2 Bases

In this chapter, we will begin by expanding upon the topics of span and linear independence from Section 1.3 to build the concept of a basis of a vector space. This will allow us to revisit our geometric tools from Section 1.2. We'll see that there are (usually) perfectly reasonable ways to think about and make use of the geometry of *any* vector space. First, we have a bit of organizing to do.

We'll need some basic definitions. [1]

1: We leave it to the reader to determine for themselves whether that qualifies as wordplay.

2.1 Introduction to Bases

The following definition builds off of our discussion of the span of a set of vectors in both Sections 1.3 and 1.4.

Definition 2.1.1 *Let V be a vector space and $\vec{v}_1, \ldots, \vec{v}_p \in V$. We say the vectors $\vec{v}_1, \ldots, \vec{v}_p$* ***span*** *$V$ if*

$$\operatorname{Span}\{\vec{v}_1, \ldots, \vec{v}_p\} = V.$$

In this situation, we call $\{\vec{v}_1, \ldots, \vec{v}_p\}$ a ***spanning set*** *for V.*

Compare this with Definition 1.3.2.[2] The difference is subtle. Basically, all we've done here is make verb and adjective versions of the noun "span."

2: Did you notice citations like this are hyperlinked? You're welcome.

Example 2.1.1 Let's consider the vectors

$$\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Is the set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ a spanning set for $\mathbb{R}^2$?

To answer this question, we should first try to better understand $\operatorname{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Because $\vec{v}_1 = \vec{v}_2 - \vec{v}_3$, we know $\operatorname{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \operatorname{Span}\{\vec{v}_2, \vec{v}_3\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_3\} = \operatorname{Span}\{\vec{v}_1, \vec{v}_2\}$. While it is not necessary to reduce our set, it will simplify the algebra involved in the problem since two vectors are easier to work with than three.

Let's now show that $\operatorname{Span}\{\vec{v}_1, \vec{v}_2\}$ is a spanning set for $\mathbb{R}^2$. Since $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^2$, we know automatically that $\operatorname{Span}\{\vec{v}_1, \vec{v}_2\} \subseteq \mathbb{R}^2$. So we only need to check that $\mathbb{R}^2 \subseteq \text{Span}\{\vec{v}_1, \vec{v}_2\}$ to have the desired set equality. To do this, we choose a general vector in $\mathbb{R}^2$ and show it can be obtained as a linear combination of $\vec{v}_1$ and $\vec{v}_2$. For our general vector, we choose

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

where $x_1, x_2 \in \mathbb{R}$. Then we need to find $a, b \in \mathbb{R}$ such that $a\vec{v}_1 + b\vec{v}_2 = \vec{x}$. That is,

$$a \begin{bmatrix} -1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

This gives us the equations $-a = x_1$ and $a + b = x_2$. Our goal is to find a and b in terms of x_1 and x_2. We quickly have $a = -x_1$, and we can substitute and rearrange to get $b = x_1 + x_2$. This gives us a way to write any vector in $\mathbb{R}^2$ as a linear combination of $\vec{v}_1$ and $\vec{v}_2$, so $\mathbb{R}^2 \subseteq \text{Span}\{\vec{v}_1, \vec{v}_2\}$. Since we started by arguing that $\text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \text{Span}\{\vec{v}_1, \vec{v}_2\}$, we now know that $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is also a spanning set.

Exploration 37 Use the equations solved for in the example above to write

$$\begin{bmatrix} 2 \\ 10 \end{bmatrix}$$

as a linear combination of

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Spanning sets may be a convenient way to describe a vector space, but this method doesn't preclude us from doing silly things like

$$\mathbb{R}^2 = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 1,234 \end{bmatrix} \right\}.$$

It's probably a good time to decide what makes a spanning set a "good" spanning set, where by good, we mean avoiding inefficiencies like the preceding example. Surely we don't need more than one thousand vectors to span $\mathbb{R}^2$! What exactly do we need then? Well, in Example 2.1.1 we used the fact that the set was linearly dependent to reduce it to a more manageable size, and we can do this in general. That means linearly independent sets are more desirable, since they cannot be reduced down further without changing the span.

Definition 2.1.2 *Let V be a vector space. A finite set of vectors $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ is a* ***basis*** *for V if*

(a) *$\mathcal{B}$ is linearly independent, and*
(b) *$\mathcal{B}$ spans V.*

Any set satisfying the above definition is a basis. However, when we are using a specific basis in this text, we will implicitly impose an ordering on the elements of the set as determined by the order they are listed. This is sometimes referred to as an "ordered basis," but we will not use this terminology. Just know that for us, if the set $\{\vec{v}_1, \vec{v}_2\}$ is a basis, then the set $\{\vec{v}_2, \vec{v}_1\}$ is also a basis. However, the orderings on these two bases[3] are different.

3: The plural of "basis" is "bases," pronounced with a long e.

Now, we've claimed that a basis is better than just any old spanning set, and it is. However, before we tell you why, let's spend some time with the definition and some specific examples. How exactly do you show a set meets these conditions?

…So You Think Your Set's a Basis

Example 2.1.2 Let $\mathcal{B} = \{\vec{p}_1 = 1 + x, \vec{p}_2 = x + x^2, \vec{p}_3 = x\}$. This is a basis for the vector space $\mathbb{P}_2$. To show this using the definition of basis, we'd need to show that $\mathcal{B}$ is linearly independent and that it spans $\mathbb{P}_2$. Let's verify!

- *Linearly Independent:* Suppose $a\vec{p}_1 + b\vec{p}_2 + c\vec{p}_3 = \vec{0}$. Then we have $a(1+x)+b(x+x^2)+c(x) = 0$ which simplifies to $a+(a+b+c)x + bx^2 = 0$. For this equation to be true, the coefficients on each of the terms must be zero. This can only happen when $a = b = 0$. Thus, we also see that $c = 0$ since $a + b + c = 0$. The only solution is then $a = b = c = 0$, where all coefficients are zero.
- *Spans* $\mathbb{P}_2$. Let $a_0 + a_1x + a_2x^2$ be any polynomial in $\mathbb{P}_2$. Then we have

$$\begin{aligned} &a_0\vec{p}_1 + a_2\vec{p}_2 + (a_1 - a_2 - a_0)\vec{p}_3 \\ &\quad = a_0(1 + x) + a_2(x + x^2) + (a_1 - a_2 - a_0)x \\ &\quad = a_0 + a_1x + a_2x^2 \end{aligned}$$

  This gives us a recipe for writing any vector of $\mathbb{P}_2$ as a linear combination of the vectors in $\mathcal{B}$, so $\mathcal{B}$ spans $\mathbb{P}_2$!

Exploration 38 Let's do another one! Let

$$\mathcal{B} = \left\{ \vec{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

Verify this is a basis for $\mathbb{R}^2$.

- Argue that $\mathcal{B}$ is linearly independent.

- Now we show $\mathcal{B}$ spans $\mathbb{R}^2$. Let

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

  be any vector in $\mathbb{R}^2$. Find a "recipe" for writing $\vec{x}$ as a linear combination of $\vec{b}_1$ and $\vec{b}_2$.

Exploration 39 Let $\mathcal{B} = \{\vec{p}_1 = 1 + x, \vec{p}_2 = x + x^2\}$. Note that $\mathcal{D}$ is linearly independent.[4] Show that this is not a basis by finding a polynomial in $\mathbb{P}_2$ that is not a linear combination of $\vec{p}_1$ and $\vec{p}_2$. This will mean it fails to span all of $\mathbb{P}_2$.

4: This is actually seen in the example above since any subset of a linearly independent set will still be linearly independent.

Oh yeah, then why not prove it?

No, *you* prove it.

Exercise?

Deal.

Standard Basis Vectors

We've found bases now for $\mathbb{R}^2$ and $\mathbb{P}_2$, but neither of these bases is actually the most preferred basis for its space. (Of course, preferences are debatable…)[5] Instead, let's define the *standard basis* for each of these spaces, or really, for the more general vector spaces $\mathbb{R}^n$ and $\mathbb{P}_n$.

5: What if I have a mathematically rigorous way to quantify preferences?

Do you though?

Not yet, but I'm working on it.

- *Standard Basis for* $\mathbb{R}^n$ Let $\vec{e}_i$ be the vector with 0 in every entry except for the ith entry, which is 1. For example, in $\mathbb{R}^3$,

$$\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Then the standard basis for $\mathbb{R}^n$ is

$$\mathcal{E} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$$

where each $\vec{e}_i \in \mathbb{R}^n$. To be thorough, we should say a bit about verifying this is a basis. Even though its name is *standard basis*, such things shouldn't be taken for granted. The key thing to note is that for any $a_1, \ldots, a_n \in \mathbb{R}$,

$$a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Thus, the only way that $a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n = \vec{0}$ is if each $a_i = 0$. Also, any vector in $\mathbb{R}^n$ can be written as a linear combination of the vectors in $\mathcal{B}$ by using the ith entry of the vector for a_i, the coefficient of $\vec{e}_i$ in the sum. For example,

$$\begin{bmatrix} 7 \\ 2 \\ 3 \end{bmatrix} = 7\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 7\vec{e}_1 + 2\vec{e}_2 + 3\vec{e}_3.$$

- *Standard Basis for* $\mathbb{P}_n$ The standard basis for $\mathbb{P}_n$ is

$$\{1, x, x^2, ..., x^n\}.$$

Again, we should pause to convince ourselves this is a basis. Have you convinced yourself…? Good! Share your thoughts about this below:

Another Method to Show a Set Spans

Suppose you have a set S that you suspect is a basis for some vector space V. To verify it's a basis as we've illustrated, there are then two things to show. First, you need to argue that the set is linearly independent. We have a method for how to do this from Section 1.3, which is what we did in Example 2.1.2 as well. However, you also need to show it spans the vector space. Perhaps you noticed in the previous examples that this part can be a bit more difficult. Here are two strategies:

- *Method 1:* Argue that any general element of V can be obtained as a linear combination of the vectors in S.

- *Method 2:* Argue that a known basis for V is a subset of $\text{Span}\{S\}$.

The first method is what we employed in Example 2.1.2 and Exploration 38. It has the advantage of giving a recipe for how to write any vector in V as a linear combination of the vectors in S. The disadvantage is that sometimes the number of variables involved can get a bit overwhelming. Let's talk a bit about the second method now. It relies on the same logic we used in Example 2.1.1 in which we reasoned that if a subset is a spanning set, then the whole set must also be a spanning set. Let's formalize that.

Theorem 2.1.1 *Suppose* $S = \{\vec{v}_1, \ldots, \vec{v}_n\}$ *and* $U = \{\vec{u}_1, \ldots, \vec{u}_k\}$ *are subsets of a vector space* V.
(a) *If* $U \subseteq \text{Span}\{S\}$, *then* $\text{Span}\{U\} \subseteq \text{Span}\{S\}$.
(b) *If* $U \subseteq \text{Span}\{S\}$ *and* $\text{Span}\{U\} = V$, *then* $\text{Span}\{S\} = V$.

Before proceeding to our proof, we must welcome a new notation. The small problem we have is that we will need a bunch of scalars for each of a bunch of different vectors, and there simply aren't enough letters... or subscripts. Thus, we must implement a second subscript. For example, if given an ith vector $\vec{v}_i$ that requires j scalars (for some reason), we could (that is, we will) use the notation $a_{i1}, \ldots, a_{ij}$ for these j scalars. This is nice[6] because it also indicates these scalars' association with this given ith vector $\vec{v}_i$. Now, on to the proof.

6: I object to the description of this as "nice." Having two subscripts is annoying.

Those subscripts are actually necessary, unlike, for example, having two unicorns.

You would be so bored without me.

PROOF. Suppose $U = \{\vec{u}_1, \ldots, \vec{u}_k\}$ is a subset of $\text{Span}\{S\}$. Note that the span of a set of vectors is all linear combinations of the vectors in that set. Thus, every vector in U is a linear combination of the vectors in S. More specifically, we have

$$\begin{aligned}\vec{u}_1 &= a_{11}\vec{v}_1 + a_{12}\vec{v}_2 + \cdots + a_{1n}\vec{v}_n\\ \vec{u}_2 &= a_{21}\vec{v}_1 + a_{22}\vec{v}_2 + \cdots + a_{2n}\vec{v}_n\\ &\vdots\\ \vec{u}_k &= a_{k1}\vec{v}_1 + a_{k2}\vec{v}_2 + \cdots + a_{kn}\vec{v}_n\end{aligned}$$

Suppose now that $\vec{x} \in \text{Span}\{U\}$. Then

$$\begin{aligned}\vec{x} &= b_1\vec{u}_1 + \cdots b_k\vec{u}_k \\ &= b_1(a_{11}\vec{v}_1 + a_{12}\vec{v}_2 \cdots + a_{1n}\vec{v}_2) + \cdots + b_k(a_{k1}\vec{v}_1 + a_{k2}\vec{v}_2 + \cdots + a_{kn}\vec{v}_n) \\ &= (b_1a_{11} + b_2a_{21} \cdots + b_ka_{k1})\vec{v}_1 + \cdots + (b_1a_{1n} + b_2a_{2n} \cdots + b_ka_{kn})\vec{v}_n.\end{aligned}$$

Thus, $\vec{x} \in \text{Span}\{S\}$ which tells us $\text{Span}\{U\} \subseteq \text{Span}\{S\}$. Now, suppose U is a spanning set. That means $\text{Span}\{U\} = V$, and thus, $V \subseteq \text{Span}\{S\}$. Since we always have $\text{Span}\{S\} \subseteq V$, we know then that $\text{Span}\{S\} = V$ and S must also be a spanning set. □

This theorem gives us the tools for our second strategy. We know a basis is also a spanning set. Thus, if we want to check that $S = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is a spanning set for V when we already know a basis for V, we can check that the basis is contained in $\text{Span}\{S\}$. If V is $\mathbb{R}^n$ or $\mathbb{P}_n$, a convenient basis to look for is the standard one.

Example 2.1.3 Let's show that $S = \{1, 1+x, 1+x+x^2\}$ is a spanning set for $\mathbb{P}_2$ by showing the standard basis $\{1, x, x^2\} \subseteq \text{Span}\{S\}$.

- First, we must argue that $1 \in \text{Span}\{S\}$. Well, wait. We know $1 \in S$, so there's nothing to check here. Great!

- Now, we need to argue that $x \in \text{Span}\{S\}$. This one requires a bit more thought, but it's still not bad. We need to find coefficients $a, b, c \in \mathbb{R}$ such that
$$a(1) + b(1+x) + c(1+x+x^2) = x.$$
This simplifies to the equations $a + b = 0$, $b + c = 1$, and $c = 0$. Thus, we have $a = -1$, $b = 1$, and $c = 0$. That is, $x = (-1)(1) + (1+x) \in \text{Span}\{S\}$.

- Finally, we can argue that $x^2 \in \text{Span}\{S\}$. Again, we need to find coefficients $a, b, c \in \mathbb{R}$ such that
$$a(1) + b(1+x) + c(1+x+x^2) = x^2.$$
This simplifies to the equations $a + b = 0$, $b + c = 0$, and $c = 1$. Thus, we have $a = 0$, $b = -1$, and $c = 1$. That is, $x^2 = (-1)(1+x) + (1+x+x^2) \in \text{Span}\{S\}$.

Now that we've checked that the standard basis is contained in $\text{Span}\{S\}$, we can conclude that S is a spanning set.

Now, if you are trying to use this method for say $\mathbb{R}^{10}$ or $\mathbb{P}_9$, it might take a while to show all 10 basis vectors are in the span of the set. This method is less efficient than Method 1 in many cases, but sometimes it is preferred just because of the fewer variables involved.

Exploration 40 Complete the argument that S from Example 2.1.3 is a basis by showing the vectors are linearly independent.

Exploration 41 Use Method 2 to argue

$$S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

is a spanning set for $\mathbb{R}^2$.

Finding a Basis

Alright. We've defined a basis. We've seen some examples of bases, and we've discussed how to verify a set is a basis. What's left? Well, there are several facts we can state about bases and also things they tell us about their respective vector spaces. Much of the next section will be involved with stating and proving these facts.[7] For now, we will focus on a very useful theorem that helps us to *find* a basis for a given vector space.

7: Those facts better be theorems. Bubbles gets testy when you don't do proofs.

Part of the statement below is a repeat from Theorem 1.3.2

Theorem 2.1.2 *Let V be a vector space, let $S = \{\vec{v}_1, \ldots, \vec{v}_p\} \subset V$, and let S span V.*

(a) *If one of the vectors $\vec{v}_k \in S$ is a linear combination of the other vectors in S, then the set formed from S by removing $\vec{v}_k$ still spans V.*

(b) *If $V \neq \{\vec{0}\}$, then some subset of S is a basis for V.*

Proof. Let's prove the first statement. Suppose $\vec{v}_k \in S$ is a linear combination of the other vectors in S; that is

$$\vec{v}_k = a_1\vec{v}_1 + \cdots a_{k-1}\vec{v}_{k-1} + a_{k+1}\vec{v}_{k+1} + \cdots + a_p\vec{v}_p.$$

Since S spans V, for any $\vec{v} \in V$, there exist scalars $b_1, \ldots, b_p \in \mathbb{R}$ such that

$$\begin{aligned} \vec{v} &= b_1\vec{v}_1 + \cdots + b_k\vec{v}_k + \cdots + b_p\vec{v}_p \\ &= b_1\vec{v}_1 + \cdots + b_k(a_1\vec{v}_1 + \cdots + a_{k-1}\vec{v}_{k-1} + a_{k+1}\vec{v}_{k+1} + \cdots + a_p\vec{v}_p) + \\ &\quad \cdots + b_p\vec{v}_p \\ &= (b_1 + b_ka_1)\vec{v}_1 + \cdots + (b_{k-1} + b_ka_{k-1})\vec{v}_{k-1} + (b_{k+1} + b_ka_{k+1})\vec{v}_{k+1} + \\ &\quad \cdots + (b_p + b_ka_p)\vec{v}_p. \end{aligned}$$

Thus, any vector in V can be written as a linear combination of the vectors in the set formed from S by removing $\vec{v}_k$. It follows that the set formed from S by removing $\vec{v}_k$ spans V.

Let's move now to the second statement. If the set formed from S by removing $\vec{v}_k$ is linearly independent, then we have a basis for V and we are done. If it is linearly dependent, then one of the vectors in the set is a linear combination of the others by Theorem 1.3.1. In this case, we start this whole procedure again; repeating this as many times as necessary, the set will eventually be linearly independent since the set is finite and a set with one nonzero vector is trivially linearly independent. Also, from part one of this theorem, we know the set still spans V and is therefore a basis of V. $\square$

This theorem gives us a nice recipe for finding a basis for any vector space when we have a spanning set. Just remove the linearly dependent vectors one at a time until there are none. Nice! Let's try it!

Example 2.1.4 Recall from the beginning of this section:

$$\mathbb{R}^2 = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 1,234 \end{bmatrix} \right\}.$$

We're already told the set

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 1,234 \end{bmatrix} \right\}$$

spans $\mathbb{R}^2$, but it would be easy to check that for any vector $\vec{x} \in \mathbb{R}^2$, we have

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \dots + 0 \begin{bmatrix} 0 \\ 1,234 \end{bmatrix}.$$

By Theorem 2.1.2, all we have to do is throw out linearly dependent vectors from S until it is linearly independent, and we'll have a basis. Note that for any integer $2 \leq k \leq 1,234$, we have

$$\begin{bmatrix} 0 \\ k \end{bmatrix} = k \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

so the last $1,232$ vectors in S are linear combinations of the second vector. By Theorem 2.1.2, we can throw them all out of S and the resulting set still spans $\mathbb{R}^2$. Specifically,

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

still spans $\mathbb{R}^2$. These vectors are linearly independent (since there are two of them and they're not scalar multiples of each other), so this is a basis[8] for $\mathbb{R}^2$.

8: It's the standard basis for $\mathbb{R}^2$!

Bases of Subspaces

Theorem 2.1.2 is also very helpful[9] when finding the basis for a subspace. Let's see an example illustrating this.

9: Like me!
Debatable.

Example 2.1.5 First, let's define a few vectors to work with for this example. Let

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}. \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \text{ and } \vec{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 4 \\ 1 \end{bmatrix}.$$

Now, we can define $H = \text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$, which is a subspace of $\mathbb{R}^4$. Because of the way H is described to us, we are starting with a spanning set. So we just need to reduce it down to a linearly independent set without changing the span. To do this, we need to find a vector that can be written as a linear combination of other vectors.

The vector $\vec{v}_1$ is a good place to start. Suppose the there exist scalars $a_1, a_2, a_3 \in \mathbb{R}$ such that

$$\vec{v}_1 = a_1\vec{v}_2 + a_2\vec{v}_3 + a_3\vec{v}_4.$$

That is,

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 1 \\ 3 \\ 1 \end{bmatrix} + a_3 \begin{bmatrix} 1 \\ 1 \\ 4 \\ 1 \end{bmatrix}.$$

Then we would have the equations $a_1 + a_2 + a_3 = 1$ from the top row of the vectors and also $a_1 + a_2 + a_3 = 2$ from the last row of the vectors. Both of these can't be true, so $\vec{v}_1$ is not a linear combination of the remaining vectors.

