Linear Combinations, Spanning, and Linear Independence

We have seen that there are two operations defined on a given vector space V:

1. vector addition of two vectors, and
2. scalar multiplication of a vector by a scalar.

The most fundamental way to combine vectors in a vector space is by employing these two operations on some collection of vectors and scalars. The definition below makes this idea precise:

Definition 2.3. A *linear combination* of a list $v_1, \ldots, v_m$ of vectors in a vector space V over a field $\mathbb{F}$ is a vector of the form

$$\alpha_1 v_1 + \ldots + \alpha_m v_m,$$

where $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$.

Example. As a simple example, consider the vectors

$$w = \begin{pmatrix} 3 & -5 \\ 0 & 2 \end{pmatrix}$$

and

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ and } e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

in the vector space $\mathcal{U}_2(\mathbb{R})$ of 2×2 real upper triangular matrices. We can think of w as a linear combination of the vectors e_{11}, e_{12}, and e_{22}: indeed, it is quite easy to see that

$$w = 3e_{11} - 5e_{12} + 2e_{22}.$$

In a sense, we can use e_{11}, e_{12}, and e_{22}, together with the two operations of addition and scalar multiplication in $\mathcal{U}_2(\mathbb{R})$, to "build" w.

Example. Describe the set of *all* linear combinations of e_{11}, e_{12}, and e_{22}.

It should be clear that you can use the vectors e_{11}, e_{12}, and e_{22} to build *any* other vector in $\mathcal{U}_2(\mathbb{R})$: if

$$u = \begin{pmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{pmatrix}$$

is a vector in $\mathcal{U}_2(\mathbb{R})$, then we can think of u as the linear combination

$$u = u_{11}e_{11} + u_{12}e_{12} + u_{22}e_{22}.$$

On the other hand, it is easy to see that *any* linear combination of the vectors e_{11}, e_{12}, and e_{22} will be another vector in $\mathcal{U}_2(\mathbb{R})$. Thus there is a sense in which these three vectors give you *all the information you need to know about* $\mathcal{U}_2(\mathbb{R})$.

Span

In the example above, we saw that we could use three vectors to build *all of* $\mathcal{U}_2(\mathbb{R})$. This leads us to a definition:

Definition 2.5. Let $S = (v_1, v_2, \ldots, v_m)$ be a list of vectors in a vector space V. The *span* of S, denoted $\operatorname{span}(S)$ or $\operatorname{span}(v_1, v_2, \ldots, v_m)$, is the set of all linear combinations of vectors in S.

The span of the empty list () is defined to be $\{\mathbf{0}\}$, that is $\operatorname{span}() = \{\mathbf{0}\}$.

Remark. Most literature refers to the span of a *set*, as opposed to the span of a *list* (the difference being that a set is unorderd, while a list has an order). However, order is invariably important when referring to some set of vectors that span, so we solve the problem by thinking of our collection of vectors as a list (even though I will often refer to "spanning sets").

Example. Describe the span of the list (e_{11}, e_{12}, e_{22}).

We saw above that the set of all linear combinations of vectors in the list above is all of $\mathcal{U}_2(\mathbb{R})$; we write

$$\operatorname{span}(e_{11}, e_{12}, e_{22}) = \mathcal{U}_2(\mathbb{R}).$$

It turns out that spans of sets are extremely important in the context of vector spaces:

Theorem 2.7. Let $S = (v_1, v_2, \ldots, v_m)$ be a list of vectors in a vector space V. Then:

1. The span of S is a subspace W of V.
2. $W = \operatorname{span}(S)$ is the smallest subspace of V that contains all of the vectors in S.

