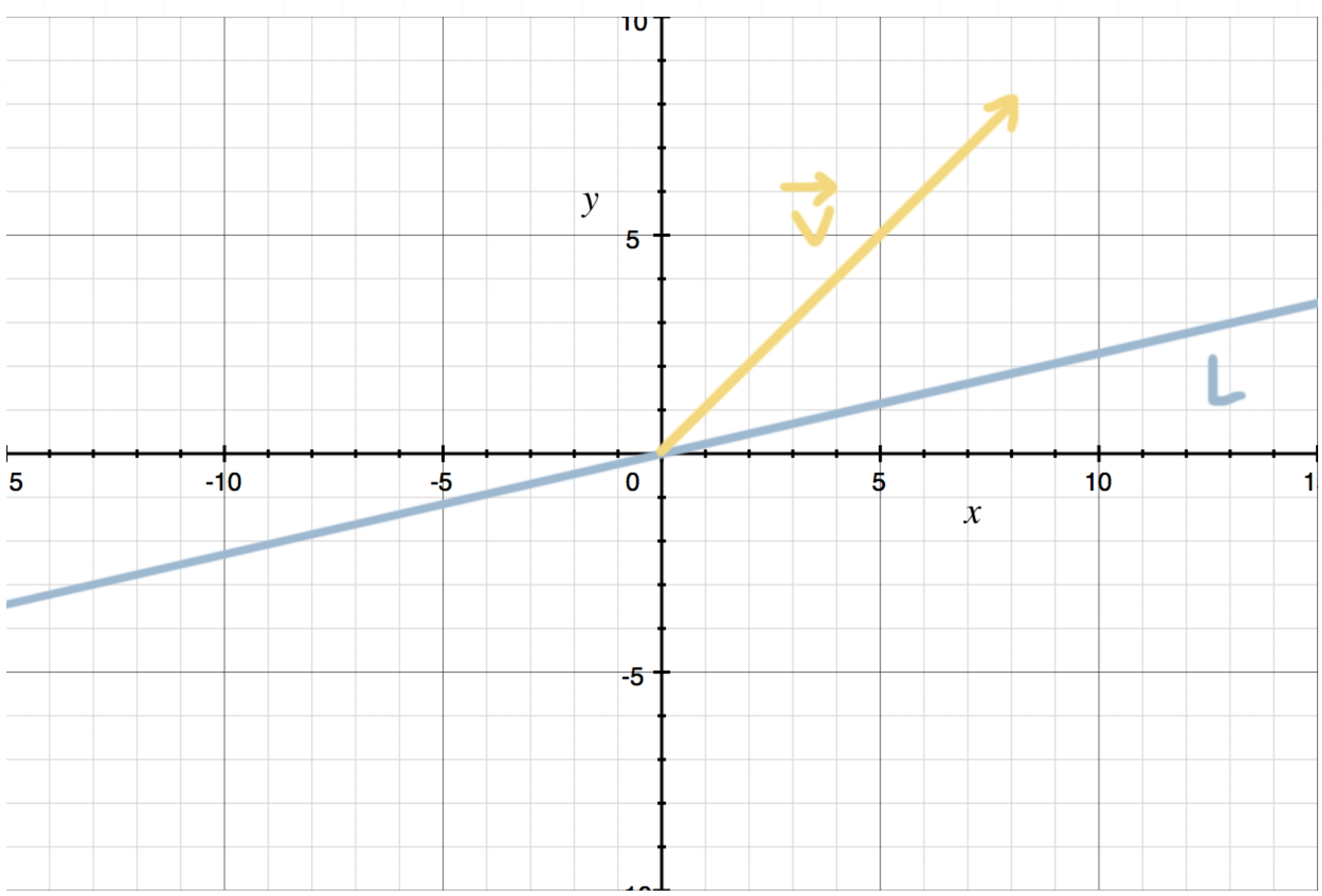
Projections as linear transformations

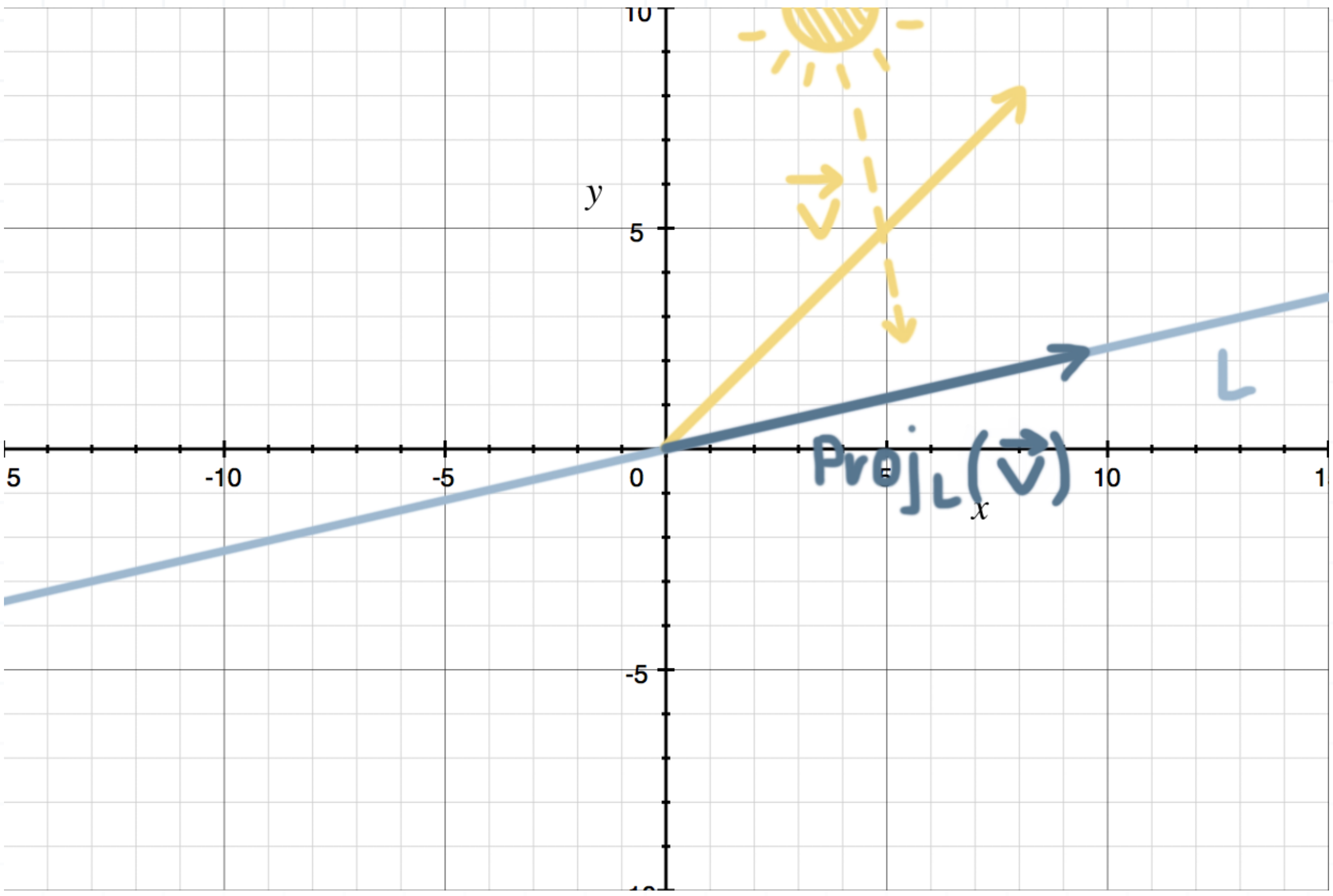
Now we want to start looking at some applications of linear transformations. In this lesson, we want to define a projection (of a vector onto a line), and show how that projection can always be expressed as a linear transformation.

The projection of a vector onto a line

Let's say you're given some line L in two-dimensional space, and a vector $\vec{v}$, like this:



Think of the **projection** of $\vec{v}$ onto L, which we write as $\text{Proj}_L(\vec{v})$, as the shadow that $\vec{v}$ casts onto L. When we imagine this shadow, think about shining a light that's above the vector $\vec{v}$ but also perpendicular to L.



