# Vector Spaces and Linear Maps

## Notes for students of MA2008

In our course we have 5 or 6 weeks to cover the basics of the theory of vector spaces and linear maps between them, and it is therefore essential that you have a good book with worthwhile exercises.

The first year linear algebra course MA1010 is not an essential prerequisite, though it does provide some of the motivation for the abstract definitions.

## Acknowledgement

This file is simply a cut-down version of the free online linear algebra text book by Jim Hefferon, available from

`http://joshua.smcvt.edu/linearalgebra/`

I have taken the most important sections from the book by Hefferon, which I will be following approximately in the lectures. If you do not mind printing 440 pages instead of 100, then I recommend you download the original instead!

Chapter Two

# Vector Spaces

The first chapter began by introducing Gauss' method and finished with a fair understanding, keyed on the Linear Combination Lemma, of how it finds the solution set of a linear system. Gauss' method systematically takes linear combinations of the rows. With that insight, we now move to a general study of linear combinations.

We need a setting for this study. At times in the first chapter, we've combined vectors from $\mathbb{R}^2$, at other times vectors from $\mathbb{R}^3$, and at other times vectors from even higher-dimensional spaces. Thus, our first impulse might be to work in $\mathbb{R}^n$, leaving $n$ unspecified. This would have the advantage that any of the results would hold for $\mathbb{R}^2$ and for $\mathbb{R}^3$ and for many other spaces, simultaneously.

But, if having the results apply to many spaces at once is advantageous then sticking only to $\mathbb{R}^n$'s is overly restrictive. We'd like the results to also apply to combinations of row vectors, as in the final section of the first chapter. We've even seen some spaces that are not just a collection of all of the same-sized column vectors or row vectors. For instance, we've seen a solution set of a homogeneous system that is a plane, inside of $\mathbb{R}^3$. This solution set is a closed system in the sense that a linear combination of these solutions is also a solution. But it is not just a collection of all of the three-tall column vectors; only some of them are in this solution set.

We want the results about linear combinations to apply anywhere that linear combinations are sensible. We shall call any such set a *vector space*. Our results, instead of being phrased as "Whenever we have a collection in which we can sensibly take linear combinations ...", will be stated as "In any vector space ...".

Such a statement describes at once what happens in many spaces. The step up in abstraction from studying a single space at a time to studying a class of spaces can be hard to make. To understand its advantages, consider this analogy. Imagine that the government made laws one person at a time: "Leslie Jones can't jay walk." That would be a bad idea; statements have the virtue of economy when they apply to many cases at once. Or, suppose that they ruled, "Kim Ke must stop when passing the scene of an accident." Contrast that with, "Any doctor must stop when passing the scene of an accident." More general statements, in some ways, are clearer.

## I Definition of Vector Space

We shall study structures with two operations, an addition and a scalar multiplication, that are subject to some simple conditions. We will reflect more on the conditions later, but on first reading notice how reasonable they are. For instance, surely any operation that can be called an addition (e.g., column vector addition, row vector addition, or real number addition) will satisfy all the conditions in (1) below.

## I.1 Definition and Examples

**1.1 Definition** A *vector space* (over $\mathbb{R}$) consists of a set $V$ along with two operations '+' and '·' subject to these conditions.

Where $v, w \in V$, (1) their *vector sum* $v + w$ is an element of $V$. If $u, v, w \in V$ then (2) $v + w = w + v$ and (3) $(v + w) + u = v + (w + u)$. (4) There is a *zero vector* $0 \in V$ such that $v + 0 = v$ for all $v \in V$. (5) Each $v \in V$ has an *additive inverse* $w \in V$ such that $w + v = 0$.

If $r, s$ are *scalars*, members of $\mathbb{R}$, and $v, w \in V$ then (6) each *scalar multiple* $r \cdot v$ is in $V$. If $r, s \in \mathbb{R}$ and $v, w \in V$ then (7) $(r + s) \cdot v = r \cdot v + s \cdot v$, and (8) $r \cdot (v + w) = r \cdot v + r \cdot w$, and (9) $(rs) \cdot v = r \cdot (s \cdot v)$, and (10) $1 \cdot v = v$.

**1.2 Remark** Because it involves two kinds of addition and two kinds of multiplication, that definition may seem confused. For instance, in condition (7) '$(r + s) \cdot v = r \cdot v + s \cdot v$', the first '+' is the real number addition operator while the '+' to the right of the equals sign represents vector addition in the structure $V$. These expressions aren't ambiguous because, e.g., $r$ and $s$ are real numbers so '$r + s$' can only mean real number addition.

The best way to go through the examples below is to check all ten conditions in the definition. That check is written out at length in the first example. Use it as a model for the others. Especially important are the first condition '$v + w$ is in $V$' and the sixth condition '$r \cdot v$ is in $V$'. These are the *closure* conditions. They specify that the addition and scalar multiplication operations are always sensible — they are defined for every pair of vectors, and every scalar and vector, and the result of the operation is a member of the set (see Example 1.4).

**1.3 Example** The set $\mathbb{R}^2$ is a vector space if the operations '+' and '·' have their usual meaning.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \qquad r \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} rx_1 \\ rx_2 \end{pmatrix}$$

We shall check all of the conditions.

There are five conditions in item (1). For (1), closure of addition, note that for any $v_1, v_2, w_1, w_2 \in \mathbb{R}$ the result of the sum

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix}$$

is a column array with two real entries, and so is in $\mathbb{R}^2$. For (2), that addition of vectors commutes, take all entries to be real numbers and compute

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} = \begin{pmatrix} w_1 + v_1 \\ w_2 + v_2 \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

(the second equality follows from the fact that the components of the vectors are real numbers, and the addition of real numbers is commutative). Condition (3), associativity of vector addition, is similar.

$$\begin{aligned} (\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}) + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} &= \begin{pmatrix} (v_1 + w_1) + u_1 \\ (v_2 + w_2) + u_2 \end{pmatrix} \\ &= \begin{pmatrix} v_1 + (w_1 + u_1) \\ v_2 + (w_2 + u_2) \end{pmatrix} \\ &= \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + (\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}) \end{aligned}$$

For the fourth condition we must produce a zero element — the vector of zeroes is it.

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

For (5), to produce an additive inverse, note that for any $v_1, v_2 \in \mathbb{R}$ we have

$$\begin{pmatrix}-v_1\\-v_2\end{pmatrix}+\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

so the first vector is the desired additive inverse of the second.

The checks for the five conditions having to do with scalar multiplication are just as routine. For (6), closure under scalar multiplication, where $r, v_1, v_2 \in \mathbb{R}$,

$$r\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}rv_1\\rv_2\end{pmatrix}$$

is a column array with two real entries, and so is in $\mathbb{R}^2$. Next, this checks (7).

$$(r+s)\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}(r+s)v_1\\(r+s)v_2\end{pmatrix}=\begin{pmatrix}rv_1+sv_1\\rv_2+sv_2\end{pmatrix}=r\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}+s\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}$$

For (8), that scalar multiplication distributes from the left over vector addition, we have this.

$$r\cdot(\begin{pmatrix}v_1\\v_2\end{pmatrix}+\begin{pmatrix}w_1\\w_2\end{pmatrix})=\begin{pmatrix}r(v_1+w_1)\\r(v_2+w_2)\end{pmatrix}=\begin{pmatrix}rv_1+rw_1\\rv_2+rw_2\end{pmatrix}=r\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}+r\cdot\begin{pmatrix}w_1\\w_2\end{pmatrix}$$

The ninth

$$(rs)\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}(rs)v_1\\(rs)v_2\end{pmatrix}=\begin{pmatrix}r(sv_1)\\r(sv_2)\end{pmatrix}=r\cdot(s\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix})$$

and tenth conditions are also straightforward.

$$1\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}1v_1\\1v_2\end{pmatrix}=\begin{pmatrix}v_1\\v_2\end{pmatrix}$$

In a similar way, each $\mathbb{R}^n$ is a vector space with the usual operations of vector addition and scalar multiplication. (In $\mathbb{R}^1$, we usually do not write the members as column vectors, i.e., we usually do not write '$(\pi)$'. Instead we just write '$\pi$'.)

**1.4 Example** This subset of $\mathbb{R}^3$ that is a plane through the origin

$$P=\{\begin{pmatrix}x\\y\\z\end{pmatrix}\mid x+y+z=0\}$$

is a vector space if '+' and '$\cdot$' are interpreted in this way.

$$\begin{pmatrix}x_1\\y_1\\z_1\end{pmatrix}+\begin{pmatrix}x_2\\y_2\\z_2\end{pmatrix}=\begin{pmatrix}x_1+x_2\\y_1+y_2\\z_1+z_2\end{pmatrix}\qquad r\cdot\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}rx\\ry\\rz\end{pmatrix}$$

The addition and scalar multiplication operations here are just the ones of $\mathbb{R}^3$, reused on its subset $P$. We say that $P$ *inherits* these operations from $\mathbb{R}^3$. This example of an addition in $P$

$$\begin{pmatrix}1\\1\\-2\end{pmatrix}+\begin{pmatrix}-1\\0\\1\end{pmatrix}=\begin{pmatrix}0\\1\\-1\end{pmatrix}$$

illustrates that $P$ is closed under addition. We've added two vectors from $P$—that is, with the property that the sum of their three entries is zero—and the result is a vector also in $P$. Of course, this example of closure is not a proof of closure. To prove that $P$ is closed under addition, take two elements of $P$

$$\begin{pmatrix}x_1\\y_1\\z_1\end{pmatrix}\quad\begin{pmatrix}x_2\\y_2\\z_2\end{pmatrix}$$

(membership in $P$ means that $x_1 + y_1 + z_1 = 0$ and $x_2 + y_2 + z_2 = 0$), and observe that their sum

$$\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}$$

is also in $P$ since its entries add $(x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) = (x_1 + y_1 + z_1) + (x_2 + y_2 + z_2)$ to $0$. To show that $P$ is closed under scalar multiplication, start with a vector from $P$

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

(so that $x + y + z = 0$) and then for $r \in \mathbb{R}$ observe that the scalar multiple

$$r \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} rx \\ ry \\ rz \end{pmatrix}$$

satisfies that $rx + ry + rz = r(x + y + z) = 0$. Thus the two closure conditions are satisfied. Verification of the other conditions in the definition of a vector space are just as straightforward.

**1.5 Example** Example 1.3 shows that the set of all two-tall vectors with real entries is a vector space. Example 1.4 gives a subset of an $\mathbb{R}^n$ that is also a vector space. In contrast with those two, consider the set of two-tall columns with entries that are integers (under the obvious operations). This is a subset of a vector space, but it is not itself a vector space. The reason is that this set is not closed under scalar multiplication, that is, it does not satisfy condition (6). Here is a column with integer entries, and a scalar, such that the outcome of the operation

$$0.5 \cdot \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1.5 \end{pmatrix}$$

is not a member of the set, since its entries are not all integers.

**1.6 Example** The singleton set

$$\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \}$$

is a vector space under the operations

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad r \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

that it inherits from $\mathbb{R}^4$.

A vector space must have at least one element, its zero vector. Thus a one-element vector space is the smallest one possible.

**1.7 Definition** A one-element vector space is a *trivial* space.

Warning! The examples so far involve sets of column vectors with the usual operations. But vector spaces need not be collections of column vectors, or even of row vectors. Below are some other types of vector spaces. The term ‘vector space’ does not mean ‘collection of columns of reals’. It means something more like ‘collection in which any linear combination is sensible’.

**1.8 Example** Consider $\mathcal{P}_3 = \{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_0, \ldots, a_3 \in \mathbb{R}\}$, the set of polynomials of degree three or less (in this book, we'll take constant polynomials, including the zero polynomial, to be of degree zero). It is a vector space under the operations

$$(a_0 + a_1x + a_2x^2 + a_3x^3) + (b_0 + b_1x + b_2x^2 + b_3x^3) \\ = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3$$

and

$$r \cdot (a_0 + a_1x + a_2x^2 + a_3x^3) = (ra_0) + (ra_1)x + (ra_2)x^2 + (ra_3)x^3$$

(the verification is easy). This vector space is worthy of attention because these are the polynomial operations familiar from high school algebra. For instance, $3 \cdot (1 - 2x + 3x^2 - 4x^3) - 2 \cdot (2 - 3x + x^2 - (1/2)x^3) = -1 + 7x^2 - 11x^3$.

Although this space is not a subset of any $\mathbb{R}^n$, there is a sense in which we can think of $\mathcal{P}_3$ as "the same" as $\mathbb{R}^4$. If we identify these two spaces's elements in this way

$$a_0 + a_1x + a_2x^2 + a_3x^3 \quad \text{corresponds to} \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

then the operations also correspond. Here is an example of corresponding additions.

$$\begin{array}{rr} & 1 - 2x + 0x^2 + 1x^3 \\ + & 2 + 3x + 7x^2 - 4x^3 \\ \hline & 3 + 1x + 7x^2 - 3x^3 \end{array} \quad \text{corresponds to} \quad \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ 7 \\ -4 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 7 \\ -3 \end{pmatrix}$$

Things we are thinking of as "the same" add to "the same" sum. Chapter Three makes precise this idea of vector space correspondence. For now we shall just leave it as an intuition.

**1.9 Example** The set $\mathcal{M}_{2\times2}$ of $2\times2$ matrices with real number entries is a vector space under the natural entry-by-entry operations.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} w & x \\ y & z \end{pmatrix} = \begin{pmatrix} a+w & b+x \\ c+y & d+z \end{pmatrix} \qquad r \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ra & rb \\ rc & rd \end{pmatrix}$$

As in the prior example, we can think of this space as "the same" as $\mathbb{R}^4$.

**1.10 Example** The set $\{f \mid f\colon \mathbb{N} \to \mathbb{R}\}$ of all real-valued functions of one natural number variable is a vector space under the operations

$$(f_1 + f_2)\,(n) = f_1(n) + f_2(n) \qquad (r \cdot f)\,(n) = r\,f(n)$$

so that if, for example, $f_1(n) = n^2 + 2\sin(n)$ and $f_2(n) = -\sin(n) + 0.5$ then $(f_1 + 2f_2)\,(n) = n^2 + 1$.

We can view this space as a generalization of Example 1.3 — instead of 2-tall vectors, these functions are like infinitely-tall vectors.

$$\begin{array}{c|c} n & f(n) = n^2 + 1 \\ \hline 0 & 1 \\ 1 & 2 \\ 2 & 5 \\ 3 & 10 \\ \vdots & \vdots \end{array} \quad \text{corresponds to} \quad \begin{pmatrix} 1 \\ 2 \\ 5 \\ 10 \\ \vdots \end{pmatrix}$$

Addition and scalar multiplication are component-wise, as in Example 1.3. (We can formalize "infinitely-tall" by saying that it means an infinite sequence, or that it means a function from $\mathbb{N}$ to $\mathbb{R}$.)

**1.11 Example** The set of polynomials with real coefficients

$$\{a_0 + a_1x + \cdots + a_nx^n \bigm| n \in \mathbb{N} \text{ and } a_0, \ldots, a_n \in \mathbb{R}\}$$

makes a vector space when given the natural '+'

$$(a_0 + a_1x + \cdots + a_nx^n) + (b_0 + b_1x + \cdots + b_nx^n) \\ = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$

and '·'.

$$r \cdot (a_0 + a_1x + \ldots a_nx^n) = (ra_0) + (ra_1)x + \ldots (ra_n)x^n$$

This space differs from the space $\mathcal{P}_3$ of Example 1.8. This space contains not just degree three polynomials, but degree thirty polynomials and degree three hundred polynomials, too. Each individual polynomial of course is of a finite degree, but the set has no single bound on the degree of all of its members.

This example, like the prior one, can be thought of in terms of infinite-tuples. For instance, we can think of $1 + 3x + 5x^2$ as corresponding to $(1, 3, 5, 0, 0, \ldots)$. However, don't confuse this space with the one from Example 1.10. Each member of this set has a bounded degree, so under our correspondence there are no elements from this space matching $(1, 2, 5, 10, \ldots)$. The vectors in this space correspond to infinite-tuples that end in zeroes.

**1.12 Example** The set $\{f \bigm| f\colon \mathbb{R} \to \mathbb{R}\}$ of all real-valued functions of one real variable is a vector space under these.

$$(f_1 + f_2)\,(x) = f_1(x) + f_2(x) \qquad (r \cdot f)\,(x) = r\,f(x)$$

The difference between this and Example 1.10 is the domain of the functions.

**1.13 Example** The set $F = \{a\cos\theta + b\sin\theta \bigm| a, b \in \mathbb{R}\}$ of real-valued functions of the real variable $\theta$ is a vector space under the operations

$$(a_1\cos\theta + b_1\sin\theta) + (a_2\cos\theta + b_2\sin\theta) = (a_1 + a_2)\cos\theta + (b_1 + b_2)\sin\theta$$

and

$$r \cdot (a\cos\theta + b\sin\theta) = (ra)\cos\theta + (rb)\sin\theta$$

inherited from the space in the prior example. (We can think of $F$ as "the same" as $\mathbb{R}^2$ in that $a\cos\theta + b\sin\theta$ corresponds to the vector with components $a$ and $b$.)

**1.14 Example** The set

$$\{f\colon \mathbb{R} \to \mathbb{R} \bigm| \frac{d^2f}{dx^2} + f = 0\}$$

is a vector space under the, by now natural, interpretation.

$$(f + g)\,(x) = f(x) + g(x) \qquad (r \cdot f)\,(x) = r\,f(x)$$

In particular, notice that closure is a consequence:

$$\frac{d^2(f + g)}{dx^2} + (f + g) = (\frac{d^2f}{dx^2} + f) + (\frac{d^2g}{dx^2} + g)$$

and

$$\frac{d^2(rf)}{dx^2} + (rf) = r(\frac{d^2f}{dx^2} + f)$$

of basic Calculus. This turns out to equal the space from the prior example — functions satisfying this differential equation have the form $a\cos\theta + b\sin\theta$ — but this description suggests an extension to solutions sets of other differential equations.

**1.15 Example** The set of solutions of a homogeneous linear system in $n$ variables is a vector space under the operations inherited from $\mathbb{R}^n$. For closure under addition, if

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \qquad w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

both satisfy the condition that their entries add to 0 then $v + w$ also satisfies that condition: $c_1(v_1 + w_1) + \cdots + c_n(v_n + w_n) = (c_1v_1 + \cdots + c_nv_n) + (c_1w_1 + \cdots + c_nw_n) = 0$. The checks of the other conditions are just as routine.

As we've done in those equations, we often omit the multiplication symbol '$\cdot$'. We can distinguish the multiplication in '$c_1v_1$' from that in '$rv$' since if both multiplicands are real numbers then real-real multiplication must be meant, while if one is a vector then scalar-vector multiplication must be meant.

The prior example has brought us full circle since it is one of our motivating examples.

**1.16 Remark** Now, with some feel for the kinds of structures that satisfy the definition of a vector space, we can reflect on that definition. For example, why specify in the definition the condition that $1 \cdot v = v$ but not a condition that $0 \cdot v = 0$?

One answer is that this is just a definition — it gives the rules of the game from here on, and if you don't like it, put the book down and walk away.

Another answer is perhaps more satisfying. People in this area have worked hard to develop the right balance of power and generality. This definition has been shaped so that it contains the conditions needed to prove all of the interesting and important properties of spaces of linear combinations. As we proceed, we shall derive all of the properties natural to collections of linear combinations from the conditions given in the definition.

The next result is an example. We do not need to include these properties in the definition of vector space because they follow from the properties already listed there.

**1.17 Lemma** In any vector space $V$, for any $v \in V$ and $r \in \mathbb{R}$, we have (1) $0 \cdot v = 0$, and (2) $(-1 \cdot v) + v = 0$, and (3) $r \cdot 0 = 0$.

Proof. For (1), note that $v = (1 + 0) \cdot v = v + (0 \cdot v)$. Add to both sides the additive inverse of $v$, the vector $w$ such that $w + v = 0$.

$$\begin{aligned} w + v &= w + v + 0 \cdot v \\ 0 &= 0 + 0 \cdot v \\ 0 &= 0 \cdot v \end{aligned}$$

The second item is easy: $(-1 \cdot v) + v = (-1 + 1) \cdot v = 0 \cdot v = 0$ shows that we can write '$-v$' for the additive inverse of $v$ without worrying about possible confusion with $(-1) \cdot v$.

For (3), this $r \cdot 0 = r \cdot (0 \cdot 0) = (r \cdot 0) \cdot 0 = 0$ will do. QED

We finish with a recap.

Our study in Chapter One of Gaussian reduction led us to consider collections of linear combinations. So in this chapter we have defined a vector space to be a structure in which we can form such combinations, expressions of the form $c_1 \cdot v_1 + \cdots + c_n \cdot v_n$ (subject to simple conditions on the addition and scalar multiplication operations). In a phrase: vector spaces are the right context in which to study linearity.

Finally, a comment. From the fact that it forms a whole chapter, and especially because that chapter is the first one, a reader could come to think that the study of linear systems is our purpose. The truth is, we will not so much use vector spaces in the study of linear systems as we will instead have linear systems start us on the study of vector spaces. The wide variety of examples from this subsection shows that the study of vector spaces is interesting and important in its own right, aside from how it helps us understand linear systems. Linear systems won't go away. But from now on our primary objects of study will be vector spaces.

## Exercises

**1.18** Give the zero vector from each of these vector spaces.

**(a)** The space of degree three polynomials under the natural operations

**(b)** The space of $2\times 4$ matrices

**(c)** The space $\{f\colon [0..1]\to\mathbb{R} \mid f \text{ is continuous}\}$

**(d)** The space of real-valued functions of one natural number variable

✓ **1.19** Find the additive inverse, in the vector space, of the vector.

**(a)** In $\mathcal{P}_3$, the vector $-3-2x+x^2$.

**(b)** In the space $2\times 2$,

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

**(c)** In $\{ae^x+be^{-x} \mid a,b\in\mathbb{R}\}$, the space of functions of the real variable $x$ under the natural operations, the vector $3e^x-2e^{-x}$.

✓ **1.20** Show that each of these is a vector space.

**(a)** The set of linear polynomials $\mathcal{P}_1=\{a_0+a_1x \mid a_0,a_1\in\mathbb{R}\}$ under the usual polynomial addition and scalar multiplication operations.

**(b)** The set of $2\times 2$ matrices with real entries under the usual matrix operations.

**(c)** The set of three-component row vectors with their usual operations.

**(d)** The set

$$L=\{\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}\in\mathbb{R}^4 \mid x+y-z+w=0\}$$

under the operations inherited from $\mathbb{R}^4$.

✓ **1.21** Show that each of these is not a vector space. (*Hint.* Start by listing two members of each set.)

**(a)** Under the operations inherited from $\mathbb{R}^3$, this set

$$\{\begin{pmatrix} x \\ y \\ z \end{pmatrix}\in\mathbb{R}^3 \mid x+y+z=1\}$$

**(b)** Under the operations inherited from $\mathbb{R}^3$, this set

$$\{\begin{pmatrix} x \\ y \\ z \end{pmatrix}\in\mathbb{R}^3 \mid x^2+y^2+z^2=1\}$$

**(c)** Under the usual matrix operations,

$$\{\begin{pmatrix} a & 1 \\ b & c \end{pmatrix} \mid a,b,c\in\mathbb{R}\}$$

**(d)** Under the usual polynomial operations,

$$\{a_0+a_1x+a_2x^2 \mid a_0,a_1,a_2\in\mathbb{R}^+\}$$

where $\mathbb{R}^+$ is the set of reals greater than zero

**(e)** Under the inherited operations,

$$\{\begin{pmatrix} x \\ y \end{pmatrix}\in\mathbb{R}^2 \mid x+3y=4 \text{ and } 2x-y=3 \text{ and } 6x+4y=10\}$$

**1.22** Define addition and scalar multiplication operations to make the complex numbers a vector space over $\mathbb{R}$.

✓ **1.23** Is the set of rational numbers a vector space over $\mathbb{R}$ under the usual addition and scalar multiplication operations?

**1.24** Show that the set of linear combinations of the variables $x,y,z$ is a vector space under the natural addition and scalar multiplication operations.

**1.25** Prove that this is not a vector space: the set of two-tall column vectors with real entries subject to these operations.

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}+\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}=\begin{pmatrix} x_1-x_2 \\ y_1-y_2 \end{pmatrix} \qquad r\cdot\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} rx \\ ry \end{pmatrix}$$

**1.26** Prove or disprove that $\mathbb{R}^3$ is a vector space under these operations.

**(a)** $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}+\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and $r\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} rx \\ ry \\ rz \end{pmatrix}$

**(b)** $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and $r \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

✓ **1.27** For each, decide if it is a vector space; the intended operations are the natural ones.

**(a)** The *diagonal* 2×2 matrices

$$\{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mid a, b \in \mathbb{R}\}$$

**(b)** This set of 2×2 matrices

$$\{\begin{pmatrix} x & x+y \\ x+y & y \end{pmatrix} \mid x, y \in \mathbb{R}\}$$

**(c)** This set

$$\{\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \in \mathbb{R}^4 \mid x + y + w = 1\}$$

**(d)** The set of functions $\{f\colon \mathbb{R} \to \mathbb{R} \mid df/dx + 2f = 0\}$

**(e)** The set of functions $\{f\colon \mathbb{R} \to \mathbb{R} \mid df/dx + 2f = 1\}$

✓ **1.28** Prove or disprove that this is a vector space: the real-valued functions $f$ of one real variable such that $f(7) = 0$.

✓ **1.29** Show that the set $\mathbb{R}^+$ of positive reals is a vector space when '$x + y$' is interpreted to mean the product of $x$ and $y$ (so that $2 + 3$ is 6), and '$r \cdot x$' is interpreted as the $r$-th power of $x$.

**1.30** Is $\{(x, y) \mid x, y \in \mathbb{R}\}$ a vector space under these operations?

**(a)** $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $r \cdot (x, y) = (rx, y)$

**(b)** $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $r \cdot (x, y) = (rx, 0)$

**1.31** Prove or disprove that this is a vector space: the set of polynomials of degree greater than or equal to two, along with the zero polynomial.

**1.32** At this point "the same" is only an intuition, but nonetheless for each vector space identify the $k$ for which the space is "the same" as $\mathbb{R}^k$.

**(a)** The 2×3 matrices under the usual operations

**(b)** The $n \times m$ matrices (under their usual operations)

**(c)** This set of 2×2 matrices

$$\{\begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \mid a, b, c \in \mathbb{R}\}$$

**(d)** This set of 2×2 matrices

$$\{\begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \mid a + b + c = 0\}$$

✓ **1.33** Using + to represent vector addition and $\cdot$ for scalar multiplication, restate the definition of vector space.

✓ **1.34** Prove these.

**(a)** Any vector is the additive inverse of the additive inverse of itself.

**(b)** Vector addition left-cancels: if $v, s, t \in V$ then $v + s = v + t$ implies that $s = t$.

**1.35** The definition of vector spaces does not explicitly say that $0 + v = v$ (it instead says that $v + 0 = v$). Show that it must nonetheless hold in any vector space.

✓ **1.36** Prove or disprove that this is a vector space: the set of all matrices, under the usual operations.

**1.37** In a vector space every element has an additive inverse. Can some elements have two or more?

**1.38** **(a)** Prove that every point, line, or plane thru the origin in $\mathbb{R}^3$ is a vector space under the inherited operations.

**(b)** What if it doesn't contain the origin?

✓ **1.39** Using the idea of a vector space we can easily reprove that the solution set of a homogeneous linear system has either one element or infinitely many elements. Assume that $v \in V$ is not 0.

**(a)** Prove that $r \cdot v = 0$ if and only if $r = 0$.

**(b)** Prove that $r_1 \cdot v = r_2 \cdot v$ if and only if $r_1 = r_2$.

**(c)** Prove that any nontrivial vector space is infinite.

**(d)** Use the fact that a nonempty solution set of a homogeneous linear system is a vector space to draw the conclusion.

**1.40** Is this a vector space under the natural operations: the real-valued functions of one real variable that are differentiable?

**1.41** A *vector space over the complex numbers* $\mathbb{C}$ has the same definition as a vector space over the reals except that scalars are drawn from $\mathbb{C}$ instead of from $\mathbb{R}$. Show that each of these is a vector space over the complex numbers. (Recall how complex numbers add and multiply: $(a_0 + a_1 i) + (b_0 + b_1 i) = (a_0 + b_0) + (a_1 + b_1)i$ and $(a_0 + a_1 i)(b_0 + b_1 i) = (a_0 b_0 - a_1 b_1) + (a_0 b_1 + a_1 b_0)i$.)

**(a)** The set of degree two polynomials with complex coefficients

**(b)** This set

$$\{\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \bigm| a, b \in \mathbb{C} \text{ and } a + b = 0 + 0i\}$$

**1.42** Find a property shared by all of the $\mathbb{R}^n$'s not listed as a requirement for a vector space.

✓ **1.43** **(a)** Prove that a sum of four vectors $v_1, \ldots, v_4 \in V$ can be associated in any way without changing the result.

$$\begin{aligned}((v_1 + v_2) + v_3) + v_4 &= (v_1 + (v_2 + v_3)) + v_4 \\ &= (v_1 + v_2) + (v_3 + v_4) \\ &= v_1 + ((v_2 + v_3) + v_4) \\ &= v_1 + (v_2 + (v_3 + v_4))\end{aligned}$$

This allows us to simply write '$v_1 + v_2 + v_3 + v_4$' without ambiguity.

**(b)** Prove that any two ways of associating a sum of any number of vectors give the same sum. (*Hint.* Use induction on the number of vectors.)

**1.44** For any vector space, a subset that is itself a vector space under the inherited operations (e.g., a plane through the origin inside of $\mathbb{R}^3$) is a *subspace*.

**(a)** Show that $\{a_0 + a_1 x + a_2 x^2 \bigm| a_0 + a_1 + a_2 = 0\}$ is a subspace of the vector space of degree two polynomials.

**(b)** Show that this is a subspace of the $2 \times 2$ matrices.

$$\{\begin{pmatrix} a & b \\ c & 0 \end{pmatrix} \bigm| a + b = 0\}$$

**(c)** Show that a nonempty subset $S$ of a real vector space is a subspace if and only if it is closed under linear combinations of pairs of vectors: whenever $c_1, c_2 \in \mathbb{R}$ and $s_1, s_2 \in S$ then the combination $c_1 v_1 + c_2 v_2$ is in $S$.

## I.2 Subspaces and Spanning Sets

One of the examples that led us to introduce the idea of a vector space was the solution set of a homogeneous system. For instance, we've seen in Example 1.4 such a space that is a planar subset of $\mathbb{R}^3$. There, the vector space $\mathbb{R}^3$ contains inside it another vector space, the plane.

**2.1 Definition** For any vector space, a *subspace* is a subset that is itself a vector space, under the inherited operations.

**2.2 Example** The plane from the prior subsection,

$$P = \{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \bigm| x + y + z = 0\}$$

is a subspace of $\mathbb{R}^3$. As specified in the definition, the operations are the ones that are inherited from the larger space, that is, vectors add in $P$ as they add in $\mathbb{R}^3$

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}$$

and scalar multiplication is also the same as it is in $\mathbb{R}^3$. To show that $P$ is a subspace, we need only note that it is a subset and then verify that it is a space. Checking that $P$ satisfies the conditions in the definition of a vector space is routine. For instance, for closure under addition, just note that if the summands satisfy that $x_1 + y_1 + z_1 = 0$ and $x_2 + y_2 + z_2 = 0$ then the sum satisfies that $(x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) = (x_1 + y_1 + z_1) + (x_2 + y_2 + z_2) = 0$.

**2.3 Example** The $x$-axis in $\mathbb{R}^2$ is a subspace where the addition and scalar multiplication operations are the inherited ones.

$$\begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ 0 \end{pmatrix} \qquad r \cdot \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} rx \\ 0 \end{pmatrix}$$

As above, to verify that this is a subspace, we simply note that it is a subset and then check that it satisfies the conditions in definition of a vector space. For instance, the two closure conditions are satisfied: (1) adding two vectors with a second component of zero results in a vector with a second component of zero, and (2) multiplying a scalar times a vector with a second component of zero results in a vector with a second component of zero.

**2.4 Example** Another subspace of $\mathbb{R}^2$ is

$$\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\}$$

its trivial subspace.

Any vector space has a trivial subspace $\{\vec{0}\,\}$. At the opposite extreme, any vector space has itself for a subspace. These two are the *improper* subspaces. Other subspaces are *proper*.

**2.5 Example** The condition in the definition requiring that the addition and scalar multiplication operations must be the ones inherited from the larger space is important. Consider the subset $\{1\}$ of the vector space $\mathbb{R}^1$. Under the operations $1+1=1$ and $r \cdot 1 = 1$ that set is a vector space, specifically, a trivial space. But it is not a subspace of $\mathbb{R}^1$ because those aren't the inherited operations, since of course $\mathbb{R}^1$ has $1+1=2$.

**2.6 Example** All kinds of vector spaces, not just $\mathbb{R}^n$'s, have subspaces. The vector space of cubic polynomials $\{a + bx + cx^2 + dx^3 \mid a, b, c, d \in \mathbb{R}\}$ has a subspace comprised of all linear polynomials $\{m + nx \mid m, n \in \mathbb{R}\}$.

**2.7 Example** Another example of a subspace not taken from an $\mathbb{R}^n$ is one from the examples following the definition of a vector space. The space of all real-valued functions of one real variable $f\colon \mathbb{R} \to \mathbb{R}$ has a subspace of functions satisfying the restriction $(d^2 f/dx^2) + f = 0$.

**2.8 Example** Being vector spaces themselves, subspaces must satisfy the closure conditions. The set $\mathbb{R}^+$ is not a subspace of the vector space $\mathbb{R}^1$ because with the inherited operations it is not closed under scalar multiplication: if $\vec{v} = 1$ then $-1 \cdot \vec{v} \notin \mathbb{R}^+$.

The next result says that Example 2.8 is prototypical. The only way that a subset can fail to be a subspace (if it is nonempty and the inherited operations are used) is if it isn't closed.

**2.9 Lemma** For a nonempty subset $S$ of a vector space, under the inherited operations, the following are equivalent statements.*

(1) $S$ is a subspace of that vector space

(2) $S$ is closed under linear combinations of pairs of vectors: for any vectors $\vec{s}_1, \vec{s}_2 \in S$ and scalars $r_1, r_2$ the vector $r_1\vec{s}_1 + r_2\vec{s}_2$ is in $S$

(3) $S$ is closed under linear combinations of any number of vectors: for any vectors $\vec{s}_1, \ldots, \vec{s}_n \in S$ and scalars $r_1, \ldots, r_n$ the vector $r_1\vec{s}_1 + \cdots + r_n\vec{s}_n$ is in $S$.

Briefly, the way that a subset gets to be a subspace is by being closed under linear combinations.

Proof. 'The following are equivalent' means that each pair of statements are equivalent.

$$(1) \iff (2) \qquad (2) \iff (3) \qquad (3) \iff (1)$$

*More information on equivalence of statements is in the appendix.

We will show this equivalence by establishing that (1) $\implies$ (3) $\implies$ (2) $\implies$ (1). This strategy is suggested by noticing that (1) $\implies$ (3) and (3) $\implies$ (2) are easy and so we need only argue the single implication (2) $\implies$ (1).

For that argument, assume that $S$ is a nonempty subset of a vector space $V$ and that $S$ is closed under combinations of pairs of vectors. We will show that $S$ is a vector space by checking the conditions.

The first item in the vector space definition has five conditions. First, for closure under addition, if $s_1, s_2 \in S$ then $s_1 + s_2 \in S$, as $s_1 + s_2 = 1 \cdot s_1 + 1 \cdot s_2$. Second, for any $s_1, s_2 \in S$, because addition is inherited from $V$, the sum $s_1 + s_2$ in $S$ equals the sum $s_1 + s_2$ in $V$, and that equals the sum $s_2 + s_1$ in $V$ (because $V$ is a vector space, its addition is commutative), and that in turn equals the sum $s_2 + s_1$ in $S$. The argument for the third condition is similar to that for the second. For the fourth, consider the zero vector of $V$ and note that closure of $S$ under linear combinations of pairs of vectors gives that (where $s$ is any member of the nonempty set $S$) $0 \cdot s + 0 \cdot s = 0$ is in $S$; showing that $0$ acts under the inherited operations as the additive identity of $S$ is easy. The fifth condition is satisfied because for any $s \in S$, closure under linear combinations shows that the vector $0 \cdot 0 + (-1) \cdot s$ is in $S$; showing that it is the additive inverse of $s$ under the inherited operations is routine.

The checks for item (2) are similar and are saved for Exercise 32. QED

We usually show that a subset is a subspace with (2) $\implies$ (1).

**2.10 Remark** At the start of this chapter we introduced vector spaces as collections in which linear combinations are "sensible". The above result speaks to this.

The vector space definition has ten conditions but eight of them — the conditions not about closure — simply ensure that referring to the operations as an 'addition' and a 'scalar multiplication' is sensible. The proof above checks that these eight are inherited from the surrounding vector space provided that the nonempty set $S$ satisfies Theorem 2.9's statement (2) (e.g., commutativity of addition in $S$ follows right from commutativity of addition in $V$). So, in this context, this meaning of "sensible" is automatically satisfied.

In assuring us that this first meaning of the word is met, the result draws our attention to the second meaning of "sensible". It has to do with the two remaining conditions, the closure conditions. Above, the two separate closure conditions inherent in statement (1) are combined in statement (2) into the single condition of closure under all linear combinations of two vectors, which is then extended in statement (3) to closure under combinations of any number of vectors. The latter two statements say that we can always make sense of an expression like $r_1 s_1 + r_2 s_2$, without restrictions on the $r$'s — such expressions are "sensible" in that the vector described is defined and is in the set $S$.

This second meaning suggests that a good way to think of a vector space is as a collection of unrestricted linear combinations. The next two examples take some spaces and describe them in this way. That is, in these examples we parametrize, just as we did in Chapter One to describe the solution set of a homogeneous linear system.

**2.11 Example** This subset of $\mathbb{R}^3$

$$S = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x - 2y + z = 0 \}$$

is a subspace under the usual addition and scalar multiplication operations of column vectors (the check that it is nonempty and closed under linear combinations of two vectors is just like the one in Example 2.2). To parametrize, we can take $x - 2y + z = 0$ to be a one-equation linear system and expressing the leading variable in terms of the free variables $x = 2y - z$.

$$S = \{ \begin{pmatrix} 2y - z \\ y \\ z \end{pmatrix} \mid y, z \in \mathbb{R} \} = \{ y \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid y, z \in \mathbb{R} \}$$

Now the subspace is described as the collection of unrestricted linear combinations of those two vectors. Of course, in either description, this is a plane through the origin.

**2.12 Example** This is a subspace of the $2\times 2$ matrices

$$L = \{\begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \bigm| a+b+c=0\}$$

(checking that it is nonempty and closed under linear combinations is easy). To parametrize, express the condition as $a=-b-c$.

$$L = \{\begin{pmatrix} -b-c & 0 \\ b & c \end{pmatrix} \bigm| b,c\in\mathbb{R}\} = \{b\begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} + c\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \bigm| b,c\in\mathbb{R}\}$$

As above, we've described the subspace as a collection of unrestricted linear combinations (by coincidence, also of two elements).

Parametrization is an easy technique, but it is important. We shall use it often.

**2.13 Definition** The *span* (or *linear closure*) of a nonempty subset $S$ of a vector space is the set of all linear combinations of vectors from $S$.

$$\langle S\rangle = \{c_1s_1+\cdots+c_ns_n \bigm| c_1,\ldots,c_n\in\mathbb{R} \text{ and } s_1,\ldots,s_n\in S\}$$

The span of the empty subset of a vector space is the trivial subspace.

No notation for the span is completely standard. The square brackets used here are common, but so are 'span($S$)' and 'sp($S$)'.

**2.14 Remark** In Chapter One, after we showed that the solution set of a homogeneous linear system can be written as $\{c_1\beta_1+\cdots+c_k\beta_k \bigm| c_1,\ldots,c_k\in\mathbb{R}\}$, we described that as the set 'generated' by the $\beta$'s. We now have the technical term; we call that the 'span' of the set $\{\beta_1,\ldots,\beta_k\}$.

Recall also the discussion of the "tricky point" in that proof. The span of the empty set is defined to be the set $\{0\}$ because we follow the convention that a linear combination of no vectors sums to $0$. Besides, defining the empty set's span to be the trivial subspace is a convienence in that it keeps results like the next one from having annoying exceptional cases.

**2.15 Lemma** In a vector space, the span of any subset is a subspace.

PROOF. Call the subset $S$. If $S$ is empty then by definition its span is the trivial subspace. If $S$ is not empty then by Lemma 2.9 we need only check that the span $\langle S\rangle$ is closed under linear combinations. For a pair of vectors from that span, $v=c_1s_1+\cdots+c_ns_n$ and $w=c_{n+1}s_{n+1}+\cdots+c_ms_m$, a linear combination

$$p\cdot(c_1s_1+\cdots+c_ns_n)+r\cdot(c_{n+1}s_{n+1}+\cdots+c_ms_m) \\ = pc_1s_1+\cdots+pc_ns_n+rc_{n+1}s_{n+1}+\cdots+rc_ms_m$$

($p$, $r$ scalars) is a linear combination of elements of $S$ and so is in $\langle S\rangle$ (possibly some of the $s_i$'s forming $v$ equal some of the $s_j$'s from $w$, but it does not matter). QED

The converse of the lemma holds: any subspace is the span of some set, because a subspace is obviously the span of the set of its members. Thus a subset of a vector space is a subspace if and only if it is a span. This fits the intuition that a good way to think of a vector space is as a collection in which linear combinations are sensible.

