Section 6.3

Orthogonality and the Gram-Schmidt Process

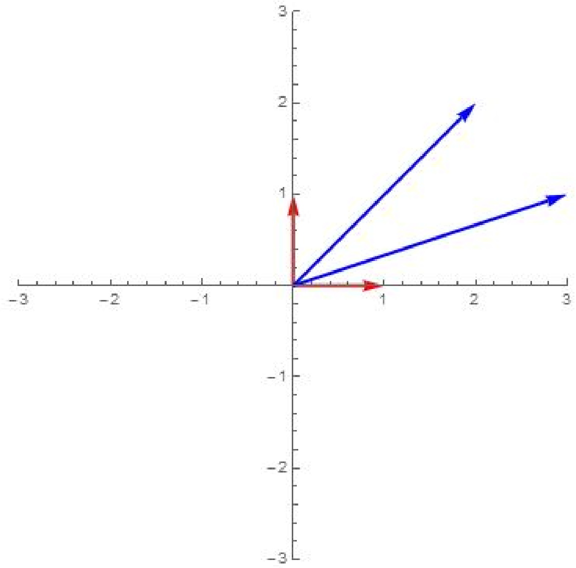
In Chapter 4, we spent a great deal of time studying the problem of finding a basis for a vector space. We know that a basis for a vector space can potentially be chosen in many different ways: for example, the sets

$$B_1 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

and

$$B_2 = \left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\}$$

both form bases for $\mathbb{R}^2$. The vectors from the first basis are graphed below in red, and the vectors from the second basis are graphed in blue:



There are two things that are particularly nice about the first basis

$$B_1 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

as compared with the second basis:

1. The vectors in B_1 are orthogonal with respect to the usual inner product (dot product) on $\mathbb{R}^2$.

2. Both of the vectors in B_1 are unit vectors (length 1) with respect to the dot product.

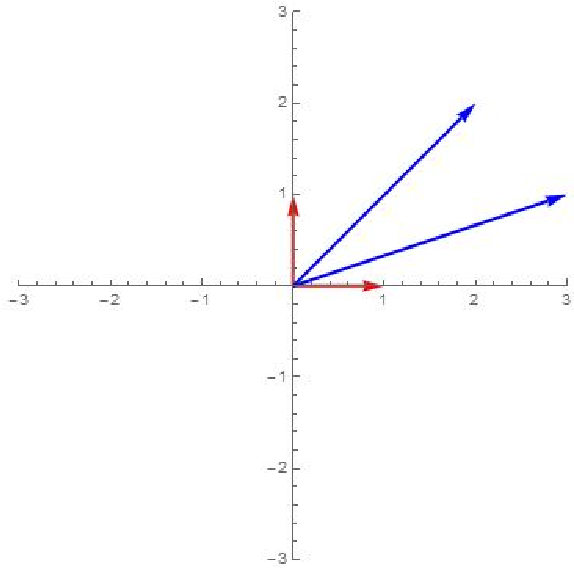
Since a basis gives us a means for representing the coordinates or location of any vector in the space, bases with the two properties listed above are especially advantageous; in this section, we will describe a method for locating bases of a vector space that have these two properties.

Orthogonal and Orthonormal Sets

It is now a convenient time to develop some vocabulary to describe the types of basis vectors that we are looking for:

Definition 1. A set of vectors in a (real) inner product space V equipped with the inner product $\langle \cdot, \cdot \rangle$ is called *orthogonal* if each distinct pair of vectors is orthogonal with respect to $\langle \cdot, \cdot \rangle$. If in addition each vector in the set has norm 1 with respect to the inner product, the set is called *orthonormal*.

Returning to our example of two different basis for $\mathbb{R}^2$, it is easy to see that the red vectors below form an orthonormal set:



They are both orthogonal and length 1, and we say that they form an orthonormal basis for $\mathbb{R}^2$.

However, the blue vectors are neither orthogonal nor length 1, thus they form neither an orthonormal nor an orthogonal basis for $\mathbb{R}^2$.

Fortunately, even if vector $\mathbf{v}$ is not a unit vector, there is a simple algorithm for finding a unit vector preserving the direction of $\mathbf{v}$: if $\mathbf{v}$ is not the zero vector, so that

$$||\mathbf{v}|| \neq 0,$$

then the vector

$$\mathbf{u_v} = \frac{1}{||\mathbf{v}||}\mathbf{v}$$

is a unit vector, and has the same direction as $\mathbf{v}$.

