1 Eigenvalues and eigenvectors of a linear operator

In this section we consider a vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$. By $\mathcal{L}(\mathcal{V})$ we denote the vector space $\mathcal{L}(\mathcal{V},\mathcal{V})$ of all linear operators on $\mathcal{V}$. The vector space $\mathcal{L}(\mathcal{V})$ 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 $\mathcal{A}$ over a field $\mathbb{F}$ is an *algebra* over $\mathbb{F}$ if the following conditions are satisfied:
(a) there exist a binary operation $\cdot : \mathcal{A}\times\mathcal{A}\to\mathcal{A}$.
(b) (*associativity*) for all $x,y,z\in\mathcal{A}$ we have $(x\cdot y)\cdot z = x\cdot(y\cdot z)$.
(c) (*right-distributivity*) for all $x,y,z\in\mathcal{A}$ we have $(x+y)\cdot z = x\cdot z + y\cdot z$.
(d) (*left-distributivity*) for all $x,y,z\in\mathcal{A}$ we have $z\cdot(x+y) = z\cdot x + z\cdot y$.
(e) (*respect for scaling*) for all $x,y\in\mathcal{A}$ 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 $\mathcal{L}(\mathcal{V})$ is the identity operator $I_{\mathcal{V}}$.

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

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

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

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

Recall that by definition of a span a nonzero $S\in\mathcal{L}(\mathcal{V})$ belongs to $\mathcal{A}_T$ 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

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

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 $T\in\mathcal{L}(\mathcal{V})$ we define

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

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

Proof. It is not difficult to prove that $\Xi_T : \mathbb{F}[z] \to \mathcal{L}(\mathcal{V})$ is linear. We will prove that $\Xi_T : \mathbb{F}[z] \to \mathcal{L}(\mathcal{V})$ 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

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

This proves the multiplicative property of Ξ_T.

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

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

Proof. Let $Q, S \in \mathcal{A}_T$. Since $\mathcal{A}_T$ 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 $QS \in \operatorname{ran}(\Xi_T) = \mathcal{A}_T$ and $QS = SQ$. That is $\mathcal{A}_T$ is closed with respect to the operator composition and the operator composition on $\mathcal{A}_T$ is commutative. □

Corollary 1.4. *Let $\mathcal{V}$ be a complex vector space and let $T \in \mathcal{L}(\mathcal{V})$ be a nonzero operator. Then for every $p \in \mathbb{C}[z]$ such that $\deg p \geq 1$ 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. □

Lemma 1.5. *Let $n \in \mathbb{N}$ and $S_1, \ldots, S_n \in \mathcal{L}(\mathcal{V})$. If $S_1, \ldots, S_n$ are all injective, then $S_1 \cdots S_n$ is injective.*

Proof. We proceed by Mathematical Induction. The base step is trivial. It is useful to prove the implication for $n = 2$. Assume that $S, T \in \mathcal{L}(\mathcal{V})$ are injective and let $u, v \in \mathcal{V}$ 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 $S_1, \ldots, S_m \in \mathcal{L}(\mathcal{V})$ are all injective. (This is the inductive hypothesis.) Now assume that $S_1, \ldots, S_m, S_{m+1} \in \mathcal{L}(\mathcal{V})$ 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 1.6. *Let $\mathcal{V}$ be a nontrivial finite dimensional vector space over $\mathbb{C}$. Let $T \in \mathcal{L}(\mathcal{V})$. Then there exists a $\lambda \in \mathbb{C}$ and $v \in \mathcal{V}$ such that $v \neq 0_v$ and $Tv = \lambda v$.*

Proof. The claim of the theorem is trivial if T is a scalar multiple of the identity operator. So, assume that $T \in \mathcal{L}(\mathcal{V})$ is not a scalar multiple of the identity operator.

Since $\mathcal{L}(\mathcal{V})$ is finite dimensional and $\mathbb{C}[z]$ is infinite dimensional, by the Rank-nullity theorem the operator Ξ_T is not injective. Thus $\operatorname{nul}(\Xi_T) \neq \{0_v\}$. Hence, there exists a $p \in \mathbb{C}[z]$ such that $p \neq 0_{\mathbb{C}[z]}$ and $\Xi_T(p) = p(T) = 0_{\mathcal{L}(\mathcal{V})}$. Since $p \neq 0_{\mathbb{C}[z]}$ then $\deg p \geq 0$. Note that if $\deg p = 0$ then $p(z) = c$ for some $c \in \mathbb{C}$ for all $z \in \mathbb{C}$. Thus $\Xi_T(p) = p(T) = cI_{\mathcal{V}}$. This is not possible since we assume that T is not a scalar multiple of the identity. Hence $\deg p > 0$. By Corollary 1.4 there exists $\alpha \neq 0$ and $z_1, \ldots, z_m \in \mathbb{C}$ such that

$$0_{\mathcal{L}(\mathcal{V})} = \Xi_T(p) = p(T) = \alpha(T - z_1 I) \cdots (T - z_m I).$$

Since $0_{\mathcal{L}(\mathcal{V})}$ is not injective, Lemma 1.5 implies that there exists $j \in \{1, \ldots, m\}$ such that $T - z_j I$ is not injective. That is, there exists $v \in \mathcal{V}$, $v \neq 0_{\mathcal{V}}$ such that

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

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

Remark 1.7. Note that the proof in the textbook is different. The proof in the textbook is somewhat more elementary since it does not use the Rank-nullity theorem.

