Chapter Three

Maps Between Spaces

I Isomorphisms

In the examples following the definition of a vector space we developed the intuition that some spaces are "the same" as others. For instance, the space of two-tall column vectors and the space of two-wide row vectors are not equal because their elements — column vectors and row vectors — are not equal, but we have the idea that these spaces differ only in how their elements appear. We will now make this idea precise.

This section illustrates a common aspect of a mathematical investigation. With the help of some examples, we've gotten an idea. We will next give a formal definition, and then we will produce some results backing our contention that the definition captures the idea. We've seen this happen already, for instance, in the first section of the Vector Space chapter. There, the study of linear systems led us to consider collections closed under linear combinations. We defined such a collection as a vector space, and we followed it with some supporting results.

Of course, that definition wasn't an end point, instead it led to new insights such as the idea of a basis. Here too, after producing a definition, and supporting it, we will get two surprises (pleasant ones). First, we will find that the definition applies to some unforeseen, and interesting, cases. Second, the study of the definition will lead to new ideas. In this way, our investigation will build a momentum.

I.1 Definition and Examples

We start with two examples that suggest the right definition.

1.1 Example Consider the example mentioned above, the space of two-wide row vectors and the space of two-tall column vectors. They are "the same" in that if we associate the vectors that have the same components, e.g.,

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \quad \longleftrightarrow \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

then this correspondence preserves the operations, for instance this addition

$$\begin{pmatrix}1 & 2\end{pmatrix} + \begin{pmatrix}3 & 4\end{pmatrix} = \begin{pmatrix}4 & 6\end{pmatrix} \quad\longleftrightarrow\quad \begin{pmatrix}1\\2\end{pmatrix} + \begin{pmatrix}3\\4\end{pmatrix} = \begin{pmatrix}4\\6\end{pmatrix}$$

and this scalar multiplication.

$$5 \cdot \begin{pmatrix}1 & 2\end{pmatrix} = \begin{pmatrix}5 & 10\end{pmatrix} \quad\longleftrightarrow\quad 5 \cdot \begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}5\\10\end{pmatrix}$$

More generally stated, under the correspondence

$$\begin{pmatrix}a_0 & a_1\end{pmatrix} \quad\longleftrightarrow\quad \begin{pmatrix}a_0\\a_1\end{pmatrix}$$

both operations are preserved:

$$\begin{pmatrix}a_0 & a_1\end{pmatrix} + \begin{pmatrix}b_0 & b_1\end{pmatrix} = \begin{pmatrix}a_0+b_0 & a_1+b_1\end{pmatrix} \longleftrightarrow \begin{pmatrix}a_0\\a_1\end{pmatrix} + \begin{pmatrix}b_0\\b_1\end{pmatrix} = \begin{pmatrix}a_0+b_0\\a_1+b_1\end{pmatrix}$$

and

$$r \cdot \begin{pmatrix}a_0 & a_1\end{pmatrix} = \begin{pmatrix}ra_0 & ra_1\end{pmatrix} \quad\longleftrightarrow\quad r \cdot \begin{pmatrix}a_0\\a_1\end{pmatrix} = \begin{pmatrix}ra_0\\ra_1\end{pmatrix}$$

(all of the variables are real numbers).

1.2 Example Another two spaces we can think of as "the same" are $\mathcal{P}_2$, the space of quadratic polynomials, and $\mathbb{R}^3$. A natural correspondence is this.

$$a_0 + a_1x + a_2x^2 \quad\longleftrightarrow\quad \begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix} \qquad (\text{e.g., } 1 + 2x + 3x^2 \longleftrightarrow \begin{pmatrix}1\\2\\3\end{pmatrix})$$

The structure is preserved: corresponding elements add in a corresponding way

$$\begin{array}{r} a_0 + a_1x + a_2x^2 \\ +\; b_0 + b_1x + b_2x^2 \\ \hline (a_0+b_0) + (a_1+b_1)x + (a_2+b_2)x^2 \end{array} \quad\longleftrightarrow\quad \begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix} + \begin{pmatrix}b_0\\b_1\\b_2\end{pmatrix} = \begin{pmatrix}a_0+b_0\\a_1+b_1\\a_2+b_2\end{pmatrix}$$

and scalar multiplication corresponds also.

$$r \cdot (a_0 + a_1x + a_2x^2) = (ra_0) + (ra_1)x + (ra_2)x^2 \quad\longleftrightarrow\quad r \cdot \begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix} = \begin{pmatrix}ra_0\\ra_1\\ra_2\end{pmatrix}$$

1.3 Definition An *isomorphism* between two vector spaces V and W is a map $f\colon V \to W$ that

(1) is a correspondence: f is one-to-one and onto;*

(2) *preserves structure:* if $\vec{v}_1, \vec{v}_2 \in V$ then

$$f(\vec{v}_1 + \vec{v}_2) = f(\vec{v}_1) + f(\vec{v}_2)$$

and if $\vec{v} \in V$ and $r \in \mathbb{R}$ then

$$f(r\vec{v}) = r\, f(\vec{v})$$

(we write $V \cong W$, read "V is isomorphic to W", when such a map exists).

("Morphism" means map, so "isomorphism" means a map expressing sameness.)

1.4 Example The vector space $G = \{c_1 \cos\theta + c_2 \sin\theta \bigm| c_1, c_2 \in \mathbb{R}\}$ of functions of θ is isomorphic to the vector space $\mathbb{R}^2$ under this map.

$$c_1 \cos\theta + c_2 \sin\theta \stackrel{f}{\longmapsto} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

We will check this by going through the conditions in the definition.

We will first verify condition (1), that the map is a correspondence between the sets underlying the spaces.

To establish that f is one-to-one, we must prove that $f(\vec{a}) = f(\vec{b})$ only when $\vec{a} = \vec{b}$. If

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

then, by the definition of f,

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

from which we can conclude that $a_1 = b_1$ and $a_2 = b_2$ because column vectors are equal only when they have equal components. We've proved that $f(\vec{a}) = f(\vec{b})$ implies that $\vec{a} = \vec{b}$, which shows that f is one-to-one.

To check that f is onto we must check that any member of the codomain $\mathbb{R}^2$ the image of some member of the domain G. But that's clear — any

$$\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$$

is the image under f of $x\cos\theta + y\sin\theta \in G$.

Next we will verify condition (2), that f preserves structure.

*More information on one-to-one and onto maps is in the appendix.

This computation shows that f preserves addition.

$$\begin{aligned}
f\big(\,(a_1\cos\theta+a_2\sin\theta)+(b_1\cos\theta+b_2\sin\theta)\,\big)&\\
&=f\big(\,(a_1+b_1)\cos\theta+(a_2+b_2)\sin\theta\,\big)\\
&=\begin{pmatrix}a_1+b_1\\a_2+b_2\end{pmatrix}\\
&=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}b_1\\b_2\end{pmatrix}\\
&=f(a_1\cos\theta+a_2\sin\theta)+f(b_1\cos\theta+b_2\sin\theta)
\end{aligned}$$

A similar computation shows that f preserves scalar multiplication.

$$\begin{aligned}
f\big(\,r\cdot(a_1\cos\theta+a_2\sin\theta)\,\big)&=f(\,ra_1\cos\theta+ra_2\sin\theta\,)\\
&=\begin{pmatrix}ra_1\\ra_2\end{pmatrix}\\
&=r\cdot\begin{pmatrix}a_1\\a_2\end{pmatrix}\\
&=r\cdot\,f(a_1\cos\theta+a_2\sin\theta)
\end{aligned}$$

With that, conditions (1) and (2) are verified, so we know that f is an isomorphism and we can say that the spaces are isomorphic $G\cong\mathbb{R}^2$.

1.5 Example Let V be the space $\{c_1x+c_2y+c_3z \mid c_1,c_2,c_3\in\mathbb{R}\}$ of linear combinations of three variables x, y, and z, under the natural addition and scalar multiplication operations. Then V is isomorphic to $\mathcal{P}_2$, the space of quadratic polynomials.

To show this we will produce an isomorphism map. There is more than one possibility; for instance, here are four.

$$c_1x+c_2y+c_3z\qquad\begin{array}{ll}
\stackrel{f_1}{\longmapsto} & c_1+c_2x+c_3x^2\\
\stackrel{f_2}{\longmapsto} & c_2+c_3x+c_1x^2\\
\stackrel{f_3}{\longmapsto} & -c_1-c_2x-c_3x^2\\
\stackrel{f_4}{\longmapsto} & c_1+(c_1+c_2)x+(c_1+c_3)x^2
\end{array}$$

The first map is the more natural correspondence in that it just carries the coefficients over. However, below we shall verify that the second one is an isomorphism, to underline that there are isomorphisms other than just the obvious one (showing that f_1 is an isomorphism is Exercise 12).

To show that f_2 is one-to-one, we will prove that if $f_2(c_1x+c_2y+c_3z)=f_2(d_1x+d_2y+d_3z)$ then $c_1x+c_2y+c_3z=d_1x+d_2y+d_3z$. The assumption that $f_2(c_1x+c_2y+c_3z)=f_2(d_1x+d_2y+d_3z)$ gives, by the definition of f_2, that $c_2+c_3x+c_1x^2=d_2+d_3x+d_1x^2$. Equal polynomials have equal coefficients, so $c_2=d_2$, $c_3=d_3$, and $c_1=d_1$. Thus $f_2(c_1x+c_2y+c_3z)=f_2(d_1x+d_2y+d_3z)$ implies that $c_1x+c_2y+c_3z=d_1x+d_2y+d_3z$ and therefore f_2 is one-to-one.

The map f_2 is onto because any member $a+bx+cx^2$ of the codomain is the image of some member of the domain, namely it is the image of $cx+ay+bz$. For instance, $2+3x-4x^2$ is $f_2(-4x+2y+3z)$.

The computations for structure preservation are like those in the prior example. This map preserves addition

$$\begin{aligned}
&f_2\bigl((c_1x+c_2y+c_3z)+(d_1x+d_2y+d_3z)\bigr)\\
&\qquad\qquad = f_2\bigl((c_1+d_1)x+(c_2+d_2)y+(c_3+d_3)z\bigr)\\
&\qquad\qquad = (c_2+d_2)+(c_3+d_3)x+(c_1+d_1)x^2\\
&\qquad\qquad = (c_2+c_3x+c_1x^2)+(d_2+d_3x+d_1x^2)\\
&\qquad\qquad = f_2(c_1x+c_2y+c_3z)+f_2(d_1x+d_2y+d_3z)
\end{aligned}$$

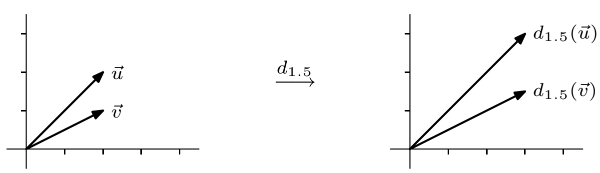
and scalar multiplication.

$$\begin{aligned}
f_2\bigl(r\cdot(c_1x+c_2y+c_3z)\bigr) &= f_2(rc_1x+rc_2y+rc_3z)\\
&= rc_2+rc_3x+rc_1x^2\\
&= r\cdot(c_2+c_3x+c_1x^2)\\
&= r\cdot\ f_2(c_1x+c_2y+c_3z)
\end{aligned}$$

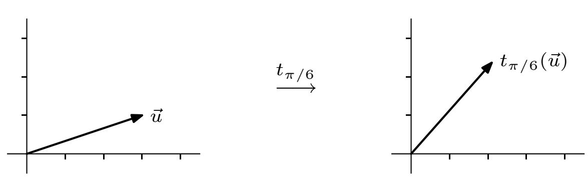
Thus f_2 is an isomorphism and we write $V\cong\mathcal{P}_2$.

We are sometimes interested in an isomorphism of a space with itself, called an *automorphism*. An identity map is an automorphism. The next two examples show that there are others.

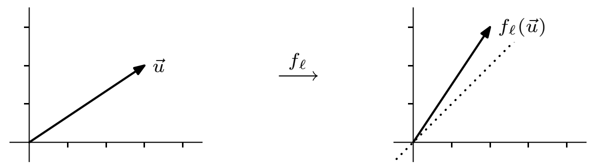
1.6 Example A *dilation* map $d_s\colon\mathbb{R}^2\to\mathbb{R}^2$ that multiplies all vectors by a nonzero scalar s is an automorphism of $\mathbb{R}^2$.



A *rotation* or *turning map* $t_\theta\colon\mathbb{R}^2\to\mathbb{R}^2$ that rotates all vectors through an angle θ is an automorphism.



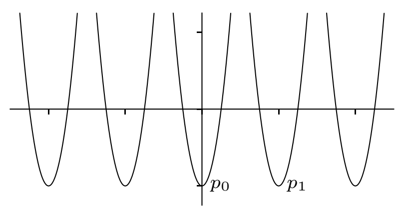
A third type of automorphism of $\mathbb{R}^2$ is a map $f_\ell\colon\mathbb{R}^2\to\mathbb{R}^2$ that *flips* or *reflects* all vectors over a line ℓ through the origin.



See Exercise 29.

1.7 Example Consider the space $\mathcal{P}_5$ of polynomials of degree 5 or less and the map f that sends a polynomial $p(x)$ to $p(x-1)$. For instance, under this map $x^2 \mapsto (x-1)^2 = x^2-2x+1$ and $x^3+2x \mapsto (x-1)^3+2(x-1) = x^3-3x^2+5x-3$. This map is an automorphism of this space; the check is Exercise 21.

This isomorphism of $\mathcal{P}_5$ with itself does more than just tell us that the space is "the same" as itself. It gives us some insight into the space's structure. For instance, below is shown a family of parabolas, graphs of members of $\mathcal{P}_5$. Each has a vertex at $y=-1$, and the left-most one has zeroes at -2.25 and -1.75, the next one has zeroes at -1.25 and -0.75, etc.



Geometrically, the substitution of $x-1$ for x in any function's argument shifts its graph to the right by one. Thus, $f(p_0)=p_1$ and f's action is to shift all of the parabolas to the right by one. Notice that the picture before f is applied is the same as the picture after f is applied, because while each parabola moves to the right, another one comes in from the left to take its place. This also holds true for cubics, etc. So the automorphism f gives us the insight that P_5 has a certain horizontal-homogeneity; this space looks the same near $x=1$ as near $x=0$.

As described in the preamble to this section, we will next produce some results supporting the contention that the definition of isomorphism above captures our intuition of vector spaces being the same.

Of course the definition itself is persuasive: a vector space consists of two components, a set and some structure, and the definition simply requires that the sets correspond and that the structures correspond also. Also persuasive are the examples above. In particular, Example 1.1, which gives an isomorphism between the space of two-wide row vectors and the space of two-tall column vectors, dramatizes our intuition that isomorphic spaces are the same in all relevant respects. Sometimes people say, where $V \cong W$, that "W is just V painted green" — any differences are merely cosmetic.

Further support for the definition, in case it is needed, is provided by the following results that, taken together, suggest that all the things of interest in a vector space correspond under an isomorphism. Since we studied vector spaces to study linear combinations, "of interest" means "pertaining to linear combinations". Not of interest is the way that the vectors are presented typographically (or their color!).

As an example, although the definition of isomorphism doesn't explicitly say that the zero vectors must correspond, it is a consequence of that definition.

1.8 Lemma An isomorphism maps a zero vector to a zero vector.

Proof. Where $f\colon V \to W$ is an isomorphism, fix any $\vec{v} \in V$. Then $f(\vec{0}_V) = f(0 \cdot \vec{v}) = 0 \cdot f(\vec{v}) = \vec{0}_W$. QED

The definition of isomorphism requires that sums of two vectors correspond and that so do scalar multiples. We can extend that to say that all linear combinations correspond.

1.9 Lemma For any map $f\colon V \to W$ between vector spaces these statements are equivalent.

(1) f preserves structure

$$f(\vec{v}_1 + \vec{v}_2) = f(\vec{v}_1) + f(\vec{v}_2) \quad\text{and}\quad f(c\vec{v}) = c\,f(\vec{v})$$

(2) f preserves linear combinations of two vectors

$$f(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1 f(\vec{v}_1) + c_2 f(\vec{v}_2)$$

(3) f preserves linear combinations of any finite number of vectors

$$f(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = c_1 f(\vec{v}_1) + \cdots + c_n f(\vec{v}_n)$$

Proof. Since the implications (3) $\Longrightarrow$ (2) and (2) $\Longrightarrow$ (1) are clear, we need only show that (1) $\Longrightarrow$ (3). Assume statement (1). We will prove statement (3) by induction on the number of summands n.

The one-summand base case, that $f(c\vec{v}_1) = c\,f(\vec{v}_1)$, is covered by the assumption of statement (1).

For the inductive step assume that statement (3) holds whenever there are k or fewer summands, that is, whenever $n = 1$, or $n = 2$, …, or $n = k$. Consider the $k+1$-summand case. The first half of (1) gives

$$f(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v}_{k+1}) = f(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k) + f(c_{k+1}\vec{v}_{k+1})$$

by breaking the sum along the final '+'. Then the inductive hypothesis lets us break up the k-term sum.

$$= f(c_1\vec{v}_1) + \cdots + f(c_k\vec{v}_k) + f(c_{k+1}\vec{v}_{k+1})$$

Finally, the second half of statement (1) gives

$$= c_1\,f(\vec{v}_1) + \cdots + c_k\,f(\vec{v}_k) + c_{k+1}\,f(\vec{v}_{k+1})$$

when applied $k+1$ times. QED

In addition to adding to the intuition that the definition of isomorphism does indeed preserve the things of interest in a vector space, that lemma's second item is an especially handy way of checking that a map preserves structure.

We close with a summary. The material in this section augments the chapter on Vector Spaces. There, after giving the definition of a vector space, we informally looked at what different things can happen. Here, we defined the relation '$\cong$' between vector spaces and we have argued that it is the right way to split the collection of vector spaces into cases because it preserves the features of interest in a vector space — in particular, it preserves linear combinations. That is, we have now said precisely what we mean by 'the same', and by 'different', and so we have precisely classified the vector spaces.

Exercises

✓ **1.10** Verify, using Example 1.4 as a model, that the two correspondences given before the definition are isomorphisms.
(a) Example 1.1 **(b)** Example 1.2

✓ **1.11** For the map $f\colon \mathcal{P}_1 \to \mathbb{R}^2$ given by

$$a + bx \stackrel{f}{\longmapsto} \begin{pmatrix} a-b \\ b \end{pmatrix}$$

Find the image of each of these elements of the domain.
(a) $3-2x$ **(b)** $2+2x$ **(c)** x

Show that this map is an isomorphism.

1.12 Show that the natural map f_1 from Example 1.5 is an isomorphism.

HW ✓ **1.13** Decide whether each map is an isomorphism (of course, if it is an isomorphism then prove it and if it isn't then state a condition that it fails to satisfy).

(a) $f\colon \mathcal{M}_{2\times 2} \to \mathbb{R}$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc$$

(b) $f\colon \mathcal{M}_{2\times 2} \to \mathbb{R}^4$ given by

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

(c) $f\colon \mathcal{M}_{2\times 2} \to \mathcal{P}_3$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto c + (d+c)x + (b+a)x^2 + ax^3$$

(d) $f\colon \mathcal{M}_{2\times 2} \to \mathcal{P}_3$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto c + (d+c)x + (b+a+1)x^2 + ax^3$$

1.14 Show that the map $f\colon \mathbb{R}^1 \to \mathbb{R}^1$ given by $f(x) = x^3$ is one-to-one and onto. Is it an isomorphism?

✓ **1.15** Refer to Example 1.1. Produce two more isomorphisms (of course, that they satisfy the conditions in the definition of isomorphism must be verified).

1.16 Refer to Example 1.2. Produce two more isomorphisms (and verify that they satisfy the conditions).

✓ **1.17** Show that, although $\mathbb{R}^2$ is not itself a subspace of $\mathbb{R}^3$, it is isomorphic to the xy-plane subspace of $\mathbb{R}^3$.

1.18 Find two isomorphisms between $\mathbb{R}^{16}$ and $\mathcal{M}_{4\times4}$.

✓ **1.19** For what k is $\mathcal{M}_{m\times n}$ isomorphic to $\mathbb{R}^k$?

1.20 For what k is $\mathcal{P}_k$ isomorphic to $\mathbb{R}^n$?

1.21 Prove that the map in Example 1.7, from $\mathcal{P}_5$ to $\mathcal{P}_5$ given by $p(x) \mapsto p(x-1)$, is a vector space isomorphism.

1.22 Why, in Lemma 1.8, must there be a $\vec{v} \in V$? That is, why must V be nonempty?

1.23 Are any two trivial spaces isomorphic?

1.24 In the proof of Lemma 1.9, what about the zero-summands case (that is, if n is zero)?

1.25 Show that any isomorphism $f\colon \mathcal{P}_0 \to \mathbb{R}^1$ has the form $a \mapsto ka$ for some nonzero real number k.

✓ **1.26** These prove that isomorphism is an equivalence relation.
(a) Show that the identity map $\mathrm{id}\colon V \to V$ is an isomorphism. Thus, any vector space is isomorphic to itself.
(b) Show that if $f\colon V \to W$ is an isomorphism then so is its inverse $f^{-1}\colon W \to V$. Thus, if V is isomorphic to W then also W is isomorphic to V.
(c) Show that a composition of isomorphisms is an isomorphism: if $f\colon V \to W$ is an isomorphism and $g\colon W \to U$ is an isomorphism then so also is $g \circ f\colon V \to U$. Thus, if V is isomorphic to W and W is isomorphic to U, then also V is isomorphic to U.

1.27 Suppose that $f\colon V \to W$ preserves structure. Show that f is one-to-one if and only if the unique member of V mapped by f to $\vec{0}_W$ is $\vec{0}_V$.

1.28 Suppose that $f\colon V \to W$ is an isomorphism. Prove that the set $\{\vec{v}_1, \dots, \vec{v}_k\} \subseteq V$ is linearly dependent if and only if the set of images $\{f(\vec{v}_1), \dots, f(\vec{v}_k)\} \subseteq W$ is linearly dependent.

✓ **1.29** Show that each type of map from Example 1.6 is an automorphism.
(a) Dilation d_s by a nonzero scalar s.
(b) Rotation t_θ through an angle θ.
(c) Reflection f_ℓ over a line through the origin.
Hint. For the second and third items, polar coordinates are useful.

1.30 Produce an automorphism of $\mathcal{P}_2$ other than the identity map, and other than a shift map $p(x) \mapsto p(x-k)$.

1.31 **(a)** Show that a function $f\colon \mathbb{R}^1 \to \mathbb{R}^1$ is an automorphism if and only if it has the form $x \mapsto kx$ for some $k \neq 0$.
(b) Let f be an automorphism of $\mathbb{R}^1$ such that $f(3) = 7$. Find $f(-2)$.
(c) Show that a function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is an automorphism if and only if it has the form

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

for some $a, b, c, d \in \mathbb{R}$ with $ad - bc \neq 0$. *Hint.* Exercises in prior subsections have shown that

$$\begin{pmatrix} b \\ d \end{pmatrix} \text{ is not a multiple of } \begin{pmatrix} a \\ c \end{pmatrix}$$

if and only if $ad - bc \neq 0$.

(d) Let f be an automorphism of $\mathbb{R}^2$ with

$$f(\begin{pmatrix}1\\3\end{pmatrix})=\begin{pmatrix}2\\-1\end{pmatrix}\quad\text{and}\quad f(\begin{pmatrix}1\\4\end{pmatrix})=\begin{pmatrix}0\\1\end{pmatrix}.$$

Find

$$f(\begin{pmatrix}0\\-1\end{pmatrix}).$$

1.32 Refer to Lemma 1.8 and Lemma 1.9. Find two more things preserved by isomorphism.

1.33 We show that isomorphisms can be tailored to fit in that, sometimes, given vectors in the domain and in the range we can produce an isomorphism associating those vectors.

(a) Let $B=\langle\vec{\beta}_1,\vec{\beta}_2,\vec{\beta}_3\rangle$ be a basis for $\mathcal{P}_2$ so that any $\vec{p}\in\mathcal{P}_2$ has a unique representation as $\vec{p}=c_1\vec{\beta}_1+c_2\vec{\beta}_2+c_3\vec{\beta}_3$, which we denote in this way.

$$\mathrm{Rep}_B(\vec{p})=\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}$$

Show that the $\mathrm{Rep}_B(\cdot)$ operation is a function from $\mathcal{P}_2$ to $\mathbb{R}^3$ (this entails showing that with every domain vector $\vec{v}\in\mathcal{P}_2$ there is an associated image vector in $\mathbb{R}^3$, and further, that with every domain vector $\vec{v}\in\mathcal{P}_2$ there is at most one associated image vector).

(b) Show that this $\mathrm{Rep}_B(\cdot)$ function is one-to-one and onto.

(c) Show that it preserves structure.

(d) Produce an isomorphism from $\mathcal{P}_2$ to $\mathbb{R}^3$ that fits these specifications.

$$x+x^2\mapsto\begin{pmatrix}1\\0\\0\end{pmatrix}\quad\text{and}\quad 1-x\mapsto\begin{pmatrix}0\\1\\0\end{pmatrix}$$

1.34 Prove that a space is n-dimensional if and only if it is isomorphic to $\mathbb{R}^n$. *Hint.* Fix a basis B for the space and consider the map sending a vector over to its representation with respect to B.

1.35 *(Requires the subsection on Combining Subspaces, which is optional.)* Let U and W be vector spaces. Define a new vector space, consisting of the set $U\times W=\{(\vec{u},\vec{w})\mid\vec{u}\in U\text{ and }\vec{w}\in W\}$ along with these operations.

$$(\vec{u}_1,\vec{w}_1)+(\vec{u}_2,\vec{w}_2)=(\vec{u}_1+\vec{u}_2,\vec{w}_1+\vec{w}_2)\quad\text{and}\quad r\cdot(\vec{u},\vec{w})=(r\vec{u},r\vec{w})$$

This is a vector space, the *external direct sum* of U and W.

(a) Check that it is a vector space.

(b) Find a basis for, and the dimension of, the external direct sum $\mathcal{P}_2\times\mathbb{R}^2$.

(c) What is the relationship among $\dim(U)$, $\dim(W)$, and $\dim(U\times W)$?

(d) Suppose that U and W are subspaces of a vector space V such that $V=U\oplus W$ (in this case we say that V is the *internal direct sum* of U and W). Show that the map $f\colon U\times W\to V$ given by

$$(\vec{u},\vec{w})\overset{f}{\longmapsto}\vec{u}+\vec{w}$$

is an isomorphism. Thus if the internal direct sum is defined then the internal and external direct sums are isomorphic.

