Gram-Schmidt process for change of basis

The **Gram-Schmidt process** is a process that turns the basis of a subspace into an orthonormal basis for the same subspace. Of course, as we've seen, this is extremely useful.

Since working with an orthonormal basis can, for certain applications, be much simpler than working with a non-orthonormal basis, we can save a lot of time by changing the basis of the subspace into an orthonormal basis.

The Gram-Schmidt process changes the basis by dealing with one basis vector at a time, and it works regardless of how many vectors form the basis for the subspace.

How to use Gram-Schmidt

Let's say that a non-orthonormal basis of the subspace V is given by $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$. In other words,

$$V = \text{Span}(\vec{v}_1, \vec{v}_2, \vec{v}_3)$$

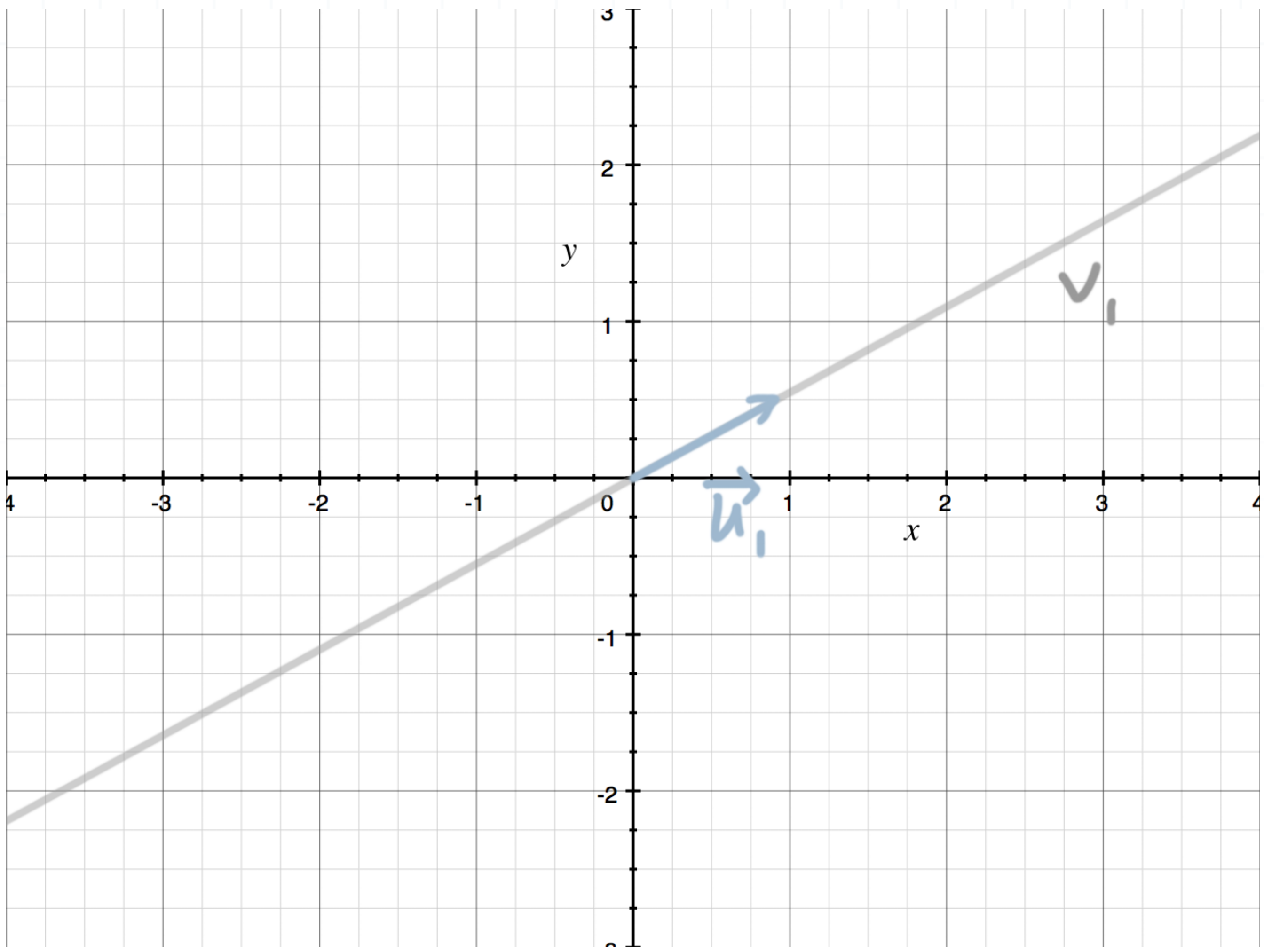
Our first step is to normalize $\vec{v}_1$ (make its length equal to 1). We can make $\vec{v}_1$ normal using

