3 Linear Transformations

In this chapter, we're going to finally begin exploring functions on vector spaces.[1] Linear algebra is the study of linear transformations on vector spaces. We did vector spaces, so once we do linear transformations, we should be done, right?[2] You will not be surprised that the implications for everything currently known that springs forth from the definition of a linear transformation *alone* can fill an entire chapter. Indeed, the entanglement and conceptual symbiosis of these two main characters in our story, vector spaces and linear transformations, could fill volumes.

1: It feels like this should be the end, right?

2: Yes!

Alas, we have but this one volume and this short time together to celebrate these two glorious concepts and their relationship with each other. Thus, we'll just have to hit the high points. In doing so, we'll find that linear transformations, while being an incredibly diverse and flexible type of function, can be understood *almost entirely* in a very systematic and concrete way.

As has become our custom at the beginning of chapters, though, we are getting a bit ahead of ourselves. Let's take a small step back to better prepare ourselves for the coming of linear transformations. Let's talk more about functions…

3.1 More Fun with Functions

There are a few more fun facts about functions for which a fresh look would be good.[3] You should be familiar, in some way, with each of the ideas we'll cover in this section, but you may not have previously seen this level of formality.[4] You are encouraged to go through this material very carefully and thoroughly; it will serve you well in future sections.

3: Favorable? Fine? Funicular? Well, at least two of those.

4: …and silliness. That should be expected by now, right?

Onto Functions

Definition 3.1.1 *For sets A and B and a function $f\colon A \to B$, the function f is* ***onto*** *if for every element $b \in B$, there is an element $a \in A$ such that f relates a to b, that is, $f(a) = b$.*

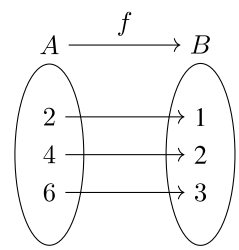
Besides the formal definition of onto, we also have a more geometric characterization that will be extremely useful from time to time.

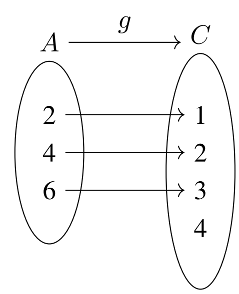
Theorem 3.1.1 *For sets A and B and a function $f\colon A \to B$, the function f is* ***onto*** *if and only if*

$$\operatorname{ran}(f) = \operatorname{codom}(f).$$

The proof of this theorem follows almost directly from Definition 3.1.1. Note that it *almost* follows.[5]

5:  Exercise!





Example 3.1.1 Let's see some examples!

- Let $A = \{2, 4, 6\}$ and $B = \{1, 2, 3\}$. Define the function $f\colon A \to B$ by the rule $f(a) = a/2$ for every $a \in A$. This function is onto because every element in B is half of one of the elements in A. If we had instead mapped $g\colon A \to C$ where $C = \{1, 2, 3, 4\}$ using the same rule as for f, then g would not be onto. This really illustrates how much the codomain has to do with this property.
- Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by the rule

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1.$$

  This map is onto. Let's convince ourselves.
  - First, let's look at a specific example. Consider 5. It's a real number, so it's in the codomain. What does f map to 5? Yep,

$$f\left(\begin{bmatrix} 5 \\ 0 \end{bmatrix}\right) = 5.$$

    Also,

$$f\left(\begin{bmatrix} 5 \\ 1 \end{bmatrix}\right) = 5.$$

    In fact,

$$f\left(\begin{bmatrix} 5 \\ x_2 \end{bmatrix}\right) = 5$$

    where x_2 is anything in $\mathbb{R}$. Let's be honest, there was nothing special about 5. Sometimes it's just nice to see how these things look with numbers, but to actually prove or verify that f is onto, we'll need to use a general element from the codomain.
  - Let $y \in \mathbb{R}$. Then

$$f\left(\begin{bmatrix} y \\ 0 \end{bmatrix}\right) = y.$$

    Since y is a general element of the codomain, we can conclude that f is onto. Note that we just needed to find *one* element in the domain that mapped to our general element y. As in our previous discussion, we could actually find infinitely many, but according to our definition of onto, that's overkill because we only need one.[6]

6:  There is still value in knowing precisely what in the domain maps to an element in the range. We'll see this again later.

- Let $f: \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x + 2$. Let's show this is onto! Let $y \in \mathbb{R}$ denote any element in the codomain. We need to find an x in the domain that f maps to y. That is, we need to solve $f(x) = y$ for a value of x. Then $f(x) = x + 2$, so we have $x + 2 = y$. Solving this for x we see $x = y - 2$. Well $f(y - 2) = (y - 2) + 2 = y$, so $y - 2$ is the correct x value to satisfy $f(x) = y$. Because we were able to do this for a general y in the codomain, we know f is onto.

Exploration 63 Consider the function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2 + 1$. This function is not onto. Find a value in $\mathbb{R}$ that is not in $\operatorname{ran}(f)$.

Exploration 64 Show that the map $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$ is onto. First, let y represent a general element in the codomain $\mathbb{R}$. Now, what must x be so that $f(x) = y$?

Exploration 65 For each function below, determine whether it is onto. If you are having trouble deciding, ask yourself whether there's anything *not* in the range of this function.[7] If you can't think of anything, see if there's an input that you could use to give you any desired output.

7:  How about an ice cream cone? I bet that's not in the range! Oh, maybe you should only consider things that are also in the codomain, too…

- Let $f: \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 x_2.$$

- Let $g: \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 x_2 \\ x_1 x_2 \end{bmatrix}.$$

One-to-one Functions

The definition companion to onto is one-to-one. This is an often misunderstood concept; at the heart of these misunderstandings is the oversimplification of this concept to something having to do with horizontal and vertical lines. We shall require a thorough understanding of this particular concept in a very broad context, so let's get the formal definition and just say no more about that standard oversimplification.

Definition 3.1.2 *For sets A and B and a function $f\colon A \to B$, the function f is* ***one-to-one*** *if for any $b \in \operatorname{ran}(f)$, we have $a_1 = a_2 \in A$ if $f(a_1) = b$ and $f(a_2) = b$.*

This definition is just requiring that if f maps both a_1 and a_2 to b, then it must be that $a_1 = a_2$. That way, no more than one element from the domain can be mapped by f to an element in the range of f.

Theorem 3.1.2 *For sets A and B and a function $f\colon A \to B$, the function f is* ***one-to-one*** *if and only if for all $a_1, a_2 \in A$,*

$$f(a_1) = f(a_2) \text{ implies } a_1 = a_2.$$

Note that in the definition, the universal quantifier (the "for any" bit) is on the elements of B, but in the equivalent definition of one-to-one given in Theorem 3.1.2, the universal quantifier is on the elements of A. This makes this theorem a bit trickier than Theorem 3.1.1.

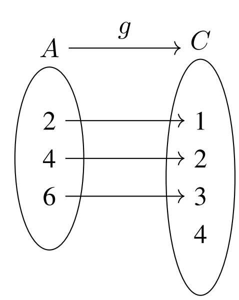
PROOF. Let's suppose the function $f\colon A \to B$ is one-to-one by our definition. That is, for any $b \in \operatorname{ran}(f)$, we have $a_1 = a_2$ whenever $f(a_1) = b$ and $f(a_2) = b$. Now, let a_1 and a_2 be in A and suppose they have the property that $f(a_1) = f(a_2)$. Then, $f(a_1) = f(a_2) = b$ for some $b \in B$. Thus, $a_1 = a_2$ by the definition of one-to-one.

Now, suppose we know the function $f\colon A \to B$ has the property that for all $a_1, a_2 \in A$, $f(a_1) = f(a_2)$ implies $a_1 = a_2$. Let $b \in \operatorname{ran}(f)$. Then we know there must be some $a_1 \in A$ such that $f(a_1) = b$. Suppose we also have $a_2 \in A$ such that $f(a_2) = b$. Then since $f(a_1) = b = f(a_2)$, we know $a_1 = a_2$. So this function is one-to-one by our definition. □

Example 3.1.2 Let's revisit some familiar functions, but now, we can ask whether they are one-to-one!

- Let $A = \{2, 4, 6\}$ and $C = \{1, 2, 3, 4\}$. Define the function $g\colon A \to C$ by the rule $g(a) = a/2$ for every $a \in A$. This function is one-to-one! To see this, suppose $c, d \in A$ are such that $g(c) = g(d)$. Using the definition of g, we see $c/2 = d/2$, which can be simplified to see $c = d$.
- Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by the rule

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1.$$



This map was onto, but it is *not* one-to-one. Recall that we had options about what mapped to 5. We saw that

$$f\left(\begin{bmatrix} 5 \\ 0 \end{bmatrix}\right) = 5 = f\left(\begin{bmatrix} 5 \\ 1 \end{bmatrix}\right).$$

This example with specific numbers is actually enough to show that the function does not satisfy our definition!

- Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x+2$. This one was onto, and it is also one-to-one! Suppose $a, b \in \mathbb{R}$ are such that $f(a) = f(b)$. then we know $a + 2 = b + 2$. This says $a = b$.

Exploration 66 Determine whether the following functions are one-to-one. If you are having trouble deciding, pick something in the range of the function. Ask yourself whether there are multiple inputs to get that same output.

- Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 x_2.$$

- Let $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 \end{bmatrix}.$$

As you probably expect, one-to-one and onto are both very nice properties for a function to have. Additionally, functions with both are really quite grand. We'll see why shortly.

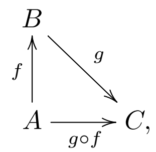
Composition of Functions

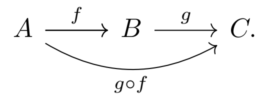
We saw composition of functions appear briefly in Section 2.3 to define inner product on vector spaces. Let us now have the formal definition.

Definition 3.1.3 *Let A, B, and C be sets and $f\colon A \to B$ and $g\colon B \to C$ be functions. The* ***composition of the functions*** *f* ***and*** *g is the function $(g \circ f)\colon A \to C$ such that $(a, c) \in g \circ f$ if and only if there is a $b \in B$ such that $(a, b) \in f$ and $(b, c) \in g$. That is, for any $a \in A$,*

$$(g \circ f)(a) = g(f(a)).$$

As was mentioned in Section 2.3, what makes the composition of functions work is the fact that outputs of f are inputs of g. This can be seen in both of the commuting diagrams below:





Example 3.1.3 Define

$$S = \left\{ \frac{z}{2} : z \in \mathbb{Z} \right\}.$$

Let $f\colon \mathbb{Z} \to S$ be defined by $f(z) = \dfrac{z}{2}$ and $g\colon S \to \mathbb{Z}$ be defined by $g(x) = 2x$. Then $g \circ f\colon \mathbb{Z} \to \mathbb{Z}$ is the map sending an integer z to itself since $(g \circ f)(z) = g(f(z)) = g\left(\dfrac{z}{2}\right) = 2\left(\dfrac{z}{2}\right) = z$.

Exploration 67 Let $A = \{1,2,3,4\}$, $B = \{2,4,6,8\}$, $C = \{0,1\}$. Let $f\colon A \to B$ be given by

$$f(a) = 2a \text{ for any } a \in A.$$

Let $g\colon B \to C$ be given by

$$g(b) = 0 \text{ for all } b \in B.$$

- ► Consider $g \circ f\colon A \to C$. We see $g(f(1)) = g(2) = 0$ and $g(f(2)) = g(4) = 0$. Since g maps everything to 0, this is the map from A to C which maps everything in A to 0.

Let $k\colon C \to B$ be defined by

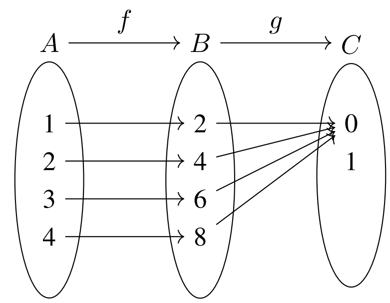
$$k(0) = 2 \text{ and } k(1) = 6.$$

- ► Consider $g \circ k\colon C \to C$. We see $g(k(0)) = g(2) = 0$ and $g(k(1)) = g(6) = 0$. Then $g \circ k$ is the map from C to itself sending both elements of C to 0.
- ► Consider $k \circ g\colon B \to B$. What is $k(g(2))$? What about $k(g(4))$? Describe $k \circ g$.

Let $h\colon A \to C$ be given by

$$h(a) = 0 \text{ if } a \in A \text{ is even and } h(a) = 1 \text{ if } a \in A \text{ is odd.}$$

- ► Consider $k \circ h\colon A \to B$. Describe this map.



Here's a picture for $g \circ f$. You should try drawing similar ones for the other functions mentioned.

Example 3.1.4 Suppose $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ are both one-to-one functions. We can show that the composition $f \circ g$ must also be one-to-one! Suppose we have $a_1, a_2 \in \mathbb{R}$ such that $(f \circ g)(a_1) = (f \circ g)(a_2)$. Then we know $f(g(a_1)) = f(g(a_2))$. Since we know f is one-to-one, we know it must be true that $g(a_1) = g(a_2)$. This is because $g(a_1)$ and $g(a_2)$ are both inputs to f with the same output from f. Now, we can use the fact that g

is one-to-one to say that $a_1 = a_2$ since $g(a_1) = g(a_2)$. This is what we needed! We started with two inputs to $f \circ g$ that had the same output from $f \circ g$ and were able to argue that the two inputs were really the same.

Exploration 68 Suppose $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ are both onto. Must $f \circ g$ be onto?

Invertible Functions

A problem that comes up a *lot* in mathematics and science is whether or not a process (often modeled by a function) can be undone in a reasonable way. For example, if I were to replace every song file on the internet with a single song file, then I strongly suspect many people would want a way to undo that process. Could the "undoing" be done with a function? That's not immediately clear, but the forward process certainly could: I hereby relate every song file on the internet to Jimmy Buffet's "Pencil Thin Mustache." [8] You're welcome world. Oh? What's that? You want a different function that relates "Pencil Thin Mustache" back to every song that used to be on the internet? Ha! Good luck with that! While there is a "Buffeting" function, I'm afraid the "Un-Buffeting" is not functional.[9]

Suppose instead that we related every song on the internet to itself, played backwards.[10] Well, if we did it again, we'd be back where we started. This is an example of a process that can be modeled by what we call an *invertible* function.

Definition 3.1.4 *A function $f\colon A \to B$ is* ***invertible*** *if there is another function $g\colon B \to A$ such that*

- *for all $a \in A$, $(g \circ f)(a) = a$, and*
- *for all $b \in B$, $(f \circ g)(b) = b$.*

If such a function exists, we call it the ***inverse of*** *f, and denote it f^{-1}.*

Example 3.1.5 Let's recall some maps from an earlier example. Define

$$S = \left\{ \frac{z}{2} \colon z \in \mathbb{Z} \right\}.$$

Let $f\colon \mathbb{Z} \to S$ be defined by $f(z) = \dfrac{z}{2}$ and $g\colon S \to \mathbb{Z}$ be defined by $g(x) = 2x$. Then we saw $g \circ f\colon \mathbb{Z} \to \mathbb{Z}$ has the property that $g \circ f(z) = z$ for any $z \in \mathbb{Z}$. We can also construct $f \circ g\colon S \to S$, and we see that $f \circ g = f(g(a)) = f(2a) = \dfrac{2a}{2} = a$ for any $a \in S$. Thus, $g = f^{-1}$.

8: This is a perfectly reasonable function with domain and codomain being all song files on the internet, but range just the single specific song file for "Pencil Thin Mustache."

9: This is not actually intended as a pun. A relation that is a function is often called a "functional" relation. However, we gladly accept the dual meaning here.

10: Disclaimer: The authors do not endorse the playing of songs backwards or the following of any nefarious instructions heard when doing so.

Example 3.1.6 Let's see an example of a function that is not invertible. Let $h\colon \mathbb{R} \to \mathbb{R}^2$ be defined by

$$h(x) = \begin{bmatrix} x \\ x \end{bmatrix}.$$

Note that h is not onto because any vector of the form

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

with $x \neq y$ will not be in the range. Suppose h has an inverse that we will call $k\colon \mathbb{R}^2 \to \mathbb{R}$. Then $(h \circ k)(\vec{v}) = \vec{v}$ for any $\vec{v} \in \mathbb{R}^2$. Let's consider then the vector

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbb{R}^2.$$

If k exists, then we have

$$(h \circ k)\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = h\left(k\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

This says, however, that

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \operatorname{ran}(h),$$

which is not true. Thus, the inverse k does not exist and h is not invertible.

Example 3.1.7 Let's see another function that fails to be invertible. Consider the function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ that relates a vector $\vec{x} \in \mathbb{R}^2$ to the vector with the same first coordinate but 0 for the second coordinate. That is, for any vector in $\mathbb{R}^2$,

$$\varphi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}.$$

The domain and codomain of φ is $\mathbb{R}^2$, but the range of φ is the horizontal line

$$\left\{\begin{bmatrix} x_1 \\ 0 \end{bmatrix} : x_1 \in \mathbb{R}\right\}.$$

Thus, the function φ collapses all of $\mathbb{R}^2$ (a plane!) to a single line. That doesn't sound like one-to-one behavior, so let's prove this function is not one-to-one. Note that

$$\varphi\left(\begin{bmatrix} 42 \\ 10 \end{bmatrix}\right) = \varphi\left(\begin{bmatrix} 42 \\ 11 \end{bmatrix}\right) = \begin{bmatrix} 42 \\ 0 \end{bmatrix}.$$

Since φ maps two vectors to the same vector in the codomain, it is not one-to-one. Why does this matter? Suppose g is a function that we'd like to be an inverse of φ. By Definition 3.1.4, for any vector $\vec{x} \in \mathbb{R}^2$, we must have $(g \circ \varphi)(\vec{x}) = \vec{x}$. However, that means

$$(g \circ \varphi)\left(\begin{bmatrix} 42 \\ 10 \end{bmatrix}\right) = \begin{bmatrix} 42 \\ 10 \end{bmatrix} \text{ and } (g \circ \varphi)\left(\begin{bmatrix} 42 \\ 11 \end{bmatrix}\right) = \begin{bmatrix} 42 \\ 11 \end{bmatrix}.$$

This first equation tells us

$$g\left(\varphi\left(\begin{bmatrix} 42 \\ 10 \end{bmatrix}\right)\right) = g\left(\begin{bmatrix} 42 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 42 \\ 10 \end{bmatrix},$$

and the second tells us that

$$g\left(\varphi\left(\begin{bmatrix} 42 \\ 11 \end{bmatrix}\right)\right) = g\left(\begin{bmatrix} 42 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 42 \\ 11 \end{bmatrix}.$$

Hang on. Those two vectors aren't equal. In order for g to be an inverse for φ, it has to map one vector to two different vectors. That is not how any respectable function behaves![11]

11: Here, you should take a respectable function to mean any function. Yes, we have just claimed all functions are respectable, even the Buffetting one from earlier.

The main idea here is that if you have a function that relates elements of A to elements of B, does there exist another function that unrelates them? More specifically, is there another function g that relates B to A so that if f relates $a_0 \in A$ to $b_0 \in B$, then g relates b_0 back to a_0. Sounds pretty easy, right? The part that makes this nontrivial is all that "for all" business. This property of "unrelating" has to hold for every element in A and every element in B. That is a very rigid condition! Fortunately, as with one-to-one and onto, there is an alternate characterization of invertibility.

In Example 3.1.6 we saw a function that failed to be invertible because it was not onto. In Example 3.1.7, we saw a function that failed to be invertible because it was not one-to-one. Perhaps then this theorem will not come as a surprise.[12]

12: Oh, this makes me sad. I love a surprise...

Theorem 3.1.3 *A function $f\colon A \to B$ is invertible if and only if f is both one-to-one and onto.*

Exploration 69 Let's start this proof together!

PROOF. First, suppose $f\colon A \to B$ is invertible. Then f^{-1} exists. Let's use f^{-1} to show that f is both one-to-one and onto.

One-to-one: Let $a, b \in A$ be such that $f(a) = f(b)$. We need to show $a = b$. To do this, let's consider the expression $f^{-1}(f(a))$. Since $f(a) = f(b)$, we have $f^{-1}(f(a)) = f^{-1}(f(b))$. Now, why does this tell us $a = b$?

Onto: Let $b \in B$. Then we need to find an element $a \in A$ that maps to b under f. Use f^{-1} to find a.

Now, we need to prove the other direction. Suppose that f is both one-to-one and onto. We need to define a function $g\colon B \to A$ so that f has an inverse. Since we know that f is onto B, we can write any element of B as $f(a)$ for some $a \in A$. Moreover, since f is one-to-one, we know for any $b \in B$ that there is exactly one $a \in A$ such that $b = f(a)$. Thus, we can define $g\colon B \to A$ to be the relation sending $f(a)$ to a. Now, we need to verify that this g is actually a function. Suppose $g(b) = a_1$ and $g(b) = a_2$. By the definition of g, this says $f(a_1) = b = f(a_2)$. Since f is one-to-one, we know $a_1 = a_2$, so g is a valid function. Note that we needed f to be both onto and one-to-one in order for g to be a function with domain B. We now just need to verify that g is the inverse of f. Our first condition is that for all $a \in A$, $(g \circ f)(a) = a$. We see this quickly since $g(\circ f)(a) = g(f(a)) = a$ by definition. Next we have that second condition: for all $b \in B$, $(f \circ g)(b) = b$. Let $b \in B$, then there is some $a \in A$ such that $f(a) = b$ since f is onto. So $f(g(b)) = f(g(f(a))) = f(a) = b$. Thus, the second condition is also satisfied by this function g because f is onto. □

Exploration 70 Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x + 2$. We saw in our previous examples that this function is both one-to-one and onto. So it's invertible! Can you find the inverse?

Recall our Buffeting function from earlier? Well, we can very quickly convince ourselves that this function is not one-to-one.[13] Thus, we have the following fun corollary to Theorem 3.1.3.

13: We can also convince ourselves that this function was ridiculous.

Corollary 3.1.4 *There is no function to undo the Buffeting.*

Functions Between Finite Sets

Something a little special happens when talking about the properties of one-to-one and onto when the function is between two finite sets.

Theorem 3.1.5 *Let A and B be finite sets and suppose $f\colon A \to B$ is a function. Let n denote the number of elements in A and m denote the number of elements in B.*

(a) *If $n > m$, then f is not one-to-one.*

(b) *If $n < m$, then f is not onto.*

(c) *If $n = m$, then f is either both one-to-one and onto or f is neither one-to-one nor onto.*

PROOF. First, suppose $n > m$. Then there are more elements in the domain of f than in its codomain. If f is one-to-one, then we would see each element in A map to a distinct element of B. However, there just aren't enough elements of B for this to happen!

