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 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 span(v) for some $v \in V$.
  - If 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

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

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

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

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

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

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

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

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, $a_1(\lambda_k - \lambda_1) = \cdots = a_{k-1}(\lambda_k - \lambda_{k-1}) = 0$. 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 $a_1 = \cdots = a_{k-1} = 0$. But this implies that

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

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 $T : V \to W$ to $U \subset V$): The function $T|_U : U \to W$ defined by $T|_U(u) = Tu$ for all $u \in U$. *Denoted by* $\boldsymbol{T|_U}$.
  - 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 $TT = T^2 \in \mathcal{L}(V)$.
- $\boldsymbol{T^m}$: The operator $T^m \in \mathcal{L}(V)$ defined by

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

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

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

  for any $m, n \in \mathbb{Z}$ if T is invertible and for any $m, n \in \mathbb{N}$ if T is not invertible.
- $\boldsymbol{p(T)}$: 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 $z \in \mathbb{F}$.
- $f : \mathcal{P}(\mathbb{F}) \to \mathcal{L}(V)$ 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) $(pq)(T) = p(T)q(T)$.

  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 $\operatorname{len}(v, Tv, T^2v, \dots, T^nv) > \dim V$, 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

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

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 $a_0 = \cdots = a_n = 0$, 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

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

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

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

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

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

  - 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).

[1]Note 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 $v_1, \ldots, v_n$ 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]

9/22: • Conditions on T that imply that $\mathcal{M}(T)$ is upper triangular.

Theorem 5.6. *Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Then the following are equivalent.*

(*a*) *The matrix of T with respect to $v_1, \ldots, v_n$ is upper triangular.*

(*b*) $Tv_j \in \operatorname{span}(v_1, \ldots, v_j)$ *for each* $j = 1, \ldots, n$.

(*c*) $\operatorname{span}(v_1, \ldots, v_j)$ *is invariant under T for each* $j = 1, \ldots, n$.

Proof. We will first prove that (a) implies (b); the proof in the reverse direction will be symmetric. We will then prove that (b) implies (c), and wrap up by quickly showing that (c) implies (b), and. Let's begin.

Suppose first that the matrix of T with respect to $v_1, \ldots, v_n$ is upper triangular. Let $j \in \{1, \ldots, n\}$ be arbitrary. By the definition of $\mathcal{M}(T)$, $Tv_j = A_{1,j}v_1 + \cdots + A_{n,j}v_n$. Additionally, since $\mathcal{M}(T)$ is upper triangular, we know that $A_{j+1,j} = \cdots = A_{n,j} = 0$. Therefore, $Tv_j = A_{1,j}v_1 + \cdots + A_{j,j}v_j \in \operatorname{span}(v_1, \ldots, v_j)$, as desired. The proof is symmetric in the other direction.

Now suppose that $Tv_j \in \operatorname{span}(v_1, \ldots, v_j)$ for each $j = 1, \ldots, n$. Fix $j \in \{1, \ldots, n\}$. Let $v \in \operatorname{span}(v_1, \ldots, v_j)$. It follows since $Tv_i \in \operatorname{span}(v_1, \ldots, v_i) \subset \operatorname{span}(v_1, \ldots, v_j)$ for all $i = 1, \ldots, j$ by hypothesis that

$$\begin{aligned}Tv &= T(a_1v_1 + \cdots + a_jv_j)\\ &= a_1Tv_1 + \cdots + a_jTv_j\\ &\in \operatorname{span}(v_1, \ldots, v_j)\end{aligned}$$

as desired. On the other hand, suppose $\operatorname{span}(v_1, \ldots, v_j)$ is invariant under T for each $j = 1, \ldots, n$. Fix $j \in \{1, \ldots, n\}$. Then by the definition of invariance, $v_j \in \operatorname{span}(v_1, \ldots, v_j)$ implies that $Tv_j \in \operatorname{span}(v_1, \ldots, v_j)$, as desired. ■

- We now prove that over $\mathbb{C}$ (notably not over $\mathbb{R}$), every operator has an upper-triangular matrix.

Theorem 5.7. *Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V.*[3]

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

[3]Note that Axler (2015) gives two proofs, but only one (the shorter one using more advanced topics) is transcribed here.

Proof. We induct on $n = \dim V$. For the base case $n = 1$, the desired result holds since any one-dimensional matrix is upper-triangular by definition. Now suppose inductively that we have proven the result for $n - 1$; we wish to prove it for n. By Theorem 5.5, we may let v_1 be an eigenvector of T. Thus, if we let $U = \operatorname{span}(v_1)$, we know that U is an invariant subspace of T with $\dim U = 1$.

