Math Circle Worksheet
Fun with Matrices 2 a.k.a. Basic Linear Algebra

Wesley Hamilton

4/15/2017

1 Introduction

Last week we learned how to work with matrices. This week, we'll explore the heart of linear algebra: linear transformations. These are the nicest kinds of transformations on the plane, or things that you plug a vector into and get a vector back. The fact that these transformations are *linear* means it doesn't matter if we add two vectors and then apply the transformation, or apply the transformation to each vector and then add the results.

The goal of this worksheet is to understand rotations of the plane. We'll start by playing around with a few transformations and getting a feel for what they do to points in the plane, after which we'll focus on rotations. An interesting (possibly surprising?) connection to complex numbers will pop out along the way.

2 Matrices as Transformations

The way most (at least pure) mathematicians think of matrices are as transformations acting on some space[1]. In this section we'll explore that idea, and see how matrices act as certain kinds of transformations of the plane.

2.1 The geometry of the Euclidean plane

We usually describe a point in the plane by two numbers, its coordinates relative to the x- and y- axis. For example, the point $\begin{pmatrix} 5 \\ -1 \end{pmatrix}$ will sit five units to the right of the origin, and one unit below the x-axis. With our matrix glasses on, we think of this point as

$$\begin{pmatrix} 5 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (-1) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

[1] 3Blue1Brown on YouTube has a *great* video series about linear algebra, and it's accessible to anyone willing to learn! Here's a link to the preview video of the series: https://www.youtube.com/watch?v=kjBOesZCoqc. Alternatively, search for "3b1b linear algebra" in YouTube and the series should pop up right away.

This highlights the "going to the right five units, going down one unit" interpretation, since the vectors $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ act as a basis for our coordinate system. In general, a **basis** for the plane is a pair of vectors that you can use to describe any point in the plane. Since these are the usual vectors we use as a basis, we call them the **standard basis vectors**.

We can choose another basis for our coordinate system though, and that would change how we'd tell someone to "go to the right" or "go up." For example, if instead of using the standard basis $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$, we could use the vectors $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}$. Using this new basis, we would write

$$\begin{pmatrix}5\\-1\end{pmatrix} = 2\begin{pmatrix}1\\1\end{pmatrix} + 3\begin{pmatrix}1\\-1\end{pmatrix}.$$

This tell us that, to get to the point $\begin{pmatrix}5\\-1\end{pmatrix}$, we need to travel two units in the $\begin{pmatrix}1\\1\end{pmatrix}$ direction and three units in the $\begin{pmatrix}1\\-1\end{pmatrix}$ direction.

How do we find the 2 and 3 we needed above? The "brute-force" method is to solve for a and b in the equation

$$\begin{pmatrix}5\\-1\end{pmatrix} = a\begin{pmatrix}1\\1\end{pmatrix} + b\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}a\\a\end{pmatrix} + \begin{pmatrix}b\\-b\end{pmatrix} = \begin{pmatrix}a+b\\a-b\end{pmatrix}.$$

This tells us that, for the vectors on the very left and very right to be equal, we need $5 = a + b$ and $-1 = a - b$. Solving that system gives $a = 2, b = 3$.

Exercise 2.1 Draw the vectors $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}$ on a plane. How do these basis vectors compare with the standard basis vectors?

Exercise 2.2 How would you express the vector $\begin{pmatrix}5\\-1\end{pmatrix}$ in terms of the basis $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}$?

Exercise 2.3 Would the vectors $\begin{pmatrix}1\\-3\end{pmatrix}, \begin{pmatrix}-2\\6\end{pmatrix}$ make a good set of basis vectors? In other words, would we be able to describe the location of, say, the point $\begin{pmatrix}5\\-1\end{pmatrix}$ using $\begin{pmatrix}1\\-3\end{pmatrix}, \begin{pmatrix}-2\\6\end{pmatrix}$? Why or why not? If you're stuck try drawing a picture.

Exercise 2.4 Compute the determinants of the matrices $\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}, \begin{pmatrix}1 & 1\\1 & 0\end{pmatrix}$, and $\begin{pmatrix}1 & -3\\-2 & 6\end{pmatrix}$. What do you notice? How does this connect to the previous exercises?

Exercise 2.5 What do you think is the connection between the determinant and suitable basis vectors?

