Transformation matrices and the image of the subset

Matrices can be extremely useful when it comes to describing what would otherwise be complex transformations in space.

For instance, let's say you want to know what happens if you take every point in the coordinate plane, shift it up vertically by 2 units, and stretch it out horizontally by 5 units.

Up to now, we don't have a simple way to describe this change mathematically. But that's where transformation matrices come in. They allow us to organize the transformation information into a matrix, which will tell us exactly how every vector in the space should move.

In the last lesson, we learned that a transformation was basically just a function that takes in a vector and transforms it into another vector. In this lesson we want to look at turning a transformation into a matrix, so let's start just by looking at what happens when we use a transformation to move a vector.

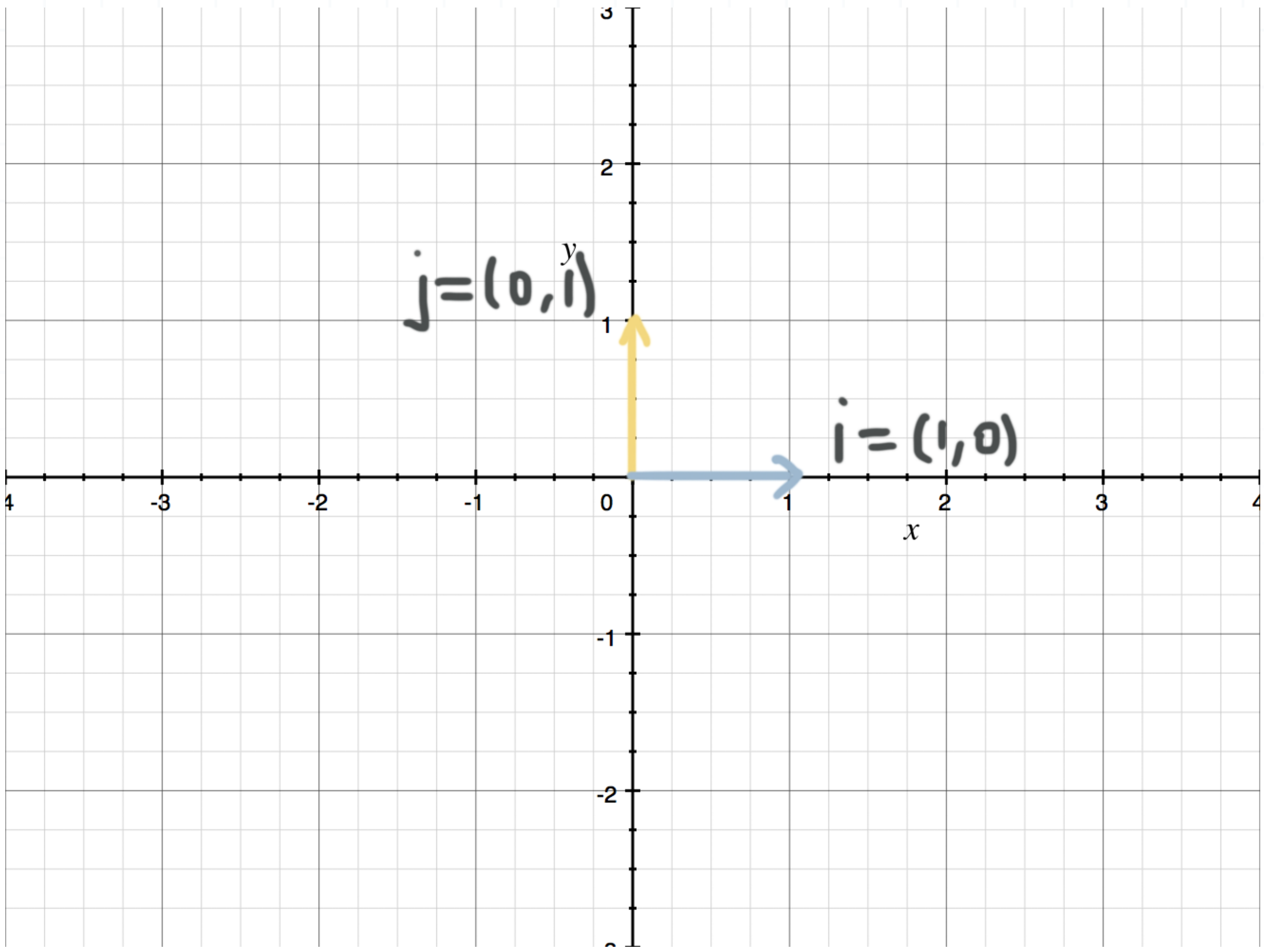
Understanding the transformation matrix

Think about transformations as a series of shifts, stretches, compressions, rotations, etc., that move a vector, or set of vectors, from one position to another. For instance, let's say we have the transformation matrix

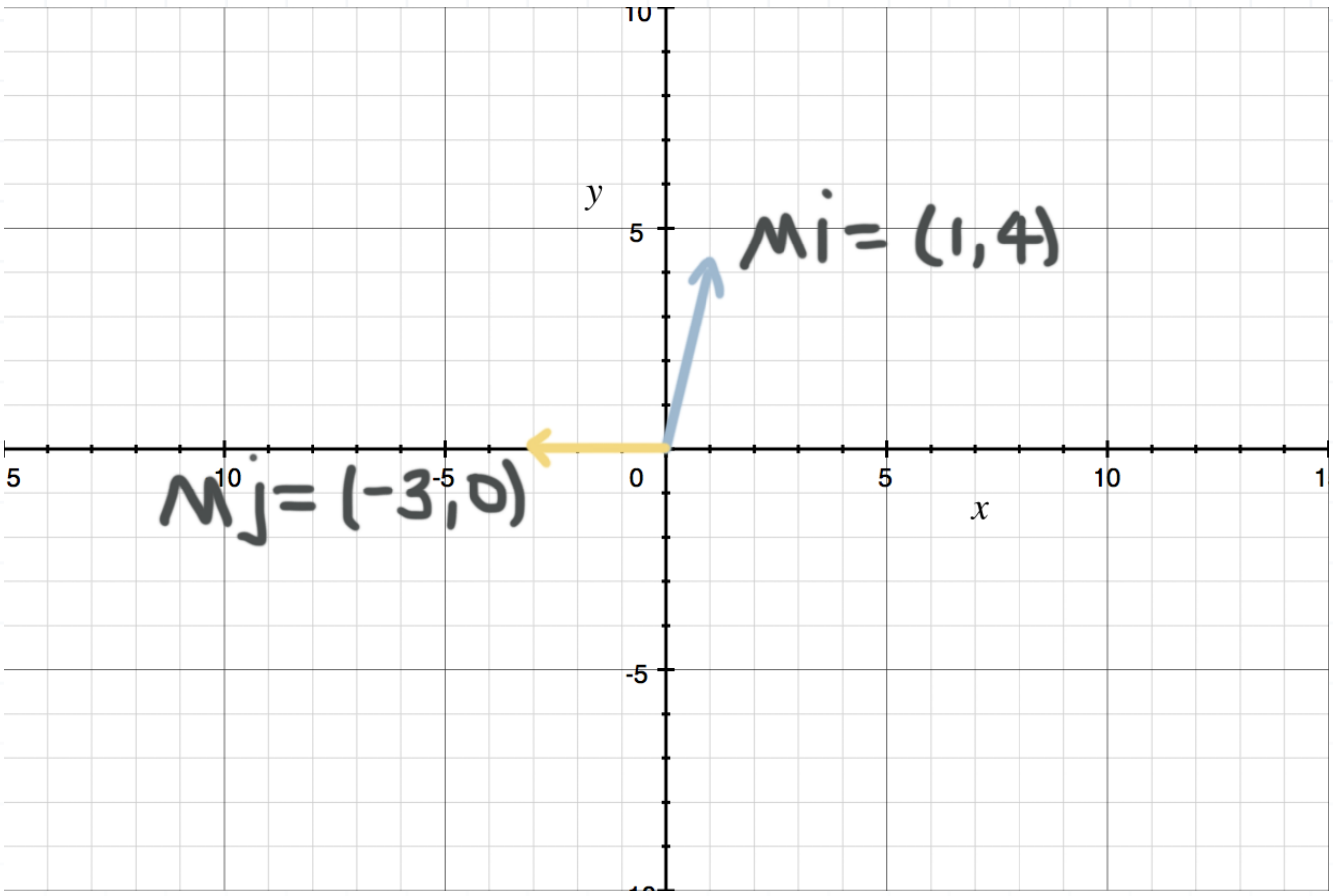
$$M = \begin{bmatrix} 1 & -3 \\ 4 & 0 \end{bmatrix}$$

