Eigensystem of a linear operator

Branko Ćurgus

February 9, 2020 at 23:01

1 Algebra of linear operators

In this section we consider a vector space $\mathscr{V}$ over a scalar field $\mathbb{F}$. By $\mathscr{L}(\mathscr{V})$ we denote the vector space $\mathscr{L}(\mathscr{V}, \mathscr{V})$ of all linear operators on $\mathscr{V}$. The vector space $\mathscr{L}(\mathscr{V})$ with the composition of operators as an additional binary operation is an algebra in the sense of the following definition.

Definition 1.1. A vector space $\mathscr{A}$ over a field $\mathbb{F}$ is an *algebra* over $\mathbb{F}$ if the following conditions are satisfied:

(a) there exist a binary operation $\cdot : \mathscr{A} \times \mathscr{A} \to \mathscr{A}$.

(b) (*associativity*) for all $x, y, z \in \mathscr{A}$ we have $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

(c) (*right-distributivity*) for all $x, y, z \in \mathscr{A}$ we have $(x+y) \cdot z = x \cdot z + y \cdot z$.

(d) (*left-distributivity*) for all $x, y, z \in \mathscr{A}$ we have $z \cdot (x+y) = z \cdot x + z \cdot y$.

(e) (*respect for scaling*) for all $x, y \in \mathscr{A}$ and all $\alpha \in \mathbb{F}$ we have $\alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y)$.

The binary operation in an algebra is often referred to as *multiplication*.

The multiplicative identity in the algebra $\mathscr{L}(\mathscr{V})$ is the identity operator $I_{\mathscr{V}}$.

For $T \in \mathscr{L}(\mathscr{V})$ we recursively define nonnegative integer powers of T by $T^0 = I_{\mathscr{V}}$ and for all $n \in \mathbb{N}$ $T^n = T \circ T^{n-1}$.

For $T \in \mathscr{L}(\mathscr{V})$, set

$$\mathscr{A}_T = \operatorname{span}\bigl\{ T^k : k \in \mathbb{N} \cup \{0\} \bigr\}.$$

Clearly $\mathscr{A}_T$ is a subspace of $\mathscr{L}(\mathscr{V})$. Moreover, we will see below that $\mathscr{A}_T$ is a commutative subalgebra of $\mathscr{L}(\mathscr{V})$.

Recall that by definition of a span a nonzero $S \in \mathscr{L}(\mathscr{V})$ belongs to $\mathscr{A}_T$ if and only if $\exists m \in \mathbb{N} \cup \{0\}$ and $\alpha_0, \alpha_1, \ldots, \alpha_m \in \mathbb{F}$ such that $a_m \neq 0$ and

$$S = \sum_{k=0}^{m} \alpha_k T^k. \tag{1}$$

The last expression reminds us of a polynomial over $\mathbb{F}$. Recall that by $\mathbb{F}[z]$ we denote the algebra of all polynomials over $\mathbb{F}$. That is

$$\mathbb{F}[z] = \left\{ \sum_{j=0}^{n} \alpha_j z^j \ : \ n \in \mathbb{N} \cup \{0\}, \ (\alpha_0, \ldots, \alpha_n) \in \mathbb{F}^{n+1} \right\}.$$

Next we recall the definition of the multiplication in the algebra $\mathbb{F}[z]$. Let $m, n \in \mathbb{N} \cup \{0\}$ and

$$p(z) = \sum_{i=0}^{m} \alpha_i z^i \in \mathbb{F}[z] \qquad \text{and} \qquad q(z) = \sum_{j=0}^{n} \beta_j z^j \in \mathbb{F}[z]. \tag{2}$$

Then by definition

$$(pq)(z) = \sum_{k=0}^{m+n} \left(\sum_{\substack{i+j=k \\ i \in \{0,\ldots,m\} \\ j \in \{0,\ldots,n\}}} \alpha_i \beta_j \right) z^k.$$

Since the multiplication in $\mathbb{F}$ is commutative, it follows that $pq = qp$. That is $\mathbb{F}[z]$ is a commutative algebra.

The obvious alikeness of the expression (1) and the expression for the polynomial p in (2) is the motivation for the following definition. For a fixed $T \in \mathscr{L}(\mathscr{V})$ we define

$$\Xi_T : \mathbb{F}[z] \to \mathscr{L}(\mathscr{V})$$

by setting

$$\Xi_T(p) = \sum_{i=0}^{m} \alpha_i T^i \qquad \text{where} \qquad p(z) = \sum_{i=0}^{m} \alpha_i z^i \in \mathbb{F}[z]. \tag{3}$$

It is common to write $p(T)$ for $\Xi_T(p)$.

Theorem 1.2. *Let $T \in \mathscr{L}(\mathscr{V})$. The function $\Xi_T : \mathbb{F}[z] \to \mathscr{L}(\mathscr{V})$ defined in (3) is an algebra homomorphism. The range of Ξ_T is $\mathscr{A}_T$.*

Proof. It is not difficult to prove that $\Xi_T : \mathbb{F}[z] \to \mathscr{L}(\mathscr{V})$ is linear. We will prove that $\Xi_T : \mathbb{F}[z] \to \mathscr{L}(\mathscr{V})$ is multiplicative, that is, for all $p, q \in \mathbb{F}[z]$ we have $\Xi_T(pq) = \Xi_T(p)\Xi_T(q)$. To prove this let $p, q \in \mathbb{F}[z]$ be arbitrary and given in (2). Then

$$\begin{aligned}
\Xi_T(p)\Xi_T(q) &= \left(\sum_{i=0}^{m} \alpha_i T^i\right)\left(\sum_{j=0}^{n} \beta_j T^j\right) && \text{(by definition in (3))}\\
&= \sum_{i=0}^{m}\sum_{j=0}^{n} \alpha_i\beta_j T^{i+j} && \text{(since } \mathscr{L}(\mathscr{V}) \text{ is an algebra)}\\
&= \sum_{k=0}^{m+n}\left(\sum_{\substack{i+j=k\\ i\in\{0,\ldots,m\}\\ j\in\{0,\ldots,n\}}} \alpha_i\beta_j\right)T^k && \text{(since } \mathscr{L}(\mathscr{V}) \text{ is a vector space)}\\
&= \Xi_T(pq) && \text{(by definition in (3)).}
\end{aligned}$$

This proves the multiplicative property of Ξ_T.

