Eigenvalues and Eigenvectors of Matrices and Transformations

In this section we will introduce the concept of eigenvalues and eigenvectors of a transformation. We begin with an illustrative example.

Define

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$

by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 + 2x_2 \\ x_2 \end{pmatrix}.$$

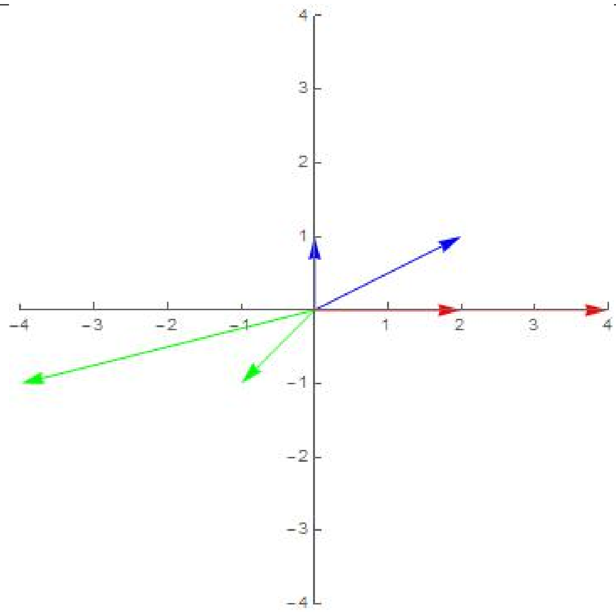
Geometrically, we can think of T as a "shear"; indeed, T leaves the 2nd coordinate of each vector alone, but stretches out the first coordinate. Let's calculate the images of vectors

$$v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix},\ \text{and } v_3 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

under the action of T:

$$\begin{aligned} T(v_1) &= T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\ &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ T(v_2) &= T\left(\begin{pmatrix} -1 \\ -1 \end{pmatrix}\right) \\ &= \begin{pmatrix} -4 \\ 1 \end{pmatrix} \\ T(v_3) &= T\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) \\ &= \begin{pmatrix} 4 \\ 0 \end{pmatrix}. \end{aligned}$$

The vectors above and their images are graphed below; vectors v_1 and $T(v_1)$ are graphed in blue; v_2 and $T(v_2)$ are in green; and v_3 and $T(v_3)$ are in red:



There is something interesting about the action of T on vector v_3; indeed,

$$T(v_3) = T\left(\begin{pmatrix}2\\0\end{pmatrix}\right) = \begin{pmatrix}4\\0\end{pmatrix} = 2T(v_3),$$

that is, T simply *scaled* vector v_3 (unlike the other two vectors, which were also rotated).

The vector v_3 above is our first in-class example of an *eigenvector*, and the scalar 2 is called an *eigenvalue* for T. We introduce the definitions below:

Definitions 5.5/5.7. Let $T : V \to V$ be a linear operator (that is, a transformation from V to V). A number $\lambda \in \mathbb{F}$ is called an *eigenvalue* of T if there is a vector $v \in V$, $v \neq \mathbf{0}$, so that $T(v) = \lambda v$, and vector v is called an *eigenvector* corresponding to λ.

Example. Given *any* vector space V, the transformation

$$T_{\mathbf{0}} : V \to V$$

given by

$$T_{\mathbf{0}}(v) = \mathbf{0}$$

has eigenvalue $\lambda = 0$, since for any vector $v \in V$,

$$T(v) = 0 \cdot v = \mathbf{0}.$$

Every nonzero vector $v \in V$ is an eigenvector associated with $\lambda = 0$ (the zero vector $\mathbf{0}$ is not called an eigenvector).

Example. Let V be *any* vector space, and let

$$T_I : V \to V$$

be defined by

$$T_I(v) = v.$$

The scalar $1 \in \mathbb{F}$ is the only unique eigenvalue for T, since

$$T_I(v) = 1 \cdot v = v.$$

Any nonzero vector v is an eigenvector associated with $\lambda = 1$.

Example. Let $T : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ be given by

$$T(X) = X + X^\top.$$

Find all eigenvalues and associated eigenvectors for T.

