

# Vectors, mappings, and matrices

We discuss Vectors, mappings, and matrices

In real life, there is most often more than one variable. We wish to organize dealing with multiple variables in a consistent manner, and in particular organize dealing with linear equations and linear mappings, as those both rather useful and rather easy to handle. Mathematicians joke that “to an engineer every problem is linear, and everything is a matrix.” And well, they (the engineers) are not wrong. Quite often, solving an engineering problem is figuring out the right finite-dimensional linear problem to solve, which is then solved with some matrix manipulation. Most importantly, linear problems are the ones that we know how to solve, and we have many tools to solve them. For engineers, mathematicians, physicists, and anybody in a technical field it is absolutely vital to learn linear algebra.

As motivation, suppose we wish to solve

$$\begin{aligned}x - y &= 2, \\ 2x + y &= 4,\end{aligned}$$

for  $x$  and  $y$ , that is, find numbers  $x$  and  $y$  such that the two equations are satisfied. Let us perhaps start by adding the equations together to find

$$x + 2x - y + y = 2 + 4, \quad \text{or} \quad 3x = 6.$$

In other words,  $x = 2$ . Once we have that, we plug in  $x = 2$  into the first equation to find  $2 - y = 2$ , so  $y = 0$ . OK, that was easy. What is all this fuss about linear equations. Well, try doing this if you have 5000 unknowns<sup>1</sup>. Also, we may have such equations not of just numbers, but of functions and derivatives of functions in differential equations. Clearly we need a more systematic way of doing things. A nice consequence of making things systematic and simpler to write down is that it becomes easier to have computers do the work for us. Computers are rather stupid, they do not think, but are very good at doing lots of repetitive tasks precisely, as long as we figure out a systematic way for them to perform the tasks.

## Vectors and operations on vectors

Consider  $n$  real numbers as an  $n$ -tuple:

$$(x_1, x_2, \dots, x_n).$$

The set of such  $n$ -tuples is the so-called  $n$ -dimensional space, often denoted by  $\mathbb{R}^n$ . Sometimes we call this the  $n$ -dimensional *euclidean space*<sup>2</sup>. In two dimensions,  $\mathbb{R}^2$  is called the *cartesian plane*<sup>3</sup>, and in three dimensions, it is the same “3-dimensional space” that is dealt with in multivariable calculus. Each such  $n$ -tuple represents a point in the  $n$ -dimensional space. For example, the point  $(1, 2)$  in the plane  $\mathbb{R}^2$  is one unit to the right and two units up from the origin.

When we do algebra with these  $n$ -tuples of numbers we call them *vectors*<sup>4</sup>. Mathematicians are keen on separating what is a vector and what is a point of the space or in the plane, and it turns out to be an important distinction, however, for the purposes of linear algebra we can think of everything being represented by a vector. A way to think of a vector, which is especially useful in calculus and differential equations, is an arrow. It is an object that has a *direction* and a *magnitude*. For example, the vector  $(1, 2)$  is the arrow from the origin to the point  $(1, 2)$  in the plane. The magnitude is the length of the arrow. See **Normally a reference to a previous figure goes here.** If we think of vectors as arrows, the arrow doesn’t always have to start at the origin. If we do move it around, however, it should always keep the same direction and the same magnitude.

As vectors are arrows, when we want to give a name to a vector, we draw a little arrow above it:

$$\vec{x}$$

---

Learning outcomes: Express  $n$ -tuples of numbers as vectors Perform operations on vectors Understand how linear maps on vectors give rise to matrices.

Author(s): Matthew Charnley and Jason Nowell

<sup>1</sup>One of the downsides of making everything look like a linear problem is that the number of variables tends to become huge.

<sup>2</sup>Named after the ancient Greek mathematician **Euclid of Alexandria** (around 300 BC), possibly the most famous of mathematicians; even small towns often have Euclid Street or Euclid Avenue.

<sup>3</sup>Named after the French mathematician **René Descartes** (1596–1650). It is “cartesian” as his name in Latin is Renatus Cartesius.

<sup>4</sup>A common notation to distinguish vectors from points is to write  $(1, 2)$  for the point and  $\langle 1, 2 \rangle$  for the vector. We write both as  $(1, 2)$ .