II Homomorphisms

The definition of isomorphism has two conditions. In this section we will consider the second one, that the map must preserve the algebraic structure of the space. We will focus on this condition by studying maps that are required only to preserve structure; that is, maps that are not required to be correspondences.

Experience shows that this kind of map is tremendously useful in the study of vector spaces. For one thing, as we shall see in the second subsection below, while isomorphisms describe how spaces are the same, these maps describe how spaces can be thought of as alike.

II.1 Definition

1.1 Definition A function between vector spaces $h\colon V \to W$ that preserves the operations of addition

$$\text{if } \vec{v}_1, \vec{v}_2 \in V \text{ then } h(\vec{v}_1 + \vec{v}_2) = h(\vec{v}_1) + h(\vec{v}_2)$$

and scalar multiplication

$$\text{if } \vec{v} \in V \text{ and } r \in \mathbb{R} \text{ then } h(r \cdot \vec{v}) = r \cdot h(\vec{v})$$

is a *homomorphism* or *linear map*.

1.2 Example The projection map $\pi\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \end{pmatrix}$$

is a homomorphism. It preserves addition

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

and scalar multiplication.

$$\pi(r \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}) = \pi(\begin{pmatrix} rx_1 \\ ry_1 \\ rz_1 \end{pmatrix}) = \begin{pmatrix} rx_1 \\ ry_1 \end{pmatrix} = r \cdot \pi(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix})$$

This map is not an isomorphism since it is not one-to-one. For instance, both $\vec{0}$ and $\vec{e}_3$ in $\mathbb{R}^3$ are mapped to the zero vector in $\mathbb{R}^2$.

1.3 Example Of course, the domain and codomain might be other than spaces of column vectors. Both of these are homomorphisms.

(1) $f_1\colon \mathcal{P}_2 \to \mathcal{P}_3$ given by

$$a_0 + a_1x + a_2x^2 \;\mapsto\; a_0x + (a_1/2)x^2 + (a_2/3)x^3$$

(2) $f_2\colon M_{2\times2} \to \mathbb{R}$ given by

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

The verifications are straightforward.

1.4 Example Between any two spaces there is a *zero homomorphism*, mapping every vector in the domain to the zero vector in the codomain.

1.5 Example These two suggest why we use the term 'linear map'.

(1) The map $g\colon \mathbb{R}^3 \to \mathbb{R}$ given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{g}{\longmapsto} 3x + 2y - 4.5z$$

is linear (i.e., is a homomorphism). In contrast, the map $\hat{g}\colon \mathbb{R}^3 \to \mathbb{R}$ given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\hat{g}}{\longmapsto} 3x + 2y - 4.5z + 1$$

is not; for instance,

$$\hat{g}(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = 4 \quad \text{while} \quad \hat{g}(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}) + \hat{g}(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = 5$$

(to show that a map is not linear we need only produce one example of a linear combination that is not preserved).

(2) The first of these two maps $t_1, t_2\colon \mathbb{R}^3 \to \mathbb{R}^2$ is linear while the second is not.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{t_1}{\longmapsto} \begin{pmatrix} 5x - 2y \\ x + y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{t_2}{\longmapsto} \begin{pmatrix} 5x - 2y \\ xy \end{pmatrix}$$

Finding an example that the second fails to preserve structure is easy.

What distinguishes the homomorphisms is that the coordinate functions are linear combinations of the arguments. See also Exercise 23.

Obviously, any isomorphism is a homomorphism — an isomorphism is a homomorphism that is also a correspondence. So, one way to think of the 'homomorphism' idea is that it is a generalization of 'isomorphism', motivated by the observation that many of the properties of isomorphisms have only to do with the map's structure preservation property and not to do with it being a correspondence. As examples, these two results from the prior section do not use one-to-one-ness or onto-ness in their proof, and therefore apply to any homomorphism.

1.6 Lemma A homomorphism sends a zero vector to a zero vector.

1.7 Lemma Each of these is a necessary and sufficient condition for $f\colon V \to W$ to be a homomorphism.

(1) $f(c_1 \cdot \vec{v}_1 + c_2 \cdot \vec{v}_2) = c_1 \cdot f(\vec{v}_1) + c_2 \cdot f(\vec{v}_2)$ for any $c_1, c_2 \in \mathbb{R}$ and $\vec{v}_1, \vec{v}_2 \in V$

(2) $f(c_1 \cdot \vec{v}_1 + \cdots + c_n \cdot \vec{v}_n) = c_1 \cdot f(\vec{v}_1) + \cdots + c_n \cdot f(\vec{v}_n)$ for any $c_1, \ldots, c_n \in \mathbb{R}$ and $\vec{v}_1, \ldots, \vec{v}_n \in V$

Part (1) is often used to check that a function is linear.

1.8 Example The map $f\colon \mathbb{R}^2 \to \mathbb{R}^4$ given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \overset{f}{\longmapsto} \begin{pmatrix} x/2 \\ 0 \\ x+y \\ 3y \end{pmatrix}$$

satisfies (1) of the prior result

$$\begin{pmatrix} r_1(x_1/2) + r_2(x_2/2) \\ 0 \\ r_1(x_1+y_1) + r_2(x_2+y_2) \\ r_1(3y_1) + r_2(3y_2) \end{pmatrix} = r_1 \begin{pmatrix} x_1/2 \\ 0 \\ x_1+y_1 \\ 3y_1 \end{pmatrix} + r_2 \begin{pmatrix} x_2/2 \\ 0 \\ x_2+y_2 \\ 3y_2 \end{pmatrix}$$

and so it is a homomorphism.

However, some of the results that we have seen for isomorphisms fail to hold for homomorphisms in general. Consider the theorem that an isomorphism between spaces gives a correspondence between their bases. Homomorphisms do not give any such correspondence; Example 1.2 shows that there is no such correspondence, and another example is the zero map between any two nontrivial spaces. Instead, for homomorphisms a weaker but still very useful result holds.

1.9 Theorem A homomorphism is determined by its action on a basis. That is, if $\langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ is a basis of a vector space V and $\vec{w}_1, \ldots, \vec{w}_n$ are (perhaps not distinct) elements of a vector space W then there exists a homomorphism from V to W sending $\vec{\beta}_1$ to $\vec{w}_1$, …, and $\vec{\beta}_n$ to $\vec{w}_n$, and that homomorphism is unique.

Proof. We will define the map by associating $\vec{\beta}_1$ with $\vec{w}_1$, etc., and then extending linearly to all of the domain. That is, where $\vec{v} = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n$, the map $h\colon V \to W$ is given by $h(\vec{v}) = c_1\vec{w}_1 + \cdots + c_n\vec{w}_n$. This is well-defined because, with respect to the basis, the representation of each domain vector $\vec{v}$ is unique.

This map is a homomorphism since it preserves linear combinations; where $\vec{v_1} = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n$ and $\vec{v_2} = d_1\vec{\beta}_1 + \cdots + d_n\vec{\beta}_n$, we have this.

$$\begin{aligned} h(r_1\vec{v}_1 + r_2\vec{v}_2) &= h((r_1c_1 + r_2d_1)\vec{\beta}_1 + \cdots + (r_1c_n + r_2d_n)\vec{\beta}_n) \\ &= (r_1c_1 + r_2d_1)\vec{w}_1 + \cdots + (r_1c_n + r_2d_n)\vec{w}_n \\ &= r_1h(\vec{v}_1) + r_2h(\vec{v}_2) \end{aligned}$$

And, this map is unique since if $\hat{h}\colon V \to W$ is another homomorphism such that $\hat{h}(\vec{\beta}_i) = \vec{w}_i$ for each i then h and $\hat{h}$ agree on all of the vectors in the domain.

$$\begin{aligned} \hat{h}(\vec{v}) &= \hat{h}(c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n) \\ &= c_1\hat{h}(\vec{\beta}_1) + \cdots + c_n\hat{h}(\vec{\beta}_n) \\ &= c_1\vec{w}_1 + \cdots + c_n\vec{w}_n \\ &= h(\vec{v}) \end{aligned}$$

Thus, h and $\hat{h}$ are the same map. QED

1.10 Example This result says that we can construct a homomorphism by fixing a basis for the domain and specifying where the map sends those basis vectors. For instance, if we specify a map $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ that acts on the standard basis $\mathcal{E}_2$ in this way

$$h(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad\text{and}\quad h(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} -4 \\ 4 \end{pmatrix}$$

then the action of h on any other member of the domain is also specified. For instance, the value of h on this argument

$$h(\begin{pmatrix} 3 \\ -2 \end{pmatrix}) = h(3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}) = 3 \cdot h(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) - 2 \cdot h(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 5 \\ -5 \end{pmatrix}$$

is a direct consequence of the value of h on the basis vectors.

Later in this chapter we shall develop a scheme, using matrices, that is convienent for computations like this one.

Just as the isomorphisms of a space with itself are useful and interesting, so too are the homomorphisms of a space with itself.

1.11 Definition A linear map from a space into itself $t\colon V \to V$ is a *linear transformation.*

1.12 Remark In this book we use 'linear transformation' only in the case where the codomain equals the domain, but it is widely used in other texts as a general synonym for 'homomorphism'.

1.13 Example The map on $\mathbb{R}^2$ that projects all vectors down to the x-axis

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$$

is a linear transformation.

1.14 Example The derivative map $d/dx\colon \mathcal{P}_n \to \mathcal{P}_n$

$$a_0 + a_1x + \cdots + a_nx^n \stackrel{d/dx}{\longmapsto} a_1 + 2a_2x + 3a_3x^2 + \cdots + na_nx^{n-1}$$

is a linear transformation, as this result from calculus notes: $d(c_1f + c_2g)/dx = c_1\,(df/dx) + c_2\,(dg/dx)$.

1.15 Example The matrix transpose map

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

is a linear transformation of $\mathcal{M}_{2\times2}$. Note that this transformation is one-to-one and onto, and so in fact it is an automorphism.

We finish this subsection about maps by recalling that we can linearly combine maps. For instance, for these maps from $\mathbb{R}^2$ to itself

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{f}{\longmapsto} \begin{pmatrix} 2x \\ 3x - 2y \end{pmatrix} \quad\text{and}\quad \begin{pmatrix} x \\ y \end{pmatrix} \stackrel{g}{\longmapsto} \begin{pmatrix} 0 \\ 5x \end{pmatrix}$$

the linear combination $5f - 2g$ is also a map from R^2 to itself.

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{5f-2g}{\longmapsto} \begin{pmatrix} 10x \\ 5x - 10y \end{pmatrix}$$

1.16 Lemma For vector spaces V and W, the set of linear functions from V to W is itself a vector space, a subspace of the space of all functions from V to W. It is denoted $\mathcal{L}(V, W)$.

Proof. This set is non-empty because it contains the zero homomorphism. So to show that it is a subspace we need only check that it is closed under linear combinations. Let $f, g\colon V \to W$ be linear. Then their sum is linear

$$\begin{aligned}(f + g)(c_1\vec{v}_1 + c_2\vec{v}_2) &= c_1f(\vec{v}_1) + c_2f(\vec{v}_2) + c_1g(\vec{v}_1) + c_2g(\vec{v}_2) \\ &= c_1\bigl(f + g\bigr)(\vec{v}_1) + c_2\bigl(f + g\bigr)(\vec{v}_2)\end{aligned}$$

and any scalar multiple is also linear.

$$\begin{aligned}(r \cdot f)(c_1\vec{v}_1 + c_2\vec{v}_2) &= r(c_1f(\vec{v}_1) + c_2f(\vec{v}_2)) \\ &= c_1(r \cdot f)(\vec{v}_1) + c_2(r \cdot f)(\vec{v}_2)\end{aligned}$$

Hence $\mathcal{L}(V, W)$ is a subspace. QED

We started this section by isolating the structure preservation property of isomorphisms. That is, we defined homomorphisms as a generalization of isomorphisms. Some of the properties that we studied for isomorphisms carried over unchanged, while others were adapted to this more general setting.

It would be a mistake, though, to view this new notion of homomorphism as derived from, or somehow secondary to, that of isomorphism. In the rest of this chapter we shall work mostly with homomorphisms, partly because any statement made about homomorphisms is automatically true about isomorphisms, but more because, while the isomorphism concept is perhaps more natural, experience shows that the homomorphism concept is actually more fruitful and more central to further progress.

Exercises

HW ✓ **1.17** Decide if each $h\colon \mathbb{R}^3 \to \mathbb{R}^2$ is linear.

(a) $h(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} x \\ x+y+z \end{pmatrix}$ **(b)** $h(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ **(c)** $h(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(d) $h(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} 2x+y \\ 3y-4z \end{pmatrix}$

✓ **1.18** Decide if each map $h\colon \mathcal{M}_{2\times 2} \to \mathbb{R}$ is linear.

(a) $h(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = a + d$

(b) $h(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = ad - bc$

(c) $h(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = 2a + 3b + c - d$

(d) $h(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = a^2 + b^2$

✓ **1.19** Show that these two maps are homomorphisms.

(a) $d/dx\colon \mathcal{P}_3 \to \mathcal{P}_2$ given by $a_0 + a_1x + a_2x^2 + a_3x^3$ maps to $a_1 + 2a_2x + 3a_3x^2$

(b) $\int\colon \mathcal{P}_2 \to \mathcal{P}_3$ given by $b_0 + b_1x + b_2x^2$ maps to $b_0x + (b_1/2)x^2 + (b_2/3)x^3$

Are these maps inverse to each other?

1.20 Is (perpendicular) projection from $\mathbb{R}^3$ to the xz-plane a homomorphism? Projection to the yz-plane? To the x-axis? The y-axis? The z-axis? Projection to the origin?

1.21 Show that, while the maps from Example 1.3 preserve linear operations, they are not isomorphisms.

1.22 Is an identity map a linear transformation?

✓ **1.23** Stating that a function is 'linear' is different than stating that its graph is a line.

(a) The function $f_1\colon \mathbb{R} \to \mathbb{R}$ given by $f_1(x) = 2x - 1$ has a graph that is a line. Show that it is not a linear function.

(b) The function $f_2\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto x + 2y$$

does not have a graph that is a line. Show that it is a linear function.

✓ **1.24** Part of the definition of a linear function is that it respects addition. Does a linear function respect subtraction?

1.25 Assume that h is a linear transformation of V and that $\langle\vec{\beta}_1,\dots,\vec{\beta}_n\rangle$ is a basis of V. Prove each statement.

(a) If $h(\vec{\beta}_i)=\vec{0}$ for each basis vector then h is the zero map.

(b) If $h(\vec{\beta}_i)=\vec{\beta}_i$ for each basis vector then h is the identity map.

(c) If there is a scalar r such that $h(\vec{\beta}_i)=r\cdot\vec{\beta}_i$ for each basis vector then $h(\vec{v})=r\cdot\vec{v}$ for all vectors in V.

✓ **1.26** Consider the vector space $\mathbb{R}^+$ where vector addition and scalar multiplication are not the ones inherited from $\mathbb{R}$ but rather are these: $a+b$ is the product of a and b, and $r\cdot a$ is the r-th power of a. (This was shown to be a vector space in an earlier exercise.) Verify that the natural logarithm map $\ln\colon\mathbb{R}^+\to\mathbb{R}$ is a homomorphism between these two spaces. Is it an isomorphism?

✓ **1.27** Consider this transformation of $\mathbb{R}^2$.

$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x/2\\y/3\end{pmatrix}$$

Find the image under this map of this ellipse.

$$\{\begin{pmatrix}x\\y\end{pmatrix}\bigm|(x^2/4)+(y^2/9)=1\}$$

✓ **1.28** Imagine a rope wound around the earth's equator so that it fits snugly (suppose that the earth is a sphere). How much extra rope must be added to raise the circle to a constant six feet off the ground?

✓ **1.29** Verify that this map $h\colon\mathbb{R}^3\to\mathbb{R}$

$$\begin{pmatrix}x\\y\\z\end{pmatrix}\mapsto\begin{pmatrix}x\\y\\z\end{pmatrix}\cdot\begin{pmatrix}3\\-1\\-1\end{pmatrix}=3x-y-z$$

is linear. Generalize.

1.30 Show that every homomorphism from $\mathbb{R}^1$ to $\mathbb{R}^1$ acts via multiplication by a scalar. Conclude that every nontrivial linear transformation of $\mathbb{R}^1$ is an isomorphism. Is that true for transformations of $\mathbb{R}^2$? $\mathbb{R}^n$?

1.31 **(a)** Show that for any scalars $a_{1,1},\dots,a_{m,n}$ this map $h\colon\mathbb{R}^n\to\mathbb{R}^m$ is a homomorphism.

$$\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}\mapsto\begin{pmatrix}a_{1,1}x_1+\dots+a_{1,n}x_n\\\vdots\\a_{m,1}x_1+\dots+a_{m,n}x_n\end{pmatrix}$$

(b) Show that for each i, the i-th derivative operator d^i/dx^i is a linear transformation of $\mathcal{P}_n$. Conclude that for any scalars $c_k,\dots,c_0$ this map is a linear transformation of that space.

$$f\mapsto\frac{d^k}{dx^k}f+c_{k-1}\frac{d^{k-1}}{dx^{k-1}}f+\dots+c_1\frac{d}{dx}f+c_0f$$

1.32 Lemma 1.16 shows that a sum of linear functions is linear and that a scalar multiple of a linear function is linear. Show also that a composition of linear functions is linear.

✓ **1.33** Where $f\colon V\to W$ is linear, suppose that $f(\vec{v}_1)=\vec{w}_1,\ \dots,\ f(\vec{v}_n)=\vec{w}_n$ for some vectors $\vec{w}_1,\ \dots,\ \vec{w}_n$ from W.

(a) If the set of $\vec{w}$'s is independent, must the set of $\vec{v}$'s also be independent?

(b) If the set of $\vec{v}$'s is independent, must the set of $\vec{w}$'s also be independent?
(c) If the set of $\vec{w}$'s spans W, must the set of $\vec{v}$'s span V?
(d) If the set of $\vec{v}$'s spans V, must the set of $\vec{w}$'s span W?

1.34 Generalize Example 1.15 by proving that the matrix transpose map is linear. What is the domain and codomain?

1.35 **(a)** Where $\vec{u}, \vec{v} \in \mathbb{R}^n$, the line segment connecting them is defined to be the set $\ell = \{t \cdot \vec{u} + (1-t) \cdot \vec{v} \mid t \in [0..1]\}$. Show that the image, under a homomorphism h, of the segment between $\vec{u}$ and $\vec{v}$ is the segment between $h(\vec{u})$ and $h(\vec{v})$.
(b) A subset of $\mathbb{R}^n$ is *convex* if, for any two points in that set, the line segment joining them lies entirely in that set. (The inside of a sphere is convex while the skin of a sphere is not.) Prove that linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ preserve the property of set convexity.

✓ **1.36** Let $h\colon \mathbb{R}^n \to \mathbb{R}^m$ be a homomorphism.
(a) Show that the image under h of a line in $\mathbb{R}^n$ is a (possibly degenerate) line in $\mathbb{R}^n$.
(b) What happens to a k-dimensional linear surface?

1.37 Prove that the restriction of a homomorphism to a subspace of its domain is another homomorphism.

1.38 Assume that $h\colon V \to W$ is linear.
(a) Show that the *rangespace* of this map $\{h(\vec{v}) \mid \vec{v} \in V\}$ is a subspace of the codomain W.
(b) Show that the *nullspace* of this map $\{\vec{v} \in V \mid h(\vec{v}) = \vec{0}_W\}$ is a subspace of the domain V.
(c) Show that if U is a subspace of the domain V then its image $\{h(\vec{u}) \mid \vec{u} \in U\}$ is a subspace of the codomain W. This generalizes the first item.
(d) Generalize the second item.

1.39 Consider the set of isomorphisms from a vector space to itself. Is this a subspace of the space $\mathcal{L}(V, V)$ of homomorphisms from the space to itself?

1.40 Does Theorem 1.9 need that $\langle\vec{\beta}_1, \dots, \vec{\beta}_n\rangle$ is a basis? That is, can we still get a well-defined and unique homomorphism if we drop either the condition that the set of $\vec{\beta}$'s be linearly independent, or the condition that it span the domain?

1.41 Let V be a vector space and assume that the maps $f_1, f_2\colon V \to \mathbb{R}^1$ are linear.
(a) Define a map $F\colon V \to \mathbb{R}^2$ whose component functions are the given linear ones.

$$\vec{v} \mapsto \begin{pmatrix} f_1(\vec{v}) \\ f_2(\vec{v}) \end{pmatrix}$$

Show that F is linear.
(b) Does the converse hold — is any linear map from V to $\mathbb{R}^2$ made up of two linear component maps to $\mathbb{R}^1$?
(c) Generalize.

II.2 Rangespace and Nullspace

Isomorphisms and homomorphisms both preserve structure. The difference is that homomorphisms needn't be onto and needn't be one-to-one. This means that homomorphisms are a more general kind of map — subject to fewer restrictions than isomorphisms. We will now examine what can happen with homomorphisms that the extra conditions forestall with isomorphisms.

We first consider the effect of dropping the onto requirement, of not requiring as part of the definition that a homomorphism be onto its codomain. Of course, being a function, a homomorphism is onto some set, namely its range. The next result says that this range set is a vector space.

2.1 Lemma Under a homomorphism, the image of any subspace of the domain is a subspace of the codomain. In particular, the image of the entire space, the range of the homomorphism, is a subspace of the codomain.

Proof. Let $h\colon V \to W$ be linear and let S be a subspace of the domain V. The image $h(S)$ is a subset of the codomain W. It is nonempty because S is nonempty and thus to show that $h(S)$ is a subspace of W we need only show that it is closed under linear combinations of two vectors. If $h(\vec{s}_1)$ and $h(\vec{s}_2)$ are members of $h(S)$ then $c_1\cdot h(\vec{s}_1)+c_2\cdot h(\vec{s}_2) = h(c_1\cdot\vec{s}_1)+h(c_2\cdot\vec{s}_2) = h(c_1\cdot\vec{s}_1+c_2\cdot\vec{s}_2)$ is also a member of $h(S)$ because it is the image of $c_1\cdot\vec{s}_1+c_2\cdot\vec{s}_2$ from S. QED

2.2 Definition The *rangespace* of a homomorphism $h\colon V \to W$ is

$$\mathscr{R}(h) = \{h(\vec{v}) \mid \vec{v} \in V\}$$

sometimes denoted $h(V)$. The dimension of the rangespace is the map's *rank*.

(We shall soon see the connection between the rank of a map and the rank of a matrix.)

2.3 Example Recall that the derivative map $d/dx\colon \mathcal{P}_3 \to \mathcal{P}_3$ given by $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_1 + 2a_2x + 3a_3x^2$ is linear. The rangespace $\mathscr{R}(d/dx)$ is the set of quadratic polynomials $\{r + sx + tx^2 \mid r, s, t \in \mathbb{R}\}$. Thus, the rank of this map is three.

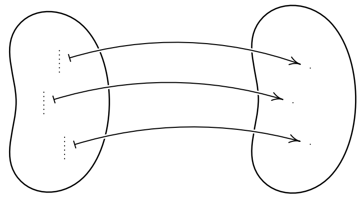
2.4 Example With this homomorphism $h\colon \mathcal{M}_{2\times 2} \to \mathcal{P}_3$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a + b + 2d) + 0x + cx^2 + cx^3$$

an image vector in the range can have any constant term, must have an x coefficient of zero, and must have the same coefficient of x^2 as of x^3. That is, the rangespace is $\mathscr{R}(h) = \{r + 0x + sx^2 + sx^3 \mid r, s \in \mathbb{R}\}$ and so the rank is two.

The prior result shows that, in passing from the definition of isomorphism to the more general definition of homomorphism, omitting the 'onto' requirement doesn't make an essential difference. Any homomorphism is onto its rangespace.

However, omitting the 'one-to-one' condition does make a difference. A homomorphism may have many elements of the domain that map to one element of the codomain. Below is a "bean " sketch of a many-to-one map between sets.* It shows three elements of the codomain that are each the image of many members of the domain.

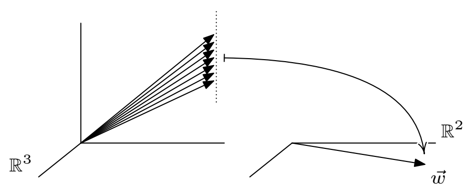


Recall that for any function $h\colon V \to W$, the set of elements of V that are mapped to $\vec{w} \in W$ is the *inverse image* $h^{-1}(\vec{w}) = \{\vec{v} \in V \mid h(\vec{v}) = \vec{w}\}$. Above, the three sets of many elements on the left are inverse images.

2.5 Example Consider the projection $\pi\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \end{pmatrix}$$

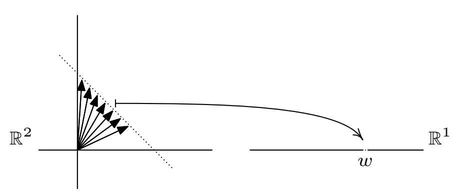
which is a homomorphism that is many-to-one. In this instance, an inverse image set is a vertical line of vectors in the domain.



2.6 Example This homomorphism $h\colon \mathbb{R}^2 \to \mathbb{R}^1$

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{h}{\longmapsto} x + y$$

is also many-to-one; for a fixed $w \in \mathbb{R}^1$, the inverse image $h^{-1}(w)$



is the set of plane vectors whose components add to w.

The above examples have only to do with the fact that we are considering functions, specifically, many-to-one functions. They show the inverse images as sets of vectors that are related to the image vector $\vec{w}$. But these are more than

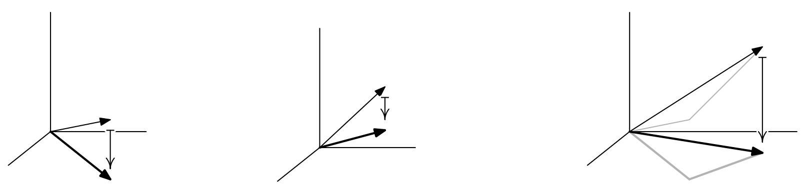
* More information on many-to-one maps is in the appendix.

just arbitrary functions, they are homomorphisms; what do the two preservation conditions say about the relationships?

In generalizing from isomorphisms to homomorphisms by dropping the one-to-one condition, we lose the property that we've stated intuitively as: the domain is "the same as" the range. That is, we lose that the domain corresponds perfectly to the range in a one-vector-by-one-vector way. What we shall keep, as the examples below illustrate, is that a homomorphism describes a way in which the domain is "like", or "analgous to", the range.

2.7 Example We think of $\mathbb{R}^3$ as being like $\mathbb{R}^2$, except that vectors have an extra component. That is, we think of the vector with components x, y, and z as like the vector with components x and y. In defining the projection map π, we make precise which members of the domain we are thinking of as related to which members of the codomain.