What we want to understand is how the entries in M work together to transform a vector (or vector set).

In a 2×2 transformation matrix, the first column (in this case $(1,4)$) tells us where the standard basis vector $\mathbf{i} = (1,0)$ will land after the transformation. The second column (in this case $(-3,0)$) tells us where the standard basis vector $\mathbf{j} = (0,1)$ will land after the transformation. In other words, given the vectors $\mathbf{i} = (1,0)$ in light blue and $\mathbf{j} = (0,1)$ in yellow,



the transformation matrix M changes $(1,0)$ into $(1,4)$, and changes $(0,1)$ into $(-3,0)$.

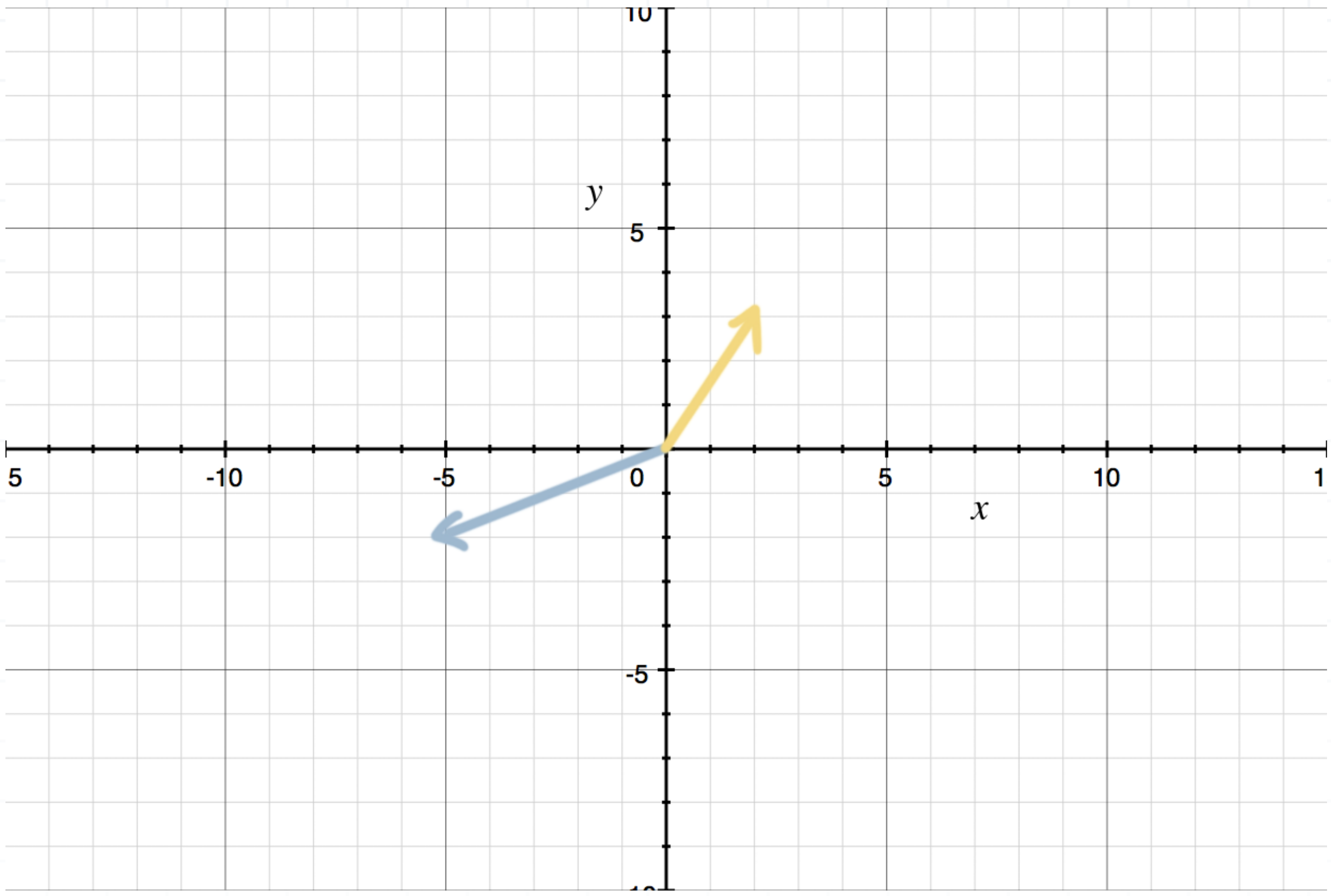


The vector (or vector set) before any transformation has been applied is called the **preimage**, and the vector (or vector set) after the transformation has been applied is called the **image**.

With this in mind, let's do an example where we start with the preimage and image and work backwards to find the transformation matrix.

Example

The graph shows the light blue vector $(1,0)$ and the yellow vector $(0,1)$, after a transformation has been applied. Find the transformation matrix that did the transformation.