Instead, we can check whether $\vec{v}_2$ is a linear combination of $\vec{v}_3$ and $\vec{v}_4$. (We don't need to consider $\vec{v}_1$ here because we've already ruled it out.) Let $b_1, b_2 \in \mathbb{R}$ be such that

$$\vec{v}_2 = b_1\vec{v}_3 + b_2\vec{v}_4.$$

That is,

$$\begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 1 \\ 3 \\ 1 \end{bmatrix} + b_2 \begin{bmatrix} 1 \\ 1 \\ 4 \\ 1 \end{bmatrix}.$$

Then we have the equations $1 = b_1 + b_2$ and $2 = 3b_1 + 4b_2$, which have the mutual solution $b_1 = 2$ and $b_2 = -1$, so

$$\vec{v}_2 = 2\vec{v}_3 - \vec{v}_4.$$

Now by Theorem 2.1.2, we can remove one of these three vectors and still have a spanning set. Thus, $H = \text{Span}\{\vec{v}_1, \vec{v}_3, \vec{v}_4\}$. Now, is the set $\{\vec{v}_1, \vec{v}_3, \vec{v}_4\}$ linearly independent? Yes! We can conclude this fairly quickly by our previous work. We know $\vec{v}_1$ is not a linear combination of $\vec{v}_3$ and $\vec{v}_4$, and we also see from inspection that $\vec{v}_3$ and $\vec{v}_4$ are not scalar multiples of each other. We have found a basis!

Exploration 42 In Example 2.1.5 above, we found a basis for the subspace H of $\mathbb{R}^4$. Can you find a different basis for H?

What about when a subspace is described differently?[10] Well, in that case, we start with finding a spanning set and then proceed like Example 2.1.5.

10: There's nothing wrong with being different.

How is that a substantive contribution to the narrative?

Exploration 43 Let's consider the subspace K of $\mathbb{R}^4$ defined below.

$$K = \left\{ \begin{bmatrix} a+b+c+d \\ a+b+c+d \\ a+2b+3c+4d \\ 2a+b+c+d \end{bmatrix} : a, b, c, d \in \mathbb{R} \right\}.$$

Find a set of vectors that span K. (Hint: These should look familiar!) Then use this spanning set to find a basis for K.

Section Highlights

- Determine when the span of a set of vectors is a *spanning set* for a vector space. See Definition 2.1.1.
- There are multiple methods for determining whether a set is a spanning set. See Examples 2.1.1 and 2.1.3.
- A *basis* for a vector space (Definition 2.1.2) is a linearly independent spanning set.
- Any spanning set can be reduced to a basis by carefully removing vectors that are linear combinations of other basis vectors. See Example 2.1.5.

Exercises for Section 2.1

2.1.1. The following sets are not bases for $\mathbb{R}^2$. Determine whether they fail to be a spanning set, fail to be linearly independent, or both.

(a) $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

(b) $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right\}$

(c) $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$

2.1.2. The following sets are not a basis for $\mathbb{R}^3$. Determine whether they fail to be a spanning set, fail to be linearly independent, or both.

(a) $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$

(b) $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right\}$

(c) $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

2.1.3. The following sets are not a basis for $\mathbb{P}_1$. Determine whether they fail to be a spanning set, fail to be linearly independent, or both.

(a) $\{1 + x\}$

(b) $\{1 + x, -1 - x\}$

(c) $\{1 + x, -1 - x, x\}$

2.1.4. The set $\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$ is not a basis for $\mathbb{R}^3$, even though it spans $\mathbb{R}^3$. Use the procedure from Theorem 2.1.2 to find a basis for $\mathbb{R}^3$.

2.1.5. Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$.

(a) Use Method 1 to show this is a spanning set for $\mathbb{R}^2$.

(b) Use Method 2 to show this is a spanning set for $\mathbb{R}^2$.

2.1.6. Let $\mathcal{B} = \left\{ 1 + x, 1 + x^2, 1 + x + x^2 \right\}$.

(a) Use Method 1 to show this is a spanning set for $\mathbb{P}_2$.

(b) Use Method 2 to show this is a spanning set for $\mathbb{P}_2$.

2.1.7. Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$.

(a) Show $\mathcal{B}$ is a basis for $\mathbb{R}^2$.

(b) Find a way to write $\begin{bmatrix} 4 \\ 6 \end{bmatrix}$ as a linear combination of the vectors in $\mathcal{B}$.

(c) Find a way to write $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as a linear combination of the vectors in $\mathcal{B}$.

2.1.8. Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

(a) Show $\mathcal{B}$ is a basis for $\mathbb{R}^3$.

(b) Find a way to write $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as a linear combination of the vectors in $\mathcal{B}$.

2.1.9. Show $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^3$. How do you write $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ as a linear combination of these basis vectors?

2.1.10. Show $\mathcal{B} = \{1 + x, x\}$ is a basis for $\mathbb{P}_1$.

2.1.11. Show $\mathcal{B} = \left\{5 + x, 1 + x + x^2, x^2\right\}$ is a basis for $\mathbb{P}_2$.

2.1.12. Let $\mathcal{B} = \left\{1 + x, 2x, x + x^2\right\}$.

(a) Show $\mathcal{B}$ is a basis for $\mathbb{P}_2$.

(b) Find 1 as a linear combination of these basis vectors.

(c) Find $1 + x + x^2$ as a linear combination of these basis vectors.

2.1.13. If $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a basis for a vector space V, show that $\{\vec{v}_1, \vec{v}_1 + \vec{v}_2, \vec{v}_1 + \vec{v}_2 + \vec{v}_3\}$ is also a basis. (Hint: Method 2 may be a helpful way to show this is a spanning set.)

2.1.14. Let $\vec{p}_1 = x^2 + 1$, $\vec{p}_2 = x^2 - 1$, and $\vec{p}_3 = 3$. Show that $\{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$ is a linearly dependent set, and find a basis for Span $\{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$.

2.1.15. As mentioned in the section, any nonempty subset of a linearly independent set is itself linearly independent. Let's show this.
Let $S = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be a subset of some vector space V and suppose $\widehat{S} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is a subset of S. Suppose S is a linearly independent set.

- ► Suppose $\widehat{S}$ is not linearly independent. Then there are scalars $a_1, \ldots, a_k \in \mathbb{R}$ not all zero such that

$$a_1\vec{v}_1 + \cdots a_k\vec{v}_k = \vec{0}$$

Explain how this violates the definition of linear independence for the set S. This allows us to conclude $\widehat{S}$ is also linearly independent since otherwise it can't be true that S was linearly independent.

2.1.16. The subset H below is a subspace of $\mathbb{R}^4$. Find a basis for H.

$$H = \left\{ \begin{bmatrix} x \\ -y \\ x \\ y \end{bmatrix} : x, y \in \mathbb{R} \right\}.$$

2.1.17. The subset J below is a subspace of $\mathbb{R}^4$. Find a basis for J.

$$J = \left\{ \begin{bmatrix} 2x + 3y + z \\ -y \\ x \\ y + z \end{bmatrix} : x, y, z \in \mathbb{R} \right\}.$$

2.1.18. The subset K below is a subspace of $\mathbb{R}^4$. Find a basis for K.

$$K = \left\{ \begin{bmatrix} x + y + 2z \\ -y - 2z \\ x \\ y + 2z \end{bmatrix} : x, y, z \in \mathbb{R} \right\}.$$

2.1.19. Find a basis for each subspace below.

(a) $\text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$

(b) $\text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

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

2.1.20. Let

$$\vec{v}_1 = \begin{bmatrix} 0 \\ -5 \\ 1 \\ 2 \end{bmatrix}, \qquad \vec{v}_2 = \begin{bmatrix} 3 \\ 4 \\ -2 \\ 5 \end{bmatrix}, \quad \text{and} \quad \vec{v}_3 = \begin{bmatrix} 2 \\ 1 \\ -1 \\ 4 \end{bmatrix}.$$

and

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \qquad \vec{u}_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 6 \end{bmatrix}$$

(a) Check that $\vec{v}_1 + 2\vec{v}_2 - 3\vec{v}_3 = \vec{0}$.

(b) Find a basis for $H_1 = \operatorname{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

(c) Let $H_2 = \operatorname{Span}\{\vec{u}_1, \vec{u}_2\}$. Find a basis for $H_1 + H_2$.

2.1.21. Show that $\{1, i\}$ is a basis for $\mathbb{C}$.

2.1.22. Show that $\{1 + i, i\}$ is also a basis for $\mathbb{C}$.

2.1.23. In the Section 1.1 Exercises, you showed $V = \{a \colon a \in \mathbb{R}, a \geq 0\} = (0, \infty)$ is a vector space with addition given by $a \boxplus b = ab$ and scalar multiplication given by $ka = a^k$ for any $k \in \mathbb{R}$, $a, b \in V$. Find a basis for this vector space.

2.2 More Fun with Bases

In the previous section, we introduced the concept of a basis for a vector space and we focused on finding bases or showing a suspected basis is a basis. Now, we will see several results related to a basis for a vector space. Some of these will help us in our search to find a basis or to tell whether a set is a basis. Overall, they will show us how finding a basis for a vector space reveals an intrinsic property of the vector space, called dimension.[11]

11: Cool! Will I be able to use this to visit parallel universes?

Yes.

Really?!

. . .

How Large Can An Independent Set Be?

We'll start with a motivating example.

Example 2.2.1 Let's consider the set

$$S = \left\{ \vec{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \vec{u}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}.$$

While we could show the set S is linearly dependent in the same manner as we have in the past, we'll take a different approach for this example. Recall that the standard basis for $\mathbb{R}^2$ is

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

We can write each of the vectors in S as a linear combination of $\vec{e}_1$ and $\vec{e}_2$. In particular,

$$\begin{aligned} \vec{u}_1 &= \vec{e}_1 + \vec{e}_2, \\ \vec{u}_2 &= \vec{e}_1 - \vec{e}_2, \text{ and} \\ \vec{u}_3 &= \vec{e}_1 + 2\vec{e}_2. \end{aligned}$$

Now, we can solve for $\vec{e}_1$ in our first equation to get

$$\vec{e}_1 = \vec{u}_1 - \vec{e}_2.$$

Then we have

$$\begin{aligned} \vec{u}_2 &= \vec{e}_1 - \vec{e}_2 \\ &= (\vec{u}_1 - \vec{e}_2) - \vec{e}_2 \\ &= -\vec{u}_1 - 2\vec{e}_2. \end{aligned}$$

Rearranging this we see that

$$\vec{e}_2 = \frac{\vec{u}_1 + \vec{u}_2}{2}.$$

Then

$$\begin{aligned}\vec{u}_3 &= \vec{u}_1 - \vec{e}_2 + 2\vec{e}_2 \\ &= \vec{u}_1 + \vec{e}_2 \\ &= \vec{u}_1 + \frac{\vec{u}_1 + \vec{u}_2}{2} \\ &= \frac{3\vec{u}_1 + \vec{u}_2}{2} \\ &= \frac{3}{2}\vec{u}_1 + \frac{1}{2}\vec{u}_2.\end{aligned}$$

Since $\vec{u}_3$ is a linear combination of $\vec{u}_1$ and $\vec{u}_2$, this shows the set S is linearly dependent by Theorem 1.3.1. This method relied only on the fact that the set of vectors was larger than the size of the standard basis. Below we will generalize this example to prove this is always enough to conclude that such a set is linearly dependent.

Theorem 2.2.1 *If a vector space V has a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$, then any set in V containing more than p vectors is linearly dependent.*

PROOF. Let $S = \{\vec{u}_1, \ldots, \vec{u}_{p+1}\} \subseteq V$. If any of these vectors are the zero vector, we are done, so assume that $\vec{u}_i \neq \vec{0}$ for $1 \leq i \leq p+1$. We shall attempt to write one of the vectors $\vec{u}_i$ as a linear combination of the other vectors in S. Since $\mathcal{B}$ is a basis for V and $\vec{u}_1 \in V$, there are weights such that

$$\vec{u}_1 = a_{11}\vec{v}_1 + \cdots + a_{1p}\vec{v}_p. \tag{2.1}$$

Since $\vec{u}_1 \neq \vec{0}$, we know $a_{1j} \neq 0$ for some $1 \leq j \leq p$. Suppose without loss of generality that $a_{11} \neq 0$. Then we may solve for $\vec{v}_1$ in Equation 2.1:

$$\vec{v}_1 = \frac{1}{a_{11}}\vec{u}_1 - \frac{a_{12}}{a_{11}}\vec{v}_2 - \cdots - \frac{a_{1p}}{a_{11}}\vec{v}_p. \tag{2.2}$$

Thus, $\vec{v}_1 \in \text{Span}\{\vec{u}_1, \vec{v}_2, \ldots, \vec{v}_p\}$. Again, since $\mathcal{B}$ is a basis for V and $\vec{u}_2 \in V$, there are weights such that

$$\vec{u}_2 = a_{21}\vec{v}_1 + \cdots + a_{2p}\vec{v}_p. \tag{2.3}$$

Since $\vec{v}_1 \in \text{Span}\{\vec{u}_1, \vec{v}_2, \ldots, \vec{v}_p\}$, we can find weights such that $\vec{v}_1 = b_{21}\vec{u}_1 + b_{22}\vec{v}_2 + \cdots + b_{2p}\vec{v}_p$. In fact, we already found these weights in Equation 2.2! Then

$$\begin{aligned}\vec{u}_2 &= a_{21}\vec{v}_1 + \cdots + a_{2p}\vec{v}_p \\ &= a_{21}\left(b_{21}\vec{u}_1 + b_{22}\vec{v}_2 + \cdots + b_{2p}\vec{v}_p\right) + a_{22}\vec{v}_2 + \cdots + a_{2p}\vec{v}_p \\ &= a_{21}b_{21}\vec{u}_1 + (a_{21}b_{22} + a_{22})\vec{v}_2 + \cdots + (a_{21}b_{2p} + a_{2p})\vec{v}_p \\ &= c_{21}\vec{u}_1 + c_{22}\vec{v}_2 + \cdots + c_{2p}\vec{v}_p, \qquad (2.4)\end{aligned}$$

Where the weights, $c_{21}, \ldots, c_{2p}$, are defined by the line in Equation 2.4 that immediately precedes them. What have we done here?[12] We can write $\vec{u}_2$ as a linear combination of the vectors in $\mathcal{B}$, but we can find different weights to replace $\vec{v}_1$ in our linear combination with $\vec{u}_1$. This is one step on the way to writing one of the vectors in S as a linear combination of the others. If $c_{22} = \cdots = c_{2p} = 0$, then $\vec{u}_2$ is a linear combination of $\vec{u}_1$. In that case,

12: What have we done? Was it good?

Yes. Pay attention!

the set S is linearly dependent by Theorem 1.3.1, and we are done. Otherwise, $c_{2j} \neq 0$ for some $2 \leq j \leq p$, so suppose without loss of generality that $c_{22} \neq 0$. We'll have to do this procedure again: Solve for $\vec{v}_2$ in Equation 2.4 to show $\vec{v}_3 \in \text{Span}\{\vec{u}_1, \vec{u}_2, \vec{v}_3, \ldots, \vec{v}_p\}$; then use that to show $\vec{u}_3 \in \text{Span}\{\vec{u}_1, \vec{u}_2, \vec{v}_3, \ldots, \vec{v}_p\}$.

How many times could we do this? One of two things must happen; either $\vec{u}_i$, for some $2 \leq i \leq p$, will be written as a linear combination of $\vec{u}_1, \ldots, \vec{u}_{i-1}$ (in which case S would be linearly dependent by Theorem 1.3.1), or we will have to run this procedure p times, writing each $\vec{v}_j$, for $1 \leq j \leq p$, with linear combinations of vectors in S. That is, for each $1 \leq j \leq p$,

$$\vec{v}_j = c_{1j}\vec{u}_1 + \cdots + c_{pj}\vec{u}_p = \sum_{i=1}^{p} c_{ij}\vec{u}_i. \tag{2.5}$$

In this case, we still have one more vector $\vec{u}_{p+1} \in S$. Since $\vec{u}_{p+1} \in V$, we can write $\vec{u}_{p+1}$ as a linear combination of vectors in $\mathcal{B}$ and then substitute the linear combinations in Equation 2.5 for each $\vec{v}_j$:

$$\begin{aligned} \vec{u}_{p+1} &= a_1\vec{v}_1 + \cdots + a_p\vec{v}_p \\ &= a_1\left(\sum_{i=1}^{p} c_{i1}\vec{u}_i\right) + \cdots + a_p\left(\sum_{i=1}^{p} c_{ip}\vec{u}_i\right). \end{aligned}$$

Thus, $\vec{u}_{p+1}$ is a linear combination of the other vectors in S, so by Theorem 1.3.1, S is linearly dependent. □

Now, let's consider how this works for us in $\mathbb{R}^4$.[13] According to the theorem we just proved, any set of five vectors in $\mathbb{R}^4$ is linearly dependent (because the standard basis for $\mathbb{R}^4$ has four vectors). But what about four vectors?

13: No particular reason to choose 4 here, just feeling like a 4 apparently.

Example 2.2.2 Let's consider two different sets of four vectors in $\mathbb{R}^4$. Here's the first one:

$$S_1 = \left\{ \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

This set is linearly dependent. To see this, note that $\vec{v}_1 = \vec{v}_2 + \vec{v}_3$. Here's another set:

$$S_2 = \left\{ \vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \vec{u}_3 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

This set is linearly independent. If $a\vec{u}_1 + b\vec{u}_2 + c\vec{u}_3 + d\vec{u}_4 = \vec{0}$, then we get the equations $a = 0$, $b + d = 0$, $a + d = 0$, $c + d = 0$. The only real numbers satisfying all of these are $a = b = c = d = 0$.

From these examples, we see that while our theorem says every set of vectors with 5 vectors in $\mathbb{R}^4$ must be linearly dependent, sets with 4 vector in $\mathbb{R}^4$ can go either way, dependent or independent.

Dimension

Let's make an observation. From our explorations and examples in Section 2.1, we've seen that $\mathbb{R}^2$ has bases

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \text{ and } \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

From this, we see that a vector space does not have a unique basis.[14] Notice something else, though: each of these bases has two vectors. We also saw that $\mathbb{P}_2$ has bases $\{1 + x, x + x^2, x\}$ and $\{1, x, x^2\}$. Again, these bases have the same number of vectors. Well, we were due for some good news; that's always true!

14: This seems very inconvenient, but we will eventually make this work to our advantage.

Theorem 2.2.2 *If a vector space V has a basis with p vectors, then every basis for V must contain exactly p vectors.*

PROOF. Let $\mathcal{B}$ be a basis for V with p vectors. Suppose $\mathcal{B}_0$ is another basis for V. We know that $\mathcal{B}_0$ cannot have more than p vectors since then, by Theorem 2.2.1, $\mathcal{B}_0$ would be linearly dependent. We also know that $\mathcal{B}_0$ cannot have fewer than p vectors since then, by Theorem 2.2.1, $\mathcal{B}$ would be linearly dependent. Thus, $\mathcal{B}_0$ also has exactly p vectors. $\square$

Note that we didn't really talk about spans in that proof. However, there's a nice corollary about spanning sets that comes from this result. Since any set that is a basis must have the same number of vectors, we know any linearly independent set that has too few vectors to be a basis must fail to span. Otherwise, we would contradict this result!

Corollary 2.2.3 *If a vector space V has a basis with p vectors, then every spanning set for V must contain at least p vectors.*

Every vector space has many different bases;[15] Theorem 2.2.2 guarantees that they all have the same number of vectors. Good. This makes the following well-defined:

15: Actually, most of the vector spaces over $\mathbb{R}$ we will see have infinitely many bases!

Definition 2.2.1 *Let V be a vector space. The* ***dimension of*** *V, denoted $\dim V$, is the number of vectors in a basis for V.*

The term "dimension" is likely not new to you. People regularly make use of three spacial dimensions, and movie-going often offers us a choice between two and three dimensional experiences. What's new here is giving it a formal definition that relies on a basis. We can now say that $\mathbb{R}^n$ has dimension n and also that $\mathbb{P}_n$ has dimension $n + 1$. This tells us immediately how large a basis for these spaces should be. This also gives us a nice way to talk about the subspaces in these vector spaces. The only subspace of any vector space with dimension 0 is the trivial subspace $\{\vec{0}\}$. In $\mathbb{R}^n$, the subspaces with dimension 1 are the lines through the origin, and the subspaces with dimension 2 are the planes through the origin. While we're talking about subspaces, here's a formal statement about their dimensions.

Theorem 2.2.4 *If H is a subspace of a vector space V, then $\dim H \leq \dim V$. Moreover, the only subspace of V with dimension $\dim V$ is V itself.*

Well, it would be lovely to prove this theorem. We need something else though first.

