5 Isomorphisms and Similarity

5.1 Isomorphisms

In the previous section, we looked at a lot of linear transformations. In particular, we saw the coordinate map $[\cdot]_E : U \to \mathbb{R}^n$ which sends a vector $\mathbf{u} \in U$ to its coordinates with respect to E. We observed that this mapping in fact goes both ways: if we have a vector we can compute the coordinates, and if we have coordinates we can compute the vectors. Functions like this are very important and have a special name.

Definition 5.1. Let $f : U \to V$ be a function. If there is a $g : V \to U$ such that $g(f(\mathbf{u})) = \mathbf{u}$ for all $u \in U$, and $f(g(\mathbf{v})) = \mathbf{v}$ for all $\mathbf{v} \in V$, then we say that $g = f^{-1}$ is the *inverse* of f, and that f is *invertible*.

If f is an invertible linear transformation, we say that f is an *isomorphism* between U and V.

If U and V are vector spaces, we say they are *isomorphic* if there exists an isomorphism from U to V. We write $U \cong V$.

Remark 5.2. We will see that if two spaces are isomorphic, we can treat them as essentially the same. This does not mean they are the same; $\mathbb{R}^5$ is not the same thing as the space of degree-four polynomials.

But if $U \cong V$ then they are the same *as vector spaces*, because if we want to do something to U, we can instead map it to V, do it there, and then map it back.

Example 5.3. Let U be a vector space with basis $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, and let $f : U \to \mathbb{R}^n$ be defined by $f(\mathbf{u}) = [\mathbf{u}]_E$. Then f is invertible, and the inverse of f is given by the function $g(a_1, \dots, a_n) = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n$. To prove this, we check two things.

For any $\mathbf{u} \in U$ we can write $\mathbf{u} = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n$. Then

$$g(f(\mathbf{u})) = g(f(a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n)) = g(a_1, \dots, a_n) = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n = \mathbf{u}.$$

Similarly, for any $(a_1, \dots, a_n) \in \mathbb{R}^n$ we have

$$f(g(a_1, \dots, a_n)) = f(a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n) = (a_1, \dots, a_n).$$

Thus g is the inverse of f by definition.

Example 5.4. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $f(x, y) = (x + y, x - y)$. Then f is invertible, with inverse $g(a, b) = \left(\frac{a+b}{2}, \frac{a-b}{2}\right)$. To prove this we check:

$$g(f(x, y)) = g(x + y, x - y) = \left(\frac{(x + y) + (x - y)}{2}, \frac{(x + y) - (x - y)}{2}\right) = (x, y)$$
$$f(g(a, b)) = f\left(\frac{a + b}{2}, \frac{a - b}{2}\right) = \left(\frac{a + b}{2} + \frac{a - b}{2}, \frac{a + b}{2} - \frac{a - b}{2}\right) = (a, b).$$

Thus g is the inverse of f by definition.

So far we can check whether a given g is the inverse of f, but we don't have a good way of determining if a function is invertible. In order to do that, we have to recall a few properties of functions. (Discrete Math covers these topics in much more detail; we need much less knowledge and information about them).

Definition 5.5.

- A function f is *one-to-one* or *injective* if it has the property that: if $f(x) = f(y)$ then $x = y$. This tells us that anything in the image of f is only in the image once.
- A function $f : A \to B$ is *onto* or *surjective* if the image of f is B. That is, f is onto if for every $b \in B$ there is an $a \in A$ with $f(a) = b$. This tells us we can reach every element of the codomain from some element of the domain.
- A function f is *bijective* if it is both one-to-one and onto.

Proposition 5.6. *Let $L : U \to V$ be a linear transformation of vector spaces. Then:*

1. *L is one-to-one if and only if* $\ker(L) = \{\mathbf{0}\}$.
2. *L is invertible if and only if L is bijective.*

Proof.

1. See Homework 8.
2. Suppose L is bijective. Define a transformation $T : V \to U$ as follows: let $\mathbf{v} \in V$. Then L is onto, so by definition there is some $\mathbf{u} \in U$ such that $L(\mathbf{u}) = \mathbf{v}$. Since L is one-to-one there is only one such element, since if $L(\mathbf{u}) = L(\mathbf{u}_1)$ then $\mathbf{u} = \mathbf{u}_1$ by definition of one-to-one. Define $T(\mathbf{v}) = \mathbf{u}$.

