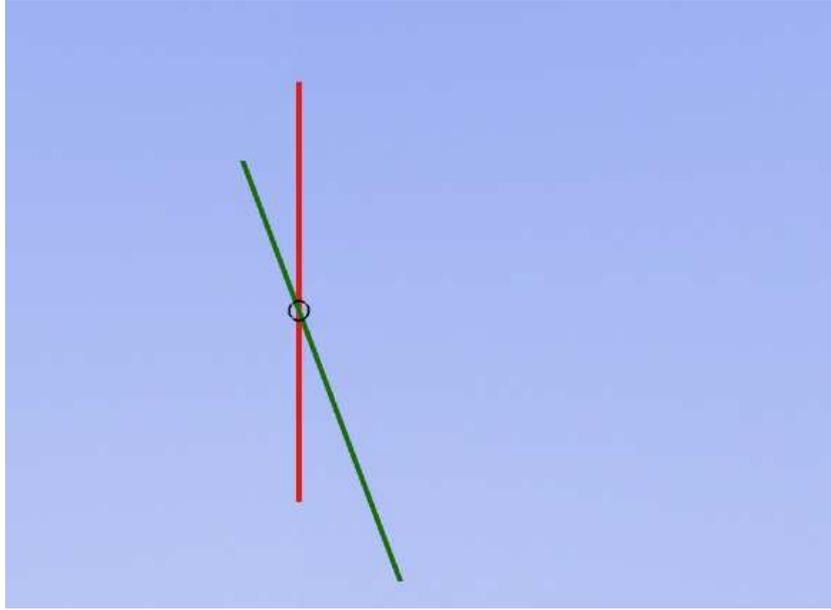


Representing Line Intersection as a System of Linear Equations

Andy Giese

June 12, 2016



Introduction

In a previous post, I outlined an analytical solution for intersecting lines and ellipses. In this post I'm doing much the same thing but rather with lines on lines. I'll point out why the normal slope-intercept form for a line is a poor representation, and what we can do about that.

In computer graphics, or just geometry in general, you often find yourself in a scenario where you want to know if and how two or more objects intersect. For example, in the latest shooting game you're playing perhaps a bullet is represented as a sphere and a target is represented as a disc. You want to know if the bullet you've fired has struck the target, and if so, was it a bulls-eye Figure (1)?



Figure 1: Where did the line hit the target?

In this post, I'll step you through one way we can accomplish discovering intersection points between two lines, being sure to carefully walk through each step of the calculation so that you don't get lost.

Slope-Intercept Form

We're going to fly in the face of many approaches to line-line intersection here and try to ultimately wind up with a system of linear equations to solve. This is different from the "usual" way of finding intersections by crafting an equation to represent each geometric body, and then somehow setting those equations equal to each other. For example, if we used the familiar slope-intercept form to represent a line, we'd end up representing the lines as

$$l_1 : y = m_1x + b_1 \tag{1}$$

$$l_2 : y = m_2x + b_2 \tag{2}$$

Then afterwards we could solve for x by assigning the equations to each other:

$$m_1x + b_1 = m_2x + b_2$$

And with some algebraic fiddling, we could get x , and then take that x and insert it into either line equation in Eqn. 1 to get y .

But here are some questions to think about.

A problem

Slope-intercept form assumes two things:

1. Every line has a slope
2. Every line has a y-intercept

This is all well and good for most lines. Even horizontal lines have a slope of 0, and a y-intercept somewhere. Our problem is with perfectly vertical lines (Figure 2). What is the slope of a vertical line, since the "run" part of rise-over-run, is zero? Vertical lines don't touch the y-axis either unless they're collinear with it, and even then there wouldn't be just a single y-intercept.

