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

Discussion 9B

1. Gram-Schmidt Algorithm

Let's apply Gram-Schmidt orthonormalization to a set of three linearly independent vectors $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$.

(a) Find unit vector $\vec{q}_1$ such that $\text{span}(\{\vec{q}_1\}) = \text{span}(\{\vec{s}_1\})$.

Answer: Since $\text{span}(\{\vec{s}_1\})$ is a one dimensional vector space, the unit vector that spans the same vector space would just be the normalized vector which points in the same direction as $\vec{s}_1$. Therefore

$$\vec{q}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}.$$

(b) Given $\vec{q}_1$ from the previous step, find $\vec{q}_2$ such that $\text{span}(\{\vec{q}_1, \vec{q}_2\}) = \text{span}(\{\vec{s}_1, \vec{s}_2\})$ and $\vec{q}_2$ is orthogonal to $\vec{q}_1$.

Answer: We want to find the projection of $\vec{s}_2$ onto $\vec{q}_1$. The last time we worked with this type of projections is when we discussed least squares. We will try to work in that framework. A detailed discussion of least squares is in 16A Note 23.

Specifically, least squares solves the problem $A\vec{x} \approx \vec{b}$ by finding the closest point $A\vec{x}$ to $\vec{b}$ among all vectors $\vec{x}$. Since $A\vec{x}$ ranges over all vectors in $\text{col}(A)$, this is equivalent to finding the closest $\vec{b}_*$ in $\text{col}(A)$ to $\vec{b}$. We know

$$\vec{x} = \left(A^\top A\right)^{-1} A^\top \vec{b} \implies \vec{b}_* = A\left(A^\top A\right)^{-1} A^\top \vec{b}.$$

Thus, to project a vector into the column space of A, we multiply it by the matrix $A\left(A^\top A\right)^{-1} A^\top$.

In particular, we want to project $\vec{s}_2$ onto $\vec{q}_1$. A matrix A such that $\text{span}(\vec{q}_1) = \text{col}(A)$ is just $A = \vec{q}_1$. Since $\vec{q}_1$ is normalized,

$$A\left(A^\top A\right)^{-1} A^\top = \vec{q}_1(\underbrace{\vec{q}_1^\top \vec{q}_1}_{1})^{-1}\vec{q}_1^\top = \vec{q}_1\vec{q}_1^\top \implies A\left(A^\top A\right)^{-1} A^\top \vec{s}_2 = \vec{q}_1\vec{q}_1^\top \vec{s}_2 = \left(\vec{q}_1^\top \vec{s}_2\right)\vec{q}_1.$$

And subtracting out this projection gets us only the component orthogonal to $\vec{q}_1$:

$$\vec{z}_2 = \vec{s}_2 - \left(\vec{q}_1^\top \vec{s}_2\right)\vec{q}_1.$$

Note that this projection formula *only works if $\vec{q}_1$ is normalized*. Now, normalizing to get $\vec{q}_2$, we have $\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|}$.

(c) Confirm that $\text{span}(\{\vec{q}_1, \vec{q}_2\}) = \text{span}(\{\vec{s}_1, \vec{s}_2\})$.

Answer: We will show that each vector in $\{\vec{q}_1, \vec{q}_2\}$ can be written as a linear combination of $\{\vec{s}_1, \vec{s}_2\}$, and that each vector in $\{\vec{s}_1, \vec{s}_2\}$ can be written as a linear combination of $\{\vec{q}_1, \vec{q}_2\}$. From there, we use the fact that linear combinations of linear combinations of vectors in a set are themselves linear combinations of vectors in the set, which proves the claim.

