

## CHAPTER 3

# *Linear Maps*

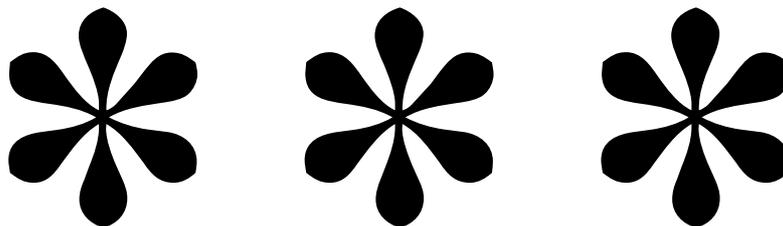
So far our attention has focused on vector spaces. No one gets excited about vector spaces. The interesting part of linear algebra is the subject to which we now turn—linear maps.

Let's review our standing assumptions:

Recall that  $\mathbf{F}$  denotes  $\mathbf{R}$  or  $\mathbf{C}$ .  
Recall also that  $V$  is a vector space over  $\mathbf{F}$ .

In this chapter we will frequently need another vector space in addition to  $V$ . We will call this additional vector space  $W$ :

Let's agree that for the rest of this chapter  
 $W$  will denote a vector space over  $\mathbf{F}$ .



## *Definitions and Examples*

Some mathematicians use the term **linear transformation**, which means the same as linear map.

A **linear map** from  $V$  to  $W$  is a function  $T: V \rightarrow W$  with the following properties:

### additivity

$$T(u + v) = Tu + Tv \text{ for all } u, v \in V;$$

### homogeneity

$$T(av) = a(Tv) \text{ for all } a \in \mathbf{F} \text{ and all } v \in V.$$

Note that for linear maps we often use the notation  $Tv$  as well as the more standard functional notation  $T(v)$ .

The set of all linear maps from  $V$  to  $W$  is denoted  $\mathcal{L}(V, W)$ . Let's look at some examples of linear maps. Make sure you verify that each of the functions defined below is indeed a linear map:

### zero

In addition to its other uses, we let the symbol  $0$  denote the function that takes each element of some vector space to the additive identity of another vector space. To be specific,  $0 \in \mathcal{L}(V, W)$  is defined by

$$0v = 0.$$

Note that the  $0$  on the left side of the equation above is a function from  $V$  to  $W$ , whereas the  $0$  on the right side is the additive identity in  $W$ . As usual, the context should allow you to distinguish between the many uses of the symbol  $0$ .

### identity

The **identity map**, denoted  $I$ , is the function on some vector space that takes each element to itself. To be specific,  $I \in \mathcal{L}(V, V)$  is defined by

$$Iv = v.$$

### differentiation

Define  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  by

$$Tp = p'.$$

The assertion that this function is a linear map is another way of stating a basic result about differentiation:  $(f + g)' = f' + g'$  and  $(af)' = af'$  whenever  $f, g$  are differentiable and  $a$  is a constant.

**integration**

Define  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathbf{R})$  by

$$Tp = \int_0^1 p(x) dx.$$

The assertion that this function is linear is another way of stating a basic result about integration: the integral of the sum of two functions equals the sum of the integrals, and the integral of a constant times a function equals the constant times the integral of the function.

**multiplication by  $x^2$** 

Define  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  by

$$(Tp)(x) = x^2 p(x)$$

for  $x \in \mathbf{R}$ .

**backward shift**

Recall that  $\mathbf{F}^\infty$  denotes the vector space of all sequences of elements of  $\mathbf{F}$ . Define  $T \in \mathcal{L}(\mathbf{F}^\infty, \mathbf{F}^\infty)$  by

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots).$$

**from  $\mathbf{F}^n$  to  $\mathbf{F}^m$** 

Define  $T \in \mathcal{L}(\mathbf{R}^3, \mathbf{R}^2)$  by

$$T(x, y, z) = (2x - y + 3z, 7x + 5y - 6z).$$

More generally, let  $m$  and  $n$  be positive integers, let  $a_{j,k} \in \mathbf{F}$  for  $j = 1, \dots, m$  and  $k = 1, \dots, n$ , and define  $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$  by

$$T(x_1, \dots, x_n) = (a_{1,1}x_1 + \dots + a_{1,n}x_n, \dots, a_{m,1}x_1 + \dots + a_{m,n}x_n).$$

Later we will see that every linear map from  $\mathbf{F}^n$  to  $\mathbf{F}^m$  is of this form.

Suppose  $(v_1, \dots, v_n)$  is a basis of  $V$  and  $T: V \rightarrow W$  is linear. If  $v \in V$ , then we can write  $v$  in the form

$$v = a_1 v_1 + \dots + a_n v_n.$$

The linearity of  $T$  implies that

*Though linear maps are pervasive throughout mathematics, they are not as ubiquitous as imagined by some confused students who seem to think that  $\cos$  is a linear map from  $\mathbf{R}$  to  $\mathbf{R}$  when they write "identities" such as  $\cos 2x = 2 \cos x$  and  $\cos(x + y) = \cos x + \cos y$ .*

$$T\mathbf{v} = a_1T\mathbf{v}_1 + \cdots + a_nT\mathbf{v}_n.$$

In particular, the values of  $T\mathbf{v}_1, \dots, T\mathbf{v}_n$  determine the values of  $T$  on arbitrary vectors in  $V$ .

Linear maps can be constructed that take on arbitrary values on a basis. Specifically, given a basis  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  of  $V$  and any choice of vectors  $\mathbf{w}_1, \dots, \mathbf{w}_n \in W$ , we can construct a linear map  $T: V \rightarrow W$  such that  $T\mathbf{v}_j = \mathbf{w}_j$  for  $j = 1, \dots, n$ . There is no choice of how to do this—we must define  $T$  by

$$T(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n,$$

where  $a_1, \dots, a_n$  are arbitrary elements of  $\mathbf{F}$ . Because  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  is a basis of  $V$ , the equation above does indeed define a function  $T$  from  $V$  to  $W$ . You should verify that the function  $T$  defined above is linear and that  $T\mathbf{v}_j = \mathbf{w}_j$  for  $j = 1, \dots, n$ .

Now we will make  $\mathcal{L}(V, W)$  into a vector space by defining addition and scalar multiplication on it. For  $S, T \in \mathcal{L}(V, W)$ , define a function  $S + T \in \mathcal{L}(V, W)$  in the usual manner of adding functions:

$$(S + T)\mathbf{v} = S\mathbf{v} + T\mathbf{v}$$

for  $\mathbf{v} \in V$ . You should verify that  $S + T$  is indeed a linear map from  $V$  to  $W$  whenever  $S, T \in \mathcal{L}(V, W)$ . For  $a \in \mathbf{F}$  and  $T \in \mathcal{L}(V, W)$ , define a function  $aT \in \mathcal{L}(V, W)$  in the usual manner of multiplying a function by a scalar:

$$(aT)\mathbf{v} = a(T\mathbf{v})$$

for  $\mathbf{v} \in V$ . You should verify that  $aT$  is indeed a linear map from  $V$  to  $W$  whenever  $a \in \mathbf{F}$  and  $T \in \mathcal{L}(V, W)$ . With the operations we have just defined,  $\mathcal{L}(V, W)$  becomes a vector space (as you should verify). Note that the additive identity of  $\mathcal{L}(V, W)$  is the zero linear map defined earlier in this section.