The light blue vector now points to $(-5,-2)$. Since the first column of the transformation matrix represents where the unit vector $(1,0)$ lands after a transformation, we can fill in the first column of the transformation matrix.

$$\begin{bmatrix} -5 & \\ -2 & \end{bmatrix}$$

The yellow vector now points to $(2,3)$. Since the second column of a transformation matrix represents where the unit vector $(0,1)$ lands after a transformation, we can fill in the second column of the transformation matrix.

$$\begin{bmatrix} -5 & 2 \\ -2 & 3 \end{bmatrix}$$

This is the matrix that describes the transformation happening in coordinate space when $(1,0)$ moves to $(-5,-2)$ and when $(0,1)$ moves to $(2,3)$.

We can double-check this by multiplying the transformation matrix we just found by the matrix of column vectors of $\mathbf{i}$ and $\mathbf{j}$.

$$\begin{bmatrix}-5 & 2\\-2 & 3\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

$$\begin{bmatrix}-5(1)+2(0) & -5(0)+2(1)\\-2(1)+3(0) & -2(0)+3(1)\end{bmatrix}$$

$$\begin{bmatrix}-5+0 & 0+2\\-2+0 & 0+3\end{bmatrix}$$

$$\begin{bmatrix}-5 & 2\\-2 & 3\end{bmatrix}$$

There's a really important conclusion that we can draw from this last example problem. The transformation matrix we found doesn't just transform the individual vectors $\mathbf{i}=(1,0)$ and $\mathbf{j}=(0,1)$, it models the transformation of every vector in the coordinate plane!

Therefore, armed with this transformation matrix, you can now figure out the transformed location of any other vector in the plane. For instance, let's say you want to know what happens to $\vec{v}=(5,3)$ under the same transformation. Simply multiply the transformation matrix by the vector $\vec{v}$:

$$\begin{bmatrix} -5 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} -5(5) + 2(3) \\ -2(5) + 3(3) \end{bmatrix}$$

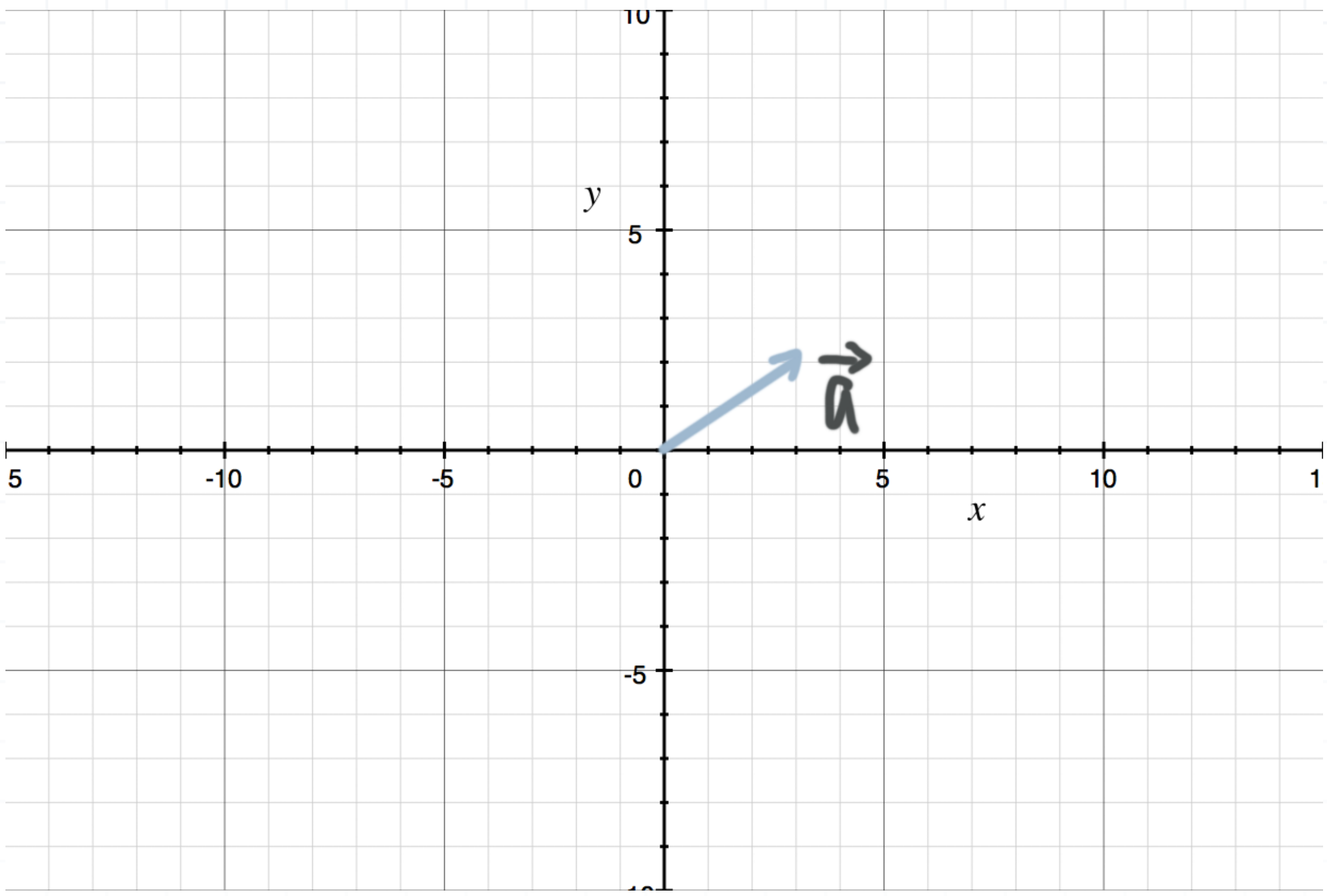
$$\begin{bmatrix} -25 + 6 \\ -10 + 9 \end{bmatrix}$$

$$\begin{bmatrix} -19 \\ -1 \end{bmatrix}$$

Under this transformation, $\overrightarrow{v} = (5,3)$ transforms to $T(\overrightarrow{v}) = (-19, -1)$. So the transformation matrix transforms the entire coordinate plane, all with one simple matrix!

Transforming vectors

Let's say we have the vector $\overrightarrow{a} = (3,2)$.



We can apply a transformation matrix to the vector, and doing so will change it into a transformed vector. Let's say we use the transformation matrix

$$M = \begin{bmatrix} -2 & 4 \\ 3 & 0 \end{bmatrix}$$

and apply it to the vector $\vec{a}$. Then the transformation of $\vec{a}$ by M will be the multiplication of M by $\vec{a}$.