If λ is an eigenvalue for T, then there is a nonzero matrix X so that

$$\lambda X = X + X^\top.$$

There are two cases to consider:

1. $\lambda \neq 0$: Then

$$X = \frac{1}{\lambda}(X + X^\top).$$

   Notice that

$$(X + X^\top)^\top = X + X^\top,$$

   that is $\frac{1}{\lambda}(X + X^\top)$ (and thus X) is symmetric. Of course, if X is symmetric then $X = X^\top$, so we have

$$\begin{aligned} X &= \frac{1}{\lambda}(X + X^\top) \\ &= \frac{1}{\lambda}(X + X) \\ &= \frac{2}{\lambda}X. \end{aligned}$$

   In order to guarantee equality, we must have $\lambda = 2$, so $\lambda = 2$ is an eigenvalue of T corresponding to any $X \in \mathcal{M}_2(\mathbb{R})$ so that X is symmetric.

2. $\lambda = 0$: Then

$$\mathbf{0} = X + X^\top$$

   so that

$$X = -X^\top,$$

   that is X is skew symmetric. Thus $\lambda = 0$ is an eigenvalue of T corresponding to any $X \in \mathcal{M}_2(\mathbb{R})$ so that X is skew-symmetric.

Example. Let

$$A = \begin{pmatrix} 3 & 3 \\ 3 & -5 \end{pmatrix}$$

and define $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$T_A(x) = Ax.$$

Find all eigenvalues of T_A and describe the associated eigenvectors.

We wish to find $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^2$ so that

$$Ax = \lambda x.$$

This is equivalent to solving the matrix equation

$$(A - \lambda I)x = \mathbf{0}$$

for x; we could, of course, proceed by row reducing

$$A - \lambda I = \begin{pmatrix} 3-\lambda & 3 \\ 3 & -5-\lambda \end{pmatrix}.$$

However, this will be a bit troublesome. We can solve the problem using a nice observation: the system

$$(A - \lambda I)x = \mathbf{0}$$

has only the trivial solution $x = \mathbf{0}$ *if and only if* $\det(A - \lambda I) \neq 0$. Of course, $x = \mathbf{0}$ is not an eigenvector.

So we actually wish to find all scalars λ so that

$$(A - \lambda I)x = \mathbf{0}$$

has trivial solutions, which occurs if and only if $\det(A - \lambda I) = 0$. Thus we look for λ with this property:

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{pmatrix} 3-\lambda & 3 \\ 3 & -5-\lambda \end{pmatrix} \\ &= -(3-\lambda)(5+\lambda) - 9 \\ &= -(15 - 2\lambda - \lambda^2) - 9 \\ &= -15 + 2\lambda + \lambda^2 - 9 \\ &= \lambda^2 + 2\lambda - 24 \\ &= (\lambda + 6)(\lambda - 4). \end{aligned}$$

Thus the only scalars so that $\det(A - \lambda I) = 0$ are $\lambda = -6$ and $\lambda = 4$. These are the only eigenvalues for T_A.

To find the associated eigenvectors, we look for vectors x so that

$$Ax = 4x \text{ or } Ax = -6x.$$

If $Ax = 4x$, then we have

$$\begin{pmatrix} 3 & 3 \\ 3 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4x_1 \\ 4x_2 \end{pmatrix},$$

which results in the system of equations

$$\begin{aligned} 3x_1 + 3x_2 &= 4x_1 \\ 3x_1 - 5x_2 &= 4x_2. \end{aligned}$$

Simplifying, we have

$$\begin{aligned} -x_1 + 3x_2 &= 0 \\ 3x_1 - 9x_2 &= 0; \end{aligned}$$

row-reducing the resulting augmented matrix, we have

$$\begin{aligned} \left(\begin{array}{cc|c} -1 & 3 & 0 \\ 3 & -9 & 0 \end{array}\right) &\to \left(\begin{array}{cc|c} 1 & -3 & 0 \\ 3 & -9 & 0 \end{array}\right) \\ &\to \left(\begin{array}{cc|c} 1 & -3 & 0 \\ 0 & 0 & 0 \end{array}\right). \end{aligned}$$

The system thus has infinitely many solutions; parameterizing $x_2 = t$, we see that any vector of the form

$$x = \begin{pmatrix} 3t \\ t \end{pmatrix}$$

is an eigenvector associated with $\lambda = 4$.

