The Elements of Linear Algebra
Version 2018S

Theron J. Hitchman

January 19, 2018

Introduction: To The Student

So. I am writing this book. (Let's not pretend I am done. You can see the state it is in.) There are dozens (hundreds?) of introductory linear algebra books, so it is pretty reasonable to ask why I am putting in the effort, and, in the meantime, causing this much pain. I should explain.

I have taught linear algebra many times and I have liked some books, but never loved one. The closest match for what I wanted to teach is Strang's *Introduction to Linear Algebra*, and I am sure that people who have read that will see some influences here. But my students never seemed to connect with Prof. Strang's enthusiastic, stream-of-consciousness prose. And over time, I found that the things I need to emphasize for my students just don't match with that text, or any other.

In addition, most textbooks assume a certain class structure: lectures accompanied by weekly homework, with some exams. I don't want to run our course that way.

So this book is my solution. It is my attempt to make a thing which matches how I want our class to run.

It is important to read this book actively. If you haven't learned how to read a math text before, there are some key ideas:

Time Mathematics is often technical and tricky. It takes time to absorb. Plan to give yourself lots of time to read and think. And don't be surprised if you have to read some section more than once. (This is not a novel. As much as I see it as a story, it won't sweep you away.)

Examples In the interest of brevity, I have streamlined the exposition. In particular, there are no examples. **The point is that you should make your own.** This is so important a skill that it is basically a mathematical super-power. Whenever you come across an idea, if you understand it or not, you should make some very explicit examples and consider them carefully.

Questions As part of your *active* engagement with the text, you will find things that don't quite make sense, yet. This is normal. The mathematician's best approach then is to (1) write down a specific question or two about the confusing bit, and (2) talk to other people about it. You are fortunate that you have an instructor and classmates to talk to. Make lists of questions and try to get them answered!

The real beauty in linear algebra is the tight set of connections between algebra and geometry. I hope you enjoy it.

Chapter 1

Orientation and Preliminaries

1.1 The Three Viewpoints, Algebraically

Linear algebra is about solving systems of linear equations. It is also about the geometry of vectors, and it is also about matrices and transformations of one space into another. Somehow, it is all of those things all at the same time, because those are one and the same. Linear algebra is like a gem with many facets, and the beauty of the subject comes out when you learn how to see the light it casts in different directions.

Let's preview the most basic linear algebra situation and three useful geometric ways to look at it.

A Start: the Row Picture

Consider the following three examples. Each of the first two is a *system of two linear equations in two unknowns*. The third is a *system of three equations in two unknowns*. Because they are familiar, we are using x and y as the names of our unknowns, but we could use other symbols.

$$\left\{ \begin{array}{rcrcr} x & + & 2y & = & 3 \\ x & + & y & = & 1 \end{array} \right. \tag{A}$$

$$\left\{ \begin{array}{rcrcr} x & + & 2y & = & 3 \\ 3x & + & 6y & = & 4 \end{array} \right. \tag{B}$$

$$\left\{\begin{array}{rcrcl} x & + & 2y & = & 3 \\ x & + & y & = & 1 \\ 3x & + & 6y & = & 4 \end{array}\right. \tag{C}$$

These are typical examples, though they are "small." In applications, a system of linear equations might have thousands of equations, thousands of unknowns, or both!

Reading Exercise. *Make a few other examples of systems of linear equations. How are your examples different from those above? How are they the same?*

Clearly we are looking at equations, so we can ask the kinds of questions that mathematicians always ask when they see equations:

1. Does the system have any solutions at all? Can we decide before we do a lot of work?

2. If it has solutions, how many solutions are there? Is there a way to tell when the system will have a single unique solution?

3. What is the collection of all solutions? Is there a good way to understand this collection geometrically?

4. Is there a good algorithm for finding solutions? That is, is there an easily-followed-and-not-too-slow computational process for solving the system? Can we make a computer do the work for us?

5. If there is not a solution to the system, can we find an approximate solution? Is that a reasonable thing to compute? What should "approximate solution" mean in this setting?

Our first goal is to learn to answer these questions for any system of linear equations.

Reading Exercise. *Can you answer any of the questions in our list for any of the examples above? Or for any of your new examples?*

Consider how each of these systems is organized by being lined up in rows. Each row is a single equation, and that equation defines a collection of points in the xy-plane. Since the equation is linear, the collection happens to make a line. (That is why we call them linear equations.) So, we can make a geometric model of each system that involves the interaction of some

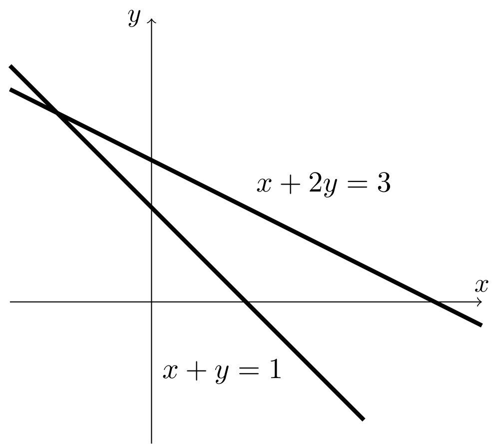


Figure 1.1: The row picture for system (A)

lines in the xy-plane. This geometric model is the *row picture* view of linear algebra.

Here is the row picture for the system (A). In this, we have labeled the two lines with their corresponding equations. The fact that these lines appear to meet gives us a way to talk about the possible solutions to the system (A).

Reading Exercise. *Make row pictures for the examples you designed above. Can you use these pictures to address any of our questions for your examples?*

A Second Look: the Column Picture

Notice that each of our systems is carefully set on the page so that the unknowns line up in columns, too. We can use that! Let's agree to bundle the coefficients together in columns so that they become objects. Just put each column of coefficients in a set of parentheses. We call these vertical-stacks-of-numbers objects *vectors*, and we reorganize each of our systems into a *linear combination of vectors equation.* The three examples above get reworked to look like this:

$$x\begin{pmatrix}1\\1\end{pmatrix}+y\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}3\\1\end{pmatrix} \tag{A}$$

$$x\begin{pmatrix}1\\3\end{pmatrix}+y\begin{pmatrix}2\\6\end{pmatrix}=\begin{pmatrix}3\\4\end{pmatrix} \tag{B}$$

$$x \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} + y \begin{pmatrix} 2 \\ 1 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \tag{C}$$

First, we will have to make sense of the algebra of combining vectors like this. This will become the concept of *linear combination*. (We'll come back to the details, soon.) Note that in the first two examples, the vectors all have two entries, but in the third example, the vectors have three entries. These entries are called *coordinates*. Sometimes we talk about the *shape* of a vector, by which we mean the number of coordinates that vector has.

Then we can make a different sort of picture, as *column picture*. Let us do this for system B. We interpret a vector $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ with two components as an arrow which goes from the origin to the point in the plane with Cartesian coordinates (a, b). Drawing all three of the vectors from our situation at once, we have the column picture for system (B).

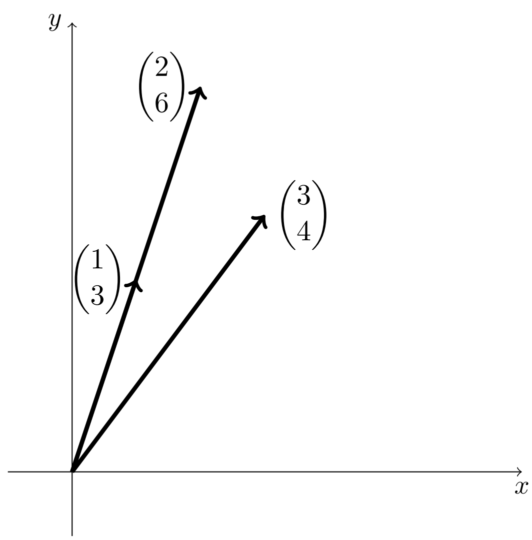


Figure 1.2: The column picture for system (B)

Notice that the two vectors from the left-hand side of our equation point in the same direction, but the one from the right-hand side points in another direction. This can help us reason about our equation.

Reading Exercise. *Translate your examples of systems of equations into the form of a linear combination equation. Can you also draw their column pictures?*

Often, a linear combination of vectors equation is written in this compact form

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = w,$$

where the symbols $a_1, \ldots, a_k$ represent numbers, and the symbols $v_1, \ldots, v_k$ and w represent vectors all having the same shape. This is visually much simpler than the full system.

