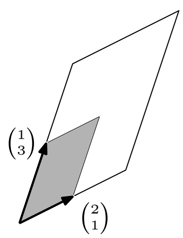
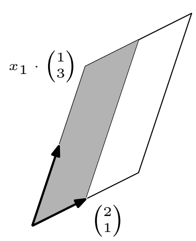
Linear Algebra

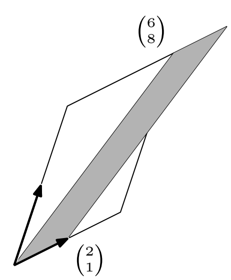
Jim Hefferon



$$\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}$$



$$\begin{vmatrix} x \cdot 1 & 2 \\ x \cdot 3 & 1 \end{vmatrix}$$



$$\begin{vmatrix} 6 & 2 \\ 8 & 1 \end{vmatrix}$$

Notation

$\mathbb{R}$	real numbers
$\mathbb{N}$	natural numbers: $\{0, 1, 2, \ldots\}$
$\mathbb{C}$	complex numbers
$\{\ldots \mid \ldots\}$	set of ... such that ...
$\langle \ldots \rangle$	sequence; like a set but order matters
V, W, U	vector spaces
$\vec{v}, \vec{w}$	vectors
$\vec{0}$, $\vec{0}_V$	zero vector, zero vector of V
B, D	bases
$\mathcal{E}_n = \langle \vec{e}_1, \ldots, \vec{e}_n \rangle$	standard basis for $\mathbb{R}^n$
$\vec{\beta}, \vec{\delta}$	basis vectors
$\mathrm{Rep}_B(\vec{v})$	matrix representing the vector
$\mathcal{P}_n$	set of n-th degree polynomials
$\mathcal{M}_{n\times m}$	set of $n\times m$ matrices
$[S]$	span of the set S
$M \oplus N$	direct sum of subspaces
$V \cong W$	isomorphic spaces
h, g	homomorphisms, linear maps
H, G	matrices
t, s	transformations; maps from a space to itself
T, S	square matrices
$\mathrm{Rep}_{B,D}(h)$	matrix representing the map h
$h_{i,j}$	matrix entry from row i, column j
$\lvert T \rvert$	determinant of the matrix T
$\mathscr{R}(h), \mathscr{N}(h)$	rangespace and nullspace of the map h
$\mathscr{R}_\infty(h), \mathscr{N}_\infty(h)$	generalized rangespace and nullspace

Lower case Greek alphabet

name	character	name	character	name	character
alpha	α	iota	ι	rho	ρ
beta	β	kappa	κ	sigma	σ
gamma	γ	lambda	λ	tau	τ
delta	δ	mu	μ	upsilon	υ
epsilon	ϵ	nu	ν	phi	ϕ
zeta	ζ	xi	ξ	chi	χ
eta	η	omicron	o	psi	ψ
theta	θ	pi	π	omega	ω

Cover. This is Cramer's Rule for the system $x + 2y = 6$, $3x + y = 8$. The size of the first box is the determinant shown (the absolute value of the size is the area). The size of the second box is x times that, and equals the size of the final box. Hence, x is the final determinant divided by the first determinant.

Contents

* *Note:* starred subsections are optional.

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 '$\cdot$' such that

(1) if $\vec{v}, \vec{w} \in V$ then their *vector sum* $\vec{v} + \vec{w}$ is in V and

- $\vec{v} + \vec{w} = \vec{w} + \vec{v}$
- $(\vec{v} + \vec{w}) + \vec{u} = \vec{v} + (\vec{w} + \vec{u})$ (where $\vec{u} \in V$)
- there is a *zero vector* $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$ for all $\vec{v} \in V$
- each $\vec{v} \in V$ has an *additive inverse* $\vec{w} \in V$ such that $\vec{w} + \vec{v} = \vec{0}$

(2) if r, s are *scalars* (members of $\mathbb{R}$) and $\vec{v}, \vec{w} \in V$ then each *scalar multiple* $r \cdot \vec{v}$ is in V and

- $(r + s) \cdot \vec{v} = r \cdot \vec{v} + s \cdot \vec{v}$
- $r \cdot (\vec{v} + \vec{w}) = r \cdot \vec{v} + r \cdot \vec{w}$
- $(rs) \cdot \vec{v} = r \cdot (s \cdot \vec{v})$
- $1 \cdot \vec{v} = \vec{v}$.

1.2 Remark Because it involves two kinds of addition and two kinds of multiplication, that definition may seem confused. For instance, in '$(r + s) \cdot \vec{v} = r \cdot \vec{v} + s \cdot \vec{v}$', 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 of the conditions in the definition. That check is written out in the first example. Use it as a model for the others. Especially important are the two: '$\vec{v} + \vec{w}$ is in V' and '$r \cdot \vec{v}$ is in V'. These are the *closure* conditions. They specify that the addition and scalar multiplication operations are always sensible — they must be defined for every pair of vectors, and every scalar and vector, and the result of the operation must be 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 '$\cdot$' 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 in the definition.

There are five conditions in item (1). First, for 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 $\mathbb{R}^2$. Second, to show 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). The third condition, 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 we must produce a zero element — the vector of zeroes is it.

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

Fifth, 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 in item (2) are just as routine. First, for 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 $\mathbb{R}^2$. This checks the second condition.

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

For the third condition, that scalar multiplication distributes from the left over vector addition, the check is also straightforward.

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

The fourth

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

and fifth conditions are also easy.

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

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 '(π)'. Instead we just write 'π'.)

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 '·' 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

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

(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. The checks for 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 requirement (2) in the definition. 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

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

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 $\{f \mid f\colon \mathbb{N} \to \mathbb{R}\}$ 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 by thinking of these functions as "the same" as infinitely-tall vectors:

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

with addition and scalar multiplication are component-wise, as before. (The "infinitely-tall" vector can be formalized as an infinite sequence, or just as a function from $\mathbb{N}$ to $\mathbb{R}$, in which case the above correspondence is an equality.)

1.10 Example The set of polynomials with real coefficients

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

makes a vector space when given the natural '+'

$$\begin{aligned}(a_0 + a_1x + \cdots + a_nx^n) &+ (b_0 + b_1x + \cdots + b_nx^n) \\ &= (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n\end{aligned}$$

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 1.9. 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.

1.11 Example The set $\{f \mid f\colon \mathbb{R} \to \mathbb{R}\}$ 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 1.9 is the domain of the functions.

1.12 Example The set $F = \{a\cos\theta + b\sin\theta \mid a, b \in \mathbb{R}\}$ of real-valued functions of the real variable θ 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.)

1.13 Example The set

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

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:

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

and

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

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.

1.14 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

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

both satisfy the condition that their entries add to zero then $\vec{v} + \vec{w}$ also satisfies that condition: $c_1(v_1 + w_1) + \cdots + c_n(v_n + w_n) = (c_1 v_1 + \cdots + c_n v_n) + (c_1 w_1 + \cdots + c_n w_n) = 0$. 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 '$c_1 v_1$' from that in '$r\vec{v}$' 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.

1.15 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 $1 \cdot \vec{v} = \vec{v}$ but not a condition that $0 \cdot \vec{v} = \vec{0}$?

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, and so that it does not contain extra conditions that only bar as examples spaces where those properties occur. 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.

1.16 Lemma In any vector space V,

(1) $0 \cdot \vec{v} = \vec{0}$

(2) $(-1 \cdot \vec{v}) + \vec{v} = \vec{0}$

(3) $r \cdot \vec{0} = \vec{0}$

for any $\vec{v} \in V$ and $r \in \mathbb{R}$.

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

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

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

For the third one, this $r \cdot \vec{0} = r \cdot (0 \cdot \vec{0}) = (r \cdot 0) \cdot \vec{0} = \vec{0}$ will do. QED

We finish this subsection with a recap, and a comment.

Chapter One studied Gaussian reduction. That led us to study collections of linear combinations. We have named any such structure a 'vector space'. In a phrase, the point of this material is that 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

1.17 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 2×4 matrices

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

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

✓ **1.18** 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 of 2×2 matrices with real number entries under the usual matrix addition and scalar multiplication,

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

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

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

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

(b) The set of 2×2 matrices with real entries under the usual matrix operations.

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

(d) The set

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

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

✓ **1.20** 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

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

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

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

(c) Under the usual matrix operations,

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

(d) Under the usual polynomial operations,

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

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

(e) Under the inherited operations,

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

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

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

1.23 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.

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

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

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

✓ **1.26** 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 $\{f\colon \mathbb{R} \to \mathbb{R} \mid df/dx + 2f = 1\}$

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

✓ **1.28** 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.

1.29 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 $r(x, y) = (rx, y)$

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

1.30 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.

1.31 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

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

(d) This set of 2×2 matrices

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

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

✓ **1.33** Prove these.

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

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

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

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

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

1.37 **(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?

✓ **1.38** 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 $\vec{v} \in V$ is not $\vec{0}$.

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

(b) Prove that $r_1 \cdot \vec{v} = r_2 \cdot \vec{v}$ 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.

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

1.40 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

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

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

✓ **1.42** **(a)** Prove that a sum of four vectors $\vec{v}_1, \ldots, \vec{v}_4 \in V$ can be associated in any way without changing the result.

$$\begin{aligned}((\vec{v}_1 + \vec{v}_2) + \vec{v}_3) + \vec{v}_4 &= (\vec{v}_1 + (\vec{v}_2 + \vec{v}_3)) + \vec{v}_4 \\ &= (\vec{v}_1 + \vec{v}_2) + (\vec{v}_3 + \vec{v}_4) \\ &= \vec{v}_1 + ((\vec{v}_2 + \vec{v}_3) + \vec{v}_4) \\ &= \vec{v}_1 + (\vec{v}_2 + (\vec{v}_3 + \vec{v}_4))\end{aligned}$$

This allows us to simply write '$\vec{v}_1 + \vec{v}_2 + \vec{v}_3 + \vec{v}_4$' 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.)

1.43 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 $\{a_0 + a_1x + a_2x^2 \bigm| a_0 + a_1 + a_2 = 0\}$ is a subspace of the vector space of degree two polynomials.

(b) Show that this is a subspace of the 2×2 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 $\vec{s}_1, \vec{s}_2 \in S$ then the combination $c_1\vec{v}_1 + c_2\vec{v}_2$ 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_3 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.