With these ideas in mind, we record a definition:

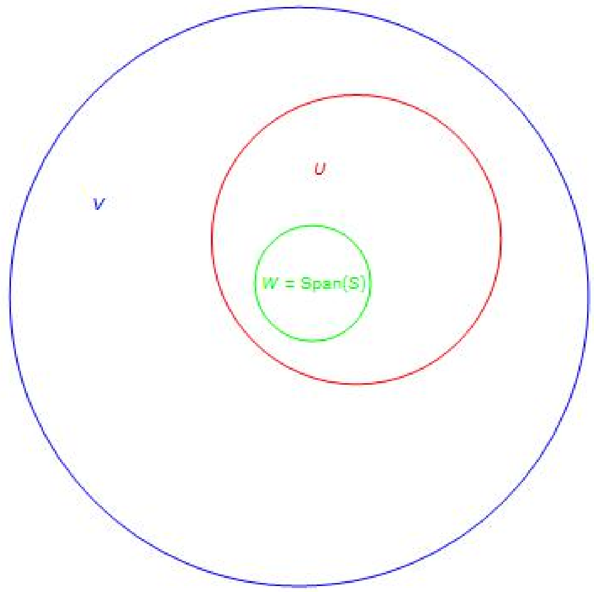
Definition 2.8. If there is a list of vectors $(v_1,\ v_2,\ \ldots,\ v_n)$ so that

$$V = \operatorname{span}(v_1,\ v_2,\ \ldots,\ v_n),$$

we say that the vectors $v_1,\ v_2,\ \ldots,\ v_n$ *span* V.

The theorem above tells us several useful facts. First of all, it guarantees that the set that we create when we create the span of a list of vectors is *automatically* a subspace of the original vector space.

Second, the theorem guarantees that if U is a subspace of V different from $W = \operatorname{span}(S)$ and U contains all of the vectors in S, then U contains *all* of the vectors in W and is a larger vector space than is W:



Key Point. The key point of the theorem is that *spans of sets are vector subspaces.* In a sense, the theorem gives us the power to build many different subspaces of a vector space V: just choose your favorite vectors from V, and take the span (set of all linear combinations) of those vectors. You automatically get a subspace.

Example 1

In a previous section, we saw that the set $\mathcal{U}_2(\mathbb{R})$ of all real upper triangular 2×2 matrices is a subspace of the vector space $\mathcal{M}_2(\mathbb{R})$ of all real 2×2 matrices, by checking that $\mathcal{U}_\in(\mathbb{R})$ was closed under the operations of addition and scalar multiplication.

Theorem 2.7 gives us another way to check that $\mathcal{U}_2(\mathbb{R})$ is a subspace: we know that

$$\mathcal{U}_2(\mathbb{R}) = \operatorname{span}(e_{11}, e_{12}, e_{22}).$$

Since the vectors e_{11}, e_{12}, and e_{22} are also vectors in the vector space $\mathcal{M}_2(\mathbb{R})$, Theorem 2.7 says that their span $\mathcal{U}_2(\mathbb{R})$ is a subspace.

Example 2

The standard unit vectors

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

in $\mathbb{R}_n$ actually span $\mathbb{R}^n$; any n dimensional vector can be written as a linear combination of the vectors e_1, e_2,..., e_n.

Example 3

In the vector space $\mathcal{P}_5(\mathbb{F})$ of all polynomials over $\mathbb{F}$ of degree at most 5, the list

$$S = (1, x, x^2, x^3, x^4, x^5)$$

spans P_5, because *any* vector in $\mathcal{P}_5(\mathbb{F})$ (i.e., any polynomial of degree at most 5) can be written as a linear combination of the vectors in S.

Example 4

In section 1 of this unit, we saw that the set $\mathfrak{sl}(2, \mathbb{R})$ of 2×2 trace 0 matrices is a vector space. Recall that every 2×2 trace 0 matrix has form

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix}.$$

With this in mind, it shoud be clear that the vectors

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

span $\mathfrak{sl}(2, \mathbb{R})$.

Example 5: A set that does not span

Let's return once more to the set $\mathcal{U}_2(\mathbb{R})$ of all upper triangular 2×2 matrices. We saw earlier that the three vectors

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ and } e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

span $\mathcal{U}_2(\mathbb{R})$.