The fact that $\mathscr{A}_T$ is the range of Ξ_T is obvious. □

Corollary 1.3. *Let $T \in \mathscr{L}(\mathscr{V})$. The subspace $\mathscr{A}_T$ of $\mathscr{L}(\mathscr{V})$ is a commutative subalgebra of $\mathscr{L}(\mathscr{V})$.*

Proof. Let $Q, S \in \mathscr{A}_T$. Since $\mathscr{A}_T$ is the range of Ξ_T there exist $p, q \in \mathbb{F}[z]$ such that $Q = \Xi_T(p)$ and $S = \Xi_T(q)$. Then, since Ξ_T is an algebra homomorphism we have

$$QS = \Xi_T(p)\Xi_T(p) = \Xi_T(pq) = \Xi_T(qp) = \Xi_T(q)\Xi_T(p) = SQ.$$

This sequence of equalities shows that $QS \in \operatorname{ran}(\Xi_T) = \mathscr{A}_T$ and $QS = SQ$. That is $\mathscr{A}_T$ is closed with respect to the operator composition and the operator composition on $\mathscr{A}_T$ is commutative. □

Corollary 1.4. *Let $\mathscr{V}$ be a complex vector space and let $T \in \mathscr{L}(\mathscr{V})$ be a nonzero operator. Then for every $p \in \mathbb{C}[z]$ such that $m = \deg p \geq 1$ there exist a nonzero $\alpha \in \mathbb{C}$ and $z_1, \ldots, z_m \in \mathbb{C}$ such that*

$$\Xi_T(p) = p(T) = \alpha(T - z_1 I) \cdots (T - z_m I).$$

Proof. Let $p \in \mathbb{C}[z]$ such that $m = \deg p \geq 1$. Then there exist $\alpha_0, \ldots, \alpha_m \in \mathbb{C}$ such that $\alpha_m \neq 0$ such that

$$p(z) = \sum_{k=0}^{m} \alpha_j z^j.$$

By the Fundamental Theorem of Algebra there exist nonzero $\alpha \in \mathbb{C}$ and $z_1, \ldots, z_m \in \mathbb{C}$ such that

$$p(z) = \alpha(z - z_1) \cdots (z - z_m).$$

Here $\alpha = \alpha_m$ and $z_1, \ldots, z_m$ are the roots of p. Since Ξ_T is an algebra homomorphism we have

$$p(T) = \Xi_T(p) = \alpha\, \Xi_T(z - z_1) \;\cdots\; \Xi_T(z - z_m) = \alpha(T - z_1 I) \;\cdots\; (T - z_m I).$$

This completes the proof. □

2 Existence of an eigenvalue

Lemma 2.1. *Let $n \in \mathbb{N}$ and $S_1, \ldots, S_n \in \mathscr{L}(\mathscr{V})$. If $S_1, \ldots, S_n$ are all injections, then the composition $S_1 \cdots S_n$ is an injection.*

Proof. We proceed by Mathematical Induction. The base step is trivial. It is useful to prove the implication for $n = 2$. Assume that $S, T \in \mathscr{L}(\mathscr{V})$ are injective and let $u, v \in \mathscr{V}$ be such that $u \neq v$. Then, since T is injective, $Tu \neq Tv$. Since S is injective, $S(Tu) \neq S(Tv)$. Thus, ST is injective.

Next we prove the inductive step. Let $m \in \mathbb{N}$ and assume that $S_1 \cdots S_m$ is injective whenever $S_1, \ldots, S_m \in \mathscr{L}(\mathscr{V})$ are all injective. (This is the inductive hypothesis.) Now assume that $S_1, \ldots, S_m, S_{m+1} \in \mathscr{L}(\mathscr{V})$ are all injective. By the inductive hypothesis the operator $S = S_1 \cdots S_m$ is injective. Since by assumption $T = S_{m+1}$ is injective, the already proved claim for $n = 2$ yields that

$$ST = S_1 \;\cdots\; S_m\, S_{m+1}$$

is injective. This completes the proof. □

Theorem 2.2. *Let $\mathscr{V}$ be a nontrivial finite dimensional vector space over $\mathbb{C}$. Let $T \in \mathscr{L}(\mathscr{V})$. Then there exists a $\lambda \in \mathbb{C}$ and $v \in \mathscr{V}$ such that $v \neq 0_{\mathscr{V}}$ and $Tv = \lambda v$.*

Proof. The claim of the theorem is trivial if $T = 0_{\mathscr{L}(\mathscr{V})}$. So, assume that $T \in \mathscr{L}(\mathscr{V})$ is a nonzero operator.

Let $n = \dim \mathscr{V}$ and let $u \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$. Now consider the vectors

$$u, Tu, T^2u, \ldots, T^n u. \tag{4}$$

If two of these vectors coincide, say $k, l \in \{0, \ldots, n\}$, $k < l$ are such that $T^k u = T^l u$, setting $\alpha_j = 0$ for $j \in \{0, \ldots, n\} \setminus \{k, l\}$ and $\alpha_k = 1$ and $\alpha_l = -1$ we obtain a nontrivial linear combination of the vectors in (4).

