3D *Invertibility and Isomorphisms*

Invertible Linear Maps

We begin this section by defining the notions of invertible and inverse in the context of linear maps.

3.59 definition: *invertible, inverse*

- 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 operator on V and TS equals the identity operator 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 operator on V and the second I is the identity operator on W).

The definition above mentions "an inverse". However, the next result shows that we can change this terminology to "the inverse".

3.60 *inverse is unique*

An invertible linear map has a unique inverse.

Proof Suppose $T \in \mathcal{L}(V, W)$ is invertible and S_1 and S_2 are inverses of T. Then

$$S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2.$$

Thus $S_1 = S_2$. ∎

Now that we know that the inverse is unique, we can give it a notation.

3.61 notation: T^{-1}

If T is invertible, then its inverse is denoted by T^{-1}. In other words, 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 $TT^{-1} = I$.

3.62 example: *inverse of a linear map from* $\mathbf{R}^3$ *to* $\mathbf{R}^3$

Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ is defined by $T(x, y, z) = (-y, x, 4z)$. Thus T is a counterclockwise rotation by 90° in the xy-plane and a stretch by a factor of 4 in the direction of the z-axis.

Hence the inverse map $T^{-1} \in \mathcal{L}(\mathbf{R}^3)$ is the clockwise rotation by 90° in the xy-plane and a stretch by a factor of $\frac{1}{4}$ in the direction of the z-axis:

$$T^{-1}(x, y, z) = \left(y, -x, \tfrac{1}{4}z\right).$$

The next result shows that a linear map is invertible if and only if it is one-to-one and onto.

3.63 *invertibility $\Longleftrightarrow$ injectivity and surjectivity*

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 T is invertible. To show that T is injective, suppose $u, v \in V$ 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 $w = T\big(T^{-1}w\big)$, which shows that w is in the range of T. Thus $\operatorname{range} T = W$. Hence T is surjective, completing this direction of the proof.

Now suppose T is injective and surjective. We want to prove that T is invertible. For each $w \in W$, define $S(w)$ to be the unique element of V such that $T\big(S(w)\big) = w$ (the existence and uniqueness of such an element follow from the surjectivity and injectivity of T). The definition of S implies that $T \circ S$ equals the identity operator on W.

To prove that $S \circ T$ equals the identity operator on V, let $v \in V$. Then

$$T\big((S \circ T)v\big) = (T \circ S)(Tv) = I(Tv) = Tv.$$

This equation implies that $(S \circ T)v = v$ (because T is injective). Thus $S \circ T$ equals the identity operator on V.

To complete the proof, we need to show that S is linear. To do this, suppose $w_1, w_2 \in W$. Then

$$T\big(S(w_1) + S(w_2)\big) = T\big(S(w_1)\big) + T\big(S(w_2)\big) = w_1 + w_2.$$

Thus $S(w_1) + S(w_2)$ is the unique element of V that T maps to $w_1 + w_2$. By the definition of S, this implies that $S(w_1 + w_2) = S(w_1) + S(w_2)$. Hence S satisfies the additive property required for linearity.

The proof of homogeneity is similar. Specifically, if $w \in W$ and $\lambda \in \mathbf{F}$, then

$$T\big(\lambda S(w)\big) = \lambda T\big(S(w)\big) = \lambda w.$$

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

For a linear map from a vector space to 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, as illustrated by the next example, which uses two familiar linear maps from Example 3.3.

3.64 example: *neither injectivity nor surjectivity implies invertibility*

- The multiplication by x^2 linear map from $\mathcal{P}(\mathbf{R})$ to $\mathcal{P}(\mathbf{R})$ (see 3.3) is injective but it is not invertible because it is not surjective (the polynomial 1 is not in the range).
- The backward shift linear map from $\mathbf{F}^\infty$ to $\mathbf{F}^\infty$ (see 3.3) is surjective but it is not invertible because it is not injective [the vector $(1, 0, 0, 0, \dots)$ is in the null space].

In view of the example above, the next result is remarkable—it states that for a linear map from a finite-dimensional vector space to a vector space of the same dimension, either injectivity or surjectivity alone implies the other condition. Note that the hypothesis below that $\dim V = \dim W$ is automatically satisfied in the important special case where V is finite-dimensional and $W = V$.