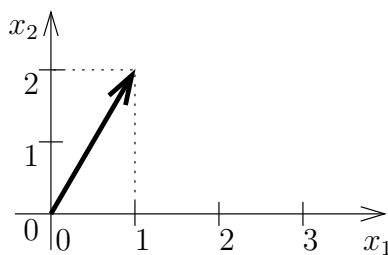


Figure 1: The vector  $(1,2)$  drawn as an arrow from the origin to the point  $(1,2)$ .

Another popular notation is  $\mathbf{x}$ , although we will use the little arrows. It may be easy to write a bold letter in a book, but it is not so easy to write it by hand on paper or on the board. Mathematicians often don't even write the arrows. A mathematician would write  $x$  and just remember that  $x$  is a vector and not a number. Just like you remember that Bob is your uncle, and you don't have to keep repeating "Uncle Bob" and you can just say "Bob." In this book, however, we will call Bob "Uncle Bob" and write vectors with the little arrows.

The *magnitude* can be computed using Pythagorean theorem. The vector  $(1,2)$  drawn in the figure has magnitude  $\sqrt{1^2 + 2^2} = \sqrt{5}$ . The magnitude is denoted by  $\|\vec{x}\|$ , and, in any number of dimensions, it can be computed in the same way:

$$\|\vec{x}\| = \|(x_1, x_2, \dots, x_n)\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

For reasons that will become clear in the next section, we often write vectors as so-called *column vectors*:

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

Don't worry. It is just a different way of writing the same thing, and it will be useful later. For example, the vector  $(1,2)$  can be written as

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

The fact that we write arrows above vectors allows us to write several vectors  $\vec{x}_1$ ,  $\vec{x}_2$ , etc., without confusing these with the components of some other vector  $\vec{x}$ .

So where is the *algebra* from *linear algebra*? Well, arrows can be added, subtracted, and multiplied by numbers. First we consider *addition*. If we have two arrows, we simply move along one, and then along the other. See **Normally a reference to a previous figure goes here..**

It is rather easy to see what it does to the numbers that represent the vectors. Suppose we want to add  $(1,2)$  to  $(2,-3)$  as in the figure. So we travel along  $(1,2)$  and then we travel along  $(2,-3)$  in the sense of "tip-to-tail" addition that you may have seen in previous classes. What we did was travel one unit right, two units up, and then we travelled two units right, and three units down (the negative three). That means that we ended up at  $(1+2, 2+(-3)) = (3, -1)$ . And that's how addition always works:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

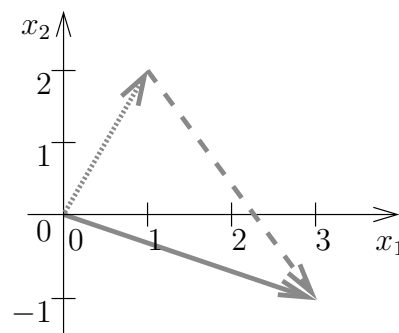


Figure 2: Adding the vectors  $(1,2)$ , drawn dotted, and  $(2,-3)$ , drawn dashed. The result,  $(3,-1)$ , is drawn as a solid arrow.

*Subtracting* is similar. What  $\vec{x} - \vec{y}$  means visually is that we first travel along  $\vec{x}$ , and then we travel backwards along  $\vec{y}$ . See **Normally a reference to a previous figure goes here..** It is like adding  $\vec{x} + (-\vec{y})$  where  $-\vec{y}$  is the arrow we obtain by erasing the arrow head from one side and drawing it on the other side, that is, we reverse the direction. In terms of the numbers, we simply go backwards in both directions, so we negate both numbers. For example, if  $\vec{y}$  is  $(-2, 1)$ , then  $-\vec{y}$  is  $(2, -1)$ .

Another intuitive thing to do to a vector is to *scale* it. We represent this by multiplication of a number with a vector. Because of this, when we wish to distinguish between vectors and numbers, we call the numbers *scalars*. For example, suppose we want to travel three times further. If the vector is  $(1, 2)$ , traveling 3 times further means going 3 units to the right and 6 units up, so we get the vector  $(3, 6)$ . We just multiply each number in the vector by 3. If  $\alpha$  is a number, then

$$\alpha \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}.$$

Scaling (by a positive number) multiplies the magnitude and leaves direction untouched. The magnitude of  $(1, 2)$  is  $\sqrt{5}$ . The magnitude of 3 times  $(1, 2)$ , that is,  $(3, 6)$ , is  $3\sqrt{5}$ .

