UC Irvine Math 121B Fall 2023
Linear Algebra II

Professor: Vladimir Baranovsky
Teaching Assistant: Fei Xiang
Notes: Timothy Cho

December 2023
Lecture Note Series #**3**

Introduction

These notes come from both the lecture and the discussion. Sections are numbered chronologically (not by content) using the following scheme by taking the section number modulo 10:

Date	Lecture	Discussion
Monday	0	1
Tuesday	2	3
Wednesday	4	5
Thursday	6	7
Friday	8	9

Additionally, the first digit (first two if the section number is three digits long) denotes the week that the lecture/discussion occurred in. It should be noted that not every lecture is recorded in these notes: some lectures were skipped, but despite this the notes should be comprehensible.

The text used was *Linear Algebra Done Right*, 3e, by Sheldon Axler. Numbers in [brackets] refer to sections in this text. Homework exercises assigned this term are found at the back of these notes.

10 Basic Notions (I)

We should be familiar with this theorem from Math 121A.

Theorem 10.1 (Rank-Nullity Theorem). *Let $T : V \to W$ be a linear map, where V is finite-dimensional. Then*

$$\dim V = \dim\ker T + \dim\operatorname{im} T.$$

Proof. Choose a basis $\{u_i\}_1^n$ of $\ker T$, and extend it to a basis $\{u_i\}_1^n \cup \{v_j\}_1^k$ of V. Set $\langle v_1, \ldots, v_k\rangle =: V'$, so that $V = V' \oplus \ker T$. By the First Isomorphism Theorem, $V' \cong \operatorname{im} T$, so the claim is proven by checking dimensions. $\square$

Now, take the basis of V as given in the above proof, and choose this basis of W:

$$\{Tv_i\}_1^k \cup \{w_j\}_{k+1}^{k+m},$$

so that $\dim W = k + m$. Then the matrix of T is

$$M = \begin{pmatrix} Tv_1 & Tv_2 & \cdots & Tv_k & 0 & 0 & \cdots & 0\end{pmatrix}.$$

Our main topic of study in this course is the *diagonalizability* of such matrices M: i.e., is there a “nice” basis such that T has a very simple matrix M? Of course, this is not always the case, so we give conditions to see when this is possible.

13 Basic Notions (II)

The following should be review from Math 121A. Let $\mathbb{F}$ denote a field.

Proposition 13.1. *Given a matrix $A \in M_{m\times n}(\mathbb{F})$, we have $\operatorname{rank}(A) \leq \min(m, n)$.*

Proof. Consider the associated transformation $T_A : \mathbb{F}^n \to \mathbb{F}^m$. $\square$

The following definitions will be used often.

Definition 13.2. Let $\{x_i\}_1^n$ be a set of vectors. We say that $\{x_i\}$ is *linearly independent* if

$$a_1x_1 + \cdots + a_nx_n = 0 \text{ implies } a_i = 0$$

for all $i \leq n$.

Definition 13.3. Let V be a vector space. A set $B \subseteq V$ is a *basis* of V is B is linearly independent and $\langle B \rangle = V$.

Example 13.4. Let $T : V \to W$, and define $G_T = \{(v, Tv) : v \in V\} \subseteq V \times W$. We claim that T is linear if and only if $G_T \leq V \times W$.

Proof. ($\implies$): Suppose T is linear. We verify the subspace criterion for G_T. We notice that if $(u, Tu), (v, Tv) \in G_T$, by linearity of T we have

$$(u, Tu) + (v, Tv) = (u + v, Tu + Tv) = (u + v, T(u + v)) \in G_T,$$

so G_T is additively closed. Note that $(0, T(0)) = (0, 0) \in G_T$, so G_T has an additive identity. Finally, by linearity of T we have for every $c \in \mathbb{F}$ and $(u, Tu) \in G_T$,

$$c(u, Tu) = (cu, Tu) = (cu, T(cu)) \in G_T,$$

so G_T is scalar-closed. Hence $G_T \leq V \times W$.

($\impliedby$): Suppose $G_T \leq V \times W$. Pick vectors $u, v \in V$. We note that $(u, Tu), (v, Tv) \in G_T$, so that by additive closure $(u, Tu) + (v, Tv) = (u + v, Tu + Tv) \in G_T$. By definition of G_T, we see that $T(u + v) = Tu + Tv$, so T is an additive group homomorphism.

Now, choose $\alpha \in \mathbb{F}$. Then by scalar closure, we have $(\alpha u, \alpha Tu) \in G_T$. But this implies $T(\alpha u) = \alpha Tu$, hence T is linear. $\square$

Here is an exercise in using the First Isomorphism Theorem.

Example 13.5. Let $\varphi : V \to \mathbb{F}$ be linear and suppose $\varphi \neq 0$. We show that $\dim(V / \ker \varphi) = 1$.

