

Lecture 4

Acting on \mathbb{R}^2

4.1 Goals

- (a) Understanding we can represent elements of the x-y plane as column vectors.
- (b) Seeing that each 2-by-2 matrix acts on \mathbb{R}^2
- (c) Understanding why linearity allows us to understand matrices through their columns
- (d) Practicing how to visualize a 2-by-2 matrix's action by using images of squares

4.2 \mathbb{R}^2

Recall that \mathbb{R}^2 is the Euclidean plane. As a set, it consists of all ordered pairs of real numbers. For example, the following are all elements of \mathbb{R}^2 :

$$(0, 0) \quad (\pi, -1) \quad (\sqrt{2}, 5) \quad (0, 7) \quad (-2, -2).$$

You may be used to writing elements of \mathbb{R}^2 as (x, y) . You are also probably used to drawing the set \mathbb{R}^2 as the x-y plane. In this class, and in most advanced math classes, we'll often say \mathbb{R}^2 (this is pronounced "R2," as in the Star Wars droid R2-D2) instead of "the xy-plane."

Remark 4.2.1. More generally, if X and Y are sets, one can define the *direct product*, or *Cartesian product* of X and Y as the set whose elements are exactly ordered pairs (x, y) with $x \in X$ and $Y \in Y$. The notation $X \times Y$ is often used for this direct product. (This is because—if X and Y are finite sets—then the number of elements of $X \times Y$ is the product of the number of elements in X and in Y . The multiplicative notation invokes this phenomenon.) When $X = Y$, we often write X^2 .

Indeed, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.

4.2.1 Elements of \mathbb{R}^2 as column vector

To specify an element of \mathbb{R}^2 is to specify two real numbers. There is another mathematical object specified by two real numbers: a 2-by-1 matrix. Using this fact, we will often choose to write an element of \mathbb{R}^2 (that is, a point on the x-y plane) as such a matrix.

Example 4.2.2. The point $(7, 3)$ in the x-y plane can be written as the matrix

$$\begin{pmatrix} 7 \\ 3 \end{pmatrix}.$$

We will call such vertical matrices *column vectors*. Writing elements of \mathbb{R}^2 as matrices, or as column vectors, is quite convenient: It becomes clear that matrices can act on elements of \mathbb{R}^2 .

Notation 4.2.3. We will often denote a column vector by the bolded letter \mathbf{x} . When you hand-write this notation, you may write it as \underline{x} or as \vec{x} . We will denote the entries of \mathbf{x} as x_1, x_2 .

Example 4.2.4. If $\mathbf{x} \in \mathbb{R}^2$, we will write

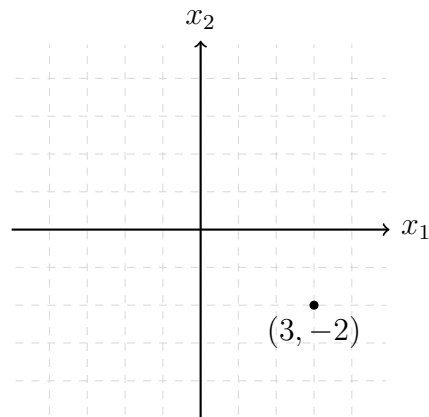
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Remark 4.2.5. You are most likely used to write (x, y) for a point on the x-y plane. But letters are precious, so we'll often choose instead to write (x_1, x_2) for an element of the x-y plane. So, confusingly, each of the following is a way to indicate an element of \mathbb{R}^2 :

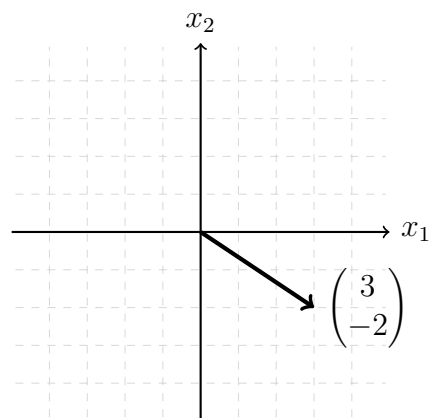
$$(x, y) \quad \begin{pmatrix} x \\ y \end{pmatrix} \quad (x_1, x_2) \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mathbf{x}.$$