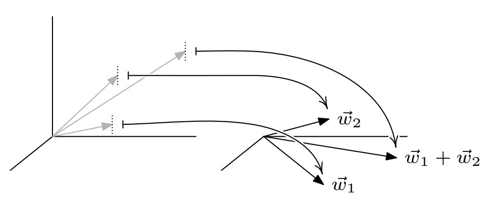
Understanding in what way the preservation conditions in the definition of homomorphism show that the domain elements are like the codomain elements is easiest if we draw $\mathbb{R}^2$ as the xy-plane inside of $\mathbb{R}^3$. (Of course, $\mathbb{R}^2$ is a set of two-tall vectors while the xy-plane is a set of three-tall vectors with a third component of zero, but there is an obvious correspondence.) Then, $\pi(\vec{v})$ is the "shadow" of $\vec{v}$ in the plane and the preservation of addition property says that



$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \text{ above } \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \text{ plus } \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \text{ above } \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \text{ equals } \begin{pmatrix} x_1 + y_1 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix} \text{ above } \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}$$

Briefly, the shadow of a sum $\pi(\vec{v}_1 + \vec{v}_2)$ equals the sum of the shadows $\pi(\vec{v}_1) + \pi(\vec{v}_2)$. (Preservation of scalar multiplication has a similar interpretation.)

Redrawing by separating the two spaces, moving the codomain $\mathbb{R}^2$ to the right, gives an uglier picture but one that is more faithful to the "bean" sketch.



Again in this drawing, the vectors that map to $\vec{w}_1$ lie in the domain in a vertical line (only one such vector is shown, in gray). Call any such member of this inverse image a "$\vec{w}_1$ vector". Similarly, there is a vertical line of "$\vec{w}_2$ vectors" and a vertical line of "$\vec{w}_1 + \vec{w}_2$ vectors". Now, π has the property that if $\pi(\vec{v}_1) = \vec{w}_1$ and $\pi(\vec{v}_2) = \vec{w}_2$ then $\pi(\vec{v}_1 + \vec{v}_2) = \pi(\vec{v}_1) + \pi(\vec{v}_2) = \vec{w}_1 + \vec{w}_2$.

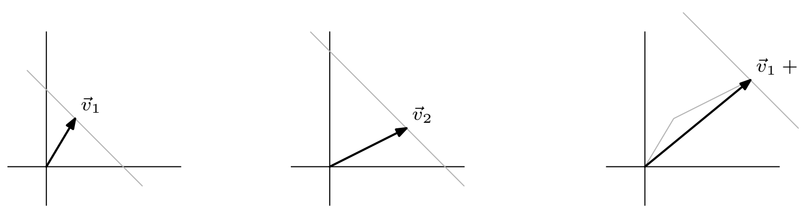
This says that the vector classes add, in the sense that any $\vec{w}_1$ vector plus any $\vec{w}_2$ vector equals a $\vec{w}_1 + \vec{w}_2$ vector, (A similar statement holds about the classes under scalar multiplication.)

Thus, although the two spaces $\mathbb{R}^3$ and $\mathbb{R}^2$ are not isomorphic, π describes a way in which they are alike: vectors in $\mathbb{R}^3$ add as do the associated vectors in $\mathbb{R}^2$ — vectors add as their shadows add.

2.8 Example A homomorphism can be used to express an analogy between spaces that is more subtle than the prior one. For the map

$$\begin{pmatrix} x \\ y \end{pmatrix} \overset{h}{\longmapsto} x + y$$

from Example 2.8 fix two numbers w_1, w_2 in the range $\mathbb{R}$. A $\vec{v}_1$ that maps to w_1 has components that add to w_1, that is, the inverse image $h^{-1}(w_1)$ is the set of vectors with endpoint on the diagonal line $x + y = w_1$. Call these the "w_1 vectors". Similarly, we have the "w_2 vectors" and the "$w_1 + w_2$ vectors". Then the addition preservation property says that



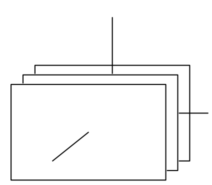
a "w_1 vector" plus a "w_2 vector" equals a "$w_1 + w_2$ vector".

Restated, if a w_1 vector is added to a w_2 vector then the result is mapped by h to a $w_1 + w_2$ vector. Briefly, the image of a sum is the sum of the images. Even more briefly, $h(\vec{v}_1 + \vec{v}_2) = h(\vec{v}_1) + h(\vec{v}_2)$. (The preservation of scalar multiplication condition has a similar restatement.)

2.9 Example The inverse images can be structures other than lines. For the linear map $h\colon \mathbb{R}^3 \to \mathbb{R}^2$

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

the inverse image sets are planes $x = 0$, $x = 1$, etc., perpendicular to the x-axis.



We won't describe how every homomorphism that we will use is an analogy because the formal sense that we make of "alike in that ..." is 'a homomorphism exists such that ...'. Nonetheless, the idea that a homomorphism between two spaces expresses how the domain's vectors fall into classes that act like the the range's vectors is a good way to view homomorphisms.

Another reason that we won't treat all of the homomorphisms that we see as above is that many vector spaces are hard to draw (e.g., a space of polynomials). However, there is nothing bad about gaining insights from those spaces that we are able to draw, especially when those insights extend to all vector spaces. We derive two such insights from the three examples 2.7, 2.8, and 2.9.

First, in all three examples, the inverse images are lines or planes, that is, linear surfaces. In particular, the inverse image of the range's zero vector is a line or plane through the origin — a subspace of the domain.

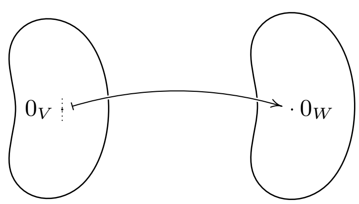
2.10 Lemma For any homomorphism, the inverse image of a subspace of the range is a subspace of the domain. In particular, the inverse image of the trivial subspace of the range is a subspace of the domain.

Proof. Let $h\colon V\to W$ be a homomorphism and let S be a subspace of the rangespace h. Consider $h^{-1}(S)=\{\vec{v}\in V \bigm| h(\vec{v})\in S\}$, the inverse image of the set S. It is nonempty because it contains $\vec{0}_V$, since $h(\vec{0}_V)=\vec{0}_W$, which is an element S, as S is a subspace. To show that $h^{-1}(S)$ is closed under linear combinations, let $\vec{v}_1$ and $\vec{v}_2$ be elements, so that $h(\vec{v}_1)$ and $h(\vec{v}_2)$ are elements of S, and then $c_1\vec{v}_1+c_2\vec{v}_2$ is also in the inverse image because $h(c_1\vec{v}_1+c_2\vec{v}_2)=c_1h(\vec{v}_1)+c_2h(\vec{v}_2)$ is a member of the subspace S. QED

2.11 Definition The *nullspace* or *kernel* of a linear map $h\colon V\to W$ is the inverse image of 0_W

$$\mathscr{N}(h)=h^{-1}(\vec{0}_W)=\{\vec{v}\in V \bigm| h(\vec{v})=\vec{0}_W\}.$$

The dimension of the nullspace is the map's *nullity*.



2.12 Example The map from Example 2.3 has this nullspace $\mathscr{N}(d/dx)=\{a_0+0x+0x^2+0x^3 \bigm| a_0\in\mathbb{R}\}$.

2.13 Example The map from Example 2.4 has this nullspace.

$$\mathscr{N}(h)=\{\begin{pmatrix}a&b\\0&-(a+b)/2\end{pmatrix} \bigm| a,b\in\mathbb{R}\}$$

Now for the second insight from the above pictures. In Example 2.7, each of the vertical lines is squashed down to a single point — π, in passing from the domain to the range, takes all of these one-dimensional vertical lines and "zeroes them out", leaving the range one dimension smaller than the domain. Similarly, in Example 2.8, the two-dimensional domain is mapped to a one-dimensional range by breaking the domain into lines (here, they are diagonal lines), and compressing each of those lines to a single member of the range. Finally, in Example 2.9, the domain breaks into planes which get "zeroed out", and so the map starts with a three-dimensional domain but ends with a one-dimensional range — this map "subtracts" two from the dimension. (Notice that, in this third example, the codomain is two-dimensional but the range of the map is only one-dimensional, and it is the dimension of the range that is of interest.)

2.14 Theorem A linear map's rank plus its nullity equals the dimension of its domain.

Proof. Let $h\colon V\to W$ be linear and let $B_N=\langle\vec{\beta}_1,\dots,\vec{\beta}_k\rangle$ be a basis for the nullspace. Extend that to a basis $B_V=\langle\vec{\beta}_1,\dots,\vec{\beta}_k,\vec{\beta}_{k+1},\dots,\vec{\beta}_n\rangle$ for the entire domain. We shall show that $B_R=\langle h(\vec{\beta}_{k+1}),\dots,h(\vec{\beta}_n)\rangle$ is a basis for the rangespace. Then counting the size of these bases gives the result.

To see that B_R is linearly independent, consider the equation $c_{k+1}h(\vec{\beta}_{k+1})+\cdots+c_nh(\vec{\beta}_n)=\vec{0}_W$. This gives that $h(c_{k+1}\vec{\beta}_{k+1}+\cdots+c_n\vec{\beta}_n)=\vec{0}_W$ and so $c_{k+1}\vec{\beta}_{k+1}+\cdots+c_n\vec{\beta}_n$ is in the nullspace of h. As B_N is a basis for this nullspace, there are scalars $c_1,\dots,c_k\in\mathbb{R}$ satisfying this relationship.

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

But B_V is a basis for V so each scalar equals zero. Therefore B_R is linearly independent.

To show that B_R spans the rangespace, consider $h(\vec{v})\in\mathscr{R}(h)$ and write $\vec{v}$ as a linear combination $\vec{v}=c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n$ of members of B_V. This gives $h(\vec{v})=h(c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n)=c_1h(\vec{\beta}_1)+\cdots+c_kh(\vec{\beta}_k)+c_{k+1}h(\vec{\beta}_{k+1})+\cdots+c_nh(\vec{\beta}_n)$ and since $\vec{\beta}_1$, ..., $\vec{\beta}_k$ are in the nullspace, we have that $h(\vec{v})=\vec{0}+\cdots+\vec{0}+c_{k+1}h(\vec{\beta}_{k+1})+\cdots+c_nh(\vec{\beta}_n)$. Thus, $h(\vec{v})$ is a linear combination of members of B_R, and so B_R spans the space. QED

2.15 Example Where $h\colon\mathbb{R}^3\to\mathbb{R}^4$ is

$$\begin{pmatrix}x\\y\\z\end{pmatrix}\stackrel{h}{\longmapsto}\begin{pmatrix}x\\0\\y\\0\end{pmatrix}$$

the rangespace and nullspace are

$$\mathscr{R}(h)=\{\begin{pmatrix}a\\0\\b\\0\end{pmatrix}\bigm| a,b\in\mathbb{R}\}\quad\text{and}\quad\mathscr{N}(h)=\{\begin{pmatrix}0\\0\\z\end{pmatrix}\bigm| z\in\mathbb{R}\}$$

and so the rank of h is two while the nullity is one.

2.16 Example If $t\colon \mathbb{R} \to \mathbb{R}$ is the linear transformation $x \mapsto -4x$, then the range is $\mathscr{R}(t) = \mathbb{R}^1$, and so the rank of t is one and the nullity is zero.

2.17 Corollary The rank of a linear map is less than or equal to the dimension of the domain. Equality holds if and only if the nullity of the map is zero.

We know that an isomorphism exists between two spaces if and only if their dimensions are equal. Here we see that for a homomorphism to exist, the dimension of the range must be less than or equal to the dimension of the domain. For instance, there is no homomorphism from $\mathbb{R}^2$ onto $\mathbb{R}^3$. There are many homomorphisms from $\mathbb{R}^2$ into $\mathbb{R}^3$, but none is onto all of three-space.

The rangespace of a linear map can be of dimension strictly less than the dimension of the domain (Example 2.3's derivative transformation on $\mathcal{P}_3$ has a domain of dimension four but a range of dimension three). Thus, under a homomorphism, linearly independent sets in the domain may map to linearly dependent sets in the range (for instance, the derivative sends $\{1, x, x^2, x^3\}$ to $\{0, 1, 2x, 3x^2\}$). That is, under a homomorphism, independence may be lost. In contrast, dependence stays.

2.18 Lemma Under a linear map, the image of a linearly dependent set is linearly dependent.

Proof. Suppose that $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}_V$, with some c_i nonzero. Then, because $h(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = c_1h(\vec{v}_1) + \cdots + c_nh(\vec{v}_n)$ and because $h(\vec{0}_V) = \vec{0}_W$, we have that $c_1h(\vec{v}_1) + \cdots + c_nh(\vec{v}_n) = \vec{0}_W$ with some nonzero c_i. QED

When is independence not lost? One obvious sufficient condition is when the homomorphism is an isomorphism. This condition is also necessary; see Exercise 35. We will finish this subsection comparing homomorphisms with isomorphisms by observing that a one-to-one homomorphism is an isomorphism from its domain onto its range.

2.19 Definition A linear map that is one-to-one is *nonsingular*.

(In the next section we will see the connection between this use of 'nonsingular' for maps and its familiar use for matrices.)

2.20 Example This nonsingular homomorphism $\iota\colon \mathbb{R}^2 \to \mathbb{R}^3$

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{\iota}{\longmapsto} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

gives the obvious correspondence between $\mathbb{R}^2$ and the xy-plane inside of $\mathbb{R}^3$.

The prior observation allows us to adapt some results about isomorphisms to this setting.

2.21 Theorem In an n-dimensional vector space V, these

(1) h is nonsingular, that is, one-to-one

(2) h has a linear inverse

(3) $\mathscr{N}(h) = \{\vec{0}\,\}$, that is, $\text{nullity}(h) = 0$

(4) $\text{rank}(h) = n$

(5) if $\langle \vec{\beta}_1, \dots, \vec{\beta}_n \rangle$ is a basis for V then $\langle h(\vec{\beta}_1), \dots, h(\vec{\beta}_n) \rangle$ is a basis for $\mathscr{R}(h)$

are equivalent statements about a linear map $h\colon V \to W$.

Proof. We will first show that (1) $\iff$ (2). We will then show that (1) $\implies$ (3) $\implies$ (4) $\implies$ (5) $\implies$ (2).

For (1) $\implies$ (2), suppose that the linear map h is one-to-one, and so has an inverse. The domain of that inverse is the range of h and so a linear combination of two members of that domain has the form $c_1 h(\vec{v}_1) + c_2 h(\vec{v}_2)$. On that combination, the inverse h^{-1} gives this.

$$\begin{aligned}
h^{-1}(c_1 h(\vec{v}_1) + c_2 h(\vec{v}_2)) &= h^{-1}(h(c_1\vec{v}_1 + c_2\vec{v}_2)) \\
&= h^{-1} \circ h\,(c_1\vec{v}_1 + c_2\vec{v}_2) \\
&= c_1\vec{v}_1 + c_2\vec{v}_2 \\
&= c_1 h^{-1} \circ h\,(\vec{v}_1) + c_2 h^{-1} \circ h\,(\vec{v}_2) \\
&= c_1 \cdot h^{-1}(h(\vec{v}_1)) + c_2 \cdot h^{-1}(h(\vec{v}_2))
\end{aligned}$$

Thus the inverse of a one-to-one linear map is automatically linear. But this also gives the (1) $\implies$ (2) implication, because the inverse itself must be one-to-one.

Of the remaining implications, (1) $\implies$ (3) holds because any homomorphism maps $\vec{0}_V$ to $\vec{0}_W$, but a one-to-one map sends at most one member of V to $\vec{0}_W$.

Next, (3) $\implies$ (4) is true since rank plus nullity equals the dimension of the domain.

For (4) $\implies$ (5), to show that $\langle h(\vec{\beta}_1), \dots, h(\vec{\beta}_n) \rangle$ is a basis for the rangespace we need only show that it is a spanning set, because by assumption the range has dimension n. Consider $h(\vec{v}) \in \mathscr{R}(h)$. Expressing $\vec{v}$ as a linear combination of basis elements produces $h(\vec{v}) = h(c_1\vec{\beta}_1 + c_2\vec{\beta}_2 + \cdots + c_n\vec{\beta}_n)$, which gives that $h(\vec{v}) = c_1 h(\vec{\beta}_1) + \cdots + c_n h(\vec{\beta}_n)$, as desired.

Finally, for the (5) $\implies$ (2) implication, assume that $\langle \vec{\beta}_1, \dots, \vec{\beta}_n \rangle$ is a basis for V so that $\langle h(\vec{\beta}_1), \dots, h(\vec{\beta}_n) \rangle$ is a basis for $\mathscr{R}(h)$. Then every $\vec{w} \in \mathscr{R}(h)$ a the unique representation $\vec{w} = c_1 h(\vec{\beta}_1) + \cdots + c_n h(\vec{\beta}_n)$. Define a map from $\mathscr{R}(h)$ to V by

$$\vec{w} \;\mapsto\; c_1\vec{\beta}_1 + c_2\vec{\beta}_2 + \cdots + c_n\vec{\beta}_n$$

(uniqueness of the representation makes this well-defined). Checking that it is linear and that it is the inverse of h are easy. QED

We've now seen that a linear map shows how the structure of the domain is like that of the range. Such a map can be thought to organize the domain space into inverse images of points in the range. In the special case that the map is one-to-one, each inverse image is a single point and the map is an isomorphism between the domain and the range.

Exercises

HW ✓ **2.22** Let $h\colon \mathcal{P}_3 \to \mathcal{P}_4$ be given by $p(x) \mapsto x \cdot p(x)$. Which of these are in the nullspace? Which are in the rangespace?
(a) x^3 **(b)** 0 **(c)** 7 **(d)** $12x - 0.5x^3$ **(e)** $1 + 3x^2 - x^3$

HW ✓ **2.23** Find the nullspace, nullity, rangespace, and rank of each map.
(a) $h\colon \mathbb{R}^2 \to \mathcal{P}_3$ given by

$$\begin{pmatrix} a \\ b \end{pmatrix} \mapsto a + ax + ax^2$$

(b) $h\colon \mathcal{M}_{2\times 2} \to \mathbb{R}$ given by

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

(c) $h\colon \mathcal{M}_{2\times 2} \to \mathcal{P}_2$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + b + c + dx^2$$

(d) the zero map $Z\colon \mathbb{R}^3 \to \mathbb{R}^4$

✓ **2.24** Find the nullity of each map.
(a) $h\colon \mathbb{R}^5 \to \mathbb{R}^8$ of rank five **(b)** $h\colon \mathcal{P}_3 \to \mathcal{P}_3$ of rank one
(c) $h\colon \mathbb{R}^6 \to \mathbb{R}^3$, an onto map **(d)** $h\colon \mathcal{M}_{3\times 3} \to \mathcal{M}_{3\times 3}$, onto

✓ **2.25** What is the nullspace of the differentiation transformation $d/dx\colon \mathcal{P}_n \to \mathcal{P}_n$? What is the nullspace of the second derivative, as a transformation of $\mathcal{P}_n$? The k-th derivative?

2.26 Example 2.7 restates the first condition in the definition of homomorphism as 'the shadow of a sum is the sum of the shadows'. Restate the second condition in the same style.

2.27 For the homomorphism $h\colon \mathcal{P}_3 \to \mathcal{P}_3$ given by $h(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0 + (a_0 + a_1)x + (a_2 + a_3)x^3$ find these.
(a) $\mathscr{N}(h)$ **(b)** $h^{-1}(2 - x^3)$ **(c)** $h^{-1}(1 + x^2)$

✓ **2.28** For the map $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(\begin{pmatrix} x \\ y \end{pmatrix}) = 2x + y$$

sketch these inverse image sets: $f^{-1}(-3)$, $f^{-1}(0)$, and $f^{-1}(1)$.

✓ **2.29** Each of these transformations of $\mathcal{P}_3$ is nonsingular. Find the inverse function of each.
(a) $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_0 + a_1x + 2a_2x^2 + 3a_3x^3$
(b) $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_0 + a_2x + a_1x^2 + a_3x^3$
(c) $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto a_1 + a_2x + a_3x^2 + a_0x^3$
(d) $a_0+a_1x+a_2x^2+a_3x^3 \mapsto a_0+(a_0+a_1)x+(a_0+a_1+a_2)x^2+(a_0+a_1+a_2+a_3)x^3$

2.30 Describe the nullspace and rangespace of a transformation given by $\vec{v} \mapsto 2\vec{v}$.

2.31 List all pairs $(\text{rank}(h), \text{nullity}(h))$ that are possible for linear maps from $\mathbb{R}^5$ to $\mathbb{R}^3$.

2.32 Does the differentiation map $d/dx\colon \mathcal{P}_n \to \mathcal{P}_n$ have an inverse?

✓ **2.33** Find the nullity of the map $h\colon \mathcal{P}_n \to \mathbb{R}$ given by

$$a_0 + a_1x + \cdots + a_nx^n \mapsto \int_{x=0}^{x=1} a_0 + a_1x + \cdots + a_nx^n \, dx.$$

2.34 **(a)** Prove that a homomorphism is onto if and only if its rank equals the dimension of its codomain.
(b) Conclude that a homomorphism between vector spaces with the same dimension is one-to-one if and only if it is onto.

2.35 Show that a linear map is nonsingular if and only if it preserves linear independence.

2.36 Corollary 2.17 says that for there to be an onto homomorphism from a vector space V to a vector space W, it is necessary that the dimension of W be less than or equal to the dimension of V. Prove that this condition is also sufficient; use Theorem 1.9 to show that if the dimension of W is less than or equal to the dimension of V, then there is a homomorphism from V to W that is onto.

2.37 Let $h\colon V \to \mathbb{R}$ be a homomorphism, but not the zero homomorphism. Prove that if $\langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ is a basis for the nullspace and if $\vec{v} \in V$ is not in the nullspace then $\langle \vec{v}, \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ is a basis for the entire domain V.

✓ **2.38** Recall that the nullspace is a subset of the domain and the rangespace is a subset of the codomain. Are they necessarily distinct? Is there a homomorphism that has a nontrivial intersection of its nullspace and its rangespace?

2.39 Prove that the image of a span equals the span of the images. That is, where $h\colon V \to W$ is linear, prove that if S is a subset of V then $h([S])$ equals $[h(S)]$. This generalizes Lemma 2.1 since it shows that if U is any subspace of V then its image $\{h(\vec{u}) \mid \vec{u} \in U\}$ is a subspace of W, because the span of the set U is U.

✓ **2.40** **(a)** Prove that for any linear map $h\colon V \to W$ and any $\vec{w} \in W$, the set $h^{-1}(\vec{w})$ has the form

$$\{\vec{v} + \vec{n} \mid \vec{n} \in \mathscr{N}(h)\}$$

for $\vec{v} \in V$ with $h(\vec{v}) = \vec{w}$ (if h is not onto then this set may be empty). Such a set is a *coset* of $\mathscr{N}(h)$ and is denoted $\vec{v} + \mathscr{N}(h)$.
(b) Consider the map $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{t}{\longmapsto} \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

for some scalars a, b, c, and d. Prove that t is linear.
(c) Conclude from the prior two items that for any linear system of the form

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$$

the solution set can be written (the vectors are members of $\mathbb{R}^2$)

$$\{\vec{p} + \vec{h} \mid \vec{h} \text{ satisfies the associated homogeneous system}\}$$

where $\vec{p}$ is a particular solution of that linear system (if there is no particular solution then the above set is empty).
(d) Show that this map $h\colon \mathbb{R}^n \to \mathbb{R}^m$ is linear

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{pmatrix}$$

for any scalars $a_{1,1}, \ldots, a_{m,n}$. Extend the conclusion made in the prior item.

(e) Show that the k-th derivative map is a linear transformation of $\mathcal{P}_n$ for each k. Prove that this map is a linear transformation of that space

$$f \mapsto \frac{d^k}{dx^k}f + c_{k-1}\frac{d^{k-1}}{dx^{k-1}}f + \cdots + c_1\frac{d}{dx}f + c_0 f$$

for any scalars c_k, ..., c_0. Draw a conclusion as above.

2.41 Prove that for any transformation $t\colon V \to V$ that is rank one, the map given by composing the operator with itself $t \circ t\colon V \to V$ satisfies $t \circ t = r \cdot t$ for some real number r.

2.42 Show that for any space V of dimension n, the *dual space*

$$\mathcal{L}(V, \mathbb{R}) = \{h\colon V \to \mathbb{R} \bigm| h \text{ is linear}\}$$

is isomorphic to $\mathbb{R}^n$. It is often denoted V^*. Conclude that $V^* \cong V$.

2.43 Show that any linear map is the sum of maps of rank one.

2.44 Is 'is homomorphic to' an equivalence relation? (*Hint:* the difficulty is to decide on an appropriate meaning for the quoted phrase.)

2.45 Show that the rangespaces and nullspaces of powers of linear maps $t\colon V \to V$ form descending

$$V \supseteq \mathscr{R}(t) \supseteq \mathscr{R}(t^2) \supseteq \ldots$$

and ascending

$$\{\vec{0}\} \subseteq \mathscr{N}(t) \subseteq \mathscr{N}(t^2) \subseteq \ldots$$

chains. Also show that if k is such that $\mathscr{R}(t^k) = \mathscr{R}(t^{k+1})$ then all following rangespaces are equal: $\mathscr{R}(t^k) = \mathscr{R}(t^{k+1}) = \mathscr{R}(t^{k+2}) \ldots$. Similarly, if $\mathscr{N}(t^k) = \mathscr{N}(t^{k+1})$ then $\mathscr{N}(t^k) = \mathscr{N}(t^{k+1}) = \mathscr{N}(t^{k+2}) = \ldots$.

III Computing Linear Maps

The prior section shows that a linear map is determined by its action on a basis. In fact, the equation

$$h(\vec{v}) = h(c_1 \cdot \vec{\beta}_1 + \cdots + c_n \cdot \vec{\beta}_n) = c_1 \cdot h(\vec{\beta}_1) + \cdots + c_n \cdot h(\vec{\beta}_n)$$

shows that, if we know the value of the map on the vectors in a basis, then we can compute the value of the map on any vector $\vec{v}$ at all. We just need to findi the c's to express $\vec{v}$ with respect to the basis.

This section gives the scheme that computes, from the representation of a vector in the domain $\mathrm{Rep}_B(\vec{v})$, the representation of that vector's image in the codomain $\mathrm{Rep}_D(h(\vec{v}))$, using the representations of $h(\vec{\beta}_1)$, ..., $h(\vec{\beta}_n)$.