Next, suppose $n < m$. Then this function has more elements in its codomain than its domain. In order for f to be onto, we must have $\operatorname{ran}(f) = \operatorname{codom}(f)$. However, to be a function, f must map each element of A to only one element of B. This means the largest $\operatorname{ran}(f)$ can be is n, so f cannot be onto.

Lastly, let's suppose $n = m$. If f is onto, then every element in B is mapped to by an element in A. Because these sets have the same size, the map is forced to be one-to-one since it can only send each element of A to one distinct element in B. Suppose now that f is one-to-one. Then it will send each element in A to a different element in B. Since these sets have the same size, we must use every element in B in this mapping, so f is also onto. We've just argued that f is one-to-one if and only if it is onto if $n = m$. This gives our result. □

Since almost all of our vector spaces are infinite sets,[14] you may be wondering why we bothered to tell you about Theorem 3.1.5. You'll see in a few sections

14: All except the trivial one, $\{\vec{0}\}$, in fact.

Nicky! What are *you* doing here?!

It looked like you needed help, so here I am. Helping.

that there is actually a wonderful analog to this theorem in the case of functions between vector spaces.

Section Highlights

- A function is one-to-one if every element in the range is mapped to exactly once. In terms of function input from the domain and output from the range, this means each output has exactly one input that maps to it. See Definition 3.1.2 and Theorem 3.1.2
- A function is onto if its range is the entirety of the codomain. This means every possible element in the codomain is an output mapped to by some input in the domain. See Definition 3.1.1 and Theorem 3.1.1.
- Function composition is a way to combine functions. See Definition 3.1.3.
- If a function is both one-to-one and onto, then it is invertible. This means that there exists an inverse function with which it composes to form the identity map. See Definition 3.1.4 and Theorem 3.1.3.

Exercises for Section 3.1

3.1.1. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + x_2.$$

(a) Is f onto? Prove or give a counterexample.

(b) Is f one-to-one? Prove or give a counterexample.

3.1.2. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + 3.$$

(a) Is f onto? Prove or give a counterexample.

(b) Is f one-to-one? Prove or give a counterexample.

3.1.3. For each function below, determine whether it is one-to-one.

(a) $f\colon \{a, b, c\} \to \{1, 2, 3, 4\}$ defined by $\{(a, 1), (b, 2), (c, 1)\}$

(b) $f\colon \{a, b, c\} \to \{1, 2, 3, 4\}$ defined by $\{(a, 1), (b, 2), (c, 4)\}$

(c) $f\colon \{a, b, c\} \to \{1, 2, 3\}$ defined by $\{(a, 1), (b, 2), (c, 3)\}$

(d) $f\colon \{a, b, c, d\} \to \{1, 2, 3\}$ defined by $\{(a, 1), (b, 2), (c, 3), (d, 2)\}$

(e) $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \left(\frac{x+1}{4}\right)$

(f) $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1$

(g) $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + 3x_2$

(h) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 \\ x_1x_2 \end{bmatrix}$

(i) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 \\ x_1 \end{bmatrix}$

(j) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + 3 \end{bmatrix}$

(k) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 + 3 \end{bmatrix}$

(l) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ 2x_1 + 2x_2 \end{bmatrix}$

(m) $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ 2x_1 + 2x_2 \\ x_2 \end{bmatrix}$

3.1.4.For each function below, determine whether it is onto.

(a) $f\colon \{a, b, c\} \to \{1, 2, 3, 4\}$ defined by $\{(a, 1), (b, 2), (c, 1)\}$

(b) $f\colon \{a, b, c\} \to \{1, 2, 3, 4\}$ defined by $\{(a, 1), (b, 2), (c, 4)\}$

(c) $f\colon \{a, b, c\} \to \{1, 2, 3\}$ defined by $\{(a, 1), (b, 2), (c, 3)\}$

(d) $f\colon \{a, b, c, d\} \to \{1, 2, 3\}$ defined by $\{(a, 1), (b, 2), (c, 3), (d, 2)\}$

(e) $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \left(\frac{x+1}{4}\right)$

(f) $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1$

(g) $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + 3x_2$

(h) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 \\ x_1 x_2 \end{bmatrix}$

(i) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 \\ x_1 \end{bmatrix}$

(j) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + 3 \end{bmatrix}$

(k) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 + 3 \end{bmatrix}$

(l) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ 2x_1 + 2x_2 \end{bmatrix}$

(m) $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ 2x_1 + 2x_2 \\ x_2 \end{bmatrix}$

3.1.5.Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x + 3$. Show that f is one-to-one and onto. Find f^{-1}.

3.1.6.Let $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_3 \\ 2x_1 + x_2 \\ x_2 \end{bmatrix}$. Show that f is both one-to-one and onto. Find f^{-1}.

3.1.7.A constant function maps every element in the domain to the same element in the codomain. Give an example of

(a) a domain, codomain, and constant function that is onto and not invertible;

(b) a domain, codomain, and constant function that is onto and invertible.

3.1.8.Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + x_2.$$

Does f have the property that $f(\vec{v} + \vec{u}) = f(\vec{v}) + f(\vec{u})$ for any $\vec{v}, \vec{u} \in \mathbb{R}^2$?

3.1.9.Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + 3.$$

Does f have the property that $f(\vec{v} + \vec{u}) = f(\vec{v}) + f(\vec{u})$ for any $\vec{v}, \vec{u} \in \mathbb{R}^2$?

3.1.10.Suppose $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ are both onto. Must $f + g$ be onto? Give an example to support your claim.

3.1.11.Let A, B, and C be sets and $f\colon A \to B$ and $g\colon B \to C$ be functions.

(a) Suppose $g \circ f$ is one-to-one. Then the function f is one-to-one. Let's see why. Suppose $a_1, a_2 \in A$ and $f(a_1) = f(a_2)$. We want to show $a_1 = a_2$, and we know we need to involve $g \circ f$ to do this. Let's consider $g \circ f(a_1) = g(f(a_1))$ and $g \circ f(a_2) = g(f(a_2))$. How are these related?

Now use the fact that $g \circ f$ is one-to-one to conclude $a_1 = a_2$.

(b) Suppose g is onto. It's not necessarily true that $g \circ f$ is onto. Can you come up with an example where g is onto but $g \circ f$ is not?

3.1.12.Let A, B, and C be sets and $f\colon A \to B$ and $g\colon B \to C$ be functions.

(a) Suppose f is one-to-one and g is onto.

(i) Is $g \circ f$ one-to-one? Prove or give a counterexample.

(ii) Is $g \circ f$ onto? Prove or give a counterexample.

(b) Suppose $g \circ f$ is onto and g is one-to-one.

(i) Is g onto? Prove or give a counterexample.

(ii) Is $g \circ f$ one-to-one? Prove or give a counterexample.

3.2 Linear Transformations

We've now talked quite a bit about functions, but very little about functions in relation[15] to vector spaces.[16] It's time to fix that. It's time to celebrate. It's time to finally formally familiarize ourselves with the functions that best fit this narrative we've been following. We're all here to hear about linear transformations on vector spaces. Well, we may never actually be done with amateur word play, but the long wait is over. Get ready. It's linear transformation time.

15: Pun not intended, but it's there anyway.

16: Well, except for those inner product ones, but that was ages ago now.

Respect the Operations

Let's think about what it means for a function to map one vector space to another by starting with an example.

Let $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$f\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = \begin{bmatrix} a+2 \\ b+2 \\ c+2 \end{bmatrix}.$$

This seems like a perfectly good function between vector spaces. Commence exploration!

Exploration 71 To get a sense for what this function does, let us experiment and see what happens with a few specific vectors.

- ▶ Compute $f\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right)$ and $f\left(\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}\right)$.

- ▶ Compute $2f\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = f\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) + f\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right)$.

From these computations, we see that for a vector $\vec{v} \in \mathbb{R}^3$, this function has

$$f(\vec{v}+\vec{v}) \neq f(\vec{v}) + f(\vec{v}) \qquad \text{and} \quad f(2\vec{v}) \neq 2f(\vec{v}).$$

While f is a perfectly good function between the *set* of vectors in $\mathbb{R}^3$ and the *set* of vectors in $\mathbb{R}^3$, it is not a useful function between the *vector spaces*. We have $\vec{v} \in \mathbb{R}^3$ and $f(\vec{v}) \in \mathbb{R}^3$; since $\mathbb{R}^3$ is closed under vector addition, we also have that

$$\vec{v}+\vec{v} \in \mathbb{R}^3 \quad \text{and} \quad f(\vec{v}) + f(\vec{v}) \in \mathbb{R}^3.$$

Since our function related $\vec{v}$ to $f(\vec{v})$, our function would, ideally, relate $\vec{v}+\vec{v}$ to $f(\vec{v}) + f(\vec{v})$; that is, it would be very nice if our function related our notion of vector addition in the domain to our notion of vector addition in the codomain. However, our function relates $\vec{v}+\vec{v}$ to $f(\vec{v}+\vec{v})$, and as we've already shown, $f(\vec{v}+\vec{v}) \neq f(\vec{v})+f(\vec{v})$. In order for our function to relate the vector space structure in the domain to the vector space structure in the codomain, it needs to behave well with respect to the vector space operations: vector addition and scalar multiplication.

There is a very useful class of functions for which this respect of operations is cleverly guaranteed.

Definition 3.2.1 *A function $f\colon V \to W$, where V and W are vector spaces, is called a* ***linear transformation*** *if for any vectors $\vec{v}, \vec{u} \in V$ and any scalar $a \in \mathbb{R}$,*

- $f(\vec{v}+\vec{u}) = f(\vec{v})+f(\vec{u})$ *and*
- $f(a\vec{v}) = af(\vec{v})$.

Linear transformations[17] resolve the issue of how vector addition and scalar multiplication are defined in the range of a function (function first or operation first) by saying it doesn't matter. For linear transformations, the two vectors are equal:

- $f(\vec{v}+\vec{u}) = f(\vec{v})+f(\vec{u})$; sum and then function, or function then sum; it's the same either way.
- $f(a\vec{v}) = af(\vec{v})$; scale and then function, or function then scale; it's the same either way.

17:  Yay! That's on the cover of the book!

Linear transformations are often said to "preserve" vector addition and scalar multiplication for this reason. If this seems overly restrictive, keep in mind that these are the only operations on V, and making sure those work correctly in the range is really the only restriction we've imposed. Speaking of the "range," you'll find this arrangement has its advantages:

Definition 3.2.2 *Let V and W be vector spaces and $f\colon V \to W$ be a linear transformation. The* ***image of*** *f is the set of vectors $\vec{w} \in W$ such that there is a vector $\vec{v} \in V$ with $\vec{w} = f(\vec{v})$. We shall use the notation*

$$\operatorname{Imag} f = \{\vec{w} \in W : \vec{w} = f(\vec{v}) \text{ for some } \vec{v} \in V\}.$$

You're probably asking yourself, "Isn't that just the range of f?" Yeah, it is. That's true. However, for any function $f : V \to W$ where V and W are just sets and $v \in V$, we often refer to $f(v)$ as the image of v. The idea is that the image of the *point* v is the *point* $f(v)$. Now, when V and W are vector spaces and f is a linear transformation, we extend this term to encompass all of the range. Thus, the image of a vector space V is the vector space $f(V)$ when f is a linear transformation. The word "image" is commonly used instead of "range" when a function is the type that preserves the algebraic structure of the domain.

Theorem 3.2.1 *Let V and W be vector spaces and $f\colon V \to W$ be a linear transformation. Then the image of f, $\operatorname{Imag}(f)$, is a subspace of W.*

Exploration 72 Let's prove this one together!

PROOF. We know that $\text{Imag } f$ is a subset of W, so we just need to verify closure for addition and scalar multiplication and also that it contains $\vec{0}$.

- Let $\vec{v}$ and $\vec{u}$ be vectors in V. Use the properties of the linear transformation f to show $f(\vec{v}) + f(\vec{u})$ is in $\text{Imag } f$. (To show a vector is in $\text{Imag } f$, find a way to write it as $f(\text{a vector in } V)$.)

- Let $\vec{v} \in V$ and $a \in \mathbb{R}$. Show $af(\vec{v}) \in \text{Imag } f$.

- The fact that $\vec{0} \in \text{Imag } f$ is actually related to an exercise from all the way back in Chapter 0. Linear transformations have the property that $f(a\vec{v}) = af(\vec{v})$ for any $a \in \mathbb{R}$ and any $\vec{v} \in V$. Thus, if $\vec{v} = \vec{0}$ and $a = 0$, we have $f(\vec{0}) = f(0\vec{0}) = 0f(\vec{0}) = \vec{0}$. Thus, $\vec{0} \in \text{Imag } f$ since any linear transformation maps $\vec{0}$ to $\vec{0}$.

□

Examples Abound

Example 3.2.1 Let's start with a fairly straightforward function. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + x_2$$

where $x_1, x_2 \in \mathbb{R}$. Let's verify that this is a linear transformation. We need the axioms in the definition to hold for *all* vectors, so we need general vectors. Let $\vec{x}, \vec{y} \in \mathbb{R}^2$. Then

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

for some $x_1, x_2, y_1, y_2 \in \mathbb{R}$. Then

$$\begin{aligned} f(\vec{x} + \vec{y}) &= f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = f\left(\begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}\right) \\ &= (x_1 + y_1) + (x_2 + y_2) = (x_1 + x_2) + (y_1 + y_2) \\ &= f(\vec{x}) + f(\vec{y}). \end{aligned}$$

Thus, f preserves vector addition. Suppose $a \in \mathbb{R}$. We will show now that f preserves scalar multiplication. We will use the same vector $\vec{x}$ as above.[18]

$$\begin{aligned} f(a\vec{x}) &= f\left(a\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = f\left(\begin{bmatrix} ax_1 \\ ax_2 \end{bmatrix}\right) \\ &= ax_1 + ax_2 = a(x_1 + x_2) = af(\vec{x}). \end{aligned}$$

Thus, f is a linear transformation.

18:  Now wait a minute. Weren't we using $\vec{v}$ and $\vec{u}$ for our vectors?

 Yeah, we were! What's with all these $\vec{x}$'s and $\vec{y}$'s?

 Settle down. Using x for domain elements is pretty standard. Now that we're dealing with functions, it makes sense to change our vector naming convention.

Example 3.2.2 Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be given by $T(\vec{x}) = \alpha\vec{x}$ for some fixed real number $\alpha > 1$. This function rescales vectors in $\mathbb{R}^2$ by a factor of α. To see this in action, let S be the following square in $\mathbb{R}^2$:

$$S = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \colon 0 \le x_1 \le 1 \text{ and } 0 \le x_2 \le 1 \right\}.$$

Figure 3.1 shows the image of S under the function T.
Let's verify that T is a linear transformation; we already know T is a function, so we just need to verify the two properties of a linear transformation. Let $\vec{x}, \vec{y} \in \mathbb{R}^2$ and $k \in \mathbb{R}$. Note that

$$\begin{aligned} T(\vec{x}+\vec{y}) &= \alpha(\vec{x}+\vec{y}) = \alpha\vec{x} + \alpha\vec{y} = T(\vec{x}) + T(\vec{y}) \text{ and} \\ T(k\vec{x}) &= \alpha(k\vec{x}) = (\alpha k)\vec{x} = (k\alpha)\vec{x} = k(\alpha\vec{x}) = kT(\vec{x}). \end{aligned}$$

It follows that T is a linear transformation.

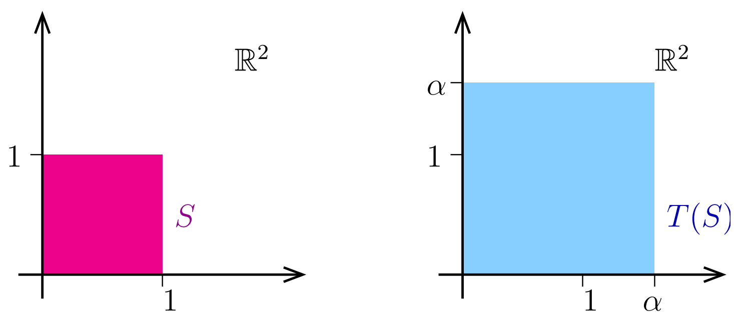


FIGURE 3.1. T is a linear transformation that rescales vectors by $\alpha > 1$.

Example 3.2.3 Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the function such that for any

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \quad \text{we define} \quad T(\vec{x}) = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ 4 \end{bmatrix}.$$

This perhaps appears more complicated, but all we're doing here is assigning vectors in $\mathbb{R}^3$ to a specific linear combination of vectors in $\mathbb{R}^2$ by using the components of our domain vectors as the weights of the linear combination. Not only does it turn out this is also a linear transformation, this also is an easy and convenient way to define a linear transformation. We should probably remember this example! Again, T is a function, so we just verify the two properties of a linear transformation. Let $\vec{x}, \vec{y} \in \mathbb{R}^2$, where

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{and} \quad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},$$

and $a \in \mathbb{R}$. This is gonna be a little gross. Note first that

$$\begin{aligned} T(\vec{x}+\vec{y}) &= T\left(\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ x_3+y_3 \end{bmatrix}\right) \\ &= (x_1+y_1)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + (x_2+y_2)\begin{bmatrix} 3 \\ 1 \end{bmatrix} + (x_3+y_3)\begin{bmatrix} 5 \\ 4 \end{bmatrix} \\ &= \left(x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} 5 \\ 4 \end{bmatrix}\right) \\ &\quad + \left(y_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y_2\begin{bmatrix} 3 \\ 1 \end{bmatrix} + y_3\begin{bmatrix} 5 \\ 4 \end{bmatrix}\right) \\ &= T(\vec{x}) + T(\vec{y}). \end{aligned}$$

Similarly,

$$\begin{aligned} T(a\vec{x}) = T\left(\begin{bmatrix} ax_1 \\ ax_2 \\ ax_3 \end{bmatrix}\right) &= ax_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + ax_2\begin{bmatrix} 3 \\ 1 \end{bmatrix} + ax_3\begin{bmatrix} 5 \\ 4 \end{bmatrix} \\ &= a\left(x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} 5 \\ 4 \end{bmatrix}\right) \\ &= aT(\vec{x}). \end{aligned}$$

Exploration 73 Recall Example 3.1.7 from Section 3.1. The function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ relates a vector $\vec{x} \in \mathbb{R}^2$ to the vector with the same first coordinate but 0 for the second coordinate. That is, for any vector in $\mathbb{R}^2$,

$$\varphi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}.$$

We showed this was a function that is not invertible already. Let's show that it's a linear transformation. We'll need some general vectors in $\mathbb{R}^2$ to start. Let $\vec{x}, \vec{y} \in \mathbb{R}^2$. Then

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

for some $x_1, x_2, y_1, y_2 \in \mathbb{R}$.

- First, verify that $\varphi(\vec{x}+\vec{y}) = \varphi(\vec{x}) + \varphi(\vec{y})$.

- Second, verify that $\varphi(a\vec{x}) = a\varphi(\vec{x})$ for any $a \in \mathbb{R}$.

Example 3.2.4 We've seen several examples of functions that are linear transformations. Let's see another one that is *not* a linear transformation.

Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1x_2$$

where $x_1, x_2 \in \mathbb{R}$. We can show this fails to be a linear transformation in multiple ways. First, we can show it fails to preserve vector addition.

$$f\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = f\left(\begin{bmatrix} 3 \\ 5 \end{bmatrix}\right) = (3)(5) = 15,$$

but

$$f\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) + f\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = (2)(3) + (1)(2) = 8.$$

Next, note that

$$f\left(5\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) = f\left(\begin{bmatrix} 10 \\ 15 \end{bmatrix}\right) = (10)(15) = 150,$$

but

$$5f\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) = 5(2)(3) = 30.$$

Thus, f also does not preserve scalar multiplication. Failing either condition is enough though to see that it is not a linear transformation.

Example 3.2.5 Let $T\colon \mathbb{P}_2 \to \mathbb{R}^2$ be defined by

$$T(\vec{p}) = \begin{bmatrix} p(1) \\ p(0) \end{bmatrix}.$$

Let's show T is a linear transformation; this one is particularly interesting because T has a completely different domain and codomain. Let $\vec{p} = ax^2 + bx + c$ and $\vec{q} = dx^2 + ex + f$ be arbitrary vectors in $\mathbb{P}_2$. Note that

$$T(\vec{p}) = \begin{bmatrix} a+b+c \\ c \end{bmatrix} \quad \text{and} \quad T(\vec{q}) = \begin{bmatrix} d+e+f \\ f \end{bmatrix}.$$

Checking vector addition, we have

$$\begin{aligned} T(\vec{p}+\vec{q}) &= T((ax^2+bx+c)+(dx^2+ex+f)) \\ &= T((a+d)x^2+(b+e)x+(c+f)) \\ &= \begin{bmatrix} (a+d)+(b+e)+(c+f) \\ c+f \end{bmatrix} \\ &= \begin{bmatrix} (a+b+c)+(d+e+f) \\ c+f \end{bmatrix} \\ &= \begin{bmatrix} a+b+c \\ c \end{bmatrix} + \begin{bmatrix} d+e+f \\ f \end{bmatrix} = T(\vec{p}) + T(\vec{q}). \end{aligned}$$

All that remains is to check scalar multiplication; let $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} T(\alpha\vec{p}) = T(\alpha(ax^2+bx+c)) &= T((\alpha a)x^2+(\alpha b)x+(\alpha c)) \\ &= \begin{bmatrix} \alpha a+\alpha b+\alpha c \\ \alpha c \end{bmatrix} \\ &= \begin{bmatrix} \alpha(a+b+c) \\ \alpha c \end{bmatrix} \\ &= \alpha \begin{bmatrix} a+b+c \\ c \end{bmatrix} = \alpha T(\vec{p}). \end{aligned}$$

Exploration 74 Let $T\colon \mathbb{P}_2 \to \mathbb{R}^2$ be defined by

$$T(\vec{p}) = \begin{bmatrix} p(2) \\ p(0) \end{bmatrix}.$$

Is T a linear transformation? Justify your response.

Some Noteworthy Examples

Until now, we've focused on examples to illustrate how we verify a function is a linear transformation. Now, we'll see some examples of common linear transformations that will be important for us in future sections.