We can quickly check that each of these claims is true: first, let's calculate the length of $\mathbf{u_v}$:

$$\begin{aligned} ||\mathbf{u_v}|| &= ||\frac{1}{||\mathbf{v}||}\mathbf{v}|| \\ &= \frac{1}{||\mathbf{v}||}||\mathbf{v}|| \quad \text{since } \frac{1}{||\mathbf{v}||} \text{ is a nonnegative scalar} \\ &= 1, \end{aligned}$$

so that $\mathbf{u_v}$ is indeed a unit vector.

Let's check that $\mathbf{u_v}$ and $\mathbf{v}$ have the same direction by calculating the angle θ between them: recall that

$$\begin{aligned} \theta &= \cos^{-1}\left(\frac{\langle \mathbf{u_v}, \mathbf{v}\rangle}{||\mathbf{u}||\ ||\mathbf{v}||}\right) \\ &= \cos^{-1}\left(\frac{\langle \frac{1}{||\mathbf{v}||}\mathbf{v}, \mathbf{v}\rangle}{||\mathbf{v}||}\right) \\ &= \cos^{-1}\left(\frac{\langle \mathbf{v}, \mathbf{v}\rangle}{||\mathbf{v}||^2}\right) \\ &= \cos^{-1}\left(\frac{\langle \mathbf{v}, \mathbf{v}\rangle}{\left(\sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}\right)^2}\right) \\ &= \cos^{-1}\left(\frac{\langle \mathbf{v}, \mathbf{v}\rangle}{\langle \mathbf{v}, \mathbf{v}\rangle}\right) \\ &= \cos^{-1} 1 \\ &= 0. \end{aligned}$$

Since the angle θ between $\mathbf{u_v}$ and $\mathbf{v}$ is 0, the vectors clearly have the same direction.

The process of finding a unit vector with the same direction as $\mathbf{v}$ is called *normalizing* $\mathbf{v}$.

Key Point. We can normalize any nonzero vector $\mathbf{v}$ by calculating

$$\mathbf{u_v} = \frac{1}{||\mathbf{v}||}\mathbf{v}.$$

The normalized vector $\mathbf{u_v}$ is the unit vector in the same direction as $\mathbf{v}$.

Regardless of their length, orthogonal vectors are extremely important for the following reason:

Theorem 1.6.3. If $S = \{\mathbf{v_1},\ \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is an orthogonal set of (nonzero) vectors in a inner product space V, then the S is a linearly independent set.

The theorem leads to a helpful observation:

Key Point. A set of n orthogonal vectors in an n dimensional inner product space V is a basis for V.

Example

The vectors

$$\mathbf{f} = f(x) = 2 + x^2, \ \mathbf{g} = g(x) = 2x, \ \text{and } \mathbf{h} = h(x) = -1 + 2x^2$$

form a basis for P_2.

1. Is the basis an orthogonal basis under the usual inner product on P_2?
2. Is the basis an orthonormal basis?
3. If it is orthogonal but not orthonormal, use the vectors above to find a basis for P_2 that *is* orthonormal.

Recall that the standard inner product on P_2 is defined on vectors

$$\mathbf{f} = f(x) = a_0 + a_1x + a_2x^2 \text{ and } \mathbf{g} = g(x) = b_0 + b_1x + b_2x^2$$

in P_2 by

$$\langle \mathbf{f}, \mathbf{g} \rangle = a_0b_0 + a_1b_1 + a_2b_2.$$

To answer all of the questions above, we'll need to calculate the lengths of each vector and the angles between each pair of vectors; since both calculations involve the inner products, we record each pair of inner products on the chart below:

$\langle \cdot, \cdot \rangle$	$\mathbf{f}$	$\mathbf{g}$	$\mathbf{h}$
$\mathbf{f}$	5	0	0
$\mathbf{g}$	0	4	0
$\mathbf{h}$	0	0	5

Since the length of a vector $\mathbf{f}$ is given by

$$||\mathbf{f}|| = \sqrt{\langle \mathbf{f}, \mathbf{f} \rangle},$$

we see that

$$||\mathbf{f}|| = \sqrt{5}, \ ||\mathbf{g}|| = 2, \ \text{and } ||\mathbf{h}|| = \sqrt{5}.$$

The angle between a pair of vectors is given by

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{f}, \mathbf{g} \rangle}{||\mathbf{f}|| \ ||\mathbf{g}||} \right);$$

however, inspecting the chart above, we see that

$$\langle \mathbf{f}, \mathbf{g} \rangle = \langle \mathbf{g}, \mathbf{h} \rangle = \langle \mathbf{f}, \mathbf{h} \rangle = 0.$$

Since $\cos^{-1} 0 = \frac{\pi}{2}$, the three vectors are clearly mutually orthogonal.