We first show that each vector in $\{\vec{q}_1, \vec{q}_2\}$ can be written as a linear combination of $\{\vec{s}_1, \vec{s}_2\}$. In particular, by construction,

$$\vec{q}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}$$

so $\vec{q}_1$ is a linear combination of $\{\vec{s}_1, \vec{s}_2\}$. And, unrolling the Gram-Schmidt algorithm,

$$\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} = \frac{\vec{s}_2 - (\vec{q}_1^\top \vec{s}_2)\vec{q}_1}{\left\|\vec{s}_2 - (\vec{q}_1^\top \vec{s}_2)\vec{q}_1\right\|} = \underbrace{\frac{1}{\left\|\vec{s}_2 - (\vec{q}_1^\top \vec{s}_2)\vec{q}_1\right\|}}_{\text{a scalar}} \vec{s}_2 - \underbrace{\frac{\vec{q}_1^\top \vec{s}_2}{\left\|\vec{s}_2 - (\vec{q}_1^\top \vec{s}_2)\vec{q}_1\right\|}}_{\text{another scalar}} \vec{s}_1$$

so $\vec{q}_2$ is a linear combination of $\{\vec{s}_1, \vec{s}_2\}$.

We now show that each vector in $\{\vec{s}_1, \vec{s}_2\}$ can be written as a linear combination of $\{\vec{q}_1, \vec{q}_2\}$. Indeed,

$$\vec{s}_1 = \|\vec{s}_1\| \, \vec{q}_1$$

so $\vec{s}_1$ is a linear combination of $\{\vec{q}_1, \vec{q}_2\}$. And from the same unrolling as above, some algebra gets

$$\vec{s}_2 = \underbrace{\left\|\vec{s}_2 - (\vec{q}_1^\top \vec{s}_2)\vec{q}_1\right\|}_{\text{a scalar}} \vec{q}_2 + \underbrace{(\vec{q}_1^\top \vec{s}_2)}_{\text{another scalar}} \vec{q}_1.$$

Thus $\vec{s}_2$ is a linear combination of $\{\vec{q}_1, \vec{q}_2\}$ and we are done.

(d) What would happen if $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$ were *not* linearly independent, but rather $\vec{s}_1$ were a multiple of $\vec{s}_2$?

Answer: If $\vec{s}_2$ is a multiple of $\vec{s}_1$, then $\vec{z}_2 = 0$. This means that the projection of $\vec{s}_2$ onto $\text{span}(\{\vec{s}_1\})$ is just $\vec{s}_2$, so we have found an orthonormal basis for $\text{span}(\{\vec{s}_1, \vec{s}_2\})$, in particular the basis $\{\vec{q}_1\}$. Hence, we can move onto $\vec{s}_3$ and continue the algorithm from there.

(e) Now given $\vec{q}_1$ and $\vec{q}_2$ in the previous steps, find $\vec{q}_3$ such that $\text{span}(\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}) = \text{span}(\{\vec{s}_1, \vec{s}_2, \vec{s}_3\})$, and $\vec{q}_3$ is orthogonal to both $\vec{q}_1$ and $\vec{q}_2$, and finally $\|\vec{q}_3\| = 1$.

Answer: We want to project $\vec{s}_3$ onto $\text{span}(\{\vec{q}_1, \vec{q}_2\})$. A matrix A such that $\text{span}(\{\vec{q}_1, \vec{q}_2\}) = \text{col}(A)$ is just $A = \begin{bmatrix}\vec{q}_1 & \vec{q}_2\end{bmatrix}$. Since $\{\vec{q}_1, \vec{q}_2\}$ are orthonormal, the projection matrix is

$$\begin{aligned}
A\left(A^\top A\right)^{-1} A^\top &= \begin{bmatrix}\vec{q}_1 & \vec{q}_2\end{bmatrix} \left(\begin{bmatrix}\vec{q}_1^\top \\ \vec{q}_2^\top\end{bmatrix} \begin{bmatrix}\vec{q}_1 & \vec{q}_2\end{bmatrix} \right)^{-1} \begin{bmatrix}\vec{q}_1^\top \\ \vec{q}_2^\top\end{bmatrix} \\
&= \begin{bmatrix}\vec{q}_1 & \vec{q}_2\end{bmatrix} \begin{bmatrix}\vec{q}_1^\top \vec{q}_1 & \vec{q}_1^\top \vec{q}_2 \\ \vec{q}_2^\top \vec{q}_1 & \vec{q}_2^\top \vec{q}_2\end{bmatrix}^{-1} \begin{bmatrix}\vec{q}_1^\top \\ \vec{q}_2^\top\end{bmatrix} \\
&= \begin{bmatrix}\vec{q}_1 & \vec{q}_2\end{bmatrix} \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix}^{-1} \begin{bmatrix}\vec{q}_1^\top \\ \vec{q}_2^\top\end{bmatrix} \\
&= \begin{bmatrix}\vec{q}_1 & \vec{q}_2\end{bmatrix} \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} \begin{bmatrix}\vec{q}_1^\top \\ \vec{q}_2^\top\end{bmatrix} \\
&= \begin{bmatrix}\vec{q}_1 & \vec{q}_2\end{bmatrix} \begin{bmatrix}\vec{q}_1^\top \\ \vec{q}_2^\top\end{bmatrix} \\
&= \vec{q}_1\vec{q}_1^\top + \vec{q}_2\vec{q}_2^\top.
\end{aligned}$$