Vertical lines are better represented as a function of x . Like

$$x = 42$$

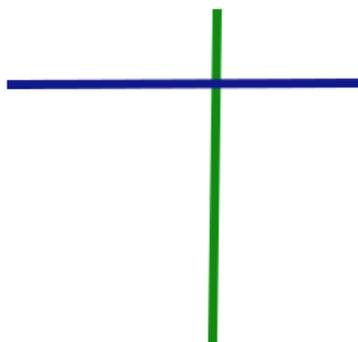


Figure 2: A vertical line segment (green) crosses a horizontal line (blue)

the y part isn't even 0 here, it just doesn't exist! That's because there are infinite values of y for our given x .

One solution

Vertical lines are the bane of slope-intercept form's existence. If both lines were vertical, we could test for that and then not bother with testing for intersection, but what if just one of them were vertical? How would we check for the intersection point then?

Well, let's examine an alternative representation of a line. Those of you who have ever taken a linear algebra course should be familiar with it:

$$ax + by = c \tag{3}$$

Where a , b , and c are known constant values, and we're solving for x and y .

Okay, what the heck is that? In the next section I'll explain exactly how this solves our vertical line problem, but first I need to demonstrate to you how we can even *get* a line into that format.

Two-point form: an alternative representation for a line

“What? We don't always use slope-intercept form??? All my teachers have lied to me!” In fact, there are many ways we can represent a line, and like any tool there's a time and a place for each of them.

The representation we're interested in here is called **Two-point form**. And we can derive it if we already have two points on the line (which is

common if you have a bunch of line-segments in the plane, like sketch strokes).

Given: We have two points on the line, $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, we can represent a line in the following form, parameterized on x and y :

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1} (x - x_1) \quad (4)$$

Still doesn't seem to help us, does it? We've gotten rid of the intercept, but we still have a slope, which becomes a big problem when $x_2 - x_1 = 0$. So let's fix that but multiplying both sides of the equation by $(x_2 - x_1)$:

$$(x_2 - x_1)(y - y_1) = (y_2 - y_1)(x - x_1) \quad (5)$$

(This is called **Symmetric form**) Let's manipulate Equation 5 so that x and y (no subscripts) appear only once. Start by multiplying everything through:

$$x_2y - x_1y - x_2y_1 + x_1y_1 = xy_2 - xy_1 - x_1y_2 + x_1y_1$$

We notice that the x_1y_1 terms can cancel out:

$$x_2y - x_1y - x_2y_1 = xy_2 - xy_1 - x_1y_2$$

Now we rearrange the left side in terms of y :

$$y(x_2 - x_1) - x_2y_1 = xy_2 - xy_1 - x_1y_2$$

And the right side in terms of x :

$$y(x_2 - x_1) - x_2y_1 = x(y_2 - y_1) - x_1y_2$$

Subtract the y term from both sides:

$$-x_2y_1 = x(y_2 - y_1) - y(x_2 - x_1) - x_1y_2$$

and add the x_1y_2 term:

$$x_1y_2 - x_2y_1 = x(y_2 - y_1) - y(x_2 - x_1)$$

Let's move the equals sign to the other side:

$$x(y_2 - y_1) - y(x_2 - x_1) = x_1y_2 - x_2y_1$$

And move the x and y to the other side of the parentheses:

$$(y_2 - y_1)x - (x_2 - x_1)y = x_1y_2 - x_2y_1$$

Notice now how similar this equation is to Equation 3.
We can define:

$$\begin{aligned}a &= y_2 - y_1 \\b &= -(x_2 - x_1) \\c &= x_1y_2 - x_2y_1\end{aligned}$$

Distributing the negative through for b :

$$\begin{aligned}a &= y_2 - y_1 \\b &= x_1 - x_2 \\c &= x_1y_2 - x_2y_1\end{aligned} \tag{6}$$

To be left with the same equation (restated here):

$$ax + by = c$$

Line-line intersection via solving a system of linear equations

Now that we can represent a single line as a linear equation in two variables, x and y , we can represent the intersection of two lines as a system of linear equations in two variables:

$$\begin{aligned}a_1x + b_1y &= c_1 \\a_2x + b_2y &= c_2\end{aligned}$$

Where we compute a_1, a_2, b_1, b_2, c_1 , and c_2 by using the two-point form mathematics from the previous section. With two equations and two unknowns, we can compute x and y .