   Then for any $\mathbf{v} \in V$, we have $L(T(\mathbf{v})) = L(\mathbf{u}) = \mathbf{v}$, and for any $\mathbf{u} \in U$ we have $T(L(\mathbf{u})) = T(\mathbf{v}) = \mathbf{u}$. Thus by definition, $T = L^{-1}$.

Conversely, suppose L is invertible, and let $T = L^{-1}$. Suppose $L(\mathbf{u}) = L(\mathbf{v})$. Then $T(L(\mathbf{u})) = T(L(\mathbf{v}))$, but $T(L(\mathbf{u})) = \mathbf{u}$ and $T(L(\mathbf{v})) = \mathbf{v}$, so $\mathbf{u} = \mathbf{v}$, and by definition L is one-to-one.

Let $\mathbf{v} \in V$. Then $T(\mathbf{v}) \in U$, and $L(T(\mathbf{v})) = \mathbf{v}$, so $\mathbf{v} \in L(U)$ for any $\mathbf{v}\ inV$. Thus L is onto by definition, and since it is one-to-one and onto, it is bijective.

□

Remark 5.7. Another way to think of the second result is that "onto" guarantees that every element $\mathbf{v} \in V$ has an inverse, and "one-to-one" guarantees that no element has more than one, so the inverse function is actually well-defined as a function.

We can contrast with, say, the function $f(x) = x^2$. This function is not one-to-one, so when we ask for the inverse or square root of 4, we get two possible answers. (This function isn't linear, but we didn't actually use linearity anywhere in the previous proof, and in fact it works for all functions).

Proposition 5.6 gives us an easy way to check if a linear function is injective, and if we also check that it is surjective we can easily see whether it is invertible. We can always check surjectivity directly, and on your homework you will, but we'd like to make this easier as well. To do that we want to convert the Rank-Nullity Theorem to discuss all linear transformations. First we need to lay some groundwork.

Proposition 5.8. *Let $L : U \to V$ be a linear transformation.*

1. *If L is one-to-one, then the image of a linearly independent set is linearly independent.*
2. *If L is onto, then the image of a spanning set is a spanning set.*
3. *If L is bijective, then the image of a basis is a basis.*

Proof.

1. Let $S = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be linearly independent. Suppose $a_1 L(\mathbf{u}_1) + \cdots + a_n L(\mathbf{u}_n) = \mathbf{0}$. Then by linearity, $L(a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n) = \mathbf{0}$, and since L is injective, this means that $a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0}$. Finally, since S is linearly independent, this implies that $a_1 = \cdots = a_n = 0$.
2. Suppose $S = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a spanning set for U. We wish to prove that the set $L(S) = \{L(\mathbf{u}_1), \dots, L(\mathbf{u}_n)\}$ is a spanning set for V.

Let $\mathbf{v} \in V$. Then since S is onto, there is a $\mathbf{u} \in U$ such that $L(\mathbf{u}) = \mathbf{v}$. Since S is a spanning set, we can write $\mathbf{u} = a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n$, and then we have

$$\begin{aligned} \mathbf{v} = L(\mathbf{u}) &= L(a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n) \\ &= a_1L(\mathbf{u}_1) + \cdots + a_nL(\mathbf{u}_n) \end{aligned}$$

and thus $\mathbf{v}$ is a linear combination of elements of $L(S)$. So $L(S)$ spans V.

3. This follows from the previous two results. If L is bijective then it is one-to-one and onto, and if B is a basis then it is linearly independent and spanning, and its image under L is also linearly independent and spans, and thus $L(B)$ is a basis.

□

Corollary 5.9. *If* $U \cong V$ *then* $\dim U = \dim V$.

Proof. Let $B = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis for U, so $\dim U = n$. Then by Proposition 5.8 $L(B) = \{L(\mathbf{e}_1), \ldots, L(\mathbf{e}_n)\}$ is a basis for V, so $\dim V = n = \dim U$. □

Theorem 5.10 (Rank-Nullity Theorem)**.** *If* U, V *are finite-dimensional, then* $\dim U = \dim \ker(L) + \dim L(U)$.