4.2.2 Drawing points on \mathbb{R}^2 as vectors

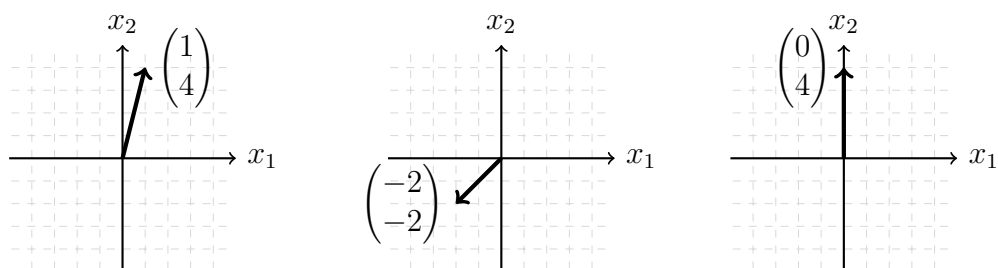
You are most likely used to drawing \mathbb{R}^2 , and a point in \mathbb{R}^2 , as follows:



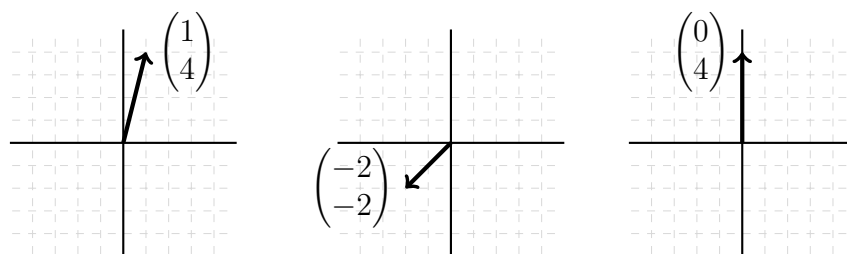
For various reasons, we will instead use the convention of drawing an *arrow* from the origin to the point we want to indicate. When we do this, we will think of this arrow as a visual depiction of a column vector. So the point $(3, -2)$ above will be drawn as follows when thought of as a vector:



Here are some more drawings of vectors:



We may often be lazy and delete the x_1, x_2 labels from our axes, along with the arrowheads of the axes. So the above pictures may equally well be drawn as follows:



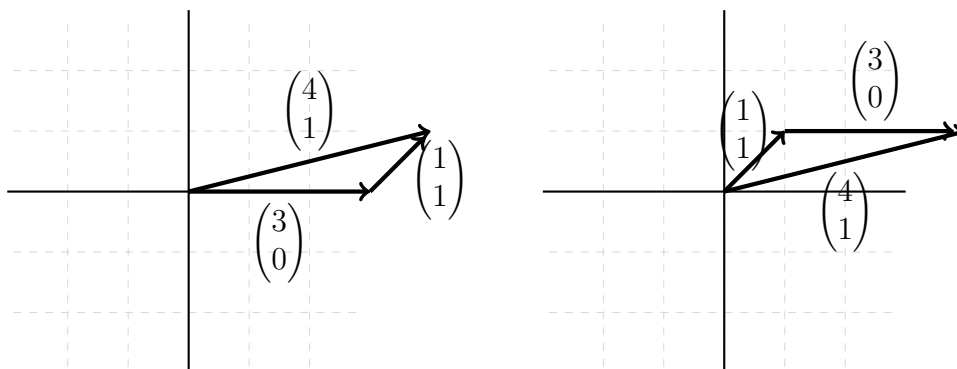
4.2.3 Adding and scaling column vectors

Given two column vectors, you can add them:

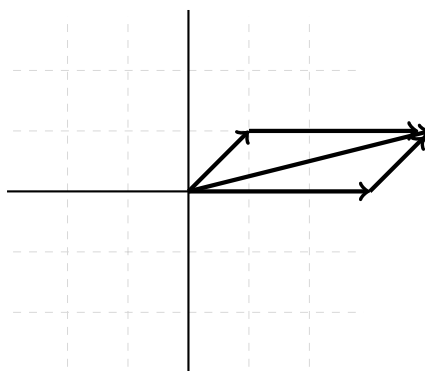
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}.$$

(This is usual matrix addition.) We can visualize this as well, by putting one vector's tail at the head of another. (The "head-to-tail" method.) The

picture below shows how $\begin{pmatrix} 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$.



Of course, as shown on the right, we could obtain the same vector by first drawing the $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ vector, then adding on the $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$ vector. This is because addition of matrices (and hence column vectors) is commutative. So sometimes, people also draw a parallelogram to indicate the vector sum:

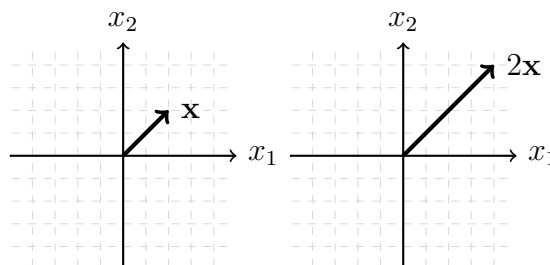


You can also scale a column vector. More precisely, fix a real number λ and a column vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Then we define

$$\lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}.$$

In other words, we multiply all the coordinates of \mathbf{x} by λ .

This is called *scaling* for a good reason. Multiplying a vector by λ “scales” the vector by a factor of λ .



4.3 2-by-2 matrices act on \mathbb{R}^2

We saw in last lecture that for any integer $n \geq 1$ and for any ring R , $M_n(R)$ is also a ring. From hereon, we will focus on the case $n = 2$ and $R = \mathbb{R}$. (That is, the case of 2-by-2 matrices whose entries are real numbers.)

Let's fix a 2-by-2 matrix with real entries. We'll call the matrix A and write its entries as

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then A determines a function from \mathbb{R}^2 to \mathbb{R}^2 as follows:

$$\mathbf{x} \mapsto A\mathbf{x}$$

where $A\mathbf{x}$ denotes matrix multiplication.

Notation 4.3.1. The arrow \mapsto is an arrow with a vertical tail. It should be read as “maps to.” This arrow is used to describe a function's input and output. The tail of the arrow is closest to the input, while the head of the arrow is closest to the output. Hence, the notation $\mathbf{x} \mapsto A\mathbf{x}$ describes a function which, given an input called \mathbf{x} , outputs $A\mathbf{x}$.

We often say that \mathbf{x} is “sent to” $A\mathbf{x}$.

Explicitly, given an element of \mathbb{R}^2 , which we think of as a column vector \mathbf{x} , we obtain another element of \mathbb{R}^2 by multiplying \mathbf{x} on the left with A . If we give \mathbf{x} coordinates x_1, x_2 , this can be explicitly written out:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}.$$