III.1 Representing Linear Maps with Matrices

1.1 Example Consider a map h with domain $\mathbb{R}^2$ and codomain $\mathbb{R}^3$ (fixing

$$B = \langle \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix} \rangle \quad \text{and} \quad D = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \rangle$$

as the bases for these spaces) that is determined by this action on the vectors in the domain's basis.

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} \stackrel{h}{\longmapsto} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 4 \end{pmatrix} \stackrel{h}{\longmapsto} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

To compute the action of this map on any vector at all from the domain, we first express $h(\vec{\beta}_1)$ and $h(\vec{\beta}_2)$ with respect to the codomain's basis:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{so} \quad \mathrm{Rep}_D(h(\vec{\beta}_1)) = \begin{pmatrix} 0 \\ -1/2 \\ 1 \end{pmatrix}_D$$

and

$$\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{so} \quad \mathrm{Rep}_D(h(\vec{\beta}_2)) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}_D$$

(these are easy to check). Then, as described in the preamble, for any member $\vec{v}$ of the domain, we can express the image $h(\vec{v})$ in terms of the $h(\vec{\beta})$'s.

$$\begin{aligned}
h(\vec{v}) &= h(c_1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 1 \\ 4 \end{pmatrix}) \\
&= c_1 \cdot h(\begin{pmatrix} 2 \\ 0 \end{pmatrix}) + c_2 \cdot h(\begin{pmatrix} 1 \\ 4 \end{pmatrix}) \\
&= c_1 \cdot (0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}) + c_2 \cdot (1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}) \\
&= (0c_1 + 1c_2) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-\frac{1}{2}c_1 - 1c_2) \cdot \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} + (1c_1 + 0c_2) \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}$$

Thus,

$$\text{with } \mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \text{ then } \mathrm{Rep}_D(\,h(\vec{v})\,) = \begin{pmatrix} 0c_1 + 1c_2 \\ -(1/2)c_1 - 1c_2 \\ 1c_1 + 0c_2 \end{pmatrix}.$$

For instance,

$$\text{with } \mathrm{Rep}_B(\begin{pmatrix} 4 \\ 8 \end{pmatrix}) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_B \text{ then } \mathrm{Rep}_D(\,h(\begin{pmatrix} 4 \\ 8 \end{pmatrix})\,) = \begin{pmatrix} 2 \\ -5/2 \\ 1 \end{pmatrix}.$$

We will express computations like the one above with a matrix notation.

$$\begin{pmatrix} 0 & 1 \\ -1/2 & -1 \\ 1 & 0 \end{pmatrix}_{B,D} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}_B = \begin{pmatrix} 0c_1 + 1c_2 \\ (-1/2)c_1 - 1c_2 \\ 1c_1 + 0c_2 \end{pmatrix}_D$$

In the middle is the argument $\vec{v}$ to the map, represented with respect to the domain's basis B by a column vector with components c_1 and c_2. On the right is the value $h(\vec{v})$ of the map on that argument, represented with respect to the codomain's basis D by a column vector with components $0c_1 + 1c_2$, etc. The matrix on the left is the new thing. It consists of the coefficients from the vector on the right, 0 and 1 from the first row, $-1/2$ and -1 from the second row, and 1 and 0 from the third row.

This notation simply breaks the parts from the right, the coefficients and the c's, out separately on the left, into a vector that represents the map's argument and a matrix that we will take to represent the map itself.

1.2 Definition Suppose that V and W are vector spaces of dimensions n and m with bases B and D, and that $h\colon V \to W$ is a linear map. If

$$\mathrm{Rep}_D(h(\vec{\beta}_1)) = \begin{pmatrix} h_{1,1} \\ h_{2,1} \\ \vdots \\ h_{m,1} \end{pmatrix}_D \quad \ldots \quad \mathrm{Rep}_D(h(\vec{\beta}_n)) = \begin{pmatrix} h_{1,n} \\ h_{2,n} \\ \vdots \\ h_{m,n} \end{pmatrix}_D$$

then

$$\mathrm{Rep}_{B,D}(h) = \begin{pmatrix} h_{1,1} & h_{1,2} & \ldots & h_{1,n} \\ h_{2,1} & h_{2,2} & \ldots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \ldots & h_{m,n} \end{pmatrix}_{B,D}$$

is the *matrix representation of h with respect to B, D.*

Briefly, the vectors representing the $h(\vec{\beta})$'s are adjoined to make the matrix representing the map.

$$\mathrm{Rep}_{B,D}(h) = \left(\begin{array}{c|c|c} \vdots & & \vdots \\ \mathrm{Rep}_D(\,h(\vec{\beta}_1)\,) & \cdots & \mathrm{Rep}_D(\,h(\vec{\beta}_n)\,) \\ \vdots & & \vdots \end{array}\right)$$

Observe that the number of columns n of the matrix is the dimension of the domain of the map, and the number of rows m is the dimension of the codomain.

1.3 Example If $h\colon \mathbb{R}^3 \to \mathcal{P}_1$ is given by

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \overset{h}{\longmapsto} (2a_1 + a_2) + (-a_3)x$$

then where

$$B = \langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \rangle \quad \text{and} \quad D = \langle 1 + x, -1 + x \rangle$$

the action of h on B is given by

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \overset{h}{\longmapsto} -x \qquad \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \overset{h}{\longmapsto} 2 \qquad \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \overset{h}{\longmapsto} 4$$

and a simple calculation gives

$$\mathrm{Rep}_D(-x) = \begin{pmatrix} -1/2 \\ -1/2 \end{pmatrix}_D \quad \mathrm{Rep}_D(2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}_D \quad \mathrm{Rep}_D(4) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}_D$$

showing that this is the matrix representing h with respect to the bases.

$$\operatorname{Rep}_{B,D}(h) = \begin{pmatrix} -1/2 & 1 & 2 \\ -1/2 & -1 & -2 \end{pmatrix}_{B,D}$$

We will use lower case letters for a map, upper case for the matrix, and lower case again for the entries of the matrix. Thus for the map h, the matrix representing it is H, with entries $h_{i,j}$.

1.4 Theorem Assume that V and W are vector spaces of dimensions m and n with bases B and D, and that $h\colon V \to W$ is a linear map. If h is represented by

$$\operatorname{Rep}_{B,D}(h) = \begin{pmatrix} h_{1,1} & h_{1,2} & \ldots & h_{1,n} \\ h_{2,1} & h_{2,2} & \ldots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \ldots & h_{m,n} \end{pmatrix}_{B,D}$$

and $\vec{v} \in V$ is represented by

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

then the representation of the image of $\vec{v}$ is this.

$$\operatorname{Rep}_D(\,h(\vec{v})\,) = \begin{pmatrix} h_{1,1}c_1 + h_{1,2}c_2 + \cdots + h_{1,n}c_n \\ h_{2,1}c_1 + h_{2,2}c_2 + \cdots + h_{2,n}c_n \\ \vdots \\ h_{m,1}c_1 + h_{m,2}c_2 + \cdots + h_{m,n}c_n \end{pmatrix}_D$$

Proof. Exercise 28. QED

We will think of the matrix $\operatorname{Rep}_{B,D}(h)$ and the vector $\operatorname{Rep}_B(\vec{v})$ as combining to make the vector $\operatorname{Rep}_D(h(\vec{v}))$.

1.5 Definition The *matrix-vector product* of a $m{\times}n$ matrix and a $n{\times}1$ vector is this.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ & \vdots & & \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} a_{1,1}c_1 + a_{1,2}c_2 + \cdots + a_{1,n}c_n \\ a_{2,1}c_1 + a_{2,2}c_2 + \cdots + a_{2,n}c_n \\ \vdots \\ a_{m,1}c_1 + a_{m,2}c_2 + \cdots + a_{m,n}c_n \end{pmatrix}$$

The point of Definition 1.2 is to generalize Example 1.1, that is, the point of the definition is Theorem 1.4, that the matrix describes how to get from the representation of a domain vector with respect to the domain's basis to the representation of its image in the codomain with respect to the codomain's basis. With Definition 1.5, we can restate this as: application of a linear map is represented by the matrix-vector product of the map's representative and the vector's representative.

1.6 Example With the matrix from Example 1.3 we can calculate where that map sends this vector.

$$\vec{v}=\begin{pmatrix}4\\1\\0\end{pmatrix}$$

This vector is represented, with respect to the domain basis B, by

$$\mathrm{Rep}_B(\vec{v})=\begin{pmatrix}0\\1/2\\2\end{pmatrix}_B$$

and so this is the representation of the value $h(\vec{v})$ with respect to the codomain basis D.

$$\begin{aligned}\mathrm{Rep}_D(h(\vec{v}))&=\begin{pmatrix}-1/2&1&2\\-1/2&-1&-2\end{pmatrix}_{B,D}\begin{pmatrix}0\\1/2\\2\end{pmatrix}_B\\&=\begin{pmatrix}(-1/2)\cdot 0+1\cdot(1/2)+2\cdot 2\\(-1/2)\cdot 0-1\cdot(1/2)-2\cdot 2\end{pmatrix}_D=\begin{pmatrix}9/2\\-9/2\end{pmatrix}_D\end{aligned}$$

To find $h(\vec{v})$ itself, not its representation, take $(9/2)(1+x)-(9/2)(-1+x)=9$.

1.7 Example Let $\pi\colon\mathbb{R}^3\to\mathbb{R}^2$ be projection onto the xy-plane. To give a matrix representing this map, we first fix bases.

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

For each vector in the domain's basis, we find its image under the map.

$$\begin{pmatrix}1\\0\\0\end{pmatrix}\overset{\pi}{\longmapsto}\begin{pmatrix}1\\0\end{pmatrix}\qquad\begin{pmatrix}1\\1\\0\end{pmatrix}\overset{\pi}{\longmapsto}\begin{pmatrix}1\\1\end{pmatrix}\qquad\begin{pmatrix}-1\\0\\1\end{pmatrix}\overset{\pi}{\longmapsto}\begin{pmatrix}-1\\0\end{pmatrix}$$

Then we find the representation of each image with respect to the codomain's basis

$$\mathrm{Rep}_D(\begin{pmatrix}1\\0\end{pmatrix})=\begin{pmatrix}1\\-1\end{pmatrix}\quad\mathrm{Rep}_D(\begin{pmatrix}1\\1\end{pmatrix})=\begin{pmatrix}0\\1\end{pmatrix}\quad\mathrm{Rep}_D(\begin{pmatrix}-1\\0\end{pmatrix})=\begin{pmatrix}-1\\1\end{pmatrix}$$

(these are easily checked). Finally, adjoining these representations gives the matrix representing π with respect to B,D.

$$\mathrm{Rep}_{B,D}(\pi)=\begin{pmatrix}1&0&-1\\-1&1&1\end{pmatrix}_{B,D}$$

We can illustrate Theorem 1.4 by computing the matrix-vector product representing the following statement about the projection map.

$$\pi(\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

Representing this vector from the domain with respect to the domain's basis

$$\mathrm{Rep}_B(\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}) = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}_B$$

gives this matrix-vector product.

$$\mathrm{Rep}_D(\,\pi(\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix})\,) = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix}_{B,D} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}_B = \begin{pmatrix} 0 \\ 2 \end{pmatrix}_D$$

Expanding this representation into a linear combination of vectors from D

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

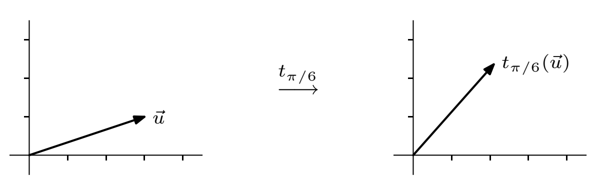
checks that the map's action is indeed reflected in the operation of the matrix. (We will sometimes compress these three displayed equations into one

$$\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}_B \overset{h}{\underset{H}{\longmapsto}} \begin{pmatrix} 0 \\ 2 \end{pmatrix}_D = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

in the course of a calculation.)

We now have two ways to compute the effect of projection, the straightforward formula that drops each three-tall vector's third component to make a two-tall vector, and the above formula that uses representations and matrix-vector multiplication. Compared to the first way, the second way might seem complicated. However, it has advantages. The next example shows that giving a formula for some maps is simplified by this new scheme.

1.8 Example To represent a *rotation* map $t_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ that turns all vectors in the plane counterclockwise through an angle θ



we start by fixing bases. Using $\mathcal{E}_2$ both as a domain basis and as a codomain basis is natural, Now, we find the image under the map of each vector in the domain's basis.

$$\begin{pmatrix}1\\0\end{pmatrix} \stackrel{t_\theta}{\longmapsto} \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} \qquad \begin{pmatrix}0\\1\end{pmatrix} \stackrel{t_\theta}{\longmapsto} \begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}$$

Then we represent these images with respect to the codomain's basis. Because this basis is $\mathcal{E}_2$, vectors are represented by themselves. Finally, adjoining the representations gives the matrix representing the map.

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(t_\theta) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

The advantage of this scheme is that just by knowing how to represent the image of the two basis vectors, we get a formula that tells us the image of any vector at all; here a vector rotated by $\theta = \pi/6$.

$$\begin{pmatrix}3\\-2\end{pmatrix} \stackrel{t_{\pi/6}}{\longmapsto} \begin{pmatrix}\sqrt{3}/2 & -1/2\\ 1/2 & \sqrt{3}/2\end{pmatrix}\begin{pmatrix}3\\-2\end{pmatrix} \approx \begin{pmatrix}3.598\\-0.232\end{pmatrix}$$

(Again, we are using the fact that, with respect to $\mathcal{E}_2$, vectors represent themselves.)

We have already seen the addition and scalar multiplication operations of matrices and the dot product operation of vectors. Matrix-vector multiplication is a new operation in the arithmetic of vectors and matrices. Nothing in Definition 1.5 requires us to view it in terms of representations. We can get some insight into this operation by turning away from what is being represented, and instead focusing on how the entries combine.

1.9 Example In the definition the width of the matrix equals the height of the vector. Hence, the first product below is defined while the second is not.

$$\begin{pmatrix}1 & 0 & 0\\ 4 & 3 & 1\end{pmatrix}\begin{pmatrix}1\\0\\2\end{pmatrix} = \begin{pmatrix}1\\6\end{pmatrix} \qquad \begin{pmatrix}1 & 0 & 0\\ 4 & 3 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$$

One reason that this product is not defined is purely formal: the definition requires that the sizes match, and these sizes don't match. Behind the formality, though, is a reason why we will leave it undefined — the matrix represents a map with a three-dimensional domain while the vector represents a member of a two-dimensional space.

A good way to view a matrix-vector product is as the dot products of the rows of the matrix with the column vector.

$$\begin{pmatrix} & \vdots & & \\ a_{i,1} & a_{i,2} & \ldots & a_{i,n}\\ & \vdots & & \end{pmatrix}\begin{pmatrix}c_1\\c_2\\ \vdots\\ c_n\end{pmatrix} = \begin{pmatrix}\vdots\\ a_{i,1}c_1 + a_{i,2}c_2 + \ldots + a_{i,n}c_n\\ \vdots\end{pmatrix}$$

Looked at in this row-by-row way, this new operation generalizes dot product.

Matrix-vector product can also be viewed column-by-column.

$$\begin{pmatrix} h_{1,1} & h_{1,2} & \ldots & h_{1,n} \\ h_{2,1} & h_{2,2} & \ldots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \ldots & h_{m,n} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} h_{1,1}c_1 + h_{1,2}c_2 + \cdots + h_{1,n}c_n \\ h_{2,1}c_1 + h_{2,2}c_2 + \cdots + h_{2,n}c_n \\ \vdots \\ h_{m,1}c_1 + h_{m,2}c_2 + \cdots + h_{m,n}c_n \end{pmatrix}$$
$$= c_1 \begin{pmatrix} h_{1,1} \\ h_{2,1} \\ \vdots \\ h_{m,1} \end{pmatrix} + \cdots + c_n \begin{pmatrix} h_{1,n} \\ h_{2,n} \\ \vdots \\ h_{m,n} \end{pmatrix}$$

1.10 Example

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \end{pmatrix}$$

The result has the columns of the matrix weighted by the entries of the vector. This way of looking at it brings us back to the objective stated at the start of this section, to compute $h(c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n)$ as $c_1 h(\vec{\beta}_1) + \cdots + c_n h(\vec{\beta}_n)$.

We began this section by noting that the equality of these two enables us to compute the action of h on any argument knowing only $h(\vec{\beta}_1)$, ..., $h(\vec{\beta}_n)$. We have developed this into a scheme to compute the action of the map by taking the matrix-vector product of the matrix representing the map and the vector representing the argument. In this way, any linear map is represented with respect to some bases by a matrix. In the next subsection, we will show the converse, that any matrix represents a linear map.

Exercises

✓ **1.11** Multiply the matrix

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

by each vector (or state "not defined").

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

1.12 Perform, if possible, each matrix-vector multiplication.

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

HW ✓ **1.13** Solve this matrix equation.

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 3 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \\ 4 \end{pmatrix}$$

✓ **1.14** For a homomorphism from $\mathcal{P}_2$ to $\mathcal{P}_3$ that sends

$$1 \mapsto 1 + x, \quad x \mapsto 1 + 2x, \quad \text{and} \quad x^2 \mapsto x - x^3$$

where does $1 - 3x + 2x^2$ go?

HW ✓ **1.15** Assume that $h\colon \mathbb{R}^2 \to \mathbb{R}^3$ is determined by this action.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

Using the standard bases, find

(a) the matrix representing this map;

(b) a general formula for $h(\vec{v})$.

✓ **1.16** Let $d/dx\colon \mathcal{P}_3 \to \mathcal{P}_3$ be the derivative transformation.

(a) Represent d/dx with respect to B, B where $B = \langle 1, x, x^2, x^3 \rangle$.

(b) Represent d/dx with respect to B, D where $D = \langle 1, 2x, 3x^2, 4x^3 \rangle$.

✓ **1.17** Represent each linear map with respect to each pair of bases.

(a) $d/dx\colon \mathcal{P}_n \to \mathcal{P}_n$ with respect to B, B where $B = \langle 1, x, \ldots, x^n \rangle$, given by

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mapsto a_1 + 2a_2 x + \cdots + n a_n x^{n-1}$$

(b) $\int\colon \mathcal{P}_n \to \mathcal{P}_{n+1}$ with respect to B_n, B_{n+1} where $B_i = \langle 1, x, \ldots, x^i \rangle$, given by

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mapsto a_0 x + \frac{a_1}{2} x^2 + \cdots + \frac{a_n}{n+1} x^{n+1}$$

(c) $\int_0^1\colon \mathcal{P}_n \to \mathbb{R}$ with respect to $B, \mathcal{E}_1$ where $B = \langle 1, x, \ldots, x^n \rangle$ and $\mathcal{E}_1 = \langle 1 \rangle$, given by

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mapsto a_0 + \frac{a_1}{2} + \cdots + \frac{a_n}{n+1}$$

(d) $\text{eval}_3\colon \mathcal{P}_n \to \mathbb{R}$ with respect to $B, \mathcal{E}_1$ where $B = \langle 1, x, \ldots, x^n \rangle$ and $\mathcal{E}_1 = \langle 1 \rangle$, given by

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mapsto a_0 + a_1 \cdot 3 + a_2 \cdot 3^2 + \cdots + a_n \cdot 3^n$$

(e) $\text{slide}_{-1}\colon \mathcal{P}_n \to \mathcal{P}_n$ with respect to B, B where $B = \langle 1, x, \ldots, x^n \rangle$, given by

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mapsto a_0 + a_1 \cdot (x+1) + \cdots + a_n \cdot (x+1)^n$$

1.18 Represent the identity map on any nontrivial space with respect to B, B, where B is any basis.

1.19 Represent, with respect to the natural basis, the transpose transformation on the space $\mathcal{M}_{2\times2}$ of 2×2 matrices.

1.20 Assume that $B = \langle \vec{\beta}_1, \vec{\beta}_2, \vec{\beta}_3, \vec{\beta}_4 \rangle$ is a basis for a vector space. Represent with respect to B, B the transformation that is determined by each.

(a) $\vec{\beta}_1 \mapsto \vec{\beta}_2$, $\vec{\beta}_2 \mapsto \vec{\beta}_3$, $\vec{\beta}_3 \mapsto \vec{\beta}_4$, $\vec{\beta}_4 \mapsto \vec{0}$

(b) $\vec{\beta}_1 \mapsto \vec{\beta}_2$, $\vec{\beta}_2 \mapsto \vec{0}$, $\vec{\beta}_3 \mapsto \vec{\beta}_4$, $\vec{\beta}_4 \mapsto \vec{0}$

(c) $\vec{\beta}_1 \mapsto \vec{\beta}_2$, $\vec{\beta}_2 \mapsto \vec{\beta}_3$, $\vec{\beta}_3 \mapsto \vec{0}$, $\vec{\beta}_4 \mapsto \vec{0}$

1.21 Example 1.8 shows how to represent the rotation transformation of the plane with respect to the standard basis. Express these other transformations also with respect to the standard basis.

(a) the *dilation* map d_s, which multiplies all vectors by the same scalar s

(b) the *reflection* map f_ℓ, which reflects all all vectors across a line ℓ through the origin

HW ✓ **1.22** Consider a linear transformation of $\mathbb{R}^2$ determined by these two.

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

(a) Represent this transformation with respect to the standard bases.

(b) Where does the transformation send this vector?

$$\begin{pmatrix} 0 \\ 5 \end{pmatrix}$$

(c) Represent this transformation with respect to these bases.

$$B = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle \qquad D = \langle \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rangle$$

(d) Using B from the prior item, represent the transformation with respect to B, B.

1.23 Suppose that $h\colon V \to W$ is nonsingular so that by Theorem 2.21, for any basis $B = \langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle \subset V$ the image $h(B) = \langle h(\vec{\beta}_1), \ldots, h(\vec{\beta}_n) \rangle$ is a basis for W.

(a) Represent the map h with respect to $B, h(B)$.

(b) For a member $\vec{v}$ of the domain, where the representation of $\vec{v}$ has components $c_1, \ldots, c_n$, represent the image vector $h(\vec{v})$ with respect to the image basis $h(B)$.

1.24 Give a formula for the product of a matrix and $\vec{e}_i$, the column vector that is all zeroes except for a single one in the i-th position.

✓ **1.25** For each vector space of functions of one real variable, represent the derivative transformation with respect to B, B.

(a) $\{a \cos x + b \sin x \bigm| a, b \in \mathbb{R}\}$, $B = \langle \cos x, \sin x \rangle$

(b) $\{a e^x + b e^{2x} \bigm| a, b \in \mathbb{R}\}$, $B = \langle e^x, e^{2x} \rangle$

(c) $\{a + bx + ce^x + dxe^{2x} \bigm| a, b, c, d \in \mathbb{R}\}$, $B = \langle 1, x, e^x, xe^x \rangle$

1.26 Find the range of the linear transformation of $\mathbb{R}^2$ represented with respect to the standard bases by each matrix.

(a) $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ **(b)** $\begin{pmatrix} 0 & 0 \\ 3 & 2 \end{pmatrix}$ **(c)** a matrix of the form $\begin{pmatrix} a & b \\ 2a & 2b \end{pmatrix}$

✓ **1.27** Can one matrix represent two different linear maps? That is, can $\text{Rep}_{B,D}(h) = \text{Rep}_{\hat{B},\hat{D}}(\hat{h})$?

1.28 Prove Theorem 1.4.

✓ **1.29** Example 1.8 shows how to represent rotation of all vectors in the plane through an angle θ about the origin, with respect to the standard bases.

(a) Rotation of all vectors in three-space through an angle θ about the x-axis is a transformation of $\mathbb{R}^3$. Represent it with respect to the standard bases. Arrange the rotation so that to someone whose feet are at the origin and whose head is at $(1, 0, 0)$, the movement appears clockwise.

(b) Repeat the prior item, only rotate about the y-axis instead. (Put the person's head at $\vec{e}_2$.)

(c) Repeat, about the z-axis.

(d) Extend the prior item to $\mathbb{R}^4$. (*Hint:* 'rotate about the z-axis' can be restated as 'rotate parallel to the xy-plane'.)

1.30 (Schur's Triangularization Lemma)

(a) Let U be a subspace of V and fix bases $B_U \subseteq B_V$. What is the relationship between the representation of a vector from U with respect to B_U and the representation of that vector (viewed as a member of V) with respect to B_V?

(b) What about maps?

(c) Fix a basis $B = \langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ for V and observe that the spans

$$[\{\vec{0}\}] = \{\vec{0}\} \subset [\{\vec{\beta}_1\}] \subset [\{\vec{\beta}_1, \vec{\beta}_2\}] \subset \quad \cdots \quad \subset [B] = V$$

form a strictly increasing chain of subspaces. Show that for any linear map $h\colon V \to W$ there is a chain $W_0 = \{\vec{0}\} \subseteq W_1 \subseteq \cdots \subseteq W_m = W$ of subspaces of W such that

$$h([\{\vec{\beta}_1, \dots, \vec{\beta}_i\}]) \subset W_i$$

for each i.

(d) Conclude that for every linear map $h\colon V \to W$ there are bases B, D so the matrix representing h with respect to B, D is upper-triangular (that is, each entry $h_{i,j}$ with $i > j$ is zero).

(e) Is an upper-triangular representation unique?

III.2 Any Matrix Represents a Linear Map

The prior subsection shows that the action of a linear map h is described by a matrix H, with respect to appropriate bases, in this way.

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_B \xmapsto[H]{h} \begin{pmatrix} h_{1,1}v_1 + \cdots + h_{1,n}v_n \\ \vdots \\ h_{m,1}v_1 + \cdots + h_{m,n}v_n \end{pmatrix}_D = h(\vec{v})$$

In this subsection, we will show the converse, that each matrix represents a linear map.

Recall that, in the definition of the matrix representation of a linear map, the number of columns of the matrix is the dimension of the map's domain and the number of rows of the matrix is the dimension of the map's codomain. Thus, for instance, a 2×3 matrix cannot represent a map from $\mathbb{R}^5$ to $\mathbb{R}^4$. The next result says that, beyond this restriction on the dimensions, there are no other limitations: the 2×3 matrix represents a map from any three-dimensional space to any two-dimensional space.

2.1 Theorem Any matrix represents a homomorphism between vector spaces of appropriate dimensions, with respect to any pair of bases.