In the next part I will slowly walk through how we will solve this system using basic techniques from linear algebra (Warning! Lots of math incoming!).

The idea behind solving a system of equations like this is to get it into something called *row-echelon form*, which is a fancy way of saying “I want the coefficient on x in the top equation to be 1, and the coefficient of y in the bottom equation to be 1, with other coefficients of x and y to be zero”:

$$\begin{aligned}x + (0)y &= (\text{something}) \\(0)x + y &= (\text{something else})\end{aligned}$$

Let's begin with getting the coefficient on x in the top row to be 1. First divide through the top equation by a_1 :

$$\begin{aligned}\frac{a_1}{a_1}x + \frac{b_1}{a_1}y &= \frac{c_1}{a_1} \\ a_2x + b_2y &= c_2\end{aligned}$$

Simplifying:

$$x + \frac{b_1}{a_1}y = \frac{c_1}{a_1}$$

$$a_2x + b_2y = c_2$$

Notice that at this point, we've succeeded in getting a coefficient of 1 on x in the top equation. The next thing we do is notice that, since we have a bare x in the top equation, we could multiply it by a_2 and then subtract it from the x in the bottom equation to get the bottom one's coefficient on x to be zero. However, we cannot do this solely on x . Instead, we're restricted to what are called *Elementary Row Operations* that limit what we can do while preserving the correctness of the equations. So what we have to do is multiply the *entire* top row by a_2 , and then subtract the top row from the bottom row, which looks like this:

$$\begin{aligned}x + \frac{b_1}{a_1}y &= \frac{c_1}{a_1} \\ (a_2 - a_2)x + \left(b_2 - \frac{a_2b_1}{a_1}\right)y &= \left(c_2 - \frac{a_2c_1}{a_1}\right)\end{aligned}$$

Which gives us a 0 coefficient on x in the second row!

$$\begin{aligned}x + \frac{b_1}{a_1}y &= \frac{c_1}{a_1} \\ (0)x + \left(b_2 - \frac{a_2b_1}{a_1}\right)y &= \left(c_2 - \frac{a_2c_1}{a_1}\right)\end{aligned}$$

Let's simplify the $\left(b_2 - \frac{a_2b_1}{a_1}\right)$ and $\left(c_2 - \frac{a_2c_1}{a_1}\right)$ terms in the second row to have a common denominator:

$$\begin{aligned}x + \frac{b_1}{a_1}y &= \frac{c_1}{a_1} \\ (0)x + \left(\frac{a_1b_2 - a_2b_1}{a_1}\right)y &= \left(\frac{a_1c_2 - a_2c_1}{a_1}\right)\end{aligned}$$

The next step is to get a coefficient of 1 on the y term in the second row by dividing through the row by $\left(\frac{a_1 b_2 - a_2 b_1}{a_1}\right)$, which results in us replacing the coefficient on y with a 1, and then multiplying the third term by the flipped fraction (remember that when we divide two fractions, we flip the second one and then multiply):

$$\begin{aligned} x + \frac{b_1}{a_1}y &= \frac{c_1}{a_1} \\ (0)x + y &= \left(\frac{a_1 c_2 - a_2 c_1}{a_1} \times \frac{a_1}{a_1 b_2 - a_2 b_1}\right) \end{aligned}$$

We can cancel the a_1 terms in the multiplication, and the resulting equation becomes:

$$\begin{aligned} x + \frac{b_1}{a_1}y &= \frac{c_1}{a_1} \\ (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1} \end{aligned}$$