3.65 *injectivity is equivalent to surjectivity (if* $\dim V = \dim W < \infty$*)*

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

$$T \text{ is invertible} \iff T \text{ is injective} \iff T \text{ is surjective.}$$

Proof The fundamental theorem of linear maps (3.21) states that

3.66
$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

If T is injective (which by 3.15 is equivalent to the condition $\dim \operatorname{null} T = 0$), then the equation above implies that

$$\dim \operatorname{range} T = \dim V - \dim \operatorname{null} T = \dim V = \dim W,$$

which implies that T is surjective (by 2.39).

Conversely, if T is surjective, then 3.66 implies that

$$\dim \operatorname{null} T = \dim V - \dim \operatorname{range} T = \dim V - \dim W = 0,$$

which implies that T is injective.

Thus we have shown that T is injective if and only if T is surjective. Thus if T is either injective or surjective, then T is both injective and surjective, which implies that T is invertible. Hence T is invertible if and only if T is injective if and only if T is surjective. ∎

The next example illustrates the power of the previous result. Although it is possible to prove the result in the example below without using linear algebra, the proof using linear algebra is cleaner and easier.

3.67 example: *there exists a polynomial p such that $\big((x^2+5x+7)p\big)'' = q$*

The linear map

$$p \mapsto \big((x^2+5x+7)p\big)''$$

from $\mathcal{P}(\mathbf{R})$ to itself is injective, as you can show. Thus we are tempted to use 3.65 to show that this map is surjective. However, Example 3.64 shows that the magic of 3.65 does not apply to the infinite-dimensional vector space $\mathcal{P}(\mathbf{R})$. We will get around this problem by restricting attention to the finite-dimensional vector space $\mathcal{P}_m(\mathbf{R})$.

Suppose $q \in \mathcal{P}(\mathbf{R})$. There exists a nonnegative integer m such that $q \in \mathcal{P}_m(\mathbf{R})$. Define $T\colon \mathcal{P}_m(\mathbf{R}) \to \mathcal{P}_m(\mathbf{R})$ by

$$Tp = \big((x^2+5x+7)p\big)''.$$

Multiplying a nonzero polynomial by (x^2+5x+7) increases the degree by 2, and then differentiating twice reduces the degree by 2. Thus T is indeed a linear map from $\mathcal{P}_m(\mathbf{R})$ to itself.

Every polynomial whose second derivative equals 0 is of the form $ax+b$, where $a, b \in \mathbf{R}$. Thus $\operatorname{null} T = \{0\}$. Hence T is injective.

Thus T is surjective (by 3.65), which means that there exists a polynomial $p \in \mathcal{P}_m(\mathbf{R})$ such that $\big((x^2+5x+7)p\big)'' = q$, as claimed in the title of this example.

Exercise 35 in Section 6A gives a similar but more spectacular example of using 3.65.

The hypothesis in the result below that $\dim V = \dim W$ holds in the important special case in which V is finite-dimensional and $W = V$. Thus in that case, the equation $ST = I$ implies that $ST = TS$, even though we do not have multiplicative commutativity of arbitrary linear maps from V to V.

3.68 $ST = I \iff TS = I$ *(on vector spaces of the same dimension)*

Suppose V and W are finite-dimensional vector spaces of the same dimension, $S \in \mathcal{L}(W,V)$, and $T \in \mathcal{L}(V,W)$. Then $ST = I$ if and only if $TS = I$.

Proof First suppose $ST = I$. If $v \in V$ and $Tv = 0$, then

$$v = Iv = (ST)v = S(Tv) = S(0) = 0.$$

Thus T is injective (by 3.15). Because V and W have the same dimension, this implies that T is invertible (by 3.65).

Now multiply both sides of the equation $ST = I$ by T^{-1} on the right, getting

$$S = T^{-1}.$$

Thus $TS = TT^{-1} = I$, as desired.

To prove the implication in the other direction, simply reverse the roles of S and T (and V and W) in the direction we have already proved, showing that if $TS = I$, then $ST = I$. ∎

Isomorphic Vector Spaces

The next definition captures the idea of two vector spaces that are essentially the same, except for the names of their elements.

3.69 definition: *isomorphism, isomorphic*

- An *isomorphism* is an invertible linear map.
- Two vector spaces are called *isomorphic* if there is an isomorphism from one vector space onto the other one.

Think of an isomorphism $T\colon V \to W$ as relabeling $v \in V$ as $Tv \in W$. This viewpoint explains why two isomorphic vector spaces have the same vector space properties. The terms "isomorphism" and "invertible linear map" mean the same thing. Use "isomorphism" when you want to emphasize that the two spaces are essentially the same.