Definition 1.8. Let $\mathcal{V}$ be a vector space over $\mathbb{F}$, $T \in \mathcal{L}(\mathcal{V})$. A scalar $\lambda \in \mathbb{F}$ is an *eigenvalue* of T if there exists $v \in \mathcal{V}$ such that $v \neq 0$ and $Tv = \lambda v$. The subspace $\operatorname{nul}(T - \lambda I)$ of $\mathcal{V}$ is called the *eigenspace* of T corresponding to λ

Definition 1.9. Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$. Let $T \in \mathcal{L}(\mathcal{V})$. The set of all eigenvalues of T is denoted by $\sigma(T)$. It is called the *spectrum* of T.

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

Theorem 1.10. *Let $\mathcal{V}$ be a vector space over $\mathbb{F}$, $T \in \mathcal{L}(\mathcal{V})$ and $n \in \mathbb{N}$. If the following two conditions are satisfied:*

(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) $v_1, \ldots, v_n \in \mathcal{V}$ *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. 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 $v_1 \in \mathcal{V}$ 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) $\mu_1, \ldots, \mu_m \in \mathbb{F}$ are such that $\mu_i \neq \mu_j$ for all $i, j \in \{1, \ldots, m\}$ such that $i \neq j$,
(ii) $w_1, \ldots, w_m \in \mathcal{V}$ are such that $Tw_k = \mu_k w_k$ and $w_k \neq 0$ for all $k \in \{1, \ldots, m\}$,
then $\{w_1, \ldots, w_m\}$ is linearly independent.

We need to prove the following implication

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

Assume (I) and (II) in the red box. We need to prove that $\{v_1, \ldots, v_{m+1}\}$ is linearly independent.
Let $\alpha_1, \ldots, \alpha_{m+1} \in \mathbb{F}$ be such that

$$\alpha_1 v_1 + \cdots + \alpha_m v_m + \alpha_{m+1} v_{m+1} = 0. \tag{4}$$

Applying $T \in \mathcal{L}(\mathcal{V})$ to both sides of (4), using the linearity of T and assumption (II) we get

$$\alpha_1 \lambda_1 v_1 + \cdots + \alpha_m \lambda_m v_m + \alpha_{m+1} \lambda_{m+1} v_{m+1} = 0. \tag{5}$$

Multiplying both sides of (4) by λ_{m+1} we get

$$\alpha_1 \lambda_{m+1} v_1 + \cdots + \alpha_m \lambda_{m+1} v_m + \alpha_{m+1} \lambda_{m+1} v_{m+1} = 0. \tag{6}$$

Subtracting (6) from (5) 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 $j \in \{1, \ldots, m\}$, setting

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

and taking into account (II) we have

$$w_j \neq 0 \qquad \text{and} \qquad Tw_j = \lambda_j w_j \qquad \text{for all} \qquad j \in \{1, \ldots, m\}. \tag{7}$$

Thus, by (I) and (7), the scalars $\lambda_1, \ldots, \lambda_m$ and vectors $w_1, \ldots, w_m$ satisfy assumptions (i) and (ii) of the inductive hypothesis (the green box). Consequently, the vectors $w_1, \ldots, w_m$ are linearly independent. Since by (7) 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 (4) 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 $v_1, \ldots, v_{m+1}$. □

Corollary 3: Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$ and let $T \in \mathcal{L}(\mathcal{V})$. Then T has at most $n = \dim \mathcal{V}$ distinct eigenvalues.

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

Definition 1.11. Let $\mathcal{V}$ be a vector space over $\mathbb{F}$ and $T \in \mathcal{L}(V)$. A subspace $\mathcal{U}$ of $\mathcal{V}$ is called an *invariant subspace* under T if $T(\mathcal{U}) \subseteq \mathcal{U}$.

The following proposition is straightforward.

Proposition 1.12. *Let $S, T \in \mathcal{L}(\mathcal{V})$ 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 1.13. A matrix $A \in \mathbb{F}^{n\times n}$ with entries a_{ij}, $i, j \in \{1, \dots, n\}$ is called *upper triangular* if $a_{i,j} = 0$ for all $i, j \in \{1, \dots, n\}$ such that $i > j$.

Definition 1.14. Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $n = \dim \mathcal{V} > 0$. Let $T \in \mathcal{L}(\mathcal{V})$. A sequence of nontrivial subspaces $\mathcal{U}_1, \dots, \mathcal{U}_n$ of $\mathcal{V}$ such that

$$\mathcal{U}_1 \subsetneq \mathcal{U}_2 \subsetneq \cdots \subsetneq \mathcal{U}_n \tag{8}$$

and

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

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

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

form a fan for T.

Notice that (8) implies

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

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

Theorem 1.15 (Theorem 5.12)**.** *Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$ with* $\dim \mathcal{V} = n$ *and let* $T \in \mathcal{L}(\mathcal{V})$. *Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis of $\mathcal{V}$ and set*

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

The following statements are equivalent.