If the vectors in (4) are distinct, since $n = \dim \mathscr{V}$, it follows from the Steinitz Exchange Lemma that the vectors in (4) are linearly dependent.

Hence, in either case, there exist $\alpha_0, \dots, \alpha_n \in \mathbb{C}$ and $k \in \{0, \dots, n\}$ such that

$$\alpha_0 u + \alpha_1 T u + \alpha_2 T^2 u + \cdots + \alpha_n T^n u = 0_{\mathscr{V}} \tag{5}$$

and $\alpha_k \neq 0$. Since $u \neq 0_{\mathscr{V}}$ it is not possible that $\alpha_j = 0$ for all $j \in \{1, \dots, n\}$. Therefore, there exists $k \in \{1, \dots, n\}$ such that $\alpha_k \neq 0$.

Set

$$p(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \cdots + \alpha_n z^n.$$

Since there exists $k \in \{1, \dots, n\}$ such that $\alpha_k \neq 0$, we have that $m = \deg p > 0$. By the Fundamental Theorem of Algebra there exist $\alpha \neq 0$ and $z_1, \dots, z_m \in \mathbb{C}$ such that

$$p(z) = \alpha (z - z_1) \cdots (z - z_m).$$

Here $\alpha = \alpha_m$ and $z_1, \dots, z_m$ are the roots of p.

Since Ξ_T is an algebra homomorphism we have

$$p(T) = \Xi_T(p) = \alpha\, \Xi_T(z - z_1) \ \cdots\ \Xi_T(z - z_m) = \alpha (T - z_1 I)\ \cdots\ (T - z_m I).$$

Equality (5) yields that the operator $p(T)$ is not an injection. Lemma 2.1 now implies that there exists $j \in \{1, \dots, m\}$ such that $T - z_j I$ is not injective. That is, there exists $v \in \mathscr{V}$, $v \neq 0_{\mathscr{V}}$ such that

$$(T - z_j I) v = 0.$$

Setting $\lambda = z_j$ completes the proof. □

Definition 2.3. Let $\mathscr{V}$ be a vector space over $\mathbb{F}$, $T \in \mathscr{L}(\mathscr{V})$. A scalar $\lambda \in \mathbb{F}$ is an *eigenvalue* of T if there exists $v \in \mathscr{V}$ such that $v \neq 0$ and $Tv = \lambda v$. The subspace $\operatorname{nul}(T - \lambda I)$ of $\mathscr{V}$ is called the *eigenspace* of T corresponding to λ.

Definition 2.4. Let $\mathscr{V}$ be a finite dimensional vector space over $\mathbb{F}$. Let $T \in \mathscr{L}(\mathscr{V})$. The set of all eigenvalues of T is denoted by $\sigma(T)$. It is called the *spectrum* of T.

The above proof of the existence of an eigenvalue can be adopted so that we prove more. Below we prove: $T^n = 0_{\mathscr{L}(\mathscr{V})}$ if and only if $\sigma(T) = \{0\}$.

More information proof. Let $n = \dim \mathscr{V}$ and let $u \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$. Then $n \in \mathbb{N}$. First assume that $T^n = 0_{\mathscr{L}(\mathscr{V})}$. The set

$$\mathbb{K} = \bigl\{ j \in \{0, \ldots, n\} \,|\, T^j u \neq 0 \bigr\}$$

is nonempty since $0 \in \mathbb{K}$. Set $k = \min \mathbb{K}$. Then $k \in \{0, \ldots, n-1\}$, $T^k u \neq 0$ and $T\bigl(T^k u\bigr) = 0$. Thus, we can take $\lambda = 0$ and $v = T^k u$ in this case.

Now assume that $T^n \neq 0_{\mathscr{L}(\mathscr{V})}$. Then there exists $w \in \mathscr{V}$ such that $T^n w \neq 0_{\mathscr{V}}$. Consequently, none of the vectors

$$w, Tw, T^2 w, \ldots, T^n w \tag{6}$$

equals to $0_{\mathscr{V}}$.

If two of the vectors in (6) coincide, say $k, l \in \{0, \ldots, n\}$ are such that $k < l$ and $T^k u = T^l u$, setting $\alpha_j = 0$ for $j \in \{0, \ldots, n\} \setminus \{k, l\}$ and $\alpha_k = 1$ and $\alpha_l = -1$ we obtain a nontrivial linear combination of the vectors in (6).

If the vectors in (6) are distinct then we have $n+1$ vector in a vector space of the dimension $n = \dim \mathscr{V}$. The Steinitz Exchange Lemma implies that the vectors in (6) are linearly dependent.

Hence, in either case, there exist $\alpha_0, \ldots, \alpha_n \in \mathbb{C}$ such that

$$\alpha_0 u + \alpha_1 T u + \alpha_2 T^2 u + \cdots + \alpha_n T^n u = 0_{\mathscr{V}} \tag{7}$$

and at least one among $\alpha_0, \ldots, \alpha_n$ is nonzero.

Set

$$l = \min\bigl\{ i \in \{0, \ldots, n\} \,|\, \alpha_i \neq 0 \bigr\}, \qquad m = \max\bigl\{ i \in \{0, \ldots, n\} \,|\, \alpha_i \neq 0 \bigr\}.$$

Since all the vectors in (6) are nonzero we have $l < m$. Further set $k = m - l$ and

$$p(z) = \alpha_l + \alpha_{l+1} z + \cdots + \alpha_m z^k.$$

Then $k = \deg p > 0$ and (7) reads $p(T) T^l w = 0_{\mathscr{V}}$ with $T^l w \neq 0_{\mathscr{V}}$.