Proof. This follows from the rank-nullity theorem on matrices. If A is the matrix of U with respect to any bases, then we have that the number of columns of A is equal to the rank of A plus the nullity of A. But:

The number of columns of A is equal to the dimension of the domain, so is equal to $\dim U$.

We claim that the columnspace of A is isomorphic to $L(U)$, with the isomorphism given by $[\cdot]_F$. We already know that $[\cdot]_F$ is injective, so we just need to prove that $[L(U)]_F = \mathrm{Col}(A)$.

Recall the columnspace of A is precisely the vectors b such that $A\mathbf{x} = \mathbf{b}$ has a solution; thus $\mathrm{Col}(A) = \{[\mathbf{v}]_F : [\mathbf{v}]_F = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n\}$. But $[\mathbf{v}]_F = A\mathbf{x}$ for some $\mathbf{x}$ if and only if $\mathbf{v} = L(\mathbf{u})$ for some $\mathbf{u} \in U$, so $[\mathbf{v}]_F \in \mathrm{Col}(A)$ if and only if $\mathbf{v} \in L(U)$. Thus $[L(U)]_F = \mathrm{Col}(A)$, which proves our claim.

But since $L(U) \cong \mathrm{Col}(A)$, we know that $\dim L(U) = \dim \mathrm{Col}(A) = \mathrm{rk}(A)$.

Finally, by the same reasoning, the kernel of L is isomorphic to the nullspace of A, with the isomorphism given by $[\cdot]_E$. Thus the nullity of A is equal to $\dim \ker(L)$.

Since the Rank-Nullity theorem for matrices tells us that the number of columns of A is the rank of A plus the nullity of A; by the previous three identites, we have $\dim U = \dim L(U) + \dim \ker(L)$. □

Corollary 5.11. *Let $L : U \to V$ be linear. Then L is one-to-one if and only if $\dim U = \dim L(U)$, and L is onto if and only if $\dim V = \dim(U) - \dim \ker(L)$.*

If $\dim U = \dim V$, L is an isomorphism if and only if $\ker(L) = \{\mathbf{0}\}$.

Now we can easily determine if a transformation is invertible. But how do we find the inverse? Like everything in linear algebra, it's easier to do computations if we change things to be a matrix.

Proposition 5.12. *Let $f : U \to V$ be a linear transformation of finite dimensional vector spaces, and let E, F be bases for U, V respectively. Let A be the matrix of f with respect to E, F. Then f is invertible if and only if A is invertible, and the matrix of f^{-1} is A^{-1}.*

Proof. Suppose f is invertible, and that the matrix of f is A and the Let B be the matrix of f^{-1}. Then for any $\mathbf{u} \in U$,

$$[f^{-1}(f(\mathbf{u}))]_E = B[f(\mathbf{u})]_F = BA[\mathbf{u}]_E$$
$$[f^{-1}(f(\mathbf{u}))]_E = [\mathbf{u}]_E$$

and thus $BA[\mathbf{u}]_E = [\mathbf{u}]_E$ for all $\mathbf{u} \in U$. Thus $BA\mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$, and thus $BA = I_n$. So by definition $B = A^{-1}$.

Conversely, suppose the matrix of f is A, and A has an inverse A^{-1}. Let g be the function corresponding to A^{-1}, so for all $\mathbf{v} \in V$ we have $[g(\mathbf{v})]_E = A^{-1}[\mathbf{v}]_F$. Then for any $\mathbf{u} \in U, \mathbf{v} \in V$, we compute

$$[g(f(\mathbf{u}))]_E = A^{-1}[f(\mathbf{u})]_F = A^{-1}A[\mathbf{u}]_E = [\mathbf{u}]_E$$
$$[f(g(\mathbf{v}))]_F = A[g(\mathbf{v})]_E = AA^{-1}[\mathbf{v}]_F = [\mathbf{v}]_F.$$

Thus $g(f(\mathbf{u})) = \mathbf{u}$ and $f(g(\mathbf{v})) = \mathbf{v}$, so by definition $g = f^{-1}$. ☐

Corollary 5.13. *A $n \times n$ matrix is invertible if and only if its nullspace is trivial.*

This gives us a method for finding the inverse of any linear transformation: we find the matrix of the transformation, use Gaussian elimination to invert the matrix, and then return to the corresponding transformation.

