Axler: Eigenvalues and Eigenvectors

December 2023

James Pagan

Contents

1 Invariant Subspaces

1.1 Eigenvalues

Suppose $\mathbf{T} \in \mathcal{L}(V)$. A subspace U of V is called **invariant** under $\mathbf{T}$ if $\mathbf{Tu} \in U$ for all $\mathbf{u} \in U$. A number $\lambda \in \mathbb{F}$ is called an **eigenvalue** of $\mathbf{T} \in \mathcal{L}(V)$ if there exists $\mathbf{v} \in V$ such that $\mathbf{Tv} = \lambda\mathbf{v}$.

Theorem 1. *Suppose V is finite-dimensional, $\mathbf{T} \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. Then the following are equivalent:*

1. *λ is an eigenvalue of T.*
2. *$\mathbf{T} - \lambda\mathbf{I}$ is not injective.*
3. *$\mathbf{T} - \lambda\mathbf{I}$ is not surjective.*
4. *$\mathbf{T} - \lambda\mathbf{I}$ is not bijective.*

Proof. Conditions (1) and (2) are equivalent, as $\mathbf{Tv} = \lambda\mathbf{v}$ if and only if $(\mathbf{T} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$. Conditions (2), (3), and (4) are equivalent by the fact V is finite-dimensional. □

Suppose $\mathbf{T} \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$ is an eigenvalue of $\mathbf{T}$. A vector $\mathbf{v} \in V$ is called an **eigenvector** of $\mathbf{T}$ corresponding to λ if $\mathbf{v} \neq \mathbf{0}$ and $\mathbf{Tv} = \lambda\mathbf{v}$. Such eigenvectors biconditionally satisfy $\mathbf{v} \in \mathrm{null}(\mathbf{T} - \lambda\mathbf{I})$.

Theorem 2. *Suppose $\mathbf{T} \in \mathcal{L}(V)$. Then every list of eigenvectors of $\mathbf{T}$ corresponding to distinct eigenvalues of $\mathbf{T}$ is linearly independent.*

Proof. Suppose the desired result is false. Let m be the smallest positive integer such that the list of eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_m$ is dependent. As $m > 1$, there exist $\mu_1, \ldots, \mu_m \in \mathbb{F}$ — none of which are zero, by the minimality of m — such that

$$\mu_1\mathbf{v}_1 + \cdots + \mu_m\mathbf{v}_m = \mathbf{0}.$$

Applying $\mathbf{T} - \lambda\mathbf{I}$ to this equation, we find that

$$\mu_1(\lambda_1 - \lambda_m)\mathbf{v}_1 + \cdots + \mu_{m-1}(\lambda_{m-1} - \lambda_m)\mathbf{v}_m = \mathbf{0}.$$

None of the coefficents above equal zero, as the eigenvalues are distinct and $\mu_1, \ldots, \mu_m$ are nonzero. Thus, $\mathbf{v}_1, \ldots \mathbf{v}_{m-1}$ are linearly dependent — which violates the minimality of m, yielding our desired contradiction. □

The proof above is beautiful, yielding a swift execution to the following theorem:

Theorem 3. *Suppose V is finite-dimensional. Then each operator on V has at most* $\dim V$ *distinct eigenvalues.*

Proof. Suppose $\mathbf{T}$ has distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be nonzero vectors that correspond to these eigenvalues; by Lemma 2, they are independent. Then $m \leq \dim V$, as desired. $\square$

1.2 Polynomials Applied to Operators

Suppose $\mathbf{T} \in \mathcal{L}(V)$ and $m \in \mathbb{Z}_{>0}$. Then

- $\mathbf{T}^m \in \mathcal{L}(V)$ is defined to be $\mathbf{T} \cdots \mathbf{T}$ (m times).
- $\mathbf{T}^0 \in \mathcal{L}(V)$ is defined to be $\mathbf{I}$.
- $\mathbf{T}^{-m} \in \mathcal{L}(V)$ is defined to be $(\mathbf{T}^{-1})^m$, if $\mathbf{T}$ is invertible.

