Consider the linear mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

**Question:** Given a vector $\mathbf{x} \in \mathbb{R}^2$, how are $\mathbf{x}$ and its image under $T$ related geometrically?

Let's compute the images of the standard basis vectors under $T$. We have

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

As you can see, $T$ acts on the standard basis vectors by both "stretching" and "rotating" them.

**Question:** Are there any vectors other than the zero vector that undergo only stretching (or shrinking) and no rotation? if there is such a vector, say $\mathbf{x} \neq \mathbf{0}$, then it must satisfy the equation $T(\mathbf{x}) = \lambda\mathbf{x}$, where $\lambda$ is a scalar.

We have

$$T(\mathbf{x}) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix}$$

and

$$\lambda\mathbf{x} = \begin{bmatrix} \lambda x_1 \\ \lambda x_1 \end{bmatrix}.$$

Then $T(\mathbf{x}) = \lambda\mathbf{x}$ implies

$$\begin{bmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix},$$

which is equivalent to the system

$$\begin{aligned} (2-\lambda)x_1 + x_2 &= 0 \\ x_1 + (2-\lambda)x_2 &= 0 \end{aligned}$$

This homogeneous system has the matrix form

$$\begin{bmatrix} (2-\lambda) & 1 \\ 1 & (2-\lambda) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Note that the matrix $\begin{bmatrix} (2-\lambda) & 1 \\ 1 & (2-\lambda) \end{bmatrix}$ can be written as $A - \lambda I_2$. Thus, the homogeneous system has a nontrivial solution if and only if $\det(A - \lambda I_2) = 0$. As you can see, $\det(A - \lambda I_2)$ is a polynomial in $\lambda$, which leads to the following definition.

**Definition** Let $A \in \mathbb{R}^{n\times n}$. The polynomial $\det(A - \lambda I_n)$ is called the *characteristic polynomial* of $A$.

Let's go back to our previous example $\begin{bmatrix} (2-\lambda) & 1 \\ 1 & (2-\lambda) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Let's find the values of $\lambda$ for which the coefficient matrix in singular. Set $\det(A - \lambda I_2) = 0$. If you solve this equation, the roots (zeros) of the characteristic polynomial are $\lambda = 1$ and $\lambda = 3$.
When $\lambda = 1$, the coefficient matrix becomes $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, and the set of solutions to the system is the set of all scalar multiples of the vector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. That is, when $\lambda = 1$, the set of solutions to the system $\text{Span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$. This means $T$ leaves all the vectors on the line $\text{Span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$ fixed.

**Note:** $\text{Span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$ is the kernel of the coefficient matrix when $\lambda = 1$. Equivalently, $\ker(T - \lambda I) = \text{Span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$ when $\lambda = 1$.

When $\lambda = 3$, the coefficient matrix becomes $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$, and the set of solutions to the system is the set of all scalar multiples of the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. That is, when $\lambda = 3$, the set of solutions to the system span$\left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$. This means $T$ stretches every vector on the line span$\left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$ by a factor of 3.

**Note:** Span$\left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$ is the kernel of the coefficient matrix when $\lambda = 3$. Equivalently, $\ker(T - \lambda I) = \text{Span}\left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$ when $\lambda = 3$.