Theorem 3.2.2 *Let V be a vector space with basis $\mathcal{B}$ and dimension n. The function $\varphi_{\mathcal{B}}\colon V \to \mathbb{R}^n$ that relates vectors in V to their coordinate vector relative to $\mathcal{B}$ in $\mathbb{R}^n$ is a linear transformation; that is, the function given by*

$$\varphi_{\mathcal{B}}(\vec{v}) = [\vec{v}]_{\mathcal{B}}$$

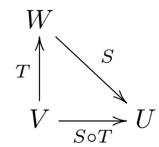
*is a linear transformation. This function is sometimes called the **coordinate mapping**.*

PROOF. In the proof of the Pythagorean Theorem in Section 2.3 we established that $[\vec{v}+\vec{u}]_{\mathcal{B}} = [\vec{v}]_{\mathcal{B}} + [\vec{u}]_{\mathcal{B}}$ for any $\vec{v}, \vec{u} \in V$. We just now need to see that $[\alpha\vec{v}]_{\mathcal{B}} = \alpha\,[\vec{v}]_{\mathcal{B}}$ for any scalar α and any $\vec{v} \in V$. However, this seems more like an exercise at this point. Expect to see this for homework. ☐

It turns out, quite conveniently, that compositions of linear transformations are also linear transformations.

Theorem 3.2.3 *Let V, W, and U be vector spaces, $T\colon V \to W$ a linear transformation, and $S\colon W \to U$ a linear transformation. Then the composition $S \circ T\colon V \to U$ is a linear transformation.*

This can be seen in the commuting diagram below:



PROOF. To show $S \circ T$ is a linear transformation, we need to show it respects vector addition and that it respects scalar multiplication. Let $\vec{v}_1$ and $\vec{v}_2$ be vectors in V, and let $a \in \mathbb{R}$. Then since S and T are both linear transformations,

$$S \circ T(\vec{v}_1+\vec{v}_2) = S(T(\vec{v}_1+\vec{v}_2)) = S(T(\vec{v}_1)+T(\vec{v}_2)) = S(T(\vec{v}_1))+S(T(\vec{v}_2)).$$

Similarly, we have

$$S \circ T(a\vec{v}_1) = S(T(a\vec{v}_1)) = S(aT(\vec{v}_1)) = aS(T(\vec{v}_1)).$$

Thus, the composition of two linear transformations is a linear transformation. □

While not all linear transformations are invertible, something interesting happens when they are. Inverses of linear transformations are also linear transformations.

Theorem 3.2.4 *Suppose V and W are vector spaces and $T: V \to W$ is a linear transformation. If T is invertible with inverse T^{-1}, then T^{-1} is a linear transformation.*

PROOF. Recall from Theorem 3.1.3 that T is invertible if and only if it is both one-to-one and onto. Also, from the definition of invertible, we know for any $\vec{v} \in V$ that $(T^{-1} \circ T)(\vec{v}) = \vec{v}$. Suppose then that $\vec{u}, \vec{w}$ are in W. Since T is onto, there exist some $\vec{x}, \vec{y} \in V$ such that $\vec{u} = T(\vec{x})$ and $\vec{w} = T(\vec{y})$. Because T^{-1} is the inverse of T, we must also have that $T^{-1}(\vec{u}) = \vec{x}$ and $T^{-1}(\vec{w}) = \vec{y}$. Thus

$$\begin{aligned} T^{-1}(\vec{u}+\vec{w}) &= T^{-1}(T(\vec{x})+T(\vec{w})) = T^{-1}(T(\vec{x}+\vec{y})) \\ &= (T^{-1} \circ T)(\vec{x}+\vec{y}) = \vec{x}+\vec{y} \\ &= T^{-1}(\vec{u}) + T^{-1}(\vec{w}), \end{aligned}$$

and we see T^{-1} preserves vector addition. Let $a \in \mathbb{R}$. Then we have

$$T^{-1}(a\vec{u}) = T^{-1}(aT(\vec{x})) = T^{-1}(T(a\vec{x})) = (T^{-1} \circ T)(a\vec{x}) = a\vec{x} = aT^{-1}(\vec{u}).$$

Thus, T^{-1} is a linear transformation because T is a linear transformation. □

We'll talk quite a bit more about these invertible linear transformations in the next section.

Linear Transformations and Bases

We spent quite a bit of time in Chapter 2 convincing you how great it is to have a basis, but we had to save one of the best things about them until now. Bases are extremely useful in helping us to understand linear transformations.

Theorem 3.2.5 *Suppose V is a vector space with a spanning set $\mathcal{P} = \{\vec{v}_1, \ldots, \vec{v}_k\}$. Any linear transformation $T\colon V \to W$ is determined by $T(\vec{v}_1), \ldots, T(\vec{v}_k)$. That is, for any $\vec{v} \in V$, we can realize $T(\vec{v})$ as a linear combination of the vectors $T(\vec{v}_1), \ldots, T(\vec{v}_k)$.*

PROOF. This theorem follows directly from the definitions of spanning set and linear transformation. Since $\mathcal{P}$ is a spanning set of V, we know there are coefficients $c_1, \ldots, c_k$ such that $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ for any $\vec{v} \in V$ and

$$T(\vec{v}) = T(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k) = c_1T(\vec{v}_1) + \cdots + c_kT(\vec{v}_k).$$

□

Corollary 3.2.6 *Suppose V is a vector space with a spanning set $\mathcal{P} = \{\vec{v}_1, \ldots, \vec{v}_k\}$. If $T\colon V \to W$ is a linear transformation, then* $\operatorname{Imag} T = \operatorname{Span}\{T(\vec{v}_1), \ldots, T(\vec{v}_k)\}$.

Corollary 3.2.7 *Suppose V is a vector space with a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$. Any linear transformation $T\colon V \to W$ is determined by $T(\vec{v}_1), \ldots, T(\vec{v}_n)$.*

Example 3.2.6 First, let's consider the vectors

$$\vec{v} = \begin{bmatrix} 1 \\ -4 \end{bmatrix} \quad \text{and} \quad \vec{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Since these are two linearly independent vectors in $\mathbb{R}^2$, we see that $\{\vec{v}, \vec{u}\}$ is a basis for $\mathbb{R}^2$. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that

$$T(\vec{v}) = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad T(\vec{u}) = \begin{bmatrix} -5 \\ 7 \end{bmatrix}.$$

First, let us find the images under T of $2\vec{v}$ and $4\vec{u}$. It seems pretty straightforward using our properties of a linear transformation.

$$\begin{aligned} T(2\vec{v}) &= 2T(\vec{v}) = 2\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}, \text{ and} \\ T(4\vec{u}) &= 4T(\vec{u}) = 4\begin{bmatrix} -5 \\ 7 \end{bmatrix} = \begin{bmatrix} -20 \\ 28 \end{bmatrix}. \end{aligned}$$

Similarly, we can also use linearity to find $2\vec{v} + 4\vec{u}$:

$$T(2\vec{v} + 4\vec{u}) = T(2\vec{v}) + T(4\vec{u}) = \begin{bmatrix} 6 \\ 2 \end{bmatrix} + \begin{bmatrix} -20 \\ 28 \end{bmatrix} = \begin{bmatrix} -14 \\ 30 \end{bmatrix}$$

We saw in Theorem 3.2.5 that the linear transformation is determined by what it does on a basis, so T is determined by $T(\vec{u})$ and $T(\vec{v})$.

Now, suppose we want to know what T maps $\vec{x} = \begin{bmatrix} 3 \\ 10 \end{bmatrix}$ to? We need to find the coordinate vector for $\vec{x}$! That is, we need to solve for $a, b \in \mathbb{R}$ such that $\vec{x} = a\vec{v} + b\vec{u}$. This is equivalent to

$$\begin{bmatrix} 3 \\ 10 \end{bmatrix} = a\begin{bmatrix} 1 \\ -4 \end{bmatrix} + b\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

which gives us the two equations $3 = a + 2b$ and $10 = -4a + 3b$. Solving these gives us $a = -1$ and $b = 2$. So $\vec{x} = -\vec{v} + 2\vec{u}$. Now we know how to find $T(\vec{x})$!

$$T(\vec{x}) = T(-\vec{v}+2\vec{u}) = -T(\vec{v})+2T(\vec{u}) = -\begin{bmatrix} 3 \\ 1 \end{bmatrix}+2\begin{bmatrix} -5 \\ 7 \end{bmatrix} = \begin{bmatrix} -13 \\ 13 \end{bmatrix}.$$

This is the procedure we could use to find where T sends any vector, but of course, don't expect the coordinate vector weights to always be quite so nice.

Respect the Kernel

We've seen that for a linear transformation $T\colon V \to W$, where V and W are vector spaces, $\mathrm{Imag}\ T$ is a subspace of the codomain, W. That's great for the codomain. Oh, to be guaranteed a subspace! How very nice, indeed. Perhaps we should try to do the same for the domain.

Exploration 75 Recall again the function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined in Example 3.1.7 by

$$\varphi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix},$$

which we saw was a linear transformation in Exploration 73. Give an example of a nonzero vector that maps to $\vec{0}$.
Can you describe all the vectors that map to $\vec{0}$?

Definition 3.2.3 *Let V and W be vector spaces and $T\colon V \to W$ be a linear transformation. The **kernel of** T is the set of vectors $\vec{v} \in V$ such that $T(\vec{v}) = \vec{0}$. We shall use the notation*

$$\mathrm{Ker}\ T = \{\vec{v} \in V : T(\vec{v}) = \vec{0}\}.$$

Theorem 3.2.8 *Let V and W be vector spaces and $T\colon V \to W$ be a linear transformation. Then the kernel of T is a subspace of V.*

Exploration 76 PROOF. Let V and W be vector spaces and $T\colon V \to W$ be a linear transformation. We know that $\mathrm{Ker}\ T$ is a subset of V, so we just need to verify closure for addition and scalar multiplication and also that it contains $\vec{0}$.

- Let $\vec{v}$ and $\vec{u}$ be vectors in $\mathrm{Ker}\ T$, so $T(\vec{v}) = \vec{0}$ and $T(\vec{u}) = \vec{0}$. Show $T(\vec{v} + \vec{u}) = \vec{0}$ so that $\vec{v} + \vec{u}$ is in $\mathrm{Ker}\ T$.

- Let $\vec{v} \in \mathrm{Ker}\ T$ and $a \in \mathbb{R}$. Then $T(\vec{v}) = \vec{0}$. Show $T(a\vec{v}) = \vec{0}$ so that $a\vec{v} \in \mathrm{Ker}\ T$.

- In our proof that $\text{Imag } T$ is a subspace, we established that a linear transformation always maps $\vec{0}$ to $\vec{0}$. Thus, $\vec{0} \in \text{Ker } T$.

□

Hooray! A linear transformations guarantees a subspace in the domain *and* another one in the codomain, each of which will be useful for us.

Example 3.2.7 Let's look at some of the linear transformations we've seen already and see what their kernels are.

- Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be given by $T(\vec{x}) = \alpha\vec{x}$ for some fixed real number $\alpha > 1$. For this linear transformation, we only have $\alpha\vec{x} = \vec{0}$ if $\vec{x} = \vec{0}$ since $\alpha \neq 0$. The kernel is then just the zero vector space.
- Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the function such that for any

$$\begin{aligned} \vec{x} &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \quad \text{we define} \\ T(\vec{x}) &= x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ 4 \end{bmatrix}. \end{aligned}$$

Let's determine the kernel for this one as well. This means we need to solve

$$\vec{0} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ 4 \end{bmatrix}. \tag{3.1}$$

Since we know any set of three vectors in $\mathbb{R}^2$ must be linearly dependent, we know there must be nontrivial solutions to this equation. Note that

$$-7 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}.$$

Thus we can rewrite our Equation 3.1 as

$$\begin{aligned} \vec{0} &= x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3 \left(-7 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right) \\ &= (x_1 - 7x_3) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (x_2 + 4x_3) \begin{bmatrix} 3 \\ 1 \end{bmatrix} \end{aligned}$$

The vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ are linearly independent, so the only solution to this is given by

$$x_1 - 7x_3 = 0 \quad \text{and} \quad x_2 + 4x_3 = 0 \tag{3.2}$$

We can rearrange these to get $x_1 = 7x_3$ and $x_2 = -4x_3$. Thus,

$$\text{Ker } T = \left\{ \begin{bmatrix} 7x_3 \\ -4x_3 \\ x_3 \end{bmatrix} : x_3 \in \mathbb{R} \right\} = \text{Span} \left\{ \begin{bmatrix} 7 \\ -4 \\ 1 \end{bmatrix} \right\}.$$

Exploration 77 Define $T\colon \mathbb{P}_2 \to \mathbb{R}^2$ by

$$T(\vec{p}) = \begin{bmatrix} p(1) \\ p(0) \end{bmatrix}.$$

What must $\mathrm{Ker}\ T$ be?

Now that we've spent some time talking about these special functions called linear transformations, we should spend some time talking about how they relate to the concepts of Section 3.1. We'll do that in the next section because this seems like enough for now.[19]

19: It is. It is enough.

Section Highlights

- A linear transformation is a function between vector spaces that preserves vector space structure. In particular, the range of a linear transformation, called its image, is a vector space. See the discussion before Definition 3.2.1 and Theorem 3.2.1.
- A function between vector spaces is a linear transformation if it preserves the operations of vector addition and scalar multiplication. This means that for $f\colon V \to W$ to be a linear transformation, it must satisfy $f(\vec{x}+\vec{y}) = f(\vec{x})+f(\vec{y})$ and $f(a\vec{x}) = af(\vec{x})$ where the operations on the left of the equals are in V and the operations on the right are in W for any vectors $\vec{x}$ and $\vec{y}$ in V. See Definition 3.2.1.
- The kernel of a linear transformation, T, is the subspace in the domain consisting of the vectors mapped to the zero vector by T. See Definition 3.2.3 and Theorem 3.2.8. See Example 3.2.7 on how to compute the kernel.

Exercises for Section 3.2

3.2.1. Consider the function $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}.$$

Show that T is a linear transformation.

3.2.2. Let $T\colon \mathbb{R}^5 \to \mathbb{R}^5$ be defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}\right) = \begin{bmatrix} 5x_1 + 4 \\ x_2 + x_3 \\ x_3 - 1 \\ x_1 \\ x_2 \end{bmatrix}.$$

Show that T is *not* a linear transformation.

3.2.3. All of the functions between vector spaces below fail to be linear transformations. Give specific examples illustrating why they fail.

(a) $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + 5$

(b) $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 x_2$

(c) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 \\ x_1 x_2 \end{bmatrix}$

(d) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + 3 \end{bmatrix}$

(e) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1^3 \end{bmatrix}$

3.2.4. Consider the map $T\colon \mathbb{R}^4 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + x_3 + x_4 \\ x_1 + x_2 - x_3 - x_4 \end{bmatrix}.$$

Show this is a linear transformation.

3.2.5. Let $T\colon \mathbb{R}^5 \to \mathbb{R}^5$ be defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}\right) = \begin{bmatrix} 5x_1 \\ x_2 + x_3 \\ x_3 \\ x_1 \\ x_2 \end{bmatrix}.$$

(a) Show that T is a linear transformation.

(b) Find $\operatorname{Ker} T$.

3.2.6. Consider the function $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ 0 \end{bmatrix}.$$

(a) Show that T is a linear transformation.

(b) Find $\operatorname{Ker} T$.

3.2.7. Consider the function $T\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + 3x_2.$$

(a) Show that T is a linear transformation.

(b) Find $\operatorname{Ker} T$.

3.2.8. Let $T\colon \mathbb{P}_2 \to \mathbb{R}^4$ be defined by

$$T\left(a_0 + a_1 x + a_2 x^2\right) = \begin{bmatrix} a_0 + a_1 \\ a_0 - a_1 \\ a_1 - a_2 \\ a_1 + a_2 \end{bmatrix}.$$

(a) Show T is a linear transformation.

(b) Find $\operatorname{Ker} T$.

3.2.9. Let $T\colon \mathbb{P}_2 \to \mathbb{R}^4$ be defined by

$$T\left(a_0 + a_1 x + a_2 x^2\right) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ (a_2)^2 \end{bmatrix}.$$

Show that T is *not* a linear transformation.

3.2.10. Let $\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$ is a basis for $\mathbb{R}^2$. Suppose $T : \mathbb{R}^2 \to \mathbb{P}_2$ is the linear transformation determined by

$$T(\vec{v}_1) = 1 + x \quad \text{and} \quad T(\vec{v}_2) = x + x^2.$$

(a) Find $T(\vec{v}_1 + 3\vec{v}_2)$.

(b) Suppose $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Find $T(\vec{x})$.

(c) Find $T(\vec{e}_2)$.

(d) Find $T\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix}\right)$.

3.2.11. Suppose $T : \mathbb{P}_2 \to \mathbb{R}^2$ is the linear transformation determined by

$$T(1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad T(1+x) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad T(1+x^2) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(a) Find $T(1 + x + x^2)$.

(b) Find $T(x)$.

(c) Find $T(x^2)$.

3.2.12. Show that any nonzero constant function between vector spaces is not a linear transformation.

3.2.13. Let V be a vector space with basis $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$. Complete the proof from Theorem 3.2.2 that the coordinate mapping is a linear transformation by showing $[\alpha\vec{v}]_{\mathcal{B}} = \alpha [\vec{v}]_{\mathcal{B}}$ for any real number α and any $\vec{v} \in V$.

3.2.14. Let V be a vector space such that $\dim V = 4$; let $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} \subset V$ be linearly independent and $W = \operatorname{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Show that the function $f\colon V \to V$ that relates $\vec{v} \in V$ to $\operatorname{proj}_W(\vec{v})$ is a linear transformation.

3.3 One-to-one and Onto Linear Transformations

In Section 3.1, we learned about when a function is one-to-one, onto, and invertible. Let's revisit all of these concepts in the more specific context of linear transformations. We'll begin with one-to-one.

Example 3.3.1 Consider the function $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}.$$

This is a linear transformation.[20] Moreover, it is one-to-one. To show this from the definition, suppose $T(\vec{v}) = T(\vec{u})$ for some $\vec{v}, \vec{u} \in \mathbb{R}^2$. Then, we know that if

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{and} \quad \vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

then

$$T(\vec{v}) = \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix} \quad \text{and} \quad T(\vec{u}) = \begin{bmatrix} u_1 \\ u_2 \\ 0 \end{bmatrix}.$$

Thus $T(\vec{v}) = T(\vec{u})$ means

$$\begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ 0 \end{bmatrix}.$$

From this we see that $v_1 = u_1$ and $v_2 = u_2$, so $\vec{v} = \vec{u}$; this tells us, by definition, that T is one-to-one.

Now, let's consider Ker T. If $\vec{x} \in$ Ker T, then $T(\vec{x}) = \vec{0}$. Then

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Thus, the only possibility for $\vec{x}$ is the vector $\vec{0} \in \mathbb{R}^2$. In fact, this is always the kernel of a one-to-one function, and this can even be used to tell whether a function is one-to-one.

20: Does this look familiar? Could this have been an exercise in the last section?

This example illustrates a convenient fact that is true in general.

Theorem 3.3.1 *Let V and W be vector spaces. A linear transformation $T\colon V \to W$ is one-to-one if and only if* Ker $T = \{\vec{0}\}$.

PROOF. Suppose first that the linear transformation $T\colon V \to W$ is one-to-one. Then there is a unique element of V which maps to $\vec{0}$ in W. Since we know $T(\vec{0}) = \vec{0}$ whenever T is a linear transformation, Ker $T = \{\vec{0}\}$.

Now suppose Ker $T = \{\vec{0}\}$. Let $\vec{v}_1$ and $\vec{v}_2$ be vectors in V such that $T(\vec{v}_1) = T(\vec{v}_2)$. We need to show that $\vec{v}_1 = \vec{v}_2$ in order for T to be one-to-one. Since

T is a linear transformation, we have

$$\begin{aligned} T(\vec{v}_1) =& T(\vec{v}_2) \\ T(\vec{v}_1) - T(\vec{v}_2) =& \vec{0} \\ T(\vec{v}_1 - \vec{v}_2) =& \vec{0}. \end{aligned}$$

Thus, $\vec{v}_1 - \vec{v}_2 \in \text{Ker } T = \{\vec{0}\}$, so $\vec{v}_1 - \vec{v}_2 = \vec{0}$. We then see $\vec{v}_1 = \vec{v}_2$, as desired. ☐

Exploration 78 Consider our favorite linear transformation, $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$, defined by

$$\varphi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}.$$

Is this function one-to-one?

Now we should consider what it means for a linear transformation to be onto. In particular, Theorem 3.1.1 from Section 3.1 can be restated for linear transformations.

Theorem 3.3.2 *Let V and W be vector spaces. A linear transformation $T\colon V \to W$ is onto if and only if* $\text{Imag } T = W$.

Example 3.3.2 Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation such that for any

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \quad \text{we define} \quad T(\vec{x}) = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ 4 \end{bmatrix}.$$

We know this is not one-to-one since $\text{Ker } T \neq \{\vec{0}\}$.[21] Let's see if it is onto, though. Based on the definition of the function, we know x_1, x_2, and x_3 can be any real numbers. Thus,

$$\text{Imag } T = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}\right\}.$$

Since this is a subspace of dimension 2 in $\mathbb{R}^2$, we see that it is all of $\mathbb{R}^2$. Thus, T is onto.

21:  This was done in the previous section.

Exploration 79 Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 + x_2$$

where $x_1, x_2 \in \mathbb{R}$. We showed in the previous section that this was a linear transformation. Now, find $\text{Ker } f$ and $\text{Imag } f$ to determine whether it is one-to-one, onto, both, or neither.

Isomorphisms

Now that we have discussed one-to-one linear transformations and onto linear transformations, we should talk about when a linear transformation has *both* of these properties. In Section 3.1, we called such functions invertible. We have a special name for an invertible linear transformation.[22]

22:  No, the name is not "Ricky," but maybe it should be...

Definition 3.3.1 *Let V and W be vector spaces. A linear transformation $T: V \to W$ is called an* ***isomorphism*** *if it is both one-to-one and onto. When such a linear transformation exists, we say V and W are* ***isomorphic*** *vector spaces, and denote this by $V \cong W$.*

The notation here of $V \cong W$ suggests that there is a sense that V and W are "equal" if they are isomorphic. This is essentially true, at least as vector spaces. A one-to-one and onto map between sets means they are in many ways interchangeable, and the fact here that such a map preserves the vector space structure means that they have identical structure as vector spaces. However, to claim that isomorphic is a form of equality, there are some properties of equality that should be satisfied. [23]

23:  The properties outlined here are those needed to form an *equivalence relation* on a set. What we've actually saying here is that being isomorphic gives an equivalence relation on the set of all vector spaces.

Theorem 3.3.3 (a) *If V is any vector space, then $V \cong V$.*
(b) *If V and W are vector spaces such that $V \cong W$, then $W \cong V$.*
(c) *If V, W, and U are vector spaces such that $V \cong W$ and $W \cong U$, then $V \cong U$.*

These properties might just come in handy later. Rather than prove these all here, we'll do it in the Appendix.