Proceeding in a similar fashion for $\lambda = -6$, we wish to find $x \in \mathbb{R}^2$ so that

$$\begin{aligned} 3x_1 + 3x_2 &= -6x_1 \\ 3x_1 - 5x_2 &= -6x_2 \end{aligned}$$

or

$$\begin{aligned} 9x_1 + 3x_2 &= 0 \\ 3x_1 + x_2 &= 0. \end{aligned}$$

Again row reducing, we have

$$\begin{aligned} \left(\begin{array}{cc|c} 9 & 3 & 0 \\ 3 & 1 & 0 \end{array}\right) &\to \left(\begin{array}{cc|c} 1 & 1/3 & 0 \\ 3 & 1 & 0 \end{array}\right) \\ &\to \left(\begin{array}{cc|c} 1 & 1/3 & 0 \\ 0 & 0 & 0 \end{array}\right). \end{aligned}$$

Once more, we have infinitely many solutions; parameterizing $x_2 = t$, we see that any vector of the form

$$x = \begin{pmatrix} -t/3 \\ t \end{pmatrix}$$

is an eigenvector associated with $\lambda = -6$.

Observation. It is easy to see that, if $v \neq \mathbf{0}$ is an eigenvector of T associated with λ, then αv is also an eigenvector associated with λ for all $\alpha \neq 0$. Similarly, if v_1 and v_2 are eigenvectors associated with the same eigenvalue λ, then $v_1 + v_2$ is also an eigenvector associated with λ. Thus if λ is an eigenvalue,

$$V_\lambda = \{v \in V | T(v) = \lambda v\}$$

is a subspace of V (as you have proved in the context of matrices in a recent homework assignment).

Eigenvalues and Operators

We know how to add linear transformations (and thus operators), a fact that we can use to quickly determine whether or not a scalar λ is an eigenvalue. Indeed,

$$\begin{aligned} T(x) = \lambda x \quad &\iff \quad T(x) = \lambda T_I(x) \\ &\iff \quad T(x) - \lambda T_I(x) = \mathbf{0} \\ &\iff \quad (T - \lambda T_I)(x) = \mathbf{0} \\ &\iff \quad x \in \text{null}\,(T - \lambda T_I). \end{aligned}$$

This is actually a proof of $(a) \iff (b)$ in the theorem below; the remaining equivalences follow immediately from Theorem 3.69:

Theorem 5.6. Let V be finite dimensional, $T : V \to V$ a linear operator, and $\lambda \in \mathbb{F}$. The following are equivalent:

(a) λ is an eigenvalue of T;

(b) $T - \lambda T_I$ is not injective;

(c) $T - \lambda T_I$ is not surjective;

(d) $T - \lambda T_I$ is not invertible.

We may be curious as to the relationships among the eigenvectors for a linear transformation T; the theorem below provides a partial answer.

Theorem 5.10. Let $T : V \to V$ be a linear operator on the finite dimensional vector space V. If $\lambda_1, \ldots, \lambda_n$ are *distinct* eigenvalues of T, and if $v_1, \ldots, v_n$ are vectors so that v_i is an eigenvector associated with λ_i, then the list $(v_1,\ v_2,\ \ldots,\ v_n)$ is an independent list.

Sketch of Proof. Proceed by induction: show that, if v_1 and v_2 are eigenvectors for T and are also dependent, then they must be associated with the same eigenvalue.

For the inductive hypothesis, let $v_1, \ldots, v_n$ be any eigenvectors associated with unique eigenvalues, so that $(v_1,\ \ldots,\ v_n)$ is an independent list. Let v_{n+1} be any eigenvector of T in span $(v_1,\ \ldots,\ v_n)$, and show that v_{n+1} must be associated with one of the eigenvalues $\lambda_1, \ldots, \lambda_n$.

The next theorem follows immediately:

Theorem 5.13. A linear operator on an n dimensional vector space V has at most n distinct eigenvalues.

Proof. V can have at most n linearly independent vectors in any list. Since any independent list of k eigenvectors has k distinct associated eigenvalues, V has at most n distinct eigenvalues.

Remark. While n is an upper bound on the number of distinct eigenvalues of an operator T on an n dimensional space V, T could certainly have fewer than n distinct eigenvalues. For example, we saw that the only eigenvalue of the operator T_I is $\lambda = 1$, regardless of the dimension of the space V.

