

# Chapter 2

Ben Zelditch

August 14, 2017

## Linear Transformations

Linear transformations (or maps) are the most important topic in linear algebra. In fact, linear algebra basically is the study of linear transformations between vector spaces. There are a few different ways one can think about them and we'll look at all of them in this chapter.

## Linear Transformations

Let's first define what a linear transformation is. Let  $V, W$  be two vector spaces. A linear transformation  $T : V \rightarrow W$  is a function with the following two properties:

### Additivity

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

### Homogeneity

$$T(\alpha v) = \alpha T(v) \text{ for all } \alpha \in \mathbb{R} \text{ and } v \in V$$

That's it. Pretty simple, right? Note that homogeneity implies that  $T(0) = 0$  for *any* linear map  $T$  because

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

From now on we'll use the notations  $T(v)$  and  $Tv$  interchangeably. Most textbooks do this already, but it's worth mentioning again that they mean the same thing to clear up any confusion.

The set of all linear maps from  $V$  to  $W$  is denoted  $\mathcal{L}(V, W)$ . The set of all linear maps from a vector space  $V$  to itself is denoted  $\mathcal{L}(V) = \mathcal{L}(V, V)$  and is called the set of all linear *operators* on  $V$ . Linear operators are a very important special case of linear transformations that we'll look into later.

Once again, let  $T \in \mathcal{L}(V, W)$  be a linear map. Let  $v \in V$  be a vector and let  $b_1, b_2, \dots, b_n$  be a basis of  $V$ . Remember from Chapter 1 that we can write

$$v = a_1b_1 + a_2b_2 + \dots + a_nb_n$$

for some  $a_1, a_2, \dots, a_n \in \mathbb{R}$ . By the properties of linear transformations that we looked at above, this implies that

$$Tv = T(a_1b_1 + a_2b_2 + \dots + a_nb_n) = a_1Tb_1 + a_2Tb_2 + \dots + a_nTb_n.$$

This is huge. This fact is telling us that a linear map  $T : V \rightarrow W$  is *uniquely* determined by where it maps the basis vectors. Better yet, we can *define* a linear transformation simply by defining where the transformation maps the basis vectors. Since any other vector in the vector space can be written uniquely in terms of the basis vectors, its image under the transformation must be uniquely determined as well. It's that easy.

## Linear Transformations as Vector Spaces

Let  $S, T \in \mathcal{L}(V, W)$  be two linear maps. We can make  $\mathcal{L}(V, W)$  into a vector space by defining addition and scalar multiplication on it. Let's define  $S + T \in \mathcal{L}(V, W)$  to be

$$(S + T)v = Sv + Tv$$

for  $v \in V$  and scalar multiplication to be

$$(\alpha T)v = \alpha(Tv)$$

for  $\alpha \in \mathbb{R}$ . With these two definitions alone, we can now think of  $\mathcal{L}(V, W)$  as a vector space. Going a step further, we can define the *product* of two linear maps as well. First let  $U$  be a vector space. Then for  $T \in \mathcal{L}(U, V)$  and  $S \in \mathcal{L}(V, W)$  we have

$$(ST)u = S(Tu)$$

for  $u \in U$ . In other words,  $ST$  is just the usual composition  $S \circ T$ . When  $S$  and  $T$  are both linear, so is  $ST$ . Note that linear map multiplication is associative and distributive, but *not* commutative. In general,  $ST \neq TS$ .

## Null Spaces and Ranges

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

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

The null space is also called the *kernel*. As an example, suppose that  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is the map that shifts a two-dimensional vector backward:

$$T((x_1, x_2)^T) = (x_2, 0)^T.$$

Then  $\text{null } T = \{(x_1, 0)^T : x_1 \in \mathbb{R}\}$ , i.e. the null space of  $T$  is the  $x$ -axis. Observe that  $\text{null } T$  is a subspace of  $\mathbb{R}^2$ . In fact, this is always true! For any linear map  $T \in \mathcal{L}(V, W)$ , we have that  $\text{null } T \subseteq V$  is a subspace of  $V$ .

For  $T \in \mathcal{L}(V, W)$  we define the *range* of  $T$ , denoted  $\text{range } T$ , to be the subset of  $W$  consisting of vectors of the form  $Tv$  for some  $v \in V$ :

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

Sometimes the word *image* is used instead of range, but they mean the same thing. For example, if we take  $T$  to be the backwards shift map that we looked at above, then  $\text{range } T = \{(x_2, 0)^T : x_2 \in \mathbb{R}\}$ . In other words, the range of  $T$  is the  $x$ -axis. At first glance this may seem confusing because the null space of  $T$  is also the  $x$ -axis, but remember: the range of a linear map  $T$  is a subset of the *co-domain* and the null space is a subset of the *domain*. Similar to the null space, we have that  $\text{range } T \subseteq W$  is a subspace of  $W$ .

Now that we're familiar with null spaces and ranges, we can state a few facts about them. Let's look at the most important one first. For  $T \in \mathcal{L}(V, W)$  we have

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

Surprising, isn't it? At first glance it isn't clear why the dimension of the domain of a linear map should be related to the dimension of the codomain in such a simple way, but it's true. In fact, the proof of this theorem isn't very difficult! It relies solely on the properties of bases that we've looked at here.

For a sanity check, let's look at the backwards shift map again. Recall that the domain of the map was  $\mathbb{R}^2$ , so its dimension was two. In addition, we determined that the null space and the range were the  $x$ -axis, each of which was one dimensional. Since  $2 = 1 + 1$ , the theorem holds.