Now if we only consider the list (e_{11}, e_{12}) (i.e., throw away e_{22}), I claim that we *no longer* have a spanning list of $\mathcal{U}_2(\mathbb{R})$. It is actually quite easy to see that the list (e_{11}, e_{12}) does not span $\mathcal{U}_2(\mathbb{R})$: for example, there is no way to find a linear combination of

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

that will yield the vector

$$w = \begin{pmatrix} 4 & 1 \\ 0 & -2 \end{pmatrix}.$$

Since there are elements of $\mathcal{U}_2(\mathbb{R})$ that are not linear combinations of e_{11} and e_{12}, the list (e_{11}, e_{12}) does not span $\mathcal{U}_2(\mathbb{R})$.

Example

Do the vectors

$$v_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

span $\mathbb{R}^2$?

Determining if v_1 and v_2 span $\mathbb{R}^2$ amounts to checking that *every* vector in $\mathbb{R}^2$ can be written as a linear combination of v_1 and v_2.

Given an arbitrary vector

$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

in $\mathbb{R}^2$, we hope to be able to find scalars k_1 and k_2 so that

$$b = k_1 v_1 + k_2 v_2;$$

alternatively, we can write this equation as

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 3k_1 \\ k_1 \end{pmatrix} + \begin{pmatrix} 2k_2 \\ 2k_2 \end{pmatrix}.$$

In equation form, we want to show that the system

$$\begin{aligned} 3k_1 + 2k_2 &= b_1 \\ k_1 + 2k_2 &= b_2 \end{aligned}$$

is consistent for any real numbers b_1 and b_2. Fortunately, we learned a theorem in Unit 1, Section 10 that will allow us to check this quite easily:

Theorem. Let A be an $n \times n$ matrix. Then the following are equivalent:

- $Ax = b$ is consistent for every $n \times 1$ matrix b.
- $\det A \neq 0$.

Thinking of our system

$$\begin{aligned} 3k_1 + 2k_2 &= b_1 \\ k_1 + 2k_2 &= b_2 \end{aligned}$$

in matrix form as

$$\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} b_2 \\ b_2 \end{pmatrix},$$

we see that the system will be consistent for every vector b in $\mathbb{R}^2$, as desired, *if and only if* the determinant of the coefficient matrix for the system is nonzero. Let's check the determinant:

$$\begin{aligned} \det \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} &= 3 \cdot 2 - 2 \cdot 1 \\ &= 6 - 2 \\ &= 4. \end{aligned}$$

Since the determinant of the coefficient matrix for the system is nonzero, the system is consistent for *any* vector $\mathbf{b}$ in $\mathbb{R}^2$. In other words, no matter how we choose the vector $\mathbf{b}$, we will always be able to find scalars k_1 and k_2 so that

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 3k_1 \\ k_1 \end{pmatrix} + \begin{pmatrix} 2k_2 \\ 2k_2 \end{pmatrix}.$$

Thus the vectors $\mathbf{v_1}$ and $\mathbf{v_2}$ span $\mathbb{R}^2$, so that we may use them to "build" all of $\mathbb{R}^2$.

For example, if we choose

$$\mathbf{b} = \begin{pmatrix} -7 \\ 3 \end{pmatrix},$$

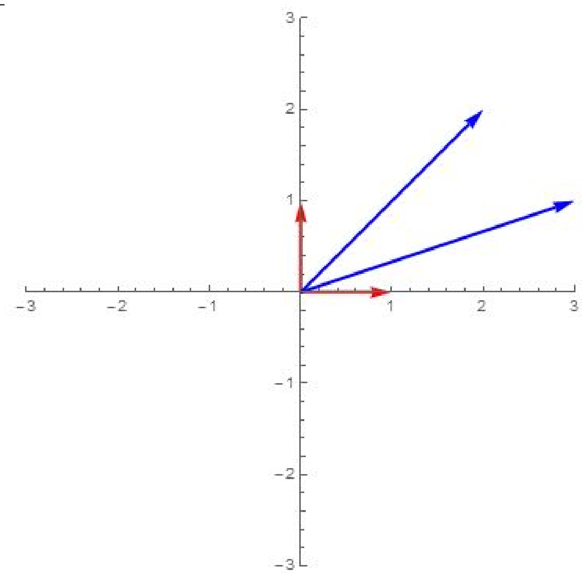
we can write $\mathbf{b}$ as the linear combination

$$\begin{pmatrix} -7 \\ 3 \end{pmatrix} = -5 \begin{pmatrix} 3 \\ 1 \end{pmatrix} + 4 \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