We are now ready to answer each of the questions posed above:

1. *Is the basis an orthogonal basis under the usual inner product on P_2?*

   Since the basis vectors are mutually orthogonal, the basis is an orthogonal basis.

2. *Is the basis an orthonormal basis?*

   None of the basis vectors are unit vectors, so it is orthogonal but not orthonormal.

3. *If it is orthogonal but not orthonormal, use the vectors above to find a basis for P_2 that* is *orthonormal.*

   We can use the vectors above to create unit vectors using the formula

   $$\mathbf{u_v} = \frac{\mathbf{v}}{||\mathbf{v}||}.$$

   Since $||\mathbf{f}|| = \sqrt{5}$, $||\mathbf{g}|| = 2$, and $||\mathbf{h}|| = \sqrt{5}$, the vectors

   $$\mathbf{u_f} = \frac{2}{\sqrt{5}} + \frac{1}{\sqrt{5}}x^2, \ \mathbf{u_g} = x, \text{ and } \mathbf{u_h} = -\frac{1}{\sqrt{5}} + \frac{2}{\sqrt{5}}x^2$$

   are unit vectors. Since they preserve the directions of $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{h}$ respectively, $\mathbf{u_f}$, $\mathbf{u_g}$, and $\mathbf{u_h}$ are mutually orthogonal (and so are linearly independent as well as a basis for the three dimensional P_2). Since they are also unit vectors, they form an orthonormal basis for P_2.

Examples of Orthonormal Bases

Standard Basis for $\mathbb{R}^n$

The standard basis vectors

$$\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \text{ and } \mathbf{e_n} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

of $\mathbb{R}^n$ are mutually orthogonal unit vectors under the standard inner product (dot product) on $\mathbb{R}^n$, thus form an orthonormal basis for $\mathbb{R}^n$.

Standard Basis for P_n

The standard basis vectors

$$1, \ x, \ x^2, \ldots, \text{ and } x^n$$

of the vector space P_n of all polynomials of degree no more than n are mutually orthogonal unit vectors under the standard inner product on P_n defined by

$$\langle \mathbf{f}, \mathbf{g} \rangle = a_0 b_0 + a_1 b_1 + \ldots + a_n b_n.$$

An Orthonormal Basis for $\mathcal{M}_n$

Let $\mathbf{e_{ij}}$ indicate the $n \times n$ matrix whose i, j entry is 1, and all of whose other entries are 0. For example, $\mathbf{e_{23}}$ is given by

$$\mathbf{e_{32}} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Under the standard inner product on $\mathcal{M}_n$,

$$\langle U, V \rangle = \operatorname{tr}(U^\top V),$$

the set of all matrices of the form $\mathbf{e_{ij}}$ for $1 \leq i, j, \leq n$ forms an orthonormal basis for $\mathcal{M}_n$.

Coordinates and Orthogonal/Orthonormal Bases

In Section 4.4, we learned that if the (ordered) set $S = \{\mathbf{v_1}, \ \dots, \ \mathbf{v_n}\}$ of vectors in V forms a basis for V, so that *any* vector $\mathbf{u}$ in V is a linear combination of the basis vectors, such as

$$\mathbf{u} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n},$$

then we can use the coefficients c_1, c_2, etc. of the basis vectors to write the coordinates of $\mathbf{u}$ under the basis as

$$(\mathbf{u})_S \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

As we saw in Section 4.4, it can often be complicated to find the desired coordinates for a given vector. However, *if* the basis being used for the coordinate representation is orthogonal or orthonormal, there is a simple formula for finding the coordinates of a given vector:

Theorem 6.3.2. Let $S = \{\mathbf{v_1}, \ \dots, \ \mathbf{v_n}\}$ be a basis for an inner product space V under the inner product $\langle \cdot, \cdot \rangle$.