When the scalar is negative, then when we multiply a vector by it, the vector is not only scaled, but it also switches direction. So multiplying  $(1, 2)$  by  $-3$  means we should go 3 times further but in the opposite direction, so 3 units to the left and 6 units down, or in other words,  $(-3, -6)$ . As we mentioned above,  $-\vec{y}$  is a reverse of  $\vec{y}$ , and this is the same as  $(-1)\vec{y}$ .

In **Normally a reference to a previous figure goes here..**, you can see a couple of examples of what scaling a vector means visually.

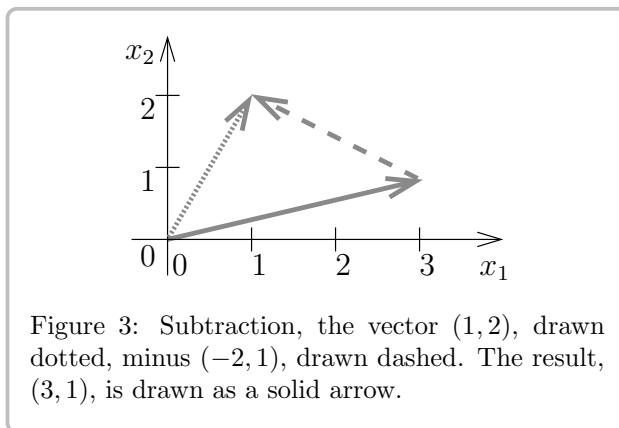


Figure 3: Subtraction, the vector  $(1, 2)$ , drawn dotted, minus  $(-2, 1)$ , drawn dashed. The result,  $(3, 1)$ , is drawn as a solid arrow.

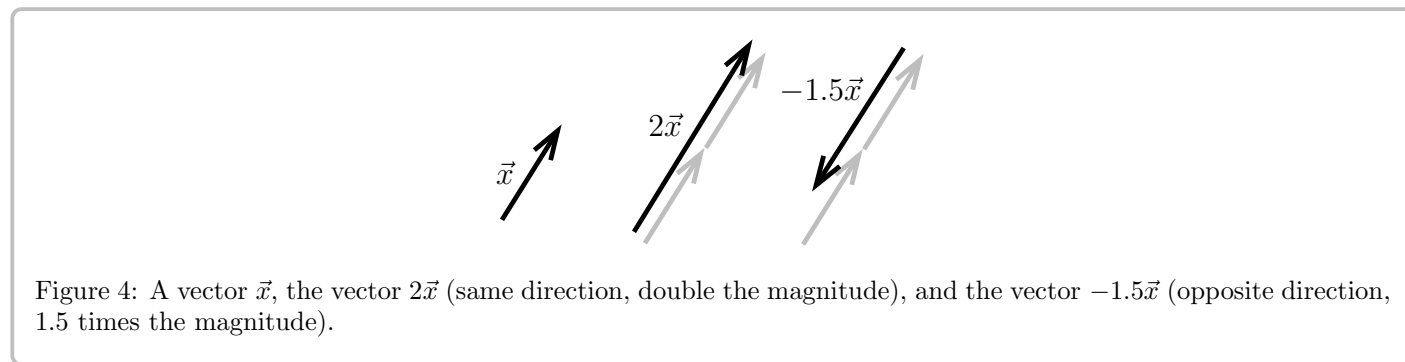


Figure 4: A vector  $\vec{x}$ , the vector  $2\vec{x}$  (same direction, double the magnitude), and the vector  $-1.5\vec{x}$  (opposite direction, 1.5 times the magnitude).

We put all of these operations together to work out more complicated expressions. Let us compute a small example:

$$3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} -4 \\ -1 \end{bmatrix} - 3 \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 3(1) + 2(-4) - 3(-2) \\ 3(2) + 2(-1) - 3(2) \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

As we said a vector is a direction and a magnitude. Magnitude is easy to represent, it is just a number. The *direction* is usually given by a vector with magnitude one. We call such a vector a *unit vector*. That is,  $\vec{u}$  is a unit vector when  $\|\vec{u}\| = 1$ .