The reason this theorem is so useful is because we often know two of the terms in the equality, but not the other (typically we don't know the dimension of the kernel), so it allows us to quickly find the third. It also allows us to deduce additional facts about linear maps. For example, suppose that  $V$  and  $W$  are finite-dimensional vector spaces such that  $\dim V > \dim W$ . Then *no* linear map from  $V$  to  $W$  is one-to-one. This is a pretty strong statement, but its proof is remarkably straightforward given the theorem. Here it is:

*Proof.* Let  $T \in \mathcal{L}(V, W)$ . Then

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

$$\geq \dim V - \dim W$$

$$> 0.$$

The second line follows from the fact that  $\text{range } T \subseteq W$  so  $\dim \text{range } T \leq \dim W$ . Since  $\dim \text{null } T > 0$ , we know that there exist vectors other than the zero vector that get mapped to zero, so  $T$  can't be one-to-one.  $\square$

Using this theorem, we proved a very strong statement about linear transformations as a whole in just three lines. That's pretty cool.

## The Matrix of a Linear Map

Matrices: the bane of every engineering student's existence at one point or another. Up until now we've been looking at linear algebra through the lens of abstract linear transformations and, as a result, we haven't been able to work

through many concrete examples. To use what we've learned about linear maps in a real-world setting we need matrices. Let's quickly run through the basics of matrices before going any further. An  $m \times n$  matrix  $A$  is an  $m \times n$  rectangular grid of real numbers that looks like this:

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

We'll use  $a_{ij}$  to denote the entry of  $A$  in the  $i$ th row and  $j$ th column. Recall that we can add and multiply matrices (assuming the dimensions work out) and that we can multiply a vector by a matrix. Let's look at how these work. If we take  $A$  to be the matrix above and let  $v = (v_1, v_2, \dots, v_n)^T$  be a vector in  $\mathbb{R}^n$  then

$$Av = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + v_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + v_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Matrix multiplication is more cumbersome to write out completely, but if  $A$  is an  $m \times n$  matrix and  $B$  is an  $n \times k$  matrix then  $AB$  is an  $m \times k$  matrix whose entry in the  $i$ th row and  $j$ th column is

$$ab_{ij} = \sum_{k=1}^n a_{ik}b_{kj}.$$

So that's how matrix multiplication works. It's worth working through a few examples on one's own because the formulas make it look more complicated than it is.

This brings us to the meat of this section: linear transformations as matrices. Remember from the first section of this chapter that if  $v \in V$  was a vector in a vector space  $V$ , if  $b_1, b_2, \dots, b_n$  was a basis of  $V$ , and if  $T : V \rightarrow W$  was a linear map from  $V$  to another vector space  $W$  then

$$Tv = T(v_1b_1 + v_2b_2 + \cdots + v_nb_n) = v_1Tb_1 + v_2Tb_2 + \cdots + v_nTb_n$$

for some  $v_1, v_2, \dots, v_n \in \mathbb{R}$ .

Now compare this last expression to the one right above for the product  $Av$ . They look almost exactly the same!

That's because they are. If we construct a matrix  $A$  such that  $k$ th column of  $A$  is equal to  $Tb_k$ , the image of the  $k$ th basis vector of  $V$  under  $T$ , then the expression

$$Av = v_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + v_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + v_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

corresponds exactly to the vector  $Tv$  (with respect to this particular basis).

This is a big deal. It means that there's a *one-to-one correspondence* between matrices and linear transformations with respect to a given basis. Given a linear transformation  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  and a specified basis  $b_1, b_2, \dots, b_n$ , there exists a  $n \times m$  matrix  $M(T)$  that corresponds precisely to  $T$ . Similarly, given a  $n \times m$  matrix  $A$  there exists some linear transformation  $L(A) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  that corresponds precisely to  $A$ . Note that this implies that  $\dim \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = nm$  and that, more generally,  $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$ . All of this means that matrices and linear maps are essentially interchangeable and that there is no meaningful difference between the two when working through problems. A good way to think about it this: Linear maps are the fundamental objects of study of linear algebra and matrices are the tools we use to study them.

One more thing: a linear transformation, or its corresponding matrix, is always defined *with respect to a basis*. This is crucial to understand when learning about linear maps and matrices, but people often overlook it. Since there are many choices of a basis for a given vector space, there are just as many choices of the *representation* of a given linear map. In other words, two different matrices may correspond to the same linear map if they are with respect to two different bases. For a particular basis, though, there is only one matrix representation per linear map and vice versa. Unless stated otherwise, all linear maps and matrices that we'll look at here will be with respect to the standard basis  $e_1, e_2, \dots, e_n$ .

## Invertibility

This brings us to the last section of this chapter: invertibility. A linear map  $T \in \mathcal{L}(V, W)$  is said to be *invertible* if there exists another linear map  $S \in \mathcal{L}(W, V)$  such that  $TS$  is the identity map on  $V$  and  $ST$  is the identity map on  $W$ . We call  $S$  the *inverse* of  $T$  and we often denote it by  $T^{-1}$ .

We won't prove it here, but an important fact about inverse maps is that they're unique, i.e. if  $S$  and  $S'$  are inverses of  $T$  then  $S = S'$ . Another important fact is that a linear map  $T$  is invertible if and only if it is both one-to-one and onto. This actually implies that  $\dim V = \dim W$  must hold for  $T$  to be invertible as well.

We say that two vector spaces  $V$  and  $W$  are *isomorphic*, denoted  $V \cong W$ , if there exists an invertible linear map from one onto the other. Intuitively, this means that  $V$  and  $W$  are the "same" vector space up to renaming of the vectors. In other words, if you took each vector  $v \in V$  and gave it a new name, but still preserved the relationships between the vectors, then that new renamed vector space could be called  $W$ . Isomorphisms pop up pretty frequently in math when showing that two vector spaces are of the same size or are essentially the same, so it's good to become familiar with them early on.

## Summary