Example 5.14. Let $L(x, y, z) = (x + y + z, 2x + 3y + 2z, x + 5y + 4z)$. What is L^{-1}?

The matrix for L (with respect to the standard basis) is $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 1 & 5 & 4 \end{bmatrix}$. So we compute

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 1 & 5 & 4 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 4 & 3 & -1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 3 & -1 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 3 & 7 & -4 & 1 \end{array}\right]$$
$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 3 & -1 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 7/3 & -4/3 & 1/3 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2/3 & 1/3 & -1/3 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 7/3 & -4/3 & 1/3 \end{array}\right]$$

Since A is invertible, this tells us that L is also invertible. And from A^{-1} we can see that

$$L^{-1}(a, b, c) = (2a/3 + b/3 - c/3, -2a + b, 7a/3 - 4b/3 + c/3).$$

We can check by multiplying the original matrices together and seeing that we get the identity, or by computing $L^{-1}(L(x, y, z))$ and confirming that we get $(x, y, z$.

Example 5.15. Let $E : \mathcal{P}_2(x) \to \mathbb{R}^3$ be given by $E(f(x)) = (f(-1), f(0), f(1))$. Can we find an inverse for this function?

Let $\{1, x, x^2\}$ be the basis for $\mathcal{P}_2(x)$, and use the standard basis for $\mathbb{R}^3$. Then the matrix of this transformation is $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. To find the inverse, we compute

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \end{array}\right]$$
$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 & -2 & 1 \end{array}\right]$$
$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1/2 & -1 & 1/2 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 1/2 & -1 & 1/2 \end{array}\right].$$

We can double-check our work, again, by multiplying out the original matrices.

What have we concluded? If we have a quadratic polynomial such that $f(-1) = a, f(0) = b, f(1) = 3$, then we must have

$$f(x) = b + (c/2 - a/2)x + (a/2 + c/2 - b)x^2.$$

Thus we can use this technique to find the (minimal degree) polynomial that goes through a given set of points.

Example 5.16. Let $D : \mathbb{P}_3(x) \to \mathbb{P}_3(x)$ be given by the derivative map. Is this function invertible?

The function is not invertible, since it has non-trivial kernel. We can also see this by writing down the matrix relative to the obvious basis: $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Since there is a row of all zeroes, the rows are not linearly independent, so the matrix is not invertible.

Let's tweak things a bit. Let $Q = \{ax + bx^2 + cx^3\} \subset \mathcal{P}_3(x)$, and let $D : Q \to \mathcal{P}_2(x)$ be given by the derivative. Then if we let $E = \{x, x^2, x^3\}$, $F = \{1, x, x^2\}$ be bases for the domain and codomain, we see the matrix is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ whose inverse is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$. Thus this function is invertible; in fact we see the inverse sends $a + bx + cx^2 \mapsto ax + bx^2/2 + cx^3/3$, which you should recognize as an integral.

5.2 Change of Basis

In the last section we said that two vector spaces are isomorphic if they're essentially the same, at least from the perspective of vector spaces. (Polynomials are not the same thing as lists of real numbers. But they are the same as far as the vector space structure goes). From this we should immediately assume that every space is isomorphic to *itself.*

Proposition 5.17. *Let V be a vector space. Then $V \cong V$. In particular, the identity map Id_V defined by $Id_V(\mathbf{v}) = \mathbf{v}$ is an isomorphism from V to V.*

But two isomorphic vector spaces can have more than one isomorphism between them. In fact, any non-trivial vector space has infinitely many isomorphisms from itself to itself, and these isomorphisms are extremely useful.