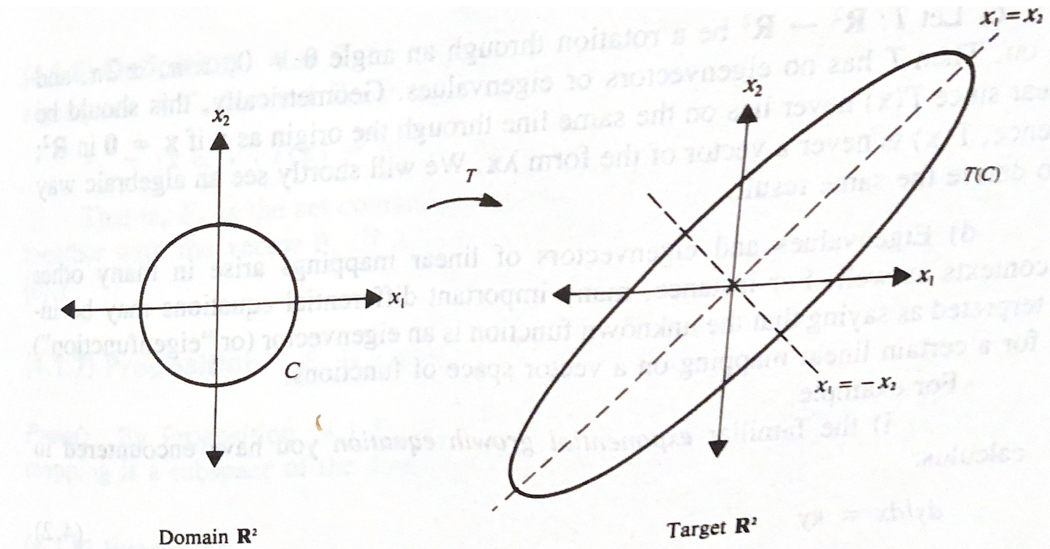
What does $T$ do to the unit circe? Let $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, where $x_1^2 + x_2^2 = 1$. By a previous computation, we know that

$$T(\mathbf{x}) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix}.$$

Where does the image lie? A straightforward computation yields that the point $(2x_1 + x_2, x_1 + 2x_2)$ satisfies the equation

$$5x_1^2 - 8x_1x_2 + 5x_2^2 = 9.$$

The equation $5x_1^2 - 8x_1x_2 + 5x_2^2 = 9$ describes a rotated ellipse with semimajor axis 3 along the line $x_1 = x_2$ and semiminor axis along the line $x_1 = -x_2$. The map $T$ takes the unit circle to the ellipse as shown in the following figure:




$T$ will take other circles centered at the origin to ellipses concentric to the one above.

What have we done so far? We have determined the vectors other than the zero vector that undergo only stretching (or shrinking) and no rotation under a linear map. In general, given a vector space $V$ and a linear mapping $T : V \to V$, we will want to study $T$ by finding all vectors $\mathbf{x}$ satisfying equations $T(\mathbf{x}) = \lambda \mathbf{x}$ for some scalar $\lambda$.

The following two definitions follow from our example.

**Definition** Let $V :\to V$ be a linear mapping.

(1) A vector $\mathbf{x} \in V$ is called an *eigenvector* of $T$ if $\mathbf{x} \neq \mathbf{0}$ and there exists a scalar $\lambda \in \mathbb{R}$ such that $T(\mathbf{x}) = \lambda \mathbf{x}$.
(2) If $\mathbf{x}$ is an eigenvector of $T$ and $T(\mathbf{x}) = \lambda \mathbf{x}$, the scalar $\lambda$ is called the *eigenvalue* of $T$ corresponding to $\mathbf{x}$. That is, an eigenvalue of $T$ is a scalar $\lambda$ for which there exists a nonzero vector $\mathbf{x}$ such that $T(\mathbf{x}) = \lambda \mathbf{x}$.

**Example 1** In the previous example, the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of $T$ with eigenvalue $\lambda = 3$, whereas the vector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = 1$. Any nonzero scalar multiple of either of these two vectors is also an eigenvector for the mapping $T$.

**Example 2** Recall the orthogonal projection map we discussed in class. Let $\mathbf{x} \in \mathbb{R}^2$. Then the orthogonal projection of $\mathbf{x}$ onto the line spanned by the vector $\mathbf{w}$ is a map from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by

$$\text{proj}_L(\mathbf{x}) = \left( \frac{\mathbf{x} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \right) \mathbf{w}$$

Let $L$ be the $x$-axis, and choose $\mathbf{w}$ to be the unit vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then

- Any vector of the form $\begin{bmatrix} a \\ 0 \end{bmatrix}$, where $a \neq 0$, is projected to itself. That is,

$$\text{proj}_L\left(\begin{bmatrix} a \\ 0 \end{bmatrix}\right) = \begin{bmatrix} a \\ 0 \end{bmatrix}.$$

  This means, any vector $\begin{bmatrix} a \\ 0 \end{bmatrix}$, where $a \neq 0$, is an eigenvector of $\text{proj}_L(\cdot)$ with eigenvalue $\lambda = 1$.