Usually it makes no sense to multiply together two elements of a vector space, but for some pairs of linear maps a useful product exists. We will need a third vector space, so suppose  $U$  is a vector space over  $\mathbf{F}$ . If  $T \in \mathcal{L}(U, V)$  and  $S \in \mathcal{L}(V, W)$ , then we define  $ST \in \mathcal{L}(U, W)$  by

$$(ST)(\mathbf{v}) = S(T\mathbf{v})$$

for  $\mathbf{v} \in U$ . In other words,  $ST$  is just the usual composition  $S \circ T$  of two functions, but when both functions are linear, most mathematicians

write  $ST$  instead of  $S \circ T$ . You should verify that  $ST$  is indeed a linear map from  $U$  to  $W$  whenever  $T \in \mathcal{L}(U, V)$  and  $S \in \mathcal{L}(V, W)$ . Note that  $ST$  is defined only when  $T$  maps into the domain of  $S$ . We often call  $ST$  the **product** of  $S$  and  $T$ . You should verify that it has most of the usual properties expected of a product:

#### associativity

$(T_1 T_2) T_3 = T_1 (T_2 T_3)$  whenever  $T_1, T_2$ , and  $T_3$  are linear maps such that the products make sense (meaning that  $T_3$  must map into the domain of  $T_2$ , and  $T_2$  must map into the domain of  $T_1$ ).

#### identity

$TI = T$  and  $IT = T$  whenever  $T \in \mathcal{L}(V, W)$  (note that in the first equation  $I$  is the identity map on  $V$ , and in the second equation  $I$  is the identity map on  $W$ ).

#### distributive properties

$(S_1 + S_2)T = S_1T + S_2T$  and  $S(T_1 + T_2) = ST_1 + ST_2$  whenever  $T, T_1, T_2 \in \mathcal{L}(U, V)$  and  $S, S_1, S_2 \in \mathcal{L}(V, W)$ .

Multiplication of linear maps is not commutative. In other words, it is not necessarily true that  $ST = TS$ , even if both sides of the equation make sense. For example, if  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  is the differentiation map defined earlier in this section and  $S \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  is the multiplication by  $x^2$  map defined earlier in this section, then

$$((ST)p)(x) = x^2 p'(x) \quad \text{but} \quad ((TS)p)(x) = x^2 p'(x) + 2xp(x).$$

In other words, multiplying by  $x^2$  and then differentiating is not the same as differentiating and then multiplying by  $x^2$ .

## *Null Spaces and Ranges*

For  $T \in \mathcal{L}(V, W)$ , the **null space** of  $T$ , denoted  $\text{null } T$ , is the subset of  $V$  consisting of those vectors that  $T$  maps to 0:

$$\text{null } T = \{v \in V : Tv = 0\}.$$

Let's look at a few examples from the previous section. In the differentiation example, we defined  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  by  $Tp = p'$ . The

*Some mathematicians use the term **kernel** instead of null space.*

only functions whose derivative equals the zero function are the constant functions, so in this case the null space of  $T$  equals the set of constant functions.

In the multiplication by  $x^2$  example, we defined  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  by  $(Tp)(x) = x^2p(x)$ . The only polynomial  $p$  such that  $x^2p(x) = 0$  for all  $x \in \mathbf{R}$  is the 0 polynomial. Thus in this case we have

$$\text{null } T = \{0\}.$$

In the backward shift example, we defined  $T \in \mathcal{L}(\mathbf{F}^\infty, \mathbf{F}^\infty)$  by

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots).$$

Clearly  $T(x_1, x_2, x_3, \dots)$  equals 0 if and only if  $x_2, x_3, \dots$  are all 0. Thus in this case we have

$$\text{null } T = \{(a, 0, 0, \dots) : a \in \mathbf{F}\}.$$

The next proposition shows that the null space of any linear map is a subspace of the domain. In particular, 0 is in the null space of every linear map.

**3.1 Proposition:** *If  $T \in \mathcal{L}(V, W)$ , then  $\text{null } T$  is a subspace of  $V$ .*

PROOF: Suppose  $T \in \mathcal{L}(V, W)$ . By additivity, we have

$$T(0) = T(0 + 0) = T(0) + T(0),$$

which implies that  $T(0) = 0$ . Thus  $0 \in \text{null } T$ .

If  $u, v \in \text{null } T$ , then

$$T(u + v) = Tu + Tv = 0 + 0 = 0,$$

and hence  $u + v \in \text{null } T$ . Thus  $\text{null } T$  is closed under addition.

If  $u \in \text{null } T$  and  $a \in \mathbf{F}$ , then

$$T(au) = aTu = a0 = 0,$$

and hence  $au \in \text{null } T$ . Thus  $\text{null } T$  is closed under scalar multiplication.

We have shown that  $\text{null } T$  contains 0 and is closed under addition and scalar multiplication. Thus  $\text{null } T$  is a subspace of  $V$ . ■

A linear map  $T: V \rightarrow W$  is called **injective** if whenever  $u, v \in V$  and  $Tu = Tv$ , we have  $u = v$ . The next proposition says that we can check whether a linear map is injective by checking whether  $0$  is the only vector that gets mapped to  $0$ . As a simple application of this proposition, we see that of the three linear maps whose null spaces we computed earlier in this section (differentiation, multiplication by  $x^2$ , and backward shift), only multiplication by  $x^2$  is injective.

Many mathematicians use the term **one-to-one**, which means the same as injective.

**3.2 Proposition:** *Let  $T \in \mathcal{L}(V, W)$ . Then  $T$  is injective if and only if  $\text{null } T = \{0\}$ .*

**PROOF:** First suppose that  $T$  is injective. We want to prove that  $\text{null } T = \{0\}$ . We already know that  $\{0\} \subset \text{null } T$  (by 3.1). To prove the inclusion in the other direction, suppose  $v \in \text{null } T$ . Then

$$T(v) = 0 = T(0).$$

Because  $T$  is injective, the equation above implies that  $v = 0$ . Thus  $\text{null } T = \{0\}$ , as desired.

To prove the implication in the other direction, now suppose that  $\text{null } T = \{0\}$ . We want to prove that  $T$  is injective. To do this, suppose  $u, v \in V$  and  $Tu = Tv$ . Then

$$0 = Tu - Tv = T(u - v).$$

Thus  $u - v$  is in  $\text{null } T$ , which equals  $\{0\}$ . Hence  $u - v = 0$ , which implies that  $u = v$ . Hence  $T$  is injective, as desired. ■

For  $T \in \mathcal{L}(V, W)$ , the **range** of  $T$ , denoted  $\text{range } T$ , is the subset of  $W$  consisting of those vectors that are of the form  $Tv$  for some  $v \in V$ :

$$\text{range } T = \{Tv : v \in V\}.$$

Some mathematicians use the word **image**, which means the same as range.

For example, if  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  is the differentiation map defined by  $Tp = p'$ , then  $\text{range } T = \mathcal{P}(\mathbf{R})$  because for every polynomial  $q \in \mathcal{P}(\mathbf{R})$  there exists a polynomial  $p \in \mathcal{P}(\mathbf{R})$  such that  $p' = q$ .