By the Fundamental Theorem of Algebra there exist $z_1, \ldots, z_k \in \mathbb{C}$ such that

$$p(z) = \alpha_m (z - z_1) \cdots (z - z_k).$$

Here $z_1, \ldots, z_k$ are the roots of p. Since $\alpha_l \neq 0$, that is the constant coefficient of p is nonzero, we have that $z_i \neq 0$ for all $i \in \{1, \ldots, k\}$.

Since Ξ_T is an algebra homomorphism we have

$$p(T) = \Xi_T(p) = \alpha_l \, \Xi_T(z - z_1) \cdots \Xi_T(z - z_k) = \alpha_l (T - z_1 I) \cdots (T - z_k I).$$

Since $p(T)T^l w = 0_{\mathscr{V}}$ and $T^l w \neq 0_{\mathscr{V}}$ the operator $p(T)$ is not an injection. Lemma 2.1 now implies that there exists $j \in \{1, \dots, k\}$ such that $T - z_j I$ is not injective. That is, there exists $v \in \mathscr{V}$, $v \neq 0_{\mathscr{V}}$ such that

$$(T - z_j I)v = 0.$$

As all the roots of p are nonzero, setting $\lambda = z_j \neq 0$ completes the proof. □

The next theorem can be stated in English simply as: Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Theorem 2.5. *Let $\mathscr{V}$ be a vector space over $\mathbb{F}$, $T \in \mathscr{L}(\mathscr{V})$ and $n \in \mathbb{N}$. Assume*

(a) *$\lambda_1, \dots, \lambda_n \in \mathbb{F}$ are such that $\lambda_i \neq \lambda_j$ for all $i, j \in \{1, \dots, n\}$ such that $i \neq j$,*

(b) *$v_1, \dots, v_n \in \mathscr{V}$ are such that $Tv_k = \lambda_k v_k$ and $v_k \neq 0$ for all $k \in \{1, \dots, n\}$.*

Then $\{v_1, \dots, v_n\}$ is linearly independent.

Proof. We will prove this by using the mathematical induction on n. For the base case, we will prove the claim for $n = 1$. Let $\lambda_1 \in \mathbb{F}$ and let $v_1 \in \mathscr{V}$ be such that $v_1 \neq 0$ and $Tv_1 = \lambda_1 v_1$. Since $v_1 \neq 0$, we conclude that $\{v_1\}$ is linearly independent.

Next we prove the inductive step. Let $m \in \mathbb{N}$ be arbitrary. The inductive hypothesis is the assumption that the following implication holds.

> If the following two conditions are satisfied:
>
> (i) $\mu_1, \dots, \mu_m \in \mathbb{F}$ are such that $\mu_i \neq \mu_j$ for all $i, j \in \{1, \dots, m\}$ such that $i \neq j$,
>
> (ii) $w_1, \dots, w_m \in \mathscr{V}$ are such that $Tw_k = \mu_k w_k$ and $w_k \neq 0$ for all $k \in \{1, \dots, m\}$,
>
> then $\{w_1, \dots, w_m\}$ is linearly independent.

We need to prove the following implication

> If the following two conditions are satisfied:
>
> (I) $\lambda_1, \dots, \lambda_{m+1} \in \mathbb{F}$ are such that $\lambda_i \neq \lambda_j$ for all $i, j \in \{1, \dots, m+1\}$ such that $i \neq j$,
>
> (II) $v_1, \dots, v_{m+1} \in \mathscr{V}$ are such that $Tv_k = \lambda_k v_k$ and $v_k \neq 0$ for all $k \in \{1, \dots, m+1\}$,
>
> then $\{v_1, \dots, v_{m+1}\}$ is linearly independent.

Assume (I) and (II) in the red box. We need to prove that $\{v_1, \dots, v_{m+1}\}$ is linearly independent.

Let $\alpha_1, \dots, \alpha_{m+1} \in \mathbb{F}$ be such that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m + \alpha_{m+1} v_{m+1} = 0. \tag{8}$$

Applying $T \in \mathscr{L}(\mathscr{V})$ to both sides of (8), using the linearity of T and assumption (II) we get

$$\alpha_1 \lambda_1 v_1 + \cdots + \alpha_m \lambda_m v_m + \alpha_{m+1} \lambda_{m+1} v_{m+1} = 0. \tag{9}$$

Multiplying both sides of (8) by λ_{m+1} we get

$$\alpha_1 \lambda_{m+1} v_1 + \cdots + \alpha_m \lambda_{m+1} v_m + \alpha_{m+1} \lambda_{m+1} v_{m+1} = 0. \tag{10}$$

Subtracting (10) from (9) we get

$$\alpha_1(\lambda_1 - \lambda_{m+1}) v_1 + \cdots + \alpha_m(\lambda_m - \lambda_{m+1}) v_m = 0.$$

Since by assumption (I) we have $\lambda_j - \lambda_{m+1} \neq 0$ for all $j \in \{1, \dots, m\}$, setting

$$w_j = (\lambda_j - \lambda_{m+1}) v_j, \qquad j \in \{1, \dots, m\},$$

and taking into account (II) we have

$$w_j \neq 0 \qquad \text{and} \qquad T w_j = \lambda_j w_j \qquad \text{for all} \qquad j \in \{1, \dots, m\}. \tag{11}$$

Thus, by (I) and (11), the scalars $\lambda_1, \dots, \lambda_m$ and vectors $w_1, \dots, w_m$ satisfy assumptions (i) and (ii) of the inductive hypothesis (the green box). Consequently, the vectors $w_1, \dots, w_m$ are linearly independent. Since by (11) we have

$$\alpha_1 w_1 + \cdots + \alpha_m w_m = 0,$$

it follows that $\alpha_1 = \cdots = \alpha_m = 0$. Substituting these values in (8) we get $\alpha_{m+1} v_{m+1} = 0$. Since by (II), $v_{m+1} \neq 0$ we conclude that $\alpha_{m+1} = 0$. This completes the proof of the linear independence of $v_1, \dots, v_{m+1}$. □

A different proof follows.