A Third Look: Matrices and Transformations

As we passed from the system of linear equations to the linear combination of vectors equation, we managed to clean up our representation by getting rid of some messy duplication of structural symbols. Now, we will do it again. We take advantage of the alignment of rows and columns at the same time, and rewrite our equations like this, as *matrix-vector equations*:

$$\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \tag{A}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} \tag{B}$$

$$\begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \tag{C}$$

In each case, we have put the two unknowns into a vector, $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$. The new objects we have introduced are two-dimensional arrays of numbers, called *matrices*. Note that each matrix is constructed by taking the column vectors we found in the linear combination equation and setting them next to each other as columns. The first two examples,

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 & 1 \\ 51 & 17 \end{pmatrix},$$

are 2×2 *square matrices* because they have two rows and two columns. The third example,

$$\begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 6 \end{pmatrix},$$

is a 3×2 *matrix*, because it has three rows and two columns. In general, the number of rows corresponds to the number of equations, and the number of columns corresponds to the number of unknowns.

Reading Exercise. *Translate your examples into matrix-vector equations.*

Again, we have to figure out what the algebra of something like "a matrix times a vector" should mean, and sort out the geometry of that. The idea is that the matrix represents a kind of *function*, or *transformation*, that takes in vectors of a particular kind, and then gives you back vectors of a (possibly) different kind. In our third example, the matrix takes in vectors like $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ with two components, and then gives back vectors with three components. Since we can represent vectors with two components by arrows in the plane, and vectors with three components by vectors in space, we can make a picture like Figure 1.3.

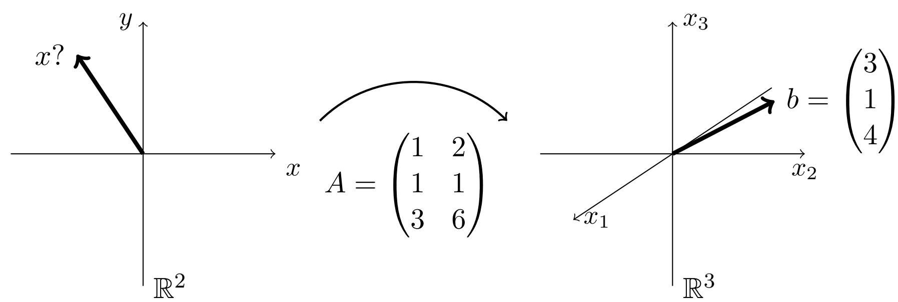


Figure 1.3: The transformational view of system (C)

Often, a matrix-vector equation is written in the ultra-compact form

$$Ax = b,$$

where A is a matrix, b is some known vector, and x is the unknown vector we seek. This form looks simple, because it hides all of the complexity in the abstract matrix and vector objects.

Exercises

We introduced many concepts and new words in this section, and most of them imprecisely. This should make you feel a bit uneasy, but we will start cleaning up by being more careful in the next section. For now, focus on the shapes and structures of the three (otherwise mysterious) algebraic representations introduced in this section.

Exercise 1. Systems of linear equations come up in all sorts of situations. Here is a typical one. When expanded out and rearranged a little, the equation of a circle in the plane takes the form

$$x^2 + y^2 + ax + by + c = 0.$$

Suppose that you have a circle which goes through the three points below. Set up the system of linear equations which should help you find the equation of your circle more exactly. (There is no need to solve the system.)

$$P = (2, 3), \qquad Q = (-4, 2), \qquad R = (7, 1)$$

How many equations do you have? How many unknowns do you have?

Translate your system of linear equations into a linear combination of vectors equation. What shape to the vectors take? Next, translate your system into a matrix-vector equation of the form $Ax = b$. What shape is the matrix A? What shape is the vector x? What shape is the vector b?

Exercise 2. Make an example of a system of two equations in three unknowns. (It's okay. Just pick something you find interesting.)

Translate your example into a linear combination of vectors equation. What shapes are your vectors? Now translate your system into a matrix-vector equation of the form $Ax = b$. What shape is the matrix A? What shape is the vector x? What shape is the vector b?

Exercise 3. Make up a matrix-vector equation $Ax = b$ that has a 4×2 matrix. That is, the matrix should have 4 rows and 2 columns.

Translate your matrix-vector equation into a linear combination of vectors equation. How many vectors does this have? What shape are they? Then translate your matrix-vector equation into a system of linear equations. How many equations do you have? How many unknowns are there?

Exercise 4. Make an example of a linear combination of vectors equation that has 5 vectors of the left-hand side of the equal sign, each of which is a stack of 3 numbers.

Translate your equation into a matrix-vector equation. What shape do all of the pieces have? Translate your equation into a system of linear equations. How many equations are there? How many unknowns are there?

Exercise 5. Think about the ideas discussed in this section. What questions do you have? What do you wonder about?

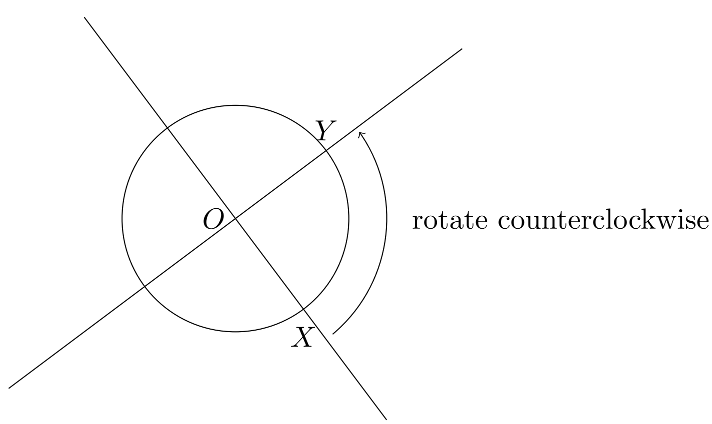
1.2 The Space $\mathbb{R}^m$: Points, Vectors, and Vectors

Our goal here is to introduce the fundamental object of linear algebra, the vector. The word has slightly different meanings depending on context, so we will to sort this out carefully first in two dimensions, where we can draw the best pictures. Then we will extend the ideas to the general situation.

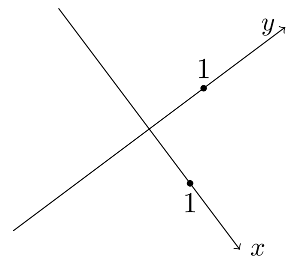
The Idea of a Vector

Let's recall the idea of *the plane* from classical geometry: the plane is like a flat sheet of drawing paper, which extends indefinitely in all directions without bound.

You have likely seen the idea of *Cartesian coordinates* on the plane before. To be clear, let's set things down carefully. In the plane, we choose a pair of perpendicular lines which meet at a point O. This special point is called the *origin*. Then, we choose a point X on one of the lines and draw the circle centered at O which passes through X. Note that this circle meets our two lines in two points each, four points total, one of which is the point X. Then we rotate from X around the circle by a quarter turn counterclockwise until we hit one of the points on the other line. We call this new point Y. Are you drawing with me? Here is my picture so far.



We call the line OX the *x-axis* and the line containing Y the *y-axis*. Here comes the amazing part: we declare the circle we used to be of *unit size*, and make the lines OX and OY into number lines! The important part is that the point O should represent 0 on both number lines, and the points X and Y should each represent 1 on their lines. So, instead of marking things with O, X, and Y, we put down marks where X and Y are and label them with 1's, and add little arrows marked with x and y near the positive "ends" of the lines OX and OY, respectively.



Note that above I have done something a bit silly and let the picture just fall on the paper in an unusual way. I really mean unusual as "not usual." The usual way arranges the lines on the paper to match our expected horizontal (x) and vertical (y) directions. This isn't actually required, but it is what everyone always does. The typical picture looks like Figure 1.4.

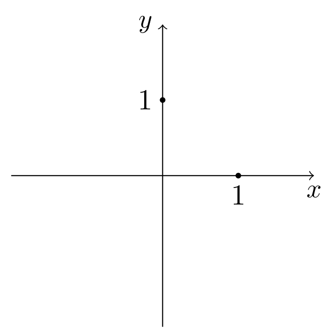


Figure 1.4: The Standard Cartesian Coordinate System

Now suppose we have some point in the plane, let's call it P. We can describe the location of P relative to our two lines in a simple way. First, we draw a line through P which is parallel to the y-axis and perpendicular to the x-axis. The foot of this perpendicular hits the x-axis at some point A. But this point A is part of the number line OX, so it has an associated real number, which we will call a. So the point A is instead marked with the label a from this number line.