You might recall that the concept of vector spaces being essentially the same has come up for us before. In Section 1.2, we equated the appropriate vector space of arrow vectors with each of the vector spaces $\mathbb{R}^1$, $\mathbb{R}^2$, and $\mathbb{R}^3$. In truth, these are *isomorphic* vector spaces. The elements in these sets and the ways that they are described are very different, except they are really the same *as vector spaces*; that is, they are *isomorphic*. We also talked in Section 1.4 about how any plane through the origin in $\mathbb{R}^3$ "looks like" $\mathbb{R}^2$. In truth, each of these is isomorphic to $\mathbb{R}^2$.

Here's an important one; the coordinate mapping is *always* an isomorphism.

Theorem 3.3.4 *Let V be a vector space of dimension n with basis $\mathcal{B}$. The coordinate mapping $\varphi_{\mathcal{B}}: V \to \mathbb{R}^n$ defined by $\varphi_{\mathcal{B}}(\vec{v}) = [\vec{v}]_{\mathcal{B}}$ is an isomorphism.*

PROOF. From Theorem 3.2.2, we know $\varphi_{\mathcal{B}}$ is a linear transformation. We need now to show that it is one-to-one and onto. Since Theorem 2.3.1 tells us that each element in V is represented uniquely on the basis $\mathcal{B}$, we know $\varphi_{\mathcal{B}}$ is one-to-one. By Theorem 3.2.1, $\varphi_{\mathcal{B}}$ is onto if and only if $\text{Imag}\ \varphi_{\mathcal{B}} = \mathbb{R}^n$. By the definition of $\varphi_{\mathcal{B}}$, we already have that $\text{Imag}\ \varphi_{\mathcal{B}} \subseteq \mathbb{R}^n$. Since $\text{Span}\,\{\mathcal{B}\} = V$, every possible linear combination of the elements of $\mathcal{B}$ must be in V. Thus, every vector in $\mathbb{R}^n$ appears in $\text{Imag}\ \varphi_{\mathcal{B}}$. It follows that $\text{Imag}\ \varphi_{\mathcal{B}} = \mathbb{R}^n$. □

Let's think about what this theorem tells us for a moment. Suppose V is a vector space of dimension n. Then the coordinate mapping gives us an isomorphism between V and $\mathbb{R}^n$. This is actually incredibly important; it tells us that every vector space of dimension n is in essence "the same" as $\mathbb{R}^n$. This shouldn't make you think other vector spaces are not important, but it means a thorough understanding of how things work in $\mathbb{R}^n$ can be useful for predicting how they work in other real vector spaces.

There's something else that should be apparent from this; the dimensions of isomorphic vector spaces match. Let's formalize this.

Theorem 3.3.5 *Let V and W be vector spaces, and suppose $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis of V. If $T\colon V \to W$ is an isomorphism, then $\widehat{\mathcal{B}} = \{T(\vec{v}_1), \ldots, T(\vec{v}_n)\}$ is a basis of W.*

We'll walk through the proof of this fact in the exercises. The above theorem tells us that isomorphic vector spaces have bases of the same size. Thus, we have this useful corollary.

Corollary 3.3.6 *Two real vector spaces are isomorphic if and only if they have the same dimension.*

PROOF. We see from Theorem 3.3.5 that two isomorphic vector spaces will have the same dimensions since they have the same size bases. The other direction, where we begin by assuming two vector spaces have the same dimension, is given to us from Theorem 3.3.4 and Theorem 3.3.3 since they would both be isomorphic to $\mathbb{R}^n$ for the same n. □

We mentioned above that any plane through the origin in $\mathbb{R}^n$ is actually isomorphic to $\mathbb{R}^2$, but what's the isomorphism? Let's see an explicit example of this.

Example 3.3.3 Depending on your background, you may have seen planes in $\mathbb{R}^3$ described differently than the span of two vectors. For instance, the solutions to the equation $3x+2y-z=0$ in $\mathbb{R}^3$ form a plane. Let's translate that plane into more of our language.

$$3x + 2y - z = 0 \quad \text{says} \quad z = 3x + 2y.$$

Thus, the plane is all vectors in the set

$$\left\{ \begin{bmatrix} x \\ y \\ 3x+2y \end{bmatrix} : x, y \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} y \colon x, y \in \mathbb{R} \right\}$$
$$= \operatorname{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

Since this is now a vector space of dimension 2, we know it must be isomorphic to $\mathbb{R}^2$ and a specific isomorphism here can be given by

$$T\colon \text{Span}\left\{\begin{bmatrix}1\\0\\3\end{bmatrix}, \begin{bmatrix}0\\1\\2\end{bmatrix}\right\} \to \mathbb{R}^2 \quad \text{defined by}$$

$$T\left(\begin{bmatrix}1\\0\\3\end{bmatrix}\right) = \begin{bmatrix}1\\0\end{bmatrix} \quad \text{and} \quad T\left(\begin{bmatrix}0\\1\\2\end{bmatrix}\right) = \begin{bmatrix}0\\1\end{bmatrix}.$$

What if the plane we want to consider does not go through the origin? Well, it is then not a subspace of $\mathbb{R}^3$, but there is a way to define our operations of vector addition and scalar multiplication so that it is a vector space. For an example of such operations, see the Appendix. This vector space would again be isomorphic to $\mathbb{R}^2$.

Let's take note of something from that previous example. Once we had a basis for our plane in $\mathbb{R}^3$, we defined our isomorphism by mapping each basis vector in the space to one of the standard basis vectors in $\mathbb{R}^2$. This is really the easiest way to define an isomorphism.

Theorem 3.3.7 *Suppose V and W are vector spaces of the same dimension. Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be any basis for V. A linear transformation $T\colon V \to W$ is an isomorphism if and only if T maps each basis vector in $\{\vec{v}_1, \ldots, \vec{v}_n\}$ to a distinct vector in a basis $\{\vec{w}_1, \ldots, \vec{w}_n\}$ of W.*

PROOF. This is just the combination of Theorem 3.3.5 and Theorem 3.2.5. □

Exploration 80 Let $V = \text{Span}\left\{1+x, x^2\right\}$ in $\mathbb{P}_2$. Just like we did in Example 3.3.3, define an isomorphism T from V to $\mathbb{R}^2$. Say explicitly what $T(1+x)$ and $T(x^2)$ are.

Exploration 81 Again let $V = \text{Span}\left\{1+x, x^2\right\}$ in $\mathbb{P}_2$. Also, let $W = \text{Span}\{1, x\}$ in $\mathbb{P}_2$. We know that these vector spaces both have dimension 2, so they should be isomorphic. Use the given bases to give an isomorphism T from V to W. Say explicitly what $T(1+x)$ and $T(x^2)$ are.

Exploration 82 Can you find an example of a function between vector spaces that is both one-to-one and onto but is *not* an isomorphism? Hint: There's one somewhere in the previous section.

Rank-Nullity Theorem

There is a theorem known as the *Rank-Nullity Theorem* that we should probably talk about now. We name it this because that's what all the cool textbooks call it. For now, you should just assume the theorem is due to amateur mathematicians named Ronnie Rank and Noether Nullity.[24] This is, of course, a lie,[25] but the mathematical definitions of the words rank and nullity will come later.

Theorem 3.3.8 *Let V and W be inner product spaces and $T\colon V \to W$ be a linear transformation between them. Then*

$$(\text{Ker } T)^{\perp} \cong \text{Imag } T.$$

Corollary 3.3.9 *Let V and W be inner product spaces with $\dim V = \dim W$ and $T\colon V \to W$ be a linear transformation between them. Then*

$$\begin{aligned} (\text{Ker } T)^{\perp} &\cong \text{Imag } T \text{ and} \\ \text{Ker } T &\cong (\text{Imag } T)^{\perp}. \end{aligned}$$

Again, the various subspaces to which we will refer in the proof can be seen in Figure 3.2.

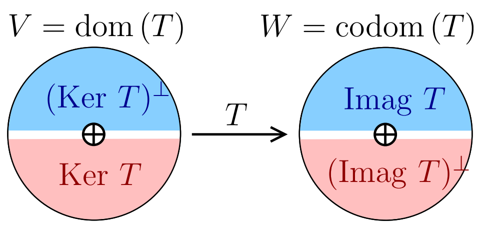


FIGURE 3.2. The linear transformation $T\colon V \to W$ generates four subspaces: Ker T and $(\text{Ker } T)^{\perp}$ in the domain, V, and Imag T and $(\text{Imag } T)^{\perp}$ in the codomain, W. The top two subspaces are isomorphic., and when $\dim V = \dim W$, the bottom subspaces are also isomorphic.

PROOF OF THEOREM 3.3.8. This proof has several elements. We'll use bullets to keep track.

- We know from Theorem 3.2.8 that Ker T is a subspace of V. Thus, from Corollary 2.5.2, we know $V = \text{Ker } T \oplus (\text{Ker } T)^{\perp}$. So every element of V can be written as $\vec{n} + \vec{v}$ for some $\vec{n} \in \text{Ker } T$ and some $\vec{v} \in (\text{Ker } T)^{\perp}$.
- To show these vector spaces are isomorphic, we need an isomorphism. We'll use T restricted to $(\text{Ker } T)^{\perp}$, and we'll call this map $\overline{T}$ to keep it straight. To be clear,

$$\overline{T}\colon (\text{Ker } T)^{\perp} \longrightarrow \text{Imag } T$$

is defined by

$$\overline{T}(\vec{x}) = T(\vec{x})$$

24:  They were also full-time mimes; this was the gig that paid the bills.

25:  The bit about who proved the Rank-Nullity Theorem was a lie. We, however, choose to believe that in an infinite universe, there exist both Ronnie Rank and Noether Nullity, full-time mimes. Oh yeah, they're unicorns, too.

for any $\vec{x} \in (\text{Ker } T)^\perp$. This is a function since T is a function, and it is a linear transformation since T is a linear transformation. We need to show now that it is an isomorphism. More specifically, we need to show this map is one-to-one and onto.

- To show $\overline{T}$ is onto, let $\vec{y} \in \text{Imag } T$. Then there is some $\vec{u} \in V$ such that $T(\vec{u}) = \vec{y}$. We know from above that

$$\vec{u} = \vec{n}_u + \vec{v}_u$$

for some $\vec{n}_u \in \text{Ker } T$ and some $\vec{v}_u \in (\text{Ker } T)^\perp$. Thus,

$$\vec{y} = T(\vec{u}) = T(\vec{n}_u + \vec{v}_u) = T(\vec{n}_u) + T(\vec{v}_u) = \vec{0} + T(\vec{v}_u) = \overline{T}(\vec{v}_u).$$

This tells us that $\overline{T}$ maps onto $\text{Imag } T$.

- We need lastly to show that $\overline{T}$ is one-to-one. We could do this from the definition, but Theorem 3.3.1 says we need only establish that $\text{Ker } \overline{T} = \{\vec{0}\}$. Suppose $\vec{x} \in \text{Ker } \overline{T}$. Then $\overline{T}(\vec{x}) = T(\vec{x}) = \vec{0}$ and $\vec{x} \in \text{Ker } T$. Since $\text{Ker } T \cap (\text{Ker } T)^\perp = \{\vec{0}\}$, this says $\vec{x} = \vec{0}$ and therefore $\overline{T}$ is one-to-one.

□

The following is more of a corollary to the above result, but all the other textbooks give it a fancy-sounding name. So we'll make it a theorem. No, it's not just because all the other books are doing it... No, we wouldn't jump off a bridge if all the other books did. Look, it's just that we wouldn't want to miss an opportunity to sound really fancy! Whatever. Just call it Theorem 3.3.10 if you want. Can we just drop it now?[26]

Theorem 3.3.10 (Rank-Nullity Theorem) *Let V and W be vector spaces and $T\colon V \to W$ be a linear transformation between them. Then* $\dim V = \dim \text{Ker } T + \dim \text{Imag } T$.

PROOF. We know from Corollary 3.3.6 that $(\text{Ker } T)^\perp$ and $\text{Imag } T$ have the same dimension. From Corollary 2.5.2 of the Orthogonal Decomposition theorem and Theorem 2.5.3, we have that

$$\dim V = \dim \text{Ker } T + \dim (\text{Ker } T)^\perp = \dim \text{Ker } T + \dim \text{Imag } T.$$

□

Let's talk about those words, *rank* and *nullity*.

Definition 3.3.2 *The **rank** of a linear transformation is the dimension of its image. The **nullity** of a linear transformation is the dimension of its kernel.*

Thus, the Rank-Nullity Theorem is aptly named.[27] Let's see how the theorem can be useful.

26:  Seriously though, you should know the name; we expect you to want to talk about Linear Algebra with many people throughout your lifetime, and sadly, many have not been fortunate enough to learn from this text. Therefore, you need to know what everyone else means when they refer to the Rank-Nullity Theorem.

27:  It was originally named (gesticulate wildly with your front hooves) by Ronnie and Noether.

Example 3.3.4 Let $T\colon \mathbb{R}^5 \to \mathbb{R}^5$ be defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}\right) = \begin{bmatrix} 5x_1 \\ x_2 + x_3 \\ x_3 \\ x_1 \\ x_2 \end{bmatrix}.$$

Since x_1, x_2, x_3 can be any real numbers, we get that

$$\begin{aligned} \text{Imag } T &= \left\{ \begin{bmatrix} 5x_1 \\ x_2 + x_3 \\ x_3 \\ x_1 \\ x_2 \end{bmatrix} : x_1, x_2, x_3 \in \mathbb{R} \right\} \\ &= \text{Span}\left\{ \begin{bmatrix} 5 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\} \end{aligned}$$

We can see then that $\dim \text{Imag } T = 3$. We can then conclude $\text{Ker } T$ has dimension 2. This helps us to find $\text{Ker } T$ because if we can find two linearly independent vectors in the kernel, we know they form a basis of the kernel.

$$\text{Ker } T = \text{Span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Exploration 83 Consider the linear transformation $T\colon \mathbb{R}^4 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + x_3 + x_4 \\ x_1 + x_2 - x_3 - x_4 \end{bmatrix}.$$

Convince yourself that this map is onto by identifying vectors in $\mathbb{R}^4$ that map to the basis

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

for $\mathbb{R}^2$.

Now, the dimension of $\text{Ker } T$ must be 2 from the Rank-Nullity Theorem. Find two linearly independent vectors in the kernel.

Another Useful Theorem

At the end of Section 3.1, we proved Theorem 3.1.5 about the properties of one-to-one and onto when a function is defined between two finite sets. Then in Theorem 3.2.5, we saw that linear transformations are completely determined by what they do on a spanning set. This result implies linear transformations between finite dimensional vector spaces are similar in some way to functions between finite sets, and we can now prove a linear transformation version of Theorem 3.1.5.

Theorem 3.3.11 *Suppose $T\colon V \to W$ is a linear transformation between finite dimensional vector spaces V and W.*

(a) *If $\dim V > \dim W$, then T is not one-to-one.*

(b) *If $\dim V < \dim W$, then T is not onto.*

(c) *If $\dim V = \dim W$, then either T is both one-to-one and onto or T is neither one-to-one nor onto.*

PROOF. Suppose first that $\dim V > \dim W$. Since we know $\operatorname{Imag} T$ is a subspace of W, we know $\dim \operatorname{Imag} T \leq \dim W$ from Theorem 2.2.4. Together, this says $\dim \operatorname{Imag} T < \dim V$. Then rearranging the Rank-Nullity Theorem tells us that $\dim \operatorname{Ker} T = \dim V - \dim \operatorname{Imag} T$ and $\dim \operatorname{Ker} T$ must be nonzero since the inequality $\dim \operatorname{Imag} T < \dim V$ is strict. Theorem 3.3.1 says T then is not one-to-one.

Suppose next that $\dim V < \dim W$. Again, we turn to the Rank-Nullity Theorem to see that $\dim \operatorname{Imag} T = \dim V - \dim \operatorname{Ker} T$. From this we see that $\dim \operatorname{Imag} T \leq \dim V$ and we can combine the inequalities to see that $\dim \operatorname{Imag} T < \dim W$. We know from Theorem 3.3.2 that T is not onto then since $\operatorname{Imag} T = W$ would mean they have the same dimension.

Suppose lastly that $\dim V = \dim W$. If T is onto, then we know $\dim W = \dim \operatorname{Imag} T$. Then the Rank-Nullity Theorem says

$$\dim W = \dim V = \dim \operatorname{Imag} T + \dim \operatorname{Ker} T = \dim W + \dim \operatorname{Ker} T.$$

This can only be true if $\dim \operatorname{Ker} T = 0$. In that case, T is one-to-one by Theorem 3.3.1. Now, if we begin by assuming that T is one-to-one, we know $\dim \operatorname{Ker} T = 0$ and the Rank-Nullity Theorem says $\dim W = \dim V = \dim \operatorname{Imag} T$. This means $W = \operatorname{Imag} T$ and T is onto. We have argued that T is onto if and only if it is one-to-one. The result follows. □

Section Highlights

- A linear transformation is one-to-one if and only if its kernel is exactly the zero vector. See Theorem 3.3.1.
- For a linear transformation, T, the Rank-Nullity Theorem tells us that $\dim \operatorname{dom}(T) = \dim \operatorname{Ker} T + \dim \operatorname{Imag} T$. See Theorem 3.3.10.
- For any linear transformation T, $\dim \operatorname{Ker} T$ can be used to determine whether the function is one-to-one, onto, both, or neither. In particular:

- T is one-to-one if and only if $\dim \mathrm{Ker}\ T = 0$;
- T is onto if and only if $\dim \mathrm{Imag}\ T = \dim \mathrm{codom}\,(T)$, and by the Rank-Nullity Theorem, $\dim \mathrm{Imag}\ T = \dim \mathrm{dom}\,(T) - \dim \mathrm{Ker}\ T$.

See Theorem 3.3.10, Theorem 3.3.1, and Theorem 3.3.2.

- ▶ If there is a one-to-one and onto linear transformation between two vector spaces, we say the vector spaces are isomorphic, a form of equivalence for vector spaces. See Definition 3.3.1.
- ▶ Any n-dimensional vector space is isomorphic to $\mathbb{R}^n$ via the coordinate mapping. See Theorem 3.3.4.

Exercises for Section 3.3

3.3.1.Determine whether the linear transformation is one-to-one, onto, both or neither. If it is not onto, find $\text{Imag}\ T$. If it is not one-to-one, find $\text{Ker}\ T$.

(a) $T\colon \mathbb{P}_2 \to \mathbb{R}^2$ defined by

$$T(a_0 + a_1x + a_2x^2) = \begin{bmatrix} a_0 + a_1 \\ a_2 \end{bmatrix}$$

(b) $T\colon \mathbb{P}_2 \to \mathbb{R}^3$ defined by

$$T(a_0 + a_1x + a_2x^2) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

(c) $T\colon \mathbb{P}_2 \to \mathbb{R}^3$ defined by

$$T(a_0 + a_1x + a_2x^2) = \begin{bmatrix} a_0 + a_1 \\ 2a_1 \\ a_1 + a_2 \end{bmatrix}$$

(d) $T\colon \mathbb{P}_2 \to \mathbb{R}^4$ defined by

$$T(a_0 + a_1x + a_2x^2) = \begin{bmatrix} a_0 \\ a_1 + a_2 \\ a_1 + a_2 \\ a_2 \end{bmatrix}$$

(e) $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 \\ x_2 \end{bmatrix}$$

(f) $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ x_1 \\ x_2 \end{bmatrix}$$

(g) $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

3.3.2.Determine whether the linear transformation is one-to-one, onto, both or neither.

(a) $T\colon \mathbb{P}_2 \to \mathbb{R}^4$ defined by

$$T(a_0 + a_1 x + a_2 x^2) = \begin{bmatrix} a_0 - a_1 \\ a_1 - a_2 \\ a_2 \\ a_1 \end{bmatrix}$$

(b) $T\colon \mathbb{P}_2 \to \mathbb{R}^4$ defined by

$$T(a_0 + a_1 x + a_2 x^2) = \begin{bmatrix} a_1 \\ a_1 - a_2 \\ a_2 \\ a_1 \end{bmatrix}$$

3.3.3.Let $T\colon \mathbb{P}_2 \to \mathbb{R}^3$ defined by

$$T(a_0 + a_1 x + a_2 x^2) = \begin{bmatrix} a_0 - a_1 \\ a_1 - a_2 \\ a_2 - a_0 \end{bmatrix}.$$

(a) Which of the following vectors are in $\operatorname{Ker} T$?

$$1 - x, \qquad 1 + x + x^2, \qquad x - x^2, \qquad 1 - x - x^2$$

(b) Which of the following vectors are in $\operatorname{Imag} T$?

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

3.3.4.Find the dimension of $\operatorname{Ker} T$ using the Rank-Nullity Theorem.

- $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$$

- $T\colon \mathbb{R}^5 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}\right) = \begin{bmatrix} x_3 \\ x_1 \end{bmatrix}$$

- $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_1 \end{bmatrix}$$

3.3.5. Suppose $T\colon V \to W$ is an isomorphism of vector spaces and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for V.

(a) Use the fact that T is onto to show that $\{T(\vec{v}_1), \ldots, T(\vec{v}_n)\}$ is a spanning set of W.

(b) Use the fact that T is one-to-one to argue that $\{T(\vec{v}_1), \ldots, T(\vec{v}_n)\}$ is a linearly independent set.

This completes the proof of Theorem 3.3.5.

3.3.6. Let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 47x_2 \\ x_2 - 48x_3 \\ x_3 \end{bmatrix}.$$

Calculate $T(\vec{e}_1)$, $T(\vec{e}_2)$, and $T(\vec{e}_3)$, and explain how this proves that T is an isomorphism.

3.3.7. Let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ and $S\colon \mathbb{R}^n \to \mathbb{R}^n$.

(a) Is it possible that $\dim \operatorname{Ker}\, T < \dim \operatorname{Ker}\,(S \circ T)$? Explain.

(b) Is it possible that $\dim \operatorname{Ker}\, S < \dim \operatorname{Ker}\,(S \circ T)$? Explain.

(c) Is it possible that $\dim \operatorname{Ker}\,(S \circ T) < \dim \operatorname{Ker}\, T$? Explain.

(d) Is it possible that $\dim \operatorname{Ker}\,(S \circ T) < \dim \operatorname{Ker}\, S$? Explain.

3.3.8. Let $V = \operatorname{Span}\left\{x, 1 + x^2\right\}$ in $\mathbb{P}_2$. Find two distinct isomorphisms T and S from V to $\mathbb{R}^2$.

