

Chapter 5

Eigenvalues, Eigenvectors, and Invariant Subspaces

5.A Invariant Subspaces

- 9/8: • Let $T \in \mathcal{L}(V)$, and let V be decomposable into a direct sum of proper subspaces as follows.

$$V = U_1 \oplus \cdots \oplus U_m$$

- To understand T , we need only understand each **restriction** of T to a U_j .
- Since $T|_{U_j}$ may not map U_j onto itself in every case, to use operator-based tools, we need to consider only direct sum decompositions into subspaces that T maps onto themselves, or **invariant subspace**.
- **Invariant subspace** (of V under T): A subspace U of V such that $u \in U$ implies $Tu \in U$, where $T \in \mathcal{L}(V)$.
 - In other words, U is invariant under T iff $T|_U \in \mathcal{L}(U)$.
- Some invariant subspaces under $T \in \mathcal{L}(V)$: $\{0\}$, V , null T , and range T .
- **Invariant subspace problem**: The most famous unsolved problem in functional analysis, dealing with invariant subspaces of operators on infinite-dimensional vector spaces.
- To begin our study of invariant subspaces, we consider the simplest possible type of invariant subspace: those with dimension 1.
- Every 1-dimensional subspace of V is of the form $\text{span}(v)$ for some $v \in V$.
 - If $\text{span}(v)$ is invariant under $T \in \mathcal{L}(V)$, then $Tv \in \text{span}(v)$.
 - If $Tv \in \text{span}(v)$, then there exists $\lambda \in \mathbb{F}$ such that $Tv = \lambda v$.
- **Eigenvalue** (of T): A number $\lambda \in \mathbb{F}$ such that there exists a nonzero vector $v \in V$ satisfying the equation $Tv = \lambda v$. *Also known as **characteristic value**.*
- “ T has a 1-dimensional invariant subspace if and only if T has an eigenvalue” (Axler, 2015, p. 134).
- We now give some conditions λ can satisfy to be deemed an eigenvalue.

Theorem 5.1. *Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, $I \in \mathcal{L}(V)$ is the identity operator on V , and $\lambda \in \mathbb{F}$. Then the following are equivalent.*

- (a) λ is an eigenvalue of T .

- (b) $T - \lambda I$ is not injective.
- (c) $T - \lambda I$ is not surjective.
- (d) $T - \lambda I$ is not invertible.

Proof. Suppose first that λ is an eigenvalue of T . Then

$$\begin{aligned}Tv &= \lambda v \\Tv &= \lambda Iv \\Tv - \lambda Iv &= 0 \\(T - \lambda I)v &= 0\end{aligned}$$

for some $v \in V$ such that $v \neq 0$. It follows that $v \in \text{null}(T - \lambda I)$, so by Theorem 3.4, $T - \lambda I$ is not injective, as desired. The proof is symmetric in the other direction. Therefore, conditions (a) and (b) are equivalent.

To prove that (a), (b), (c), and (d) are equivalent at this point, it will suffice to show that (b), (c), and (d) are equivalent. But we have this by Theorem 3.12, as desired. ■

- **Eigenvector** (of T): A nonzero vector $v \in V$ such that there exists a $\lambda \in \mathbb{F}$ satisfying the equation $Tv = \lambda v$.
- Since $Tv = \lambda v$ iff $(T - \lambda I)v = 0$, “a vector $v \in V$ with $v \neq 0$ is an eigenvector of T corresponding to λ if and only if $v \in \text{null}(T - \lambda I)$ ” (Axler, 2015, p. 135).
- Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Theorem 5.2. Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T and v_1, \dots, v_m are corresponding eigenvectors. Then v_1, \dots, v_m is linearly independent.

Proof. Suppose for the sake of contradiction that v_1, \dots, v_m is linearly dependent. Then by the Linear Dependence Lemma, we may let k be the smallest positive integer such that $v_k \in \text{span}(v_1, \dots, v_{k-1})$. It follows that

$$v_k = a_1 v_1 + \dots + a_{k-1} v_{k-1}$$

for some $a_1, \dots, a_{k-1} \in \mathbb{F}$. Thus, applying T , we have that

$$\begin{aligned}Tv_k &= a_1 T v_1 + \dots + a_{k-1} T v_{k-1} \\ \lambda_k v_k &= a_1 \lambda_1 v_1 + \dots + a_{k-1} \lambda_{k-1} v_{k-1}\end{aligned}$$