It can be difficult to determine whether two mathematical structures (such as groups or topological spaces) are essentially the same, differing only in the names of the elements of underlying sets. However, the next result shows that we need to look at only a single number (the dimension) to determine whether two vector spaces are isomorphic.

3.70 *dimension shows whether vector spaces are isomorphic*

Two finite-dimensional vector spaces over $\mathbf{F}$ 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 isomorphism T from V onto W. Because T is invertible, we have $\operatorname{null} T = \{0\}$ and $\operatorname{range} T = W$. Thus

$$\dim \operatorname{null} T = 0 \quad \text{and} \quad \dim \operatorname{range} T = \dim W.$$

The formula

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

(the fundamental theorem of linear maps, which is 3.21) 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 of the same dimension. Let $v_1, \dots, v_n$ be a basis of V and $w_1, \dots, w_n$ be a basis of W. Let $T \in \mathcal{L}(V, W)$ be defined by

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n.$$

Then T is a well-defined linear map because $v_1, \dots, v_n$ is a basis of V. Also, T is surjective because $w_1, \dots, w_n$ spans W. Furthermore, $\operatorname{null} T = \{0\}$ because $w_1, \dots, w_n$ is linearly independent. Thus T is injective. Because T is injective and surjective, it is an isomorphism (see 3.63). Hence V and W are isomorphic. ■

The previous result implies that each finite-dimensional vector space V is isomorphic to $\mathbf{F}^n$, where $n = \dim V$. For example, if m is a nonnegative integer, then $\mathcal{P}_m(\mathbf{F})$ is isomorphic to $\mathbf{F}^{m+1}$.

Recall that the notation $\mathbf{F}^{m,n}$ denotes the vector space of m-by-n matrices with entries in $\mathbf{F}$. 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 $\mathcal{M}(T) \in \mathbf{F}^{m,n}$. Thus once bases have been fixed for V and W, $\mathcal{M}$ becomes a function from $\mathcal{L}(V, W)$ to $\mathbf{F}^{m,n}$. Notice that 3.35 and 3.38 show that $\mathcal{M}$ is a linear map. This linear map is actually an isomorphism, as we now show.

Every finite-dimensional vector space is isomorphic to some $\mathbf{F}^n$. Thus why not just study $\mathbf{F}^n$ instead of more general vector spaces? To answer this question, note that an investigation of $\mathbf{F}^n$ would soon lead to other vector spaces. For example, we would encounter the null space and range of linear maps. Although each of these vector spaces is isomorphic to some $\mathbf{F}^m$, thinking of them that way often adds complexity but no new insight.

3.71 *$\mathcal{L}(V, W)$ and $\mathbf{F}^{m,n}$ are isomorphic*

Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. Then $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V, W)$ and $\mathbf{F}^{m,n}$.

Proof We already noted that $\mathcal{M}$ is linear. We need to prove that $\mathcal{M}$ is injective and surjective.

We begin with injectivity. If $T \in \mathcal{L}(V, W)$ and $\mathcal{M}(T) = 0$, then $Tv_k = 0$ for each $k = 1, \dots, n$. Because $v_1, \dots, v_n$ is a basis of V, this implies $T = 0$. Thus $\mathcal{M}$ is injective (by 3.15).

To prove that $\mathcal{M}$ is surjective, suppose $A \in \mathbf{F}^{m,n}$. By the linear map lemma (3.4), there exists $T \in \mathcal{L}(V, W)$ such that

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

for each $k = 1, \dots, n$. Because $\mathcal{M}(T)$ equals A, the range of $\mathcal{M}$ equals $\mathbf{F}^{m,n}$, as desired. ■

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

3.72 $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$

Suppose 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 The desired result follows from 3.71, 3.70, and 3.40. ■

Linear Maps Thought of as Matrix Multiplication

Previously we defined the matrix of a linear map. Now we define the matrix of a vector.

3.73 definition: *matrix of a vector, $\mathcal{M}(v)$*

Suppose $v \in V$ and $v_1, \dots, v_n$ is a basis of V. The *matrix of* v with respect to this basis is the n-by-1 matrix

$$\mathcal{M}(v) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

where $b_1, \dots, b_n$ are the scalars such that

$$v = b_1 v_1 + \cdots + b_n v_n.$$

The matrix $\mathcal{M}(v)$ of a vector $v \in V$ depends on the basis $v_1, \dots, v_n$ of V, as well as on v. However, the basis should be clear from the context and thus it is not included in the notation.