Proposition 5.18. *Let U, V be vector spaces, let $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be a basis for U, and $L : U \to V$ be a linear map. Then $L(E)$ spans $L(U)$, and L is an isomorphism if and only if $L(E)$ is a basis for V.*

Corollary 5.19. *Let U, V be vector spaces, with $\dim U = \dim V$. Then $U \cong V$.*

Proof. Let $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be a basis for U, and $F = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ a basis for V; we know they have the same number of elements since $\dim U = \dim V$. Define a linear map $L : U \to V$ by linearly extending $L(\mathbf{e}_i) = \mathbf{f}_i$; that is, we define

$$L(a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n) = a_1\mathbf{f}_1 + \cdots + a_n\mathbf{f}_n.$$

This map sends a basis to a basis, and thus is an isomorphism. □

Corollary 5.20. *A linear map from V to V is an isomorphism if and only if it sends a basis to a basis.*

Definition 5.21. We call such an isomorphism a *change of basis map*. The matrix of such an isomorphism is called *transition matrix*.

Example 5.22. We know that $E = \{(1,0,0),(0,1,0),(0,0,1)\}$ and $F = \{(1,0,0),(1,1,0),(1,1,1)\}$ are both bases for $\mathbb{R}^3$. Define $L : \mathbb{R}^3 \to \mathbb{R}^3$ by setting

$$\begin{aligned} L(1,0,0) &= (1,0,0) \\ L(0,1,0) &= (1,1,0) \\ L(0,0,1) &= (1,1,1) \\ A &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Then L is an isomorphism and a change of basis from E to F.

We can find the inverse in one of two ways. One is to use row reduction as usual, but that takes effort. Instead, we can note that the inverse is just the map that sends F to E:

$$\begin{aligned} L^{-1}(1,0,0) &= (1,0,0) \\ L^{-1}(1,1,0) &= (0,1,0) = (1,1,0) - (1,0,0) \\ L^{-1}(1,1,1) &= (0,0,1) = (1,1,1) - (1,1,0) \\ A^{-1} &= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

By multiplication or substitution we can check that this is definitely an inverse.

We can use this to write a vector, given to us in one basis, in another basis. But it turns out that this behaves *backwards* from what we'd expect.

Suppose we have bases $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $F = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ for a vector space V, and we have some vector $\mathbf{u} = a_1\mathbf{f}_1 + \cdots + a_n\mathbf{f}_n$.

Then we can write each $f_i = c_{1i}\mathbf{e}_1 + \cdots + c_{ni}\mathbf{e}_n$, and we have that $(c_{1i}, \dots, c_{ni})$ is the ith column vector of the matrix A that sends E to F. We have

$$\begin{aligned}
\mathbf{u} &= a_1\mathbf{f}_1 + \cdots + a_n\mathbf{f}_n \\
&= a_1(c_{11}\mathbf{e}_1 + \cdots + c_{n1}\mathbf{e}_n) + \cdots + a_n(c_{1n}\mathbf{e}_1 + c_{nn}\mathbf{e}_n) \\
&= (a_1c_{11} + \cdots + a_nc_{1n})\mathbf{e}_1 + \cdots + (a_1c_{n1} + \cdots + a_nc_{nn})\mathbf{e}_n \\
[\mathbf{u}]_E &= \begin{bmatrix} a_1c_{11} + \cdots + a_nc_{1n} \\ \vdots \\ a_1c_{n1} + \cdots + a_nc_{nn} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = A[\mathbf{u}]_F.
\end{aligned}$$

Why do I say this is weird? Because A is the matrix that sends E to F; but it is also the matrix that tranforms coordinates in F to coordinates in E. (The fancy word for this is "contravariant"). For this reason, we say that the matrix that sends E to F is in fact the *transition matrix from F to E.*

Example 5.23 (continued)**.** Suppose we'd like to take the vector $\mathbf{u} = 2(1, 0, 0) + 3(1, 1, 0) + 5(1, 1, 1)$ and find its coordinates in the standard basis. We have

$$\begin{aligned}
\mathbf{u} &= 2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2(\mathbf{e}_1) + 3(\mathbf{e}_1 + \mathbf{e}_2) + 5(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) \\
&= (2 + 3 + 5)\mathbf{e}_1 + (3 + 5)\mathbf{e}_2 + 5\mathbf{e}_3 \qquad \text{or} \\
[\mathbf{u}]_E &= A[\mathbf{u}]_F = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 10 \\ 8 \\ 5 \end{bmatrix}.
\end{aligned}$$

In many ways it's more useful to do things the other way. Suppose we have the vector $(5, 2, 7)$ and want to express it as a linear combination of elements of F. Then we need the transition matrix from E to F, which is A^{-1}. So we have

$$A^{-1}\begin{bmatrix} 5 \\ 2 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 7 \end{bmatrix} = \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix}.$$

Thus
$$\begin{bmatrix} 5 \\ 2 \\ 7 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 5 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 7 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Example 5.24. Let's represent the polynomial $a + bx + cx^2 \in \mathcal{P}_3(x)$ as a linear combination of $F = \{1, 2x, 4x^2 - 2\}$.