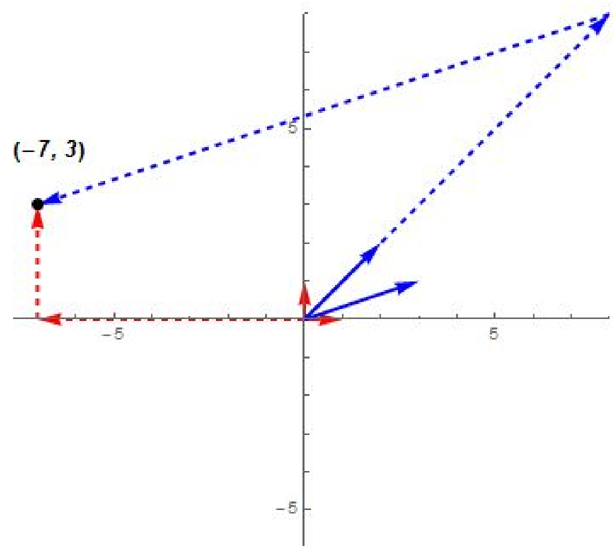
Key Point. At this point, we know of *two different* lists that span $\mathbb{R}^2$:

$$\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \text{ and } \left(\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right).$$

In a sense, these two sets give us *two different ways* to build $\mathbb{R}^2$. The two alternate spanning sets are graphed below in $\mathbb{R}^2$:



As indicated above, we can use either spanning set to build any vector in $\mathbb{R}^2$ that we like, as illustrated below with the vector $\begin{pmatrix} -7 & 3 \end{pmatrix}$:



Finite and Infinite Dimensional Vector Spaces

We will soon see that the ideas of *spanning* and *linear combination* are closely tied to the idea of the "size" of a vector space. Thinking back to Example 2 above in $\mathbb{R}^n$, we saw that there is a set of 2 vectors that spans $\mathbb{R}^2$; a set of 3 vectors that spans $\mathbb{R}^3$; etc.

Of course, we also think of $\mathbb{R}^2$ as a "smaller" space than $\mathbb{R}^3$ (indeed, $\mathbb{R}^2$ is embedded in $\mathbb{R}^3$); we will soon see that the relative sizes of spanning sets correspond in a natural way to the relative sizes of the spaces themselves (we will discuss this idea further in the next few sections).

At this point, we would like to make an observation: *there are vector spaces that are not spanned by any finite set.*

Example. Let $\mathcal{P}(\mathbb{R})$ be the vector space of *all* polynomials over $\mathbb{R}$, and let S be any (finite) list of vectors in $\mathcal{P}(\mathbb{R})$. The list is finite, so there is a vector in the list of highest degree, say $\alpha_m x^m + \ldots + \alpha_1 x + \alpha_0$.

Clearly the polynomial x^{m+1} is *not* a linear combination of vectors from S, so no finite set of vectors spans $\mathcal{P}(\mathbb{R})$.

With this example in mind, we record a definition, and follow the definition with a reasonable assumption:

Definitions 2.10/2.15. A vector space V is called *finite dimensional* if there is a finite list of vectors spanning V. Otherwise, V is *infinite dimensional.*

Example. We saw above that there is no finite list spanning $\mathcal{P}(\mathbb{R})$, so it is an example of an infinite dimensional vector space. In fact, the spaces $\mathbb{R}(-\infty, \infty)$ and $C(-\infty, \infty)$ of real-valued functions and continuous functions with domain $(-\infty, \infty)$ are both infinite dimensional.

On the other hand, $\mathcal{P}_n(\mathbb{F})$ is finite dimensional for any $n < \infty$, as are $\mathbb{R}^n$ and $\mathcal{M}_{mn}$.