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

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 $\vec{s}_1, \vec{s}_2 \in S$ then $\vec{s}_1 + \vec{s}_2 \in S$, as $\vec{s}_1 + \vec{s}_2 = 1 \cdot \vec{s}_1 + 1 \cdot \vec{s}_2$. Second, for any $\vec{s}_1, \vec{s}_2 \in S$, because addition is inherited from V, the sum $\vec{s}_1 + \vec{s}_2$ in S equals the sum $\vec{s}_1 + \vec{s}_2$ in V, and that equals the sum $\vec{s}_2 + \vec{s}_1$ in V (because V is a vector space, its addition is commutative), and that in turn equals the sum $\vec{s}_2 + \vec{s}_1$ 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 $\vec{s}$ is any member of the nonempty set S) $0 \cdot \vec{s} + 0 \cdot \vec{s} = \vec{0}$ is in S; showing that $\vec{0}$ acts under the inherited operations as the additive identity of S is easy. The fifth condition is satisfied because for any $\vec{s} \in S$, closure under linear combinations shows that the vector $0 \cdot \vec{0} + (-1) \cdot \vec{s}$ is in S; showing that it is the additive inverse of $\vec{s}$ 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.

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

The vector space definition has ten conditions but eight of them, the ones stated there with the '•' bullets, simply ensure that referring to the operations as an 'addition' and a 'scalar multiplication' is sensible. The proof above checks that if the nonempty set S satisfies statement (2) then inheritance of the operations from the surrounding vector space brings with it the inheritance of these eight properties also (i.e., 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. 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 $r_1\vec{s}_1 + r_2\vec{s}_2$, 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 paramatrize, 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 paramatrize, 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 2×2 matrices

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

(checking that it is nonempty and closed under linear combinations is easy). To paramatrize, 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).

Paramatrization 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.

$$[S] = \{c_1\vec{s}_1 + \cdots + c_n\vec{s}_n \bigm| c_1, \ldots, c_n \in \mathbb{R} \text{ and } \vec{s}_1, \ldots, \vec{s}_n \in S\}$$

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 written as $\{c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \bigm| c_1, \ldots, c_k \in \mathbb{R}\}$, we described that as the set 'generated' by the $\vec{\beta}$'s. We now have the technical term; we call that the 'span' of the set $\{\vec{\beta}_1, \ldots, \vec{\beta}_k\}$.

Recall also the discussion of the "tricky point" in that proof. The span of the empty set is defined to be the set $\{\vec{0}\}$ because we follow the convention that a linear combination of no vectors sums to $\vec{0}$. 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 $[S]$ is closed under linear combinations. For a pair of vectors from that span, $\vec{v} = c_1\vec{s}_1 + \cdots + c_n\vec{s}_n$ and $\vec{w} = c_{n+1}\vec{s}_{n+1} + \cdots + c_m\vec{s}_m$, a linear combination

$$\begin{aligned} &p \cdot (c_1\vec{s}_1 + \cdots + c_n\vec{s}_n) + r \cdot (c_{n+1}\vec{s}_{n+1} + \cdots + c_m\vec{s}_m) \\ &\qquad = pc_1\vec{s}_1 + \cdots + pc_n\vec{s}_n + rc_{n+1}\vec{s}_{n+1} + \cdots + rc_m\vec{s}_m \end{aligned}$$

(p, r scalars) is a linear combination of elements of S and so is in $[S]$ (possibly some of the $\vec{s}_i$'s forming $\vec{v}$ equal some of the $\vec{s}_j$'s from $\vec{w}$, 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 $\vec{v}$, the set $\{r \cdot \vec{v} \mid r \in \mathbb{R}\}$ is a subspace of V. For instance, for any vector $\vec{v} \in \mathbb{R}^3$, the line through the origin containing that vector, $\{k\vec{v} \mid k \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$. This is true even when $\vec{v}$ 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 $\mathbb{R}^2$.

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

Tocheck 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

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

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 paramatrize 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 subspace 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?

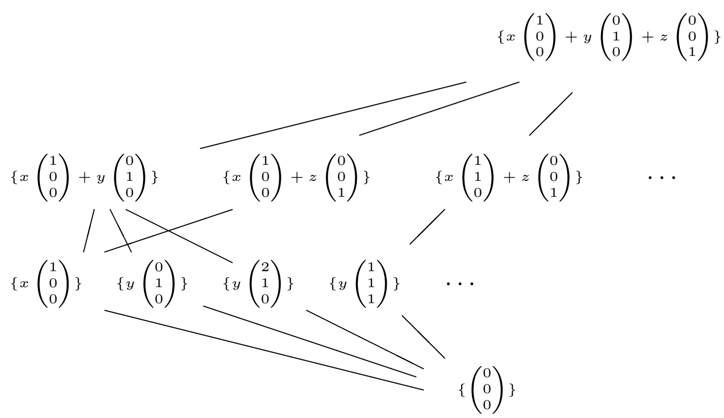
We are asking: for which members $a_2x^2 + a_1x + a_0$ of $\mathcal{P}_2$ are there c_1 and c_2 such that $a_2x^2 + a_1x + a_0 = c_1(3x - x^2) + c_2(2x)$? Since polynomials are equal if and only if their coefficients are equal, we are looking for conditions on a_2, a_1, and a_0 satisfying these.

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

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, paramatrize 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 so, paramatrize 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 $\mathcal{M}_{2\times2}$

2.23 Which of these are members of the span $[\{\cos^2 x, \sin^2 x\}]$ 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** Paramatrize each subspace's description. Then express each subspace as a span.

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

(b) This subset of $\mathcal{M}_{2\times2}$

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

(c) This subset of $\mathcal{M}_{2\times2}$

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

(d) The subset $\{a+bx+cx^3 \bigm| a-2b+c=0\}$ 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.* Paramatrize each.)

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

(b) $\{\begin{pmatrix}x\\y\\z\end{pmatrix} \bigm| 3x+2y+z=0\}$ 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) $\{a_0 + a_1 x + a_2 x^2 + a_3 x^3 \mid a_0 + a_1 = 0 \text{ and } a_2 - a_3 = 0\}$ 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 $\vec{v}$ that is an element of a vector space V, the set $\{r \cdot \vec{v} \mid r \in \mathbb{R}\}$ is a subspace of V. (This is of course, simply the span of the singleton set $\{\vec{v}\}$.) 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 $[[S]] = [S]$. (*Hint.* Members of $[S]$ are linear combinations of members of S. Members of $[[S]]$ 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 ong 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 thru 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 this 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,

$$[S]=[S \cup\{\vec{v}\}] \quad \text{if and only if} \quad \vec{v} \in[S]$$

for any $\vec{v} \in V$.

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

For the right to left implication assume that $\vec{v} \in[S]$ to show that $[S]=[S \cup \{\vec{v}\}]$ by mutual inclusion. The inclusion $[S] \subseteq[S \cup\{\vec{v}\}]$ is obvious. For the other inclusion $[S] \supseteq[S \cup\{\vec{v}\}]$, write an element of $[S \cup\{\vec{v}\}]$ as $d_0 \vec{v}+d_1 \vec{s}_1+\cdots+d_m \vec{s}_m$ and substitute $\vec{v}$'s expansion as a linear combination of members of the same set $d_0(c_0 \vec{t}_0+\cdots+c_k \vec{t}_k)+d_1 \vec{s}_1+\cdots+d_m \vec{s}_m$. 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 $[S \cup\{\vec{v}\}]$ is also a member of $[S]$. QED

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

$$\vec{v}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix}2\\1\\0\end{pmatrix}$$

the spans $[\{\vec{v}_1, \vec{v}_2\}]$ and $[\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}]$ are equal since $\vec{v}_3$ is in the span $[\{\vec{v}_1, \vec{v}_2\}]$.

The lemma says that if we have a spanning set then we can remove a $\vec{v}$ to get a new set S with the same span if and only if $\vec{v}$ 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

$$\vec{s}_0 = c_1\vec{s}_1 + c_2\vec{s}_2 + \cdots + c_n\vec{s}_n$$

visually sets $\vec{s}_0$ off from the other vectors, algebraically there is nothing special in that equation about $\vec{s}_0$. For any $\vec{s}_i$ with a coefficient c_i that is nonzero, we can rewrite the relationship to set off $\vec{s}_i$.

$$\vec{s}_i = (1/c_i)\vec{s}_0 + (-c_1/c_i)\vec{s}_1 + \cdots + (-c_n/c_i)\vec{s}_n$$

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 $\vec{s}_0, \vec{s}_1, \ldots, \vec{s}_n$ 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 $\vec{s}_1, \ldots, \vec{s}_n \in S$ the only linear relationship among those vectors

$$c_1\vec{s}_1 + \cdots + c_n\vec{s}_n = \vec{0} \qquad c_1, \ldots, c_n \in \mathbb{R}$$

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 $\vec{s}_i$ can be written as a linear combination of the other vectors from S so there is no linear relationship where some of the $\vec{s}$'s have nonzero coefficients. If S is not linearly independent then some $\vec{s}_i$ is a linear combination $\vec{s}_i = c_1\vec{s}_1 + \cdots + c_{i-1}\vec{s}_{i-1} + c_{i+1}\vec{s}_{i+1} + \cdots + c_n\vec{s}_n$ of other vectors from S, and subtracting $\vec{s}_i$ from both sides of that equation gives a linear relationship involving a nonzero coefficient, namely the -1 in front of $\vec{s}_i$. 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

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

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

$$\vec{v}_1 = \begin{pmatrix}3\\4\\5\end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix}2\\9\\2\end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix}4\\18\\4\end{pmatrix}$$

the set $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is linearly dependent because this is a relationship

$$0 \cdot \vec{v}_1 + 2 \cdot \vec{v}_2 - 1 \cdot \vec{v}_3 = \vec{0}$$

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 $\vec{v}_1 = c_2\vec{v}_2 + c_3\vec{v}_3$ 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 $\vec{v}_2 = c_1\vec{v}_1 + c_3\vec{v}_3$ 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 $0\cdot\vec{v}_1 + 0\cdot\vec{v}_2 + 1\cdot\vec{0} = \vec{0}$, 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, $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ 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 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 $S = \{\vec{s}_1, \ldots, \vec{s}_n\}$ 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 $\vec{s}_i$ that is a linear combination of the others. Discard that vector — define the set $S_1 = S - \{\vec{s}_i\}$. By Lemma 1.1, the span does not shrink $[S_1] = [S]$.