$$\vec{u}_1 = \frac{\vec{v}_1}{||\vec{v}_1||}$$

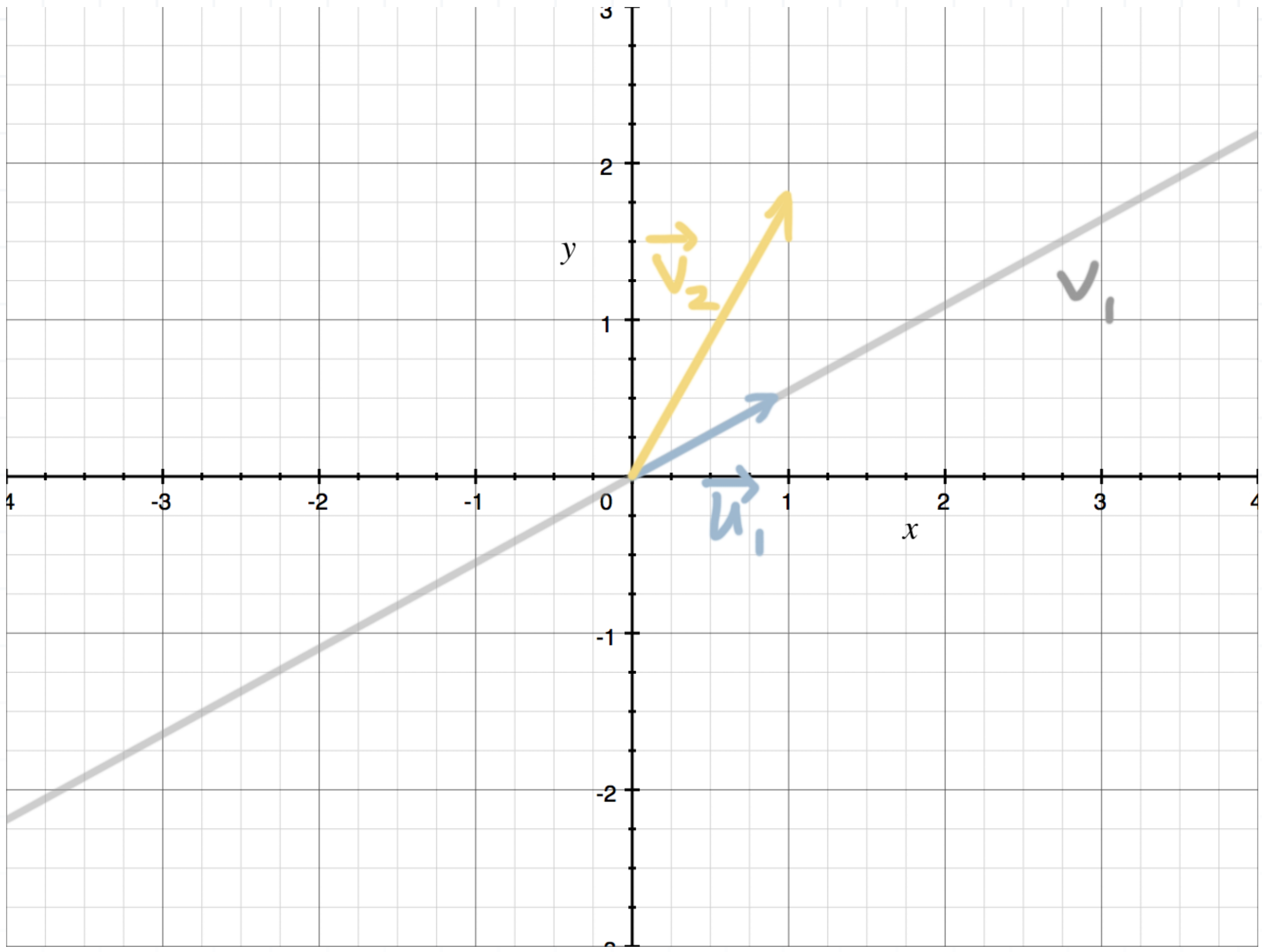
where $||\vec{v}_1||$ is the length of $\vec{v}_1$. Then with $\vec{v}_1$ normalized, the basis of V can be formed by $\vec{u}_1$, $\vec{v}_2$, and $\vec{v}_3$.