Similarly, we draw a line through P parallel to the x-axis and perpendicular to the y-axis. The foot of this perpendicular hits the y-axis at some point B, which is part of the number line OY. We denote the number associated to B by b. Again, the point is labeled with the number b from the number line.

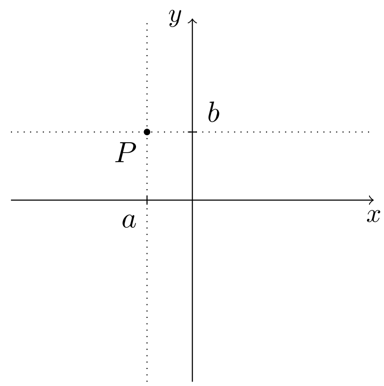


Figure 1.5: A point P and its coordinates (a, b).

So, to identify the point P, we can instead give the pair of numbers a and b. These numbers are called the *coordinates* of P. Of course, the order of the coordinates matters, so we make what we call an *ordered pair* (a, b) to keep things straight, where the x-coordinate comes first, and the y-coordinate comes second. Note that in Figure 1.5 the x-coordinate a is negative, but the y-coordinate b is postive.

This whole process is reversible, too. If we pick a pair of numbers c and d, in order, then we can find a point in the plane which corresponds, and we can do it unambiguously. Find the spot labeled c on the x-axis number line and construct a line perpendicular to the x-axis through this point. Similarly, find the spot labeled d on the y-axis and construct a line perpendicular to the y-axis through this point. The two lines you just drew will meet in exactly one point Q, and Q will have coordinates (c, d).

This setup of coordinates on the plane allows us to formalize a wonderful and useful idea from physics, too. Physicists use the concept of a *vector* to describe something (like the wind) which has both magnitude or size (like how fast the air is moving) and direction (which way the air is going). Usually, vectors are drawn as little arrows: the arrow has a direction, and it has a length which represents its magnitude. It is possible to draw vectors which have the same direction but different lengths, and vice versa.

We can use coordinates to describe vectors in the plane, too. Here's how: A physicist's vector v is some arrow in the plane. That arrow has an initial point P, called its *tail*, and a final point Q, called its *head*. We can write the coordinates of these points as $P = (a, b)$ and $Q = (c, d)$.

Then the coordinates of v are taken to be the numbers $c - a$ and $d - b$, which we interpret how much v acts in the directions parallel to the x-axis

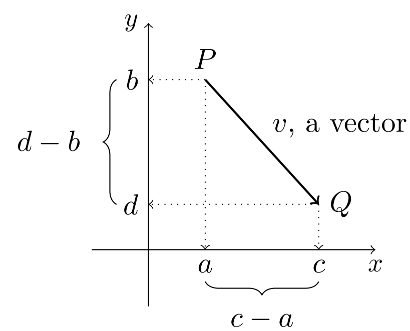


Figure 1.6: Coordinates for a physicist's vector v.

and the y-axis, respectively. Note that in Figure 1.6, the y-coordinate is negative, since $b > d$.

These coordinates have a hint of algebraic manipulation in them. Those subtractions line up almost like we could write $v = (c - a, d - b) = (c, d) - (a, b) = Q - P$. But v is a vector, and if we write it like $(c - a, d - b)$, it looks like the notation for a point. We should not do that because it could get confusing. Furthermore, that "equation" would mean that we are subtracting points and creating a vector, which is also weird. Still, there is something to it. We will return to this idea soon.

For now, let's focus on a bit of ambiguity in the physicist's idea of a vector. Where should that vector be? That is, given the coordinates of a vector in the plane, it is not clear where to draw it! We can slide a vector around the plane, and as long as we keep it parallel to the original, the coordinates won't change. So, unlike the coordinates of a point, the coordinates of a vector do not uniquely specify the vector.

The mathematician's special fix is this: we simply declare all our vectors to have their initial points, their tails, at the point O, the origin of our coordinate system. That curtails some of the (admittedly useful) freedom in the physics notion, but it also lets us be more clear.

It pays to keep in mind that the physicists conception of the vector V with coordinates $c - a$ and $d - b$ could be one of many different arrows, while the mathematician's vector v is the arrow from the point $O = (0, 0)$ to the point $(c - a, d - b)$.

Now we have circled back around to a muddle. If a mathematician's vector is always based at O, we only need to specify where the head of the vector is... which is just a point. So, how is a vector supposed to be different from a point, again?

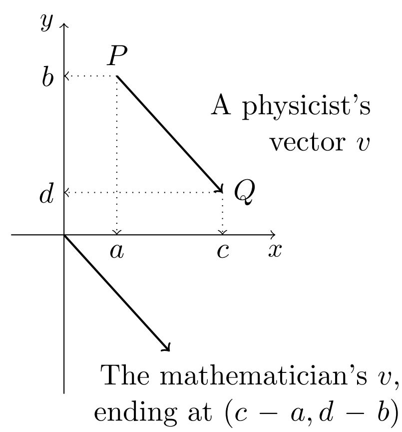


Figure 1.7: The physicist's vector vs. the mathematician's vector.

This confusion of three different, shifting, partially-overlapping interpretations causes some trouble to the new learner. Professionals tend to pass back and forth between these and use them flexibly to get results. Once you have gotten used to the ideas, you will, too. You should watch out for these instances where the words point and vector get interchanged. If they cause you trouble, remember that we have three different things, which are closely related.

For now, the simplest way to handle things is like this:

- Ignore the physicist's version of the word vector as much as possible.
- A point is a location in the plane, and represented by coordinates in the form of an ordered pair of numbers (a, b).
- A vector is an arrow based at the origin, which can be specified by the coordinates of its head. To keep this separate from the idea of a point, we will write it differently, with the numbers stacked vertically like this: $\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$.
- Always remember that for each point in the plane, there is a unique mathematician's vector which corresponds, and vice versa.

With this in mind, we make our first official definition.

Definition 1. A 2-vector is a vertical stack of 2 real numbers, like so:

$$v = \begin{pmatrix} a \\ b \end{pmatrix}.$$

The collection of all such 2-vectors is called *the plane*, and written with this notation: $\mathbb{R}^2$.

The notation $\mathbb{R}^2$ is often read "arr-two," and many people use that language interchangeably with "the plane." Also, most of the time we will just say "vector," rather than "2-vector."

Vector Algebra

Let's return to that glimpse of subtraction we saw in Figure 1.6. We saw there that for points $P = (a, b)$ and $Q = (c, d)$, the vector v from P to Q has coordinates $c - a$ and $d - b$. This looks almost like we subtracted the points to get $Q - P = v$. Can we use that? The weird part is that it mixes up points and vectors. So, we change viewpoints, and instead think of P and Q as (mathematician's) vectors. To keep things clear, let's introduce new labels.

$$p = \begin{pmatrix} a \\ b \end{pmatrix}, \quad q = \begin{pmatrix} c \\ d \end{pmatrix}$$

If we put these together on the plane with the physicist's vector v from p to q and the mathematician's v we see a wonderful triangle, and an extra vector. So we see a way to talk about subtracting vectors: Given two vectors

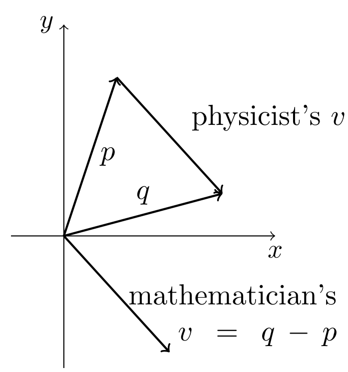


Figure 1.8: Subtraction of vectors

p and q as above, their *difference* is the vector

$$V = Q - P = \begin{pmatrix} c \\ d \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c - a \\ d - b \end{pmatrix}.$$

Geometrically, we draw the arrow from p to q and then translate it down so that its tail is at the origin $O = (0, 0)$. Of course, the order of p and q in this operation matters. If we switch them, we get an arrow pointing in the opposite direction.

If we can subtract vectors, surely we can add vectors. How would that work? Algebraically, if $v = q - p$, then we expect $q = v + p$ by rearranging. That would mean

$$q = \begin{pmatrix} c \\ d \end{pmatrix} = v + p = \begin{pmatrix} c-a \\ d-b \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix},$$

which all fits. It looks like addition should be defined coordinate-by-coordinate.