As another example, if  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  is the linear map of multiplication by  $x^2$  defined by  $(Tp)(x) = x^2p(x)$ , then the range of  $T$  is the set of polynomials of the form  $a_2x^2 + \cdots + a_mx^m$ , where  $a_2, \dots, a_m \in \mathbf{R}$ .

The next proposition shows that the range of any linear map is a subspace of the target space.

**3.3 Proposition:** *If  $T \in \mathcal{L}(V, W)$ , then  $\text{range } T$  is a subspace of  $W$ .*

**PROOF:** Suppose  $T \in \mathcal{L}(V, W)$ . Then  $T(0) = 0$  (by 3.1), which implies that  $0 \in \text{range } T$ .

If  $w_1, w_2 \in \text{range } T$ , then there exist  $v_1, v_2 \in V$  such that  $Tv_1 = w_1$  and  $Tv_2 = w_2$ . Thus

$$T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2,$$

and hence  $w_1 + w_2 \in \text{range } T$ . Thus  $\text{range } T$  is closed under addition.

If  $w \in \text{range } T$  and  $a \in \mathbf{F}$ , then there exists  $v \in V$  such that  $Tv = w$ . Thus

$$T(av) = aTv = aw,$$

and hence  $aw \in \text{range } T$ . Thus  $\text{range } T$  is closed under scalar multiplication.

We have shown that  $\text{range } T$  contains 0 and is closed under addition and scalar multiplication. Thus  $\text{range } T$  is a subspace of  $W$ . ■

*Many mathematicians use the term **onto**, which means the same as surjective.*

A linear map  $T: V \rightarrow W$  is called **surjective** if its range equals  $W$ . For example, the differentiation map  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  defined by  $Tp = p'$  is surjective because its range equals  $\mathcal{P}(\mathbf{R})$ . As another example, the linear map  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathcal{P}(\mathbf{R}))$  defined by  $(Tp)(x) = x^2p(x)$  is not surjective because its range does not equal  $\mathcal{P}(\mathbf{R})$ . As a final example, you should verify that the backward shift  $T \in \mathcal{L}(\mathbf{F}^\infty, \mathbf{F}^\infty)$  defined by

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$$

is surjective.

Whether a linear map is surjective can depend upon what we are thinking of as the target space. For example, fix a positive integer  $m$ . The differentiation map  $T \in \mathcal{L}(\mathcal{P}_m(\mathbf{R}), \mathcal{P}_m(\mathbf{R}))$  defined by  $Tp = p'$  is not surjective because the polynomial  $x^m$  is not in the range of  $T$ . However, the differentiation map  $T \in \mathcal{L}(\mathcal{P}_m(\mathbf{R}), \mathcal{P}_{m-1}(\mathbf{R}))$  defined by  $Tp = p'$  is surjective because its range equals  $\mathcal{P}_{m-1}(\mathbf{R})$ , which is now the target space.

The next theorem, which is the key result in this chapter, states that the dimension of the null space plus the dimension of the range of a linear map on a finite-dimensional vector space equals the dimension of the domain.

**3.4 Theorem:** *If  $V$  is finite dimensional and  $T \in \mathcal{L}(V, W)$ , then range  $T$  is a finite-dimensional subspace of  $W$  and*

$$\dim V = \dim \text{null } T + \dim \text{range } T.$$

PROOF: Suppose that  $V$  is a finite-dimensional vector space and  $T \in \mathcal{L}(V, W)$ . Let  $(u_1, \dots, u_m)$  be a basis of null  $T$ ; thus  $\dim \text{null } T = m$ . The linearly independent list  $(u_1, \dots, u_m)$  can be extended to a basis  $(u_1, \dots, u_m, w_1, \dots, w_n)$  of  $V$  (by 2.12). Thus  $\dim V = m + n$ , and to complete the proof, we need only show that range  $T$  is finite dimensional and  $\dim \text{range } T = n$ . We will do this by proving that  $(Tw_1, \dots, Tw_n)$  is a basis of range  $T$ .

Let  $v \in V$ . Because  $(u_1, \dots, u_m, w_1, \dots, w_n)$  spans  $V$ , we can write

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_n w_n,$$

where the  $a$ 's and  $b$ 's are in  $\mathbf{F}$ . Applying  $T$  to both sides of this equation, we get

$$Tv = b_1 Tw_1 + \cdots + b_n Tw_n,$$

where the terms of the form  $Tu_j$  disappeared because each  $u_j \in \text{null } T$ . The last equation implies that  $(Tw_1, \dots, Tw_n)$  spans range  $T$ . In particular, range  $T$  is finite dimensional.

To show that  $(Tw_1, \dots, Tw_n)$  is linearly independent, suppose that  $c_1, \dots, c_n \in \mathbf{F}$  and

$$c_1 Tw_1 + \cdots + c_n Tw_n = 0.$$

Then

$$T(c_1 w_1 + \cdots + c_n w_n) = 0,$$

and hence

$$c_1 w_1 + \cdots + c_n w_n \in \text{null } T.$$

Because  $(u_1, \dots, u_m)$  spans null  $T$ , we can write

$$c_1 w_1 + \cdots + c_n w_n = d_1 u_1 + \cdots + d_m u_m,$$

where the  $d$ 's are in  $\mathbf{F}$ . This equation implies that all the  $c$ 's (and  $d$ 's) are 0 (because  $(u_1, \dots, u_m, w_1, \dots, w_n)$  is linearly independent). Thus  $(Tw_1, \dots, Tw_n)$  is linearly independent and hence is a basis for range  $T$ , as desired. ■

Now we can show that no linear map from a finite-dimensional vector space to a “smaller” vector space can be injective, where “smaller” is measured by dimension.

**3.5 Corollary:** *If  $V$  and  $W$  are finite-dimensional vector spaces such that  $\dim V > \dim W$ , then no linear map from  $V$  to  $W$  is injective.*

PROOF: Suppose  $V$  and  $W$  are finite-dimensional vector spaces such that  $\dim V > \dim W$ . Let  $T \in \mathcal{L}(V, W)$ . Then

$$\begin{aligned}\dim \text{null } T &= \dim V - \dim \text{range } T \\ &\geq \dim V - \dim W \\ &> 0,\end{aligned}$$

where the equality above comes from 3.4. We have just shown that  $\dim \text{null } T > 0$ . This means that  $\text{null } T$  must contain vectors other than 0. Thus  $T$  is not injective (by 3.2). ■

The next corollary, which is in some sense dual to the previous corollary, shows that no linear map from a finite-dimensional vector space to a “bigger” vector space can be surjective, where “bigger” is measured by dimension.