It is easy to verify that $\mathbf{T}^{n+m} = \mathbf{T}^n \mathbf{T}^n$ and $(\mathbf{T}^n)^m = \mathbf{T}^{nm}$. Now, suppose $\mathbf{T} \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$ is a polynomial of the form

$$p(z) = a_m p^m + \cdots + a_1 p + a_0.$$

Then $p(\mathbf{T})$ is the operator on V defined by

$$p(\mathbf{T}) = a_m \mathbf{T}^m + \cdots + a_1 \mathbf{T} + a_0 \mathbf{I}.$$

If $p, q \in \mathcal{P}(\mathbb{F})$, we further define $pq(z) = p(z)q(z)$ for all $z \in \mathbb{F}$. Order here is irrelevant:

Theorem 4. $(pq)(\mathbf{T}) = p(\mathbf{T})q(\mathbf{T})$ *and* $p(\mathbf{T})q(\mathbf{T}) = q(\mathbf{T})p(\mathbf{T})$.

Proof. Suppose $p(z) = \sum_{i=0}^{n} a_i z_i$ and $q(z) = \sum_{j=0}^{m} b_j z_j$. Then $(pq)(z) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j z^{i+j}$, so

$$\begin{aligned}
(pz)(\mathbf{T}) &= \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j \mathbf{T}^{i+j} \\
&= \left(\sum_{i=0}^{n} a_i \mathbf{T}^i \right) \left(\sum_{j=0}^{m} b_j \mathbf{T}^j \right) \\
&= p(\mathbf{T})q(\mathbf{T}).
\end{aligned}$$

For the second result, see that $p(\mathbf{T})q(\mathbf{T}) = (pq)(\mathbf{T}) = (qp)(\mathbf{T}) = q(\mathbf{T})p(\mathbf{T})$. $\square$

Theorem 5. *Suppose* $\mathbf{T} \in \mathcal{L}(V)$ *and* $p \in \mathcal{P}(\mathbb{F})$. *Then* $\operatorname{null} p(\mathbf{T})$ *and* $\operatorname{range} p(\mathbf{T})$ *are invariant subspaces under* $\mathbf{T}$.

Proof. Clearly, $\operatorname{null} \mathbf{T}$ and $\operatorname{range} \mathbf{T}$ are invariant under the operator $\mathbf{T}$. Now, suppose $\mathbf{u} \in \operatorname{null} p(\mathbf{T})$; then

$$(p(\mathbf{T}))(\mathbf{T}\mathbf{u}) = \mathbf{T}(p(\mathbf{T})(\mathbf{u})) = \mathbf{T}(\mathbf{0}) = \mathbf{0},$$

so $\mathbf{T}\mathbf{u} \in \operatorname{null} p(\mathbf{T})$, and $\operatorname{null} p(\mathbf{T})$ is invariant. Clearly $\mathbf{u} \in \operatorname{range} p(\mathbf{T})$ implies that

$$p(\mathbf{T})(\mathbf{T}\mathbf{u}) = \mathbf{T}(p(\mathbf{T})\mathbf{u}) \in \operatorname{range} \mathbf{T};$$

we conclude that the null space and range of $p(\mathbf{T})$ are invariant under $\mathbf{T}$. $\square$

2 The Minimal Polynomial

2.1 Existence of Eigenvalues on Complex Vector Spaces

Theorem 6. *Every operator on a nonzero complex vector space* V *with finite dimension* n *has an eigenvalue.*

Proof. Choose $\mathbf{v} \in V$, such that $\mathbf{v} \neq 0$. Then

$$\mathbf{v}, \mathbf{T}\mathbf{v}, \mathbf{T}^2\mathbf{v}, \ldots, \mathbf{T}^n\mathbf{v}$$

is a dependent list. Hence there exists a linear combination of these vectors equal to $\mathbf{0}$. Simplifying, we find a nonconstant polynomial p of *smallest degree* such that

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

By the Fundamental Theorem of Algebra, this polynomial has a root λ. Then

$$p(z) = (z - \lambda)q(z)$$

for some polynomial q. Then using the multiplicative properties of polynomials,

$$\mathbf{0} = p(\mathbf{T})\mathbf{v} = (\mathbf{T} - \lambda\mathbf{I})(q(\mathbf{T})\mathbf{v}).$$

As q has degree smaller than p, the expression $q(\mathbf{T})\mathbf{v}$ is never the zero vector. Thus, the above equation implies that λ is an eigenvector of $\mathbf{T}$ with eigenvector $q(\mathbf{T})\mathbf{v}$. $\square$

The theorem above fails if $\mathbb{C}$ is replaced with $\mathbb{R}$ or if V is infinite dimensional.