The last exercise should lead to a conjecture which turns out to be true, though we won't prove it here.

Intuitively, in the plane a basis of vectors should be a collection of two vectors that point in different directions, a left/right and up/down direction, or something similar. If the two vectors actually point in the same direction (like in exercise 3.4), we won't be able to move in two directions, so only the points that lay on that line can be described by the basis. The concept underlying this is **linear independence**, and the determinant is a great way for detecting whether a pair of vectors (in the plane) is linearly independent.

2.2 How do matrices act on points?

Now, we'll start to treat matrices as *maps*, instead of collections of numbers. This interpretation of matrices is central to the next two sections.

Exercise 2.6 What is $F \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$? Plot $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $F \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ on the same graph. Now plot the points $F^2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $F^3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, etc. What do you notice? Where are the points going?

Exercise 2.7 Do the same as in the last exercise, this time with the point $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Where do these points head off to?

Exercise 2.8 * Finally, do the same exercise but with the point $\begin{pmatrix} \phi_2 \\ 1 \end{pmatrix}$, where ϕ_2 satisfies ${\phi_2}^2 - \phi_2 - 1 = 0$. The exact value of ϕ_2 is $\frac{1-\sqrt{5}}{2}$ and is approximately -0.618. What happens in this case? This vector is actually one of the eigenvectors of F.

If you were to do this for a lot of points, and draw a line connecting the points in each orbit, you might get something that looks like the figure below. This is one kind of behaviour that matrices, acting as transformations, can give.

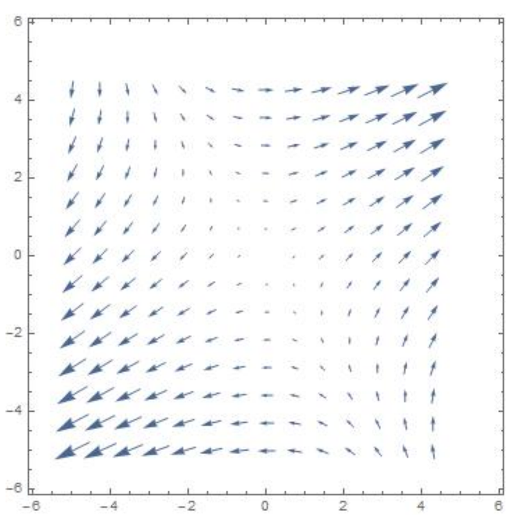

Figure 1: The behaviour of the map F.

Exercise 2.9 Next, try the same procedure as above but with the matrix $R_{90^\circ} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Some good points to test it on are $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Why do you think I called it R_{90°?

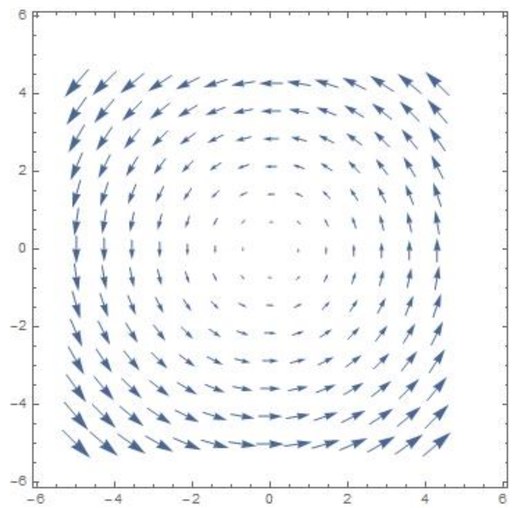

Figure 2: The behaviour of the map R_{90°.

Exercise 2.10 Try the same thing, this time with the matrices $Pr_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $Pr_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Any points are good points to test these maps on! What happens in each case? Why do you think I called them Pr_1 and Pr_2?

2.3 (Extra) Why are they called linear transformations?

As mentioned in the introduction, linear transformations are nice because it doesn't matter if you add vectors and then apply the transformation, or apply the transformation and then add the vectors. Let's make that a bit more precise.

A transformation T of the plane is **linear** if the following two conditions hold:

1. $T \cdot \left(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right) = T \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + T \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$;

2. $T \cdot \begin{pmatrix} ca_1 \\ ca_2 \end{pmatrix} = c \left(T \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right)$.