**3.6 Corollary:** *If  $V$  and  $W$  are finite-dimensional vector spaces such that  $\dim V < \dim W$ , then no linear map from  $V$  to  $W$  is surjective.*

PROOF: Suppose  $V$  and  $W$  are finite-dimensional vector spaces such that  $\dim V < \dim W$ . Let  $T \in \mathcal{L}(V, W)$ . Then

$$\begin{aligned}\dim \text{range } T &= \dim V - \dim \text{null } T \\ &\leq \dim V \\ &< \dim W,\end{aligned}$$

where the equality above comes from 3.4. We have just shown that  $\dim \text{range } T < \dim W$ . This means that  $\text{range } T$  cannot equal  $W$ . Thus  $T$  is not surjective. ■

The last two corollaries have important consequences in the theory of linear equations. To see this, fix positive integers  $m$  and  $n$ , and let  $a_{j,k} \in \mathbf{F}$  for  $j = 1, \dots, m$  and  $k = 1, \dots, n$ . Define  $T: \mathbf{F}^n \rightarrow \mathbf{F}^m$  by

$$T(x_1, \dots, x_n) = \left( \sum_{k=1}^n a_{1,k}x_k, \dots, \sum_{k=1}^n a_{m,k}x_k \right).$$

Now consider the equation  $Tx = 0$  (where  $x \in \mathbf{F}^n$  and the 0 here is the additive identity in  $\mathbf{F}^m$ , namely, the list of length  $m$  consisting of all 0's). Letting  $x = (x_1, \dots, x_n)$ , we can rewrite the equation  $Tx = 0$  as a system of homogeneous equations:

$$\begin{aligned} \sum_{k=1}^n a_{1,k}x_k &= 0 \\ &\vdots \\ \sum_{k=1}^n a_{m,k}x_k &= 0. \end{aligned}$$

We think of the  $a$ 's as known; we are interested in solutions for the variables  $x_1, \dots, x_n$ . Thus we have  $m$  equations and  $n$  variables. Obviously  $x_1 = \dots = x_n = 0$  is a solution; the key question here is whether any other solutions exist. In other words, we want to know if null  $T$  is strictly bigger than  $\{0\}$ . This happens precisely when  $T$  is not injective (by 3.2). From 3.5 we see that  $T$  is not injective if  $n > m$ . Conclusion: a homogeneous system of linear equations in which there are more variables than equations must have nonzero solutions.

With  $T$  as in the previous paragraph, now consider the equation  $Tx = c$ , where  $c = (c_1, \dots, c_m) \in \mathbf{F}^m$ . We can rewrite the equation  $Tx = c$  as a system of inhomogeneous equations:

$$\begin{aligned} \sum_{k=1}^n a_{1,k}x_k &= c_1 \\ &\vdots \\ \sum_{k=1}^n a_{m,k}x_k &= c_m. \end{aligned}$$

As before, we think of the  $a$ 's as known. The key question here is whether for every choice of the constant terms  $c_1, \dots, c_m \in \mathbf{F}$ , there exists at least one solution for the variables  $x_1, \dots, x_n$ . In other words, we want to know whether range  $T$  equals  $\mathbf{F}^m$ . From 3.6 we see that  $T$  is not surjective if  $n < m$ . Conclusion: an inhomogeneous system of linear equations in which there are more equations than variables has no solution for some choice of the constant terms.

*Homogeneous, in this context, means that the constant term on the right side of each equation equals 0.*

*These results about homogeneous systems with more variables than equations and inhomogeneous systems with more equations than variables are often proved using Gaussian elimination. The abstract approach taken here leads to cleaner proofs.*

## The Matrix of a Linear Map

We have seen that if  $(v_1, \dots, v_n)$  is a basis of  $V$  and  $T: V \rightarrow W$  is linear, then the values of  $Tv_1, \dots, Tv_n$  determine the values of  $T$  on arbitrary vectors in  $V$ . In this section we will see how matrices are used as an efficient method of recording the values of the  $Tv_j$ 's in terms of a basis of  $W$ .

Let  $m$  and  $n$  denote positive integers. An  $m$ -by- $n$  **matrix** is a rectangular array with  $m$  rows and  $n$  columns that looks like this:

$$3.7 \quad \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}.$$

Note that the first index refers to the row number and the second index refers to the column number. Thus  $a_{3,2}$  refers to the entry in the third row, second column of the matrix above. We will usually consider matrices whose entries are elements of  $\mathbf{F}$ .

Let  $T \in \mathcal{L}(V, W)$ . Suppose that  $(v_1, \dots, v_n)$  is a basis of  $V$  and  $(w_1, \dots, w_m)$  is a basis of  $W$ . For each  $k = 1, \dots, n$ , we can write  $Tv_k$  uniquely as a linear combination of the  $w$ 's:

$$3.8 \quad Tv_k = a_{1,k}w_1 + \cdots + a_{m,k}w_m,$$

where  $a_{j,k} \in \mathbf{F}$  for  $j = 1, \dots, m$ . The scalars  $a_{j,k}$  completely determine the linear map  $T$  because a linear map is determined by its values on a basis. The  $m$ -by- $n$  matrix 3.7 formed by the  $a$ 's is called the **matrix** of  $T$  with respect to the bases  $(v_1, \dots, v_n)$  and  $(w_1, \dots, w_m)$ ; we denote it by

$$\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m)).$$

If the bases  $(v_1, \dots, v_n)$  and  $(w_1, \dots, w_m)$  are clear from the context (for example, if only one set of bases is in sight), we write just  $\mathcal{M}(T)$  instead of  $\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$ .

As an aid to remembering how  $\mathcal{M}(T)$  is constructed from  $T$ , you might write the basis vectors  $v_1, \dots, v_n$  for the domain across the top and the basis vectors  $w_1, \dots, w_m$  for the target space along the left, as follows:

$$\begin{array}{cccc}
 & \nu_1 & \dots & \nu_k & \dots & \nu_n \\
 \begin{array}{c} w_1 \\ \vdots \\ w_m \end{array} & \left[ \begin{array}{cccc} & & & a_{1,k} \\ & & & \vdots \\ & & & a_{m,k} \end{array} \right] & & & 
 \end{array}$$

Note that in the matrix above only the  $k^{\text{th}}$  column is displayed (and thus the second index of each displayed  $a$  is  $k$ ). The  $k^{\text{th}}$  column of  $\mathcal{M}(T)$  consists of the scalars needed to write  $T\nu_k$  as a linear combination of the  $w$ 's. Thus the picture above should remind you that  $T\nu_k$  is retrieved from the matrix  $\mathcal{M}(T)$  by multiplying each entry in the  $k^{\text{th}}$  column by the corresponding  $w$  from the left column, and then adding up the resulting vectors.

*With respect to any choice of bases, the matrix of the 0 linear map (the linear map that takes every vector to 0) consists of all 0's.*

If  $T$  is a linear map from  $\mathbf{F}^n$  to  $\mathbf{F}^m$ , then unless stated otherwise you should assume that the bases in question are the standard ones (where the  $k^{\text{th}}$  basis vector is 1 in the  $k^{\text{th}}$  slot and 0 in all the other slots). If you think of elements of  $\mathbf{F}^m$  as columns of  $m$  numbers, then you can think of the  $k^{\text{th}}$  column of  $\mathcal{M}(T)$  as  $T$  applied to the  $k^{\text{th}}$  basis vector. For example, if  $T \in \mathcal{L}(\mathbf{F}^2, \mathbf{F}^3)$  is defined by

$$T(x, y) = (x + 3y, 2x + 5y, 7x + 9y),$$

then  $T(1, 0) = (1, 2, 7)$  and  $T(0, 1) = (3, 5, 9)$ , so the matrix of  $T$  (with respect to the standard bases) is the 3-by-2 matrix

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{bmatrix}.$$

Suppose we have bases  $(\nu_1, \dots, \nu_n)$  of  $V$  and  $(w_1, \dots, w_m)$  of  $W$ . Thus for each linear map from  $V$  to  $W$ , we can talk about its matrix (with respect to these bases, of course). Is the matrix of the sum of two linear maps equal to the sum of the matrices of the two maps?