Thus,

$$A\left(A^\top A\right)^{-1}A^\top\vec{s}_3 = \vec{q}_1\vec{q}_1^\top\vec{s}_3 + \vec{q}_2\vec{q}_2^\top\vec{s}_3 = \left(\vec{q}_1^\top\vec{s}_3\right)\vec{q}_1 + \left(\vec{q}_2^\top\vec{s}_3\right)\vec{q}_2.$$

Note that this projection formula *only works if* $\{\vec{q}_1, \vec{q}_2\}$ *are orthonormal*. Then

$$\vec{z}_3 = \vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2$$

is orthogonal to $\vec{q}_1$ and $\vec{q}_2$, hence orthogonal to any vector in $\text{span}(\{\vec{q}_1, \vec{q}_2\})$. Normalizing, we have $\vec{q}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|}$.

(f) (Optional.) Confirm that $\text{span}(\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}) = \text{span}(\{\vec{s}_1, \vec{s}_2, \vec{s}_3\})$.

Answer: We only have to show that $\vec{q}_3$ can be written as a linear combination of $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$, and that $\vec{s}_3$ can be written as a linear combination of $\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$. The rest follows from part (c), taking the coefficients of $\vec{s}_3$ or $\vec{q}_3$ to be 0 in every linear combination.

To write $\vec{q}_3$ as a linear combination of $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$, we unroll the Gram-Schmidt algorithm:

$$\begin{aligned}\vec{q}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|} &= \frac{\vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2}{\left\|\vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2\right\|} \\ &= \underbrace{\frac{1}{\left\|\vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2\right\|}}_{\text{a scalar}}\vec{s}_3 - \underbrace{\frac{\vec{s}_3^\top\vec{q}_1}{\left\|\vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2\right\|}}_{\text{another scalar}}\vec{q}_1 \\ &\quad - \underbrace{\frac{\vec{s}_3^\top\vec{q}_2}{\left\|\vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2\right\|}}_{\text{another scalar}}\vec{q}_2\end{aligned}$$

Since we have already shown $\vec{q}_1$ and $\vec{q}_2$ are linear combinations of vectors in $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$, we see that $\vec{q}_3$ is a linear combination of $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$.

To write $\vec{s}_3$ as a linear combination of $\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$, we again unroll the algorithm, this time doing another simplification:

$$\begin{aligned}\vec{q}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|} &= \frac{\vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2}{\left\|\vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2\right\|} \\ \vec{s}_3 &= \underbrace{\left\|\vec{s}_3 - \left(\vec{s}_3^\top\vec{q}_1\right)\vec{q}_1 - \left(\vec{s}_3^\top\vec{q}_2\right)\vec{q}_2\right\|}_{\text{a scalar}}\vec{q}_3 + \underbrace{\left(\vec{s}_3^\top\vec{q}_1\right)}_{\text{another scalar}}\vec{q}_1 + \underbrace{\left(\vec{s}_3^\top\vec{q}_2\right)}_{\text{another scalar}}\vec{q}_2.\end{aligned}$$

Hence $\vec{s}_3$ is a linear combination of $\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$ and the proof is complete.

(g) Let's extend this algorithm to n linearly independent vectors. That is, given an input $\{\vec{s}_1, \ldots, \vec{s}_n\}$, write the algorithm to calculate the orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$, where $\text{span}(\{\vec{s}_1, \ldots, \vec{s}_n\}) = \text{span}(\{\vec{q}_1, \ldots, \vec{q}_n\})$.

Hint: How would you calculate the i^{th} vector, $\vec{q}_i$?