The first condition is the one already mentioned, while the second is new(ish). The second condition tells you that if you apply a transformation to a scaled vector, then you get the same thing as if you applied the transformation to the non-scaled vector, then scale that. Let's work a few examples to get a better feel for these conditions.

Exercise 2.11 Consider the transformation $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Check that $A \cdot \left(\begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix} \right) = A \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + A \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

Exercise 2.12 Now check that $A \cdot \begin{pmatrix} ca_1 \\ ca_2 \end{pmatrix} = c \left(A \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right)$.

Exercise 2.13 Next, consider the transformation $B = \begin{pmatrix} 6 & 3 \\ 2 & 1 \end{pmatrix}$. What is $B \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix}$? Using this, and the fact that B is linear, show that $B \cdot \begin{pmatrix} 1 + a_1 \\ -2 + a_2 \end{pmatrix} = B \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ for any real numbers a_1, a_2.

It turns out that linear transformations on the plane and 2×2 matrices are the exact same thing; this means that every 2×2 matrix is a linear transformation, and if you're given any linear transformation then you can find a matrix representation for it. Not every transformation on the plane is linear though, as we'll now see with a fairly simple example.

Exercise 2.14 Let T_a be translation by $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, so that

$$T_a \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 + a_1 \\ v_2 + a_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Show that T_a is *not* linear. Hint: Compute $c \left(T_a \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right)$ and $T_a \cdot \begin{pmatrix} cv_1 \\ cv_2 \end{pmatrix}$. Are these the same?

3 Complex Numbers as Matrices

In this section, we'll explore a connection between complex numbers and 2×2 matrices. If you've never seen complex or imaginary numbers before then don't worry!

3.1 Two important matrices

As before we have $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and we'll define $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. It's not a coincidence that J is the same as R_{45° from before, as we'll soon find out.

Exercise 3.1 What is $J^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$?

In otherwords, we've found a solution J to the (matrix) equation $x^2 + I = 0$. If you've played with polynomials before, you should recognize that the corresponding equation $x^2 + 1 = 0$ doesn't have any solutions in the real numbers. This is because no matter which number we plug in for x, say $x = 6, -7$, or even $-1,000$, x^2 will be ≥ 0. Adding one then gives $x^2 + 1 \geq 1$ for every real x, so $x^2 + 1 = 0$ has no solutions in the reals.

This is the point where people define the imaginary number $i = \sqrt{-1}$ to provide a solution to $x^2 + 1 = 0$, from which complex numbers are born. In our case we've done a similar thing, except instead of defining a new type of number, we've just produced a matrix of a certain form! Also, I'm calling this matrix "$\sqrt{-I}$" $= J$ because (capital) i is already being used.

Complex numbers are of the form $a_1 + a_2 i$, where a and b are both real numbers. With this in mind, we'll define our "complex matrices" as 2×2 matrices of the form $a_1 I + a_2 J$, where a_1, a_2 are real numbers. Thus, the matrices in this collection will all look like

$$a_1 I + a_2 J = \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix}.$$

Exercise 3.2 Is the sum of two "complex matrices" a "complex matrix"? Why or why not?

Exercise 3.3 What is the product of two "complex matrices" $a_1 I + a_2 J$ and $b_1 I + b_2 J$? Is it also a "complex matrix"?

Exercise 3.4 For what a_1, a_2 is the matrix $a_1 I + a_2 J$ invertible? Remember to look at the determinant. For these a_1, a_2, what is the inverse matrix? Express it in the form $b_1 I + b_2 J$.

Exercise 3.5 Using the previous exercises, what do you think is the connection between these "complex matrices" and complex numbers?

3.2 "Complex Matrices" as Rotations

In this last bit we'll further explore rotation matrices, and how these "complex matrices" act as rotations.

To start recall that $\sin(0) = 0$ and $\cos(0) = 1$, as well as $\sin(90°) = 1$ and $\cos(90°) = 0$. Let's use these and rewrite

$$I = R_{0°} = \begin{pmatrix} \cos(0) & -\sin(0) \\ \sin(0) & \cos(0) \end{pmatrix}, \quad R_{90°} = \begin{pmatrix} \cos(90°) & -\sin(90°) \\ \sin(90°) & \cos(90°) \end{pmatrix}.$$

Motivated by this, let's define a rotation matrix to be of the form

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

where θ can be any angle.