$$V = \text{Span}(\vec{u}_1, \vec{v}_2, \vec{v}_3)$$

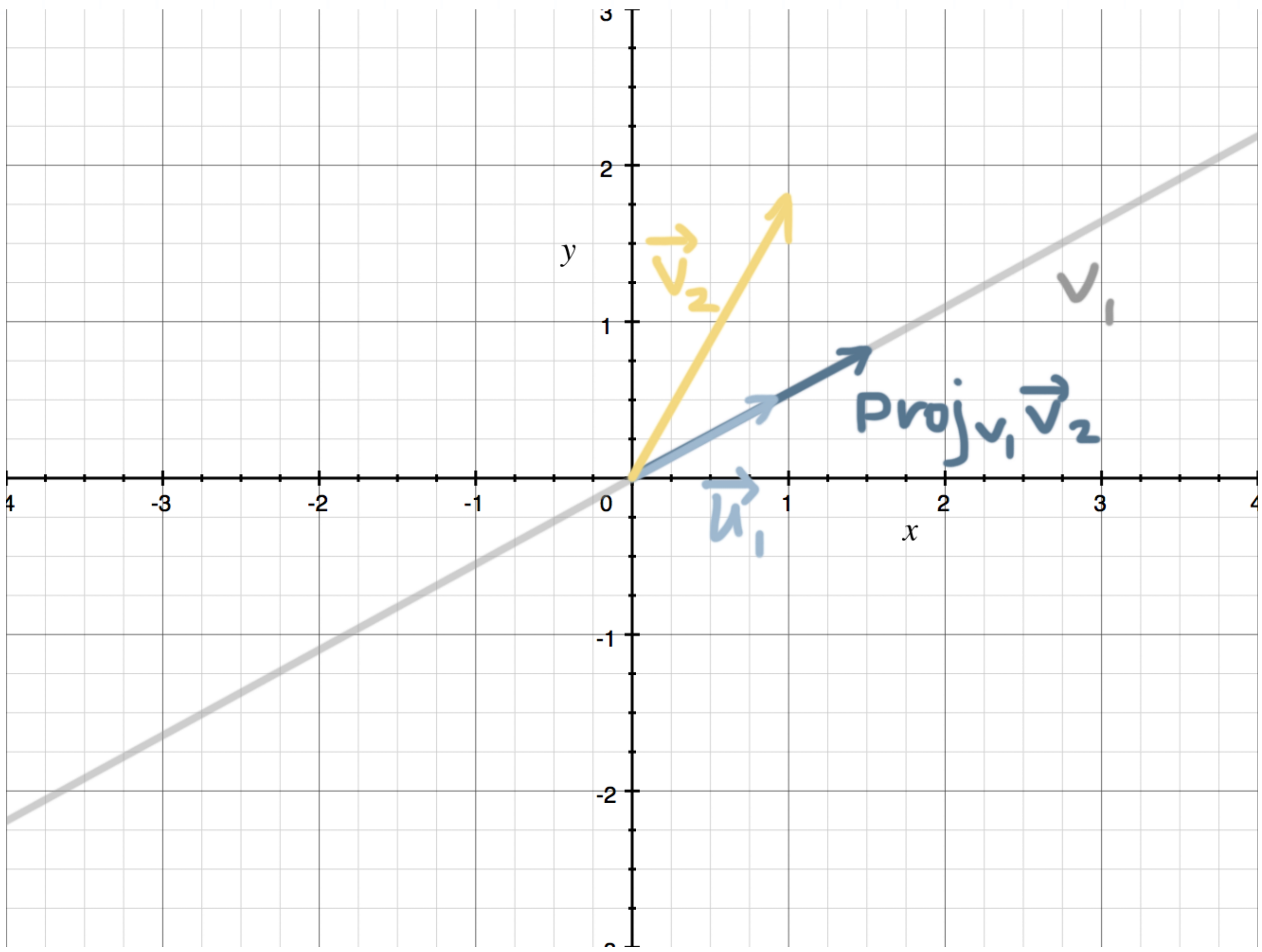
The next step is to replace $\vec{v}_2$ with with a vector that's both orthogonal to $\vec{u}_1$, and normal. To do that, we need to think about the span of just $\vec{u}_1$, which we'll call V_1. Generally, we could imagine that as



