

Eigenvalues and Eigenvectors

Isaiah Lankham, Bruno Nachtergaele, Anne Schilling
(February 12, 2007)

In this section we are going to study linear maps $T : V \rightarrow V$ from a vector space to itself. These linear maps are also called operators on V . We are interested in the question when there is a basis for V such that T has a particularly nice form, like being diagonal or upper triangular. This quest leads us to the notion of eigenvalues and eigenvectors of linear operators, which is one of the most important concepts in linear algebra and beyond. For example, quantum mechanics is based on the study of eigenvalues and eigenvectors of operators.

1 Invariant subspaces

To begin our study we will look at subspaces U of V that have special properties under an operator $T \in \mathcal{L}(V, V)$.

Definition 1. Let V be a finite-dimensional vector space over \mathbb{F} with $\dim V \geq 1$, and let $T \in \mathcal{L}(V, V)$ be an operator in V . Then a subspace $U \subset V$ is called an **invariant subspace** under T if

$$Tu \in U \quad \text{for all } u \in U.$$

That is if

$$TU = \{Tu \mid u \in U\} \subset U.$$

Example 1. The subspaces $\text{null } T$ and $\text{range } T$ are invariant subspaces under T . To see this, let $u \in \text{null } T$. This means that $Tu = 0$. But since $0 \in \text{null } T$ this implies that $Tu = 0 \in \text{null } T$. Similarly, let $u \in \text{range } T$. Since $Tv \in \text{range } T$ for all $v \in V$, we certainly also have that $Tu \in \text{range } T$.

An important special case is the case of one-dimensional invariant subspaces of an operator $T \in \mathcal{L}(V, V)$. If $\dim U = 1$, then there exists a nonzero vector $u \in V$ such that

$$U = \{au \mid a \in \mathbb{F}\}.$$

In this case we must have

$$Tu = \lambda u \quad \text{for some } \lambda \in \mathbb{F}.$$

This motivates the definition of eigenvectors and eigenvalues of a linear operator T .

2 Eigenvalues

Definition 2. Let $T \in \mathcal{L}(V, V)$. Then $\lambda \in \mathbb{F}$ is an **eigenvalue** of T if there exists a nonzero vector $u \in V$ such that

$$Tu = \lambda u.$$

The vector u is called the **eigenvector** (with eigenvalue λ) of T .

Finding the eigenvalues and eigenvectors of linear operators is one of the most important problems in linear algebra. We will see later that they have many uses and applications. For example all of quantum mechanics is based on eigenvalues and eigenvectors of operators.

Example 2.

1. Let T be the zero map defined by $T(v) = 0$ for all $v \in V$. Then every vector $u \neq 0$ is an eigenvector of T with eigenvalue 0.
2. Let I be the identity map defined by $I(v) = v$ for all $v \in V$. Then every vector $u \neq 0$ is an eigenvector of T with eigenvalue 1.
3. The projection $P : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ defined by $P(x, y, z) = (x, y, 0)$ has eigenvalues 0 and 1. The vector $(0, 0, 1)$ is an eigenvector with eigenvalue 0 and $(1, 0, 0)$ and $(0, 1, 0)$ are eigenvectors with eigenvalue 1.
4. Take the operator $R : \mathbb{F}^2 \rightarrow \mathbb{F}^2$ defined by $R(x, y) = (-y, x)$. When $\mathbb{F} = \mathbb{R}$, then R can be interpreted as counterclockwise rotation by 90° . From this interpretation it is clear that no vector in \mathbb{R}^2 is left invariant (up to a scalar multiplication). Hence for $\mathbb{F} = \mathbb{R}$ the operator R has no eigenvalues. For $\mathbb{F} = \mathbb{C}$ the situation is different! Then $\lambda \in \mathbb{C}$ is an eigenvalue if

$$R(x, y) = (-y, x) = \lambda(x, y),$$

so that $y = -\lambda x$ and $x = \lambda y$. This implies that $y = -\lambda^2 y$ or $\lambda^2 = -1$. The solutions are hence $\lambda = \pm i$. One can check that $(1, -i)$ is an eigenvector with eigenvalue i and $(1, i)$ is an eigenvector with eigenvalue $-i$.