2.2 Eigenvalues and the Minimal Polynomial

Theorem 7. *Suppose V is finite-dimensional and $\mathbf{T} \in \mathcal{L}(V)$. Then there is a unique monic polynomial $p \in \mathcal{P}(\mathbb{F})$ of smallest degree such that $p(\mathbf{T}) = \mathbf{0}$. Furthermore,* $\deg p \leq \dim V$.

Proof. We proceed via strong induction. If $\dim V = 0$, then the constant polynomial 1 suffices – thus, we assume the existence, uniqueness, and degree of the polynomial p for $\dim V \in \{0, \ldots, n-1\}$.

Let $\dim V = n$ and select some nonzero $\mathbf{v} \in V$. Consider the family of vectors

$$\mathbf{v}, \mathbf{T}\mathbf{v}, \mathbf{T}^2\mathbf{v}, \ldots, \mathbf{T}^n\mathbf{v}.$$

It has length $n+1$, so it must be dependent. Then by the Linear Dependence Lemma, there exists an integer m such that $\mathbf{v}, \mathbf{T}\mathbf{v}, \ldots \mathbf{T}^{m-1}\mathbf{v}$ is independent but $\mathbf{v}, \mathbf{T}\mathbf{v}, \ldots, \mathbf{T}^m\mathbf{v}$ is not — namely, that there exist scalars $c_0, \ldots, c_{m-1} \in \mathbb{F}$ such that

$$\mathbf{T}^m\mathbf{v} + c_{m-1}\mathbf{T}^{m-1}\mathbf{v} + \cdots + c_1\mathbf{T}\mathbf{v} + c_0\mathbf{v} = \mathbf{0}.$$

Define the monic polynomial $q \in \mathcal{P}_m(\mathbb{F})$ by $q(z) = z^m + c_{m-1}z^{m-1} + \cdots + c_1 z + c_0$; then the above equation reads $q(\mathbf{T})\mathbf{v} = \mathbf{0}$. Now, realize that for all $k \in \{0, \ldots, m-1\}$,

$$q(\mathbf{T})(\mathbf{T}^k\mathbf{v}) = \mathbf{T}^k(q(\mathbf{T})\mathbf{v}) = \mathbf{T}^k\mathbf{0} = \mathbf{0}.$$

Hence, $\dim \operatorname{null} q(\mathbf{T}) \geq m$. Then by the Fundamental Theorem of Linear Maps,

$$\dim \operatorname{range} q(\mathbf{T}) = \dim V - \dim \operatorname{null} q(\mathbf{T}) \leq n - m.$$

Then because $\dim \operatorname{range} q(\mathbf{T}) < n$, our inductive hypothesis applies to the vector space $\dim \operatorname{range} q(\mathbf{T})$ and the operator $T \mid_{\operatorname{range} q(\mathbf{T})}$. We deduce the existence of a unique monic polynomial s of smallest degree with

$$s(\mathbf{T} \mid_{\operatorname{range} q(\mathbf{T})}) = \mathbf{0} \qquad \text{and} \qquad \deg s \leq n - m.$$

We claim that $(sq)(\mathbf{T}) = \mathbf{0}$. For all $\mathbf{v} \in V$, realize that $q(\mathbf{T})\mathbf{v} \in \operatorname{range} q(\mathbf{T})$; thus,

$$(sq)(\mathbf{T})\mathbf{v} = s(\mathbf{T})(q(\mathbf{T})\mathbf{T}) = \mathbf{0}.$$

Furthermore, the degree of sq satisfies the desired requirement:

$$\deg sq = \deg s + \deg q \leq (n-m) + m = n.$$

We have identified a monic polynomial of degree at most n which when applied to T returns the zero operator. Thus, there exist a monic polynomial of *smallest degree* with this property; all that remains to be proven is its uniqueness.