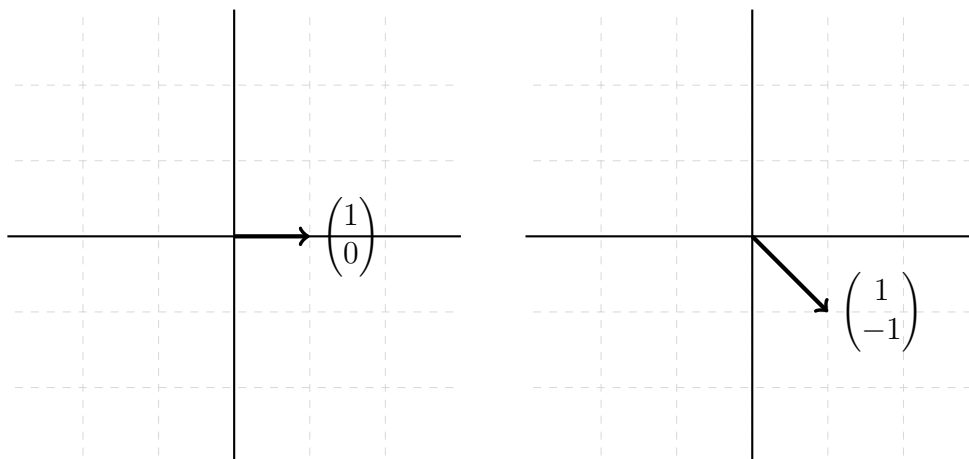
Example 4.3.2. Let

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}.$$

Then the column vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is sent to the column vector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, as shown by the following computation:

$$A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 1 \cdot 0 \\ -1 \cdot 1 + 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Once we perform this computation, we can try to graphically represent this input-output relation as follows:

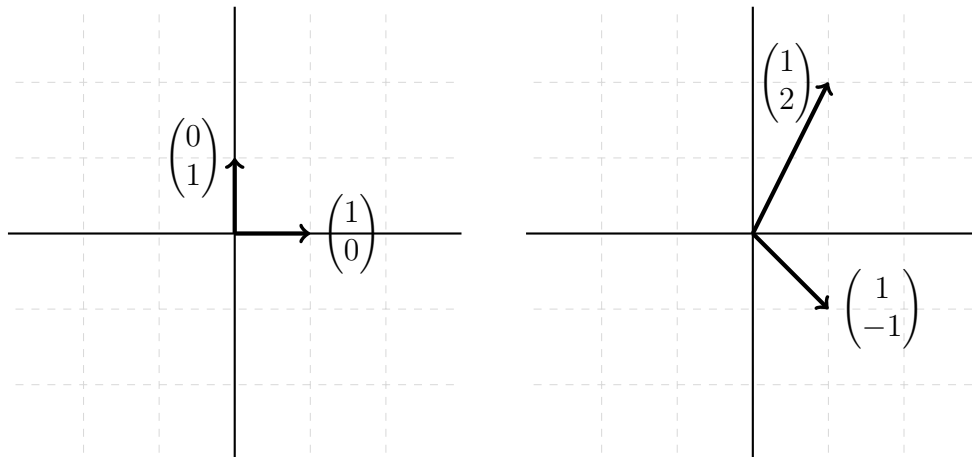


(The matrix A sends the vector on the left to the vector on the right.) Likewise, $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ is sent to $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$, as shown by the following computation:

$$A \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 1 \cdot 1 \\ -1 \cdot 2 + 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

I'll leave it to you to check that $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is sent to $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. We can then draw the

effect of A as follows:



(A takes the vectors on the left to the vectors on the right.)

4.4 Understanding how a matrix acts on \mathbb{R}^2

Let A be a matrix.

Proposition 4.4.1 (Matrices are linear transformations). Let A be a 2-by-2 matrix and fix $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Further, let λ be a real number. Then

$$A(\lambda\mathbf{x}) = \lambda A(\mathbf{x}) \quad \text{and} \quad A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}.$$

(Note that \mathbf{x} and \mathbf{y} refer to two elements of \mathbb{R}^2 —in particular, \mathbf{y} is not a coordinate, but a column vector.)

We'll prove the proposition shortly, but let's talk about the upshot. The proposition tells us that we can understand what a matrix does on all of \mathbb{R}^2 fairly straightforwardly. After all, any element of \mathbb{R}^2 is of the form