Right now this question does not make sense because, though we have defined the sum of two linear maps, we have not defined the sum of two matrices. Fortunately the obvious definition of the sum of two matrices has the right properties. Specifically, we define addition of matrices of the same size by adding corresponding entries in the matrices:

$$\begin{aligned} \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{bmatrix} \\ = \begin{bmatrix} a_{1,1} + b_{1,1} & \dots & a_{1,n} + b_{1,n} \\ \vdots & & \vdots \\ a_{m,1} + b_{m,1} & \dots & a_{m,n} + b_{m,n} \end{bmatrix}. \end{aligned}$$

You should verify that with this definition of matrix addition,

$$\mathbf{3.9} \quad \mathcal{M}(T + S) = \mathcal{M}(T) + \mathcal{M}(S)$$

whenever  $T, S \in \mathcal{L}(V, W)$ .

Still assuming that we have some bases in mind, is the matrix of a scalar times a linear map equal to the scalar times the matrix of the linear map? Again the question does not make sense because we have not defined scalar multiplication on matrices. Fortunately the obvious definition again has the right properties. Specifically, we define the product of a scalar and a matrix by multiplying each entry in the matrix by the scalar:

$$c \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} = \begin{bmatrix} ca_{1,1} & \dots & ca_{1,n} \\ \vdots & & \vdots \\ ca_{m,1} & \dots & ca_{m,n} \end{bmatrix}.$$

You should verify that with this definition of scalar multiplication on matrices,

$$\mathbf{3.10} \quad \mathcal{M}(cT) = c\mathcal{M}(T)$$

whenever  $c \in \mathbf{F}$  and  $T \in \mathcal{L}(V, W)$ .

Because addition and scalar multiplication have now been defined for matrices, you should not be surprised that a vector space is about to appear. We need only a bit of notation so that this new vector space has a name. The set of all  $m$ -by- $n$  matrices with entries in  $\mathbf{F}$  is denoted by  $\text{Mat}(m, n, \mathbf{F})$ . You should verify that with addition and scalar multiplication defined as above,  $\text{Mat}(m, n, \mathbf{F})$  is a vector space. Note that the additive identity in  $\text{Mat}(m, n, \mathbf{F})$  is the  $m$ -by- $n$  matrix all of whose entries equal 0.

Suppose  $(v_1, \dots, v_n)$  is a basis of  $V$  and  $(w_1, \dots, w_m)$  is a basis of  $W$ . Suppose also that we have another vector space  $U$  and that  $(u_1, \dots, u_p)$

is a basis of  $U$ . Consider linear maps  $S: U \rightarrow V$  and  $T: V \rightarrow W$ . The composition  $TS$  is a linear map from  $U$  to  $W$ . How can  $\mathcal{M}(TS)$  be computed from  $\mathcal{M}(T)$  and  $\mathcal{M}(S)$ ? The nicest solution to this question would be to have the following pretty relationship:

$$3.11 \quad \mathcal{M}(TS) = \mathcal{M}(T)\mathcal{M}(S).$$

So far, however, the right side of this equation does not make sense because we have not yet defined the product of two matrices. We will choose a definition of matrix multiplication that forces the equation above to hold. Let's see how to do this.

Let

$$\mathcal{M}(T) = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \quad \text{and} \quad \mathcal{M}(S) = \begin{bmatrix} b_{1,1} & \dots & b_{1,p} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,p} \end{bmatrix}.$$

For  $k \in \{1, \dots, p\}$ , we have

$$\begin{aligned} TSu_k &= T\left(\sum_{r=1}^n b_{r,k}v_r\right) \\ &= \sum_{r=1}^n b_{r,k}Tv_r \\ &= \sum_{r=1}^n b_{r,k} \sum_{j=1}^m a_{j,r}w_j \\ &= \sum_{j=1}^m \left(\sum_{r=1}^n a_{j,r}b_{r,k}\right)w_j. \end{aligned}$$

Thus  $\mathcal{M}(TS)$  is the  $m$ -by- $p$  matrix whose entry in row  $j$ , column  $k$  equals  $\sum_{r=1}^n a_{j,r}b_{r,k}$ .

Now it's clear how to define matrix multiplication so that 3.11 holds. Namely, if  $A$  is an  $m$ -by- $n$  matrix with entries  $a_{j,k}$  and  $B$  is an  $n$ -by- $p$  matrix with entries  $b_{j,k}$ , then  $AB$  is defined to be the  $m$ -by- $p$  matrix whose entry in row  $j$ , column  $k$ , equals

$$\sum_{r=1}^n a_{j,r}b_{r,k}.$$

In other words, the entry in row  $j$ , column  $k$ , of  $AB$  is computed by taking row  $j$  of  $A$  and column  $k$  of  $B$ , multiplying together corresponding entries, and then summing. Note that we define the product of two

*You probably learned this definition of matrix multiplication in an earlier course, although you may not have seen this motivation for it.*

*You should find an example to show that matrix multiplication is not commutative. In other words,  $AB$  is not necessarily equal to  $BA$ , even when both are defined.*

matrices only when the number of columns of the first matrix equals the number of rows of the second matrix.

As an example of matrix multiplication, here we multiply together a 3-by-2 matrix and a 2-by-4 matrix, obtaining a 3-by-4 matrix:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 6 & 5 & 4 & 3 \\ 2 & 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 10 & 7 & 4 & 1 \\ 26 & 19 & 12 & 5 \\ 42 & 31 & 20 & 9 \end{bmatrix}.$$

Suppose  $(\nu_1, \dots, \nu_n)$  is a basis of  $V$ . If  $\nu \in V$ , then there exist unique scalars  $b_1, \dots, b_n$  such that

$$3.12 \quad \nu = b_1\nu_1 + \dots + b_n\nu_n.$$

The *matrix* of  $\nu$ , denoted  $\mathcal{M}(\nu)$ , is the  $n$ -by-1 matrix defined by

$$3.13 \quad \mathcal{M}(\nu) = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Usually the basis is obvious from the context, but when the basis needs to be displayed explicitly use the notation  $\mathcal{M}(\nu, (\nu_1, \dots, \nu_n))$  instead of  $\mathcal{M}(\nu)$ .

For example, the matrix of a vector  $x \in \mathbf{F}^n$  with respect to the standard basis is obtained by writing the coordinates of  $x$  as the entries in an  $n$ -by-1 matrix. In other words, if  $x = (x_1, \dots, x_n) \in \mathbf{F}^n$ , then

$$\mathcal{M}(x) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

The next proposition shows how the notions of the matrix of a linear map, the matrix of a vector, and matrix multiplication fit together. In this proposition  $\mathcal{M}(T\nu)$  is the matrix of the vector  $T\nu$  with respect to the basis  $(w_1, \dots, w_m)$  and  $\mathcal{M}(\nu)$  is the matrix of the vector  $\nu$  with respect to the basis  $(\nu_1, \dots, \nu_n)$ , whereas  $\mathcal{M}(T)$  is the matrix of the linear map  $T$  with respect to the bases  $(\nu_1, \dots, \nu_n)$  and  $(w_1, \dots, w_m)$ .