Remark. Unless otherwise specified, we will assume throughout this course that all of our vector spaces are finite dimensional–i.e. that each one has a finite list of vectors spanning it.

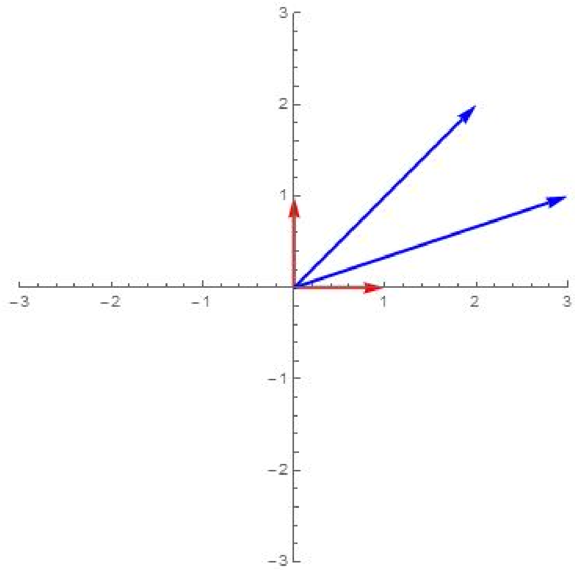
Linear Independence

We saw above that, if a list $S = (v_1, v_2, \ldots, v_m)$ of vectors from a vector space V spans V, then every vector in V can be written as a linear combination of the vectors in S. In a sense, we can use the vectors in S, along with the operations of addition and scalar multiplication on V, to "build" any vector we want from V.

We looked at an example above in $\mathbb{R}^2$: we saw that both of the lists

$$S = (e_1,\ e_2) = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \text{ and } T = (v_1,\ v_2) = \left(\begin{pmatrix} 3 \\ 1 \end{pmatrix},\ \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right)$$

span $\mathbb{R}^2$; the vectors from S is graphed below in red, and those from T are graphed in blue.



Now I claim that the list of *four* vectors

$$R = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right)$$

is *also* a spanning list. Of course, my claim is quite easy to check: since the list R includes all of the vectors from the spanning lists S and T, we can certainly write any vector in $\mathbb{R}^2$ as a linear combination of the vectors in R.

Unfortunately, this spanning list causes us some problems. As an example, consider the vector

$$w = \begin{pmatrix} 6 \\ 4 \end{pmatrix}.$$

Since w is in $\mathbb{R}^2$, it is a linear combination of the vectors in the spanning set R, say

$$\begin{aligned} w &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ &= e_1 + e_2 + v_1 + v_2. \end{aligned}$$

However, we could also write w as a linear combination of the vectors in R as

$$\begin{aligned} w &= 6\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= 6e_1 + 4e_2 + 0v_1 + 0v_2, \end{aligned}$$

or even as

$$\begin{aligned} w &= -4\begin{pmatrix} 1 \\ 0 \end{pmatrix} - 6\begin{pmatrix} 0 \\ 1 \end{pmatrix} + 5\begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ &= -4e_1 - 6e_2 + 0v_1 + 5v_2. \end{aligned}$$

In a sense, this particular spanning set for $\mathbb{R}^2$ introduces ambiguity: it gives us multiple ways to write a particular vector in $\mathbb{R}^2$ as a linear combination of vectors in the spanning set.

Now, if we restrict our lists to just

$$S = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

or alternatively to

$$T = \left(\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right),$$

we won't have this difficulty: a vector in $\mathbb{R}^2$ can be written in *only one way* as a linear combination of vectors in S, and in only one way as a linear combination of vectors in T.

So there is a sense in which the spanning list

$$R = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right)$$

for $\mathbb{R}^2$ is just too big: it gives us too much flexibility, introducing ambiguity into the way we build vectors (should we write w as $w = e_1 + e_2 + v_1 + v_2$, or as $w = -4e_1 - 6e_2 + 5v_2$?)

Key Point. Saying that a list S spans a vector space V is the same as saying that we can use the vectors in S to build any vector that we like from V; there is a sense in which S is all we need to know in order to understand V.