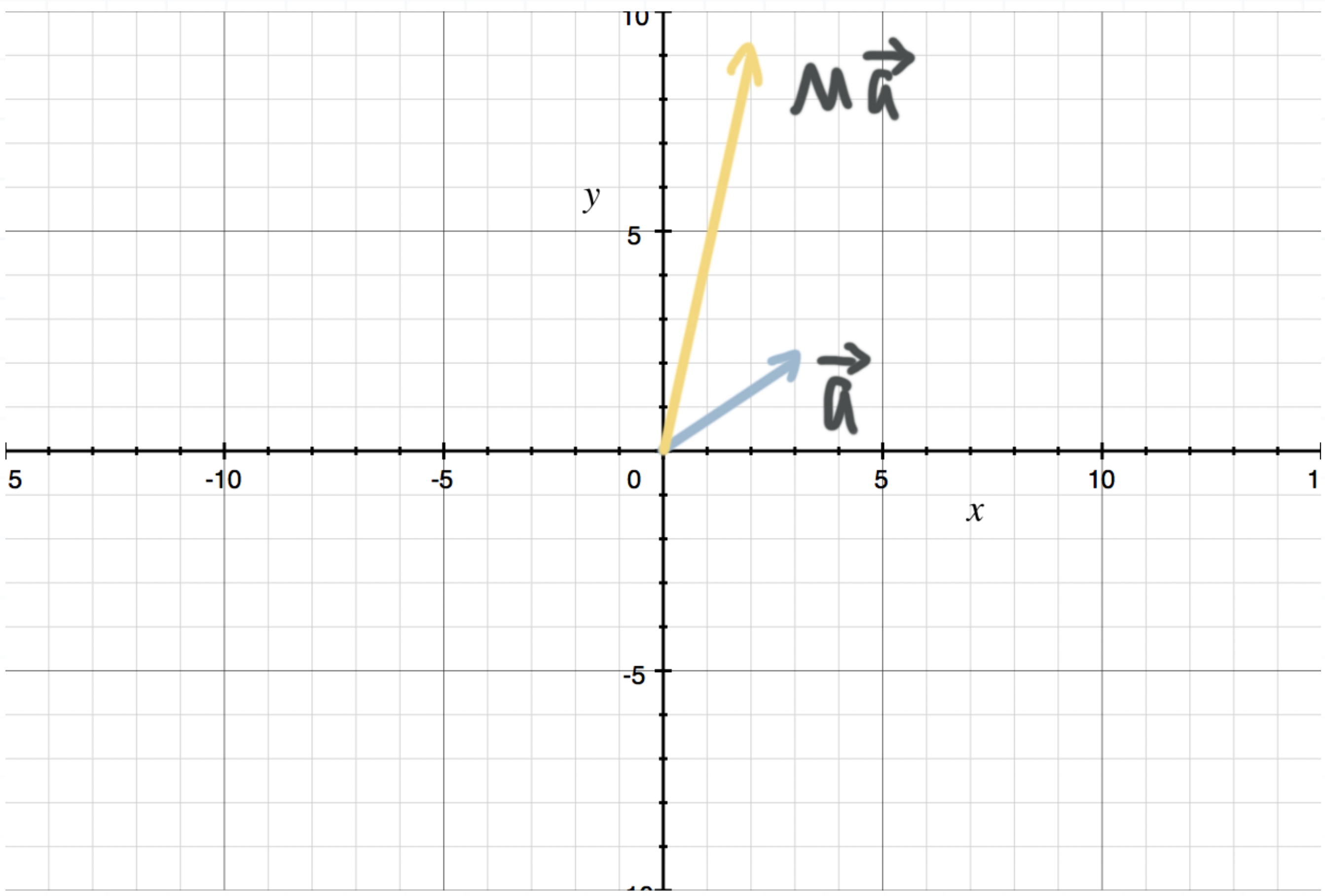
$$M\vec{a} = \begin{bmatrix} -2 & 4 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$M\vec{a} = \begin{bmatrix} -2(3) + 4(2) \\ 3(3) + 0(2) \end{bmatrix}$$

$$M\vec{a} = \begin{bmatrix} -6 + 8 \\ 9 + 0 \end{bmatrix}$$

$$M\vec{a} = \begin{bmatrix} 2 \\ 9 \end{bmatrix}$$

So matrix M transforms $\vec{a} = (3,2)$ into $M\vec{a} = (2,9)$.