1. If S is an orthogonal basis and $\mathbf{u}$ is any vector in V, then $\mathbf{u}$ may be written as the linear combination

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v_1} \rangle}{||\mathbf{v_1}||^2}\mathbf{v_1} + \frac{\langle \mathbf{u}, \mathbf{v_2} \rangle}{||\mathbf{v_2}||^2}\mathbf{v_2} + \ldots + \frac{\langle \mathbf{u}, \mathbf{v_n} \rangle}{||\mathbf{v_n}||^2}\mathbf{v_n},$$

and the coordinates of **u** under this basis are given by the matrix

$$\begin{pmatrix} \frac{\langle \mathbf{u}, \mathbf{v_1} \rangle}{||\mathbf{v_1}||^2} \\ \frac{\langle \mathbf{u}, \mathbf{v_n} \rangle}{||\mathbf{v_n}||^2} \\ \vdots \\ \frac{\langle \mathbf{u}, \mathbf{v_n} \rangle}{||\mathbf{v_n}||^2} \end{pmatrix}.$$

2. If S is an orthonormal basis and **u** is any vector in V, then **u** may be written as the linear combination

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v_1} \rangle \mathbf{v_1} + \langle \mathbf{u}, \mathbf{v_2} \rangle \mathbf{v_2} + \ldots + \langle \mathbf{u}, \mathbf{v_n} \rangle \mathbf{v_n},$$

and the coordinates of **u** under this basis are given by the matrix

$$\begin{pmatrix} \langle \mathbf{u}, \mathbf{v_1} \rangle \\ \langle \mathbf{u}, \mathbf{v_n} \rangle \\ \vdots \\ \langle \mathbf{u}, \mathbf{v_n} \rangle \end{pmatrix}.$$

Example

Find the coordinates of $\mathbf{p} = p(x) = -2 + x - 2x^2$ under the orthonormal basis

$$\{\mathbf{u_f} = \frac{2}{\sqrt{5}} + \frac{1}{\sqrt{5}}x^2, \ \mathbf{u_g} = x, \mathbf{u_h} = -\frac{1}{\sqrt{5}} + \frac{2}{\sqrt{5}}x^2\}.$$

Since the given basis is orthonormal, the coordinates of **p** are given by the matrix

$$\begin{pmatrix} \langle \mathbf{p}, \mathbf{u_f} \rangle \\ \langle \mathbf{p}, \mathbf{u_g} \rangle \\ \langle \mathbf{p}, \mathbf{u_h} \rangle \end{pmatrix}.$$

Let's make the calculations:

$$\begin{aligned}\langle \mathbf{p}, \mathbf{u_f} \rangle &= \langle -2 + x - 2x^2, \frac{2}{\sqrt{5}} + \frac{1}{\sqrt{5}}x^2 \rangle \\ &= -2 \cdot \frac{2}{\sqrt{5}} + 1 \cdot 0 + (-2) \cdot (\frac{1}{\sqrt{5}}) \\ &= -\frac{4}{\sqrt{5}} - \frac{2}{\sqrt{5}} \\ &= -\frac{6}{\sqrt{5}};\end{aligned}$$

$$\begin{aligned}\langle \mathbf{p}, \mathbf{u_g} \rangle &= \langle -2 + x - 2x^2, x \rangle \\ &= -2 \cdot 0 + 1 \cdot 1 - 2 \cdot 0 \\ &= 1;\end{aligned}$$

$$\begin{aligned}\langle \mathbf{p}, \mathbf{u_h} \rangle &= \langle -2 + x - 2x^2, -\frac{1}{\sqrt{5}} + \frac{2}{\sqrt{5}}x^2 \rangle \\ &= -2 \cdot (-\frac{1}{\sqrt{5}}) + 1 \cdot 0 - 2 \cdot (\frac{2}{\sqrt{5}}) \\ &= \frac{2}{\sqrt{5}} - \frac{4}{\sqrt{5}} \\ &= -\frac{2}{\sqrt{5}}.\end{aligned}$$

Thus the coordinates of $\mathbf{p}$ with respect to this basis are given by

$$\begin{pmatrix} \langle \mathbf{p}, \mathbf{u_f} \rangle \\ \langle \mathbf{p}, \mathbf{u_g} \rangle \\ \langle \mathbf{p}, \mathbf{u_h} \rangle \end{pmatrix} = \begin{pmatrix} -\frac{6}{\sqrt{5}} \\ 1 \\ -\frac{2}{\sqrt{5}} \end{pmatrix}.$$

Finding an Orthonormal/Orthogonal Basis for an Inner Product Space

The next theorem describes a truly beautiful fact about inner product spaces:

Theorem 6.3.5. Every (finite-dimensional) inner product space has an orthonormal basis.

Of course, the theorem leads to a question: Given an inner product space V, how do we *find* an orthonormal basis for it?

The following algorithm (one of the more important and well known algorithms in the field), called the Gram-Schmidt Process, answers this question. Given any basis for the inner product space of interest, the Gram-Schmidt Process converts it into an orthogonal (or orthonormal, if desired) basis.