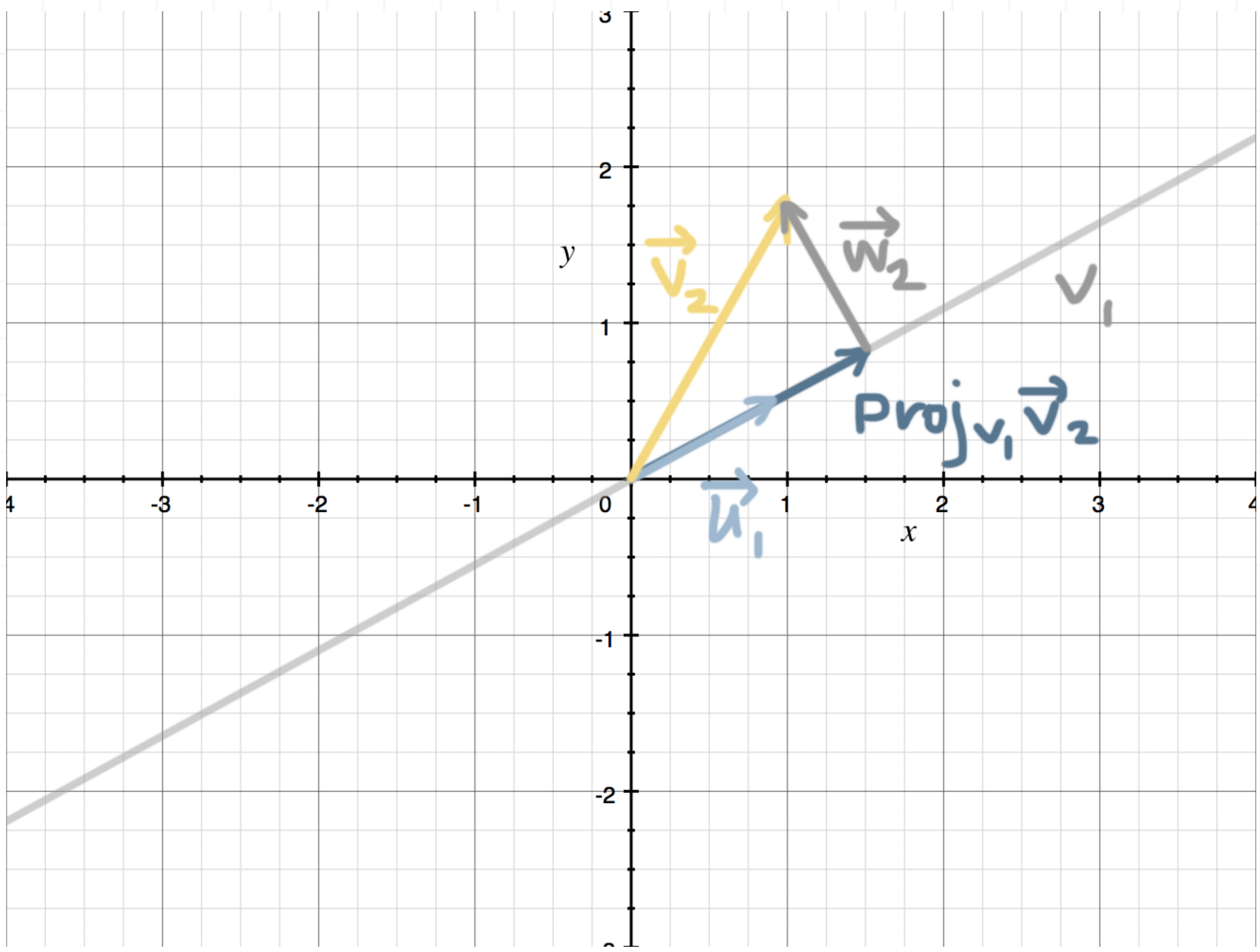
Let's imagine that $\vec{v}_2$ is another vector:



We could sketch the projection of $\vec{v}_2$ onto V_1, $\text{Proj}_{V_1}\vec{v}_2$:



Then $\vec{w}_2$ is the vector that connects $\text{Proj}_{V_1}\vec{v}_2$ to $\vec{v}_2$, so $\vec{w}_2 = \vec{v}_2 - \text{Proj}_{V_1}\vec{v}_2$,