Taken together, Lemma 2.9 and Lemma 2.15 show that the span of a subset $S$ of a vector space is the smallest subspace containing all the members of $S$.

**2.16 Example** In any vector space $V$, for any vector $v$, the set $\{r\cdot v \bigm| r\in\mathbb{R}\}$ is a subspace of $V$. For instance, for any vector $v\in\mathbb{R}^3$, the line through the origin containing that vector, $\{kv \bigm| k\in\mathbb{R}\}$ is a subspace of $\mathbb{R}^3$. This is true even when $v$ is the zero vector, in which case the subspace is the degenerate line, the trivial subspace.

**2.17 Example** The span of this set is all of $\mathbb{R}^2$.

$$\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\}$$

To check this we must show that any member of $\mathbb{R}^2$ is a linear combination of these two vectors. So we ask: for which vectors (with real components $x$ and $y$) are there scalars $c_1$ and $c_2$ such that this holds?

$$c_1\begin{pmatrix}1\\1\end{pmatrix}+c_2\begin{pmatrix}1\\-1\end{pmatrix}=\begin{pmatrix}x\\y\end{pmatrix}$$

Gauss' method

$$\begin{array}{rl} c_1+c_2=x \\ c_1-c_2=y \end{array} \xrightarrow{-\rho_1+\rho_2} \begin{array}{rl} c_1+\quad c_2= & x \\ -2c_2= & -x+y \end{array}$$

with back substitution gives $c_2=(x-y)/2$ and $c_1=(x+y)/2$. These two equations show that for any $x$ and $y$ that we start with, there are appropriate coefficients $c_1$ and $c_2$ making the above vector equation true. For instance, for $x=1$ and $y=2$ the coefficients $c_2=-1/2$ and $c_1=3/2$ will do. That is, any vector in $\mathbb{R}^2$ can be written as a linear combination of the two given vectors.

Since spans are subspaces, and we know that a good way to understand a subspace is to parametrize its description, we can try to understand a set's span in that way.

**2.18 Example** Consider, in $\mathcal{P}_2$, the span of the set $\{3x-x^2,2x\}$. By the definition of span, it is the set of unrestricted linear combinations of the two $\{c_1(3x-x^2)+c_2(2x)\mid c_1,c_2\in\mathbb{R}\}$. Clearly polynomials in this span must have a constant term of zero. Is that necessary condition also sufficient?

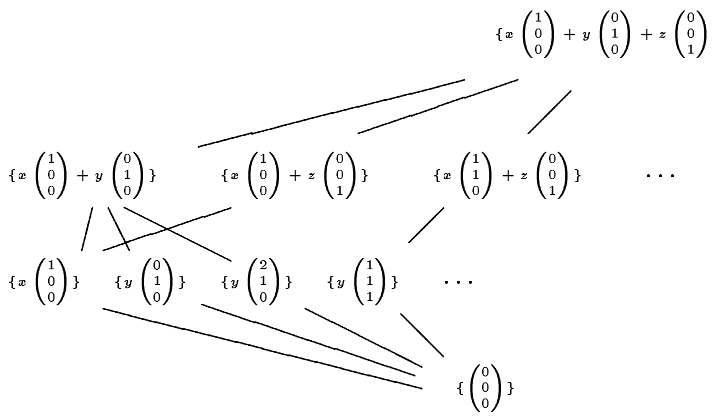
We are asking: for which members $a_2x^2+a_1x+a_0$ of $\mathcal{P}_2$ are there $c_1$ and $c_2$ such that $a_2x^2+a_1x+a_0=c_1(3x-x^2)+c_2(2x)$? Since polynomials are equal if and only if their coefficients are equal, we are looking for conditions on $a_2$, $a_1$, and $a_0$ satisfying these.

$$\begin{array}{rl} -c_1 \qquad & =a_2 \\ 3c_1+2c_2 & =a_1 \\ 0 & =a_0 \end{array}$$

Gauss' method gives that $c_1=-a_2$, $c_2=(3/2)a_2+(1/2)a_1$, and $0=a_0$. Thus the only condition on polynomials in the span is the condition that we knew of — as long as $a_0=0$, we can give appropriate coefficients $c_1$ and $c_2$ to describe the polynomial $a_0+a_1x+a_2x^2$ as in the span. For instance, for the polynomial $0-4x+3x^2$, the coefficients $c_1=-3$ and $c_2=5/2$ will do. So the span of the given set is $\{a_1x+a_2x^2\mid a_1,a_2\in\mathbb{R}\}$.

This shows, incidentally, that the set $\{x,x^2\}$ also spans this subspace. A space can have more than one spanning set. Two other sets spanning this subspace are $\{x,x^2,-x+2x^2\}$ and $\{x,x+x^2,x+2x^2,\dots\}$. (Naturally, we usually prefer to work with spanning sets that have only a few members.)

**2.19 Example** These are the subspaces of $\mathbb{R}^3$ that we now know of, the trivial subspace, the lines through the origin, the planes through the origin, and the whole space (of course, the picture shows only a few of the infinitely many subspaces). In the next section we will prove that $\mathbb{R}^3$ has no other type of subspaces, so in fact this picture shows them all.




The subsets are described as spans of sets, using a minimal number of members, and are shown connected to their supersets. Note that these subspaces fall naturally into levels — planes on one level, lines on another, etc. — according to how many vectors are in a minimal-sized spanning set.

So far in this chapter we have seen that to study the properties of linear combinations, the right setting is a collection that is closed under these combinations. In the first subsection we introduced such collections, vector spaces, and we saw a great variety of examples. In this subsection we saw still more spaces, ones that happen to be subspaces of others. In all of the variety we've seen a commonality. Example 2.19 above brings it out: vector spaces and subspaces are best understood as a span, and especially as a span of a small number of vectors. The next section studies spanning sets that are minimal.

**Exercises**

✓ **2.20** Which of these subsets of the vector space of 2×2 matrices are subspaces under the inherited operations? For each one that is a subspace, parametrize its description. For each that is not, give a condition that fails.

**(a)** $\{\begin{pmatrix}a&0\\0&b\end{pmatrix}\bigm| a,b\in\mathbb{R}\}$

**(b)** $\{\begin{pmatrix}a&0\\0&b\end{pmatrix}\bigm| a+b=0\}$

**(c)** $\{\begin{pmatrix}a&0\\0&b\end{pmatrix}\bigm| a+b=5\}$

**(d)** $\{\begin{pmatrix}a&c\\0&b\end{pmatrix}\bigm| a+b=0, c\in\mathbb{R}\}$

✓ **2.21** Is this a subspace of $\mathcal{P}_2$: $\{a_0+a_1x+a_2x^2 \bigm| a_0+2a_1+a_2=4\}$? If it is then parametrize its description.

✓ **2.22** Decide if the vector lies in the span of the set, inside of the space.

**(a)** $\begin{pmatrix}2\\0\\1\end{pmatrix}$, $\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\}$, in $\mathbb{R}^3$

**(b)** $x-x^3$, $\{x^2, 2x+x^2, x+x^3\}$, in $\mathcal{P}_3$

**(c)** $\begin{pmatrix}0&1\\4&2\end{pmatrix}$, $\{\begin{pmatrix}1&0\\1&1\end{pmatrix},\begin{pmatrix}2&0\\2&3\end{pmatrix}\}$, in $\mathcal{M}_{2\times 2}$

**2.23** Which of these are members of the span $\langle\{\cos^2 x, \sin^2 x\}\rangle$ in the vector space of real-valued functions of one real variable?

**(a)** $f(x)=1$ **(b)** $f(x)=3+x^2$ **(c)** $f(x)=\sin x$ **(d)** $f(x)=\cos(2x)$

✓ **2.24** Which of these sets spans $\mathbb{R}^3$? That is, which of these sets has the property that any three-tall vector can be expressed as a suitable linear combination of the set's elements?

**(a)** $\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\2\\0\end{pmatrix},\begin{pmatrix}0\\0\\3\end{pmatrix}\}$ **(b)** $\{\begin{pmatrix}2\\0\\1\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\}$ **(c)** $\{\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}3\\0\\0\end{pmatrix}\}$ **(d)** $\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}3\\1\\0\end{pmatrix},\begin{pmatrix}-1\\0\\0\end{pmatrix},\begin{pmatrix}2\\1\\5\end{pmatrix}\}$

**(e)** $\{\begin{pmatrix}2\\1\\1\end{pmatrix},\begin{pmatrix}3\\0\\1\end{pmatrix},\begin{pmatrix}5\\1\\2\end{pmatrix},\begin{pmatrix}6\\0\\2\end{pmatrix}\}$

✓ **2.25** Parametrize each subspace's description. Then express each subspace as a span.

**(a)** The subset $\{\begin{pmatrix}a & b & c\end{pmatrix} \mid a-c=0\}$ of the three-wide row vectors

**(b)** This subset of $\mathcal{M}_{2\times 2}$

$$\{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \mid a+d=0\}$$

**(c)** This subset of $\mathcal{M}_{2\times 2}$

$$\{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \mid 2a-c-d=0 \text{ and } a+3b=0\}$$

**(d)** The subset $\{a+bx+cx^3 \mid a-2b+c=0\}$ of $\mathcal{P}_3$

**(e)** The subset of $\mathcal{P}_2$ of quadratic polynomials $p$ such that $p(7)=0$

✓ **2.26** Find a set to span the given subspace of the given space. (*Hint.* Parametrize each.)

**(a)** the $xz$-plane in $\mathbb{R}^3$

**(b)** $\{\begin{pmatrix}x\\ y\\ z\end{pmatrix} \mid 3x+2y+z=0\}$ in $\mathbb{R}^3$

**(c)** $\{\begin{pmatrix}x\\ y\\ z\\ w\end{pmatrix} \mid 2x+y+w=0 \text{ and } y+2z=0\}$ in $\mathbb{R}^4$

**(d)** $\{a_0+a_1x+a_2x^2+a_3x^3 \mid a_0+a_1=0 \text{ and } a_2-a_3=0\}$ in $\mathcal{P}_3$

**(e)** The set $\mathcal{P}_4$ in the space $\mathcal{P}_4$

**(f)** $\mathcal{M}_{2\times 2}$ in $\mathcal{M}_{2\times 2}$

**2.27** Is $\mathbb{R}^2$ a subspace of $\mathbb{R}^3$?

✓ **2.28** Decide if each is a subspace of the vector space of real-valued functions of one real variable.

**(a)** The *even* functions $\{f\colon \mathbb{R}\to\mathbb{R} \mid f(-x)=f(x) \text{ for all } x\}$. For example, two members of this set are $f_1(x)=x^2$ and $f_2(x)=\cos(x)$.

**(b)** The *odd* functions $\{f\colon \mathbb{R}\to\mathbb{R} \mid f(-x)=-f(x) \text{ for all } x\}$. Two members are $f_3(x)=x^3$ and $f_4(x)=\sin(x)$.

**2.29** Example 2.16 says that for any vector $v$ that is an element of a vector space $V$, the set $\{r\cdot v \mid r\in\mathbb{R}\}$ is a subspace of $V$. (This is of course, simply the span of the singleton set $\{v\}$.) Must any such subspace be a proper subspace, or can it be improper?

**2.30** An example following the definition of a vector space shows that the solution set of a homogeneous linear system is a vector space. In the terminology of this subsection, it is a subspace of $\mathbb{R}^n$ where the system has $n$ variables. What about a non-homogeneous linear system; do its solutions form a subspace (under the inherited operations)?

**2.31** Example 2.19 shows that $\mathbb{R}^3$ has infinitely many subspaces. Does every nontrivial space have infinitely many subspaces?

**2.32** Finish the proof of Lemma 2.9.

**2.33** Show that each vector space has only one trivial subspace.

✓ **2.34** Show that for any subset $S$ of a vector space, the span of the span equals the span $\langle\langle S\rangle\rangle=\langle S\rangle$. (*Hint.* Members of $\langle S\rangle$ are linear combinations of members of $S$. Members of $\langle\langle S\rangle\rangle$ are linear combinations of linear combinations of members of $S$.)

**2.35** All of the subspaces that we've seen use zero in their description in some way. For example, the subspace in Example 2.3 consists of all the vectors from $\mathbb{R}^2$ with a second component of zero. In contrast, the collection of vectors from $\mathbb{R}^2$ with a second component of one does not form a subspace (it is not closed under scalar multiplication). Another example is Example 2.2, where the condition on the vectors is that the three components add to zero. If the condition were that the three components add to one then it would not be a subspace (again, it would fail to be closed). This exercise shows that a reliance on zero is not strictly necessary. Consider the set

$$\{\begin{pmatrix}x\\ y\\ z\end{pmatrix} \mid x+y+z=1\}$$

under these operations.

$$\begin{pmatrix}x_1\\ y_1\\ z_1\end{pmatrix}+\begin{pmatrix}x_2\\ y_2\\ z_2\end{pmatrix}=\begin{pmatrix}x_1+x_2-1\\ y_1+y_2\\ z_1+z_2\end{pmatrix} \qquad r\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}rx-r+1\\ ry\\ rz\end{pmatrix}$$

**(a)** Show that it is not a subspace of $\mathbb{R}^3$. (*Hint.* See Example 2.5).

(b) Show that it is a vector space. Note that by the prior item, Lemma 2.9 can not apply.

(c) Show that any subspace of $\mathbb{R}^3$ must pass through the origin, and so any subspace of $\mathbb{R}^3$ must involve zero in its description. Does the converse hold? Does any subset of $\mathbb{R}^3$ that contains the origin become a subspace when given the inherited operations?

**2.36** We can give a justification for the convention that the sum of zero-many vectors equals the zero vector. Consider this sum of three vectors $\vec{v}_1 + \vec{v}_2 + \vec{v}_3$.

(a) What is the difference between this sum of three vectors and the sum of the first two of these three?

(b) What is the difference between the prior sum and the sum of just the first one vector?

(c) What should be the difference between the prior sum of one vector and the sum of no vectors?

(d) So what should be the definition of the sum of no vectors?

**2.37** Is a space determined by its subspaces? That is, if two vector spaces have the same subspaces, must the two be equal?

**2.38** (a) Give a set that is closed under scalar multiplication but not addition.

(b) Give a set closed under addition but not scalar multiplication.

(c) Give a set closed under neither.

**2.39** Show that the span of a set of vectors does not depend on the order in which the vectors are listed in that set.

**2.40** Which trivial subspace is the span of the empty set? Is it

$$\{\begin{pmatrix}0\\0\\0\end{pmatrix}\} \subseteq \mathbb{R}^3, \quad \text{or} \quad \{0+0x\} \subseteq \mathcal{P}_1,$$

or some other subspace?

**2.41** Show that if a vector is in the span of a set then adding that vector to the set won't make the span any bigger. Is that also 'only if'?

✓ **2.42** Subspaces are subsets and so we naturally consider how 'is a subspace of' interacts with the usual set operations.

(a) If $A, B$ are subspaces of a vector space, must $A \cap B$ be a subspace? Always? Sometimes? Never?

(b) Must $A \cup B$ be a subspace?

(c) If $A$ is a subspace, must its complement be a subspace?

(*Hint.* Try some test subspaces from Example 2.19.)

✓ **2.43** Does the span of a set depend on the enclosing space? That is, if $W$ is a subspace of $V$ and $S$ is a subset of $W$ (and so also a subset of $V$), might the span of $S$ in $W$ differ from the span of $S$ in $V$?

**2.44** Is the relation 'is a subspace of' transitive? That is, if $V$ is a subspace of $W$ and $W$ is a subspace of $X$, must $V$ be a subspace of $X$?

✓ **2.45** Because 'span of' is an operation on sets we naturally consider how it interacts with the usual set operations.

(a) If $S \subseteq T$ are subsets of a vector space, is $[S] \subseteq [T]$? Always? Sometimes? Never?

(b) If $S, T$ are subsets of a vector space, is $[S \cup T] = [S] \cup [T]$?

(c) If $S, T$ are subsets of a vector space, is $[S \cap T] = [S] \cap [T]$?

(d) Is the span of the complement equal to the complement of the span?

**2.46** Reprove Lemma 2.15 without doing the empty set separately.

**2.47** Find a structure that is closed under linear combinations, and yet is not a vector space. (*Remark.* This is a bit of a trick question.)

# II Linear Independence

The prior section shows that a vector space can be understood as an unrestricted linear combination of some of its elements — that is, as a span. For example, the space of linear polynomials $\{a + bx \mid a, b \in \mathbb{R}\}$ is spanned by the set $\{1, x\}$. The prior section also showed that a space can have many sets that span it. The space of linear polynomials is also spanned by $\{1, 2x\}$ and $\{1, x, 2x\}$.

At the end of that section we described some spanning sets as 'minimal', but we never precisely defined that word. We could take 'minimal' to mean one of two things. We could mean that a spanning set is minimal if it contains the smallest number of members of any set with the same span. With this meaning $\{1, x, 2x\}$ is not minimal because it has one member more than the other two. Or we could mean that a spanning set is minimal when it has no elements that can be removed without changing the span. Under this meaning $\{1, x, 2x\}$ is not minimal because removing the $2x$ and getting $\{1, x\}$ leaves the span unchanged.

The first sense of minimality appears to be a global requirement, in that to check if a spanning set is minimal we seemingly must look at all the spanning sets of a subspace and find one with the least number of elements. The second sense of minimality is local in that we need to look only at the set under discussion and consider the span with and without various elements. For instance, using the second sense, we could compare the span of $\{1, x, 2x\}$ with the span of $\{1, x\}$ and note that the $2x$ is a "repeat" in that its removal doesn't shrink the span.

In this section we will use the second sense of 'minimal spanning set' because of this technical convenience. However, the most important result of this book is that the two senses coincide; we will prove that in the section after this one.

## II.1 Definition and Examples

We first characterize when a vector can be removed from a set without changing the span of that set.

**1.1 Lemma** Where $S$ is a subset of a vector space $V$,

$$\langle S\rangle = \langle S \cup \{v\}\rangle \quad \text{if and only if} \quad v \in \langle S\rangle$$

for any $v \in V$.

Proof. The left to right implication is easy. If $\langle S\rangle = \langle S \cup \{v\}\rangle$ then, since $v \in \langle S \cup \{v\}\rangle$, the equality of the two sets gives that $v \in \langle S\rangle$.

For the right to left implication assume that $v \in \langle S\rangle$ to show that $\langle S\rangle = \langle S \cup \{v\}\rangle$ by mutual inclusion. The inclusion $\langle S\rangle \subseteq \langle S \cup \{v\}\rangle$ is obvious. For the other inclusion $\langle S\rangle \supseteq \langle S \cup \{v\}\rangle$, write an element of $\langle S \cup \{v\}\rangle$ as $d_0 v + d_1 s_1 + \cdots + d_m s_m$ and substitute $v$'s expansion as a linear combination of members of the same set $d_0(c_0 t_0 + \cdots + c_k t_k) + d_1 s_1 + \cdots + d_m s_m$. This is a linear combination of linear combinations and so distributing $d_0$ results in a linear combination of vectors from $S$. Hence each member of $\langle S \cup \{v\}\rangle$ is also a member of $\langle S\rangle$. QED

**1.2 Example** In $\mathbb{R}^3$, where

$$v_1 = \begin{pmatrix}1\\0\\0\end{pmatrix} \quad v_2 = \begin{pmatrix}0\\1\\0\end{pmatrix} \quad v_3 = \begin{pmatrix}2\\1\\0\end{pmatrix}$$

the spans $\langle\{v_1, v_2\}\rangle$ and $\langle\{v_1, v_2, v_3\}\rangle$ are equal since $v_3$ is in the span $\langle\{v_1, v_2\}\rangle$.

The lemma says that if we have a spanning set then we can remove a $v$ to get a new set $S$ with the same span if and only if $v$ is a linear combination of vectors from $S$. Thus, under the second sense described above, a spanning set is minimal if and only if it contains no vectors that are linear combinations of the others in that set. We have a term for this important property.

**1.3 Definition** A subset of a vector space is *linearly independent* if none of its elements is a linear combination of the others. Otherwise it is *linearly dependent*.

Here is an important observation: although this way of writing one vector as a combination of the others

$$s_0 = c_1 s_1 + c_2 s_2 + \cdots + c_n s_n$$

visually sets $s_0$ off from the other vectors, algebraically there is nothing special in that equation about $s_0$. For any $s_i$ with a coefficient $c_i$ that is nonzero, we can rewrite the relationship to set off $s_i$.

$$s_i = (1/c_i)s_0 + (-c_1/c_i)s_1 + \cdots + (-c_n/c_i)s_n$$

When we don't want to single out any vector by writing it alone on one side of the equation we will instead say that $s_0, s_1, \ldots, s_n$ are in a *linear relationship* and write the relationship with all of the vectors on the same side. The next result rephrases the linear independence definition in this style. It gives what is usually the easiest way to compute whether a finite set is dependent or independent.

**1.4 Lemma** A subset $S$ of a vector space is linearly independent if and only if for any distinct $s_1, \ldots, s_n \in S$ the only linear relationship among those vectors

$$c_1 s_1 + \cdots + c_n s_n = 0 \qquad c_1, \ldots, c_n \in \mathbb{R}$$

is the trivial one: $c_1 = 0, \ldots, c_n = 0$.

Proof. This is a direct consequence of the observation above.

If the set $S$ is linearly independent then no vector $s_i$ can be written as a linear combination of the other vectors from $S$ so there is no linear relationship where some of the $s$'s have nonzero coefficients. If $S$ is not linearly independent then some $s_i$ is a linear combination $s_i = c_1 s_1 + \cdots + c_{i-1}s_{i-1} + c_{i+1}s_{i+1} + \cdots + c_n s_n$ of other vectors from $S$, and subtracting $s_i$ from both sides of that equation gives a linear relationship involving a nonzero coefficient, namely the $-1$ in front of $s_i$. QED

**1.5 Example** In the vector space of two-wide row vectors, the two-element set $\{\begin{pmatrix}40 & 15\end{pmatrix}, \begin{pmatrix}-50 & 25\end{pmatrix}\}$ is linearly independent. To check this, set

$$c_1 \cdot \begin{pmatrix}40 & 15\end{pmatrix} + c_2 \cdot \begin{pmatrix}-50 & 25\end{pmatrix} = \begin{pmatrix}0 & 0\end{pmatrix}$$

and solving the resulting system

$$\begin{array}{rl} 40c_1 - 50c_2 = 0 \\ 15c_1 + 25c_2 = 0 \end{array} \xrightarrow{-(15/40)\rho_1+\rho_2} \begin{array}{rl} 40c_1 - \quad 50c_2 = 0 \\ (175/4)c_2 = 0 \end{array}$$

shows that both $c_1$ and $c_2$ are zero. So the only linear relationship between the two given row vectors is the trivial relationship.

In the same vector space, $\{\begin{pmatrix}40 & 15\end{pmatrix}, \begin{pmatrix}20 & 7.5\end{pmatrix}\}$ is linearly dependent since we can satisfy

$$c_1 \begin{pmatrix}40 & 15\end{pmatrix} + c_2 \cdot \begin{pmatrix}20 & 7.5\end{pmatrix} = \begin{pmatrix}0 & 0\end{pmatrix}$$

with $c_1 = 1$ and $c_2 = -2$.

**1.6 Remark** Recall the Statics example that began this book. We first set the unknown-mass objects at 40 cm and 15 cm and got a balance, and then we set the objects at $-50$ cm and 25 cm and got a balance. With those two pieces of information we could compute values of the unknown masses. Had we instead first set the unknown-mass objects at 40 cm and 15 cm, and then at 20 cm and 7.5 cm, we would not have been able to compute the values of the unknown masses (try it). Intuitively, the problem is that the $\begin{pmatrix}20 & 7.5\end{pmatrix}$ information is a "repeat" of the $\begin{pmatrix}40 & 15\end{pmatrix}$ information — that is, $\begin{pmatrix}20 & 7.5\end{pmatrix}$ is in the span of the set $\{\begin{pmatrix}40 & 15\end{pmatrix}\}$ — and so we would be trying to solve a two-unknowns problem with what is essentially one piece of information.

**1.7 Example** The set $\{1+x, 1-x\}$ is linearly independent in $\mathcal{P}_2$, the space of quadratic polynomials with real coefficients, because

$$0 + 0x + 0x^2 = c_1(1+x) + c_2(1-x) = (c_1+c_2) + (c_1-c_2)x + 0x^2$$

gives

$$\begin{array}{rl} c_1 + c_2 = 0 \\ c_1 - c_2 = 0 \end{array} \xrightarrow{-\rho_1+\rho_2} \begin{array}{rl} c_1 + c_2 = 0 \\ 2c_2 = 0 \end{array}$$

since polynomials are equal only if their coefficients are equal. Thus, the only linear relationship between these two members of $\mathcal{P}_2$ is the trivial one.

**1.8 Example** In $\mathbb{R}^3$, where

$$v_1 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \quad v_2 = \begin{pmatrix} 2 \\ 9 \\ 2 \end{pmatrix} \quad v_3 = \begin{pmatrix} 4 \\ 18 \\ 4 \end{pmatrix}$$

the set $S = \{v_1, v_2, v_3\}$ is linearly dependent because this is a relationship

$$0 \cdot v_1 + 2 \cdot v_2 - 1 \cdot v_3 = 0$$

where not all of the scalars are zero (the fact that some of the scalars are zero doesn't matter).

**1.9 Remark** That example illustrates why, although Definition 1.3 is a clearer statement of what independence is, Lemma 1.4 is more useful for computations. Working straight from the definition, someone trying to compute whether $S$ is linearly independent would start by setting $v_1 = c_2 v_2 + c_3 v_3$ and concluding that there are no such $c_2$ and $c_3$. But knowing that the first vector is not dependent on the other two is not enough. This person would have to go on to try $v_2 = c_1 v_1 + c_3 v_3$ to find the dependence $c_1 = 0$, $c_3 = 1/2$. Lemma 1.4 gets the same conclusion with only one computation.

**1.10 Example** The empty subset of a vector space is linearly independent. There is no nontrivial linear relationship among its members as it has no members.

**1.11 Example** In any vector space, any subset containing the zero vector is linearly dependent. For example, in the space $\mathcal{P}_2$ of quadratic polynomials, consider the subset $\{1+x, x+x^2, 0\}$.

One way to see that this subset is linearly dependent is to use Lemma 1.4: we have $0 \cdot v_1 + 0 \cdot v_2 + 1 \cdot 0 = 0$, and this is a nontrivial relationship as not all of the coefficients are zero. Another way to see that this subset is linearly dependent is to go straight to Definition 1.3: we can express the third member of the subset as a linear combination of the first two, namely, $c_1 v_1 + c_2 v_2 = 0$ is satisfied by taking $c_1 = 0$ and $c_2 = 0$ (in contrast to the lemma, the definition allows all of the coefficients to be zero).

(There is still another way to see that this subset is dependent that is subtler. The zero vector is equal to the trivial sum, that is, it is the sum of no vectors. So in a set containing the zero vector, there is an element that can be written as a combination of a collection of other vectors from the set, specifically, the zero vector can be written as a combination of the empty collection.)

The above examples, especially Example 1.5, underline the discussion that begins this section. The next result says that given a finite set, we can produce a linearly independent subset by discarding what Remark 1.6 calls "repeats".

**1.12 Theorem** In a vector space, any finite subset has a linearly independent subset with the same span.

Proof. If the set $S = \{s_1, \ldots, s_n\}$ is linearly independent then $S$ itself satisfies the statement, so assume that it is linearly dependent.

By the definition of dependence, there is a vector $s_i$ that is a linear combination of the others. Call that vector $v_1$. Discard it—define the set $S_1 = S - \{v_1\}$. By Lemma 1.1, the span does not shrink $\langle S_1 \rangle = \langle S \rangle$.

Now, if $S_1$ is linearly independent then we are finished. Otherwise iterate the prior paragraph: take a vector $v_2$ that is a linear combination of other members of $S_1$ and discard it to derive $S_2 = S_1 - \{v_2\}$ such that $\langle S_2 \rangle = \langle S_1 \rangle$. Repeat this until a linearly independent set $S_j$ appears; one must appear eventually because $S$ is finite and the empty set is linearly independent. (Formally, this argument uses induction on $n$, the number of elements in the starting set. Exercise 37 asks for the details.) QED

**1.13 Example** This set spans $\mathbb{R}^3$.

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

Looking for a linear relationship

$$c_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + c_2 \begin{pmatrix}0\\2\\0\end{pmatrix} + c_3 \begin{pmatrix}1\\2\\0\end{pmatrix} + c_4 \begin{pmatrix}0\\-1\\1\end{pmatrix} + c_5 \begin{pmatrix}3\\3\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$$

gives a three equations/five unknowns linear system whose solution set can be parametrized in this way.

$$\{ \begin{pmatrix}c_1\\c_2\\c_3\\c_4\\c_5\end{pmatrix} = c_3 \begin{pmatrix}-1\\-1\\1\\0\\0\end{pmatrix} + c_5 \begin{pmatrix}-3\\-3/2\\0\\0\\1\end{pmatrix} \bigm| c_3, c_5 \in \mathbb{R} \}$$

So $S$ is linearly dependent. Setting $c_3 = 0$ and $c_5 = 1$ shows that the fifth vector is a linear combination of the first two. Thus, Lemma 1.1 says that discarding the fifth vector

$$S_1 = \{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\2\\0\end{pmatrix}, \begin{pmatrix}1\\2\\0\end{pmatrix}, \begin{pmatrix}0\\-1\\1\end{pmatrix} \}$$

leaves the span unchanged $\langle S_1 \rangle = \langle S \rangle$. Now, the third vector of $S_1$ is a linear combination of the first two and we get

$$S_2 = \{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\2\\0\end{pmatrix}, \begin{pmatrix}0\\-1\\1\end{pmatrix} \}$$

with the same span as $S_1$, and therefore the same span as $S$, but with one difference. The set $S_2$ is linearly independent (this is easily checked), and so discarding any of its elements will shrink the span.

Theorem 1.12 describes producing a linearly independent set by shrinking, that is, by taking subsets. We finish this subsection by considering how linear independence and dependence, which are properties of sets, interact with the subset relation between sets.

**1.14 Lemma** Any subset of a linearly independent set is also linearly independent. Any superset of a linearly dependent set is also linearly dependent.

PROOF. This is clear. QED

Restated, independence is preserved by subset and dependence is preserved by superset.

Those are two of the four possible cases of interaction that we can consider. The third case, whether linear dependence is preserved by the subset operation, is covered by Example 1.13, which gives a linearly dependent set $S$ with a subset $S_1$ that is linearly dependent and another subset $S_2$ that is linearly independent.

That leaves one case, whether linear independence is preserved by superset. The next example shows what can happen.

**1.15 Example** In each of these three paragraphs the subset $S$ is linearly independent.

For the set

$$S = \{ \begin{pmatrix}1\\0\\0\end{pmatrix} \}$$

the span $\langle S \rangle$ is the $x$ axis. Here are two supersets of $S$, one linearly dependent and the other linearly independent.

$$\text{dependent: } \{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}-3\\0\\0\end{pmatrix} \} \qquad \text{independent: } \{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix} \}$$

Checking the dependence or independence of these sets is easy.

For

$$S = \{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix} \}$$

the span $\langle S\rangle$ is the $xy$ plane. These are two supersets.

$$\text{dependent: } \{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}3\\-2\\0\end{pmatrix} \} \qquad \text{independent: } \{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \}$$

If

$$S = \{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \}$$

then $\langle S\rangle = \mathbb{R}^3$. A linearly dependent superset is

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

but there are no linearly independent supersets of $S$. The reason is that for any vector that we would add to make a superset, the linear dependence equation

$$\begin{pmatrix}x\\y\\z\end{pmatrix} = c_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + c_2 \begin{pmatrix}0\\1\\0\end{pmatrix} + c_3 \begin{pmatrix}0\\0\\1\end{pmatrix}$$

has a solution $c_1 = x$, $c_2 = y$, and $c_3 = z$.

So, in general, a linearly independent set may have a superset that is dependent. And, in general, a linearly independent set may have a superset that is independent. We can characterize when the superset is one and when it is the other.

**1.16 Lemma** Where $S$ is a linearly independent subset of a vector space $V$,

$$S \cup \{v\} \text{ is linearly dependent} \quad \text{if and only if} \quad v \in \langle S\rangle$$

for any $v \in V$ with $v \notin S$.

Proof. One implication is clear: if $v \in \langle S\rangle$ then $v = c_1 s_1 + c_2 s_2 + \cdots + c_n s_n$ where each $s_i \in S$ and $c_i \in \mathbb{R}$, and so $0 = c_1 s_1 + c_2 s_2 + \cdots + c_n s_n + (-1)v$ is a nontrivial linear relationship among elements of $S \cup \{v\}$.

The other implication requires the assumption that $S$ is linearly independent. With $S \cup \{v\}$ linearly dependent, there is a nontrivial linear relationship $c_0 v + c_1 s_1 + c_2 s_2 + \cdots + c_n s_n = 0$ and independence of $S$ then implies that $c_0 \neq 0$, or else that would be a nontrivial relationship among members of $S$. Now rewriting this equation as $v = -(c_1/c_0)s_1 - \cdots - (c_n/c_0)s_n$ shows that $v \in \langle S\rangle$. QED

(Compare this result with Lemma 1.1. Both say, roughly, that $v$ is a "repeat" if it is in the span of $S$. However, note the additional hypothesis here of linear independence.)

**1.17 Corollary** A subset $S = \{s_1, \ldots, s_n\}$ of a vector space is linearly dependent if and only if some $s_i$ is a linear combination of the vectors $s_1$, ..., $s_{i-1}$ listed before it.

Proof. Consider $S_0 = \{\}$, $S_1 = \{s_1\}$, $S_2 = \{s_1, s_2\}$, etc. Some index $i \geq 1$ is the first one with $S_{i-1} \cup \{s_i\}$ linearly dependent, and there $s_i \in \langle S_{i-1}\rangle$. QED

Lemma 1.16 can be restated in terms of independence instead of dependence: if $S$ is linearly independent and $v \notin S$ then the set $S \cup \{v\}$ is also linearly independent if and only if $v \notin \langle S \rangle$. Applying Lemma 1.1, we conclude that if $S$ is linearly independent and $v \notin S$ then $S \cup \{v\}$ is also linearly independent if and only if $\langle S \cup \{v\} \rangle \neq \langle S \rangle$. Briefly, when passing from $S$ to a superset $S_1$, to preserve linear independence we must expand the span $\langle S_1 \rangle \supset \langle S \rangle$.

Example 1.15 shows that some linearly independent sets are maximal — have as many elements as possible — in that they have no supersets that are linearly independent. By the prior paragraph, a linearly independent sets is maximal if and only if it spans the entire space, because then no vector exists that is not already in the span.

This table summarizes the interaction between the properties of independence and dependence and the relations of subset and superset.

| | $S_1 \subset S$ | $S_1 \supset S$ |
|---|---|---|
| *$S$ independent* | $S_1$ must be independent | $S_1$ may be either |
| *$S$ dependent* | $S_1$ may be either | $S_1$ must be dependent |

In developing this table we've uncovered an intimate relationship between linear independence and span. Complementing the fact that a spanning set is minimal if and only if it is linearly independent, a linearly independent set is maximal if and only if it spans the space.

In summary, we have introduced the definition of linear independence to formalize the idea of the minimality of a spanning set. We have developed some properties of this idea. The most important is Lemma 1.16, which tells us that a linearly independent set is maximal when it spans the space.

## Exercises

✓ **1.18** Decide whether each subset of $\mathbb{R}^3$ is linearly dependent or linearly independent.

**(a)** $\{\begin{pmatrix}1\\-3\\5\end{pmatrix}, \begin{pmatrix}2\\2\\4\end{pmatrix}, \begin{pmatrix}4\\-4\\14\end{pmatrix}\}$

**(b)** $\{\begin{pmatrix}1\\7\\7\end{pmatrix}, \begin{pmatrix}2\\7\\7\end{pmatrix}, \begin{pmatrix}3\\7\\7\end{pmatrix}\}$

**(c)** $\{\begin{pmatrix}0\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\0\\4\end{pmatrix}\}$

**(d)** $\{\begin{pmatrix}9\\9\\0\end{pmatrix}, \begin{pmatrix}2\\0\\1\end{pmatrix}, \begin{pmatrix}3\\5\\-4\end{pmatrix}, \begin{pmatrix}12\\12\\-1\end{pmatrix}\}$

✓ **1.19** Which of these subsets of $\mathcal{P}_3$ are linearly dependent and which are independent?

**(a)** $\{3 - x + 9x^2, 5 - 6x + 3x^2, 1 + 1x - 5x^2\}$
**(b)** $\{-x^2, 1 + 4x^2\}$
**(c)** $\{2 + x + 7x^2, 3 - x + 2x^2, 4 - 3x^2\}$
**(d)** $\{8 + 3x + 3x^2, x + 2x^2, 2 + 2x + 2x^2, 8 - 2x + 5x^2\}$

✓ **1.20** Prove that each set $\{f, g\}$ is linearly independent in the vector space of all functions from $\mathbb{R}^+$ to $\mathbb{R}$.

**(a)** $f(x) = x$ and $g(x) = 1/x$
**(b)** $f(x) = \cos(x)$ and $g(x) = \sin(x)$
**(c)** $f(x) = e^x$ and $g(x) = \ln(x)$

✓ **1.21** Which of these subsets of the space of real-valued functions of one real variable is linearly dependent and which is linearly independent? (Note that we have abbreviated some constant functions; e.g., in the first item, the '2' stands for the constant function $f(x) = 2$.)

**(a)** $\{2, 4\sin^2(x), \cos^2(x)\}$ **(b)** $\{1, \sin(x), \sin(2x)\}$ **(c)** $\{x, \cos(x)\}$ **(d)** $\{(1+x)^2, x^2 + 2x, 3\}$
**(e)** $\{\cos(2x), \sin^2(x), \cos^2(x)\}$ **(f)** $\{0, x, x^2\}$

**1.22** Does the equation $\sin^2(x)/\cos^2(x) = \tan^2(x)$ show that this set of functions $\{\sin^2(x), \cos^2(x), \tan^2(x)\}$ is a linearly dependent subset of the set of all real-valued functions with domain the interval $(-\pi/2..\pi/2)$ of real numbers between $-\pi/2$ and $\pi/2$)?

**1.23** Why does Lemma 1.4 say "distinct"?

✓ **1.24** Show that the nonzero rows of an echelon form matrix form a linearly independent set.

✓ **1.25** **(a)** Show that if the set $\{u, v, w\}$ is linearly independent set then so is the set $\{u, u+v, u+v+w\}$.
**(b)** What is the relationship between the linear independence or dependence of the set $\{u, v, w\}$ and the independence or dependence of $\{u-v, v-w, w-u\}$?

**1.26** Example 1.10 shows that the empty set is linearly independent.
**(a)** When is a one-element set linearly independent?
**(b)** How about a set with two elements?

**1.27** In any vector space $V$, the empty set is linearly independent. What about all of $V$?

**1.28** Show that if $\{x, y, z\}$ is linearly independent then so are all of its proper subsets: $\{x, y\}$, $\{x, z\}$, $\{y, z\}$, $\{x\}$,$\{y\}$, $\{z\}$, and $\{\}$. Is that 'only if' also?

**1.29** **(a)** Show that this

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

is a linearly independent subset of $\mathbb{R}^3$.
**(b)** Show that

$$\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$$

is in the span of $S$ by finding $c_1$ and $c_2$ giving a linear relationship.

$$c_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$$

Show that the pair $c_1, c_2$ is unique.
**(c)** Assume that $S$ is a subset of a vector space and that $v$ is in $\langle S \rangle$, so that $v$ is a linear combination of vectors from $S$. Prove that if $S$ is linearly independent then a linear combination of vectors from $S$ adding to $v$ is unique (that is, unique up to reordering and adding or taking away terms of the form $0 \cdot s$). Thus $S$ as a spanning set is minimal in this strong sense: each vector in $\langle S \rangle$ is "hit" a minimum number of times — only once.
**(d)** Prove that it can happen when $S$ is not linearly independent that distinct linear combinations sum to the same vector.

**1.30** Prove that a polynomial gives rise to the zero function if and only if it is the zero polynomial. (*Comment.* This question is not a Linear Algebra matter, but we often use the result. A polynomial gives rise to a function in the obvious way: $x \mapsto c_n x^n + \cdots + c_1 x + c_0$.)

**1.31** Return to Section 1.2 and redefine point, line, plane, and other linear surfaces to avoid degenerate cases.

**1.32** **(a)** Show that any set of four vectors in $\mathbb{R}^2$ is linearly dependent.
**(b)** Is this true for any set of five? Any set of three?
**(c)** What is the most number of elements that a linearly independent subset of $\mathbb{R}^2$ can have?

✓ **1.33** Is there a set of four vectors in $\mathbb{R}^3$, any three of which form a linearly independent set?

**1.34** Must every linearly dependent set have a subset that is dependent and a subset that is independent?

**1.35** In $\mathbb{R}^4$, what is the biggest linearly independent set you can find? The smallest? The biggest linearly dependent set? The smallest? ('Biggest' and 'smallest' mean that there are no supersets or subsets with the same property.)

