularizes Q (note the middle matrix of Q, which we can call T_2, is block upper-triangular). We can then write that:

$$Q = U_2 \underbrace{\begin{bmatrix} \lambda_2 & \vec{b}^\top \\ \vec{0} & P \end{bmatrix}}_{T_2} U_2^\top \tag{27}$$

We had an expression for Q previously; $R^\top M R$. We can equate the two representations and simplify:

$$U_2 T_2 U_2^\top = R^\top M R \tag{28}$$
$$T_2 = U_2^\top R^\top M R U_2 \tag{29}$$
$$= (RU_2)^\top M R U_2 \tag{30}$$

We know that T_2 is an upper triangular matrix, so what the final equation above indicates is that the new orthonormal basis that upper triangularizes M *better than the original R basis*, is the RU_2 basis. That is, instead of using $\begin{bmatrix} \vec{v}_1 & R \end{bmatrix}$, we want $\begin{bmatrix} \vec{v}_1 & RU_2 \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}$.

Motivated by this, we start with

$$M = U \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & U_2 T_2 U_2^\top \end{bmatrix} U^\top \tag{31}$$
$$= U \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2 \end{bmatrix} \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2^\top \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & U_2 T_2 U_2^\top \end{bmatrix} \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2 \end{bmatrix} \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2^\top \end{bmatrix} U^\top \tag{32}$$

Here we have used the fact that U_2 is an orthonormal matrix and that $\begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2 \end{bmatrix} \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2^\top \end{bmatrix} = I$

$$\therefore M = U \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{a}^\top U_2 \\ \vec{0} & U_2^\top U_2 T_2 U_2^\top U_2 \end{bmatrix} \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2^\top \end{bmatrix} U^\top \tag{33}$$
$$= \begin{bmatrix} \vec{v}_1 & R \end{bmatrix} \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{a}^\top U_2 \\ \vec{0} & T_2 \end{bmatrix} \begin{bmatrix} 1 & \vec{0} \\ \vec{0} & U_2^\top \end{bmatrix} \begin{bmatrix} \vec{v}_1 & R \end{bmatrix}^\top \tag{34}$$

$$= \begin{bmatrix} \vec{v}_1 & RU_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{a}^\top U_2 \\ \vec{0} & T_2 \end{bmatrix} \begin{bmatrix} \vec{v}_1 & RU_2 \end{bmatrix}^\top \tag{35}$$

Defining the notation as $\vec{a}^\top U_2 = \breve{\vec{a}}^\top$, we can finally write that:

$$M = \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix} \begin{bmatrix} \lambda_1 & \breve{a}_1 & \breve{\vec{a}}_{\text{rest}}^\top \\ 0 & \lambda_2 & \vec{b}^\top \\ \vec{0} & \vec{0} & P \end{bmatrix} \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}^\top \tag{36}$$

We can be precise and write that $\breve{\vec{a}}_{\text{rest}}^\top = \breve{\vec{a}}_{2:n-1}^\top$.

The numerical results are:

$$Q = \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \\ \frac{-\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \\ \frac{-\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}^\top \tag{37}$$

$$M = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}^\top \tag{38}$$

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

Solution: To show that the matrix $A = \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}$ is orthonormal, we want to show that the columns are mutually orthogonal, and all columns are unit vectors.

Orthogonality: 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 $\vec{v}_1^\top RY = \vec{0}^\top$ since $\vec{v}_1^\top R = \vec{0}^\top$. As for the orthogonality of $R\vec{v}_2$ and RY, we can see that

$$(R\vec{v}_2)^\top RY = \vec{v}_2^\top R^\top RY = \vec{v}_2^\top Y = \vec{0}^\top \tag{39}$$