Eigenspaces are important examples of invariant subspaces. Let $T \in \mathcal{L}(V, V)$ and let $\lambda \in \mathbb{F}$ be an eigenvalue of T . Then

$$V_\lambda = \{v \in V \mid Tv = \lambda v\}$$

is called an **eigenspace** of T . Equivalently

$$V_\lambda = \text{null}(T - \lambda I).$$

Note that $V_\lambda \neq \{0\}$ since λ is an eigenvalue if and only if there exists a nonzero vector $u \in V$ such that $Tu = \lambda u$. We can reformulate this by saying that:

$\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if the operator $T - \lambda I$ is not injective.

Or, since we know that for finite-dimensional vector spaces the notion of injectivity, surjectivity and invertibility are equivalent we can say that

$\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if the operator $T - \lambda I$ is not surjective.

$\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if the operator $T - \lambda I$ is not invertible.

Theorem 1. *Let $T \in \mathcal{L}(V, V)$ and let $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ be m distinct eigenvalues of T with corresponding nonzero eigenvectors v_1, \dots, v_m . Then (v_1, \dots, v_m) is linearly independent.*

Proof. Suppose that (v_1, \dots, v_m) is linearly dependent. Then by the Linear Dependence Lemma there exists a $k \in \{2, \dots, m\}$ such that

$$v_k \in \text{span}(v_1, \dots, v_{k-1})$$

and (v_1, \dots, v_{k-1}) is linearly independent. This means that there exist scalars $a_1, \dots, a_{k-1} \in \mathbb{F}$ such that

$$v_k = a_1 v_1 + \dots + a_{k-1} v_{k-1}. \quad (1)$$

Applying T to both sides yields, using that v_j is an eigenvector with eigenvalue λ_j ,

$$\lambda_k v_k = a_1 \lambda_1 v_1 + \dots + a_{k-1} \lambda_{k-1} v_{k-1}.$$

Subtracting λ_k times equation (1) from this, we obtain

$$0 = (\lambda_k - \lambda_1) a_1 v_1 + \dots + (\lambda_k - \lambda_{k-1}) a_{k-1} v_{k-1}.$$

Since (v_1, \dots, v_{k-1}) is linearly independent, we must have $(\lambda_k - \lambda_j) a_j = 0$ for all $j = 1, 2, \dots, k-1$. By assumption all eigenvalues are distinct, so that $\lambda_k - \lambda_j \neq 0$, which implies that $a_j = 0$ for all $j = 1, 2, \dots, k-1$. But then by (1), $v_k = 0$ which contradicts the assumption that all eigenvectors are nonzero. Hence (v_1, \dots, v_m) is linearly independent. \square

Corollary 2. *Any operator $T \in \mathcal{L}(V, V)$ has at most $\dim V$ distinct eigenvalues.*

Proof. Let $\lambda_1, \dots, \lambda_m$ be distinct eigenvalues of T . Let v_1, \dots, v_m be the corresponding nonzero eigenvectors. By Theorem 1 the list (v_1, \dots, v_m) is linearly independent. Hence $m \leq \dim V$. \square

3 Diagonal matrices

Note that if T has $n = \dim V$ distinct eigenvalues, then there exists a basis (v_1, \dots, v_n) of V such that

$$Tv_j = \lambda_j v_j \quad \text{for all } j = 1, 2, \dots, n.$$

Then any $v \in V$ can be written as a linear combination of v_1, \dots, v_n as $v = a_1 v_1 + \dots + a_n v_n$. Applying T to this we obtain

$$Tv = \lambda_1 a_1 v_1 + \dots + \lambda_n a_n v_n.$$

Hence the vector

$$M(v) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

is mapped to

$$M(Tv) = \begin{bmatrix} \lambda_1 a_1 \\ \vdots \\ \lambda_n a_n \end{bmatrix}.$$

This means that the matrix $M(T)$ for T with respect to the basis of eigenvectors (v_1, \dots, v_n) is **diagonal**

$$M(T) = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

We summarize the results of the above discussion in the following Proposition.