Theorem 2.2.5 *Let V be a p-dimensional vector space, where $p \geq 1$. Then*

(a) *Any linearly independent set of exactly p vectors in V is a basis for V.*

(b) *Any set of p vectors that spans V is a basis for V.*

Exploration 44 Let's do something a bit different. Let's prove this one together!

PROOF.

▶ First, let's suppose $\mathcal{D}$ is a set of exactly p linearly independent vectors. Well, suppose this is not a basis for V. Then, it must not span all of V. So there must be some vector $v \in V$ that is not in the span of $\mathcal{D}$. Then v is specifically not a linear combination of any of the vectors in $\mathcal{D}$, and none of the vectors in $\mathcal{D}$ are a linear combination of each other. So adding v to $\mathcal{D}$ *must* still give a linearly independent set of vectors. On the other hand, some of the results in the previous section can be used to argue that the set that is $\mathcal{D}$ with v added *must* be linearly dependent. Explain why it must be linearly dependent. It can't be *both* linearly independent and linearly dependent. What's going on here? Well, this was all built on the supposition that $\mathcal{D}$ is not a basis. Since we have arrived at something absurd based on this assumption, it must be that $\mathcal{D}$ *is* a basis. (Which is what we wanted to show!)[16]

▶ Now suppose $S \subseteq V$ is a set of exactly p vectors that spans all of V. We want to show that S is a basis for V. Since we know it spans V, all that is left is that we argue it must be linearly independent. Okay, suppose it's not. Suppose instead that it is linearly dependent. Like before, our goal will be to arrive at something absurd using this supposition. What does Theorem 2.1.2 allow us to do since this is a spanning set that is linearly dependent? How does this contradict Theorem 2.2.2?

□

16: This technique of proof is called "proof by contradiction." As a rule of thumb, it is a lovely way to include words such as "absurd", "preposterous", "silly", or "ridiculous" in your formal mathematical writing.

Formal writing like this, right?

Well. . .

That was just what we needed to prove Theorem 2.2.4!

PROOF OF THEOREM 2.2.4. Suppose H is a subspace of a vector space V with $\dim V = p$. If H is also dimension p, then it must have a basis of size p. However, from Theorem 2.2.5, we know any linearly independent set of size p in V must be a basis for p. Thus, we must have $H = V$ since any basis of H is also a basis for V. Otherwise, H must have a basis that is smaller in size than p since we know any set with more than p vectors will be linearly dependent by Theorem 2.2.1. □

Theorem 2.2.5 has already been useful, but it gets better. To be a basis for a vector space V, a set must do two things: it must be linearly independent and it must span V. That is, unless you know the dimension of V, say p, and your set of vectors already has p vectors in it. Then you only need one of the two criterion for a basis to be true! That is sometimes very helpful.

Example 2.2.3 Hey, remember that set we showed was a spanning set in Example 2.1.1? Oh, well, maybe we showed several sets were spanning sets actually since we had $\text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \text{Span}\{\vec{v}_2, \vec{v}_3\} = \text{Span}\{\vec{v}_1, \vec{v}_3\} = \text{Span}\{\vec{v}_1, \vec{v}_2\}$. Now, we can conclude, by Theorem 2.2.5, that any of those spanning sets of size two are a basis for $\mathbb{R}^2$ because they each have two vectors and span $\mathbb{R}^2$! Yay!
Let's be slightly more specific. We can all agree that $\mathbb{R}^2$ has dimension two because the standard basis $\{\vec{e}_1, \vec{e}_2\}$ for $\mathbb{R}^2$ has two vectors. By Theorem 2.2.5, any spanning set of two vectors in $\mathbb{R}^2$ is a basis for $\mathbb{R}^2$. This will work the same for $\mathbb{R}^n$ in general. Also, any linearly independent set of n vectors in $\mathbb{R}^n$ is a basis for $\mathbb{R}^n$.

Example 2.2.4 Note that $\mathbb{P}_2$ has standard basis $\{1, x, x^2\}$, so it has dimension three. While $\{1 + x, x^2\}$ is a linearly independent set in $\mathbb{P}_2$, by Theorem 2.2.5 it does not span. It does not have enough vectors to be linearly independent *and* span $\mathbb{P}_2$; a set needs exactly three vectors to be linearly independent and span (i.e. be a basis for) a three dimensional vector space.

Exploration 45 Let's show that $\{\vec{p}_1 = 1, \vec{p}_2 = 2x, \vec{p}_3 = -2 + 4x^2, \vec{p}_4 = -11x + 8x^3\}$ is a basis for $\mathbb{P}_3$! Well, first of all, it has 4 vectors, which is luckily the correct number. Now we need only show that it is either linearly independent or that it spans all of $\mathbb{P}_3$. But which one is easier? You decide!

- *Linearly Independent.* Let a, b, c, d be real numbers such that

$$a\vec{p}_1 + b\vec{p}_2 + c\vec{p}_3 + d\vec{p}_4 = \vec{0}.$$

Then

$$a + 2bx - 2c + 4cx^2 - 11dx + 8dx^3 = 0.$$

Why can we conclude that $a = b = c = d = 0$?

- *Spans* $\mathbb{P}_3$. Let $a + bx + cx^2 + dx^3$ be any vector in $\mathbb{P}_3$. Find coefficients of $\vec{p}_1$, $\vec{p}_2$, $\vec{p}_3$, and $\vec{p}_4$ so that

$$a + bx + cx^2 + dx^3 = \underline{\hspace{2em}}\vec{p}_1 + \underline{\hspace{2em}}\vec{p}_2 + \underline{\hspace{2em}}\vec{p}_3 + \underline{\hspace{2em}}\vec{p}_4.$$

  It will be easiest to first determine the coefficient of $\vec{p}_4$, then $\vec{p}_3$, then $\vec{p}_2$, and lastly $\vec{p}_1$. (This one seems harder. . .)

You may perhaps be wondering why you bothered at all to learn how to show a set is a spanning set.[17] The algebra of that is usually much trickier than showing a set is linearly independent. However, there is something extremely useful about being able to write a vector as a linear combination of basis vectors, as we'll see in just the next section.

17:  I was definitely wondering.

Before concluding our initial discussion of dimension, it is important to not that our definition of basis specifies that a basis is a finite set of vectors. Hence, any vector space for which we use this definition (basis or dimension) must be *finite dimensional*. Indeed, one should assume that for remainder of this book, all vectors spaces have are finite dimensional. Infinite dimensional vectors spaces do exist and are a lot of fun, but their fun is simply too great to be contained in this book.

Section Highlights

- A basis $\mathcal{B}$ for vector space V is a minimal spanning set in the following sense: If any vector is removed from $\mathcal{B}$, the resulting set will fail to span V.
- Any set with fewer vectors than a basis will automatically fail to span the vector space. See Corollary 2.2.3.
- A basis $\mathcal{B}$ is a maximal linearly independent set in the vector space V in the following sense. If any vector is apended to $\mathcal{B}$, the resulting set will be linearly dependent. See Theorem 2.2.1.
- Every basis for a specific vector space V has the same number of vectors. This number of vectors is the *dimension* of the vector space. See Theorem 2.2.2 and Definition 2.2.1.
- If the dimension of a vector space is known, it can simplify the determination of whether a set is a basis. See Example 2.2.3.

Exercises for Section 2.2

2.2.1. Let $\vec{x}_1, \vec{x}_2, \vec{x}_3 \in \mathbb{R}^3$ be such that $\text{Span}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\} = \mathbb{R}^3$. Explain why $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ is a basis for $\mathbb{R}^3$.

2.2.2. Let $\vec{v}_1, \ldots, \vec{v}_7 \in \mathbb{R}^7$. Suppose $H_6 = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_6\}$ and $H_7 = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_7\}$. Explain why $\mathbb{R}^7 \neq H_6$. Must it be true that $\mathbb{R}^7 = H_7$?

2.2.3. Recall that in Exercise 1.3.21 in Section 1.3 we showed that $\mathbb{R}^2 = \text{Span}\left\{\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \end{bmatrix}\right\} = H$. Here's another fun way to do that using the fact that $\mathbb{R}^2 = \text{Span}\{\vec{e}_1, \vec{e}_2\}$. I know

$$\begin{bmatrix} 7 \\ 0 \end{bmatrix} \in H,$$

so $\vec{e}_1 \in H$ as well. But that means

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} - 2\vec{e}_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \in H,$$

so $\vec{e}_2 \in H$. Since $\dim H = 2$, we know $H = \text{Span}\{\vec{e}_1, \vec{e}_2\} = \mathbb{R}^2$. Neat, eh? Adapt this procedure to show that

$$\text{Span}\left\{\begin{bmatrix} 7 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}\right\} = \mathbb{R}^3.$$

2.2.4. The following sets are all too large to be linearly independent. Find a vector that is a linear combination of the others in the set to verify this.

(a) $\{1+x, 1-x, 1+x+x^2, 2x^2\}$ in $\mathbb{P}_2$.

(b) $\{1+x^2, 1-x, 1+2x+x^2, 1-2x^2, x\}$ in $\mathbb{P}_2$.

(c) $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 1 \end{bmatrix}\right\}$ in $\mathbb{R}^2$

(d) $\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 1 \\ 1 \end{bmatrix}\right\}$ in $\mathbb{R}^3$

2.2.5. The following sets S are linearly independent but too small to be spanning sets. Find a vector that is not in $\text{Span}\{S\}$ and add it to the set to form a basis for the vector space. (Hint: If it is not a spanning set, there must be at least one standard basis vector that is missing.)

(a) $\{1+x, 1-x^2\}$ in $\mathbb{P}_2$.

(b) $\{1+x^2, 1-x\}$ in $\mathbb{P}_2$.

(c) $\left\{\begin{bmatrix} 5 \\ 1 \end{bmatrix}\right\}$ in $\mathbb{R}^2$

(d) $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

(e) $\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

(f) $\left\{ \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

2.2.6. If a set is the correct size to be a basis according to the dimension of the vector space, but it is not a basis, it must fail both conditions to be a basis. That is, it must be linearly dependent and fail to be a spanning set. Verify that each of these sets fail both conditions to be a basis.

(a) $\{1+x, 1+2x+x^2, 2+3x+x^2\}$ in $\mathbb{P}_2$.

(b) $\{1+x^2, 1-x, x+x^2\}$ in $\mathbb{P}_2$.

(c) $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \\ 0 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

(d) $\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ 5 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

(e) $\left\{ \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ -3 \\ 5 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

2.2.7. Show the sets below are bases for the given vector space.

(a) $\left\{\vec{5}\right\}$ for $\mathbb{R}$

(b) $\{5, 1+x\}$ for $\mathbb{P}_1$

(c) $\{x, 1+x\}$ for $\mathbb{P}_1$

(d) $\left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \end{bmatrix} \right\}$ for $\mathbb{R}^2$

(e) $\left\{ \begin{bmatrix} -2 \\ 5 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \end{bmatrix} \right\}$ for $\mathbb{R}^2$

(f) $\left\{x, 1+x, x^2\right\}$ for $\mathbb{P}_2$

(g) $\{1+x, x+x^2, x^2\}$ for $\mathbb{P}_2$

(h) $\left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{R}^3$

(i) $\left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{R}^3$

(j) $\left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{R}^4$

2.2.8. Find a basis for each subspace below. State the dimension of the subspace.

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

(b) $H = \text{Span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -7 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix} \right\}$

(c) $H = \text{Span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -7 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

2.2.9. Find a basis for the set of vectors in $\mathbb{R}^3$ that lie in the subspace

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : 2x + 4y - z = 0 \right\}.$$

To do this, first identify a vector in $\mathbb{R}^3$ that is not in W. Then, find two linearly independent vectors that are in W. We know W cannot be 3-dimensional because you found a vector that is not in W. Since you were able to find two linearly independent vectors that are in W, we can conclude that $\dim W = 2$ and these vectors form a basis.

2.2.10. Consider the subspaces below for $\mathbb{R}^3$.

$$U = \left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad V = \left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

(a) Show that $\dim U = 2$ and $\dim V = 2$.

(b) Since both U and V are 2-dimensional subspaces of a 3-dimensional space, they must have an intersection that is at least 1-dimensional. Why? Find $U \cap V$ to verify this.

2.3 Coordinates, Inner Products, and Orthogonality: Oh my!

At the start of this chapter, we promised to focus more on geometry. We are now ready to see how the concept of a basis for a vector space allows us to extend the geometry we already talked about in $\mathbb{R}^n$ to other vector spaces.[18]

18: What geometry?

Length and distance via an inner product.

Bases In Action: Coordinates

One of the best things about bases is that they provide a convenient way to organize vectors in a vector space. If every vector in the vector space V is a linear combination of the vectors in some set $\mathcal{B} \subseteq V$, then those linear combinations can be used as a description for the vectors. This is something you are actually very much accustomed to; consider the standard basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ for $\mathbb{R}^3$. Consider the following example:

$$\begin{aligned} \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} &= \begin{bmatrix} 7 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 8 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 9 \end{bmatrix} \\ &= 7 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 8 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 9 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= 7\vec{e}_1 + 8\vec{e}_2 + 9\vec{e}_3. \end{aligned}$$

Now we know that $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ is a basis for $\mathbb{R}^3$, so we shouldn't be surprised that we can write any vector in $\mathbb{R}^3$ as a linear combination of $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$. It also turns out that the weights on that linear combination, 7, 8, and 9 respectively, have meaning. What is the vector

$$\begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}?$$

It's 7 $\vec{e}_1$'s, 8 $\vec{e}_2$'s, and 9 $\vec{e}_3$'s. What if we used a different basis? Those weights would probably have to change, right?

Exploration 46 Let

$$\mathcal{B} = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \text{ and } \vec{x} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}.$$

Note that $\mathcal{B}$ is a basis for $\mathbb{R}^3$.[19] Write $\vec{x}$ as a linear combination using this basis.

19: Or instead of "note", maybe you should check this?

Or maybe get a group of friends together for a basis-checking party!

What if there was more than one linear combination for a particular vector and a particular basis? Well, fortunately that cannot happen:

Theorem 2.3.1 *Let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_p\}$ be a basis for a vector space V. Then for each $\vec{v} \in V$, there exists a unique set of scalars $c_1, \dots, c_p \in \mathbb{R}$ such that*

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_p\vec{v}_p. \tag{2.6}$$

PROOF. The existence of an equation such as Equation 2.6 is guaranteed by the fact every basis must span the vector space. Thus, the content here is really in showing that such an equation is unique. Suppose there is another set of scalars $d_1, \dots, d_p \in \mathbb{R}$ such that

$$\vec{v} = d_1\vec{v}_1 + \cdots + d_p\vec{v}_p. \tag{2.7}$$

Subtracting Equation 2.7 from Equation 2.6, we have

$$\vec{0} = \vec{x} - \vec{x} = (c_1 - d_1)\vec{v}_1 + \cdots + (c_p - d_p)\vec{v}_p.$$

Since $\mathcal{B}$ is a basis, we know it is linearly independent, so by Definition 1.3.3 we have that $c_i - d_i = 0$ for all $1 \le i \le p$. Thus, $c_i = d_i$ for all $1 \le i \le p$; our new batch of scalars has to be the same as the original ones. □

This is great. Given a basis for a vector space, we can write every vector in that vector space as a *unique* linear combination of the basis vectors. Then we only really need to know the weights, right? Each set of weights associated to each vector provides all the information you need for that vector.

Definition 2.3.1 *Let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_p\}$ be a basis for a vector space V, and suppose $\vec{v} \in V$. The* ***coordinates for $\vec{v}$ relative to $\mathcal{B}$*** *(or the* ***$\mathcal{B}$-coordinates of $\vec{v}$****) are the weights $c_1, \dots, c_p$ such that*

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_p\vec{v}_p. \tag{2.8}$$

The ***coordinate vector of $\vec{v}$ relative to $\mathcal{B}$*** *is*

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}. \tag{2.9}$$

Yay! Now we can use coordinates relative[20] to a basis to write any vector in any p-dimensional vector space as a vector in $\mathbb{R}^p$! Note that the order of the coordinates is determined entirely by the order of the basis vectors.

20: Some other books say "with respect to $\mathcal{B}$" rather than "relative to $\mathcal{B}$."
What if $\mathcal{B}$ is not a respectable basis?
Then those other books look very silly.

Example 2.3.1 Let

$$\vec{x} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} \in \mathbb{R}^2.$$

Using the standard basis for $\mathbb{R}^2$, $\{\vec{e}_1, \vec{e}_2\}$, we have that $\vec{x} = 4\vec{e}_1 + 0\vec{e}_2$, but this is neither interesting nor exciting. Perhaps we could use a different basis, $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$, where

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \text{ and } \vec{v}_2 = \begin{bmatrix} 5 \\ -6 \end{bmatrix}.$$

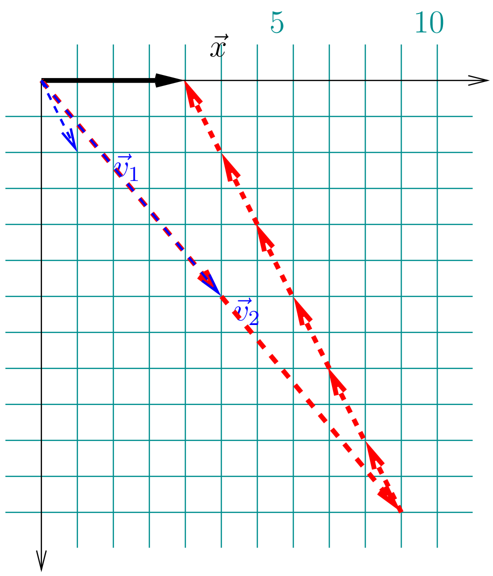


FIGURE 2.1. The vector $\vec{x} \in \mathbb{R}^2$ is shown with a solid arrow line. Using $\{\vec{v}_1, \vec{v}_2\}$ (with dashed arrow line) as a basis for $\mathbb{R}^2$, we see that $\vec{x} = -6\vec{v}_1 + 2\vec{v}_2$.

We can still write $\vec{x}$ as a linear combination of the vectors in this basis:

$$\vec{x} = -6\vec{v}_1 + 2\vec{v}_2,$$

and this can be seen in Figure 2.1. Thus, the coordinates for $\vec{x}$ relative to $\mathcal{B}$ are

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} -6 \\ 2 \end{bmatrix},$$

indicating that to get the vector $\vec{x}$ with vectors from $\mathcal{B}$, you'll need -6 $\vec{v}_1$'s and 2 $\vec{v}_2$'s. You can think of these coordinates (the weights) as an address using "the directions" given in your basis.

Exploration 47 Let's apply what we just learned to $\mathbb{P}_2$. Let

$$\mathcal{B}_1 = \{1, x, x^2\} \text{ and } \mathcal{B}_2 = \{1+x, x+x^2, x\}.$$

- Let $\vec{v} = 2 + 3x + 4x^2$. Find $[\vec{v}]_{\mathcal{B}_1}$ and $[\vec{v}]_{\mathcal{B}_2}$.

- Let $\vec{u} = a + bx + cx^2$. Find $[\vec{u}]_{\mathcal{B}_1}$ and $[\vec{u}]_{\mathcal{B}_2}$.

Let's think about what we've just done. We've used two different bases for $\mathbb{P}_2$ to write a vector in $\mathbb{P}_2$ in a way that looks like a vector in $\mathbb{R}^3$. This seems quite useful. In fact, based on what we know from Chapter 0, it seems like there could be a function from $\mathbb{P}_2$ to $\mathbb{R}^3$ floating around here. However, the vectors are different when the bases are different, so there's not just a single way to do this.

Geometry in Vector Spaces

We need to tidy up our geometric intuition a bit with respect to more general vector spaces. In the previous chapter, we learned about an inner product on $\mathbb{R}^n$ as well as the concept of distance in $\mathbb{R}^n$; we just need to generalize these definitions so they can be used for any vector space.

If only there was a way to write vectors from a vectors space V as vectors in $\mathbb{R}^n$...[21]

Yes! Very good! Coordinate vectors![22]

21: I know this. Don't tell me.

22: I was just about to say that!

Definition 2.3.2 *Let V be a vector space with basis $\mathcal{B}$ and let $[\cdot]_{\mathcal{B}} : V \to \mathbb{R}^n$ be the function that relates vectors in V to their coordinate vector relative to $\mathcal{B}$ in $\mathbb{R}^n$. The* ***inner product on** V **relative to** $\mathcal{B}$ *is the function $\cdot_{\mathcal{B}} : V \times V \to \mathbb{R}$ defined as the composition of $[\cdot]_{\mathcal{B}} \times [\cdot]_{\mathcal{B}}$ on $V \times V$ with the standard inner product $\cdot$ on $\mathbb{R}^n \times \mathbb{R}^n$ (from Section 1.2). That is, for any vectors $\vec{v}, \vec{u} \in V$, we define*

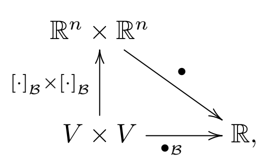
$$\vec{v} \cdot_{\mathcal{B}} \vec{u} = [\vec{v}]_{\mathcal{B}} \cdot [\vec{u}]_{\mathcal{B}} . \tag{2.10}$$

There are a couple of annoying things about this. First, the inner product on a vector space depends on the basis you're using. Unfortunately, you're just going to have to deal with that; it's a fact of life. Second, the notation is a bit obnoxious; since the inner product depends on the basis, the dot for our function name has to carry around a $\mathcal{B}$ as a subscript. We propose this compromise. If we all agree to understand that the inner product on a vector space is dependent on the basis you're using, we can just suppress the subscript $\mathcal{B}$. Then Equation (2.10) becomes

$$\vec{v} \cdot \vec{u} = [\vec{v}]_{\mathcal{B}} \cdot [\vec{u}]_{\mathcal{B}} ,$$

which definitely looks better.