Now consider V/U. It follows from the above by Theroem 3.24 that $\dim V/U = n - 1$. Consequently, we may apply the inductive hypothesis to learn that $T/U \in \mathcal{L}(V/U)$ has an upper-triangular matrix with respect to some basis $v_2 + U, \dots, v_n + U$ of V/U. It follows by 5.6 that

$$(T/U)(v_j + U) \in \operatorname{span}(v_2 + U, \dots, v_j + U)$$

for each $j = 2, \dots, n$. Thus, since $Tv_1 = \lambda v_1 \in \operatorname{span}(v_1)$ and since the above implies that $Tv_j \in \operatorname{span}(v_2, \dots, v_j) \subset \operatorname{span}(v_1, \dots, v_j)$ for all $j = 2, \dots, n$, we have that $Tv_j \in \operatorname{span}(v_1, \dots, v_j)$ for all $j = 1, \dots, n$. Therefore, by Theorem 5.6, we have that the matrix of T with respect to $v_1, \dots, v_n$ is upper triangular, as desired. ■

- We now prove an easy method for determining invertibility from an upper-triangular matrix.

Theorem 5.8. *Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then T is invertible if and only if all the entries on the diagonal of that upper-triangular matrix are nonzero.*

Proof. Let $v_1, \dots, v_n$ be the basis of V with respect to which $\mathcal{M}(T)$ is upper triangular, and let $\lambda_1, \dots, \lambda_n$ be the diagonal entries of said matrix.

Suppose first that T is invertible, and suppose for the sake of contradiction that $\lambda_j = 0$ for some $j = 1, \dots, n$. Then by the definition of $\mathcal{M}(T)$, we have that T maps $\operatorname{span}(v_1, \dots, v_j)$ into $\operatorname{span}(v_1, \dots, v_{j-1})$. Thus, since $\dim \operatorname{span}(v_1, \dots, v_j) = j$ and $\dim \operatorname{span}(v_1, \dots, v_{j-1}) = j - 1$, Theorem 3.7 implies that the restriction of T to $\operatorname{span}(v_1, \dots, v_j)$ is not injective. Consequently, by Theorem 3.4, there exists a nonzero $v \in \operatorname{span}(v_1, \dots, v_j)$ such that $Tv = 0$. But this implies that the original T is not injective, i.e., not invertible (by Theorem 3.12), contradicting our supposition that T is invertible, as desired.

Now suppose that $\lambda_j \neq 0$ for all $j = 1, \dots, n$. To prove that T is invertible, Theorem 3.18 tells us that it will suffice to show that T is surjective. To verify that $\operatorname{range} T = V$, we will demonstrate that $v_1, \dots, v_n \in \operatorname{range} T$, from which it will follow by Theorem 3.5 that $\operatorname{range} T = V$. Let's begin. Since $\mathcal{M}(T)$ is upper triangular, we have that

$$\begin{aligned} Tv_1 &= \lambda_1 v_1 + 0v_2 + \cdots + 0v_n \\ v_1 &= T(v_1/\lambda_1) \in \operatorname{range} T \end{aligned}$$

Similarly, we have that

$$\begin{aligned} Tv_2 &= av_1 + \lambda_2 v_2 + 0v_3 + \cdots + 0v_n \\ v_2 &= T(v_2/\lambda_2) - \frac{a}{\lambda_2} v_1 \in \operatorname{range} T \end{aligned}$$

since $T(v_2/\lambda_2)$ clearly and linear combinations of v_1 are elements of $\operatorname{range} T$ by the previous result. We may analogously prove that $v_3, \dots, v_n \in \operatorname{range} T$, as desired. ■

- We cannot exactly compute the eigenvalues of a linear operator from its matrix, but we can approximate them with powerful numerical methods.
- Lastly, we will show that we can use upper-triangular matrices to determine the eigenvalues of the linear operator it represents.

Theorem 5.9. *Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that upper-triangular matrix.*

Proof. Let $v_1, \dots, v_n$ be the basis of V with respect to which $\mathcal{M}(T)$ is upper triangular, and let $\lambda_1, \dots, \lambda_n$ be the diagonal entries of said matrix.

Let $\lambda \in \mathbb{F}$. Then $\mathcal{M}(T - \lambda I)$ has diagonal entries $\lambda_1 - \lambda, \dots, \lambda_n - \lambda$. Thus, by Theorem 5.8, $T - \lambda I$ is not invertible iff $\lambda = \lambda_j$ for some $j = 1, \dots, n$. But by Theorem 5.1, it follows that λ is an eigenvalue of T iff $T - \lambda I$ is not invertible iff $\lambda = \lambda_j$ for some $j = 1, \dots, n$, as desired. ∎

5.C Eigenspaces and Diagonal Matrices

- **Diagonal matrix**: A square matrix that is 0 everywhere except possibly along the diagonal.
  - Naturally, every diagonal matrix is upper triangular.
- **Eigenspace** (of $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$): The set of all eigenvectors of T corresponding to λ, along with the zero vector; in other words, the vector space $\operatorname{null}(T - \lambda I)$. *Denoted by* $\boldsymbol{E(\lambda, T)}$.
- We now show that the sum of the eigenspaces is a direct sum.

Theorem 5.10. *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose also that $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues of T. Then $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum. Furthermore,* $\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim V$.