The Gram-Schmidt Process. Let $\{\mathbf{u_1},\ \mathbf{u_2},\ \ldots,\ \mathbf{u_n}\}$ be any basis for an inner product space V with inner product $\langle\cdot,\cdot\rangle$. The following process uses this basis to build an orthogonal basis $\{\mathbf{v_1},\ \mathbf{v_2},\ \ldots,\ \mathbf{v_n}\}$ for V:

1. $\mathbf{v_1} = \mathbf{u_1}$
2. $\mathbf{v_2} = \mathbf{u_2} - \dfrac{\langle \mathbf{u_2}, \mathbf{v_1}\rangle}{||\mathbf{v_1}||^2}\mathbf{v_1}$
3. $\mathbf{v_3} = \mathbf{u_3} - \dfrac{\langle \mathbf{u_3}, \mathbf{v_1}\rangle}{||\mathbf{v_1}||^2}\mathbf{v_1} - \dfrac{\langle \mathbf{u_3}, \mathbf{v_2}\rangle}{||\mathbf{v_2}||^2}\mathbf{v_2}$

4-r. Repeat up to $\mathbf{v_r} = \ldots$.

To convert this basis to an orthonormal one, normalize each of the basis vectors $\mathbf{v_1}$, $\mathbf{v_2}$,..., $\mathbf{v_r}$.

Essentially, the algorithm works by projecting each old basis vector $\mathbf{u_i}$ onto the orthogonal subspace to that spanned by the previous orthogonal vectors $\mathbf{v_1},\ \ldots,\ \mathbf{v_{i-1}}$.

Example

Convert the basis

$$\left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix},\ \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\}$$

into an orthonormal one.

With

$$\mathbf{u_1} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \text{ and } \mathbf{u_2} = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

we begin the Gram-Schmidt Process by setting $\mathbf{v_1} = \mathbf{u_1}$.

Next, we calculate $\mathbf{v_2}$ using the formula

$$\mathbf{v_2} = \mathbf{u_2} - \frac{\langle \mathbf{u_2}, \mathbf{v_1} \rangle}{||\mathbf{v_1}||^2}\mathbf{v_1};$$

we need to find

$$\begin{aligned} \langle \mathbf{u_2}, \mathbf{v_1} \rangle &= 2 \cdot 3 + 2 \cdot 1 \\ &= 8 \end{aligned}$$

and

$$\begin{aligned} ||\mathbf{v_1}|| &= \sqrt{3^2 + 1^2} \\ &= \sqrt{10}. \end{aligned}$$

Thus we have

$$\begin{aligned} \mathbf{v_2} &= \mathbf{u_2} - \frac{\langle \mathbf{u_2}, \mathbf{v_1} \rangle}{||\mathbf{v_1}||^2}\mathbf{v_1} \\ &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \frac{8}{10}\begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \begin{pmatrix} \frac{12}{5} \\ \frac{4}{5} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{2}{5} \\ \frac{6}{5} \end{pmatrix}. \end{aligned}$$

Thus the set

$$\left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -\dfrac{2}{5} \\ \dfrac{6}{5} \end{pmatrix} \right\}$$

forms an orthogonal basis for $\mathbb{R}^2$; however, as we were asked to find an *orthonormal* basis, we need to normalize each of the vectors above by multiplying them by the reciprocals of their lengths. We have

$$\begin{aligned}
\frac{1}{||\mathbf{v_1}||}\mathbf{v_1} &= \frac{1}{\sqrt{10}}\begin{pmatrix}3\\1\end{pmatrix} \\
&= \begin{pmatrix}\dfrac{3}{\sqrt{10}}\\ \\ \dfrac{1}{\sqrt{10}}\end{pmatrix} \\
&= \begin{pmatrix}\dfrac{3\sqrt{10}}{10}\\ \\ \dfrac{\sqrt{10}}{10}\end{pmatrix},
\end{aligned}$$