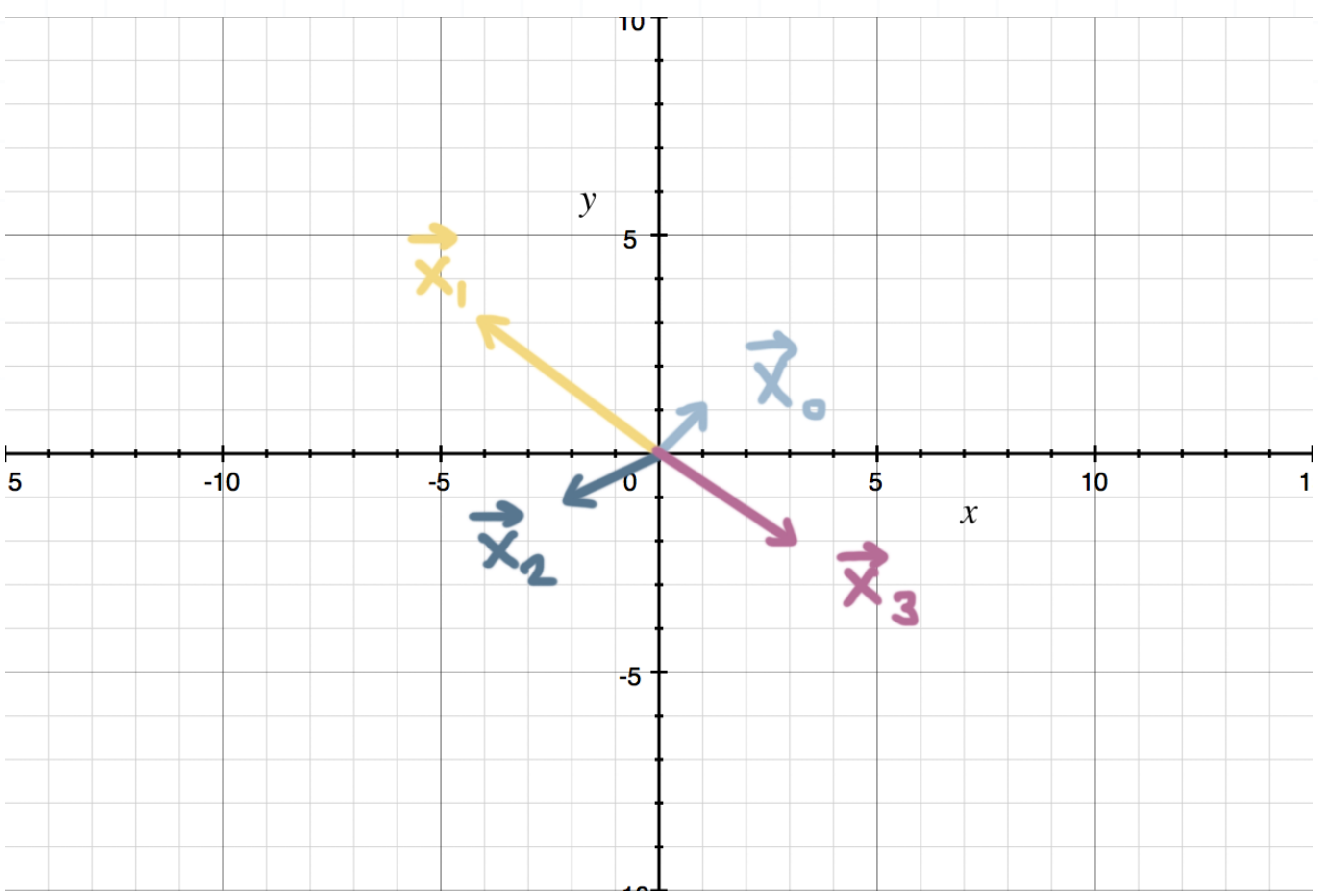
Transforming a set

If instead of being given a single vector $\vec{a}$, we'd been given a set of vectors representing the vertices of some polygon (like a triangle, quadrilateral, pentagon, hexagon, etc.), we can apply a transformation matrix to its vertex vectors, and thereby transform the figure.

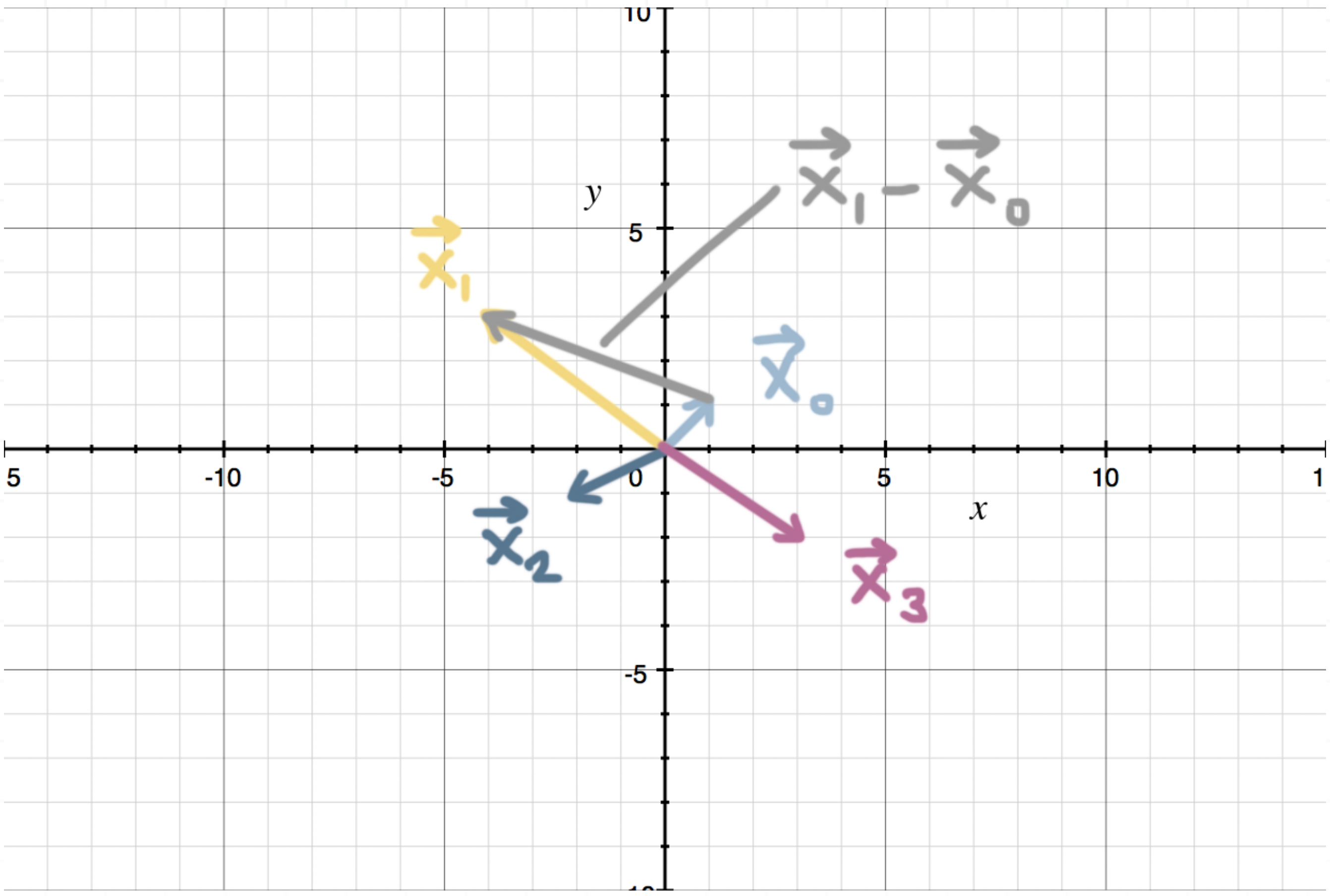
We call the vector set a **subset**, because it's a subset of the space that it's in. For example, you could have a two-dimensional figure, represented by a vector set that's a subset of $\mathbb{R}^2$. Or you could have a three-dimensional figure, represented by a vector set that's a subset of $\mathbb{R}^3$. The original subset is still called the **preimage**, and the transformed subset is still called the **image**. So when we say "the image of the subset under the transformation," we're talking about the set of vectors that define the transformed figure.

For example, let's say we want to transform the quadrilateral Q with vertices $(1,1)$, $(-4,3)$, $(-2,-1)$, and $(3,-2)$. The vertices of Q can be given by position vectors.

$$\vec{x_0} = \begin{bmatrix}1\\1\end{bmatrix}, \vec{x_1} = \begin{bmatrix}-4\\3\end{bmatrix}, \vec{x_2} = \begin{bmatrix}-2\\-1\end{bmatrix}, \vec{x_3} = \begin{bmatrix}3\\-2\end{bmatrix}$$



Now that each vertex of the quadrilateral is defined as a vector, we can describe each side of the quadrilateral in terms of those vectors. For example, we could describe the side connecting $\vec{x}_0$ to $\vec{x}_1$ as the vector $\vec{x_1} - \vec{x_0}$.