**3.14 Proposition:** *Suppose  $T \in \mathcal{L}(V, W)$  and  $(\nu_1, \dots, \nu_n)$  is a basis of  $V$  and  $(w_1, \dots, w_m)$  is a basis of  $W$ . Then*

$$\mathcal{M}(T\nu) = \mathcal{M}(T)\mathcal{M}(\nu)$$

for every  $\nu \in V$ .

PROOF: Let

$$3.15 \quad \mathcal{M}(T) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}.$$

This means, we recall, that

$$3.16 \quad T\mathbf{v}_k = \sum_{j=1}^m a_{j,k} \mathbf{w}_j$$

for each  $k$ . Let  $\mathbf{v}$  be an arbitrary vector in  $V$ , which we can write in the form 3.12. Thus  $\mathcal{M}(\mathbf{v})$  is given by 3.13. Now

$$\begin{aligned} T\mathbf{v} &= b_1 T\mathbf{v}_1 + \cdots + b_n T\mathbf{v}_n \\ &= b_1 \sum_{j=1}^m a_{j,1} \mathbf{w}_j + \cdots + b_n \sum_{j=1}^m a_{j,n} \mathbf{w}_j \\ &= \sum_{j=1}^m (a_{j,1} b_1 + \cdots + a_{j,n} b_n) \mathbf{w}_j, \end{aligned}$$

where the first equality comes from 3.12 and the second equality comes from 3.16. The last equation shows that  $\mathcal{M}(T\mathbf{v})$ , the  $m$ -by-1 matrix of the vector  $T\mathbf{v}$  with respect to the basis  $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ , is given by the equation

$$\mathcal{M}(T\mathbf{v}) = \begin{bmatrix} a_{1,1} b_1 + \cdots + a_{1,n} b_n \\ \vdots \\ a_{m,1} b_1 + \cdots + a_{m,n} b_n \end{bmatrix}.$$

This formula, along with the formulas 3.15 and 3.13 and the definition of matrix multiplication, shows that  $\mathcal{M}(T\mathbf{v}) = \mathcal{M}(T)\mathcal{M}(\mathbf{v})$ . ■

## Invertibility

A linear map  $T \in \mathcal{L}(V, W)$  is called **invertible** if there exists a linear map  $S \in \mathcal{L}(W, V)$  such that  $ST$  equals the identity map on  $V$  and  $TS$  equals the identity map on  $W$ . A linear map  $S \in \mathcal{L}(W, V)$  satisfying  $ST = I$  and  $TS = I$  is called an **inverse** of  $T$  (note that the first  $I$  is the identity map on  $V$  and the second  $I$  is the identity map on  $W$ ).

If  $S$  and  $S'$  are inverses of  $T$ , then

$$S = SI = S(TS') = (ST)S' = IS' = S',$$

so  $S = S'$ . In other words, if  $T$  is invertible, then it has a unique inverse, which we denote by  $T^{-1}$ . Rephrasing all this once more, if  $T \in \mathcal{L}(V, W)$  is invertible, then  $T^{-1}$  is the unique element of  $\mathcal{L}(W, V)$  such that  $T^{-1}T = I$  and  $TT^{-1} = I$ . The following proposition characterizes the invertible linear maps.

**3.17 Proposition:** *A linear map is invertible if and only if it is injective and surjective.*

**PROOF:** Suppose  $T \in \mathcal{L}(V, W)$ . We need to show that  $T$  is invertible if and only if it is injective and surjective.

First suppose that  $T$  is invertible. To show that  $T$  is injective, suppose that  $u, v \in V$  and  $Tu = Tv$ . Then

$$u = T^{-1}(Tu) = T^{-1}(Tv) = v,$$

so  $u = v$ . Hence  $T$  is injective.

We are still assuming that  $T$  is invertible. Now we want to prove that  $T$  is surjective. To do this, let  $w \in W$ . Then  $w = T(T^{-1}w)$ , which shows that  $w$  is in the range of  $T$ . Thus  $\text{range } T = W$ , and hence  $T$  is surjective, completing this direction of the proof.

Now suppose that  $T$  is injective and surjective. We want to prove that  $T$  is invertible. For each  $w \in W$ , define  $Sw$  to be the unique element of  $V$  such that  $T(Sw) = w$  (the existence and uniqueness of such an element follow from the surjectivity and injectivity of  $T$ ). Clearly  $TS$  equals the identity map on  $W$ . To prove that  $ST$  equals the identity map on  $V$ , let  $v \in V$ . Then

$$T(STv) = (TS)(Tv) = I(Tv) = Tv.$$

This equation implies that  $STv = v$  (because  $T$  is injective), and thus  $ST$  equals the identity map on  $V$ . To complete the proof, we need to show that  $S$  is linear. To do this, let  $w_1, w_2 \in W$ . Then

$$T(Sw_1 + Sw_2) = T(Sw_1) + T(Sw_2) = w_1 + w_2.$$

Thus  $Sw_1 + Sw_2$  is the unique element of  $V$  that  $T$  maps to  $w_1 + w_2$ . By the definition of  $S$ , this implies that  $S(w_1 + w_2) = Sw_1 + Sw_2$ . Hence  $S$  satisfies the additive property required for linearity. The proof of homogeneity is similar. Specifically, if  $w \in W$  and  $a \in \mathbb{F}$ , then

$$T(aS\mathcal{w}) = aT(S\mathcal{w}) = a\mathcal{w}.$$

Thus  $aS\mathcal{w}$  is the unique element of  $V$  that  $T$  maps to  $a\mathcal{w}$ . By the definition of  $S$ , this implies that  $S(a\mathcal{w}) = aS\mathcal{w}$ . Hence  $S$  is linear, as desired. ■

Two vector spaces are called **isomorphic** if there is an invertible linear map from one vector space onto the other one. As abstract vector spaces, two isomorphic spaces have the same properties. From this viewpoint, you can think of an invertible linear map as a relabeling of the elements of a vector space.

*The Greek word **isos** means equal; the Greek word **morph** means shape. Thus **isomorphic** literally means equal shape.*

If two vector spaces are isomorphic and one of them is finite dimensional, then so is the other one. To see this, suppose that  $V$  and  $W$  are isomorphic and that  $T \in \mathcal{L}(V, W)$  is an invertible linear map. If  $V$  is finite dimensional, then so is  $W$  (by 3.4). The same reasoning, with  $T$  replaced with  $T^{-1} \in \mathcal{L}(W, V)$ , shows that if  $W$  is finite dimensional, then so is  $V$ . Actually much more is true, as the following theorem shows.

**3.18 Theorem:** *Two finite-dimensional vector spaces are isomorphic if and only if they have the same dimension.*

**PROOF:** First suppose  $V$  and  $W$  are isomorphic finite-dimensional vector spaces. Thus there exists an invertible linear map  $T$  from  $V$  onto  $W$ . Because  $T$  is invertible, we have  $\text{null } T = \{0\}$  and  $\text{range } T = W$ . Thus  $\dim \text{null } T = 0$  and  $\dim \text{range } T = \dim W$ . The formula

$$\dim V = \dim \text{null } T + \dim \text{range } T$$

(see 3.4) thus becomes the equation  $\dim V = \dim W$ , completing the proof in one direction.