If we name a vector $\vec{x}$ that lies on L, then we could say that the projection of $\vec{v}$ onto L is just some scaled version of $\vec{x}$, because L is equal to the set of all scalar multiples of the vector $\vec{x}$. And we could therefore write the projection as

$$\text{Proj}_L(\vec{v}) = c\vec{x}$$

The vector $\text{Proj}_L(\vec{v})$ is some vector in L, where $\vec{v} - \text{Proj}_L(\vec{v})$ is orthogonal to L. And if we do some work with $\vec{x}$, $\vec{v}$, and c, we can actually determine the value of c as

$$c = \frac{\vec{v} \cdot \vec{x}}{\vec{x} \cdot \vec{x}}$$

So the projection of $\vec{v}$ onto L is given by

$$\text{Proj}_L(\vec{v}) = \left(\frac{\vec{v} \cdot \vec{x}}{\vec{x} \cdot \vec{x}}\right)\vec{x}$$

Let's do an example with another diagram so that we can see what it looks like to actually find the projection vector.

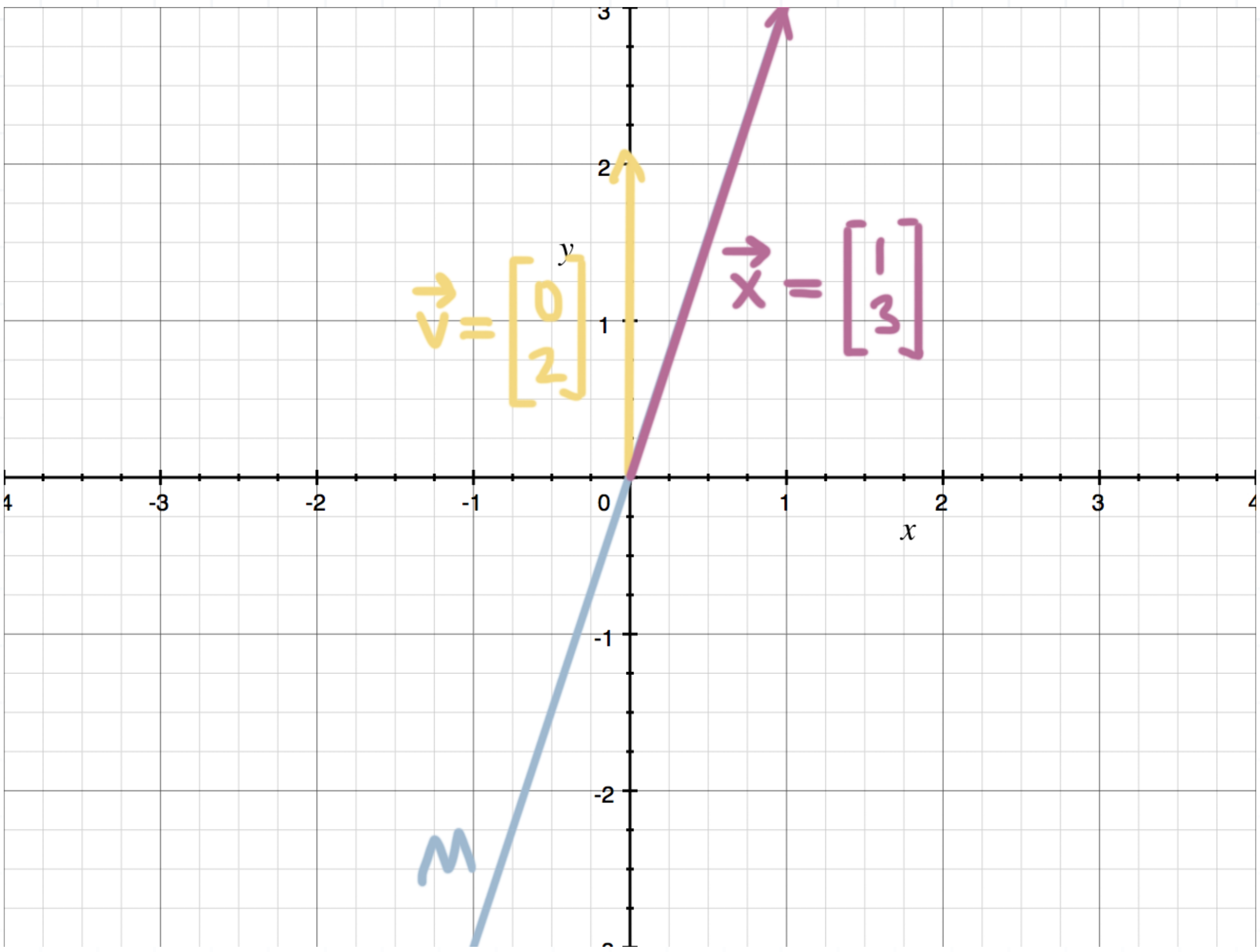
Example

Find the projection of $\vec{v}$ onto M.

$$M = \left\{ c \begin{bmatrix} 1 \\ 3 \end{bmatrix} \ \middle| \ c \in \mathbb{R} \right\}$$

$$\vec{v} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

The line M is given as all the scaled versions of the vector $\vec{x} = (1,3)$. So we could sketch M, $\vec{v}$, and $\vec{x}$ as



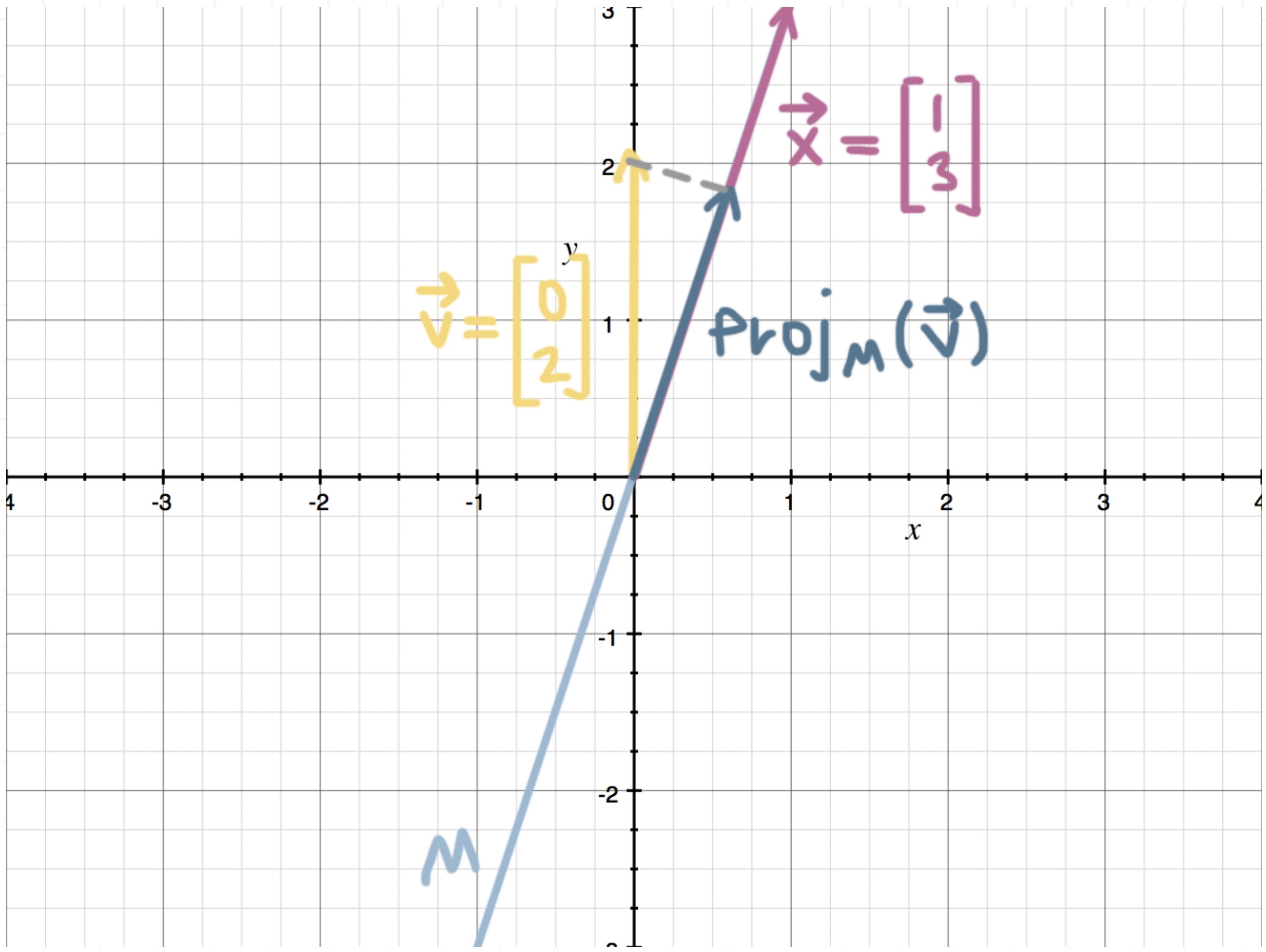
Then the projection of $\vec{v}$ onto M is given by