So if we call that side Q_0, we can define it as

$$Q_0 = \left\{\vec{x_0} + t(\vec{x_1} - \vec{x_0}) \ \mid \ 0 \le t \le 1\right\}$$

This works by limiting t to values between $t = 0$ and $t = 1$. When $t = 0$, the second term will cancel and we'll just be left with the vector $\vec{x_0}$.

$$Q_0 = \vec{x_0} + 0(\vec{x_1} - \vec{x_0})$$

$$Q_0 = \vec{x_0} + 0$$

$$Q_0 = \vec{x_0}$$

And when $t = 1$, the $\vec{x_0}$ will cancel and we'll just be left with the vector $\vec{x_1}$.

$$Q_0 = \vec{x_0} + 1(\vec{x_1} - \vec{x_0})$$

$$Q_0 = \vec{x_0} + \vec{x_1} - \vec{x_0}$$

$$Q_0 = \vec{x_1}$$

So the expression $Q_0 = \left\{\vec{x_0} + t(\vec{x_1} - \vec{x_0}) \ | \ 0 \leq t \leq 1\right\}$ will define (in terms of vectors) all the points between $\vec{x}_0$ and $\vec{x}_1$, thereby defining the side of the quadrilateral connecting those two vertices. All four sides of the quadrilateral are therefore defined like this:

$$Q_0 = \left\{\vec{x_0} + t(\vec{x_1} - \vec{x_0}) \ | \ 0 \leq t \leq 1\right\}$$

$$Q_1 = \left\{\vec{x_1} + t(\vec{x_2} - \vec{x_1}) \ | \ 0 \leq t \leq 1\right\}$$

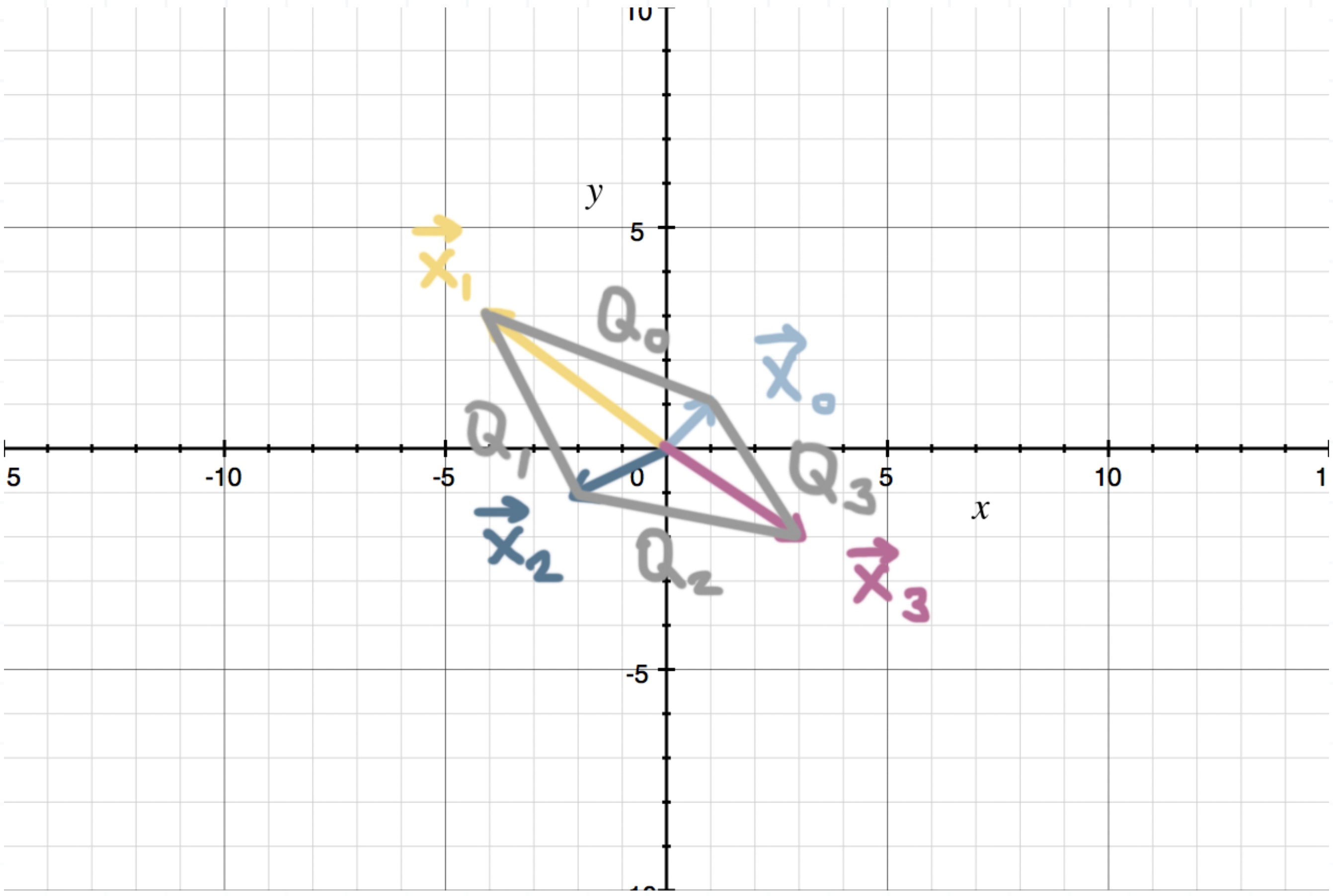
$$Q_2 = \left\{\vec{x_2} + t(\vec{x_3} - \vec{x_2}) \ | \ 0 \leq t \leq 1\right\}$$

$$Q_3 = \left\{\vec{x_3} + t(\vec{x_0} - \vec{x_3}) \ | \ 0 \leq t \leq 1\right\}$$

With the sides defined, we could say that the set of vectors that defines Q is simply

$$Q = \left\{Q_0, Q_1, Q_2, Q_3,\right\}$$

and that a sketch of Q, including each side of the quadrilateral, and the position vectors that define its vertices, is:



Now we need to pick a transformation. Let's say we want to transform the quadrilateral with the transformation T:

$$T(\vec{x}) = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Apply T to each of the position vectors that defines the vertices of Q.