We also quickly glossed over that "composition" bit in Definition 2.3.2. We'll discuss composition in greater detail in Section 3.1, but for now, just think of function composition as using the outputs of the first functions as the inputs of the second function. A diagram like this



called a **commuting diagram**, is often used to describe this situation. It is read as follows: to get the inner product from $V \times V$ to $\mathbb{R}$ (the arrow from left to right), you take coordinate vectors (the up arrow) then the standard inner product on $\mathbb{R}^n$ (the diagonal arrow).

This all probably sounds much more complicated than it actually is. Let's try some examples.

Example 2.3.2 Let $\vec{p} = x^2 - 2x + 1$ and $\vec{q} = x^2 + 2x + 6$ be vectors in $\mathbb{P}_2$. What is the inner product of $\vec{p}$ and $\vec{q}$? How does one take the inner product of two polynomials? Right. This is exactly the issue we've been solving. Before you can take an inner product, you need the coordinate vectors for $\vec{p}$ and $\vec{q}$ relative to some basis. Let's use the standard basis for $\mathbb{P}_2$: $\mathcal{B} = \{1, x, x^2\}$. Then

$$[\vec{p}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \text{ and } [\vec{q}]_{\mathcal{B}} = \begin{bmatrix} 6 \\ 2 \\ 1 \end{bmatrix}.$$

Using the coordinate vectors, we have

$$\vec{p} \cdot \vec{q} = [\vec{p}]_{\mathcal{B}} \cdot [\vec{q}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 2 \\ 1 \end{bmatrix} = 6 - 4 + 1 = 3.$$

Exploration 48 We've said that this inner product depends on the basis, since the coordinate vectors depend on the basis. Let's see this in action. Again, let $\vec{p} = x^2 - 2x + 1$ and $\vec{q} = x^2 + 2x + 6$. Back in Section 2.1, we showed that $\mathcal{B}_1 = \{1 + x, x + x^2, x\}$ also formed a basis of $\mathbb{P}_2$.

- ► Find the coordinate vectors $[\vec{p}]_{\mathcal{B}_1}$ and $[\vec{q}]_{\mathcal{B}_1}$. (You might find it helpful to look back at Example 2.1.2.)

- ► Now, use these coordinate vectors to compute $\vec{p} \cdot \vec{q}$.

  The answer's different, right? So, the inner product can depend on the basis.

Exploration 49 In the exercises of Section 2.2, we saw that $\{\vec{1}, \vec{i}\}$ forms a basis for the complex numbers $\mathbb{C} = \{a + bi \colon a, b \in \mathbb{R}\}$ when viewed as a real vector space. Use this basis to compute $(1 \vec{+} i) \cdot (-1 \vec{+} 2i)$. Keep in mind that the dot here indicates inner product of vectors and *not* multiplication of complex numbers.

This whole "just use the coordinate vector" strategy is pretty great. Let's apply it to length, too!

Definition 2.3.3 *Let V be a vector space with basis $\mathcal{B}$ and let $[\cdot]_{\mathcal{B}} : V \to \mathbb{R}^n$ be the function that relates vectors in V to their coordinate vector relative to $\mathcal{B}$ in $\mathbb{R}^n$.* ***Length** relative to* $\mathcal{B}$ *(or the* ***norm** relative to* $\mathcal{B}$*) is the function $\|\cdot\|_{\mathcal{B}} : \mathbb{V} \to \mathbb{R}$ defined by relating vectors to their length by composing the function $[\cdot]_{\mathcal{B}} \times [\cdot]_{\mathcal{B}}$ on V with $\|\cdot\|$ on $\mathbb{R}^n$. That is, for any vector $\vec{v} \in V$, we define*

$$\|\vec{v}\|_{\mathcal{B}} = \|[\vec{v}]_{\mathcal{B}}\|\,. \tag{2.11}$$

Again, we'll all agree that length is also relative to basis and regularly omit the subscript $\mathcal{B}$ on the norm of vectors in general vector spaces. As before, the idea is that if you don't have a vector in $\mathbb{R}^n$, you'll need to make it a coordinate vector before you can take its length.

Example 2.3.3 What is the norm of $\vec{p} = x^3 + 4x - 2 \in \mathbb{P}_3$? The length of this vector depends on your choice of basis. For lack of imagination, we'll use the standard basis again, $\mathcal{B} = \{1, x, x^2, x^3\}$. Then

$$[\vec{p}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 4 \\ 0 \\ 1 \end{bmatrix},$$

so

$$\|\vec{p}\| = \|[\vec{p}]_{\mathcal{B}}\| = \left\| \begin{bmatrix} -2 \\ 4 \\ 0 \\ 1 \end{bmatrix} \right\| = \sqrt{(-2)^2 + 4^2 + 0^2 + 1^2} = \sqrt{21}$$

Exploration 50 Now let's do something similar using $\mathbb{P}_2$. As before, let $\mathcal{B} = \{1, x, x^2\}$ and $\mathcal{B}_1 = \{1 + x, x + x^2, x\}$. Let $\vec{p} = x + 5$.

- Find $[\vec{p}]_{\mathcal{B}}$ and $[\vec{p}]_{\mathcal{B}_1}$.
- Find $\|\vec{p}\|_{\mathcal{B}}$.
- Find $\|\vec{p}\|_{\mathcal{B}_1}$.

Example 2.3.4 Now, we again have the concept of unit vectors. Let's find a unit vector in the direction of $\vec{q} = 5x^4 - 2x + 3 \in \mathbb{P}_4$ using the standard basis $\mathcal{B}$ for $\mathbb{P}_4$. First, the coordinate vector for $5x^4 - 2x + 3$ is

$$[\vec{q}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ 0 \\ 5 \end{bmatrix}.$$

Now, we find that the length of this vector is

$$\|\vec{q}\| = \sqrt{9+4+25} = \sqrt{38}.$$

The unit vector is then the vector with coordinate vector given by

$$\begin{bmatrix} \frac{3}{\sqrt{38}} \\ \frac{-2}{\sqrt{38}} \\ 0 \\ 0 \\ \frac{5}{\sqrt{38}} \end{bmatrix}.$$

To find the actual unit vector, we should then translate this back into a vector in $\mathbb{P}_4$. Our final answer is then

$$\frac{3}{\sqrt{38}} - \frac{2x}{\sqrt{38}} + \frac{5x^4}{\sqrt{38}}.$$

Moving forward, we'll often need to make use of inner products. Since one can choose from a lot of different bases and inner products on a vector space, we'll usually need to agree on a basis and inner product before proceeding with the use of any inner product.

Definition 2.3.4 *We call a vector space, V, together with inner product relative to basis $\mathcal{B}$ an* ***inner product space****.*

In practice, whenever you see the words "inner product space," know that we're just talking about a vectors space with a specifically chosen inner product and basis.

Now for something completely different. What does it mean for vectors to be perpendicular in $\mathbb{R}^2$? Right,[23] their directions differ by $90°$ or, better yet, $\pi/2$ radians. What about in $\mathbb{R}^3$ or, better yet, $\mathbb{R}^5$? Can you imagine what it means for "directions differ by $90°$" in $\mathbb{R}^5$? Right. The notion of perpendicular is too restricted by geometry in $\mathbb{R}^n$ (for small n!) to be useful for general vector spaces. We need something better.

23: Pun intended.

Definition 2.3.5 *Vectors $\vec{v}$ and $\vec{u}$ in vector space V with chosen basis $\mathcal{B}$ are said to be* ***orthogonal*** *if $\vec{v} \cdot \vec{u} = 0$.*

Great, but does it still work the same as perpendicular in $\mathbb{R}^n$ for small n? It does!

Theorem 2.3.2 *Two nonzero vectors $\vec{v}$ and $\vec{u}$ in $\mathbb{R}^2$ are orthogonal with respect to the standard inner product if and only if they are perpendicular.*

PROOF. We first need to better define what it means to be perpendicular in $\mathbb{R}^2$. Let α be the angle between $\vec{v}$ and $\vec{u}$. By the Law of Cosines and definition of distance,

$$\begin{aligned} \text{dist}\,(\vec{v},\vec{u})^2 &= \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\|\vec{v}\|\|\vec{u}\|\cos\alpha \quad \text{and} \\ \text{dist}\,(\vec{v},\vec{u})^2 &= \|\vec{v}-\vec{u}\|^2 \\ &= (\vec{v}-\vec{u})\cdot(\vec{v}-\vec{u}) \\ &= \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\vec{v}\cdot\vec{u}. \end{aligned}$$

This gives us two different expressions, each equal to $\operatorname{dist}(\vec{v},\vec{u})^2$. Thus we have

$$\|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\|\vec{v}\|\|\vec{u}\|\cos\alpha = \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\vec{v}\cdot\vec{u}.$$

Once we simplify this a bit, it follows that

$$\vec{v}\cdot\vec{u} = \|\vec{v}\|\|\vec{u}\|\cos\alpha.$$

Since $\vec{v}$ and $\vec{u}$ are nonzero, $\vec{v}\cdot\vec{u} = 0$ if and only if $\alpha = 90°$. □

Also, we do have a concept of the angle between vectors in $\mathbb{R}^3$. Since any two vectors in $\mathbb{R}^3$ lie in a plane together, the angle between the vectors can be determined by the angle between them specifically in that plane.[24] Our definition of orthogonality also agrees with the use of perpendicular to mean the angle between two vectors in $\mathbb{R}^3$ is $90°$.

24: The Law of Cosines also works in that plane as well!

Orthogonality is great. In $\mathbb{R}^n$ (for small n), it means the same thing as perpendicular, but it also works in any vector space. That said, for the sake of consistency, we shall only use the word "orthogonal" in this course.

Exploration 51 Let

$$\vec{v}_1 = \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Note that $\vec{v}_2\cdot\vec{v}_3 = 0$, so $\vec{v}_2$ and $\vec{v}_3$ are orthogonal to each other. Find all pairs of orthogonal vectors.

Exploration 52 Let $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. The vector $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ is orthogonal to $\vec{v}$. Find a vector $\vec{w}$ orthogonal to $\vec{v}$ so that $\{\vec{u},\vec{w}\}$ is a linearly independent set.

As was noted in Section 1.2, length in vector spaces is a generalization of absolute value in $\mathbb{R}$. It should come as little surprise then that many of your favorite facts about absolute value are true for length in vector spaces.

Theorem 2.3.3 (The Triangle Inequality) *For any two vectors $\vec{v}$ and $\vec{u}$ in vector space V,*

$$\|\vec{v}+\vec{u}\| \leq \|\vec{v}\| + \|\vec{u}\|.$$

The proof for this fact would take us a bit off track, so we've omitted it.

Theorem 2.3.4 (The Pythagorean Theorem) *For any two orthogonal vectors $\vec{v}$ and $\vec{u}$ in a vector space V,*

$$\|\vec{v}+\vec{u}\|^2 = \|\vec{v}\|^2 + \|\vec{u}\|^2$$

Note that in the theorem above, the length function is computed using the same basis for which the vectors are orthogonal.

PROOF. First, we need a fact about coordinate vectors. Let $\vec{v}$ and $\vec{u}$ be vectors in a vector space V with basis $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$. Then there are real numbers $v_1, \ldots, v_n$ and $u_1, \ldots, u_n$ so that

$$\begin{aligned}\vec{v} =& v_1\vec{b}_1 + \cdots + v_n\vec{b}_n \\ \vec{u} =& u_1\vec{b}_1 + \cdots + u_n\vec{b}_n\end{aligned}$$

Thus, $\vec{v} + \vec{u} = (v_1\vec{b}_1 + \cdots + v_n\vec{b}_n) + (u_1\vec{b}_1 + \cdots + u_n\vec{b}_n) = (v_1 + u_1)\vec{b}_1 + \cdots + (v_n + u_n)\vec{b}_n$. From this equation we see that

$$[\vec{v} + \vec{u}]_{\mathcal{B}} = [\vec{v}]_{\mathcal{B}} + [\vec{u}]_{\mathcal{B}}\,.$$

Now, we will use this to prove our theorem. Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. Then, by the definition of length in $\mathbb{R}^n$ we have

$$\begin{aligned}\|\vec{x} + \vec{y}\|^2 =& \left(\sqrt{(x_1 + y_1)^2 + \cdots + (x_n + y_n)^2}\right)^2 \\ =& (x_1 + y_1)^2 + \cdots + (x_n + y_n)^2 \\ =& (x_1^2 + 2x_1y_1 + y_1^2) + \cdots + (x_n^2 + 2x_ny_n + y_n^2) \\ =& (x_1^2 + \cdots + x_n^2) + 2(x_1y_1 + \cdots + x_ny_n) + (y_1^2 + \cdots + y_n^2) \\ =& \|\vec{x}\|^2 + 2\vec{x} \cdot \vec{y} + \|\vec{y}\|^2\end{aligned}$$

Then if we know in addition that $\vec{x}$ and $\vec{y}$ are orthogonal, then $2\vec{x} \cdot \vec{y} = 0$ and we have

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2.$$

Thus, the theorem holds for vectors in $\mathbb{R}^n$.

Now let $\vec{v}$ and $\vec{u}$ be orthogonal vectors in a vector space V with basis $\mathcal{B}$. We then have

$$\begin{aligned}\|\vec{v} + \vec{u}\|^2 =& \|\,[\vec{v} + \vec{u}]_{\mathcal{B}}\,\|^2 \\ =& \|\,[\vec{v}]_{\mathcal{B}} + [\vec{u}]_{\mathcal{B}}\,\|^2 \\ =& \|\,[\vec{v}]_{\mathcal{B}}\,\|^2 + \|\,[\vec{u}]_{\mathcal{B}}\,\|^2 \\ =& \|\vec{v}\|^2 + \|\vec{u}\|^2\end{aligned}$$

□

Section Highlights

- There is a unique way to write any vector $\vec{v}$ in a vector space V as a linear combination of the vectors in a basis for V. The coefficients from this linear combination form the *coordinate vector* for $\vec{v}$. See Theorem 2.3.1 and Definition 2.3.1.
- Coordinate vectors can be used to define the inner product like in Section 1.2 for any vector space. See Definition 2.3.2.

- The concept of perpendicular vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ is extended to general vector spaces of dimension n as orthogonality. Two vectors are said to be orthogonal if their inner product is zero. See Definition 2.3.5 and Theorem 2.3.2.

Exercises for Section 2.3

2.3.1. Let $\mathcal{B} = \left\{1 + x, 2x, x + x^2\right\}$. This is a basis for $\mathbb{P}_2$.

(a) Find $\left[1 + x + x^2\right]_{\mathcal{B}}$ and $[1]_{\mathcal{B}}$.

(b) Suppose $\vec{p} \in \mathbb{P}_2$ has coordinate vector $[\vec{p}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$. Find the polynomial $\vec{p} \in \mathbb{P}_2$.

2.3.2. Let $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \end{bmatrix} \right\}$. This is a basis for $\mathbb{R}^2$.

(a) Find $\begin{bmatrix} 4 \\ 6 \end{bmatrix}_{\mathcal{B}}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\mathcal{B}}$.

(b) Suppose $\vec{x} \in \mathbb{R}^2$ has coordinate vector $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Find $\vec{x}$.

2.3.3. Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$. This is a basis for $\mathbb{R}^3$.

(a) Find $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}}$ and $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}}$.

(b) Suppose $\vec{x} \in \mathbb{R}^3$ has coordinate vector $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$. Find $\vec{x}$.

2.3.4. Below are two bases for $\mathbb{P}_2$:

$$\mathcal{B}_1 = \left\{x, 1 + x, x^2\right\} \text{ and}$$
$$\mathcal{B}_2 = \left\{1 + x, x + x^2, x^2\right\}.$$

Find $\left[1 + 2x - 3x^2\right]_{\mathcal{B}_1}$ and $\left[1 + 2x - 3x^2\right]_{\mathcal{B}_2}$.

2.3.5. Below are two bases for $\mathbb{R}_3$:

$$\mathcal{B}_1 = \left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix} \right\} \text{ and}$$
$$\mathcal{B}_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix} \right\}.$$

Find $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}_1}$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}_2}$.

2.3.6. Let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ be a basis for a vector space V. Show that for any constants k_1 and k_2,

$$\left[k_1\vec{b}_1\right]_{\mathcal{B}} = \begin{bmatrix} k_1 \\ 0 \end{bmatrix} \quad \text{and} \quad \left[k_2\vec{b}_2\right]_{\mathcal{B}} = \begin{bmatrix} 0 \\ k_2 \end{bmatrix}.$$

2.3.7. Find the distance between

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}.$$

2.3.8. Find the distance, using the standard basis $\{1, x, x^2, x^3\}$ for $\mathbb{P}_3$, between $1 + x + x^3$ and $-2x - 2x^3$.

2.3.9. Let

$$\vec{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \text{ and } \vec{v}_4 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}.$$

Find all pairs of orthogonal vectors among $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$, and $\vec{v}_4$.

2.3.10. Let

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \text{ and } \vec{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Find all pairs of orthogonal vectors among $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$, and $\vec{v}_4$.

2.3.11. Let $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. The vector $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ is orthogonal to $\vec{v}$. Find a vector $\vec{w}$ orthogonal to $\vec{v}$ so that $\{\vec{u}, \vec{w}\}$ is a linearly independent set.

2.3.12. Let $\vec{v}, \vec{u}_1, \vec{u}_2$ be vectors in some vector space V.

(a) Suppose $\vec{v}$ is orthogonal to $\vec{u}_1$ and $\vec{u}_2$. Use properties of the dot product to argue that $\vec{v}$ is orthogonal to $\vec{u}_1 + \vec{u}_2$.

(b) Suppose $\vec{v}$ is orthogonal to $\vec{u}_1 + \vec{u}_2$. Must it be true that $\vec{v}$ is orthogonal to $\vec{u}_1$ and $\vec{u}_2$? Explain.

(c) Suppose $\vec{v}$ is orthogonal to $\vec{u}_1$ and $\vec{u}_2$. If $\vec{x} \in \text{Span}\,\{\vec{u}_1, \vec{u}_2\}$, must $\vec{v}$ be orthogonal to $\vec{x}$? Explain.

2.3.13. Let $\vec{v} \in \mathbb{R}^n$. Why is $\|\vec{v}\| \geq 0$? When is $\|\vec{v}\| = 0$? Explain.

2.3.14. Let $\mathcal{B}$ be the standard basis for $\mathbb{P}_3$ and $\mathcal{C} = \{1, 1 + x, x + x^2, x^2 + x^3\}$, which is also a basis of $\mathbb{P}^3$. Let $\vec{v} = 2 + x + x^3$ and $\vec{u} = x + 3x^2 - x^3$.

(a) Find $[\vec{v}]_{\mathcal{B}}$ and $[\vec{u}]_{\mathcal{B}}$.

(b) Find $\vec{v} \cdot \vec{u}$ with respect to $\mathcal{B}$.

(c) Are $\vec{v}$ and $\vec{u}$ orthogonal with respect to $\mathcal{B}$?

(d) Find $\|\vec{v}\|$ and $\|\vec{u}\|$ with respect to $\mathcal{B}$.

(e) Find $[\vec{v}]_{\mathcal{C}}$ and $[\vec{u}]_{\mathcal{C}}$.

(f) Find $\vec{v} \cdot \vec{u}$ with respect to $\mathcal{C}$.

(g) Are $\vec{v}$ and $\vec{u}$ orthogonal with respect to $\mathcal{C}$?

(h) Find $\|\vec{v}\|$ and $\|\vec{u}\|$ with respect to $\mathcal{C}$.

2.3.15. Let $\mathcal{B}$ be the standard basis for $\mathbb{P}_3$. Let $\vec{v} = x + 3x^2 + x^3$ and $\vec{u} = 2 + 3x - x^2$.

(a) Find $[\vec{v}]_{\mathcal{B}}$ and $[\vec{u}]_{\mathcal{B}}$.

(b) Find $\vec{v} \cdot \vec{u}$ with respect to $\mathcal{B}$.

(c) Are $\vec{v}$ and $\vec{u}$ orthogonal with respect to $\mathcal{B}$?

(d) Give a new vector $\vec{w}$ so that $\vec{w}$ is orthogonal to $\vec{v}$.

(e) Find $\|\vec{v}\|$ and $\|\vec{u}\|$ with respect to $\mathcal{B}$.