We take $E = \{1, x, x^2\}$ to be the standard basis, and if A is the transition matrix from F to E we have

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 4 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/4 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1/2 \\ 0 & 1 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/4 \end{array}\right]$$

$$A^{-1} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}.$$

Thus we have
$$[a + bx + cx^2]_F = A^{-1} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a + c/2 \\ b/2 \\ c/4 \end{bmatrix}.$$

Thus $[a + bx + cx^2]_F = (a + c/2, b/2, c/4)$ and
$$a + bx + cx^2 = (a + c/2)(1) + (b/2)(2x) + (c/4)(4x^2 - 2).$$

I want to mention one last idea here, which is the ability to paste transition matrices together. If A is the transition matrix from F to E, and B is the transition matrix from G to F, then AB is the transition matrix from G to E. This is primarily useful once we introduce inverses.

Example 5.25. Let $E = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}$, $F = \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ be two bases for $\mathbb{R}^3$. Let $\mathbf{u} = 3 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$. Let's find the coordinates of $\mathbf{u}$ with respect to F.

We could try to compute the transition matrix directly, but that requires us to do a bunch of equation solving. Instead, we notice that the transition matrix from E to the standard basis is

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

and the transition matrix from F to the standard basis is

$$B = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

We want the transition matrix from E to F so that we can convert coordinates from E to F. Thus the matrix we actually want is $B^{-1}A$: A takes us from E to the standard basis, and then B^{-1} takes us from the standard basis to F. We compute

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 2 & 1 & 1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1/2 & 1/2 & 1/2 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1/2 & -1/2 & -1/2 \\ 0 & 1 & 1 & 1/2 & 1/2 & -1/2 \\ 0 & 0 & 1 & 1/2 & 1/2 & 1/2 \end{array}\right]$$

and thus

$$B^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$B^{-1}A = \frac{1}{2}\begin{bmatrix} -2 & -1 & 1 \\ 2 & 1 & 3 \\ 4 & 3 & 1 \end{bmatrix}$$

$$[\mathbf{u}]_F = B^{-1}A[\mathbf{u}]_E = \frac{1}{2}\begin{bmatrix} -2 & -1 & 1 \\ 2 & 1 & 3 \\ 4 & 3 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -3 \\ 7 \\ 7 \end{bmatrix} = \begin{bmatrix} -3/2 \\ 7/2 \\ 7/2 \end{bmatrix}.$$

We check that, indeed,

$$\frac{-3}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{7}{2}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + \frac{7}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

5.3 Similarity

We now want to return to talking about general linear transformations, but bringing with us our new perspective on bases and changes of bases.

Let $R_{\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by a rotation ninety degrees counterclockwise. We saw earlier that with respect to the standard basis, this transformation has matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. But we can also compute the matrix with respect to, say, $F = \{(1,0),(1,1)\}$. Then we have

$$\begin{aligned} R_{\pi/2}(1,0) &= (0,1) = (1,1) - (1,0) \to (-1,1) \\ R_{\pi/2}(1,1) &= (-1,1) = (1,1) - 2(1,0) \to (-2,1) \\ B &= \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

These two matrices represent the same transformation, with respect to different bases. But they are clearly not the same matrix! What's going on here?

The answer is that we changed the coordinate system, and so our matrix changed. After we account for that, we should get the same matrix. To account for this, we need the change of basis matrix between F and the standard basis E. We have

$$U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

the transition matrix from F to the standard basis, and thus

$$U^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

is the transition matrix from the standard basis to F.