For example, the vectors  $(1, 0)$ ,  $(1/\sqrt{2}, 1/\sqrt{2})$ , and  $(0, -1)$  are all unit vectors. To represent the direction of a vector  $\vec{x}$ , we need to find the unit vector in the same direction. To do so, we simply rescale  $\vec{x}$  by the reciprocal of the magnitude, that is  $\frac{1}{\|\vec{x}\|}\vec{x}$ , or more concisely  $\frac{\vec{x}}{\|\vec{x}\|}$ .

For example, the unit vector in the direction of  $(1, 2)$  is the vector

$$\frac{1}{\sqrt{1^2 + 2^2}}(1, 2) = \left( \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right).$$

## Linear Combinations and Linear Independence

If we have vectors of a given size (say a collection of 3 component vectors), there are only two different operations we can perform on them; adding them together and multiplying them by real numbers. While this may seem fairly limited, we can take a small number of vectors and generate a lot more vectors from them. By putting these operations together, we get the definition of a linear combination.

**Definition 1.** Given row or column vectors  $\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n$ , a linear combination is an expression of the form

$$\alpha_1 \vec{y}_1 + \alpha_2 \vec{y}_2 + \dots + \alpha_n \vec{y}_n,$$

where  $\alpha_1, \alpha_2, \dots, \alpha_n$  are all scalars.

For example,  $3\vec{y}_1 + \vec{y}_2 - 5\vec{y}_3$  is a linear combination of  $\vec{y}_1$ ,  $\vec{y}_2$ , and  $\vec{y}_3$ . Another example is that  $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$  can be written as a linear combination of  $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ ,  $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ , and  $\begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$  as

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

Given any collection of vectors  $\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n$ , it is always possible to write  $\vec{0}$  as a linear combination of these vectors by choosing  $\alpha_i = 0$  for all  $i$ . Then we are adding a bunch of zeros together and so get zero. However, this may not be the only way to get zero. Whether or not it is depends heavily on the exact vectors involved.

**Definition 2.** We say the vectors  $\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n$  are linearly independent if the only way to pick  $\alpha_1, \alpha_2, \dots, \alpha_n$  to satisfy

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_n \vec{x}_n = \vec{0}$$

is  $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$ . Otherwise, we say the vectors are linearly dependent.

**Example 1.** The vectors  $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$  and  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$  are linearly independent.

**Solution:** Let's try:

$$\alpha_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ 2\alpha_1 + \alpha_2 \end{bmatrix} = \vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

So  $\alpha_1 = 0$ , and then it is clear that  $\alpha_2 = 0$  as well. In other words, the vectors are linearly independent. ┘

If a set of vectors is linearly dependent, that is, we have an expression of the form

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_n \vec{x}_n = \vec{0}$$

with some of the  $\alpha_j$ 's are nonzero, then we can solve for one vector in terms of the others. Suppose  $\alpha_1 \neq 0$ . Since  $\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_n \vec{x}_n = \vec{0}$ , then

$$\vec{x}_1 = \frac{-\alpha_2}{\alpha_1} \vec{x}_2 + \frac{-\alpha_3}{\alpha_1} \vec{x}_3 + \dots + \frac{-\alpha_n}{\alpha_1} \vec{x}_n.$$

For example,

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

and so

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

In particular, this tells us that if two vectors  $\vec{y}_1$  and  $\vec{y}_2$  are linearly dependent, then it must be the case that  $\vec{y}_1 = A\vec{y}_2$  for some constant  $A$ , namely  $A = \frac{-\alpha_2}{\alpha_1}$  where  $\alpha_1 \vec{y}_1 + \alpha_2 \vec{y}_2 = \vec{0}$ . If this is not the case, then the vectors are linearly independent. For more than two vectors, the process is more complicated and involves solving a system of linear equations, which we will deal with in § ??.

## Matrices

The next object we need to define here is a *matrix*.

**Definition 3.** In general, an  $m \times n$  matrix  $A$  is a rectangular array of  $mn$  numbers,

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

An  $m \times n$  matrix indicates that it will have  $m$  rows and  $n$  columns.