3.4 Matrices

At some point, you might have been given the impression[28] that Linear Algebra is all about matrices. Don't feel bad; matrix theory is often confused with linear algebra. We know now that Linear Algebra is the study of linear transformations on vector spaces.[29] In this section, we will begin the discussion of what a matrix is and also how it can be connected to a linear transformation.

28: Certainly not from us.

29: One of the three unicorns does, anyway.

What is... a Matrix?

Definition 3.4.1 *An $m \times n$* ***matrix*** *A is a rectangular array of numbers with m rows and n columns:*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

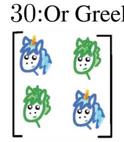
*The number a_{ij} in the ith row and jth column is called the ij**th entry**. Matrices are sometimes also written as*

$$A = [a_{ij}]_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}.$$

An $n \times n$ matrix is often called a ***square matrix****. For convenience, the set of all $m \times n$ matrices with real number entries will be denoted by $\mathcal{M}_{m\times n}$.*

The numbers in a matrix may be integers, real numbers, complex numbers, etc. The entries don't even have to be numbers! You could make a matrix of polynomials or even emojis[30] if you really wanted. In this course, we will use real numbers unless otherwise specified.[31]

30: Or Greek symbols. Or even...



31: I caught that! They're planning to specify non-real matrices at some point!

Probably in Chapter 5. I've heard everything gets a bit more complex there.

Do you think they'll change the meaning of $\mathcal{M}_{m\times n}$ then?

Quite possibly.

Another common notation (that will be particularly useful for us) is to think of an $m \times n$ matrix as n vectors from $\mathbb{R}^m$ all lined up next to each other, so

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \\ &= \left[\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} \cdots \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \right] = [\vec{a}_1\ \vec{a}_2 \cdots \vec{a}_n], \end{aligned}$$

where

$$\vec{a}_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad \text{for} \quad 1 \leq j \leq n.$$

We call $\vec{a}_j$ a *column vector* for the matrix A.

Example 3.4.1 Here's a 2×3 matrix:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = [\vec{a}_1\ \vec{a}_2\ \vec{a}_3],$$

where

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad \text{and} \quad \vec{a}_3 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Fun facts: $a_{21} = 4$ and $a_{12} = 2$.

Exploration 84 In the matrix A from Example 3.4.1, what are a_{13} and a_{22}?

Building a Linear Transformation from a Matrix

Before we can use a matrix to make a linear transformation, we will need to define a type of multiplication between a matrix and a vector.

Definition 3.4.2 *Let $A \in \mathcal{M}_{m\times n}$ with columns $\vec{a}_1, \ldots, \vec{a}_n$, and let $\vec{x} \in \mathbb{R}^n$. The product of a matrix and a vector, that is, the **product of** A **and** $\vec{x}$, is the linear combination of the columns of A with the entries of $\vec{x}$ as weights. That is,*

$$A\vec{x} = [\vec{a}_1\ \vec{a}_2 \cdots \vec{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

Example 3.4.2 Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}, \quad \text{and} \quad \vec{u} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}.$$

We can calculate $A\vec{v}$ and $B\vec{u}$; see Exploration 85.
Note, however, that $A\vec{u}$ and $B\vec{v}$ are not defined. For $A\vec{u}$, the matrix A has three columns, while $\vec{u}$ has only two components; thus, there is no way to form a linear combination. Similarly for $B\vec{v}$, the matrix B has only two columns, and $\vec{v}$ has three components.

Exploration 85 Let

$$A = \begin{bmatrix} 1 & 0 & 5 \\ 2 & -2 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & -1 \end{bmatrix},$$

$$\vec{v} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad \text{and} \quad \vec{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Calculate $A\vec{v}$ and $B\vec{u}$.

$$\begin{aligned} A\vec{v} &= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = 7\begin{bmatrix} 1 \\ 4 \end{bmatrix} + 8\begin{bmatrix} 2 \\ 5 \end{bmatrix} + 9\begin{bmatrix} 3 \\ 6 \end{bmatrix} \\ &= \begin{bmatrix} 7 \\ 28 \end{bmatrix} + \begin{bmatrix} 16 \\ 40 \end{bmatrix} + \begin{bmatrix} 27 \\ 54 \end{bmatrix} = \begin{bmatrix} 50 \\ 122 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} B\vec{u} &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \end{bmatrix} = 7\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + 8\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \\ &= \begin{bmatrix} 7 \\ 21 \\ 35 \end{bmatrix} + \begin{bmatrix} 16 \\ 32 \\ 48 \end{bmatrix} = \begin{bmatrix} 23 \\ 53 \\ 83 \end{bmatrix} \end{aligned}$$

With this product definition in hand, we are able to make any matrix into a linear transformation.[32]

32:  Read that last sentence again. It's a big deal.

Theorem 3.4.1 *Let $A \in \mathcal{M}_{m\times n}$, and define $T\colon \mathbb{R}^n \to \mathbb{R}^m$ by*

$$T(\vec{x}) = A\vec{x}.$$

Then T is a linear transformation.

When a linear transformation is defined as multiplication by a specific matrix, we often use the following terminology and notation.

Definition 3.4.3 *Let $A \in \mathcal{M}_{m\times n}$, and define $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ by*

$$T_A(\vec{x}) = A\vec{x}.$$

We will call T_A the ***linear transformation induced by*** *A.*

PROOF. We need to show that for any vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ and any scalar $\alpha \in \mathbb{R}$, we have

$$T_A(\vec{x}+\vec{y}) = T_A(\vec{x}) + T_A(\vec{y}) \quad \text{and} \quad T_A(\alpha\vec{x}) = \alpha T_A(\vec{x}).$$

We may write $A = [\vec{a}_1 \cdots \vec{a}_n]$. First note that

$$\begin{aligned} T_A(\vec{x}+\vec{y}) = A(\vec{x}+\vec{y}) &= [\vec{a}_1 \cdots \vec{a}_n] \begin{bmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{bmatrix} \\ &= (x_1+y_1)\vec{a}_1 + \cdots + (x_n+y_n)\vec{a}_n \\ &= (x_1\vec{a}_1 + \cdots + x_n\vec{a}_n) + (y_1\vec{a}_1 + \cdots + y_n\vec{a}_n) \\ &= [\vec{a}_1 \cdots \vec{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + [\vec{a}_1 \cdots \vec{a}_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \\ &= A\vec{x} + A\vec{y} = T_A(\vec{x}) + T_A(\vec{y}). \end{aligned}$$

Now, we need to show that for any scalar $\alpha \in \mathbb{R}$, we have $T_A(\alpha\vec{x}) = \alpha T_A(\vec{x})$. To see this,

$$\begin{aligned} T_A(\alpha\vec{x}) = A(\alpha\vec{x}) &= [\vec{a}_1 \cdots \vec{a}_n]\left(\alpha \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) \\ &= [\vec{a}_1 \cdots \vec{a}_n]\begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix} \\ &= (\alpha x_1)\vec{a}_1 + \cdots + (\alpha x_n)\vec{a}_n \\ &= \alpha(x_1\vec{a}_1 + \cdots + x_n\vec{a}_n) \\ &= \alpha A\vec{x} = \alpha T_A(\vec{x}). \end{aligned}$$

□

Example 3.4.3 Let

$$A = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \text{and} \quad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$$

and define $T_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $T_A(\vec{x}) = A\vec{x}$. Find the images $T_A(\vec{u})$ and $T_A(\vec{v})$. Here our linear transformation is defined as multiplication by a matrix A. We can compute

$$T_A(\vec{u}) = A\vec{u} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = (1)\begin{bmatrix} 3 \\ 0 \end{bmatrix} + (-2)\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix},$$

and both $\vec{u}$ and its image under T_A can be seen in Figure 3.3.
For the arbitrary vector $\vec{v}$, we have

$$T_A(\vec{v}) = A\vec{v} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1\begin{bmatrix} 3 \\ 0 \end{bmatrix} + v_2\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 3v_1 \\ -v_2 \end{bmatrix},$$

and now we have a nice formula for the image under T_A of any vector in $\mathbb{R}^2$.

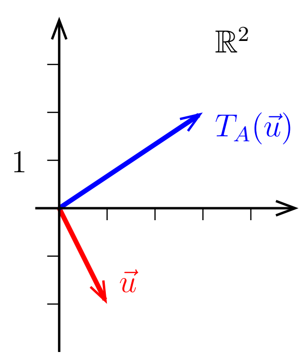


FIGURE 3.3. The vector $\vec{u} \in \mathbb{R}^2$ and its image $T_A(\vec{u}) \in \mathbb{R}^2$.

Exploration 86 Let's use our matrix B from earlier to define a linear transformation. Specifically, let

$$B = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix},$$

and define $T_B\colon \mathbb{R}^2 \to \mathbb{R}^3$ by $T_B(\vec{x}) = B\vec{x}$.

- What is $T_B\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$?

- What is $T_B\left(\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right)$?

Example 3.4.4 Suppose we wanted to define a map $G\colon \mathbb{R}^9 \to \mathbb{R}^{14}$ by $G(\vec{x}) = A\vec{x}$? How many rows and columns must A have for this linear transformation to be well-defined? Since

$$A\vec{x} = [\vec{a}_1 \cdots \vec{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n,$$

It follows that $\vec{x}$ needs as many components as A has columns. Since $\vec{x} \in \mathbb{R}^9$, we see that A must have nine columns. Moreover, $A\vec{x}$ is a linear combination of the column vectors $\vec{a}_j$. Since $A\vec{x} \in \mathbb{R}^{14}$, we must also have $\vec{a}_j \in \mathbb{R}^{14}$ for each $1 \leq j \leq 9$. It follows that A must have fourteen rows. Thus, $A \in \mathcal{M}_{14\times 9}$.

Revisiting Image and Kernel

Now that we've seen that a matrix can define a linear transformation, we can talk about the kernel and image of such a function. Let's start with an example.

Example 3.4.5 Let

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

and define $T_A\colon \mathbb{R}^4 \to \mathbb{R}^2$ by $T_A(\vec{x}) = A\vec{x}$. We shall find Ker T_A; that is, we would like to describe the set

$$\text{Ker } T_A = \left\{\vec{x} \in \mathbb{R}^4\colon T_A(\vec{x}) = \vec{0}\right\} = \left\{\vec{x} \in \mathbb{R}^4\colon A\vec{x} = \vec{0}\right\}.$$

Solving $A\vec{x} = \vec{0}$ means solving

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \text{where} \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

After doing the matrix-vector multiplication, this becomes

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ -2 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

or

$$\begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ x_2 - 2x_3 + 4x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Now we just have a system of two equation in four variables:

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

Solving for x_2 in the second component, and substituting it into the first equation, we have

$$\begin{aligned} x_1 &= -7x_3 + 8x_4 \\ x_2 &= 2x_3 - 4x_4. \end{aligned}$$

These are the conditions on the vector $\vec{x}$ that must be satisfied in order for $A\vec{x} = \vec{0}$. Thus

$$\text{Ker } T_A = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in \mathbb{R}^4 \colon x_1 = -7x_3 + 8x_4 \text{ and } x_2 = 2x_3 - 4x_4 \right\}.$$

However, this is not very satisfying. We can, in fact, clean this up quite a bit. Note that

$$\begin{aligned} &\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in \mathbb{R}^4 \colon x_1 = -7x_3 + 8x_4 \text{ and } x_2 = 2x_3 - 4x_4 \right\} \\ &= \left\{ \begin{bmatrix} -7x_3 + 8x_4 \\ 2x_3 - 4x_4 \\ x_3 \\ x_4 \end{bmatrix} \in \mathbb{R}^4 \colon x_3, x_4 \in \mathbb{R} \right\} \\ &= \left\{ x_3 \begin{bmatrix} -7 \\ 2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 8 \\ -4 \\ 0 \\ 1 \end{bmatrix} \colon x_3, x_4 \in \mathbb{R} \right\}. \end{aligned}$$

It follows that

$$\text{Ker } T_A = \text{Span} \left\{ \begin{bmatrix} -7 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 8 \\ -4 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

It turns out we really only needed the matrix A to tell us what T_A does to a vector in $\mathbb{R}^4$. Since T_A is completely determined by the matrix A, the following definition makes sense.

Definition 3.4.4 *Suppose $A \in \mathcal{M}_{m\times n}$ and $T_A \colon \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation induced by A. We then define* ***the kernel of the matrix*** *A,* *denoted* $\text{Ker } A$*, to be* $\text{Ker } T_A$*. That is,*

$$\text{Ker } A = \{\vec{x} \in \mathbb{R}^n \colon A\vec{x} = \vec{0}\} = \text{Ker } T_A.$$

Sometimes $\text{Ker } A$ *is called the* ***nullspace of the matrix*** A*.*

Example 3.4.6 Define $A = \begin{bmatrix} -1 & 0 & 4 \\ 0 & -1 & 2 \end{bmatrix}$. Then we can find $\text{Ker } A$ by solving for all vectors

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

such that $A\vec{x} = \vec{0}$. That is,

$$\begin{bmatrix} -1 & 0 & 4 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This gives us the equations

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

which simplify to $x_1 = 4x_3$ and $x_2 = 2x_3$. So

$$\text{Ker } A = \left\{ \begin{bmatrix} 4x_3 \\ 2x_3 \\ x_3 \end{bmatrix} : x_3 \in \mathbb{R} \right\} = \text{Span} \left\{ \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Exploration 87 Define $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. Find $\text{Ker } A$.

Now that we've seen that we can extend our definition of the kernel of a linear transformation to a matrix, let's look at an example of the image of a linear transformation defined by a matrix. Let's start with a familiar example.

Example 3.4.7 Let

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

and define $T_A \colon \mathbb{R}^4 \to \mathbb{R}^2$ by $T_A(\vec{x}) = A\vec{x}$. This time, we will find $\text{Imag } T_A$. Let

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

We can use this general vector to see what a general element of $\text{Imag } T_A$ must look like. That is,

$$T_A(\vec{x}) = A\vec{x} = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$= x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ -2 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 4 \end{bmatrix}.$$

From this equation, we conclude

$$\text{Imag } T_A = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix} \right\}.$$

Since this is definitely a spanning set for $\mathbb{R}^2$, we can say $\text{Imag } T_A = \mathbb{R}^2$. Now, how does this relate to the matrix A? Well, it's the span of the column vectors of A! This is true in general.

Definition 3.4.5 *Let $A = [\vec{a}_1 \cdots \vec{a}_n] \in \mathcal{M}_{m\times n}$. The* ***column space of*** *A, denoted $\text{Col } A$, is the span of the column vectors $\vec{a}_j$ for $1 \leq j \leq n$. That is,*

$$\text{Col } A = \text{Span}\{\vec{a}_1, \ldots, \vec{a}_n\}.$$

Exploration 88 Find $\text{Col } A$ where

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

Theorem 3.4.2 *Let $A \in \mathcal{M}_{m\times n}$ and suppose $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation induced by A. Then $\text{Imag } T_A = \text{Col } A$.*

PROOF. Let $\vec{x} \in \mathbb{R}^n$ and suppose $A = [\vec{a}_1 \cdots \vec{a}_n]$ where $a_j \in \mathbb{R}^m$ for all $1 \leq j \leq n$. We know

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

for some $x_i \in \mathbb{R}$ for each $1 \leq i \leq n$. Thus,

$$T_A(\vec{x}) = A\vec{x} = [\vec{a}_1 \cdots \vec{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

From this equation, we see that a vector $\vec{y} \in \mathbb{R}^m$ is in $\text{Imag } T_A$ if and only if $\vec{y} \in \text{Col } A$, and the sets are equal. □

One-to-one and Onto for T_A

Recall that we say a linear transformation is onto if its codomain and image are equal. Thus, we have the following corollary.

Corollary 3.4.3 *Let $A \in \mathcal{M}_{m\times n}$ and suppose $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation induced by A. Then T_A is onto if and only if the columns of A span $\mathbb{R}^m$.*

Now that we've brought up the concept of T_A being onto, what about determining when it's one-to-one?

Theorem 3.4.4 *Let $A \in \mathcal{M}_{m\times n}$ and suppose $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation induced by A. Then T_A is one-to-one if and only if the columns of A are linearly independent.*

PROOF. The columns of A are linearly dependent if and only if there are scalars (not all 0) $c_1, \ldots, c_n$ such that $c_1\vec{a}_1 + \cdots c_n\vec{a}_n = \vec{0}$. This is true if and only if

$$A\vec{c} = [\vec{a}_1 \cdots \vec{a}_n]\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \vec{0}.$$

Then we have a nonzero vector $\vec{c} \in \operatorname{Ker} A$. By Theorem 3.3.1, this is true if and only if T_A is *not* one-to-one. □

Example 3.4.8 Define the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\vec{x}) = A\vec{x}$, and

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

Since the columns of A fail to span $\mathbb{R}^3$, we have from Corollary 3.6.3 that $\operatorname{Imag} T$ has dimension two. Since $\mathbb{R}^3$ has dimension three, it is not possible that $\operatorname{Imag} T = \mathbb{R}^3$, so by Theorem 3.1.1, T is not onto.
However, since we can see the two column vectors are linearly independent, we know that T is one-to-one.

Exploration 89 Define the linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $T(\vec{x}) = A\vec{x}$, where

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

Is this linear transformation onto? Is it one-to-one?

Section Highlights

- A matrix is a rectangular array of numbers. See Definition 3.4.1.

- A matrix, A, with m rows and n columns, induces a linear transformation, $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$. See Theorem 3.4.1 and Definition 3.4.3.
- The span of the column vectors in a matrix A is called the column space of A and denoted $\mathrm{Col}\ A$, and $\mathrm{Imag}\ T_A = \mathrm{Col}\ A$. See Definition 3.6.1 and Theorem 3.4.2.
- The kernel of matrix A, denoted $\mathrm{Ker}\ A$, is defined as $\mathrm{Ker}\ T_A$. See Definition 3.4.4.
- The linear transformation T_A is one-to-one if and only if the columns of the matrix A are linearly independent. See Theorem 3.6.2.
- The linear transformation T_A is onto if and only if the columns of A span the codomain. See Corollary 3.6.3.

Exercises for Section 3.4

3.4.1.Let $\vec{x} \in \mathbb{R}^3$, $\vec{y} \in \mathbb{R}^4$, $\vec{z} \in \mathbb{R}^5$, and $A \in \mathcal{M}_{m\times n}$. What must m and n be…

(a) …so that $A\vec{x} = \vec{y}$?

(b) …so that $A\vec{x} = \vec{z}$?

(c) …so that $A\vec{z} = \vec{y}$?

3.4.2.Consider the vectors below:

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} \quad \vec{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} \quad \vec{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \\ 3 \\ 1 \end{bmatrix}$$

For each matrix A below, exactly one of $A\vec{x}$, $A\vec{y}$, $A\vec{u}$, and $A\vec{v}$ can be computed. Compute the one that is defined.

(a) $A = \begin{bmatrix} 1 & 3 \\ -2 & 5 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & -2 \\ 1 & 5 \\ -2 & 3 \\ 1 & 1 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 3 & 2 \\ -2 & 5 & 3 \end{bmatrix}$

(d) $A = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 2 & 3 \\ -2 & 0 & 3 \end{bmatrix}$

(e) $A = \begin{bmatrix} 1 & 3 & 2 & 1 \\ -2 & 5 & 0 & 2 \end{bmatrix}$

(f) $A = \begin{bmatrix} 1 & 3 & 2 & 3 & 1 \\ -2 & 1 & 3 & 2 & 0 \\ -1 & 1 & 2 & 0 & 0 \end{bmatrix}$

3.4.3.Let $T_A\colon \mathbb{R}^4 \to \mathbb{R}^2$ be the linear transformation induced by A where

$$A = \begin{bmatrix} 1 & 1 & -2 & 3 \\ 1 & -1 & 4 & -6 \end{bmatrix}.$$

Find $\operatorname{Ker} A$.

3.4.4.Let

$$A = \begin{bmatrix} 2 & -2 & 3 \\ 0 & 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 3 & -5 \end{bmatrix},$$

$$\vec{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad \text{and} \quad \vec{u} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

(a) Compute $A\vec{u}$ and $B\vec{v}$.

(b) Let $T_A\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation induced by A. Note that means $T_A(\vec{x}) = A\vec{x}$ for any $\vec{x} \in \mathbb{R}^3$.

(i) Find $\text{Imag } T_A = \text{Col } A$.

(ii) Find $\text{Ker } T_A = \text{Ker } A$.

(iii) Is T_A one-to-one?

(iv) Is T_A onto?

(c) Define $T_B\colon \mathbb{R}^2 \to \mathbb{R}^3$ to be the linear transformation induced by B.

(i) Find $\text{Imag } T_B = \text{Col } B$.

(ii) Find $\text{Ker } T_B = \text{Ker } B$.

(iii) Is T_B one-to-one?

(iv) Is T_B onto?

3.4.5.Let

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

$$\vec{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad \text{and} \quad \vec{u} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

(a) Compute $A\vec{u}$ and $B\vec{v}$.

(b) Let $T_A\colon \mathbb{R}^3 \to \mathbb{R}^4$ be the linear transformation induced by A. Note that means $T_A(\vec{x}) = A\vec{x}$ for any $\vec{x} \in \mathbb{R}^3$.

(i) Find $\text{Imag } T_A = \text{Col } A$.

(ii) Find $\text{Ker } T_A = \text{Ker } A$.

(iii) Is T_A one-to-one?

(iv) Is T_A onto?

(c) Define $T_B\colon \mathbb{R}^2 \to \mathbb{R}^2$ to be the linear transformation induced by B.

(i) Find $\text{Imag } T_B = \text{Col } B$.

(ii) Find $\text{Ker } T_B = \text{Ker } B$.

(iii) Is T_B one-to-one?

(iv) Is T_B onto?

3.4.6. For each matrix A below, do the following in order:

(a) Find the the domain and codomain of T_A.

(b) Find a basis for $\mathrm{Ker}\ A$ and compute $\dim \mathrm{Ker}\ A$.

(c) Use the Rank-Nullity Theorem to compute $\dim \mathrm{Col}\ A$.

(d) Find a basis for $\mathrm{Col}\ A$.

(e) Determine whether T_A is one-to-one, onto, both, or neither.

(f) Based on your previous answers, are the columns of A linearly independent?