Now, if S_1 is linearly independent then we are finished. Otherwise iterate the prior paragraph: take another vector, $\vec{v}_2$, this time one that is a linear combination of other members of S_1, and discard it to derive $S_2 = S_1 - \{\vec{v}_2\}$ such that $[S_2] = [S_1]$. 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 paramatrized 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 $[S_1] = [S]$. Now, the third vector of K_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 $[S]$ 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 $[S]$ 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 $[S] = \mathbb{R}^3$. 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,

$$S \cup \{\vec{v}\} \text{ is linearly dependent} \quad \text{if and only if} \quad \vec{v} \in [S]$$

for any $\vec{v} \in V$ with $\vec{v} \notin S$.

Proof. One implication is clear: if $\vec{v} \in [S]$ then $\vec{v} = c_1\vec{s}_1 + c_2\vec{s}_2 + \cdots + c_n\vec{s}_n$ where each $\vec{s}_i \in S$ and $c_i \in \mathbb{R}$, and so $\vec{0} = c_1\vec{s}_1 + c_2\vec{s}_2 + \cdots + c_n\vec{s}_n + (-1)\vec{v}$ is a nontrivial linear relationship among elements of $S \cup \{\vec{v}\}$.

The other implication requires the assumption that S is linearly independent. With $S \cup \{\vec{v}\}$ linearly dependent, there is a nontrivial linear relationship $c_0\vec{v} + c_1\vec{s}_1 + c_2\vec{s}_2 + \cdots + c_n\vec{s}_n = \vec{0}$ 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 $\vec{v} = -(c_1/c_0)\vec{s}_1 - \cdots - (c_n/c_0)\vec{s}_n$ shows that $\vec{v} \in [S]$. QED

(Compare this result with Lemma 1.1. Both say, roughly, that $\vec{v}$ 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 $S = \{\vec{s}_1, \ldots, \vec{s}_n\}$ of a vector space is linearly dependent if and only if some $\vec{s}_i$ is a linear combination of the vectors $\vec{s}_1$, …, $\vec{s}_{i-1}$ listed before it.

Proof. Consider $S_0 = \{\}$, $S_1 = \{\vec{s_1}\}$, $S_2 = \{\vec{s}_1, \vec{s}_2\}$, etc. Some index $i \geq 1$ is the first one with $S_{i-1} \cup \{\vec{s}_i\}$ linearly dependent, and there $\vec{s}_i \in [S_{i-1}]$. QED

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

Example 1.15 shows that some linearly independent sets are maximal — have as many elements as possible — in that they have no supesets 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 $(-\pi/2..\pi/2)$?

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 $\{\vec{u},\vec{v},\vec{w}\}$ linearly independent set then so is the set $\{\vec{u},\vec{u}+\vec{v},\vec{u}+\vec{v}+\vec{w}\}$.
(b) What is the relationship between the linear independence or dependence of the set $\{\vec{u},\vec{v},\vec{w}\}$ and the independence or dependence of $\{\vec{u}-\vec{v},\vec{v}-\vec{w},\vec{w}-\vec{u}\}$?

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 $\{\vec{x},\vec{y},\vec{z}\}$ is linearly independent then so are all of its proper subsets: $\{\vec{x},\vec{y}\}$, $\{\vec{x},\vec{z}\}$, $\{\vec{y},\vec{z}\}$, $\{\vec{x}\}$,$\{\vec{y}\}$, $\{\vec{z}\}$, 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 $\vec{v}$ is in $[S]$, so that $\vec{v}$ 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 $\vec{v}$ is unique (that is, unique up to reordering and adding or taking away terms of the form $0\cdot\vec{s}$). Thus S as a spanning set is minimal in this strong sense: each vector in $[S]$ 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: $x\mapsto c_nx^n+\cdots+c_1x+c_0$.)

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

$$c_1\begin{pmatrix}2\\4\end{pmatrix} + c_2\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \implies \begin{array}{l}2c_1 + 1c_2 = 0\\4c_1 + 1c_2 = 0\end{array} \implies c_1 = c_2 = 0$$

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

$$\begin{array}{l}2c_1 + 1c_2 = x\\4c_1 + 1c_2 = y\end{array} \implies c_2 = 2x - y \text{ and } c_1 = (y - x)/2$$

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.

* More information on sequences is in the appendix.

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 $\mathbb{R}^2$.

1.6 Example Consider the space $\{a \cdot \cos\theta + b \cdot \sin\theta \mid a, b \in \mathbb{R}\}$ of function of the real variable θ. 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 paramatrizing.

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

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 $\mathcal{M}_{2\times 2}$

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

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 '$0\cdot\vec{\beta}$'.

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 $0\vec{\beta}_i$ terms, so that the two sums combine the same $\vec{\beta}$'s in the same order: $\vec{v} = c_1\vec{\beta}_1 + c_2\vec{\beta}_2 + \cdots + c_n\vec{\beta}_n$ and $\vec{v} = d_1\vec{\beta}_1 + d_2\vec{\beta}_2 + \cdots + d_n\vec{\beta}_n$. Now

$$c_1\vec{\beta}_1 + c_2\vec{\beta}_2 + \cdots + c_n\vec{\beta}_n = d_1\vec{\beta}_1 + d_2\vec{\beta}_2 + \cdots + d_n\vec{\beta}_n$$

holds if and only if

$$(c_1 - d_1)\vec{\beta}_1 + \cdots + (c_n - d_n)\vec{\beta}_n = \vec{0}$$

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 $\vec{v}$ with respect to B* is the column vector of the coefficients used to express $\vec{v}$ as a linear combination of the basis vectors:

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

where $B = \langle\vec{\beta}_1, \dots, \vec{\beta}_n\rangle$ and $\vec{v} = c_1\vec{\beta}_1 + c_2\vec{\beta}_2 + \cdots + c_n\vec{\beta}_n$. The c's are the *coordinates of $\vec{v}$ with respect to B*.

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 $\vec{v} = \binom{3}{2}$' 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.

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

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.

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

✓ **1.20** Find a basis for $\mathcal{M}_{2\times 2}$, the space of 2×2 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 2×2 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 $\langle\vec{\beta}_1, \vec{\beta}_2, \vec{\beta}_3\rangle$ be a basis for a vector space.
(a) Show that $\langle c_1\vec{\beta}_1, c_2\vec{\beta}_2, c_3\vec{\beta}_3\rangle$ 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 $\langle\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3\rangle$ is a basis where $\vec{\alpha}_i = \vec{\beta}_1 + \vec{\beta}_i$.

1.28 Find one vector $\vec{v}$ that will make each into a basis for the space.

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

✓ **1.29** Where $\langle\vec{\beta}_1, \dots, \vec{\beta}_n\rangle$ is a basis, show that in this equation

$$c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k = c_{k+1}\vec{\beta}_{k+1} + \cdots + c_n\vec{\beta}_n$$

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 2×2 matrices.
(b) Find a basis for the space of symmetric 3×3 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

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

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 $\{a_2x^2+a_1x+a_0 \mid 2a_2-a_0=a_1\}$, 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 $B = \langle\vec{\beta}_1, \ldots, \vec{\beta}_n\rangle$ is a basis for a vector space, and that for the vector $\vec{v}$ the relationship $\vec{v} = c_1\vec{\beta}_1 + c_2\vec{\beta}_2 + \cdots + c_n\vec{\beta}_n$ has $c_i \neq 0$. Then exchanging $\vec{\beta}_i$ for $\vec{v}$ yields another basis for the space.

Proof. Call the outcome of the exchange $\hat{B} = \langle\vec{\beta}_1, \ldots, \vec{\beta}_{i-1}, \vec{v}, \vec{\beta}_{i+1}, \ldots, \vec{\beta}_n\rangle$.

We first show that $\hat{B}$ is linearly independent. Any relationship $d_1\vec{\beta}_1 + \cdots + d_i\vec{v} + \cdots + d_n\vec{\beta}_n = \vec{0}$ among the members of $\hat{B}$, after substitution for $\vec{v}$,

$$d_1\vec{\beta}_1 + \cdots + d_i \cdot (c_1\vec{\beta}_1 + \cdots + c_i\vec{\beta}_i + \cdots + c_n\vec{\beta}_n) + \cdots + d_n\vec{\beta}_n = \vec{0} \qquad (*)$$

gives a linear relationship among the members of B. The basis B is linearly independent, so the coefficient d_ic_i of $\vec{\beta}_i$ 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 $[\hat{B}] \subseteq [B]$, is easy; any member $d_1\vec{\beta}_1 + \cdots + d_i\vec{v} + \cdots + d_n\vec{\beta}_n$ of $[\hat{B}]$ can be written $d_1\vec{\beta}_1 + \cdots + d_i \cdot (c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n) + \cdots + d_n\vec{\beta}_n$, which is a linear combination of linear combinations of members of B, and hence is in $[B]$. For the $[B] \subseteq [\hat{B}]$ half of the argument, recall that when $\vec{v} = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n$ with $c_i \neq 0$, then the equation can be rearranged to $\vec{\beta}_i = (-c_1/c_i)\vec{\beta}_1 + \cdots + (-1/c_i)\vec{v} + \cdots + (-c_n/c_i)\vec{\beta}_n$. Now, consider any member $d_1\vec{\beta}_1 + \cdots + d_i\vec{\beta}_i + \cdots + d_n\vec{\beta}_n$ of $[B]$, substitute for $\vec{\beta}_i$ 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 $[\hat{B}]$. 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 $B = \langle\vec{\beta}_1, \ldots, \vec{\beta}_n\rangle$ of minimal size. We will show that any other basis $D = \langle\vec{\delta}_1, \vec{\delta}_2, \ldots\rangle$ 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 $\vec{\delta}_1$ is in the space, so $\vec{\delta}_1$ is a nontrivial linear combination of elements of B. By the Exchange Lemma, $\vec{\delta}_1$ can be swapped for a vector from B, resulting in a basis B_1, where one element is $\vec{\delta}$ and all of the $n-1$ other elements are $\vec{\beta}$'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 $\vec{\delta}_{k+1}$. 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 $\vec{\beta}_i$ or else that representation would be a nontrivial linear relationship among elements of the linearly independent set D. Exchange $\vec{\delta}_{k+1}$ for $\vec{\beta}_i$ to get a new basis B_{k+1} with one $\vec{\delta}$ more and one $\vec{\beta}$ fewer than the previous basis B_k.