Matrices, just like vectors, are generally written with square brackets on the outside, although some books will use parentheses for this. The convention for notation is that matrices will be denoted by capital letters ( $A$ ) and the individual *entries* of the matrix, the numbers that make it up, will be denoted using lowercase letters ( $a_{ij}$ ) where the first number  $i$  indicates which row of the matrix we are talking about, and the second number  $j$  indicates which column. For example, in the matrix

$$A = \begin{bmatrix} 1 & 4 & 0 \\ -2 & 3 & 1 \\ 2 & 0 & 5 \end{bmatrix},$$

we could talk about the entire matrix using  $A$ , but would also have that  $a_{21} = -2$  and  $a_{33} = 5$ .

Note that an  $m \times 1$  matrix is just a column vector, so in terms of the basic structure, matrices are an extension of vectors. However, they can be used for so much more, as we will see in future sections.

Another way to view matrices is as a set of column vectors all laid out side-by-side. If we have  $\vec{v}_1$ ,  $\vec{v}_2$  and  $\vec{v}_3$ , three different four component vectors, we can form a  $4 \times 3$  matrix  $B$  as

$$B = [\vec{v}_1 \mid \vec{v}_2 \mid \vec{v}_3]$$

that uses each of the given vectors as a column of the matrix. In this case, the vertical lines are used to indicate that this is actually a matrix, because each of the entries given there are vectors, not just individual numbers. If we wanted to write a  $1 \times 3$  matrix this way, these vertical lines will not be included.

We will go into more properties of matrices and the operations we can perform on them in § ???. To conclude this section though, we will look at one other way that matrices come about, and that is as the representation of a linear map.

## Linear mappings and matrices

A *vector-valued function*  $F$  is a rule that takes a vector  $\vec{x}$  and returns another vector  $\vec{y}$ . For example,  $F$  could be a scaling that doubles the size of vectors:

$$F(\vec{x}) = 2\vec{x}.$$

For example,

$$F\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right) = 2\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

If  $F$  is a mapping that takes vectors in  $\mathbb{R}^2$  to  $\mathbb{R}^2$  (such as the above), we write

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^2.$$

The words *function* and *mapping* are used rather interchangeably, although more often than not, *mapping* is used when talking about a vector-valued function, and the word *function* is often used when the function is scalar-valued.

A beginning student of mathematics (and many a seasoned mathematician), that sees an expression such as

$$f(3x + 8y)$$

yearns to write

$$3f(x) + 8f(y).$$

After all, who hasn't wanted to write  $\sqrt{x+y} = \sqrt{x} + \sqrt{y}$  or something like that at some point in their mathematical lives. Wouldn't life be simple if we could do that? Of course we can't always do that (for example, not with the square roots!) It turns out there are many functions where we can do exactly the above. Such functions are called *linear*.

A mapping  $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$  is called *linear* if

$$F(\vec{x} + \vec{y}) = F(\vec{x}) + F(\vec{y}),$$

for any vectors  $\vec{x}$  and  $\vec{y}$ , and also

$$F(\alpha\vec{x}) = \alpha F(\vec{x}),$$

for any scalar  $\alpha$ . The  $F$  we defined above that doubles the size of all vectors is linear. Let us check:

$$F(\vec{x} + \vec{y}) = 2(\vec{x} + \vec{y}) = 2\vec{x} + 2\vec{y} = F(\vec{x}) + F(\vec{y}),$$

and also

$$F(\alpha\vec{x}) = 2\alpha\vec{x} = \alpha 2\vec{x} = \alpha F(\vec{x}).$$

We also call a linear function a *linear transformation*. If you want to be really fancy and impress your friends, you can call it a *linear operator*.

When a mapping is linear we often do not write the parentheses. We write simply

$$F\vec{x}$$

instead of  $F(\vec{x})$ . We do this because linearity means that the mapping  $F$  behaves like multiplying  $\vec{x}$  by “something.” That something is a matrix.

Now how does a matrix  $A$  relate to a linear mapping? Well a matrix tells you where certain special vectors go. Let’s give a name to those certain vectors. The *standard basis vectors* of  $\mathbb{R}^n$  are

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

For example, in  $\mathbb{R}^3$  these vectors are

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

You may recall from calculus of several variables that these are sometimes called  $\vec{i}$ ,  $\vec{j}$ ,  $\vec{k}$ .