✓ **1.36** Linear independence and linear dependence are properties of sets. We can thus naturally ask how those properties act with respect to the familiar elementary set relations and operations. In this body of this subsection we have covered the subset and superset relations. We can also consider the operations of intersection, complementation, and union.
**(a)** How does linear independence relate to intersection: can an intersection of linearly independent sets be independent? Must it be?
**(b)** How does linear independence relate to complementation?
**(c)** Show that the union of two linearly independent sets need not be linearly independent.
**(d)** Characterize when the union of two linearly independent sets is linearly independent, in terms of the intersection of the span of each.

✓ **1.37** For Theorem 1.12,
**(a)** fill in the induction for the proof;

**(b)** give an alternate proof that starts with the empty set and builds a sequence of linearly independent subsets of the given finite set until one appears with the same span as the given set.

**1.38** With a little calculation we can get formulas to determine whether or not a set of vectors is linearly independent.

**(a)** Show that this subset of $\mathbb{R}^2$

$$\{\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\}$$

is linearly independent if and only if $ad - bc \neq 0$.

**(b)** Show that this subset of $\mathbb{R}^3$

$$\{\begin{pmatrix} a \\ d \\ g \end{pmatrix}, \begin{pmatrix} b \\ e \\ h \end{pmatrix}, \begin{pmatrix} c \\ f \\ i \end{pmatrix}\}$$

is linearly independent iff $aei + bfg + cdh - hfa - idb - gec \neq 0$.

**(c)** When is this subset of $\mathbb{R}^3$

$$\{\begin{pmatrix} a \\ d \\ g \end{pmatrix}, \begin{pmatrix} b \\ e \\ h \end{pmatrix}\}$$

linearly independent?

**(d)** This is an opinion question: for a set of four vectors from $\mathbb{R}^4$, must there be a formula involving the sixteen entries that determines independence of the set? (You needn't produce such a formula, just decide if one exists.)

✓ **1.39** **(a)** Prove that a set of two perpendicular nonzero vectors from $\mathbb{R}^n$ is linearly independent when $n > 1$.

**(b)** What if $n = 1$? $n = 0$?

**(c)** Generalize to more than two vectors.

**1.40** Consider the set of functions from the open interval $(-1..1)$ to $\mathbb{R}$.

**(a)** Show that this set is a vector space under the usual operations.

**(b)** Recall the formula for the sum of an infinite geometric series: $1 + x + x^2 + \cdots = 1/(1-x)$ for all $x \in (-1..1)$. Why does this not express a dependence inside of the set $\{g(x) = 1/(1-x), f_0(x) = 1, f_1(x) = x, f_2(x) = x^2, \ldots\}$ (in the vector space that we are considering)? (*Hint.* Review the definition of linear combination.)

**(c)** Show that the set in the prior item is linearly independent.

This shows that some vector spaces exist with linearly independent subsets that are infinite.

**1.41** Show that, where $S$ is a subspace of $V$, if a subset $T$ of $S$ is linearly independent in $S$ then $T$ is also linearly independent in $V$. Is that 'only if'?

# III Basis and Dimension

The prior section ends with the statement that a spanning set is minimal when it is linearly independent and a linearly independent set is maximal when it spans the space. So the notions of minimal spanning set and maximal independent set coincide. In this section we will name this idea and study its properties.

## III.1 Basis

**1.1 Definition** A *basis* for a vector space is a sequence of vectors that form a set that is linearly independent and that spans the space.

We denote a basis with angle brackets $\langle \vec{\beta}_1, \vec{\beta}_2, \ldots \rangle$ to signify that this collection is a sequence* — the order of the elements is significant. (The requirement that a basis be ordered will be needed, for instance, in Definition 1.13.)

**1.2 Example** This is a basis for $\mathbb{R}^2$.

$$\langle \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$$

It is linearly independent

$$c_1 \begin{pmatrix} 2 \\ 4 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad \begin{aligned} 2c_1 + 1c_2 &= 0 \\ 4c_1 + 1c_2 &= 0 \end{aligned} \quad \Longrightarrow \quad c_1 = c_2 = 0$$

and it spans $\mathbb{R}^2$.

$$\begin{aligned} 2c_1 + 1c_2 &= x \\ 4c_1 + 1c_2 &= y \end{aligned} \quad \Longrightarrow \quad c_2 = 2x - y \text{ and } c_1 = (y - x)/2$$

**1.3 Example** This basis for $\mathbb{R}^2$

$$\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix} \rangle$$

differs from the prior one because the vectors are in a different order. The verification that it is a basis is just as in the prior example.

**1.4 Example** The space $\mathbb{R}^2$ has many bases. Another one is this.

$$\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle$$

The verification is easy.

**1.5 Definition** For any $\mathbb{R}^n$,

$$\mathcal{E}_n = \langle \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \rangle$$

is the *standard* (or *natural*) basis. We denote these vectors by $\vec{e}_1, \ldots, \vec{e}_n$.

(Calculus books refer to $\mathbb{R}^2$'s standard basis vectors $\vec{\imath}$ and $\vec{\jmath}$ instead of $\vec{e}_1$ and $\vec{e}_2$, and they refer to $\mathbb{R}^3$'s standard basis vectors $\vec{\imath}$, $\vec{\jmath}$, and $\vec{k}$ instead of $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$.) Note that the symbol '$\vec{e}_1$' means something different in a discussion of $\mathbb{R}^3$ than it means in a discussion of $\mathbb{R}^2$.

*More information on sequences is in the appendix.

**1.6 Example** Consider the space $\{a\cdot\cos\theta+b\cdot\sin\theta \mid a,b\in\mathbb{R}\}$ of functions of the real variable $\theta$. This is a natural basis.

$$\langle 1\cdot\cos\theta+0\cdot\sin\theta, 0\cdot\cos\theta+1\cdot\sin\theta\rangle=\langle\cos\theta,\sin\theta\rangle$$

Another, more generic, basis is $\langle\cos\theta-\sin\theta, 2\cos\theta+3\sin\theta\rangle$. Verfication that these two are bases is Exercise 22.

**1.7 Example** A natural basis for the vector space of cubic polynomials $\mathcal{P}_3$ is $\langle 1,x,x^2,x^3\rangle$. Two other bases for this space are $\langle x^3,3x^2,6x,6\rangle$ and $\langle 1,1+x,1+x+x^2,1+x+x^2+x^3\rangle$. Checking that these are linearly independent and span the space is easy.

**1.8 Example** The trivial space $\{\vec{0}\}$ has only one basis, the empty one $\langle\rangle$.

**1.9 Example** The space of finite-degree polynomials has a basis with infinitely many elements $\langle 1,x,x^2,\ldots\rangle$.

**1.10 Example** We have seen bases before. In the first chapter we described the solution set of homogeneous systems such as this one

$$\begin{aligned} x+y\quad\; -w&=0\\ z+w&=0\end{aligned}$$

by parametrizing.

$$\{\begin{pmatrix}-1\\1\\0\\0\end{pmatrix}y+\begin{pmatrix}1\\0\\-1\\1\end{pmatrix}w \mid y,w\in\mathbb{R}\}$$

That is, we described the vector space of solutions as the span of a two-element set. We can easily check that this two-vector set is also linearly independent. Thus the solution set is a subspace of $\mathbb{R}^4$ with a two-element basis.

**1.11 Example** Parameterization helps find bases for other vector spaces, not just for solution sets of homogeneous systems. To find a basis for this subspace of $\mathcal{M}_{2\times2}$

$$\{\begin{pmatrix}a&b\\c&0\end{pmatrix} \mid a+b-2c=0\}$$

we rewrite the condition as $a=-b+2c$.

$$\{\begin{pmatrix}-b+2c&b\\c&0\end{pmatrix} \mid b,c\in\mathbb{R}\}=\{b\begin{pmatrix}-1&1\\0&0\end{pmatrix}+c\begin{pmatrix}2&0\\1&0\end{pmatrix} \mid b,c\in\mathbb{R}\}$$

Thus, this is a natural candidate for a basis.

$$\langle\begin{pmatrix}-1&1\\0&0\end{pmatrix},\begin{pmatrix}2&0\\1&0\end{pmatrix}\rangle$$

The above work shows that it spans the space. To show that it is linearly independent is routine.

Consider again Example 1.2. It involves two verifications.

In the first, to check that the set is linearly independent we looked at linear combinations of the set's members that total to the zero vector $c_1\vec{\beta}_1+c_2\vec{\beta}_2=\binom{0}{0}$. The resulting calculation shows that such a combination is unique, that $c_1$ must be $0$ and $c_2$ must be $0$.

The second verification, that the set spans the space, looks at linear combinations that total to any member of the space $c_1\vec{\beta}_1+c_2\vec{\beta}_2=\binom{x}{y}$. In Example 1.2 we noted only that the resulting calculation shows that such a combination exists, that for each $x,y$ there is a $c_1,c_2$. However, in fact the calculation also shows that the combination is unique: $c_1$ must be $(y-x)/2$ and $c_2$ must be $2x-y$.

That is, the first calculation is a special case of the second. The next result says that this holds in general for a spanning set: the combination totaling to the zero vector is unique if and only if the combination totaling to any vector is unique.

**1.12 Theorem** In any vector space, a subset is a basis if and only if each vector in the space can be expressed as a linear combination of elements of the subset in a unique way.

We consider combinations to be the same if they differ only in the order of summands or in the addition or deletion of terms of the form '$0 \cdot \beta$'.

PROOF. By definition, a sequence is a basis if and only if its vectors form both a spanning set and a linearly independent set. A subset is a spanning set if and only if each vector in the space is a linear combination of elements of that subset in at least one way.

Thus, to finish we need only show that a subset is linearly independent if and only if every vector in the space is a linear combination of elements from the subset in at most one way. Consider two expressions of a vector as a linear combination of the members of the basis. We can rearrange the two sums, and if necessary add some $0\beta_i$ terms, so that the two sums combine the same $\beta$'s in the same order: $v = c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$ and $v = d_1\beta_1 + d_2\beta_2 + \cdots + d_n\beta_n$. Now

$$c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n = d_1\beta_1 + d_2\beta_2 + \cdots + d_n\beta_n$$

holds if and only if

$$(c_1 - d_1)\beta_1 + \cdots + (c_n - d_n)\beta_n = 0$$

holds, and so asserting that each coefficient in the lower equation is zero is the same thing as asserting that $c_i = d_i$ for each $i$. QED

**1.13 Definition** In a vector space with basis $B$ the *representation of $v$ with respect to $B$* is the column vector of the coefficients used to express $v$ as a linear combination of the basis vectors:

$$\text{Rep}_B(v) = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

where $B = \langle \beta_1, \ldots, \beta_n \rangle$ and $v = c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$. The $c$'s are the *coordinates of $v$ with respect to $B$*.

We will later do representations in contexts that involve more than one basis. To help with the book-keeping, we shall often attach a subscript $B$ to the column vector.

**1.14 Example** In $\mathcal{P}_3$, with respect to the basis $B = \langle 1, 2x, 2x^2, 2x^3 \rangle$, the representation of $x + x^2$ is

$$\text{Rep}_B(x + x^2) = \begin{pmatrix} 0 \\ 1/2 \\ 1/2 \\ 0 \end{pmatrix}_B$$

(note that the coordinates are scalars, not vectors). With respect to a different basis $D = \langle 1 + x, 1 - x, x + x^2, x + x^3 \rangle$, the representation

$$\text{Rep}_D(x + x^2) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}_D$$

is different.

**1.15 Remark** This use of column notation and the term 'coordinates' has both a down side and an up side.

The down side is that representations look like vectors from $\mathbb{R}^n$, which can be confusing when the vector space we are working with is $\mathbb{R}^n$, especially since we sometimes omit the subscript base. We must then infer the intent from the context. For example, the phrase 'in $\mathbb{R}^2$, where $v = \binom{3}{2}$' refers to the plane vector that, when in canonical position, ends at $(3,2)$. To find the coordinates of that vector with respect to the basis

$$B = \langle \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\2\end{pmatrix} \rangle$$

we solve

$$c_1 \begin{pmatrix}1\\1\end{pmatrix} + c_2 \begin{pmatrix}0\\2\end{pmatrix} = \begin{pmatrix}3\\2\end{pmatrix}$$

to get that $c_1 = 3$ and $c_2 = 1/2$. Then we have this.

$$\text{Rep}_B(v) = \begin{pmatrix}3\\-1/2\end{pmatrix}$$

Here, although we've ommited the subscript $B$ from the column, the fact that the right side is a representation is clear from the context.

The up side of the notation and the term 'coordinates' is that they generalize the use that we are familiar with: in $\mathbb{R}^n$ and with respect to the standard basis $\mathcal{E}_n$, the vector starting at the origin and ending at $(v_1, \ldots, v_n)$ has this representation.

$$\text{Rep}_{\mathcal{E}_n}(\begin{pmatrix}v_1\\ \vdots \\ v_n\end{pmatrix}) = \begin{pmatrix}v_1\\ \vdots \\ v_n\end{pmatrix}_{\mathcal{E}_n}$$

Our main use of representations will come in the third chapter. The definition appears here because the fact that every vector is a linear combination of basis vectors in a unique way is a crucial property of bases, and also to help make two points. First, we fix an order for the elements of a basis so that coordinates can be stated in that order. Second, for calculation of coordinates, among other things, we shall restrict our attention to spaces with bases having only finitely many elements. We will see that in the next subsection.

### Exercises

✓ **1.16** Decide if each is a basis for $\mathbb{R}^3$.

**(a)** $\langle \begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}3\\2\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \rangle$ **(b)** $\langle \begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}3\\2\\1\end{pmatrix} \rangle$ **(c)** $\langle \begin{pmatrix}0\\2\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\5\\0\end{pmatrix} \rangle$ **(d)** $\langle \begin{pmatrix}0\\2\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\3\\0\end{pmatrix} \rangle$

✓ **1.17** Represent the vector with respect to the basis.

**(a)** $\begin{pmatrix}1\\2\end{pmatrix}$, $B = \langle \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix} \rangle \subseteq \mathbb{R}^2$

**(b)** $x^2 + x^3$, $D = \langle 1, 1+x, 1+x+x^2, 1+x+x^2+x^3 \rangle \subseteq \mathcal{P}_3$

**(c)** $\begin{pmatrix}0\\-1\\0\\1\end{pmatrix}$, $\mathcal{E}_4 \subseteq \mathbb{R}^4$

**1.18** Find a basis for $\mathcal{P}_2$, the space of all quadratic polynomials. Must any such basis contain a polynomial of each degree: degree zero, degree one, and degree two?

**1.19** Find a basis for the solution set of this system.

$$\begin{aligned} x_1 - 4x_2 + 3x_3 - \phantom{2}x_4 &= 0 \\ 2x_1 - 8x_2 + 6x_3 - 2x_4 &= 0 \end{aligned}$$

✓ **1.20** Find a basis for $\mathcal{M}_{2\times2}$, the space of $2\times2$ matrices.

✓ **1.21** Find a basis for each.

**(a)** The subspace $\{a_2x^2 + a_1x + a_0 \mid a_2 - 2a_1 = a_0\}$ of $\mathcal{P}_2$

**(b)** The space of three-wide row vectors whose first and second components add to zero

**(c)** This subspace of the $2\times2$ matrices

$$\{\begin{pmatrix}a & b\\ 0 & c\end{pmatrix} \mid c - 2b = 0\}$$

**1.22** Check Example 1.6.

✓ **1.23** Find the span of each set and then find a basis for that span.

**(a)** $\{1+x, 1+2x\}$ in $\mathcal{P}_2$ **(b)** $\{2-2x, 3+4x^2\}$ in $\mathcal{P}_2$

✓ **1.24** Find a basis for each of these subspaces of the space $\mathcal{P}_3$ of cubic polynomials.

**(a)** The subspace of cubic polynomials $p(x)$ such that $p(7)=0$

**(b)** The subspace of polynomials $p(x)$ such that $p(7)=0$ and $p(5)=0$

**(c)** The subspace of polynomials $p(x)$ such that $p(7)=0$, $p(5)=0$, and $p(3)=0$

**(d)** The space of polynomials $p(x)$ such that $p(7)=0$, $p(5)=0$, $p(3)=0$, and $p(1)=0$

**1.25** We've seen that it is possible for a basis to remain a basis when it is reordered. Must it remain a basis?

**1.26** Can a basis contain a zero vector?

✓ **1.27** Let $\langle \beta_1, \beta_2, \beta_3 \rangle$ be a basis for a vector space.

**(a)** Show that $\langle c_1\beta_1, c_2\beta_2, c_3\beta_3 \rangle$ is a basis when $c_1, c_2, c_3 \neq 0$. What happens when at least one $c_i$ is 0?

**(b)** Prove that $\langle \alpha_1, \alpha_2, \alpha_3 \rangle$ is a basis where $\alpha_i = \beta_1 + \beta_i$.

**1.28** Find one vector $v$ that will make each into a basis for the space.

**(a)** $\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v \rangle$ in $\mathbb{R}^2$ **(b)** $\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, v \rangle$ in $\mathbb{R}^3$ **(c)** $\langle x, 1+x^2, v \rangle$ in $\mathcal{P}_2$

✓ **1.29** Where $\langle \beta_1, \ldots, \beta_n \rangle$ is a basis, show that in this equation

$$c_1\beta_1 + \cdots + c_k\beta_k = c_{k+1}\beta_{k+1} + \cdots + c_n\beta_n$$

each of the $c_i$'s is zero. Generalize.

**1.30** A basis contains some of the vectors from a vector space; can it contain them all?

**1.31** Theorem 1.12 shows that, with respect to a basis, every linear combination is unique. If a subset is not a basis, can linear combinations be not unique? If so, must they be?

✓ **1.32** A square matrix is *symmetric* if for all indices $i$ and $j$, entry $i, j$ equals entry $j, i$.

**(a)** Find a basis for the vector space of symmetric $2 \times 2$ matrices.

**(b)** Find a basis for the space of symmetric $3 \times 3$ matrices.

**(c)** Find a basis for the space of symmetric $n \times n$ matrices.

✓ **1.33** We can show that every basis for $\mathbb{R}^3$ contains the same number of vectors.

**(a)** Show that no linearly independent subset of $\mathbb{R}^3$ contains more than three vectors.

**(b)** Show that no spanning subset of $\mathbb{R}^3$ contains fewer than three vectors. (*Hint.* Recall how to calculate the span of a set and show that this method, when applied to two vectors, cannot yield all of $\mathbb{R}^3$.)

**1.34** One of the exercises in the Subspaces subsection shows that the set

$$\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \bigm| x+y+z=1\}$$

is a vector space under these operations.

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1+x_2-1 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix} \qquad r\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} rx-r+1 \\ ry \\ rz \end{pmatrix}$$

Find a basis.

## III.2 Dimension

In the prior subsection we defined the basis of a vector space, and we saw that a space can have many different bases. For example, following the definition of a basis, we saw three different bases for $\mathbb{R}^2$. So we cannot talk about "the" basis for a vector space. True, some vector spaces have bases that strike us as more natural than others, for instance, $\mathbb{R}^2$'s basis $\mathcal{E}_2$ or $\mathbb{R}^3$'s basis $\mathcal{E}_3$ or $\mathcal{P}_2$'s basis $\langle 1, x, x^2 \rangle$. But, for example in the space $\{a_2x^2 + a_1x + a_0 \bigm| 2a_2 - a_0 = a_1\}$, no particular basis leaps out at us as the most natural one. We cannot, in general, associate with a space any single basis that best describes that space.

We can, however, find something about the bases that is uniquely associated with the space. This subsection shows that any two bases for a space have the same number of elements. So, with each space we can associate a number, the number of vectors in any of its bases.

This brings us back to when we considered the two things that could be meant by the term 'minimal spanning set'. At that point we defined 'minimal' as linearly independent, but we noted that another reasonable interpretation of the term is that a spanning set is 'minimal' when it has the fewest number of elements of any set with the same span. At the end of this subsection, after we have shown that all bases have the same number of elements, then we will have shown that the two senses of 'minimal' are equivalent.

Before we start, we first limit our attention to spaces where at least one basis has only finitely many members.

**2.1 Definition** A vector space is *finite-dimensional* if it has a basis with only finitely many vectors.

(One reason for sticking to finite-dimensional spaces is so that the representation of a vector with respect to a basis is a finitely-tall vector, and so can be easily written.) From now on we study only finite-dimensional vector spaces. We shall take the term 'vector space' to mean 'finite-dimensional vector space'. Other spaces are interesting and important, but they lie outside of our scope.

To prove the main theorem we shall use a technical result.

**2.2 Lemma (Exchange Lemma)** Assume that $B = \langle \beta_1, \ldots, \beta_n \rangle$ is a basis for a vector space, and that for the vector $v$ the relationship $v = c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$ has $c_i \neq 0$. Then exchanging $\beta_i$ for $v$ yields another basis for the space.

Proof. Call the outcome of the exchange $\hat{B} = \langle \beta_1, \ldots, \beta_{i-1}, v, \beta_{i+1}, \ldots, \beta_n \rangle$.

We first show that $\hat{B}$ is linearly independent. Any relationship $d_1\beta_1 + \cdots + d_iv + \cdots + d_n\beta_n = 0$ among the members of $\hat{B}$, after substitution for $v$,

$$d_1\beta_1 + \cdots + d_i \cdot (c_1\beta_1 + \cdots + c_i\beta_i + \cdots + c_n\beta_n) + \cdots + d_n\beta_n = 0 \qquad (*)$$

gives a linear relationship among the members of $B$. The basis $B$ is linearly independent, so the coefficient $d_ic_i$ of $\beta_i$ is zero. Because $c_i$ is assumed to be nonzero, $d_i = 0$. Using this in equation $(*)$ above gives that all of the other $d$'s are also zero. Therefore $\hat{B}$ is linearly independent.

We finish by showing that $\hat{B}$ has the same span as $B$. Half of this argument, that $\langle \hat{B} \rangle \subseteq \langle B \rangle$, is easy; any member $d_1\beta_1 + \cdots + d_iv + \cdots + d_n\beta_n$ of $\langle \hat{B} \rangle$ can be written $d_1\beta_1 + \cdots + d_i \cdot (c_1\beta_1 + \cdots + c_n\beta_n) + \cdots + d_n\beta_n$, which is a linear combination of linear combinations of members of $B$, and hence is in $\langle B \rangle$. For the $\langle B \rangle \subseteq \langle \hat{B} \rangle$ half of the argument, recall that when $v = c_1\beta_1 + \cdots + c_n\beta_n$ with $c_i \neq 0$, then the equation can be rearranged to $\beta_i = (-c_1/c_i)\beta_1 + \cdots + (1/c_i)v + \cdots + (-c_n/c_i)\beta_n$. Now, consider any member $d_1\beta_1 + \cdots + d_i\beta_i + \cdots + d_n\beta_n$ of $\langle B \rangle$, substitute for $\beta_i$ its expression as a linear combination of the members of $\hat{B}$, and recognize (as in the first half of this argument) that the result is a linear combination of linear combinations, of members of $\hat{B}$, and hence is in $\langle \hat{B} \rangle$. QED

**2.3 Theorem** In any finite-dimensional vector space, all of the bases have the same number of elements.

Proof. Fix a vector space with at least one finite basis. Choose, from among all of this space's bases, one $B = \langle \beta_1, \ldots, \beta_n \rangle$ of minimal size. We will show that any other basis $D = \langle \delta_1, \delta_2, \ldots \rangle$ also has the same number of members, $n$. Because $B$ has minimal size, $D$ has no fewer than $n$ vectors. We will argue that it cannot have more than $n$ vectors.

The basis $B$ spans the space and $\delta_1$ is in the space, so $\delta_1$ is a nontrivial linear combination of elements of $B$. By the Exchange Lemma, $\delta_1$ can be swapped for a vector from $B$, resulting in a basis $B_1$, where one element is $\delta$ and all of the $n-1$ other elements are $\beta$'s.

The prior paragraph forms the basis step for an induction argument. The inductive step starts with a basis $B_k$ (for $1 \leq k < n$) containing $k$ members of $D$ and $n-k$ members of $B$. We know that $D$ has at least $n$ members so there is a $\delta_{k+1}$. Represent it as a linear combination of elements of $B_k$. The key point: in that representation, at least one of the nonzero scalars must be associated with a $\beta_i$ or else that representation would be a nontrivial linear relationship among elements of the linearly independent set $D$. Exchange $\delta_{k+1}$ for $\beta_i$ to get a new basis $B_{k+1}$ with one $\delta$ more and one $\beta$ fewer than the previous basis $B_k$.

Repeat the inductive step until no $\beta$'s remain, so that $B_n$ contains $\delta_1, \ldots, \delta_n$. Now, $D$ cannot have more than these $n$ vectors because any $\delta_{n+1}$ that remains would be in the span of $B_n$ (since it is a basis) and hence would be a linear combination of the other $\delta$'s, contradicting that $D$ is linearly independent. QED

**2.4 Definition** The *dimension* of a vector space is the number of vectors in any of its bases.

**2.5 Example** Any basis for $\mathbb{R}^n$ has $n$ vectors since the standard basis $\mathcal{E}_n$ has $n$ vectors. Thus, this definition generalizes the most familiar use of term, that $\mathbb{R}^n$ is $n$-dimensional.

**2.6 Example** The space $\mathcal{P}_n$ of polynomials of degree at most $n$ has dimension $n+1$. We can show this by exhibiting any basis — $\langle 1, x, \ldots, x^n \rangle$ comes to mind — and counting its members.

**2.7 Example** A trivial space is zero-dimensional since its basis is empty.

Again, although we sometimes say 'finite-dimensional' as a reminder, in the rest of this book all vector spaces are assumed to be finite-dimensional. An instance of this is that in the next result the word 'space' should be taken to mean 'finite-dimensional vector space'.

**2.8 Corollary** No linearly independent set can have a size greater than the dimension of the enclosing space.

Proof. Inspection of the above proof shows that it never uses that $D$ spans the space, only that $D$ is linearly independent. QED

**2.9 Example** Recall the subspace diagram from the prior section showing the subspaces of $\mathbb{R}^3$. Each subspace shown is described with a minimal spanning set, for which we now have the term 'basis'. The whole space has a basis with three members, the plane subspaces have bases with two members, the line subspaces have bases with one member, and the trivial subspace has a basis with zero members. When we saw that diagram we could not show that these are the only subspaces that this space has. We can show it now. The prior corollary proves that the only subspaces of $\mathbb{R}^3$ are either three-, two-, one-, or zero-dimensional. Therefore, the diagram indicates all of the subspaces. There are no subspaces somehow, say, between lines and planes.

**2.10 Corollary** Any linearly independent set can be expanded to make a basis.

Proof. If a linearly independent set is not already a basis then it must not span the space. Adding to it a vector that is not in the span preserves linear independence. Keep adding, until the resulting set does span the space, which the prior corollary shows will happen after only a finite number of steps. QED

**2.11 Corollary** Any spanning set can be shrunk to a basis.

Proof. Call the spanning set $S$. If $S$ is empty then it is already a basis (the space must be a trivial space). If $S = \{0\}$ then it can be shrunk to the empty basis, thereby making it linearly independent, without changing its span.

Otherwise, $S$ contains a vector $s_1$ with $s_1 \neq 0$ and we can form a basis $B_1 = \langle s_1 \rangle$. If $\langle B_1 \rangle = \langle S \rangle$ then we are done.

If not then there is a $s_2 \in \langle S \rangle$ such that $s_2 \notin \langle B_1 \rangle$. Let $B_2 = \langle s_1, s_2 \rangle$; if $\langle B_2 \rangle = \langle S \rangle$ then we are done.

We can repeat this process until the spans are equal, which must happen in at most finitely many steps. QED

**2.12 Corollary** In an $n$-dimensional space, a set of $n$ vectors is linearly independent if and only if it spans the space.

Proof. First we will show that a subset with $n$ vectors is linearly independent if and only if it is a basis. 'If' is trivially true — bases are linearly independent. 'Only if' holds because a linearly independent set can be expanded to a basis, but a basis has $n$ elements, so this expansion is actually the set that we began with.

To finish, we will show that any subset with $n$ vectors spans the space if and only if it is a basis. Again, 'if' is trivial. 'Only if' holds because any spanning set can be shrunk to a basis, but a basis has $n$ elements and so this shrunken set is just the one we started with. QED

The main result of this subsection, that all of the bases in a finite-dimensional vector space have the same number of elements, is the single most important result in this book because, as Example 2.9 shows, it describes what vector spaces and subspaces there can be. We will see more in the next chapter.

**2.13 Remark** The case of infinite-dimensional vector spaces is somewhat controversial. The statement 'any infinite-dimensional vector space has a basis' is known to be equivalent to a statement called the Axiom of Choice (see [Blass 1984]). Mathematicians differ philosophically on whether to accept or reject this statement as an axiom on which to base mathematics (although, the great majority seem to accept it). Consequently the question about infinite-dimensional vector spaces is still somewhat up in the air. (A discussion of the Axiom of Choice can be found in the Frequently Asked Questions list for the Usenet group `sci.math`. Another accessible reference is [Rucker].)

## Exercises

*Assume that all spaces are finite-dimensional unless otherwise stated.*

✓ **2.14** Find a basis for, and the dimension of, $\mathcal{P}_2$.

**2.15** Find a basis for, and the dimension of, the solution set of this system.

$$\begin{aligned} x_1 - 4x_2 + 3x_3 - \phantom{2}x_4 &= 0 \\ 2x_1 - 8x_2 + 6x_3 - 2x_4 &= 0 \end{aligned}$$

✓ **2.16** Find a basis for, and the dimension of, $\mathcal{M}_{2\times2}$, the vector space of $2\times2$ matrices.

**2.17** Find the dimension of the vector space of matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

subject to each condition.

**(a)** $a, b, c, d \in \mathbb{R}$
**(b)** $a - b + 2c = 0$ and $d \in \mathbb{R}$
**(c)** $a + b + c = 0$, $a + b - c = 0$, and $d \in \mathbb{R}$

✓ **2.18** Find the dimension of each.
**(a)** The space of cubic polynomials $p(x)$ such that $p(7) = 0$
**(b)** The space of cubic polynomials $p(x)$ such that $p(7) = 0$ and $p(5) = 0$
**(c)** The space of cubic polynomials $p(x)$ such that $p(7) = 0$, $p(5) = 0$, and $p(3) = 0$
**(d)** The space of cubic polynomials $p(x)$ such that $p(7) = 0$, $p(5) = 0$, $p(3) = 0$, and $p(1) = 0$

**2.19** What is the dimension of the span of the set $\{\cos^2\theta, \sin^2\theta, \cos 2\theta, \sin 2\theta\}$? This span is a subspace of the space of all real-valued functions of one real variable.

**2.20** Find the dimension of $\mathbb{C}^{47}$, the vector space of 47-tuples of complex numbers.

**2.21** What is the dimension of the vector space $\mathcal{M}_{3\times5}$ of $3\times5$ matrices?

✓ **2.22** Show that this is a basis for $\mathbb{R}^4$.

$$\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \rangle$$

(The results of this subsection can be used to simplify this job.)

**2.23** Refer to Example 2.9.
**(a)** Sketch a similar subspace diagram for $\mathcal{P}_2$.
**(b)** Sketch one for $\mathcal{M}_{2\times2}$.

✓ **2.24** Where $S$ is a set, the functions $f\colon S \to \mathbb{R}$ form a vector space under the natural operations: the sum $f + g$ is the function given by $f + g\,(s) = f(s) + g(s)$ and the scalar product is given by $r \cdot f\,(s) = r \cdot f(s)$. What is the dimension of the space resulting for each domain?
**(a)** $S = \{1\}$ **(b)** $S = \{1, 2\}$ **(c)** $S = \{1, \ldots, n\}$

**2.25** (See Exercise 24.) Prove that this is an infinite-dimensional space: the set of all functions $f\colon \mathbb{R} \to \mathbb{R}$ under the natural operations.

**2.26** (See Exercise 24.) What is the dimension of the vector space of functions $f\colon S \to \mathbb{R}$, under the natural operations, where the domain $S$ is the empty set?

**2.27** Show that any set of four vectors in $\mathbb{R}^2$ is linearly dependent.

**2.28** Show that the set $\langle \alpha_1, \alpha_2, \alpha_3 \rangle \subset \mathbb{R}^3$ is a basis if and only if there is no plane through the origin containing all three vectors.

**2.29** **(a)** Prove that any subspace of a finite dimensional space has a basis.
**(b)** Prove that any subspace of a finite dimensional space is finite dimensional.

**2.30** Where is the finiteness of $B$ used in Theorem 2.3?

✓ **2.31** Prove that if $U$ and $W$ are both three-dimensional subspaces of $\mathbb{R}^5$ then $U \cap W$ is non-trivial. Generalize.

**2.32** Because a basis for a space is a subset of that space, we are naturally led to how the property 'is a basis' interacts with set operations.

**(a)** Consider first how bases might be related by 'subset'. Assume that $U, W$ are subspaces of some vector space and that $U \subseteq W$. Can there exist bases $B_U$ for $U$ and $B_W$ for $W$ such that $B_U \subseteq B_W$? Must such bases exist?
For any basis $B_U$ for $U$, must there be a basis $B_W$ for $W$ such that $B_U \subseteq B_W$?
For any basis $B_W$ for $W$, must there be a basis $B_U$ for $U$ such that $B_U \subseteq B_W$?
For any bases $B_U, B_W$ for $U$ and $W$, must $B_U$ be a subset of $B_W$?

**(b)** Is the intersection of bases a basis? For what space?

**(c)** Is the union of bases a basis? For what space?

**(d)** What about complement?

(*Hint.* Test any conjectures against some subspaces of $\mathbb{R}^3$.)

✓ **2.33** Consider how 'dimension' interacts with 'subset'. Assume $U$ and $W$ are both subspaces of some vector space, and that $U \subseteq W$.

**(a)** Prove that $\dim(U) \leq \dim(W)$.

**(b)** Prove that equality of dimension holds if and only if $U = W$.

**(c)** Show that the prior item does not hold if they are infinite-dimensional.

? **2.34** For any vector $\vec{v}$ in $\mathbb{R}^n$ and any permutation $\sigma$ of the numbers 1, 2, ..., $n$ (that is, $\sigma$ is a rearrangement of those numbers into a new order), define $\sigma(\vec{v})$ to be the vector whose components are $v_{\sigma(1)}$, $v_{\sigma(2)}$, ..., and $v_{\sigma(n)}$ (where $\sigma(1)$ is the first number in the rearrangement, etc.). Now fix $\vec{v}$ and let $V$ be the span of $\{\sigma(\vec{v}) \mid \sigma \text{ permutes } 1, \ldots, n\}$. What are the possibilities for the dimension of $V$? [Wohascum no. 47]

## III.3 Vector Spaces and Linear Systems

We will now reconsider linear systems and Gauss' method, aided by the tools and terms of this chapter. We will make three points.

For the first point, recall the first chapter's Linear Combination Lemma and its corollary: if two matrices are related by row operations $A \longrightarrow \cdots \longrightarrow B$ then each row of $B$ is a linear combination of the rows of $A$. That is, Gauss' method works by taking linear combinations of rows. Therefore, the right setting in which to study row operations in general, and Gauss' method in particular, is the following vector space.

**3.1 Definition** The *row space* of a matrix is the span of the set of its rows. The *row rank* is the dimension of the row space, the number of linearly independent rows.

**3.2 Example** If

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}$$

then Rowspace$(A)$ is this subspace of the space of two-component row vectors.

$$\{c_1 \cdot \begin{pmatrix} 2 & 3 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 4 & 6 \end{pmatrix} \mid c_1, c_2 \in \mathbb{R}\}$$

The linear dependence of the second on the first is obvious and so we can simplify this description to $\{c \cdot \begin{pmatrix} 2 & 3 \end{pmatrix} \mid c \in \mathbb{R}\}$.

**3.3 Lemma** If the matrices $A$ and $B$ are related by a row operation

$$A \xrightarrow{\rho_i \leftrightarrow \rho_j} B \quad \text{or} \quad A \xrightarrow{k\rho_i} B \quad \text{or} \quad A \xrightarrow{k\rho_i + \rho_j} B$$

(for $i \neq j$ and $k \neq 0$) then their row spaces are equal. Hence, row-equivalent matrices have the same row space, and hence also, the same row rank.

Proof. By the Linear Combination Lemma's corollary, each row of $B$ is in the row space of $A$. Further, $\text{Rowspace}(B) \subseteq \text{Rowspace}(A)$ because a member of the set $\text{Rowspace}(B)$ is a linear combination of the rows of $B$, which means it is a combination of a combination of the rows of $A$, and hence, by the Linear Combination Lemma, is also a member of $\text{Rowspace}(A)$.

For the other containment, recall that row operations are reversible: $A \longrightarrow B$ if and only if $B \longrightarrow A$. With that, $\text{Rowspace}(A) \subseteq \text{Rowspace}(B)$ also follows from the prior paragraph, and so the two sets are equal. QED

Thus, row operations leave the row space unchanged. But of course, Gauss' method performs the row operations systematically, with a specific goal in mind, echelon form.

**3.4 Lemma** The nonzero rows of an echelon form matrix make up a linearly independent set.

Proof. A result in the first chapter, Lemma III.??, states that in an echelon form matrix, no nonzero row is a linear combination of the other rows. This is a restatement of that result into new terminology. QED

Thus, in the language of this chapter, Gaussian reduction works by eliminating linear dependences among rows, leaving the span unchanged, until no nontrivial linear relationships remain (among the nonzero rows). That is, Gauss' method produces a basis for the row space.

**3.5 Example** From any matrix, we can produce a basis for the row space by performing Gauss' method and taking the nonzero rows of the resulting echelon form matrix. For instance,

$$\begin{pmatrix} 1 & 3 & 1 \\ 1 & 4 & 1 \\ 2 & 0 & 5 \end{pmatrix} \xrightarrow[-2\rho_1+\rho_3]{-\rho_1+\rho_2} \xrightarrow{6\rho_2+\rho_3} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

produces the basis $\langle \begin{pmatrix} 1 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 3 \end{pmatrix} \rangle$ for the row space. This is a basis for the row space of both the starting and ending matrices, since the two row spaces are equal.

Using this technique, we can also find bases for spans not directly involving row vectors.

**3.6 Definition** The *column space* of a matrix is the span of the set of its columns. The *column rank* is the dimension of the column space, the number of linearly independent columns.

Our interest in column spaces stems from our study of linear systems. An example is that this system

$$\begin{aligned} c_1 + 3c_2 + 7c_3 &= d_1 \\ 2c_1 + 3c_2 + 8c_3 &= d_2 \\ c_2 + 2c_3 &= d_3 \\ 4c_1 \qquad + 4c_3 &= d_4 \end{aligned}$$

has a solution if and only if the vector of $d$'s is a linear combination of the other column vectors,

$$c_1 \begin{pmatrix} 1 \\ 2 \\ 0 \\ 4 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ 3 \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 7 \\ 8 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix}$$

meaning that the vector of $d$'s is in the column space of the matrix of coefficients.

**3.7 Example** Given this matrix,

$$\begin{pmatrix} 1 & 3 & 7 \\ 2 & 3 & 8 \\ 0 & 1 & 2 \\ 4 & 0 & 4 \end{pmatrix}$$

to get a basis for the column space, temporarily turn the columns into rows and reduce.

$$\begin{pmatrix}1&2&0&4\\3&3&1&0\\7&8&2&4\end{pmatrix}\xrightarrow[-7\rho_1+\rho_3]{-3\rho_1+\rho_2}\xrightarrow{-2\rho_2+\rho_3}\begin{pmatrix}1&2&0&4\\0&-3&1&-12\\0&0&0&0\end{pmatrix}$$

Now turn the rows back to columns.

$$\langle\begin{pmatrix}1\\2\\0\\4\end{pmatrix},\begin{pmatrix}0\\-3\\1\\-12\end{pmatrix}\rangle$$

The result is a basis for the column space of the given matrix.

**3.8 Definition** The *transpose* of a matrix is the result of interchanging the rows and columns of that matrix. That is, column $j$ of the matrix $A$ is row $j$ of $A^{\text{trans}}$, and vice versa.

So the instructions for the prior example are "transpose, reduce, and transpose back".

We can even, at the price of tolerating the as-yet-vague idea of vector spaces being "the same", use Gauss' method to find bases for spans in other types of vector spaces.

**3.9 Example** To get a basis for the span of $\{x^2+x^4, 2x^2+3x^4, -x^2-3x^4\}$ in the space $\mathcal{P}_4$, think of these three polynomials as "the same" as the row vectors $\begin{pmatrix}0&0&1&0&1\end{pmatrix}$, $\begin{pmatrix}0&0&2&0&3\end{pmatrix}$, and $\begin{pmatrix}0&0&-1&0&-3\end{pmatrix}$, apply Gauss' method

$$\begin{pmatrix}0&0&1&0&1\\0&0&2&0&3\\0&0&-1&0&-3\end{pmatrix}\xrightarrow[\rho_1+\rho_3]{-2\rho_1+\rho_2}\xrightarrow{2\rho_2+\rho_3}\begin{pmatrix}0&0&1&0&1\\0&0&0&0&1\\0&0&0&0&0\end{pmatrix}$$

and translate back to get the basis $\langle x^2+x^4, x^4\rangle$. (As mentioned earlier, we will make the phrase "the same" precise at the start of the next chapter.)

Thus, our first point in this subsection is that the tools of this chapter give us a more conceptual understanding of Gaussian reduction.

For the second point of this subsection, consider the effect on the column space of this row reduction.

$$\begin{pmatrix}1&2\\2&4\end{pmatrix}\xrightarrow{-2\rho_1+\rho_2}\begin{pmatrix}1&2\\0&0\end{pmatrix}$$

The column space of the left-hand matrix contains vectors with a second component that is nonzero. But the column space of the right-hand matrix is different because it contains only vectors whose second component is zero. It is this knowledge that row operations can change the column space that makes next result surprising.