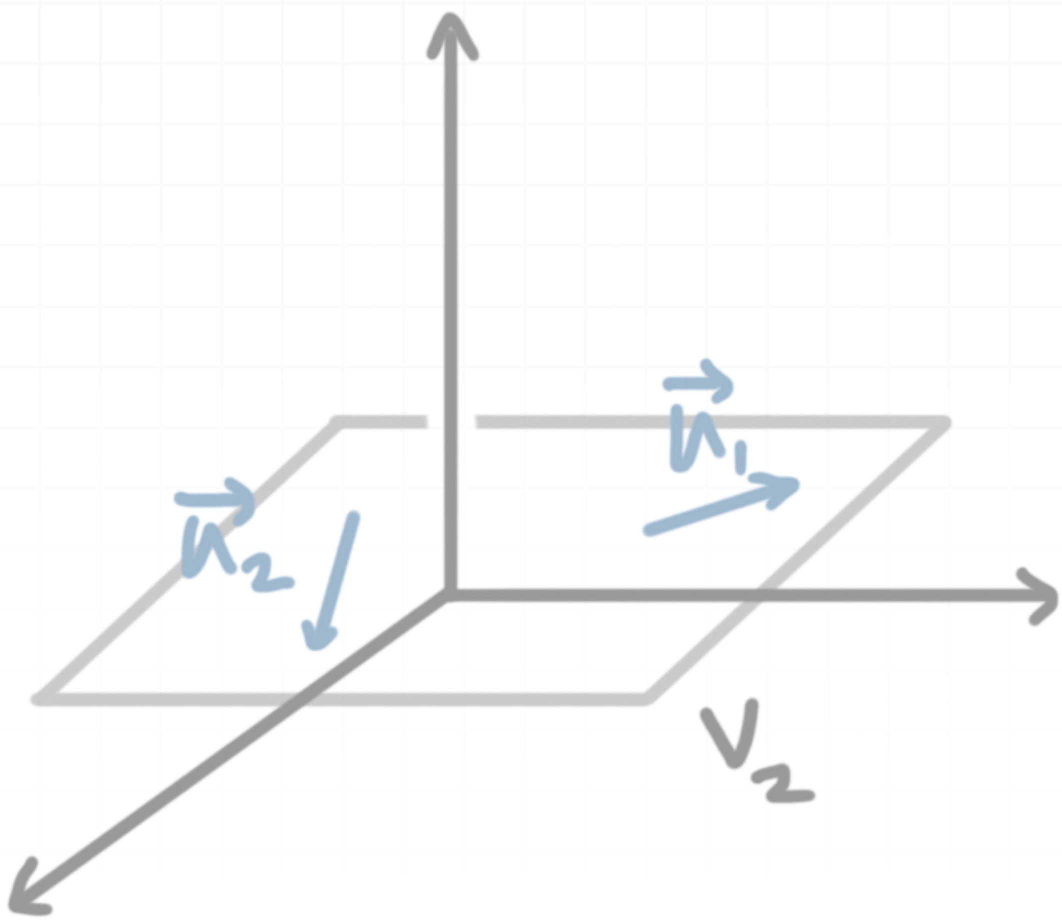
and $\vec{w}_2$ is orthogonal to $\vec{u}_1$. But realize that V_1 is an orthonormal subspace. Since there's only one vector $\vec{u}_1$ that forms the basis for V_1, every vector in the basis is orthogonal to every other vector (because there are no other vectors), and $\vec{u}_1$ is normal, so the basis of V_1 is orthonormal. Which means that we could rewrite the projection $\text{Proj}_{V_1}\vec{v}_2$ as