$$\text{Proj}_M(\vec{v}) = \left(\frac{\vec{v}\cdot\vec{x}}{\vec{x}\cdot\vec{x}}\right)\vec{x}$$

$$\text{Proj}_M(\vec{v}) = \frac{\begin{bmatrix}0 & 2\end{bmatrix}\cdot\begin{bmatrix}1\\3\end{bmatrix}}{\begin{bmatrix}1 & 3\end{bmatrix}\cdot\begin{bmatrix}1\\3\end{bmatrix}}\cdot\begin{bmatrix}1\\3\end{bmatrix} = \frac{0(1)+2(3)}{1(1)+3(3)}\cdot\begin{bmatrix}1\\3\end{bmatrix}$$

$$\text{Proj}_M(\vec{v}) = \frac{6}{10}\cdot\begin{bmatrix}1\\3\end{bmatrix} = \frac{3}{5}\cdot\begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}\frac{3}{5}\\\frac{9}{5}\end{bmatrix} = \begin{bmatrix}0.6\\1.8\end{bmatrix}$$

We can sketch the projection, $\text{Proj}_M(\vec{v})$,



and we can see how it confirms our idea of the projection as the shadow of $\vec{v}$ onto M.

Normalizing the vector that defines the line

At this point, you might be wondering how we can always get a correct value for the projection, when the vector we use to define the line could change to any vector that's collinear with the line.

After all, in the last example, we defined the line with $\vec{x} = (1,3)$, but we could just have easily defined the line M as $\vec{x} = (2,6)$, $\vec{x} = (-1, -3)$, or any other of the infinitely many vectors that lie along M.

While the projection formula still works, regardless of the vector we pick to lie along M, this brings up the interesting point that we could have picked a unit vector instead. In other words, given a vector $\vec{x}$ that lies along the line, we could normalize it to its corresponding unit vector using

$$\hat{u} = \frac{1}{||\vec{x}||}\vec{x}$$

If we do, then we can actually simplify the projection formula. First, we know that a vector dotted with itself is equivalent to the square of that vector's length. So the projection formula can be rewritten as

$$\text{Proj}_L(\vec{v}) = \left(\frac{\vec{v} \cdot \vec{x}}{\vec{x} \cdot \vec{x}}\right)\vec{x}$$

$$\text{Proj}_L(\vec{v}) = \left(\frac{\vec{v} \cdot \vec{x}}{||\vec{x}||^2}\right)\vec{x}$$