(a) *$M_{\mathcal{B}}^{\mathcal{B}}(T)$ is upper-triangular.*
(b) *$Tv_k \in \mathcal{V}_k$ for all $k \in \{1, \dots, n\}$.*
(c) *$T\mathcal{V}_k \subseteq \mathcal{V}_k$ for all $k \in \{1, \dots, n\}$.*
(d) *$\mathcal{B}$ is a fan basis corresponding to T.*

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

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

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

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

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

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

Thus, (b) is proved.

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

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

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

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

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

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

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

(b) $\Rightarrow$ (c). Suppose $Tv_k \in \mathcal{V}_k = \mathit{span}\{v_1, \dots, v_k\}$ for all $k \in \{1, \dots, n\}$. Let $v \in \mathcal{V}_k$. 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,

$$Tv \in \operatorname{span}\{Tv_1, \dots, Tv_k\}. \tag{9}$$

Since

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

we have

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

Together with (9), this proves (c).

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

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

Theorem 1.16. *Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $\dim \mathcal{V} = n$, and let $T \in \mathcal{L}(\mathcal{V})$. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of $\mathcal{V}$ such that $M_{\mathcal{B}}^{\mathcal{B}}(T)$ 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

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

Then

$$\mathcal{V}_1 \subsetneq \mathcal{V}_2 \subsetneq \ldots \subsetneq \mathcal{V}_n \tag{10}$$

and by Theorem 1.15, $T\mathcal{V}_k \subseteq \mathcal{V}_k$.

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

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

Since T is not injective, $\operatorname{nul} T \neq \{0_{\mathcal{V}}\}$. Thus by the Rank-Nullity Theorem, $\operatorname{ran} T \subsetneq \mathcal{V} = \mathcal{V}_n$. Since $T\mathcal{V}_n = \operatorname{ran} T$, it follows that $T\mathcal{V}_n \subsetneq \mathcal{V}_n$. 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 $\dim \mathcal{V}_1 = 1$, but since $T\mathcal{V}_1 \subsetneq \mathcal{V}_1$ it must be that $\dim T\mathcal{V}_1 = 0$. Thus $T\mathcal{V}_1 = \{0_{\mathcal{V}}\}$, so $Tv_1 = 0_v$. Hence $C_{\mathcal{B}}(T) = [0 \cdots 0]^\top$ and so $a_{jj} = 0$. If $j > 1$, then $j - 1 \in \{1, \ldots, n\}$ but $j - 1 \notin \mathbb{K}$. By Theorem 1.15, $T\mathcal{V}_{j-1} \subseteq \mathcal{V}_{j-1}$ and, since $j - 1 \notin \mathbb{K}$, $T\mathcal{V}_{j-1} \subsetneq \mathcal{V}_{j-1}$ is not true. Hence $T\mathcal{V}_{j-1} = \mathcal{V}_{j-1}$. Since $j \in \mathbb{K}$, we have $T\mathcal{V}_j \subsetneq \mathcal{V}_j$. Now we have

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

Consequently,

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

which implies $\dim\bigl(T\mathcal{V}_j\bigr) = j - 1$ and therefore $T\mathcal{V}_j = \mathcal{V}_{j-1}$. 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 $M_{\mathcal{B}}^{\mathcal{B}}$ 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

$$Tv_j = a_{1j} v_1 + \cdots + a_{j-1,j} v_{j-1} + 0 v_j \in \mathcal{V}_{j-1}. \tag{11}$$

By Theorem 1.15 and (10) we have

$$Tv_i \in \mathcal{V}_i \subseteq \mathcal{V}_{j-1} \qquad \text{for all} \qquad i \in \{1, \ldots, j-1\}. \tag{12}$$

Now (11) and (12) imply $Tv_i \in \mathcal{V}_{j-1}$ for all $i \in \{1, \ldots, j\}$ and consequently $T\mathcal{V}_j \subseteq \mathcal{V}_j$. To complete the proof, we apply the Rank-Nullity theorem to the restriction $T|_{\mathcal{V}_j}$ of T to the subspace $\mathcal{V}_j$:

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

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

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

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

Corollary 1.17 (Theorem 5.16)**.** *Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $\dim \mathcal{V} = n$, and let $T \in \mathcal{L}(\mathcal{V})$. Let $\mathcal{B}$ be a basis of $\mathcal{V}$ such that $M_{\mathcal{B}}^{\mathcal{B}}(T)$ 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 $j \in \{1, \dots, n\}$ 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 1.16 and The equivalence (d) $\Leftrightarrow$ (e) is should have been proved in high school. □

Theorem 1.18. *Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$ with $\dim \mathcal{V} = n$, and let $T \in \mathcal{L}(\mathcal{V})$. Let $\mathcal{B}$ be a basis of $\mathcal{V}$ such that $M_{\mathcal{B}}^{\mathcal{B}}(T)$ is upper triangular with diagonal entries a_{jj}, $j \in \{1, \dots, n\}$. Then*

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

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

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

Here I_n denotes the identity matrix in $\mathbb{F}^{n\times n}$. As $M_{\mathcal{B}}^{\mathcal{B}}(T)$ and $M_{\mathcal{B}}^{\mathcal{B}}(I) = I_n$ are upper triangular, $M_{\mathcal{B}}^{\mathcal{B}}(T-\lambda I)$ 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 1.16 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 1.16 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 1.19. Theorem 1.18 is identical to Theorem 5.18 in the textbook.