**3.10 Lemma** Row operations do not change the column rank.

PROOF. Restated, if $A$ reduces to $B$ then the column rank of $B$ equals the column rank of $A$.

We will be done if we can show that row operations do not affect linear relationships among columns (e.g., if the fifth column is twice the second plus the fourth before a row operation then that relationship still holds afterwards), because the column rank is just the size of the largest set of unrelated columns. But this is exactly the first theorem of this book: in a relationship among columns,

$$c_1\cdot\begin{pmatrix}a_{1,1}\\a_{2,1}\\\vdots\\a_{m,1}\end{pmatrix}+\cdots+c_n\cdot\begin{pmatrix}a_{1,n}\\a_{2,n}\\\vdots\\a_{m,n}\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$$

row operations leave unchanged the set of solutions $(c_1,\ldots,c_n)$. QED

Another way, besides the prior result, to state that Gauss' method has something to say about the column space as well as about the row space is to consider again Gauss-Jordan reduction. Recall that it ends with the reduced echelon form of a matrix, as here.

$$\begin{pmatrix} 1 & 3 & 1 & 6 \\ 2 & 6 & 3 & 16 \\ 1 & 3 & 1 & 6 \end{pmatrix} \longrightarrow \cdots \longrightarrow \begin{pmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Consider the row space and the column space of this result. Our first point made above says that a basis for the row space is easy to get: simply collect together all of the rows with leading entries. However, because this is a reduced echelon form matrix, a basis for the column space is just as easy: take the columns containing the leading entries, that is, $\langle e_1, e_2 \rangle$. (Linear independence is obvious. The other columns are in the span of this set, since they all have a third component of zero.) Thus, for a reduced echelon form matrix, bases for the row and column spaces can be found in essentially the same way — by taking the parts of the matrix, the rows or columns, containing the leading entries.

**3.11 Theorem** The row rank and column rank of a matrix are equal.

Proof. First bring the matrix to reduced echelon form. At that point, the row rank equals the number of leading entries since each equals the number of nonzero rows. Also at that point, the number of leading entries equals the column rank because the set of columns containing leading entries consists of some of the $e_i$'s from a standard basis, and that set is linearly independent and spans the set of columns. Hence, in the reduced echelon form matrix, the row rank equals the column rank, because each equals the number of leading entries.

But Lemma 3.3 and Lemma 3.10 show that the row rank and column rank are not changed by using row operations to get to reduced echelon form. Thus the row rank and the column rank of the original matrix are also equal. QED

**3.12 Definition** The *rank* of a matrix is its row rank or column rank.

So our second point in this subsection is that the column space and row space of a matrix have the same dimension. Our third and final point is that the concepts that we've seen arising naturally in the study of vector spaces are exactly the ones that we have studied with linear systems.

**3.13 Theorem** For linear systems with $n$ unknowns and with matrix of coefficients $A$, the statements

(1) the rank of $A$ is $r$

(2) the space of solutions of the associated homogeneous system has dimension $n - r$

are equivalent.

So if the system has at least one particular solution then for the set of solutions, the number of parameters equals $n - r$, the number of variables minus the rank of the matrix of coefficients.

Proof. The rank of $A$ is $r$ if and only if Gaussian reduction on $A$ ends with $r$ nonzero rows. That's true if and only if echelon form matrices row equivalent to $A$ have $r$-many leading variables. That in turn holds if and only if there are $n - r$ free variables. QED

**3.14 Remark** [Munkres] Sometimes that result is mistakenly remembered to say that the general solution of an $n$ unknown system of $m$ equations uses $n - m$ parameters. The number of equations is not the relevant figure, rather, what matters is the number of independent equations (the number of equations in a maximal independent set). Where there are $r$ independent equations, the general solution involves $n - r$ parameters.

**3.15 Corollary** Where the matrix $A$ is $n \times n$, the statements

(1) the rank of $A$ is $n$

(2) $A$ is nonsingular

(3) the rows of $A$ form a linearly independent set

(4) the columns of $A$ form a linearly independent set

(5) any linear system whose matrix of coefficients is $A$ has one and only one solution
are equivalent.

PROOF. Clearly (1) $\iff$ (2) $\iff$ (3) $\iff$ (4). The last, (4) $\iff$ (5), holds because a set of $n$ column vectors is linearly independent if and only if it is a basis for $\mathbb{R}^n$, but the system

$$c_1 \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{pmatrix} + \cdots + c_n \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{pmatrix}$$

has a unique solution for all choices of $d_1, \ldots, d_n \in \mathbb{R}$ if and only if the vectors of $a$'s form a basis. QED

## Exercises

**3.16** Transpose each.

**(a)** $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ **(c)** $\begin{pmatrix} 1 & 4 & 3 \\ 6 & 7 & 8 \end{pmatrix}$ **(d)** $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ **(e)** $\begin{pmatrix} -1 & -2 \end{pmatrix}$

✓ **3.17** Decide if the vector is in the row space of the matrix.

**(a)** $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \end{pmatrix}$ **(b)** $\begin{pmatrix} 0 & 1 & 3 \\ -1 & 0 & 1 \\ -1 & 2 & 7 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$

✓ **3.18** Decide if the vector is in the column space.

**(a)** $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ **(b)** $\begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & 4 \\ 1 & -3 & -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

✓ **3.19** Find a basis for the row space of this matrix.

$$\begin{pmatrix} 2 & 0 & 3 & 4 \\ 0 & 1 & 1 & -1 \\ 3 & 1 & 0 & 2 \\ 1 & 0 & -4 & -1 \end{pmatrix}$$

✓ **3.20** Find the rank of each matrix.

**(a)** $\begin{pmatrix} 2 & 1 & 3 \\ 1 & -1 & 2 \\ 1 & 0 & 3 \end{pmatrix}$ **(b)** $\begin{pmatrix} 1 & -1 & 2 \\ 3 & -3 & 6 \\ -2 & 2 & -4 \end{pmatrix}$ **(c)** $\begin{pmatrix} 1 & 3 & 2 \\ 5 & 1 & 1 \\ 6 & 4 & 3 \end{pmatrix}$ **(d)** $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

✓ **3.21** Find a basis for the span of each set.

**(a)** $\{\begin{pmatrix} 1 & 3 \end{pmatrix}, \begin{pmatrix} -1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 1 \end{pmatrix}\} \subseteq \mathcal{M}_{1\times 2}$

**(b)** $\{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ -3 \end{pmatrix}\} \subseteq \mathbb{R}^3$

**(c)** $\{1 + x, 1 - x^2, 3 + 2x - x^2\} \subseteq \mathcal{P}_3$

**(d)** $\{\begin{pmatrix} 1 & 0 & 1 \\ 3 & 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \end{pmatrix}, \begin{pmatrix} -1 & 0 & -5 \\ -1 & -1 & -9 \end{pmatrix}\} \subseteq \mathcal{M}_{2\times 3}$

**3.22** Which matrices have rank zero? Rank one?

✓ **3.23** Given $a, b, c \in \mathbb{R}$, what choice of $d$ will cause this matrix to have the rank of one?

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

**3.24** Find the column rank of this matrix.

$$\begin{pmatrix} 1 & 3 & -1 & 5 & 0 & 4 \\ 2 & 0 & 1 & 0 & 4 & 1 \end{pmatrix}$$

**3.25** Show that a linear system with at least one solution has at most one solution if and only if the matrix of coefficients has rank equal to the number of its columns.

✓ **3.26** If a matrix is $5\times 9$, which set must be dependent, its set of rows or its set of columns?

**3.27** Give an example to show that, despite that they have the same dimension, the row space and column space of a matrix need not be equal. Are they ever equal?

**3.28** Show that the set $\{(1,-1,2,-3),(1,1,2,0),(3,-1,6,-6)\}$ does not have the same span as $\{(1,0,1,0),(0,2,0,3)\}$. What, by the way, is the vector space?

✓ **3.29** Show that this set of column vectors

$$\left\{\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix} \;\middle|\; \text{there are } x,\ y, \text{ and } z \text{ such that } \begin{array}{rl} 3x+2y+4z&=d_1\\ x\phantom{+2y}\ \ -\ z&=d_2\\ 2x+2y+5z&=d_3\end{array}\right\}$$

is a subspace of $\mathbb{R}^3$. Find a basis.

**3.30** Show that the transpose operation is *linear*:

$$(rA+sB)^{\text{trans}}=rA^{\text{trans}}+sB^{\text{trans}}$$

for $r,s\in\mathbb{R}$ and $A,B\in\mathcal{M}_{m\times n}$,

✓ **3.31** In this subsection we have shown that Gaussian reduction finds a basis for the row space.

**(a)** Show that this basis is not unique — different reductions may yield different bases.

**(b)** Produce matrices with equal row spaces but unequal numbers of rows.

**(c)** Prove that two matrices have equal row spaces if and only if after Gauss-Jordan reduction they have the same nonzero rows.

**3.32** Why is there not a problem with Remark 3.14 in the case that $r$ is bigger than $n$?

**3.33** Show that the row rank of an $m\times n$ matrix is at most $m$. Is there a better bound?

✓ **3.34** Show that the rank of a matrix equals the rank of its transpose.

**3.35** True or false: the column space of a matrix equals the row space of its transpose.

✓ **3.36** We have seen that a row operation may change the column space. Must it?

**3.37** Prove that a linear system has a solution if and only if that system's matrix of coefficients has the same rank as its augmented matrix.

**3.38** An $m\times n$ matrix has *full row rank* if its row rank is $m$, and it has *full column rank* if its column rank is $n$.

**(a)** Show that a matrix can have both full row rank and full column rank only if it is square.

**(b)** Prove that the linear system with matrix of coefficients $A$ has a solution for any $d_1,\ \ldots,\ d_n$'s on the right side if and only if $A$ has full row rank.

**(c)** Prove that a homogeneous system has a unique solution if and only if its matrix of coefficients $A$ has full column rank.

**(d)** Prove that the statement "if a system with matrix of coefficients $A$ has any solution then it has a unique solution" holds if and only if $A$ has full column rank.

**3.39** How would the conclusion of Lemma 3.3 change if Gauss' method is changed to allow multiplying a row by zero?

✓ **3.40** What is the relationship between $\text{rank}(A)$ and $\text{rank}(-A)$? Between $\text{rank}(A)$ and $\text{rank}(kA)$? What, if any, is the relationship between $\text{rank}(A)$, $\text{rank}(B)$, and $\text{rank}(A+B)$?

## III.4 Combining Subspaces

*This subsection is optional. It is required only for the last sections of Chapter Three and Chapter Five and for occasional exercises, and can be passed over without loss of continuity.*

This chapter opened with the definition of a vector space, and the middle consisted of a first analysis of the idea. This subsection closes the chapter by finishing the analysis, in the sense that 'analysis' means "method of determining the ... essential features of something by separating it into parts" [Macmillan Dictionary].

A common way to understand things is to see how they can be built from component parts. For instance, we think of $\mathbb{R}^3$ as put together, in some way, from the $x$-axis, the $y$-axis, and $z$-axis. In this subsection we will make this precise; we will describe how to decompose a vector space into a combination of some of its subspaces. In developing this idea of subspace combination, we will keep the $\mathbb{R}^3$ example in mind as a benchmark model.

Subspaces are subsets and sets combine via union. But taking the combination operation for subspaces to be the simple union operation isn't what we want. For one thing, the union of the $x$-axis, the $y$-axis, and $z$-axis is not all of $\mathbb{R}^3$, so the benchmark model would be left out. Besides, union is all wrong for this reason: a union of subspaces need not be a subspace (it need not be closed; for instance, this $\mathbb{R}^3$ vector

$$\begin{pmatrix}1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\1\\0\end{pmatrix}+\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}1\\1\\1\end{pmatrix}$$

is in none of the three axes and hence is not in the union). In addition to the members of the subspaces, we must at least also include all of the linear combinations.

**4.1 Definition** Where $W_1, \ldots, W_k$ are subspaces of a vector space, their *sum* is the span of their union $W_1 + W_2 + \cdots + W_k = \langle W_1 \cup W_2 \cup \ldots W_k \rangle$.

(The notation, writing the '+' between sets in addition to using it between vectors, fits with the practice of using this symbol for any natural accumulation operation.)

**4.2 Example** The $\mathbb{R}^3$ model fits with this operation. Any vector $w \in \mathbb{R}^3$ can be written as a linear combination $c_1v_1 + c_2v_2 + c_3v_3$ where $v_1$ is a member of the $x$-axis, etc., in this way

$$\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}=1\cdot\begin{pmatrix}w_1\\0\\0\end{pmatrix}+1\cdot\begin{pmatrix}0\\w_2\\0\end{pmatrix}+1\cdot\begin{pmatrix}0\\0\\w_3\end{pmatrix}$$

and so $\mathbb{R}^3 = x\text{-axis} + y\text{-axis} + z\text{-axis}$.

**4.3 Example** A sum of subspaces can be less than the entire space. Inside of $\mathcal{P}_4$, let $L$ be the subspace of linear polynomials $\{a + bx \mid a, b \in \mathbb{R}\}$ and let $C$ be the subspace of purely-cubic polynomials $\{cx^3 \mid c \in \mathbb{R}\}$. Then $L + C$ is not all of $\mathcal{P}_4$. Instead, it is the subspace $L + C = \{a + bx + cx^3 \mid a, b, c \in \mathbb{R}\}$.

**4.4 Example** A space can be described as a combination of subspaces in more than one way. Besides the decomposition $\mathbb{R}^3 = x\text{-axis} + y\text{-axis} + z\text{-axis}$, we can also write $\mathbb{R}^3 = xy\text{-plane} + yz\text{-plane}$. To check this, note that any $w \in \mathbb{R}^3$ can be written as a linear combination of a member of the $xy$-plane and a member of the $yz$-plane; here are two such combinations.

$$\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}=1\cdot\begin{pmatrix}w_1\\w_2\\0\end{pmatrix}+1\cdot\begin{pmatrix}0\\0\\w_3\end{pmatrix}\qquad\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}=1\cdot\begin{pmatrix}w_1\\w_2/2\\0\end{pmatrix}+1\cdot\begin{pmatrix}0\\w_2/2\\w_3\end{pmatrix}$$

The above definition gives one way in which a space can be thought of as a combination of some of its parts. However, the prior example shows that there is at least one interesting property of our benchmark model that is not captured by the definition of the sum of subspaces. In the familiar decomposition of $\mathbb{R}^3$, we often speak of a vector's '$x$ part' or '$y$ part' or '$z$ part'. That is, in this model, each vector has a unique decomposition into parts that come from the parts making up the whole space. But in the decomposition used in Example 4.4, we cannot refer to the "$xy$ part" of a vector — these three sums

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

all describe the vector as comprised of something from the first plane plus something from the second plane, but the "$xy$ part" is different in each.

That is, when we consider how $\mathbb{R}^3$ is put together from the three axes "in some way", we might mean "in such a way that every vector has at least one decomposition", and that leads to the definition above. But if we take it to mean "in such a way that every vector has one and only one decomposition" then we need another condition on combinations. To see what this condition is, recall that vectors are uniquely represented in terms of a basis. We can use this to break a space into a sum of subspaces such that any vector in the space breaks uniquely into a sum of members of those subspaces.

**4.5 Example** The benchmark is $\mathbb{R}^3$ with its standard basis $\mathcal{E}_3 = \langle e_1, e_2, e_3 \rangle$. The subspace with the basis $B_1 = \langle e_1 \rangle$ is the $x$-axis. The subspace with the basis $B_2 = \langle e_2 \rangle$ is the $y$-axis. The subspace with the basis $B_3 = \langle e_3 \rangle$ is the $z$-axis. The fact that any member of $\mathbb{R}^3$ is expressible as a sum of vectors from these subspaces

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$$

is a reflection of the fact that $\mathcal{E}_3$ spans the space — this equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

has a solution for any $x, y, z \in \mathbb{R}$. And, the fact that each such expression is unique reflects that fact that $\mathcal{E}_3$ is linearly independent — any equation like the one above has a unique solution.

**4.6 Example** We don't have to take the basis vectors one at a time, the same idea works if we conglomerate them into larger sequences. Consider again the space $\mathbb{R}^3$ and the vectors from the standard basis $\mathcal{E}_3$. The subspace with the basis $B_1 = \langle e_1, e_3 \rangle$ is the $xz$-plane. The subspace with the basis $B_2 = \langle e_2 \rangle$ is the $y$-axis. As in the prior example, the fact that any member of the space is a sum of members of the two subspaces in one and only one way

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix}$$

is a reflection of the fact that these vectors form a basis — this system

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}) + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

has one and only one solution for any $x, y, z \in \mathbb{R}$.

These examples illustrate a natural way to decompose a space into a sum of subspaces in such a way that each vector decomposes uniquely into a sum of vectors from the parts. The next result says that this way is the only way.

**4.7 Definition** The *concatenation* of the sequences $B_1 = \langle \beta_{1,1}, \dots, \beta_{1,n_1} \rangle$, ..., $B_k = \langle \beta_{k,1}, \dots, \beta_{k,n_k} \rangle$ is their adjoinment.

$$B_1 {}^\frown B_2 {}^\frown \cdots {}^\frown B_k = \langle \beta_{1,1}, \dots, \beta_{1,n_1}, \beta_{2,1}, \dots, \beta_{k,n_k} \rangle$$

**4.8 Lemma** Let $V$ be a vector space that is the sum of some of its subspaces $V = W_1 + \cdots + W_k$. Let $B_1$, ..., $B_k$ be any bases for these subspaces. Then the following are equivalent.

(1) For every $v \in V$, the expression $v = w_1 + \cdots + w_k$ (with $w_i \in W_i$) is unique.
(2) The concatenation $B_1 {}^\frown \cdots {}^\frown B_k$ is a basis for $V$.
(3) The nonzero members of $\{w_1, \dots, w_k\}$ (with $w_i \in W_i$) form a linearly independent set — among nonzero vectors from different $W_i$'s, every linear relationship is trivial.

Proof. We will show that (1) $\implies$ (2), that (2) $\implies$ (3), and finally that (3) $\implies$ (1). For these arguments, observe that we can pass from a combination of $w$'s to a combination of $\beta$'s

$$\begin{aligned} d_1 w_1 + \cdots + d_k w_k \\ &= d_1(c_{1,1}\beta_{1,1} + \cdots + c_{1,n_1}\beta_{1,n_1}) + \cdots + d_k(c_{k,1}\beta_{k,1} + \cdots + c_{k,n_k}\beta_{k,n_k}) \\ &= d_1 c_{1,1} \cdot \beta_{1,1} + \cdots + d_k c_{k,n_k} \cdot \beta_{k,n_k} \qquad (*) \end{aligned}$$

and vice versa.

For (1) $\implies$ (2), assume that all decompositions are unique. We will show that $B_1{}^\frown\cdots{}^\frown B_k$ spans the space and is linearly independent. It spans the space because the assumption that $V = W_1 + \cdots + W_k$ means that every $v$ can be expressed as $v = w_1 + \cdots + w_k$, which translates by equation $(*)$ to an expression of $v$ as a linear combination of the $\beta$'s from the concatenation. For linear independence, consider this linear relationship.

$$0 = c_{1,1}\beta_{1,1} + \cdots + c_{k,n_k}\beta_{k,n_k}$$

Regroup as in $(*)$ (that is, take $d_1, \ldots, d_k$ to be 1 and move from bottom to top) to get the decomposition $0 = w_1 + \cdots + w_k$. Because of the assumption that decompositions are unique, and because the zero vector obviously has the decomposition $0 = 0 + \cdots + 0$, we now have that each $w_i$ is the zero vector. This means that $c_{i,1}\beta_{i,1} + \cdots + c_{i,n_i}\beta_{i,n_i} = 0$. Thus, since each $B_i$ is a basis, we have the desired conclusion that all of the $c$'s are zero.

For (2) $\implies$ (3), assume that $B_1{}^\frown\cdots{}^\frown B_k$ is a basis for the space. Consider a linear relationship among nonzero vectors from different $W_i$'s,

$$0 = \cdots + d_i w_i + \cdots$$

in order to show that it is trivial. (The relationship is written in this way because we are considering a combination of nonzero vectors from only some of the $W_i$'s; for instance, there might not be a $w_1$ in this combination.) As in $(*)$, $0 = \cdots + d_i(c_{i,1}\beta_{i,1} + \cdots + c_{i,n_i}\beta_{i,n_i}) + \cdots = \cdots + d_i c_{i,1} \cdot \beta_{i,1} + \cdots + d_i c_{i,n_i} \cdot \beta_{i,n_i} + \cdots$ and the linear independence of $B_1{}^\frown\cdots{}^\frown B_k$ gives that each coefficient $d_i c_{i,j}$ is zero. Now, $w_i$ is a nonzero vector, so at least one of the $c_{i,j}$'s is not zero, and thus $d_i$ is zero. This holds for each $d_i$, and therefore the linear relationship is trivial.

Finally, for (3) $\implies$ (1), assume that, among nonzero vectors from different $W_i$'s, any linear relationship is trivial. Consider two decompositions of a vector $v = w_1 + \cdots + w_k$ and $v = u_1 + \cdots + u_k$ in order to show that the two are the same. We have

$$0 = (w_1 + \cdots + w_k) - (u_1 + \cdots + u_k) = (w_1 - u_1) + \cdots + (w_k - u_k)$$

which violates the assumption unless each $w_i - u_i$ is the zero vector. Hence, decompositions are unique. QED

**4.9 Definition** A collection of subspaces $\{W_1, \ldots, W_k\}$ is *independent* if no nonzero vector from any $W_i$ is a linear combination of vectors from the other subspaces $W_1, \ldots, W_{i-1}, W_{i+1}, \ldots, W_k$.

**4.10 Definition** A vector space $V$ is the *direct sum* (or *internal direct sum*) of its subspaces $W_1, \ldots, W_k$ if $V = W_1 + W_2 + \cdots + W_k$ and the collection $\{W_1, \ldots, W_k\}$ is independent. We write $V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$.

**4.11 Example** The benchmark model fits: $\mathbb{R}^3 = x\text{-axis} \oplus y\text{-axis} \oplus z\text{-axis}$.

**4.12 Example** The space of $2\times 2$ matrices is this direct sum.

$$\{\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \mid a, d \in \mathbb{R}\} \oplus \{\begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} \mid b \in \mathbb{R}\} \oplus \{\begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} \mid c \in \mathbb{R}\}$$

It is the direct sum of subspaces in many other ways as well; direct sum decompositions are not unique.

**4.13 Corollary** The dimension of a direct sum is the sum of the dimensions of its summands.

Proof. In Lemma 4.8, the number of basis vectors in the concatenation equals the sum of the number of vectors in the subbases that make up the concatenation. QED

The special case of two subspaces is worth mentioning separately.

**4.14 Definition** When a vector space is the direct sum of two of its subspaces, then they are said to be *complements*.

**4.15 Lemma** A vector space $V$ is the direct sum of two of its subspaces $W_1$ and $W_2$ if and only if it is the sum of the two $V = W_1 + W_2$ and their intersection is trivial $W_1 \cap W_2 = \{\vec{0}\,\}$.

Proof. Suppose first that $V = W_1 \oplus W_2$. By definition, $V$ is the sum of the two. To show that the two have a trivial intersection, let $v$ be a vector from $W_1 \cap W_2$ and consider the equation $v = v$. On the left side of that equation is a member of $W_1$, and on the right side is a linear combination of members (actually, of only one member) of $W_2$. But the independence of the spaces then implies that $v = 0$, as desired.

For the other direction, suppose that $V$ is the sum of two spaces with a trivial intersection. To show that $V$ is a direct sum of the two, we need only show that the spaces are independent — no nonzero member of the first is expressible as a linear combination of members of the second, and vice versa. This is true because any relationship $\vec{w}_1 = c_1\vec{w}_{2,1} + \cdots + d_k\vec{w}_{2,k}$ (with $\vec{w}_1 \in W_1$ and $\vec{w}_{2,j} \in W_2$ for all $j$) shows that the vector on the left is also in $W_2$, since the right side is a combination of members of $W_2$. The intersection of these two spaces is trivial, so $\vec{w}_1 = \vec{0}$. The same argument works for any $\vec{w}_2$. QED

**4.16 Example** In the space $\mathbb{R}^2$, the $x$-axis and the $y$-axis are complements, that is, $\mathbb{R}^2 = x\text{-axis} \oplus y\text{-axis}$. A space can have more than one pair of complementary subspaces; another pair here are the subspaces consisting of the lines $y = x$ and $y = 2x$.

**4.17 Example** In the space $F = \{a\cos\theta + b\sin\theta \mid a, b \in \mathbb{R}\}$, the subspaces $W_1 = \{a\cos\theta \mid a \in \mathbb{R}\}$ and $W_2 = \{b\sin\theta \mid b \in \mathbb{R}\}$ are complements. In addition to the fact that a space like $F$ can have more than one pair of complementary subspaces, inside of the space a single subspace like $W_1$ can have more than one complement — another complement of $W_1$ is $W_3 = \{b\sin\theta + b\cos\theta \mid b \in \mathbb{R}\}$.

**4.18 Example** In $\mathbb{R}^3$, the $xy$-plane and the $yz$-planes are not complements, which is the point of the discussion following Example 4.4. One complement of the $xy$-plane is the $z$-axis. A complement of the $yz$-plane is the line through $(1, 1, 1)$.

**4.19 Example** Following Lemma 4.15, here is a natural question: is the simple sum $V = W_1 + \cdots + W_k$ also a direct sum if and only if the intersection of the subspaces is trivial? The answer is that if there are more than two subspaces then having a trivial intersection is not enough to guarantee unique decomposition (i.e., is not enough to ensure that the spaces are independent). In $\mathbb{R}^3$, let $W_1$ be the $x$-axis, let $W_2$ be the $y$-axis, and let $W_3$ be this.

$$W_3 = \{\begin{pmatrix} q \\ q \\ r \end{pmatrix} \mid q, r \in \mathbb{R}\}$$

The check that $\mathbb{R}^3 = W_1 + W_2 + W_3$ is easy. The intersection $W_1 \cap W_2 \cap W_3$ is trivial, but decompositions aren't unique.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y - x \\ 0 \end{pmatrix} + \begin{pmatrix} x \\ x \\ z \end{pmatrix} = \begin{pmatrix} x - y \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} y \\ y \\ z \end{pmatrix}$$

(This example also shows that this requirement is also not enough: that all pairwise intersections of the subspaces be trivial. See Exercise 30.)

In this subsection we have seen two ways to regard a space as built up from component parts. Both are useful; in particular, in this book the direct sum definition is needed to do the Jordan Form construction in the fifth chapter.

### Exercises

✓ **4.20** Decide if $\mathbb{R}^2$ is the direct sum of each $W_1$ and $W_2$.

**(a)** $W_1 = \{\begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R}\}$, $W_2 = \{\begin{pmatrix} x \\ x \end{pmatrix} \mid x \in \mathbb{R}\}$

**(b)** $W_1 = \{\begin{pmatrix} s \\ s \end{pmatrix} \mid s \in \mathbb{R}\}$, $W_2 = \{\begin{pmatrix} s \\ 1.1s \end{pmatrix} \mid s \in \mathbb{R}\}$

**(c)** $W_1 = \mathbb{R}^2$, $W_2 = \{\vec{0}\}$

**(d)** $W_1 = W_2 = \{\begin{pmatrix} t \\ t \end{pmatrix} \mid t \in \mathbb{R}\}$

**(e)** $W_1 = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R}\}$, $W_2 = \{\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix} \mid y \in \mathbb{R}\}$

✓ **4.21** Show that $\mathbb{R}^3$ is the direct sum of the $xy$-plane with each of these.
 **(a)** the $z$-axis
 **(b)** the line

$$\{\begin{pmatrix} z \\ z \\ z \end{pmatrix} \bigm| z \in \mathbb{R}\}$$

**4.22** Is $\mathcal{P}_2$ the direct sum of $\{a + bx^2 \bigm| a, b \in \mathbb{R}\}$ and $\{cx \bigm| c \in \mathbb{R}\}$?

✓ **4.23** In $\mathcal{P}_n$, the *even* polynomials are the members of this set

$$\mathcal{E} = \{p \in \mathcal{P}_n \bigm| p(-x) = p(x) \text{ for all } x\}$$

and the *odd* polynomials are the members of this set.

$$\mathcal{O} = \{p \in \mathcal{P}_n \bigm| p(-x) = -p(x) \text{ for all } x\}$$

Show that these are complementary subspaces.

**4.24** Which of these subspaces of $\mathbb{R}^3$

$$W_1\text{: the } x\text{-axis},\quad W_2\text{: the } y\text{-axis},\quad W_3\text{: the } z\text{-axis},$$
$$W_4\text{: the plane } x + y + z = 0,\quad W_5\text{: the } yz\text{-plane}$$

can be combined to
 **(a)** sum to $\mathbb{R}^3$? **(b)** direct sum to $\mathbb{R}^3$?

✓ **4.25** Show that $\mathcal{P}_n = \{a_0 \bigm| a_0 \in \mathbb{R}\} \oplus \ldots \oplus \{a_n x^n \bigm| a_n \in \mathbb{R}\}$.

**4.26** What is $W_1 + W_2$ if $W_1 \subseteq W_2$?

**4.27** Does Example 4.5 generalize? That is, is this true or false: if a vector space $V$ has a basis $\langle \beta_1, \ldots, \beta_n \rangle$ then it is the direct sum of the spans of the one-dimensional subspaces $V = [\{\beta_1\}] \oplus \ldots \oplus [\{\beta_n\}]$?

**4.28** Can $\mathbb{R}^4$ be decomposed as a direct sum in two different ways? Can $\mathbb{R}^1$?

**4.29** This exercise makes the notation of writing '+' between sets more natural. Prove that, where $W_1, \ldots, W_k$ are subspaces of a vector space,

$$W_1 + \cdots + W_k = \{w_1 + w_2 + \cdots + w_k \bigm| w_1 \in W_1, \ldots, w_k \in W_k\},$$

and so the sum of subspaces is the subspace of all sums.

**4.30** (Refer to Example 4.19. This exercise shows that the requirement that pariwise intersections be trivial is genuinely stronger than the requirement only that the intersection of all of the subspaces be trivial.) Give a vector space and three subspaces $W_1$, $W_2$, and $W_3$ such that the space is the sum of the subspaces, the intersection of all three subspaces $W_1 \cap W_2 \cap W_3$ is trivial, but the pairwise intersections $W_1 \cap W_2$, $W_1 \cap W_3$, and $W_2 \cap W_3$ are nontrivial.

✓ **4.31** Prove that if $V = W_1 \oplus \ldots \oplus W_k$ then $W_i \cap W_j$ is trivial whenever $i \neq j$. This shows that the first half of the proof of Lemma 4.15 extends to the case of more than two subspaces. (Example 4.19 shows that this implication does not reverse; the other half does not extend.)

**4.32** Recall that no linearly independent set contains the zero vector. Can an independent set of subspaces contain the trivial subspace?

✓ **4.33** Does every subspace have a complement?

✓ **4.34** Let $W_1, W_2$ be subspaces of a vector space.
 **(a)** Assume that the set $S_1$ spans $W_1$, and that the set $S_2$ spans $W_2$. Can $S_1 \cup S_2$ span $W_1 + W_2$? Must it?
 **(b)** Assume that $S_1$ is a linearly independent subset of $W_1$ and that $S_2$ is a linearly independent subset of $W_2$. Can $S_1 \cup S_2$ be a linearly independent subset of $W_1 + W_2$? Must it?

**4.35** When a vector space is decomposed as a direct sum, the dimensions of the subspaces add to the dimension of the space. The situation with a space that is given as the sum of its subspaces is not as simple. This exercise considers the two-subspace special case.
 **(a)** For these subspaces of $\mathcal{M}_{2\times2}$ find $W_1 \cap W_2$, $\dim(W_1 \cap W_2)$, $W_1 + W_2$, and $\dim(W_1 + W_2)$.

$$W_1 = \{\begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix} \bigm| c, d \in \mathbb{R}\} \qquad W_2 = \{\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \bigm| b, c \in \mathbb{R}\}$$

 **(b)** Suppose that $U$ and $W$ are subspaces of a vector space. Suppose that the sequence $\langle \beta_1, \ldots, \beta_k \rangle$ is a basis for $U \cap W$. Finally, suppose that the prior sequence has been expanded to give a sequence $\langle \mu_1, \ldots, \mu_j, \beta_1, \ldots, \beta_k \rangle$ that is a basis for $U$, and a sequence $\langle \beta_1, \ldots, \beta_k, \omega_1, \ldots, \omega_p \rangle$ that is a basis for $W$. Prove that this sequence

$$\langle \mu_1, \ldots, \mu_j, \beta_1, \ldots, \beta_k, \omega_1, \ldots, \omega_p \rangle$$

is a basis for for the sum $U + W$.
 **(c)** Conclude that $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$.

**(d)** Let $W_1$ and $W_2$ be eight-dimensional subspaces of a ten-dimensional space. List all values possible for $\dim(W_1 \cap W_2)$.

**4.36** Let $V = W_1 \oplus \ldots \oplus W_k$ and for each index $i$ suppose that $S_i$ is a linearly independent subset of $W_i$. Prove that the union of the $S_i$'s is linearly independent.

**4.37** A matrix is *symmetric* if for each pair of indices $i$ and $j$, the $i, j$ entry equals the $j, i$ entry. A matrix is *antisymmetric* if each $i, j$ entry is the negative of the $j, i$ entry.

**(a)** Give a symmetric $2\times 2$ matrix and an antisymmetric $2\times 2$ matrix. (*Remark.* For the second one, be careful about the entries on the diagonal.)

**(b)** What is the relationship between a square symmetric matrix and its transpose? Between a square antisymmetric matrix and its transpose?

**(c)** Show that $\mathcal{M}_{n\times n}$ is the direct sum of the space of symmetric matrices and the space of antisymmetric matrices.

**4.38** Let $W_1, W_2, W_3$ be subspaces of a vector space. Prove that $(W_1 \cap W_2) + (W_1 \cap W_3) \subseteq W_1 \cap (W_2 + W_3)$. Does the inclusion reverse?

**4.39** The example of the $x$-axis and the $y$-axis in $\mathbb{R}^2$ shows that $W_1 \oplus W_2 = V$ does not imply that $W_1 \cup W_2 = V$. Can $W_1 \oplus W_2 = V$ and $W_1 \cup W_2 = V$ happen?

✓ **4.40** Our model for complementary subspaces, the $x$-axis and the $y$-axis in $\mathbb{R}^2$, has one property not used here. Where $U$ is a subspace of $\mathbb{R}^n$ we define the *orthocomplement* of $U$ to be

$$U^{\perp} = \{v \in \mathbb{R}^n \bigm| v \cdot u = 0 \text{ for all } u \in U\}$$

(read "$U$ perp").

**(a)** Find the orthocomplement of the $x$-axis in $\mathbb{R}^2$.

**(b)** Find the orthocomplement of the $x$-axis in $\mathbb{R}^3$.

**(c)** Find the orthocomplement of the $xy$-plane in $\mathbb{R}^3$.

**(d)** Show that the orthocomplement of a subspace is a subspace.

**(e)** Show that if $W$ is the orthocomplement of $U$ then $U$ is the orthocomplement of $W$.

**(f)** Prove that a subspace and its orthocomplement have a trivial intersection.

**(g)** Conclude that for any $n$ and subspace $U \subseteq \mathbb{R}^n$ we have that $\mathbb{R}^n = U \oplus U^{\perp}$.

**(h)** Show that $\dim(U) + \dim(U^{\perp})$ equals the dimension of the enclosing space.

✓ **4.41** Consider Corollary 4.13. Does it work both ways—that is, supposing that $V = W_1 + \cdots + W_k$, is $V = W_1 \oplus \ldots \oplus W_k$ if and only if $\dim(V) = \dim(W_1) + \cdots + \dim(W_k)$?

**4.42** We know that if $V = W_1 \oplus W_2$ then there is a basis for $V$ that splits into a basis for $W_1$ and a basis for $W_2$. Can we make the stronger statement that every basis for $V$ splits into a basis for $W_1$ and a basis for $W_2$?

**4.43** We can ask about the algebra of the '+' operation.

**(a)** Is it commutative; is $W_1 + W_2 = W_2 + W_1$?

**(b)** Is it associative; is $(W_1 + W_2) + W_3 = W_1 + (W_2 + W_3)$?

**(c)** Let $W$ be a subspace of some vector space. Show that $W + W = W$.

**(d)** Must there be an identity element, a subspace $I$ such that $I + W = W + I = W$ for all subspaces $W$?

**(e)** Does left-cancelation hold: if $W_1 + W_2 = W_1 + W_3$ then $W_2 = W_3$? Right cancelation?

**4.44** Consider the algebraic properties of the direct sum operation.

**(a)** Does direct sum commute: does $V = W_1 \oplus W_2$ imply that $V = W_2 \oplus W_1$?

**(b)** Prove that direct sum is associative: $(W_1 \oplus W_2) \oplus W_3 = W_1 \oplus (W_2 \oplus W_3)$.

**(c)** Show that $\mathbb{R}^3$ is the direct sum of the three axes (the relevance here is that by the previous item, we needn't specify which two of the threee axes are combined first).

**(d)** Does the direct sum operation left-cancel: does $W_1 \oplus W_2 = W_1 \oplus W_3$ imply $W_2 = W_3$? Does it right-cancel?

**(e)** There is an identity element with respect to this operation. Find it.

**(f)** Do some, or all, subspaces have inverses with respect to this operation: is there a subspace $W$ of some vector space such that there is a subspace $U$ with the property that $U \oplus W$ equals the identity element from the prior item?

Chapter Three

# Linear Maps between Vector Spaces

# I Homomorphisms

The definition of isomorphism has two conditions. In this section we will consider the second one, that the map must preserve the algebraic structure of the space. We will focus on this condition by studying maps that are required only to preserve structure; that is, maps that are not required to be correspondences.

Experience shows that this kind of map is tremendously useful in the study of vector spaces. For one thing, as we shall see in the second subsection below, while isomorphisms describe how spaces are the same, these maps describe how spaces can be thought of as alike.

## I.1 Definition

**1.1 Definition** A function between vector spaces $h\colon V \to W$ that preserves the operations of addition

$$\text{if } \vec{v}_1, \vec{v}_2 \in V \text{ then } h(\vec{v}_1 + \vec{v}_2) = h(\vec{v}_1) + h(\vec{v}_2)$$

and scalar multiplication

$$\text{if } \vec{v} \in V \text{ and } r \in \mathbb{R} \text{ then } h(r \cdot \vec{v}) = r \cdot h(\vec{v})$$

is a *homomorphism* or *linear map*.

**1.2 Example** The projection map $\pi\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \end{pmatrix}$$

is a homomorphism. It preserves addition

$$\pi(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}) = \pi(\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}) = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} = \pi(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}) + \pi(\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix})$$

and scalar multiplication.

$$\pi(r \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}) = \pi(\begin{pmatrix} rx_1 \\ ry_1 \\ rz_1 \end{pmatrix}) = \begin{pmatrix} rx_1 \\ ry_1 \end{pmatrix} = r \cdot \pi(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix})$$

This map is not an isomorphism since it is not one-to-one. For instance, both $\vec{0}$ and $\vec{e}_3$ in $\mathbb{R}^3$ are mapped to the zero vector in $\mathbb{R}^2$.

**1.3 Example** Of course, the domain and codomain might be other than spaces of column vectors. Both of these are homomorphisms; the verifications are straightforward.

(1) $f_1\colon \mathcal{P}_2 \to \mathcal{P}_3$ given by

$$a_0 + a_1 x + a_2 x^2 \;\mapsto\; a_0 x + (a_1/2)x^2 + (a_2/3)x^3$$

(2) $f_2\colon \mathcal{M}_{2\times 2} \to \mathbb{R}$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + d$$

**1.4 Example** Between any two spaces there is a *zero homomorphism*, mapping every vector in the domain to the zero vector in the codomain.

**1.5 Example** These two suggest why we use the term 'linear map'.

(1) The map $g\colon \mathbb{R}^3 \to \mathbb{R}$ given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \overset{g}{\longmapsto} 3x + 2y - 4.5z$$

is linear (i.e., is a homomorphism). In contrast, the map $\hat{g}\colon \mathbb{R}^3 \to \mathbb{R}$ given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \overset{\hat{g}}{\longmapsto} 3x + 2y - 4.5z + 1$$

is not; for instance,

$$\hat{g}(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = 4 \quad \text{while} \quad \hat{g}(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}) + \hat{g}(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = 5$$

(to show that a map is not linear we need only produce one example of a linear combination that is not preserved).

(2) The first of these two maps $t_1, t_2\colon \mathbb{R}^3 \to \mathbb{R}^2$ is linear while the second is not.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \overset{t_1}{\longmapsto} \begin{pmatrix} 5x - 2y \\ x + y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \overset{t_2}{\longmapsto} \begin{pmatrix} 5x - 2y \\ xy \end{pmatrix}$$

Finding an example that the second fails to preserve structure is easy.

What distinguishes the homomorphisms is that the coordinate functions are linear combinations of the arguments. See also Exercise 23.

Obviously, any isomorphism is a homomorphism — an isomorphism is a homomorphism that is also a correspondence. So, one way to think of the 'homomorphism' idea is that it is a generalization of 'isomorphism', motivated by the observation that many of the properties of isomorphisms have only to do with the map's structure preservation property and not to do with it being a correspondence. As examples, these two results from the prior section do not use one-to-one-ness or onto-ness in their proof, and therefore apply to any homomorphism.