Definition 2 (Addition of Vectors)**.** Let $p = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ and $q = \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right)$ be two vectors in $\mathbb{R}^2$. Their sum is the vector

$$p + q = \begin{pmatrix} a+c \\ b+d \end{pmatrix}.$$

Theorem 3. Addition of vectors satisfies the same rules as addition of real numbers:

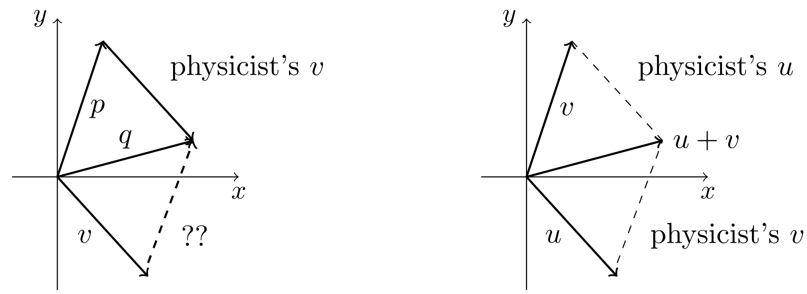
- when adding more than two vectors, it doesn't matter which operations you do first: $(p+q)+r = p+(q+r)$;
- one can add in either order $p+q = q+p$;
- the vector 0 corresponding to the origin O is a "zero" since $p + 0 = 0 + p = p$;
- for each vector p, there is an opposite vector $-p$ so that $p + (-p) = 0$.

Reading Exercise. *Remember that you are supposed to read actively. You can draw all of these pictures and try out all of these things with specific examples that you invent. You should check the statements in Theorem 3 by making examples and working out the details. Can you also draw the pictures which go with your examples?*

But what about subtracting geometrically? In Figure 1.8, I have a strong desire to complete the figure by joining the loose end of v to the rest of the figure. If we draw the arrow from the head of v to the head of q, we get Figure 1.9a.

What should the label on the dashed vector in Figure 1.9a be? Just as the physicist's v and the mathematician's v are parallel, this new vector is parallel to the mathematician's vector p. So the dashed vector must be a physicist's version of p. Then we see $q = v + p$.

Now we know how to add geometrically: to add two vectors u and v, translate v until its tail is on the head of u, then draw a new vector $u + v$ as the vector from the tail of u to the head of this translated v. It just



(a) Subtraction of vectors (b) Geometric Addition of Vectors

Figure 1.9: Some algebra of vectors

repurposes the structure of Figure 1.9a. This is called the *parallelogram rule* for addition of vectors.

There is another useful operation on vectors called *scalar multiplication*. The terminology comes from physics (again) where a *scalar* quantity is just a number, and not a vector. So "scalar multiplication" means to multiply a vector by a scalar.

Definition 4 (Scalar Multiplication for vectors)**.** Let $p = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ be a vector in $\mathbb{R}^2$, and let λ be a real number. Then the *scalar multiple* λp is defined to be

$$\lambda p = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix}$$

If you have never seen the symbol λ before, it is an old Greek letter pronounced "lamb-duh." It is traditional to use it in linear algebra in lots of places. Welcome to the λ-club. Oh, there are other such letters, like μ, which is pronounced "myoo."

Again, this operation has some important similarities to the familiar multiplication of numbers, but because it combines a scalar (a number) with a vector (not a number, exactly) to produce another vector (again, not a number) things are a little different.

Theorem 5. Suppose that p and q are vectors, and λ and μ are numbers. Scalar multiplication has the following properties:

- Scalar multiplication distributes over vector addition: $\lambda(p + q) = \lambda p + \lambda q$;
- Scalar multiplication distributes over scalar addition: $(\lambda + \mu)p = \lambda p + \mu q$;

- Scalar multiplication and regular multiplication can be done in either order: $\lambda(\mu p) = (\lambda\mu)p$;
- if $\lambda = 0$, then $\lambda p = 0p = 0$ is the zero vector.
- if n is a counting number, then np is the same thing as adding together n copies of p. In particular, $1p = p$.

This Theorem, like the last one, says that some natural properties you hope will still work really do still work. When you study **Modern Algebra**, making lists of these kinds of properties will be really useful. So far, we have collected up the properties that describe a *vector space*.

What is the geometry of scalar multiplication? It corresponds to stretching (or shrinking) a vector, without changing its direction. Let λ be a non-zero number. Since

$$\lambda p = \lambda \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix},$$

we see that the ratio of the two coordinates of a vector doesn't change under scalar multiplication. This means that p and λp point in the same direction. One can see this by considering similar triangles with sides parallel to the x-axis, the y-axis, and p.

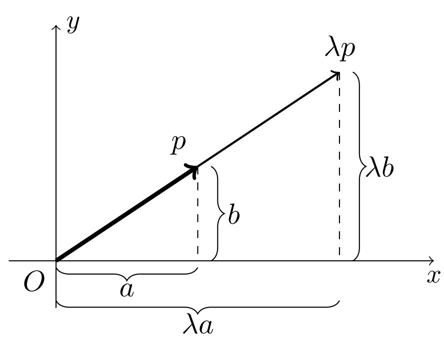


Figure 1.10: Similar triangles and scalar multiplication, $\lambda > 1$

The triangles in Figure 1.10 are similar: their corresponding horizontal and vertical sides are parallel, and those pairs of sides have a common ratio. We learn that p and λp lie in the same line.

This is important! Later we will describe lines in the plane, and we have just discovered how scalar multiplication relates to those lines which pass through the origin, O.

By the way, this picture helps explain the terminology. The vector λp is a rescaled version of p. A *scalar* is a thing which *scales* vectors.

The General Case

Now we can give the fully general definition.

Definition 6. Let m be a counting number. We define an m-vector to be a vertical stack of m real numbers, like so:

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}.$$

The individual entries u_i of u are called its *components*. The collection of all m-vectors is called *m-space*, and denoted $\mathbb{R}^m$.

The symbol $\mathbb{R}^m$ is usually read as "arr-em."

Definition 7 (Linear Combinations of vectors)**.** Suppose that u and v are two m-vectors. Their *sum* $u + v$ is defined by adding the individual components in corresponding positions. If λ is a number, then the *scalar multiple* λu is defined by multiplying each of the components of u by the number λ.

Suppose that v_1, v_2, $\dots$, v_k are all m-vectors, and that a_1, a_2, $\dots$, a_k are all real numbers. Then the vector

$$a_1v_1 + a_2v_2 + \cdots + a_kv_k$$

is called a *linear combination of the vectors v_1, v_2, $\dots$, v_k with weights a_1, a_2, $\dots$, a_k.*

It is much harder to think through the geometry of $\mathbb{R}^m$ when m is large, but the algebra works in much the same way.

Theorem 8. *Fix a natural number m. The results of Theorem 3 and Theorem 5 about the algebra of vectors hold for vectors in $\mathbb{R}^m$, too.*

Now that we have a little bit of algebraic structure for vectors, we can form equations.

Definition 9. An equation of the form $\lambda_1u_1 + \dots \lambda_nu_n = w$, where all of the vectors u_i and w are known, but the scalars λ_i are unknown, is called a *linear combination of vectors equation*. A *solution* to such an equation is a collection of scalars which make the equation true.

Exercises

Exercise 6. Write down three distinct points in the plane in proper notation. Plot those three points on a single diagram.

Exercise 7. Write down three distinct vectors in the plane in proper notation. Your three points from this task should NOT match any of the three points from the previous task. Plot those three vectors on a single diagram.

Exercise 8. Find the sum of your three vectors from the last exercise. Then, choose some order of those three vectors so that they are u_1, u_2 and u_3, and compute the linear combination

$$3u_1 - 2u_2 + (1/2)u_3.$$

Exercise 9. Let's consider the vectors $u = \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right)$, $v = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$, and $w = \left(\begin{smallmatrix} -3 \\ 1 \end{smallmatrix}\right)$. Compute all of the vectors in this list:

$$\frac{u+v}{2}, v-u, v-w, u + \left(\frac{v-u}{3}\right), u + \left(\frac{3(v-u)}{4}\right)$$

Then make a single diagram which contains u, v, w and all of those vectors from the list, plotted as accurately as you can.

What do you notice? Is anything interesting going on?

Exercise 10. Consider the vector $u = \left(\begin{smallmatrix} -4 \\ 2 \end{smallmatrix}\right)$. Find a vector v which has the property that $u + v$ is the zero vector, or explain why this is not possible.

Exercise 11. For now, keep $u = \left(\begin{smallmatrix} -4 \\ 2 \end{smallmatrix}\right)$. Let $v = \left(\begin{smallmatrix} 3 \\ 5 \end{smallmatrix}\right)$. How many solutions does the linear combination of vectors equation $\lambda u = v$ have?