If we want to perform the operation $R_{\pi/2}$ on the vectors of F, we can use the matrix B that we found. Alternatively, we can transform our vectors into E-coordinates, use the matrix A, and then transform back into F-coordinates. This operation would be given by $U^{-1}AU$. We calculate that

$$U^{-1}AU = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}.$$

This is the same as the matrix B, as it should be.

Definition 5.26. If A and B are $n \times n$ matrices, we say they are *similar* if there is some invertible matrix U such that $B = U^{-1}AU$. We write $A \sim B$.

Proposition 5.27. *Let $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}, F = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be two bases for V, and let $L : V \to V$ be a linear function. Let U be the transition matrix from F to E.*

If A is the matrix representing L with respect to E, and B is the matrix representing L with respect to F, then $B = U^{-1}AU$.

Example 5.28. Let $D : \mathcal{P}_2(x) \to \mathcal{P}_2(x)$ be the differentiation operator. Let's find the matrix of D with respect to $E = \{1, x, x^2\}$ and with respect to $F = \{1, 2x, 4x^2 - 2\}$.

We've already seen that the matrix of D with spect to E is $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

We can work out the matrix with respect to F directly:

$$\begin{aligned} D(1) &= 0 \to (0,0,0) \\ D(2x) &= 2 \to (2,0,0) \\ D(4x^2 - 2) &= 8x \to (0,4,0) \\ B &= \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Alternatively, we could recall that the change of basis matrices between E and F:

$$\begin{aligned} E \to F : \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} &= U^{-1} \\ F \to E : \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} &= U. \end{aligned}$$

So we can compute the matrix B for D by saying

$$\begin{aligned} B = U^{-1}AU &= \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Example 5.29. Let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be given by $L(x, y, z) = (x + 3y + z, 2x - y + 3z, y - z)$. Find the matrix of L with respect to $\{(4, 1, 2), (3, 0, 1), (1, -1, 0)\}$, and show it is similar to the matrix with respect to the standard basis.

We have

$$\begin{aligned} L(1,0,0) &= (1,2,0) \\ L(0,1,0) &= (3,-1,1) \\ L(0,0,1) &= (1,3,-1) \\ A &= \begin{bmatrix} 1 & 3 & 1 \\ 2 & -1 & 3 \\ 0 & 1 & -1 \end{bmatrix}. \end{aligned}$$

We can compute the change of basis matrices. If U is the matrix from F to E, then we have

$$U = \begin{bmatrix} 4 & 3 & 1 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$$

and

$$\left[\begin{array}{ccc|ccc} 4 & 3 & 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 4 & 3 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 3 & 5 & 1 & -4 & 0 \\ 0 & 1 & 2 & 0 & -2 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & -2 & 1 \\ 0 & 3 & 5 & 1 & -4 & 0 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & -2 & 1 \\ 0 & 0 & -1 & 1 & 2 & -3 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & -2 & 1 \\ 0 & 0 & 1 & -1 & -2 & 3 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -1 & 3 \\ 0 & 1 & 0 & 2 & 2 & -5 \\ 0 & 0 & 1 & -1 & -2 & 3 \end{array}\right]$$

so we have

$$U^{-1} = \begin{bmatrix} -1 & -1 & 3 \\ 2 & 2 & -5 \\ -1 & -2 & 3 \end{bmatrix}.$$

Thus to find the matrix with respect to F, we can compute

$$B = U^{-1}AU = \begin{bmatrix} -1 & -1 & 3 \\ 2 & 2 & -5 \\ -1 & -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 2 & -1 & 3 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & 3 & 1 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} -3 & 1 & -7 \\ 6 & -1 & 13 \\ -5 & 2 & -10 \end{bmatrix} \begin{bmatrix} 4 & 3 & 1 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -25 & -16 & -4 \\ 49 & 31 & 7 \\ -38 & -25 & -7 \end{bmatrix}.$$

There's not a really efficient way to determine whether two matrices are similar in general, although we have a few tools that can tell us two matrices are *not* similar.

Proposition 5.30. *Let $A, B \in M_{n\times n}$ with $A \sim B$. Then:*

- *The rank of A is equal to the rank of B.*
- *The nullity of A is equal to the nullity of B.*
- *A is invertible if and only if B is invertible.*

We'll add to this list in section 7. But first we want to take a detour to do a bit of geometry.