**1.6 Lemma** A homomorphism sends a zero vector to a zero vector.

**1.7 Lemma** Each of these is a necessary and sufficient condition for $f\colon V \to W$ to be a homomorphism.

(1) $f(c_1 \cdot \vec{v}_1 + c_2 \cdot \vec{v}_2) = c_1 \cdot f(\vec{v}_1) + c_2 \cdot f(\vec{v}_2)$ for any $c_1, c_2 \in \mathbb{R}$ and $\vec{v}_1, \vec{v}_2 \in V$

(2) $f(c_1 \cdot \vec{v}_1 + \cdots + c_n \cdot \vec{v}_n) = c_1 \cdot f(\vec{v}_1) + \cdots + c_n \cdot f(\vec{v}_n)$ for any $c_1, \ldots, c_n \in \mathbb{R}$ and $\vec{v}_1, \ldots, \vec{v}_n \in V$

Part (1) is often used to check that a function is linear.

**1.8 Example** The map $f\colon \mathbb{R}^2 \to \mathbb{R}^4$ given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \overset{f}{\longmapsto} \begin{pmatrix} x/2 \\ 0 \\ x + y \\ 3y \end{pmatrix}$$

satisfies (1) of the prior result

$$\begin{pmatrix} r_1(x_1/2) + r_2(x_2/2) \\ 0 \\ r_1(x_1 + y_1) + r_2(x_2 + y_2) \\ r_1(3y_1) + r_2(3y_2) \end{pmatrix} = r_1 \begin{pmatrix} x_1/2 \\ 0 \\ x_1 + y_1 \\ 3y_1 \end{pmatrix} + r_2 \begin{pmatrix} x_2/2 \\ 0 \\ x_2 + y_2 \\ 3y_2 \end{pmatrix}$$

and so it is a homomorphism.

However, some of the results that we have seen for isomorphisms fail to hold for homomorphisms in general. Consider the theorem that an isomorphism between spaces gives a correspondence between their bases. Homomorphisms do not give any such correspondence; Example 1.2 shows that there is no such correspondence, and another example is the zero map between any two nontrivial spaces. Instead, for homomorphisms a weaker but still very useful result holds.

**1.9 Theorem** A homomorphism is determined by its action on a basis. That is, if $\langle\beta_1,\dots,\beta_n\rangle$ is a basis of a vector space $V$ and $w_1,\dots,w_n$ are (perhaps not distinct) elements of a vector space $W$ then there exists a homomorphism from $V$ to $W$ sending $\beta_1$ to $w_1$, …, and $\beta_n$ to $w_n$, and that homomorphism is unique.

Proof. We will define the map by associating $\beta_1$ with $w_1$, etc., and then extending linearly to all of the domain. That is, where $v = c_1\beta_1 + \cdots + c_n\beta_n$, the map $h\colon V\to W$ is given by $h(v) = c_1w_1 + \cdots + c_nw_n$. This is well-defined because, with respect to the basis, the representation of each domain vector $v$ is unique.

This map is a homomorphism since it preserves linear combinations; where $v_1 = c_1\beta_1 + \cdots + c_n\beta_n$ and $v_2 = d_1\beta_1 + \cdots + d_n\beta_n$, we have this.

$$\begin{aligned} h(r_1v_1 + r_2v_2) &= h((r_1c_1 + r_2d_1)\beta_1 + \cdots + (r_1c_n + r_2d_n)\beta_n) \\ &= (r_1c_1 + r_2d_1)w_1 + \cdots + (r_1c_n + r_2d_n)w_n \\ &= r_1h(v_1) + r_2h(v_2) \end{aligned}$$

And, this map is unique since if $\hat{h}\colon V\to W$ is another homomorphism such that $\hat{h}(\beta_i) = w_i$ for each $i$ then $h$ and $\hat{h}$ agree on all of the vectors in the domain.

$$\begin{aligned} \hat{h}(v) &= \hat{h}(c_1\beta_1 + \cdots + c_n\beta_n) \\ &= c_1\hat{h}(\beta_1) + \cdots + c_n\hat{h}(\beta_n) \\ &= c_1w_1 + \cdots + c_nw_n \\ &= h(v) \end{aligned}$$

Thus, $h$ and $\hat{h}$ are the same map. QED

**1.10 Example** This result says that we can construct a homomorphism by fixing a basis for the domain and specifying where the map sends those basis vectors. For instance, if we specify a map $h\colon\mathbb{R}^2\to\mathbb{R}^2$ that acts on the standard basis $\mathcal{E}_2$ in this way

$$h(\begin{pmatrix}1\\0\end{pmatrix}) = \begin{pmatrix}-1\\1\end{pmatrix} \quad\text{and}\quad h(\begin{pmatrix}0\\1\end{pmatrix}) = \begin{pmatrix}-4\\4\end{pmatrix}$$

then the action of $h$ on any other member of the domain is also specified. For instance, the value of $h$ on this argument

$$h(\begin{pmatrix}3\\-2\end{pmatrix}) = h(3\cdot\begin{pmatrix}1\\0\end{pmatrix} - 2\cdot\begin{pmatrix}0\\1\end{pmatrix}) = 3\cdot h(\begin{pmatrix}1\\0\end{pmatrix}) - 2\cdot h(\begin{pmatrix}0\\1\end{pmatrix}) = \begin{pmatrix}5\\-5\end{pmatrix}$$

is a direct consequence of the value of $h$ on the basis vectors.

Later in this chapter we shall develop a scheme, using matrices, that is convienent for computations like this one.

Just as the isomorphisms of a space with itself are useful and interesting, so too are the homomorphisms of a space with itself.

**1.11 Definition** A linear map from a space into itself $t\colon V\to V$ is a *linear transformation.*

**1.12 Remark** In this book we use 'linear transformation' only in the case where the codomain equals the domain, but it is widely used in other texts as a general synonym for 'homomorphism'.

**1.13 Example** The map on $\mathbb{R}^2$ that projects all vectors down to the $x$-axis

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$$

is a linear transformation.

**1.14 Example** The derivative map $d/dx\colon \mathcal{P}_n \to \mathcal{P}_n$

$$a_0 + a_1 x + \cdots + a_n x^n \overset{d/dx}{\longmapsto} a_1 + 2a_2 x + 3a_3 x^2 + \cdots + n a_n x^{n-1}$$

is a linear transformation, as this result from calculus notes: $d(c_1 f + c_2 g)/dx = c_1\,(df/dx) + c_2\,(dg/dx)$.

**1.15 Example** The matrix transpose map

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

is a linear transformation of $\mathcal{M}_{2\times 2}$. Note that this transformation is one-to-one and onto, and so in fact it is an automorphism.

We finish this subsection about maps by recalling that we can linearly combine maps. For instance, for these maps from $\mathbb{R}^2$ to itself

$$\begin{pmatrix} x \\ y \end{pmatrix} \overset{f}{\longmapsto} \begin{pmatrix} 2x \\ 3x - 2y \end{pmatrix} \quad\text{and}\quad \begin{pmatrix} x \\ y \end{pmatrix} \overset{g}{\longmapsto} \begin{pmatrix} 0 \\ 5x \end{pmatrix}$$

the linear combination $5f - 2g$ is also a map from $R^2$ to itself.

$$\begin{pmatrix} x \\ y \end{pmatrix} \overset{5f-2g}{\longmapsto} \begin{pmatrix} 10x \\ 5x - 10y \end{pmatrix}$$

**1.16 Lemma** For vector spaces $V$ and $W$, the set of linear functions from $V$ to $W$ is itself a vector space, a subspace of the space of all functions from $V$ to $W$. It is denoted $\mathcal{L}(V, W)$.

Proof. This set is non-empty because it contains the zero homomorphism. So to show that it is a subspace we need only check that it is closed under linear combinations. Let $f, g\colon V \to W$ be linear. Then their sum is linear

$$\begin{aligned} (f + g)(c_1 v_1 + c_2 v_2) &= c_1 f(v_1) + c_2 f(v_2) + c_1 g(v_1) + c_2 g(v_2) \\ &= c_1 \big(f + g\big)(v_1) + c_2 \big(f + g\big)(v_2) \end{aligned}$$

and any scalar multiple is also linear.

$$\begin{aligned} (r \cdot f)(c_1 v_1 + c_2 v_2) &= r(c_1 f(v_1) + c_2 f(v_2)) \\ &= c_1 (r \cdot f)(v_1) + c_2 (r \cdot f)(v_2) \end{aligned}$$

Hence $\mathcal{L}(V, W)$ is a subspace. QED

We started this section by isolating the structure preservation property of isomorphisms. That is, we defined homomorphisms as a generalization of isomorphisms. Some of the properties that we studied for isomorphisms carried over unchanged, while others were adapted to this more general setting.

It would be a mistake, though, to view this new notion of homomorphism as derived from, or somehow secondary to, that of isomorphism. In the rest of this chapter we shall work mostly with homomorphisms, partly because any statement made about homomorphisms is automatically true about isomorphisms, but more because, while the isomorphism concept is perhaps more natural, experience shows that the homomorphism concept is actually more fruitful and more central to further progress.

### Exercises

✓ **1.17** Decide if each $h\colon \mathbb{R}^3 \to \mathbb{R}^2$ is linear.

**(a)** $h(\begin{pmatrix}x\\y\\z\end{pmatrix})=\begin{pmatrix}x\\x+y+z\end{pmatrix}$ **(b)** $h(\begin{pmatrix}x\\y\\z\end{pmatrix})=\begin{pmatrix}0\\0\end{pmatrix}$ **(c)** $h(\begin{pmatrix}x\\y\\z\end{pmatrix})=\begin{pmatrix}1\\1\end{pmatrix}$ **(d)** $h(\begin{pmatrix}x\\y\\z\end{pmatrix})=\begin{pmatrix}2x+y\\3y-4z\end{pmatrix}$

✓ **1.18** Decide if each map $h\colon \mathcal{M}_{2\times 2}\to\mathbb{R}$ is linear.

**(a)** $h(\begin{pmatrix}a&b\\c&d\end{pmatrix})=a+d$

**(b)** $h(\begin{pmatrix}a&b\\c&d\end{pmatrix})=ad-bc$

**(c)** $h(\begin{pmatrix}a&b\\c&d\end{pmatrix})=2a+3b+c-d$

**(d)** $h(\begin{pmatrix}a&b\\c&d\end{pmatrix})=a^2+b^2$

✓ **1.19** Show that these two maps are homomorphisms.

**(a)** $d/dx\colon \mathcal{P}_3\to\mathcal{P}_2$ given by $a_0+a_1x+a_2x^2+a_3x^3$ maps to $a_1+2a_2x+3a_3x^2$

**(b)** $\int\colon \mathcal{P}_2\to\mathcal{P}_3$ given by $b_0+b_1x+b_2x^2$ maps to $b_0x+(b_1/2)x^2+(b_2/3)x^3$

Are these maps inverse to each other?

**1.20** Is (perpendicular) projection from $\mathbb{R}^3$ to the $xz$-plane a homomorphism? Projection to the $yz$-plane? To the $x$-axis? The $y$-axis? The $z$-axis? Projection to the origin?

**1.21** Show that, while the maps from Example 1.3 preserve linear operations, they are not isomorphisms.

**1.22** Is an identity map a linear transformation?

✓ **1.23** Stating that a function is 'linear' is different than stating that its graph is a line.

**(a)** The function $f_1\colon\mathbb{R}\to\mathbb{R}$ given by $f_1(x)=2x-1$ has a graph that is a line. Show that it is not a linear function.

**(b)** The function $f_2\colon\mathbb{R}^2\to\mathbb{R}$ given by

$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto x+2y$$

does not have a graph that is a line. Show that it is a linear function.

✓ **1.24** Part of the definition of a linear function is that it respects addition. Does a linear function respect subtraction?

**1.25** Assume that $h$ is a linear transformation of $V$ and that $\langle\beta_1,\dots,\beta_n\rangle$ is a basis of $V$. Prove each statement.

**(a)** If $h(\beta_i)=0$ for each basis vector then $h$ is the zero map.

**(b)** If $h(\beta_i)=\beta_i$ for each basis vector then $h$ is the identity map.

**(c)** If there is a scalar $r$ such that $h(\beta_i)=r\cdot\beta_i$ for each basis vector then $h(v)=r\cdot v$ for all vectors in $V$.

✓ **1.26** Consider the vector space $\mathbb{R}^+$ where vector addition and scalar multiplication are not the ones inherited from $\mathbb{R}$ but rather are these: $a+b$ is the product of $a$ and $b$, and $r\cdot a$ is the $r$-th power of $a$. (This was shown to be a vector space in an earlier exercise.) Verify that the natural logarithm map $\ln\colon\mathbb{R}^+\to\mathbb{R}$ is a homomorphism between these two spaces. Is it an isomorphism?

✓ **1.27** Consider this transformation of $\mathbb{R}^2$.

$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x/2\\y/3\end{pmatrix}$$

Find the image under this map of this ellipse.

$$\{\begin{pmatrix}x\\y\end{pmatrix}\bigm|(x^2/4)+(y^2/9)=1\}$$

✓ **1.28** Imagine a rope wound around the earth's equator so that it fits snugly (suppose that the earth is a sphere). How much extra rope must be added to raise the circle to a constant six feet off the ground?

✓ **1.29** Verify that this map $h\colon\mathbb{R}^3\to\mathbb{R}$

$$\begin{pmatrix}x\\y\\z\end{pmatrix}\mapsto\begin{pmatrix}x\\y\\z\end{pmatrix}\cdot\begin{pmatrix}3\\-1\\-1\end{pmatrix}=3x-y-z$$

is linear. Generalize.

**1.30** Show that every homomorphism from $\mathbb{R}^1$ to $\mathbb{R}^1$ acts via multiplication by a scalar. Conclude that every nontrivial linear transformation of $\mathbb{R}^1$ is an isomorphism. Is that true for transformations of $\mathbb{R}^2$? $\mathbb{R}^n$?

**1.31** (a) Show that for any scalars $a_{1,1},\ldots,a_{m,n}$ this map $h\colon \mathbb{R}^n\to\mathbb{R}^m$ is a homomorphism.

$$\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\mapsto\begin{pmatrix}a_{1,1}x_1+\cdots+a_{1,n}x_n\\ \vdots\\ a_{m,1}x_1+\cdots+a_{m,n}x_n\end{pmatrix}$$

(b) Show that for each $i$, the $i$-th derivative operator $d^i/dx^i$ is a linear transformation of $\mathcal{P}_n$. Conclude that for any scalars $c_k,\ldots,c_0$ this map is a linear transformation of that space.

$$f\mapsto\frac{d^k}{dx^k}f+c_{k-1}\frac{d^{k-1}}{dx^{k-1}}f+\cdots+c_1\frac{d}{dx}f+c_0f$$

**1.32** Lemma 1.16 shows that a sum of linear functions is linear and that a scalar multiple of a linear function is linear. Show also that a composition of linear functions is linear.

✓ **1.33** Where $f\colon V\to W$ is linear, suppose that $f(\vec{v}_1)=\vec{w}_1,\ \ldots,\ f(\vec{v}_n)=\vec{w}_n$ for some vectors $\vec{w}_1,\ \ldots,\ \vec{w}_n$ from $W$.
(a) If the set of $\vec{w}$'s is independent, must the set of $\vec{v}$'s also be independent?
(b) If the set of $\vec{v}$'s is independent, must the set of $\vec{w}$'s also be independent?
(c) If the set of $\vec{w}$'s spans $W$, must the set of $\vec{v}$'s span $V$?
(d) If the set of $\vec{v}$'s spans $V$, must the set of $\vec{w}$'s span $W$?

**1.34** Generalize Example 1.15 by proving that the matrix transpose map is linear. What is the domain and codomain?

**1.35** (a) Where $\vec{u},\vec{v}\in\mathbb{R}^n$, the line segment connecting them is defined to be the set $\ell=\{t\cdot\vec{u}+(1-t)\cdot\vec{v}\mid t\in[0..1]\}$. Show that the image, under a homomorphism $h$, of the segment between $\vec{u}$ and $\vec{v}$ is the segment between $h(\vec{u})$ and $h(\vec{v})$.
(b) A subset of $\mathbb{R}^n$ is *convex* if, for any two points in that set, the line segment joining them lies entirely in that set. (The inside of a sphere is convex while the skin of a sphere is not.) Prove that linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ preserve the property of set convexity.

✓ **1.36** Let $h\colon\mathbb{R}^n\to\mathbb{R}^m$ be a homomorphism.
(a) Show that the image under $h$ of a line in $\mathbb{R}^n$ is a (possibly degenerate) line in $\mathbb{R}^n$.
(b) What happens to a $k$-dimensional linear surface?

**1.37** Prove that the restriction of a homomorphism to a subspace of its domain is another homomorphism.

**1.38** Assume that $h\colon V\to W$ is linear.
(a) Show that the *rangespace* of this map $\{h(\vec{v})\mid\vec{v}\in V\}$ is a subspace of the codomain $W$.
(b) Show that the *nullspace* of this map $\{\vec{v}\in V\mid h(\vec{v})=\vec{0}_W\}$ is a subspace of the domain $V$.
(c) Show that if $U$ is a subspace of the domain $V$ then its image $\{h(\vec{u})\mid\vec{u}\in U\}$ is a subspace of the codomain $W$. This generalizes the first item.
(d) Generalize the second item.

**1.39** Consider the set of isomorphisms from a vector space to itself. Is this a subspace of the space $\mathcal{L}(V,V)$ of homomorphisms from the space to itself?

**1.40** Does Theorem 1.9 need that $\langle\vec{\beta}_1,\ldots,\vec{\beta}_n\rangle$ is a basis? That is, can we still get a well-defined and unique homomorphism if we drop either the condition that the set of $\vec{\beta}$'s be linearly independent, or the condition that it span the domain?

**1.41** Let $V$ be a vector space and assume that the maps $f_1,f_2\colon V\to\mathbb{R}^1$ are linear.
(a) Define a map $F\colon V\to\mathbb{R}^2$ whose component functions are the given linear ones.

$$\vec{v}\mapsto\begin{pmatrix}f_1(\vec{v})\\ f_2(\vec{v})\end{pmatrix}$$

Show that $F$ is linear.
(b) Does the converse hold—is any linear map from $V$ to $\mathbb{R}^2$ made up of two linear component maps to $\mathbb{R}^1$?
(c) Generalize.

## I.2 Rangespace and Nullspace

Isomorphisms and homomorphisms both preserve structure. The difference is that homomorphisms needn't be onto and needn't be one-to-one. This means that homomorphisms are a more general kind of map, subject to fewer restrictions than isomorphisms. We will examine what can happen with homomorphisms that is prevented by the extra restrictions satisfied by isomorphisms.

We first consider the effect of dropping the onto requirement, of not requiring as part of the definition that a homomorphism be onto its codomain. For instance, the injection map $\iota\colon \mathbb{R}^2 \to \mathbb{R}^3$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

is not an isomorphism because it is not onto. Of course, being a function, a homomorphism is onto some set, namely its range; the map $\iota$ is onto the $xy$-plane subset of $\mathbb{R}^3$.

**2.1 Lemma** Under a homomorphism, the image of any subspace of the domain is a subspace of the codomain. In particular, the image of the entire space, the range of the homomorphism, is a subspace of the codomain.

Proof. Let $h\colon V \to W$ be linear and let $S$ be a subspace of the domain $V$. The image $h(S)$ is a subset of the codomain $W$. It is nonempty because $S$ is nonempty and thus to show that $h(S)$ is a subspace of $W$ we need only show that it is closed under linear combinations of two vectors. If $h(s_1)$ and $h(s_2)$ are members of $h(S)$ then $c_1 \cdot h(s_1) + c_2 \cdot h(s_2) = h(c_1 \cdot s_1) + h(c_2 \cdot s_2) = h(c_1 \cdot s_1 + c_2 \cdot s_2)$ is also a member of $h(S)$ because it is the image of $c_1 \cdot s_1 + c_2 \cdot s_2$ from $S$. QED

**2.2 Definition** The *rangespace* of a homomorphism $h\colon V \to W$ is

$$\mathscr{R}(h) = \{h(v) \mid v \in V\}$$

sometimes denoted $h(V)$. The dimension of the rangespace is the map's *rank*.

(We shall soon see the connection between the rank of a map and the rank of a matrix.)

**2.3 Example** Recall that the derivative map $d/dx\colon \mathcal{P}_3 \to \mathcal{P}_3$ given by $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_1 + 2a_2x + 3a_3x^2$ is linear. The rangespace $\mathscr{R}(d/dx)$ is the set of quadratic polynomials $\{r + sx + tx^2 \mid r, s, t \in \mathbb{R}\}$. Thus, the rank of this map is three.

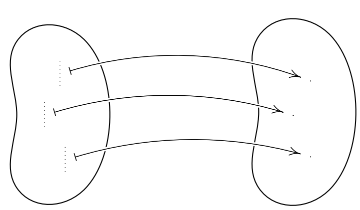
**2.4 Example** With this homomorphism $h\colon \mathcal{M}_{2\times 2} \to \mathcal{P}_3$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a + b + 2d) + 0x + cx^2 + cx^3$$

an image vector in the range can have any constant term, must have an $x$ coefficient of zero, and must have the same coefficient of $x^2$ as of $x^3$. That is, the rangespace is $\mathscr{R}(h) = \{r + 0x + sx^2 + sx^3 \mid r, s \in \mathbb{R}\}$ and so the rank is two.

The prior result shows that, in passing from the definition of isomorphism to the more general definition of homomorphism, omitting the 'onto' requirement doesn't make an essential difference. Any homomorphism is onto its rangespace.

However, omitting the 'one-to-one' condition does make a difference. A homomorphism may have many elements of the domain that map to one element of the codomain. Below is a "bean " sketch of a many-to-one map between sets.* It shows three elements of the codomain that are each the image of many members of the domain.



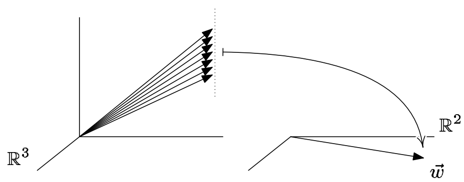
*More information on many-to-one maps is in the appendix.

Recall that for any function $h: V \to W$, the set of elements of $V$ that are mapped to $w \in W$ is the *inverse image* $h^{-1}(w) = \{v \in V \mid h(v) = w\}$. Above, the three sets of many elements on the left are inverse images.

**2.5 Example** Consider the projection $\pi: \mathbb{R}^3 \to \mathbb{R}^2$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \end{pmatrix}$$

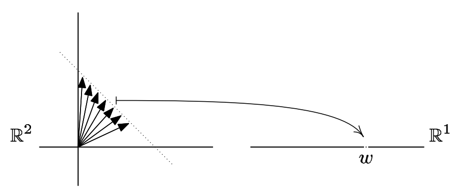
which is a homomorphism that is many-to-one. In this instance, an inverse image set is a vertical line of vectors in the domain.




**2.6 Example** This homomorphism $h: \mathbb{R}^2 \to \mathbb{R}^1$

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{h}{\longmapsto} x + y$$

is also many-to-one; for a fixed $w \in \mathbb{R}^1$, the inverse image $h^{-1}(w)$



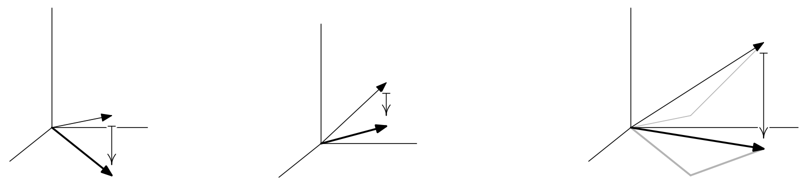

is the set of plane vectors whose components add to $w$.

The above examples have only to do with the fact that we are considering functions, specifically, many-to-one functions. They show the inverse images as sets of vectors that are related to the image vector $w$. But these are more than just arbitrary functions, they are homomorphisms; what do the two preservation conditions say about the relationships?

In generalizing from isomorphisms to homomorphisms by dropping the one-to-one condition, we lose the property that we've stated intuitively as: the domain is "the same as" the range. That is, we lose that the domain corresponds perfectly to the range in a one-vector-by-one-vector way. What we shall keep, as the examples below illustrate, is that a homomorphism describes a way in which the domain is "like", or "analgous to", the range.

**2.7 Example** We think of $\mathbb{R}^3$ as being like $\mathbb{R}^2$, except that vectors have an extra component. That is, we think of the vector with components $x$, $y$, and $z$ as like the vector with components $x$ and $y$. In defining the projection map $\pi$, we make precise which members of the domain we are thinking of as related to which members of the codomain.

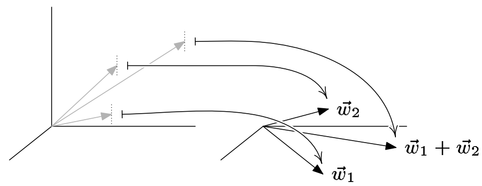
Understanding in what way the preservation conditions in the definition of homomorphism show that the domain elements are like the codomain elements is easiest if we draw $\mathbb{R}^2$ as the $xy$-plane inside of $\mathbb{R}^3$. (Of course, $\mathbb{R}^2$ is a set of two-tall vectors while the $xy$-plane is a set of three-tall vectors with a third component of zero, but there is an obvious correspondence.) Then, $\pi(v)$ is the "shadow" of $v$ in the plane and the preservation of addition property says that



$$\begin{pmatrix}x_1\\y_1\\z_1\end{pmatrix} \text{ above } \begin{pmatrix}x_1\\y_1\end{pmatrix} \text{ plus } \begin{pmatrix}x_2\\y_2\\z_2\end{pmatrix} \text{ above } \begin{pmatrix}x_2\\y_2\end{pmatrix} \text{ equals } \begin{pmatrix}x_1+y_1\\y_1+y_2\\z_1+z_2\end{pmatrix} \text{ above } \begin{pmatrix}x_1+x_2\\y_1+y_2\end{pmatrix}$$

Briefly, the shadow of a sum $\pi(v_1+v_2)$ equals the sum of the shadows $\pi(v_1)+\pi(v_2)$. (Preservation of scalar multiplication has a similar interpretation.)

Redrawing by separating the two spaces, moving the codomain $\mathbb{R}^2$ to the right, gives an uglier picture but one that is more faithful to the "bean" sketch.



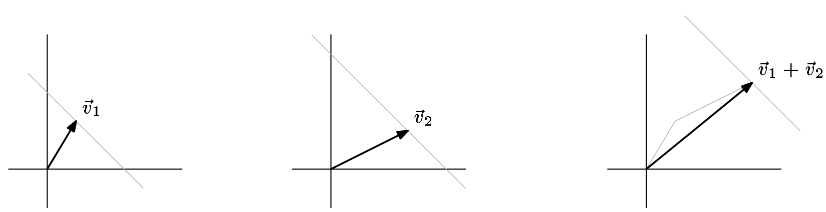

Again in this drawing, the vectors that map to $w_1$ lie in the domain in a vertical line (only one such vector is shown, in gray). Call any such member of this inverse image a "$w_1$ vector". Similarly, there is a vertical line of "$w_2$ vectors" and a vertical line of "$w_1+w_2$ vectors". Now, $\pi$ has the property that if $\pi(v_1)=w_1$ and $\pi(v_2)=w_2$ then $\pi(v_1+v_2)=\pi(v_1)+\pi(v_2)=w_1+w_2$. This says that the vector classes add, in the sense that any $w_1$ vector plus any $w_2$ vector equals a $w_1+w_2$ vector, (A similar statement holds about the classes under scalar multiplication.)

Thus, although the two spaces $\mathbb{R}^3$ and $\mathbb{R}^2$ are not isomorphic, $\pi$ describes a way in which they are alike: vectors in $\mathbb{R}^3$ add as do the associated vectors in $\mathbb{R}^2$ — vectors add as their shadows add.

**2.8 Example** A homomorphism can be used to express an analogy between spaces that is more subtle than the prior one. For the map

$$\begin{pmatrix}x\\y\end{pmatrix} \stackrel{h}{\longmapsto} x+y$$

from Example 2.6 fix two numbers $w_1, w_2$ in the range $\mathbb{R}$. A $v_1$ that maps to $w_1$ has components that add to $w_1$, that is, the inverse image $h^{-1}(w_1)$ is the set of vectors with endpoint on the diagonal line $x+y=w_1$. Call these the "$w_1$ vectors". Similarly, we have the "$w_2$ vectors" and the "$w_1+w_2$ vectors". Then the addition preservation property says that



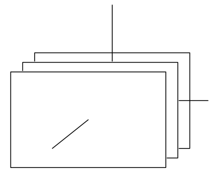

a "$w_1$ vector" plus a "$w_2$ vector" equals a "$w_1+w_2$ vector".

Restated, if a $w_1$ vector is added to a $w_2$ vector then the result is mapped by $h$ to a $w_1+w_2$ vector. Briefly, the image of a sum is the sum of the images. Even more briefly, $h(v_1+v_2)=h(v_1)+h(v_2)$. (The preservation of scalar multiplication condition has a similar restatement.)

**2.9 Example** The inverse images can be structures other than lines. For the linear map $h\colon\mathbb{R}^3\to\mathbb{R}^2$

$$\begin{pmatrix}x\\y\\z\end{pmatrix} \mapsto \begin{pmatrix}x\\x\end{pmatrix}$$

the inverse image sets are planes $x = 0$, $x = 1$, etc., perpendicular to the $x$-axis.



We won't describe how every homomorphism that we will use is an analogy because the formal sense that we make of "alike in that ..." is 'a homomorphism exists such that ...'. Nonetheless, the idea that a homomorphism between two spaces expresses how the domain's vectors fall into classes that act like the the range's vectors is a good way to view homomorphisms.

Another reason that we won't treat all of the homomorphisms that we see as above is that many vector spaces are hard to draw (e.g., a space of polynomials). However, there is nothing bad about gaining insights from those spaces that we are able to draw, especially when those insights extend to all vector spaces. We derive two such insights from the three examples 2.7, 2.8, and 2.9.

First, in all three examples, the inverse images are lines or planes, that is, linear surfaces. In particular, the inverse image of the range's zero vector is a line or plane through the origin — a subspace of the domain.

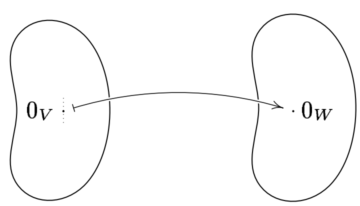
**2.10 Lemma** For any homomorphism, the inverse image of a subspace of the range is a subspace of the domain. In particular, the inverse image of the trivial subspace of the range is a subspace of the domain.

Proof. Let $h\colon V \to W$ be a homomorphism and let $S$ be a subspace of the rangespace $h$. Consider $h^{-1}(S) = \{v \in V \mid h(v) \in S\}$, the inverse image of the set $S$. It is nonempty because it contains $0_V$, since $h(0_V) = 0_W$, which is an element $S$, as $S$ is a subspace. To show that $h^{-1}(S)$ is closed under linear combinations, let $v_1$ and $v_2$ be elements, so that $h(v_1)$ and $h(v_2)$ are elements of $S$, and then $c_1v_1 + c_2v_2$ is also in the inverse image because $h(c_1v_1 + c_2v_2) = c_1h(v_1) + c_2h(v_2)$ is a member of the subspace $S$. QED

**2.11 Definition** The *nullspace* or *kernel* of a linear map $h\colon V \to W$ is the inverse image of $0_W$

$$\mathscr{N}(h) = h^{-1}(0_W) = \{v \in V \mid h(v) = 0_W\}.$$

The dimension of the nullspace is the map's *nullity*.




**2.12 Example** The map from Example 2.3 has this nullspace $\mathscr{N}(d/dx) = \{a_0 + 0x + 0x^2 + 0x^3 \mid a_0 \in \mathbb{R}\}$.

**2.13 Example** The map from Example 2.4 has this nullspace.

$$\mathscr{N}(h) = \{\begin{pmatrix} a & b \\ 0 & -(a+b)/2 \end{pmatrix} \mid a, b \in \mathbb{R}\}$$

Now for the second insight from the above pictures. In Example 2.7, each of the vertical lines is squashed down to a single point — $\pi$, in passing from the domain to the range, takes all of these one-dimensional vertical lines and "zeroes them out", leaving the range one dimension smaller than the domain. Similarly, in Example 2.8, the two-dimensional domain is mapped to a one-dimensional range by breaking the domain into lines (here, they are diagonal lines), and compressing each of those lines to a single member of the range. Finally, in Example 2.9, the domain breaks into planes which get "zeroed out", and so the map starts with a three-dimensional domain but ends with a one-dimensional range — this map "subtracts" two from the dimension. (Notice that, in this third example, the codomain is two-dimensional but the range of the map is only one-dimensional, and it is the dimension of the range that is of interest.)

**2.14 Theorem** A linear map's rank plus its nullity equals the dimension of its domain.

Proof. Let $h\colon V \to W$ be linear and let $B_N = \langle \beta_1, \ldots, \beta_k \rangle$ be a basis for the nullspace. Extend that to a basis $B_V = \langle \beta_1, \ldots, \beta_k, \beta_{k+1}, \ldots, \beta_n \rangle$ for the entire domain. We shall show that $B_R = \langle h(\beta_{k+1}), \ldots, h(\beta_n) \rangle$ is a basis for the rangespace. Then counting the size of these bases gives the result.

To see that $B_R$ is linearly independent, consider the equation $c_{k+1}h(\beta_{k+1}) + \cdots + c_n h(\beta_n) = 0_W$. This gives that $h(c_{k+1}\beta_{k+1} + \cdots + c_n\beta_n) = 0_W$ and so $c_{k+1}\beta_{k+1} + \cdots + c_n\beta_n$ is in the nullspace of $h$. As $B_N$ is a basis for this nullspace, there are scalars $c_1, \ldots, c_k \in \mathbb{R}$ satisfying this relationship.

$$c_1\beta_1 + \cdots + c_k\beta_k = c_{k+1}\beta_{k+1} + \cdots + c_n\beta_n$$

But $B_V$ is a basis for $V$ so each scalar equals zero. Therefore $B_R$ is linearly independent.

To show that $B_R$ spans the rangespace, consider $h(v) \in \mathscr{R}(h)$ and write $v$ as a linear combination $v = c_1\beta_1 + \cdots + c_n\beta_n$ of members of $B_V$. This gives $h(v) = h(c_1\beta_1 + \cdots + c_n\beta_n) = c_1h(\beta_1) + \cdots + c_kh(\beta_k) + c_{k+1}h(\beta_{k+1}) + \cdots + c_nh(\beta_n)$ and since $\beta_1, \ldots, \beta_k$ are in the nullspace, we have that $h(v) = 0 + \cdots + 0 + c_{k+1}h(\beta_{k+1}) + \cdots + c_nh(\beta_n)$. Thus, $h(v)$ is a linear combination of members of $B_R$, and so $B_R$ spans the space. QED

**2.15 Example** Where $h\colon \mathbb{R}^3 \to \mathbb{R}^4$ is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{h}{\longmapsto} \begin{pmatrix} x \\ 0 \\ y \\ 0 \end{pmatrix}$$

the rangespace and nullspace are

$$\mathscr{R}(h) = \{ \begin{pmatrix} a \\ 0 \\ b \\ 0 \end{pmatrix} \mid a, b \in \mathbb{R} \} \quad \text{and} \quad \mathscr{N}(h) = \{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \mid z \in \mathbb{R} \}$$

and so the rank of $h$ is two while the nullity is one.

**2.16 Example** If $t\colon \mathbb{R} \to \mathbb{R}$ is the linear transformation $x \mapsto -4x$, then the range is $\mathscr{R}(t) = \mathbb{R}^1$, and so the rank of $t$ is one and the nullity is zero.

**2.17 Corollary** The rank of a linear map is less than or equal to the dimension of the domain. Equality holds if and only if the nullity of the map is zero.

We know that an isomorphism exists between two spaces if and only if their dimensions are equal. Here we see that for a homomorphism to exist, the dimension of the range must be less than or equal to the dimension of the domain. For instance, there is no homomorphism from $\mathbb{R}^2$ onto $\mathbb{R}^3$. There are many homomorphisms from $\mathbb{R}^2$ into $\mathbb{R}^3$, but none is onto all of three-space.

The rangespace of a linear map can be of dimension strictly less than the dimension of the domain (Example 2.3's derivative transformation on $\mathcal{P}_3$ has a domain of dimension four but a range of dimension three). Thus, under a homomorphism, linearly independent sets in the domain may map to linearly dependent sets in the range (for instance, the derivative sends $\{1, x, x^2, x^3\}$ to $\{0, 1, 2x, 3x^2\}$). That is, under a homomorphism, independence may be lost. In contrast, dependence stays.

**2.18 Lemma** Under a linear map, the image of a linearly dependent set is linearly dependent.

Proof. Suppose that $c_1v_1 + \cdots + c_nv_n = 0_V$, with some $c_i$ nonzero. Then, because $h(c_1v_1 + \cdots + c_nv_n) = c_1h(v_1) + \cdots + c_nh(v_n)$ and because $h(0_V) = 0_W$, we have that $c_1h(v_1) + \cdots + c_nh(v_n) = 0_W$ with some nonzero $c_i$. QED

When is independence not lost? One obvious sufficient condition is when the homomorphism is an isomorphism. This condition is also necessary; see Exercise 35. We will finish this subsection comparing homomorphisms with isomorphisms by observing that a one-to-one homomorphism is an isomorphism from its domain onto its range.

**2.19 Definition** A linear map that is one-to-one is *nonsingular*.

(In the next section we will see the connection between this use of 'nonsingular' for maps and its familiar use for matrices.)

**2.20 Example** This nonsingular homomorphism $\iota\colon \mathbb{R}^2 \to \mathbb{R}^3$

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{\iota}{\longmapsto} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

gives the obvious correspondence between $\mathbb{R}^2$ and the $xy$-plane inside of $\mathbb{R}^3$.

The prior observation allows us to adapt some results about isomorphisms to this setting.

**2.21 Theorem** In an $n$-dimensional vector space $V$, these:
(1) $h$ is nonsingular, that is, one-to-one
(2) $h$ has a linear inverse
(3) $\mathscr{N}(h) = \{0\}$, that is, $\text{nullity}(h) = 0$
(4) $\text{rank}(h) = n$
(5) if $\langle \beta_1, \ldots, \beta_n \rangle$ is a basis for $V$ then $\langle h(\beta_1), \ldots, h(\beta_n) \rangle$ is a basis for $\mathscr{R}(h)$
are equivalent statements about a linear map $h\colon V \to W$.

Proof. We will first show that (1) $\iff$ (2). We will then show that (1) $\implies$ (3) $\implies$ (4) $\implies$ (5) $\implies$ (2).

For (1) $\implies$ (2), suppose that the linear map $h$ is one-to-one, and so has an inverse. The domain of that inverse is the range of $h$ and so a linear combination of two members of that domain has the form $c_1h(v_1) + c_2h(v_2)$. On that combination, the inverse $h^{-1}$ gives this.

$$\begin{aligned} h^{-1}(c_1h(v_1) + c_2h(v_2)) &= h^{-1}(h(c_1v_1 + c_2v_2)) \\ &= h^{-1} \circ h\,(c_1v_1 + c_2v_2) \\ &= c_1v_1 + c_2v_2 \\ &= c_1h^{-1} \circ h\,(v_1) + c_2h^{-1} \circ h\,(v_2) \\ &= c_1 \cdot h^{-1}(h(v_1)) + c_2 \cdot h^{-1}(h(v_2)) \end{aligned}$$

Thus the inverse of a one-to-one linear map is automatically linear. But this also gives the (1) $\implies$ (2) implication, because the inverse itself must be one-to-one.

Of the remaining implications, (1) $\implies$ (3) holds because any homomorphism maps $0_V$ to $0_W$, but a one-to-one map sends at most one member of $V$ to $0_W$.

Next, (3) $\implies$ (4) is true since rank plus nullity equals the dimension of the domain.

For (4) $\implies$ (5), to show that $\langle h(\beta_1), \ldots, h(\beta_n) \rangle$ is a basis for the rangespace we need only show that it is a spanning set, because by assumption the range has dimension $n$. Consider $h(v) \in \mathscr{R}(h)$. Expressing $v$ as a linear combination of basis elements produces $h(v) = h(c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n)$, which gives that $h(v) = c_1h(\beta_1) + \cdots + c_nh(\beta_n)$, as desired.

Finally, for the (5) $\implies$ (2) implication, assume that $\langle \beta_1, \ldots, \beta_n \rangle$ is a basis for $V$ so that $\langle h(\beta_1), \ldots, h(\beta_n) \rangle$ is a basis for $\mathscr{R}(h)$. Then every $w \in \mathscr{R}(h)$ a the unique representation $w = c_1h(\beta_1) + \cdots + c_nh(\beta_n)$. Define a map from $\mathscr{R}(h)$ to $V$ by

$$w \;\mapsto\; c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$$

(uniqueness of the representation makes this well-defined). Checking that it is linear and that it is the inverse of $h$ are easy. QED

We've now seen that a linear map shows how the structure of the domain is like that of the range. Such a map can be thought to organize the domain space into inverse images of points in the range. In the special case that the map is one-to-one, each inverse image is a single point and the map is an isomorphism between the domain and the range.

## Exercises

✓ **2.22** Let $h\colon \mathcal{P}_3 \to \mathcal{P}_4$ be given by $p(x) \mapsto x \cdot p(x)$. Which of these are in the nullspace? Which are in the rangespace?