Proof. For the matrix

$$H = \begin{pmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,n} \\ h_{2,1} & h_{2,2} & \dots & h_{2,n} \\ & \vdots & & \\ h_{m,1} & h_{m,2} & \dots & h_{m,n} \end{pmatrix}$$

fix any n-dimensional domain space V and any m-dimensional codomain space W. Also fix bases $B = \langle \vec{\beta}_1, \dots, \vec{\beta}_n \rangle$ and $D = \langle \vec{\delta}_1, \dots, \vec{\delta}_m \rangle$ for those spaces. Define a function $h\colon V \to W$ by: where $\vec{v}$ in the domain is represented as

$$\mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_B$$

then its image $h(\vec{v})$ is the member the codomain represented by

$$\mathrm{Rep}_D(\,h(\vec{v})\,) = \begin{pmatrix} h_{1,1}v_1 + \cdots + h_{1,n}v_n \\ \vdots \\ h_{m,1}v_1 + \cdots + h_{m,n}v_n \end{pmatrix}_D$$

that is, $h(\vec{v}) = h(v_1\vec{\beta}_1 + \cdots + v_n\vec{\beta}_n)$ is defined to be $(h_{1,1}v_1 + \cdots + h_{1,n}v_n)\cdot\vec{\delta}_1 + \cdots + (h_{m,1}v_1 + \cdots + h_{m,n}v_n)\cdot\vec{\delta}_m$. (This is well-defined by the uniqueness of the representation $\mathrm{Rep}_B(\vec{v})$.)

Observe that h has simply been defined to make it the map that is represented with respect to B, D by the matrix H. So to finish, we need only check that h is linear. If $\vec{v}, \vec{u} \in V$ are such that

$$\mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \text{and} \quad \mathrm{Rep}_B(\vec{u}) = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

and $c, d \in \mathbb{R}$ then the calculation

$$\begin{aligned} h(c\vec{v} + d\vec{u}) &= \bigl(h_{1,1}(cv_1 + du_1) + \cdots + h_{1,n}(cv_n + du_n)\bigr)\cdot\vec{\delta}_1 + \\ &\qquad \cdots + \bigl(h_{m,1}(cv_1 + du_1) + \cdots + h_{m,n}(cv_n + du_n)\bigr)\cdot\vec{\delta}_m \\ &= c\cdot h(\vec{v}) + d\cdot h(\vec{u}) \end{aligned}$$

provides this verification. QED

2.2 Example Which map the matrix represents depends on which bases are used. If

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_1 = D_1 = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle, \quad \text{and} \quad B_2 = D_2 = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle,$$

then $h_1 \colon \mathbb{R}^2 \to \mathbb{R}^2$ represented by H with respect to B_1, D_1 maps

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}_{B_1} \quad \mapsto \quad \begin{pmatrix} c_1 \\ 0 \end{pmatrix}_{D_1} = \begin{pmatrix} c_1 \\ 0 \end{pmatrix}$$

while $h_2 \colon \mathbb{R}^2 \to \mathbb{R}^2$ represented by H with respect to B_2, D_2 is this map.

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} c_2 \\ c_1 \end{pmatrix}_{B_2} \quad \mapsto \quad \begin{pmatrix} c_2 \\ 0 \end{pmatrix}_{D_2} = \begin{pmatrix} 0 \\ c_2 \end{pmatrix}$$

These two are different. The first is projection onto the x axis, while the second is projection onto the y axis.

So not only is any linear map described by a matrix but any matrix describes a linear map. This means that we can, when convenient, handle linear maps entirely as matrices, simply doing the computations, without have to worry that a matrix of interest does not represent a linear map on some pair of spaces of interest. (In practice, when we are working with a matrix but no spaces or bases have been specified, we will often take the domain and codomain to be $\mathbb{R}^n$ and $\mathbb{R}^m$ and use the standard bases. In this case, because the representation is transparent — the representation with respect to the standard basis of $\vec{v}$ is $\vec{v}$ — the column space of the matrix equals the range of the map. Consequently, the column space of H is often denoted by $\mathscr{R}(H)$.)

With the theorem, we have characterized linear maps as those maps that act in this matrix way. Each linear map is described by a matrix and each matrix describes a linear map. We finish this section by illustrating how a matrix can be used to tell things about its maps.

2.3 Theorem The rank of a matrix equals the rank of any map that it represents.

Proof. Suppose that the matrix H is $m\times n$. Fix domain and codomain spaces V and W of dimension n and m, with bases $B=\langle\vec{\beta}_1,\dots,\vec{\beta}_n\rangle$ and D. Then H represents some linear map h between those spaces with respect to these bases whose rangespace

$$\begin{aligned}\{h(\vec{v}) \bigm| \vec{v}\in V\} &= \{h(c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n) \bigm| c_1,\dots,c_n\in\mathbb{R}\}\\ &= \{c_1h(\vec{\beta}_1)+\cdots+c_nh(\vec{\beta}_n) \bigm| c_1,\dots,c_n\in\mathbb{R}\}\end{aligned}$$

is the span $[\{h(\vec{\beta}_1),\dots,h(\vec{\beta}_n)\}]$. The rank of h is the dimension of this rangespace.

The rank of the matrix is its column rank (or its row rank; the two are equal). This is the dimension of the column space of the matrix, which is the span of the set of column vectors $[\{\mathrm{Rep}_D(h(\vec{\beta}_1)),\dots,\mathrm{Rep}_D(h(\vec{\beta}_n))\}]$.

To see that the two spans have the same dimension, recall that a representation with respect to a basis gives an isomorphism $\mathrm{Rep}_D\colon W\to\mathbb{R}^m$. Under this isomorphism, there is a linear relationship among members of the rangespace if and only if the same relationship holds in the column space, e.g, $\vec{0}=c_1h(\vec{\beta}_1)+\cdots+c_nh(\vec{\beta}_n)$ if and only if $\vec{0}=c_1\mathrm{Rep}_D(h(\vec{\beta}_1))+\cdots+c_n\mathrm{Rep}_D(h(\vec{\beta}_n))$. Hence, a subset of the rangespace is linearly independent if and only if the corresponding subset of the column space is linearly independent. This means that the size of the largest linearly independent subset of the rangespace equals the size of the largest linearly independent subset of the column space, and so the two spaces have the same dimension. QED

2.4 Example Any map represented by

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

must, by definition, be from a three-dimensional domain to a four-dimensional codomain. In addition, because the rank of this matrix is two (we can spot this by eye or get it with Gauss' method), any map represented by this matrix has a two-dimensional rangespace.

2.5 Corollary Let h be a linear map represented by a matrix H. Then h is onto if and only if the rank of H equals the number of its rows, and h is one-to-one if and only if the rank of H equals the number of its columns.

Proof. For the first half, the dimension of the rangespace of h is the rank of h, which equals the rank of H by the theorem. Since the dimension of the codomain of h is the number of columns in H, if the rank of H equals the number of columns, then the dimension of the rangespace equals the dimension of the codomain. But a subspace with the same dimension as its superspace must equal that superspace (a basis for the rangespace is a linearly independent subset of the codomain, whose size is equal to the dimension of the codomain, and so this set is a basis for the codomain).

For the second half, a linear map is one-to-one if and only if it is an isomorphism between its domain and its range, that is, if and only if its domain has the same dimension as its range. But the number of columns in h is the dimension of h's domain, and by the theorem the rank of H equals the dimension of h's range. QED

The above results end any confusion caused by our use of the word 'rank' to mean apparently different things when applied to matrices and when applied to maps. We can also justify the dual use of 'nonsingular'. We've defined a matrix to be nonsingular if it is square and is the matrix of coefficients of a linear system with a unique solution, and we've defined a linear map to be nonsingular if it is one-to-one.

2.6 Corollary A square matrix represents nonsingular maps if and only if it is a nonsingular matrix. Thus, a matrix represents an isomorphism if and only if it is square and nonsingular.

Proof. Immediate from the prior result. QED

2.7 Example Any map from $\mathbb{R}^2$ to $\mathcal{P}_1$ represented with respect to any pair of bases by

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

is nonsingular because this matrix has rank two.

2.8 Example Any map $g\colon V \to W$ represented by

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$$

is not nonsingular because this matrix is not nonsingular.

We've now seen that the relationship between maps and matrices goes both ways: fixing bases, any linear map is represented by a matrix and any matrix describes a linear map. That is, by fixing spaces and bases we get a correspondence between maps and matrices. In the rest of this chapter we will explore this correspondence. For instance, we've defined for linear maps the operations of addition and scalar multiplication and we shall see what the corresponding matrix operations are. We shall also see the matrix operation that represent the map operation of composition. And, we shall see how to find the matrix that represents a map's inverse.

Exercises

✓ **2.9** Decide if the vector is in the column space of the matrix.

(a) $\begin{pmatrix} 2 & 1 \\ 2 & 5 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \end{pmatrix}$ **(b)** $\begin{pmatrix} 4 & -8 \\ 2 & -4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ **(c)** $\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$

✓ **2.10** Decide if each vector lies in the range of the map from $\mathbb{R}^3$ to $\mathbb{R}^2$ represented with respect to the standard bases by the matrix.

(a) $\begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ **(b)** $\begin{pmatrix} 2 & 0 & 3 \\ 4 & 0 & 6 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

HW ✓ **2.11** Consider this matrix, representing a transformation of $\mathbb{R}^2$, and these bases for that space.

$$\frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \qquad B = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle \quad D = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle$$

(a) To what vector in the codomain is the first member of B mapped?
(b) The second member?
(c) Where is a general vector from the domain (a vector with components x and y) mapped? That is, what transformation of $\mathbb{R}^2$ is represented with respect to B, D by this matrix?

2.12 What transformation of $F = \{a\cos\theta + b\sin\theta \mid a, b \in \mathbb{R}\}$ is represented with respect to $B = \langle \cos\theta - \sin\theta, \sin\theta \rangle$ and $D = \langle \cos\theta + \sin\theta, \cos\theta \rangle$ by this matrix?

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

HW ✓ **2.13** Decide if $1 + 2x$ is in the range of the map from $\mathbb{R}^3$ to $\mathcal{P}_2$ represented with respect to $\mathcal{E}_3$ and $\langle 1, 1 + x^2, x \rangle$ by this matrix.

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

2.14 Example 2.8 gives a matrix that is nonsingular, and is therefore associated with maps that are nonsingular.

(a) Find the set of column vectors representing the members of the nullspace of any map represented by this matrix.
(b) Find the nullity of any such map.
(c) Find the set of column vectors representing the members of the rangespace of any map represented by this matrix.
(d) Find the rank of any such map.
(e) Check that rank plus nullity equals the dimension of the domain.

✓ **2.15** Because the rank of a matrix equals the rank of any map it represents, if one matrix represents two different maps $H = \mathrm{Rep}_{B,D}(h) = \mathrm{Rep}_{\hat{B},\hat{D}}(\hat{h})$ (where $h, \hat{h}\colon V \to W$) then the dimension of the rangespace of h equals the dimension of the rangespace of $\hat{h}$. Must these equal-dimensioned rangespaces actually be the same?

✓ **2.16** Let V be an n-dimensional space with bases B and D. Consider a map that sends, for $\vec{v} \in V$, the column vector representing $\vec{v}$ with respect to B to the column vector representing $\vec{v}$ with respect to D. Show that is a linear transformation of $\mathbb{R}^n$.

2.17 Example 2.2 shows that changing the pair of bases can change the map that a matrix represents, even though the domain and codomain remain the same. Could the map ever not change? Is there a matrix H, vector spaces V and W, and associated pairs of bases B_1, D_1 and B_2, D_2 (with $B_1 \neq B_2$ or $D_1 \neq D_2$ or both) such that the map represented by H with respect to B_1, D_1 equals the map represented by H with respect to B_2, D_2?

✓ **2.18** A square matrix is a *diagonal* matrix if it is all zeroes except possibly for the entries on its upper-left to lower-right diagonal — its $1,1$ entry, its $2,2$ entry, etc. Show that a linear map is an isomorphism if there are bases such that, with respect to those bases, the map is represented by a diagonal matrix with no zeroes on the diagonal.

HW **2.19** Describe geometrically the action on $\mathbb{R}^2$ of the map represented with respect to the standard bases $\mathcal{E}_2, \mathcal{E}_2$ by this matrix.

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

Do the same for these.

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

2.20 The fact that for any linear map the rank plus the nullity equals the dimension of the domain shows that a necessary condition for the existence of a homomorphism between two spaces, onto the second space, is that there be no gain in dimension. That is, where $h\colon V \to W$ is onto, the dimension of W must be less than or equal to the dimension of V.

(a) Show that this (strong) converse holds: no gain in dimension implies that there is a homomorphism and, further, any matrix with the correct size and correct rank represents such a map.

(b) Are there bases for $\mathbb{R}^3$ such that this matrix

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

represents a map from $\mathbb{R}^3$ to $\mathbb{R}^3$ whose range is the xy plane subspace of $\mathbb{R}^3$?

2.21 Let V be an n-dimensional space and suppose that $\vec{x} \in \mathbb{R}^n$. Fix a basis B for V and consider the map $h_{\vec{x}}\colon V \to \mathbb{R}$ given $\vec{v} \mapsto \vec{x} \cdot \mathrm{Rep}_B(\vec{v})$ by the dot product.

(a) Show that this map is linear.

(b) Show that for any linear map $g\colon V \to \mathbb{R}$ there is an $\vec{x} \in \mathbb{R}^n$ such that $g = h_{\vec{x}}$.

(c) In the prior item we fixed the basis and varied the $\vec{x}$ to get all possible linear maps. Can we get all possible linear maps by fixing an $\vec{x}$ and varying the basis?

2.22 Let V, W, X be vector spaces with bases B, C, D.

(a) Suppose that $h\colon V \to W$ is represented with respect to B, C by the matrix H. Give the matrix representing the scalar multiple rh (where $r \in \mathbb{R}$) with respect to B, C by expressing it in terms of H.

(b) Suppose that $h, g\colon V \to W$ are represented with respect to B, C by H and G. Give the matrix representing $h + g$ with respect to B, C by expressing it in terms of H and G.

(c) Suppose that $h\colon V \to W$ is represented with respect to B, C by H and $g\colon W \to X$ is represented with respect to C, D by G. Give the matrix representing $g \circ h$ with respect to B, D by expressing it in terms of H and G.

IV Matrix Operations

The prior section shows how matrices represent linear maps. A good strategy, on seeing a new idea, is to explore how it interacts with some already-established ideas. In the first subsection we will ask how the representation of the sum of two maps $f+g$ is related to the representations of the two maps, and how the representation of a scalar product $r \cdot h$ of a map is related to the representation of that map. In later subsections we will see how to represent map composition and map inverse.

IV.1 Sums and Scalar Products

Recall that for two maps f and g with the same domain and codomain, the map sum $f+g$ has this definition.

$$\vec{v} \xmapsto{f+g} f(\vec{v}) + g(\vec{v})$$

The easiest way to see how the representations of the maps combine to represent the map sum is with an example.

1.1 Example Suppose that $f, g\colon \mathbb{R}^2 \to \mathbb{R}^3$ are represented with respect to the bases B and D by these matrices.

$$F = \mathrm{Rep}_{B,D}(f) = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 1 & 0 \end{pmatrix}_{B,D} \qquad G = \mathrm{Rep}_{B,D}(g) = \begin{pmatrix} 0 & 0 \\ -1 & -2 \\ 2 & 4 \end{pmatrix}_{B,D}$$

Then, for any $\vec{v} \in V$ represented with respect to B, computation of the representation of $f(\vec{v}) + g(\vec{v})$

$$\begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -1 & -2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1v_1 + 3v_2 \\ 2v_1 + 0v_2 \\ 1v_1 + 0v_2 \end{pmatrix} + \begin{pmatrix} 0v_1 + 0v_2 \\ -1v_1 - 2v_2 \\ 2v_1 + 4v_2 \end{pmatrix}$$

gives this representation of $f+g\,(\vec{v})$.

$$\begin{pmatrix} (1+0)v_1 + (3+0)v_2 \\ (2-1)v_1 + (0-2)v_2 \\ (1+2)v_1 + (0+4)v_2 \end{pmatrix} = \begin{pmatrix} 1v_1 + 3v_2 \\ 1v_1 - 2v_2 \\ 3v_1 + 4v_2 \end{pmatrix}$$

Thus, the action of $f+g$ is described by this matrix-vector product.

$$\begin{pmatrix} 1 & 3 \\ 1 & -2 \\ 3 & 4 \end{pmatrix}_{B,D} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_B = \begin{pmatrix} 1v_1 + 3v_2 \\ 1v_1 - 2v_2 \\ 3v_1 + 4v_2 \end{pmatrix}_D$$

This matrix is the entry-by-entry sum of original matrices, e.g., the $1,1$ entry of $\mathrm{Rep}_{B,D}(f+g)$ is the sum of the $1,1$ entry of F and the $1,1$ entry of G.

Representing a scalar multiple of a map works the same way.

1.2 Example If t is a transformation represented by

$$\mathrm{Rep}_{B,D}(t) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}_{B,D} \quad \text{so that} \quad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_B \mapsto \begin{pmatrix} v_1 \\ v_1 + v_2 \end{pmatrix}_D = t(\vec{v})$$

then the scalar multiple map $5t$ acts in this way.

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_B \longmapsto \begin{pmatrix} 5v_1 \\ 5v_1 + 5v_2 \end{pmatrix}_D = 5 \cdot t(\vec{v})$$

Therefore, this is the matrix representing $5t$.

$$\mathrm{Rep}_{B,D}(5t) = \begin{pmatrix} 5 & 0 \\ 5 & 5 \end{pmatrix}_{B,D}$$

1.3 Definition The *sum* of two same-sized matrices is their entry-by-entry sum. The *scalar multiple* of a matrix is the result of entry-by-entry scalar multiplication.

1.4 Remark These extend the vector addition and scalar multiplication operations that we defined in the first chapter.

1.5 Theorem Let $h, g \colon V \to W$ be linear maps represented with respect to bases B, D by the matrices H and G, and let r be a scalar. Then the map $h + g \colon V \to W$ is represented with respect to B, D by $H + G$, and the map $r \cdot h \colon V \to W$ is represented with respect to B, D by rH.

Proof. Exercise 8; generalize the examples above. QED

A notable special case of scalar multiplication is multiplication by zero. For any map $0 \cdot h$ is the zero homomorphism and for any matrix $0 \cdot H$ is the zero matrix.

1.6 Example The zero map from any three-dimensional space to any two-dimensional space is represented by the 2×3 zero matrix

$$Z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

no matter which domain and codomain bases are used.

Exercises

✓ **1.7** Perform the indicated operations, if defined.

(a) $\begin{pmatrix} 5 & -1 & 2 \\ 6 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 1 & 4 \\ 3 & 0 & 5 \end{pmatrix}$

(b) $6 \cdot \begin{pmatrix} 2 & -1 & -1 \\ 1 & 2 & 3 \end{pmatrix}$

(c) $\begin{pmatrix}2&1\\0&3\end{pmatrix}+\begin{pmatrix}2&1\\0&3\end{pmatrix}$

(d) $4\begin{pmatrix}1&2\\3&-1\end{pmatrix}+5\begin{pmatrix}-1&4\\-2&1\end{pmatrix}$

(e) $3\begin{pmatrix}2&1\\3&0\end{pmatrix}+2\begin{pmatrix}1&1&4\\3&0&5\end{pmatrix}$

1.8 Prove Theorem 1.5.
(a) Prove that matrix addition represents addition of linear maps.
(b) Prove that matrix scalar multiplication represents scalar multiplication of linear maps.

✓ **1.9** Prove each, where the operations are defined, where G, H, and J are matrices, where Z is the zero matrix, and where r and s are scalars.
(a) Matrix addition is commutative $G+H=H+G$.
(b) Matrix addition is associative $G+(H+J)=(G+H)+J$.
(c) The zero matrix is an additive identity $G+Z=G$.
(d) $0\cdot G=Z$
(e) $(r+s)G=rG+sG$
(f) Matrices have an additive inverse $G+(-1)\cdot G=Z$.
(g) $r(G+H)=rG+rH$
(h) $(rs)G=r(sG)$

1.10 Fix domain and codomain spaces. In general, one matrix can represent many different maps with respect to different bases. However, prove that a zero matrix represents only a zero map. Are there other such matrices?

✓ **1.11** Let V and W be vector spaces of dimensions n and m. Show that the space $\mathcal{L}(V,W)$ of linear maps from V to W is isomorphic to $\mathcal{M}_{m\times n}$.

✓ **1.12** Show that it follows from the prior questions that for any six transformations $t_1,\ldots,t_6\colon\mathbb{R}^2\to\mathbb{R}^2$ there are scalars $c_1,\ldots,c_6\in\mathbb{R}$ such that $c_1t_1+\cdots+c_6t_6$ is the zero map. (*Hint:* this is a bit of a misleading question.)

HW **1.13** The *trace* of a square matrix is the sum of the entries on the main diagonal (the $1,1$ entry plus the $2,2$ entry, etc.; we will see the significance of the trace in Chapter Five). Show that $\text{trace}(H+G)=\text{trace}(H)+\text{trace}(G)$. Is there a similar result for scalar multiplication?

HW **1.14** Recall that the *transpose* of a matrix M is another matrix, whose i,j entry is the j,i entry of M. Verify these identities.
(a) $(G+H)^{\text{trans}}=G^{\text{trans}}+H^{\text{trans}}$
(b) $(r\cdot H)^{\text{trans}}=r\cdot H^{\text{trans}}$

✓ **1.15** A square matrix is *symmetric* if each i,j entry equals the j,i entry, that is, if the matrix equals its transpose.
(a) Prove that for any H, the matrix $H+H^{\text{trans}}$ is symmetric. Does every symmetric matrix have this form?
(b) Prove that the set of $n\times n$ symmetric matrices is a subspace of $\mathcal{M}_{n\times n}$.

✓ **1.16** **(a)** How does matrix rank interact with scalar multiplication — can a scalar product of a rank n matrix have rank less than n? Greater?
(b) How does matrix rank interact with matrix addition — can a sum of rank n matrices have rank less than n? Greater?

IV.2 Matrix Multiplication

After representing addition and scalar multiplication of linear maps in the prior subsection, the natural next map operation to consider is composition.

2.1 Lemma A composition of linear maps is linear.

PROOF. *(This argument has appeared earlier, as part of the proof that isomorphism is an equivalence relation between spaces.)* Let $h\colon V\to W$ and $g\colon W\to U$ be linear. The calculation

$$\begin{aligned} g\circ h\,\bigl(c_1\cdot\vec{v}_1+c_2\cdot\vec{v}_2\bigr) &= g\bigl(\,h(c_1\cdot\vec{v}_1+c_2\cdot\vec{v}_2)\,\bigr) = g\bigl(\,c_1\cdot h(\vec{v}_1)+c_2\cdot h(\vec{v}_2)\,\bigr) \\ &= c_1\cdot g\bigl(h(\vec{v}_1)\bigr)+c_2\cdot g\bigl(h(\vec{v}_2)\bigr) = c_1\cdot(g\circ h)(\vec{v}_1)+c_2\cdot(g\circ h)(\vec{v}_2) \end{aligned}$$

shows that $g\circ h\colon V\to U$ preserves linear combinations. QED

To see how the representation of the composite arises out of the representations of the two compositors, consider an example.

2.2 Example Let $h\colon\mathbb{R}^4\to\mathbb{R}^2$ and $g\colon\mathbb{R}^2\to\mathbb{R}^3$, fix bases $B\subset\mathbb{R}^4$, $C\subset\mathbb{R}^2$, $D\subset\mathbb{R}^3$, and let these be the representations.

$$H=\operatorname{Rep}_{B,C}(h)=\begin{pmatrix}4&6&8&2\\5&7&9&3\end{pmatrix}_{B,C}\qquad G=\operatorname{Rep}_{C,D}(g)=\begin{pmatrix}1&1\\0&1\\1&0\end{pmatrix}_{C,D}$$

To represent the composition $g\circ h\colon\mathbb{R}^4\to\mathbb{R}^3$ we fix a $\vec{v}$, represent h of $\vec{v}$, and then represent g of that. The representation of $h(\vec{v})$ is the product of h's matrix and $\vec{v}$'s vector.

$$\operatorname{Rep}_C(\,h(\vec{v})\,)=\begin{pmatrix}4&6&8&2\\5&7&9&3\end{pmatrix}_{B,C}\begin{pmatrix}v_1\\v_2\\v_3\\v_4\end{pmatrix}_B=\begin{pmatrix}4v_1+6v_2+8v_3+2v_4\\5v_1+7v_2+9v_3+3v_4\end{pmatrix}_C$$

The representation of $g(\,h(\vec{v})\,)$ is the product of g's matrix and $h(\vec{v})$'s vector.

$$\begin{aligned}\operatorname{Rep}_D(\,g(h(\vec{v}))\,)&=\begin{pmatrix}1&1\\0&1\\1&0\end{pmatrix}_{C,D}\begin{pmatrix}4v_1+6v_2+8v_3+2v_4\\5v_1+7v_2+9v_3+3v_4\end{pmatrix}_C\\&=\begin{pmatrix}1\cdot(4v_1+6v_2+8v_3+2v_4)+1\cdot(5v_1+7v_2+9v_3+3v_4)\\0\cdot(4v_1+6v_2+8v_3+2v_4)+1\cdot(5v_1+7v_2+9v_3+3v_4)\\1\cdot(4v_1+6v_2+8v_3+2v_4)+0\cdot(5v_1+7v_2+9v_3+3v_4)\end{pmatrix}_D\end{aligned}$$

Distributing and regrouping on the v's gives

$$=\begin{pmatrix}(1\cdot4+1\cdot5)v_1+(1\cdot6+1\cdot7)v_2+(1\cdot8+1\cdot9)v_3+(1\cdot2+1\cdot3)v_4\\(0\cdot4+1\cdot5)v_1+(0\cdot6+1\cdot7)v_2+(0\cdot8+1\cdot9)v_3+(0\cdot2+1\cdot3)v_4\\(1\cdot4+0\cdot5)v_1+(1\cdot6+0\cdot7)v_2+(1\cdot8+0\cdot9)v_3+(1\cdot2+0\cdot3)v_4\end{pmatrix}_D$$

which we recognizing as the result of this matrix-vector product.

$$= \begin{pmatrix} 1\cdot 4+1\cdot 5 & 1\cdot 6+1\cdot 7 & 1\cdot 8+1\cdot 9 & 1\cdot 2+1\cdot 3 \\ 0\cdot 4+1\cdot 5 & 0\cdot 6+1\cdot 7 & 0\cdot 8+1\cdot 9 & 0\cdot 2+1\cdot 3 \\ 1\cdot 4+0\cdot 5 & 1\cdot 6+0\cdot 7 & 1\cdot 8+0\cdot 9 & 1\cdot 2+0\cdot 3 \end{pmatrix}_{B,D} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}_D$$

Thus, the matrix representing $g\circ h$ has the rows of G combined with the columns of H.