Proof. The First Isomorphism Theorem tells us that $V / \ker \varphi \cong \operatorname{im} \varphi$. Since $\varphi \neq 0$, φ must have rank 1. But $\dim_{\mathbb{F}} \mathbb{F} = 1$, so φ is surjective. This implies $\operatorname{im} \varphi = \mathbb{F}$, so $\dim(V / \ker \varphi) = \dim \mathbb{F} = 1$. $\square$

14 Eigenvalues and Eigenvectors

Let $T : V \to V$be a linear map with $\dim V < \infty$. Assume there exists a basis $\{e_i\}_1^n$ of V such that T has a diagonal matrix $A_T := \operatorname{diag}(a_1, \ldots, a_n)$; i.e., $Te_i = a_ie_i$ for all $i \leq n$. This is equivalent to saying $(T - a_iI)e_i = 0$, so we make the following definition.

Definition 14.1. Let $v \neq 0$ and let $T \in \mathcal{L}(V)$, where V has ground field $\mathbb{F}$. If there exists some $\lambda \in \mathbb{F}$ such that $Tv = \lambda v$, then λ is an *eigenvalue* of T, and v is a λ-*eigenvector* of T.

The following theorem thus follows immediately from our discussion above.

Theorem 14.2. *Let $T : V \to V$. Then T is diagonalizable if and only if V has a basis of eigenvectors of T.*

We recall the following facts about eigenvectors.

Proposition 14.3. *Let $v \in V$ and $\lambda \in \mathbb{F}$, and let $T \in \mathcal{L}(V)$. If v is a λ-eigenvector of T, then the following hold:*

1. *$v \in \ker(T - \lambda I)$;*
2. *$T - \lambda I$ is not injective;*
3. *$T - \lambda I$ is not surjective;*
4. *$T - \lambda I$ is not bijective.*

Remark 14.4. Note that eigenvectors may not exist. Take $V = \mathbb{R}^2$ over $\mathbb{R}$, and let $T \in \mathcal{L}(\mathbb{R}^2)$ be defined by the matrix

$$A_T = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

which is a rotation matrix with angle α. Then, T only has eigenvectors if $\alpha \in \pi\mathbb{Z}$.

Theorem 14.5. *If $T : V \to V$ is linear, and $\{\lambda_i\}_1^n$ are distinct eigenvalues with corresponding eigenvectors $\{v_i\}_1^n$, then $\{v_i\}_1^n$ is linearly independent.*

Proof. For contradiction, suppose that $\{v_i\}_1^n$ is linearly dependent. Let k be the smallest index such that $v_k \in \langle v_1, \dots, v_{k-1}\rangle$. This is equivalent to saying $v_k = a_1v_1 + \cdots + a_{k-1}v_{k-1}$ for some nontrivial constants a_j, $j \leq k-1$. Hence

$$Tv_k = a_1Tv_1 + \cdots + a_{k-1}Tv_{k-1} \iff \lambda_k v_k = a_1\lambda_1 v_1 + \cdots + a_{k-1}\lambda_{k-1}v_{k-1}.$$

This implies

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

but each $\lambda_k - \lambda_j$ is nonzero due to distinctness. But this violates the minimality of k if the a_j are indeed nontrivial, so the set $\{v_i\}_1^n$ must be linearly independent. □

We thus have some corollaries:

Corollary 14.6. *If $\dim V = n$, then $T \in \mathcal{L}(V)$ has at most n distinct eigenvalues.*

Corollary 14.7. *If $\dim V = n$ and $T \in \mathcal{L}(V)$ has n eigenvalues, then V has a basis of eigenvectors of T.*

17 Invariant Subspaces: Worked Examples

[Invariant subspaces are defined in Definition 18.1, but this discussion section was ahead of the lecture.]

Example 17.1. Let $T \in \mathcal{L}(V)$, and let $\{U_i\}_1^n$ be a set of subspaces of V invariant under T. Prove that $U_1 + \cdots + U_n$ is invariant under T.

Proof. From each U_i, pick a $u_i \in U_i$. Then $Tu_i \in U_i$ by invariance of each U_i, so that by linearity, $T\sum u_i = \sum Tu_i \in \sum U_i$, so we are done. □

Example 17.2. Define $T \in \mathcal{L}(\mathbb{R}^2)$ by $(x, y) \mapsto (y, x)$ Find all eigenvalues and eigenvectors of T.

Solution. Suppose $(x, y) \neq 0$ is an eigenvector of T with eigenvalue λ. Then $T(x, y) = \lambda(x, y) = (y, x)$, so we have the system of equations $y = \lambda x$, $x = \lambda y$, so $y = \lambda^2 y$. We consider two cases.

Case I: $y = 0$. In this case, we have $x = \lambda y = 0$, which is impossible as $(x, y) \neq 0$.

Case II: $y \neq 0$. In this case, we obtain $\lambda^2 = 1 \implies \lambda = \pm 1$. If $\lambda = 1$, we have $y = x$, so $\lambda = 1$ has the associated eigenvectors $\langle (1, 1) \rangle$. Similarly, $\lambda = -1$ has the associated eigenvectors $\langle (1, -1) \rangle$. •

Example 17.3. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ is such that every subspace of V with dimension $(\dim V - 1)$ is invariant under T. Prove that T is a scalar multiple of the identity operator I.