At this point, we've solved for y . That is,

$$y = \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1}$$

We *could* substitute for y in the top row to solve for x , but a more linear-algebra-ish way would be to perform another elementary row operation — multiply the bottom row by $\frac{b_1}{a_1}$, and then subtract it from the top row. Here's what that looks like:

$$\begin{aligned} x + \left(\frac{b_1}{a_1} - \frac{b_1}{a_1}\right)y &= \left(\frac{c_1}{a_1} - \frac{b_1(a_1 c_2 - a_2 c_1)}{a_1(a_1 b_2 - a_2 b_1)}\right) \\ (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1} \end{aligned}$$

Now we've achieved a coefficient of 0 on the y term in the top row!

$$\begin{aligned} x + (0)y &= \left(\frac{c_1}{a_1} - \frac{b_1(a_1 c_2 - a_2 c_1)}{a_1(a_1 b_2 - a_2 b_1)}\right) \\ (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1} \end{aligned}$$

We've essentially solved for x at this point, but the term on the other side of the $=$ is a little ugly, and we can simplify it. Let's start by making the two parts of the term have the same denominator, which means we need to multiply $\frac{c_1}{a_1}$ by $\frac{(a_1 b_2 - a_2 b_1)}{(a_1 b_2 - a_2 b_1)}$ (which is really just multiplying by 1!):

$$\begin{aligned}
 x + (0)y &= \left(\frac{a_1 b_2 c_1 - a_2 b_1 c_1}{a_1 (a_1 b_2 - a_2 b_1)} - \frac{b_1 (a_1 c_2 - a_2 c_1)}{a_1 (a_1 b_2 - a_2 b_1)} \right) \\
 (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1}
 \end{aligned}$$

Distributing the b_1 term in the numerator, and combining the terms because they have the same denominator:

$$\begin{aligned}
 x + (0)y &= \frac{a_1 b_2 c_1 - a_2 b_1 c_1 - (a_1 b_1 c_2 - a_2 b_1 c_1)}{a_1 (a_1 b_2 - a_2 b_1)} \\
 (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1}
 \end{aligned}$$

When we distribute the $-$ in the numerator, the term $-a_2 b_1 c_1$ becomes positive:

$$\begin{aligned}
 x + (0)y &= \frac{a_1 b_2 c_1 - a_2 b_1 c_1 - a_1 b_1 c_2 + a_2 b_1 c_1}{a_1 (a_1 b_2 - a_2 b_1)} \\
 (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1}
 \end{aligned}$$

Which leaves us with both $-a_2 b_1 c_1$ and $+a_2 b_1 c_1$ in the numerator, which cancel:

$$\begin{aligned}
 x + (0)y &= \frac{a_1 b_2 c_1 - a_1 b_1 c_2}{a_1 (a_1 b_2 - a_2 b_1)} \\
 (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1}
 \end{aligned}$$

We can factor a_1 out in the numerator:

$$\begin{aligned}
 x + (0)y &= \frac{a_1 (b_2 c_1 - b_1 c_2)}{a_1 (a_1 b_2 - a_2 b_1)} \\
 (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1}
 \end{aligned}$$

Finally, the a_1 terms in the numerator and denominator cancel, leaving us with:

$$\begin{aligned}
 x + (0)y &= \frac{b_2 c_1 - b_1 c_2}{a_1 b_2 - a_2 b_1} \\
 (0)x + y &= \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1}
 \end{aligned}$$

Now we're ready to say we've solved the linear system for x and y , leaving us with

$$\begin{aligned}x &= \frac{b_2c_1 - b_1c_2}{a_1b_2 - a_2b_1} \\y &= \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}\end{aligned}\tag{7}$$

Substituting the values from Equation 6 into Equation 7 yields:

$$\begin{aligned}x &= \frac{(x_3 - x_4)(x_1y_2 - x_2y_1) - (x_1 - x_2)(x_3y_4 - x_4y_3)}{(y_2 - y_1)(x_3 - x_4) - (y_4 - y_3)(x_1 - x_2)} \\y &= \frac{(y_2 - y_1)(x_3y_4 - x_4y_3) - (y_4 - y_3)(x_1y_2 - x_2y_1)}{(y_2 - y_1)(x_3 - x_4) - (y_4 - y_3)(x_1 - x_2)}\end{aligned}\tag{8}$$

This formulation is identical to the one you'll find on Wikipedia (although they've arranged the denominators slightly differently, but still mathematically equivalent).