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

Normality : To check for normality (i.e all vectors are unit length), we can consider the inner products of each element with itself:

$$\vec{v}_1^\top \vec{v}_1 = 1 \tag{40}$$

$$(R\vec{v}_2)^\top R\vec{v}_2 = \vec{v}_2^\top R^\top R\vec{v}_2 \tag{41}$$

$$= \vec{v}_2^\top \vec{v}_2 = 1 \tag{42}$$

$$(RY)^\top RY = Y^\top R^\top RY \tag{43}$$

$$= Y^\top Y = I. \tag{44}$$

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

(g) *(Practice)* 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?**

Solution:

In class we've seen a recursive algorithm for upper-triangularization

Algorithm 1 UpperTriangularize

Require: matrix M
if dim(M) == 1 **then**
  return ([1])
else
  $\vec{v}_1$ = eigenvector(M)
  R = GramSchmidtRest($\vec{v}_1$) ▷ Create the rest of an orthonormal matrix given $\vec{v}_1$
  Compute $B = R^\top M R$ ▷ $(n-1) \times (n-1)$ matrix
  U^1 = UpperTriangularize(B)
  $U = \begin{bmatrix} \vec{v}_1 & RU^1 \end{bmatrix}$
  return (U)
end if

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

$$M_n = \begin{bmatrix} \vec{v}_1 & R_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & M_{n-1} \end{bmatrix} \begin{bmatrix} \vec{v}_1 & R_n \end{bmatrix}^\top \tag{45}$$

$$= U_n A_n U_n^\top, \tag{46}$$

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 $U_{n,\text{final}}$:

$$U_{i,\text{final}} = \begin{bmatrix} \vec{v}_{n-i+1} & R_i U_{i-1,\text{final}} \end{bmatrix} \tag{47}$$

Further, here in this part we see something more. Namely that we can actually do this in a single loop — the recursion can be transformed into a tail recursion. The key is that we can advance to get the the next vector in the basis directly – it is $R\vec{v}_2$.

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:

$$M = UTU^\top \tag{48}$$

$$T = U^\top MU. \tag{49}$$

Algorithm 2 UpperTriangularizeLoop

Require: matrix M

CurrentMatrix = M
$U = []$ ▷ Need a place to accumulate the result
R =Identity(M) ▷ Same dimension as M to start, but this will accumulate the transformation
while dim(CurrentMatrix) > 0 **do**
    $\vec{v}$ = eigenvector(CurrentMatrix) ▷ Get one that is normalized
    U =columnstack($U, R\vec{v}$) ▷ Add the new vector to the basis
    if dim(CurrentMatrix) == 1 **then**
        CurrentMatrix $= []$
    else
        Y = GramSchmidtRest($\vec{v}$) ▷ Create an orthonormal matrix given $\vec{v}$
        CurrentMatrix= $Y^\top \text{CurrentMatrix } Y$ ▷ One smaller than before
        $R = RY$ ▷ Update translation to original coordinates
    end if
end while
return (U)

(h) *(Practice)* **Show that the characteristic polynomial of square matrix M is the same as that of the square matrix UMU^{-1} for any invertible U.** You should use the key property $\det(AB) = \det(A)\det(B)$ for square matrices.

Solution: 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

$$\det\left(UMU^{-1} - \lambda I\right) = \det\left(UMU^{-1} - \lambda UU^{-1}\right) \tag{50}$$

$$= \det\left(U(M - \lambda I)U^{-1}\right) \tag{51}$$

$$= \det(U)\det(M - \lambda I)\det\left(U^{-1}\right). \tag{52}$$

Recognizing that $\det(U) \cdot \det\left(U^{-1}\right) = 1$ we can simplify eq. (52) to:

$$\implies \det\left(UMU^{-1} - \lambda I\right) = \det(M - \lambda I). \tag{53}$$

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

Contributors:

- Neelesh Ramachandran.
- Yuxun Zhou.
- Edward Wang.
- Anant Sahai.
- Sanjit Batra.

- Pavan Bhargava.
- Kumar Krishna Agrawal.

