Eigensystem of a linear operator

Branko Ćurgus

February 20, 2020 at 09:25

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

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

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

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

Then by definition

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

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

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

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

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

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. □

co-linfact **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. $\square$

2 Existence of an eigenvalue

We will need the following lemma about injections.

le-inj-n **Lemma 2.1.** *Let $n \in \mathbb{N}$, let A be a nonempty set and let $f_1, \ldots, f_n \in A^A$. If for all $k \in \{1, \ldots, n\}$ f_k is an injection, then the composition $f_1 \circ \cdots \circ f_n$ 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 $f, g \in A^A$ are injections. Let $s, t \in A$ be such that $s \neq t$. Then, since g is an injection, $g(s) \neq g(t)$. Since f is injective, $f(g(x)) \neq f(g(t))$. Thus, $f \circ g$ is injective.

Next we prove the inductive step. Let $m \in \mathbb{N}$ and assume that $f_1 \circ \cdots \circ f_m$ is an injection whenever $f_1, \ldots, f_m \in A^A$ are all injections. (This is the inductive hypothesis.) Now assume that $f_1, \ldots, f_m, f_{m+1} \in A^A$ are all injections. By the inductive hypothesis the function $f = f_1 \circ \cdots \circ f_m$ is an injection. Since by assumption $g = f_{m+1}$ is an injection, the already proved claim for $n = 2$ yields that

$$f \circ g = f_1 \circ \cdots \circ f_m \circ f_{m+1}$$

is an injection. This completes the proof. $\square$

Definition 2.2. 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 2.3. 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.

Our first goal is to prove that for an operator T defined on a vector space $\mathscr{V}$ over $\mathbb{C}$ we have $\sigma(T) \neq \emptyset$. However, even before proving that we can prove the Spectral Mapping Theorem.

th-smt **Theorem 2.4.** *Let $\mathscr{V}$ be a nontrivial finite-dimensional vector space over $\mathbb{C}$ and let $T \in \mathscr{L}(\mathscr{V})$. For a nonconstant $p \in \mathbb{C}[z]$ we have*

$$\sigma\bigl(p(T)\bigr) = p\bigl(\sigma(T)\bigr).$$

Specifically we have: If λ is an eigenvalue of T with a corresponding eigenvector v, then $p(\lambda)$ is an eigenvalue of $p(T)$ with the same eigenvector v

Proof. The equality is obvious if the polynomial p is constant. Assume that $\deg p = m \in \mathbb{N}$ and let the coefficients of p be $\alpha_0, \cdots, \alpha_m \in \mathbb{C}$. Let $\lambda \in \sigma(T)$. Then there exists $v \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ such that

$$Tv = \lambda v.$$

Calculate

$$p(T)v = \Xi_T(p)v = \left(\sum_{k=0}^{m} \alpha_k T^k\right) v = \sum_{k=0}^{m} \alpha_k T^k v = \sum_{k=0}^{m} \alpha_k \lambda^k v = p(\lambda)v.$$

Thus $p(\lambda)$ is an eigenvalue of $p(T)$. This proves the specific statement in the theorem and the inclusion

$$p\bigl(\sigma(T)\bigr) \subseteq \sigma\bigl(p(T)\bigr).$$

To prove the reverse inclusion let $\mu \in p\bigl(\sigma(T)\bigr)$. Then there exists $w \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ such that

$$p(T)w = \mu w.$$

Set $q(z) = p(z) - \mu$. Then $q(T)w = 0_{\mathscr{V}}$ and since $w \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ the operator $q(T)$ is not an injection. By the Fundamental Theorem of Algebra there exist $\alpha, z_1, \ldots, z_m \in \mathbb{C}$ such that $\alpha \neq 0$ and

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

Since Ξ_T is an algebra homomorphism we have

$$\begin{aligned} q(T) &= \Xi_T(q) \\ &= \alpha\, \Xi_T(z - z_1) \cdots \Xi_T(z - z_m) \\ &= \alpha(T - z_1 I) \cdots (T - z_m I). \end{aligned}$$

That is, $q(T)$ is a composition of $m+1$ operators. Since $q(T)$ is not an injection, Lemma 2.1 yields that there exists $k \in \{1,\dots,m\}$ such that the operator $T - z_k I$ is not an injection. This implies that $z_k \in \sigma(T)$. Set $\lambda = z_k \in \sigma(T)$. Then $q(\lambda) = 0$, that is $p(\lambda) - \mu = 0$. Thus we have proved that for arbitrary $\mu \in p\bigl(\sigma(T)\bigr)$ there exists $\lambda \in \sigma(T)$ such that $\mu = p(\lambda)$. This proves

$$\sigma\bigl(p(T)\bigr) \subseteq p\bigl(\sigma(T)\bigr)$$

and completes the proof. □

Using the method of the proof of the preceding theorem one can prove.