Of course, we want to avoid any ambiguity in our understanding of V, so we would like to choose the list S to be *as small as possible.* We quantify what we mean when we say that a set is "too big" below with a discussion of *linear independence.*

Linearly Independent Lists

The reason that the spanning set

$$R = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right)$$

is "too big" is that *we can use some of the vectors in R to build other vectors in R.* For example, you should check that e_1 can be written as the linear combination

$$e_1 = \frac{1}{2}v_1 - \frac{1}{4}v_2.$$

So in a sense, we could throw out e_1 from our spanning list without losing any information: anytime we want to write e_1, we could just as well write $\frac{1}{2}v_1 - \frac{1}{4}v_2$.

With these ideas in mind, we introduce the idea of a list of linearly independent vectors:

Definitions 2.17/2.19. A nonempty list $S = (v_1, v_2, \ldots, v_m)$ of one or more vectors in V is said to be *linearly independent* if the only choice of scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ so that

$$\alpha_1 v_1 + \ldots + \alpha_m v_m = 0$$

is

$$\alpha_1 = \ldots = \alpha_m = 0.$$

Otherwise, the list of vectors is *linearly dependent*.

For purposes of convenience, we choose to call the empty list () independent.

Key Point. Saying that a set of vectors is linearly independent is the same as saying that they are "unique", or perhaps even *essential*; you can't build one of them using the others, so losing one of the vectors in the list would result in a loss of information about the list and its span.

Altering the observation from the example above just a bit, we see that

$$e_1 - \frac{1}{2}v_1 + \frac{1}{4}v_2 = \mathbf{0};$$

thus

$$R = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right)$$

is a list of dependent vectors.

However, we will see momentarily that lists

$$S = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

and

$$T = \left(\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right)$$

are both linearly independent.

Remark. Any (finite) set containing the $\mathbf{0}$ vector is automatically linearly dependent.

Examples of Linearly Independent Vectors

Example 1

In $\mathbb{R}^n$, the standard unit vectors

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

(which span $\mathbb{R}^n$) are also linearly independent.

To see why, suppose that $\alpha_1, \ldots, a_n$ are constants so that

$$\alpha_1 e_1 + \ldots + \alpha_n e_n = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Then clearly $\alpha_1 = \ldots = \alpha_n = 0$, and the vectors are linearly independent.

Example 2

The list

$$T = \left(\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \; \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right)$$

of vectors in $\mathbb{R}^2$ is an independent list: indeed, suppose that

$$\alpha_1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Then we have

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \alpha_1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 3\alpha_1 \\ \alpha_1 \end{pmatrix} + \begin{pmatrix} 2\alpha_2 \\ 2\alpha_2 \end{pmatrix} \\ &= \begin{pmatrix} 3\alpha_1 + 2\alpha_2 \\ \alpha_1 + 2\alpha_2 \end{pmatrix}. \end{aligned}$$

In other words, we would like to know what values for α_1 and α_2 will make the system

$$\begin{aligned} 3\alpha_1 + 2\alpha_2 &= 0 \\ \alpha_1 + 2\alpha_2 &= 0 \end{aligned}$$

consistent.

Once again, we return to a theorem from Unit 1, Section 10:

Theorem. Let A be an $n \times n$ matrix. Then the following are equivalent:

- $Ax = \mathbf{0}$ has only the trivial solution.
- $\det A \neq 0$.

Thinking of our system

$$\begin{aligned} 3\alpha_1 + 2\alpha_2 &= 0 \\ \alpha_1 + 2\alpha_2 &= 0 \end{aligned}$$

in matrix form as

$$\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

we see that we should check the determinant of the coefficient matrix:

$$\begin{aligned} \det \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} &= 3 \cdot 2 - 2 \cdot 1 \\ &= 6 - 2 \\ &= 4. \end{aligned}$$

Since the determinant of the coefficient matrix is nonzero, the system

$$\begin{aligned} 3\alpha_1 + 2\alpha_2 &= 0 \\ \alpha_1 + 2\alpha_2 &= 0 \end{aligned}$$

has *only* the trivial solution $\alpha_1 = \alpha_2 = 0$: in other words, the only way to make the statements true is to choose $\alpha_1 = 0$ and $\alpha_2 = 0$.