**(a)** $x^3$ **(b)** $0$ **(c)** $7$ **(d)** $12x - 0.5x^3$ **(e)** $1 + 3x^2 - x^3$

✓ **2.23** Find the nullspace, nullity, rangespace, and rank of each map.

**(a)** $h\colon \mathbb{R}^2 \to \mathcal{P}_3$ given by

$$\begin{pmatrix} a \\ b \end{pmatrix} \mapsto a + ax + ax^2$$

**(b)** $h\colon \mathcal{M}_{2\times 2} \to \mathbb{R}$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + d$$

**(c)** $h\colon \mathcal{M}_{2\times 2} \to \mathcal{P}_2$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + b + c + dx^2$$

**(d)** the zero map $Z\colon \mathbb{R}^3 \to \mathbb{R}^4$

✓ **2.24** Find the nullity of each map.

**(a)** $h\colon \mathbb{R}^5 \to \mathbb{R}^8$ of rank five **(b)** $h\colon \mathcal{P}_3 \to \mathcal{P}_3$ of rank one **(c)** $h\colon \mathbb{R}^6 \to \mathbb{R}^3$, an onto map
**(d)** $h\colon \mathcal{M}_{3\times 3} \to \mathcal{M}_{3\times 3}$, onto

✓ **2.25** What is the nullspace of the differentiation transformation $d/dx\colon \mathcal{P}_n \to \mathcal{P}_n$? What is the nullspace of the second derivative, as a transformation of $\mathcal{P}_n$? The $k$-th derivative?

**2.26** Example 2.7 restates the first condition in the definition of homomorphism as 'the shadow of a sum is the sum of the shadows'. Restate the second condition in the same style.

**2.27** For the homomorphism $h\colon \mathcal{P}_3 \to \mathcal{P}_3$ given by $h(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0 + (a_0 + a_1)x + (a_2 + a_3)x^3$ find these.

**(a)** $\mathscr{N}(h)$ **(b)** $h^{-1}(2 - x^3)$ **(c)** $h^{-1}(1 + x^2)$

✓ **2.28** For the map $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(\begin{pmatrix} x \\ y \end{pmatrix}) = 2x + y$$

sketch these inverse image sets: $f^{-1}(-3)$, $f^{-1}(0)$, and $f^{-1}(1)$.

✓ **2.29** Each of these transformations of $\mathcal{P}_3$ is nonsingular. Find the inverse function of each.

**(a)** $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_0 + a_1x + 2a_2x^2 + 3a_3x^3$
**(b)** $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_0 + a_2x + a_1x^2 + a_3x^3$
**(c)** $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_1 + a_2x + a_3x^2 + a_0x^3$
**(d)** $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_0 + (a_0 + a_1)x + (a_0 + a_1 + a_2)x^2 + (a_0 + a_1 + a_2 + a_3)x^3$

**2.30** Describe the nullspace and rangespace of a transformation given by $v \mapsto 2v$.

**2.31** List all pairs $(\text{rank}(h), \text{nullity}(h))$ that are possible for linear maps from $\mathbb{R}^5$ to $\mathbb{R}^3$.

**2.32** Does the differentiation map $d/dx\colon \mathcal{P}_n \to \mathcal{P}_n$ have an inverse?

✓ **2.33** Find the nullity of the map $h\colon \mathcal{P}_n \to \mathbb{R}$ given by

$$a_0 + a_1x + \cdots + a_nx^n \mapsto \int_{x=0}^{x=1} a_0 + a_1x + \cdots + a_nx^n\,dx.$$

**2.34** **(a)** Prove that a homomorphism is onto if and only if its rank equals the dimension of its codomain.
**(b)** Conclude that a homomorphism between vector spaces with the same dimension is one-to-one if and only if it is onto.

**2.35** Show that a linear map is nonsingular if and only if it preserves linear independence.

**2.36** Corollary 2.17 says that for there to be an onto homomorphism from a vector space $V$ to a vector space $W$, it is necessary that the dimension of $W$ be less than or equal to the dimension of $V$. Prove that this condition is also sufficient; use Theorem 1.9 to show that if the dimension of $W$ is less than or equal to the dimension of $V$, then there is a homomorphism from $V$ to $W$ that is onto.

**2.37** Let $h\colon V \to \mathbb{R}$ be a homomorphism, but not the zero homomorphism. Prove that if $\langle \beta_1, \ldots, \beta_n \rangle$ is a basis for the nullspace and if $v \in V$ is not in the nullspace then $\langle v, \beta_1, \ldots, \beta_n \rangle$ is a basis for the entire domain $V$.

✓ **2.38** Recall that the nullspace is a subset of the domain and the rangespace is a subset of the codomain. Are they necessarily distinct? Is there a homomorphism that has a nontrivial intersection of its nullspace and its rangespace?

**2.39** Prove that the image of a span equals the span of the images. That is, where $h\colon V \to W$ is linear, prove that if $S$ is a subset of $V$ then $h(\langle S \rangle)$ equals $\langle h(S) \rangle$. This generalizes Lemma 2.1 since it shows that if $U$ is any subspace of $V$ then its image $\{h(u) \mid u \in U\}$ is a subspace of $W$, because the span of the set $U$ is $U$.

✓ **2.40** **(a)** Prove that for any linear map $h\colon V \to W$ and any $w \in W$, the set $h^{-1}(w)$ has the form

$$\{v + n \mid n \in \mathscr{N}(h)\}$$

for $v \in V$ with $h(v) = w$ (if $h$ is not onto then this set may be empty). Such a set is a *coset* of $\mathscr{N}(h)$ and is denoted $v + \mathscr{N}(h)$.

**(b)** Consider the map $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{t}{\longmapsto} \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

for some scalars $a$, $b$, $c$, and $d$. Prove that $t$ is linear.

**(c)** Conclude from the prior two items that for any linear system of the form

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$$

the solution set can be written (the vectors are members of $\mathbb{R}^2$)

$$\{p + h \mid h \text{ satisfies the associated homogeneous system}\}$$

where $p$ is a particular solution of that linear system (if there is no particular solution then the above set is empty).

**(d)** Show that this map $h\colon \mathbb{R}^n \to \mathbb{R}^m$ is linear

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{pmatrix}$$

for any scalars $a_{1,1}, \ldots, a_{m,n}$. Extend the conclusion made in the prior item.

**(e)** Show that the $k$-th derivative map is a linear transformation of $\mathcal{P}_n$ for each $k$. Prove that this map is a linear transformation of that space

$$f \mapsto \frac{d^k}{dx^k} f + c_{k-1} \frac{d^{k-1}}{dx^{k-1}} f + \cdots + c_1 \frac{d}{dx} f + c_0 f$$

for any scalars $c_k, \ldots, c_0$. Draw a conclusion as above.

**2.41** Prove that for any transformation $t\colon V \to V$ that is rank one, the map given by composing the operator with itself $t \circ t\colon V \to V$ satisfies $t \circ t = r \cdot t$ for some real number $r$.

**2.42** Show that for any space $V$ of dimension $n$, the *dual space*

$$\mathcal{L}(V, \mathbb{R}) = \{h\colon V \to \mathbb{R} \mid h \text{ is linear}\}$$

is isomorphic to $\mathbb{R}^n$. It is often denoted $V^*$. Conclude that $V^* \cong V$.

**2.43** Show that any linear map is the sum of maps of rank one.

**2.44** Is 'is homomorphic to' an equivalence relation? (*Hint:* the difficulty is to decide on an appropriate meaning for the quoted phrase.)

**2.45** Show that the rangespaces and nullspaces of powers of linear maps $t\colon V \to V$ form descending

$$V \supseteq \mathscr{R}(t) \supseteq \mathscr{R}(t^2) \supseteq \ldots$$

and ascending

$$\{\vec{0}\} \subseteq \mathscr{N}(t) \subseteq \mathscr{N}(t^2) \subseteq \ldots$$

chains. Also show that if $k$ is such that $\mathscr{R}(t^k) = \mathscr{R}(t^{k+1})$ then all following rangespaces are equal: $\mathscr{R}(t^k) = \mathscr{R}(t^{k+1}) = \mathscr{R}(t^{k+2}) \ldots$. Similarly, if $\mathscr{N}(t^k) = \mathscr{N}(t^{k+1})$ then $\mathscr{N}(t^k) = \mathscr{N}(t^{k+1}) = \mathscr{N}(t^{k+2}) = \ldots$.

# II Computing Linear Maps

The prior section shows that a linear map is determined by its action on a basis. In fact, the equation

$$h(v) = h(c_1 \cdot \beta_1 + \cdots + c_n \cdot \beta_n) = c_1 \cdot h(\beta_1) + \cdots + c_n \cdot h(\beta_n)$$

shows that, if we know the value of the map on the vectors in a basis, then we can compute the value of the map on any vector $v$ at all. We just need to find the $c$'s to express $v$ with respect to the basis.

This section gives the scheme that computes, from the representation of a vector in the domain $\mathrm{Rep}_B(v)$, the representation of that vector's image in the codomain $\mathrm{Rep}_D(h(v))$, using the representations of $h(\beta_1)$, ..., $h(\beta_n)$.

## II.1 Representing Linear Maps with Matrices

**1.1 Example** Consider a map $h$ with domain $\mathbb{R}^2$ and codomain $\mathbb{R}^3$ (fixing

$$B = \langle \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix} \rangle \quad \text{and} \quad D = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \rangle$$

as the bases for these spaces) that is determined by this action on the vectors in the domain's basis.

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} \stackrel{h}{\longmapsto} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 4 \end{pmatrix} \stackrel{h}{\longmapsto} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

To compute the action of this map on any vector at all from the domain, we first express $h(\beta_1)$ and $h(\beta_2)$ with respect to the codomain's basis:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{so} \quad \mathrm{Rep}_D(h(\beta_1)) = \begin{pmatrix} 0 \\ -1/2 \\ 1 \end{pmatrix}_D$$

and

$$\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{so} \quad \mathrm{Rep}_D(h(\beta_2)) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}_D$$

(these are easy to check). Then, as described in the preamble, for any member $v$ of the domain, we can express the image $h(v)$ in terms of the $h(\beta)$'s.

$$\begin{aligned}
h(v) &= h(c_1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 1 \\ 4 \end{pmatrix}) \\
&= c_1 \cdot h(\begin{pmatrix} 2 \\ 0 \end{pmatrix}) + c_2 \cdot h(\begin{pmatrix} 1 \\ 4 \end{pmatrix}) \\
&= c_1 \cdot (0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}) + c_2 \cdot (1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}) \\
&= (0c_1 + 1c_2) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-\frac{1}{2}c_1 - 1c_2) \cdot \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + (1c_1 + 0c_2) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}$$

Thus,

$$\text{with } \mathrm{Rep}_B(v) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \text{ then } \mathrm{Rep}_D(\,h(v)\,) = \begin{pmatrix} 0c_1 + 1c_2 \\ -(1/2)c_1 - 1c_2 \\ 1c_1 + 0c_2 \end{pmatrix}.$$

For instance,

$$\text{with } \mathrm{Rep}_B(\begin{pmatrix} 4 \\ 8 \end{pmatrix}) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_B \text{ then } \mathrm{Rep}_D(\,h(\begin{pmatrix} 4 \\ 8 \end{pmatrix})\,) = \begin{pmatrix} 2 \\ -5/2 \\ 1 \end{pmatrix}.$$

We will express computations like the one above with a matrix notation.

$$\begin{pmatrix} 0 & 1 \\ -1/2 & -1 \\ 1 & 0 \end{pmatrix}_{B,D} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}_B = \begin{pmatrix} 0c_1 + 1c_2 \\ (-1/2)c_1 - 1c_2 \\ 1c_1 + 0c_2 \end{pmatrix}_D$$

In the middle is the argument $v$ to the map, represented with respect to the domain's basis $B$ by a column vector with components $c_1$ and $c_2$. On the right is the value $h(v)$ of the map on that argument, represented with respect to the codomain's basis $D$ by a column vector with components $0c_1 + 1c_2$, etc. The matrix on the left is the new thing. It consists of the coefficients from the vector on the right, 0 and 1 from the first row, $-1/2$ and $-1$ from the second row, and 1 and 0 from the third row.

This notation simply breaks the parts from the right, the coefficients and the $c$'s, out separately on the left, into a vector that represents the map's argument and a matrix that we will take to represent the map itself.

**1.2 Definition** Suppose that $V$ and $W$ are vector spaces of dimensions $n$ and $m$ with bases $B$ and $D$, and that $h\colon V \to W$ is a linear map. If

$$\mathrm{Rep}_D(h(\beta_1)) = \begin{pmatrix} h_{1,1} \\ h_{2,1} \\ \vdots \\ h_{m,1} \end{pmatrix}_D \quad \ldots \quad \mathrm{Rep}_D(h(\beta_n)) = \begin{pmatrix} h_{1,n} \\ h_{2,n} \\ \vdots \\ h_{m,n} \end{pmatrix}_D$$

then

$$\mathrm{Rep}_{B,D}(h) = \begin{pmatrix} h_{1,1} & h_{1,2} & \ldots & h_{1,n} \\ h_{2,1} & h_{2,2} & \ldots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \ldots & h_{m,n} \end{pmatrix}_{B,D}$$

is the *matrix representation of $h$ with respect to $B, D$.*

Briefly, the vectors representing the $h(\beta)$'s are adjoined to make the matrix representing the map.

$$\mathrm{Rep}_{B,D}(h) = \left(\begin{array}{c|c|c} \vdots & & \vdots \\ \mathrm{Rep}_D(\,h(\beta_1)\,) & \cdots & \mathrm{Rep}_D(\,h(\beta_n)\,) \\ \vdots & & \vdots \end{array}\right)$$

Observe that the number of columns $n$ of the matrix is the dimension of the domain of the map, and the number of rows $m$ is the dimension of the codomain.

**1.3 Example** If $h\colon \mathbb{R}^3 \to \mathcal{P}_1$ is given by

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \stackrel{h}{\longmapsto} (2a_1 + a_2) + (-a_3)x$$

then where

$$B = \langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \rangle \quad \text{and} \quad D = \langle 1+x, -1+x \rangle$$

the action of $h$ on $B$ is given by

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \stackrel{h}{\longmapsto} -x \qquad \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \stackrel{h}{\longmapsto} 2 \qquad \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \stackrel{h}{\longmapsto} 4$$

and a simple calculation gives

$$\text{Rep}_D(-x) = \begin{pmatrix} -1/2 \\ -1/2 \end{pmatrix}_D \quad \text{Rep}_D(2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}_D \quad \text{Rep}_D(4) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}_D$$

showing that this is the matrix representing $h$ with respect to the bases.

$$\text{Rep}_{B,D}(h) = \begin{pmatrix} -1/2 & 1 & 2 \\ -1/2 & -1 & -2 \end{pmatrix}_{B,D}$$

We will use lower case letters for a map, upper case for the matrix, and lower case again for the entries of the matrix. Thus for the map $h$, the matrix representing it is $H$, with entries $h_{i,j}$.

**1.4 Theorem** Assume that $V$ and $W$ are vector spaces of dimensions $m$ and $n$ with bases $B$ and $D$, and that $h\colon V \to W$ is a linear map. If $h$ is represented by

$$\text{Rep}_{B,D}(h) = \begin{pmatrix} h_{1,1} & h_{1,2} & \ldots & h_{1,n} \\ h_{2,1} & h_{2,2} & \ldots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \ldots & h_{m,n} \end{pmatrix}_{B,D}$$

and $v \in V$ is represented by

$$\text{Rep}_B(v) = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}_B$$

then the representation of the image of $v$ is this.

$$\text{Rep}_D(\,h(v)\,) = \begin{pmatrix} h_{1,1}c_1 + h_{1,2}c_2 + \cdots + h_{1,n}c_n \\ h_{2,1}c_1 + h_{2,2}c_2 + \cdots + h_{2,n}c_n \\ \vdots \\ h_{m,1}c_1 + h_{m,2}c_2 + \cdots + h_{m,n}c_n \end{pmatrix}_D$$

Proof. Exercise 28. QED

We will think of the matrix $\text{Rep}_{B,D}(h)$ and the vector $\text{Rep}_B(v)$ as combining to make the vector $\text{Rep}_D(h(v))$.

**1.5 Definition** The *matrix-vector product* of a $m\times n$ matrix and a $n\times 1$ vector is this.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ & \vdots & & \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} a_{1,1}c_1 + a_{1,2}c_2 + \cdots + a_{1,n}c_n \\ a_{2,1}c_1 + a_{2,2}c_2 + \cdots + a_{2,n}c_n \\ \vdots \\ a_{m,1}c_1 + a_{m,2}c_2 + \cdots + a_{m,n}c_n \end{pmatrix}$$

The point of Definition 1.2 is to generalize Example 1.1, that is, the point of the definition is Theorem 1.4, that the matrix describes how to get from the representation of a domain vector with respect to the domain's basis to the representation of its image in the codomain with respect to the codomain's basis. With Definition 1.5, we can restate this as: application of a linear map is represented by the matrix-vector product of the map's representative and the vector's representative.

**1.6 Example** With the matrix from Example 1.3 we can calculate where that map sends this vector.

$$v = \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}$$

This vector is represented, with respect to the domain basis $B$, by

$$\text{Rep}_B(v) = \begin{pmatrix} 0 \\ 1/2 \\ 2 \end{pmatrix}_B$$

and so this is the representation of the value $h(v)$ with respect to the codomain basis $D$.

$$\begin{aligned} \text{Rep}_D(h(v)) &= \begin{pmatrix} -1/2 & 1 & 2 \\ -1/2 & -1 & -2 \end{pmatrix}_{B,D} \begin{pmatrix} 0 \\ 1/2 \\ 2 \end{pmatrix}_B \\ &= \begin{pmatrix} (-1/2)\cdot 0 + 1\cdot(1/2) + 2\cdot 2 \\ (-1/2)\cdot 0 - 1\cdot(1/2) - 2\cdot 2 \end{pmatrix}_D = \begin{pmatrix} 9/2 \\ -9/2 \end{pmatrix}_D \end{aligned}$$

To find $h(v)$ itself, not its representation, take $(9/2)(1+x) - (9/2)(-1+x) = 9$.

**1.7 Example** Let $\pi\colon \mathbb{R}^3 \to \mathbb{R}^2$ be projection onto the $xy$-plane. To give a matrix representing this map, we first fix bases.

$$B = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \rangle \qquad D = \langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$$

For each vector in the domain's basis, we find its image under the map.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Then we find the representation of each image with respect to the codomain's basis

$$\text{Rep}_D(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{Rep}_D(\begin{pmatrix} 1 \\ 1 \end{pmatrix}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{Rep}_D(\begin{pmatrix} -1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

(these are easily checked). Finally, adjoining these representations gives the matrix representing $\pi$ with respect to $B, D$.

$$\text{Rep}_{B,D}(\pi) = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix}_{B,D}$$

We can illustrate Theorem 1.4 by computing the matrix-vector product representing the following statement about the projection map.

$$\pi(\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

Representing this vector from the domain with respect to the domain's basis

$$\text{Rep}_B(\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}) = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}_B$$

gives this matrix-vector product.

$$\operatorname{Rep}_D(\pi(\begin{pmatrix}2\\1\\1\end{pmatrix}))=\begin{pmatrix}1&0&-1\\-1&1&1\end{pmatrix}_{B,D}\begin{pmatrix}1\\2\\1\end{pmatrix}_B=\begin{pmatrix}0\\2\end{pmatrix}_D$$

Expanding this representation into a linear combination of vectors from $D$

$$0\cdot\begin{pmatrix}2\\1\end{pmatrix}+2\cdot\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}2\\2\end{pmatrix}$$

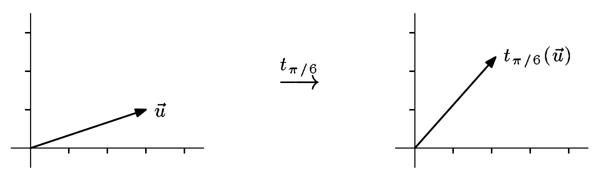
checks that the map's action is indeed reflected in the operation of the matrix. (We will sometimes compress these three displayed equations into one

$$\begin{pmatrix}2\\2\\1\end{pmatrix}=\begin{pmatrix}1\\2\\1\end{pmatrix}_B\overset{h}{\underset{H}{\longmapsto}}\begin{pmatrix}0\\2\end{pmatrix}_D=\begin{pmatrix}2\\2\end{pmatrix}$$

in the course of a calculation.)

We now have two ways to compute the effect of projection, the straightforward formula that drops each three-tall vector's third component to make a two-tall vector, and the above formula that uses representations and matrix-vector multiplication. Compared to the first way, the second way might seem complicated. However, it has advantages. The next example shows that giving a formula for some maps is simplified by this new scheme.

**1.8 Example** To represent a *rotation* map $t_\theta\colon\mathbb{R}^2\to\mathbb{R}^2$ that turns all vectors in the plane counterclockwise through an angle $\theta$




we start by fixing bases. Using $\mathcal{E}_2$ both as a domain basis and as a codomain basis is natural, Now, we find the image under the map of each vector in the domain's basis.

$$\begin{pmatrix}1\\0\end{pmatrix}\overset{t_\theta}{\longmapsto}\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}\qquad\begin{pmatrix}0\\1\end{pmatrix}\overset{t_\theta}{\longmapsto}\begin{pmatrix}-\sin\theta\\\cos\theta\end{pmatrix}$$

Then we represent these images with respect to the codomain's basis. Because this basis is $\mathcal{E}_2$, vectors are represented by themselves. Finally, adjoining the representations gives the matrix representing the map.

$$\operatorname{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(t_\theta)=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}$$

The advantage of this scheme is that just by knowing how to represent the image of the two basis vectors, we get a formula that tells us the image of any vector at all; here a vector rotated by $\theta=\pi/6$.

$$\begin{pmatrix}3\\-2\end{pmatrix}\overset{t_{\pi/6}}{\longmapsto}\begin{pmatrix}\sqrt{3}/2&-1/2\\1/2&\sqrt{3}/2\end{pmatrix}\begin{pmatrix}3\\-2\end{pmatrix}\approx\begin{pmatrix}3.598\\-0.232\end{pmatrix}$$

(Again, we are using the fact that, with respect to $\mathcal{E}_2$, vectors represent themselves.)

We have already seen the addition and scalar multiplication operations of matrices and the dot product operation of vectors. Matrix-vector multiplication is a new operation in the arithmetic of vectors and matrices. Nothing in Definition 1.5 requires us to view it in terms of representations. We can get some insight into this operation by turning away from what is being represented, and instead focusing on how the entries combine.

**1.9 Example** In the definition the width of the matrix equals the height of the vector. Hence, the first product below is defined while the second is not.

$$\begin{pmatrix} 1 & 0 & 0 \\ 4 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 4 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

One reason that this product is not defined is purely formal: the definition requires that the sizes match, and these sizes don't match. Behind the formality, though, is a reason why we will leave it undefined — the matrix represents a map with a three-dimensional domain while the vector represents a member of a two-dimensional space.

A good way to view a matrix-vector product is as the dot products of the rows of the matrix with the column vector.

$$\begin{pmatrix} & \vdots & & \\ a_{i,1} & a_{i,2} & \ldots & a_{i,n} \\ & \vdots & & \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \vdots \\ a_{i,1}c_1 + a_{i,2}c_2 + \ldots + a_{i,n}c_n \\ \vdots \end{pmatrix}$$

Looked at in this row-by-row way, this new operation generalizes dot product.

Matrix-vector product can also be viewed column-by-column.

$$\begin{aligned} \begin{pmatrix} h_{1,1} & h_{1,2} & \ldots & h_{1,n} \\ h_{2,1} & h_{2,2} & \ldots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \ldots & h_{m,n} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} &= \begin{pmatrix} h_{1,1}c_1 + h_{1,2}c_2 + \cdots + h_{1,n}c_n \\ h_{2,1}c_1 + h_{2,2}c_2 + \cdots + h_{2,n}c_n \\ \vdots \\ h_{m,1}c_1 + h_{m,2}c_2 + \cdots + h_{m,n}c_n \end{pmatrix} \\ &= c_1 \begin{pmatrix} h_{1,1} \\ h_{2,1} \\ \vdots \\ h_{m,1} \end{pmatrix} + \cdots + c_n \begin{pmatrix} h_{1,n} \\ h_{2,n} \\ \vdots \\ h_{m,n} \end{pmatrix} \end{aligned}$$

**1.10 Example**

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \end{pmatrix}$$

The result has the columns of the matrix weighted by the entries of the vector. This way of looking at it brings us back to the objective stated at the start of this section, to compute $h(c_1\beta_1 + \cdots + c_n\beta_n)$ as $c_1h(\beta_1) + \cdots + c_nh(\beta_n)$.

We began this section by noting that the equality of these two enables us to compute the action of $h$ on any argument knowing only $h(\beta_1)$, ..., $h(\beta_n)$. We have developed this into a scheme to compute the action of the map by taking the matrix-vector product of the matrix representing the map and the vector representing the argument. In this way, any linear map is represented with respect to some bases by a matrix. In the next subsection, we will show the converse, that any matrix represents a linear map.

### Exercises

✓ **1.11** Multiply the matrix

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

by each vector (or state "not defined").

**(a)** $\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ **(b)** $\begin{pmatrix} -2 \\ -2 \end{pmatrix}$ **(c)** $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

**1.12** Perform, if possible, each matrix-vector multiplication.

(a) $\begin{pmatrix} 2 & 1 \\ 3 & -1/2 \end{pmatrix}\begin{pmatrix} 4 \\ 2 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 & 0 \\ -2 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$

✓ **1.13** Solve this matrix equation.

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 3 \\ 1 & -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \\ 4 \end{pmatrix}$$

✓ **1.14** For a homomorphism from $\mathcal{P}_2$ to $\mathcal{P}_3$ that sends

$$1 \mapsto 1+x, \quad x \mapsto 1+2x, \quad \text{and} \quad x^2 \mapsto x - x^3$$

where does $1 - 3x + 2x^2$ go?

✓ **1.15** Assume that $h\colon \mathbb{R}^2 \to \mathbb{R}^3$ is determined by this action.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

Using the standard bases, find

**(a)** the matrix representing this map;

**(b)** a general formula for $h(\vec{v})$.

✓ **1.16** Let $d/dx\colon \mathcal{P}_3 \to \mathcal{P}_3$ be the derivative transformation.

**(a)** Represent $d/dx$ with respect to $B, B$ where $B = \langle 1, x, x^2, x^3 \rangle$.

**(b)** Represent $d/dx$ with respect to $B, D$ where $D = \langle 1, 2x, 3x^2, 4x^3 \rangle$.

✓ **1.17** Represent each linear map with respect to each pair of bases.

**(a)** $d/dx\colon \mathcal{P}_n \to \mathcal{P}_n$ with respect to $B, B$ where $B = \langle 1, x, \ldots, x^n \rangle$, given by

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \mapsto a_1 + 2a_2x + \cdots + na_nx^{n-1}$$

**(b)** $\int\colon \mathcal{P}_n \to \mathcal{P}_{n+1}$ with respect to $B_n, B_{n+1}$ where $B_i = \langle 1, x, \ldots, x^i \rangle$, given by

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \mapsto a_0x + \frac{a_1}{2}x^2 + \cdots + \frac{a_n}{n+1}x^{n+1}$$

**(c)** $\int_0^1\colon \mathcal{P}_n \to \mathbb{R}$ with respect to $B, \mathcal{E}_1$ where $B = \langle 1, x, \ldots, x^n \rangle$ and $\mathcal{E}_1 = \langle 1 \rangle$, given by

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \mapsto a_0 + \frac{a_1}{2} + \cdots + \frac{a_n}{n+1}$$

**(d)** $\text{eval}_3\colon \mathcal{P}_n \to \mathbb{R}$ with respect to $B, \mathcal{E}_1$ where $B = \langle 1, x, \ldots, x^n \rangle$ and $\mathcal{E}_1 = \langle 1 \rangle$, given by

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \mapsto a_0 + a_1 \cdot 3 + a_2 \cdot 3^2 + \cdots + a_n \cdot 3^n$$

**(e)** $\text{slide}_{-1}\colon \mathcal{P}_n \to \mathcal{P}_n$ with respect to $B, B$ where $B = \langle 1, x, \ldots, x^n \rangle$, given by

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \mapsto a_0 + a_1 \cdot (x+1) + \cdots + a_n \cdot (x+1)^n$$

**1.18** Represent the identity map on any nontrivial space with respect to $B, B$, where $B$ is any basis.

**1.19** Represent, with respect to the natural basis, the transpose transformation on the space $\mathcal{M}_{2\times2}$ of 2×2 matrices.

**1.20** Assume that $B = \langle \beta_1, \beta_2, \beta_3, \beta_4 \rangle$ is a basis for a vector space. Represent with respect to $B, B$ the transformation that is determined by each.

**(a)** $\beta_1 \mapsto \beta_2$, $\beta_2 \mapsto \beta_3$, $\beta_3 \mapsto \beta_4$, $\beta_4 \mapsto 0$

**(b)** $\beta_1 \mapsto \beta_2$, $\beta_2 \mapsto 0$, $\beta_3 \mapsto \beta_4$, $\beta_4 \mapsto 0$

**(c)** $\beta_1 \mapsto \beta_2$, $\beta_2 \mapsto \beta_3$, $\beta_3 \mapsto 0$, $\beta_4 \mapsto 0$

**1.21** Example 1.8 shows how to represent the rotation transformation of the plane with respect to the standard basis. Express these other transformations also with respect to the standard basis.

**(a)** the *dilation* map $d_s$, which multiplies all vectors by the same scalar $s$

**(b)** the *reflection* map $f_\ell$, which reflects all all vectors across a line $\ell$ through the origin

✓ **1.22** Consider a linear transformation of $\mathbb{R}^2$ determined by these two.

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

**(a)** Represent this transformation with respect to the standard bases.

**(b)** Where does the transformation send this vector?

$$\begin{pmatrix} 0 \\ 5 \end{pmatrix}$$

**(c)** Represent this transformation with respect to these bases.

$$B = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle \qquad D = \langle \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rangle$$

**(d)** Using $B$ from the prior item, represent the transformation with respect to $B, B$.

**1.23** Suppose that $h\colon V \to W$ is nonsingular so that by Theorem 2.21, for any basis $B = \langle \beta_1, \dots, \beta_n \rangle \subset V$ the image $h(B) = \langle h(\beta_1), \dots, h(\beta_n) \rangle$ is a basis for $W$.

**(a)** Represent the map $h$ with respect to $B, h(B)$.

**(b)** For a member $v$ of the domain, where the representation of $v$ has components $c_1$, …, $c_n$, represent the image vector $h(v)$ with respect to the image basis $h(B)$.

**1.24** Give a formula for the product of a matrix and $e_i$, the column vector that is all zeroes except for a single one in the $i$-th position.

✓ **1.25** For each vector space of functions of one real variable, represent the derivative transformation with respect to $B, B$.

**(a)** $\{a\cos x + b\sin x \mid a, b \in \mathbb{R}\}$, $B = \langle \cos x, \sin x \rangle$

**(b)** $\{ae^x + be^{2x} \mid a, b \in \mathbb{R}\}$, $B = \langle e^x, e^{2x} \rangle$

**(c)** $\{a + bx + ce^x + dxe^x \mid a, b, c, d \in \mathbb{R}\}$, $B = \langle 1, x, e^x, xe^x \rangle$

**1.26** Find the range of the linear transformation of $\mathbb{R}^2$ represented with respect to the standard bases by each matrix.

**(a)** $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ **(b)** $\begin{pmatrix} 0 & 0 \\ 3 & 2 \end{pmatrix}$ **(c)** a matrix of the form $\begin{pmatrix} a & b \\ 2a & 2b \end{pmatrix}$

✓ **1.27** Can one matrix represent two different linear maps? That is, can $\text{Rep}_{B,D}(h) = \text{Rep}_{\hat{B},\hat{D}}(\hat{h})$?

**1.28** Prove Theorem 1.4.

✓ **1.29** Example 1.8 shows how to represent rotation of all vectors in the plane through an angle $\theta$ about the origin, with respect to the standard bases.

**(a)** Rotation of all vectors in three-space through an angle $\theta$ about the $x$-axis is a transformation of $\mathbb{R}^3$. Represent it with respect to the standard bases. Arrange the rotation so that to someone whose feet are at the origin and whose head is at $(1, 0, 0)$, the movement appears clockwise.

**(b)** Repeat the prior item, only rotate about the $y$-axis instead. (Put the person's head at $e_2$.)

**(c)** Repeat, about the $z$-axis.

**(d)** Extend the prior item to $\mathbb{R}^4$. (*Hint:* 'rotate about the $z$-axis' can be restated as 'rotate parallel to the $xy$-plane'.)

**1.30** (Schur's Triangularization Lemma)

**(a)** Let $U$ be a subspace of $V$ and fix bases $B_U \subseteq B_V$. What is the relationship between the representation of a vector from $U$ with respect to $B_U$ and the representation of that vector (viewed as a member of $V$) with respect to $B_V$?

**(b)** What about maps?

**(c)** Fix a basis $B = \langle \beta_1, \dots, \beta_n \rangle$ for $V$ and observe that the spans

$$\langle \{\vec{0}\} \rangle = \{\vec{0}\} \subset \langle \{\beta_1\} \rangle \subset \langle \{\beta_1, \beta_2\} \rangle \subset \quad \cdots \quad \subset \langle B \rangle = V$$

form a strictly increasing chain of subspaces. Show that for any linear map $h\colon V \to W$ there is a chain $W_0 = \{\vec{0}\} \subseteq W_1 \subseteq \cdots \subseteq W_m = W$ of subspaces of $W$ such that

$$h(\langle \{\beta_1, \dots, \beta_i\} \rangle) \subset W_i$$

for each $i$.

**(d)** Conclude that for every linear map $h\colon V \to W$ there are bases $B, D$ so the matrix representing $h$ with respect to $B, D$ is upper-triangular (that is, each entry $h_{i,j}$ with $i > j$ is zero).

**(e)** Is an upper-triangular representation unique?

## II.2 Any Matrix Represents a Linear Map

The prior subsection shows that the action of a linear map $h$ is described by a matrix $H$, with respect to appropriate bases, in this way.

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_B \xmapsto[H]{h} \begin{pmatrix} h_{1,1}v_1 + \cdots + h_{1,n}v_n \\ \vdots \\ h_{m,1}v_1 + \cdots + h_{m,n}v_n \end{pmatrix}_D = h(v)$$

In this subsection, we will show the converse, that each matrix represents a linear map.

Recall that, in the definition of the matrix representation of a linear map, the number of columns of the matrix is the dimension of the map's domain and the number of rows of the matrix is the dimension of the map's codomain. Thus, for instance, a $2\times3$ matrix cannot represent a map from $\mathbb{R}^5$ to $\mathbb{R}^4$. The next result says that, beyond this restriction on the dimensions, there are no other limitations: the $2\times3$ matrix represents a map from any three-dimensional space to any two-dimensional space.

**2.1 Theorem** Any matrix represents a homomorphism between vector spaces of appropriate dimensions, with respect to any pair of bases.

Proof. For the matrix

$$H = \begin{pmatrix} h_{1,1} & h_{1,2} & \ldots & h_{1,n} \\ h_{2,1} & h_{2,2} & \ldots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \ldots & h_{m,n} \end{pmatrix}$$

fix any $n$-dimensional domain space $V$ and any $m$-dimensional codomain space $W$. Also fix bases $B = \langle \beta_1, \ldots, \beta_n \rangle$ and $D = \langle \delta_1, \ldots, \delta_m \rangle$ for those spaces. Define a function $h \colon V \to W$ by: where $v$ in the domain is represented as

$$\text{Rep}_B(v) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_B$$

then its image $h(v)$ is the member the codomain represented by

$$\text{Rep}_D(\,h(v)\,) = \begin{pmatrix} h_{1,1}v_1 + \cdots + h_{1,n}v_n \\ \vdots \\ h_{m,1}v_1 + \cdots + h_{m,n}v_n \end{pmatrix}_D$$

that is, $h(v) = h(v_1\beta_1 + \cdots + v_n\beta_n)$ is defined to be $(h_{1,1}v_1 + \cdots + h_{1,n}v_n)\cdot\delta_1 + \cdots + (h_{m,1}v_1 + \cdots + h_{m,n}v_n)\cdot\delta_m$. (This is well-defined by the uniqueness of the representation $\text{Rep}_B(v)$.)

Observe that $h$ has simply been defined to make it the map that is represented with respect to $B, D$ by the matrix $H$. So to finish, we need only check that $h$ is linear. If $v, u \in V$ are such that

$$\text{Rep}_B(v) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \text{and} \quad \text{Rep}_B(u) = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

and $c, d \in \mathbb{R}$ then the calculation

$$\begin{aligned} h(cv + du) &= \bigl(h_{1,1}(cv_1 + du_1) + \cdots + h_{1,n}(cv_n + du_n)\bigr)\cdot\delta_1 + \\ &\qquad \cdots + \bigl(h_{m,1}(cv_1 + du_1) + \cdots + h_{m,n}(cv_n + du_n)\bigr)\cdot\delta_m \\ &= c\cdot h(v) + d\cdot h(u) \end{aligned}$$

provides this verification. QED

**2.2 Example** Which map the matrix represents depends on which bases are used. If

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_1 = D_1 = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle, \quad \text{and} \quad B_2 = D_2 = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle,$$

then $h_1\colon \mathbb{R}^2 \to \mathbb{R}^2$ represented by $H$ with respect to $B_1, D_1$ maps

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}_{B_1} \quad \mapsto \quad \begin{pmatrix} c_1 \\ 0 \end{pmatrix}_{D_1} = \begin{pmatrix} c_1 \\ 0 \end{pmatrix}$$

while $h_2\colon \mathbb{R}^2 \to \mathbb{R}^2$ represented by $H$ with respect to $B_2, D_2$ is this map.

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} c_2 \\ c_1 \end{pmatrix}_{B_2} \quad \mapsto \quad \begin{pmatrix} c_2 \\ 0 \end{pmatrix}_{D_2} = \begin{pmatrix} 0 \\ c_2 \end{pmatrix}$$

These two are different. The first is projection onto the $x$ axis, while the second is projection onto the $y$ axis.

So not only is any linear map described by a matrix but any matrix describes a linear map. This means that we can, when convenient, handle linear maps entirely as matrices, simply doing the computations, without have to worry that a matrix of interest does not represent a linear map on some pair of spaces of interest. (In practice, when we are working with a matrix but no spaces or bases have been specified, we will often take the domain and codomain to be $\mathbb{R}^n$ and $\mathbb{R}^m$ and use the standard bases. In this case, because the representation is transparent — the representation with respect to the standard basis of $v$ is $v$ — the column space of the matrix equals the range of the map. Consequently, the column space of $H$ is often denoted by $\mathscr{R}(H)$.)

With the theorem, we have characterized linear maps as those maps that act in this matrix way. Each linear map is described by a matrix and each matrix describes a linear map. We finish this section by illustrating how a matrix can be used to tell things about its maps.

**2.3 Theorem** The rank of a matrix equals the rank of any map that it represents.

Proof. Suppose that the matrix $H$ is $m \times n$. Fix domain and codomain spaces $V$ and $W$ of dimension $n$ and $m$, with bases $B = \langle \beta_1, \dots, \beta_n \rangle$ and $D$. Then $H$ represents some linear map $h$ between those spaces with respect to these bases whose rangespace

$$\begin{aligned} \{h(v) \mid v \in V\} &= \{h(c_1\beta_1 + \cdots + c_n\beta_n) \mid c_1, \dots, c_n \in \mathbb{R}\} \\ &= \{c_1 h(\beta_1) + \cdots + c_n h(\beta_n) \mid c_1, \dots, c_n \in \mathbb{R}\} \end{aligned}$$

is the span $\langle \{h(\beta_1), \dots, h(\beta_n)\} \rangle$. The rank of $h$ is the dimension of this rangespace.

The rank of the matrix is its column rank (or its row rank; the two are equal). This is the dimension of the column space of the matrix, which is the span of the set of column vectors $\langle \{\mathrm{Rep}_D(h(\beta_1)), \dots, \mathrm{Rep}_D(h(\beta_n))\} \rangle$.

To see that the two spans have the same dimension, recall that a representation with respect to a basis gives an isomorphism $\mathrm{Rep}_D\colon W \to \mathbb{R}^m$. Under this isomorphism, there is a linear relationship among members of the rangespace if and only if the same relationship holds in the column space, e.g, $0 = c_1 h(\beta_1) + \cdots + c_n h(\beta_n)$ if and only if $0 = c_1 \mathrm{Rep}_D(h(\beta_1)) + \cdots + c_n \mathrm{Rep}_D(h(\beta_n))$. Hence, a subset of the rangespace is linearly independent if and only if the corresponding subset of the column space is linearly independent. This means that the size of the largest linearly independent subset of the rangespace equals the size of the largest linearly independent subset of the column space, and so the two spaces have the same dimension. QED

**2.4 Example** Any map represented by

$$\begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

must, by definition, be from a three-dimensional domain to a four-dimensional codomain. In addition, because the rank of this matrix is two (we can spot this by eye or get it with Gauss' method), any map represented by this matrix has a two-dimensional rangespace.

**2.5 Corollary** Let $h$ be a linear map represented by a matrix $H$. Then $h$ is onto if and only if the rank of $H$ equals the number of its rows, and $h$ is one-to-one if and only if the rank of $H$ equals the number of its columns.