Proof. Assume for contradiction that $T \neq \lambda I$ for any $\lambda \in \mathbb{F}$. Then there exists some $v \in V \setminus 0$ such that Tv is not a multiple of v, i.e., $\{v, Tv\}$ is linearly independent. Hence, we extend this set to a basis β of V (letting $v =: v_1$ and $Tv =: v_2$): $\beta = \{v_i\}_1^n$, where $\dim V =: n$. Now $\langle v_1, v_3, v_4, \ldots, v_n \rangle$ is a subspace of dimension $n - 1$, but by linear independence of β, $Tv = v_2 \notin \langle v_1, v_3, v_4, \ldots, v_n \rangle$. But $v = v_1$, so $Tv_1 = Tv = v_2$, but we have a contradiction as by invariance we have $Tv_1 \in \langle v_1, v_3, v_4, \ldots, v_n \rangle$. Hence, $T = \lambda I$ for some $\lambda \in \mathbb{F}$. □

18 Invariant Subspaces (I)

We now define what it means for a subspace to be *invariant*.

Definition 18.1. Let $T \in \mathcal{L}(V)$. A subspace $W \leq V$ is *T-invariant* if $T(W) \leq W$; i.e., if $Tw \in W$ for all $w \in W$.

Clearly, 0 and V are T-invariant for any $T \in \mathcal{L}(V)$.

Example 18.2. Let $T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $\langle (1, 0) \rangle$ is T-invariant.

Suppose $W < V$ is T-invariant. Fix $W = \langle v_1, \ldots, v_k \rangle$, where the v_i form a basis of W. We extend this to a basis $\{v_i\}_{i=1}^n$ of V. Since $Tv_i \in W$ for all $i \leq k$, we know that $Tv_i \in \langle v_1, \ldots, v_k \rangle$, so the matrix of T is of this form:

$$T = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}.$$

We call matrices of this form *block-upper triangular*.

Definition 18.3. Let $T \in \mathcal{L}(V)$, and let $W \leq V$ be T-invariant. Then, we define the *restriction of T onto W* by $T|_W : W \to W$ by $w \mapsto Tw$ for all $w \in W$.

Definition 18.4. Let $T \in \mathcal{L}(V)$, and let $W \leq V$. We define the *quotient operator of T by W* by $T/W : V/W \to V/W$ by $v + W \mapsto Tv + W$ for all $v \in V$.

The next lemma is useful later.

Lemma 18.5. *Let $T \in \mathcal{L}(V)$. Then $\ker T$ and $\operatorname{im} T$ are invariant under T.*

Proof. First, take $v \in \ker T$. Then $Tv = 0$, but $0 \in \ker T$, so $\ker T$ is T-invariant. Now, take $v \in \operatorname{im} T$. Clearly, $Tv \in \operatorname{im} T$. □

It of course follows that if $T \in \mathcal{L}(V)$ has eigenvalue λ, then the eigenspace corresponding to λ, $\ker(T - \lambda I)$, is invariant. Next, we define an important class of linear transformations.

Definition 18.6. Let $T \in \mathcal{L}(V)$, and fix $f(x) = a_0 + a_1 x + \cdots + a_n x^n$. We define

$$f(T) := \sum_{i=0}^{n} a_i T^i,$$

where we agree to let $T^0 = I$, the identity operator.

We note that if $f, g \in \mathbb{F}[x]$, then we certainly have $f(T)g(T) = g(T)f(T)$; i.e., the ring $\mathbb{F}[T]$ is a commutative ring where addition is given by function addition, and composition is the ring multiplication.

Lemma 18.7. *Let $T \in \mathcal{L}(V)$, and let $f(x) \in \mathbb{F}[x]$. Then $\ker f(T)$ and $\operatorname{im} f(T)$ are T-invariant.*

Notice that this is *not* the same statement as Lemma 18.5, which would assert that $\ker f(T)$ and $\operatorname{im} f(T)$ are $f(T)$-invariant. We now look at the proof.

Proof. Pick $v \in \ker f(T)$. Then $f(T)v = 0$. Applying T to both sides, we see that $Tf(T)v = T(0) = 0$, but $0 = Tf(T)v = f(T)Tv = f(T)(Tv)$, so $Tv \in \ker f(T)$.