Proposition 3. *If $T \in \mathcal{L}(V, V)$ has $\dim V$ distinct eigenvalues, then $M(T)$ is diagonal with respect to some basis of V . Moreover, V has a basis consisting of eigenvectors of T .*

4 Existence of eigenvalues

In what follows we want to study the question of when an operator T has any eigenvalue. To answer this question we will use polynomials $p(z) \in \mathcal{P}(\mathbb{F})$ evaluated on operators $T \in \mathcal{L}(V, V)$ or equivalently on square matrices $\mathbb{F}^{n \times n}$. More explicitly, for a polynomial

$$p(z) = a_0 + a_1 z + \dots + a_k z^k$$

we can associate the operator

$$p(T) = a_0I_V + a_1T + \cdots + a_kT^k.$$

Note that for $p, q \in \mathcal{P}(\mathbb{F})$ we have

$$(pq)(T) = p(T)q(T) = q(T)p(T).$$

The results of this section will be for complex vector spaces. The reason for this is that the proof of the existence of eigenvalues relies on the Fundamental Theorem of Algebra, which makes a statement about the existence of zeroes of polynomials over the complex numbers.

Theorem 4. *Let $V \neq \{0\}$ be a finite-dimensional vector space over \mathbb{C} and $T \in \mathcal{L}(V, V)$. Then T has at least one eigenvalue.*

Proof. Let $v \in V$, $v \neq 0$ and consider

$$(v, Tv, T^2v, \dots, T^nv),$$

where $n = \dim V$. Since the list contains $n + 1$ vectors, it must be linearly dependent. Hence there exist scalars $a_0, a_1, \dots, a_n \in \mathbb{C}$, not all zero, such that

$$0 = a_0v + a_1Tv + a_2T^2v + \cdots + a_nT^nv.$$

Let m be largest such that $a_m \neq 0$. Since $v \neq 0$ we must have $m > 0$ (but possibly $m = n$). Consider the polynomial

$$p(z) = a_0 + a_1z + \cdots + a_mz^m = c(z - \lambda_1) \cdots (z - \lambda_m),$$

where $c, \lambda_1, \dots, \lambda_m \in \mathbb{C}$ and $c \neq 0$. Note that the polynomial factors because of the Fundamental Theorem of Algebra, which says that every polynomial over \mathbb{C} has at least one zero. Call this zero λ_1 . Then $p(z) = (z - \lambda_1)\tilde{p}(z)$ where $\tilde{p}(z)$ is a polynomial of degree $m - 1$. Continue the process until $p(z)$ is completely factored.

Therefore

$$\begin{aligned} 0 &= a_0v + a_1Tv + a_2T^2v + \cdots + a_nT^nv = p(T)v \\ &= c(T - \lambda_1I)(T - \lambda_2I) \cdots (T - \lambda_mI)v, \end{aligned}$$

so that at least one of the factors $T - \lambda_jI$ must be noninjective. In other words, this λ_j is an eigenvalue of T . \square

Note that the proof of this Theorem only uses basic concepts about linear maps, not

determinants or characteristic polynomials as many other proofs of the same result! This is the same approach as in the textbook of Axler, "Linear Algebra Done Right".

Theorem 4 does not hold for real vector spaces. We have already seen in Example 2 that the rotation operator R on \mathbb{R}^2 has no eigenvalues.

5 Upper triangular matrices

As before let V be a complex vector space.