Thus the vectors in the list

$$T = \left(\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right)$$

are linearly independent.

Example 3

We saw above that the list

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right)$$

does not span the vector space $\mathcal{U}_2(\mathbb{R})$ of all real upper triangular 2×2 matrices.

However, it is quite easy to see that this set *is* linearly independent: the only way to make

$$\begin{pmatrix} k_1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & k_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

is by choosing $k_1 = k_2 = 0$.

Example 4

The vectors

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

span $\mathfrak{sl}(2, \mathbb{R})$; I claim that they are also linearly independent. This fact is easy to see: if

$$\begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix} + \begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \gamma & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

then clearly

$$\alpha = \beta = \gamma = 0.$$

Example of a List of Dependent Vectors

I claim that the vectors

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \end{pmatrix}$$

are *not* linearly independent in $\mathbb{R}^4$. We can check that this is true by inspecting the coefficient matrix of the system of equations

$$\begin{aligned} \alpha_1 + 0\alpha_2 + 0\alpha_3 + 0\alpha_4 &= 0 \\ 0\alpha_1 + 0\alpha_2 + 0\alpha_3 + 0\alpha_4 &= 0 \\ \alpha_1 + \alpha_2 - 2\alpha_3 + 0\alpha_4 &= 0 \\ 3\alpha_1 + 0\alpha_2 + \alpha_3 + 3\alpha_4 &= 0. \end{aligned}$$

Again, we resort to the theorem from Unit 1, Section 10:

Theorem. Let A be an $n \times n$ matrix. Then the following are equivalent:

- $Ax = \mathbf{0}$ has only the trivial solution.
- $\det A \neq 0$.

Since the statements above are equivalent, we know that there are nontrivial solutions α_1, α_2, α_3, and α_4 to the system if and only if the determinant of the coefficient matrix is 0. The coefficient matrix is given by

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 3 & 0 & 1 & 3 \end{pmatrix}.$$

Since the matrix is lower triangular, its determinant is just the product of its diagonal entries; we have

$$\det A = 1 \cdot 0 \cdot (-2) \cdot 3 = 0,$$

which indicates that there are nontrivial solutions to the system. Thus the vectors are linearly dependent.

Relative Sizes of Spanning Lists and Independent Lists

We have considered the list

$$S = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right)$$

of vectors in $\mathcal{U}_2(\mathbb{R})$ several times; in particular, we have seen that the list is independent, but does not span $\mathcal{U}_2(\mathbb{R})$.

On the other hand, the list

$$S' = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

certainly spans $\mathcal{U}_2(\mathbb{R})$; however, it is not an independent list (you should check)!

The relative sizes of these lists is important to note when considering the phenomena described above. Notice that S is a relatively small list, while S' is relatively large. Smaller lists are more likely to be *independent*, whereas larger lists are more likely to span. In a sense, the more vectors you add to a list, the more likely you are to introduce some dependence relations (thus preventing the list from being independent); but adding more vectors means that you are also more likely to be able to build the entire space (i.e., span).

These ideas are made concrete by the following theorem; recall that the notation $|\cdot|$, when applied to a list or set, refers to the number of items in the set or "length" of the list.

Theorem 2.23. Let V be a finite dimensional vector space. Suppose that I is a list of independent vectors in the space, and that S is a list of vectors that span V. Then

$$|I| \leq |S|.$$

Before we look at a proof of the theorem, let us discuss a helpful lemma:

Linear Dependence Lemma. Suppose that $v_1, \ldots, v_m$ is a linearly dependent list in V. Then there is a j, $1 \leq j \leq m$, so that

1. $v_j \in \operatorname{span}(v_1, v_2, \ldots v_{j-1})$, and
2. $\operatorname{span}(v_1, v_2, \ldots v_m) = \operatorname{span}(v_1, v_2, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$; that is, removing v_j from the list does not affect the span of the list.