Theorem 1.20 (Theorem 5.13)**.** *Let $\mathcal{V}$ be a nonzero finite dimensional complex vector space. If $\dim \mathcal{V} = n$ and $T \in \mathcal{L}(\mathcal{V})$, then there exists a basis $\mathcal{B}$ of $\mathcal{V}$ such that $M_{\mathcal{B}}^{\mathcal{B}}(T)$ is upper-triangular.*

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

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

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 $\dim \mathcal{U} = k$ and $S \in \mathcal{L}(\mathcal{U})$, then there exists a basis $\mathcal{A}$ of $\mathcal{U}$ such that $M_{\mathcal{A}}^{\mathcal{A}}(S)$ is upper-triangular.

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

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

To prove the red implication assume that $\dim \mathcal{V} = m+1$ and $T \in \mathcal{L}(\mathcal{V})$. By Theorem 1.6 the operator T has an eigenvalue. Let λ be an eigenvalue of T. Set $\mathcal{U} = \operatorname{ran}(T - \lambda I)$. Because $(T - \lambda I)$ is not injective, it is not surjective, and thus $k = \dim(\mathcal{U}) < \dim(\mathcal{V}) = m + 1$. That is $k \in \{1, \dots, m\}$.

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

By the inductive hypothesis (the green box), there exists a basis $\mathcal{A} = \{u_1, \ldots, u_k\}$ of $\mathcal{U}$ such that $M_{\mathcal{A}}^{\mathcal{A}}(S)$ 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 $\mathcal{A}$ to a basis $\mathcal{B} = \{u_1, \ldots, u_k, v_1, \ldots, v_{n-k}\}$ of $\mathcal{V}$. Since

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

where $(T - \lambda I)v_j \in \mathcal{U}$, we have

$$Tv_j \in \operatorname{span}\{u_1, \ldots, u_m, v_j\} \subseteq \operatorname{span}\{u_1, \ldots, u_m, v_1, \ldots, v_j\} \qquad \text{for all} \qquad j \in \{1, \ldots, n-k\}.$$

By Theorem 1.15, $M_{\mathcal{B}}^{\mathcal{B}}(T)$ is upper-triangular. □

2 Inner Product Spaces

We will first introduce several "dot-product-like" objects. We start with the most general.

Definition 2.1. Let $\mathcal{V}$ be a vector space over a scalar field $\mathbb{F}$. A function

$$[\,\cdot\,,\cdot\,] : \mathcal{V} \times \mathcal{V} \to \mathbb{F}$$

is a *sesquilinear form* on $\mathcal{V}$ if the following two conditions are satisfied.

(a) (linearity in the first variable) $\forall\, \alpha, \beta \in \mathbb{F} \quad \forall\, u, v, w \in \mathcal{V} \quad [\alpha u + \beta v, w] = \alpha[u, w] + \beta[v, w]$.
(b) (anti-linearity in the second variable) $\forall\, \alpha, \beta \in \mathbb{F} \quad \forall\, u, v, w \in \mathcal{V} \quad [u, \alpha v + \beta w] = \overline{\alpha}[u, v] + \overline{\beta}[u, w]$.

Example 2.2. Let $M \in \mathbb{C}^{n\times n}$ be arbitrary. Then

$$[\mathbf{x}, \mathbf{y}] = (M\mathbf{x}) \cdot \mathbf{y}, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{C}^n,$$

is a sesquilinear form on the complex vector space $\mathbb{C}^n$. Here $\cdot$ denotes the usual dot product in $\mathbb{C}$.

Theorem 2.3. *Let $\mathcal{V}$ be a vector space over a scalar field $\mathbb{F}$ and let $[\,\cdot\,,\cdot\,] : \mathcal{V} \times \mathcal{V} \to \mathbb{F}$ be a sesquilinear form on $\mathcal{V}$. If $\mathrm{i} \in \mathbb{F}$, then*

$$[u, v] = \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \big[u + \mathrm{i}^k v, u + \mathrm{i}^k v\big] \tag{13}$$

for all $u, v \in \mathcal{V}$.

Corollary 2.4. *Let $\mathcal{V}$ be a vector space over a scalar field $\mathbb{F}$ and let $[\,\cdot\,,\cdot\,] : \mathcal{V} \times \mathcal{V} \to \mathbb{F}$ be a sesquilinear form on $\mathcal{V}$. If $\mathrm{i} \in \mathbb{F}$ and $[v, v] = 0$ for all $v \in \mathcal{V}$, then $[u, v] = 0$ for all $u, v \in \mathcal{V}$.*

Definition 2.5. Let $\mathcal{V}$ be a vector space over a scalar field $\mathbb{F}$. A sesquilinear form $[\,\cdot\,,\cdot\,] : \mathcal{V} \times \mathcal{V} \to \mathbb{F}$ is *hermitian* if

(c) (hermiticity) $\forall\, u, v \in \mathcal{V} \quad \overline{[u, v]} = [v, u]$.

A hermitian sesquilinear form is also called an *inner product*.