Answer:

We extend this same process to n vectors. Note that each part of this algorithm is motivated; to get $\vec{z}_i$, we take $\vec{s}_i$ and subtract off the projection onto $\vec{q}_1, \ldots, \vec{q}_{i-1}$. Then $\vec{z}_i$ is orthogonal to each of $\vec{q}_1, \ldots, \vec{q}_{i-1}$. Then we normalize $\vec{z}_i$ to get $\vec{q}_i$, so that $\{\vec{q}_1, \ldots, \vec{q}_i\}$ are truly orthonormal.

Formally, the algorithm is this:

Inputs

- A set of linearly independent vectors $\{\vec{s}_1, \ldots, \vec{s}_n\}$.

Outputs

- An orthonormal set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$, where span($\{\vec{s}_1, \ldots, \vec{s}_n\}$) = span($\{\vec{q}_1, \ldots, \vec{q}_n\}$).

Gram Schmidt Procedure

- compute $\vec{q}_1$:

$$\vec{q}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}$$

- for $i = 2 \ldots n$:
  i. Compute vector $\vec{z}_i$, such that $\{\vec{q}_1, \ldots, \vec{q}_{i-1}, \vec{z}_i\}$ is orthogonal:

$$\vec{z}_i = \vec{s}_i - \sum_{j=1}^{i-1} \left(\vec{q}_j^\top \vec{s}_i\right) \vec{q}_j$$

  ii. Normalize to compute $\vec{q}_i$:

$$\vec{q}_i = \frac{\vec{z}_i}{\|\vec{z}_i\|}$$

Each part of this algorithm is conceptually the first thing you would try; but the end result looks pretty complicated, and definitely very powerful!

2. The Order of Gram-Schmidt

If we are performing the Gram-Schmidt method on a set of vectors, does the order in which we take the vectors matter? It turns out that it does. There are infinitely many bases (and thus orthonormal bases) for $\mathbb{R}^3$, and Gram-Schmidt can output several of them depending on the order of the inputs.

Let's try it out. Consider the set of vectors

$$\vec{v}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix}1\\1\\0\end{bmatrix} \qquad \vec{v}_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

(a) Perform Gram-Schmidt on these vectors first in the order $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3$.

Answer: If we start with $\vec{v}_1$ we get

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

$$\vec{z}_2 = \vec{v}_2 - \left(\vec{v}_2^\top \vec{q}_1\right)\vec{q}_1 = \vec{v}_2 - \vec{q}_1 = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

$$\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} = \vec{z}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

$$\vec{z}_3 = \vec{v}_3 - \left(\vec{v}_3^\top \vec{q}_1\right)\vec{q}_1 - \left(\vec{v}_3^\top \vec{q}_2\right)\vec{q}_2 = \vec{v}_3 - \vec{q}_1 - \vec{q}_2 = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

$$\vec{q}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|} = \vec{z}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

so

$$\{\vec{q}_1, \vec{q}_2, \vec{q}_3\} = \left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$$

The fact that Gram-Schmidt outputs the standard basis is just a coincidence. It is possible to get a basis for $\mathbb{R}^3$ from Gram-Schmidt that is not the standard basis, as we see in the next part.

(b) Now perform Gram-Schmidt on these vectors in the order $\vec{v}_3,\ \vec{v}_2,\ \vec{v}_1$. Do you get the same result?

Answer: If we write the basis starting with $\vec{v}_3$.

$$\vec{q}_1 = \frac{\vec{v}_3}{\|\vec{v}_3\|} = \frac{1}{\sqrt{3}}\vec{v}_3 = \begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

$$\vec{z}_2 = \vec{v}_2 - \left(\vec{v}_2^\top \vec{q}_1\right)\vec{q}_1 = \vec{v}_2 - \frac{2}{\sqrt{3}}\vec{q}_1 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$$

$$\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} = \frac{1}{\sqrt{2/3}}\vec{z}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -\sqrt{2/3} \end{bmatrix}$$

$$\vec{z}_3 = \vec{v}_1 - \left(\vec{v}_1^\top \vec{q}_1\right)\vec{q}_1 - \left(\vec{v}_1^\top \vec{q}_2\right)\vec{q}_2 = \vec{v}_1 - \frac{1}{\sqrt{3}}\vec{q}_1 - \frac{1}{\sqrt{6}}\vec{q}_2 = \begin{bmatrix} 1/2 \\ -1/2 \\ 0 \end{bmatrix}$$

$$\vec{q}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|} = \frac{1}{1/\sqrt{2}}\vec{z}_3 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}$$

So

$$\{\vec{q}_1, \vec{q}_2, \vec{q}_3\} = \left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -\sqrt{2/3} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} \right\}$$

The next part is optional.