3.74 example: *matrix of a vector*

- The matrix of the polynomial $2 - 7x + 5x^3 + x^4$ with respect to the standard basis of $\mathcal{P}_4(\mathbf{R})$ is
$$\begin{pmatrix} 2 \\ -7 \\ 0 \\ 5 \\ 1 \end{pmatrix}.$$
- 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
$$\mathcal{M}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Occasionally we want to think of elements of V as relabeled to be n-by-1 matrices. Once a basis $v_1, \dots, v_n$ is chosen, the function $\mathcal{M}$ that takes $v \in V$ to $\mathcal{M}(v)$ is an isomorphism of V onto $\mathbf{F}^{n,1}$ that implements this relabeling.

Recall that if A is an m-by-n matrix, then $A_{\cdot,k}$ denotes the k^{th} column of A, thought of as an m-by-1 matrix. In the next result, $\mathcal{M}(Tv_k)$ is computed with respect to the basis $w_1, \dots, w_m$ of W.

3.75 $\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$.

Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. Let $1 \le k \le n$. Then the k^{th} column of $\mathcal{M}(T)$, which is denoted by $\mathcal{M}(T)_{\cdot,k}$, equals $\mathcal{M}(Tv_k)$.

Proof The desired result follows immediately from the definitions of $\mathcal{M}(T)$ and $\mathcal{M}(Tv_k)$. ∎

The next result shows how the notions of the matrix of a linear map, the matrix of a vector, and matrix multiplication fit together.

3.76 *linear maps act like matrix multiplication*

Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. Then

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

Proof Suppose $v = b_1 v_1 + \cdots + b_n v_n$, where $b_1, \dots, b_n \in \mathbf{F}$. Thus

3.77
$$Tv = b_1 T v_1 + \cdots + b_n T v_n.$$

Hence

$$\begin{aligned}
\mathcal{M}(Tv) &= b_1 \mathcal{M}(Tv_1) + \cdots + b_n \mathcal{M}(Tv_n) \\
&= b_1 \mathcal{M}(T)_{\cdot,1} + \cdots + b_n \mathcal{M}(T)_{\cdot,n} \\
&= \mathcal{M}(T)\mathcal{M}(v),
\end{aligned}$$

where the first equality follows from 3.77 and the linearity of $\mathcal{M}$, the second equality comes from 3.75, and the last equality comes from 3.50. ∎

Each m-by-n matrix A induces a linear map from $\mathbf{F}^{n,1}$ to $\mathbf{F}^{m,1}$, namely the matrix multiplication function that takes $x \in \mathbf{F}^{n,1}$ to $Ax \in \mathbf{F}^{m,1}$. The result above can be used to think of every linear map (from a finite-dimensional vector space to another finite-dimensional vector space) as a matrix multiplication map after suitable relabeling via the isomorphisms given by $\mathcal{M}$. Specifically, if $T \in \mathcal{L}(V, W)$ and we identify $v \in V$ with $\mathcal{M}(v) \in \mathbf{F}^{n,1}$, then the result above says that we can identify Tv with $\mathcal{M}(T)\mathcal{M}(v)$.

Because the result above allows us to think (via isomorphisms) of each linear map as multiplication on $\mathbf{F}^{n,1}$ by some matrix A, keep in mind that the specific matrix A depends not only on the linear map but also on the choice of bases. One of the themes of many of the most important results in later chapters will be the choice of a basis that makes the matrix A as simple as possible.

In this book, we concentrate on linear maps rather than on matrices. However, sometimes thinking of linear maps as matrices (or thinking of matrices as linear maps) gives important insights that we will find useful.

Notice that no bases are in sight in the statement of the next result. Although $\mathcal{M}(T)$ in the next result depends on a choice of bases of V and W, the next result shows that the column rank of $\mathcal{M}(T)$ is the same for all such choices (because range T does not depend on a choice of basis).

3.78 *dimension of* range T *equals column rank of* $\mathcal{M}(T)$

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then dim range T equals the column rank of $\mathcal{M}(T)$.