Let $[\,\cdot\,,\cdot\,]$ be an inner product on $\mathcal{V}$. The hermiticity of $[\,\cdot\,,\cdot\,]$ implies that $\overline{[v, v]} = [v, v]$ for all $v \in \mathcal{V}$. Thus $[v, v] \in \mathbb{R}$ for all $v \in \mathcal{V}$. The natural trichotomy that arises is the motivation for the following definition.

Definition 2.6. An inner product $[\,\cdot\,,\cdot\,]$ on $\mathcal{V}$ is called *nonnegative* if $[v,v] \geq 0$ for all $v \in \mathcal{V}$, it is called *nonpositive* if $[v,v] \leq 0$ for all $v \in \mathcal{V}$, and it is called *indefinite* if there exist $u \in \mathcal{V}$ and $v \in \mathcal{V}$ such that $[u,u] < 0$ and $[v,v] > 0$.

The following implication that you might have learned in high school will be useful below.

Theorem 2.7 (High School Theorem). *Let a,b,c be real numbers. Assume $a \geq 0$. Then the following implication holds:*

$$\forall\, x \in \mathbb{Q} \quad ax^2 + bx + c \geq 0 \qquad \Rightarrow \qquad b^2 - 4ac \leq 0. \tag{14}$$

Theorem 2.8 (Cauchy-Bunyakovsky-Schwartz Inequality). *Let $\mathcal{V}$ be a vector space over $\mathbb{F}$ and let $\langle\,\cdot\,,\cdot\,\rangle$ be a nonnegative inner product on $\mathcal{V}$. Then*

$$\forall\, u,v \in \mathcal{V} \quad |\langle u,v\rangle|^2 \leq \langle u,u\rangle\langle v,v\rangle. \tag{15}$$

The equality occurs in (15) *if and only if there exists $\alpha,\beta \in \mathbb{F}$ not both* 0 *such that* $\langle \alpha u + \beta v, \alpha u + \beta v\rangle = 0$.

Proof. Let $u,v \in \mathcal{V}$ be arbitrary. Since $\langle\cdot,\cdot\rangle$ is nonnegative we have

$$\forall\, t \in \mathbb{Q} \qquad \big\langle u + t\langle u,v\rangle v, u + t\langle u,v\rangle v\big\rangle \geq 0. \tag{16}$$

Since $\langle\cdot,\cdot\rangle$ is a sesquilinear hermitian form on $\mathcal{V}$, (16) is equivalent to

$$\forall\, t \in \mathbb{Q} \qquad \langle u,u\rangle + 2t|\langle u,v\rangle|^2 + t^2|\langle u,v\rangle|^2\langle v,v\rangle \geq 0. \tag{17}$$

As $\langle v,v\rangle \geq 0$, the High School Theorem applies and (17) implies

$$4|\langle u,v\rangle|^4 - 4|\langle u,v\rangle|^2\langle u,u\rangle\langle v,v\rangle \leq 0. \tag{18}$$

Again, since $\langle u,u\rangle \geq 0$ and $\langle v,v\rangle \geq 0$, (18) is equivalent to

$$|\langle u,v\rangle|^2 \leq \langle u,u\rangle\langle v,v\rangle.$$

Since $u,v \in \mathcal{V}$ were arbitrary, (15) is proved. □

Corollary 2.9. *Let $\mathcal{V}$ be a vector space over $\mathbb{F}$ and let $\langle\,\cdot\,,\cdot\,\rangle$ be a nonnegative inner product on $\mathcal{V}$. Then the following two implications are equivalent.*

(i) *If $v \in \mathcal{V}$ and $\langle u,v\rangle = 0$ for all $u \in \mathcal{V}$, then $v = 0$.*
(ii) *If $v \in \mathcal{V}$ and $\langle v,v\rangle = 0$, then $v = 0$.*

Proof. Assume that the implication (i) holds and let $v \in \mathcal{V}$ be such that $\langle v,v\rangle = 0$. Let $u \in \mathcal{V}$ be arbitrary. By the the CBS inequality

$$|\langle u,v\rangle|^2 \leq \langle u,u\rangle\langle v,v\rangle = 0.$$

Thus, $\langle u,v\rangle = 0$ for all $u \in \mathcal{V}$. By (i) we conclude $v = 0$. This proves (ii).

The converse is trivial. However, here is a proof. Assume that the implication (ii) holds. To prove (i), let $v \in \mathcal{V}$ and assume $\langle u,v\rangle = 0$ for all $u \in \mathcal{V}$. Setting $u = v$ we get $\langle v,v\rangle = 0$. Now (ii) yields $v = 0$. □

Definition 2.10. Let $\mathcal{V}$ be a vector space over a scalar field $\mathbb{F}$. An inner product $[\,\cdot\,,\cdot\,]$ on $\mathcal{V}$ is *nondegenerate* if the following implication holds

(d) (nondegenerecy) $u \in \mathcal{V}$ and $[u,v] = 0$ for all $v \in \mathcal{V}$ implies $u = 0$.

It follows from Corollary 2.9 that a nonnegative inner product $\langle\,\cdot\,,\cdot\,\rangle$ on $\mathcal{V}$ is nondegenerate if and only if $\langle v, v\rangle = 0$ implies $v = 0$. A nonnegative nondegenerate inner product is also called *positive definite inner product.* Since this is the most often encountered inner product we give its definition as it commonly given in textbooks.