Similarly, let $v \in \operatorname{im} f(T)$. Then there exists some $u \in V$ with $f(T)u = v$. Applying T to both sides, we see that $Tv = Tf(T)u = f(T)Tu = f(T)(Tu) \in \operatorname{im} f(T)$. □

Theorem 18.8. *Let V be a vector space of dimension n over $\mathbb{C}$, and take $T \in \mathcal{L}(V)$. Then T has at least one eigenvalue.*

Proof. Take some $W < V$, and some $w \in W \setminus 0$. Since $\dim W < n$, the set $\{w, Tw, \ldots, T^n w\}$ is linearly dependent, so there exist some complex numbers a_j, $0 \le j \le n$ such that

$$0 = a_0 w + a_1 Tw + \cdots + a_n T^n w.$$

Letting $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{C}[x]$, we see that $0 = f(T)(w)$. By the Fundamental Theorem of Algebra, write

$$f(x) = a_n \prod_{j=1}^{n} (x - \lambda_j),$$

so that

$$f(T)(w) = a_n \prod_{j=1}^{n} (T - \lambda_j I)(w) = 0.$$

If each $(T - \lambda_j I)$ were invertible, we would have $f(T)(w) \neq 0$. Hence, there exists some $T - \lambda_k I$, uninvertible, so that $(T - \lambda_k I)w = 0$. Hence, λ_k is an eigenvalue for T. □

20 Invariant Subspaces (II)

Example 20.1. Take $T \in \mathcal{L}(\mathbb{C}^2)$ to be defined by the rotation matrix $T = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. Then we verify that if $(x, y) \neq 0$ is an eigenvector with eigenvalue λ,

$$T(x, y) = (x\cos\alpha - y\sin\alpha, x\sin\alpha + y\cos\alpha) = (\lambda x, \lambda y).$$

Solving this system yields $\lambda^2 - 2\lambda\cos\alpha + 1 = 0$, which is a quadratic in λ, so we recover the eigenvalues $\lambda_{1,2} = \cos\alpha \pm i\sin\alpha$.

Theorem 20.2. *Let V be a complex vector space with dimension n, and let $T \in \mathcal{L}(V)$. Then there exists a basis $\{v_i\}_1^n \subset V$ in which T has an upper-triangular matrix:*

$$M_T = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Proof. We prove by induction on the dimension n. If $n = 1$, then we are done. Assume that the theorem holds for all vector spaces with dimension at most $n-1$, $n \geq 2$. Suppose $\dim V = n$, so by Theorem 18.8, T has a λ_1-eigenvector $v_1 \neq 0$. Extend to a basis $\{v_i\}_{i=1}^n$ of V; the matrix of T under this basis has the form

$$A_T = \begin{pmatrix} \lambda_1 & * \\ 0 & B \end{pmatrix}.$$

Now, by linear independence, note that $V = \langle v_1 \rangle \oplus \langle v_2, \ldots, v_n \rangle$. If $U := \langle v_1 \rangle$, then we can see that T/U has matrix B in this basis. But $\dim(T/U) = n - 1$, so the inductive hypothesis applies, so there is a basis $\{w_i + U\}_{i=2}^n$ of V/U such that T/U has an upper triangular matrix. Now $V/U \cong \langle v_2, \ldots, v_m \rangle$, so using the basis $\{v_1\} \cup \{w_i\}_2^n$ finishes the proof. $\square$

23 Polynomial Operators: Worked Examples

Example 23.1. Let $S, T \in \mathcal{L}(V)$, and let S be invertible, and fix $p \in \mathbb{F}[x]$. Prove that $p(STS^{-1}) = Sp(T)S^{-1}$.

Proof. Choose an arbitrary $v \in V$, and write $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$. Now

$$p(STS^{-1})v = \sum_{i=0}^n a_i (STS^{-1})^i v,$$

but we note that $(STS^{-1})^i = ST^iS^{-1}$, so the above simplifies to (by linearity)

$$p(STS^{-1})v = \sum_{i=0}^n a_i ST^iS^{-1}v = \left[S \left(\sum_{i=0}^n a_i T^i \right) S^{-1} \right] v = Sp(T)S^{-1}v,$$

which completes the proof. $\square$

Example 23.2. Let $T \in \mathcal{L}(V)$, and take $p \in \mathbb{C}[x]$ and some $\alpha \in \mathbb{C}$. Prove that α is an eigenvalue of $p(T)$ if and only if $\alpha = p(\lambda)$ for some eigenvalue λ of T.

Proof. ($\Longleftarrow$): Suppose $\alpha = p(\lambda)$, where λ is an eigenvalue of T. Let v be a λ-eigenvector of T, so that $Tv = \lambda v$. By linearity, $p(T)v = \sum a_i T^i v = \sum a_i \lambda^i v = p(\lambda)v = \alpha v$, so v is an α-eigenvector of $p(T)$.

($\Longrightarrow$): Suppose α is an eigenvalue of $p(T)$. By algebraic closure of $\mathbb{C}$, write

$$p(x) - \alpha = c \prod_{i=1}^n (x - \lambda_i),$$

for some $c \in \mathbb{C}$, so that

$$p(T) - \alpha I = c \prod_{i=1}^{n} (T - \lambda_i I).$$

By a similar argument to the proof of Theorem 18.8, there exists one λ_j such that $(T - \lambda_j I)v = 0$, so $\lambda := \lambda_j$ is an eigenvalue for T. Now $p(\lambda) - \alpha = 0$ by construction, so $p(\lambda) = \alpha$. $\square$

Example 23.3. Let V have dimension $n > 1$. Let $T \in \mathcal{L}(V)$. Show that $\mathbb{F}[T] \neq \mathcal{L}(V)$.