Repeat the inductive step until no $\vec{\beta}$'s remain, so that B_n contains $\vec{\delta}_1, \ldots, \vec{\delta}_n$. Now, D cannot have more than these n vectors because any $\vec{\delta}_{n+1}$ 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 $\vec{\delta}$'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 $S = \{\vec{0}\}$ then it can be shrunk to the empty basis, thereby making it linearly independent, without changing its span.

Otherwise, S contains a vector $\vec{s}_1$ with $\vec{s}_1 \neq \vec{0}$ and we can form a basis $B_1 = \langle \vec{s}_1 \rangle$. If $[B_1] = [S]$ then we are done.

If not then there is a $\vec{s}_2 \in [S]$ such that $\vec{s}_2 \notin [B_1]$. Let $B_2 = \langle \vec{s}_1, \vec{s}_2 \rangle$; if $[B_2] = [S]$ 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 that this expansion is actually the set 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, P_2.

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

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

✓ **2.16** Find a basis for, and the dimension of, $\mathcal{M}_{2\times2}$, the vector space of 2×2 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 3×5 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** Observe that, where S is a set, the functions $f\colon S \to \mathbb{R}$ form a vector space under the natural operations: $f + g\,(s) = f(s) + g(s)$ and $r \cdot f\,(s) = r \cdot f(s)$. 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 $\langle \vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3 \rangle \subset \mathbb{R}^3$ 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 σ of the numbers 1, 2, …, n (that is, σ 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, Rowspace(B) $\subseteq$ Rowspace(A) because a member of the set 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 Rowspace(A).

For the other containment, recall that row operations are reversible: $A \longrightarrow B$ if and only if $B \longrightarrow A$. With that, Rowspace(A) $\subseteq$ 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.2.5, 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^{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, $\langle \vec{e}_1, \vec{e}_2 \rangle$. (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 $\vec{e}_i$'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

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

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 5×9, 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

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

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 $W_1+W_2+\cdots+W_k=[W_1\cup W_2\cup\ldots W_k]$.

(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 $\vec{w}\in\mathbb{R}^3$ can be written as a linear combination $c_1\vec{v}_1+c_2\vec{v}_2+c_3\vec{v}_3$ where $\vec{v}_1$ 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 $\{a+bx\bigm| a,b\in\mathbb{R}\}$ and let C be the subspace of purely-cubic polynomials $\{cx^3\bigm| c\in\mathbb{R}\}$. Then $L+C$ is not all of $\mathcal{P}_4$. Instead, it is the subspace $L+C=\{a+bx+cx^3\bigm| a,b,c\in\mathbb{R}\}$.

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, we simply note that any $\vec{w}\in\mathbb{R}^3$ can be written

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

as a linear combination of a member of the xy-plane and a member of the yz-plane.

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 $\mathcal{E}_3=\langle\vec{e}_1,\vec{e}_2,\vec{e}_3\rangle$. The subspace with the basis $B_1=\langle\vec{e}_1\rangle$ is the x-axis. The subspace with the basis $B_2=\langle\vec{e}_2\rangle$ is the y-axis. The subspace with the basis $B_3=\langle\vec{e}_3\rangle$ 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 $B_1 = \langle \vec{e}_1, \vec{e}_3 \rangle$ is the xz-plane. The subspace with the basis $B_2 = \langle \vec{e}_2 \rangle$ 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 $B_1 = \langle \vec{\beta}_{1,1}, \ldots, \vec{\beta}_{1,n_1} \rangle$, …, $B_k = \langle \vec{\beta}_{k,1}, \ldots, \vec{\beta}_{k,n_k} \rangle$ is their adjoinment.

$$B_1 {}^\frown B_2 {}^\frown \cdots {}^\frown B_k = \langle \vec{\beta}_{1,1}, \ldots, \vec{\beta}_{1,n_1}, \vec{\beta}_{2,1}, \ldots, \vec{\beta}_{k,n_k} \rangle$$

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 $\vec{v} \in V$, the expression $\vec{v} = \vec{w}_1 + \cdots + \vec{w}_k$ (with $\vec{w}_i \in W_i$) is unique.
(2) The concatenation $B_1 {}^\frown \cdots {}^\frown B_k$ is a basis for V.
(3) The nonzero members of $\{\vec{w}_1, \ldots, \vec{w}_k\}$ (with $\vec{w}_i \in W_i$) 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 $\vec{w}$'s to a combination of $\vec{\beta}$'s

$$\begin{aligned} &d_1\vec{w}_1 + \cdots + d_k\vec{w}_k \\ &\quad = d_1(c_{1,1}\vec{\beta}_{1,1} + \cdots + c_{1,n_1}\vec{\beta}_{1,n_1}) + \cdots + d_k(c_{k,1}\vec{\beta}_{k,1} + \cdots + c_{k,n_k}\vec{\beta}_{k,n_k}) \\ &\quad = d_1c_{1,1} \cdot \vec{\beta}_{1,1} + \cdots + d_kc_{k,n_k} \cdot \vec{\beta}_{k,n_k} \end{aligned} \qquad (*)$$

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 $\vec{v}$

can be expressed as $\vec{v} = \vec{w}_1 + \cdots + \vec{w}_k$, which translates by equation $(*)$ to an expression of $\vec{v}$ as a linear combination of the $\vec{\beta}$'s from the concatenation. For linear independence, consider this linear relationship.

$$\vec{0} = c_{1,1}\vec{\beta}_{1,1} + \cdots + c_{k,n_k}\vec{\beta}_{k,n_k}$$

Regroup as in $(*)$ (that is, take $d_1, \ldots, d_k$ to be 1 and move from bottom to top) to get the decomposition $\vec{0} = \vec{w}_1 + \cdots + \vec{w}_k$. Because of the assumption that decompositions are unique, and because the zero vector obviously has the decomposition $\vec{0} = \vec{0} + \cdots + \vec{0}$, we now have that each $\vec{w}_i$ is the zero vector. This means that $c_{i,1}\vec{\beta}_{i,1} + \cdots + c_{i,n_i}\vec{\beta}_{i,n_i} = \vec{0}$. 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,

$$\vec{0} = \cdots + d_i\vec{w}_i + \cdots$$

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 $\vec{w}_1$ in this combination.) As in $(*)$, $\vec{0} = \cdots + d_i(c_{i,1}\vec{\beta}_{i,1} + \cdots + c_{i,n_i}\vec{\beta}_{i,n_i}) + \cdots = \cdots + d_ic_{i,1}\cdot\vec{\beta}_{i,1} + \cdots + d_ic_{i,n_i}\cdot\vec{\beta}_{i,n_i} + \cdots$ and the linear independence of $B_1{}^\frown \cdots {}^\frown B_k$ gives that each coefficient $d_ic_{i,j}$ is zero. Now, $\vec{w}_i$ is a nonzero vector, so at least one of the $c_{i,j}$'s is 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 $\vec{v} = \vec{w}_1 + \cdots + \vec{w}_k$ and $\vec{v} = \vec{u}_1 + \cdots + \vec{u}_k$ in order to show that the two are the same. We have

$$\vec{0} = (\vec{w}_1 + \cdots + \vec{w}_k) - (\vec{u}_1 + \cdots + \vec{u}_k) = (\vec{w}_1 - \vec{u}_1) + \cdots + (\vec{w}_k - \vec{u}_k)$$

which violates the assumption unless each $\vec{w}_i - \vec{u}_i$ 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 2×2 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 $\vec{v}$ be a vector from $W_1 \cap W_2$ and consider the equation $\vec{v} = \vec{v}$. 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 $\vec{v} = \vec{0}$, 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.

$$W_3 = \{\begin{pmatrix} q \\ q \\ r \end{pmatrix} \bigm| q, r \in \mathbb{R}\}$$

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) $W_1 = \{\begin{pmatrix} x \\ 0 \end{pmatrix} \bigm| x \in \mathbb{R}\}$, $W_2 = \{\begin{pmatrix} x \\ x \end{pmatrix} \bigm| x \in \mathbb{R}\}$

(b) $W_1 = \{\begin{pmatrix} s \\ s \end{pmatrix} \bigm| s \in \mathbb{R}\}$, $W_2 = \{\begin{pmatrix} s \\ 1.1s \end{pmatrix} \bigm| s \in \mathbb{R}\}$

(c) $W_1 = \mathbb{R}^2$, $W_2 = \{\vec{0}\}$

(d) $W_1 = W_2 = \{\begin{pmatrix} t \\ t \end{pmatrix} \bigm| t \in \mathbb{R}\}$

(e) $W_1 = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} x \\ 0 \end{pmatrix} \bigm| x \in \mathbb{R}\}$, $W_2 = \{\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix} \bigm| y \in \mathbb{R}\}$

✓ **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 $\mathcal{P}_n = \{a_0 \mid a_0 \in \mathbb{R}\} \oplus \ldots \oplus \{a_n x^n \mid a_n \in \mathbb{R}\}$.

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 $\langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ then it is the direct sum of the spans of the one-dimensional subspaces $V = [\{\vec{\beta}_1\}] \oplus \ldots \oplus [\{\vec{\beta}_n\}]$?

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,

$$W_1 + \cdots + W_k = \{\vec{w}_1 + \vec{w}_2 + \cdots + \vec{w}_k \mid \vec{w}_1 \in W_1, \ldots, \vec{w}_k \in W_k\},$$

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 pairwise 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 $\dim(W_1 + W_2)$.

$$W_1 = \{\begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix} \mid c, d \in \mathbb{R}\} \qquad W_2 = \{\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \mid b, c \in \mathbb{R}\}$$

(b) Suppose that U and W are subspaces of a vector space. Suppose that the sequence $\langle \vec{\beta}_1, \ldots, \vec{\beta}_k \rangle$ is a basis for $U \cap W$. Finally, suppose that the prior sequence has been expanded to give a sequence $\langle \vec{\mu}_1, \ldots, \vec{\mu}_j, \vec{\beta}_1, \ldots, \vec{\beta}_k \rangle$ that is a basis for U, and a sequence $\langle \vec{\beta}_1, \ldots, \vec{\beta}_k, \vec{\omega}_1, \ldots, \vec{\omega}_p \rangle$ that is a basis for W. Prove that this sequence

$$\langle \vec{\mu}_1, \ldots, \vec{\mu}_j, \vec{\beta}_1, \ldots, \vec{\beta}_k, \vec{\omega}_1, \ldots, \vec{\omega}_p \rangle$$

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 2×2 matrix and an antisymmetric 2×2 matrix. (*Remark.* For the second one, be careful about the entries on the diagional.)