2.3.16. Note that $\mathcal{B}_1 = \{1, 1+x, x+x^2, x^3\}$ is a basis for $\mathbb{P}_3$. Let $\vec{v} = x + 3x^2 + x^3$ and $\vec{u} = 2 + 3x - x^2$.

(a) Find $[\vec{v}]_{\mathcal{B}_1}$ and $[\vec{u}]_{\mathcal{B}_1}$.

(b) Find $\vec{v} \cdot \vec{u}$ with respect to $\mathcal{B}_1$.

(c) Are $\vec{v}$ and $\vec{u}$ orthogonal with respect to $\mathcal{B}_1$?

(d) Give a new vector $\vec{w}$ so that $\vec{w}$ is orthogonal to $\vec{v}$.

(e) Find $\|\vec{v}\|$ and $\|\vec{u}\|$ with respect to $\mathcal{B}_1$.

2.3.17. Let S be the unit cube in $\mathbb{R}^3$. That is, let S be the cube with corners $(0,0,0)$, $(0,1,0)$, $(1,0,0)$, $(1,1,0)$, $(0,0,1)$, $(0,1,1)$, $(1,0,1)$, and $(1,1,1)$. There are four different diagonals in S (segments that connect one corner of S to another and go through the center of S). By subtracting corners, we can write these diagonals as vectors in $\mathbb{R}^3$. Show that any two of these diagonals are *not* orthogonal.

2.3.18. Let V be a vector space with basis $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$. Define $T\colon V \to \mathbb{R}^n$ by $T(\vec{v}) = [\vec{v}]_{\mathcal{B}}$. Such a mapping T is often called a *coordinate mapping*. Show that for any $\vec{v}_1, \vec{v}_2 \in V$ and any $a \in \mathbb{R}$, we have

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) \quad \text{and} \quad T(a\vec{v}_1) = aT(\vec{v}_1).$$

2.3.19. Let $\vec{v}, \vec{u} \in \mathbb{R}^n$. Following techniques similar to the proof of the Pythagorean Theorem, show $\|\vec{v}+\vec{u}\|^2 + \|\vec{v}-\vec{u}\|^2 = 2\|\vec{v}\|^2 + 2\|\vec{u}\|^2$. This is called the Parallelogram Law.

2.3.20. Let $\vec{v}, \vec{u} \in \mathbb{R}^n$. Use the Triangle Inequality to show $\big|\|\vec{v}\| - \|\vec{u}\|\big| \leq \|\vec{v}-\vec{u}\|$ using the fact that $\vec{v} = \vec{v}-\vec{u}+\vec{u}$. This is called the Reverse Triangle Inequality.

2.4 Orthogonal Sets

Here we will extend the notion of orthogonality to sets; it will end up being very useful. When we say useful here, we mean it will produce a formula that can save you a *lot* of time.

Definition 2.4.1 *Let V be an inner product space and W be a subspace of V. If a vector $\vec{v} \in V$ is orthogonal to every vector in W, then we say $\vec{v}$ is orthogonal to W. The set of all vectors $\vec{v} \in V$ that are orthogonal to W is called the* ***orthogonal complement of*** *W. The orthogonal complement of W is denoted $W^{\perp}$.*

Note that orthogonality is defined in terms of an inner product, which depends on the basis one uses. Thus, orthogonality always depends on the choice of basis in a vector space, so we'll almost always be working in inner product spaces (where we have a chosen basis and inner product) when orthogonality is relevant. Also for clarification, the notation $W^{\perp}$ is usually pronounced "W perp."[25]

Example 2.4.1 Let

$$W = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right\}.$$

This is a subspace in $\mathbb{R}^3$. In particular, this is a line through the origin in $\mathbb{R}^3$. Thus, the vectors orthogonal to this line form a plane. Specifically, they form a plane through the origin since the zero vector is orthogonal to every vector in $\mathbb{R}^3$. Now let us determine this plane.

Let $\vec{v} \in W$. Since $\vec{0}$ is orthogonal to every vector, we should specify that $\vec{v} \neq \vec{0}$. Then for some $c \in \mathbb{R}$ with $c \neq 0$,

$$\vec{v} = \begin{bmatrix} c \\ 2c \\ 3c \end{bmatrix}.$$

Let

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

be a vector orthogonal (using the standard basis for $\mathbb{R}^3$) to $\vec{v}$. We know

$$\begin{bmatrix} c \\ 2c \\ 3c \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = cx + 2cy + 3cz = 0.$$

Thus,

$$W^{\perp} = \left\{\begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + 2y + 3z = 0\right\}.$$

Since $c \neq 0$, we were able to algebraically clean up our condition and require simply that $x + 2y + 3z = 0$.

25: As in "purple?"
No.
As in "perpetrator?"
No.
"Porpoise?"
It's not pronounced "porp!"
It is now.

However, you may perhaps feel unsatisfied with this description of $W^\perp$. We can also describe $W^\perp$ as

$$\text{Span}\left\{ \begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} \right\},$$

where the appropriate vectors were determined by substituting into the equation $x + 2y + 3z = 0$. We know these two vectors are linearly independent since they are not scalar multiples of each other. Thus, they must span our 2-dimensional plane. This feels more like a description we would typically see in this book.

We saw in the example that a way to find the orthogonal complement of a subspace W is to consider a general element of W and determine what it means for some other vector to be orthogonal to that vector. However, we can actually streamline this a bit. Rather than taking a general element, we can instead separately consider each element in a spanning set of W. In the example above, this would mean it was enough to just find all vectors orthogonal to

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

which amounts to not including the scalar multiple c. Since we simplified out our c anyway, you can see that was exactly what we needed!

Theorem 2.4.1 *Let W be a subspace of an inner product space V, and let $S = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be a set of vectors that spans W. Then $\vec{v} \in W^\perp$ if and only if $\vec{v}$ is orthogonal to every vector in S.*

PROOF. Suppose $\vec{v} \in W^\perp$. Then $\vec{v}$ is orthogonal to every vector in W, including those in S. Now suppose $\vec{v}$ is orthogonal to every vector in S. Since S spans W, any vector $\vec{w} \in W$ can be written as a linear combination of the vectors in S:

$$\vec{w} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n.$$

Since $\vec{v}$ is orthogonal to $\vec{v}_i$ for $1 \leq i \leq n$,

$$\begin{aligned} \vec{v} \cdot \vec{w} &= \vec{v} \cdot (a_1\vec{v}_1 + \cdots + a_n\vec{v}_n) \\ &= \vec{v} \cdot (a_1\vec{v}_1) + \cdots + \vec{v} \cdot (a_n\vec{v}_n) \\ &= a_1(\vec{v} \cdot \vec{v}_1) + \cdots + a_n(\vec{v} \cdot \vec{v}_n) \\ &= a_1(0) + \cdots + a_n(0) \\ &= 0. \end{aligned}$$

Thus, $\vec{v} \in W^\perp$. It follows that $\vec{v} \in W^\perp$ if and only if $\vec{v}$ is orthogonal to every vector in S. □

Theorem 2.4.2 *If W is a subspace of and inner product space V, then $W^\perp$ is a subspace of V.*

Exploration 53 To show $W^\perp$ is a subspace, we will verify the three axioms of being a subspace.

- First, we know $\vec{0} \in W^\perp$ because the zero vector is orthogonal to every vector.
- Now, let $\vec{v}$ and $\vec{u}$ be in $W^\perp$. Let $\vec{w} \in W$. To show $W^\perp$ is closed under addition, we need to verify that $(\vec{v}+\vec{u})\cdot\vec{w} = 0$. Use a property of the inner product to argue that this must be true.
- Lastly, we need to show that $W^\perp$ is closed under scalar multiplication. Let $k \in \mathbb{R}$. Suppose as before that $\vec{v} \in W^\perp$ and $\vec{w} \in W$. Then we need to verify $(k\vec{v}) \cdot \vec{w} = 0$. Again, use a property of the inner product to show this.

Example 2.4.2 Let $\vec{v}$ and $\vec{u}$ be any two linearly independent vectors in $\mathbb{R}^3$ with basis $\mathcal{B}$. Then $W = \text{Span}\{\vec{v}, \vec{u}\}$ is a two-dimensional subspace of $\mathbb{R}^3$ that can be visualized as a plane through the origin in $\mathbb{R}^3$. What do we know about $W^\perp$? It's a subspace of $\mathbb{R}^3$, it must contain the zero vector, and all of its vectors must be orthogonal to W, which is a plane through the origin. What does it mean to be orthogonal to all the vectors in a plane in $\mathbb{R}^3$? As an arrow vector, it would have to stick straight out of the plane. Imagine, for example, that a tabletop is the plane, and you set a marker or pen standing up on its end on the table; an arrow vector orthogonal to that plane would point in the same direction as the marker (or pen). It could also point directly down. It could also be any length! Our set of orthogonal vectors does, as we've noted, have to contain the zero vector, too. Thus, $W^\perp$ is the line through the origin orthogonal to W. Moreover, $\dim W^\perp = 1$.

Exploration 54 In Example 2.4.1, we saw the orthogonal complement of a 1-dimensional subspace in $\mathbb{R}^3$ had dimension 2. Then in Example 2.4.2, we used our geometric understanding to note that the orthogonal complement of a 2-dimensional subspace of $\mathbb{R}^3$ is 1-dimensional. Now let $\vec{v}$ and $\vec{u}$ be any two linearly independent vectors in $\mathbb{R}^4$. Again, $W = \text{Span}\{\vec{v}, \vec{u}\}$ is a two-dimensional subspace of $\mathbb{R}^4$, but it's harder to visualize this time. What is $W^\perp$, and what is $\dim W^\perp$?

Well, if you are having trouble figuring this out, let's see if a more specific example will help. Use the standard basis for $\mathbb{R}^4$, and let

$$W_0 = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\} \text{ and } \vec{v} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in W_0^\perp.$$

Then we know

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = a = 0, \quad \text{and} \quad \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = b = 0.$$

Using this information, can you determine $W_0^{\perp}$?

Well, these examples do seem to be following a pattern. However, examples are not enough to conclude a theorem. Let's instead make a "conjecture." This is really just a guess based on some evidence. Hopefully, we'll be able to prove this conjecture in a future section.[26]

26: Perhaps even the next section.

Conjecture 1 *Let W be a subspace of an inner product space V such that $\dim V = n$ and $\dim W = p$. Then $\dim W^{\perp} = n - p$.*

Orthogonal Sets and Bases

Definition 2.4.2 *If S is a set of vectors in an inner product space such that all pairs of vectors in S are orthogonal, then S is said to be an* ***orthogonal set****.*

Theorem 2.4.3 *Orthogonal sets of nonzero vectors are linearly independent.*

PROOF. We will use the definition of linear independence to show an orthogonal set $S = \{\vec{v}_1, \ldots, \vec{v}_p\}$ of nonzero vectors is linearly independent. Suppose there are scalars $a_1, \ldots, a_p$ such that $a_1\vec{v}_1 + \cdots + a_p\vec{v}_p = \vec{0}$. Taking the inner product of both sides of this equation with $\vec{v}_i$ for some $1 \leq i \leq p$, we have

$$\begin{aligned} 0 = \vec{v}_i \cdot \vec{0} &= \vec{v}_i \cdot (a_1\vec{v}_1 + \cdots + a_p\vec{v}_p) \\ &= a_1\vec{v}_i \cdot \vec{v}_1 + \cdots + a_i\vec{v}_i \cdot \vec{v}_i + \cdots + a_p\vec{v}_i \cdot \vec{v}_p \\ &= a_i\vec{v}_i \cdot \vec{v}_i, \end{aligned}$$

since $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$ (by definition of orthogonal set). Since $\vec{v}_i \neq \vec{0}$, we know $\vec{v}_i \cdot \vec{v}_i \neq 0$. Thus, $a_i = 0$. Since this argument holds for any $1 \leq i \leq p$, it follows that S is linearly independent. □

Sometimes a another useful result follows immediately from a theorem. We call such a result a corollary to the theorem. The following is a corollary to Theorem 2.4.3.

Corollary 2.4.4 *If S is an orthogonal set in an inner product space V, then S is a basis for the subspace* $\text{Span}\,\{S\}$.

Orthogonality is a property enjoy by sets of vectors in a variety of contexts. In fact, it's such an enjoyable property, it gets its own definition with one of favorite sets of vectors:

Definition 2.4.3 *An* ***orthogonal basis*** *for a subspace W is a basis for W that is also an orthogonal set.*

What makes an orthogonal basis so enjoyable? We're glad you asked.

Theorem 2.4.5 *Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ be an orthogonal basis for a subspace W of an inner product space V. Then for any $\vec{w} \in W$,*

$$\vec{w} = \left(\frac{\vec{w}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\right)\vec{v}_1 + \left(\frac{\vec{w}\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2}\right)\vec{v}_2 + \cdots + \left(\frac{\vec{w}\cdot\vec{v}_p}{\vec{v}_p\cdot\vec{v}_p}\right)\vec{v}_p,$$

where the inner product is taken with respect to $\mathcal{B}$.

It's no surprise that $\vec{w}$ can be written as a linear combination of the vectors in $\mathcal{B}$ because $\mathcal{B}$ is a basis. The innovation here is the weights; when $\mathcal{B}$ is an *orthogonal* basis, we get a formula for the weights. That's extremely enjoyable. Oh, we should prove it, though.

PROOF. This is a good one. Since $\mathcal{B}$ is a basis, we know there are weights c_i for $1 \leq i \leq p$ such that

$$\vec{w} = c_1\vec{v}_1 + \cdots + c_p\vec{v}_p.$$

To solve for c_i for $1 \leq i \leq p$, apply the dot product with $\vec{v}_i$ to both sides of this equation. We then have

$$\begin{aligned}
\vec{w}\cdot\vec{v}_i &= (c_1\vec{v}_1 + \cdots + c_p\vec{v}_p)\cdot\vec{v}_i \\
&= c_1(\vec{v}_1\cdot\vec{v}_i) + \cdots + c_i(\vec{v}_i\cdot\vec{v}_i) + \cdots + c_p(\vec{v}_p\cdot\vec{v}_i) \\
&= c_1(0) + \cdots + c_i(\vec{v}_i\cdot\vec{v}_i) + \cdots + c_p(0) \\
&= c_i\vec{v}_i\cdot\vec{v}_i,
\end{aligned}$$

where all but one of the dot products is zero because $\mathcal{B}$ is an orthogonal set. Then just solve for c_i; this can be done for each $1 \leq i \leq p$. □

Exploration 55 Let's do an example together to illustrate that theorem!

- ▶ Using the standard basis for $\mathbb{R}^3$, show that the set $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, where

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and } \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

  is an orthogonal set. This is readily verified using the definition of inner product. Find $\vec{v}_1\cdot\vec{v}_2$, $\vec{v}_1\cdot\vec{v}_3$ and $\vec{v}_2\cdot\vec{v}_3$. (Hopefully, you get 0 for each one!)

- ▶ Now, we want to write

$$\vec{x} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

  as a linear combination of the vectors in S. According to Theorem 2.4.5, we may write

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 \text{ where } c_i = \frac{\vec{x}\cdot\vec{v}_i}{\vec{v}_i\cdot\vec{v}_i},$$

  where $1 \leq i \leq 3$. We compute each of the c_i below:

$$\begin{aligned} c_1 &= \frac{\vec{x}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1} = \frac{4(1)+5(2)+6(-5)}{1(1)+2(2)+(-5)(-5)} = \frac{-16}{30}, \\ c_2 &= \frac{\vec{x}\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2} = \frac{4(-2)+5(1)+6(0)}{(-2)(-2)+1(1)+0(0)} = \frac{-3}{5}, \\ c_3 &= \frac{\vec{x}\cdot\vec{v}_3}{\vec{v}_3\cdot\vec{v}_3} = \frac{4(1)+5(2)+6(1)}{1(1)+2(2)+1(1)} = \frac{20}{6}. \end{aligned}$$

Thus,

$$\vec{x} = -\frac{8}{15}\vec{v}_1 - \frac{3}{5}\vec{v}_2 + \frac{10}{3}\vec{v}_3.$$

► Now that we've seen how it's done, let's write

$$\vec{y} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

as a linear combination of the vectors in S. Find each c_i below

$$\begin{aligned} c_1 &= \frac{\vec{y}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1} = \underline{\hspace{6cm}} \\ c_2 &= \frac{\vec{y}\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2} = \underline{\hspace{6cm}}, \\ c_3 &= \frac{\vec{y}\cdot\vec{v}_3}{\vec{v}_3\cdot\vec{v}_3} = \underline{\hspace{6cm}}. \end{aligned}$$

Now, use your answers for c_1, c_2, and c_3 to fill in the blanks below

$$\vec{y} = \underline{\hspace{2.5cm}}\vec{v}_1 + \underline{\hspace{2.5cm}}\vec{v}_2 + \underline{\hspace{2.5cm}}\vec{v}_3.$$

If you have a vector space V, having a basis is very nice. Having an orthogonal basis is even better. Shall we introduce a superlative?

Definition 2.4.4 *If S is an orthogonal set of vectors in an inner product space such that any vector in S is a unit vector, then S is said to be an* ***orthonormal set****. If S happens to be an orthogonal basis and any vector in S is a unit vector, then S is said to be an* ***orthonormal basis****.*

An orthonormal set is just an orthogonal set of unit vectors. Why should we care? In what way is this better?[27]

27: How is being normal better than being "-gonal?" You should never strive for normality. Always be exceptional.

What are you talking about?!

Theorem 2.4.6 *Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ be an orthonormal basis for subspace W of inner product space V. Then for any $\vec{w} \in W$,*

$$\vec{w} = (\vec{w}\cdot\vec{v}_1)\,\vec{v}_1 + (\vec{w}\cdot\vec{v}_2)\,\vec{v}_2 + \cdots + (\vec{w}\cdot\vec{v}_n)\,\vec{v}_n,$$

where the inner product is taken with respect to $\mathcal{B}$.

This theorem follows immediately from the proof of Theorem 2.4.5 once you realize that $\vec{v}_i \cdot \vec{v}_i = 1$ for all $1 \le i \le p$ since each $\vec{v}_i$ is a unit vector (since $\mathcal{B}$ is an orthonormal basis).

Example 2.4.3 The set $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, where

$$\vec{v}_1 = \begin{bmatrix} 1/\sqrt{30} \\ 2/\sqrt{30} \\ -5/\sqrt{30} \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}, \quad \text{and } \vec{v}_3 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

is an orthonormal set. One can see this by noting that

$$\vec{v}_1 = \frac{1}{\sqrt{30}} \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}, \quad \vec{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and } \vec{v}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix};$$

then the fact that this is an orthogonal set follows from properties of inner product and Exploration 55. For example,

$$\begin{aligned} \vec{v}_1 \cdot \vec{v}_2 &= \left(\frac{1}{\sqrt{30}} \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix} \right) \cdot \left(\frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right) \\ &= \left(\frac{1}{\sqrt{30}} \frac{1}{\sqrt{5}} \right) \left(\begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right) = \left(\frac{1}{\sqrt{30}} \frac{1}{\sqrt{5}} \right) 0 = 0. \end{aligned}$$

The remaining pairs of vectors yield similar calculations and could readily be checked. However, we must also verify that each of these vectors is a unit vector:

$$\begin{aligned} \|\vec{v}_1\| &= \left\| \frac{1}{\sqrt{30}} \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix} \right\| = \frac{1}{\sqrt{30}} \left\| \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix} \right\| = \frac{1}{\sqrt{30}} \sqrt{30} = 1, \\ \|\vec{v}_2\| &= \left\| \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\| = \frac{1}{\sqrt{5}} \left\| \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\| = \frac{1}{\sqrt{5}} \sqrt{5} = 1, \\ \|\vec{v}_3\| &= \left\| \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\| = \frac{1}{\sqrt{6}} \left\| \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\| = \frac{1}{\sqrt{6}} \sqrt{6} = 1. \end{aligned}$$

Then we can write

$$\vec{x} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

as a linear combination of the vectors in S. According to Theorem 2.4.6, we may write

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 \text{ where } c_i = \vec{x} \cdot \vec{v}_i,$$

where $1 \le i \le 3$. Since

$$\begin{aligned} c_1 &= \vec{x} \cdot \vec{v}_1 = 4(1/\sqrt{30}) + 5(2/\sqrt{30}) + 6(-5/\sqrt{30}) = \frac{-16}{\sqrt{30}}, \\ c_2 &= \vec{x} \cdot \vec{v}_2 = 4(-2/\sqrt{5}) + 5(1/\sqrt{5}) + 6(0) = \frac{-3}{\sqrt{5}}, \\ c_3 &= \vec{x} \cdot \vec{v}_3 = 4(1/\sqrt{6}) + 5(2/\sqrt{6}) + 6(1/\sqrt{6}) = \frac{20}{\sqrt{6}}. \end{aligned}$$

Thus,

$$\vec{x} = -\frac{16}{\sqrt{30}}\vec{v}_1 - \frac{3}{\sqrt{5}}\vec{v}_2 + \frac{20}{\sqrt{6}}\vec{v}_3.$$

Note that this agrees with what we found in Exploration 55:

$$\begin{aligned}\vec{x} &= -\frac{16}{\sqrt{30}}\frac{1}{\sqrt{30}}\begin{bmatrix}1\\2\\-5\end{bmatrix} - \frac{3}{\sqrt{5}}\frac{1}{\sqrt{5}}\begin{bmatrix}-2\\1\\0\end{bmatrix} + \frac{20}{\sqrt{6}}\frac{1}{\sqrt{6}}\begin{bmatrix}1\\2\\1\end{bmatrix}\\ &= -\frac{8}{15}\begin{bmatrix}1\\2\\-5\end{bmatrix} - \frac{3}{5}\begin{bmatrix}-2\\1\\0\end{bmatrix} + \frac{10}{3}\begin{bmatrix}1\\2\\1\end{bmatrix}.\end{aligned}$$

Exploration 56 Write

$$\vec{y} = \begin{bmatrix}1\\-1\\0\end{bmatrix}$$

as a linear combination of the orthonormal basis vectors from the example above.