Suppose $p, q \in \mathcal{P}(\mathbb{F})$ are two monic polynomials of the smallest degree m such that $p(\mathbf{T}) = q(\mathbf{T}) = \mathbf{0}$. Then $(p - q)(\mathbf{T}) = \mathbf{0}$ and $\deg(p - q) < m$; if we simply divide $p - q$ by its leading coefficent, weis a polynomial multiple of the minimal polynomial of find a monic polynomial that contradicts the minimality of m. Hence $p = q$, which completes the proof. □

For a finite-dimensional vector space V and operator $\mathbf{T} \in \mathcal{L}(V)$, the **minimal polynomial** of $\mathbf{T}$ is the unique monic polynomial $p \in \mathcal{P}(\mathbb{F})$ such that $p(\mathbf{T}) = \mathbf{0}$.

Theorem 8. *The zeroes of the minimal polynomial of* $\mathbf{T}$ *are the eigenvalues of* $\mathbf{T}$*.*

Proof. Let p be the minimal polynomial of $\mathbf{T}$. Suppose that $\lambda \in \mathbb{F}$ is a zero of p; then for some monic polynomial $p \in \mathcal{P}(\mathbb{F})$,

$$p(z) = (z - \lambda)q(z).$$

Because $p(\mathbf{T}) = \mathbf{0}$, we have that for all $\mathbf{v} \in V$,

$$\mathbf{0} = (\mathbf{T} - \lambda\mathbf{I})(q(\mathbf{T})\mathbf{v}).$$

As q has smaller degree than the minimal polynomial, there exists $\mathbf{w} \in V$ such that $q(\mathbf{T})\mathbf{w} \neq \mathbf{0}$; the above equation implies that $q(\mathbf{T})\mathbf{w}$ must be an eigenvector of $\mathbf{T}$ with eigenvalue λ.

Now, suppose that $\lambda \in \mathbb{F}$ is an eigenvalue of $\mathbf{T}$. Then for some $\mathbf{T}\mathbf{v} = \lambda\mathbf{v}$ for some nonzero $\mathbf{v} \in V$; iterated applications yield that $\mathbf{T}^k\mathbf{v} = \lambda^k\mathbf{v}$ for all $k \in \mathbb{Z}_{\geq 0}$, so

$$p(\mathbf{T})\mathbf{v} = p(\lambda)\mathbf{v}.$$

the left-hand side is $\mathbf{0}$; then $p(\lambda)$ must be zero, and λ is a root of the minimal polynomial. □

If V is a finite-dimensional vector space over $\mathbb{C}$ and $\mathbf{T} \in \mathcal{L}(\mathbb{F})$, then the minimal polynomial of $\mathbf{T}$ has the form

$$(z - \lambda_1) \cdots (z - \lambda_m),$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{T}$, possibly with repetitions.

Theorem 9. *Suppose V is finite dimensional, $\mathbf{T} = \mathcal{L}(V)$, and $q \in \mathcal{P}(\mathbb{F})$. Then $q(\mathbf{T}) = \mathbf{0}$ if and only if q is a polynomial multiple of the minimal polynomial of $\mathbf{T}$.*

Proof. Let the minimal polynomial of $\mathbf{T}$ be p. As $\deg q \geq \deg p$, we may divide them to deduce the existence of $s, r \in \mathcal{P}(\mathbb{F})$ such that

$$q = ps + r,$$

where $\deg r < \deg p$. Then $q(\mathbf{T}) = p(\mathbf{T})s(\mathbf{T}) + r(\mathbf{T})$. Because $q(\mathbf{T}) = p(\mathbf{T}) = \mathbf{0}$, this equation simplifies to

$$\mathbf{0} = r(\mathbf{T}).$$

If r was nonzero, then we could divide by its leading coefficent to yield a polynomial that contradics the minimality of p. Thus $r = 0$.

If q is a polynomial multiple of p, then there exists $s \in \mathcal{P}(\mathbb{F})$ such that $q = ps$. Then

$$\mathbf{0} = p(\mathbf{T})s(\mathbf{T}) = q(\mathbf{T}),$$

which completes the proof. □

The next result is a nice consequence of the above.