How many solutions does the linear combination of vectors equation $\lambda u + \mu v = 0$ have? (Here, treat 0 as the zero vector.)

Exercise 12. We still use the notation u for the vector $u = \left(\begin{smallmatrix} -4 \\ 2 \end{smallmatrix}\right)$, but now use v for the vector $v = \left(\begin{smallmatrix} 28 \\ -14 \end{smallmatrix}\right)$. How many solutions does the linear combination of vectors equation $\lambda u = v$ have?

How many solutions does the linear combination of vectors equation $\lambda u + \mu v = 0$ have? (Again, treat 0 as the zero vector.)

Exercise 13. Find the midpoint between the points $P = (4, -2)$ and $Q = (3, 5)$. Then find the two points which divide the segment PQ into thirds.

How can vectors make this simpler than it first appears?

Challenge 14. Suppose you are given three points in the plane. Let's call them P, Q, and R. How can you use vectors to (quickly) determine if these three points are collinear?

Exercise 15. Write down four 3-vectors of your choosing, where none has any coordinate equal to 0. Call these vectors u_1, u_2, u_3, and u_4. Now choose any four non-zero scalars you like, call them λ_1, λ_2, λ_3 and λ_4. Compute the linear combination

$$\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 + \lambda_4 u_4.$$

1.3 The Idea of a Subset

As our study progresses, we will encounter many examples of collections of things. Often those collections will have other interesting collections inside of them. We will also encounter this phenomenon in reverse, where a collection which interests us actually lies inside some other collection, which is somehow larger. It will be convenient for us to have some basic language for these relationships, so we take that up now.

A Set and its Elements

The way a mathematician talks about a collection of things is to use the word *set*. What is a *set*? Well, a *set* is a collection of things. (There are some really deep intellectual waters nearby this concept, but we will avoid those this semester. The basic idea will be enough to get us through.) What is important is that a set has *elements*.

In fact, the only thing we want to formalize here is the idea of membership. Suppose that we have two mathematical objects. The first, S, is required to be a set. The second, a, need not be a set, though it might be. We will say that *a is an element of S* and write the notation

$$a \in S$$

when a happens to be one of the things in the collection S.

A simple example will go far. Suppose that S is the collection consisting of the numbers 1, 2, 3 and 4. Then S is a set, and the numbers 1, 2, 3, and 4 are the elements of S. But the number 5 is not an element of S, because it is not part of the collection.

It will help us to have notation for describing sets. The standard way to do it is to describe a set using curly braces, like so:

$$S = \{1, 2, 3, 4\}$$

The above sentence should be read "S is the set consisting of the elements 1, 2, 3, and 4." Here, the curly braces are a visual signal of the beginning and the end of the description of the set. They also serve as a visual metaphor: mathematicians tend to think of a set as a type of container which holds other things. Because we know that 3 is an element of this set S, but 5 is not an element of S, we write

$$3 \in S \text{ and } 5 \notin S,$$

where the slash through the $\in$ symbol changes the meaning from "is an element of" to "is not an element of."

Sometimes it is convenient to list all of the elements in a set, but often it is not. In those cases, we use a modification of the notation above. First, we set up some notation for the possible elements, then we write a vertical bar, and then we write down a description telling what has to be true for that object before the bar to be an element. Take note that this description might be something written in an sentence, or something written with mathematical symbols.

Again, some examples will help. The set of all real numbers, the set of integers, the set of 2d vectors, the set of positive real numbers, and the set of points in the plane which lie on a standard parabola can be written as in these examples below.

$$\begin{aligned}
\mathbb{R} &= \{x \mid x \text{ is a real number}\} \\
\mathbb{Z} &= \{x \in \mathbb{R} \mid x \text{ is an integer}\} \\
\mathbb{R}^2 &= \{v = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \mid x, y \in \mathbb{R}\} \\
P &= \{x \in \mathbb{R} \mid x > 0\} \\
Q &= \{p = (x, y) \mid x, y \in \mathbb{R} \text{ and } y = x^2\}
\end{aligned}$$

The description to the right of the vertical bar is very important for deciding if something is an element of the set or not. It gives exactly the criterion for testing. If your object makes the statement true, it is an element of the set. If your object makes the statement false, then it is not an element of the set. For example, it should be clear that the following statements are true:

$$(5, 10) \notin Q, \quad (-3, 9) \in Q.$$

Subsets

We can build on the relationship between a set and its elements to consider a kind of containment relationship between two sets.

Definition 10. Let S and T be two sets. We say that S *is a subset of* T when for each element x of S, we have that x is also an element of T. If S is a subset of T, we will use the notation

$$S \subset T.$$

In our list of examples above, we have several subset relationships. For example, since every integer is a real number, we know that $\mathbb{Z} \subset \mathbb{R}$. Similarly, $P \subset \mathbb{R}$. But there are integers which are not positive numbers, so it is not the case that $\mathbb{Z}$ is a subset of P.

Exercises

Exercise 16. Use our notation to write down a description of the set E which consists of all even integers.

Exercise 17. Rewrite the set definition below as a sentence in plain English.

$$C = \left\{ v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| \, x_1^2 + x_2^2 = 2 \right\}$$

It turns out that C is a subset of some familiar set. Which one is it?

Exercise 18. Let C be the set from the last exercise. Find three examples of vectors in $\mathbb{R}^2$ which are elements of the set C.

Then find three example of vectors in $\mathbb{R}^2$ which are not elements of C.

Exercise 19. Let $v \in \mathbb{R}^2$ be the vector $v = \left(\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right)$. We are interested in the subset of $\mathbb{R}^2$ which consists of vectors w so that v and w make a right angle. First, write out a description of this set in proper notation. Then try to describe what this set looks like in common terms.

Exercise 20. Repeat the last exercise, but instead work in $\mathbb{R}^3$ and with the vector v below.

$$v = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$$

1.4 The Dot Product and Geometry

A lot of the basic geometry of $\mathbb{R}^n$ is captured by a mysterious object called the *dot product*. First, we will show how this can be understood in $\mathbb{R}^2$, and then we will generalize everything to $\mathbb{R}^n$ for $n \geq 2$.

Lengths and Angles in $\mathbb{R}^2$

Given two vectors, their dot product is a real number. This number is a strange measure of how much the vectors are alike, and captures information about both lengths and angles. The formal definition is this.

Definition 11 (The Dot Product in $\mathbb{R}^2$)**.** Let $u = \left(\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right)$ and $v = \left(\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right)$ be vectors in $\mathbb{R}^2$. The *dot product* of u and v is the number

$$u \cdot v = u_1 v_1 + u_2 v_2.$$

The clearest geometric information we can pull out of the dot product is about lengths.

Theorem 12 (Lengths in $\mathbb{R}^2$)**.** Let $u = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \in \mathbb{R}^2$ be a vector. The length of u is equal to the square root of the dot product of u with itself. That is, the length of u is

$$\sqrt{u \cdot u} = \sqrt{a^2 + b^2}$$

Proof. The key comes from considering Figure 1.11. There we see the vector u in $\mathbb{R}^2$ is the hypotenuse of a right triangle having its legs parallel to the two coordinate axes.

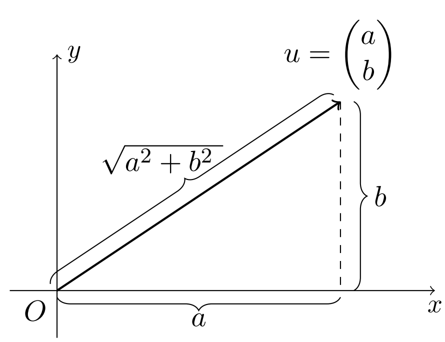


Figure 1.11: The Norm of a vector

These two legs then have lengths which are equal to our two coordinates, respectively. So we can compute the length of the hypotenuse (our vector) by the Pythagorean Theorem, which completes our proof. □

The length of a vector is a useful concept, but mathematicians often use another name for it. New learners often find this confusing, so beware!

Definition 13 (Norm, Unit Vector)**.** Let $u = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ be a vector in $\mathbb{R}^2$. The *norm* of u is the number

$$\|u\| = \sqrt{u \cdot u} = \sqrt{a^2 + b^2}\,.$$

We say that u is a *unit vector* when it has a norm of $\|u\| = 1$.