To prove the other direction, suppose  $V$  and  $W$  are finite-dimensional vector spaces with the same dimension. Let  $(\mathcal{v}_1, \dots, \mathcal{v}_n)$  be a basis of  $V$  and  $(\mathcal{w}_1, \dots, \mathcal{w}_n)$  be a basis of  $W$ . Let  $T$  be the linear map from  $V$  to  $W$  defined by

$$T(a_1\mathcal{v}_1 + \dots + a_n\mathcal{v}_n) = a_1\mathcal{w}_1 + \dots + a_n\mathcal{w}_n.$$

Then  $T$  is surjective because  $(\mathcal{w}_1, \dots, \mathcal{w}_n)$  spans  $W$ , and  $T$  is injective because  $(\mathcal{w}_1, \dots, \mathcal{w}_n)$  is linearly independent. Because  $T$  is injective and

surjective, it is invertible (see 3.17), and hence  $V$  and  $W$  are isomorphic, as desired. ■

*Because every finite-dimensional vector space is isomorphic to some  $\mathbf{F}^n$ , why bother with abstract vector spaces? To answer this question, note that an investigation of  $\mathbf{F}^n$  would soon lead to vector spaces that do not equal  $\mathbf{F}^n$ . For example, we would encounter the null space and range of linear maps, the set of matrices  $\text{Mat}(n, n, \mathbf{F})$ , and the polynomials  $\mathcal{P}_n(\mathbf{F})$ . Though each of these vector spaces is isomorphic to some  $\mathbf{F}^m$ , thinking of them that way often adds complexity but no new insight.*

The last theorem implies that every finite-dimensional vector space is isomorphic to some  $\mathbf{F}^n$ . Specifically, if  $V$  is a finite-dimensional vector space and  $\dim V = n$ , then  $V$  and  $\mathbf{F}^n$  are isomorphic.

If  $(v_1, \dots, v_n)$  is a basis of  $V$  and  $(w_1, \dots, w_m)$  is a basis of  $W$ , then for each  $T \in \mathcal{L}(V, W)$ , we have a matrix  $\mathcal{M}(T) \in \text{Mat}(m, n, \mathbf{F})$ . In other words, once bases have been fixed for  $V$  and  $W$ ,  $\mathcal{M}$  becomes a function from  $\mathcal{L}(V, W)$  to  $\text{Mat}(m, n, \mathbf{F})$ . Notice that 3.9 and 3.10 show that  $\mathcal{M}$  is a linear map. This linear map is actually invertible, as we now show.

**3.19 Proposition:** *Suppose that  $(v_1, \dots, v_n)$  is a basis of  $V$  and  $(w_1, \dots, w_m)$  is a basis of  $W$ . Then  $\mathcal{M}$  is an invertible linear map between  $\mathcal{L}(V, W)$  and  $\text{Mat}(m, n, \mathbf{F})$ .*

**PROOF:** We have already noted that  $\mathcal{M}$  is linear, so we need only prove that  $\mathcal{M}$  is injective and surjective (by 3.17). Both are easy. Let's begin with injectivity. If  $T \in \mathcal{L}(V, W)$  and  $\mathcal{M}(T) = 0$ , then  $Tv_k = 0$  for  $k = 1, \dots, n$ . Because  $(v_1, \dots, v_n)$  is a basis of  $V$ , this implies that  $T = 0$ . Thus  $\mathcal{M}$  is injective (by 3.2).

To prove that  $\mathcal{M}$  is surjective, let

$$A = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix}$$

be a matrix in  $\text{Mat}(m, n, \mathbf{F})$ . Let  $T$  be the linear map from  $V$  to  $W$  such that

$$Tv_k = \sum_{j=1}^m a_{j,k} w_j$$

for  $k = 1, \dots, n$ . Obviously  $\mathcal{M}(T)$  equals  $A$ , and so the range of  $\mathcal{M}$  equals  $\text{Mat}(m, n, \mathbf{F})$ , as desired. ■

An obvious basis of  $\text{Mat}(m, n, \mathbf{F})$  consists of those  $m$ -by- $n$  matrices that have 0 in all entries except for a 1 in one entry. There are  $mn$  such matrices, so the dimension of  $\text{Mat}(m, n, \mathbf{F})$  equals  $mn$ .

Now we can determine the dimension of the vector space of linear maps from one finite-dimensional vector space to another.

**3.20 Proposition:** *If  $V$  and  $W$  are finite dimensional, then  $\mathcal{L}(V, W)$  is finite dimensional and*

$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W).$$

PROOF: This follows from the equation  $\dim \text{Mat}(m, n, \mathbf{F}) = mn$ , 3.18, and 3.19. ■

A linear map from a vector space to itself is called an **operator**. If we want to specify the vector space, we say that a linear map  $T: V \rightarrow V$  is an operator on  $V$ . Because we are so often interested in linear maps from a vector space into itself, we use the notation  $\mathcal{L}(V)$  to denote the set of all operators on  $V$ . In other words,  $\mathcal{L}(V) = \mathcal{L}(V, V)$ .

Recall from 3.17 that a linear map is invertible if it is injective and surjective. For a linear map of a vector space into itself, you might wonder whether injectivity alone, or surjectivity alone, is enough to imply invertibility. On infinite-dimensional vector spaces neither condition alone implies invertibility. We can see this from some examples we have already considered. The multiplication by  $x^2$  operator (from  $\mathcal{P}(\mathbf{R})$  to itself) is injective but not surjective. The backward shift (from  $\mathbf{F}^\infty$  to itself) is surjective but not injective. In view of these examples, the next theorem is remarkable—it states that for maps from a finite-dimensional vector space to itself, either injectivity or surjectivity alone implies the other condition.

*The deepest and most important parts of linear algebra, as well as most of the rest of this book, deal with operators.*

**3.21 Theorem:** *Suppose  $V$  is finite dimensional. If  $T \in \mathcal{L}(V)$ , then the following are equivalent:*

- (a)  $T$  is invertible;
- (b)  $T$  is injective;
- (c)  $T$  is surjective.

PROOF: Suppose  $T \in \mathcal{L}(V)$ . Clearly (a) implies (b).

Now suppose (b) holds, so that  $T$  is injective. Thus  $\text{null } T = \{0\}$  (by 3.2). From 3.4 we have

$$\begin{aligned} \dim \text{range } T &= \dim V - \dim \text{null } T \\ &= \dim V, \end{aligned}$$

which implies that  $\text{range } T$  equals  $V$  (see Exercise 11 in Chapter 2). Thus  $T$  is surjective. Hence (b) implies (c).

Now suppose (c) holds, so that  $T$  is surjective. Thus  $\text{range } T = V$ . From 3.4 we have

$$\begin{aligned}\dim \text{null } T &= \dim V - \dim \text{range } T \\ &= 0,\end{aligned}$$

which implies that  $\text{null } T$  equals  $\{0\}$ . Thus  $T$  is injective (by 3.2), and so  $T$  is invertible (we already knew that  $T$  was surjective). Hence (c) implies (a), completing the proof. ■

## Exercises

1. Show that every linear map from a one-dimensional vector space to itself is multiplication by some scalar. More precisely, prove that if  $\dim V = 1$  and  $T \in \mathcal{L}(V, V)$ , then there exists  $a \in \mathbf{F}$  such that  $Tv = av$  for all  $v \in V$ .

2. Give an example of a function  $f: \mathbf{R}^2 \rightarrow \mathbf{R}$  such that

$$f(av) = af(v)$$

for all  $a \in \mathbf{R}$  and all  $v \in \mathbf{R}^2$  but  $f$  is not linear.