Proof. Since V is n-dimensional, we see that for every $v \in V$, the set $\{v, Tv, \dots, T^n v\}$ is linearly dependent. Hence $T^n = a_0 I + a_1 T + \cdots + a_n T^n$ for some $a_i \in \mathbb{F}$. Hence $\dim \mathbb{F}[T] \leq n$, but $\dim \mathcal{L}(V) = n^2 > n$, as $n > 1$. Hence $\mathbb{F}[T] \neq \mathcal{L}(V)$. $\square$

24 Diagonalizability of Matrices

Let $\dim V = n$, and let $T \in \mathcal{L}(V)$. Assume that T has the upper-triangular matrix with the numbers λ_i on the diagonal. From here, it is not too hard to see that the λ_i are exactly the eigenvalues of T by an argument similar to that of Theorem 18.8. We state the following theorem.

Theorem 24.1. *Let $\dim V = n$, and $T \in \mathcal{L}(V)$. Then T is invertible if and only if 0 is not an eigenvalue for T.*

Proof. If T is not invertible, then there exists some $v \neq 0$ such that $Tv = 0 = 0v$. Hence v is a 0-eigenvalue for T. Conversely, if 0 is an eigenvalue for T, then $T - 0I = T$ is not invertible by Proposition 14.3. $\square$

Example 24.2. Consider the matrix $A = \begin{pmatrix} 1 & x & y \\ 0 & 2 & z \\ 0 & 0 & 3 \end{pmatrix} = \operatorname{diag}(1,2,3) \cdot \begin{pmatrix} 1 & x & y \\ 0 & 1 & z/2 \\ 0 & 0 & 1 \end{pmatrix}$. Reading off the diagonal, we see that 0 is not an eigenvalue, so A is invertible. Hence, we calculate its inverse. Note that

$$A^{-1} = \begin{pmatrix} 1 & x & y \\ 0 & 1 & z/2 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \operatorname{diag}\left(1, \frac{1}{2}, \frac{1}{3}\right), \text{ and}$$

$$\begin{pmatrix} 1 & x & y \\ 0 & 1 & z/2 \\ 0 & 0 & 1 \end{pmatrix} = I + \begin{pmatrix} 0 & x & y \\ 0 & 0 & z/2 \\ 0 & 0 & 0 \end{pmatrix} =: I + W,$$

so we can check that $(I + W)^{-1} = I - W + W^2$ (later on, we will see that this means that W has *nilpotency degree* 3). Hence $A^{-1} = \boxed{(I - W + W^2) \operatorname{diag}(1, 1/2, 1/3)}$.

From the discussion earlier, we pull out the following corollary.

Corollary 24.3. *If T has an upper-triangular matrix in some basis, then the eigenvalues of T are the diagonal entries.*

Now, let us define the notion of an *eigenspace*.

Definition 24.4. Let $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{C}$. Then the *λ-eigenspace of T* is the subspace $E(\lambda, T) := \ker(T - \lambda I)$.

Example 24.5. Suppose T has the matrix $\operatorname{diag}(8,5,5)$ in some basis. Then $E(8,T) = \langle(1,0,0)\rangle$ and $E(5,T) = \langle(0,1,0),(0,0,1)\rangle$. In contrast, $E(-33,T) = 0$.

Now, fix $T \in \mathcal{L}(V)$, and let T have distinct eigenvalues $\lambda_1,\ldots,\lambda_k$. Then each of the $E(\lambda_i,T)$, $i \leq k$, are nontrivial. We have the following lemma.

Lemma 24.6. *If T has distinct eigenvalues $\lambda_1,\ldots,\lambda_k$, then the sum $E(\lambda_1,T)+E(\lambda_2,T)+\cdots+E(\lambda_k,T)$ is direct.*

Proof. Suppose $v \in \sum v_i = \sum w_i$, where $v_i, w_i \in E(\lambda_i,T)$. Then $0 = \sum(v_i - w_i)$, so if $v_i - w_i \neq 0$, then $v_i - w_i$ is an λ_i-eigenvector of T. But eigenvectors with distinct eigenvalues are linearly independent, so each $v_i - w_i$ must be 0. Hence, the sum $v = \sum v_i$ is uniquely determined, so the sum of the spaces $E(\lambda_i,T)$ is direct. □

27 Eigenvalues and Matrices: Worked Examples

Example 27.1. Let $\dim V = n < \infty$, and let $T \in \mathcal{L}(V)$ have n distinct eigenvalues. Also, let $S \in \mathcal{L}(V)$ have the same eigen*vectors* as T. Show that S and T commute.

Proof. Let $\{\lambda_i\}_1^n$ be the distinct set of eigenvalues for T, and let $\{v_i\}_1^n$ be a set of λ_i-eigenvectors of T. Now, the v_i are eigenvalues for S as well, so to each v_i, let μ_i be its S-eigenvalue. [Notice that the μ_i are not necessarily distinct.]