A caveat

Ironically, despite all that extra math, we'll still have problems with vertical lines, but only when those lines are *parallel*, which means there are either infinitely many solutions (lines are collinear) or zero solutions (lines are parallel but not collinear, like in Figure 3).

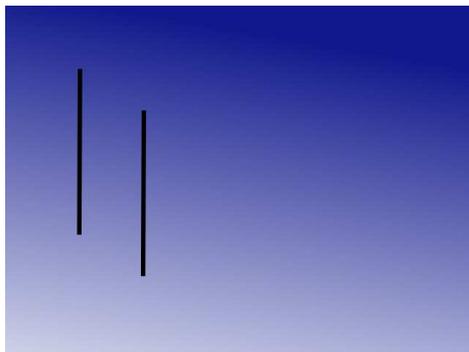


Figure 3: Two vertical lines would be problematic only because they're parallel

Checking for parallel lines is fortunately pretty simple.

Checking for parallel lines

Let's say we have two line segments floating around the Cartesian plane like in Figure 4. To check for whether these are parallel, first imagine translating the lines so that they both have an endpoint at $(0,0)$ (Figure 5).

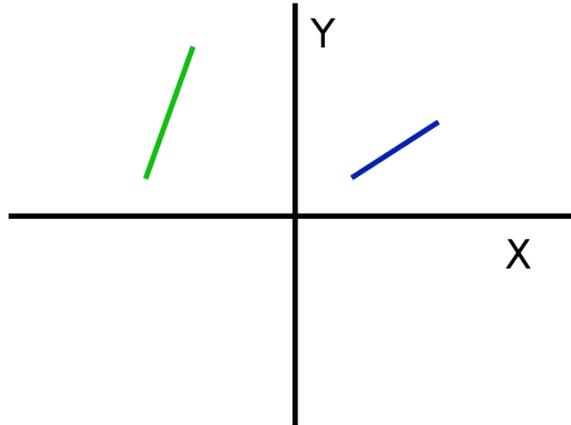


Figure 4: Two line segments in the x-y plane

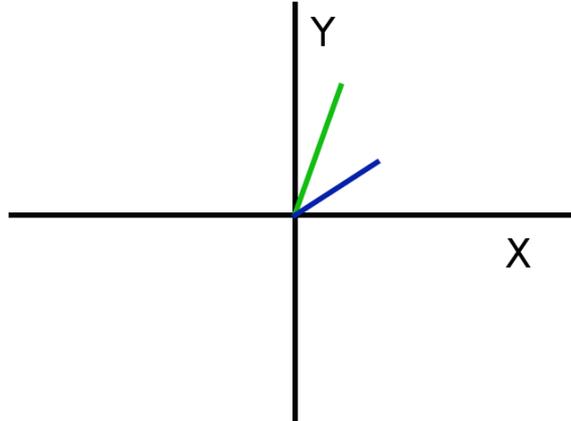


Figure 5: We've moved the lines so they start at the origin

There will be some angle θ between the two lines (Figure 6).

If this angle is 0, or π radians (180°) then $\sin \theta$ will be 0.

Given that we are representing our lines using two points, we can use those points to create *vectors*, which are like our lines from the origin $(0,0)$ that have a length and a magnitude.