$$\vec{w}_2 = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1)\vec{u}_1$$

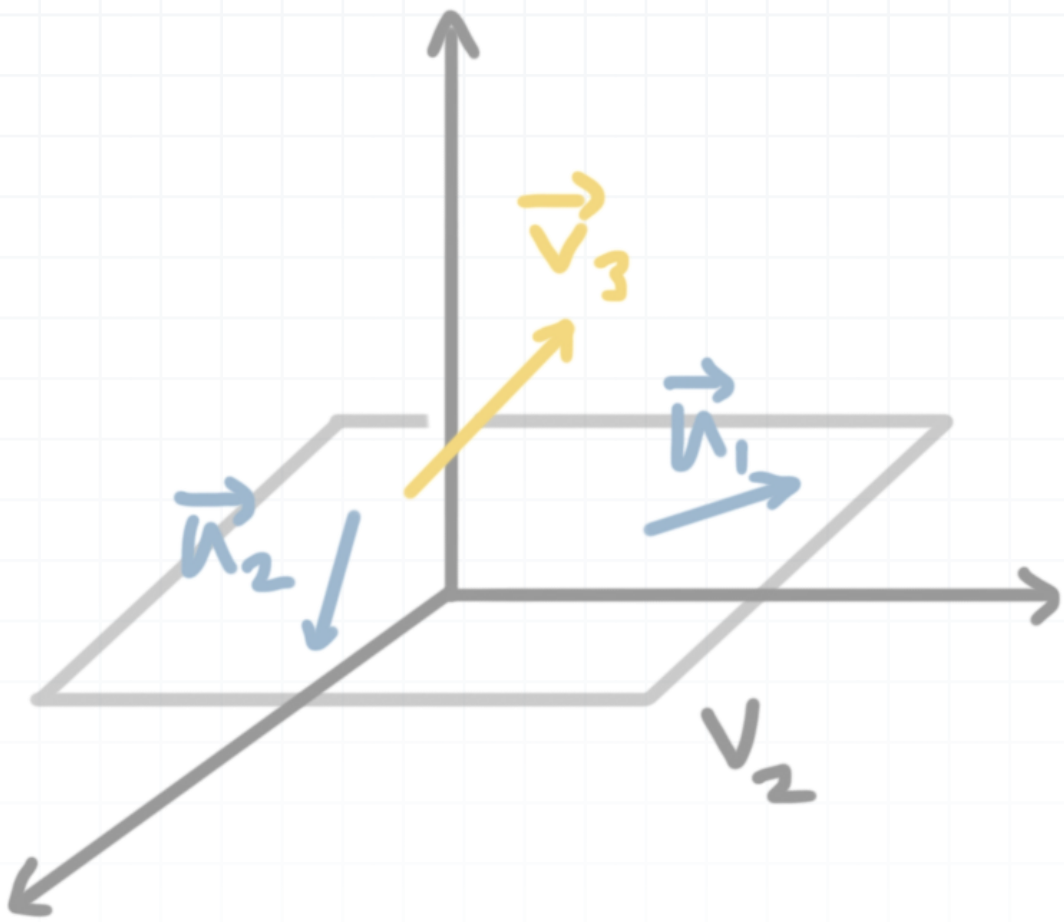
This will give us a vector $\vec{w}_2$ that we can use in place of $\vec{v}_2$. Once we normalize $\vec{w}_2$ the same way we normalized $\vec{v}_1$, we'll call it $\vec{u}_2$, and then we'll be able to say that the basis of V can be formed by $\vec{u}_1$, $\vec{u}_2$, and $\vec{v}_3$.

$$V = \text{Span}(\vec{u}_1, \vec{u}_2, \vec{v}_3)$$

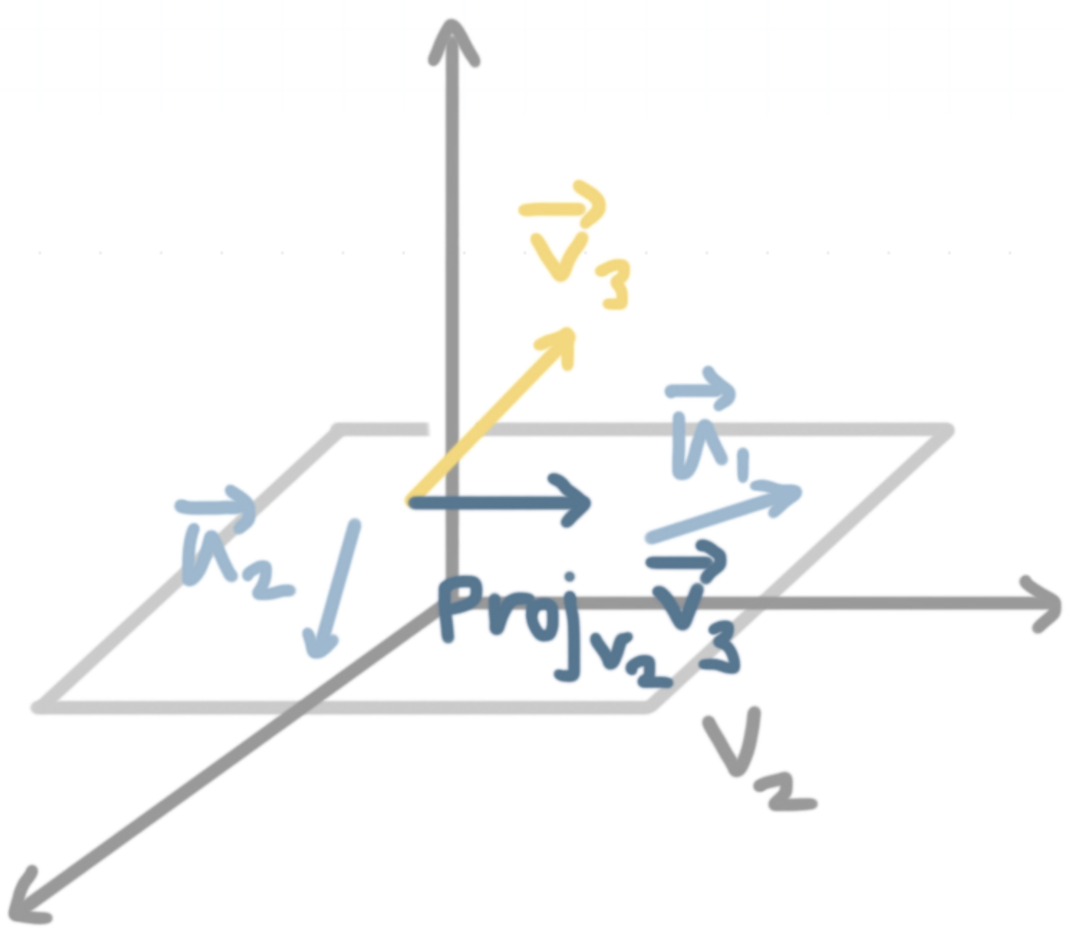
And we would just keep repeating this process with every basis vector. In this case, for the last basis vector $\vec{v}_3$, we'd think about the span of $\vec{u}_1$ and $\vec{u}_2$, which we'll call V_2. The subspace V_2 will be a plane, whereas the subspace V_1 was a line. Generally, we could imagine that as