2.3 Definition The *matrix-multiplicative product* of the $m\times r$ matrix G and the $r\times n$ matrix H is the $m\times n$ matrix P, where

$$p_{i,j} = g_{i,1}h_{1,j} + g_{i,2}h_{2,j} + \cdots + g_{i,r}h_{r,j}$$

that is, the i,j-th entry of the product is the dot product of the i-th row and the j-th column.

$$GH = \begin{pmatrix} & \vdots & & \\ g_{i,1} & g_{i,2} & \ldots & g_{i,r} \\ & \vdots & & \end{pmatrix} \begin{pmatrix} & h_{1,j} & \\ \ldots & h_{2,j} & \ldots \\ & \vdots & \\ & h_{r,j} & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \ldots & p_{i,j} & \ldots \\ & \vdots & \end{pmatrix}$$

2.4 Example The matrices from Example 2.2 combine in this way.

$$\begin{pmatrix} 1\cdot 4+1\cdot 5 & 1\cdot 6+1\cdot 7 & 1\cdot 8+1\cdot 9 & 1\cdot 2+1\cdot 3 \\ 0\cdot 4+1\cdot 5 & 0\cdot 6+1\cdot 7 & 0\cdot 8+1\cdot 9 & 0\cdot 2+1\cdot 3 \\ 1\cdot 4+0\cdot 5 & 1\cdot 6+0\cdot 7 & 1\cdot 8+0\cdot 9 & 1\cdot 2+0\cdot 3 \end{pmatrix} = \begin{pmatrix} 9 & 13 & 17 & 5 \\ 5 & 7 & 9 & 3 \\ 4 & 6 & 8 & 2 \end{pmatrix}$$

2.5 Example

$$\begin{pmatrix} 2 & 0 \\ 4 & 6 \\ 8 & 2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} = \begin{pmatrix} 2\cdot 1+0\cdot 5 & 2\cdot 3+0\cdot 7 \\ 4\cdot 1+6\cdot 5 & 4\cdot 3+6\cdot 7 \\ 8\cdot 1+2\cdot 5 & 8\cdot 3+2\cdot 7 \end{pmatrix} = \begin{pmatrix} 2 & 6 \\ 34 & 54 \\ 18 & 38 \end{pmatrix}$$

2.6 Theorem A composition of linear maps is represented by the matrix product of the representatives.

Proof. *(This argument parallels Example 2.2.)* Let $h\colon V\to W$ and $g\colon W\to X$ be represented by H and G with respect to bases $B\subset V$, $C\subset W$, and $D\subset X$, of sizes n, r, and m. For any $\vec{v}\in V$, the k-th component of $\mathrm{Rep}_C(\,h(\vec{v})\,)$ is

$$h_{k,1}v_1 + \cdots + h_{k,n}v_n$$

and so the i-th component of $\mathrm{Rep}_D(\,g\circ h\,(\vec{v})\,)$ is this.

$$g_{i,1}\cdot(h_{1,1}v_1+\cdots+h_{1,n}v_n) + g_{i,2}\cdot(h_{2,1}v_1+\cdots+h_{2,n}v_n) \\ +\cdots+ g_{i,r}\cdot(h_{r,1}v_1+\cdots+h_{r,n}v_n)$$

Distribute and regroup on the v's.

$$\begin{aligned} &= (g_{i,1}h_{1,1} + g_{i,2}h_{2,1} + \cdots + g_{i,r}h_{r,1}) \cdot v_1 \\ &\qquad\qquad + \cdots + (g_{i,1}h_{1,n} + g_{i,2}h_{2,n} + \cdots + g_{i,r}h_{r,n}) \cdot v_n \end{aligned}$$

Finish by recognizing that the coefficient of each v_j

$$g_{i,1}h_{1,j} + g_{i,2}h_{2,j} + \cdots + g_{i,r}h_{r,j}$$

matches the definition of the i, j entry of the product GH. QED

The theorem is an example of a result that supports a definition. We can picture what the definition and theorem together say with this *arrow diagram* ('wrt' abbreviates 'with respect to').

$$\begin{tikzcd} & W_{wrt\ C} \arrow[dr, "g", "G"'] & \\ V_{wrt\ B} \arrow[ur, "h", "H"'] \arrow[rr, "g \circ h", "GH"'] & & X_{wrt\ D} \end{tikzcd}$$

Above the arrows, the maps show that the two ways of going from V to X, straight over via the composition or else by way of W, have the same effect

$$\vec{v} \xmapsto{g \circ h} g(h(\vec{v})) \qquad \vec{v} \xmapsto{h} h(\vec{v}) \xmapsto{g} g(h(\vec{v}))$$

(this is just the definition of composition). Below the arrows, the matrices indicate that the product does the same thing — multiplying GH into the column vector $\mathrm{Rep}_B(\vec{v})$ has the same effect as multiplying the column first by H and then multiplying the result by G.

$$\mathrm{Rep}_{B,D}(g \circ h) = GH = \mathrm{Rep}_{C,D}(g)\,\mathrm{Rep}_{B,C}(h)$$

The definition of the matrix-matrix product operation does not restrict us to view it as a representation of a linear map composition. We can get insight into this operation by studying it as a mechanical procedure. The striking thing is the way that rows and columns combine.

One aspect of that combination is that the sizes of the matrices involved is significant. Briefly, $m \times r$ times $r \times n$ equals $m \times n$.

2.7 Example This product is not defined

$$\begin{pmatrix} -1 & 2 & 0 \\ 0 & 10 & 1.1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

because the number of columns on the left does not equal the number of rows on the right.

In terms of the underlying maps, the fact that the sizes must match up reflects the fact that matrix multiplication is defined only when a corresponding function composition

$$\text{dimension } n \text{ space} \xrightarrow{h} \text{dimension } r \text{ space} \xrightarrow{g} \text{dimension } m \text{ space}$$

is possible.

2.8 Remark The order in which these things are written can be confusing. In the '$m \times r$ times $r \times n$ equals $m \times n$' equation, the number written first m is the dimension of g's codomain and is thus the number that appears last in the map dimension description above. The explanation is that while f is done first and then g is applied, that composition is written $g \circ f$, from the notation '$g(f(\vec{v}))$'. (Some people try to lessen confusion by reading '$g \circ f$' aloud as "g following f".) That order then carries over to matrices: $g \circ f$ is represented by GF.

Another aspect of the way that rows and columns combine in the matrix product operation is that in the definition of the i, j entry

$$p_{i,j} = g_{i,\boxed{1}} h_{\boxed{1},j} + g_{i,\boxed{2}} h_{\boxed{2},j} + \cdots + g_{i,\boxed{r}} h_{\boxed{r},j}$$

the boxed subscripts on the g's are column indicators while those on the h's indicate rows. That is, summation takes place over the columns of G but over the rows of H; left is treated differently than right, so GH may be unequal to HG. Matrix multiplication is not commutative.

2.9 Example Matrix multiplication hardly ever commutes. Test that by multiplying randomly chosen matrices both ways.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix} \qquad \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$$

2.10 Example Commutativity can fail more dramatically:

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 23 & 34 & 0 \\ 31 & 46 & 0 \end{pmatrix}$$

while

$$\begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

isn't even defined.

2.11 Remark The fact that matrix multiplication is not commutative may be puzzling at first sight, perhaps just because most algebraic operations in elementary mathematics are commutative. But on further reflection, it isn't so surprising. After all, matrix multiplication represents function composition, which is not commutative — if $f(x) = 2x$ and $g(x) = x + 1$ then $g \circ f(x) = 2x + 1$ while $f \circ g(x) = 2(x + 1) = 2x + 2$. True, this g is not linear and we might have hoped that linear functions commute, but this perspective shows that the failure of commutativity for matrix multiplication fits into a larger context.

Except for the lack of commutativity, matrix multiplication is algebraically well-behaved. Below are some nice properties and more are in Exercise 23 and Exercise 24.

2.12 Theorem If F, G, and H are matrices, and the matrix products are defined, then the product is associative $(FG)H = F(GH)$ and distributes over matrix addition $F(G+H) = FG + FH$ and $(G+H)F = GF + HF$.

Proof. Associativity holds because matrix multiplication represents function composition, which is associative: the maps $(f \circ g) \circ h$ and $f \circ (g \circ h)$ are equal as both send $\vec{v}$ to $f(g(h(\vec{v})))$.

Distributivity is similar. For instance, the first one goes $f \circ (g+h)\,(\vec{v}) = f\big(\,(g+h)(\vec{v})\,\big) = f\big(\,g(\vec{v}) + h(\vec{v})\,\big) = f(g(\vec{v})) + f(h(\vec{v})) = f \circ g(\vec{v}) + f \circ h(\vec{v})$ (the third equality uses the linearity of f). QED

2.13 Remark We could alternatively prove that result by slogging through the indices. For example, associativity goes: the i,j-th entry of $(FG)H$ is

$$\begin{aligned}
&(f_{i,1}g_{1,1} + f_{i,2}g_{2,1} + \cdots + f_{i,r}g_{r,1})h_{1,j} \\
&\quad + (f_{i,1}g_{1,2} + f_{i,2}g_{2,2} + \cdots + f_{i,r}g_{r,2})h_{2,j} \\
&\quad\ \vdots \\
&\quad + (f_{i,1}g_{1,s} + f_{i,2}g_{2,s} + \cdots + f_{i,r}g_{r,s})h_{s,j}
\end{aligned}$$

(where F, G, and H are $m \times r$, $r \times s$, and $s \times n$ matrices), distribute

$$\begin{aligned}
&f_{i,1}g_{1,1}h_{1,j} + f_{i,2}g_{2,1}h_{1,j} + \cdots + f_{i,r}g_{r,1}h_{1,j} \\
&\quad + f_{i,1}g_{1,2}h_{2,j} + f_{i,2}g_{2,2}h_{2,j} + \cdots + f_{i,r}g_{r,2}h_{2,j} \\
&\quad\ \vdots \\
&\quad + f_{i,1}g_{1,s}h_{s,j} + f_{i,2}g_{2,s}h_{s,j} + \cdots + f_{i,r}g_{r,s}h_{s,j}
\end{aligned}$$

and regroup around the f's

$$\begin{aligned}
&f_{i,1}(g_{1,1}h_{1,j} + g_{1,2}h_{2,j} + \cdots + g_{1,s}h_{s,j}) \\
&\quad + f_{i,2}(g_{2,1}h_{1,j} + g_{2,2}h_{2,j} + \cdots + g_{2,s}h_{s,j}) \\
&\quad\ \vdots \\
&\quad + f_{i,r}(g_{r,1}h_{1,j} + g_{r,2}h_{2,j} + \cdots + g_{r,s}h_{s,j})
\end{aligned}$$

to get the i,j entry of $F(GH)$.

Contrast these two ways of verifying associativity, the one in the proof and the one just above. The argument just above is hard to understand in the sense that, while the calculations are easy to check, the arithmetic seems unconnected to any idea (it also essentially repeats the proof of Theorem 2.6 and so is inefficient). The argument in the proof is shorter, clearer, and says why this property "really" holds. This illustrates the comments made in the preamble to the chapter on vector spaces — at least some of the time an argument from higher-level constructs is clearer.

We have now seen how the representation of the composition of two linear maps is derived from the representations of the two maps. We have called the combination the product of the two matrices. This operation is extremely important. Before we go on to study how to represent the inverse of a linear map, we will explore it some more in the next subsection.

Exercises

HW ✓ **2.14** Compute, or state "not defined".

(a) $\begin{pmatrix} 3 & 1 \\ -4 & 2 \end{pmatrix}\begin{pmatrix} 0 & 5 \\ 0 & 0.5 \end{pmatrix}$ **(b)** $\begin{pmatrix} 1 & 1 & -1 \\ 4 & 0 & 3 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ 3 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & -7 \\ 7 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & 5 \\ -1 & 1 & 1 \\ 3 & 8 & 4 \end{pmatrix}$ **(d)** $\begin{pmatrix} 5 & 2 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 3 & -5 \end{pmatrix}$

✓ **2.15** Where

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 5 & 2 \\ 4 & 4 \end{pmatrix} \quad C = \begin{pmatrix} -2 & 3 \\ -4 & 1 \end{pmatrix}$$

compute or state 'not defined'.

(a) AB **(b)** $(AB)C$ **(c)** BC **(d)** $A(BC)$

2.16 Which products are defined?

(a) 3×2 times 2×3 **(b)** 2×3 times 3×2 **(c)** 2×2 times 3×3
(d) 3×3 times 2×2

✓ **2.17** Give the size of the product or state "not defined".

(a) a 2×3 matrix times a 3×1 matrix
(b) a 1×12 matrix times a 12×1 matrix
(c) a 2×3 matrix times a 2×1 matrix
(d) a 2×2 matrix times a 2×2 matrix

✓ **2.18** Find the system of equations resulting from starting with

$$\begin{aligned} h_{1,1}x_1 + h_{1,2}x_2 + h_{1,3}x_3 &= d_1 \\ h_{2,1}x_1 + h_{2,2}x_2 + h_{2,3}x_3 &= d_2 \end{aligned}$$

and making this change of variable (i.e., substitution).

$$\begin{aligned} x_1 &= g_{1,1}y_1 + g_{1,2}y_2 \\ x_2 &= g_{2,1}y_1 + g_{2,2}y_2 \\ x_3 &= g_{3,1}y_1 + g_{3,2}y_2 \end{aligned}$$

2.19 As Definition 2.3 points out, the matrix product operation generalizes the dot product. Is the dot product of a $1\times n$ row vector and a $n\times 1$ column vector the same as their matrix-multiplicative product?

✓ **2.20** Represent the derivative map on $\mathcal{P}_n$ with respect to B, B where B is the natural basis $\langle 1, x, \ldots, x^n\rangle$. Show that the product of this matrix with itself is defined; what the map does it represent?

2.21 Show that composition of linear transformations on $\mathbb{R}^1$ is commutative. Is this true for any one-dimensional space?

2.22 Why is matrix multiplication not defined as entry-wise multiplication? That would be easier, and commutative too.

✓ **2.23** **(a)** Prove that $H^pH^q = H^{p+q}$ and $(H^p)^q = H^{pq}$ for positive integers p, q.
(b) Prove that $(rH)^p = r^p \cdot H^p$ for any positive integer p and scalar $r \in \mathbb{R}$.

✓ **2.24** **(a)** How does matrix multiplication interact with scalar multiplication: is $r(GH) = (rG)H$? Is $G(rH) = r(GH)$?

(b) How does matrix multiplication interact with linear combinations: is $F(rG+sH)=r(FG)+s(FH)$? Is $(rF+sG)H=rFH+sGH$?

2.25 We can ask how the matrix product operation interacts with the transpose operation.

(a) Show that $(GH)^{\text{trans}}=H^{\text{trans}}G^{\text{trans}}$.

(b) A square matrix is *symmetric* if each i,j entry equals the j,i entry, that is, if the matrix equals its own transpose. Show that the matrices HH^{trans} and $H^{\text{trans}}H$ are symmetric.

✓ **2.26** Rotation of vectors in $\mathbb{R}^3$ about an axis is a linear map. Show that linear maps do not commute by showing geometrically that rotations do not commute.

2.27 In the proof of Theorem 2.12 some maps are used. What are the domains and codomains?

2.28 How does matrix rank interact with matrix multiplication?

(a) Can the product of rank n matrices have rank less than n? Greater?

(b) Show that the rank of the product of two matrices is less than or equal to the minimum of the rank of each factor.

2.29 Is 'commutes with' an equivalence relation among $n\times n$ matrices?

✓ **2.30** *(This will be used in the Matrix Inverses exercises.)* Here is another property of matrix multiplication that might be puzzling at first sight.

(a) Prove that the composition of the projections $\pi_x,\pi_y\colon\mathbb{R}^3\to\mathbb{R}^3$ onto the x and y axes is the zero map despite that neither one is itself the zero map.

(b) Prove that the composition of the derivatives d^2/dx^2, $d^3/dx^3\colon\mathcal{P}_4\to\mathcal{P}_4$ is the zero map despite that neither is the zero map.

(c) Give a matrix equation representing the first fact.

(d) Give a matrix equation representing the second.

When two things multiply to give zero despite that neither is zero, each is said to be a *zero divisor*.

2.31 Show that, for square matrices, $(S+T)(S-T)$ need not equal S^2-T^2.

✓ **2.32** Represent the identity transformation $\text{id}\colon V\to V$ with respect to B,B for any basis B. This is the *identity matrix* I. Show that this matrix plays the role in matrix multiplication that the number 1 plays in real number multiplication: $HI=IH=H$ (for all matrices H for which the product is defined).

2.33 In real number algebra, quadratic equations have at most two solutions. That is not so with matrix algebra. Show that the 2×2 matrix equation $T^2=I$ has more than two solutions, where I is the identity matrix (this matrix has ones in its $1,1$ and $2,2$ entries and zeroes elsewhere; see Exercise 32).

2.34 **(a)** Prove that for any 2×2 matrix T there are scalars $c_0,\ldots,c_4$ such that the combination $c_4T^4+c_3T^3+c_2T^2+c_1T+I$ is the zero matrix (where I is the 2×2 identity matrix, with ones in its $1,1$ and $2,2$ entries and zeroes elsewhere; see Exercise 32).

(b) Let $p(x)$ be a polynomial $p(x)=c_nx^n+\cdots+c_1x+c_0$. If T is a square matrix we define $p(T)$ to be the matrix $c_nT^n+\cdots+c_1T+I$ (where I is the appropriately-sized identity matrix). Prove that for any square matrix there is a polynomial such that $p(T)$ is the zero matrix.

(c) The *minimal polynomial* $m(x)$ of a square matrix is the polynomial of least degree, and with leading coefficient 1, such that $m(T)$ is the zero matrix. Find the minimal polynomial of this matrix.

$$\begin{pmatrix}\sqrt{3}/2 & -1/2\\ 1/2 & \sqrt{3}/2\end{pmatrix}$$

(This is the representation with respect to $\mathcal{E}_2, \mathcal{E}_2$, the standard basis, of a rotation through $\pi/6$ radians counterclockwise.)

2.35 The infinite-dimensional space $\mathcal{P}$ of all finite-degree polynomials gives a memorable example of the non-commutativity of linear maps. Let $d/dx\colon \mathcal{P} \to \mathcal{P}$ be the usual derivative and let $s\colon \mathcal{P} \to \mathcal{P}$ be the *shift* map.

$$a_0 + a_1x + \cdots + a_nx^n \stackrel{s}{\longmapsto} 0 + a_0x + a_1x^2 + \cdots + a_nx^{n+1}$$

Show that the two maps don't commute $d/dx \circ s \neq s \circ d/dx$; in fact, not only is $(d/dx \circ s) - (s \circ d/dx)$ not the zero map, it is the identity map.

2.36 Recall the notation for the sum of the sequence of numbers $a_1, a_2, \ldots, a_n$.

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n$$

In this notation, the i, j entry of the product of G and H is this.

$$p_{i,j} = \sum_{k=1}^{r} g_{i,k}h_{k,j}$$

Using this notation,

(a) reprove that matrix multiplication is associative;

(b) reprove Theorem 2.6.

IV.3 Mechanics of Matrix Multiplication

In this subsection we consider matrix multiplication as a mechanical process, putting aside for the moment any implications about the underlying maps. As described earlier, the striking thing about matrix multiplication is the way rows and columns combine. The i, j entry of the matrix product is the dot product of row i of the left matrix with column j of the right one. For instance, here a second row and a third column combine to make a $2, 3$ entry.

$$\begin{pmatrix} 1 & 1 \\ \boxed{0 \quad 1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 4 & 6 & \boxed{\begin{matrix} 8 \\ 9 \end{matrix}} & 2 \\ 5 & 7 & & 3 \end{pmatrix} = \begin{pmatrix} 9 & 13 & 17 & 5 \\ 5 & 7 & \boxed{9} & 3 \\ 4 & 6 & 8 & 2 \end{pmatrix}$$

We can view this as the left matrix acting by multiplying its rows, one at a time, into the columns of the right matrix. Of course, another perspective is that the right matrix uses its columns to act on the left matrix's rows. Below, we will examine actions from the left and from the right for some simple matrices.

The first case, the action of a zero matrix, is very easy.

3.1 Example Multiplying by an appropriately-sized zero matrix from the left or from the right

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

results in a zero matrix.

After zero matrices, the matrices whose actions are easiest to understand are the ones with a single nonzero entry.

3.2 Definition A matrix with all zeroes except for a one in the i, j entry is an i, j *unit* matrix.

3.3 Example This is the $1, 2$ unit matrix with three rows and two columns, multiplying from the left.

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 7 & 8 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Acting from the left, an i, j unit matrix copies row j of the multiplicand into row i of the result. From the right an i, j unit matrix copies column i of the multiplicand into column j of the result.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 4 \\ 0 & 7 \end{pmatrix}$$

3.4 Example Rescaling these matrices simply rescales the result. This is the action from the left of the matrix that is twice the one in the prior example.

$$\begin{pmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 14 & 16 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

And this is the action of the matrix that is minus three times the one from the prior example.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -3 \\ 0 & -12 \\ 0 & -21 \end{pmatrix}$$

Next in complication are matrices with two nonzero entries. There are two cases. If a left-multiplier has entries in different rows then their actions don't interact.

3.5 Example

$$\begin{aligned}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} &= \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \right) \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 14 & 16 & 18 \\ 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & 3 \\ 14 & 16 & 18 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

But if the left-multiplier's nonzero entries are in the same row then that row of the result is a combination.

3.6 Example

$$\begin{aligned}
\begin{pmatrix}1&0&2\\0&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix} &= (\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}+\begin{pmatrix}0&0&2\\0&0&0\\0&0&0\end{pmatrix})\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix} \\
&= \begin{pmatrix}1&2&3\\0&0&0\\0&0&0\end{pmatrix}+\begin{pmatrix}14&16&18\\0&0&0\\0&0&0\end{pmatrix} \\
&= \begin{pmatrix}15&18&21\\0&0&0\\0&0&0\end{pmatrix}
\end{aligned}$$

Right-multiplication acts in the same way, with columns.

These observations about matrices that are mostly zeroes extend to arbitrary matrices.

3.7 Lemma In a product of two matrices G and H, the columns of GH are formed by taking G times the columns of H

$$G\cdot\left(\begin{array}{c|c|c}\vdots&&\vdots\\ \vec{h}_1&\cdots&\vec{h}_n\\ \vdots&&\vdots\end{array}\right)=\left(\begin{array}{c|c|c}\vdots&&\vdots\\ G\cdot\vec{h}_1&\cdots&G\cdot\vec{h}_n\\ \vdots&&\vdots\end{array}\right)$$

and the rows of GH are formed by taking the rows of G times H

$$\left(\begin{array}{c}\cdots\ \vec{g}_1\ \cdots\\ \hline \vdots\\ \hline \cdots\ \vec{g}_r\ \cdots\end{array}\right)\cdot H=\left(\begin{array}{c}\cdots\ \vec{g}_1\cdot H\ \cdots\\ \hline \vdots\\ \hline \cdots\ \vec{g}_r\cdot H\ \cdots\end{array}\right)$$

(ignoring the extra parentheses).

Proof. We will show the 2×2 case and leave the general case as an exercise.

$$GH=\begin{pmatrix}g_{1,1}&g_{1,2}\\g_{2,1}&g_{2,2}\end{pmatrix}\begin{pmatrix}h_{1,1}&h_{1,2}\\h_{2,1}&h_{2,2}\end{pmatrix}=\begin{pmatrix}g_{1,1}h_{1,1}+g_{1,2}h_{2,1}&g_{1,1}h_{1,2}+g_{1,2}h_{2,2}\\g_{2,1}h_{1,1}+g_{2,2}h_{2,1}&g_{2,1}h_{1,2}+g_{2,2}h_{2,2}\end{pmatrix}$$

The right side of the first equation in the result

$$\left(G\begin{pmatrix}h_{1,1}\\h_{2,1}\end{pmatrix}\;\middle|\;G\begin{pmatrix}h_{1,2}\\h_{2,2}\end{pmatrix}\right)=\left(\begin{pmatrix}g_{1,1}h_{1,1}+g_{1,2}h_{2,1}\\g_{2,1}h_{1,1}+g_{2,2}h_{2,1}\end{pmatrix}\;\middle|\;\begin{pmatrix}g_{1,1}h_{1,2}+g_{1,2}h_{2,2}\\g_{2,1}h_{1,2}+g_{2,2}h_{2,2}\end{pmatrix}\right)$$

is indeed the same as the right side of GH, except for the extra parentheses (the ones marking the columns as column vectors). The other equation is similarly easy to recognize. QED

An application of those observations is that there is a matrix that just copies out the rows and columns.

3.8 Definition The *main diagonal* (or *principle diagonal* or *diagonal*) of a square matrix goes from the upper left to the lower right.

3.9 Definition An *identity matrix* is square and has with all entries zero except for ones in the main diagonal.

$$I_{n\times n}=\begin{pmatrix}1&0&\ldots&0\\0&1&\ldots&0\\&\vdots&&\\0&0&\ldots&1\end{pmatrix}$$

3.10 Example The 3×3 identity leaves its multiplicand unchanged both from the left

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}2&3&6\\1&3&8\\-7&1&0\end{pmatrix}=\begin{pmatrix}2&3&6\\1&3&8\\-7&1&0\end{pmatrix}$$

and from the right.

$$\begin{pmatrix}2&3&6\\1&3&8\\-7&1&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}2&3&6\\1&3&8\\-7&1&0\end{pmatrix}$$

3.11 Example So does the 2×2 identity matrix.

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

In short, an identity matrix is the identity element of the set of $n\times n$ matrices with respect to the operation of matrix multiplication.

We next see two ways to generalize the identity matrix.

The first is that if the ones are relaxed to arbitrary reals, the resulting matrix will rescale whole rows or columns.

3.12 Definition A *diagonal matrix* is square and has zeros off the main diagonal.

$$\begin{pmatrix}a_{1,1}&0&\ldots&0\\0&a_{2,2}&\ldots&0\\&\vdots&&\\0&0&\ldots&a_{n,n}\end{pmatrix}$$

3.13 Example From the left, the action of multiplication by a diagonal matrix is to rescales the rows.

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

From the right such a matrix rescales the columns.

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 3 & 4 & -2 \\ 6 & 4 & -4 \end{pmatrix}$$

The second generalization of identity matrices is that we can put a single one in each row and column in ways other than putting them down the diagonal.

3.14 Definition A *permutation matrix* is square and is all zeros except for a single one in each row and column.

3.15 Example From the left these matrices permute rows.

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

From the right they permute columns.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 1 \\ 5 & 6 & 4 \\ 8 & 9 & 7 \end{pmatrix}$$

We finish this subsection by applying these observations to get matrices that perform Gauss' method and Gauss-Jordan reduction.