Proof. To prove that $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum, suppose that $u_1 + \cdots + u_m = 0$, where $u_j \in E(\lambda_j, T)$ for all $j = 1, \dots, m$. Naturally, each u_j is an eigenvector corresponding to λ_j or zero. But since Theorem 5.2 implies that $u_1, \dots, u_j$ is linearly independent if any u_j is nonzero, we must have $u_1 = \cdots = u_m = 0$.

Additionally, we have that

$$\begin{aligned}\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) &= \dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)) \\ &\leq \dim V\end{aligned}$$

as desired. ∎

- **Diagonalizable** ($T \in \mathcal{L}(V)$): An operator $T \in \mathcal{L}(V)$ such that the operator has a diagonal matrix with respect to some basis of V.
- We now give some conditions equivalent to diagonalizability.

Theorem 5.11. *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent.*

(a) *T is diagonalizable.*

(b) *V has a basis consisting of eigenvectors of T.*

(c) *There exist one-dimensional subspaces $U_1, \dots, U_n$ of V, each invariant under T, such that $V = U_1 \oplus \cdots \oplus U_n$.*

(d) *$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.*

(e) *$\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.*

Proof. We will first briefly justify the equivalence of (a) and (b). We will then show that (b) implies (c) and vice versa. Lastly, we will show that (b) implies (d), (d) implies (e), and (e) implies (b). Let's begin.

Suppose that T is diagonalizable. Then there exists a basis $v_1, \dots, v_n$ of V such that the matrix of T with respect to this basis is diagonal. Let $\lambda_1, \dots, \lambda_n$ be the diagonal entries of this matrix. Then by the definition of $\mathcal{M}(T)$, $Tv_j = \lambda_j v_j$ for all $j = 1, \dots, n$. It follows that each v_j in the basis $v_1, \dots, v_n$ of V is an eigenvector of T, as desired. The proof is symmetric in the reverse direction.

Suppose that $v_1, \dots, v_n$ is a basis of V consisting of eigenvectors of T. Let $U_j = \operatorname{span}(v_j)$ for each $j = 1, \dots, n$. Then each U_j is a one-dimensional subspace of V. Additionally, for any $av_j \in U_j$, we have that $T(av_j) = aTv_j = a\lambda_j v_j \in U_j$, proving that each U_j is invariant under T. Finally, to show that $V = U_1 \oplus \cdots \oplus U_n$, it will suffice to show that each $v \in V$ can be written uniquely as a sum $u_1 + \cdots + u_n$, where $u_j \in U_j$ for each $j = 1, \dots, n$. Let $v \in V$ be arbitrary. Since $v_1, \dots, v_n$ is a basis of V, Theorem 2.5 v can be written uniquely in the form $a_1v_1 + \cdots + a_nv_n$. Let $u_j = a_jv_j$ for each $j = 1, \dots, n$. Then v can be written uniquely as a sum $u_1 + \cdots + u_n$, where $u_j \in U_j$ for each $j = 1, \dots, n$, as desired. The proof in the reverse direction is quite similar.

Suppose that $v_1, \dots, v_n$ is a basis of V consisting of eigenvectors of T. Thus, since each $v \in V$ is a linear combination of eigenvectors of T, $V = E(\lambda_1, T) + \cdots + E(\lambda_m, T)$. Therefore, Theorem 5.10 implies that $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, as desired.

Suppose $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$. Then naturally $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$, as desired.

Suppose $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$. Choose a basis of each $E(\lambda_j, T)$. Concatenate all these bases to form a list $v_1, \dots, v_n$ of eigenvectors of T. We now wish to show that $v_1, \dots, v_n$ is a basis of V. To do so, already knowing that $n = \dim V$ by hypothesis, Theorem 2.12 tells us that we need only additionally show that $v_1, \dots, v_n$ is linearly independent. Let $a_1, \dots, a_n \in \mathbb{F}$ be such that $a_1v_1 + \cdots + a_nv_n = 0$. For each $j = 1, \dots, m$, let u_j denote the sum of all the terms a_kv_k such that $v_k \in E(\lambda_j, T)$. Thus, each $u_j \in E(\lambda_j, T)$ and $u_1 + \cdots + u_m = 0$. But since $u_1, \dots, u_m$ are eigenvectors of T corresponding to distinct eigenvalues, Theorem 5.2 implies that $u_1, \dots, u_m$ are linearly independent. This combined with the previous result implies that $u_j = 0$ for each $j = 1, \dots, m$. It follows since each (zero) u_j is a linear combination of a basis of an eigenspace, each $a_j = 0$, as desired. ■

- Note that not every operator is diagonalizable.
- We now prove that if an operator has sufficiently many eigenvalues, it is diagonalizable.

  Theorem 5.12. *If $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues, then T is diagonalizable.*

  Proof. Let $m = \dim V$, and let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T. Let $v_1, \dots, v_m$ be the corresponding eigenvectors. Since $v_1, \dots, v_m$ has length equal to the dimension of V and is linearly independent (by Theorem 5.2), Theorem 2.12 implies that $v_1, \dots, v_m$ is a basis of V. Therefore, by Theorem 5.11, T is diagonalizable. ■

- Note that the converse of Theorem 5.12 is not true.