Proof. Consider operators $T - \lambda_j I$ for $j \in \{1, \dots, n\}$. Then

$$(T - \lambda_j I) v_k = \bigl(\lambda_k - \lambda_j\bigr) v_k, \quad j, k \in \{1, \dots, n\}.$$

Or, more precisely,

$$(T - \lambda_j I)v_k = \begin{cases} (\lambda_k - \lambda_j)v_k & j \neq k, \\ 0_{\mathscr{V}} & j = k \end{cases} \quad j, k \in \{1, \ldots, n\}. \tag{12}$$

Let $i, k \in \{1, \ldots, n\}$. Repeated application of (12) yields

$$\left(\prod_{j=1,\, j\neq i}^{n} (T - \lambda_j I) \right) v_k = \left(\prod_{j=1,\, j\neq i}^{n} (\lambda_k - \lambda_j) \right) v_k,$$

or, more precisely,

$$\left(\prod_{j=1,\, j\neq i}^{n} (T - \lambda_j I) \right) v_k = \begin{cases} 0_{\mathscr{V}} & k \neq i, \\ \left(\prod_{j=1,\, j\neq k}^{n} (\lambda_k - \lambda_j) \right) v_k & k = i \end{cases}. \tag{13}$$

Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ be such that

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_{\mathscr{V}}. \tag{14}$$

Let $k \in \{1, \ldots, n\}$ be arbitrary and apply the operator

$$\prod_{j=1,\, j\neq k}^{n} (T - \lambda_j I)$$

to both sides of (14). Then by (13) we get

$$\alpha_k \left(\prod_{j=1,\, j\neq k}^{n} (\lambda_k - \lambda_j) \right) v_k = 0_{\mathscr{V}}. \tag{15}$$

Since $\lambda_1, \ldots, \lambda_n$ are distinct we have

$$\prod_{j=1,\, j\neq k}^{n} (\lambda_k - \lambda_j) \neq 0,$$

and since also $v_k \neq 0_{\mathscr{V}}$, from (15) we deduce $\alpha_k = 0$. Since $k \in \{1, \ldots, n\}$ was arbitrary, the theorem is proved. □

Corollary 2.6. *Let $\mathscr{V}$ be a finite dimensional vector space over $\mathbb{F}$ and let $T \in \mathscr{L}(\mathscr{V})$. Then T has at most $n = \dim \mathscr{V}$ distinct eigenvalues.*

Proof. Let $\mathscr{B}$ be a basis of $\mathscr{V}$ where $\mathscr{B} = \{u_1, ..., u_n\}$. Then $|\mathscr{B}| = n$ and $\operatorname{span} \mathscr{B} = \mathscr{V}$. Let $\mathscr{C} = \{v_1, ..., v_m\}$ be eigenvectors corresponding to m distinct eigenvalues. Then $\mathscr{C}$ is a linearly independent set with $|\mathscr{C}| = m$. By the Steinitz Exchange Lemma, $m \leq n$. Consequently, T has at most n distinct eigenvalues. □

3 Existence of an upper-triangular matrix representation

Definition 3.1. A matrix $A \in \mathbb{F}^{n\times n}$ with entries a_{ij}, $i, j \in \{1, \ldots, n\}$ is called *upper triangular* if $a_{ij} = 0$ for all $i, j \in \{1, \ldots, n\}$ such that $i > j$.

Definition 3.2. Let $\mathscr{V}$ be a vector space over $\mathbb{F}$ and $T \in \mathscr{L}(\mathscr{V})$. A subspace $\mathscr{U}$ of $\mathscr{V}$ is called an *invariant subspace* under T if $T(\mathscr{U}) \subseteq \mathscr{U}$.

The following proposition is straightforward.

Proposition 3.3. *Let $S, T \in \mathscr{L}(\mathscr{V})$ be such that $ST = TS$. Then $\operatorname{nul} T$ is invariant under S and $\operatorname{nul} S$ is invariant under T. In particular, all eigenspaces of T are invariant under T.*

Definition 3.4. Let $\mathscr{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $n = \dim \mathscr{V} > 0$. Let $T \in \mathscr{L}(\mathscr{V})$. A sequence of nontrivial subspaces $\mathscr{U}_1, \ldots, \mathscr{U}_n$ of $\mathscr{V}$ such that

$$\mathscr{U}_1 \subsetneq \mathscr{U}_2 \subsetneq \cdots \subsetneq \mathscr{U}_n \tag{16}$$

and

$$T\mathscr{U}_k \subseteq \mathscr{U}_k \qquad \text{for all} \qquad k \in \{1, \ldots, n\}$$

is called a *fan* for T in $\mathscr{V}$. A basis $\{v_1, \ldots, v_n\}$ of $\mathscr{V}$ is called a *fan basis* corresponding to T if the subspaces

$$\mathscr{V}_k = \operatorname{span}\{v_1, \ldots, v_k\}, \qquad k \in \{1, \ldots, n\},$$

form a fan for T.