Theorem 14 (The Norm and Scalar Multiplication)**.** Let λ be a scalar, and let $u \in \mathbb{R}^2$ be a vector. Then

$$\|\lambda u\| = |\lambda|\, \|u\|$$

This conforms to our basic ideas about scalar multiplication. If we rescale a vector, we are really just changing its length by that factor. (And then there is the worry over the sign of the scalar, because lengths must not be negative.)

Proof. Suppose that the vector is $u = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$. We can do the following straightforward computation:

$$\begin{aligned} \|\lambda u\| &= \left\| \lambda \begin{pmatrix} a \\ b \end{pmatrix} \right\| = \left\| \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix} \right\| = \sqrt{(\lambda a)^2 + (\lambda b)^2} \\ &= \sqrt{\lambda^2 (a^2 + b^2)} = |\lambda| \sqrt{a^2 + b^2} \\ &= |\lambda|\, \|u\| \end{aligned}$$

Since the first and last are equal, we have established the desired result. $\square$

In a terrible linguistic collision, the most useful role for the norm of a vector is in the process of *normalizing* that vector. Beware! Mathematics has several words that are overused! Norm, normal, normalize, normalized and other forms of this are a prime example. To *normalize* a vector u, we multiply it by the scalar $\|u\|^{-1}$, and thus produce a new vector $u / \|u\|$ which points in the same direction as u, but has norm equal to 1. This is because we can use our observation just above about how norms interact with scalar multiplication to compute:

$$\left\| \|u\|^{-1} u \right\| = \|u\|^{-1} \|u\| = 1.$$

That last multiplication is just a multiplication of numbers. Note that all of that only makes sense as long as u is not the zero vector. If u is the zero vector, its norm is $\|u\| = 0$. Since it makes no sense to divide by zero, it is not possible to normalize the zero vector.

To sum up: given a non-zero vector u, it is possible to normalize it, which means to replace it with the unit vector $u/\|u\|$.

Let's turn out attention to measuring angles between vectors. To do this properly, we will need a little trigonometry. The three necessary facts are collected in the next theorem as a refresher.

Theorem 15 (Some Trigonometry)**.** The following statements are true.

1. Each point on the unit circle can be represented as a point of the form $(\cos(\alpha), \sin(\alpha))$ for some angle α between 0 and 2π, so the corresponding vector can be written as $p = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$.

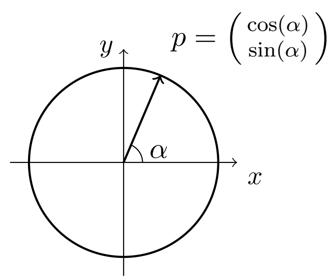


Figure 1.12: Trigonometry for points on the unit circle

2. The function $\theta \mapsto \cos(\theta)$ associates to each angle in the interval $[0, \pi]$ a unique real number in the interval $[-1, 1]$, and vice versa. In particular, this function has a sensible inverse function $\arccos : [-1, 1] \to [0, \pi]$. For us, this means that it to measure an angle, we can get away with instead finding the cosine of that angle.

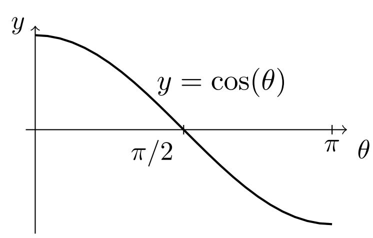


Figure 1.13: Part of the graph of $y = \cos(\theta)$

3. There is an identity on trigonometric functions that helps us deal with the the cosine of a difference of angles: If α and β are two angles, then
$$\cos(\beta - \alpha) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta).$$

You might recognize all of these facts from a trigonometry or pre-calculus class. In any case, you should take them as true. It would take us too far out of our studies to establish them now.

Theorem 16 (Angles in $\mathbb{R}^2$)**.** Let $u, v \in \mathbb{R}^2$ be two vectors. Then the angle θ between u and v is
$$\theta = \arccos\left(\frac{u \cdot v}{\|u\|\,\|v\|}\right).$$

Proof. So, we consider the angle between two vectors, u and v, in $\mathbb{R}^2$. As they are mathematician's vectors, they are both based at the origin, and naturally form an angle there. We will now apply each of the facts from Theorem 15.

Note that the angle between these vectors does not depend at all on their lengths. If we change one, or both, of the vectors by rescaling them, that will not change the directions involved, and hence will not change the angle we seek. We can get to a more uniform set up for our task by normalizing u and v. Thus, we will instead consider the unit vectors $u/\|u\|$ and $v/\|v\|$, and look for the angle θ between them as in Figure 1.14.

Since our new vectors are unit vectors, we can represent them with trigonometric functions:
$$\frac{u}{\|u\|} = \begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix}, \qquad \frac{v}{\|v\|} = \begin{pmatrix}\cos(\beta)\\ \sin(\beta)\end{pmatrix}.$$
Let's suppose that the angle β that v makes with the x-axis is larger than the angle α that u makes with the x-axis. Then we want to find the angle $\theta = \beta - \alpha$.

By angle difference formula for cosine, we have
$$\begin{aligned}\cos(\theta) &= \cos(\beta - \alpha)\\ &= \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)\\ &= \left(\frac{u}{\|u\|}\right)\cdot\left(\frac{v}{\|v\|}\right)\\ &= \frac{u \cdot v}{\|u\|\,\|v\|}.\end{aligned}$$
Now, we apply the function arccos to both sides to get the theorem. $\square$

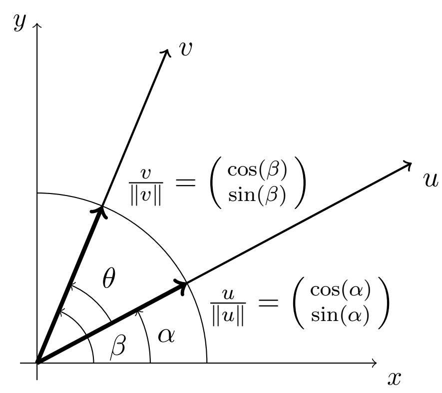


Figure 1.14: Finding the angle between two vectors.

As our study progresses, we will not have much use for measuring particular angles, but it will be very important to us to understand situations when two vectors make an angle of $\pi/2$ radians (i.e. $90°$, a right angle). It is common in geometry to call such vectors *perpendicular*. It is a fact that $\cos(\pi/2) = 0$, so that non-zero vectors u and v make an angle of $\pi/2$ when

$$0 = \cos(\pi/2) = \frac{u \cdot v}{\|u\| \, \|v\|}.$$

By clearing out the denominator of this fraction, we see that u and v make an angle of $\pi/2$ exactly when $u \cdot v = 0$. In another instance of cluttering the vocabulary list for math students, in a linear algebra context such vectors are called by yet another term.

Definition 17 (Orthogonal Vectors)**.** We say that two vectors u and v in $\mathbb{R}^2$ are *orthogonal* when $u \cdot v = 0$.

The Dot Product in $\mathbb{R}^n$

Now we will give the general definition and results. Fortunately, everything works the same.

Definition 18 (The Dot Product)**.** Let u and v be vectors in $\mathbb{R}^n$ with coordinates labeled like those below.

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

The *dot product* of u and v is the number

$$u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

Once we have the dot product available, we can work by analogy to define all of the other new words of this section for vectors in $\mathbb{R}^n$, too.

Definition 19 (Geometry in $\mathbb{R}^n$)**.** Let $u, v \in \mathbb{R}^n$ be vectors.

- The *norm of* u is $\|u\| = \sqrt{u \cdot u}$.
- The vector u is called a *unit vector* when $\|u\| = 1$.
- The *angle between* u *and* v is the number

$$\theta = \arccos\left(\frac{u \cdot v}{\|u\| \, \|v\|}\right).$$

- We say that u and v are *orthogonal* when $u \cdot v = 0$.

The next two theorems are not hard to prove, but they are tedious. In each case, one has to argue one coordinate at a time, using a relevant property for real numbers. It is best to just check them for some examples until you understand and believe them.

Theorem 20 (Algebra of the Dot Product)**.** Let u, v, and w be vectors in $\mathbb{R}^n$ and let λ and μ be scalars. Then

- The dot product is *symmetric*: $u \cdot v = v \cdot u$;
- The dot product *distributes* over linear combinations

$$u \cdot (\lambda v + \mu w) = \lambda(u \cdot v) + \mu(u \cdot w)$$

Theorem 21 (Algebra of the Norm)**.** Let $u \in \mathbb{R}^n$ be a vector, and let λ be a scalar. Then

- $\|u\| \geq 0$.
- $\|u\| = 0$ if, and only if, u is the zero vector.
- $\|\lambda u\| = |\lambda| \, \|u\|$

Exercises

Exercise 21. Choose three different vectors in $\mathbb{R}^2$ which have neither of their components equal to zero. Call these vectors u, v, and w.

a) Compute the norms of u, v, and w.
b) Compute the dot products $u \cdot v$, $v \cdot w$, and $u \cdot w$.
c) Find unit vectors u', v', and w' which point in the same directions as u, v, and w, respectively.
d) Find the angles between each of the pairs, u and v, u and w, v and w in radians.