We will not give a detailed proof of the lemma, but will discuss it briefly. Clearly if the set is dependent, then there are constants $\alpha_1, \ldots \alpha_n$, not all 0, so that $\alpha_1 v_1 + \ldots + \alpha_m v_m = \mathbf{0}$. We may assume that $\alpha_m \neq 0$ (reorder the list if necessary). Then

$$v_m = -\frac{\alpha_1}{\alpha_m} v_1 - \ldots - \frac{\alpha_{m-1}}{\alpha_m} v_{m-1},$$

so that $v_m \in \operatorname{span}(v_1, \ \ldots, \ v_{m-1})$.

For part (2), suppose that w is a linear combination of the vectors in the list; simply replace v_m by

$$-\frac{\alpha_1}{\alpha_m} v_1 - \ldots - \frac{\alpha_{m-1}}{\alpha_m} v_{m-1},$$

so that w is a linear combination of the remaining vectors.

Key Point. If v is a linear combination of other vectors in a list, removing v from the list will not affect the span of the list.

Proof of Theorem 2.23. Given the independent list

$$I = (u_1, \ u_2, \ \ldots, \ u_m)$$

and the (finite) spanning list

$$S = (w_1, \ w_2, \ \ldots, \ w_n),$$

we wish to show that $m \leq n$. We proceed in an iterative fashion:

Step 1:

Create the list

$$\hat{S}_1 = (u_1, \ w_1, \ \ldots, \ w_n).$$

Vector u_1 is in the span of S, so $\hat{S}_1$ is a dependent list with the same span as that of S, that is,

$$\operatorname{span}(\hat{S}) = \operatorname{span}(S) = V.$$

Now u_1 is a linear combination of the vectors in S, say

$$u_1 = \alpha_1 w_1 + \ldots + \alpha_n w_n;$$

we may assume that $\alpha_n \neq 0$ (reorder the list if necessary). Thus w_n is a linear combination of vectors in the list

$$S_1 = (u_1, \ w_1, \ w_2, \ \ldots, \ w_{n-1}),$$

created from $\hat{S}$ by deleting w_n. By applying the Linear Dependence Lemma to the dependent set $\hat{S}_1$, we see that removing w_n from the list does not change its span. We see that

$$\operatorname{span}(S_1) = \operatorname{span}(\hat{S}_1) = V.$$

Step 2:

Vector u_2 is in the span of S_1, so the list

$$\hat{S}_2 = (u_1,\ u_2,\ w_1,\ \ldots,\ w_{n-1})$$

is dependent (and spans V). Now u_2 is a linear combination of the vectors in the list, say

$$u_2 = \beta_1 u_1 + \alpha_1 w_1 + \ldots + \alpha_{n-1} w_{n-1}.$$

Now the list I is independent, so

$$u_2 - \beta_1 u_1 \neq 0.$$

Thus *at least one* $\alpha_i \neq 0$, say α_{n-1} (again, reorder the list if necessary). Thus w_{n-1} is a linear combination of the vectors in the list

$$S_2 = (u_1,\ u_2,\ w_1,\ w_2,\ \ldots,\ w_{n-2}),$$

created from the dependent list $\hat{S}_2$ by removing w_{n-1}. Again, we have merely removed a vector that is a linear combination of the others from a dependent list, so the Linear Dependence Lemma applies:

$$\text{span}\,(S_2) = \text{span}\,(\hat{S}_2) = V.$$

Steps 3-m:

Iterate the process from step 2 of adding u_i to S_{i-1} and deleting an appropriate w_j to create the new spanning list S_i. At each step, there must be at least one w_j available for deletion, because if not we would have

$$u_i = \beta_1 u_1 + \ldots + \beta_{i-1} u_{i-1},$$

violating the independence of I. Thus we will not run out of vectors w_j *before* we run out of vectors u_i, implying that

$$|I| \leq |S|.$$

We record without proof one more theorem in this section, which will become useful later:

Theorem 2.26. Every subspace of a finite dimensional vector space is finite dimensional.

In other words, if V has a finite spanning list, then so does any subspace of V.