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

So the proposition tells us that

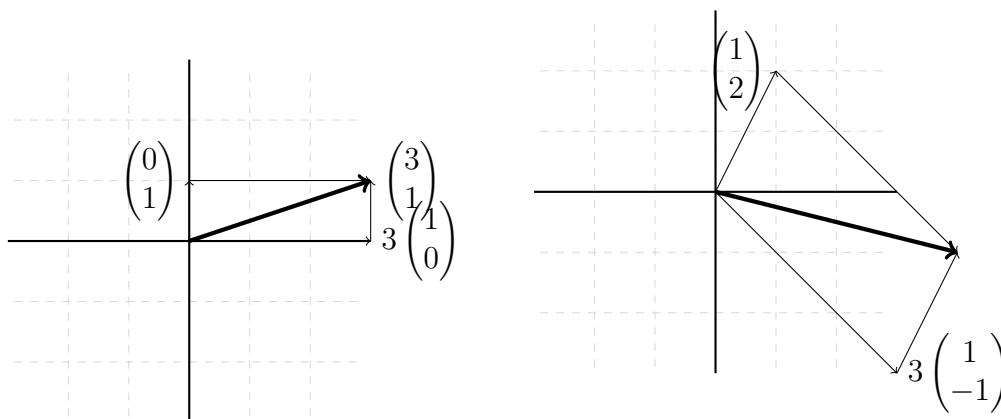
$$A\mathbf{x} = x_1 A \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 A \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In other words, if you understand what A does to $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and to $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, you understand what A does to all other vectors.

Example 4.4.2. Let

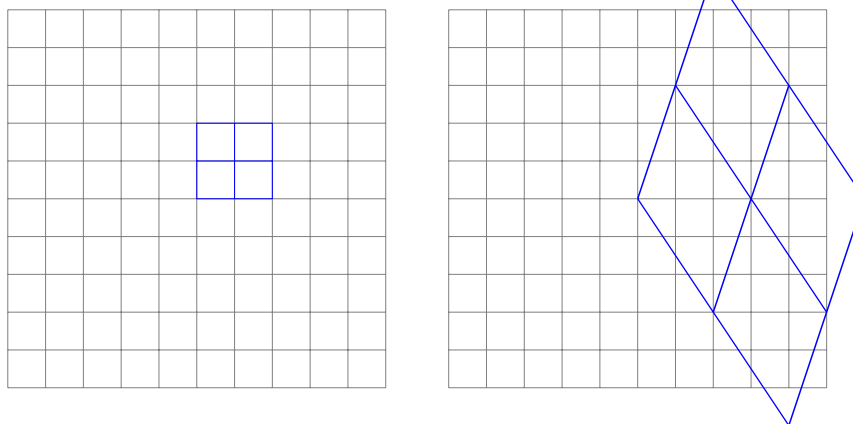
$$A = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}.$$

Let's try to understand what A does to the vector $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Well, we can write $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ as $3 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. See the drawing:



As you know, the vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ generate a square grid in \mathbb{R}^2 .

When we apply a matrix A , the vectors $A \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $A \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ often generate a new grid. This is a very good way to visualize what A is doing to \mathbb{R}^2 :



The above picture shows how the matrix A , when inputting the blue grid on the left, outputs the blue grid on the right. (The standard square grid is shown also in gray, on both images.) As you can see, A not only turns squares into tilted parallelograms, it also increases the area of each square in the grid considerably.

Proof of Proposition 4.4.1. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

$$\begin{aligned} A(\lambda\mathbf{x}) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix} \\ &= \begin{pmatrix} a \cdot \lambda x_1 + b \cdot \lambda x_2 \\ c \cdot \lambda x_1 + d \cdot \lambda x_2 \end{pmatrix} \\ &= \begin{pmatrix} \lambda(ax_1 + bx_2) \\ \lambda(cx_1 + dx_2) \end{pmatrix} \\ &= \lambda \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} \\ &= \lambda A\mathbf{x}. \end{aligned}$$

Letting $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ we have

$$\begin{aligned} A(\mathbf{x} + \mathbf{y}) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \\ &= \begin{pmatrix} a(x_1 + y_1) + b(x_2 + y_2) \\ c(x_1 + y_1) + d(x_2 + y_2) \end{pmatrix} \\ &= \begin{pmatrix} (ax_1 + bx_2) + (ay_1 + by_2) \\ (cx_1 + dx_2) + (cy_1 + dy_2) \end{pmatrix} \\ &= \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} + \begin{pmatrix} ay_1 + by_2 \\ cy_1 + dy_2 \end{pmatrix} \\ &= A\mathbf{x} + A\mathbf{y}. \end{aligned}$$

□

4.5 Exercises

Exercise 4.5.1. Choose any real number θ . Convince yourself that

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

is a matrix that rotates \mathbb{R}^2 by θ radians about the origin, counterclockwise.