If we've normalized $\vec{x}$, then the length of $\vec{x}$ is 1, which means the projection formula becomes

$$\text{Proj}_L(\vec{v}) = \left(\frac{\vec{v} \cdot \vec{x}}{1^2}\right)\vec{x}$$

$$\text{Proj}_L(\vec{v}) = (\vec{v} \cdot \vec{x})\vec{x}$$

Furthermore, we've changed $\vec{x}$ into the unit vector $\hat{u}$, so the projection formula is

$$\text{Proj}_L(\vec{v}) = (\vec{v} \cdot \hat{u})\hat{u}$$

So let's say we continue on with the previous example. We were told that the line M was defined by the vector $\vec{x}=(1,3)$. We could have normalized that vector as

$$||\vec{x}||=\sqrt{1^2+3^2}=\sqrt{1+9}=\sqrt{10}$$

Then the unit vector associated with $\vec{x}=(1,3)$ would be

$$\hat{u}=\frac{1}{\sqrt{10}}\begin{bmatrix}1\\3\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{10}}\\\frac{3}{\sqrt{10}}\end{bmatrix}$$

And the projection of $\vec{v}=(0,2)$ onto M would have been

$$\text{Proj}_M(\vec{v})=(\vec{v}\cdot\hat{u})\hat{u}$$

$$\text{Proj}_M(\vec{v})=\left(\begin{bmatrix}0 & 2\end{bmatrix}\cdot\begin{bmatrix}\frac{1}{\sqrt{10}}\\\frac{3}{\sqrt{10}}\end{bmatrix}\right)\begin{bmatrix}\frac{1}{\sqrt{10}}\\\frac{3}{\sqrt{10}}\end{bmatrix}$$

$$\text{Proj}_M(\vec{v})=\left(0\left(\frac{1}{\sqrt{10}}\right)+2\left(\frac{3}{\sqrt{10}}\right)\right)\begin{bmatrix}\frac{1}{\sqrt{10}}\\\frac{3}{\sqrt{10}}\end{bmatrix}$$

$$\text{Proj}_M(\vec{v})=\frac{6}{\sqrt{10}}\begin{bmatrix}\frac{1}{\sqrt{10}}\\\frac{3}{\sqrt{10}}\end{bmatrix}$$

$$\text{Proj}_M(\vec{v}) = \begin{bmatrix} \frac{6}{10} \\ \frac{18}{10} \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{9}{5} \end{bmatrix} = \begin{bmatrix} 0.6 \\ 1.8 \end{bmatrix}$$

The projection is a linear transformation

As it turns out, the projection of a vector $\vec{v}$ onto the line L is always a linear transformation. Since it's a linear transformation, of course that means that it's closed under addition and closed under scalar multiplication. It's closed under addition because the projection of the sum of two vectors is equivalent to the projection of each vector individually.

$$\text{Proj}_L(\vec{a} + \vec{b}) = ((\vec{a} + \vec{b}) \cdot \hat{u})\hat{u}$$

$$\text{Proj}_L(\vec{a} + \vec{b}) = (\vec{a} \cdot \hat{u} + \vec{b} \cdot \hat{u})\hat{u}$$

$$\text{Proj}_L(\vec{a} + \vec{b}) = (\vec{a} \cdot \hat{u})\hat{u} + (\vec{b} \cdot \hat{u})\hat{u}$$

$$\text{Proj}_L(\vec{a} + \vec{b}) = \text{Proj}_L(\vec{a}) + \text{Proj}_L(\vec{b})$$

And it's closed under scalar multiplication because the projection of the product of a scalar and a vector is equivalent to the product of that scalar and the projection of the vector.

$$\text{Proj}_L(c\vec{a}) = (c\vec{a} \cdot \hat{u})\hat{u}$$

$$\text{Proj}_L(c\vec{a}) = c(\vec{a} \cdot \hat{u})\hat{u}$$

$$\text{Proj}_L(c\vec{a}) = c\text{Proj}_L(\vec{a})$$

Because a projection is always a linear transformation, that means we must be able to express it as a matrix-vector product. In other words, $\text{Proj}_L(\overrightarrow{v}) = (\overrightarrow{v} \cdot \hat{u})\hat{u} = A\overrightarrow{v}$, and we just need to find the matrix A.