Perhaps you're convinced at this point that orthogonal bases are pretty great, and of course, orthonormal bases are even better. However, if you're given a random basis for a vector space, how likely do you think it is that this basis is an orthogonal basis, much less an orthonormal one? Right;[28] it is very unlikely. Thus, it would be good to develop a method for making sets of vectors into orthogonal sets. (Then we could take any basis and make a new orthogonal one!)

28: Pun intended, again.

Orthogonal Projection

Before we take on the task of making a new basis in general, we should start with the very simple case of two vectors. That is, given two distinct, linearly independent vectors $\vec{v}$ and $\vec{u}$ in a vector space V, is there a way to write $\vec{v}$ as the sum of $\vec{u}$ and some other vector $\vec{w}$? Yes!

Lemma 2.4.7 *Let $\vec{v}$ and $\vec{u}$ be vectors in a vector space V, and $\vec{w} = \vec{v} - \vec{u}$. Then*

$$\text{Span}\{\vec{v}, \vec{u}\} = \text{Span}\{\vec{u} + \vec{w}, \vec{u}\} = \text{Span}\{\vec{w}, \vec{u}\}. \tag{2.12}$$

PROOF. Note that $\vec{u}+\vec{w} = \vec{u}+(\vec{v}-\vec{u}) = \vec{v}$. Using the definition of span, we have

$$\begin{aligned}
\text{Span}\{\vec{v},\vec{u}\} &= \{a\vec{v}+b\vec{u}\colon\, a,b\in\mathbb{R}\}\\
&= \{a(\vec{u}+\vec{w})+b\vec{u}\colon\, a,b\in\mathbb{R}\} = \text{Span}\{\vec{u}+\vec{w},\vec{u}\}\\
&= \{a\vec{u}+a\vec{w}+b\vec{u}\colon\, a,b\in\mathbb{R}\}\\
&= \{a\vec{w}+(a+b)\vec{u}\colon\, a,b\in\mathbb{R}\}\\
&= \{a\vec{w}+c\vec{u}\colon\, a,c\in\mathbb{R}\} = \text{Span}\{\vec{w},\vec{u}\}\,.
\end{aligned}$$

□

Exploration 57 Let's be sure you believe this before moving on. Let

$$\vec{v} = \begin{bmatrix} 1\\ 1\\ 0 \end{bmatrix} \text{ and } \vec{u} = \begin{bmatrix} 1\\ 0\\ 2 \end{bmatrix}.$$

- ▶ Find the vector $\vec{w}$ such that $\vec{v} = \vec{u}+\vec{w}$.
- ▶ Let

  $$\vec{x} \in \text{Span}\left\{\begin{bmatrix} 1\\ 1\\ 0 \end{bmatrix}, \begin{bmatrix} 1\\ 0\\ 2 \end{bmatrix}\right\}, \text{ so } \vec{x} = \begin{bmatrix} a+b\\ a\\ 2b \end{bmatrix}$$

  for some $a,b\in\mathbb{R}$. Write $\vec{x}$ as a linear combination of $\vec{w}$ and $\vec{u}$.
  Hint: The coefficients should be a and $(a+b)$.

What does Equation (2.12) do for us? Well, as mentioned, it gives us a way to interchange elements in a spanning set. However, our goal is to ultimately replace a basis with an orthogonal basis, so what we really would like is for $\vec{u}$ and $\vec{w}$ to be orthogonal and satisfy $\vec{v} = \vec{u}+\vec{w}$. However, as we attempt this, we see that $\vec{u}$ needs a little adjusting. Let's instead require that some scalar multiple of $\vec{u}$, say $\alpha\vec{u}$, and some vector $\vec{w}$ be orthogonal such that $\vec{v} = \alpha\vec{u}+\vec{w}$. Just like before, we need $\vec{w} = \vec{v}-\alpha\vec{u}$. If you're having trouble visualizing this, refer to Figure 2.2.

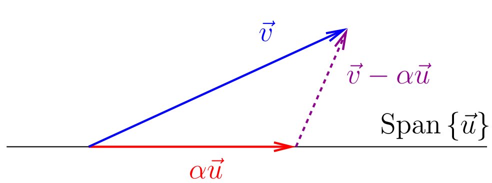


FIGURE 2.2. We need to find a scalar α such that $\vec{v}-\alpha\vec{u}$ is orthogonal to $\vec{u}$.

According to Figure 2.2, we need to "tune" α so that $\vec{v}-\alpha\vec{u}$ is orthogonal to $\vec{u}$, or

$$(\vec{v}-\alpha\vec{u})\cdot\vec{u} = 0.$$

Using a distributive property of our inner product, we see this is true if and only if $\alpha(\vec{u}\cdot\vec{u}) = \vec{v}\cdot\vec{u}$. Solving for α, we have

$$\alpha = \frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}.$$

We've just constructed a vector $\vec{v} - \frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}$ that is orthogonal to $\vec{u}$ and satisfies $\vec{v} = \vec{u} + \vec{w}$. Mission accomplished! Well, the short-term mission at least. This takes care of a basis of size 2, but we'll actually wait until the next section for that whole orthogonal basis part in general because this needs a bit more discussion.

Didn't α look a tad familiar?[29] What we've actually done here in finding $\alpha\vec{u}$ is we have found a vector to represent the part of $\vec{v}$ that is traveling in the direction of $\vec{u}$. This actually has its own name and special notation.

29:  After making you compute it in those explorations, it really should!

Definition 2.4.5 *For any two vectors $\vec{v}$ and $\vec{u}$ in an inner product space, the **orthogonal projection of** $\vec{v}$ **onto** $\vec{u}$ is*

$$\operatorname{proj}_{\vec{u}}(\vec{v}) = \frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}.$$

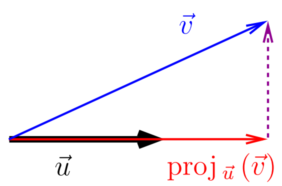


FIGURE 2.3. Here we project $\vec{v}$ onto $\vec{u}$.

Exploration 58 Let

$$\vec{v} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}, \quad \text{and } \vec{u} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

- Let's compute the orthogonal projection of $\vec{v}$ onto $\vec{u}$. First, we see $\vec{v}\cdot\vec{u} = -2-2+3 = -1$ and $u\cdot u = 1+4+1 = 6$. Then we get

$$\operatorname{proj}_{\vec{u}}(\vec{v}) = -\frac{1}{6}\vec{u} = \begin{bmatrix} -\frac{1}{6} \\ \frac{1}{3} \\ -\frac{1}{6} \end{bmatrix}$$

- Now, let's compute the orthogonal projection of $\vec{u}$ onto $\vec{v}$. We already know $\vec{v}\cdot\vec{u} = \vec{u}\cdot\vec{v} = -1$ from above. Now, compute $\vec{v}\cdot\vec{v}$. Combine

  these to find $\operatorname{proj}_{\vec{v}}(\vec{u})$.

Now, why would this be useful? Let's consider a scenario. You have a cart loaded with gold on a train track. The cart is unfortunately not self-propelled, but luckily, you also own a sturdy plow-horse and a very thick rope. The horse cannot walk on the train tracks, so he cannot pull the cart directly forward along the track... Wait. From the picture below, your plow horse is actually a unicorn, and it flies above the track so that the rope makes a $45°$ angle with the train track. Either way, it's the same basic problem, so we'll go with the unicorn.[30] See the diagram below.

30:  Is *that* unicorn drawing why we're here?

 I guess so.

 Yikes.

 Yeah. Maybe we should rethink our standards?

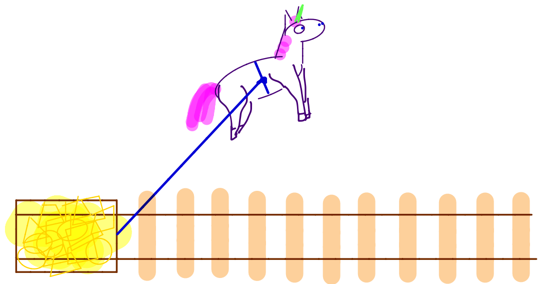


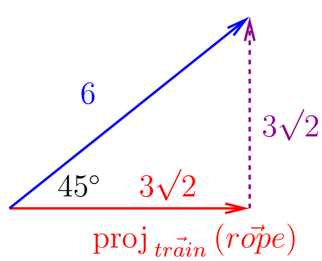
FIGURE 2.4. It's just as you pictured, right?

Suppose the forward force that must be exerted on the cart for it to move is 5 N (where N stands for Newtons, a metric unit for force). If the "horse" is able to pull with a force of 6 N, will the cart move? The way to answer this is by computing the orthogonal projection of the "rope vector" onto the "train track vector." The vector representing the rope should have length 6 to represent the force with which the unicorn can pull. Since we know the angle is $45°$, we see that the "rope vector" is the vector

$$\begin{bmatrix} 3\sqrt{2} \\ 3\sqrt{2} \end{bmatrix}$$

when we are assigning the "train track vector" to be the vector

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$



Because of our diagram, we can actually compute that

$$\text{proj}_{\vec{train}}(\vec{rope}) = \begin{bmatrix} 3\sqrt{2} \\ 0 \end{bmatrix}$$

from the picture, but we also get this using the formula in our definition. Unfortunately, this vector has length $3\sqrt{2}$, which is less than 5. Therefore, our cart will not be moving anywhere unless we either unload some of the gold or we provide additional fairy dust to our flying unicorn. Let this be a lesson. You should never fill your cart before you compute the strength of your horse.[31]

31:  ...or unicorn.

Section Highlights

- For any subspace W of an inner product space, there is an *orthogonal complement* $W^{\perp}$ that is also a subspace. See Theorem 2.4.2.
- Any vector in $W^{\perp}$ will be orthogonal to any vector in W. See Definition 2.4.1 and Theorem 2.4.1.
- Any set of vectors whose vectors are pairwise orthogonal is called an *orthogonal set*. See Definition 2.4.2.
- Any orthogonal set is linearly independent, and therefore, an orthogonal set of n vectors in an n-dimensional vector space will be an orthogonal basis. See Theorem 2.4.5 and Corollary 2.4.4.
- If a basis is orthogonal, then there is a formula that can be used to compute the coordinate vector with respect to that basis. See Theorem 2.4.5 and Exploration 55.
- An *orthogonal basis* can be turned into an orthonormal basis by scaling each vector to be a unit vector. See Definition 2.4.4 and Example 2.4.3.
- The projection of a vector $\vec{u}$ onto a vector $\vec{v}$ is the part of the vector $\vec{u}$ that is in the direction of $\vec{v}$. See Definition 2.4.5.

Exercises for Section 2.4

2.4.1.

Using the standard inner product in $\mathbb{R}^3$, determine all orthogonal subsets of this set of vectors:

$$\left\{ \vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \vec{v}_5 = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}, \vec{v}_6 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}.$$

2.4.2. Let

$$\vec{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ -2 \\ 8 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and } \vec{x} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}.$$

Show that $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an orthogonal basis for $\mathbb{R}^3$ with standard inner product. Then express $\vec{x}$ as a linear combination of $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$.

2.4.3. Let

$$\vec{v}_1 = \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 8 \\ 5 \\ -4 \end{bmatrix}, \quad \text{and } \vec{x} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}.$$

Show that $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an orthogonal basis for $\mathbb{R}^3$ standard inner product. Then express $\vec{x}$ as a linear combination of $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$.

2.4.4. Let $\vec{p}_1 = 2 + x - 2x^2$, $\vec{p}_2 = -1 + 4x + x^2$, $\vec{p}_3 = 1 + x^2$, and $\vec{q} = 8 - 4x - 3x^2$. Show that $\{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$ is an orthogonal basis for $\mathbb{P}_2$ with respect to the standard basis of $\mathbb{P}_2$. Then express $\vec{q}$ as a linear combination of $\vec{p}_1$, $\vec{p}_2$, and $\vec{p}_3$.

2.4.5. Using the standard inner product in $\mathbb{R}^3$, compute the projection from $\vec{v}$ onto $\vec{u}$, $\text{proj}_{\vec{u}}(\vec{v})$, for each pair of vectors below. Note that the answer should always be a scalar multiple of $\vec{u}$.

(a) $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ onto $\vec{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

(b) $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ onto $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

(c) $\vec{v} = \begin{bmatrix} 0 \\ 4 \\ -1 \end{bmatrix}$ onto $\vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

(d) $\vec{v} = \begin{bmatrix} 0 \\ 4 \\ -1 \end{bmatrix}$ onto $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

(e) $\vec{v} = \begin{bmatrix} 3 \\ 4 \\ -1 \end{bmatrix}$ onto $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

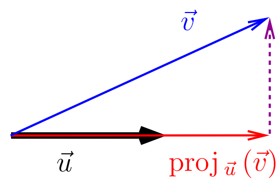
2.4.6. Using the standard basis for $\mathbb{P}_2$, compute the projection from $\vec{p}$ onto $\vec{q}$ below. Note that the answer should always be a scalar multiple of $\vec{q}$.

(a) $\vec{p} = 1 + x + x^2$ onto $\vec{q} = -1 - x$.

(b) $\vec{p} = 1 + 2x - x^2$ onto $\vec{q} = -1 - x + x^2$.

(c) $\vec{p} = x + x^2$ onto $\vec{q} = -1 - x$.

2.4.7. We consider the distance between a vector and a subspace to be the distance from the tip of the vector to the closest point in the subspace. When the subspace is just the span of a single vector, we've seen how to handle this.



As you can see from the diagram, the distance is the length of $\vec{v} - \text{proj}_{\vec{u}}(\vec{v})$, denoted $\|\vec{v} - \text{proj}_{\vec{u}}(\vec{v})\|$. Answer this question for each $\vec{v}$ and $\vec{u}$ below: How far is $\vec{v}$ from $\text{Span}\{\vec{u}\}$?

(a) $\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(b) $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(c) $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(d) $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$

2.4.8. Let $\vec{v} \in \mathbb{R}^n$. Using the standard inner product, show that the orthogonal projection of $\vec{v}$ onto $\vec{v}$ is $\vec{v}$.

2.4.9. Let $W = \text{Span}\left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$, and use the standard inner product in $\mathbb{R}^3$.

(a) Show that $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$ is an orthogonal basis for W. Note, you must confirm that $\mathcal{B} \subset W$.

(b) Find a vector in $W^\perp$ by solving for $a, b, c \in \mathbb{R}$ such that

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 0$$

(c) What must e, f, and g be so that $\mathcal{B}_1 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} e \\ f \\ g \end{bmatrix} \right\}$ is also an orthogonal basis for W? Note that you must choose e, f, and g so that $\mathcal{B}_1 \subset W$.

2.4.10. Let $W = \text{Span}\left\{\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}\right\}$, and use the standard inner product in $\mathbb{R}^3$.

(a) Find conditions to assure a vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is in $W^\perp$. Use these to write

$$W^\perp = \left\{\begin{bmatrix} \rule{2em}{0.4pt} \\ \rule{2em}{0.4pt} \\ \rule{2em}{0.4pt} \end{bmatrix} : a, b, c \in \mathbb{R}\right\}$$

(b) Use the description above to find a basis for $W^\perp$.

(c) Verify $\vec{v} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \in W^\perp$ and write it as a linear combination of your basis vectors.

2.4.11. Let $W = \text{Span}\left\{1 + x^2\right\}$ be a subspace of $\mathbb{P}_2$. Use the method from Exercise 2.4.10 to find $W^\perp$ with respect to the standard basis for $\mathbb{P}_2$.

2.4.12. Suppose $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is an orthogonal set in inner product space V. Show that $\vec{v}_1 + \vec{v}_2$ is orthogonal to $\vec{v}_3 + \vec{v}_4$.

2.4.13. Let W be a subspace of an inner product space V. Show that $W \cap W^\perp = \{\vec{0}\}$.

2.4.14. Let V be an inner product space and $\vec{v}_0 \in V$. Define $T\colon V \to V$ by $T(\vec{v}) = \text{proj}_{\vec{v}_0}(\vec{v})$. Show that for any $\vec{v}_1, \vec{v}_2 \in V$ and any $a \in \mathbb{R}$, we have

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) \quad \text{and} \quad T(a\vec{v}_1) = aT(\vec{v}_1).$$

2.5 The Gram-Schmidt Process

We're well on our way at this point to having a very efficient and geometrically intuitive way to describe vectors in a vector space. For reasons that will be obvious much later, it's often nice to be able to think about vectors relative to a given subspace. We actually just did this in a small way in the previous section![32] When we found the projection of a vector $\vec{v}$ onto a vector $\vec{u}$, we obtained a scalar multiple of our vector $\vec{u}$. Really, we related $\vec{v}$ to the subspace $\mathrm{Span}\,\{\vec{u}\}$.

32: We did?
We did! Pay attention!

Now, we'll want a way to do this in general, when our subspace is more than the span of a single vector. For example, if W is a subspace of a vector space V, then we could think of any vector in V as having two parts: the part in W and the part not in W. We're going to explore the right way to do this, where by "right" we mean orthogonal. You see, it's funny because "right" has two meanings; you may have thought we meant "the correct way," but we really meant it in both senses. Mediocre wordplay is really best when over-explained, don't you think?

Theorem 2.5.1 (The Orthogonal Decomposition Theorem) *Let W be a subspace of an inner product space V such that $\dim W = p < n = \dim V$, and suppose W has an orthogonal basis. Then any $\vec{v} \in V$ can be written uniquely as*

$$\vec{v} = \vec{w} + \vec{u},$$

where $\vec{w} \in W$ and $\vec{u} \in W^{\perp}$.

According to this theorem, we can take *any* vector in V and write it as the sum of a vector in W and some other vector orthogonal to W, and there is only one way to do this. That's actually pretty amazing; we should prove it.

PROOF. Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ be the orthogonal basis for W. The choice for vectors $\vec{w}$ and $\vec{u}$ is key here. Inspired by the fact that we could write the part of $\vec{v}$ in the direction of $\vec{v}_1$ using a projection, let's try doing that for all p vectors in $\mathcal{B}$. Define

$$\begin{aligned} \vec{w} &= \frac{\vec{v}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 + \cdots + \frac{\vec{v}\cdot\vec{v}_p}{\vec{v}_p\cdot\vec{v}_p}\vec{v}_p. \text{ and} \\ \vec{u} &= \vec{v} - \vec{w}; \end{aligned}$$

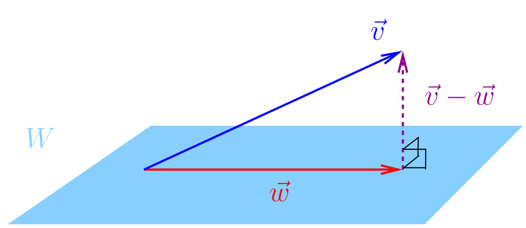


FIGURE 2.5. This is the most important picture in Section 2.5.

these are pictured in Figure 2.5. From these definitions, it is obvious that $\vec{v} = \vec{w} + \vec{u}$ and $\vec{w} \in W$. We only have to show that $\vec{u} \in W^{\perp}$. This is fun to check; note that for any $1 \leq i \leq p$, we have

$$\begin{aligned}
\vec{u} \cdot \vec{v}_i &= (\vec{v} - \vec{w}) \cdot \vec{v}_i \\
&= \vec{v} \cdot \vec{v}_i - \vec{w} \cdot \vec{v}_i \\
&= \vec{v} \cdot \vec{v}_i - \left(\frac{\vec{v} \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 + \cdots + \frac{\vec{v} \cdot \vec{v}_p}{\vec{v}_p \cdot \vec{v}_p} \vec{v}_p \right) \cdot \vec{v}_i \\
&= \vec{v} \cdot \vec{v}_i - \left(\frac{\vec{v} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i} \vec{v}_i \cdot \vec{v}_i \right) \\
&= \vec{v} \cdot \vec{v}_i - \vec{v} \cdot \vec{v}_i \ = \ 0.
\end{aligned}$$

Since $\vec{u}$ is orthogonal to every vector in $\mathcal{B}$, a basis for W, we have from Theorem 2.4.1 that $\vec{u} \in W^\perp$.