There is a way to simplify our arithmetic and get rid of the many annoying square roots and such. Recall that such trouble originated by computing $(\vec{v}_i^\top \vec{q}_j)\vec{q}_j$, and most of this comes from the $\vec{q}_j$. What if we could bypass using $\vec{q}_j$ until the very end? In fact, just by using the definition of $\vec{q}_j$ as $\vec{q}_j = \frac{\vec{z}_j}{\|\vec{z}_j\|}$ in our formula for $\vec{z}_i$, we get

$$\begin{aligned} \vec{z}_i &= \vec{v}_i - \left(\vec{v}_i^\top \vec{q}_1\right)\vec{q}_1 - \cdots - \left(\vec{v}_i^\top \vec{q}_{i-1}\right)\vec{q}_{i-1} \\ &= \vec{v}_i - \left(\vec{v}_i^\top \frac{\vec{z}_1}{\|\vec{z}_1\|}\right)\frac{\vec{z}_1}{\|\vec{z}_1\|} - \cdots - \left(\vec{v}_i^\top \frac{\vec{z}_{i-1}}{\|\vec{z}\|_{i-1}}\right)\frac{\vec{z}_{i-1}}{\|\vec{z}_{i-1}\|} \\ &= \vec{v}_i - \frac{\vec{v}_i^\top \vec{z}_1}{\|\vec{z}_1\|^2}\vec{z}_1 - \cdots - \frac{\vec{v}_i^\top \vec{z}_{i-1}}{\|\vec{z}_{i-1}\|^2}\vec{z}_{i-1}. \end{aligned}$$

Boom, no more square roots! At least, until we need to compute $\vec{q}_i$ by again taking $\vec{q}_i = \frac{\vec{z}_i}{\|\vec{z}_i\|}$, but in many cases this is simpler than using $\vec{q}_j$ for each of the $\vec{z}_i$ calculations.

It bears mentioning why exactly this formula works, and what each term represents. Our derivation actually helps figure out the new interpretation.

- Of course, $\vec{v}_i$ stays the same, as the vector we're orthogonalizing.
- Before our change, $\vec{v}_i^\top \vec{q}_j$ gave the fraction of $\vec{q}_j$ which is along $\vec{v}_i$. We point it in the direction of $\vec{q}_j$ by multiplying by $\vec{q}_j$, which is normalized. This is the projection of $\vec{v}_i$ onto $\vec{q}_j$. Correspondingly, $\frac{\vec{v}_i^\top \vec{z}_j}{\|\vec{z}_j\|}$ is the fraction of $\vec{z}_j$ which is along $\vec{v}_i$. We point it in the direction of $\vec{z}_j$ by multiplying by $\frac{\vec{z}_j}{\|\vec{z}_j\|}$, which is also normalized. Thus this term is the projection of $\vec{v}_i$ onto $\vec{z}_j$.

So really all we're doing here is taking projections onto non-normalized, but still orthogonal vectors.

(c) What would happen if we stopped after $\vec{v}_2$ in the above question? In particular, if we did Gram-Schmidt on the vectors $\{\vec{v}_3, \vec{v}_2\}$? What kind of structure would our output vectors span, and how many of them would there be?

Answer: Recall that Gram-Schmidt process is used to orthonormalize some linearly independent vectors. If we have fewer vectors than the dimension of the space we're in, then we stop after the last vector in the input. We don't get a basis for the space we're in, and *that's fine.* In this case, our two output vectors would be

$$\vec{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \qquad \vec{q}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -\sqrt{2/3} \end{bmatrix}$$

and they would span a plane in $\mathbb{R}^3$.

Contributors:

- Regina Eckert.
- Druv Pai.