(a) $A = \begin{bmatrix} 1 & 3 \\ 0 & 5 \end{bmatrix}$

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

(c) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 3 \\ 0 & 1 \end{bmatrix}$

(d) $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 3 & 1 \end{bmatrix}$

(e) $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 5 & 3 \end{bmatrix}$

(f) $A = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 0 & -4 \\ 0 & 0 & 3 \end{bmatrix}$

(g) $A = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 0 & 5 & 0 & 2 \end{bmatrix}$

(h) $A = \begin{bmatrix} 1 & 3 & 4 & 1 \\ 0 & 5 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 1 \end{bmatrix}$

(i) $A = \begin{bmatrix} 1 & 3 & 2 & 3 & 1 \\ -2 & 0 & 3 & 2 & 0 \\ -1 & 1 & 2 & 0 & 0 \end{bmatrix}$

3.4.7. Suppose we have the matrix

$$A = \begin{bmatrix} a & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

For what values of a is the induced linear transformation T_A an onto linear transformation? Why?

3.4.8. Determine by inspection of the columns whether these matrices correspond to transformations that are one-to-one. Explain your reasoning.

(a) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

3.4.9. Determine by inspection whether these matrices correspond to transformations that are onto. Explain your reasoning.

(a) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

3.4.10. Suppose $B \in \mathcal{M}_{n\times n}$ with induced linear transformation $T_B \colon \mathbb{R}^n \to \mathbb{R}^n$. Let $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$ be such that $\{\vec{x}_1, \vec{x}_2\}$ is a linearly independent set. If $B\vec{x}_1 = B\vec{x}_2 = \vec{0}$, what is the largest and smallest $\dim \mathrm{Ker}\, T_B$ and $\dim \mathrm{Imag}\, T_B$ can be? Explain your reasoning.

3.4.11. Let $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T(\vec{x}) = A\vec{x}$ where

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

For any vector $\vec{x} \in \mathbb{R}^3$, describe its image $T(\vec{x})$ geometrically.

3.5 The Matrix of a Linear Transformation

In the last section, we learned about matrices and how a matrix can be used to define a linear transformation. But can we go the other way? If we start with a linear transformation, is there a matrix that can be associated to it? Yes![33]

33: Yes?

YES! There are actually infinitely many.

That's too many. You can keep almost all of them.

Matrix Representation with the Standard Basis of $\mathbb{R}^n$

Let's start with an example.

Example 3.5.1 Let $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that for the standard basis vectors $\vec{e}_1, \vec{e}_2 \in \mathbb{R}^2$, we have

$$T(\vec{e}_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad \text{and} \quad T(\vec{e}_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}.$$

Note that for any vector $\vec{x} \in \mathbb{R}^2$, we have

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2.$$

Then using the definition of linear transformation,

$$\begin{aligned} T(\vec{x}) &= T(x_1\vec{e}_1 + x_2\vec{e}_2) \\ &= x_1T(\vec{e}_1) + x_2T(\vec{e}_2) = x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}. \end{aligned}$$

Then using the definition of the product of a matrix and a vector,

$$T(\vec{x}) = x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 & -3 \\ -7 & 8 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Thus, $T(\vec{x}) = A\vec{x}$, where

$$A = \begin{bmatrix} 5 & -3 \\ -7 & 8 \\ 2 & 0 \end{bmatrix}.$$

This strategy works in general!

Theorem 3.5.1 *Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Using the standard basis $\{\vec{e}_1, \dots, \vec{e}_n\}$ for $\mathbb{R}^n$, define*

$$A = [T(\vec{e}_1) \cdots T(\vec{e}_n)].$$

Then A is the unique matrix in $\mathcal{M}_{m\times n}$ such that $T = T_A$. That is, $T(\vec{x}) = A\vec{x}$ for any $\vec{x} \in \mathbb{R}^n$.

Because of its connection to the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, this matrix is sometimes referred to as the *standard matrix for the linear transformation T*. If you look carefully at the steps in Example 3.6.2, you will see how to prove this result. Let's walk through that together.

Exploration 90 PROOF. Let $\vec{x} \in \mathbb{R}^n$. We would normally write $\vec{x}$ as a column vector, but let's write it as

$$\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$$

instead. Note that this is the unique way to represent $\vec{x}$ in terms of the standard basis by Theorem 2.3.1. Using $\vec{x}$ in this fashion and the linearity of T, find a way to write $T(\vec{x})$ as $A\vec{x}$ for some matrix A.

Our theorem states that this matrix A is unique. How does this follows from the uniqueness result of Theorem 2.3.1?

□

Let's see an example of this theorem in action.

Example 3.5.2 Let $T\colon \mathbb{R}^3 \to \mathbb{R}^4$ by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ 0 \\ x_3 \\ 0 \end{bmatrix}.$$

Using Theorem 3.5.1, we can build A by finding the image by T of each standard basis vector for $\mathbb{R}^3$; these will be the columns of A. Specifically, we have $T(\vec{x}) = A\vec{x}$, where

$$A = [T(\vec{e}_1)\, T(\vec{e}_2)\, T(\vec{e}_3)] = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that we weren't told at the start that this was a linear transformation. However, if we verify that $T(\vec{x}) = A\vec{x}$ for any $\vec{x} \in \mathbb{R}^3$, we'll know T is a linear transformation from Theorem 3.4.1.

$$A\vec{x} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 - x_2 \\ 0 \\ x_3 \\ 0 \end{bmatrix} = T(\vec{x})$$

This tells us T is a linear transformation!

Example 3.5.3 First, let us take the frequent geometric convention to label the three coordinate directions in $\mathbb{R}^3$ as x, y, and z, respectively. Let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ be defined as the linear transformation rotating all vectors

$\pi/4$ radians around the z-axis and increasing by a factor of 2 in the z direction. Let's also assume this rotation is counterclockwise in the plane[34]

$$\left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} : x, y \in \mathbb{R} \right\}.$$

Let's use the standard basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ for the domain and codomain. Note that

$$T(\vec{e}_1) = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \\ 0 \end{bmatrix}, T(\vec{e}_2) = \begin{bmatrix} -\sqrt{2} \\ \sqrt{2} \\ 0 \end{bmatrix}, T(\vec{e}_3) = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

By Theorem 3.5.1, we have

$$A = [T(\vec{e}_1)\; T(\vec{e}_2)\; T(\vec{e}_3)] = \begin{bmatrix} \sqrt{2} & -\sqrt{2} & 0 \\ \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

34: This is the usual xy plane in $\mathbb{R}^3$.

Exploration 91 Find a linear transformation $T\colon \mathbb{R}^4 \to \mathbb{R}^4$ that interchanges the x_1 axis with the x_3 axis, maps the x_2 axis to $\vec{0}$, and does nothing to the x_4 axis. Then find a matrix representation for T using the standard basis for $\mathbb{R}^4$ in the domain and the codomain.

General Version of a Matrix Representation

Using coordinate vectors, Theorem 3.5.1 generalizes nicely for arbitrary finite dimensional vector spaces.

Theorem 3.5.2 *Let V and W be vector spaces and $T\colon V \to W$ be a linear transformation. For each pair of fixed bases, $\mathcal{B}_V$ for V and $\mathcal{B}_W$ for W, there exists a unique matrix A such that for all vectors $\vec{v} \in V$,*

$$[T(\vec{v})]_{\mathcal{B}_W} = A\, [\vec{v}]_{\mathcal{B}_V}.$$

Moreover, if $\mathcal{B}_V = \{\vec{v}_1, \ldots, \vec{v}_n\}$, then

$$A = \left[[T(\vec{v}_1)]_{\mathcal{B}_W} \cdots [T(\vec{v}_n)]_{\mathcal{B}_W} \right].$$

For any linear transformation $T\colon V \to W$, we call the matrix A obtained from Theorem 3.5.2 the **matrix representation of T relative to the bases $\mathcal{B}_V$ and $\mathcal{B}_W$**. Note that a different choice of basis for either V or W would result in a different matrix representation. Before we delve deep into the proof of the theorem, let's see it in action with an example.

Example 3.5.4 Let $T\colon \mathbb{P}_3 \to \mathbb{P}_3$ be defined by

$$T(a + bx + cx^2 + dx^3) = (a - b) + (c - d)x^2.$$

The advantage to Theorem 3.5.2 over Theorem 3.5.1 is that with the more general theorem, we can build a matrix representation for a linear transformation between any vector spaces, even if their vectors aren't column vectors in $\mathbb{R}^n$ for some positive integer n. This is done, of course, by way of coordinate vectors, so the matrix representation, A, will depend on our choice of bases. Let's use $\mathcal{B} = \{1, x, x^2, x^3\}$ as a basis for $\mathbb{P}_3$ in both the domain and codomain. Recall that a general vector $\vec{p} = a + bx + cx^2 + dx^3$ has coordinate vector

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix},$$

so, for example,

$$\left[T(x - x^3)\right]_{\mathcal{B}} = \left[-1 + x^2\right]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

To use Theorem 3.5.2, we need to find the coordinate of the image by T of each basis vector in $\mathcal{B}$. That is,

$$\begin{aligned} A &= \left[[T(1)]_{\mathcal{B}}\ [T(x)]_{\mathcal{B}}\ \left[T(x^2)\right]_{\mathcal{B}}\ \left[T(x^3)\right]_{\mathcal{B}}\right] \\ &= \left[[1]_{\mathcal{B}}\ [-1]_{\mathcal{B}}\ \left[x^2\right]_{\mathcal{B}}\ \left[-x^2\right]_{\mathcal{B}}\right] = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Great! We just built a matrix A that is a matrix representation for T. Let's think a bit about what all this was on the level of functions. We started with $T\colon \mathbb{P}_3 \to \mathbb{P}_3$. We know that $\mathbb{P}_3 \cong \mathbb{R}^4$ because the coordinate mapping $\varphi_{\mathcal{B}}\colon \mathbb{P}_3 \to \mathbb{R}^4$ is an isomorphism. You probably noticed that $A \in \mathcal{M}_{4\times 4}$, so we can define a linear transformation $T_A\colon \mathbb{R}^4 \to \mathbb{R}^4$ by $T_A(\vec{x}) = A\vec{x}$. This gives us the following commuting diagram of linear transformations:

$$\begin{array}{ccc} \mathbb{P}_3 & \xrightarrow{T} & \mathbb{P}_3 \\ {\scriptstyle \varphi_{\mathcal{B}}}\big\downarrow & & \big\downarrow{\scriptstyle \varphi_{\mathcal{B}}} \\ \mathbb{R}^4 & \xrightarrow{T_A} & \mathbb{R}^4 \end{array}$$

The way you should think about this diagram is that the coordinate isomorphisms represented by the vertical arrows allow us to translate between T and T_A. Then, we can examine all the stuff with T and $\mathbb{P}_3$ *or* we can use the matrix representation T_A and $\mathbb{R}^4$. We'll talk a bit more about diagrams like this below.

We should be ready now to talk about the proof of Theorem 3.5.2, so let's get started.

PROOF. Let's suppose V and W are vector spaces with $\dim V = n$ and $\dim W = m$, and fix some bases $\mathcal{B}_V = \{\vec{v}_1, \dots, \vec{v}_n\}$ and $\mathcal{B}_W = \{\vec{w}_1, \dots, \vec{w}_m\}$ for V and W, respectively. We know from Theorem 3.3.4 that the coordinate mapping gives an isomorphism between V and $\mathbb{R}^n$ and between W and $\mathbb{R}^m$.

This statement actually involves four different linear transformations, and they are all relevant here. Specifically, we have

- $\varphi_{\mathcal{B}_V} : V \to \mathbb{R}^n$ defined by $\varphi_{\mathcal{B}_V}(\vec{v}) = [\vec{v}]_{\mathcal{B}_V}$ for any $\vec{v} \in V$
- $\varphi^{-1}_{\mathcal{B}_V} : \mathbb{R}^n \to V$ defined by $\varphi^{-1}_{\mathcal{B}_V}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$
- $\varphi_{\mathcal{B}_W} : W \to \mathbb{R}^m$ defined by $\varphi_{\mathcal{B}_W}(\vec{w}) = [\vec{w}]_{\mathcal{B}_W}$ for any $\vec{w} \in W$
- $\varphi^{-1}_{\mathcal{B}_W} : \mathbb{R}^m \to W$ defined by $\varphi^{-1}_{\mathcal{B}_W}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}\right) = x_1\vec{w}_1 + \cdots + x_m\vec{w}_m$.

Suppose now that we have a linear transformation $T: V \to W$. We know from Theorem 3.2.3 that we can compose linear transformations to form new linear transformations. Consider then the composition

$$\varphi_{\mathcal{B}_W} \circ T \circ \varphi^{-1}_{\mathcal{B}_V} : \mathbb{R}^n \to \mathbb{R}^m.$$

Let's call this new linear transformation $\widehat{T}$. Here's a useful commuting diagram for this situation.

(3.3)
$$\underbrace{\mathbb{R}^n \xrightarrow{\varphi^{-1}_{\mathcal{B}_V}} V \xrightarrow{T} W \xrightarrow{\varphi_{\mathcal{B}_W}} \mathbb{R}^m}_{\widehat{T}}.$$

This $\widehat{T}$ is a linear transformation which has input of coordinate vectors for $\mathcal{B}_V$ and whose output is in coordinate vectors for $\mathcal{B}_W$. Also, it is determined by our linear transformation T! Now, let's compute the matrix A from Theorem 3.5.1 for $\widehat{T}$.

$$\begin{aligned} A &= \left[\widehat{T}(\vec{e}_1) \ \cdots \ \widehat{T}(\vec{e}_n)\right] \\ &= \left[\varphi_{\mathcal{B}_W} \circ T \circ \varphi^{-1}_{\mathcal{B}_V}(\vec{e}_1) \ \cdots \ \varphi_{\mathcal{B}_W} \circ T \circ \varphi^{-1}_{\mathcal{B}_V}(\vec{e}_n)\right] \\ &= \left[\varphi_{\mathcal{B}_W} \circ T(\vec{v}_1) \ \cdots \ \varphi_{\mathcal{B}_W} \circ T(\vec{v}_n)\right] \\ &= \left[[T(\vec{v}_1)]_{\mathcal{B}_W} \ \cdots \ [T(\vec{v}_n)]_{\mathcal{B}_W}\right] \end{aligned}$$

So by Theorem 3.5.1, we know $\widehat{T} = T_A$ for this matrix A. Moreover, by examining our functions we see that this says exactly that

$$[T(\vec{v})]_{\mathcal{B}_W} = A\,[\vec{v}]_{\mathcal{B}_V}.$$

□

The commuting diagram in Figure 3.4 is a convenient way to organize all of the sets and functions that came up in Theorem 3.5.1 and its proof. As long as we follow the maps in the direction they point, composing functions as we go, all that matters are the starting and ending points. As a bonus, remember that the coordinate mappings are isomorphisms, so, while they are not pictured, there are inverses for each of the coordinate mappings with arrows that point in the opposite direction. Since these vertical arrows represent isomorphisms, we can think of T and T_A as “equivalent.” They aren’t completely the same since they map between different spaces, but they will share all the linear transformation properties. [35]

35: This is the same thing that isomorphisms do for vector spaces!

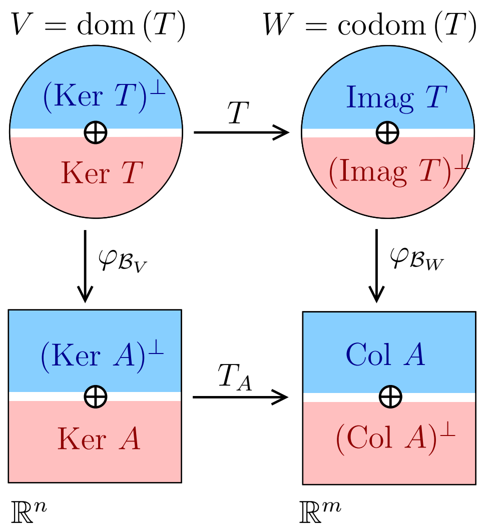


FIGURE 3.4. The linear transformation $T\colon V \to W$ generates four subspaces: $\text{Ker } T$ and $(\text{Ker } T)^{\perp}$ in the domain V and $\text{Imag } T$ and $(\text{Imag } T)^{\perp}$ in the codomain W. The matrix transformation T_A given by the matrix representation A of T also generates subspaces.

Figure 3.4 also strongly suggests that there is a natural decomposition of the domain and codomain of T_A that is analogous to the decompositions for T. This should not be terribly surprising, but we’ll put off a formal statement and proof of that fact until the next chapter.

Kernel and Image of T from A

We have said that the advantage of this matrix representation is that we can work with the matrix rather than the original linear transformation. In Section 3.4, we saw this with $\text{Ker } A$ and $\text{Ker } T_A$. Now we can define the relationship more generally between $\text{Ker } T$ and $\text{Ker } A$, where A is any matrix representation for T. Like we said before, it’s not quite true that these are equal. What *is* true is that these are isomorphic via the coordinate mapping, which is admittedly pretty close to being equal. Let’s state this formally and prove it.

Theorem 3.5.3 *Let V and W be vector spaces with bases $\mathcal{B}_V$ and $\mathcal{B}_W$, respectively. Let $T\colon V \to W$ be a linear transformation whose matrix representation relative to $\mathcal{B}_V$ and $\mathcal{B}_W$ is $A \in \mathcal{M}_{m\times n}$. Then*

$$\text{Ker } T = \{\vec{x} \in V\colon A\,[\vec{x}]_{\mathcal{B}_V} = \vec{0}\}.$$

That is, $\vec{x} \in \text{Ker } T$ if and only if $\varphi_{\mathcal{B}_V}(\vec{x}) \in \text{Ker } A$, and $\text{Ker } T \cong \text{Ker } A$ via the coordinate mapping.

PROOF. Suppose $\vec{x} \in \text{Ker } T$, so $T(\vec{x}) = \vec{0}$. Since $T = \varphi_{\mathcal{B}_W}^{-1} \circ T_A \circ \varphi_{\mathcal{B}_V}$, we know $T(\vec{x}) = \vec{0}$ if and only if

$$\begin{aligned} \left(\varphi_{\mathcal{B}_W}^{-1} \circ T_A \circ \varphi_{\mathcal{B}_V}\right)(\vec{x}) &= \vec{0} \text{ if and only if} \\ \left(T_A \circ \varphi_{\mathcal{B}_V}\right)(\vec{x}) &= \varphi_{\mathcal{B}_W}(\vec{0}) \text{ if and only if} \\ T_A\,[\vec{x}]_{\mathcal{B}_V} &= \vec{0}. \end{aligned}$$

Thus, $\vec{x} \in \text{Ker } T$ if and only if $[\vec{x}]_{\mathcal{B}_V} \in \text{Ker } T_A = \text{Ker } A$. □

We can do something similar with $\text{Imag } T$ and $\text{Col } A$. These are also isomorphic by the coordinate mapping.

Theorem 3.5.4 *Let V and W be vector spaces with bases $\mathcal{B}_V$ and $\mathcal{B}_W$, respectively. Let $T\colon V \to W$ be a linear transformation whose matrix representation with respect to $\mathcal{B}_V$ and $\mathcal{B}_W$ is $A \in \mathcal{M}_{m\times n}$. Then*

$$\text{Imag } T = \{\vec{w} \in W\colon\ [\vec{w}]_{\mathcal{B}_W} \in \text{Col } A\}.$$

That is, $\vec{w} \in \text{Imag } T$ if and only if $\varphi_{\mathcal{B}_W}(\vec{w}) \in \text{Col } A$, and $\text{Imag } T \cong \text{Col } A$ via the coordinate mapping.

The proof of Theorem 3.6.1 is similar to the proof of Theorem 3.6.10.[36]

36: Exercise!

Now that we know $\text{Ker } T \cong \text{Ker } A$ and $\text{Imag } T \cong \text{Col } A$ whenever A is a matrix representation for the matrix T, we actually can tell quite a lot about T from any matrix representation, A. For instance, if we know either $\dim \text{Ker } T$ or $\dim \text{Imag } T$, we can say whether T is one-to-one or onto. Now, we can conclude these from properties of the matrix A!

Theorem 3.5.5 *Let V and W be vector spaces. Let $T\colon V \to W$ be a linear transformation with matrix representation $A \in \mathcal{M}_{m\times n}$. Then we know the following:*

(a) *T is one-to-one if and only if the columns of A are linearly independent.*
(b) *T is onto if and only if the columns of A span $\mathbb{R}^m$.*
(c) *T is an isomorphism if and only if the columns of A form a basis for $\mathbb{R}^m$.*

This doesn't really need a proof. For the first two statements, we've just restated Corollary 3.6.3 and Theorem 3.6.2 in light of Theorems 3.6.10 and 3.6.1. Then the last statement is a combination of the first two with the definitions of a basis and an isomorphism.

More Examples!

The big takeaway here is that anything you want to know about *any* linear transformation on *any* finite dimensional vector spaces can be found using matrices and column vectors. That's actually quite amazing. We should look into that more.

Example 3.5.5 Define $T\colon \mathbb{P}_1 \to \mathbb{P}_2$ by taking the indefinite integral of vectors in $\mathbb{P}_1$ and using zero for the constant of integration. That is, for $\vec{p} = a + bx$, we have

$$T(\vec{p}) = T(a+bx) = ax + \frac{1}{2}bx^2 \in \mathbb{P}_2.$$

Since a and b are arbitrary real numbers, we see that $\mathrm{Imag}\ T$ is the set of vectors in $\mathbb{P}_2$ with zero as a constant term.
Using the standard bases $\mathcal{B}_1 = \{1, x\}$ and $\mathcal{B}_2 = \{1, x, x^2\}$ for $\mathbb{P}_1$ and $\mathbb{P}_2$ respectively, we have from Theorem 3.5.2 that $[T(\vec{p})]_{\mathcal{B}_2} = A\,[\vec{p}]_{\mathcal{B}_1}$, where

$$A = \left[[T(1)]_{\mathcal{B}_2}\ [T(x)]_{\mathcal{B}_2} \right] = \left[[x]_{\mathcal{B}_2}\ \left[\frac{1}{2}x^2\right]_{\mathcal{B}_2} \right] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{bmatrix}.$$

Then $\mathrm{Col}\ A$ is (by definition) the span of the vectors

$$\vec{a}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \vec{a}_2 = \begin{bmatrix} 0 \\ 0 \\ 1/2 \end{bmatrix}.$$

Since $\mathrm{Imag}\ T$ is the set of all vectors in $\mathbb{P}_2$ with zero as a constant term, we have

$$\begin{aligned} \{[\vec{p}]_{\mathcal{B}_2} \in \mathbb{R}^3 \colon \vec{p} \in \mathrm{Imag}\ T\} &= \left\{ a \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} : a, b \in \mathbb{R} \right\} \\ &= \mathrm{Span}\,\{\vec{e}_2, \vec{e}_3\} = \mathrm{Span}\,\{\vec{a}_1, \vec{a}_2\}, \end{aligned}$$

where $\vec{e}_i$ are the standard basis vectors in $\mathbb{R}^3$. This is precisely what we were told by Theorem 3.6.1:

$$\mathrm{Col}\ A = \{[\vec{p}]_{\mathcal{B}_2} \in \mathbb{R}^3 \colon \vec{p} \in \mathrm{Imag}\ T\}.$$

Exploration 92 Let $T\colon \mathbb{P}_2 \to \mathbb{P}_1$ be defined by $T(a+bx+cx^2) = (b-a)+ax$ for any $a+bx+cx^2 \in \mathbb{P}_2$. Find the matrix A that represents this transformation relative to the bases $\{x^2, x, 1\}$ and $\{x, 1\}$.