Proof. For the first half, the dimension of the rangespace of $h$ is the rank of $h$, which equals the rank of $H$ by the theorem. Since the dimension of the codomain of $h$ is the number of rows in $H$, if the rank of $H$ equals the number of rows, then the dimension of the rangespace equals the dimension of the codomain. But a subspace with the same dimension as its superspace must equal that superspace (a basis for the rangespace is a linearly independent subset of the codomain, whose size is equal to the dimension of the codomain, and so this set is a basis for the codomain).

For the second half, a linear map is one-to-one if and only if it is an isomorphism between its domain and its range, that is, if and only if its domain has the same dimension as its range. But the number of columns in $h$ is the dimension of $h$'s domain, and by the theorem the rank of $H$ equals the dimension of $h$'s range. QED

The above results end any confusion caused by our use of the word 'rank' to mean apparently different things when applied to matrices and when applied to maps. We can also justify the dual use of 'nonsingular'. We've defined a matrix to be nonsingular if it is square and is the matrix of coefficients of a linear system with a unique solution, and we've defined a linear map to be nonsingular if it is one-to-one.

**2.6 Corollary** A square matrix represents nonsingular maps if and only if it is a nonsingular matrix. Thus, a matrix represents an isomorphism if and only if it is square and nonsingular.

Proof. Immediate from the prior result. QED

**2.7 Example** Any map from $\mathbb{R}^2$ to $\mathcal{P}_1$ represented with respect to any pair of bases by

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$$

is nonsingular because this matrix has rank two.

**2.8 Example** Any map $g\colon V \to W$ represented by

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$$

is not nonsingular because this matrix is not nonsingular.

We've now seen that the relationship between maps and matrices goes both ways: fixing bases, any linear map is represented by a matrix and any matrix describes a linear map. That is, by fixing spaces and bases we get a correspondence between maps and matrices. In the rest of this chapter we will explore this correspondence. For instance, we've defined for linear maps the operations of addition and scalar multiplication and we shall see what the corresponding matrix operations are. We shall also see the matrix operation that represent the map operation of composition. And, we shall see how to find the matrix that represents a map's inverse.

### Exercises

✓ **2.9** Decide if the vector is in the column space of the matrix.

(a) $\begin{pmatrix} 2 & 1 \\ 2 & 5 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \end{pmatrix}$ (b) $\begin{pmatrix} 4 & -8 \\ 2 & -4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$

✓ **2.10** Decide if each vector lies in the range of the map from $\mathbb{R}^3$ to $\mathbb{R}^2$ represented with respect to the standard bases by the matrix.

**(a)** $\begin{pmatrix}1&1&3\\0&1&4\end{pmatrix}, \begin{pmatrix}1\\3\end{pmatrix}$ **(b)** $\begin{pmatrix}2&0&3\\4&0&6\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}$

✓ **2.11** Consider this matrix, representing a transformation of $\mathbb{R}^2$, and these bases for that space.

$$\frac{1}{2}\cdot\begin{pmatrix}1&1\\-1&1\end{pmatrix} \qquad B=\langle\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\rangle \quad D=\langle\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\rangle$$

**(a)** To what vector in the codomain is the first member of $B$ mapped?
**(b)** The second member?
**(c)** Where is a general vector from the domain (a vector with components $x$ and $y$) mapped? That is, what transformation of $\mathbb{R}^2$ is represented with respect to $B, D$ by this matrix?

**2.12** What transformation of $F=\{a\cos\theta+b\sin\theta \mid a,b\in\mathbb{R}\}$ is represented with respect to $B=\langle\cos\theta-\sin\theta,\sin\theta\rangle$ and $D=\langle\cos\theta+\sin\theta,\cos\theta\rangle$ by this matrix?

$$\begin{pmatrix}0&0\\1&0\end{pmatrix}$$

✓ **2.13** Decide if $1+2x$ is in the range of the map from $\mathbb{R}^3$ to $\mathcal{P}_2$ represented with respect to $\mathcal{E}_3$ and $\langle 1,1+x^2,x\rangle$ by this matrix.

$$\begin{pmatrix}1&3&0\\0&1&0\\1&0&1\end{pmatrix}$$

**2.14** Example 2.8 gives a matrix that is nonsingular, and is therefore associated with maps that are nonsingular.
**(a)** Find the set of column vectors representing the members of the nullspace of any map represented by this matrix.
**(b)** Find the nullity of any such map.
**(c)** Find the set of column vectors representing the members of the rangespace of any map represented by this matrix.
**(d)** Find the rank of any such map.
**(e)** Check that rank plus nullity equals the dimension of the domain.

✓ **2.15** Because the rank of a matrix equals the rank of any map it represents, if one matrix represents two different maps $H=\operatorname{Rep}_{B,D}(h)=\operatorname{Rep}_{\hat{B},\hat{D}}(\hat{h})$ (where $h,\hat{h}\colon V\to W$) then the dimension of the rangespace of $h$ equals the dimension of the rangespace of $\hat{h}$. Must these equal-dimensioned rangespaces actually be the same?

✓ **2.16** Let $V$ be an $n$-dimensional space with bases $B$ and $D$. Consider a map that sends, for $v\in V$, the column vector representing $v$ with respect to $B$ to the column vector representing $v$ with respect to $D$. Show that is a linear transformation of $\mathbb{R}^n$.

**2.17** Example 2.2 shows that changing the pair of bases can change the map that a matrix represents, even though the domain and codomain remain the same. Could the map ever not change? Is there a matrix $H$, vector spaces $V$ and $W$, and associated pairs of bases $B_1, D_1$ and $B_2, D_2$ (with $B_1\neq B_2$ or $D_1\neq D_2$ or both) such that the map represented by $H$ with respect to $B_1, D_1$ equals the map represented by $H$ with respect to $B_2, D_2$?

✓ **2.18** A square matrix is a *diagonal* matrix if it is all zeroes except possibly for the entries on its upper-left to lower-right diagonal — its $1,1$ entry, its $2,2$ entry, etc. Show that a linear map is an isomorphism if there are bases such that, with respect to those bases, the map is represented by a diagonal matrix with no zeroes on the diagonal.

**2.19** Describe geometrically the action on $\mathbb{R}^2$ of the map represented with respect to the standard bases $\mathcal{E}_2, \mathcal{E}_2$ by this matrix.

$$\begin{pmatrix}3&0\\0&2\end{pmatrix}$$

Do the same for these.

$$\begin{pmatrix}1&0\\0&0\end{pmatrix} \quad \begin{pmatrix}0&1\\1&0\end{pmatrix} \quad \begin{pmatrix}1&3\\0&1\end{pmatrix}$$

**2.20** The fact that for any linear map the rank plus the nullity equals the dimension of the domain shows that a necessary condition for the existence of a homomorphism between two spaces, onto the second space, is that there be no gain in dimension. That is, where $h\colon V\to W$ is onto, the dimension of $W$ must be less than or equal to the dimension of $V$.
**(a)** Show that this (strong) converse holds: no gain in dimension implies that there is a homomorphism and, further, any matrix with the correct size and correct rank represents such a map.

**(b)** Are there bases for $\mathbb{R}^3$ such that this matrix

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

represents a map from $\mathbb{R}^3$ to $\mathbb{R}^3$ whose range is the $xy$ plane subspace of $\mathbb{R}^3$?

**2.21** Let $V$ be an $n$-dimensional space and suppose that $x \in \mathbb{R}^n$. Fix a basis $B$ for $V$ and consider the map $h_x \colon V \to \mathbb{R}$ given $v \mapsto x \cdot \mathrm{Rep}_B(v)$ by the dot product.

**(a)** Show that this map is linear.

**(b)** Show that for any linear map $g \colon V \to \mathbb{R}$ there is an $x \in \mathbb{R}^n$ such that $g = h_x$.

**(c)** In the prior item we fixed the basis and varied the $x$ to get all possible linear maps. Can we get all possible linear maps by fixing an $x$ and varying the basis?

**2.22** Let $V, W, X$ be vector spaces with bases $B, C, D$.

**(a)** Suppose that $h \colon V \to W$ is represented with respect to $B, C$ by the matrix $H$. Give the matrix representing the scalar multiple $rh$ (where $r \in \mathbb{R}$) with respect to $B, C$ by expressing it in terms of $H$.

**(b)** Suppose that $h, g \colon V \to W$ are represented with respect to $B, C$ by $H$ and $G$. Give the matrix representing $h + g$ with respect to $B, C$ by expressing it in terms of $H$ and $G$.

**(c)** Suppose that $h \colon V \to W$ is represented with respect to $B, C$ by $H$ and $g \colon W \to X$ is represented with respect to $C, D$ by $G$. Give the matrix representing $g \circ h$ with respect to $B, D$ by expressing it in terms of $H$ and $G$.

# III Change of Basis

Representations, whether of vectors or of maps, vary with the bases. For instance, with respect to the two bases $\mathcal{E}_2$ and

$$B = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle$$

for $\mathbb{R}^2$, the vector $e_1$ has two different representations.

$$\text{Rep}_{\mathcal{E}_2}(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{Rep}_B(e_1) = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

Similarly, with respect to $\mathcal{E}_2, \mathcal{E}_2$ and $\mathcal{E}_2, B$, the identity map has two different representations.

$$\text{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(\text{id}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{Rep}_{\mathcal{E}_2,B}(\text{id}) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$$

With our point of view that the objects of our studies are vectors and maps, in fixing bases we are adopting a scheme of tags or names for these objects, that are convienent for computation. We will now see how to translate among these names — we will see exactly how representations vary as the bases vary.

## III.1 Changing Representations of Vectors

In converting $\text{Rep}_B(v)$ to $\text{Rep}_D(v)$ the underlying vector $v$ doesn't change. Thus, this translation is accomplished by the identity map on the space, described so that the domain space vectors are represented with respect to $B$ and the codomain space vectors are represented with respect to $D$.

$$\begin{array}{c} V_{\text{w.r.t. } B} \\ \text{id}\big\downarrow \\ V_{\text{w.r.t. } D} \end{array}$$

(The diagram is vertical to fit with the ones in the next subsection.)

**1.1 Definition** The *change of basis matrix* for bases $B, D \subset V$ is the representation of the identity map $\text{id}\colon V \to V$ with respect to those bases.

$$\text{Rep}_{B,D}(\text{id}) = \left(\begin{array}{c|c|c} \vdots & & \vdots \\ \text{Rep}_D(\beta_1) & \cdots & \text{Rep}_D(\beta_n) \\ \vdots & & \vdots \end{array}\right)$$

**1.2 Lemma** Left-multiplication by the change of basis matrix for $B, D$ converts a representation with respect to $B$ to one with respect to $D$. Conversly, if left-multiplication by a matrix changes bases $M \cdot \text{Rep}_B(v) = \text{Rep}_D(v)$ then $M$ is a change of basis matrix.

Proof. For the first sentence, for each $v$, as matrix-vector multiplication represents a map application, $\text{Rep}_{B,D}(\text{id}) \cdot \text{Rep}_B(v) = \text{Rep}_D(\text{id}(v)) = \text{Rep}_D(v)$. For the second sentence, with respect to $B, D$ the matrix $M$ represents some linear map, whose action is $v \mapsto v$, and is therefore the identity map. QED

**1.3 Example** With these bases for $\mathbb{R}^2$,

$$B = \langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle \qquad D = \langle \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$$

because

$$\mathrm{Rep}_D(\,\mathrm{id}(\begin{pmatrix}2\\1\end{pmatrix})) = \begin{pmatrix}-1/2\\3/2\end{pmatrix}_D \qquad \mathrm{Rep}_D(\,\mathrm{id}(\begin{pmatrix}1\\0\end{pmatrix})) = \begin{pmatrix}-1/2\\1/2\end{pmatrix}_D$$

the change of basis matrix is this.

$$\mathrm{Rep}_{B,D}(\mathrm{id}) = \begin{pmatrix}-1/2 & -1/2\\3/2 & 1/2\end{pmatrix}$$

We can see this matrix at work by finding the two representations of $e_2$

$$\mathrm{Rep}_B(\begin{pmatrix}0\\1\end{pmatrix}) = \begin{pmatrix}1\\-2\end{pmatrix} \qquad \mathrm{Rep}_D(\begin{pmatrix}0\\1\end{pmatrix}) = \begin{pmatrix}1/2\\1/2\end{pmatrix}$$

and checking that the conversion goes as expected.

$$\begin{pmatrix}-1/2 & -1/2\\3/2 & 1/2\end{pmatrix}\begin{pmatrix}1\\-2\end{pmatrix} = \begin{pmatrix}1/2\\1/2\end{pmatrix}$$

We finish this subsection by recognizing that the change of basis matrices are familiar.

**1.4 Lemma** A matrix changes bases if and only if it is nonsingular.

Proof. For one direction, if left-multiplication by a matrix changes bases then the matrix represents an invertible function, simply because the function is inverted by changing the bases back. Such a matrix is itself invertible, and so nonsingular.

To finish, we will show that any nonsingular matrix $M$ performs a change of basis operation from any given starting basis $B$ to some ending basis. Because the matrix is nonsingular, it will Gauss-Jordan reduce to the identity, so there are elementatry reduction matrices such that $R_r \cdots R_1 \cdot M = I$. Elementary matrices are invertible and their inverses are also elementary, so multiplying from the left first by ${R_r}^{-1}$, then by ${R_{r-1}}^{-1}$, etc., gives $M$ as a product of elementary matrices $M = {R_1}^{-1} \cdots {R_r}^{-1}$. Thus, we will be done if we show that elementary matrices change a given basis to another basis, for then ${R_r}^{-1}$ changes $B$ to some other basis $B_r$, and ${R_{r-1}}^{-1}$ changes $B_r$ to some $B_{r-1}$, ..., and the net effect is that $M$ changes $B$ to $B_1$. We will prove this about elementary matrices by covering the three types as separate cases.

Applying a row-multiplication matrix

$$M_i(k)\begin{pmatrix}c_1\\\vdots\\c_i\\\vdots\\c_n\end{pmatrix} = \begin{pmatrix}c_1\\\vdots\\kc_i\\\vdots\\c_n\end{pmatrix}$$

changes a representation with respect to $\langle\beta_1,\dots,\beta_i,\dots,\beta_n\rangle$ to one with respect to $\langle\beta_1,\dots,(1/k)\beta_i,\dots,\beta_n\rangle$ in this way.

$$v = c_1\cdot\beta_1 + \cdots + c_i\cdot\beta_i + \cdots + c_n\cdot\beta_n$$
$$\mapsto\ c_1\cdot\beta_1 + \cdots + kc_i\cdot(1/k)\beta_i + \cdots + c_n\cdot\beta_n = v$$

Similarly, left-multiplication by a row-swap matrix $P_{i,j}$ changes a representation with respect to the basis $\langle\beta_1,\dots,\beta_i,\dots,\beta_j,\dots,\beta_n\rangle$ into one with respect to the basis $\langle\beta_1,\dots,\beta_j,\dots,\beta_i,\dots,\beta_n\rangle$ in this way.

$$v = c_1\cdot\beta_1 + \cdots + c_i\cdot\beta_i + \cdots + c_j\beta_j + \cdots + c_n\cdot\beta_n$$
$$\mapsto\ c_1\cdot\beta_1 + \cdots + c_j\cdot\beta_j + \cdots + c_i\cdot\beta_i + \cdots + c_n\cdot\beta_n = v$$

And, a representation with respect to $\langle\beta_1,\dots,\beta_i,\dots,\beta_j,\dots,\beta_n\rangle$ changes via left-multiplication by a row-combination matrix $C_{i,j}(k)$ into a representation with respect to $\langle\beta_1,\dots,\beta_i - k\beta_j,\dots,\beta_j,\dots,\beta_n\rangle$

$$v = c_1\cdot\beta_1 + \cdots + c_i\cdot\beta_i + c_j\beta_j + \cdots + c_n\cdot\beta_n$$
$$\mapsto\ c_1\cdot\beta_1 + \cdots + c_i\cdot(\beta_i - k\beta_j) + \cdots + (kc_i + c_j)\cdot\beta_j + \cdots + c_n\cdot\beta_n = v$$

(the definition of reduction matrices specifies that $i \neq k$ and $k \neq 0$ and so this last one is a basis). QED

**1.5 Corollary** A matrix is nonsingular if and only if it represents the identity map with respect to some pair of bases.

In the next subsection we will see how to translate among representations of maps, that is, how to change $\mathrm{Rep}_{B,D}(h)$ to $\mathrm{Rep}_{\hat{B},\hat{D}}(h)$. The above corollary is a special case of this, where the domain and range are the same space, and where the map is the identity map.

## Exercises

✓ **1.6** In $\mathbb{R}^2$, where

$$D = \langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 4 \end{pmatrix} \rangle$$

find the change of basis matrices from $D$ to $\mathcal{E}_2$ and from $\mathcal{E}_2$ to $D$. Multiply the two.

✓ **1.7** Find the change of basis matrix for $B, D \subseteq \mathbb{R}^2$.

**(a)** $B = \mathcal{E}_2$, $D = \langle \vec{e}_2, \vec{e}_1 \rangle$ **(b)** $B = \mathcal{E}_2$, $D = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix} \rangle$ **(c)** $B = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix} \rangle$, $D = \mathcal{E}_2$

**(d)** $B = \langle \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \rangle$, $D = \langle \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \rangle$

**1.8** For the bases in Exercise 7, find the change of basis matrix in the other direction, from $D$ to $B$.

✓ **1.9** Find the change of basis matrix for each $B, D \subseteq \mathcal{P}_2$.

**(a)** $B = \langle 1, x, x^2 \rangle, D = \langle x^2, 1, x \rangle$ **(b)** $B = \langle 1, x, x^2 \rangle, D = \langle 1, 1+x, 1+x+x^2 \rangle$ **(c)** $B = \langle 2, 2x, x^2 \rangle, D = \langle 1+x^2, 1-x^2, x+x^2 \rangle$

✓ **1.10** Decide if each changes bases on $\mathbb{R}^2$. To what basis is $\mathcal{E}_2$ changed?

**(a)** $\begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix}$ **(b)** $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$ **(c)** $\begin{pmatrix} -1 & 4 \\ 2 & -8 \end{pmatrix}$ **(d)** $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$

**1.11** Find bases such that this matrix represents the identity map with respect to those bases.

$$\begin{pmatrix} 3 & 1 & 4 \\ 2 & -1 & 1 \\ 0 & 0 & 4 \end{pmatrix}$$

**1.12** Conside the vector space of real-valued functions with basis $\langle \sin(x), \cos(x) \rangle$. Show that $\langle 2\sin(x)+\cos(x), 3\cos(x) \rangle$ is also a basis for this space. Find the change of basis matrix in each direction.

**1.13** Where does this matrix

$$\begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$$

send the standard basis for $\mathbb{R}^2$? Any other bases? *Hint.* Consider the inverse.

✓ **1.14** What is the change of basis matrix with respect to $B, B$?

**1.15** Prove that a matrix changes bases if and only if it is invertible.

**1.16** Finish the proof of Lemma 1.4.

✓ **1.17** Let $H$ be a $n \times n$ nonsingular matrix. What basis of $\mathbb{R}^n$ does $H$ change to the standard basis?

✓ **1.18** **(a)** In $\mathcal{P}_3$ with basis $B = \langle 1+x, 1-x, x^2+x^3, x^2-x^3 \rangle$ we have this represenatation.

$$\mathrm{Rep}_B(1 - x + 3x^2 - x^3) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}_B$$

Find a basis $D$ giving this different representation for the same polynomial.

$$\mathrm{Rep}_D(1 - x + 3x^2 - x^3) = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}_D$$

**(b)** State and prove that any nonzero vector representation can be changed to any other.

*Hint.* The proof of Lemma 1.4 is constructive — it not only says the bases change, it shows how they change.

**1.19** Let $V, W$ be vector spaces, and let $B, \hat{B}$ be bases for $V$ and $D, \hat{D}$ be bases for $W$. Where $h\colon V \to W$ is linear, find a formula relating $\mathrm{Rep}_{B,D}(h)$ to $\mathrm{Rep}_{\hat{B},\hat{D}}(h)$.

✓ **1.20** Show that the columns of an $n \times n$ change of basis matrix form a basis for $\mathbb{R}^n$. Do all bases appear in that way: can the vectors from any $\mathbb{R}^n$ basis make the columns of a change of basis matrix?

✓ **1.21** Find a matrix having this effect.

$$\begin{pmatrix}1\\3\end{pmatrix}\mapsto\begin{pmatrix}4\\-1\end{pmatrix}$$

That is, find a $M$ that left-multiplies the starting vector to yield the ending vector. Is there a matrix having these two effects?

**(a)** $\begin{pmatrix}1\\3\end{pmatrix}\mapsto\begin{pmatrix}1\\1\end{pmatrix}$ $\begin{pmatrix}2\\-1\end{pmatrix}\mapsto\begin{pmatrix}-1\\-1\end{pmatrix}$ **(b)** $\begin{pmatrix}1\\3\end{pmatrix}\mapsto\begin{pmatrix}1\\1\end{pmatrix}$ $\begin{pmatrix}2\\6\end{pmatrix}\mapsto\begin{pmatrix}-1\\-1\end{pmatrix}$

Give a necessary and sufficient condition for there to be a matrix such that $\vec{v}_1\mapsto\vec{w}_1$ and $\vec{v}_2\mapsto\vec{w}_2$.

## III.2 Changing Map Representations

The first subsection shows how to convert the representation of a vector with respect to one basis to the representation of that same vector with respect to another basis. Here we will see how to convert the representation of a map with respect to one pair of bases to the representation of that map with respect to a different pair. That is, we want the relationship between the matrices in this arrow diagram.

$$\begin{array}{ccc} V_{\text{w.r.t. } B} & \xrightarrow[H]{h} & W_{\text{w.r.t. } D} \\ \text{id}\big\downarrow & & \text{id}\big\downarrow \\ V_{\text{w.r.t. } \hat{B}} & \xrightarrow[\hat{H}]{h} & W_{\text{w.r.t. } \hat{D}} \end{array}$$

To move from the lower-left of this diagram to the lower-right we can either go straight over, or else up to $V_B$ then over to $W_D$ and then down. Restated in terms of the matrices, we can calculate $\hat{H}=\text{Rep}_{\hat{B},\hat{D}}(h)$ either by simply using $\hat{B}$ and $\hat{D}$, or else by first changing bases with $\text{Rep}_{\hat{B},B}(\text{id})$ then multiplying by $H=\text{Rep}_{B,D}(h)$ and then changing bases with $\text{Rep}_{D,\hat{D}}(\text{id})$. This equation summarizes.

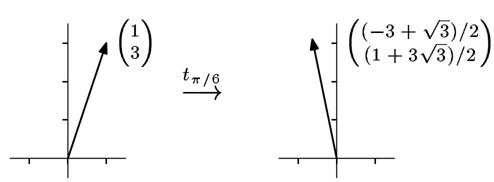
$$\hat{H}=\text{Rep}_{D,\hat{D}}(\text{id})\cdot H\cdot\text{Rep}_{\hat{B},B}(\text{id}) \qquad (*)$$

(To compare this equation with the sentence before it, remember that the equation is read from right to left because function composition is read right to left and matrix multiplication represent the composition.)

**2.1 Example** The matrix

$$T=\begin{pmatrix}\cos(\pi/6) & -\sin(\pi/6)\\ \sin(\pi/6) & \cos(\pi/6)\end{pmatrix}=\begin{pmatrix}\sqrt{3}/2 & -1/2\\ 1/2 & \sqrt{3}/2\end{pmatrix}$$

represents, with respect to $\mathcal{E}_2,\mathcal{E}_2$, the transformation $t\colon\mathbb{R}^2\to\mathbb{R}^2$ that rotates vectors $\pi/6$ radians counterclockwise.




We can translate that representation with respect to $\mathcal{E}_2,\mathcal{E}_2$ to one with respect to

$$\hat{B}=\langle\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}0\\2\end{pmatrix}\rangle \qquad \hat{D}=\langle\begin{pmatrix}-1\\0\end{pmatrix},\begin{pmatrix}2\\3\end{pmatrix}\rangle$$

by using the arrow diagram and formula (∗) above.

$$\begin{array}{ccc} \mathbb{R}^2_{\text{w.r.t. }\mathcal{E}_2} & \xrightarrow[T]{t} & \mathbb{R}^2_{\text{w.r.t. }\mathcal{E}_2} \\ \text{id}\big\downarrow & & \text{id}\big\downarrow \\ \mathbb{R}^2_{\text{w.r.t. }\hat{B}} & \xrightarrow[\hat{T}]{t} & \mathbb{R}^2_{\text{w.r.t. }\hat{D}} \end{array} \qquad \hat{T} = \text{Rep}_{\mathcal{E}_2,\hat{D}}(\text{id}) \cdot T \cdot \text{Rep}_{\hat{B},\mathcal{E}_2}(\text{id})$$

Note that $\text{Rep}_{\mathcal{E}_2,\hat{D}}(\text{id})$ can be calculated as the matrix inverse of $\text{Rep}_{\hat{D},\mathcal{E}_2}(\text{id})$.

$$\begin{aligned} \text{Rep}_{\hat{B},\hat{D}}(t) &= \begin{pmatrix} -1 & 2 \\ 0 & 3 \end{pmatrix}^{-1} \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} (5-\sqrt{3})/6 & (3+2\sqrt{3})/3 \\ (1+\sqrt{3})/6 & \sqrt{3}/3 \end{pmatrix} \end{aligned}$$

Although the new matrix is messier-appearing, the map that it represents is the same. For instance, to replicate the effect of $t$ in the picture, start with $\hat{B}$,

$$\text{Rep}_{\hat{B}}(\begin{pmatrix} 1 \\ 3 \end{pmatrix}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\hat{B}}$$

apply $\hat{T}$,

$$\begin{pmatrix} (5-\sqrt{3})/6 & (3+2\sqrt{3})/3 \\ (1+\sqrt{3})/6 & \sqrt{3}/3 \end{pmatrix}_{\hat{B},\hat{D}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\hat{B}} = \begin{pmatrix} (11+3\sqrt{3})/6 \\ (1+3\sqrt{3})/6 \end{pmatrix}_{\hat{D}}$$

and check it against $\hat{D}$

$$\frac{11+3\sqrt{3}}{6} \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \frac{1+3\sqrt{3}}{6} \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} (-3+\sqrt{3})/2 \\ (1+3\sqrt{3})/2 \end{pmatrix}$$

to see that it is the same result as above.

**2.2 Example** On $\mathbb{R}^3$ the map

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{t}{\longmapsto} \begin{pmatrix} y+z \\ x+z \\ x+y \end{pmatrix}$$

that is represented with respect to the standard basis in this way

$$\text{Rep}_{\mathcal{E}_3,\mathcal{E}_3}(t) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

can also be represented with respect to another basis

$$\text{if } B = \langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle \qquad \text{then } \text{Rep}_{B,B}(t) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

in a way that is simpler, in that the action of a diagonal matrix is easy to understand.

Naturally, we usually prefer basis changes that make the representation easier to understand. When the representation with respect to equal starting and ending bases is a diagonal matrix we say the map or matrix has been *diagonalized*. In Chaper Five we shall see which maps and matrices are diagonalizable, and where one is not, we shall see how to get a representation that is nearly diagonal.

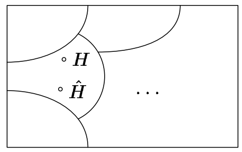
We finish this subsection by considering the easier case where representations are with respect to possibly different starting and ending bases. Recall that the prior subsection shows that a matrix changes bases if and only if it is nonsingular. That gives us another version of the above arrow diagram and equation (∗).

**2.3 Definition** Same-sized matrices $H$ and $\hat{H}$ are *matrix equivalent* if there are nonsingular matrices $P$ and $Q$ such that $\hat{H} = PHQ$.

**2.4 Corollary** Matrix equivalent matrices represent the same map, with respect to appropriate pairs of bases.

Exercise 19 checks that matrix equivalence is an equivalence relation. Thus it partitions the set of matrices into matrix equivalence classes.

All matrices:




$H$ matrix equivalent to $\hat{H}$

We can get some insight into the classes by comparing matrix equivalence with row equivalence (recall that matrices are row equivalent when they can be reduced to each other by row operations). In $\hat{H} = PHQ$, the matrices $P$ and $Q$ are nonsingular and thus each can be written as a product of elementary reduction matrices (Lemma **??**). Left-multiplication by the reduction matrices making up $P$ has the effect of performing row operations. Right-multiplication by the reduction matrices making up $Q$ performs column operations. Therefore, matrix equivalence is a generalization of row equivalence — two matrices are row equivalent if one can be converted to the other by a sequence of row reduction steps, while two matrices are matrix equivalent if one can be converted to the other by a sequence of row reduction steps followed by a sequence of column reduction steps.

Thus, if matrices are row equivalent then they are also matrix equivalent (since we can take $Q$ to be the identity matrix and so perform no column operations). The converse, however, does not hold.

**2.5 Example** These two

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

are matrix equivalent because the second can be reduced to the first by the column operation of taking $-1$ times the first column and adding to the second. They are not row equivalent because they have different reduced echelon forms (in fact, both are already in reduced form).

We will close this section by finding a set of representatives for the matrix equivalence classes.*

**2.6 Theorem** Any $m{\times}n$ matrix of rank $k$ is matrix equivalent to the $m{\times}n$ matrix that is all zeros except that the first $k$ diagonal entries are ones.

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ & & \vdots & & & & \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ & & \vdots & & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

Sometimes this is described as a *block partial-identity* form.

$$\left(\begin{array}{c|c} I & Z \\ \hline Z & Z \end{array}\right)$$

*More information on class representatives is in the appendix.

Proof. As discussed above, Gauss-Jordan reduce the given matrix and combine all the reduction matrices used there to make $P$. Then use the leading entries to do column reduction and finish by swapping columns to put the leading ones on the diagonal. Combine the reduction matrices used for those column operations into $Q$. QED

**2.7 Example** We illustrate the proof by finding the $P$ and $Q$ for this matrix.

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

First Gauss-Jordan row-reduce.

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

Then column-reduce, which involves right-multiplication.

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

Finish by swapping columns.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Finally, combine the left-multipliers together as $P$ and the right-multipliers together as $Q$ to get the $PHQ$ equation.

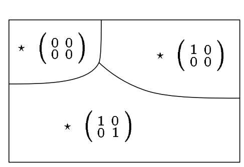
$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ 2 & 4 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0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

**2.8 Corollary** Two same-sized matrices are matrix equivalent if and only if they have the same rank. That is, the matrix equivalence classes are characterized by rank.

Proof. Two same-sized matrices with the same rank are equivalent to the same block partial-identity matrix. QED

**2.9 Example** The $2 \times 2$ matrices have only three possible ranks: zero, one, or two. Thus there are three matrix-equivalence classes.

All $2 \times 2$ matrices:



Three equivalence classes

Each class consists of all of the $2 \times 2$ matrices with the same rank. There is only one rank zero matrix, so that class has only one member, but the other two classes each have infinitely many members.

In this subsection we have seen how to change the representation of a map with respect to a first pair of bases to one with respect to a second pair. That led to a definition describing when matrices are equivalent in this way. Finally we noted that, with the proper choice of (possibly different) starting and ending bases, any map can be represented in block partial-identity form.

One of the nice things about this representation is that, in some sense, we can completely understand the map when it is expressed in this way: if the bases are $B = \langle \beta_1, \ldots, \beta_n \rangle$ and $D = \langle \delta_1, \ldots, \delta_m \rangle$ then the map sends

$$c_1\beta_1 + \cdots + c_k\beta_k + c_{k+1}\beta_{k+1} + \cdots + c_n\beta_n \quad \longmapsto \quad c_1\delta_1 + \cdots + c_k\delta_k + 0 + \cdots + 0$$

where $k$ is the map's rank. Thus, we can understand any linear map as a kind of projection.

$$\begin{pmatrix} c_1 \\ \vdots \\ c_k \\ c_{k+1} \\ \vdots \\ c_n \end{pmatrix}_B \mapsto \begin{pmatrix} c_1 \\ \vdots \\ c_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}_D$$

Of course, "understanding" a map expressed in this way requires that we understand the relationship between $B$ and $D$. However, despite that difficulty, this is a good classification of linear maps.

**Exercises**

✓ **2.10** Decide if these matrices are matrix equivalent.

**(a)** $\begin{pmatrix} 1 & 3 & 0 \\ 2 & 3 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 2 & 1 \\ 0 & 5 & -1 \end{pmatrix}$

**(b)** $\begin{pmatrix} 0 & 3 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 5 \end{pmatrix}$

**(c)** $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 2 & -6 \end{pmatrix}$

✓ **2.11** Find the canonical representative of the matrix-equivalence class of each matrix.

**(a)** $\begin{pmatrix} 2 & 1 & 0 \\ 4 & 2 & 0 \end{pmatrix}$ **(b)** $\begin{pmatrix} 0 & 1 & 0 & 2 \\ 1 & 1 & 0 & 4 \\ 3 & 3 & 3 & -1 \end{pmatrix}$

**2.12** Suppose that, with respect to

$$B = \mathcal{E}_2 \qquad D = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle$$

the transformation $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ is represented by this matrix.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Use change of basis matrices to represent $t$ with respect to each pair.

**(a)** $\hat{B} = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$, $\hat{D} = \langle \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rangle$

**(b)** $\hat{B} = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle$, $\hat{D} = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rangle$

✓ **2.13** What sizes are $P$ and $Q$ in the equation $\hat{H} = PHQ$?

✓ **2.14** Use Theorem 2.6 to show that a square matrix is nonsingular if and only if it is equivalent to an identity matrix.

✓ **2.15** Show that, where $A$ is a nonsingular square matrix, if $P$ and $Q$ are nonsingular square matrices such that $PAQ = I$ then $QP = A^{-1}$.

✓ **2.16** Why does Theorem 2.6 not show that every matrix is diagonalizable (see Example 2.2)?

**2.17** Must matrix equivalent matrices have matrix equivalent transposes?

**2.18** What happens in Theorem 2.6 if $k = 0$?

✓ **2.19** Show that matrix-equivalence is an equivalence relation.

✓ **2.20** Show that a zero matrix is alone in its matrix equivalence class. Are there other matrices like that?

**2.21** What are the matrix equivalence classes of matrices of transformations on $\mathbb{R}^1$? $\mathbb{R}^3$?

**2.22** How many matrix equivalence classes are there?

**2.23** Are matrix equivalence classes closed under scalar multiplication? Addition?

**2.24** Let $t\colon \mathbb{R}^n \to \mathbb{R}^n$ represented by $T$ with respect to $\mathcal{E}_n, \mathcal{E}_n$.

**(a)** Find $\text{Rep}_{B,B}(t)$ in this specific case.

$$T = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \qquad B = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \rangle$$

**(b)** Describe $\text{Rep}_{B,B}(t)$ in the general case where $B = \langle \beta_1, \ldots, \beta_n \rangle$.

**2.25** **(a)** Let $V$ have bases $B_1$ and $B_2$ and suppose that $W$ has the basis $D$. Where $h\colon V \to W$, find the formula that computes $\text{Rep}_{B_2,D}(h)$ from $\text{Rep}_{B_1,D}(h)$.

**(b)** Repeat the prior question with one basis for $V$ and two bases for $W$.

**2.26** **(a)** If two matrices are matrix-equivalent and invertible, must their inverses be matrix-equivalent?

**(b)** If two matrices have matrix-equivalent inverses, must the two be matrix-equivalent?

**(c)** If two matrices are square and matrix-equivalent, must their squares be matrix-equivalent?

**(d)** If two matrices are square and have matrix-equivalent squares, must they be matrix-equivalent?

✓ **2.27** Square matrices are *similar* if they represent the same transformation, but each with respect to the same ending as starting basis. That is, $\text{Rep}_{B_1,B_1}(t)$ is similar to $\text{Rep}_{B_2,B_2}(t)$.

**(a)** Give a definition of matrix similarity like that of Definition 2.3.

**(b)** Prove that similar matrices are matrix equivalent.

**(c)** Show that similarity is an equivalence relation.

**(d)** Show that if $T$ is similar to $\hat{T}$ then $T^2$ is similar to $\hat{T}^2$, the cubes are similar, etc. *Contrast with the prior exercise.*

**(e)** Prove that there are matrix equivalent matrices that are not similar.

Chapter Five

# Similarity and Diagonalization

# I Similarity

## I.1 Definition and Examples

We've defined $H$ and $\hat{H}$ to be matrix-equivalent if there are nonsingular matrices $P$ and $Q$ such that $\hat{H} = PHQ$. That definition is motivated by this diagram

$$\begin{array}{ccc} V_{\text{w.r.t. } B} & \xrightarrow[H]{h} & W_{\text{w.r.t. } D} \\ \text{id}\big\downarrow & & \text{id}\big\downarrow \\ V_{\text{w.r.t. } \hat{B}} & \xrightarrow[\hat{H}]{h} & W_{\text{w.r.t. } \hat{D}} \end{array}$$

showing that $H$ and $\hat{H}$ both represent $h$ but with respect to different pairs of bases. We now specialize that setup to the case where the codomain equals the domain, and where the codomain's basis equals the domain's basis.

$$\begin{array}{ccc} V_{\text{w.r.t. } B} & \xrightarrow{t} & V_{\text{w.r.t. } B} \\ \text{id}\big\downarrow & & \text{id}\big\downarrow \\ V_{\text{w.r.t. } D} & \xrightarrow{t} & V_{\text{w.r.t. } D} \end{array}$$

To move from the lower left to the lower right we can either go straight over, or up, over, and then down. In matrix terms,

$$\text{Rep}_{D,D}(t) = \text{Rep}_{B,D}(\text{id})\ \text{Rep}_{B,B}(t)\ \big(\text{Rep}_{B,D}(\text{id})\big)^{-1}$$

(recall that a representation of composition like this one reads right to left).

**1.1 Definition** The matrices $T$ and $S$ are *similar* if there is a nonsingular $P$ such that $T = PSP^{-1}$.

Since nonsingular matrices are square, the similar matrices $T$ and $S$ must be square and of the same size.

**1.2 Example** With these two,

$$P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad S = \begin{pmatrix} 2 & -3 \\ 1 & -1 \end{pmatrix}$$

calculation gives that $S$ is similar to this matrix.

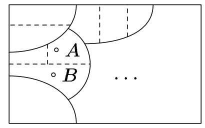
$$T = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

**1.3 Example** The only matrix similar to the zero matrix is itself: $PZP^{-1} = PZ = Z$. The only matrix similar to the identity matrix is itself: $PIP^{-1} = PP^{-1} = I$.

Since matrix similarity is a special case of matrix equivalence, if two matrices are similar then they are equivalent. What about the converse: must matrix equivalent square matrices be similar? The answer is no. The prior example shows that the similarity classes are different from the matrix equivalence classes, because the matrix equivalence class of the identity consists of all nonsingular matrices of that size. Thus, for instance, these two are matrix equivalent but not similar.

$$T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad S = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$$

So some matrix equivalence classes split into two or more similarity classes — similarity gives a finer partition than does equivalence. This picture shows some matrix equivalence classes subdivided into similarity classes.




To understand the similarity relation we shall study the similarity classes. We approach this question in the same way that we've studied both the row equivalence and matrix equivalence relations, by finding a canonical form for representatives* of the similarity classes, called Jordan form. With this canonical form, we can decide if two matrices are similar by checking whether they reduce to the same representative. We've also seen with both row equivalence and matrix equivalence that a canonical form gives us insight into the ways in which members of the same class are alike (e.g., two identically-sized matrices are matrix equivalent if and only if they have the same rank).

**Exercises**

**1.4** For

$$S=\begin{pmatrix}1&3\\-2&-6\end{pmatrix}\quad T=\begin{pmatrix}0&0\\-11/2&-5\end{pmatrix}\quad P=\begin{pmatrix}4&2\\-3&2\end{pmatrix}$$

check that $T=PSP^{-1}$.

✓ **1.5** Example 1.3 shows that the only matrix similar to a zero matrix is itself and that the only matrix similar to the identity is itself.

**(a)** Show that the $1\times 1$ matrix $(2)$, also, is similar only to itself.

**(b)** Is a matrix of the form $cI$ for some scalar $c$ similar only to itself?

**(c)** Is a diagonal matrix similar only to itself?

**1.6** Show that these matrices are not similar.

$$\begin{pmatrix}1&0&4\\1&1&3\\2&1&7\end{pmatrix}\qquad\begin{pmatrix}1&0&1\\0&1&1\\3&1&2\end{pmatrix}$$

**1.7** Consider the transformation $t\colon\mathcal{P}_2\to\mathcal{P}_2$ described by $x^2\mapsto x+1$, $x\mapsto x^2-1$, and $1\mapsto 3$.

**(a)** Find $T=\text{Rep}_{B,B}(t)$ where $B=\langle x^2,x,1\rangle$.

**(b)** Find $S=\text{Rep}_{D,D}(t)$ where $D=\langle 1,1+x,1+x+x^2\rangle$.

**(c)** Find the matrix $P$ such that $T=PSP^{-1}$.

✓ **1.8** Exhibit an nontrivial similarity relationship in this way: let $t\colon\mathbb{C}^2\to\mathbb{C}^2$ act by

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

and pick two bases, and represent $t$ with respect to then $T=\text{Rep}_{B,B}(t)$ and $S=\text{Rep}_{D,D}(t)$. Then compute the $P$ and $P^{-1}$ to change bases from $B$ to $D$ and back again.

**1.9** Explain Example 1.3 in terms of maps.

✓ **1.10** Are there two matrices $A$ and $B$ that are similar while $A^2$ and $B^2$ are not similar? [Halmos]

✓ **1.11** Prove that if two matrices are similar and one is invertible then so is the other.