3.16 Example We have seen how to produce a matrix that will rescale rows. Multiplying by this diagonal matrix rescales the second row of the other by a factor of three.

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

We have seen how to produce a matrix that will swap rows. Multiplying by this permutation matrix swaps the first and third rows.

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

To see how to perform a pivot, we observe something about those two examples. The matrix that rescales the second row by a factor of three arises in this way from the identity.

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \xrightarrow{3\rho_2} \begin{pmatrix}1&0&0\\0&3&0\\0&0&1\end{pmatrix}$$

Similarly, the matrix that swaps first and third rows arises in this way.

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \xrightarrow{\rho_1\leftrightarrow\rho_3} \begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$$

3.17 Example The 3×3 matrix that arises as

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

will, when it acts from the left, perform the pivot operation $-2\rho_2+\rho_3$.

$$\begin{pmatrix}1&0&0\\0&1&0\\0&-2&1\end{pmatrix}\begin{pmatrix}1&0&2&0\\0&1&3&-3\\0&2&1&1\end{pmatrix} = \begin{pmatrix}1&0&2&0\\0&1&3&-3\\0&0&-5&7\end{pmatrix}$$

3.18 Definition The *elementary reduction matrices* are obtained from identity matrices with one Gaussian operation. We denote them:

(1) $I \xrightarrow{k\rho_i} M_i(k)$ for $k \neq 0$;

(2) $I \xrightarrow{\rho_i\leftrightarrow\rho_j} P_{i,j}$ for $i \neq j$;

(3) $I \xrightarrow{k\rho_i+\rho_j} C_{i,j}(k)$ for $i \neq j$.

3.19 Lemma Gaussian reduction can be done through matrix multiplication.

(1) If $H \xrightarrow{k\rho_i} G$ then $M_i(k)H = G$.

(2) If $H \xrightarrow{\rho_i\leftrightarrow\rho_j} G$ then $P_{i,j}H = G$.

(3) If $H \xrightarrow{k\rho_i+\rho_j} G$ then $C_{i,j}(k)H = G$.

Proof. Clear. QED

3.20 Example This is the first system, from the first chapter, on which we performed Gauss' method.

$$\begin{aligned} 3x_3 &= 9 \\ x_1 + 5x_2 - 2x_3 &= 2 \\ (1/3)x_1 + 2x_2 \qquad &= 3 \end{aligned}$$

It can be reduced with matrix multiplication. Swap the first and third rows,

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \left(\begin{array}{ccc|c} 0 & 0 & 3 & 9 \\ 1 & 5 & -2 & 2 \\ 1/3 & 2 & 0 & 3 \end{array}\right) = \left(\begin{array}{ccc|c} 1/3 & 2 & 0 & 3 \\ 1 & 5 & -2 & 2 \\ 0 & 0 & 3 & 9 \end{array}\right)$$

triple the first row,

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|c} 1/3 & 2 & 0 & 3 \\ 1 & 5 & -2 & 2 \\ 0 & 0 & 3 & 9 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 1 & 5 & -2 & 2 \\ 0 & 0 & 3 & 9 \end{array}\right)$$

and then add -1 times the first row to the second.

$$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 1 & 5 & -2 & 2 \\ 0 & 0 & 3 & 9 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 0 & -1 & -2 & -7 \\ 0 & 0 & 3 & 9 \end{array}\right)$$

Now back substitution will give the solution.

3.21 Example Gauss-Jordan reduction works the same way. For the matrix ending the prior example, first adjust the leading entries

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 0 & -1 & -2 & -7 \\ 0 & 0 & 3 & 9 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

and to finish, clear the third column and then the second column.

$$\begin{pmatrix} 1 & -6 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 1 & 3 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

We have observed the following result, which we shall use in the next subsection.

3.22 Corollary For any matrix H there are elementary reduction matrices $R_1, \ldots, R_r$ such that $R_r \cdot R_{r-1} \cdots R_1 \cdot H$ is in reduced echelon form.

Until now we have taken the point of view that our primary objects of study are vector spaces and the maps between them, and have adopted matrices only for computational convenience. This subsection show that this point of view isn't the whole story. Matrix theory is a fascinating and fruitful area.

In the rest of this book we shall continue to focus on maps as the primary objects, but we will be pragmatic — if the matrix point of view gives some clearer idea then we shall use it.

Exercises

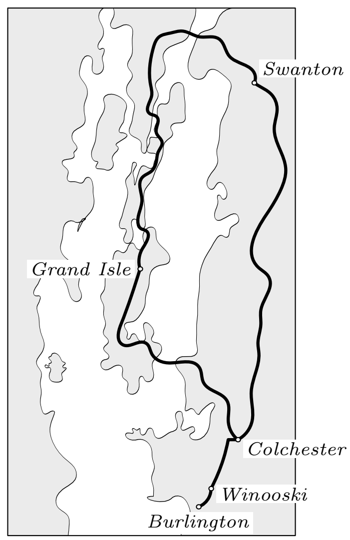
✓ **3.23** Predict the result of each multiplication by an elementary reduction matrix, and then check by multiplying it out.

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

(d) $\begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}1&-1\\0&1\end{pmatrix}$ **(e)** $\begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}$

✓ **3.24** The need to take linear combinations of rows and columns in tables of numbers arises often in practice. For instance, this is a map of part of Vermont and New York.

In part because of Lake Champlain, there are no roads directly connecting some pairs of towns. For instance, there is no way to go from Winooski to Grand Isle without going through Colchester. (Of course, many other roads and towns have been left off to simplify the graph. From top to bottom of this map is about forty miles.)



(a) The *incidence matrix* of a map is the square matrix whose i, j entry is the number of roads from city i to city j. Produce the incidence matrix of this map (take the cities in alphabetical order).

(b) A matrix is *symmetric* if it equals its transpose. Show that an incidence matrix is symmetric. (These are all two-way streets. Vermont doesn't have many one-way streets.)

(c) What is the significance of the square of the incidence matrix? The cube?

✓ **3.25** This table gives the number of hours of each type done by each worker, and the associated pay rates. Use matrices to compute the wages due.

	regular	*overtime*
Alan	40	12
Betty	35	6
Catherine	40	18
Donald	28	0

	wage
regular	\$25.00
overtime	\$45.00

(*Remark.* This illustrates, as did the prior problem, that in practice we often want to compute linear combinations of rows and columns in a context where we really aren't interested in any associated linear maps.)

3.26 Find the product of this matrix with its transpose.

$$\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}$$

HW ✓ **3.27** Prove that the diagonal matrices form a subspace of $\mathcal{M}_{n\times n}$. What is its dimension?

3.28 Does the identity matrix represent the identity map if the bases are unequal?

3.29 Show that every multiple of the identity commutes with every square matrix. Are there other matrices that commute with all square matrices?

HW **3.30** Prove or disprove: nonsingular matrices commute.

HW ✓ **3.31** Show that the product of a permutation matrix and its transpose is an identity matrix.

3.32 Show that if the first and second rows of G are equal then so are the first and second rows of GH. Generalize.

HW **3.33** Describe the product of two diagonal matrices.

HW **3.34** Write

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

as the product of two elementary reduction matrices.

✓ **3.35** Show that if G has a row of zeros then GH (if defined) has a row of zeros. Does that work for columns?

3.36 Show that the set of unit matrices forms a basis for $\mathcal{M}_{n\times m}$.

HW **3.37** Find the formula for the n-th power of this matrix.

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

HW ✓ **3.38** The *trace* of a square matrix is the sum of the entries on its diagonal (its significance appears in Chapter Five). Show that $\text{trace}(GH) = \text{trace}(HG)$.

HW ✓ **3.39** A square matrix is *upper triangular* if its only nonzero entries lie above, or on, the diagonal. Show that the product of two upper triangular matrices is upper triangular. Does this hold for lower triangular also?

3.40 A square matrix is a *Markov matrix* if each entry is between zero and one and the sum along each row is one. Prove that a product of Markov matrices is Markov.

✓ **3.41** Give an example of two matrices of the same rank with squares of differing rank.

3.42 Combine the two generalizations of the identity matrix, the one allowing entires to be other than ones, and the one allowing the single one in each row and column to be off the diagonal. What is the action of this type of matrix?

3.43 On a computer multiplications are more costly than additions, so people are interested in reducing the number of multiplications used to compute a matrix product.

(a) How many real number multiplications are needed in formula we gave for the product of a $m\times r$ matrix and a $r\times n$ matrix?

(b) Matrix multiplication is associative, so all associations yield the same result. The cost in number of multiplications, however, varies. Find the association requiring the fewest real number multiplications to compute the matrix product of a 5×10 matrix, a 10×20 matrix, a 20×5 matrix, and a 5×1 matrix.

(c) *(Very hard.)* Find a way to multiply two 2×2 matrices using only seven multiplications instead of the eight suggested by the naive approach.

? **3.44** If A and B are square matrices of the same size such that $ABAB = 0$, does it follow that $BABA = 0$? [Putnam, 1990, A-5]

3.45 Demonstrate these four assertions to get an alternate proof that column rank equals row rank. [Am. Math. Mon., Dec. 1966]

(a) $\vec{y}\cdot\vec{y}=\vec{0}$ iff $\vec{y}=\vec{0}$.
(b) $A\vec{x}=\vec{0}$ iff $A^{\text{trans}}A\vec{x}=\vec{0}$.
(c) $\dim(\mathscr{R}(A))=\dim(\mathscr{R}(A^{\text{trans}}A))$.
(d) $\text{col rank}(A)=\text{col rank}(A^{\text{trans}})=\text{row rank}(A)$.

3.46 Prove (where A is an $n\times n$ matrix and so defines a transformation of any n-dimensional space V with respect to B,B where B is a basis) that $\dim(\mathscr{R}(A)\cap\mathscr{N}(A))=\dim(\mathscr{R}(A))-\dim(\mathscr{R}(A^2))$. Conclude

(a) $\mathscr{N}(A)\subset\mathscr{R}(A)$ iff $\dim(\mathscr{N}(A))=\dim(\mathscr{R}(A))-\dim(\mathscr{R}(A^2))$;
(b) $\mathscr{R}(A)\subseteq\mathscr{N}(A)$ iff $A^2=0$;
(c) $\mathscr{R}(A)=\mathscr{N}(A)$ iff $A^2=0$ and $\dim(\mathscr{N}(A))=\dim(\mathscr{R}(A))$;
(d) $\dim(\mathscr{R}(A)\cap\mathscr{N}(A))=0$ iff $\dim(\mathscr{R}(A))=\dim(\mathscr{R}(A^2))$;
(e) *(Requires the Direct Sum subsection, which is optional.)* $V=\mathscr{R}(A)\oplus\mathscr{N}(A)$ iff $\dim(\mathscr{R}(A))=\dim(\mathscr{R}(A^2))$.

[Ackerson]

IV.4 Inverses

We now consider how to represent the inverse of a linear map.

We start by recalling some facts about function inverses.* Some functions have no inverse, or have an inverse on the left side or right side only.

4.1 Example Where $\pi\colon\mathbb{R}^3\to\mathbb{R}^2$ is the projection map

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

and $\eta\colon\mathbb{R}^2\to\mathbb{R}^3$ is the embedding

$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x\\y\\0\end{pmatrix}$$

the composition $\pi\circ\eta$ is the identity map on $\mathbb{R}^2$.

$$\begin{pmatrix}x\\y\end{pmatrix}\overset{\eta}{\longmapsto}\begin{pmatrix}x\\y\\0\end{pmatrix}\overset{\pi}{\longmapsto}\begin{pmatrix}x\\y\end{pmatrix}$$

We say π is a *left inverse map* of η or, what is the same thing, that η is a *right inverse map* of π. However, composition in the other order $\eta\circ\pi$ doesn't give the identity map — here is a vector that is not sent to itself under $\eta\circ\pi$.

$$\begin{pmatrix}0\\0\\1\end{pmatrix}\overset{\pi}{\longmapsto}\begin{pmatrix}0\\0\end{pmatrix}\overset{\eta}{\longmapsto}\begin{pmatrix}0\\0\\0\end{pmatrix}$$

* More information on function inverses is in the appendix.

In fact, the projection π has no left inverse at all. For, if f were to be a left inverse of π then we would have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \end{pmatrix} \stackrel{f}{\longmapsto} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

for all of the infinitely many z's. But no function f can send a single argument to more than one value.

(An example of a function with no inverse on either side is the zero transformation on $\mathbb{R}^2$.) Some functions have a *two-sided inverse map*, another function that is the inverse of the first, both from the left and from the right. For instance, the map given by $\vec{v} \mapsto 2 \cdot \vec{v}$ has the two-sided inverse $\vec{v} \mapsto (1/2) \cdot \vec{v}$. In this subsection we will focus on two-sided inverses. The appendix shows that a function has a two-sided inverse if and only if it is both one-to-one and onto. The appendix also shows that if a function f has a two-sided inverse then it is unique, and so it is called 'the' inverse, and is denoted f^{-1}. So our purpose in this subsection is, where a linear map h has an inverse, to find the relationship between $\text{Rep}_{B,D}(h)$ and $\text{Rep}_{D,B}(h^{-1})$ (recall that we have shown, in Theorem 2.21 of Section II of this chapter, that if a linear map has an inverse then the inverse is a linear map also).

4.2 Definition A matrix G is a *left inverse matrix* of the matrix H if GH is the identity matrix. It is a *right inverse matrix* if HG is the identity. A matrix H with a two-sided inverse is an *invertible matrix*. That two-sided inverse is called *the inverse matrix* and is denoted H^{-1}.

Because of the correspondence between linear maps and matrices, statements about map inverses translate into statements about matrix inverses.

4.3 Lemma If a matrix has both a left inverse and a right inverse then the two are equal.

4.4 Theorem A matrix is invertible if and only if it is nonsingular.

Proof. *(For both results.)* Given a matrix H, fix spaces of appropriate dimension for the domain and codomain. Fix bases for these spaces. With respect to these bases, H represents a map h. The statements are true about the map and therefore they are true about the matrix. QED

4.5 Lemma A product of invertible matrices is invertible—if G and H are invertible and if GH is defined then GH is invertible and $(GH)^{-1} = H^{-1}G^{-1}$.

Proof. *(This is just like the prior proof except that it requires two maps.)* Fix appropriate spaces and bases and consider the represented maps h and g. Note that $h^{-1}g^{-1}$ is a two-sided map inverse of gh since $(h^{-1}g^{-1})(gh) = h^{-1}(\text{id})h = h^{-1}h = \text{id}$ and $(gh)(h^{-1}g^{-1}) = g(\text{id})g^{-1} = gg^{-1} = \text{id}$. This equality is reflected in the matrices representing the maps, as required. QED

Here is the arrow diagram giving the relationship between map inverses and matrix inverses. It is a special case of the diagram for function composition and matrix multiplication.

$$\begin{tikzcd} & W_{wrt\ C} \arrow[dr, "h^{-1}", "H^{-1}"'] & \\ V_{wrt\ B} \arrow[ur, "h", "H"'] \arrow[rr, "\mathrm{id}", "I"'] & & V_{wrt\ B} \end{tikzcd}$$

Beyond its place in our general program of seeing how to represent map operations, another reason for our interest in inverses comes from solving linear systems. A linear system is equivalent to a matrix equation, as here.

$$\begin{aligned} x_1 + x_2 &= 3 \\ 2x_1 - x_2 &= 2 \end{aligned} \quad \Longleftrightarrow \quad \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \qquad (*)$$

By fixing spaces and bases (e.g., $\mathbb{R}^2, \mathbb{R}^2$ and $\mathcal{E}_2, \mathcal{E}_2$), we take the matrix H to represent some map h. Then solving the system is the same as asking: what domain vector $\vec{x}$ is mapped by h to the result $\vec{d}$? If we could invert h then we could solve the system by multiplying $\mathrm{Rep}_{D,B}(h^{-1}) \cdot \mathrm{Rep}_D(\vec{d})$ to get $\mathrm{Rep}_B(\vec{x})$.

4.6 Example We can find a left inverse for the matrix just given

$$\begin{pmatrix} m & n \\ p & q \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

by using Gauss' method to solve the resulting linear system.

$$\begin{array}{rcl} m + 2n & & = 1 \\ m - n & & = 0 \\ & p + 2q & = 0 \\ & p - q & = 1 \end{array}$$

Answer: $m = 1/3$, $n = 1/3$, $p = 2/3$, and $q = -1/3$. This matrix is actually the two-sided inverse of H, as can easily be checked. With it we can solve the system $(*)$ above by applying the inverse.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5/3 \\ 4/3 \end{pmatrix}$$

4.7 Remark Why solve systems this way, when Gauss' method takes less arithmetic (this assertion can be made precise by counting the number of arithmetic operations, as computer algorithm designers do)? Beyond its conceptual appeal of fitting into our program of discovering how to represent the various map operations, solving linear systems by using the matrix inverse has at least two advantages.

First, once the work of finding an inverse has been done, solving a system with the same coefficients but different constants is easy and fast: if we change the entries on the right of the system $(*)$ then we get a related problem

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

wtih a related solution method.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

In applications, solving many systems having the same matrix of coefficients is common.

Another advantage of inverses is that we can explore a system's sensitivity to changes in the constants. For example, tweaking the 3 on the right of the system $(*)$ to

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3.01 \\ 2 \end{pmatrix}$$

can be solved with the inverse.

$$\begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix} \begin{pmatrix} 3.01 \\ 2 \end{pmatrix} = \begin{pmatrix} (1/3)(3.01) + (1/3)(2) \\ (2/3)(3.01) - (1/3)(2) \end{pmatrix}$$

to show that x_1 changes by $1/3$ of the tweak while x_2 moves by $2/3$ of that tweak. This sort of analysis is used, for example, to decide how accurately data must be specified in a linear model to ensure that the solution has a desired accuracy.

We finish by describing the computational procedure usually used to find the inverse matrix.

4.8 Lemma A matrix is invertible if and only if it can be written as the product of elementary reduction matrices. The inverse can be computed by applying to the identity matrix the same row steps, in the same order, as are used to Gauss-Jordan reduce the invertible matrix.

Proof. A matrix H is invertible if and only if it is nonsingular and thus Gauss-Jordan reduces to the identity. By Corollary 3.22 this reduction can be done with elementary matrices $R_r \cdot R_{r-1} \dots R_1 \cdot H = I$. This equation gives the two halves of the result.

First, elementary matrices are invertible and their inverses are also elementary. Applying R_r^{-1} to the left of both sides of that equation, then R_{r-1}^{-1}, etc., gives H as the product of elementary matrices $H = R_1^{-1} \cdots R_r^{-1} \cdot I$ (the I is here to cover the trivial $r = 0$ case).

Second, matrix inverses are unique and so comparison of the above equation with $H^{-1}H = I$ shows that $H^{-1} = R_r \cdot R_{r-1} \dots R_1 \cdot I$. Therefore, applying R_1 to the identity, followed by R_2, etc., yields the inverse of H. QED

4.9 Example To find the inverse of

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$$

we do Gauss-Jordan reduction, meanwhile performing the same operations on the identity. For clerical convenience we write the matrix and the identity side-by-side, and do the reduction steps together.

$$\begin{aligned}
\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 2 & -1 & 0 & 1 \end{array}\right) &\xrightarrow{-2\rho_1+\rho_2} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & -3 & -2 & 1 \end{array}\right) \\
&\xrightarrow{-1/3\rho_2} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & 2/3 & -1/3 \end{array}\right) \\
&\xrightarrow{-\rho_2+\rho_1} \left(\begin{array}{cc|cc} 1 & 0 & 1/3 & 1/3 \\ 0 & 1 & 2/3 & -1/3 \end{array}\right)
\end{aligned}$$

This calculation has found the inverse.

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix}$$

4.10 Example This one happens to start with a row swap.

$$\begin{aligned}
\left(\begin{array}{ccc|ccc} 0 & 3 & -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right) &\xrightarrow{\rho_1\leftrightarrow\rho_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right) \\
&\xrightarrow{-\rho_1+\rho_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & -1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & -1 & 1 \end{array}\right) \\
&\quad\vdots \\
&\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/4 & 1/4 & 3/4 \\ 0 & 1 & 0 & 1/4 & 1/4 & -1/4 \\ 0 & 0 & 1 & -1/4 & 3/4 & -3/4 \end{array}\right)
\end{aligned}$$

4.11 Example A non-invertible matrix is detected by the fact that the left half won't reduce to the identity.

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{array}\right) \xrightarrow{-2\rho_1+\rho_2} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right)$$

This procedure will find the inverse of a general $n\times n$ matrix. The 2×2 case is handy.

4.12 Corollary The inverse for a 2×2 matrix exists and equals

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

if and only if $ad - bc \neq 0$.

Proof. This computation is Exercise 22. QED

We have seen here, as in the Mechanics of Matrix Multiplication subsection, that we can exploit the correspondence between linear maps and matrices. So we can fruitfully study both maps and matrices, translating back and forth to whichever helps us the most.

Over the entire four subsections of this section we have developed an algebra system for matrices. We can compare it with the familiar algebra system for the real numbers. Here we are working not with numbers but with matrices. We have matrix addition and subtraction operations, and they work in much the same way as the real number operations, except that they only combine same-sized matrices. We also have a matrix multiplication operation and an operation inverse to multiplication. These are somewhat like the familiar real number operations (associativity, and distributivity over addition, for example), but there are differences (failure of commutativity, for example). And, we have scalar multiplication, which is in some ways another extension of real number multiplication. This matrix system provides an example that algebra systems other than the elementary one can be interesting and useful.

Exercises

4.13 Supply the intermediate steps in Example 4.10.

✓ **4.14** Use Corollary 4.12 to decide if each matrix has an inverse.

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

✓ **4.15** For each invertible matrix in the prior problem, use Corollary 4.12 to find its inverse.

✓ **4.16** Find the inverse, if it exists, by using the Gauss-Jordan method. Check the answers for the 2×2 matrices with Corollary 4.12.

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

(e) $\begin{pmatrix} 0 & 1 & 5 \\ 0 & -2 & 4 \\ 2 & 3 & -2 \end{pmatrix}$ **(f)** $\begin{pmatrix} 2 & 2 & 3 \\ 1 & -2 & -3 \\ 4 & -2 & -3 \end{pmatrix}$

✓ **4.17** What matrix has this one for its inverse?

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

4.18 How does the inverse operation interact with scalar multiplication and addition of matrices?

(a) What is the inverse of rH?

(b) Is $(H+G)^{-1} = H^{-1} + G^{-1}$?

✓ **4.19** Is $(T^k)^{-1} = (T^{-1})^k$?

4.20 Is H^{-1} invertible?

4.21 For each real number θ let $t_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ be represented with respect to the standard bases by this matrix.

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Show that $t_{\theta_1+\theta_2} = t_{\theta_1} \cdot t_{\theta_2}$. Show also that ${t_\theta}^{-1} = t_{-\theta}$.

4.22 Do the calculations for the proof of Corollary 4.12.

4.23 Show that this matrix

$$H = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

has infinitely many right inverses. Show also that it has no left inverse.

4.24 In Example 4.1, how many left inverses has η?

4.25 If a matrix has infinitely many right-inverses, can it have infinitely many left-inverses? Must it have?

✓ **4.26** Assume that H is invertible and that HG is the zero matrix. Show that G is a zero matrix.

4.27 Prove that if H is invertible then the inverse commutes with a matrix $GH^{-1} = H^{-1}G$ if and only if H itself commutes with that matrix $GH = HG$.

✓ **4.28** Show that if T is square and if T^4 is the zero matrix then $(I-T)^{-1} = I + T + T^2 + T^3$. Generalize.

✓ **4.29** Let D be diagonal. Describe D^2, D^3, ... , etc. Describe D^{-1}, D^{-2}, ... , etc. Define D^0 appropriately.

4.30 Prove that any matrix row-equivalent to an invertible matrix is also invertible.

4.31 *The first question below appeared as Exercise 28.*

(a) Show that the rank of the product of two matrices is less than or equal to the minimum of the rank of each.

(b) Show that if T and S are square then $TS = I$ if and only if $ST = I$.

4.32 Show that the inverse of a permutation matrix is its transpose.

4.33 *The first two parts of this question appeared as Exercise 25.*

(a) Show that $(GH)^{\text{trans}} = H^{\text{trans}}G^{\text{trans}}$.

(b) A square matrix is *symmetric* if each i,j entry equals the j,i entry (that is, if the matrix equals its transpose). Show that the matrices HH^{trans} and $H^{\text{trans}}H$ are symmetric.

(c) Show that the inverse of the transpose is the transpose of the inverse.

(d) Show that the inverse of a symmetric matrix is symmetric.

✓ **4.34** *The items starting this question appeared as Exercise 30.*

(a) Prove that the composition of the projections $\pi_x, \pi_y \colon \mathbb{R}^3 \to \mathbb{R}^3$ is the zero map despite that neither is the zero map.

(b) Prove that the composition of the derivatives d^2/dx^2, $d^3/dx^3 \colon \mathcal{P}_4 \to \mathcal{P}_4$ is the zero map despite that neither map is the zero map.

(c) Give matrix equations representing each of the prior two items.

When two things multiply to give zero despite that neither is zero, each is said to be a *zero divisor*. Prove that no zero divisor is invertible.

4.35 In real number algebra, there are exactly two numbers, 1 and -1, that are their own multiplicative inverse. Does $H^2 = I$ have exactly two solutions for 2×2 matrices?

4.36 Is the relation 'is a two-sided inverse of' transitive? Reflexive? Symmetric?

4.37 Prove: if the sum of the elements of a square matrix is k, then the sum of the elements in each row of the inverse matrix is $1/k$. [Am. Math. Mon., Nov. 1951]

V Change of Basis

Representations, whether of vectors or of maps, vary with the bases. For instance, with respect to the two bases $\mathcal{E}_2$ and

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

for $\mathbb{R}^2$, the vector $\vec{e}_1$ has two different representations.

$$\text{Rep}_{\mathcal{E}_2}(\vec{e}_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{Rep}_B(\vec{e}_1) = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

Similarly, with respect to $\mathcal{E}_2, \mathcal{E}_2$ and $\mathcal{E}_2, B$, the identity map has two different representations.

$$\text{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(\text{id}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{Rep}_{\mathcal{E}_2,B}(\text{id}) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$$

With our point of view that the objects of our studies are vectors and maps, in fixing bases we are adopting a scheme of tags or names for these objects, that are convienent for computation. We will now see how to translate among these names — we will see exactly how representations vary as the bases vary.