Theorem 10. *Suppose V is finite-dimensional, $\mathbf{T} \in \mathcal{L}(V)$, and U is an invariant subspace of V. Then the minimal polynomial of $\mathbf{T}$ is a polynomial multiple of the minimal polynomial of $\mathbf{T}|_U$.*

Proof. Let p be the minimal polynomial of $\mathbf{T}$. Then for all $\mathbf{u} \in U$,

$$p(\mathbf{T})\mathbf{u} = \mathbf{0}.$$

We conclude that $p(\mathbf{T}|_U) = \mathbf{0}$, so p is a polynomial multiple of the minimal polynomial of $\mathbf{T}|_U$. □

Theorem 11. *Suppose V is finite-dimensional and $\mathbf{T} \in \mathcal{L}(V)$. Then T is not invertible if and only if the constant term of the minimal polynomial of T is 0.*

Proof. If $\mathbf{T}$ is not invertible, then 0 must be an eigenvalue of $\mathbf{T}$. Then 0 is a root of p, which implies that 0 does not have a constant term. The reverse of our steps holds as well. □

3 Upper-Triangular Matricies

3.1 Matrix Prerequisites

Suppose $\mathbf{T} \in \mathcal{L}(V)$. The **matrix** of $\mathbf{T}$ with respect to a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is the n-by-n matrix

$$\mathcal{M}(\mathbf{T}) = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix},$$

whose entries A are defined by

$$\mathbf{T}\mathbf{v}_k = A_{1k}\mathbf{v}_1 + \cdots + A_{nk}\mathbf{v}_n.$$

Thus, the k-th column of the matrix $\mathcal{M}(\mathbf{T})$ is formed from the coefficents used to write $\mathbf{T}\mathbf{v}_k$ as a linear combination of the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

The **diagonal** of a square matrix consists of all the entries A_{kk} for each $k \in \{1, \ldots, n\}$. If all the entires below the diagonal are 0, the square matrix is called **upper triangular**.

Theorem 12. *If $\mathbf{T} \in \mathcal{L}(V)$ and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of V, then the following conditions are equivalent:*

1. *The matrix of $\mathbf{T}$ with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is upper triangular.*
2. $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *is invariant under $\mathbf{T}$ for each $k \in \{1, \ldots, n\}$.*
3. $\mathbf{T}\mathbf{v}_k \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ *for each $k \in \{1, \ldots, n\}$*

   Proof. Suppose that (1) holds. Then for each $i \in \{1, \ldots, n\}$,

   $$\mathbf{T}\mathbf{v}_i = A_{1i}\mathbf{v}_1 + \cdots + A_{ii}\mathbf{v}_i.$$

   Let $k \in \{1, \ldots, n\}$. For all $\mathbf{w} \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, there exist scalars $\lambda_1, \ldots, \lambda_k$ such that $\mathbf{w} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_k\mathbf{v}_k$. Therefore,

   $$\begin{aligned} \mathbf{T}\mathbf{w} &= \sum_{i=1}^{k} \mathbf{T}(\lambda_i\mathbf{v}_i) \\ &= \sum_{i=1}^{k} \lambda_1(A_{1i}\mathbf{v}_1 + \ldots + A_{ii}\mathbf{v}_i) \\ &= (\lambda_1 A_{11} + \ldots + \lambda_k A_{1k})\mathbf{v}_1 + \cdots + (\lambda_k A_{kk})\mathbf{v}_k \\ &\in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k). \end{aligned}$$

Then $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is invariant under $\mathbf{T}$ for each $k \in \{1, \ldots, n\}$. If (2) holds, then setting $\mathbf{w} = \mathbf{v}_k$ achieves (3).

Now, suppose (3) holds. Then for each $k \in \{1, \ldots, n\}$, there exist scalars such that

$$\mathbf{T}\mathbf{v}_k = A_{1k}\mathbf{v}_1 + \cdots + A_{kk}\mathbf{v}_k.$$

As $\mathbf{v}_1, \ldots, \mathbf{v}_n$ constitute a basis of V, we conclude that the unique scalars that express $\mathbf{T}_k$ as a linear combination of $\mathbf{v}_k$ are $A_{1k}, \ldots, A_{kk}, 0, \ldots$ respectively. Thus, the entries of $\mathcal{M}(\mathbf{T})$ are zero below the main diagonal, implying (1). □

Theorem 13. *For some $\mathbf{T} \in \mathcal{L}(V)$, suppose that there exists a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ such that $\mathcal{M}(\mathbf{T})$ is upper triangular. Then if $\lambda_1, \ldots, \lambda_n$ are its diagonal entries, $\mathbf{T}$ satisfies the equation*

$$(\mathbf{T} - \lambda_1\mathbf{I}) \cdots (\mathbf{T} - \lambda_n\mathbf{I}) = \mathbf{0}.$$

Proof. □