Proposition 2.5. *Let $\mathscr{V}$ be a nontrivial finite-dimensional vector space over $\mathbb{C}$ with $n = \dim \mathscr{V}$ and let $T \in \mathscr{L}(\mathscr{V})$. Then $\sigma(T) = \{0\}$ if and only $T^n = 0_{\mathscr{L}(\mathscr{V})}$.*

th-ev-ex **Theorem 2.6.** *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

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

eq-uTu

If two of these vectors coincide, say $k,l \in \{0,\dots,n\}$, $k<l$ are such that $T^k u = T^l u$, setting $\alpha_j = 0$ for $j \in \{0,\dots,n\} \setminus \{k,l\}$ 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

$$\alpha_0 u + \alpha_1 Tu + \alpha_2 T^2 u + \cdots + \alpha_n T^n u = 0_{\mathscr{V}} \quad \text{and} \quad \alpha_k \neq 0. \tag{5}$$

eq-lin-com

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 $m = \deg p \geq k > 0$.

Thus we have constructed a polynomial p of positive degree for which, by (5),

$$p(T)u = 0_{\mathscr{V}} \quad \text{with} \quad u \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}.$$

Hence, $0 \in \sigma\bigl(p(T)\bigr)$. By the Spectral Mapping Theorem there exists $\lambda \in \sigma(T)$ such that $p(\lambda) = 0$. This proves that $\sigma(T)$ is a nonempty set. $\square$

Remark 2.7. In the preceding proof we have used the Spectral Mapping Theorem to make the proof compact. I believe that it is important to make proofs of fundamental theorems to be as self-contained as possible; just based on the basic principles, avoiding long citation chains. We could have easily avoided the citation of the Spectral Mapping Theorem. For this aim, the last paragraph of the preceding proof should be replaced by the following three paragraph.

By the Fundamental Theorem of Algebra there exist $\alpha \neq 0$ 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

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

Equality (5) yields that the operator $p(T)$ is not an injection. Lemma 2.1 now implies that there exists $j \in \{1, \ldots, m\}$ 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.

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

Theorem 2.8. *Let $\mathscr{V}$ be a vector space over $\mathbb{F}$, $T \in \mathscr{L}(\mathscr{V})$ and $n \in \mathbb{N}$. Assume*

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

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

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

Proof. Let $\lambda_1, \dots, \lambda_n$ be distinct eigenvalues of T and let $v_1, \dots, v_n$ be corresponding eigenvectors:

$$Tv_k = \lambda_k v_k, \quad \text{for all} \quad k \in \{1, \dots, n\}. \tag{6}$$

For each $k \in \{1, \dots, n\}$ define the polynomial

$$q_k(z) = \prod \bigl\{ (z - \lambda_j) \,:\, j \in \{1, \dots, n\} \setminus \{k\} \bigr\}.$$

Then q_k has exactly $n-1$ distinct roots $\{\lambda_1, \dots, \lambda_n\} \setminus \{\lambda_k\}$ and

$$q_k(\lambda_k) = \prod \bigl\{ (\lambda_k - \lambda_j) \,:\, j \in \{1, \dots, n\} \setminus \{k\} \bigr\} \neq 0.$$

That is

$$q_k(\lambda_j) = \begin{cases} 0 & j \neq k, \\ q_k(\lambda_k) \neq 0 & j = k, \end{cases} \qquad \text{for all} \quad j, k \in \{1, \dots, n\}. \qquad \boxed{\texttt{eq-qk1}} \tag{7}$$

By the specific statement in the Spectral Mapping Theorem we have

$$q_k(T)v_j = q_k(\lambda_j)v_j. \qquad \boxed{\texttt{eq-qk2}} \tag{8}$$

Now we are ready to prove the linear independence of $v_1, \dots, v_n$. Let $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ be such that

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_{\mathscr{V}}. \qquad \boxed{\texttt{eq-Vaughn-li}} \tag{9}$$