Definition 2.11. Let $\mathcal{V}$ be a vector space over a scalar field $\mathbb{F}$. A function $\langle\,\cdot\,,\cdot\,\rangle : \mathcal{V}\times\mathcal{V} \to \mathbb{F}$ is called a *positive definite inner product* on $\mathcal{V}$ if the following conditions are satisfied;

(a) $\forall\, u, v, w \in \mathcal{V}\;\; \forall\, \alpha, \beta \in \mathbb{F} \quad \langle \alpha u + \beta v, v\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$,
(b) $\forall\, u, v \in \mathcal{V} \quad \langle u, v\rangle = \overline{\langle v, u\rangle}$,
(c) $\forall\, v \in \mathcal{V} \quad \langle v, v\rangle \geq 0$,
(d) If $v \in \mathcal{V}$ and $\langle v, v\rangle = 0$, then $v = 0$.

■ ....■ Branko Ćurgus revised up to here. ■ ....■

Theorem 2.12. *Pythagorean Theorem*

Let $u, v \in \mathcal{V}$. Then $\langle u, v\rangle = 0 \implies \langle u + v, u + v\rangle = \langle u, u\rangle + \langle v, v\rangle$

Furthermore, if $v_1, \cdots, v_n \in \mathcal{V}$ and $\langle v_j, j_k\rangle = 0$ whenever $j \neq k$ then $\langle \sum_{j=1}^{n} v_j, \sum_{k=1}^{n} v_k\rangle = \sum_{j=1}^{n}\langle v_j, v_j\rangle$

Proof. For two vectors.

$$\begin{aligned}\langle u + v, u + v\rangle &= \langle u, u + v\rangle + \langle v, u + v\rangle \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \langle u, u\rangle + 2Re\langle u, v\rangle + \langle v, v\rangle \\ &= \langle u, u\rangle + \langle v, v\rangle\end{aligned}$$

□

November 8 (The Gram-Schmidt orthogonalization was proven the previous day)

Theorem 2.13 (Gram-Schmidt)**.** *If $\mathcal{V}$ is a finite dimensional vector space with positive definite inner product $\langle\,\cdot\,,\cdot\,\rangle$, then $\mathcal{V}$ has an orthonormal basis.*

Corollary 2.14. *If $\mathcal{V}$ is a complex vector space with positive definite inner product and $T \in \mathcal{L}(\mathcal{V})$ then there exists an orthonormal basis B such that $\mathsf{M}_B^B(T)$ is upper-triangular.*

Definition 2.15. Let $(\mathcal{V}, \langle\,\cdot\,,\cdot\,\rangle)$ be a finite dimensional positive definite inner product space and $A \subset \mathcal{V}$. We define $A^\perp = \{v \in \mathcal{V} : \langle v, a\rangle = 0\; \forall\, a \in A\}$.

Claim (Not proven in class): $A^\perp$ is a subspace of $\mathcal{V}$.

Theorem 2.16. *If $\mathcal{U}$ is a subspace of $\mathcal{V}$, then $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^\perp$.*

Proof. Let $v \in \mathcal{U}$ and $v \in \mathcal{U}^\perp$. Then $\langle v, v\rangle = 0$. Since the $\langle\,\cdot\,,\cdot\,\rangle$ is positive definite, this implies $v = 0_{\mathcal{V}}$. Note that since $\mathcal{U}$ is a subspace of $\mathcal{V}$, $\mathcal{U}$ inherits the positive definite inner product space. Thus $\mathcal{U}$ is a finite dimensional positive definite inner product space. Thus there exists an orthonormal basis of $\mathcal{U}$, $\mathcal{B} = \{u_1, u_2, \ldots u_k\}$.

Let $v \in \mathcal{V}$ be arbitrary. By the Gram-Schmidt process,

$$v = \left(\sum_{j=1}^{k}\langle v, u_j\rangle u_j\right) + \left(v - \sum_{j=1}^{k}\langle v, u_j\rangle u_j\right),$$

where the first summand is in $\mathcal{U}$ and the second summand is in $\mathcal{U}^\perp$. More succinctly, we write this as $v = w + (v - w)$ where $w = \sum_{j=1}^{k} \langle v, u_j \rangle u_j$. We prove w is unique: $u \in \mathcal{U}^\perp$ if and only if $\langle w, u_j \rangle = 0$ for all $j \in \{1, \ldots k\}$. The forward direction is trivial (from the definition of $\mathcal{U}^\perp$). To prove the reverse direction, let $u \in \mathcal{U}$ be arbitrary. Then there exist $\alpha_j \in \mathbb{F}$ such that $u = \sum_{j=1}^{k} \alpha_j u_j$. Now calculate

$$\langle w, u \rangle = \left\langle w, \sum_{j=1}^{k} \alpha_j u_j \right\rangle = \sum_{j=1}^{k} \bar{\alpha}_j \langle w, u_j \rangle = 0.$$

The last equality follows from the assumption. Thus $u \in \mathcal{U}^\perp$.