Let $T \in \mathcal{L}(V, V)$ and (v_1, \dots, v_n) a basis of V . Recall that we can associate a matrix $M(T) \in \mathbb{C}^{n \times n}$ to the operator T . By Theorem 4 we know that T has at least one eigenvalue, say $\lambda \in \mathbb{C}$. Let $v_1 \neq 0$ be an eigenvector corresponding to λ . By the Basis Extension Theorem we can extend the list (v_1) to a basis of V . Since $Tv_1 = \lambda v_1$, the first column of $M(T)$ with respect to this basis is

$$\begin{bmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

What we will show next is that we can find a basis of V such that the matrix $M(T)$ is **upper triangular**.

Definition 3. A matrix $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ is called **upper triangular** if $a_{ij} = 0$ for $i > j$.

Schematically, an upper triangular matrix has the form

$$\begin{bmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{bmatrix}$$

where the entries $*$ can be anything, but below the diagonal the matrix has zero entries.

Some of the reasons why upper triangular matrices are so fantastic are that

1. the eigenvalues are on the diagonal (as we will see later);
2. it is easy to solve the corresponding system of linear equations by back substitution.

The next Proposition tells us what upper triangularity means in terms of linear operators and invariant subspaces.

Proposition 5. *Suppose $T \in \mathcal{L}(V, V)$ and (v_1, \dots, v_n) is a basis of V . Then the following statements are equivalent:*

1. the matrix $M(T)$ with respect to the basis (v_1, \dots, v_n) is upper triangular;
2. $Tv_k \in \text{span}(v_1, \dots, v_k)$ for each $k = 1, 2, \dots, n$;
3. $\text{span}(v_1, \dots, v_k)$ is invariant under T for each $k = 1, 2, \dots, n$.

Proof. The equivalence of 1 and 2 follows easily from the definition since 2 implies that the matrix elements below the diagonal are zero.

Obviously 3 implies 2. To show that 2 implies 3 note that any vector $v \in \text{span}(v_1, \dots, v_k)$ can be written as $v = a_1v_1 + \dots + a_kv_k$. Applying T we obtain

$$Tv = a_1Tv_1 + \dots + a_kTv_k \in \text{span}(v_1, \dots, v_k)$$

since by 2 each $Tv_j \in \text{span}(v_1, \dots, v_j) \subset \text{span}(v_1, \dots, v_k)$ for $j = 1, 2, \dots, k$ and the span is a subspace of V and hence closed under addition and scalar multiplication. \square

The next theorem shows that complex vector spaces indeed have some basis for which the matrix of a given operator is upper triangular.

Theorem 6. *Let V be a finite-dimensional vector space over \mathbb{C} and $T \in \mathcal{L}(V, V)$. Then there exists a basis for V such that $M(T)$ is upper triangular with respect to this basis.*

Proof. We proceed by induction on $\dim V$. If $\dim V = 1$ there is nothing to prove.

Hence assume that $\dim V = n > 1$ and we have proved the result of the theorem for all $T \in \mathcal{L}(W, W)$, where W is a complex vector space with $\dim W \leq n - 1$. By Theorem 4 T has at least one eigenvalue λ . Define

$$U = \text{range}(T - \lambda I).$$

Note that

1. $\dim U < \dim V = n$ since λ is an eigenvalue of T and hence $T - \lambda I$ is not surjective;
2. U is an invariant subspace of T since for all $u \in U$ we have

$$Tu = (T - \lambda I)u + \lambda u$$

which implies that $Tu \in U$ since $(T - \lambda I)u \in \text{range}(T - \lambda I) = U$ and $\lambda u \in U$.

Therefore we may consider the operator $S = T|_U$, being the operator T restricted to the subspace U . By induction hypothesis there exists a basis (u_1, \dots, u_m) of U with $m \leq n - 1$ such that $M(S)$ is upper triangular with respect to (u_1, \dots, u_m) . This means that

$$Tu_j = Su_j \in \text{span}(u_1, \dots, u_j) \quad \text{for all } j = 1, 2, \dots, m.$$

Extend this to a basis $(u_1, \dots, u_m, v_1, \dots, v_k)$ of V . Then

$$Tv_j = (T - \lambda I)v_j + \lambda v_j \quad \text{for all } j = 1, 2, \dots, k.$$

Since $(T - \lambda I)v_j \in \text{range}(T - \lambda I) = U = \text{span}(u_1, \dots, u_m)$, we have that

$$Tv_j \in \text{span}(u_1, \dots, u_m, v_1, \dots, v_j) \quad \text{for all } j = 1, 2, \dots, k.$$

Hence T is upper triangular with respect to the basis $(u_1, \dots, u_m, v_1, \dots, v_k)$. \square

There are two very important facts about upper triangular matrices and their associated operators.