The last thing to prove is that the decomposition $\vec{v} = \vec{w} + \vec{u}$ is unique. To do this, we will assume there is some other decomposition and show that it has to be the same as our original decomposition. Assume $\vec{v} = \vec{w}_0 + \vec{u}_0$ for some vectors $\vec{w}_0 \in W$ and $\vec{u}_0 \in W^\perp$. Then we have $\vec{w} + \vec{u} = \vec{w}_0 + \vec{u}_0$ (since both sides are equal to v), which implies

$$\vec{w} - \vec{w}_0 = \vec{u}_0 - \vec{u}.$$

Since W is a subspace of V and $\vec{w}, \vec{w}_0 \in W$, we have $\vec{w} - \vec{w}_0 \in W$. Moreover, since $W^\perp$ is a subspace of V and $\vec{u}, \vec{u}_0 \in W^\perp$, we have $\vec{w} - \vec{w}_0 = \vec{u} - \vec{u}_0 \in W^\perp$. Since $\vec{w} - \vec{w}_0 \in W$ and $\vec{w} - \vec{w}_0 \in W^\perp$, we have

$$\|\vec{w} - \vec{w}_0\|^2 = (\vec{w} - \vec{w}_0) \cdot (\vec{w} - \vec{w}_0) = 0,$$

and this can only happen if $\vec{w} - \vec{w}_0 = 0$. Thus, $\vec{w} = \vec{w}_0$. A similar argument shows that $\vec{u} = \vec{u}_0$. Behold! The decomposition is unique. □

Corollary 2.5.2 *Let W be a subspace of a vector space V. Then $V = W \oplus W^\perp$.*

PROOF. By the Orthogonal Decomposition Theorem, every vector in V is the sum of a vector in W and a vector in $W^\perp$. Thus, $V = W + W^\perp$ by the definition of the sum of two subspaces. Also, suppose $\vec{v} \in W \cap W^\perp$. Then, $\vec{v}$ is orthogonal to itself, which means $\vec{v} \cdot \vec{v} = 0$. This only happens when $\vec{v} = \vec{0}$. Thus, the intersection is only the zero vector, and we have $V = W \oplus W^\perp$. □

In Figure 2.6, we introduce a convenient way to picture V decomposed as $W \oplus W^\perp$. Note that $W \cup W^\perp$ is not equal to V; there are vectors in V that are in neither W nor $W^\perp$; those vectors are represented by the white part of the diagram. Every vector in $V = W \oplus W^\perp$ can be written as a sum of a vector in W and a vector in $W^\perp$, though, so we've represented the whole vector space V as W and $W^\perp$, bridged by the direct sum symbol.

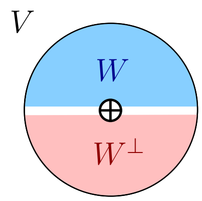


FIGURE 2.6. Here is a convenient way to picture the vector space, V, decomposed as $W \oplus W^{\perp}$ and its subspaces, W and $W^{\perp}$.

It turns out that for direct sums, like the one in Corollary 2.5.2, finding a basis for the direct sum is actually pretty easy, too.

Theorem 2.5.3 *Suppose V is an inner product space with subspace U and W such that $V = U \oplus W$. Let $\mathcal{B}_U$ be a basis for U and $\mathcal{B}_W$ be a basis for W. Then $\mathcal{B} = \mathcal{B}_U \cup \mathcal{B}_W$ is a basis for V.*

PROOF. Note that if $\mathcal{B}_U = \{\vec{u}_1, \ldots, \vec{u}_k\}$ and $\mathcal{B}_W = \{\vec{w}_1, \ldots, \vec{w}_\ell\}$, then $\mathcal{B} = \{\vec{u}_1, \ldots, \vec{u}_k, \vec{w}_1, \ldots, \vec{w}_\ell\}$. We know first of all that $\mathcal{B}$ spans V. Any vector can be written as a sum of a vector in U and a vector in W, and these vectors in turn can be written as linear combinations of vectors in $\mathcal{B}_U$ and $\mathcal{B}_W$, respectively. We also know that these vectors must be linearly independent separately. Thus, we need to consider the case where a vector in $\mathcal{B}_U$ is a linear combination of the vectors in $\mathcal{B}_W$, or vice versa. Really, the argument will be the same either way, so we'll focus on the first case. Suppose $\vec{v} \in \mathcal{B}_U$ is a linear combination of the vectors in $\mathcal{B}_W$. Then $\vec{v} \in \text{Span}\{\mathcal{B}_W\} = W$. This would mean $\vec{v} \in U \cap W$; since $U \cap W = \{\vec{0}\}$, we must have $\vec{v} = \vec{0} \in \mathcal{B}_U$, but this is not possible because a basis cannot contain the zero vector. So, $\mathcal{B}$ must be linearly independent as well as span V. So, $\mathcal{B}$ is a basis for V. □

Perhaps you will recall we made a conjecture in the last section.[33] Now we can prove it!

33: …ummmm…
Ricky! Go back and look!

Corollary 2.5.4 *Let W be a subspace of an inner product space V such that $\dim V = n$ and $\dim W = p$. Then $\dim W^{\perp} = n - p$.*

PROOF. Let $\mathcal{B}_1$ be a basis for W and $\mathcal{B}_2$ be a basis for $W^{\perp}$. Now consider $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$. From Theorem 2.5.3, we know that $\mathcal{B}$ is a basis for V since $V = W \oplus W^{\perp}$. We can conclude the statement since the dimensions are computed using the sizes of the bases. □

Corollary 2.5.5 *Let W be a subspace of an inner product space V. Suppose $\mathcal{B}_1$ is an orthogonal basis for W and $\mathcal{B}_2$ is an orthogonal basis for $W^{\perp}$. Then $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is an orthogonal basis for V.*

Another fun byproduct of Theorem 2.5.1 is that we can project any vector onto a subspace,[34] and this projection is unique (for the given subspace). Thus, we have the following definition.

34: You need an *orthogonal* basis to use the formula, though.
Right.
It's not funny anymore, Ricky.

Definition 2.5.1 *Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ be an orthogonal basis for a subspace W of an inner product space V. For any vector $\vec{v} \in V$, the* ***orthogonal projection of $\vec{v}$ onto W*** *is*

$$\operatorname{proj}_W(\vec{v}) = \frac{\vec{v}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 + \cdots + \frac{\vec{v}\cdot\vec{v}_p}{\vec{v}_p\cdot\vec{v}_p}\vec{v}_p.$$

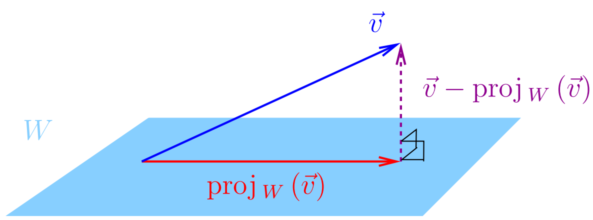


FIGURE 2.7. This is the same picture as in Figure 2.5, but now we can stop pretending that $\vec{w}$ wasn't the projection onto W all along.

Example 2.5.1 Let's see this Orthogonal Decomposition in action and use the standard inner product in $\mathbb{R}^3$. Define

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \text{and} \quad \vec{v}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

Let $W = \operatorname{Span}\{\vec{v}_1, \vec{v}_2\}$. Note that $\vec{v}_1 \cdot \vec{v}_2 = 0$, so $\{\vec{v}_1, \vec{v}_2\}$ is an orthogonal basis for W. Let's find the $\operatorname{proj}_W(\vec{v}_3)$ and use this to find the orthogonal decomposition of $\vec{v}_3$. First, we compute the coefficients.

$$\frac{\vec{v}_3\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1} = \frac{1}{3} \quad \text{and} \quad \frac{\vec{v}_3\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2} = \frac{1}{2}$$

Now, we see

$$\operatorname{proj}_W(\vec{v}_3) = \frac{1}{3}\vec{v}_1 + \frac{1}{2}\vec{v}_2.$$

This will be our $\vec{w}$ in our orthogonal decomposition. Let $\vec{u} = \vec{v}_3 - \vec{w}$. Then we have

$$\vec{w} = \frac{1}{3}\vec{v}_1 + \frac{1}{2}\vec{v}_2 = \frac{1}{3}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ -\frac{1}{3} \\ -\frac{1}{6} \end{bmatrix}$$

$$\vec{u} = \vec{v}_3 - \vec{w} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{5}{6} \\ -\frac{1}{3} \\ -\frac{1}{6} \end{bmatrix} = \begin{bmatrix} \frac{7}{6} \\ \frac{7}{3} \\ \frac{7}{6} \end{bmatrix}$$

If we've done this all correctly, we should have $\vec{u} \in W^\perp$. Let's check!

$$\vec{u} \cdot \vec{v}_1 = \begin{bmatrix} \frac{7}{6} \\ \frac{7}{3} \\ \frac{7}{6} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \frac{7}{6} - \frac{7}{3} + \frac{7}{6} = 0$$

$$\vec{u} \cdot \vec{v}_2 = \begin{bmatrix} \frac{7}{6} \\ \frac{7}{3} \\ \frac{7}{6} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{7}{6} - \frac{7}{6} = 0$$

Since $\vec{u}$ is orthogonal to the basis for W, it is orthogonal to all of W. So it's in $W^\perp$!

Example 2.5.2 Let $W = \text{Span}\{\vec{v}_1, \vec{v}_2\}$, where

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Let's find a basis for $W^\perp$. First, we should note that $\vec{v}_1$ and $\vec{v}_2$ are orthogonal. So $\dim W = 2$ and $\{\vec{v}_1, \vec{v}_2\}$ forms a basis of W. According to Corollary 2.5.4, we now know $\dim W^\perp = 1$, so we only need to find one vector. It has to be orthogonal to both $\vec{v}_1$ and $\vec{v}_2$. Thus, if the vector we seek is of the form

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

then $\vec{v}_2 \cdot \vec{x} = \vec{v}_1 \cdot \vec{x} = 0$ yields

$$x_1 - x_3 = 0 \text{ and } x_1 + x_2 + x_3 = 0.$$

From the first equation, we see that our vector $\vec{x}$ must have $x_1 = x_3$. Applying this to the second equation, we have $x_2 = -2x_1$. There are a lot of vectors that satisfy these criteria; here's one:

$$\vec{x} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

One can check that $\vec{v}_2 \cdot \vec{x} = \vec{v}_1 \cdot \vec{x} = 0$. Thus, $\vec{x} \in W^\perp$ and will serve nicely as a basis for $W^\perp$. Thus, a basis for $\mathbb{R}^3 = W \oplus W^\perp$ is given by $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \vec{x}\}$.

Exploration 59 Now that you've seen some examples, let's try an exploration. Define

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 6 \\ 6 \\ 0 \\ 0 \end{bmatrix}, \quad \text{and} \quad \vec{y} = \begin{bmatrix} 0 \\ -2 \\ 8 \\ 2 \end{bmatrix}.$$

Let $W = \text{Span}\{\vec{v}_1, \vec{v}_2\}$. Note that $\vec{v}_1 \cdot \vec{v}_2 = 0$, so this is an orthogonal basis of W. We would like to find a basis now for $W^\perp$. We'll use our Orthogonal Decomposition Theorem to do this!

- ▶ Compute $\text{proj}_W(\vec{x})$. Use this to find a vector $\vec{u}_1 \in W^\perp$.

► Compute $\operatorname{proj}_W(\vec{y})$. Use this to find a vector $\vec{u}_2 \in W^{\perp}$.

You should now have two linearly independent vectors in $W^{\perp}$, thus a basis for $W^{\perp}$. (Caution: This may not always work. You are not guaranteed that the projections of distinct vectors will be linearly independent, even if those vectors are linearly independent or orthogonal. However, in the event this fails, it could be attempted again with different vectors.)

So, we found a basis of $W^{\perp}$. Wouldn't it be better to find an *orthogonal* basis for $W^{\perp}$?

Orthogonal Basis Through the Gram-Schmidt Process

This orthogonality business all seems pretty good, but you may have noticed at this point that we've not provided a way to actually generate an orthogonal basis.[35] The good news is there is a procedure to make an orthogonal basis from any given basis, and it should seem fairly intuitive after the two vector example of the previous section. The bad news is that it's not a lot of fun to actually do.

35: We talked about this in the last section before we got sidetracked with the orthogonal projection and the unicorn example.

Theorem 2.5.6 (The Gram-Schmidt Process) *Let W be a subspace of vector space V with basis $\{\vec{v}_1, \ldots, \vec{v}_p\}$. Define*

$$\begin{aligned}
\vec{w}_1 &= \vec{v}_1, \\
\vec{w}_2 &= \vec{v}_2 - \operatorname{proj}_{\operatorname{Span}\{\vec{w}_1\}}(\vec{v}_2) \\
&= \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\vec{w}_1, \\
\vec{w}_3 &= \vec{v}_3 - \operatorname{proj}_{\operatorname{Span}\{\vec{w}_1, \vec{w}_2\}}(\vec{v}_3) \\
&= \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\vec{w}_1 - \frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\vec{w}_2, \\
&\vdots \\
\vec{w}_p &= \vec{v}_p - \operatorname{proj}_{\operatorname{Span}\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_{p-1}\}}(\vec{v}_p) \\
&= \vec{v}_p - \frac{\vec{v}_p \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\vec{w}_1 - \frac{\vec{v}_p \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\vec{w}_2 - \cdots - \frac{\vec{v}_p \cdot \vec{w}_{p-1}}{\vec{w}_{p-1} \cdot \vec{w}_{p-1}}\vec{w}_{p-1}.
\end{aligned}$$

Then $\{\vec{w}_1, \ldots, \vec{w}_p\}$ is an orthogonal set.

There isn't much to prove here. We get orthogonality from repeated application of the Orthogonal Decomposition Theorem. We get the fact that $\operatorname{Span}\{\vec{v}_1, \ldots, \vec{v}_p\} = \operatorname{Span}\{\vec{w}_1, \ldots, \vec{w}_p\}$ by a similar repeated application of what we did in Section 2.4. Just replace each vector one at a time so that the span is unchanged.

Now we can take any set of linearly independent vectors and make it orthogonal, but you should probably be warned at this point. The coefficients from the projections used in the Gram-Schmidt process are usually pretty terrible. The process actually has you feed those terrible coefficients back into a projection formula, too. Perhaps this is better experienced than described...

Example 2.5.3 Let $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, where

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \quad \text{and } \vec{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix}.$$

Find an orthogonal set $\mathcal{B}$ such that $\text{Span}\{S\} = \text{Span}\{\mathcal{B}\}$. According to Theorem 2.5.6, we're going to build an orthogonal set of vectors $\mathcal{B} = \{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ from S, and we can start by just letting $\vec{w}_1 = \vec{v}_1$. Then we use a projection to find $\vec{w}_2$:

$$\begin{aligned} \vec{w}_2 &= \vec{v}_2 - \text{proj}_{\text{Span}\{\vec{w}_1\}}(\vec{v}_2) \\ &= \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\vec{w}_1 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \frac{4+10+18}{1+4+9}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 12/7 \\ 3/7 \\ -6/7 \end{bmatrix}. \end{aligned}$$

Now that we have $\vec{w}_1$ and $\vec{w}_2$, we can use these with $\vec{v}_3$ to find $\vec{w}_3$:

$$\begin{aligned} \vec{w}_3 &= \vec{v}_3 - \text{proj}_{\text{Span}\{\vec{w}_1, \vec{w}_2\}}(\vec{v}_3) = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\vec{w}_1 - \frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\vec{w}_2 \\ &= \begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix} - \frac{7+16+0}{1+4+9}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{12 + \frac{24}{7} + 0}{\frac{144}{49} + \frac{9}{49} + \frac{36}{49}}\begin{bmatrix} 12/7 \\ 3/7 \\ -6/7 \end{bmatrix} \\ &= \begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix} - \frac{23}{14}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - 4\begin{bmatrix} 12/7 \\ 3/7 \\ -6/7 \end{bmatrix} \\ &= \begin{bmatrix} -3/2 \\ 3 \\ -3/2 \end{bmatrix}. \end{aligned}$$

Thus, $\mathcal{B} = \{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$, where

$$\vec{w}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{w}_2 = \begin{bmatrix} 12/7 \\ 3/7 \\ -6/7 \end{bmatrix}, \quad \text{and } \vec{w}_3 = \begin{bmatrix} -3/2 \\ 3 \\ -3/2 \end{bmatrix}.$$

You can check this computation by verifying that $\mathcal{B}$ is an orthogonal set and that $\text{Span}\{S\} = \text{Span}\{\mathcal{B}\}$. Also, let's note something here. Can you see a way to get a different orthogonal basis $\mathcal{B}_0$ using S and this procedure? What if the vector you started with was $\vec{v}_2$? or $\vec{v}_3$? Yes, this algorithm could be used to find multiple orthogonal bases for the same subspace.

Exploration 60 In Exploration 59, we found a basis for $W^{\perp}$, or really you found a basis. Now, make this an orthogonal basis using the Gram-Schmidt Process.

Section Highlights

- If W is any subspace of an inner product space V, then there is an orthogonal decomposition of V into $W \oplus W^{\perp}$. See Theorem 2.5.1.
- Any basis can be turned into an orthogonal basis using the Gram-Schmidt Process. See Theorem 2.5.6.
- The Gram-Schmidt Process is a recursively-defined procedure where at each step the basis vector $\vec{b}_i$ is replaced with $\vec{b}_i - \mathrm{proj}_W\left(\vec{b}_i\right)$, where W is the span of the previously constructed orthogonal vectors. See Theorem 2.5.6 and Example 2.5.3.

Exercises for Section 2.5

For fun, we will introduce a new verb for the exercise section by verbing* a noun. Henceforth, the noun "Gram-Schmidt" will also be the verb "to Gram-Schmidt" with the meaning "to perform the Gram-Schmidt process on."

2.5.1.

Below are several sets of linearly independent vectors. Use the standard inner product and Gram-Schmidt each one to produce an orthogonal set.

(a) $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix} \right\}$

(b) $\left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$

(c) $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \right\}$

(d) $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$

(e) $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

(f) $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$

(g) $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} \right\}$

(h) $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

(i) $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$

(j) $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

(k) $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}$

(l) $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}$

2.5.2. Let $\vec{p}_1 = 2+x$, $\vec{p}_2 = 2x+x^2$, and $W = \text{Span}\,\{\vec{p}_1, \vec{p}_2\}$. Use the standard basis for $\mathbb{P}_2$ and the Gram-Schmidt Process to find an orthogonal basis for W.

2.5.3. Here is a linearly dependent set of vectors:

$$\left\{ \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix} \right\}$$

Use the standard inner product to Gram-Schmidt this set. Your result cannot be a basis for $\mathbb{R}^3$ since the original vectors did not form a basis. How is this reflected in your answer?

---

*Yes, we have used the noun "verb" to implicitly introduce the act of turning a noun into a verb. This is not upsetting.

2.5.4. Suppose $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an orthogonal set of vectors using the standard inner product. Write out and simplify the formulas from the Gram-Schmidt process. What can you conclude about the outcome of the Gram-Schmidt process when the initial set is orthogonal?

2.5.5. Consider the subspace

$$W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} \right\}.$$

(a) Use the standard inner product to construct an orthogonal basis for W.

(b) Note that $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is not in W and compute $\vec{y} = \vec{x} - \text{proj}_W(\vec{x})$.

2.5.6. Let

$$\vec{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ -2 \\ 8 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and } \vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

As you might recall from the previous section, $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an orthogonal basis for $\mathbb{R}^3$ using the standard inner product. Feel free to also verify this.

(a) Write the vector $\vec{x}$ as the sum of two vectors, one in $\text{Span}\{\vec{v}_1, \vec{v}_2\}$ and one in $\text{Span}\{\vec{v}_3\}$.

(b) Write the vector $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ as the sum of two vectors, one in $\text{Span}\{\vec{v}_1\}$ and one in $\text{Span}\{\vec{v}_2, \vec{v}_3\}$.

(c) Normalize the vectors in the set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ to get an orthonormal basis for $\mathbb{R}^3$.

2.5.7. Consider the subspace

$$H = \left\{ \begin{bmatrix} a+b \\ a-b \\ b \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

(a) Construct an orthogonal basis for H using the standard inner product.

(b) Find a vector $\vec{x}$ that is not in H and compute $\vec{y} = \vec{x} - \text{proj}_H(\vec{x})$.

(c) Verify $\vec{y} \in H^{\perp}$.

2.5.8. Let W be a subspace of $\mathbb{R}^n$ with any inner product. Show that $\vec{v} \in W$ if and only if $\text{proj}_W(\vec{v}) = \vec{v}$.

2.5.9. Let W be a subspace of $\mathbb{R}^n$ with any inner product. Show that $\vec{v} \in W^{\perp}$ if and only if $\text{proj}_W(\vec{v}) = \vec{0}$.