(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

$$U^\perp = \{\vec{v} \in \mathbb{R}^n \bigm| \vec{v} \cdot \vec{u} = 0 \text{ for all } \vec{u} \in U\}$$

(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?

Topic: Fields

Linear combinations involving only fractions or only integers are much easier for computations than combinations involving real numbers, because computing with irrational numbers is awkward. Could other number systems, like the rationals or the integers, work in the place of $\mathbb{R}$ in the definition of a vector space?

Yes and no. If we take "work" to mean that the results of this chapter remain true then an analysis of which properties of the reals we have used in this chapter gives the following list of conditions an algebraic system needs in order to "work" in the place of $\mathbb{R}$.

Definition. A *field* is a set $\mathcal{F}$ with two operations '+' and '$\cdot$' such that

(1) for any $a, b \in \mathcal{F}$ the result of $a + b$ is in $\mathcal{F}$ and

- $a + b = b + a$
- if $c \in \mathcal{F}$ then $a + (b + c) = (a + b) + c$

(2) for any $a, b \in \mathcal{F}$ the result of $a \cdot b$ is in $\mathcal{F}$ and

- $a \cdot b = b \cdot a$
- if $c \in \mathcal{F}$ then $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

(3) if $a, b, c \in \mathcal{F}$ then $a \cdot (b + c) = a \cdot b + a \cdot c$

(4) there is an element $0 \in \mathcal{F}$ such that

- if $a \in \mathcal{F}$ then $a + 0 = a$
- for each $a \in \mathcal{F}$ there is an element $-a \in \mathcal{F}$ such that $(-a) + a = 0$

(5) there is an element $1 \in \mathcal{F}$ such that

- if $a \in \mathcal{F}$ then $a \cdot 1 = a$
- for each element $a \neq 0$ of $\mathcal{F}$ there is an element $a^{-1} \in \mathcal{F}$ such that $a^{-1} \cdot a = 1$.

The number system consisting of the set of real numbers along with the usual addition and multiplication operation is a field, naturally. Another field is the set of rational numbers with its usual addition and multiplication operations. An example of an algebraic structure that is not a field is the integer number system—it fails the final condition.

Some examples are surprising. The set $\{0, 1\}$ under these operations:

+	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

is a field (see Exercise 4).

We could develop Linear Algebra as the theory of vector spaces with scalars from an arbitrary field, instead of sticking to taking the scalars only from $\mathbb{R}$. In that case, almost all of the statements in this book would carry over by replacing '$\mathbb{R}$' with '$\mathcal{F}$', and thus by taking coefficients, vector entries, and matrix entries to be elements of $\mathcal{F}$ ("almost" because statements involving distances or angles are exceptions). Here are some examples; each applies to a vector space V over a field $\mathcal{F}$.

* For any $\vec{v} \in V$ and $a \in \mathcal{F}$, (i) $0 \cdot \vec{v} = \vec{0}$, and (ii) $-1 \cdot \vec{v} + \vec{v} = \vec{0}$, and (iii) $a \cdot \vec{0} = \vec{0}$.
* The span (the set of linear combinations) of a subset of V is a subspace of V.
* Any subset of a linearly independent set is also linearly independent.
* In a finite-dimensional vector space, any two bases have the same number of elements.

(Even statements that don't explicitly mention $\mathcal{F}$ use field properties in their proof.)

We won't develop vector spaces in this more general setting because the additional abstraction can be a distraction. The ideas we want to bring out already appear when we stick to the reals.

The only exception is in Chapter Five. In that chapter we must factor polynomials, so we will switch to considering vector spaces over the field of complex numbers. We will discuss this more, including a brief review of complex arithmetic, when we get there.

Exercises

1 Show that the real numbers form a field.

2 Prove that these are fields.
  (a) The rational numbers $\mathbb{Q}$ **(b)** The complex numbers $\mathbb{C}$

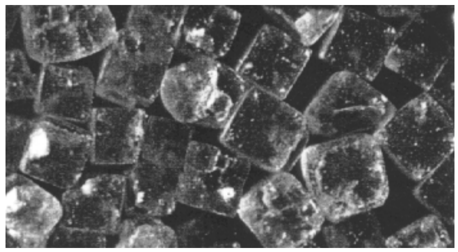
3 Give an example that shows that the integer number system is not a field.

4 Consider the set $\{0, 1\}$ subject to the operations given above. Show that it is a field.

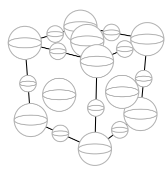
5 Give suitable operations to make the set $\{0, 1, 2\}$ a field.

Topic: Crystals

Everyone has noticed that table salt comes in little cubes.

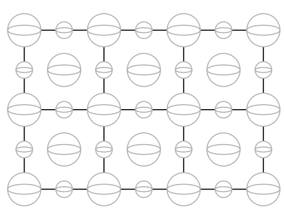


Remarkably, the explanation for the cubical external shape is the simplest one possible: the internal shape, the way the atoms lie, is also cubical. The internal structure is pictured below. Salt is sodium cloride, and the small spheres shown are sodium while the big ones are cloride. (To simplify the view, only the sodiums and clorides on the front, top, and right are shown.)



The specks of salt that we see when we spread a little out on the table consist of many repetitions of this fundamental unit. That is, these cubes of atoms stack up to make the larger cubical structure that we see. A solid, such as table salt, with a regular internal structure is a *crystal*.

We can restrict our attention to the front face. There, we have this pattern repeated many times.

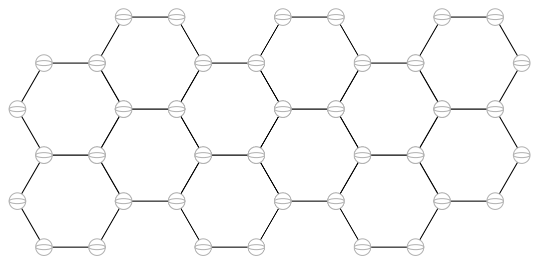


The distance between the corners of this cell is about 3.34 Ångstroms (an Ångstrom is 10^{-10} meters). Obviously that unit is unwieldly for describing points in the crystal lattice. Instead, the thing to do is to take as a unit the length of each side of the square. That is, we naturally adopt this basis.

$$\langle \begin{pmatrix} 3.34 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3.34 \end{pmatrix} \rangle$$

Then we can describe, say, the corner in the upper right of the picture above as $3\vec{\beta}_1 + 2\vec{\beta}_2$.

Another crystal from everyday experience is pencil lead. It is graphite, formed from carbon atoms arranged in this shape.



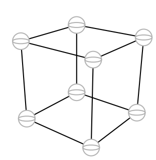
This is a single plane of graphite. A piece of graphite consists of many of these planes layered in a stack. (The chemical bonds between the planes are much weaker than the bonds inside the planes, which explains why graphite writes—it can be sheared so that the planes slide off and are left on the paper.) A convienent unit of length can be made by decomposing the hexagonal ring into three regions that are rotations of this *unit cell*.

A natural basis then would consist of the vectors that form the sides of that unit cell. The distance along the bottom and slant is 1.42 Ångstroms, so this

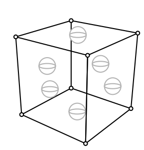
$$\langle \begin{pmatrix} 1.42 \\ 0 \end{pmatrix}, \begin{pmatrix} 1.23 \\ .71 \end{pmatrix} \rangle$$

is a good basis.

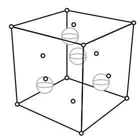
The selection of convienent bases extends to three dimensions. Another familiar crystal formed from carbon is diamond. Like table salt, it is built from cubes, but the structure inside each cube is more complicated than salt's. In addition to carbons at each corner,



there are carbons in the middle of each face.



(To show the added face carbons clearly, the corner carbons have been reduced to dots.) There are also four more carbons inside the cube, two that are a quarter of the way up from the bottom and two that are a quarter of the way down from the top.



(As before, carbons shown earlier have been reduced here to dots.) The distance along any edge of the cube is 2.18 Ångstroms. Thus, a natural basis for describing the locations of the carbons, and the bonds between them, is this.

$$\langle \begin{pmatrix} 2.18 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2.18 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2.18 \end{pmatrix} \rangle$$

Even the few examples given here show that the structures of crystals is complicated enough that some organized system to give the locations of the atoms, and how they are chemically bound, is needed. One tool for that organization is a convienent basis. This application of bases is simple, but it shows a context where the idea arises naturally. The work in this chapter just takes this simple idea and develops it.

Exercises

1 How many fundamental regions are there in one face of a speck of salt? (With a ruler, we can estimate that face is a square that is 0.1 cm on a side.)

2 In the graphite picture, imagine that we are interested in a point 5.67 Ångstroms up and 3.14 Ångstroms over from the origin.

(a) Express that point in terms of the basis given for graphite.

(b) How many hexagonal shapes away is this point from the origin?

(c) Express that point in terms of a second basis, where the first basis vector is the same, but the second is perpendicular to the first (going up the plane) and of the same length.

3 Give the locations of the atoms in the diamond cube both in terms of the basis, and in Ångstroms.

4 This illustrates how the dimensions of a unit cell could be computed from the shape in which a substance crystalizes ([Ebbing], p. 462).

(a) Recall that there are 6.022×10^{23} atoms in a mole (this is Avagadro's number). From that, and the fact that platinum has a mass of 195.08 grams per mole, calculate the mass of each atom.

(b) Platinum crystalizes in a face-centered cubic lattice with atoms at each lattice point, that is, it looks like the middle picture given above for the diamond crystal. Find the number of platinums per unit cell (hint: sum the fractions of platinums that are inside of a single cell).

(c) From that, find the mass of a unit cell.

(d) Platinum crystal has a density of 21.45 grams per cubic centimeter. From this, and the mass of a unit cell, calculate the volume of a unit cell.

(e) Find the length of each edge.
(f) Describe a natural three-dimensional basis.