Exercise 22. Fix some vector $u \in \mathbb{R}^2$. Draw a picture of u in the plane, and then shade the region of the plane which contains vectors v so that $u \cdot v > 0$.

Exercise 23. This task continues our quest for understanding the sign of a dot product geometrically.

a) Find an example of two v and w in $\mathbb{R}^2$ so that $\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right) \cdot v = 0$ and $\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right) \cdot w = 0$, or explain why such an example is not possible.
b) Let $v = \left(\begin{smallmatrix}3\\-1\end{smallmatrix}\right)$. Find an example of a pair of 2-vectors u and w such that $v \cdot u < 0$ and $v \cdot w < 0$ and $w \cdot u = 0$, or explain why no such pair of vectors can exist.
c) Find an example of three 2-vectors u, v, and w so that $u \cdot v < 0$ and $u \cdot w < 0$ and $v \cdot w < 0$, or explain why no such example exists.

Exercise 24. What shape is the set of solutions $\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)$ to the equation

$$\begin{pmatrix}3\\7\end{pmatrix} \cdot \begin{pmatrix}x\\y\end{pmatrix} = 5?$$

That is, if we look at all possible vectors $\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)$ which make the equation true, what shape does this make in the plane? Draw this shape.

What happens if we change the vector $\left(\begin{smallmatrix}3\\7\end{smallmatrix}\right)$ to some other vector? What happens if we change the number 5 to some other number?

Exercise 25.

a) Find an example of a number c so that the equation

$$\begin{pmatrix}1\\-1\end{pmatrix} \cdot \begin{pmatrix}x\\y\end{pmatrix} = c$$

has the vector $\left(\begin{smallmatrix}4\\7\end{smallmatrix}\right)$ as a solution, or explain why no such number exists.
b) Let $v = \left(\begin{smallmatrix}2\\1\end{smallmatrix}\right)$ and $w = \left(\begin{smallmatrix}-3\\4\end{smallmatrix}\right)$. Find an example of a number c so that

$$v \cdot \begin{pmatrix}1\\-1\end{pmatrix} = c \quad \text{and} \quad w \cdot \begin{pmatrix}1\\-1\end{pmatrix} = c,$$

or explain why this is not possible.

c) Let $P = \binom{-3}{4}$. Find an example of numbers c and d so that

$$\begin{pmatrix} 2 \\ -1 \end{pmatrix} \cdot P = c \quad \text{and} \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot P = d,$$

or explain why no such example is possible.

Exercise 26. Write down two vectors in $\mathbb{R}^3$ which have no coordinates equal to zero. Call them u and v. Find the following things:

- The dot product $u \cdot v$;
- The norms of u and v;
- unit vectors which point in the same directions as u and v, respectively; and
- the angle between u and v.

1.5 Lines, Especially in $\mathbb{R}^2$

We will need a good understanding of lines in $\mathbb{R}^n$. We shall introduce lines with a geometric definition using vectors, and then show how to describe a line as the image of a parametric curve. Finally, we will focus on just the plane $\mathbb{R}^2$, and see how to describe a line as the set of solutions to an equation.

Lines as Parametrized Objects

Any pair of points should define a line. In $\mathbb{R}^n$, we can do this using vectors. It will be fruitful to regularly conflate a point with the (mathematician's) vector having the point as its head, so we will do that throughout this section.

Definition 22 (Lines in $\mathbb{R}^n$)**.** Suppose that P and Q are two distinct vectors in $\mathbb{R}^n$. We say that a point X *lies on the line through P and Q* when there exists scalars λ and μ which are not both zero so that

$$\lambda(X - P) = \mu(X - Q).$$

The *line through P and Q* is the set of all such points, that is, it is the set

$$\begin{aligned} &\{X \in \mathbb{R}^n \mid X \text{ lies on the line through } P \text{ and } Q\} \\ &= \{X \in \mathbb{R}^n \mid \lambda(X - P) = \mu(X - Q), \text{for some } \lambda \text{ and } \mu \text{ where } \lambda\mu \neq 0\} \end{aligned}$$

Of course, we want P and Q to both lie on the line through P and Q. Fortunately, this is true.

Theorem 23. Let P and Q be distinct vectors in $\mathbb{R}^n$. Then both P and Q lie on the line through P and Q.

Proof. We will show that P lies on the line through P and Q. We must check the condition from the definition with X replaced by P. This means we want to know about the vectors $X - P$ and $X - Q$, when $X = P$. But $P - P = 0$, so we can just choose $\lambda = 1$ and $\mu = 0$, and get the equation

$$\lambda(X - P) = 1(P - P) = 0(P - Q) = \mu(X - Q).$$

The proof that Q lies on the line through P and Q is basically the same, with some of the letters moved around. I will leave it to you to check the details of that case. $\square$

Theorem 24 (Only One Direction Matters)**.** Let P and Q be two distinct vectors in $\mathbb{R}^n$. If $X \in \mathbb{R}^n$ lies on the line through P and Q exactly when there exists a scalar t so that $X - P = t(Q - P)$.

Proof. First suppose that X lies on the line through P and Q. Then, by definition there exists a pair of scalars λ and μ so that

$$\lambda(X - P) = \mu(X - Q). \tag{1.1}$$

Notice that we cannot have $\lambda = \mu$. The reason is that if $\lambda = \mu \neq 0$, then we can cancel them and deduce that $X - P = X - Q$, which means that $P = Q$. But we have explicitly chosen P and Q to be distinct. So, we will proceed knowing that λ and μ are different.

Rearranging equation (1.1) , we see that

$$(\lambda - \mu)X - \lambda P = -\mu Q.$$

By adding μP to both sides and gently regrouping things, we find

$$\begin{aligned}(\lambda - \mu)X - \lambda P + \mu P &= -\mu Q + \mu P\\ (\lambda - \mu)X - (\lambda - \mu)P &= -\mu(Q - P)\\ (\lambda - \mu)(X - P) &= -\mu(Q - P)\end{aligned}$$

Since $\lambda - \mu \neq 0$, we can divide by it, and we see that

$$X - P = \frac{-\mu}{\lambda - \mu}(Q - P).$$

So we can choose $t = \dfrac{-\mu}{\lambda - \mu}$ to satisfy the condition of the theorem.

Now suppose that X is a point so that $X - P = t(Q - P)$ for some scalar t. We must find the scalars λ and μ which satisfy equation (1.1). Since $X - P = t(Q - P)$, we have that

$$X = tQ + (1 - t)P \tag{1.2}$$

This means that $X - Q = (t - 1)(Q - P)$. So if we choose $\lambda = t - 1$ and $\mu = t$, we have

$$\begin{aligned}\lambda(X - P) &= (t - 1)(X - P)\\ &= (t - 1)t(Q - P)\\ &= t(X - Q)\\ &= \mu(X - Q).\end{aligned}$$

Therefore, X lies on the line through P and Q. □

The equation (1.2) is a very nice symmetric description for points X which lie on the line through P and Q. Many authors, especially those who want to do lots of geometry, emphasize this description.

Also, notice the importance this theorem places on the vector $Q - P$. This vector is useful enough for dealing with the line that it has a special name.

Definition 25 (Direction Vector)**.** Let P and Q be two distinct vectors in $\mathbb{R}^n$. The vector $Q - P$ is called a *direction vector* for the line through P and Q.

The order of the points P and Q really don't matter for defining a line, so we can see that $P - Q$ is a direction vector, too. It just points in the opposite direction. Furthermore, a line has lots of points on it, and any pair of them is good enough to describe the line. Choosing a different pair for P and Q will lead to different direction vectors, but only different by rescaling. All of the vectors will point in the same direction. So keep in mind that a line has lots of direction vectors, but only one direction.