What is $\mathrm{Ker}\ A$? Use the coordinate mapping to translate your answer back to vectors in $\mathbb{P}_2$ to find $\mathrm{Ker}\ T$.

Example 3.5.6 Let $H_D = \text{Span}\left\{\vec{b}_1 = 1 + x, \vec{b}_2 = x + x^2, \vec{b}_3 = x^3\right\}$, so $\mathcal{B}_D = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ is a basis for H_D. Similarly, let

$$H_C = \text{Span}\left\{\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}\right\},$$

so $\mathcal{B}_C = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a basis for H_C. Let $T\colon H_D \to H_C$ be the linear transformation such that $T(\vec{b}_i) = \vec{v}_i$ for each $i = 1, 2, 3$.
We know $H_D \cong \mathbb{R}^3$ and $H_C \cong \mathbb{R}^3$, so we'd like very much to build a linear transformation $T_A\colon \mathbb{R}^3 \to \mathbb{R}^3$ such that $T_A(\vec{x}) = A\left[\vec{b}\right]_{\mathcal{B}_D} = \left[T(\vec{b})\right]_{\mathcal{B}_C}$, where $A \in \mathcal{M}_{3\times 3}$. This is why we have Theorem 3.5.2:

$$\begin{aligned} A &= \left[\left[T(\vec{b}_1)\right]_{\mathcal{B}_C} \; \left[T(\vec{b}_2)\right]_{\mathcal{B}_C} \; \left[T(\vec{b}_3)\right]_{\mathcal{B}_C}\right] \\ &= \left[[v_1]_{\mathcal{B}_C} \; [v_2]_{\mathcal{B}_C} \; [v_3]_{\mathcal{B}_C}\right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Wait, what? This actually checks out, but why is our matrix so simple this time? We started with isomorphic vector spaces; note that $H_C \cong H_D$. Then our linear transformation was built on the basis of our choosing; this flexibility to choose whatever basis you like can make your life much simpler.

In that previous example, we saw how nice our matrix can be given a careful choice of the bases for the vector spaces. What if we start with a linear transformation defined by a matrix, but we'd prefer a simpler-looking matrix. Well, we could *change the basis* we are using...

Example 3.5.7 Let $A = \begin{bmatrix} -1 & -2 & -1 \\ -1 & 0 & 0 \\ 2 & 2 & 1 \end{bmatrix}$ and define $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $T(\vec{x}) = A\vec{x}$ for any $\vec{x} \in \mathbb{R}^3$. Well, the form of A is not that bad, but let's see what it looks like if we were to replace the standard basis with the basis

$$\mathcal{B} = \left\{\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}\right\}.$$

Well, how do we do this? Actually, Theorem 3.5.2 tells us exactly what to do. We are still using coordinate vectors, even though we didn't actually change vector spaces. So, we first need to find $T(\vec{v}_1)$, $T(\vec{v}_2)$, and $T(\vec{v}_3)$.

$$T(\vec{v}_1) = \begin{bmatrix} -1 & -2 & -1 \\ -1 & 0 & 0 \\ 2 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$T(\vec{v}_2) = \begin{bmatrix} -1 & -2 & -1 \\ -1 & 0 & 0 \\ 2 & 2 & 1 \end{bmatrix}\begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}$$

$$T(\vec{v}_3) = \begin{bmatrix} -1 & -2 & -1 \\ -1 & 0 & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, we need to convert all of these to coordinate vectors relative to $\mathcal{B}$. Normally, to do this, we need to find coefficients $a, b, c \in \mathbb{R}$ such that $a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3 = \vec{x}$ where $\vec{x}$ is any vector in $\mathbb{R}^3$. However, we can do these quickly by inspection. Note that $T(\vec{v}_1) = -\vec{v}_1$, $T(\vec{v}_2) = \vec{v}_2$, and $T(\vec{v}_3) = \vec{0}$. Thus, we have the matrix B for the linear transformation T relative to the basis $\mathcal{B}$.

$$B = [[T(\vec{v}_1)]_{\mathcal{B}} \ [T(\vec{v}_2)]_{\mathcal{B}} \ [T(\vec{v}_3)]_{\mathcal{B}} \ ] = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

This new matrix B is also a matrix representation for T, and from this, we can see quickly that T is neither one-to-one nor onto. We can also identify that $\vec{v}_3 \in \operatorname{Ker} T$ and that $\dim \operatorname{Imag} T = 2$. But where did this new basis come from? You'll have to wait a bit to find out, but we'll get there.

Exploration 93 Let $A = \begin{bmatrix} 4 & 2 & 1 \\ 4 & 6 & 3 \\ -8 & -8 & -4 \end{bmatrix}$ and define $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $T(\vec{x}) = A\vec{x}$ for any $\vec{x} \in \mathbb{R}^3$. Just like in the example above, find a new matrix B that represents T with respect to the basis

$$\mathcal{B} = \left\{ \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} \right\}.$$

Section Highlights

- Every linear transformation on a finite dimensional vector space can be represented by a matrix. The matrix representation depends on the bases chosen for the domain and codomain. See Theorem 3.5.2
- If $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then the matrix

  $$A = [T(\vec{e}_1) \cdots T(\vec{e}_n)]$$

  is called the standard matrix for T. This matrix is the matrix for which $T = T_A$. See Theorem 3.5.1.
- If $T\colon V \to V$ is a linear transformation and $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for the vector space V, then the matrix for T with respect to $\mathcal{B}$ is given by

  $$[[T(\vec{v}_1)]_{\mathcal{B}} \cdots [T(\vec{v}_n)]_{\mathcal{B}}].$$

  See Theorem 3.5.2.

Exercises for Section 3.5

3.5.1. Using the standard bases for $\mathbb{R}^n$ and $\mathbb{P}_n$, find a matrix A such that $T(\vec{x}) = A\vec{x}$ for each of the following linear transformations:

(a) $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 3x_3 \\ x_2 + 4x_3 \end{bmatrix}$$

(b) $T\colon \mathbb{P}_2 \to \mathbb{R}^4$ defined by

$$T(a_0 + a_1x + a_2x^2) = \begin{bmatrix} a_0 + 3a_2 \\ a_1 - a_2 \\ a_2 \\ a_1 \end{bmatrix}$$

(c) $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(d) $T\colon \mathbb{P}_2 \to \mathbb{R}^4$ defined by

$$T(a_0 + a_1x + a_2x^2) = \begin{bmatrix} a_1 \\ a_1 - a_2 \\ a_2 \\ a_1 \end{bmatrix}$$

(e) $T\colon \mathbb{R}^2 \to \mathbb{R}^5$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 3x_2 \\ 0 \\ x_1 \\ 2x_1 + x_2 \\ 4x_1 - 5x_2 \end{bmatrix}$$

(f) $T\colon \mathbb{P}_3 \to \mathbb{R}^4$ defined by

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = \begin{bmatrix} a_1 + a_3 \\ a_1 - a_2 + 4a_3 \\ 2a_2 + a_0 \\ a_1 + a_2 + a_3 \end{bmatrix}$$

(g) $T\colon \mathbb{R}^4 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 3x_2 + x_4 \\ 4x_1 - 5x_2 + x_3 \end{bmatrix}$$

(h) $T\colon \mathbb{P}_3 \to \mathbb{R}^2$ defined by

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = \begin{bmatrix} a_0 - a_1 + a_3 \\ 3a_1 - 2a_2 + 4a_3 \end{bmatrix}$$

3.5.2. Let $T\colon \mathbb{P}_3 \to \mathbb{P}_1$ be defined by $T(ax^3 + bx^2 + cx + d) = (b-a)x + (c-d)$.

(a) Using the standard bases for $\mathbb{P}_3$ and $\mathbb{P}_1$, find the matrix representation A of T.

(b) Find $\text{Ker } A$ and $\text{Col } A$.

(c) Use your answers to part $b)$ to find $\text{Ker } T$ and $\text{Imag } T$.

3.5.3. Let $T\colon \mathbb{P}_3 \to \mathbb{P}_2$ be defined by $T(ax^3 + bx^2 + cx + d) = (a-c)x^2 + (b-d)x + (c-a)$.

(a) Using the standard bases for $\mathbb{P}_3$ and $\mathbb{P}_1$, find the matrix representation A of T.

(b) Find $\text{Ker } A$ and $\text{Col } A$.

(c) Use your answers to part $b)$ to find $\text{Ker } T$ and $\text{Imag } T$.

3.5.4. Consider the subspace

$$H = \left\{ \begin{bmatrix} -2y-3z \\ y \\ z \end{bmatrix} : y, z \in \mathbb{R} \right\}$$

of $\mathbb{R}^3$. (This is the plane determined by the equation $x + 2y + 3z = 0$.)

(a) Find a basis for H.

(b) Define a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ that has $\mathrm{Imag}\ T = H$.

(c) Find the matrix representation A for T with respect to the standard bases of $\mathbb{R}^2$ and $\mathbb{R}^3$.

3.5.5. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ 2x_1 \end{bmatrix}.$$

(a) Using the standard basis for $\mathbb{R}^2$, find a matrix A such that $T(\vec{x}) = A\vec{x}$.

(b) Using the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{R}^2$, find a matrix B such that $[T(\vec{x})]_{\mathcal{B}} = B\,[\vec{x}]_{\mathcal{B}}$. That is, find the matrix representation of T relative to the basis $\mathcal{B}$.

3.5.6. Suppose $T : \mathbb{R}^3 \to \mathbb{R}^2$ is the linear transformation defined by

$$T\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad T\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Find the matrix for T with respect to the standard bases of $\mathbb{R}^2$ and $\mathbb{R}^3$.

3.5.7. Suppose $T : \mathbb{P}_2 \to \mathbb{P}_2$ is the linear transformation defined by

$$T(1+x) = 1-x \quad T(x) = 1-x+x^2 \quad T(1+x^2) = 1.$$

Find the matrix for T with respect to the standard basis of $\mathbb{P}_2$.

3.5.8. Ricky has affixed a picture of Bubbles onto the square in $\mathbb{R}^2$ whose vertices are $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$. He decides this is much too small, and he prefers that Bubbles faces the other direction.

(a) Find the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that makes the picture of Bubbles ten times bigger and reflects the image across the vertical axis. In particular, determine $T(\vec{e}_1)$ and $T(\vec{e}_2)$.

(b) Find the matrix A that represents T with respect to the standard basis of $\mathbb{R}^2$.

3.5.9. Bubbles has now also affixed a picture of Ricky onto the square in $\mathbb{R}^2$ whose vertices are $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$. He would like to find a way to animate his picture by rotating around the origin.

(a) Find the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ that rotates the picture $45°$ counterclockwise, keeping the bottom tip at $(0,0)$. In particular, determine $T(\vec{e}_1)$ and $T(\vec{e}_2)$.

(b) Find the matrix A that represents T with respect to the standard basis of $\mathbb{R}^2$.

(c) Let $\vec{x} = A\vec{e}_1$. Find $A\vec{x}$. (Repeatedly applying A would give the animation Bubbles was looking for.)

3.5.10. Fix an angle $\theta \in [0, 2\pi)$, and let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that rotates vectors about the origin by θ radians (counterclockwise). Find the matrix A that represents T with respect to the standard basis of $\mathbb{R}^2$.

3.5.11. Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is the linear transformation defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + 4x_2 \\ 2x_2 + x_3 \\ 0 \end{bmatrix}.$$

For each basis $\mathcal{B}_i$ below, find the matrix representation for T with respect to $\mathcal{B}_i$.

(a) $\mathcal{B}_1 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$

(b) $\mathcal{B}_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$

(c) $\mathcal{B}_3 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$
Note that this basis is orthogonal. Use this to simplify the computations.

3.5.12. Suppose $T : \mathbb{P}_2 \to \mathbb{P}_2$ is the linear transformation defined by

$$T\left(a_0 + a_1 x + a_2 x^2\right) = (a_0 + a_1) + 3a_2 x^2.$$

For each basis $\mathcal{B}_i$ below, find the matrix representation for T with respect to $\mathcal{B}_i$.

(a) $\mathcal{B}_1 = \left\{1 + x, 1 - x, x^2\right\}$

(b) $\mathcal{B}_2 = \left\{1, 1 - x, x + x^2\right\}$

(c) $\mathcal{B}_3 = \left\{1 + x, 1 - x + x^2, x^2\right\}$

3.5.13. Prove Theorem 3.6.1.

3.6 More Fun with Linear Transformations

Now that we have a fun and convenient way to represent our linear transformations with matrices, we should go back and think about all the things we've learned about linear transformations. Some things become a bit easier in the context of matrices, so we're devoting this section to understanding what we can learn about a linear transformation from its matrix representation.

Kernel and Image of a Matrix

Let V and W be vector spaces (with $\dim V = n$ and $\dim W = m$), and let $T\colon V \to W$ be a linear transformation. Fix bases $\mathcal{B}_V$ and $\mathcal{B}_W$ for V and W respectively, and using Theorem 3.5.2, let $A \in \mathcal{M}_{m\times n}$ be the matrix that represents T with respect to these bases. To be clear,

$$A = [[T(\vec{v}_1)]_{\mathcal{B}_W} \ldots [T(\vec{v}_n)]_{\mathcal{B}_W}] \quad \text{and} \quad T(\vec{v}) = A[\vec{v}]_{\mathcal{B}_V}$$

where $\mathcal{B}_V = \{\vec{v}_1, \ldots, \vec{v}_n\}$.

Recall from Example 3.5.4 that the matrix A determines a linear transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T_A(\vec{v}) = A\vec{v}$ for any $\vec{v} \in \mathbb{R}^n$. This transformation differs from T since we are not assuming V is $\mathbb{R}^n$, only that it is *isomorphic* to $\mathbb{R}^n$ under the coordinate mapping. Using the isomorphisms we get from the coordinate mappings $\varphi_{\mathcal{B}_V}$ and $\varphi_{\mathcal{B}_W}$, we have the following commuting diagram:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle \varphi_{\mathcal{B}_V}}\big\downarrow & & \big\downarrow{\scriptstyle \varphi_{\mathcal{B}_W}} \\ \mathbb{R}^n & \xrightarrow{T_A} & \mathbb{R}^m \end{array}$$

With all of these maps in hand, consider

$$\text{Ker } A = \{\vec{v} \in \mathbb{R}^n \colon A\vec{v} = \vec{0}\}.$$

This is something we can compute from the matrix A. By the definition of T_A, this subspace of $\mathbb{R}^n$ is Ker T_A. Its image under the inverse of the coordinate mapping $\varphi_{\mathcal{B}_V}$ is then Ker T. Thus, the representation A can identify a set in $\mathbb{R}^n$ that is isomorphic to Ker T in V, but it also actually identifies the *exact* kernel of T in V when the subspace is tracked back through the coordinate mapping. This is best illustrated by an example.

Example 3.6.1 Let $T\colon \mathbb{P}_2 \to \mathbb{R}^2$ be defined by

$$T(\vec{p}) = \begin{bmatrix} p(0) \\ 0 \end{bmatrix}.$$

We showed a similar function was linear in Example 3.2.5, and T here is also a linear transformation. It is also not hard to check that Ker T is the set of polynomials in $\mathbb{P}_2$ with 0 for the constant term. We want to show here that using the basis $\mathcal{B} = \{x^2, x, 1\}$ for V, we have

$$\text{Ker } T = \{\vec{p} \in \mathbb{P}_2 \colon T(\vec{p}) = \vec{0}\} = \{\vec{p} \in \mathbb{P}_2 \colon A\,[\vec{p}]_{\mathcal{B}} = \vec{0}\}.$$

According to Theorem 3.5.2,

$$A = [T([\vec{e}_1]_{\mathcal{B}})\ T([\vec{e}_2]_{\mathcal{B}})\ T([\vec{e}_3]_{\mathcal{B}})] = \left[T(x^2)\ T(x)\ T(1)\right] = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

For any polynomial $\vec{p} = ax^2 + bx + c \in \mathbb{P}_2$, we have

$$[\vec{p}]_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix},$$

so we must solve the equation

$$A\,[\vec{p}]_{\mathcal{B}} = \vec{0}, \quad \text{or} \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{3.4}$$

From computing the matrix multiplication, we have

$$\begin{bmatrix} c \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus, $c = 0$ while a and b are free to be any real number; that is, the solution to Equation 3.4 is

$$\left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3 : c = 0 \right\}.$$

This set of coordinate vectors corresponds to the set of vectors in $\mathrm{Ker}\ T \subset \mathbb{P}_2$.

While we can use different bases to get different matrix representations for T, they all identify the same set of vectors in V that T maps to $\vec{0}$. Because of this, we choose not to feel guilty about the following slightly abusive notation:

$$\mathrm{Ker}\ T = \mathrm{Ker}\ A.$$

Indeed, for future reference, $\mathrm{Ker}\ A$ will refer to the kernel of the linear transformation T which A represents.

Exploration 94 Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $T(\vec{x}) = A\vec{x}$ for any $\vec{x} \in \mathbb{R}^3$ where

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

Find $\mathrm{Ker}\ T$ by finding all vectors $\vec{x}$ such that

$$A\vec{x} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Exploration 95 Let $T\colon \mathbb{P}_2 \to \mathbb{P}_1$ be defined by $T(a+bx+cx^2) = (b-a)+ax$ for any $a+bx+cx^2 \in \mathbb{P}_2$. Find the matrix A that represents this transformation relative to the bases $\{x^2, x, 1\}$ and $\{x, 1\}$.

You should have gotten an A that looks very familiar. What is $\text{Ker } A$? How does this translate into finding $\text{Ker } T$?

Definition 3.6.1 *Let $A = [\vec{a}_1 \cdots \vec{a}_n] \in \mathcal{M}_{m\times n}$. The* ***column space of*** *A, denoted $\text{Col } A$, is the span of the column vectors $\vec{a}_j$ for $1 \leq j \leq n$. That is,*

$$\text{Col } A = \text{Span}\left\{\vec{a}_1, \ldots, \vec{a}_n\right\}.$$

Exploration 96 Find $\text{Col } A$ where

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

Theorem 3.6.1 *Let V and W be vector spaces with bases $\mathcal{B}_V = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and $\mathcal{B}_W = \{\vec{w}_1, \ldots, \vec{w}_m\}$, respectively. Let $T\colon V \to W$ be a linear transformation with matrix representation $A \in \mathcal{M}_{m\times n}$. Then*

$$\text{Col } A = \{[\vec{w}]_{\mathcal{B}_W} \in \mathbb{R}^m \colon \vec{w} \in \text{Imag } T\}.$$

Simply put, this theorem says that $\text{Col } A$ **is** $\text{Imag } T$ **as coordinate vectors relative to** $\mathcal{B}_W$. This will provide us with an easy way to describe $\text{Imag } T$. The proof of this theorem is similar to that of Theorem 3.6.10 at the end of this section, so we do not include it here.

Again, because of this, we choose not to feel guilty about this other slightly abusive notation:

$$\text{Imag } T = \text{Col } A.$$

For future reference, $\text{Col } A$ can be thought of as the image of the linear transformation T which A represents.

Example 3.6.2 Define $T\colon \mathbb{P}_1 \to \mathbb{P}_2$ by taking the indefinite integral of vectors in $\mathbb{P}_1$ and using zero for the constant of integration. That is, for $\vec{p} = a + bx$, we have

$$T(\vec{p}) = T(a + bx) = ax + \frac{1}{2}bx^2 \in \mathbb{P}_2.$$

Since a and b are arbitrary real numbers, we see that $\text{Imag } T$ is the set of vectors in $\mathbb{P}_2$ with zero as a constant term.

Using the standard bases $\mathcal{B}_1 = \{1, x\}$ and $\mathcal{B}_2 = \{1, x, x^2\}$ for $\mathbb{P}_1$ and $\mathbb{P}_2$ respectively, we have from Theorem 3.5.2 that $[T(\vec{p})]_{\mathcal{B}_2} = A\,[\vec{p}]_{\mathcal{B}_1}$, where

$$A = \left[[T(1)]_{\mathcal{B}_2}\ [T(x)]_{\mathcal{B}_2} \right] = \left[[x]_{\mathcal{B}_2}\ \left[\frac{1}{2}x^2\right]_{\mathcal{B}_2} \right] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{bmatrix}.$$

Then Col A is (by definition) the span of the vectors

$$\vec{a}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \vec{a}_2 = \begin{bmatrix} 0 \\ 0 \\ 1/2 \end{bmatrix}.$$

Since $\operatorname{Imag} T$ is the set of all vectors in $\mathbb{P}_2$ with zero as a constant term, we have

$$\begin{aligned} \{[\vec{p}]_{\mathcal{B}_2} \in \mathbb{R}^3 \colon \vec{p} \in \operatorname{Imag} T\} &= \left\{ a \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} : a, b \in \mathbb{R} \right\} \\ &= \operatorname{Span}\{\vec{e}_2, \vec{e}_3\} = \operatorname{Span}\{\vec{a}_1, \vec{a}_2\}, \end{aligned}$$

where $\vec{e}_i$ are the standard basis vectors in $\mathbb{R}^3$. This is precisely what we were told by Theorem 3.6.1:

$$\text{Col } A = \{[\vec{p}]_{\mathcal{B}_2} \in \mathbb{R}^3 \colon \vec{p} \in \operatorname{Imag} T\}.$$

We get some very useful results for A immediately. We'll state them for linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ for simplicity of notation, but be aware that these results hold for general vector spaces as well.

Theorem 3.6.2 *Let A be the matrix such that the linear transformation $T \colon \mathbb{R}^n \to \mathbb{R}^m$ is given by $T(\vec{x}) = A\vec{x}$. Then T is one-to-one if and only if the columns of A are linearly independent.*

PROOF. The columns of A are linearly dependent if and only if there are scalars (not all 0) $c_1, \ldots, c_n$ such that $c_1\vec{a}_1 + \cdots c_n\vec{a}_n = \vec{0}$. This is true if and only if

$$A\vec{c} = [\vec{a}_1 \cdots \vec{a}_n] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \vec{0}.$$

Then we have a nonzero vector $\vec{c} \in \operatorname{Ker} A$. By Theorem 3.3.1, this is true if and only if T is *not* one-to-one. □

Example 3.6.3 As we saw in Example 3.6.2, the linear transformation $T \colon \mathbb{P}_1 \to \mathbb{P}_2$ given by taking the indefinite integral of vectors in $\mathbb{P}_1$ and using zero for the constant of integration has matrix representation (with respect to the standard basis)

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

Since the columns of A are linearly independent, we have from Theorem 3.6.2 that T is one-to-one.