Now, the v_i form an eigenbasis of T for V, so that if $v \in V$, we have $v = \sum a_i v_i$. Hence

$$STv = ST\sum_{i=1}^{n} a_i v_i = S\sum_{i=1}^{n} a_i\lambda_i v_i = \sum_{i=1}^{n} a_i\lambda_i\mu_i v_i,$$

and we get the same thing for TSv. Hence $ST = TS$. □

Example 27.2. Let $R, T \in \mathcal{L}(\mathbb{F}^3)$ have $2,6,7$ as eigenvalues. Prove that there exists an $S \in \mathcal{L}(\mathbb{F}^3)$ such that $R = S^{-1}TS$.

Proof. Since $R, T \in \mathcal{L}(\mathbb{F}^3)$ and each has 3 distinct eigenvalues, we have the eigenbases $\{r_i\}_1^3$ (for R) and $\{t_i\}_1^3$ (for T), where each r_i, t_i has eigenvalue λ_i. Let S be the *change of basis* mapping $r_i \mapsto t_i$. We claim that S is injective (hence invertible): if $v = a_1r_1 + a_2r_2 + a_3r_3 \in \ker S$, then $Sv = 0$ forces $a_1 = a_2 = a_3 = 0$, so $v = 0$. Hence S is invertible. Now, fix $v = b_1r_1 + b_2r_2 + b_3r_3 \in V$. We can verify that indeed $S^{-1}TSv = Rv$. □

Example 27.3. Let $R, T \in \mathcal{L}(\mathbb{F}^4)$ have $2,6,7$ as their *only* eigenvalues. Does there necessarily exist some $S \in \mathcal{L}(\mathbb{F}^4)$ such that $R = S^{-1}TS$?

Solution. No: take $T = \operatorname{diag}(2,6,7,2)$, and $R = \operatorname{diag}(2,6,7,6)$. •

28 Conditions for Diagonalizability

From Lemma 24.6, we have the following corollary.

Corollary 28.1. *Let V be a vector space, and let $T \in \mathcal{L}(V)$ have $\lambda_1,\ldots,\lambda_m$ as eigenvalues. Then $\dim E(\lambda_1,T)+\cdots+\dim E(\lambda_m,T) \leq \dim V$.*

We also have the following important theorem.

Theorem 28.2 (Diagonalizability Conditions)**.** *Let $T \in \mathcal{L}(V)$ and let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, where V has dimension n. Then the following are equivalent:*

1. *T is diagonalizable.*
2. *V has a basis of eigenvectors of T.*
3. *There exist one-dimensional subspaces $U_1, \ldots, U_n \leq V$, all T-invariant, such that*
$$V = U_1 \oplus \cdots \oplus U_n.$$
4. *We have $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.*
5. *We have $n = \dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$.*

Proof. We leave $(1) \iff (2)$ as an easy verification.

$(2) \implies (3)$: Assume that (2) holds, and set $\{v_i\}_1^n$ to be an eigenbasis of T for V. Now, let $U_i := \langle v_i \rangle$. Of course, each U_i is T-invariant: if $u \in U_i$, then $u = av_i$ for some $a \in \mathbb{F}$. Hence $Tu = T(av_i) = aTv_i = a\lambda_i v_i \in U_i$. That $V = U_1 \oplus \cdots \oplus U_n$ ollows from the fact that $\langle v_i \rangle$ is a basis.

$(3) \implies (2)$: Suppose $V = U_1 \oplus \cdots \oplus U_n$, where the U_i are all one-dimensional subspaces of V as well as T-invariant. Then $U_i = \langle v_i \rangle$ for some $v_i \in V$. By T-invariance, we verify $Tv_i \in U_i \implies Tv_i = \lambda_i v_i$ for some $\lambda_i \in \mathbb{F}$. But this means that the λ_i are eigenvalues of T, which means that the v_i form an eigenbasis of T for V.

$(4) \implies (5)$ follows directly from Corollary 28.1.

$(5) \implies (2)$: Assume (5) holds. Chose a basis in $E(\lambda_i, T)$ for each $i \leq n$, so T has matrix $\lambda_i I$ when restricted to that (invariant) subspace. Combining the bases, we get a linearly independent basis of V by assuming (5), so we are done after combining the $\lambda_i I$'s. ☐

Corollary 28.3. *Let $\dim V = n$. If $T \in \mathcal{L}(V)$ has n distinct eigenvalues, then T is diagonalizable.*

Example 28.4. Let T have the matrix $M_T := \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}$. This matrix has the eigenvalues $2, 5, 8$, so it is diagonalizable.

30 Inner Products

In this section, we will assume that V is a vector space over $\mathbb{F}$, where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.

Definition 30.1. Let V be a vector space. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is an *inner product* if it satisfies these axioms for all $v, v_1, v_2, w \in V$ and $\alpha \in \mathbb{F}$:

1. $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$,
2. $\langle \alpha v, w \rangle = \alpha \langle v, w \rangle$,
3. $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ if and only if $v = 0$,
4. $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

When a vector space V is equipped with an inner product, we say that V is an *inner product space*.