3. Suppose that  $V$  is finite dimensional. Prove that any linear map on a subspace of  $V$  can be extended to a linear map on  $V$ . In other words, show that if  $U$  is a subspace of  $V$  and  $S \in \mathcal{L}(U, W)$ , then there exists  $T \in \mathcal{L}(V, W)$  such that  $Tu = Su$  for all  $u \in U$ .
4. Suppose that  $T$  is a linear map from  $V$  to  $\mathbf{F}$ . Prove that if  $u \in V$  is not in  $\text{null } T$ , then

$$V = \text{null } T \oplus \{au : a \in \mathbf{F}\}.$$

5. Suppose that  $T \in \mathcal{L}(V, W)$  is injective and  $(v_1, \dots, v_n)$  is linearly independent in  $V$ . Prove that  $(Tv_1, \dots, Tv_n)$  is linearly independent in  $W$ .
6. Prove that if  $S_1, \dots, S_n$  are injective linear maps such that  $S_1 \dots S_n$  makes sense, then  $S_1 \dots S_n$  is injective.
7. Prove that if  $(v_1, \dots, v_n)$  spans  $V$  and  $T \in \mathcal{L}(V, W)$  is surjective, then  $(Tv_1, \dots, Tv_n)$  spans  $W$ .
8. Suppose that  $V$  is finite dimensional and that  $T \in \mathcal{L}(V, W)$ . Prove that there exists a subspace  $U$  of  $V$  such that  $U \cap \text{null } T = \{0\}$  and  $\text{range } T = \{Tu : u \in U\}$ .
9. Prove that if  $T$  is a linear map from  $\mathbf{F}^4$  to  $\mathbf{F}^2$  such that

$$\text{null } T = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_1 = 5x_2 \text{ and } x_3 = 7x_4\},$$

then  $T$  is surjective.

*Exercise 2 shows that homogeneity alone is not enough to imply that a function is a linear map. Additivity alone is also not enough to imply that a function is a linear map, although the proof of this involves advanced tools that are beyond the scope of this book.*

10. Prove that there does not exist a linear map from  $\mathbf{F}^5$  to  $\mathbf{F}^2$  whose null space equals

$$\{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2 \text{ and } x_3 = x_4 = x_5\}.$$

11. Prove that if there exists a linear map on  $V$  whose null space and range are both finite dimensional, then  $V$  is finite dimensional.
12. Suppose that  $V$  and  $W$  are both finite dimensional. Prove that there exists a surjective linear map from  $V$  onto  $W$  if and only if  $\dim W \leq \dim V$ .
13. Suppose that  $V$  and  $W$  are finite dimensional and that  $U$  is a subspace of  $V$ . Prove that there exists  $T \in \mathcal{L}(V, W)$  such that  $\text{null } T = U$  if and only if  $\dim U \geq \dim V - \dim W$ .
14. Suppose that  $W$  is finite dimensional and  $T \in \mathcal{L}(V, W)$ . Prove that  $T$  is injective if and only if there exists  $S \in \mathcal{L}(W, V)$  such that  $ST$  is the identity map on  $V$ .
15. Suppose that  $V$  is finite dimensional and  $T \in \mathcal{L}(V, W)$ . Prove that  $T$  is surjective if and only if there exists  $S \in \mathcal{L}(W, V)$  such that  $TS$  is the identity map on  $W$ .
16. Suppose that  $U$  and  $V$  are finite-dimensional vector spaces and that  $S \in \mathcal{L}(V, W)$ ,  $T \in \mathcal{L}(U, V)$ . Prove that

$$\dim \text{null } ST \leq \dim \text{null } S + \dim \text{null } T.$$

17. Prove that the distributive property holds for matrix addition and matrix multiplication. In other words, suppose  $A$ ,  $B$ , and  $C$  are matrices whose sizes are such that  $A(B + C)$  makes sense. Prove that  $AB + AC$  makes sense and that  $A(B + C) = AB + AC$ .
18. Prove that matrix multiplication is associative. In other words, suppose  $A$ ,  $B$ , and  $C$  are matrices whose sizes are such that  $(AB)C$  makes sense. Prove that  $A(BC)$  makes sense and that  $(AB)C = A(BC)$ .

19. Suppose  $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$  and that

$$\mathcal{M}(T) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix},$$

where we are using the standard bases. Prove that

$$T(x_1, \dots, x_n) = (a_{1,1}x_1 + \cdots + a_{1,n}x_n, \dots, a_{m,1}x_1 + \cdots + a_{m,n}x_n)$$

for every  $(x_1, \dots, x_n) \in \mathbf{F}^n$ .

20. Suppose  $(v_1, \dots, v_n)$  is a basis of  $V$ . Prove that the function  $T: V \rightarrow \text{Mat}(n, 1, \mathbf{F})$  defined by

$$Tv = \mathcal{M}(v)$$

is an invertible linear map of  $V$  onto  $\text{Mat}(n, 1, \mathbf{F})$ ; here  $\mathcal{M}(v)$  is the matrix of  $v \in V$  with respect to the basis  $(v_1, \dots, v_n)$ .

21. Prove that every linear map from  $\text{Mat}(n, 1, \mathbf{F})$  to  $\text{Mat}(m, 1, \mathbf{F})$  is given by a matrix multiplication. In other words, prove that if  $T \in \mathcal{L}(\text{Mat}(n, 1, \mathbf{F}), \text{Mat}(m, 1, \mathbf{F}))$ , then there exists an  $m$ -by- $n$  matrix  $A$  such that  $TB = AB$  for every  $B \in \text{Mat}(n, 1, \mathbf{F})$ .
22. Suppose that  $V$  is finite dimensional and  $S, T \in \mathcal{L}(V)$ . Prove that  $ST$  is invertible if and only if both  $S$  and  $T$  are invertible.
23. Suppose that  $V$  is finite dimensional and  $S, T \in \mathcal{L}(V)$ . Prove that  $ST = I$  if and only if  $TS = I$ .
24. Suppose that  $V$  is finite dimensional and  $T \in \mathcal{L}(V)$ . Prove that  $T$  is a scalar multiple of the identity if and only if  $ST = TS$  for every  $S \in \mathcal{L}(V)$ .
25. Prove that if  $V$  is finite dimensional with  $\dim V > 1$ , then the set of noninvertible operators on  $V$  is not a subspace of  $\mathcal{L}(V)$ .

*This exercise shows that  $T$  has the form promised on page 39.*



26. Suppose  $n$  is a positive integer and  $a_{i,j} \in \mathbf{F}$  for  $i, j = 1, \dots, n$ . Prove that the following are equivalent:

- (a) The trivial solution  $x_1 = \dots = x_n = 0$  is the only solution to the homogeneous system of equations

$$\begin{aligned} \sum_{k=1}^n a_{1,k}x_k &= 0 \\ &\vdots \\ \sum_{k=1}^n a_{n,k}x_k &= 0. \end{aligned}$$

- (b) For every  $c_1, \dots, c_n \in \mathbf{F}$ , there exists a solution to the system of equations

$$\begin{aligned} \sum_{k=1}^n a_{1,k}x_k &= c_1 \\ &\vdots \\ \sum_{k=1}^n a_{n,k}x_k &= c_n. \end{aligned}$$

Note that here we have the same number of equations as variables.