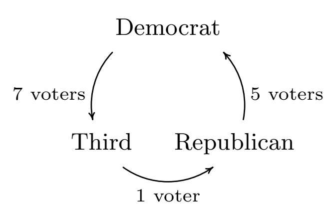
Topic: Voting Paradoxes

Imagine that a Political Science class studying the American presidential process holds a mock election. Members of the class rate, from most preferred to least preferred, the nominees from the Democratic Party, the Republican Party, and the Third Party ($>$ means 'is preferred to').

preference order	*number with that preference*
Democrat $>$ Republican $>$ Third	5
Democrat $>$ Third $>$ Republican	4
Republican $>$ Democrat $>$ Third	2
Republican $>$ Third $>$ Democrat	8
Third $>$ Democrat $>$ Republican	8
Third $>$ Republican $>$ Democrat	2
total	29

What is the preference of the group as a whole?

Overall, the group prefers the Democrat to the Republican (by five votes, seventeen of the twenty nine ranked the Democrat above the Republican versus twelve the other way). And, overall, the group prefers the Republican to the Third's nominee (by one vote, fifteen to fourteen). However, strangely enough, the group also prefers the Third to the Democrat (by seven votes; eighteen to eleven).

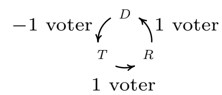


This is an example of a *voting paradox*, specifically, it is a *majority cycle*.

Voting paradoxes are studied in part because they have implications for practical politics. For instance, the class's instructor can manipulate them into choosing the Democrat by first asking them to choose between the Republican and Third candidates and then asking them to choose between the winner of that contest (the Republican) and the Democrat.

Voting paradoxes are also studied simply because they are mathematically interesting. One interesting aspect is that the overall majority cycle occurs despite that each single voter's preference list is *rational* — in a straight-line order. The majority cycle seems to arise in the aggregate without being present in the elements of that aggregate. Recently, however, linear algebra has been used [Zwicker] to argue that a tendency toward cyclic preference is actually present in each voter's preference list, and that it surfaces in the total when there is more adding of the tendency than cancelling. This Topic covers that argument.

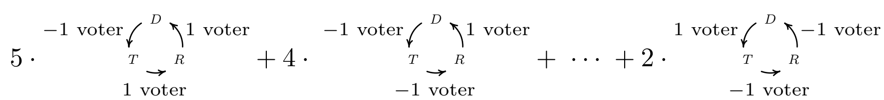
We start with a suitable description of each single voter's preference. A voter with the preference $D > R > T$ contributes to the cycle in this way.



(∗)

The negative sign is here because the arrow points from T to D while this voter's preference is the other way. Similar pictures for the other preference lists are in the table on page 150.

Conducting an election means collecting a set of $D > R > T$ preferences, a set of $T > R > D$ preferences, etc., and then somehow tallying those votes to calculate a combined, overall, winner. Many algorithms for the tallying have been proposed (all are subject to paradoxes like the one we are exploring here; see the readings cited at the end). For the Political Science mock election we did the mathematically most natural thing: we linearly combined the lists. There were five students in the set with the first preference list, four students in the set with the second, etc., so we got this.



Of course, taking linear combinations is linear algebra. The above cycle notation is inconvienient for calculations so we will instead use column vectors by taking the numbers counterclockwise from D. Thus, the class's mock election and the $D > R > T$ voter from (∗) are represented in this way.

$$\begin{pmatrix}7\\1\\5\end{pmatrix} \quad\text{and}\quad \begin{pmatrix}-1\\1\\1\end{pmatrix}$$

Now, to see how a tendency toward cycling is present in each single preference list we will decompose vote vectors into two parts, a cyclic part and an acyclic part. For the first part, we say that a vector is *purely cyclic* if it is in this subspace of $\mathbb{R}^3$.

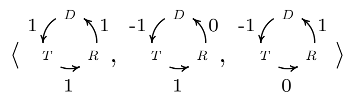
$$C = \{\begin{pmatrix}k\\k\\k\end{pmatrix} \mid k \in \mathbb{R}\} = \{k \cdot \begin{pmatrix}1\\1\\1\end{pmatrix} \mid k \in \mathbb{R}\}$$

For the second part, consider the set of vectors that are perpendicular to all of the vectors in C (read the symbol aloud as "C perp").

$$\begin{aligned} C^{\perp} &= \{\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} \mid \begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} \cdot \begin{pmatrix}k\\k\\k\end{pmatrix} = 0 \text{ for all } k \in \mathbb{R}\} \\ &= \{\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} \mid c_1 + c_2 + c_3 = 0\} = \{c_2 \begin{pmatrix}-1\\1\\0\end{pmatrix} + c_3 \begin{pmatrix}-1\\0\\1\end{pmatrix} \mid c_2, c_3 \in \mathbb{R}\} \end{aligned}$$

Like C, this set is also a subspace of $\mathbb{R}^3$ (Exercise 6). Combining the natural basis of C with the natural basis of $C^\perp$ gives this,

$$\langle \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix} \rangle \qquad \text{that is}$$



which we can check is a basis of $\mathbb{R}^3$. *(Note for readers who have covered the relevant optional section: that is, $\mathbb{R}^3$ is the direct sum of C and $C^\perp$.)*

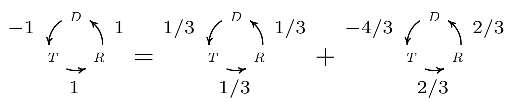
We can represent preference lists with respect to this basis and thereby decompose them into a cyclic part and an acyclic part. Consider the $D > R > T$ voter ($*$). The equations for the representation in terms of the basis

$$\begin{array}{rcl} c_1 - c_2 - c_3 &=& -1 \\ c_1 + c_2 &=& 1 \\ c_1 + c_3 &=& 1 \end{array} \xrightarrow[-\rho_1+\rho_3]{-\rho_1+\rho_2} \xrightarrow{(-1/2)\rho_2+\rho_3} \begin{array}{rcl} c_1 - c_2 - c_3 &=& -1 \\ 2c_2 + c_3 &=& 2 \\ (3/2)c_3 &=& 1 \end{array}$$

give that $c_1 = 1/3$, $c_2 = 2/3$, and $c_3 = 2/3$.

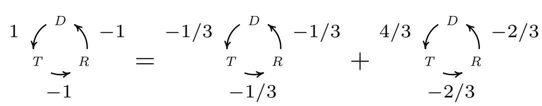
$$\begin{pmatrix}-1\\1\\1\end{pmatrix} = \frac{1}{3}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix} + \frac{2}{3}\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix} + \frac{2}{3}\cdot\begin{pmatrix}-1\\0\\1\end{pmatrix} = \begin{pmatrix}1/3\\1/3\\1/3\end{pmatrix} + \begin{pmatrix}-4/3\\2/3\\2/3\end{pmatrix}$$

The picture version



shows clearly that we have the desired decomposition into a cyclic part and and an acyclic part — the first part has three arcs of equal weight while the second part gives zero when we take a counterclockwise total. Thus, this voter's rational preference $D > R > T$ can indeed be seen to have a part that is cyclic.

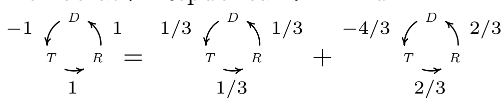
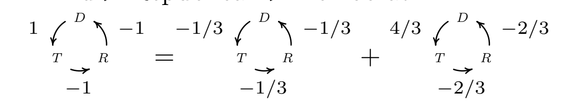
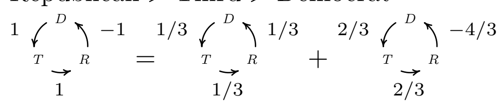
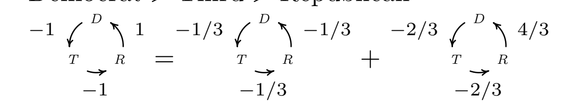
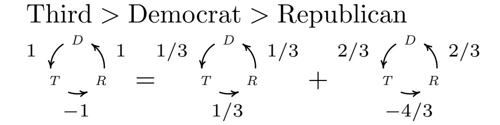
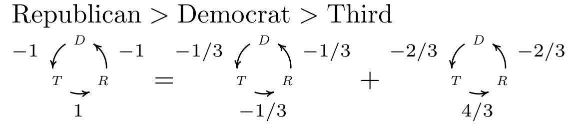
A $T > R > D$ voter is opposite to the one just considered in that the '$>$' preferences are reversed. This voter's decomposition



suggests the easily-checked fact that opposite preference lists have decompositions that are opposite. We contrast the two by saying that the $D > T > R$ voter has positive *spin* since the cycle part is with the direction that we have chosen for the arrows, while the $T > R > D >$ voter's spin is negative.

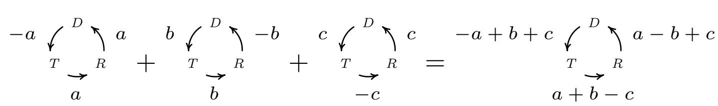
These opposite voters cancel each other in that their vote vectors add to zero. This suggests another way to calculate the linearly-combine-the-preferences winner. We could first cancel as many opposite preference lists as possible and then find the winner by combining the sets of preferences that remain.

The following table gives the six possible preference lists, where each row has a pair of opposites.

positive spin	*negative spin*
Democrat > Republican > Third 	Third > Republican > Democrat 
Republican > Third > Democrat 	Democrat > Third > Republican 
Third > Democrat > Republican 	Republican > Democrat > Third 

If we conduct the election as just described then after the cancellation of as many opposite pairs of voters as possible, there will remain three sets of preference lists, one set from the first row, one set from the second row, and one set from the third row. We will finish this Topic by proving that a voting paradox can happen only if those three sets come all from the left of the table, or all from the right. This shows that there is a connection between the majority cycle phenomenon and the decomposition that we are using — a voting paradox can happen only when the tendencies toward cyclic preference reinforce each other.

For the proof, assume that opposite preference orders have been cancelled, and we are left with one set of preference lists from each of the three rows. Consider the sum of these three (a, b, and c could be positive, negative, or zero).



A voting paradox occurs when the three numbers on the right, $a-b+c$ and $a+b-c$ and $-a+b+c$, are all nonnegative or all nonpositive. On the left, at least two of the three numbers, a and b and c, are both nonnegative or both nonpositive. We can assume that they are a and b. That makes four cases: the cycle is nonnegative and a and b are nonnegative, the cycle is nonpositive and a and b are nonpositive, etc. We will do only the first case, since the second is similar and the other two are also easy.