Corollary 26 (Parametric form of a Line)**.** Let P and Q be two distinct vectors in $\mathbb{R}^n$. The line through P and Q is the set

$$\{X \in \mathbb{R}^n \mid X = P + t(Q - P) \text{ for some scalar } t\}.$$

Proof. This condition is a simple rearrangement of the one in Theorem 24. Just add P to both sides of the description there. □

Corollary 27 (Lines through the origin)**.** A line in $\mathbb{R}^n$ can be written in the form

$$\{X \in \mathbb{R}^n \,|\, X = tV\}\,,$$

where V is some fixed vector.

Proof. Essentially, this is the case where $P = 0$. We replace $Q = Q - P$ by V. □

Note that the vector $X = P + t(Q - P)$ changes as t changes. We often think of this as defining a function of the form $\gamma : t \mapsto P + t(Q - P)$. (By the way, γ is a version of a Greek letter and is read "gamma.") Usually, the best way is to think of the number t, called the *parameter*, as representing time. As the time t changes, the vector $X = \gamma(t)$ moves around in $\mathbb{R}^n$. In the case here, the point $X = \gamma(t)$ traces out the shape of a line as time moves on.

If you prefer to think about the whole vector $\gamma(t)$, rather than just its head, it helps to imagine that vector sweeping through space as time goes one. The heads of the vectors still trace out the line, but the tails of the vectors all stay at the origin.

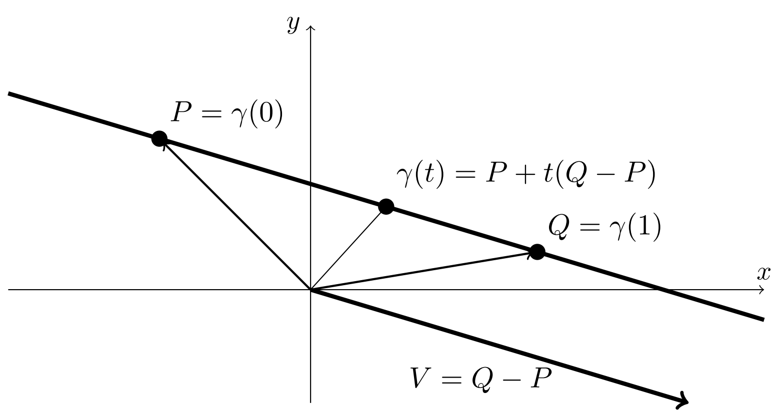


Figure 1.15: parametric line and a direction vector

You might see parametric functions in other places (like a course on calculus), but written differently. It is common to break apart a vector description into a system of component functions. We will have occasion to use this, too, so let's see how it is done. Begin with a parametric line

$$\gamma(t) = P + t(Q - P).$$

Write out the vectors $\gamma(t)$, P and Q as stacks of components,

$$\gamma(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} \qquad P = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{pmatrix} \qquad Q = \begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_n \end{pmatrix}$$

and unpack the definition using the algebra of linear combinations.

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_m \end{pmatrix} + t \begin{pmatrix} Q_1 - P_1 \\ Q_2 - P_2 \\ \vdots \\ Q_n - P_n \end{pmatrix} = \begin{pmatrix} P_1 + t(Q_1 - P_1) \\ P_2 + t(Q_2 - P_2) \\ \vdots \\ P_n + t(Q_n - P_n) \end{pmatrix}$$

Now, just read off each component, one at a time, to make a system of functions. It sounds like cheating, but you basically just erase the parentheses and then group with a big curly brace on the left instead.

$$\begin{cases} x_1 &= P_1 + t(Q_1 - P_1) \\ x_2 &= P_2 + t(Q_2 - P_2) \\ \vdots & \qquad \vdots \\ x_n &= P_n + t(Q_n - P_n) \end{cases}$$

Sometimes people will write $x_i(t)$ with the parameter t explicitly present, and sometimes they will write just x_i and leave it as understood that x_i is a function of t. I have left the t's out of the final expression because it was convenient.

Implicit Description: The Equation of a Line in $\mathbb{R}^2$

Now we narrow our attention to the plane $\mathbb{R}^2$. We will see that in this case it is possible to also describe a line as the set of points satisfying a single, simple equation.

Theorem 28. Let P and Q be two distinct vectors in $\mathbb{R}^2$. Then there are numbers a, b, and c so that any point $X = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2$ which lies on the line through P and Q must satisfy the equation

$$ax + by = c.$$

Proof. Let P and Q have components as follows:

$$P = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \qquad Q = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}$$

By the discussion at the end of the last subsection, we can find a value of t so that

$$\begin{cases} x &= P_1 + t(Q_1 - P_1) \\ y &= P_2 + t(Q_2 - P_2). \end{cases}$$

Our goal is to eliminate t from these expressions and derive a single equation relating x and y. The mechanics of this is as follows: multiply the first equation through by $Q_2 - P_2$, multiply the second equation through by $-(Q_1 - P_1)$, and then add them. In a way, we are making a "linear combination of the equations." The result is

$$\begin{aligned} (Q_2 - P_2)x - (Q_1 - P_1)y &= (Q_2 - P_2)P_1 - (Q_1 - P_1)P_2 \\ &= Q_2P_1 - Q_1P_2 \end{aligned} \tag{1.3}$$

So, we choose $a = Q_2 - P_2$, $b = -(Q_1 - P_1)$ and $c = Q_2P_1 - Q_1P_2$. Then equation (1.3) reads as $ax + by = c$. This completes the proof. □

Theorem 29. Fix numbers a, b, and c so that a and b are not both zero. Then the set of points $X = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ which satisfy the equation $ax + by = c$ is a line in $\mathbb{R}^2$.

Proof. We will work in the case where $a \neq 0$, and leave the case $a = 0$ for you to fill in.

The main idea is to turn the defining equation into a parametric description, and along the way find vectors to play the parts of P and $V = Q - P$.

We will pretend that y is our parameter by giving it a new name. So, introduce the parameter t and declare that $y = t$. Take the defining equation $ax + by = c$, isolate x and substitute in $y = t$ to remove reference to y. Then our information is represented by these two equations.

$$\begin{cases} x &= c/a - (b/a)\,t \\ y &= t \end{cases}$$

Rewrite this as a vector equation.

$$X = \begin{pmatrix} c/a \\ 0 \end{pmatrix} + t\begin{pmatrix} -b/a \\ 1 \end{pmatrix}$$

So if we choose P and Q as below, we have written exactly the parametric description of a line as in Corollary 26.

$$P = \begin{pmatrix} c/a \\ 0 \end{pmatrix} \qquad Q = \begin{pmatrix} c/a - b/a \\ 1 \end{pmatrix}$$

Therefore, our collection of points is exactly a line. This completes the proof. (You should go back and fill in how to think about the case when $a = 0$. Hint: if $a = 0$, what do you know about b?) $\square$

Looking carefully at the last two results together, we see that lines in $\mathbb{R}^2$ come with two different descriptions, a parametric description and an implicit description, but we can easily pass back and forth between them. In fact, the proofs of the last two results give us explicit methods for passing back and forth between them. In the first, we eliminate the parameter. In the second, we have to introduce a new one out of nowhere.

Theorem 30. Let a line in the plane be described as the set of solutions to the equation $ax + by = c$. A direction vector for this line is

$$V = \begin{pmatrix} -b/a \\ 1 \end{pmatrix}.$$

Proof. This follows directly from our work in the proof of Theorem 29. $\square$

While direction vectors are useful for lines in $\mathbb{R}^2$, in $\mathbb{R}^n$ we will have other, bigger objects defined by equations. In those cases, a single direction vector will not be sufficient. But we can keep things simple by changing perspective just a little.

Definition 31 (Normal Vector)**.** Suppose we have a line in $\mathbb{R}^2$ through points P and Q. A vector n is called a *normal vector* for this line if n is orthogonal to the direction vector $V = Q - P$. That is, n is a normal vector when $n \cdot (Q - P) = 0$.

Theorem 32. Let ℓ be a line in $\mathbb{R}^2$ defined as the set of solutions to the equation $ax + by = c$. Then one normal vector for ℓ is given by

$$n = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Proof. By Theorem 30, the direction vector for ℓ is $V = \left(\begin{smallmatrix} -b/a \\ 1 \end{smallmatrix}\right)$. It is straightforward to check that

$$n \cdot V = \begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} -b/a \\ 1 \end{pmatrix} = -b + b = 0.$$

Therefore, n is a normal vector for ℓ. $\square$

Note that the normal vector is pretty handy for writing out the equation of a line. If we write

$$n = \begin{pmatrix} a \\ b \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

then the equation $ax + by = c$ is the same thing as $n \cdot X = c$. This gives us a connection between the dot product and the linear equation which will be useful later.

Exercises