Proposition 7. *Suppose $T \in \mathcal{L}(V, V)$ is a linear operator and $M(T)$ is upper triangular with respect to some basis of V . Then*

1. T is invertible if and only if all entries on the diagonal of $M(T)$ are nonzero.
2. The eigenvalues of T are precisely the diagonal elements of $M(T)$.

Proof of Proposition 7 Part 1. Let (v_1, \dots, v_n) be a basis of V such that

$$M(T) = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

is upper triangular. The claim is that T is invertible if and only if $\lambda_k \neq 0$ for all $k = 1, 2, \dots, n$. Equivalently, this can be reformulated as T is not invertible if and only if $\lambda_k = 0$ for at least one $k \in \{1, 2, \dots, n\}$.

Suppose $\lambda_k = 0$. We will show that then T is not invertible. If $k = 1$, this is obvious since then $Tv_1 = 0$, which implies that $v_1 \in \text{null } T$ so that T is not injective and hence not invertible. So assume that $k > 1$. Then

$$Tv_j \in \text{span}(v_1, \dots, v_{k-1}) \quad \text{for all } j \leq k$$

since T is upper triangular and $\lambda_k = 0$. Hence we may define $S = T|_{\text{span}(v_1, \dots, v_k)}$ to be the restriction of T to the subspace $\text{span}(v_1, \dots, v_k)$

$$S : \text{span}(v_1, \dots, v_k) \rightarrow \text{span}(v_1, \dots, v_{k-1}).$$

The linear map S is not injective since the dimension of the domain is bigger than the dimension of its codomain

$$\dim \text{span}(v_1, \dots, v_k) = k > k - 1 = \dim \text{span}(v_1, \dots, v_{k-1}).$$

Hence there exists a vector $0 \neq v \in \text{span}(v_1, \dots, v_k)$ such that $Sv = Tv = 0$. This implies that T is also not injective and therefore not invertible.

Now suppose that T is not invertible. We need to show that at least one $\lambda_k = 0$. The linear map T not being invertible implies that T is not injective. Hence there exists a vector $0 \neq v \in V$ such that $Tv = 0$. We can write

$$v = a_1v_1 + \cdots + a_kv_k$$

for some k where $a_k \neq 0$. Then

$$0 = Tv = (a_1Tv_1 + \cdots + a_{k-1}Tv_{k-1}) + a_kTv_k. \quad (2)$$

Since T is upper triangular with respect to the basis (v_1, \dots, v_n) we know that $a_1Tv_1 + \cdots + a_{k-1}Tv_{k-1} \in \text{span}(v_1, \dots, v_{k-1})$. Hence (2) shows that $Tv_k \in \text{span}(v_1, \dots, v_{k-1})$, which implies that $\lambda_k = 0$. \square

Proof of Proposition 7 Part 2. Recall that $\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if the operator $T - \lambda I$ is not invertible. Let (v_1, \dots, v_n) be a basis such that $M(T)$ is upper triangular. Then

$$M(T - \lambda I) = \begin{bmatrix} \lambda_1 - \lambda & & * \\ & \ddots & \\ 0 & & \lambda_n - \lambda \end{bmatrix}$$

Hence by Part 1 of Proposition 7, $T - \lambda I$ is not invertible if and only if $\lambda = \lambda_k$ for some k . \square

6 Diagonalization of 2×2 matrices and Applications

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{F}^{2 \times 2}$, and recall that we can define a linear operator $T \in \mathcal{L}(\mathbb{F}^2)$ on \mathbb{C}^2 by setting $T(v) = Av$ for each $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{F}^2$.