We can do this in n dimensions ($\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{R}^4$, etc.), but let's assume for now that the projection is a transformation from $\mathbb{R}^2 \to \mathbb{R}^2$. Then to find A, we start with the I_2 identity matrix, and we'll call the unit vector $\hat{u} = (u_1, u_2)$. The matrix A is then

$$A = \begin{bmatrix} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} & \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \end{bmatrix}$$

$$A = \begin{bmatrix} \left(1(u_1) + 0(u_2)\right) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} & \left(0(u_1) + 1(u_2)\right) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \end{bmatrix}$$

$$A = \begin{bmatrix} u_1 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} & u_2 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \end{bmatrix}$$

$$A = \begin{bmatrix} u_1^2 & u_2 u_1 \\ u_1 u_2 & u_2^2 \end{bmatrix}$$

In other words, given any unit vector $\hat{u} = (u_1, u_2)$ along the line L, the projection of the vector $\overrightarrow{v}$ onto L, assuming that we've defined L by a unit vector $\hat{u}$, is

$$\text{Proj}_L(\overrightarrow{v}) = (\overrightarrow{v} \cdot \hat{u})\hat{u} = A\overrightarrow{v} = \begin{bmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{bmatrix} \overrightarrow{v}$$

If we go back to the example we were working with earlier, we were trying to project $\vec{v} = (0,2)$ onto M, where M was defined by the unit vector

$$\hat{u} = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{bmatrix}$$

So using the new formula for the projection as the matrix-vector product, we could say that the projection of $\vec{v}$ onto M is

$$\text{Proj}_M(\vec{v}) = \begin{bmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{bmatrix} \vec{v}$$

$$\text{Proj}_M(\vec{v}) = \begin{bmatrix} \left(\frac{1}{\sqrt{10}}\right)^2 & \left(\frac{1}{\sqrt{10}}\right)\left(\frac{3}{\sqrt{10}}\right) \\ \left(\frac{1}{\sqrt{10}}\right)\left(\frac{3}{\sqrt{10}}\right) & \left(\frac{3}{\sqrt{10}}\right)^2 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

First find the matrix A.

$$A = \begin{bmatrix} \left(\frac{1}{\sqrt{10}}\right)^2 & \left(\frac{1}{\sqrt{10}}\right)\left(\frac{3}{\sqrt{10}}\right) \\ \left(\frac{1}{\sqrt{10}}\right)\left(\frac{3}{\sqrt{10}}\right) & \left(\frac{3}{\sqrt{10}}\right)^2 \end{bmatrix}$$

$$A = \begin{bmatrix} \frac{1}{10} & \frac{3}{10} \\ \frac{3}{10} & \frac{9}{10} \end{bmatrix}$$

Then the projection of $\vec{v}=(0,2)$ onto M is

$$\text{Proj}_M(\vec{v})=\begin{bmatrix}\frac{1}{10} & \frac{3}{10}\\ \frac{3}{10} & \frac{9}{10}\end{bmatrix}\begin{bmatrix}0\\2\end{bmatrix}$$

$$\text{Proj}_M(\vec{v})=\begin{bmatrix}\frac{1}{10}(0)+\frac{3}{10}(2)\\ \frac{3}{10}(0)+\frac{9}{10}(2)\end{bmatrix}$$

$$\text{Proj}_M(\vec{v})=\begin{bmatrix}\frac{6}{10}\\ \frac{18}{10}\end{bmatrix}=\begin{bmatrix}\frac{3}{5}\\ \frac{9}{5}\end{bmatrix}=\begin{bmatrix}0.6\\1.8\end{bmatrix}$$

Not only did we get the same result for the projection as what we found earlier, but we also found a matrix for the projection of *any* vector onto M. Now, given any vector $\vec{v}$ in $\mathbb{R}^2$ that we want to project onto M, that projection will always be given by

$$\text{Proj}_M(\vec{v})=\begin{bmatrix}\frac{1}{10} & \frac{3}{10}\\ \frac{3}{10} & \frac{9}{10}\end{bmatrix}\vec{v}$$

In other words, the reason it's helpful to rewrite the projection as a matrix-vector product is because it generalizes the projection into a transformation matrix, which we can then apply to any vector we want to project onto the line.