2.6 Least Squares Applications

We've spent quite a lot of time learning to describe vector spaces in terms of bases. Then we spent quite a lot more time trying to make bases that are nice for various reasons. You probably won't be surprised to find this was all for a grand purpose. There are, in fact, many good uses for orthogonal bases, but we're going to introduce just one now. Of course by one, we mean two.

Theorem 2.6.1 *Let W be a subspace of an inner product space V, and let $\vec{v} \in V$. Then $\operatorname{proj}_W(\vec{v})$ is the closest vector in W to $\vec{v}$ in the sense that for any $\vec{w} \in W$,*

$$\|\vec{v} - \operatorname{proj}_W(\vec{v})\| \leq \|\vec{v} - \vec{w}\|. \tag{2.13}$$

Proof. Let $\vec{v} \in V$ and $\vec{w} \in W$. Then

$$\begin{aligned} \vec{v} - \vec{w} &= \vec{v} + (-\operatorname{proj}_W(v) + \operatorname{proj}_W(v)) - \vec{w} \\ &= (\vec{v} - \operatorname{proj}_W(v)) + (\operatorname{proj}_W(v) - \vec{w}) \end{aligned}$$

Thus, by the Pythagorean Theorem,

$$\begin{aligned} \|\vec{v} - \vec{w}\|^2 &= \|(\vec{v} - \operatorname{proj}_W(v)) + (\operatorname{proj}_W(v) - \vec{w})\|^2 \\ &= \|\vec{v} - \operatorname{proj}_W(v)\|^2 + \|\operatorname{proj}_W(v) - \vec{w}\|^2 \end{aligned}$$

Note now that $\vec{v} - \operatorname{proj}_W(v)$ is orthogonal to $\operatorname{proj}_W(v) - \vec{w}$ because $\operatorname{proj}_W(\vec{v}) - \vec{w} \in W$ (since W is closed under vector addition) and $\vec{v} - \operatorname{proj}_W(\vec{v})$ is orthogonal to everything in W. See Figure 2.8 for the geometric picture of this. We can rearrange now to get

$$\|\vec{v} - \vec{w}\|^2 - \|\operatorname{proj}_W(v) - \vec{w}\|^2 = \|\vec{v} - \operatorname{proj}_W(v)\|^2.$$

Since we know $\|\operatorname{proj}_W(v) - \vec{w}\|^2 \geq 0$, we have

$$\begin{aligned} \|\vec{v} - \vec{w}\|^2 &\geq \|\vec{v} - \vec{w}\|^2 - \|\operatorname{proj}_W(v) - \vec{w}\|^2 \\ &= \|\vec{v} - \operatorname{proj}_W(v)\|^2 \end{aligned}$$

Since both $\|\vec{v} - \vec{w}\|$ and $\|\vec{v} - \operatorname{proj}_W(v)\|$ are positive, this gives the desired inequality. □

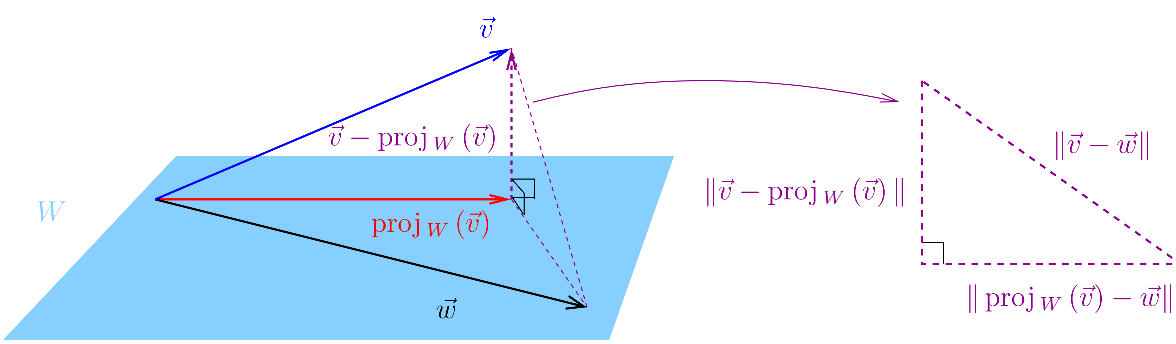


Figure 2.8: This is the most important picture in Section 2.6.

Example 2.6.1 Let

$$\vec{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \quad \vec{v}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix},$$

and $W = \text{Span}\{\vec{v}_1, \vec{v}_2\}$. How far is $\vec{y}$ from W? Let's see! According to Theorem 2.6.1, $\text{proj}_W(\vec{y})$ is the closest vector in W to $\vec{y}$, so we'll start by calculating $\text{proj}_W(\vec{y})$.

$$\begin{aligned} \text{proj}_W(\vec{y}) &= \frac{\vec{y} \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 + \frac{\vec{y} \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 \\ &= \frac{15}{30}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} + \frac{-21}{6}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}. \end{aligned}$$

Then

$$\|\vec{y} - \text{proj}_W(\vec{y})\| = \left\| \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix} \right\| = 3\sqrt{5}.$$

Exploration 61 Define

$$\vec{y} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix},$$

Let $W = \text{Span}\{\vec{v}_1, \vec{v}_2\}$. Find $\text{proj}_W(\vec{y})$.

Did you get $\text{proj}_W(\vec{y}) = \vec{y}$? Why would that happen? What does it mean that the "closest vector" to $\vec{y}$ in W is $\vec{y}$?

The idea of "the closest vector" to some subspace feels like something that could be useful in general scientific settings.[36] This concept allows us to specifically find a "line of best fit" for a set of data points in $\mathbb{R}^2$. Suppose we have an independent variable x, a dependent variable y, and have four observed data points (x, y),

36: Quite useful indeed!

$$(-2, -2), (-1, 0), (1, 2), \text{ and } (2, 1),$$

as seen in Figure 2.9.

If all of these data points were on the same line, then we would be able to find scalars m and b such that each data point satisfied the equation $y = mx + b$. As seen in Figure 2.9, though, when graphed in $\mathbb{R}^2$, these points do not lie on the same line together, so we shouldn't expect to find one m and one b to make this equation true for all four data points. This is annoying; let's try it with vectors. We could interpret the set of x coordinates for these data points as a vector $\vec{x} \in \mathbb{R}^4$ and do likewise for the y coordinates:

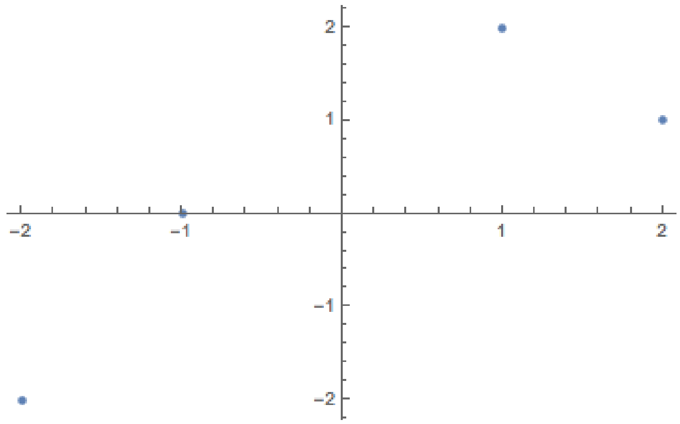


FIGURE 2.9. Here's some data.

$$\begin{matrix}(-2,-2)\\(-1,0)\\(1,2)\\(2,1)\end{matrix} \qquad \vec{x}=\begin{bmatrix}-2\\-1\\1\\2\end{bmatrix} \qquad \vec{y}=\begin{bmatrix}-2\\0\\2\\1\end{bmatrix}$$

The four data points satisfy the equation $y = mx + b$ if and only if

$$\begin{aligned}-2 &= -2m + b\\ 0 &= -1m + b\\ 2 &= 1m + b \quad \text{and}\\ 1 &= 2m + b.\end{aligned}$$

This translates into the vector equation

$$\begin{bmatrix}-2\\0\\2\\1\end{bmatrix} = m\begin{bmatrix}-2\\-1\\1\\2\end{bmatrix} + b\begin{bmatrix}1\\1\\1\\1\end{bmatrix}.$$

In other words, the four data points satisfy the equation $y = mx + b$ if and only if

$$\vec{y} \in \text{Span}\left\{\vec{x}, \vec{1}\right\}, \text{ where } \vec{1} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}.$$

However, we know for a fact that $\vec{y} \notin \text{Span}\left\{\vec{x}, \vec{1}\right\}$.[37] For simplicity of notation, let $W = \text{Span}\left\{\vec{x}, \vec{1}\right\}$. It would be very convenient right now to have the vector in W that was closest to $\vec{y}$. According to Theorem 2.6.1, $\text{proj}_W(\vec{y})$ is the vector in W that was closest to $\vec{y}$. Let us calculate $\text{proj}_W(\vec{y})$!

37: Prove it!

Note first that $\vec{x} \cdot \vec{1} = 0$, so $\{\vec{x}, \vec{1}\}$ is an orthogonal set. If you think this is one of those scams where the example in the book contains a mathematical miracle that makes everything simple, you're half right. While it does make things simpler for us, keep in mind that x was our independent variable, so we could've chosen whatever values we wanted for x's. As long as you pick

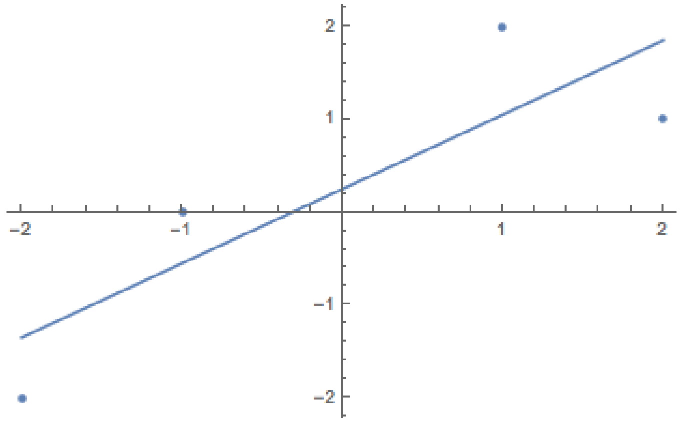


FIGURE 2.10. Here's some data with the line of best fit.

x's in pairs symmetric about 0, this procedure will yield an orthogonal set.[38] Thus, we may use Theorem 2.4.5 to find

$$\text{proj}_W(\vec{y}) = \frac{\vec{x}\cdot\vec{y}}{\vec{x}\cdot\vec{x}}\vec{x} + \frac{\vec{1}\cdot\vec{y}}{\vec{1}\cdot\vec{1}}\vec{1} = \frac{8}{10}\vec{x} + \frac{1}{4}\vec{1} = \frac{8}{10}\begin{bmatrix}-2\\-1\\1\\2\end{bmatrix} + \frac{1}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}.$$

38: Prove that, too!

While we couldn't find an m and b so that $\vec{y} = m\vec{x} + b\vec{1}$, it seems like we just did for $\text{proj}_W(\vec{y}) = m\vec{x} + b\vec{1}$. Using $m = 8/10$ and $b = 1/4$, we have an equation for a line that best approximates our four data points:

$$y = \frac{8}{10}x + \frac{1}{4}.$$

Figure 2.10 shows this line with the four data points. We call this line a **least squares approximation** for the given data.

When working with data, it's far more common that it was collected by someone else. What are we to do if we aren't fortunate enough to have an orthogonal set?

Example 2.6.2 Find the line of best fit for the following data:

$$(-2,-1), (-1,0), (0,2), \text{ and } (1,4).$$

One can quickly check that these points are not colinear, so there are no scalars m and b such that

$$\vec{y} = m\vec{x} + b\vec{1}, \quad \text{where} \quad \vec{y} = \begin{bmatrix}-1\\0\\2\\4\end{bmatrix}, \quad \vec{x} = \begin{bmatrix}-2\\-1\\0\\1\end{bmatrix}, \quad \vec{1} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}.$$

Thus, we should find the closest possible vector to $\vec{y}$ in $W = \text{Span}\left\{\vec{x}, \vec{1}\right\}$, which is $\text{proj}_W(\vec{y})$. Unfortunately, you probably noticed our basis $\{\vec{x}, \vec{1}\}$ for W is not orthogonal, so we cannot find $\text{proj}_W(\vec{y})$ as we did before. Fortunately, we have a way of making an orthogonal basis $\{\vec{w}_1, \vec{w}_2\}$ for W from $\{\vec{x}, \vec{1}\}$ by using the Gram-Schmidt process. By Theorem 2.5.6, we

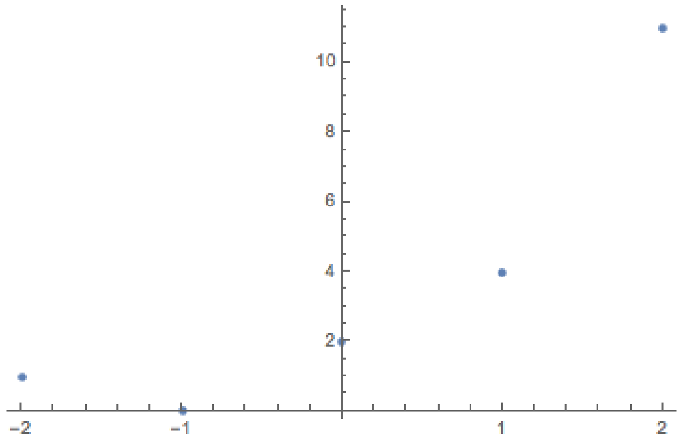


FIGURE 2.11. One last aggressively nonlinear data set.

have

$$\begin{aligned}\vec{w}_1 &= \vec{x}\\ \vec{w}_2 &= \vec{1} - \frac{\vec{1}\cdot\vec{w}_1}{\vec{w}_1\cdot\vec{w}_1}\vec{w}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{1}{3}\begin{bmatrix}-2\\-1\\0\\1\end{bmatrix} = \begin{bmatrix}1/3\\2/3\\1\\4/3\end{bmatrix}.\end{aligned}$$

For good measure, one should check that $\vec{w}_1 \cdot \vec{w}_2 = 0$ so we have an orthogonal basis. We do? Excellent! Then

$$\operatorname{proj}_W(\vec{y}) = \frac{\vec{w}_1\cdot\vec{y}}{\vec{w}_1\cdot\vec{w}_1}\vec{w}_1 + \frac{\vec{w}_2\cdot\vec{y}}{\vec{w}_2\cdot\vec{w}_2}\vec{w}_2 \ = \ \vec{w}_1 + \frac{21}{10}\vec{w}_2$$

However, to get the line of best fit, we need $\operatorname{proj}_W(\vec{y})$ in terms of $\vec{x}$ and $\vec{1}$. Since $\vec{w}_1 = \vec{x}$ and $\vec{w}_2 = \vec{1} + \frac{1}{3}\vec{x}$, we have

$$\operatorname{proj}_W(\vec{y}) = \vec{w}_1 + \frac{21}{10}\vec{w}_2 = \vec{x} + \frac{21}{10}\left(\vec{1} + \frac{1}{3}\vec{x}\right) = \frac{17}{10}\vec{x} + \frac{21}{10}\vec{1}.$$

Thus, using $m = 17/10$ and $b = 21/10$, the line $y = \frac{17}{10}x + \frac{21}{10}$ is the line of best fit, or least squares approximation, for the data.

Let's do one more, but this time with a twist (or turn, depending on your point of view).

Example 2.6.3 Find the line of best fit for the following data:

$$(-2, 1), (-1, 0), (0, 2), (1, 4) \text{ and } (2, 11).$$

One can quickly check that these points are not colinear. The first four data points are a single minus sign away from the four points in the last example. However, our new fifth point makes this set of point even farther from being on a line; see Figure 2.11.

In fact, these points look a lot more like they lie on a parabola than a line; shall we try to find the quadratic curve of best fit? Yes. Yes, we shall. We need to find scalars a, b, and c such that $y = ax^2 + bx + c$ for all five of our points. You can tell from looking at Figure 2.11 that these points don't *actually* all lie on the same parabola. Can you verify this algebraically?

Thus, we know that no such a, b, and c exist; that is there is no set of scalars a, b, and c such that $\vec{y} = a\vec{q} + b\vec{x} + c\vec{1}$, where

$$\vec{y} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 4 \\ 11 \end{bmatrix}, \vec{q} = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix}, \vec{x} = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \vec{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Wait. Where did that vector $\vec{q}$ come from? We got the vector $\vec{x}$ from the first coordinates of our data points because we had an x in the equation $y = ax^2 + bx + c$. Note also that we have an x^2 term. Thus, we need a second vector comprised of the square of the first coordinates of our data points.

Again, we should find the closest possible vector to $\vec{y}$ that is in $W = \text{Span}\left\{\vec{q}, \vec{x}, \vec{1}\right\}$, which is $\text{proj}_W(\vec{y})$. Fortunately for us, $\{\vec{q}, \vec{x}, \vec{1}\}$ is *almost* an orthogonal set; note that $\vec{x} \cdot \vec{1} = 0$ and $\vec{x} \cdot \vec{q} = 0$. Thus, to find a new orthogonal basis for W, $\{\vec{w}, \vec{x}, \vec{1}\}$, we only need to replace $\vec{q}$ with a new vector $\vec{w}$. Then

$$\vec{w} = \vec{q} - \frac{\vec{x} \cdot \vec{q}}{\vec{x} \cdot \vec{x}}\vec{x} - \frac{\vec{1} \cdot \vec{q}}{\vec{1} \cdot \vec{1}}\vec{1} = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}.$$

Then we have

$$\text{proj}_W(\vec{y}) = \frac{\vec{w} \cdot \vec{y}}{\vec{w} \cdot \vec{w}}\vec{w} + \frac{\vec{x} \cdot \vec{y}}{\vec{x} \cdot \vec{x}}\vec{x} + \frac{\vec{1} \cdot \vec{y}}{\vec{1} \cdot \vec{1}}\vec{1} = \frac{8}{7}\vec{w} + \frac{12}{5}\vec{x} + \frac{18}{5}\vec{1}.$$

It's a good thing we didn't try to guess these coefficients. This isn't quite what we need, though; $\vec{w}$ is the wrong vector. For this to be the parabola of best fit, we need $\vec{q}$, the vector of squared x's. Fortunately, we know that $\vec{w} = \vec{q} - (2)\vec{1}$. Thus,

$$\begin{aligned} \text{proj}_W(\vec{y}) &= \frac{8}{7}\vec{w} + \frac{12}{5}\vec{x} + \frac{18}{5}\vec{1} \\ &= \frac{8}{7}(\vec{q} - (2)\vec{1}) + \frac{12}{5}\vec{x} + \frac{18}{5}\vec{1} \\ &= \frac{8}{7}\vec{q} + \frac{12}{5}\vec{x} + \left(\frac{18}{5} - \frac{16}{7}\right)\vec{1} = \frac{8}{7}\vec{q} + \frac{12}{5}\vec{x} + \frac{46}{35}\vec{1} \end{aligned}$$

The quadratic equation $y = \frac{8}{7}x^2 + \frac{12}{5}x + \frac{46}{35}$ is the best quadratic least squares approximation for the given data. See Figure 2.12.

It seems like we could generalize this procedure pretty easily.

Exploration 62 Let $a, b, c, d, e \in \mathbb{R}$, and consider the data points

$$(-3, a), (-1, b), (0, c), (1, d), \text{ and } (3, e).$$

Find the line of best fit for this data. Note that the equation for your line should depend on the scalars a, b, c, d, and e.

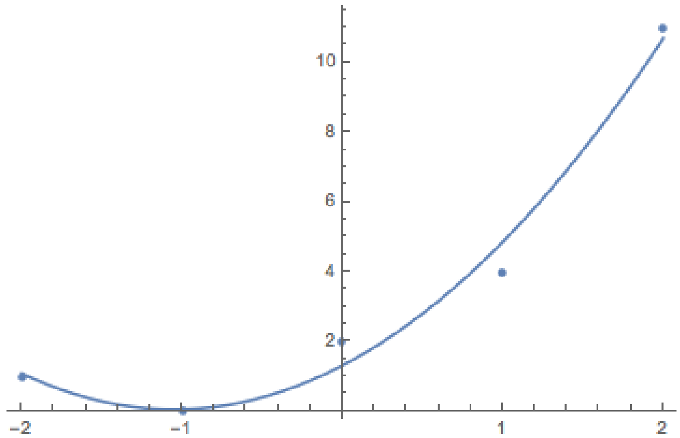


FIGURE 2.12. The last aggressively nonlinear data set with the parabola of best fit.

Exercises for Section 2.6

2.6.1. Let $W = \text{Span}\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ and $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$. Using the standard inner product, how far is $\vec{v}$ from W?

2.6.2. Let $W = \text{Span}\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ and $\vec{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Using the standard inner product, how far is $\vec{v}$ from W?

2.6.3. Let $W = \text{Span}\left\{x^2 - 2, 3x^3 + 1\right\} \subset \mathbb{P}_3$ and $\vec{p} = x + 1$. How far is $\vec{p}$ from W? (Hint: Use coordinate vectors relative to the standard basis for $\mathbb{P}_3$.)