One method for finding the eigen-information of T is to analyze the solutions of the matrix equation $Av = \lambda v$ for $\lambda \in \mathbb{F}$ and $v \in \mathbb{F}^2$. In particular, using the definition of eigenvector and eigenvalue, v is an eigenvector associated to the eigenvalue λ if and only if $Av = T(v) = \lambda v$.

A simpler method involves the equivalent matrix equation $(A - \lambda I)v = 0$, where I denotes the identity map on \mathbb{F}^2 . In particular, $0 \neq v \in \mathbb{F}^2$ is an eigenvector for T associated to the

eigenvalue $\lambda \in \mathbb{F}$ if and only if the system of linear equations

$$\left. \begin{aligned} (a - \lambda)v_1 + bv_2 &= 0 \\ cv_1 + (d - \lambda)v_2 &= 0 \end{aligned} \right\} \quad (3)$$

has a non-trivial solution. Moreover, the system of equations (3) has a non-trivial solution if and only if the polynomial $p(\lambda) = (a - \lambda)(d - \lambda) - bc$ evaluates to zero (see Homework 7, Problem 1).

In other words, the eigenvalues for T are exactly the $\lambda \in \mathbb{F}$ for which $p(\lambda) = 0$, and the eigenvectors for T associated to an eigenvalue λ are exactly the non-zero vectors $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{F}^2$ that satisfy the system of equations (3).

Example 3. Let $A = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$. Then $p(\lambda) = (-2 - \lambda)(2 - \lambda) - (-1)(5) = \lambda^2 + 1$, which is equal to zero exactly when $\lambda = \pm i$. Moreover, if $\lambda = i$, then the system of equations (3) becomes

$$\left. \begin{aligned} (-2 - i)v_1 - v_2 &= 0 \\ 5v_1 + (2 - i)v_2 &= 0 \end{aligned} \right\}$$

which is satisfied by any vector $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{C}^2$ such that $v_2 = (-2 - i)v_1$. Similarly, if $\lambda = -i$, then the system of equations (3) becomes

$$\left. \begin{aligned} (-2 + i)v_1 - v_2 &= 0 \\ 5v_1 + (2 + i)v_2 &= 0 \end{aligned} \right\}$$

which is satisfied by any vector $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{C}^2$ such that $v_2 = (-2 + i)v_1$.

It follows that, given $A = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$, the linear operator on \mathbb{C}^2 defined by $T(v) = Av$ has eigenvalues $\lambda = \pm i$, with associated eigenvectors as described above.

Example 4. Take the rotation $R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by an angle $\theta \in [0, 2\pi)$ given by the matrix

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

Then we obtain the eigenvalues by solving the polynomial equation

$$\begin{aligned} p(\lambda) &= (\cos \theta - \lambda)^2 + \sin^2 \theta \\ &= \lambda^2 - 2\lambda \cos \theta + 1 = 0, \end{aligned}$$

where we used that $\sin^2 \theta + \cos^2 \theta = 1$. Solving for λ in \mathbb{C} we obtain

$$\lambda = \cos \theta \pm \sqrt{\cos^2 \theta - 1} = \cos \theta \pm \sqrt{-\sin^2 \theta} = \cos \theta \pm i \sin \theta = e^{\pm i\theta}.$$

We see that as an operator over the real vector space \mathbb{R}^2 , the operator R_θ only has eigenvalues when $\theta = 0$ or $\theta = \pi$. However, if we interpret the vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ as a complex number $z = x_1 + ix_2$, then z is an eigenvector if $R_\theta : \mathbb{C} \rightarrow \mathbb{C}$ maps $z \mapsto \lambda z = e^{\pm i\theta} z$. We already saw in the lectures on complex numbers that multiplication by $e^{\pm i\theta}$ corresponds to rotation by the angle $\pm\theta$.