To create a vector given points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, we just subtract one from the other (for our purposes here it doesn't matter the

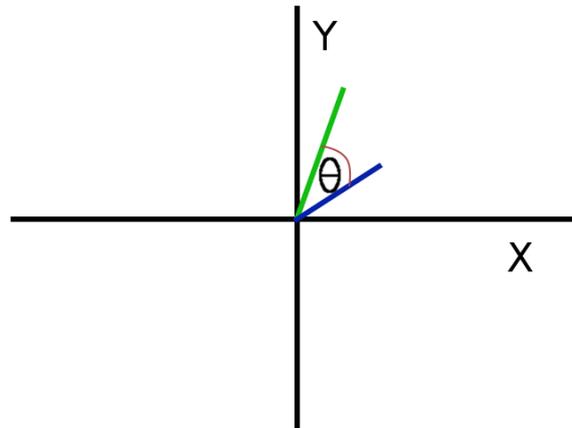


Figure 6: There is some angle, marked as θ between our lines

order!):

We draw them with little arrows to indicate a direction (Figure 7):

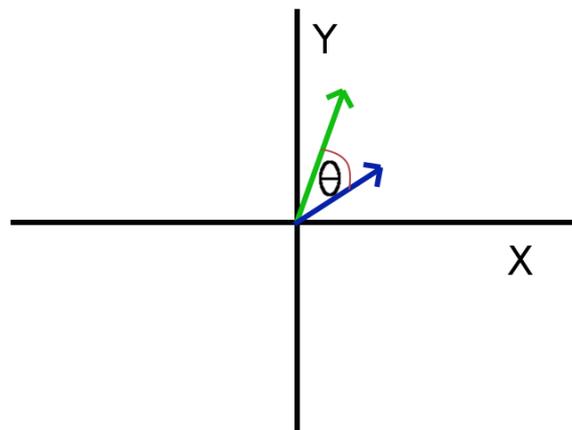


Figure 7: Our lines are now vectors

$$\vec{V}_1 = (x_2 - x_1, y_2 - y_1) \quad (9)$$

After we've created vectors for both our lines, we can take advantage of a mathematical relationship between them called the cross product to discover whether they're parallel or not. For two vectors \vec{V}_1 and \vec{V}_2 , their cross product is defined as

$$\|\vec{V}_1\| \|\vec{V}_2\| \sin \theta \quad (10)$$

Where $\|V_1\|$ and $\|V_2\|$ are the **magnitudes** of the vectors, which can be thought of as the vector's "length". For some vector $\vec{V}_1 = (X, Y)$, it's defined as

$$\|V_1\| = \sqrt{X^2 + Y^2} \quad (11)$$

However, for the sake of testing for parallel lines, what we're really interested in is the $\sin \theta$ term in the cross product, where θ is the smallest angle between the vectors (the vectors are simultaneously separated by θ and $2\pi - \theta$ radians).

Assuming our vectors don't have magnitudes of zero, when the cross product between them is zero (or really really close to zero) we consider the lines to be parallel because the angle between them must be zero. This tells us that the lines either never collide, or they're infinitely colliding because they're the same line (collinear). I leave checking for collinearity as an exercise to you.

A shortcut

One might say we solved our system of linear equations the long way. In fact, since we had exactly two equations and exactly two unknowns, we could have leveraged a mathematical technique known as Cramer's rule. The method is actually not so hard to apply, but perhaps its correctness is a little more difficult to understand.

Conclusions

Checking for line-line intersections is a harder problem than it appears to be on the surface, but once we give it a little thought it all boils down to algebra.

One application for line-line intersection testing is in computer graphics. Testing for intersections is one of the foundational subroutines for computing Binary Space-Partitioning (BSP) Trees, which are a way to efficiently represent a graphical scene, or even an individual object within a scene. BSP Trees are perhaps most famously known for their use by John Carmack in the Doom games.

In closing, we can consider a line in 2D to actually be a hyperplane that partitions the plane into areas "on one side or the other" the line. Therefore, there are analogs in 3D space where we use a 3D hyperplane (what we typically think of as a plane) to partition the space again into areas "on one side or the other" of the plane.