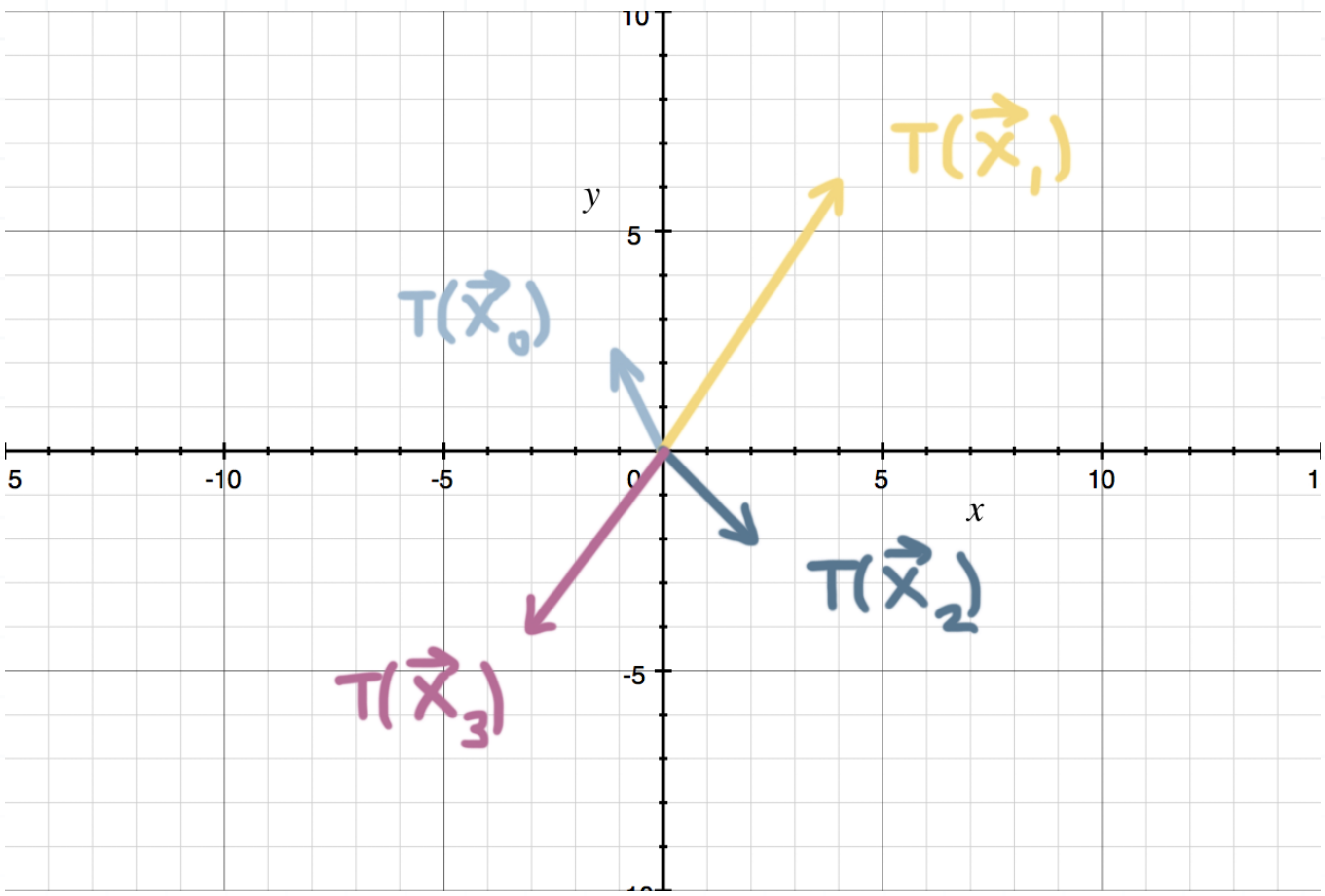
$$T(\vec{x_0}) = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1(1) + 0(1) \\ 0(1) + 2(1) \end{bmatrix} = \begin{bmatrix} -1 + 0 \\ 0 + 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$T(\vec{x_1}) = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -4 \\ 3 \end{bmatrix} = \begin{bmatrix} -1(-4) + 0(3) \\ 0(-4) + 2(3) \end{bmatrix} = \begin{bmatrix} 4 + 0 \\ 0 + 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

$$T(\vec{x_2}) = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1(-2) + 0(-1) \\ 0(-2) + 2(-1) \end{bmatrix} = \begin{bmatrix} 2 + 0 \\ 0 - 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$T(\vec{x_3}) = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -1(3) + 0(-2) \\ 0(3) + 2(-2) \end{bmatrix} = \begin{bmatrix} -3 + 0 \\ 0 - 4 \end{bmatrix} = \begin{bmatrix} -3 \\ -4 \end{bmatrix}$$

These transformed vectors are the position vectors that define the vertices of the transformed quadrilateral. If we sketch the transformed position vectors, we get



The nice thing about transformations of polygons like this one is that we only need to transform the vertices, like we just did. If a straight line connected $\vec{x}_0$ to $\vec{x}_1$ in the original figure (the preimage), then a straight line will also connect $T(\vec{x}_0)$ to $T(\vec{x}_1)$ in the transformed figure (the image). So we simply connect the vertices of the transformation in the same order in which they were connected in the original figure.

We can prove why the straight sides in the transformation follow the same pattern as the straight sides in the original figure if we apply the transformation to a side of the quadrilateral. Let's apply T to Q_0.

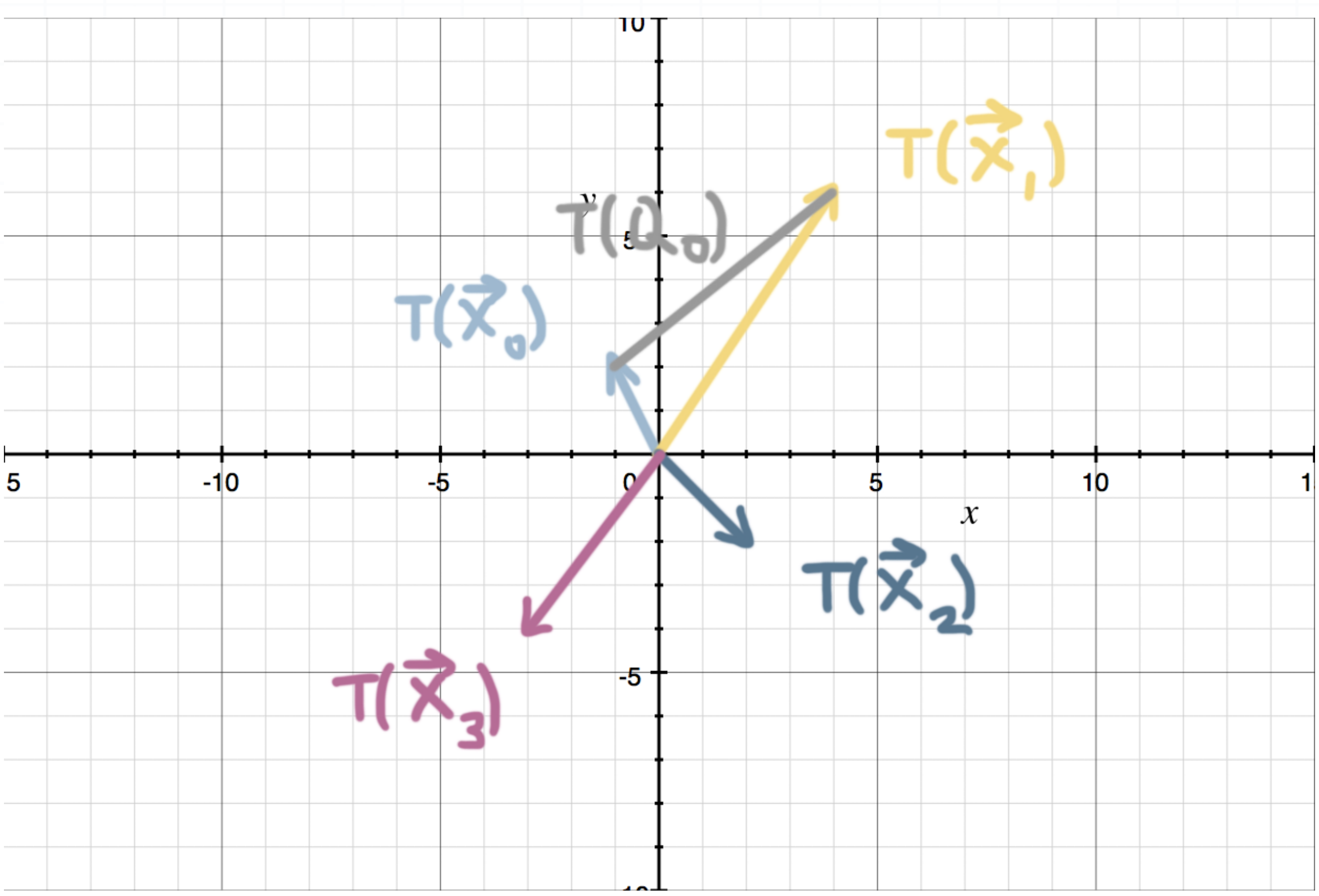
$$T(Q_0) = \left\{T(\vec{x}_0 + t(\vec{x}_1 - \vec{x}_0)) \ \ | \ \ 0 \le t \le 1\right\}$$