Now for every $i \in \{1, \ldots k\}$,

$$\langle v - w, u_i \rangle = \left\langle v - \sum_{j=1}^{k} \langle v, u_j \rangle u_j, u_i \right\rangle = \langle v, u_i \rangle - \sum_{j=1}^{k} \langle v, u_j \rangle \langle u_j, u_i \rangle = \langle v, u_i \rangle - \langle v, u_i \rangle = 0.$$

□

Definition 2.17. By the previous theorem, if $\mathcal{U}$ is a subspace of $\mathcal{V}$, then $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^\perp$ implies for all $v \in \mathcal{V}$, there exists a unique $u \in \mathcal{U}$ such that $(v - u) \in \mathcal{U}^\perp$ and $v = u + (v - u)$. This defines a function which we call the **orthogonal projection** of v onto $\mathcal{U}$ as $P_{\mathcal{U}} : \mathcal{V} \to \mathcal{U}$ such that $P_{\mathcal{U}}(v) = u$.

Since $\mathcal{U}$ is a subspace of $\mathcal{V}$, $P_{\mathcal{U}} \in \mathcal{L}(\mathcal{V})$. Furthermore, $\operatorname{ran} P_{\mathcal{U}} = \mathcal{U}$, $\operatorname{nul} P_{\mathcal{U}} = \mathcal{U}^\perp$, and $(P_{\mathcal{U}})^2 = P_{\mathcal{U}}$ (idempotent).

Proposition 2.18. *Let $\mathcal{U}$ be a subspace of $\mathcal{V}$, $v \in \mathcal{V}$ be arbitrary. Let $u_0 \in \mathcal{U}$. Then $\|v - u_0\| \leq \|v - u\|$ for every $u \in \mathcal{U}$ if and only if $P_{\mathcal{U}}(v) = u_0$ and $v - u_0 \in \mathcal{U}^\perp$.*

Proof. ($\Longleftarrow$): Assume $v \in \mathcal{V}$, $u, u_0 \in \mathcal{U}$, $v - u_0 \in \mathcal{U}^\perp$. Then $\|v - u\|^2 = \|v - u_0 + u_0 + u\|^2$, where $v - u_0 \in \mathcal{U}^\perp$ and $u_0 + u \in \mathcal{U}$. By the pythagorean theorem,

$$\|v - u_0 + u_0 + u\|^2 = \|v - u_0\|^2 + \|u_0 - u\|^2 \geq \|v - u_0\|^2.$$

($\Longrightarrow$) Assume $\|v - u_0\| \leq \|v - u\|$ for all $u \in \mathcal{U}$. We show $v - u_0 \in \mathcal{U}^\perp$. This direction of the proof was given on November 9. □

■ .... ■ Stuff from November 19, 2013 ■ .... ■

Lemma 2.19. *Let $\mathcal{V}$ be a vector space over $\mathbb{F}$ and let $\langle \cdot, \cdot \rangle$ be a positive definite inner product on $\mathcal{V}$. Let $\mathcal{U}$ be a subspace of $\mathcal{V}$ and let $T \in \mathcal{L}(\mathcal{V})$. The subspace $\mathcal{U}$ is invariant under T if and only if the subspace $\mathcal{U}^\perp$ is invariant under T^*.*

Proof. By the definition of adjoint we have

$$\langle Tu, v \rangle = \langle u, T^* v \rangle \tag{19}$$

for all $u, v \in \mathcal{V}$. Assume $T\mathcal{U} \subset \mathcal{U}$. From (19) we get

$$0 = \langle Tu, v \rangle = \langle u, T^* v \rangle \qquad \forall\, u \in \mathcal{U} \quad \text{and} \quad \forall v \in \mathcal{U}^\perp.$$

Therefore, $T^* v \in \mathcal{U}^\perp$ for all $v \in \mathcal{U}^\perp$. This proves "only if" part.

The proof of the "if" part is similar. □

In the proof of the next theorem we use δ_{ij} to represent the Kronecker delta function, that is $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

Theorem 2.20 (Spectral theorem for normal operators)**.** *Let $\mathcal{V}$ be a finite dimensional complex vector space with a positive definite inner product $\langle \cdot, \cdot \rangle$. Let $T \in \mathcal{L}(\mathcal{V})$. Then T is normal if and only if there exists an orthonormal basis of $\mathcal{V}$ which consists of eigenvectors of T.*

Proof. Set $n = \dim \mathcal{V}$. We first prove "only if" part. Assume that T is normal. Set

$$\mathbb{K} = \left\{ k \in \{1, \ldots, n\} : \begin{array}{l} \exists\, w_1, \ldots, w_k \in \mathcal{V} \quad \text{and} \quad \exists\, \lambda_1, \ldots, \lambda_k \in \mathbb{C} \\ \text{such that } \langle w_i, w_j \rangle = \delta_{ij} \text{ and } T w_j = \lambda_j w_j \\ \text{for all} \quad i, j \in \{1, \ldots, k\} \end{array} \right\}$$

Clearly $1 \in \mathbb{K}$. Since $\mathbb{K}$ is finite, $m = \max \mathbb{K}$ exists. Clearly, $m \leq n$.