Let $k \in \{1, \dots, n\}$ be arbitrary. Apply the operator $q_k(T)$ to both sides of (9) to obtain

$$\alpha_1 q_k(T)v_1 + \cdots + \alpha_n q_k(T)v_n = 0_{\mathscr{V}}. \qquad \boxed{\texttt{eq-Vaughn-li1}} \tag{10}$$

By (8) we have

$$\alpha_1 q_k(\lambda_1)v_1 + \cdots + \alpha_n(\lambda_n)v_n = 0_{\mathscr{V}}.$$

By (7) the last equality simplifies to

$$\alpha_k q_k(\lambda_k)v_k = 0_{\mathscr{V}}.$$

Since $v_k \neq 0_{\mathscr{V}}$ and $q_k(\lambda_k) \neq 0$ we deduce

$$\alpha_k = 0.$$

Since $k \in \{1, \dots, n\}$ was arbitrary the proof of linear independence is complete. $\square$

Corollary 2.9. *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}, $i, j \in \{1, \dots, n\}$ is called *upper triangular* if $a_{ij} = 0$ for all $i, j \in \{1, \dots, n\}$ 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 each eigenspaces of S is invariant under T and each eigenspaces of T is invariant under S.*

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

$$\mathscr{U}_1 \subsetneq \mathscr{U}_2 \subsetneq \cdots \subsetneq \mathscr{U}_n \tag{11}$$

eq-fan-sss

and

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

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

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

form a fan for T.

Notice that (11) implies

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

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

th-utc **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.

i-utc-1 (a) *$M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular.*

i-utc-2 (b) *For all $k \in \{1, \ldots, n\}$ we have $Tv_k \in \mathscr{V}_k$.*

i-utc-3 (c) *For all $k \in \{1, \ldots, n\}$ we have $T\mathscr{V}_k \subseteq \mathscr{V}_k$.*

i-utc-4 (d) *$\mathscr{B}$ is a fan basis corresponding to T.*

Proof. (a) $\Rightarrow$ (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) $\Rightarrow$ (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 $i \in \{j+1, \ldots, n\}$. This proves (a).

(b) $\Rightarrow$ (c). Suppose $Tv_k \in \mathscr{V}_k = \operatorname{span}\{v_1, \ldots, v_k\}$ for all $k \in \{1, \ldots, n\}$. 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,

$$Tv \in \operatorname{span}\{Tv_1, \ldots, Tv_k\}. \tag{12}$$ eq-Tv-span1

Since

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

we have

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

Together with (12), this proves (c).

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

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

th-ex-up **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 $u \in \mathscr{V}\backslash\{0_{\mathscr{V}}\}$ 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 $k \in \{1, \ldots, m\}$ 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 2.6 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. This, by Theorem 3.5, implies

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

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

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

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

th-invc

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

$$\mathscr{V}_1 \subsetneq \mathscr{V}_2 \subsetneq \ldots \subsetneq \mathscr{V}_n \tag{13}$$

eq-fan-sub

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 $C_{\mathscr{B}}(Tv_1) = [0 \cdots 0]^\top$ and so $a_{11} = 0$. 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 $M_{\mathscr{B}}^{\mathscr{B}}(T)$ 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

$$Tv_j = a_{1j} v_1 + \cdots + a_{j-1,j} v_{j-1} + 0 v_j \in \mathscr{V}_{j-1}. \tag{14}$$

eq-inv-if1

By Theorem 3.5 and (13) we have

$$Tv_i \in \mathscr{V}_i \subseteq \mathscr{V}_{j-1} \qquad \text{for all} \qquad i \in \{1, \ldots, j-1\}. \tag{15}$$

eq-inv-if2

Now (14) and (15) imply $Tv_i \in \mathscr{V}_{j-1}$ for all $i \in \{1, \dots, j\}$ 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. □

co-invc **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}, $j \in \{1, \dots, n\}$. The following statements are equivalent.*

i-invc-1 (a) *T is not injective.*

i-invc-2 (b) *T is not invertible.*

i-invc-3 (c) *0 is an eigenvalue of T.*

i-invc-4 (d) $\prod_{i=1}^{n} a_{ii} = 0$.

i-invc-5 (e) *There exists $j \in \{1, \dots, n\}$ such that $a_{jj} = 0$.*

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

th-sp-di **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}, $j \in \{1, \dots, n\}$. Then*

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

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 $\sigma(T) \subseteq \left\{a_{jj} : j \in \{1, ..., n\}\right\}$.

Next we prove $\supseteq$. Let $a_{ii} \in \left\{a_{jj} : j \in \{1, ..., n\}\right\}$ 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. $\square$

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