$$T(Q_0) = \{T(\vec{x_0}) + T(t(\vec{x_1} - \vec{x_0}))\ \mid\ 0 \le t \le 1\}$$

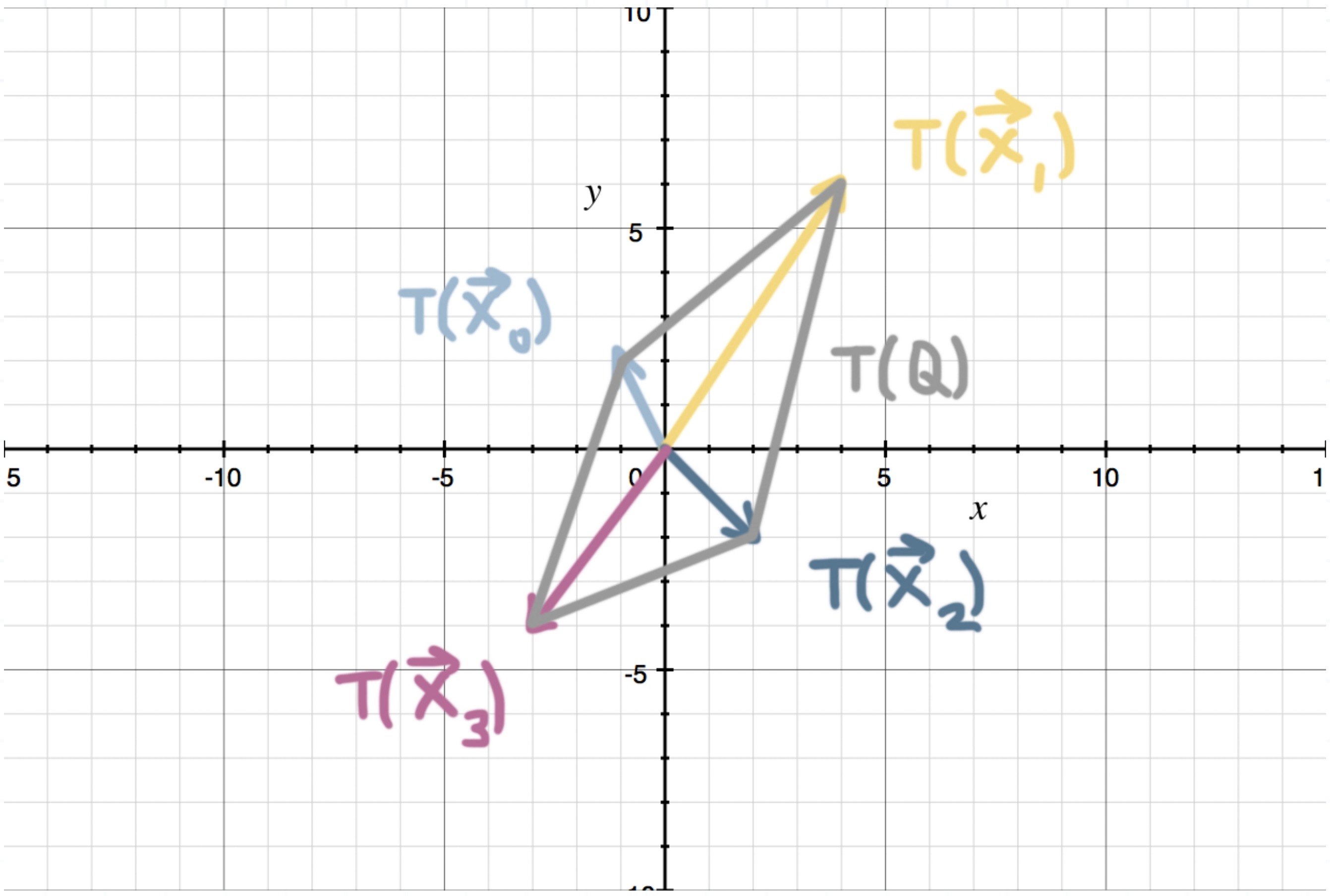
$$T(Q_0) = \{T(\vec{x_0}) + tT(\vec{x_1} - \vec{x_0})\ \mid\ 0 \le t \le 1\}$$

$$T(Q_0) = \{T(\vec{x_0}) + t(T(\vec{x_1}) - T(\vec{x_0}))\ \mid\ 0 \le t \le 1\}$$

If we read through $T(\vec{x_0}) + t(T(\vec{x_1}) - T(\vec{x_0}))$, we see that it tells us to start at the transformation of $\vec{x}_0$, and then add t multiples of the vector connecting $T(\vec{x}_1)$ to $T(\vec{x}_0)$. As long as we keep $0 \le t \le 1$, this will give us all the points along a straight line that connects $T(\vec{x}_0)$ to $T(\vec{x}_1)$. And therefore we could sketch the first side of the transformed figure.



This side specifically is called the "image of Q_0 under T." We could also apply the same transformation to the other three sides of the figure, and we would find similar results, allowing us to connect the rest of the transformed vertices. So the image of the set Q under the transformation T is



10
5
-5
-10
-5
0
5
10
x
y
T(x⃗₁)
T(x⃗₀)
T(Q)
T(x⃗₂)
T(x⃗₃)