Eigenvalues and Eigenvectors of Matrices

Since an $n \times n$ matrix A can be thought of as a linear operator on $\mathbb{F}^n$, we can talk about eigenvectors and eigenvalues for A. The definition is virtually identical to that for eigenvalues and eigenvectors of abstract transformations:

Definition. Let $A \in \mathcal{M}_n(\mathbb{F})$. A number $\lambda \in \mathbb{F}$ is called an *eigenvalue* of A if there is a vector $v \in \mathbb{F}^n$, $v \neq \mathbf{0}$, so that $Av = \lambda v$, and vector v is called an *eigenvector* corresponding to λ.

The theorems that we have discussed on eigenvalues and eigenvectors of transformations transfer immediately to eigenvalues and eigenvectors of *matrices*; we record several without proof.

Theorem. A matrix $A \in \mathcal{M}_n(\mathbb{F})$ has at most n distinct eigenvalues.

Theorem. Let $A \in \mathcal{M}_n(\mathbb{F})$ and $\lambda \in \mathbb{F}$. The following are equivalent:

1. λ is an eigenvalue of A;
2. $(A - \lambda I)x = \mathbf{0}$ has nontrivial solutions;
3. $\det(A - \lambda I) = 0$.

Notice that, if $\lambda = 0$ is an eigenvalue, then automatically $Ax = \mathbf{0}$ has nontrivial solutions. Thus we add to the list of equivalent conditions that we began in Unit 1, Section 10:

Theorem. Let A be an $n \times n$ matrix. Then the following are equivalent:

- A is invertible.
- $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
- The reduced row echelon form of A is I_n.
- $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.
- $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.
- $\det A \neq 0$.
- 0 is not an eigenvalue of A.

Characteristic Polynomial of a Matrix

In an earlier example, we found the eigenvalues of the matrix

$$A = \begin{pmatrix} 3 & 3 \\ 3 & -5 \end{pmatrix}$$

by solving the equation

$$\det(A - \lambda I) = 0.$$

This technique will actually produce the eigenvalues for *any* $n \times n$ matrix A; thus we introduce the following definition:

Definition. Given $A \in \mathcal{M}_n(\mathbb{F})$, the polynomial $p(\lambda) = \det(A - \lambda I)$ is called the *characteristic polynomial* of A, and the equation

$$\det(A - \lambda I) = 0$$

is called the *characteristic equation* of A.

Theorem. The eigenvalues of an $A \in \mathcal{M}_n(\mathbb{C})$ are precisely the solutions to its characteristic equation, that is the roots of its characteristic polynomial.

Eigenvalues of Operators and of their Associated Matrices

The discussion above leads to a natural question: do eigenvalues for a linear operator match up with the eigenvalues for its associated matrices?

The answer is yes, and is quite easy to ascertain: if $\lambda \in \mathbb{F}$ is an eigenvalue for $T : V \to V$, and $A = A_{(B,B)}$ is the matrix of A with respect to basis B of V, then we have

$$T(v) = \lambda v,$$

so that

$$\begin{aligned} A(v)_B &= (T(v))_B \\ &= (\lambda v)_B \\ &= \lambda (v)_B. \end{aligned}$$

Of course, the reverse is true as well: if A has eigenvalue λ, then λ is an eigenvalue for T. We record the observations below:

Theorem. The scalar $\lambda \in \mathbb{F}$ is an eigenvalue for the linear operator $T : V \to V$ if and only λ is an eigenvalue for the matrix A of T with respect to some basis B for V.

Since a transformation T has multiple matrix representations (one for each basis), you may be concerned about well-definedeness. That is, if A and A' are the matrices for T with respect to bases B and B' respectively, then must A and A' have the same eigenvalues?

Fortunately, the answer is yes, and is due to the following theorem:

Theorem. If A, $A' \in \mathcal{M}_n(\mathbb{F})$ are matrices so that there is an invertible matrix $X \in \mathcal{M}_n(\mathbb{F})$ with $A = XAX^{-1}$, then A and A' share eigenvalues.

Since matrices for the same transformation are similar ($A = XA'X^{-1}$), we have the following corollary:

Corollary. If A and A' are matrices for T with respect to bases B and B' respectively, then A and A' share eigenvalues.