- Any vector of the form $\begin{bmatrix} 0 \\ b \end{bmatrix}$, where $b \neq 0$, is projected to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. That is,

$$\text{proj}_L\left(\begin{bmatrix} 0 \\ b \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \cdot \begin{bmatrix} 0 \\ b \end{bmatrix}.$$

  This means, any vector $\begin{bmatrix} 0 \\ b \end{bmatrix}$, where $b \neq 0$, is an eigenvector of $\text{proj}_L(\cdot)$ with eigenvalue $\lambda = 0$.

- Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a rotation through an angle $\theta \neq 0, \pm\pi, \pm 2\pi$, and so on. The $T$ has no eigenvectors or eigenvalues. Geometrically, if you rotate a vector in $\mathbb{R}^2$ by an angle which is not a scalar multiple of $\pi$, then $T(\mathbf{x})$ never lies on the same line through the origin as $\mathbf{x}$ if $\mathbf{x} \neq \mathbf{0}$, hence $T(\mathbf{x})$ is never a vector of the form $\lambda\mathbf{x}$.

The discussion from our very first example leads to the following results and definitions.

- A vector $\mathbf{x}$ is an eigenvector of $T$ with eigenvalue $\lambda$ if and only if $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{x} \in \text{Ker}(T - \lambda I)$.

- Let $A \in \mathbb{R}^{n\times n}$. Then $\lambda \in \mathbb{R}$ is an eigenvalue of $A$ if and only if $\det(A - \lambda I) = 0$. The roots (zeros) of the characteristic polynomial are called *eigenvalues* of $A$.

- Recall the characteristic polynomial of a matrix $A$. Let $A \in \mathbb{R}^{n\times n}$. The polynomial $\det(A - \lambda I_n)$ is called the *characteristic polynomial* of $A$. We can also define the characteristic polynomial as a polynomial in variable $t$: $p(t) = \det(A - tI_n)$. Then it can be shown that $p(A) = 0$ (the $n \times n$ zero matrix). We can use this result to find the inverse of the matrix $A$ (an example will be added to the next class activity).

- Let $A \in \mathbb{R}^{n\times n}$. The characteristic polynomial $\det(A - \lambda I_n)$ is a polynomial (in variable $\lambda$) of degree exactly $n$. Then $A$ has no more than $n$ distinct eigenvalues. If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of $A$ of multiplicity $m_1, m_2, \ldots, m_k$, respectively, then $m_1 + m_2 + \cdots + m_k \leq n$.

- Let $T : V \to V$ be a linear mapping, and let $\lambda \in \mathbb{R}$. The $\lambda$-*eigenspace* of $T$, denoted $E_\lambda$, is the set

$$E_\lambda = \{\mathbf{x} \in V \mid T(\mathbf{x}) = \lambda\mathbf{x}\}.$$

  $E_\lambda$ is the collection of all the eigenvectors of $T$ associated with eigenvalue $\lambda$, together with the vector $\mathbf{0}$. If $\lambda$ is not an eigenvalue, then $E_\lambda = \{\mathbf{0}\}$. $E_\lambda = \text{Ker}(T - \lambda I)$, which is a subspace of $V$ for all $\lambda$.

**Defintion** Consider two $n \times n$ matrices $A$ and $B$. We say that $A$ is similar to $B$ if there exists an invertible matrix $Q$ such that $B = Q^{-1}AQ$.

It can be shown that similar matrices have equal characteristic polynomials. That is, if $A$ is similar to $B$, then

$$\det(A - \lambda I_n) = \det(B - \lambda I_n).$$