How would you explain to a friend that this is true?

Exercise 4.5.2. Consider the square formed by the vertices $(0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$ in \mathbb{R}^2 . For each of the following matrices, draw the image of this square under A .

(a) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

Exercise 4.5.3. Consider the square formed by the vertices $(0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$ in \mathbb{R}^2 . For each of the following matrices, draw the image of this square under A .

(a) $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

(b) $\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$

(c) $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

Exercise 4.5.4. Consider the square formed by the vertices $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$ in \mathbb{R}^2 . For each of the following matrices, draw the image of this square under A .

(a) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

(c) $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$

Exercise 4.5.5. Consider the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Prove that the square formed by the vertices $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$ is sent by A to a parallelogram of area $ad - bc$.

Exercise 4.5.6. Consider the matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Let $\mathbf{x} \in \mathbb{R}^2$. Show that $I\mathbf{x} = \mathbf{x}$.

Exercise 4.5.7. Consider the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

If $ad - bc \neq 0$, consider further the matrix

$$B = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

Prove that $AB = I$ and $BA = I$. For this reason, we will often write B as A^{-1} .

Exercise 4.5.8. A *line* in \mathbb{R}^2 is a subset $L \subset \mathbb{R}^2$ for which the following holds: There exists $\mathbf{x}_0 \in \mathbb{R}^2$ and $\mathbf{v} \in \mathbb{R}^2$ so that

$$L = \{\mathbf{x}_0 + t\mathbf{v}, t \in \mathbb{R}\}$$

Let L be a line, and let

$$A(L) = \{\mathbf{y} \in \mathbb{R}^2 \mid \mathbf{y} = A\mathbf{x} \text{ for some } \mathbf{x} \in L\}$$

be the image of L under A . Prove that $A(L)$ is also a line.

4.6 Extra Credit: Fibonacci numbers

Consider the matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

Let $\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and define inductively

$$\mathbf{x}_n = A\mathbf{x}_{n-1}.$$

For example,

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and

$$\mathbf{x}_3 = A\mathbf{x}_2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

- (a) Let F_n be the n th Fibonacci number. (I will use the convention that the Fibonacci sequence begins $0, 1, 1, 2, 3, \dots$, so $F_1 = 0, F_2 = 1$, et cetera.)

Prove that

$$\mathbf{x}_n = \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix}$$

- (b) Consider the two vectors

$$\mathbf{s}_1 = \begin{pmatrix} 1 + \sqrt{5} \\ 2 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} 1 - \sqrt{5} \\ 2 \end{pmatrix}.$$

Prove that

$$A\mathbf{s}_1 = \frac{1 + \sqrt{5}}{2}\mathbf{s}_1, \quad A\mathbf{s}_2 = \frac{1 - \sqrt{5}}{2}\mathbf{s}_2$$

- (c) Show that there exist two real numbers a_1, a_2 so that

$$\mathbf{x}_1 = a_1\mathbf{s}_1 + a_2\mathbf{s}_2.$$

- (d) More generally, prove that

$$\mathbf{x}_{n+1} = a_1 \left(\frac{1 + \sqrt{5}}{2} \right)^n \mathbf{s}_1 + a_2 \left(\frac{1 - \sqrt{5}}{2} \right)^n \mathbf{s}_2.$$

- (e) By noting that $|\frac{1-\sqrt{5}}{2}| < 1$, show that \mathbf{x}_n becomes closer and closer to the line spanned by \mathbf{s}_1 as $n \rightarrow \infty$. Conclude that, as n goes to infinity, the ratio of consecutive Fibonacci numbers approaches the golden ratio.