✓ **1.12** Show that similarity is an equivalence relation.

**1.13** Consider a matrix representing, with respect to some $B,B$, reflection across the $x$-axis in $\mathbb{R}^2$. Consider also a matrix representing, with respect to some $D,D$, reflection across the $y$-axis. Must they be similar?

**1.14** Prove that similarity preserves determinants and rank. Does the converse hold?

**1.15** Is there a matrix equivalence class with only one matrix similarity class inside? One with infinitely many similarity classes?

**1.16** Can two different diagonal matrices be in the same similarity class?

✓ **1.17** Prove that if two matrices are similar then their $k$-th powers are similar when $k>0$. What if $k\leq 0$?

✓ **1.18** Let $p(x)$ be the polynomial $c_nx^n+\cdots+c_1x+c_0$. Show that if $T$ is similar to $S$ then $p(T)=c_nT^n+\cdots+c_1T+c_0I$ is similar to $p(S)=c_nS^n+\cdots+c_1S+c_0I$.

**1.19** List all of the matrix equivalence classes of $1\times 1$ matrices. Also list the similarity classes, and describe which similarity classes are contained inside of each matrix equivalence class.

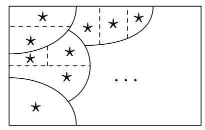
**1.20** Does similarity preserve sums?

**1.21** Show that if $T-\lambda I$ and $N$ are similar matrices then $T$ and $N+\lambda I$ are also similar.

*More information on representatives is in the appendix.

## I.2 Diagonalizability

The prior subsection defines the relation of similarity and shows that, although similar matrices are necessarily matrix equivalent, the converse does not hold. Some matrix-equivalence classes break into two or more similarity classes (the nonsingular $n \times n$ matrices, for instance). This means that the canonical form for matrix equivalence, a block partial-identity, cannot be used as a canonical form for matrix similarity because the partial-identities cannot be in more than one similarity class, so there are similarity classes without one. This picture illustrates. As earlier in this book, class representatives are shown with stars.



We are developing a canonical form for representatives of the similarity classes. We naturally try to build on our previous work, meaning first that the partial identity matrices should represent the similarity classes into which they fall, and beyond that, that the representatives should be as simple as possible. The simplest extension of the partial-identity form is a diagonal form.

**2.1 Definition** A transformation is *diagonalizable* if it has a diagonal representation with respect to the same basis for the codomain as for the domain. A *diagonalizable matrix* is one that is similar to a diagonal matrix: $T$ is diagonalizable if there is a nonsingular $P$ such that $PTP^{-1}$ is diagonal.

**2.2 Example** The matrix

$$\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$$

is diagonalizable.

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}^{-1}$$

**2.3 Example** Not every matrix is diagonalizable. The square of

$$N = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

is the zero matrix. Thus, for any map $n$ that $N$ represents (with respect to the same basis for the domain as for the codomain), the composition $n \circ n$ is the zero map. This implies that no such map $n$ can be diagonally represented (with respect to any $B, B$) because no power of a nonzero diagonal matrix is zero. That is, there is no diagonal matrix in $N$'s similarity class.

That example shows that a diagonal form will not do for a canonical form — we cannot find a diagonal matrix in each matrix similarity class. However, the canonical form that we are developing has the property that if a matrix can be diagonalized then the diagonal matrix is the canonical representative of the similarity class. The next result characterizes which maps can be diagonalized.

**2.4 Corollary** A transformation $t$ is diagonalizable if and only if there is a basis $B = \langle \beta_1, \ldots, \beta_n \rangle$ and scalars $\lambda_1, \ldots, \lambda_n$ such that $t(\beta_i) = \lambda_i \beta_i$ for each $i$.

Proof. This follows from the definition by considering a diagonal representation matrix.

$$\text{Rep}_{B,B}(t) = \left(\begin{array}{c|c|c} \vdots & & \vdots \\ \text{Rep}_B(t(\beta_1)) & \cdots & \text{Rep}_B(t(\beta_n)) \\ \vdots & & \vdots \end{array}\right) = \left(\begin{array}{c|c|c} \lambda_1 & & 0 \\ \vdots & \ddots & \vdots \\ 0 & & \lambda_n \end{array}\right)$$

This representation is equivalent to the existence of a basis satisfying the stated conditions simply by the definition of matrix representation. QED

**2.5 Example** To diagonalize

$$T = \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}$$

we take it as the representation of a transformation with respect to the standard basis $T = \text{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(t)$ and we look for a basis $B = \langle \beta_1, \beta_2 \rangle$ such that

$$\text{Rep}_{B,B}(t) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

that is, such that $t(\beta_1) = \lambda_1 \beta_1$ and $t(\beta_2) = \lambda_2 \beta_2$.

$$\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \beta_1 = \lambda_1 \cdot \beta_1 \qquad \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \beta_2 = \lambda_2 \cdot \beta_2$$

We are looking for scalars $x$ such that this equation

$$\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = x \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

has solutions $b_1$ and $b_2$, which are not both zero. Rewrite that as a linear system.

$$\begin{array}{rr} (3-x) \cdot b_1 + & 2 \cdot b_2 = 0 \\ & (1-x) \cdot b_2 = 0 \end{array} \qquad (*)$$

In the bottom equation the two numbers multiply to give zero only if at least one of them is zero so there are two possibilities, $b_2 = 0$ and $x = 1$. In the $b_2 = 0$ possibility, the first equation gives that either $b_1 = 0$ or $x = 3$. Since the case of both $b_1 = 0$ and $b_2 = 0$ is disallowed, we are left looking at the possibility of $x = 3$. With it, the first equation in $(*)$ is $0 \cdot b_1 + 2 \cdot b_2 = 0$ and so associated with 3 are vectors with a second component of zero and a first component that is free.

$$\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ 0 \end{pmatrix} = 3 \cdot \begin{pmatrix} b_1 \\ 0 \end{pmatrix}$$

That is, one solution to $(*)$ is $\lambda_1 = 3$, and we have a first basis vector.

$$\beta_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In the $x = 1$ possibility, the first equation in $(*)$ is $2 \cdot b_1 + 2 \cdot b_2 = 0$, and so associated with 1 are vectors whose second component is the negative of their first component.

$$\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ -b_1 \end{pmatrix} = 1 \cdot \begin{pmatrix} b_1 \\ -b_1 \end{pmatrix}$$

Thus, another solution is $\lambda_2 = 1$ and a second basis vector is this.

$$\beta_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

To finish, drawing the similarity diagram

$$\begin{array}{ccc} \mathbb{R}^2_{\text{w.r.t. } \mathcal{E}_2} & \xrightarrow[T]{t} & \mathbb{R}^2_{\text{w.r.t. } \mathcal{E}_2} \\ \text{id}\big\downarrow & & \text{id}\big\downarrow \\ \mathbb{R}^2_{\text{w.r.t. } B} & \xrightarrow[D]{t} & \mathbb{R}^2_{\text{w.r.t. } B} \end{array}$$

and noting that the matrix $\text{Rep}_{B,\mathcal{E}_2}(\text{id})$ is easy leads to this diagonalization.

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

In the next subsection, we will expand on that example by considering more closely the property of Corollary 2.4. This includes seeing another way, the way that we will routinely use, to find the $\lambda$'s.

## Exercises

✓ **2.6** Repeat Example 2.5 for the matrix from Example 2.2.

**2.7** Diagonalize these upper triangular matrices.

**(a)** $\begin{pmatrix} -2 & 1 \\ 0 & 2 \end{pmatrix}$ **(b)** $\begin{pmatrix} 5 & 4 \\ 0 & 1 \end{pmatrix}$

✓ **2.8** What form do the powers of a diagonal matrix have?

**2.9** Give two same-sized diagonal matrices that are not similar. Must any two different diagonal matrices come from different similarity classes?

**2.10** Give a nonsingular diagonal matrix. Can a diagonal matrix ever be singular?

✓ **2.11** Show that the inverse of a diagonal matrix is the diagonal of the the inverses, if no element on that diagonal is zero. What happens when a diagonal entry is zero?

**2.12** The equation ending Example 2.5

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

is a bit jarring because for $P$ we must take the first matrix, which is shown as an inverse, and for $P^{-1}$ we take the inverse of the first matrix, so that the two $-1$ powers cancel and this matrix is shown without a superscript $-1$.

**(a)** Check that this nicer-appearing equation holds.

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

**(b)** Is the previous item a coincidence? Or can we always switch the $P$ and the $P^{-1}$?

**2.13** Show that the $P$ used to diagonalize in Example 2.5 is not unique.

**2.14** Find a formula for the powers of this matrix *Hint*: see Exercise 8.

$$\begin{pmatrix} -3 & 1 \\ -4 & 2 \end{pmatrix}$$

✓ **2.15** Diagonalize these.

**(a)** $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ **(b)** $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

**2.16** We can ask how diagonalization interacts with the matrix operations. Assume that $t, s\colon V \to V$ are each diagonalizable. Is $ct$ diagonalizable for all scalars $c$? What about $t+s$? $t \circ s$?

✓ **2.17** Show that matrices of this form are not diagonalizable.

$$\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \qquad c \neq 0$$

**2.18** Show that each of these is diagonalizable.

**(a)** $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ **(b)** $\begin{pmatrix} x & y \\ y & z \end{pmatrix}$ $\quad x, y, z$ scalars

## I.3 Eigenvalues and Eigenvectors

In this subsection we will focus on the property of Corollary 2.4.

**3.1 Definition** A transformation $t\colon V \to V$ has a scalar *eigenvalue* $\lambda$ if there is a nonzero *eigenvector* $\zeta \in V$ such that $t(\zeta) = \lambda \cdot \zeta$.

("Eigen" is German for "characteristic of" or "peculiar to"; some authors call these *characteristic* values and vectors. No authors call them "peculiar".)

**3.2 Example** The projection map

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \qquad x, y, z \in \mathbb{C}$$

has an eigenvalue of 1 associated with any eigenvector of the form

$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

where $x$ and $y$ are non-0 scalars. On the other hand, 2 is not an eigenvalue of $\pi$ since no non-0 vector is doubled.

That example shows why the 'non-0' appears in the definition. Disallowing 0 as an eigenvector eliminates trivial eigenvalues.

**3.3 Example** The only transformation on the trivial space $\{\vec{0}\,\}$ is $\vec{0} \mapsto \vec{0}$. This map has no eigenvalues because there are no non-0 vectors $\vec{v}$ mapped to a scalar multiple $\lambda \cdot \vec{v}$ of themselves.

**3.4 Example** Consider the homomorphism $t\colon \mathcal{P}_1 \to \mathcal{P}_1$ given by $c_0 + c_1 x \mapsto (c_0 + c_1) + (c_0 + c_1)x$. The range of $t$ is one-dimensional. Thus an application of $t$ to a vector in the range will simply rescale that vector: $c + cx \mapsto (2c) + (2c)x$. That is, $t$ has an eigenvalue of 2 associated with eigenvectors of the form $c + cx$ where $c \neq 0$.

This map also has an eigenvalue of 0 associated with eigenvectors of the form $c - cx$ where $c \neq 0$.

**3.5 Definition** A square matrix $T$ has a scalar *eigenvalue* $\lambda$ associated with the non-0 *eigenvector* $\vec{\zeta}$ if $T\vec{\zeta} = \lambda \cdot \vec{\zeta}$.

**3.6 Remark** Although this extension from maps to matrices is obvious, there is a point that must be made. Eigenvalues of a map are also the eigenvalues of matrices representing that map, and so similar matrices have the same eigenvalues. But the eigenvectors are different — similar matrices need not have the same eigenvectors.

For instance, consider again the transformation $t\colon \mathcal{P}_1 \to \mathcal{P}_1$ given by $c_0 + c_1 x \mapsto (c_0 + c_1) + (c_0 + c_1)x$. It has an eigenvalue of 2 associated with eigenvectors of the form $c + cx$ where $c \neq 0$. If we represent $t$ with respect to $B = \langle 1 + 1x, 1 - 1x \rangle$

$$T = \text{Rep}_{B,B}(t) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

then 2 is an eigenvalue of $T$, associated with these eigenvectors.

$$\{\begin{pmatrix} c_0 \\ c_1 \end{pmatrix} \bigm| \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 2c_0 \\ 2c_1 \end{pmatrix}\} = \{\begin{pmatrix} c_0 \\ 0 \end{pmatrix} \bigm| c_0 \in \mathbb{C},\ c_0 \neq 0\}$$

On the other hand, representing $t$ with respect to $D = \langle 2 + 1x, 1 + 0x \rangle$ gives

$$S = \text{Rep}_{D,D}(t) = \begin{pmatrix} 3 & 1 \\ -3 & -1 \end{pmatrix}$$

and the eigenvectors of $S$ associated with the eigenvalue 2 are these.

$$\{\begin{pmatrix} c_0 \\ c_1 \end{pmatrix} \bigm| \begin{pmatrix} 3 & 1 \\ -3 & -1 \end{pmatrix}\begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 2c_0 \\ 2c_1 \end{pmatrix}\} = \{\begin{pmatrix} 0 \\ c_1 \end{pmatrix} \bigm| c_1 \in \mathbb{C},\ c_1 \neq 0\}$$

Thus similar matrices can have different eigenvectors.

Here is an informal description of what's happening. The underlying transformation doubles the eigenvectors $\vec{v} \mapsto 2 \cdot \vec{v}$. But when the matrix representing the transformation is $T = \text{Rep}_{B,B}(t)$ then it "assumes" that column vectors are representations with respect to $B$. In contrast, $S = \text{Rep}_{D,D}(t)$ "assumes" that column vectors are representations with respect to $D$. So the vectors that get doubled by each matrix look different.

The next example illustrates the basic tool for finding eigenvectors and eigenvalues.

**3.7 Example** What are the eigenvalues and eigenvectors of this matrix?

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

To find the scalars $x$ such that $T\zeta = x\zeta$ for non-0 eigenvectors $\zeta$, bring everything to the left-hand side

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} - x \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = 0$$

and factor $(T - xI)\zeta = 0$. (Note that it says $T - xI$; the expression $T - x$ doesn't make sense because $T$ is a matrix while $x$ is a scalar.) This homogeneous linear system

$$\begin{pmatrix} 1-x & 2 & 1 \\ 2 & 0-x & -2 \\ -1 & 2 & 3-x \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has a non-0 solution if and only if the matrix is singular. We can determine when that happens.

$$\begin{aligned} 0 &= |T - xI| \\ &= \begin{vmatrix} 1-x & 2 & 1 \\ 2 & 0-x & -2 \\ -1 & 2 & 3-x \end{vmatrix} \\ &= x^3 - 4x^2 + 4x \\ &= x(x-2)^2 \end{aligned}$$

The eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = 2$. To find the associated eigenvectors, plug in each eigenvalue. Plugging in $\lambda_1 = 0$ gives

$$\begin{pmatrix} 1-0 & 2 & 1 \\ 2 & 0-0 & -2 \\ -1 & 2 & 3-0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} a \\ -a \\ a \end{pmatrix}$$

for a scalar parameter $a \neq 0$ ($a$ is non-0 because eigenvectors must be non-0). In the same way, plugging in $\lambda_2 = 2$ gives

$$\begin{pmatrix} 1-2 & 2 & 1 \\ 2 & 0-2 & -2 \\ -1 & 2 & 3-2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} b \\ 0 \\ b \end{pmatrix}$$

with $b \neq 0$.

**3.8 Example** If

$$S = \begin{pmatrix} \pi & 1 \\ 0 & 3 \end{pmatrix}$$

(here $\pi$ is not a projection map, it is the number $3.14\ldots$) then

$$\left| \begin{pmatrix} \pi - x & 1 \\ 0 & 3-x \end{pmatrix} \right| = (x-\pi)(x-3)$$

so $S$ has eigenvalues of $\lambda_1 = \pi$ and $\lambda_2 = 3$. To find associated eigenvectors, first plug in $\lambda_1$ for $x$:

$$\begin{pmatrix} \pi - \pi & 1 \\ 0 & 3-\pi \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}$$

for a scalar $a \neq 0$, and then plug in $\lambda_2$:

$$\begin{pmatrix} \pi - 3 & 1 \\ 0 & 3-3 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} -b/(\pi-3) \\ b \end{pmatrix}$$

where $b \neq 0$.

**3.9 Definition** The *characteristic polynomial* of a square matrix $T$ is the determinant of the matrix $T - xI$, where $x$ is a variable. The *characteristic equation* is $|T - xI| = 0$. The characteristic polynomial of a transformation $t$ is the polynomial of any $\text{Rep}_{B,B}(t)$.

Exercise 30 checks that the characteristic polynomial of a transformation is well-defined, that is, any choice of basis yields the same polynomial.

**3.10 Lemma** A linear transformation on a nontrivial vector space has at least one eigenvalue.

Proof. Any root of the characteristic polynomial is an eigenvalue. Over the complex numbers, any polynomial of degree one or greater has a root. (This is the reason that in this chapter we've gone to scalars that are complex.) QED

Notice the familiar form of the sets of eigenvectors in the above examples.

**3.11 Definition** The *eigenspace* of a transformation $t$ associated with the eigenvalue $\lambda$ is $V_\lambda = \{\zeta \mid t(\zeta) = \lambda\zeta\} \cup \{0\}$. The eigenspace of a matrix is defined analogously.

**3.12 Lemma** An eigenspace is a subspace.

Proof. An eigenspace must be nonempty — for one thing it contains the zero vector — and so we need only check closure. Take vectors $\zeta_1, \ldots, \zeta_n$ from $V_\lambda$, to show that any linear combination is in $V_\lambda$

$$\begin{aligned} t(c_1\zeta_1 + c_2\zeta_2 + \cdots + c_n\zeta_n) &= c_1 t(\zeta_1) + \cdots + c_n t(\zeta_n) \\ &= c_1\lambda\zeta_1 + \cdots + c_n\lambda\zeta_n \\ &= \lambda(c_1\zeta_1 + \cdots + c_n\zeta_n) \end{aligned}$$

(the second equality holds even if any $\zeta_i$ is $0$ since $t(0) = \lambda \cdot 0 = 0$). QED

**3.13 Example** In Example 3.8 the eigenspace associated with the eigenvalue $\pi$ and the eigenspace associated with the eigenvalue 3 are these.

$$V_\pi = \{\begin{pmatrix} a \\ 0 \end{pmatrix} \mid a \in \mathbb{R}\} \qquad V_3 = \{\begin{pmatrix} -b/\pi - 3 \\ b \end{pmatrix} \mid b \in \mathbb{R}\}$$

**3.14 Example** In Example 3.7, these are the eigenspaces associated with the eigenvalues 0 and 2.

$$V_0 = \{\begin{pmatrix} a \\ -a \\ a \end{pmatrix} \mid a \in \mathbb{R}\}, \qquad V_2 = \{\begin{pmatrix} b \\ 0 \\ b \end{pmatrix} \mid b \in \mathbb{R}\}.$$

**3.15 Remark** The characteristic equation is $0 = x(x-2)^2$ so in some sense 2 is an eigenvalue "twice". However there are not "twice" as many eigenvectors, in that the dimension of the eigenspace is one, not two. The next example shows a case where a number, 1, is a double root of the characteristic equation and the dimension of the associated eigenspace is two.

**3.16 Example** With respect to the standard bases, this matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

represents projection.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \qquad x, y, z \in \mathbb{C}$$

Its eigenspace associated with the eigenvalue 0 and its eigenspace associated with the eigenvalue 1 are easy to find.

$$V_0 = \{\begin{pmatrix} 0 \\ 0 \\ c_3 \end{pmatrix} \mid c_3 \in \mathbb{C}\} \qquad V_1 = \{\begin{pmatrix} c_1 \\ c_2 \\ 0 \end{pmatrix} \mid c_1, c_2 \in \mathbb{C}\}$$

By the lemma, if two eigenvectors $v_1$ and $v_2$ are associated with the same eigenvalue then any linear combination of those two is also an eigenvector associated with that same eigenvalue. But, if two eigenvectors $v_1$ and $v_2$ are associated with different eigenvalues then the sum $v_1 + v_2$ need not be related to the eigenvalue of either one. In fact, just the opposite. If the eigenvalues are different then the eigenvectors are not linearly related.

**3.17 Theorem** For any set of distinct eigenvalues of a map or matrix, a set of associated eigenvectors, one per eigenvalue, is linearly independent.

Proof. We will use induction on the number of eigenvalues. If there is no eigenvalue or only one eigenvalue then the set of associated eigenvectors is empty or is a singleton set with a non-0 member, and in either case is linearly independent.

For induction, assume that the theorem is true for any set of $k$ distinct eigenvalues, suppose that $\lambda_1, \ldots, \lambda_{k+1}$ are distinct eigenvalues, and let $v_1, \ldots, v_{k+1}$ be associated eigenvectors. If $c_1 v_1 + \cdots + c_k v_k + c_{k+1} v_{k+1} = 0$ then after multiplying both sides of the displayed equation by $\lambda_{k+1}$, applying the map or matrix to both sides of the displayed equation, and subtracting the first result from the second, we have this.

$$c_1(\lambda_{k+1} - \lambda_1)v_1 + \cdots + c_k(\lambda_{k+1} - \lambda_k)v_k + c_{k+1}(\lambda_{k+1} - \lambda_{k+1})v_{k+1} = 0$$

The induction hypothesis now applies: $c_1(\lambda_{k+1} - \lambda_1) = 0, \ldots, c_k(\lambda_{k+1} - \lambda_k) = 0$. Thus, as all the eigenvalues are distinct, $c_1, \ldots, c_k$ are all 0. Finally, now $c_{k+1}$ must be 0 because we are left with the equation $v_{k+1} \neq 0$. QED

**3.18 Example** The eigenvalues of

$$\begin{pmatrix} 2 & -2 & 2 \\ 0 & 1 & 1 \\ -4 & 8 & 3 \end{pmatrix}$$

are distinct: $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 3$. A set of associated eigenvectors like

$$\{\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 9 \\ 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}\}$$

is linearly independent.

**3.19 Corollary** An $n \times n$ matrix with $n$ distinct eigenvalues is diagonalizable.

Proof. Form a basis of eigenvectors. Apply Corollary 2.4. QED

### Exercises

**3.20** For each, find the characteristic polynomial and the eigenvalues.
**(a)** $\begin{pmatrix} 10 & -9 \\ 4 & -2 \end{pmatrix}$ **(b)** $\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ **(c)** $\begin{pmatrix} 0 & 3 \\ 7 & 0 \end{pmatrix}$ **(d)** $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ **(e)** $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

✓ **3.21** For each matrix, find the characteristic equation, and the eigenvalues and associated eigenvectors.
**(a)** $\begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$

**3.22** Find the characteristic equation, and the eigenvalues and associated eigenvectors for this matrix. *Hint.* The eigenvalues are complex.

$$\begin{pmatrix} -2 & -1 \\ 5 & 2 \end{pmatrix}$$

**3.23** Find the characteristic polynomial, the eigenvalues, and the associated eigenvectors of this matrix.

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

✓ **3.24** For each matrix, find the characteristic equation, and the eigenvalues and associated eigenvectors.

**(a)** $\begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ **(b)** $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{pmatrix}$

✓ **3.25** Let $t\colon \mathcal{P}_2 \to \mathcal{P}_2$ be

$$a_0 + a_1x + a_2x^2 \mapsto (5a_0 + 6a_1 + 2a_2) - (a_1 + 8a_2)x + (a_0 - 2a_2)x^2.$$

Find its eigenvalues and the associated eigenvectors.

**3.26** Find the eigenvalues and eigenvectors of this map $t\colon \mathcal{M}_2 \to \mathcal{M}_2$.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} 2c & a+c \\ b-2c & d \end{pmatrix}$$

✓ **3.27** Find the eigenvalues and associated eigenvectors of the differentiation operator $d/dx\colon \mathcal{P}_3 \to \mathcal{P}_3$.

**3.28** Prove that the eigenvalues of a triangular matrix (upper or lower triangular) are the entries on the diagonal.

✓ **3.29** Find the formula for the characteristic polynomial of a $2\times 2$ matrix.

**3.30** Prove that the characteristic polynomial of a transformation is well-defined.

✓ **3.31** **(a)** Can any non-0 vector in any nontrivial vector space be a eigenvector? That is, given a $v \neq 0$ from a nontrivial $V$, is there a transformation $t\colon V \to V$ and a scalar $\lambda \in \mathbb{R}$ such that $t(v) = \lambda v$?

**(b)** Given a scalar $\lambda$, can any non-0 vector in any nontrivial vector space be an eigenvector associated with the eigenvalue $\lambda$?

✓ **3.32** Suppose that $t\colon V \to V$ and $T = \text{Rep}_{B,B}(t)$. Prove that the eigenvectors of $T$ associated with $\lambda$ are the non-0 vectors in the kernel of the map represented (with respect to the same bases) by $T - \lambda I$.

**3.33** Prove that if $a, \ldots, d$ are all integers and $a + b = c + d$ then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

has integral eigenvalues, namely $a + b$ and $a - c$.

✓ **3.34** Prove that if $T$ is nonsingular and has eigenvalues $\lambda_1, \ldots, \lambda_n$ then $T^{-1}$ has eigenvalues $1/\lambda_1, \ldots, 1/\lambda_n$. Is the converse true?

✓ **3.35** Suppose that $T$ is $n\times n$ and $c, d$ are scalars.

**(a)** Prove that if $T$ has the eigenvalue $\lambda$ with an associated eigenvector $v$ then $v$ is an eigenvector of $cT + dI$ associated with eigenvalue $c\lambda + d$.

**(b)** Prove that if $T$ is diagonalizable then so is $cT + dI$.

✓ **3.36** Show that $\lambda$ is an eigenvalue of $T$ if and only if the map represented by $T - \lambda I$ is not an isomorphism.

**3.37** [Strang 80]

**(a)** Show that if $\lambda$ is an eigenvalue of $A$ then $\lambda^k$ is an eigenvalue of $A^k$.

**(b)** What is wrong with this proof generalizing that? "If $\lambda$ is an eigenvalue of $A$ and $\mu$ is an eigenvalue for $B$, then $\lambda\mu$ is an eigenvalue for $AB$, for, if $Ax = \lambda x$ and $Bx = \mu x$ then $ABx = A\mu x = \mu Ax \mu\lambda x$"?

**3.38** Do matrix-equivalent matrices have the same eigenvalues?

**3.39** Show that a square matrix with real entries and an odd number of rows has at least one real eigenvalue.

**3.40** Diagonalize.

$$\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}$$

**3.41** Suppose that $P$ is a nonsingular $n\times n$ matrix. Show that the *similarity transformation* map $t_P\colon \mathcal{M}_{n\times n} \to \mathcal{M}_{n\times n}$ sending $T \mapsto PTP^{-1}$ is an isomorphism.

**? 3.42** Show that if $A$ is an $n$ square matrix and each row (column) sums to $c$ then $c$ is a characteristic root of $A$. [Math. Mag., Nov. 1967]

# Bibliography

[Abbot] Edwin Abbott, *Flatland*, Dover.

[Ackerson] R. H. Ackerson, *A Note on Vector Spaces*, American Mathematical Monthly, volume 62 number 10 (Dec. 1955), p. 721.

[Agnew] Jeanne Agnew, *Explorations in Number Theory*, Brooks/Cole, 1972.

[Am. Math. Mon., Jun. 1931] C. A. Rupp (proposer), H. T. R. Aude (solver), problem 3468, American Mathematical Monthly, volume 37 number 6 (June-July 1931), p. 355.

[Am. Math. Mon., Feb. 1933] V. F. Ivanoff (proposer), T. C. Esty (solver), problem 3529, American Mathematical Monthly, volume 39 number 2 (Feb. 1933), p. 118.

[Am. Math. Mon., Jan. 1935] W. R. Ransom (proposer), Hansraj Gupta (solver), Elementary problem 105, American Mathematical Monthly, volume 42 number 1 (Jan. 1935), p. 47.

[Am. Math. Mon., Jan. 1949] C. W. Trigg (proposer), R. J. Walker (solver), Elementary problem 813, American Mathematical Monthly, volume 56 number 1 (Jan. 1949), p. 33.

[Am. Math. Mon., Jun. 1949] Don Walter (proposer), Alex Tytun (solver), Elementary problem 834, American Mathematical Monthly, volume 56 number 6 (June-July 1949), p. 409.

[Am. Math. Mon., Nov. 1951] Albert Wilansky, *The Row-Sums of the Inverse Matrix*, American Mathematical Monthly, volume 58 number 9 (Nov. 1951), p. 614.

[Am. Math. Mon., Feb. 1953] Norman Anning (proposer), C. W. Trigg (solver), Elementary problem 1016, American Mathematical Monthly, volume 60 number 2 (Feb. 1953), p. 115.

[Am. Math. Mon., Apr. 1955] Vern Haggett (proposer), F. W. Saunders (solver), Elementary problem 1135, American Mathematical Monthly, volume 62 number 4 (Apr. 1955), p. 257.

[Am. Math. Mon., Jan. 1963] Underwood Dudley, Arnold Lebow (proposers), David Rothman (solver), Elemantary problem 1151, American Mathematical Monthly, volume 70 number 1 (Jan. 1963), p. 93.

[Am. Math. Mon., Dec. 1966] Hans Liebeck, *A Proof of the Equality of Column Rank and Row Rank of a Matrix* American Mathematical Monthly, volume 73 number 10 (Dec. 1966), p. 1114.

[Anton] Howard Anton, *Elementary Linear Algebra*, John Wiley & Sons, 1987.

[Arrow] Kenneth J. Arrow, *Social Choice and Individual Values*, Wiley, 1963.

[Ball & Coxeter] W.W. Rouse Ball, *Mathematical Recreations and Essays*, revised by H.S.M. Coxeter, MacMillan, 1962.

[Bennett] William J. Bennett, *Quantifying America's Decline*, in *Wall Street Journal*, 15 March, 1993.

[Birkhoff & MacLane] Garrett Birkhoff, Saunders MacLane, *Survey of Modern Algebra*, third edition, Macmillan, 1965.

[Blass 1984] A. Blass, *Existence of Bases Implies the Axiom of Choice*, pp. 31 – 33, Axiomatic Set Theory, J. E. Baumgartner, ed., American Mathematical Society, Providence RI, 1984.

[Bridgman] P.W. Bridgman, *Dimensional Analysis*, Yale University Press, 1931.

[Casey] John Casey, *The Elements of Euclid, Books I to VI and XI*, ninth edition, Hodges, Figgis, and Co., Dublin, 1890.

[Clark & Coupe] David H. Clark, John D. Coupe, *The Bangor Area Economy Its Present and Future*, report to the city of Bangor ME, Mar. 1967.

[Clarke] Arthur C. Clarke, *Technical Error*, Fantasy, December 1946, reprinted in Great SF Stories 8 (1946), DAW Books, 1982.

[Con. Prob. 1955] *The Contest Problem Book*, 1955 number 38.

[Coxeter] H.S.M. Coxeter, *Projective Geometry*, second edition, Springer-Verlag, 1974.

[Courant & Robbins] Richard Courant, Herbert Robbins, *What is Mathematics?*, Oxford University Press, 1978.

[Cullen] Charles G. Cullen, *Matrices and Linear Transformations*, second edition, Dover, 1990.

[Dalal, et. al.] Siddhartha R. Dalal, Edward B. Fowlkes, & Bruce Hoadley, *Lesson Learned from Challenger: A Statistical Perspective*, Stats: the Magazine for Students of Statistics, Fall 1989, p. 3.

[Davies] Thomas D. Davies, *New Evidence Places Peary at the Pole*, National Geographic Magazine, vol. 177 no. 1 (Jan. 1990), p. 44.

[de Mestre] Neville de Mestre, *The Mathematics of Projectiles in Sport*, Cambridge University Press, 1990.

[De Parville] De Parville, *La Nature*, Paris, 1884, part I, p. 285 – 286 (citation from [Ball & Coxeter]).

[Duncan] W.J. Duncan, *Method of Dimensions*, in *Encyclopaedic Dictionary of Physics* Macmillan, 1962.

[Ebbing] Darrell D. Ebbing, *General Chemistry*, fourth edition, Houghton Mifflin, 1993.

[Ebbinghaus] H. D. Ebbinghaus, *Numbers*, Springer-Verlag, 1990.

[Einstein] A. Einstein, Annals of Physics, v. 35, 1911, p. 686.

[Eggar] M.H. Eggar, *Pinhole Cameras, Perspecitve, and Projective Geometry*, American Mathematicsl Monthly, August-September 1998, p. 618 – 630.

[Feller] William Feller, *An Introduction to Probability Theory and Its Applications* (vol. 1, 3rd ed.), John Wiley, 1968.

[Finkbeiner] Daniel T. Finkbeiner III, *Introduction to Matrices and Linear Transformations*, third edition, W. H. Freeman and Company, 1978.

[Fraleigh & Beauregard] Fraleigh & Beauregard, *Linear Algebra*.

[Gardner] Martin Gardner, *The New Ambidextrous Universe*, third revised edition, W. H. Freeman and Company, 1990.

[Gardner 1957] Martin Gardner, *Mathematical Games: About the remarkable similarity between the Icosian Game and the Tower of Hanoi*, Scientific American, May 1957, p. 150 – 154.

[Gardner, 1970] Martin Gardner, *Mathematical Games, Some mathematical curiosities embedded in the solar system*, Scientific American, April 1970, p. 108 – 112.

[Gardner, 1980] Martin Gardner, *Mathematical Games, From counting votes to making votes count: the mathematics of elections*, Scientific American, October 1980.

[Gardner, 1974] Martin Gardner, *Mathematical Games, On the paradoxical situations that arise from nontransitive relations*, Scientific American, October 1974.

[Giordano, Wells, Wilde] Frank R. Giordano, Michael E. Wells, Carroll O. Wilde, *Dimensional Analysis*, UMAP Unit 526, in *UMAP Modules, 1987*, COMAP, 1987.

[Giordano, Jaye, Weir] Frank R. Giordano, Michael J. Jaye, Maurice D. Weir, *The Use of Dimensional Analysis in Mathematical Modeling*, UMAP Unit 632, in *UMAP Modules, 1986*, COMAP, 1986.

[Glazer] A.M. Glazer, *The Structure of Crystals*, Adam Hilger, 1987.

[Goult, *et al.*] R.J. Goult, R.F. Hoskins, J.A. Milner, M.J. Pratt, *Computational Methods in Linear Algebra*, Wiley, 1975.

[Graham, Knuth, Patashnik] Ronald L. Graham, Donald E. Knuth, Oren Patashnik, *Concrete Mathematics*, Addison-Wesley, 1988.

[Halmos] Paul P. Halmos, *Finite Dimensional Vector Spaces*, second edition, Van Nostrand, 1958.

[Hamming] Richard W. Hamming, *Introduction to Applied Numerical Analysis*, Hemisphere Publishing, 1971.

[Hanes] Kit Hanes, *Analytic Projective Geometry and its Applications*, UMAP Unit 710, UMAP Modules, 1990, p. 111.

[Heath] T. Heath, *Euclid's Elements*, volume 1, Dover, 1956.

[Hoffman & Kunze] Kenneth Hoffman, Ray Kunze, *Linear Algebra*, second edition, Prentice-Hall, 1971.

[Hofstadter] Douglas R. Hofstadter, *Metamagical Themas: Questing for the Essence of Mind and Pattern*, Basic Books, 1985.

[Iosifescu] Marius Iofescu, *Finite Markov Processes and Their Applications*, John Wiley, 1980.

[Kelton] Christina M.L. Kelton, *Trends on the Relocation of U.S. Manufacturing*, UMI Research Press, 1983.

[Kemeny & Snell] John G. Kemeny, J. Laurie Snell, *Finite Markov Chains*, D. Van Nostrand, 1960.

[Kemp] Franklin Kemp *Linear Equations*, American Mathematical Monthly, volume 89 number 8 (Oct. 1982), p. 608.

[Knuth] Donald E. Knuth, *The Art of Computer Programming*, Addison Wesley, 1988.

[Lang] Serge Lang, *Linear Algebra*, Addison-Wesley, 1966.

[Leontief 1951] Wassily W. Leontief, *Input-Output Economics*, Scientific American, volume 185 number 4 (Oct. 1951), p. 15.

[Leontief 1965] Wassily W. Leontief, *The Structure of the U.S. Economy*, Scientific American, volume 212 number 4 (Apr. 1965), p. 25.

[Lieberman] David Lieberman, *The Big Bucks Ad Battles Over TV's Most Expensive Minutes*, TV Guide, Jan. 26 1991, p. 11.

[Macdonald & Ridge] Kenneth Macdonald, John Ridge, *Social Mobility*, in *British Social Trends Since 1900*, A.H. Halsey, Macmillian, 1988.

[Macmillan Dictionary] William D. Halsey, Macmillan, 1979.

[Math. Mag., Sept. 1952] Dewey Duncan (proposer), W. H. Quelch (solver), Mathematics Magazine, volume 26 number 1 (Sept-Oct. 1952), p. 48.

[Math. Mag., Jan. 1957] M. S. Klamkin (proposer), Trickie T-27, Mathematics Magazine, volume 30 number 3 (Jan-Feb. 1957), p. 173.

[Math. Mag., Jan. 1963, Q237] D. L. Silverman (proposer), C. W. Trigg (solver), Quickie 237, Mathematics Magazine, volume 36 number 1 (Jan. 1963).

[Math. Mag., Jan. 1963, Q307] C. W. Trigg (proposer). Quickie 307, Mathematics Magazine, volume 36 number 1 (Jan. 1963), p. 77.

[Math. Mag., Nov. 1967] Clarence C. Morrison (proposer), Quickie, Mathematics Magazine, volume 40 number 4 (Nov. 1967), p. 232.

[Math. Mag., Jan. 1973] Marvin Bittinger (proposer), Quickie 578, Mathematics Magazine, volume 46 number 5 (Jan. 1973), p. 286, 296.

[Munkres] James R. Munkres, *Elementary Linear Algebra*, Addison-Wesley, 1964.

[Neimi & Riker] Richard G. Neimi, William H. Riker, *The Choice of Voting Systems*, Scientific American, June 1976, p. 21 – 27.

[Nering] Evar D. Nering, *Linear Algebra and Matrix Theory*, second edition, John Wiley, 1970.

[Niven & Zuckerman] I. Niven, H. Zuckerman, *An Introduction to the Theory of Numbers*, third edition, John Wiley, 1972.

[Oakley & Baker] Cletus O. Oakley, Justine C. Baker, *Least Squares and the* 3 : 40 *Mile*, Mathematics Teacher, Apr. 1977.

[Ohanian] Hans O'Hanian, *Physics*, volume one, W. W. Norton, 1985.

[Onan] Onan, *Linear Algebra*.

[Petersen] G. M. Petersen, *Area of a Triangle*, American Mathematical Monthly, volume 62 number 4 (Apr. 1955), p. 249.

[Polya] G. Polya, *Patterns of Plausible Inference*,

[Programmer's Ref.] *Microsoft Programmers Reference* Microsoft Press.

[Putnam, 1988, A-6] William Lowell Putnam Mathematical Competition, Problem A-6, 1988, solution: American Mathematical Monthly, volume 96 number 8 (Oct. 1989), p. 688.

[Putnam, 1990, A-5] William Lowell Putnam Mathematical Competition, Problem A-5, 1990.

[Rice] John R. Rice, *Numerical Mathods, Software, and Analysis*, second edition, Academic Press, 1993.

[Quine] W. V. Quine, *Methods of Logic*, fourth edition, Harvard University Press, 1982.

[Rucker] Rudy Rucker, *Infinity and the Mind*, Birkhauser, 1982.

[Ryan] Patrick J. Ryan, *Euclidean and Non-Euclidean Geometry: an Analytic Approach*, Cambridge University Press, 1986.

[Seidenberg] A. Seidenberg, *Lectures in Projective Geometry*, Van Nostrand, 1962.

[Smith] Larry Smith, *Linear Algebra*, second edition, Springer-Verlag, 1984.

[Strang 93] Gilbert Strang *The Fundamental Theorem of Linear Algebra*, American Mathematical Monthly, Nov. 1993, p. 848 – 855.

[Strang 80] Gilbert Strang, *Linear Algebra and its Applications*, second edition, Harcourt Brace Jovanovich, 1980.

[Taylor] Alan D. Taylor, *Mathematics and Politics: Strategy, Voting, Power, and Proof*, Springer-Verlag, 1995.

[Tilley] Burt Tilley, private communication, 1996.

[Trono] Tony Trono, compiler, *University of Vermont Mathematics Department High School Prize Examinations 1958-1991*, mimeographed printing, 1991.

[USSR Olympiad no. 174] *The USSR Mathematics Olympiad*, number 174.

[Weston] J. D. Weston, *Volume in Vector Spaces*, American Mathematical Monthly, volume 66 number 7 (Aug./Sept. 1959), p. 575 – 577.

[Weyl] Hermann Weyl, *Symmetry*, Princeton University Press, 1952.

[Wickens] Thomas D. Wickens, *Models for Behavior*, W.H. Freeman, 1982.

[Wilkinson 1965] Wilkinson, 1965

[Wohascum no. 2] *The Wohascum County Problem Book* problem number 2.

[Wohascum no. 47] *The Wohascum County Problem Book* problem number 47.

[Yaglom] I.M. Yaglom, *Felix Klein and Sophus Lie: Evolution of the Idea of Symmetry in the Nineteenth Century*, translated by Sergei Sossinsky, Birkhäuser, 1988.

[Zwicker] William S. Zwicker, *The Voters' Paradox, Spin, and the Borda Count*, Mathematical Social Sciences, vol. 22 (1991), p. 187 – 227.