Notice that (16) implies

$$1 \leq \dim \mathscr{U}_1 < \dim \mathscr{U}_2 < \cdots < \dim \mathscr{U}_n \leq n.$$

Consequently, if $\mathscr{U}_1, \ldots, \mathscr{U}_n$ is a fan for T we have $\dim \mathscr{U}_k = k$ for all $k \in \{1, \ldots, n\}$. In particular $\mathscr{U}_n = \mathscr{V}$.

Theorem 3.5 (Theorem 5.12)**.** *Let $\mathscr{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $\dim \mathscr{V} = n$ and let $T \in \mathscr{L}(\mathscr{V})$. Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be a basis of $\mathscr{V}$ and set*

$$\mathscr{V}_k = \operatorname{span}\{v_1, \ldots, v_k\}, \qquad k \in \{1, \ldots, n\}.$$

The following statements are equivalent.

(a) *$M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular.*

(b) *For all $k \in \{1, \ldots, n\}$ we have $Tv_k \in \mathscr{V}_k$.*

(c) *For all $k \in \{1, \ldots, n\}$ we have $T\mathscr{V}_k \subseteq \mathscr{V}_k$.*

(d) *$\mathscr{B}$ is a fan basis corresponding to T.*

Proof. (a) ⇒ (b). Assume that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper triangular. That is

$$M_{\mathscr{B}}^{\mathscr{B}}(T) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{kk} & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Let $k \in \{1, \ldots, n\}$ be arbitrary. Then, by the definition of $M_{\mathscr{B}}^{\mathscr{B}}(T)$,

$$C_{\mathscr{B}}(Tv_k) = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Consequently, by the definition of $C_{\mathscr{B}}$, we have

$$Tv_k = a_{1k}v_1 + \cdots + a_{kk}v_k \in \operatorname{span}\{v_1, \ldots, v_k\} = \mathscr{V}_k.$$

Thus, (b) is proved.

(b) ⇒ (a). Assume that $Tv_k \in \mathscr{V}_k$ for all $k \in \{1, \ldots, n\}$. Let a_{ij}, $i, j \in \{1, \ldots, n\}$, be the entries of $M_{\mathscr{B}}^{\mathscr{B}}(T)$. Let $j \in \{1, \ldots, n\}$ be arbitrary. Since $Tv_j \in \mathscr{V}_j$ there exist $\alpha_1, \ldots, \alpha_j \in \mathbb{F}$ such that

$$Tv_j = \alpha_1 v_1 + \cdots + \alpha_j v_j.$$

By the definition of $C_{\mathscr{B}}$ we have

$$C_{\mathscr{B}}(Tv_j) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

On the other side, by the definition of $M_{\mathscr{B}}^{\mathscr{B}}(T)$, we have

$$C_{\mathscr{B}}(Tv_j) = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{jj} \\ a_{j+1,j} \\ \vdots \\ a_{nj} \end{bmatrix}.$$

The last two equalities, and the fact that $C_{\mathscr{B}}$ is a function, imply $a_{ij} = 0$ for all $i \in \{j+1, \dots, n\}$. This proves (a).

(b) $\Rightarrow$ (c). Suppose $Tv_k \in \mathscr{V}_k = \operatorname{span}\{v_1, \dots, v_k\}$ for all $k \in \{1, \dots, n\}$. Let $v \in \mathscr{V}_k$. Then $v = \alpha_1 v_1 + \cdots + \alpha_k v_k$. Applying T, we get $Tv = \alpha_1 Tv_1 + \cdots + \alpha_k Tv_k$. Thus,

$$Tv \in \operatorname{span}\{Tv_1, \dots, Tv_k\}. \tag{17}$$

Since

$$Tv_j \in \mathscr{V}_j \subseteq \mathscr{V}_k \qquad \text{for all} \qquad j \in \{1, \dots, k\},$$

we have

$$\operatorname{span}\{Tv_1, \dots, Tv_k\} \subseteq \mathscr{V}_k.$$

Together with (17), this proves (c).

(c) $\Rightarrow$ (b). Suppose $T\mathscr{V}_k \subseteq \mathscr{V}_k$ for all $k \in \{1, \dots, n\}$. Then since $v_k \in \mathscr{V}_k$, we have $Tv_k \in \mathscr{V}_k$ for each $k \in \{1, \dots, n\}$.

(c) $\Leftrightarrow$ (d) follows from the definition of a fan basis corresponding to T. $\square$

Theorem 3.6 (Theorem 5.13)**.** *Let $\mathscr{V}$ be a nonzero finite dimensional complex vector space. If $\dim \mathscr{V} = n$ and $T \in \mathscr{L}(\mathscr{V})$, then there exists a basis $\mathscr{B}$ of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular.*

Proof. We proceed by the complete induction on $n = \dim(\mathscr{V})$.

The base case is trivial. Assume $\dim \mathscr{V} = 1$ and $T \in \mathscr{L}(\mathscr{V})$. Set $\mathscr{B} = \{u\}$, where $u \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ is arbitrary. Then there exists $\lambda \in \mathbb{C}$ such that $Tu = \lambda u$. Thus, $M_{\mathscr{B}}^{\mathscr{B}}(T) = \big[\lambda\big]$.

Now we prove the inductive step. Let $m \in \mathbb{N}$ be arbitrary. The inductive hypothesis is

> For every $k \in \{1, \dots, m\}$ the following implication holds: If $\dim \mathscr{U} = k$ and $S \in \mathscr{L}(\mathscr{U})$, then there exists a basis $\mathscr{A}$ of $\mathscr{U}$ such that $M_{\mathscr{A}}^{\mathscr{A}}(S)$ is upper-triangular.