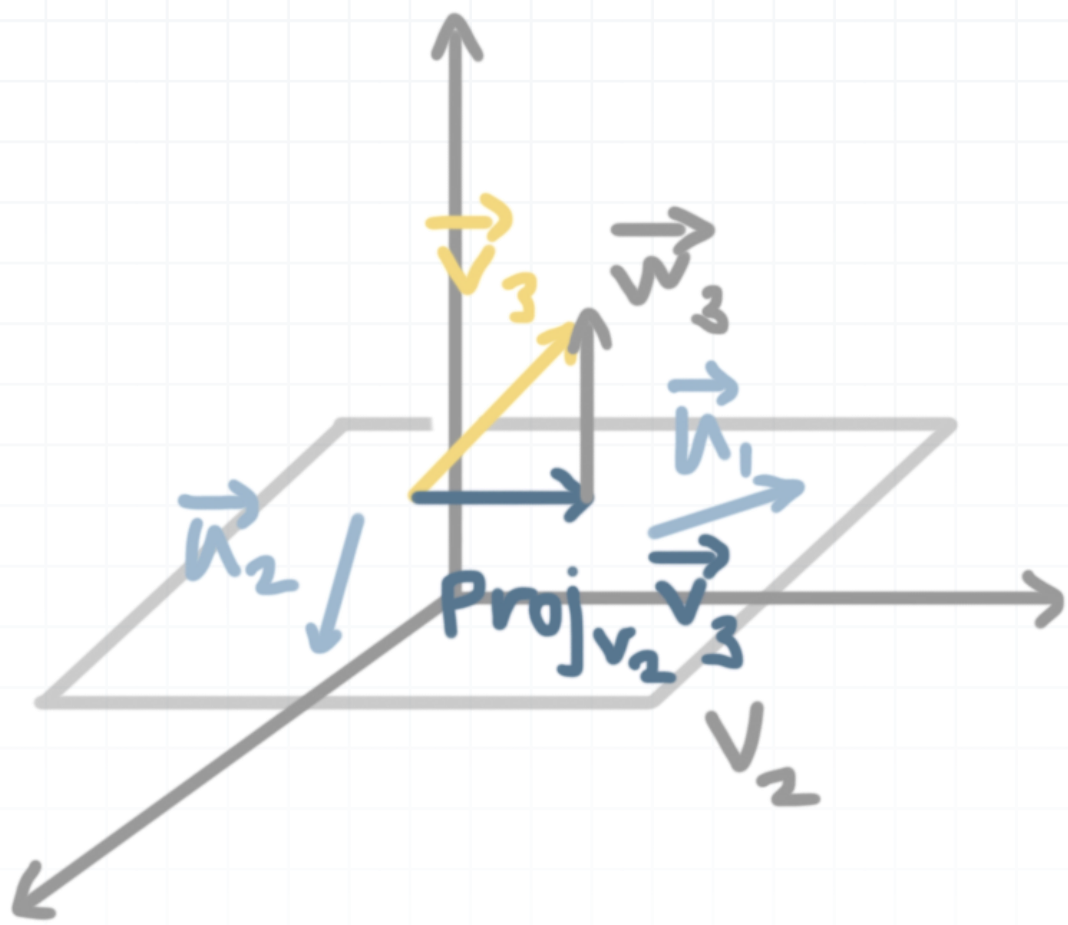
Let's imagine that $\vec{v}_3$ is another vector:



We could sketch the projection of $\vec{v}_3$ onto V_2, $\text{Proj}_{V_2}\vec{v}_3$:



Then $\vec{w}_3$ is the vector that connects $\text{Proj}_{V_2}\vec{v}_3$ to $\vec{v}_3$, so $\vec{w}_3 = \vec{v}_3 - \text{Proj}_{V_2}\vec{v}_3$,



and $\vec{w}_3$ is orthogonal to $\vec{u}_1$ and $\vec{u}_2$. But realize that V_2 is an orthonormal subspace. There are two vectors $\vec{u}_1$ and $\vec{u}_2$ that form the basis for V_2, and those two vectors are orthogonal to one another. And we know that $\vec{u}_1$ and $\vec{u}_2$ are normal, so the basis of V_2 is orthonormal. Which means that we could rewrite the projection $\text{Proj}_{V_2}\vec{v}_3$ as

$$\vec{w}_3 = \vec{v}_3 - \left[(\vec{v}_3 \cdot \vec{u}_1)\vec{u_1} + (\vec{v}_3 \cdot \vec{u}_2)\vec{u_2}\right]$$

This will give us a vector $\vec{w}_3$ that we can use in place of $\vec{v}_3$. Once we normalize $\vec{w}_3$ the same way we normalized $\vec{v}_1$ and $\vec{w}_2$, we'll call it $\vec{u}_3$, and then we'll be able to say that the basis of V can be formed by $\vec{u}_1$, $\vec{u}_2$, and $\vec{u}_3$.

$$V = \text{Span}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$$

And as you can see, if we had more basis vectors for V, we could continue this process, finding $\vec{u}_4$, $\vec{u}_5$, etc., until we've converted the entire non-orthonormal basis into an orthonormal basis.

Let's do an example so that we can see how this works with real vectors.

Example

The subspace V is a plane in $\mathbb{R}^3$. Use a Gram-Schmidt process to change the basis of V into an orthonormal basis.

$$V = \text{Span}\left(\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}-2\\1\\-5\end{bmatrix}\right)$$

Let's define $\vec{v}_1 = (1,2,0)$ and $\vec{v}_2 = (-2,1,-5)$.

$$V = \text{Span}(\vec{v}_1, \vec{v}_2)$$

Then we'll start by normalizing $\vec{v}_1$. The length of $\vec{v}_1$ is

$$||\vec{v}_1|| = \sqrt{1^2 + 2^2 + 0^2} = \sqrt{1 + 4 + 0} = \sqrt{5}$$

Then if $\vec{u}_1$ is the normalized version of $\vec{v}_1$, we can say

$$\vec{u}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\\0\end{bmatrix}$$

So we can say that V is spanned by $\vec{u}_1$ and $\vec{v}_2$.

$$V = \text{Span}(\vec{u}_1, \vec{v}_2)$$

Now all we need to do is replace $\vec{v}_2$ with a vector that's both orthogonal to $\vec{u}_1$, and normal. If we can do that, then the vector set that spans V will be orthonormal. We'll name $\vec{w}_2$ as the vector that connects $\text{Proj}_{V_1}\vec{v}_2$ to $\vec{v}_2$.

$$\vec{w}_2 = \vec{v}_2 - \text{Proj}_{V_1}\vec{v}_2$$

$$\vec{w}_2 = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1)\vec{u_1}$$

Plug in the values we already have.

$$\vec{w}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} - \left(\begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} \cdot \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right) \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} - \frac{1}{5} \left(\begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} - \frac{1}{5}(-2(1) + 1(2) - 5(0)) \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} - \frac{1}{5}(-2 + 2 - 0) \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} - \frac{1}{5}(0) \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} - 0 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}$$

This vector $\vec{w}_2$ is orthogonal to $\vec{u}_1$, but it hasn't been normalized, so let's normalize it. The length of $\vec{w}_2$ is

$$||\vec{w}_2|| = \sqrt{(-2)^2 + (1)^2 + (-5)^2}$$

$$||\vec{w}_2|| = \sqrt{4 + 1 + 25}$$

$$||\vec{w}_2|| = \sqrt{30}$$

Then the normalized version of $\vec{w}_2$ is $\vec{u}_2$:

$$\vec{u}_2 = \frac{1}{\sqrt{30}} \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}$$

Therefore, we can say that $\vec{u}_1$ and $\vec{u}_2$ form an orthonormal basis for V.

$$V = \text{Span}\left(\frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{30}} \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}\right)$$

$$V = \text{Span}\left(\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ \frac{0}{\sqrt{5}} \end{bmatrix}, \begin{bmatrix} -\frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \\ -\frac{5}{\sqrt{30}} \end{bmatrix}\right)$$

$$V = \text{Span}\left(\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \\ -\frac{5}{\sqrt{30}} \end{bmatrix}\right)$$