V.1 Changing Representations of Vectors

In converting $\text{Rep}_B(\vec{v})$ to $\text{Rep}_D(\vec{v})$ the underlying vector $\vec{v}$ doesn't change. Thus, this translation is accomplished by the identity map on the space, described so that the domain space vectors are represented with respect to B and the codomain space vectors are represented with respect to D.

$$\begin{array}{c} V_{\text{w.r.t. } B} \\ \text{id}\big\downarrow \\ V_{\text{w.r.t. } D} \end{array}$$

(The diagram is vertical to fit with the ones in the next subsection.)

1.1 Definition The *change of basis matrix* for bases $B, D \subset V$ is the representation of the identity map $\text{id}\colon V \to V$ with respect to those bases.

$$\text{Rep}_{B,D}(\text{id}) = \left(\begin{array}{c|c|c} \vdots & & \vdots \\ \text{Rep}_D(\vec{\beta}_1) & \cdots & \text{Rep}_D(\vec{\beta}_n) \\ \vdots & & \vdots \end{array}\right)$$

1.2 Lemma Left-multiplication by the change of basis matrix for B, D converts a representation with respect to B to one with respect to D. Conversely, if left-multiplication by a matrix changes bases $M \cdot \mathrm{Rep}_B(\vec{v}) = \mathrm{Rep}_D(\vec{v})$ then M is a change of basis matrix.

Proof. For the first sentence, for each $\vec{v}$, as matrix-vector multiplication represents a map application, $\mathrm{Rep}_{B,D}(\mathrm{id}) \cdot \mathrm{Rep}_B(\vec{v}) = \mathrm{Rep}_D(\,\mathrm{id}(\vec{v})\,) = \mathrm{Rep}_D(\vec{v})$. For the second sentence, with respect to B, D the matrix M represents some linear map, whose action is $\vec{v} \mapsto \vec{v}$, and is therefore the identity map. QED

1.3 Example With these bases for $\mathbb{R}^2$,

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

because

$$\mathrm{Rep}_D(\,\mathrm{id}(\begin{pmatrix} 2 \\ 1 \end{pmatrix})) = \begin{pmatrix} -1/2 \\ 3/2 \end{pmatrix}_D \qquad \mathrm{Rep}_D(\,\mathrm{id}(\begin{pmatrix} 1 \\ 0 \end{pmatrix})) = \begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix}_D$$

the change of basis matrix is this.

$$\mathrm{Rep}_{B,D}(\mathrm{id}) = \begin{pmatrix} -1/2 & -1/2 \\ 3/2 & 1/2 \end{pmatrix}$$

We can see this matrix at work by finding the two representations of $\vec{e}_2$

$$\mathrm{Rep}_B(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \qquad \mathrm{Rep}_D(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

and checking that the conversion goes as expected.

$$\begin{pmatrix} -1/2 & -1/2 \\ 3/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

We finish this subsection by recognizing that the change of basis matrices are familiar.

1.4 Lemma A matrix changes bases if and only if it is nonsingular.

Proof. For one direction, if left-multiplication by a matrix changes bases then the matrix represents an invertible function, simply because the function is inverted by changing the bases back. Such a matrix is itself invertible, and so nonsingular.

To finish, we will show that any nonsingular matrix M performs a change of basis operation from any given starting basis B to some ending basis. Because the matrix is nonsingular, it will Gauss-Jordan reduce to the identity, so there are elementatry reduction matrices such that $R_r \cdots R_1 \cdot M = I$. Elementary matrices are invertible and their inverses are also elementary, so multiplying from the left first by ${R_r}^{-1}$, then by ${R_{r-1}}^{-1}$, etc., gives M as a product of elementary matrices $M = {R_1}^{-1} \cdots {R_r}^{-1}$. Thus, we will be done if we show that elementary matrices change a given basis to another basis, for then ${R_r}^{-1}$ changes B to some other basis B_r, and ${R_{r-1}}^{-1}$ changes B_r to some B_{r-1}, ..., and the net effect is that M changes B to B_1. We will prove this about elementary matrices by covering the three types as separate cases.

Applying a row-multiplication matrix

$$M_i(k)\begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ kc_i \\ \vdots \\ c_n \end{pmatrix}$$

changes a representation with respect to $\langle \vec{\beta}_1, \ldots, \vec{\beta}_i, \ldots, \vec{\beta}_n \rangle$ to one with respect to $\langle \vec{\beta}_1, \ldots, (1/k)\vec{\beta}_i, \ldots, \vec{\beta}_n \rangle$ in this way.

$$\begin{aligned} \vec{v} = c_1 \cdot \vec{\beta}_1 + \cdots + c_i \cdot \vec{\beta}_i + \cdots + c_n \cdot \vec{\beta}_n \\ \mapsto\ c_1 \cdot \vec{\beta}_1 + \cdots + kc_i \cdot (1/k)\vec{\beta}_i + \cdots + c_n \cdot \vec{\beta}_n = \vec{v} \end{aligned}$$

Similarly, left-multiplication by a row-swap matrix $P_{i,j}$ changes a representation with respect to the basis $\langle \vec{\beta}_1, \ldots, \vec{\beta}_i, \ldots, \vec{\beta}_j, \ldots, \vec{\beta}_n \rangle$ into one with respect to the basis $\langle \vec{\beta}_1, \ldots, \vec{\beta}_j, \ldots, \vec{\beta}_i, \ldots, \vec{\beta}_n \rangle$ in this way.

$$\begin{aligned} \vec{v} = c_1 \cdot \vec{\beta}_1 + \cdots + c_i \cdot \vec{\beta}_i + \cdots + c_j\vec{\beta}_j + \cdots + c_n \cdot \vec{\beta}_n \\ \mapsto\ c_1 \cdot \vec{\beta}_1 + \cdots + c_j \cdot \vec{\beta}_j + \cdots + c_i \cdot \vec{\beta}_i + \cdots + c_n \cdot \vec{\beta}_n = \vec{v} \end{aligned}$$

And, a representation with respect to $\langle \vec{\beta}_1, \ldots, \vec{\beta}_i, \ldots, \vec{\beta}_j, \ldots, \vec{\beta}_n \rangle$ changes via left-multiplication by a row-combination matrix $C_{i,j}(k)$ into a representation with respect to $\langle \vec{\beta}_1, \ldots, \vec{\beta}_i - k\vec{\beta}_j, \ldots, \vec{\beta}_j, \ldots, \vec{\beta}_n \rangle$

$$\begin{aligned} \vec{v} = c_1 \cdot \vec{\beta}_1 + \cdots + c_i \cdot \vec{\beta}_i + c_j\vec{\beta}_j + \cdots + c_n \cdot \vec{\beta}_n \\ \mapsto\ c_1 \cdot \vec{\beta}_1 + \cdots + c_i \cdot (\vec{\beta}_i - k\vec{\beta}_j) + \cdots + (kc_i + c_j) \cdot \vec{\beta}_j + \cdots + c_n \cdot \vec{\beta}_n = \vec{v} \end{aligned}$$

(the definition of reduction matrices specifies that $i \neq k$ and $k \neq 0$ and so this last one is a basis). QED

1.5 Corollary A matrix is nonsingular if and only if it represents the identity map with respect to some pair of bases.

In the next subsection we will see how to translate among representations of maps, that is, how to change $\text{Rep}_{B,D}(h)$ to $\text{Rep}_{\hat{B},\hat{D}}(h)$. The above corollary is a special case of this, where the domain and range are the same space, and where the map is the identity map.

Exercises

✓ **1.6** In $\mathbb{R}^2$, where

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

find the change of basis matrices from D to $\mathcal{E}_2$ and from $\mathcal{E}_2$ to D. Multiply the two.

✓ **1.7** Find the change of basis matrix for $B, D \subseteq \mathbb{R}^2$.

(a) $B = \mathcal{E}_2$, $D = \langle \vec{e}_2, \vec{e}_1 \rangle$ **(b)** $B = \mathcal{E}_2$, $D = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix} \rangle$

(c) $B = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix} \rangle$, $D = \mathcal{E}_2$ **(d)** $B = \langle \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \rangle$, $D = \langle \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \rangle$

1.8 For the bases in Exercise 7, find the change of basis matrix in the other direction, from D to B.

✓ **1.9** Find the change of basis matrix for each $B, D \subseteq \mathcal{P}_2$.

(a) $B = \langle 1, x, x^2 \rangle, D = \langle x^2, 1, x \rangle$ **(b)** $B = \langle 1, x, x^2 \rangle, D = \langle 1, 1+x, 1+x+x^2 \rangle$

(c) $B = \langle 2, 2x, x^2 \rangle, D = \langle 1 + x^2, 1 - x^2, x + x^2 \rangle$

✓ **1.10** Decide if each changes bases on $\mathbb{R}^2$. To what basis is $\mathcal{E}_2$ changed?

(a) $\begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix}$ **(b)** $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$ **(c)** $\begin{pmatrix} -1 & 4 \\ 2 & -8 \end{pmatrix}$ **(d)** $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$

1.11 Find bases such that this matrix represents the identity map with respect to those bases.

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

1.12 Conside the vector space of real-valued functions with basis $\langle \sin(x), \cos(x) \rangle$. Show that $\langle 2\sin(x)+\cos(x), 3\cos(x) \rangle$ is also a basis for this space. Find the change of basis matrix in each direction.

1.13 Where does this matrix

$$\begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$$

send the standard basis for $\mathbb{R}^2$? Any other bases? *Hint.* Consider the inverse.

✓ **1.14** What is the change of basis matrix with respect to B, B?

1.15 Prove that a matrix changes bases if and only if it is invertible.

1.16 Finish the proof of Lemma 1.4.

✓ **1.17** Let H be a $n \times n$ nonsingular matrix. What basis of $\mathbb{R}^n$ does H change to the standard basis?

✓ **1.18** **(a)** In $\mathcal{P}_3$ with basis $B = \langle 1 + x, 1 - x, x^2 + x^3, x^2 - x^3 \rangle$ we have this represenatation.

$$\mathrm{Rep}_B(1 - x + 3x^2 - x^3) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}_B$$

Find a basis D giving this different representation for the same polynomial.

$$\mathrm{Rep}_D(1 - x + 3x^2 - x^3) = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}_D$$

(b) State and prove that any nonzero vector representation can be changed to any other.

Hint. The proof of Lemma 1.4 is constructive — it not only says the bases change, it shows how they change.

1.19 Let V, W be vector spaces, and let $B, \hat{B}$ be bases for V and $D, \hat{D}$ be bases for W. Where $h\colon V \to W$ is linear, find a formula relating $\text{Rep}_{B,D}(h)$ to $\text{Rep}_{\hat{B},\hat{D}}(h)$.

✓ **1.20** Show that the columns of an $n\times n$ change of basis matrix form a basis for $\mathbb{R}^n$. Do all bases appear in that way: can the vectors from any $\mathbb{R}^n$ basis make the columns of a change of basis matrix?

✓ **1.21** Find a matrix having this effect.

$$\begin{pmatrix}1\\3\end{pmatrix} \mapsto \begin{pmatrix}4\\-1\end{pmatrix}$$

That is, find a M that left-multiplies the starting vector to yield the ending vector. Is there a matrix having these two effects?

(a) $\begin{pmatrix}1\\3\end{pmatrix} \mapsto \begin{pmatrix}1\\1\end{pmatrix}$ $\begin{pmatrix}2\\-1\end{pmatrix} \mapsto \begin{pmatrix}-1\\-1\end{pmatrix}$ **(b)** $\begin{pmatrix}1\\3\end{pmatrix} \mapsto \begin{pmatrix}1\\1\end{pmatrix}$ $\begin{pmatrix}2\\6\end{pmatrix} \mapsto \begin{pmatrix}-1\\-1\end{pmatrix}$

Give a necessary and sufficient condition for there to be a matrix such that $\vec{v}_1 \mapsto \vec{w}_1$ and $\vec{v}_2 \mapsto \vec{w}_2$.

V.2 Changing Map Representations

The first subsection shows how to convert the representation of a vector with respect to one basis to the representation of that same vector with respect to another basis. Here we will see how to convert the representation of a map with respect to one pair of bases to the representation of that map with respect to a different pair. That is, we want the relationship between the matrices in this arrow diagram.

$$\begin{array}{ccc}
V_{\text{w.r.t. }B} & \xrightarrow[H]{h} & W_{\text{w.r.t. }D} \\
\text{id}\big\downarrow & & \text{id}\big\downarrow \\
V_{\text{w.r.t. }\hat{B}} & \xrightarrow[\hat{H}]{h} & W_{\text{w.r.t. }\hat{D}}
\end{array}$$

To move from the lower-left of this diagram to the lower-right we can either go straight over, or else up to V_B then over to W_D and then down. Restated in terms of the matrices, we can calculate $\hat{H} = \text{Rep}_{\hat{B},\hat{D}}(h)$ either by simply using $\hat{B}$ and $\hat{D}$, or else by first changing bases with $\text{Rep}_{\hat{B},B}(\text{id})$ then multiplying by $H = \text{Rep}_{B,D}(h)$ and then changing bases with $\text{Rep}_{D,\hat{D}}(\text{id})$. This equation summarizes.

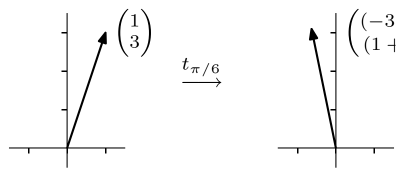
$$\hat{H} = \text{Rep}_{D,\hat{D}}(\text{id}) \cdot H \cdot \text{Rep}_{\hat{B},B}(\text{id}) \qquad (*)$$

(To compare this equation with the sentence before it, remember that the equation is read from right to left because function composition is read right to left and matrix multiplication represent the composition.)

2.1 Example The matrix

$$T=\begin{pmatrix}\cos(\pi/6)&-\sin(\pi/6)\\ \sin(\pi/6)&\cos(\pi/6)\end{pmatrix}=\begin{pmatrix}\sqrt{3}/2&-1/2\\ 1/2&\sqrt{3}/2\end{pmatrix}$$

represents, with respect to $\mathcal{E}_2,\mathcal{E}_2$, the transformation $t\colon\mathbb{R}^2\to\mathbb{R}^2$ that rotates vectors $\pi/6$ radians counterclockwise.



We can translate that representation with respect to $\mathcal{E}_2,\mathcal{E}_2$ to one with respect to

$$\hat{B}=\langle\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}0\\2\end{pmatrix}\rangle \qquad \hat{D}=\langle\begin{pmatrix}-1\\0\end{pmatrix}\begin{pmatrix}2\\3\end{pmatrix}\rangle$$

by using the arrow diagram and formula $(*)$ above.

$$\begin{tikzcd}
\mathbb{R}^2_{\text{w.r.t. }\mathcal{E}_2} \arrow[r, "t", "T"'] \arrow[d, "\mathrm{id}"'] & \mathbb{R}^2_{\text{w.r.t. }\mathcal{E}_2} \arrow[d, "\mathrm{id}"'] \\
\mathbb{R}^2_{\text{w.r.t. }\hat{B}} \arrow[r, "t", "\hat{T}"'] & \mathbb{R}^2_{\text{w.r.t. }\hat{D}}
\end{tikzcd}
\qquad \hat{T}=\mathrm{Rep}_{\mathcal{E}_2,\hat{D}}(\mathrm{id})\cdot T\cdot\mathrm{Rep}_{\hat{B},\mathcal{E}_2}(\mathrm{id})$$

Note that $\mathrm{Rep}_{\mathcal{E}_2,\hat{D}}(\mathrm{id})$ can be calculated as the matrix inverse of $\mathrm{Rep}_{\hat{D},\mathcal{E}_2}(\mathrm{id})$.

$$\begin{aligned}\mathrm{Rep}_{\hat{B},\hat{D}}(t)&=\begin{pmatrix}-1&2\\0&3\end{pmatrix}^{-1}\begin{pmatrix}\sqrt{3}/2&-1/2\\1/2&\sqrt{3}/2\end{pmatrix}\begin{pmatrix}1&0\\1&2\end{pmatrix}\\ &=\begin{pmatrix}(5-\sqrt{3})/6&(3+2\sqrt{3})/3\\(1+\sqrt{3})/6&\sqrt{3}/3\end{pmatrix}\end{aligned}$$

Although the new matrix is messier-appearing, the map that it represents is the same. For instance, to replicate the effect of t in the picture, start with $\hat{B}$,

$$\mathrm{Rep}_{\hat{B}}(\begin{pmatrix}1\\3\end{pmatrix})=\begin{pmatrix}1\\1\end{pmatrix}_{\hat{B}}$$

apply $\hat{T}$,

$$\begin{pmatrix}(5-\sqrt{3})/6&(3+2\sqrt{3})/3\\(1+\sqrt{3})/6&\sqrt{3}/3\end{pmatrix}_{\hat{B},\hat{D}}\begin{pmatrix}1\\1\end{pmatrix}_{\hat{B}}=\begin{pmatrix}(11+3\sqrt{3})/6\\(1+3\sqrt{3})/6\end{pmatrix}_{\hat{D}}$$

and check it against $\hat{D}$

$$\frac{11+3\sqrt{3}}{6}\cdot\begin{pmatrix}-1\\0\end{pmatrix}+\frac{1+3\sqrt{3}}{6}\cdot\begin{pmatrix}2\\3\end{pmatrix}=\begin{pmatrix}(-3+\sqrt{3})/2\\(1+3\sqrt{3})/2\end{pmatrix}$$

to see that it is the same result as above.

2.2 Example On $\mathbb{R}^3$ the map

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \overset{t}{\longmapsto} \begin{pmatrix} y+z \\ x+z \\ x+y \end{pmatrix}$$

that is represented with respect to the standard basis in this way

$$\mathrm{Rep}_{\mathcal{E}_3,\mathcal{E}_3}(t) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

can also be represented with respect to another basis

$$\text{if } B = \langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle \qquad \text{then } \mathrm{Rep}_{B,B}(t) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

in a way that is simpler, in that the action of a diagonal matrix is easy to understand.

Naturally, we usually prefer basis changes that make the representation easier to understand. When the representation with respect to equal starting and ending bases is a diagonal matrix we say the map or matrix has been *diagonalized*. In Chaper Five we shall see which maps and matrices are diagonalizable, and where one is not, we shall see how to get a representation that is nearly diagonal.

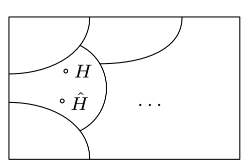
We finish this subsection by considering the easier case where representations are with respect to possibly different starting and ending bases. Recall that the prior subsection shows that a matrix changes bases if and only if it is nonsingular. That gives us another version of the above arrow diagram and equation (∗).

2.3 Definition Same-sized matrices H and $\hat{H}$ are *matrix equivalent* if there are nonsingular matrices P and Q such that $\hat{H} = PHQ$.

2.4 Corollary Matrix equivalent matrices represent the same map, with respect to appropriate pairs of bases.

Exercise 19 checks that matrix equivalence is an equivalence relation. Thus it partitions the set of matrices into matrix equivalence classes.

All matrices:



H matrix equivalent to $\hat{H}$

We can get some insight into the classes by comparing matrix equivalence with row equivalence (recall that matrices are row equivalent when they can be reduced to each other by row operations). In $\hat{H} = PHQ$, the matrices P and Q are nonsingular and thus each can be written as a product of elementary reduction matrices (Lemma 4.8). Left-multiplication by the reduction matrices making up P has the effect of performing row operations. Right-multiplication by the reduction matrices making up Q performs column operations. Therefore, matrix equivalence is a generalization of row equivalence — two matrices are row equivalent if one can be converted to the other by a sequence of row reduction steps, while two matrices are matrix equivalent if one can be converted to the other by a sequence of row reduction steps followed by a sequence of column reduction steps.

Thus, if matrices are row equivalent then they are also matrix equivalent (since we can take Q to be the identity matrix and so perform no column operations). The converse, however, does not hold.

2.5 Example These two

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

are matrix equivalent because the second can be reduced to the first by the column operation of taking -1 times the first column and adding to the second. They are not row equivalent because they have different reduced echelon forms (in fact, both are already in reduced form).

We will close this section by finding a set of representatives for the matrix equivalence classes.*

2.6 Theorem Any $m \times n$ matrix of rank k is matrix equivalent to the $m \times n$ matrix that is all zeros except that the first k diagonal entries are ones.

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ & \vdots & & & & & \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ & \vdots & & & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

Sometimes this is described as a *block partial-identity* form.

$$\left(\begin{array}{c|c} I & Z \\ \hline Z & Z \end{array}\right)$$

* More information on class representatives is in the appendix.

Proof. As discussed above, Gauss-Jordan reduce the given matrix and combine all the reduction matrices used there to make P. Then use the leading entries to do column reduction and finish by swapping columns to put the leading ones on the diagonal. Combine the reduction matrices used for those column operations into Q. QED

2.7 Example We illustrate the proof by finding the P and Q for this matrix.

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

First Gauss-Jordan row-reduce.

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

Then column-reduce, which involves right-multiplication.

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

Finish by swapping columns.

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

Finally, combine the left-multipliers together as P and the right-multipliers together as Q to get the PHQ equation.

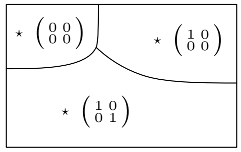
$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ 2 & 4 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0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

2.8 Corollary Two same-sized matrices are matrix equivalent if and only if they have the same rank. That is, the matrix equivalence classes are characterized by rank.

Proof. Two same-sized matrices with the same rank are equivalent to the same block partial-identity matrix. QED

2.9 Example The 2×2 matrices have only three possible ranks: zero, one, or two. Thus there are three matrix-equivalence classes.

All 2×2 matrices:



Three equivalence classes

Each class consists of all of the 2×2 matrices with the same rank. There is only one rank zero matrix, so that class has only one member, but the other two classes each have infinitely many members.

In this subsection we have seen how to change the representation of a map with respect to a first pair of bases to one with respect to a second pair. That led to a definition describing when matrices are equivalent in this way. Finally we noted that, with the proper choice of (possibly different) starting and ending bases, any map can be represented in block partial-identity form.

One of the nice things about this representation is that, in some sense, we can completely understand the map when it is expressed in this way: if the bases are $B = \langle \vec{\beta}_1, \dots, \vec{\beta}_n \rangle$ and $D = \langle \vec{\delta}_1, \dots, \vec{\delta}_m \rangle$ then the map sends

$$c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k + c_{k+1}\vec{\beta}_{k+1} + \cdots + c_n\vec{\beta}_n \longmapsto c_1\vec{\delta}_1 + \cdots + c_k\vec{\delta}_k + \vec{0} + \cdots + \vec{0}$$

where k is the map's rank. Thus, we can understand any linear map as a kind of projection.

$$\begin{pmatrix} c_1 \\ \vdots \\ c_k \\ c_{k+1} \\ \vdots \\ c_n \end{pmatrix}_B \mapsto \begin{pmatrix} c_1 \\ \vdots \\ c_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}_D$$

Of course, "understanding" a map expressed in this way requires that we understand the relationship between B and D. However, despite that difficulty, this is a good classification of linear maps.

Exercises

✓ **2.10** Decide if these matrices are matrix equivalent.

(a) $\begin{pmatrix} 1 & 3 & 0 \\ 2 & 3 & 0 \end{pmatrix}$, $\begin{pmatrix} 2 & 2 & 1 \\ 0 & 5 & -1 \end{pmatrix}$

(b) $\begin{pmatrix} 0 & 3 \\ 1 & 1 \end{pmatrix}$, $\begin{pmatrix} 4 & 0 \\ 0 & 5 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$, $\begin{pmatrix} 1 & 3 \\ 2 & -6 \end{pmatrix}$

✓ **2.11** Find the canonical representative of the matrix-equivalence class of each matrix.

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

2.12 Suppose that, with respect to

$$B = \mathcal{E}_2 \qquad D = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle$$

the transformation $t\colon \mathbb{R}^2 \to \mathbb{R}^2$ is represented by this matrix.

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

Use change of basis matrices to represent t with respect to each pair.

(a) $\hat{B} = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$, $\hat{D} = \langle \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rangle$

(b) $\hat{B} = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle$, $\hat{D} = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rangle$

✓ **2.13** What size are P and Q?

✓ **2.14** Use Theorem 2.6 to show that a square matrix is nonsingular if and only if it is equivalent to an identity matrix.

✓ **2.15** Show that, where A is a nonsingular square matrix, if P and Q are nonsingular square matrices such that $PAQ = I$ then $QP = A^{-1}$.

✓ **2.16** Why does Theorem 2.6 not show that every matrix is diagonalizable (see Example 2.2)?

2.17 Must matrix equivalent matrices have matrix equivalent transposes?

2.18 What happens in Theorem 2.6 if $k = 0$?

✓ **2.19** Show that matrix-equivalence is an equivalence relation.

✓ **2.20** Show that a zero matrix is alone in its matrix equivalence class. Are there other matrices like that?

2.21 What are the matrix equivalence classes of matrices of transformations on $\mathbb{R}^1$? $\mathbb{R}^3$?

2.22 How many matrix equivalence classes are there?

2.23 Are matrix equivalence classes closed under scalar multiplication? Addition?

2.24 Let $t\colon \mathbb{R}^n \to \mathbb{R}^n$ represented by T with respect to $\mathcal{E}_n, \mathcal{E}_n$.

(a) Find $\text{Rep}_{B,B}(t)$ in this specific case.

$$T = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \qquad B = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \rangle$$

(b) Describe $\text{Rep}_{B,B}(t)$ in the general case where $B = \langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$.

2.25 **(a)** Let V have bases B_1 and B_2 and suppose that W has the basis D. Where $h\colon V \to W$, find the formula that computes $\text{Rep}_{B_2,D}(h)$ from $\text{Rep}_{B_1,D}(h)$.

(b) Repeat the prior question with one basis for V and two bases for W.

2.26 **(a)** If two matrices are matrix-equivalent and invertible, must their inverses be matrix-equivalent?

(b) If two matrices have matrix-equivalent inverses, must the two be matrix-equivalent?

(c) If two matrices are square and matrix-equivalent, must their squares be matrix-equivalent?

(d) If two matrices are square and have matrix-equivalent squares, must they be matrix-equivalent?

✓ **2.27** Square matrices are *similar* if they represent the same transformation, but each with respect to the same ending as starting basis. That is, $\mathrm{Rep}_{B_1,B_1}(t)$ is similar to $\mathrm{Rep}_{B_2,B_2}(t)$.

(a) Give a definition of matrix similarity like that of Definition 2.3.

(b) Prove that similar matrices are matrix equivalent.

(c) Show that similarity is an equivalence relation.

(d) Show that if T is similar to $\hat{T}$ then T^2 is similar to $\hat{T}^2$, the cubes are similar, etc. *Contrast with the prior exercise.*

(e) Prove that there are matrix equivalent matrices that are not similar.