It turns out that rotation matrices do in fact correspond to rotations, which you may have noticed in exercise 4.9. Let's check that $R_{45°}$ is a rotation by 45° manually. For reference, $\sin(45°) = \cos(45°) = \frac{1}{\sqrt{2}}$.

Exercise 3.6 Where should the point $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ be sent by $R_{45°}$? (You don't need to do any matrix multiplication at this point, think unit circle). Now compute $R_{45°} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$; is it where it should be?

A fact we'll need is composition of maps is the same as matrix multiplication. In other words, suppose we have two maps A and B. First we send the point $\begin{pmatrix}1\\0\end{pmatrix}$ to the point $B\cdot\begin{pmatrix}1\\0\end{pmatrix}$, and then we send $B\cdot\begin{pmatrix}1\\0\end{pmatrix}$ to $A\cdot\left(B\cdot\begin{pmatrix}1\\0\end{pmatrix}\right)$. Another way of seeing this is by drawing some arrows:

$$\begin{pmatrix}1\\0\end{pmatrix}\longrightarrow B\cdot\begin{pmatrix}1\\0\end{pmatrix}\longrightarrow A\cdot\left(B\cdot\begin{pmatrix}1\\0\end{pmatrix}\right).$$

This, in fact, is the same as sending the point $\begin{pmatrix}1\\0\end{pmatrix}$ to the point $(A\cdot B)\cdot\begin{pmatrix}1\\0\end{pmatrix}$, where we first compute the product $A\cdot B$, and then multiply that with $\begin{pmatrix}1\\0\end{pmatrix}$. The diagram in this case becomes

$$\begin{pmatrix}1\\0\end{pmatrix}\longrightarrow (A\cdot B)\cdot\begin{pmatrix}1\\0\end{pmatrix}.$$

With this in mind we can see that a rotation by θ_1 degrees, then θ_2 degrees, is the same as if we had rotated by $\theta_1+\theta_2$ degrees! Exactly what our experience tells us should happen!

Exercise 3.7 Check this with $\theta_1=\theta_2=45°$. In other words, show that $R_{45°}\cdot R_{45°}=R_{45°+45°}=R_{90°}$.

Exercise 3.8 Use this fact (that $R_{\theta_1}\cdot R_{\theta_2}=R_{\theta_1+\theta_2}$) to re-derive the sine and cosine angle addition formulas:

$$\cos(\theta_1+\theta_2)=\cos(\theta_1)\cos(\theta_2)-\sin(\theta_1)\sin(\theta_2),$$
$$\sin(\theta_1+\theta_2)=\sin(\theta_1)\cos(\theta_2)+(\cos\theta_1)\sin(\theta_2).$$

Coming back to the "complex" case, notice that we can write

$$R_\theta=\cos\theta I+\sin\theta J,$$

which bears remarkable resemblance to **Euler's formula**[2] $e^{i\theta}=\cos\theta+i\sin\theta$.

Exercise 3.9 Check explicitly that $R_\theta=\cos\theta I+\sin\theta J$ by writing out the matrices for the left- and right-hand sides.

This suggests[3] that the matrix equivalent of $e^{i\theta}$ is precisely R_θ, so multiplying a complex number by $e^{i\theta}$ amounts to rotating that complex number in the plane by an angle θ. This gives another proof of the formula $e^{i\pi}=-1$, where π corresponds to our 180°, after identifying the real number 1 with the vector $\begin{pmatrix}1\\0\end{pmatrix}$ and the imaginary number i with the vector

[2]In Euler's formula, the θ is given by **radians** and not degrees, but for this worksheet that distinction isn't important (outside of this worksheet it is though!).

[3]3B1B has another great video on Euler's formula, apply titled "Euler's formula with introductory group theory."

$\begin{pmatrix}0\\1\end{pmatrix}$: we first compute $\sin(180°) = 0, \cos(180°) = -1$, so

$$\begin{aligned} e^{i\pi} = e^{i\pi} \cdot 1 \sim R_{180^\circ} \cdot \begin{pmatrix}1\\0\end{pmatrix} &= \begin{pmatrix}-1 & 0\\0 & -1\end{pmatrix} \cdot \begin{pmatrix}1\\0\end{pmatrix} \\ &= \begin{pmatrix}-1\\0\end{pmatrix} \\ &\sim -1. \end{aligned}$$