We should note that over $\mathbb{R}$, the conjugation is inconsequential, and so the inner product over $\mathbb{R}$ is commutative. Also, property (3) suggests a way to measure a vector's "length" in V.

Definition 30.2. Let V be an inner product space. Then the *norm* of a vector is the function $\|\cdot\| : V \to \mathbb{F}$ given by $\|v\| := \sqrt{\langle v, v\rangle}$.

Here are some basic properties of the inner product.

Proposition 30.3. *Let V be an inner product space. Then the following hold for all $u, v, w \in V$ and $\beta \in \mathbb{F}$:*

1. $\langle v, \beta w\rangle = \bar{\beta}\langle v, w\rangle$,
2. $\langle 0, w\rangle = 0 = \langle v, 0\rangle$.
3. $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$.

Proof. (1): We have $\langle v, \beta w\rangle = \overline{\langle \beta w, v\rangle} = \overline{\beta \langle w, v\rangle} = \bar{\beta}\overline{\langle w, v\rangle} = \bar{\beta}\langle v, w\rangle$.

(2): Write $\langle 0, w\rangle = \langle v - v, w\rangle = \langle v, w\rangle - \langle v, w\rangle = 0$.

(3): We have $\langle u, v + w\rangle = \overline{\langle v + w, u\rangle} = \overline{\langle v, u\rangle + \langle w, u\rangle} = \langle u, v\rangle + \langle u, w\rangle$. □

From this, we get two substantial consequences.

Theorem 30.4 (Pythagorean Theorem)**.** *If $\langle u, v\rangle = 0$, then $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.*

Proof. Expand $\|u, v\|^2 = \langle u + v, u + v\rangle = \langle u, u\rangle + \langle v, u\rangle + \langle u, v\rangle + \langle v, v\rangle$. Now, the "cross terms" vanish by assumption, so $\|u, v\|^2 = \langle u, u\rangle + \langle v, v\rangle = \|u\|^2 + \|v\|^2$. □

Theorem 30.5 (Cauchy-Schwarz Inequality)**.** *For every $u, v \in V$, where V is an inner product space, we have $|\langle u, v\rangle| \le \|u\| \cdot \|v\|$.*

Proof. Write an *orthogonal decomposition* of u:

$$u = \frac{\langle u, v\rangle}{\|v\|^2} v + w,$$

for some $w \in V$. Now $\langle u, v\rangle = \dfrac{\langle u, v\rangle}{\|v\|^2} \cdot \langle v, v\rangle + \langle w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$, so we see that $\langle w, v\rangle = 0$. Applying the Pythagorean Theorem, we see

$$\|u\|^2 = \left\| \frac{\langle u, v\rangle}{\|v\|^2} v \right\|^2 + \|w\|^2 = \left| \frac{\langle u, v\rangle}{\|v\|^2} \right|^2 \|v\|^2 + \|w\|^2 \ge \frac{|\langle u, v\rangle|^2}{\|v\|^4} \|v\|^2 = \frac{|\langle u, v\rangle|^2}{\|v\|^2}.$$

Now solving and taking square roots finishes the proof. □

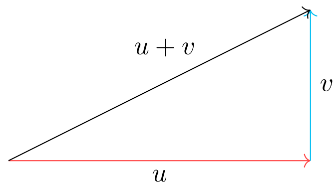
In the case where V is an inner product space over $\mathbb{R}$, we observe $1 \ge \left| \dfrac{\langle u, v\rangle}{\|u\| \, \|v\|} \right| =: x$, so we can view $x = \cos\varphi$ for some $\varphi \in \mathbb{R}$.

Example 30.6. If $\{x_i\}_1^n, \{y_i\}_1^n \subset \mathbb{R}$, then the Cauchy-Schwarz inequality implies (with the dot product) that

$$\left| \sum_{i=1}^n x_i y_i \right|^2 \le \left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{i=1}^n y_i^2 \right).$$

Finally, we introduce one more theorem, which makes sense when we draw a diagram.

Theorem 30.7 (Triangle Inequality)**.** *Let V be an inner product space. Then for any $u, v \in V$, we have $\|u+v\| \le \|u\| + \|v\|$.*



Proof. Recall that $\|u+v\|^2 = \langle u, u\rangle + \langle v, v\rangle + \langle u, v\rangle + \overline{\langle u, v\rangle} = \|u\|^2 + \|v\|^2 + 2\operatorname{Re}\langle u, v\rangle$. Now

$$\begin{aligned}\|u\|^2 + \|v\|^2 + 2\operatorname{Re}\langle u, v\rangle &\le \|u\|^2 + \|v\|^2 + 2\left|\langle u, v\rangle\right| \\ &\le \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \text{ (by Cauchy-Schwarz)} \\ &= (\|u\| + \|v\|)^2 .\end{aligned}$$

Taking square roots finishes the proof. □

33 Inner Products: Worked Examples

In this section, we will assume V is an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Example 33.1. Let $u, v \in V$ with $\|u\| = 3$, $\|u+v\| = 4$, and $\|u-v\| = 6$. What is $\|v\|$?