Proof Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. The linear map that takes $w \in W$ to $\mathcal{M}(w)$ is an isomorphism from W onto the space $\mathbf{F}^{m,1}$ of m-by-1 column vectors. The restriction of this isomorphism to range T [which equals $\operatorname{span}(Tv_1, \dots, Tv_n)$ by Exercise 10 in Section 3B] is an isomorphism from range T onto $\operatorname{span}\big(\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n)\big)$. For each $k \in \{1, \dots, n\}$, the m-by-1 matrix $\mathcal{M}(Tv_k)$ equals column k of $\mathcal{M}(T)$. Thus

$$\dim \operatorname{range} T = \text{the column rank of } \mathcal{M}(T),$$

as desired. ∎

Change of Basis

In Section 3C we defined the matrix

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

of a linear map T from V to a possibly different vector space W, where $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W. For linear maps from a vector space to itself, we usually use the same basis for both the domain vector space and the target vector space. When using a single basis in both capacities, we often write the basis only once. In other words, if $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V, then the notation $\mathcal{M}\big(T, (v_1, \dots, v_n)\big)$ is defined by the equation

$$\mathcal{M}\big(T, (v_1, \dots, v_n)\big) = \mathcal{M}\big(T, (v_1, \dots, v_n), (v_1, \dots, v_n)\big).$$

If the basis $v_1, \dots, v_n$ is clear from the context, then we can write just $\mathcal{M}(T)$.

3.79 definition: *identity matrix, I*

Suppose n is a positive integer. The n-by-n matrix

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

with 1's on the diagonal (the entries where the row number equals the column number) and 0's elsewhere is called the *identity matrix* and is denoted by I.

In the definition above, the 0 in the lower left corner of the matrix indicates that all entries below the diagonal are 0, and the 0 in the upper right corner indicates that all entries above the diagonal are 0.

With respect to each basis of V, the matrix of the identity operator $I \in \mathcal{L}(V)$ is the identity matrix I. Note that the symbol I is used to denote both the identity operator and the identity matrix. The context indicates which meaning of I is intended. For example, consider the equation $\mathcal{M}(I) = I$; on the left side I denotes the identity operator, and on the right side I denotes the identity matrix.

If A is a square matrix (with entries in **F**, as usual) of the same size as I, then $AI = IA = A$, as you should verify.

3.80 definition: *invertible, inverse,* A^{-1}

A square matrix A is called *invertible* if there is a square matrix B of the same size such that $AB = BA = I$; we call B the *inverse* of A and denote it by A^{-1}.

Some mathematicians use the terms ***nonsingular*** *and* ***singular****, which mean the same as invertible and non-invertible.*

The same proof as used in 3.60 shows that if A is an invertible square matrix, then there is a unique matrix B such that $AB = BA = I$ (and thus the notation $B = A^{-1}$ is justified).

If A is an invertible matrix, then $\left(A^{-1}\right)^{-1} = A$ because

$$A^{-1}A = AA^{-1} = I.$$

Also, if A and C are invertible square matrices of the same size, then AC is invertible and $(AC)^{-1} = C^{-1}A^{-1}$ because

$$\begin{aligned}(AC)\left(C^{-1}A^{-1}\right) &= A\left(CC^{-1}\right)A^{-1} \\ &= AIA^{-1} \\ &= AA^{-1} \\ &= I,\end{aligned}$$

and similarly $\left(C^{-1}A^{-1}\right)(AC) = I$.

The next result holds because we defined matrix multiplication to make it true—see 3.43 and the material preceding it. Now we are just being more explicit about the bases involved.

3.81 *matrix of product of linear maps*

Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$. If $u_1, \ldots, u_m$ is a basis of U, $v_1, \ldots, v_n$ is a basis of V, and $w_1, \ldots, w_p$ is a basis of W, then

$$\begin{aligned}&\mathcal{M}\left(ST, (u_1, \ldots, u_m), (w_1, \ldots, w_p)\right) = \\ &\qquad \mathcal{M}\left(S, (v_1, \ldots, v_n), (w_1, \ldots, w_p)\right)\mathcal{M}\left(T, (u_1, \ldots, u_m), (v_1, \ldots, v_n)\right).\end{aligned}$$

The next result deals with the matrix of the identity operator I with respect to two different bases. Note that the k^{th} column of $\mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)$ consists of the scalars needed to write u_k as a linear combination of the basis $v_1, \dots, v_n$.

In the statement of the next result, I denotes the identity operator from V to V. In the proof, I also denotes the n-by-n identity matrix.