Next we will prove that $k \in \mathbb{K}$ and $k < n$ implies that $k + 1 \in \mathbb{K}$. Assume $k \in \mathbb{K}$ and $k < n$. Let $w_1, \ldots, w_k \in \mathcal{V}$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ be such that $\langle w_i, w_j \rangle = \delta_{ij}$ and $T w_j = \lambda_j w_j$ for all $i, j \in \{1, \ldots, k\}$. Set

$$\mathcal{W} = \operatorname{span}\{w_1, \ldots, w_k\}.$$

Since $w_1, \ldots, w_k$ are eigenvectors of T we have $T\mathcal{W} \subseteq \mathcal{W}$. By Lemma 2.19, $T^*\bigl(\mathcal{W}^\perp\bigr) \subseteq \mathcal{W}^\perp$. Thus, $T^*|_{\mathcal{W}^\perp} \in \mathcal{L}\bigl(\mathcal{W}^\perp\bigr)$. Since $\dim \mathcal{W} = k < n$ we have $\dim\bigl(\mathcal{W}^\perp\bigr) = n - k \geq 1$. Since $\mathcal{W}^\perp$ is a complex vector space the operator $T^*|_{\mathcal{W}^\perp}$ has an eigenvalue μ with the corresponding unit eigenvector u. Clearly, $u \in \mathcal{W}^\perp$ and $T^* u = \mu u$. Since T^* is normal, we have $T u = \overline{\mu} u$. Since $u \in \mathcal{W}^\perp$ and $T u = \overline{\mu} u$, setting $w_{k+1} = u$ and $\lambda_{k+1} = \overline{\mu}$ we have

$$\langle w_i, w_j \rangle = \delta_{ij} \quad \text{and} \quad T w_j = \lambda_j w_j \quad \text{for all} \quad i, j \in \{1, \ldots, k, k+1\}.$$

Thus $k + 1 \in \mathbb{K}$. Consequently, $k < m$. Thus, for $k \in \mathbb{K}$, we have proved the implication

$$k < n \qquad \Rightarrow \qquad k < m.$$

The contrapositive of this implication is: For $k \in \mathbb{K}$, we have

$$k \geq m \qquad \Rightarrow \qquad k \geq n.$$

In particular, for $m \in \mathbb{K}$ we have $m = m$ implies $m \geq n$. Since $m \leq n$ is also true, this proves that $m = n$. That is, $n \in \mathbb{K}$. This implies that there exist $u_1, \ldots, u_n \in \mathcal{V}$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that $\langle u_i, u_j \rangle = \delta_{ij}$ and $T u_j = \lambda_j u_j$ for all $i, j \in \{1, \ldots, n\}$.

Since $u_1, \ldots, u_n$ are orthonormal, they are linearly independent. Since $n = \dim \mathcal{V}$, it turns out that $u_1, \ldots, u_n$ form a basis of $\mathcal{V}$. This completes the proof.

To prove the converse assume that there exist an orthonormal basis of $\mathcal{V}$ which consist of eigenvectors of T. That is, assume that there exists $u_1, \ldots, u_n \in \mathcal{V}$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that $\langle u_i, u_j \rangle = \delta_{ij}$ and $T u_j = \lambda_j u_j$ for all $i, j \in \{1, \ldots, n\}$.

Let $j \in \{1, \ldots, n\}$ be arbitrary. Since $u_1, \ldots, u_n$ form an orthonormal basis we have

$$\begin{aligned} T^* u_j &= \langle T^* u_j, u_1 \rangle\, u_1 + \langle T^* u_j, u_2 \rangle\, u_2 + \cdots + \langle T^* u_j, u_n \rangle\, u_n \\ &= \langle u_j, T u_1 \rangle\, u_1 + \langle u_j, T u_2 \rangle\, u_2 + \cdots + \langle u_j, T u_n \rangle\, u_n \\ &= \langle u_j, \lambda_1 u_1 \rangle\, u_1 + \langle u_j, \lambda_2 u_2 \rangle\, u_2 + \cdots + \langle u_j, \lambda_n u_n \rangle\, u_n \\ &= \overline{\lambda_1} \langle u_j, u_1 \rangle\, u_1 + \overline{\lambda_2} \langle u_j, u_2 \rangle\, u_2 + \cdots + \overline{\lambda_n} \langle u_j, u_n \rangle\, u_n \\ &= \overline{\lambda_j} u_j. \end{aligned}$$

Thus, $T^* u_j = \overline{\lambda_j} u_j$ for all $j \in \{1, \ldots, n\}$. Consequently,

$$T T^* u_j = T\bigl(\overline{\lambda_j} u_j\bigr) = \overline{\lambda_j} T u_j = \overline{\lambda_j} \lambda_j u_j = |\lambda_j|^2 u_j,$$

and also

$$T^*Tu_j = T^*\big(\lambda_j u_j\big) = \lambda_j T^* u_j = \lambda_j \overline{\lambda_j} u_j = |\lambda_j|^2 u_j.$$

Thus, $TT^*u_j = T^*Tu_j$ for all $j \in \{1, \ldots, n\}$. Since $u_1, \ldots, u_n$ form a basis of $\mathcal{V}$ this implies $TT^*v = T^*Tv$ for all $v \in \mathcal{V}$, that is, T is normal. $\square$