If we multiply the first equation by λ_k and subtract the above equation from it, we get that

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

But since k is the smallest positive integer j such that $v_j \in \text{span}(v_1, \dots, v_{j-1})$, we know that v_1, \dots, v_{k-1} are linearly independent. Thus, $a_1(\lambda_k - \lambda_1) = \dots = a_{k-1}(\lambda_k - \lambda_{k-1}) = 0$. More specifically, since all eigenvalues are distinct (i.e., $\lambda_k - \lambda_j \neq 0$ for any $j = 1, \dots, k-1$), we must have that $a_1 = \dots = a_{k-1} = 0$. But this implies that

$$\begin{aligned}v_k &= a_1 v_1 + \dots + a_{k-1} v_{k-1} \\ &= 0\end{aligned}$$

contradicting the fact that v_k , as an eigenvector, is nonzero. ■

- We now put a bound on the number of eigenvalues.

Theorem 5.3. Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

Proof. Let $T \in \mathcal{L}(V)$ have distinct eigenvalues $\lambda_1, \dots, \lambda_m$ and corresponding eigenvectors v_1, \dots, v_m . Then by Theorem 5.2, v_1, \dots, v_m is linearly independent. It follows by Theorem 2.3 that $m \leq \dim V$ ■

- **Restriction operator** (of $T : V \rightarrow W$ to $U \subset V$): The function $T|_U : U \rightarrow W$ defined by $T|_U(u) = Tu$ for all $u \in U$. Denoted by $T|_U$.
 - The fact that U is invariant under T is what allows us to consider $T|_U$ to be in $\mathcal{L}(U)$ as opposed to just $\mathcal{L}(V)$.
- **Quotient operator**: The operator $T/U \in \mathcal{L}(V/U)$ defined by $(T/U)(v + U) = Tv + U$ for all $v \in V$.
- Axler (2015) verifies that the restriction operator and the quotient operator actually *are* operators, in general, as defined.

5.B Eigenvectors and Upper-Triangular Matrices

- 9/10:
- If an operator $T \in \mathcal{L}(V)$, then $TT = T^2 \in \mathcal{L}(V)$.
 - T^m : The operator $T^m \in \mathcal{L}(V)$ defined by

$$T^m = \underbrace{T \cdots T}_{m \text{ times}}$$

where $T \in \mathcal{L}(V)$, $m \in \mathbb{N}$.

- T^0 : The identity operator $I \in \mathcal{L}(V)$, where $T \in \mathcal{L}(V)$.
- T^{-m} : The operator $T^{-m} \in \mathcal{L}(V)$ defined by

$$T^{-m} = (T^{-1})^m$$

where $T \in \mathcal{L}(V)$ is invertible with inverse T^{-1} , and $m \in \mathbb{N}$.

- It follows from these definitions that

$$T^m T^n = T^{m+n} \qquad (T^m)^n = T^{mn}$$

for any $m, n \in \mathbb{Z}$ if T is invertible and for any $m, n \in \mathbb{N}$ if T is not invertible.

- $p(T)$: The operator defined by

$$p(T) = a_0 I + a_1 T + \cdots + a_m T^m$$

where $T \in \mathcal{L}(V)$, and $p \in \mathcal{P}(\mathbb{F})$ is defined by $p(z) = a_0 + a_1 z + \cdots + a_m z^m$ for all $z \in \mathbb{F}$.

- $f : \mathcal{P}(\mathbb{F}) \rightarrow \mathcal{L}(V)$ defined by $p \mapsto p(T)$ is linear.
- **Product** (of $p, q \in \mathcal{P}(\mathbb{F})$): The polynomial $pq \in \mathcal{P}(\mathbb{F})$ defined by $(pq)(z) = p(z)q(z)$ for all $z \in \mathbb{F}$.
- Multiplicative properties of $p(T)$.

Theorem 5.4. Suppose $p, q \in \mathcal{P}(\mathbb{F})$ and $T \in \mathcal{L}(V)$. Then

$$(a) \quad (pq)(T) = p(T)q(T).$$

Proof. Suppose $p(z) = \sum_{j=0}^m a_j z^j$ and $q(z) = \sum_{k=0}^n b_k z^k$ for all $z \in \mathbb{F}$. Then

$$(pq)(z) = \sum_{j=0}^m \sum_{k=0}^n a_j b_k z^{j+k}$$

so

$$\begin{aligned}
 (pq)(T) &= \sum_{j=0}^m \sum_{k=0}^n a_j b_k T^{j+k} \\
 &= \left(\sum_{j=0}^m a_j T^j \right) \left(\sum_{k=0}^n b_k T^k \right) \\
 &= p(T)q(T)
 \end{aligned}$$

as desired. ■

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

Proof. It follows from part (a) that $p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T)$, as desired. ■

- We now prove a central result concerning operators on complex vector spaces.