To complete the inductive step, we need to prove the implication:

If $\dim \mathscr{V} = m+1$ and $T \in \mathscr{L}(\mathscr{V})$, then there exists a basis $\mathscr{B}$ of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular.

To prove the red implication assume that $\dim \mathscr{V} = m+1$ and $T \in \mathscr{L}(\mathscr{V})$. By Theorem 2.2 the operator T has an eigenvalue. Let λ be an eigenvalue of T. Set $\mathscr{U} = \operatorname{ran}(T - \lambda I)$. Because $(T - \lambda I)$ is not injective, it is not surjective, and thus $k = \dim(\mathscr{U}) < \dim(\mathscr{V}) = m+1$. That is $k \in \{1, \ldots, m\}$.

Moreover, $T\mathscr{U} \subseteq \mathscr{U}$. To show this, let $u \in \mathscr{U}$. Then $Tu = (T - \lambda I)u + \lambda u$. Since $(T - \lambda I)u \in \mathscr{U}$ and $\lambda u \in \mathscr{U}$, $Tu \in \mathscr{U}$. Hence, $S = T|_{\mathscr{U}}$ is an operator on $\mathscr{U}$.

By the inductive hypothesis (the green box), there exists a basis $\mathscr{A} = \{u_1, \ldots, u_k\}$ of $\mathscr{U}$ such that $M_{\mathscr{A}}^{\mathscr{A}}(S)$ is upper-triangular. That is,

$$Tu_j = Su_j \in \operatorname{span}\{u_1, \ldots, u_j\} \qquad \text{for all} \qquad j \in \{1, \ldots, k\}.$$

Extend $\mathscr{A}$ to a basis $\mathscr{B} = \{u_1, \ldots, u_k, v_1, \ldots, v_{n-k}\}$ of $\mathscr{V}$. Since

$$Tv_j = (T - \lambda I)v_j + \lambda v_j, \qquad j \in \{1, \ldots, n-k\},$$

where $(T - \lambda I)v_j \in \mathscr{U}$, for all $j \in \{1, \ldots, n-k\}$ we have

$$Tv_j \in \operatorname{span}\{u_1, \ldots, u_m, v_j\} \subseteq \operatorname{span}\{u_1, \ldots, u_m, v_1, \ldots, v_j\}.$$

By Theorem 3.5 $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular. □

Theorem 3.7. *Let $\mathscr{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $\dim \mathscr{V} = n$, and let $T \in \mathscr{L}(\mathscr{V})$. Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be a basis of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper triangular with diagonal entries a_{jj}, $j \in \{1, \ldots, n\}$. Then T is not injective if and only if there exists $j \in \{1, \ldots, n\}$ such that $a_{jj} = 0$.*

Proof. In this proof we set

$$\mathscr{V}_k = \operatorname{span}\{v_1, ..., v_k\}, \qquad k \in \{1, ..., n\}.$$

Then

$$\mathscr{V}_1 \subsetneq \mathscr{V}_2 \subsetneq \ldots \subsetneq \mathscr{V}_n \tag{18}$$

and by Theorem 3.5, $T\mathscr{V}_k \subseteq \mathscr{V}_k$.

We first prove the "only if" part. Assume that T is not injective. Consider the set

$$\mathbb{K} = \bigl\{k \in \{1, ..., n\} \,:\, T\mathscr{V}_k \subsetneq \mathscr{V}_k\bigr\}$$

Since T is not injective, $\operatorname{nul} T \neq \{0_{\mathscr{V}}\}$. Thus by the Rank-Nullity Theorem, $\operatorname{ran} T \subsetneq \mathscr{V} = \mathscr{V}_n$. Since $T\mathscr{V}_n = \operatorname{ran} T$, it follows that $T\mathscr{V}_n \subsetneq \mathscr{V}_n$. Therefore $n \in \mathbb{K}$. Hence the set $\mathbb{K}$ is a nonempty set of positive integers. Hence, by the Well-Ordering principle $\min \mathbb{K}$ exists. Set $j = \min \mathbb{K}$.

If $j = 1$, then $\dim \mathscr{V}_1 = 1$, but since $T\mathscr{V}_1 \subsetneq \mathscr{V}_1$ it must be that $\dim T\mathscr{V}_1 = 0$. Thus $T\mathscr{V}_1 = \{0_{\mathscr{V}}\}$, so $Tv_1 = 0_v$. Hence $C_{\mathscr{B}}(T) = [0 \cdots 0]^\top$ and so $a_{jj} = 0$. If $j > 1$, then $j-1 \in \{1, \ldots, n\}$ but $j - 1 \notin \mathbb{K}$. By Theorem 3.5, $T\mathscr{V}_{j-1} \subseteq \mathscr{V}_{j-1}$ and, since $j-1 \notin \mathbb{K}$, $T\mathscr{V}_{j-1} \subsetneq \mathscr{V}_{j-1}$ is not true. Hence $T\mathscr{V}_{j-1} = \mathscr{V}_{j-1}$. Since $j \in \mathbb{K}$, we have $T\mathscr{V}_j \subsetneq \mathscr{V}_j$. Now we have

$$\mathscr{V}_{j-1} = T\mathscr{V}_{j-1} \subseteq T\mathscr{V}_j \subsetneq \mathscr{V}_j.$$

Consequently,

$$j - 1 = \dim \mathscr{V}_{j-1} \leq \dim\bigl(T\mathscr{V}_j\bigr) < \dim \mathscr{V}_j = j,$$

which implies $\dim\bigl(T\mathscr{V}_j\bigr) = j - 1$ and therefore $T\mathscr{V}_j = \mathscr{V}_{j-1}$. This implies that there exist $\alpha_1, \ldots, \alpha_{j-1} \in \mathbb{F}$ such that

$$Tv_j = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1}.$$

By the definition of $M_{\mathscr{B}}^{\mathscr{B}}$ this implies that $a_{jj} = 0$.