The reason these are called a *basis* is that every other vector can be written as a *linear combination* of them. For example, in  $\mathbb{R}^3$  the vector  $(4, 5, 6)$  can be written as

$$4\vec{e}_1 + 5\vec{e}_2 + 6\vec{e}_3 = 4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 5 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 6 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

Keep this idea of linear combinations of vectors in mind; we’ll see a lot more of it later.

So how does a matrix represent a linear mapping? Well, the columns of the matrix are the vectors where  $A$  as a linear mapping takes  $\vec{e}_1$ ,  $\vec{e}_2$ , etc. For example, consider

$$M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

As a linear mapping  $M: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  takes  $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$  to  $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$  and  $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  to  $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ . In other words,

$$M\vec{e}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \text{and} \quad M\vec{e}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

More generally, if we have an  $n \times m$  matrix  $A$ , that is we have  $n$  rows and  $m$  columns, then the mapping  $A: \mathbb{R}^m \rightarrow \mathbb{R}^n$  takes  $\vec{e}_j$  to the  $j^{\text{th}}$  column of  $A$ . For example,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{bmatrix}$$

represents a mapping from  $\mathbb{R}^5$  to  $\mathbb{R}^3$  that does

$$A\vec{e}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}, \quad A\vec{e}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix}, \quad A\vec{e}_3 = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}, \quad A\vec{e}_4 = \begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix}, \quad A\vec{e}_5 = \begin{bmatrix} a_{15} \\ a_{25} \\ a_{35} \end{bmatrix}.$$

But what if I have another vector  $\vec{x}$ ? Where does it go? Well we use linearity. First write the vector as a linear combination of the standard basis vectors:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 + x_4\vec{e}_4 + x_5\vec{e}_5.$$

Then

$$A\vec{x} = A(x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 + x_4\vec{e}_4 + x_5\vec{e}_5) = x_1A\vec{e}_1 + x_2A\vec{e}_2 + x_3A\vec{e}_3 + x_4A\vec{e}_4 + x_5A\vec{e}_5.$$

If we know where  $A$  takes all the basis vectors, we know where it takes all vectors.

As an example, suppose  $M$  is the  $2 \times 2$  matrix from above, and suppose we wish to find

$$M \begin{bmatrix} -2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 0.1 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 0.1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -1.8 \\ -5.6 \end{bmatrix}.$$

Every linear mapping from  $\mathbb{R}^m$  to  $\mathbb{R}^n$  can be represented by an  $n \times m$  matrix. You just figure out where it takes the standard basis vectors. Conversely, every  $n \times m$  matrix represents a linear mapping. Hence, we may think of matrices being linear mappings, and linear mappings being matrices.

Or can we? In this book we study mostly linear differential operators, and linear differential operators are linear mappings, although they are not acting on  $\mathbb{R}^n$ , but on an infinite-dimensional space of functions:

$$Lf = g$$

for a function  $f$  we get a function  $g$ , and  $L$  is linear in the sense that

$$L(f + h) = Lf + Lh, \quad \text{and} \quad L(\alpha f) = \alpha Lf.$$

for any number (scalars)  $\alpha$  and all functions  $f$  and  $h$ .

So the answer is not really. But if we consider vectors in finite-dimensional spaces  $\mathbb{R}^n$  then yes, every linear mapping is a matrix. We have mentioned at the beginning of this section, that we can “make everything a vector.” That’s not strictly true, but it is true approximately. Those “infinite-dimensional” spaces of functions can be approximated by a finite-dimensional space, and then linear operators are just matrices. So approximately, this is true. And as far as actual computations that we can do on a computer, we can work only with finitely many dimensions anyway. If you ask a computer or your calculator to plot a function, it samples the function at finitely many points and then connects the dots<sup>5</sup>. It does not actually give you infinitely many values. So the way that you have been using the computer or your calculator so far has already been a certain approximation of the space of functions by a finite-dimensional space.

<sup>5</sup>In Matlab, you may have noticed that to plot a function, we take a vector of inputs, ask Matlab to compute the corresponding vector of values of the function, and then we ask it to plot the result.