So assume that the cycle is nonnegative and that a and b are nonnegative. The conditions $0 \leq a-b+c$ and $0 \leq -a+b+c$ add to give that $0 \leq 2c$, which implies that c is also nonnegative, as desired. That ends the proof.

This result only says that having all three spin in the same direction is a necessary condition for a majority cycle. It is not also a sufficient condition; see Exercise 2.

Voting theory and associated topics are the subject of current research. The are many surprising and intriguing results, most notably the one produced by

K. Arrow [Arrow], who won the Nobel Prize in part for this work, showing, essentially, that no voting system is entirely fair. For more information, some good introductory articles are [Gardner, 1970], [Gardner, 1974], [Gardner, 1980], and [Neimi & Riker]. A quite readable recent book is [Taylor]. The material of this Topic is largely drawn from [Zwicker]. *(Author's Note: I would like to thank Professor Zwicker for his kind and illuminating discussions.)*

Exercises

1 Perform the cancellations of opposite preference orders for the Political Science class's mock election. Are all of the remaining preferences from the left of the table or from the right?

2 The necessary condition that is proved above — a voting paradox can happen only if all three preference lists remaining after cancellation have the same spin — is not also sufficient.

(a) Continuing the positive cycle case considered in the proof, use the two inequalities $0 \leq a - b + c$ and $0 \leq -a + b + c$ to show that $|a - b| \leq c$.

(b) Also show that $c \leq a + b$, and hence that $|a - b| \leq c \leq a + b$.

(c) Give an example of a vote where there is a majority cycle, and addition of one more voter with the same spin causes the cycle to go away.

(d) Can the opposite happen; can addition of one voter with a "wrong" spin cause a cycle to appear?

(e) Give a condition that is both necessary and sufficient to get a majority cycle.

3 A one-voter election cannot have a majority cycle because of the requirement that we've imposed that the voter's list must be rational.

(a) Show that a two-voter election may have a majority cycle. (We consider the group preference a majority cycle if all three group totals are nonnegative or if all three are nonpositive — that is, we allow some zero's in the group preference.)

(b) Show that for any number of voters greater than one, there is an election involving that many voters that results in a majority cycle.

4 Here is a reasonable way in which a voter could have a cyclic preference. Suppose that this voter ranks each candidate on each of three criteria.

(a) Draw up a table with the rows labelled 'Democrat', 'Republican', and 'Third', and the columns labelled 'character', 'experience', and 'policies'. Inside each column, rank some candidate as most preferred, rank another as in the middle, and rank the remaining one as least preferred.

(b) In this ranking, is the Democrat preferred to the Republican in (at least) two out of three criteria, or vice versa? Is the Republican preferred to the Third?

(c) Does the table that was just constructed have a cyclic preference order? If not, make one that does.

So it is possible for a voter to have a cyclic preference among candidates. The paradox described above, however, is that even if each voter has a straight-line preference list, there can still be a cyclic group preference.

5 Compute the values in the table of decompositions.

6 Let U be a subspace of $\mathbb{R}^3$. Prove that the set $U^\perp = \{\vec{v} \mid \vec{v} \cdot \vec{u} = 0 \text{ for all } \vec{u} \in U\}$ of vectors that are perpendicular to each vector in U is also a subspace of $\mathbb{R}^3$.

Topic: Dimensional Analysis

"You can't add apples and oranges," the old saying goes. It reflects our experience that in applications the quantities have units and keeping track of those units is worthwhile. Everyone has done calculations such as this one that use the units as a check.

$$60\,\frac{\text{sec}}{\text{min}}\cdot 60\,\frac{\text{min}}{\text{hr}}\cdot 24\,\frac{\text{hr}}{\text{day}}\cdot 365\,\frac{\text{day}}{\text{year}} = 31\,536\,000\,\frac{\text{sec}}{\text{year}}$$

However, the idea of including the units can be taken beyond bookkeeping. It can be used to draw conclusions about what relationships are possible among the physical quantities.

To start, consider the physics equation: distance $= 16\cdot(\text{time})^2$. If the distance is in feet and the time is in seconds then this is a true statement about falling bodies. However it is not correct in other unit systems; for instance, it is not correct in the meter-second system. We can fix that by making the 16 a *dimensional constant*.

$$\text{dist} = 16\,\frac{\text{ft}}{\text{sec}^2}\cdot(\text{time})^2$$

For instance, the above equation holds in the yard-second system.

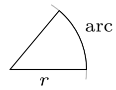
$$\text{distance in yards} = 16\,\frac{(1/3)\,\text{yd}}{\text{sec}^2}\cdot(\text{time in sec})^2 = \frac{16}{3}\,\frac{\text{yd}}{\text{sec}^2}\cdot(\text{time in sec})^2$$

So our first point is that by "including the units" we mean that we are restricting our attention to equations that use dimensional constants.

By using dimensional constants, we can be vague about units and say only that all quantities are measured in combinations of some units of length L, mass M, and time T. We shall refer to these three as *dimensions* (these are the only three dimensions that we shall need in this Topic). For instance, velocity could be measured in feet/second or fathoms/hour, but in all events it involves some unit of length divided by some unit of time so the *dimensional formula* of velocity is L/T. Similarly, the dimensional formula of density is M/L^3. We shall prefer using negative exponents over the fraction bars and we shall include the dimensions with a zero exponent, that is, we shall write the dimensional formula of velocity as $L^1M^0T^{-1}$ and that of density as $L^{-3}M^1T^0$.

In this context, "You can't add apples to oranges" becomes the advice to check that all of an equation's terms have the same dimensional formula. An example is this version of the falling body equation: $d-gt^2=0$. The dimensional formula of the d term is $L^1M^0T^0$. For the other term, the dimensional formula of g is $L^1M^0T^{-2}$ (g is the dimensional constant given above as $16\,\text{ft}/\text{sec}^2$) and the dimensional formula of t is $L^0M^0T^1$, so that of the entire gt^2 term is $L^1M^0T^{-2}(L^0M^0T^1)^2 = L^1M^0T^0$. Thus the two terms have the same dimensional formula. An equation with this property is *dimensionally homogeneous*.

Quantities with dimensional formula $L^0M^0T^0$ are *dimensionless*. For example, we measure an angle by taking the ratio of the subtended arc to the radius



which is the ratio of a length to a length $L^1M^0T^0/L^1M^0T^0$ and thus angles have the dimensional formula $L^0M^0T^0$.

The classic example of using the units for more than bookkeeping, using them to draw conclusions, considers the formula for the period of a pendulum.

$$p = \text{–some expression involving the length of the string, etc.–}$$

The period is in units of time $L^0M^0T^1$. So the quantities on the other side of the equation must have dimensional formulas that combine in such a way that their L's and M's cancel and only a single T remains. The table on page 154 has the quantities that an experienced investigator would consider possibly relevant. The only dimensional formulas involving L are for the length of the string and the acceleration due to gravity. For the L's of these two to cancel, when they appear in the equation they must be in ratio, e.g., as $(\ell/g)^2$, or as $\cos(\ell/g)$, or as $(\ell/g)^{-1}$. Therefore the period is a function of ℓ/g.

This is a remarkable result: with a pencil and paper analysis, before we ever took out the pendulum and made measurements, we have determined something about the relationship among the quantities.

To do dimensional analysis systematically, we need to know two things (arguments for these are in [Bridgman], Chapter II and IV). The first is that each equation relating physical quantities that we shall see involves a sum of terms, where each term has the form

$$m_1^{p_1} m_2^{p_2} \cdots m_k^{p_k}$$

for numbers m_1, ..., m_k that measure the quantities.

For the second, observe that an easy way to construct a dimensionally homogeneous expression is by taking a product of dimensionless quantities or by adding such dimensionless terms. Buckingham's Theorem states that any complete relationship among quantities with dimensional formulas can be algebraically manipulated into a form where there is some function f such that

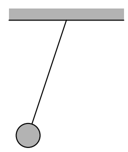
$$f(\Pi_1, \ldots, \Pi_n) = 0$$

for a complete set $\{\Pi_1, \ldots, \Pi_n\}$ of dimensionless products. (The first example below describes what makes a set of dimensionless products 'complete'.) We usually want to express one of the quantities, m_1 for instance, in terms of the others, and for that we will assume that the above equality can be rewritten

$$m_1 = m_2^{-p_2} \cdots m_k^{-p_k} \cdot \hat{f}(\Pi_2, \ldots, \Pi_n)$$

where $\Pi_1 = m_1 m_2^{p_2} \cdots m_k^{p_k}$ is dimensionless and the products Π_2, ..., Π_n don't involve m_1 (as with f, here $\hat{f}$ is just some function, this time of $n-1$ arguments). Thus, to do dimensional analysis we should find which dimensionless products are possible.

For example, consider again the formula for a pendulum's period.



quantity	*dimensional formula*
period p	$L^0M^0T^1$
length of string ℓ	$L^1M^0T^0$
mass of bob m	$L^0M^1T^0$
acceleration due to gravity g	$L^1M^0T^{-2}$
arc of swing θ	$L^0M^0T^0$

By the first fact cited above, we expect the formula to have (possibly sums of terms of) the form $p^{p_1}\ell^{p_2}m^{p_3}g^{p_4}\theta^{p_5}$. To use the second fact, to find which combinations of the powers $p_1, \ldots, p_5$ yield dimensionless products, consider this equation.

$$(L^0M^0T^1)^{p_1}(L^1M^0T^0)^{p_2}(L^0M^1T^0)^{p_3}(L^1M^0T^{-2})^{p_4}(L^0M^0T^0)^{p_5} = L^0M^0T^0$$

It gives three conditions on the powers.

$$\begin{aligned} p_2 \qquad + \quad p_4 &= 0 \\ p_3 \qquad\qquad &= 0 \\ p_1 \qquad\qquad - 2p_4 &= 0 \end{aligned}$$

Note that p_3 is 0 and so the mass of the bob does not affect the period. Gaussian reduction and parametrization of that system gives this

$$\{\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{pmatrix} = \begin{pmatrix} 1 \\ -1/2 \\ 0 \\ 1/2 \\ 0 \end{pmatrix} p_1 + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} p_5 \;\big|\; p_1, p_5 \in \mathbb{R}\}$$

(we've taken p_1 as one of the parameters in order to express the period in terms of the other quantities).