and

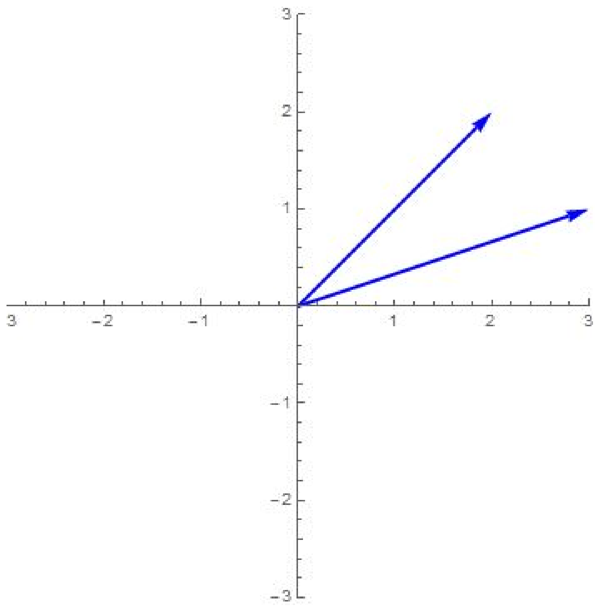
$$\begin{aligned}
\frac{1}{||\mathbf{v_2}||}\mathbf{v_2} &= \frac{5\sqrt{10}}{20}\begin{pmatrix}-\dfrac{2}{5}\\ \\ \dfrac{6}{5}\end{pmatrix} \\
&= \begin{pmatrix}-\dfrac{\sqrt{10}}{10}\\ \\ \dfrac{3\sqrt{10}}{10}\end{pmatrix}.
\end{aligned}$$

Thus the set

$$\left\{\begin{pmatrix}\dfrac{3\sqrt{10}}{10}\\ \\ \dfrac{\sqrt{10}}{10}\end{pmatrix}, \begin{pmatrix}-\dfrac{\sqrt{10}}{10}\\ \\ \dfrac{3\sqrt{10}}{10}\end{pmatrix}\right\}$$

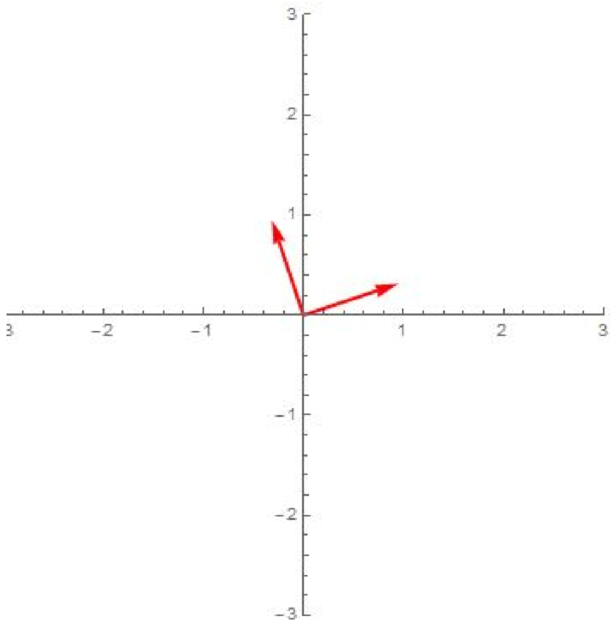
is an orthonormal basis for $\mathbb{R}^2$.

The two vectors from the *original* basis are graphed below:



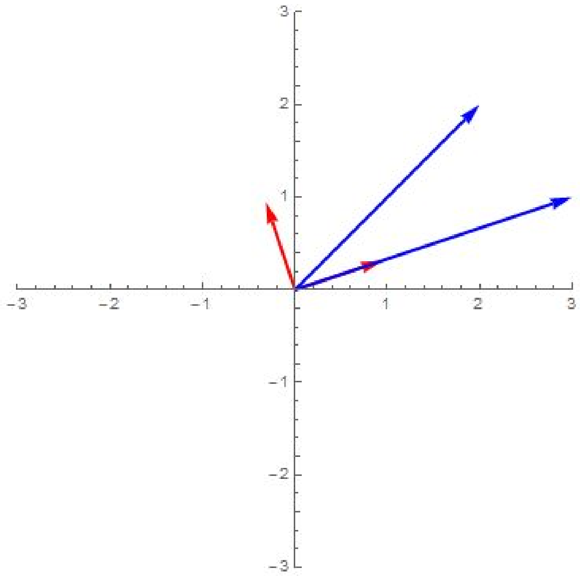
As we saw earlier, they are neither orthogonal nor orthonormal.

The new basis vectors under the Gram-Schmidt Process are graphed in red:



Notice that they are orthogonal, and that they appear to be the same length (unit vectors).

Finally, the original basis and the new basis are graphed together:



Notice that one of the new basis vectors is just a scaled-down version of one of the old vectors, whereas the second new basis vector has been moved so that it is orthogonal.