Next we prove the "if" part. Assume that there exists $j \in \{1, ..., n\}$ such that $a_{jj} = 0$. Then

$$Tv_j = a_{1j}v_1 + \cdots + a_{j-1,j}v_{j-1} + 0v_j \in \mathscr{V}_{j-1}. \tag{19}$$

By Theorem 3.5 and (18) we have

$$Tv_i \in \mathscr{V}_i \subseteq \mathscr{V}_{j-1} \qquad \text{for all} \qquad i \in \{1, \ldots, j-1\}. \tag{20}$$

Now (19) and (20) imply $Tv_i \in \mathscr{V}_{j-1}$ for all $i \in \{1, \ldots, j\}$ and consequently $T\mathscr{V}_j \subseteq \mathscr{V}_{j-1}$. To complete the proof, we apply the Rank-Nullity theorem to the restriction $T|_{\mathscr{V}_j}$ of T to the subspace $\mathscr{V}_j$:

$$\dim \operatorname{nul}\bigl(T|_{\mathscr{V}_j}\bigr) + \dim \operatorname{ran}\bigl(T|_{\mathscr{V}_j}\bigr) = j.$$

Since $T\mathscr{V}_j \subseteq \mathscr{V}_{j-1}$ implies $\dim \operatorname{ran}\bigl(T|_{\mathscr{V}_j}\bigr) \leq j - 1$, we conclude

$$\dim \operatorname{nul}\bigl(T|_{\mathscr{V}_j}\bigr) \geq 1.$$

Thus $\operatorname{nul}\bigl(T|_{\mathscr{V}_j}\bigr) \neq \{0_{\mathscr{V}}\}$, that is, there exists $v \in \mathscr{V}_j$ such that $v \neq 0$ and $Tv = T|_{\mathscr{V}_j} v = 0$. This proves that T is not invertible. $\square$

Corollary 3.8 (Theorem 5.16)**.** *Let $\mathscr{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $\dim \mathscr{V} = n$, and let $T \in \mathscr{L}(\mathscr{V})$. Let $\mathscr{B}$ be a basis of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper triangular with diagonal entries a_{jj}, $j \in \{1, \ldots, n\}$. The following statements are equivalent.*

(a) *T is not injective.*

(b) *T is not invertible.*

(c) *0 is an eigenvalue of T.*

(d) $\prod_{i=1}^{n} a_{ii} = 0$.

(e) *There exists $j \in \{1, \ldots, n\}$ such that $a_{jj} = 0$.*

Proof. The equivalence (a) $\Leftrightarrow$ (b) follows from the Rank-nullity theorem and it has been proved earlier. The equivalence (a) $\Leftrightarrow$ (c) is almost trivial. The equivalence (a) $\Leftrightarrow$ (e) was proved in Theorem 3.7 and The equivalence (d) $\Leftrightarrow$ (e) is should have been proved in high school. □

Theorem 3.9. *Let $\mathscr{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $\dim \mathscr{V} = n$, and let $T \in \mathscr{L}(\mathscr{V})$. Let $\mathscr{B}$ be a basis of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper triangular with diagonal entries a_{jj}, $j \in \{1, \ldots, n\}$. Then*

$$\sigma(T) = \big\{ a_{jj} : j \in \{1, ..., n\} \big\}.$$

Proof. Notice that $M_{\mathscr{B}}^{\mathscr{B}} : \mathscr{L}(V) \to \mathbb{F}^{n\times n}$ is a linear operator. Therefore

$$M_{\mathscr{B}}^{\mathscr{B}}(T - \lambda I) = M_{\mathscr{B}}^{\mathscr{B}}(T) - \lambda M_{\mathscr{B}}^{\mathscr{B}}(I) = M_{\mathscr{B}}^{\mathscr{B}}(T) - \lambda I_n.$$

Here I_n denotes the identity matrix in $\mathbb{F}^{n\times n}$. As $M_{\mathscr{B}}^{\mathscr{B}}(T)$ and $M_{\mathscr{B}}^{\mathscr{B}}(I) = I_n$ are upper triangular, $M_{\mathscr{B}}^{\mathscr{B}}(T - \lambda I)$ is upper triangular as well with diagonal entries $a_{jj} - \lambda$, $j \in \{1, ..., n\}$.

To prove a set equality we need to prove two inclusions.

First we prove $\subseteq$. Let $\lambda \in \sigma(T)$. Because λ is an eigenvalue, $T - \lambda I$ is not injective. Because $T - \lambda I$ is not injective, by Theorem 3.7 one of its diagonal entries is zero. So there exists $i \in \{1, ..., n\}$ such that $a_{ii} - \lambda = 0$. Thus $\lambda = a_{ii}$. So $\sigma(T) \subseteq \big\{ a_{jj} : j \in \{1, ..., n\} \big\}$.

Next we prove $\supseteq$. Let $a_{ii} \in \big\{ a_{jj} : j \in \{1, ..., n\} \big\}$ be arbitrary. Then $a_{ii} - a_{ii} = 0$. By Theorem 3.7 and the note at the beginning of this proof $T - a_{ii} I$ is not injective. This implies that a_{ii} is an eigenvalue of T. Thus $a_{ii} \in \sigma(T)$. This completes the proof. □

Remark 3.10. Theorem 3.9 is identical to Theorem 5.18 in the textbook.