Here is the linear algebra. The set of dimensionless products contains all terms $p^{p_1}\ell^{p_2}m^{p_3}a^{p_4}\theta^{p_5}$ subject to the conditions above. This set forms a vector space under the '+' operation of multiplying two such products and the '·' operation of raising such a product to the power of the scalar (see Exercise 5). The term 'complete set of dimensionless products' in Buckingham's Theorem means a basis for this vector space.

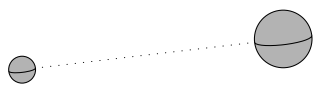
We can get a basis by first taking $p_1 = 1$, $p_5 = 0$ and then $p_1 = 0$, $p_5 = 1$. The associated dimensionless products are $\Pi_1 = p\ell^{-1/2}g^{1/2}$ and $\Pi_2 = \theta$. Because the set $\{\Pi_1, \Pi_2\}$ is complete, Buckingham's Theorem says that

$$p = \ell^{1/2}g^{-1/2} \cdot \hat{f}(\theta) = \sqrt{\ell/g} \cdot \hat{f}(\theta)$$

where $\hat{f}$ is a function that we cannot determine from this analysis (a first year physics text will show by other means that for small angles it is approximately the constant function $\hat{f}(\theta) = 2\pi$).

Thus, analysis of the relationships that are possible between the quantities with the given dimensional formulas has produced a fair amount of information: a pendulum's period does not depend on the mass of the bob, and it rises with the square root of the length of the string.

For the next example we try to determine the period of revolution of two bodies in space orbiting each other under mutual gravitational attraction. An experienced investigator could expect that these are the relevant quantities.



quantity	*dimensional formula*
period p	$L^0M^0T^1$
mean separation r	$L^1M^0T^0$
first mass m_1	$L^0M^1T^0$
second mass m_2	$L^0M^1T^0$
grav. constant G	$L^3M^{-1}T^{-2}$

To get the complete set of dimensionless products we consider the equation

$$(L^0M^0T^1)^{p_1}(L^1M^0T^0)^{p_2}(L^0M^1T^0)^{p_3}(L^0M^1T^0)^{p_4}(L^3M^{-1}T^{-2})^{p_5} = L^0M^0T^0$$

which results in a system

$$\begin{aligned} p_2 \qquad\qquad + 3p_5 &= 0 \\ p_3 + p_4 - p_5 &= 0 \\ p_1 \qquad\qquad\qquad - 2p_5 &= 0 \end{aligned}$$

with this solution.

$$\{\begin{pmatrix} 1 \\ -3/2 \\ 1/2 \\ 0 \\ 1/2 \end{pmatrix} p_1 + \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} p_4 \bigm| p_1, p_4 \in \mathbb{R}\}$$

As earlier, the linear algebra here is that the set of dimensionless products of these quantities forms a vector space, and we want to produce a basis for that space, a 'complete' set of dimensionless products. One such set, gotten from setting $p_1 = 1$ and $p_4 = 0$, and also setting $p_1 = 0$ and $p_4 = 1$ is $\{\Pi_1 = pr^{-3/2}m_1^{1/2}G^{1/2},\ \Pi_2 = m_1^{-1}m_2\}$. With that, Buckingham's Theorem says that any complete relationship among these quantities is stateable this form.

$$p = r^{3/2}m_1^{-1/2}G^{-1/2} \cdot \hat{f}(m_1^{-1}m_2) = \frac{r^{3/2}}{\sqrt{Gm_1}} \cdot \hat{f}(m_2/m_1)$$

Remark. An important application of the prior formula is when m_1 is the mass of the sun and m_2 is the mass of a planet. Because m_1 is very much greater than m_2, the argument to $\hat{f}$ is approximately 0, and we can wonder whether this part of the formula remains approximately constant as m_2 varies. One way to see that it does is this. The sun is so much larger than the planet that the mutual rotation is approximately about the sun's center. If we vary the planet's mass m_2 by a factor of x (e.g., Venus's mass is $x = 0.815$ times Earth's mass), then the force of attraction is multiplied by x, and x times the force acting on x times the mass gives, since $F = ma$, the same acceleration, about the same center (approximately). Hence, the orbit will be the same and so its period will be the same, and thus the right side of the above equation also remains unchanged (approximately). Therefore, $\hat{f}(m_2/m_1)$ is approximately constant as m_2 varies. This is Kepler's Third Law: the square of the period of a planet is proportional to the cube of the mean radius of its orbit about the sun.

The final example was one of the first explicit applications of dimensional analysis. Lord Raleigh considered the speed of a wave in deep water and suggested these as the relevant quantities.

quantity	*dimensional formula*
velocity of the wave v	$L^1M^0T^{-1}$
density of the water d	$L^{-3}M^1T^0$
acceleration due to gravity g	$L^1M^0T^{-2}$
wavelength λ	$L^1M^0T^0$

The equation

$$(L^1M^0T^{-1})^{p_1}(L^{-3}M^1T^0)^{p_2}(L^1M^0T^{-2})^{p_3}(L^1M^0T^0)^{p_4} = L^0M^0T^0$$

gives this system

$$\begin{array}{rcrcrcrcl} p_1 & - & 3p_2 & + & p_3 & + & p_4 & = & 0 \\ & & p_2 & & & & & = & 0 \\ -p_1 & & & - & 2p_3 & & & = & 0 \end{array}$$

with this solution space

$$\{\begin{pmatrix} 1 \\ 0 \\ -1/2 \\ -1/2 \end{pmatrix} p_1 \bigm| p_1 \in \mathbb{R}\}$$

(as in the pendulum example, one of the quantities d turns out not to be involved in the relationship). There is one dimensionless product, $\Pi_1 = vg^{-1/2}\lambda^{-1/2}$, and so v is $\sqrt{\lambda g}$ times a constant ($\hat{f}$ is constant since it is a function of no arguments).

As the three examples above show, dimensional analysis can bring us far toward expressing the relationship among the quantities. For further reading, the classic reference is [Bridgman]—this brief book is delightful. Another source is [Giordano, Wells, Wilde]. A description of dimensional analysis's place in modeling is in [Giordano, Jaye, Weir].

Exercises

1 Consider a projectile, launched with initial velocity v_0, at an angle θ. An investigation of this motion might start with the guess that these are the relevant quantities. [de Mestre]

quantity	*dimensional formula*
horizontal position x	$L^1M^0T^0$
vertical position y	$L^1M^0T^0$
initial speed v_0	$L^1M^0T^{-1}$
angle of launch θ	$L^0M^0T^0$
acceleration due to gravity g	$L^1M^0T^{-2}$
time t	$L^0M^0T^1$

(a) Show that $\{gt/v_0, gx/v_0^2, gy/v_0^2, \theta\}$ is a complete set of dimensionless products. (*Hint.* This can be done by finding the appropriate free variables in the linear system that arises, but there is a shortcut that uses the properties of a basis.)

(b) These two equations of motion for projectiles are familiar: $x = v_0\cos(\theta)t$ and $y = v_0\sin(\theta)t - (g/2)t^2$. Manipulate each to rewrite it as a relationship among the dimensionless products of the prior item.

2 [Einstein] conjectured that the infrared characteristic frequencies of a solid may be determined by the same forces between atoms as determine the solid's ordanary elastic behavior. The relevant quantities are

quantity	*dimensional formula*
characteristic frequency ν	$L^0M^0T^{-1}$
compressibility k	$L^1M^{-1}T^2$
number of atoms per cubic cm N	$L^{-3}M^0T^0$
mass of an atom m	$L^0M^1T^0$

Show that there is one dimensionless product. Conclude that, in any complete relationship among quantities with these dimensional formulas, k is a constant times $\nu^{-2}N^{-1/3}m^{-1}$. This conclusion played an important role in the early study of quantum phenomena.

3 The torque produced by an engine has dimensional formula $L^2M^1T^{-2}$. We may first guess that it depends on the engine's rotation rate (with dimensional formula $L^0M^0T^{-1}$), and the volume of air displaced (with dimensional formula $L^3M^0T^0$). [Giordano, Wells, Wilde]

(a) Try to find a complete set of dimensionless products. What goes wrong?

(b) Adjust the guess by adding the density of the air (with dimensional formula $L^{-3}M^1T^0$). Now find a complete set of dimensionless products.

4 Dominoes falling make a wave. We may conjecture that the wave speed v depends on the the spacing d between the dominoes, the height h of each domino, and the acceleration due to gravity g. [Tilley]

(a) Find the dimensional formula for each of the four quantities.

(b) Show that $\{\Pi_1 = h/d, \Pi_2 = dg/v^2\}$ is a complete set of dimensionless products.

(c) Show that if h/d is fixed then the propagation speed is proportional to the square root of d.

5 Prove that the dimensionless products form a vector space under the $\vec{+}$ operation of multiplying two such products and the $\vec{\cdot}$ operation of raising such the product to the power of the scalar. (The vector arrows are a precaution against confusion.) That is, prove that, for any particular homogeneous system, this set of products of powers of $m_1, \ldots, m_k$

$$\{m_1^{p_1} \ldots m_k^{p_k} \bigm| p_1, \ldots, p_k \text{ satisfy the system}\}$$

is a vector space under:

$$m_1^{p_1} \ldots m_k^{p_k} \vec{+} m_1^{q_1} \ldots m_k^{q_k} = m_1^{p_1+q_1} \ldots m_k^{p_k+q_k}$$

and

$$r \vec{\cdot} (m_1^{p_1} \ldots m_k^{p_k}) = m_1^{rp_1} \ldots m_k^{rp_k}$$

(assume that all variables represent real numbers).

6 The advice about apples and oranges is not right. Consider the familiar equations for a circle $C = 2\pi r$ and $A = \pi r^2$.

(a) Check that C and A have different dimensional formulas.

(b) Produce an equation that is not dimensionally homogeneous (i.e., it adds apples and oranges) but is nonetheless true of any circle.

(c) The prior item asks for an equation that is complete but not dimensionally homogeneous. Produce an equation that is dimensionally homogeneous but not complete.

(Just because the old saying isn't strictly right, doesn't keep it from being a useful strategy. Dimensional homogeneity is often used as a check on the plausibility of equations used in models. For an argument that any complete equation can easily be made dimensionally homogeneous, see [Bridgman], Chapter I, especially page 15.)