3.82 *matrix of identity operator with respect to two bases*

Suppose that $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then the matrices

$$\mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big) \quad \text{and} \quad \mathcal{M}\big(I, (v_1, \dots, v_n), (u_1, \dots, u_n)\big)$$

are invertible, and each is the inverse of the other.

Proof In 3.81, replace w_k with u_k, and replace S and T with I, getting

$$I = \mathcal{M}\big(I, (v_1, \dots, v_n), (u_1, \dots, u_n)\big)\mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big).$$

Now interchange the roles of the u's and v's, getting

$$I = \mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)\mathcal{M}\big(I, (v_1, \dots, v_n), (u_1, \dots, u_n)\big).$$

These two equations above give the desired result. ∎

3.83 example: *matrix of identity operator on* $\mathbf{F}^2$ *with respect to two bases*

Consider the bases $(4, 2), (5, 3)$ and $(1, 0), (0, 1)$ of $\mathbf{F}^2$. Because $I(4, 2) = 4(1, 0) + 2(0, 1)$ and $I(5, 3) = 5(1, 0) + 3(0, 1)$, we have

$$\mathcal{M}\Big(I, \big((4, 2), (5, 3)\big), \big((1, 0), (0, 1)\big)\Big) = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}.$$

The inverse of the matrix above is

$$\begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix},$$

as you should verify. Thus 3.82 implies that

$$\mathcal{M}\Big(I, \big((1, 0), (0, 1)\big), \big((4, 2), (5, 3)\big)\Big) = \begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}.$$

Our next result shows how the matrix of T changes when we change bases. In the next result, we have two different bases of V, each of which is used as a basis for the domain space and as a basis for the target space. Recall our shorthand notation that allows us to display a basis only once when it is used in both capacities:

$$\mathcal{M}\big(T, (u_1, \dots, u_n)\big) = \mathcal{M}\big(T, (u_1, \dots, u_n), (u_1, \dots, u_n)\big).$$

3.84 *change-of-basis formula*

Suppose $T \in \mathcal{L}(V)$. Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Let

$$A = \mathcal{M}\big(T, (u_1, \dots, u_n)\big) \quad \text{and} \quad B = \mathcal{M}\big(T, (v_1, \dots, v_n)\big)$$

and $C = \mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)$. Then

$$A = C^{-1}BC.$$

Proof In 3.81, replace w_k with u_k and replace S with I, getting

3.85 $$A = C^{-1}\mathcal{M}\big(T, (u_1, \dots, u_n), (v_1, \dots, v_n)\big),$$

where we have used 3.82.

Again use 3.81, this time replacing w_k with v_k. Also replace T with I and replace S with T, getting

$$\mathcal{M}\big(T, (u_1, \dots, u_n), (v_1, \dots, v_n)\big) = BC.$$

Substituting the equation above into 3.85 gives the equation $A = C^{-1}BC$. ■

The proof of the next result is left as an exercise.

3.86 *matrix of inverse equals inverse of matrix*

Suppose that $v_1, \dots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$ is invertible. Then $\mathcal{M}(T^{-1}) = \big(\mathcal{M}(T)\big)^{-1}$, where both matrices are with respect to the basis $v_1, \dots, v_n$.

Exercises 3D

1 Suppose $T \in \mathcal{L}(V, W)$ is invertible. Show that T^{-1} is invertible and

$$\big(T^{-1}\big)^{-1} = T.$$

2 Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$ are both invertible linear maps. Prove that $ST \in \mathcal{L}(U, W)$ is invertible and that $(ST)^{-1} = T^{-1}S^{-1}$.

3 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent.

(a) T is invertible.

(b) $Tv_1, \dots, Tv_n$ is a basis of V for every basis $v_1, \dots, v_n$ of V.

(c) $Tv_1, \dots, Tv_n$ is a basis of V for some basis $v_1, \dots, v_n$ of V.

4 Suppose V is finite-dimensional and $\dim V > 1$. Prove that the set of noninvertible linear maps from V to itself is not a subspace of $\mathcal{L}(V)$.

5 Suppose V is finite-dimensional, U is a subspace of V, and $S \in \mathcal{L}(U, V)$. Prove that there exists an invertible linear map T from V to itself such that $Tu = Su$ for every $u \in U$ if and only if S is injective.

6 Suppose that W is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} S = \operatorname{null} T$ if and only if there exists an invertible $E \in \mathcal{L}(W)$ such that $S = ET$.