Theorem 3.6.3 *Let A be the matrix such that the linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is given by $T(\vec{x}) = A\vec{x}$. The linearly independent columns of A form a basis for* $\operatorname{Imag} T$.

PROOF. Let $\vec{x} \in \mathbb{R}^n$, so $T(\vec{x})$ is an arbitrary vector in $\operatorname{Imag} T$. Since

$$T(\vec{x}) = A\vec{x} = x_1\vec{a}_1 + \cdots x_n\vec{a}_n,$$

it follows that $\vec{x} \in \operatorname{Span}\{\vec{a}_1, \ldots, \vec{a}_n\}$, so $\operatorname{Imag} T \subseteq \operatorname{Span}\{\vec{a}_1, \ldots, \vec{a}_n\}$. Now pick $\vec{y} \in \operatorname{Span}\{\vec{a}_1, \ldots, \vec{a}_n\}$, so $\vec{y} = c_1\vec{a}_1 + \cdots + c_n\vec{a}_n$ for some scalars $c_1, \ldots, c_n$. It follows that

$$\vec{y} = [\vec{a}_1 \cdots \vec{a}_n] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = A\vec{c},$$

so $\vec{y} \in \operatorname{Imag} T$. Thus, $\operatorname{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} \subseteq \operatorname{Imag} T$, so $\operatorname{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} = \operatorname{Imag} T$. The result then follows from Theorem 2.1.2, the fact that a subset of a spanning set is a basis. $\square$

Corollary 3.6.4 *Let A be the matrix such that the linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is given by $T(\vec{x}) = A\vec{x}$. Then T is onto if and only if the columns of A span $\mathbb{R}^m$.*

Example 3.6.4 Define the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\vec{x}) = A\vec{x}$, and

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

Since the columns of A fail to span $\mathbb{R}^3$, we have from Theorem 3.6.3 that $\operatorname{Imag} T$ has dimension two. Since $\mathbb{R}^3$ has dimension three, it is not possible that $\operatorname{Imag} T = \mathbb{R}^3$, so by Theorem 3.1.1, T is not onto.

Exploration 97 Define the linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $T(\vec{x}) = A\vec{x}$, and

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

Is this linear transformation onto?

Exploration 98 Suppose V and W are vector spaces with the property that $\dim W > \dim V$. Let $T\colon V \to W$ be a linear transformation. Use Theorem 3.6.3 and an argument similar to the one in the example above to explain why T is not onto.

Subspaces Induced by Matrix Representations

Not only do matrices allow us to discover more about our favorite subspaces generated by linear transformations, they naturally generate some new subspaces we've not yet discussed. Before we get to those, though, there's another tool we'll need.

Definition 3.6.2 *Let $A \in \mathcal{M}_{m\times n}$. The* ***transpose of*** *A, denoted A^T, is the matrix in $\mathcal{M}_{n\times m}$ derived from A by making the jth column of A into the jth row for each $1 \le j \le n$.*

Example 3.6.5 Let

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}.$$

Then

$$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \text{and} \quad B^T = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}.$$

Theorem 3.6.5 *Let $A, B \in \mathcal{M}_{m\times n}$ and $\alpha \in \mathbb{R}$. Then*
- $(A+B)^T = A^T + B^T$,
- $(\alpha A)^T = \alpha A^T$, *and*
- $(A^T)^T = A$.

PROOF. These are obvious.[37] □

37: Prove it!

Corollary 3.6.6 *Let $T_T\colon \mathcal{M}_{m\times n} \to \mathcal{M}_{n\times m}$ be the function that relates any matrix $A \in \mathcal{M}_{m\times n}$ to $A^T \in \mathcal{M}_{n\times m}$. Then T_T is a linear transformation.*

Exploration 99 Let $T_T\colon \mathcal{M}_{m\times n} \to \mathcal{M}_{n\times m}$ be the linear transformation that relates any matrix $A \in \mathcal{M}_{m\times n}$ to $A^T \in \mathcal{M}_{n\times m}$. Find a matrix representation for T_T when $m = 3$ and $n = 2$ using the bases in Figure 3.5.

Hint: The answer is

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

$$\begin{aligned}
\mathcal{B}_{3\times 2} &= \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \right. \\
&\quad \left. \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \\
\mathcal{B}_{2\times 3} &= \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \right. \\
&\quad \left. \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}
\end{aligned}$$

Figure 3.5: These are basis for $\mathcal{M}_{3\times 2}$ and $\mathcal{M}_{2\times 3}$

Speaking of rows of a matrix,

Definition 3.6.3 *For a matrix $A \in \mathcal{M}_{m\times n}$, let $\vec{r}_i$ be the vector formed from the ith row of A for each $1 \leq i \leq m$. The* ***row space of*** *A, denoted $\text{Row } A$, is the span of these row vectors. That is,*

$$\text{Row } A = \text{Span}\left\{\vec{r}_1, \ldots, \vec{r}_m\right\}.$$

Theorem 3.6.7 *For a matrix $A \in \mathcal{M}_{m\times n}$, $\text{Row } A$ is a subspace of $\mathbb{R}^m$.*

PROOF. This follows from Theorem 3.6.3 by taking the transpose of your matrix. □

Now for something really cool. We have fun subspaces of domains and codomains for linear transformations (the kernel and image, respectively), but what about the rest of the domain and codomain? You would not be shocked to find that the orthogonal complement of the kernel is a subspace of the domain[38], and the orthogonal complement of the image is a subspace of the codomain (Theorem 2.4.2). What is surprising is that these orthogonal complements are also given by the matrix representation for the linear transformation. Behold!

38: This is even isomorphic to the image from a theorem in Section 3.3.

Theorem 3.6.8 *Let A be the matrix such that the linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is given by $T(\vec{x}) = A\vec{x}$. Then*

$$\text{Ker } A = (\text{Row } A)^{\perp} \quad \textit{and} \quad \text{Imag } A = \left(\text{Ker } A^T\right)^{\perp}.$$

See Figure 4.3 for some geometric intuition.

PROOF. Let $A = [a_{ij}]$ for $1 \leq i \leq m$ and $1 \leq j \leq n$, let $\vec{r}_i$ for $1 \leq i \leq m$ be the row vectors of A, and let $\vec{a}_j$ for $1 \leq j \leq n$ be the column vectors of A.

Then

$$\begin{aligned}
& \vec{x} \in \text{Ker } A \\
\Leftrightarrow\ & A\vec{x} = \vec{0} \\
\Leftrightarrow\ & x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{0} \\
\Leftrightarrow\ & x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \\
\Leftrightarrow\ & x_1 a_{i1} + \cdots + x_n a_{in} = 0 \text{ for } 1 \le i \le n \\
\Leftrightarrow\ & \vec{x} \cdot \vec{r}_i = 0 \text{ for } 1 \le i \le n.
\end{aligned}$$

Thus, $\vec{x} \in \text{Ker } A$ if and only if it is orthogonal to every row of A. Since $\text{Row } A$ is the span of the rows of A, the result follows. A similar argument proves $(\text{Ker } A^T)^\perp = \text{Col } A$. □

Exploration 100 Consider the matrix

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

We spent some time earlier considering $\text{Col } A$ and $\text{Ker } A$. Find $\text{Row } A$ and $\text{Ker } A^T$.

Check that the claims of the theorem line up as expected.

Corollary 3.6.9 *Let A be the matrix such that the linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is given by $T(\vec{x}) = A\vec{x}$. Then*

$$\text{dom}\,(T) = \text{Ker } A \oplus \text{Row } A \quad \textit{and} \quad \text{codom}\,(T) = \text{Imag } A \oplus \text{Ker } A^T.$$

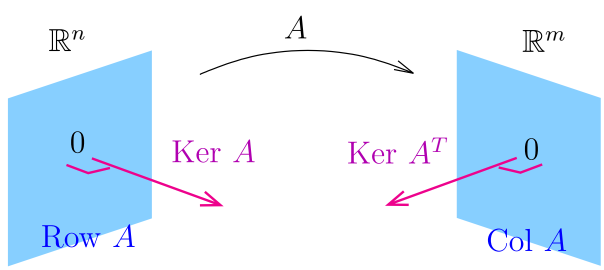


FIGURE 3.6. Some people refer to this as "the splits" of $\text{dom}\,(T)$ and $\text{codom}\,(T)$.

At the beginning of this section, we spent some time relating $\mathrm{Ker}\ A$ to $\mathrm{Ker}\ T$. Let's state this formally and prove it.

Theorem 3.6.10 *Let V and W be vector spaces with bases $\mathcal{B}_V = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and $\mathcal{B}_W = \{\vec{w}_1, \ldots, \vec{w}_m\}$, respectively. Let $T\colon V \to W$ be a linear transformation with matrix representation $A \in \mathcal{M}_{m\times n}$. Then*

$$\mathrm{Ker}\ T = \{\vec{v} \in V\colon A\,[\vec{v}]_{\mathcal{B}_V} = \vec{0}\}.$$

PROOF. Suppose $\vec{v} \in \mathrm{Ker}\ T$, so $T(\vec{x}) = \vec{0}$, and there are weights $a_1, \ldots, a_n$ such that $\vec{x} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$, that is,

$$[\vec{x}]_{\mathcal{B}_V} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

By Theorem 3.2.2, the coordinate mapping is a linear transformation, so it maps the zero vector to the zero vector. Thus $T(\vec{x}) = \vec{0}$ if and only if $[T(\vec{x})]_{\mathcal{B}_W} = \vec{0}$. Moreover, since T and the coordinate mapping are both linear,

$$\begin{aligned} & [T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n)]_{\mathcal{B}_W} = \vec{0} \\ \Leftrightarrow\quad & a_1\,[T(\vec{v}_1)]_{\mathcal{B}_W} + \cdots + a_n\,[T(\vec{v}_n)]_{\mathcal{B}_W} = \vec{0} \\ \Leftrightarrow\quad & \big[[T(\vec{v}_1)]_{\mathcal{B}_W} \cdots [T(\vec{v}_n)]_{\mathcal{B}_W}\big] \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \vec{0} \\ \Leftrightarrow\quad & A\,[\vec{x}]_{\mathcal{B}_V} = \vec{0}. \end{aligned}$$

□

Exercises for Section 3.6

3.6.1.n Define the linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by $T(\vec{x}) = A\vec{x}$, and

$$A = \begin{bmatrix} a & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

For what values of a is this an onto linear transformation?

3.6.2. Prove Theorem 4.4.8. That is, let $A, B \in \mathcal{M}_{m\times n}$ and $\alpha \in \mathbb{R}$. Prove

(a) $(A + B)^T = A^T + B^T$,

(b) $(\alpha A)^T = \alpha A^T$, and

(c) $(A^T)^T = A$.

3.6.3. Let $T\colon \mathbb{P}_3 \to \mathbb{P}_1$ by $T(ax^3 + bx^2 + cx + d) = (b - a)x + (c - d)$. Using the standard bases for $\mathbb{P}_3$ and $\mathbb{P}_1$, find the matrix representation of T and use this to find $\mathrm{Ker}\ A$ and $\mathrm{Col}\ A$. What are $\dim \mathrm{Ker}\ A^T$ and $\dim \mathrm{Row}\ A$?

3.6.4. Let $T\colon \mathbb{P}_3 \to \mathbb{P}_2$ by $T(ax^3 + bx^2 + cx + d) = (a - c)x^2 + (b - d)x + (c - a)$. Using the standard bases for $\mathbb{P}_3$ and $\mathbb{P}_2$, find the matrix representation of T and use this to find $\mathrm{Ker}\ A$ and $\mathrm{Col}\ A$. What are $\dim \mathrm{Ker}\ A^T$ and $\dim \mathrm{Row}\ A$?

3.6.5. Determine by inspection of the columns whether these matrices correspond to transformations that are one-to-one.

(a) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

3.6.6. Determine by inspection whether these matrices correspond to transformations that are onto.

(a) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

3.6.7.Let A be any $n \times m$ matrix. Prove $(\text{Ker } A^T)^\perp = \text{Col } A$.

3.6.8.Find $\text{Ker } A$, $\text{Row } A$, $\text{Col } A$, and $\text{Ker } A^T$ when A is the matrix below.

(a) $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & -1 & 0 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

3.7 Applications of Linear Tranformations

Computer Graphics and Animation

See the picture of Ricky's beautiful unicorn hooves in Figure 3.7. Yes, Ricky likes to wear zebra leg warmers.[39] A single vector would be a reasonable model[40] of Ricky's leg if unicorn legs did not bend. However, since Ricky has both knees and ankles, we should use three vectors; again, see Figure 3.7.

39: Doesn't everyone?

40: As we'll see, it's actually a very reasonable model.

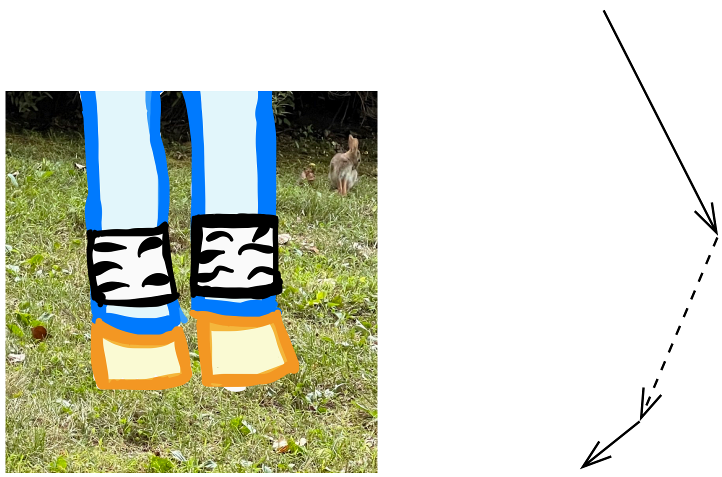


FIGURE 3.7. There are unicorn feet to the left and a vector impression of a unicorn leg to the right. The middle vector is dashed to simulate Ricky's beautiful zebra leg warmers; they're arranged to suggest a very horse-like galloping motion.

Our goal here is to animate our picture of Ricky's leg; more specifically, we will animate the vector representation of Ricky's leg. What should this entail? We should be able to translate, rotate, and scale the image. In this context, scaling might be weird. Legs usually stay the same size, but perhaps unicorns have some strange, little-known, femur-stretching powers. All kidding aside, scaling is actually extremely useful in images to give the impression that something is getting nearer or farther away. That's three things we'll need to do then: translation, rotation, and scaling.

When animating an image using a computer, sometimes it's better to use a vector (to represent a unicorn femur, for example), and sometimes it makes more sense to use a point (to represent a hoof, for example). For the entirety of this book, we've used different types of brackets or fonts to indicate whether something is a point or a vector. Since it's relatively difficult to explain this distinction to a computer, an additional component is often added to the vectors; we put a 1 in that component if our object is a vector and a 0 if it's a point. For example,

$$\text{The vector} \quad \vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{is} \quad \vec{v} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}.$$

$$\text{The point} \quad v = (x, y, z) \quad \text{is} \quad v = \begin{bmatrix} x \\ y \\ z \\ 0 \end{bmatrix}.$$

Since vector convention has vectors rooted at the origin and leg convention has bones connected to other bones, we need to translate our second and third leg vectors so they are appropriately attached to the preceding leg vector. This is another handy[41] feature of the point/vector distinction; they allow us to *translate* our vectors in space. If you are concerned that translation is not a linear transformation, your concern is very well placed. Any translation by a nonzero vector is *not* a linear transformation. Oh, hey. We should prove that.

41:  No pun intended. Seriously! We promise!

Exploration 101 Let $A \in \mathcal{M}_{m\times n}$, and let $\vec{v} \in \mathbb{R}^m$ be nonzero. Define $T\colon \mathbb{R}^n \to \mathbb{R}^m$ by $T(\vec{x}) = A\vec{x} + \vec{v}$. Show that T is not a linear transformation.

Looks like we're gonna need a new definition.

Definition 3.7.1 *An **affine transformation** is a linear transformation composed with a translation.*

To avoid ambiguity, when doing an affine transformation, we will use the point notation for the translation part of the transformation.

Example 3.7.1 Let $M \in \mathcal{M}_{m\times n}$, and let b be a point in $\mathbb{R}^m$. Then $T\colon \mathbb{R}^n \to \mathbb{R}^m$ by $T(\vec{x}) = M\vec{x} + b$ is an affine transformation. Moreover, T is a linear transformation if and only if b is the point at the origin. Lastly, it's important to note that the "+" in the equation is *not* vector addition; b isn't even a vector! The duplicitous plus, in this very specific context, indicates simply to "do translation by b when you're done with your matrix multiplication by M."

Let's get back to vector unicorn legs. See Figure 3.8. Note that the vector $\vec{t}+\vec{f}$ is not attached to the end of $\vec{f}$, it's way too long, *and* it's pointing in the wrong direction. However, the *translated* vector $\vec{t}+f$ has none of the aforementioned problems. Hooray for affine transformations!

Now we can do translations; that was the hard part. We've actually done rotation already; see Exercise 3.5.10 in Section 3.5. Indeed, to rotate a vector

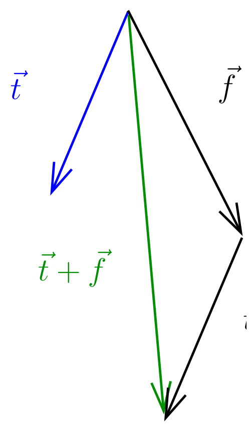


FIGURE 3.8. The shin bone $\vec{t}$ is connected to the thigh bone $\vec{f}$. The thigh bone $\vec{f}$ is connected to the hip bone (not pictured). We have chosen the letters "f" and "t" for *femur* and *tibia* because we don't know anything about biology, no one knows anything about unicorn biology, and couldn't be bothered to do a quick internet search.

$\vec{v} \in \mathbb{R}^3$ counterclockwise around the z-axis by an angle θ, we multiply

$$\vec{v} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{by} \quad \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The last row and column are there to accommodate our vector/point notational convention. More generally, if we'd like to rotate a vector $\vec{v} \in \mathbb{R}^3$ counterclockwise around an arbitrary unit vector $\vec{u}$ by an angle θ, we multiply

$$\vec{v} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

by

$$\begin{bmatrix} \cos\theta + u_1^2(1-\cos\theta) & u_1u_2(1-\cos\theta) - u_3\sin\theta & u_1u_3(1-\cos\theta) + u_2\sin\theta & 0 \\ u_1u_2(1-\cos\theta) + u_3\sin\theta & \cos\theta + u_2^2(1-\cos\theta) & u_2u_3(1-\cos\theta) - u_1\sin\theta & 0 \\ u_1u_3(1-\cos\theta) - u_2\sin\theta & u_2u_3(1-\cos\theta) + u_1\sin\theta & \cos\theta + u_3^2(1-\cos\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

That's not hard to verify, but it is pretty gross and annoying.

The only thing remaining to do is scaling, and this, like rotation, is a linear transformation, so it can be represented as a matrix transformation.

Exploration 102 Let $a, b, c \in \mathbb{R}$ and $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by scaling the first coordinate of vectors by a, the second coordinate by b, and the third coordinate by c. Find a matrix representation for T.

Alright! We now have a way to do translation, rotation, and scaling of our vector model of Ricky's leg. This is gonna be great. As we decided earlier, our vector model of Ricky's leg has three vectors, but now we label them carefully using affine transformations so that they are appropriately connected to each other. For example, since we want $\vec{t}$ to connect to the end of vector $\vec{f}$, we use an affine transformation to translate it to the end of $\vec{f}$, giving us $\vec{t} + f$. Similarly, to appropriately place $\vec{a}$, we use $\vec{a} + t + f$; see Figure 3.9.

Let A, B, and C be rotation matrices that rotate vectors counterclockwise about the origin by θ_1, θ_2, and θ_3 radians respectively. If we use A to rotate $\vec{f}$ by θ_1 radians, we must according rotate the translation parts of each affine transformation we used that involved $\vec{f}$. However, we might also want $\vec{t}$ to move, so we could apply B to $\vec{t}$ and the translation parts of each affine transformation we used that involved $\vec{t}$. Lastly, we can apply C to the little ankle vector $\vec{a}$. Again, see Figure 3.9. By appropriately moving the angles θ_1, θ_2, and θ_3, we can make this vector leg gallop (or spin in wildly unnatural ways if so desired).

In practice, we would then construct three more vector unicorn legs and use translations to place them on a vector unicorn body. At this point, we leave it to the professionals, but we know, by the power of linear algebra, we could do it were we so inclined. In the late 1970's and early 1980's many video games used just vectors to create "graphics" of tanks and asteroids and the like. More recently, carefully rendered, realistic images are associated to each vector. Indeed, we could overlay Ricky's zebra leg-warmer covered tibia over the vector $\vec{t}$, and that would be a significant improvement for realism's sake over the blue vector. Again, we leave this to the professionals.

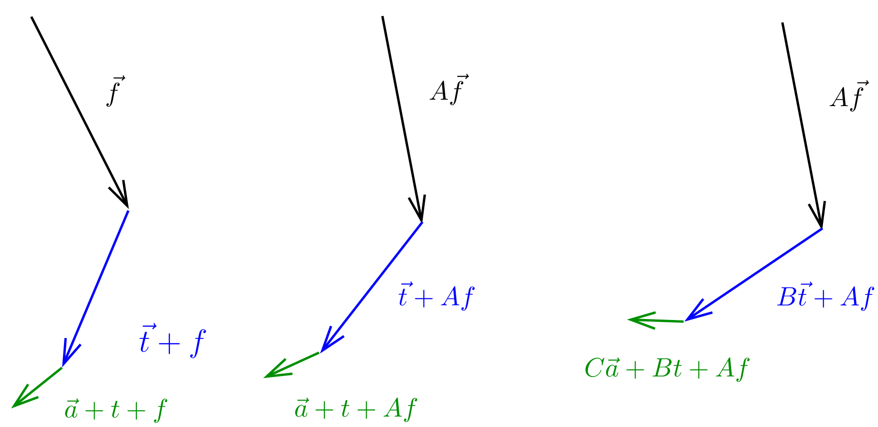


Figure 3.9: Various transformations of various leg vectors