Theorem 5.5. *Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.*

Proof. Let V be nonzero complex vector space of dimension n , let $T \in \mathcal{L}(V)$, and let $v \in V$ be nonzero. Then since $\text{len}(v, Tv, T^2v, \dots, T^nv) > \dim V$, Theorem 2.3 implies that v, Tv, \dots, T^nv is not linearly independent. Thus, there exist $a_0, \dots, a_n \in \mathbb{F}$, not all 0, such that

$$0 = a_0v + a_1Tv + \dots + a_nT^nv$$

Additionally, we have that a_1, \dots, a_n do not all equal zero (suppose they did; then $0 = a_0v$, so $a_0 = 0$ because $v \neq 0$; this would imply that $a_0 = \dots = a_n = 0$, a contradiction). Thus, if we consider the (nonconstant, by the previous result) polynomial with coefficients equal to the a 's, Theorem 4.6 implies that

$$a_0 + a_1z + \dots + a_nz^n = c(z - \lambda_1) \cdots (z - \lambda_m)$$

for all $z \in \mathbb{C}$, where $c \in \mathbb{C}$ is nonzero and each $\lambda_j \in \mathbb{C}^{[1]}$. It follows that

$$\begin{aligned}
 0 &= a_0v + a_1Tv + \dots + a_nT^nv \\
 &= (a_0I + a_1T + \dots + a_nT^n)v \\
 &= c(T - \lambda_1I) \cdots (T - \lambda_mI)v
 \end{aligned}$$
Theorem 5.4

Each $T - \lambda_jI$ is a linear operator in its own right. As a linear map, we have by Theorem 3.2 that $(T - \lambda_jI)0 = 0$ for all $j = 1, \dots, m$. However, $v \neq 0$, meaning that there exists $1 \leq k \leq m$ such that $T - \lambda_kI$ maps its nonzero argument, which we may call v' and know to be the result of all of the operators to its right being applied successively to v via function composition, to 0 as well; every operator to the left of $T - \lambda_kI$ will then map the result to 0, resulting in the final equivalence. Thus, $(T - \lambda_kI)v' = (T - \lambda_kI)0$ but $v' \neq 0$ for this operator, meaning that it is not injective. But if $T - \lambda_kI$ is not injective, Theorem 5.1 implies that λ_k is an eigenvalue of T , as desired. ■

- **Matrix** (of $T \in \mathcal{L}(V)$ with respect to the basis v_1, \dots, v_n of V): The n -by- n matrix $\mathcal{M}(T)$ whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}v_1 + \dots + A_{n,k}v_n$$

– If the basis is not clear from context, then the notation $\mathcal{M}(T, (v_1, \dots, v_n))$ is used.

- “A central goal of linear algebra is to show that given an operator $T \in \mathcal{L}(V)$, there exists a basis of V with respect to which T has a reasonably simple [sparse] matrix” (Axler, 2015, p. 146).

¹Note that $m \neq n$ necessarily, because a_n may equal 0.

- For example, given just Theorems 3.16 and 5.5, we know that we can choose a basis of V such that the first column of $\mathcal{M}(T)$ for an arbitrary T will have an eigenvalue λ of T in the first row and zeros everywhere else. Specifically, let v_1 be the eigenvector corresponding to λ (guaranteed to exist by Theorem 5.5). Then by Theorem 3.16,

$$\begin{aligned}\mathcal{M}(T)_{\cdot,1} &= \mathcal{M}(Tv_1) \\ &= \mathcal{M}(\lambda v_1) \\ &= \lambda \mathcal{M}(v_1) + 0\mathcal{M}(v_2) + \cdots + 0\mathcal{M}(v_n)\end{aligned}$$

where we extend v_1 into a basis v_1, \dots, v_n of V .

- **Diagonal** (of a square matrix): The entries along the line from the upper left corner to the bottom right corner.
- **Upper-triangular matrix**: A matrix such that all the entries below the diagonal equal 0. *Denoted by*

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}^{[2]}$$

²We often use $*$ to denote matrix entries that we do not know about or that are irrelevant to the questions being discussed.