7 Suppose that V is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} S = \operatorname{range} T$ if and only if there exists an invertible $E \in \mathcal{L}(V)$ such that $S = TE$.

8 Suppose V and W are finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that there exist invertible $E_1 \in \mathcal{L}(V)$ and $E_2 \in \mathcal{L}(W)$ such that $S = E_2 T E_1$ if and only if $\dim \operatorname{null} S = \dim \operatorname{null} T$.

9 Suppose V is finite-dimensional and $T \colon V \to W$ is a surjective linear map of V onto W. Prove that there is a subspace U of V such that $T|_U$ is an isomorphism of U onto W.

Here $T|_U$ means the function T restricted to U. Thus $T|_U$ is the function whose domain is U, with $T|_U$ defined by $T|_U(u) = Tu$ for every $u \in U$.

10 Suppose V and W are finite-dimensional and U is a subspace of V. Let

$$\mathcal{E} = \{T \in \mathcal{L}(V, W) : U \subseteq \operatorname{null} T\}.$$

(a) Show that $\mathcal{E}$ is a subspace of $\mathcal{L}(V, W)$.
(b) Find a formula for $\dim \mathcal{E}$ in terms of $\dim V$, $\dim W$, and $\dim U$.

Hint: Define $\Phi \colon \mathcal{L}(V, W) \to \mathcal{L}(U, W)$ by $\Phi(T) = T|_U$. What is $\operatorname{null} \Phi$*? What is* $\operatorname{range} \Phi$*?*

11 Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that

$$ST \text{ is invertible} \iff S \text{ and } T \text{ are invertible.}$$

12 Suppose V is finite-dimensional and $S, T, U \in \mathcal{L}(V)$ and $STU = I$. Show that T is invertible and that $T^{-1} = US$.

13 Show that the result in Exercise 12 can fail without the hypothesis that V is finite-dimensional.

14 Prove or give a counterexample: If V is a finite-dimensional vector space and $R, S, T \in \mathcal{L}(V)$ are such that RST is surjective, then S is injective.

15 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_m$ is a list in V such that $Tv_1, \dots, Tv_m$ spans V. Prove that $v_1, \dots, v_m$ spans V.

16 Prove that every linear map from $\mathbf{F}^{n,1}$ to $\mathbf{F}^{m,1}$ is given by a matrix multiplication. In other words, prove that if $T \in \mathcal{L}(\mathbf{F}^{n,1}, \mathbf{F}^{m,1})$, then there exists an m-by-n matrix A such that $Tx = Ax$ for every $x \in \mathbf{F}^{n,1}$.

17 Suppose V is finite-dimensional and $S \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by

$$\mathcal{A}(T) = ST$$

for $T \in \mathcal{L}(V)$.

(a) Show that $\dim \operatorname{null} \mathcal{A} = (\dim V)(\dim \operatorname{null} S)$.

(b) Show that $\dim \operatorname{range} \mathcal{A} = (\dim V)(\dim \operatorname{range} S)$.

18 Show that V and $\mathcal{L}(\mathbf{F}, V)$ are isomorphic vector spaces.

19 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has the same matrix with respect to every basis of V if and only if T is a scalar multiple of the identity operator.

20 Suppose $q \in \mathcal{P}(\mathbf{R})$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbf{R})$ such that

$$q(x) = (x^2 + x)p''(x) + 2xp'(x) + p(3)$$

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

21 Suppose n is a positive integer and $A_{j,k} \in \mathbf{F}$ for all $j, k = 1, \dots, n$. Prove that the following are equivalent (note that in both parts below, the number of equations equals the number of variables).

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

$$\sum_{k=1}^{n} A_{1,k} x_k = 0$$
$$\vdots$$
$$\sum_{k=1}^{n} A_{n,k} x_k = 0.$$

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

$$\sum_{k=1}^{n} A_{1,k} x_k = c_1$$
$$\vdots$$
$$\sum_{k=1}^{n} A_{n,k} x_k = c_n.$$

22 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Prove that

$$\mathcal{M}(T, (v_1, \dots, v_n)) \text{ is invertible} \iff T \text{ is invertible.}$$

23 Suppose that $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Let $T \in \mathcal{L}(V)$ be such that $Tv_k = u_k$ for each $k = 1, \dots, n$. Prove that

$$\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n)).$$

24 Suppose A and B are square matrices of the same size and $AB = I$. Prove that $BA = I$.