Solution. We have $\|u+v\|^2 = \|u\|^2 + \|v\|^2 + \langle u, v\rangle + \langle v, u\rangle$, and similarly $\|u-v\|^2 = \|u\|^2 + \|v\|^2 - (\langle u, v\rangle + \langle v, u\rangle)$. Hence, combining gives

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2 \implies 36 + 16 = 2(9) + 2\|v\|^2 ,$$

so that $\|v\| = \boxed{\sqrt{17}}$. •

Example 33.2. Prove that if $a_i, b_i \in \mathbb{R}$, we have $\left(\sum_{i=1}^n a_i b_i\right)^2 \le \left(\sum_{i=1}^n i a_i^2\right)\left(\sum_{i=1}^n \frac{b_i^2}{i}\right)$.

Proof. Write $\left(\sum_{i=1}^n a_i b_i\right)^2 = \left(\sum_{i=1}^n \sqrt{i} a_i \frac{b_i}{\sqrt{i}}\right)^2$. Now, the Cauchy-Schwarz inequality implies

$$\left(\sum a_i b_i\right)^2 \le \left(\sum (\sqrt{i} a_i)^2\right)\left(\sum \left(\frac{b_i}{\sqrt{i}}\right)^2\right) = \left(\sum i a_i^2\right)\left(\sum \frac{b_i^2}{i}\right),$$

which completes the proof. □

Example 33.3. Let $u, v \in V$. Prove that $\langle u, v\rangle = 0$ if and only if $\|u\| \le \|u + av\|$ for every $a \in \mathbb{F}$.

Proof. ($\implies$): Since $\langle u, v\rangle = 0$, we apply Pythagoras and write

$$\|u\|^2 \le \|u\|^2 + \|av\|^2 = \|u + av\|^2 ,$$

which completes this implication.

($\Longleftarrow$): Suppose $\|u\| \leq \|u + av\|$ for every $a \in \mathbb{F}$. Write

$$\begin{aligned}\|u\|^2 \leq \|u+av\|^2 = \langle u+av, u+av\rangle = \|u\|^2 + |a|^2\|v\|^2 + \langle u, av\rangle + \langle av, u\rangle \\ \implies \|u\|^2 \leq \|u\|^2 + |a|^2\|v\|^2 + 2\operatorname{Re}(\bar{a}\langle u, v\rangle) \\ \implies |a|^2\|v\|^2 \geq -2\operatorname{Re}(\bar{a}\langle u, v\rangle).\end{aligned}$$

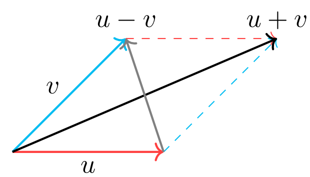
Since this holds for all $a \in \mathbb{F}$, we can set $a = -\langle u, v\rangle / \|v\|^2$ (the *projection length*), so that $2|\langle u, v\rangle|^2 \leq |\langle u, v\rangle|^2$, implying $\langle u, v\rangle = 0$. □

34 Inner Product Spaces and Orthonormal Bases

Again, we will assume that V is an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Theorem 34.1 (Parallelogram Law)**.** *For all $u, v \in V$, we have*

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



Proof. We expand by the definition of the norm:

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

which completes the proof. □

[6B] Orthogonal Bases

Definition 34.2. Let V be an inner product space. We say that $u, v \in V$ are *orthogonal* if $\langle u, v\rangle = 0$.

Definition 34.3. We say that $\{e_i\}_1^k$ is an *orthonormal system* of vectors if $\langle e_i, e_j\rangle = 0$ if $i \neq j$, and $\|e_i\| = 1$ for all $i \leq k$.

That is, a list of vectors is orthonormal if they all have the same length 1, and they are all mutually orthogonal to each other.

Example 34.4. Let $V = \mathbb{R}^3$. Certainly, $\{(1,0,0), (0,1,0), (0,0,1)\}$ is a list of orthonormal vectors under the standard dot product, but we can check that

$$\left\{\frac{1}{\sqrt{3}}(1,1,1), \frac{1}{\sqrt{2}}(-1,1,0), \frac{1}{\sqrt{6}}(1,1,-2)\right\}$$

is an orthonormal list as well.

Orthonormality plays nicely with linear independence.

Lemma 34.5. *An orthonormal system of vectors is linearly independent.*

Proof. Suppose that $\{e_i\}_1^k \subset V$ is an orthonormal system, and suppose there exist scalars $a_i \in \mathbb{F}$ with $\sum a_i e_i = 0$. Then for any $j \leq n$, we have

$$0 = \left\langle \sum_{i=1}^{n} a_i e_i, e_j \right\rangle = \sum_{i=1}^{n} \langle a_i e_i, e_j \rangle = \sum_{i=1}^{n} a_i \langle e_i, e_j \rangle,$$

and by orthonormality, this simplifies down to $0 = a_j \langle e_j, e_j \rangle = a_j \cdot 1 = 0$. Hence, our list is linearly independent. □

Lemma 34.6. *If $\dim V =: n$ and $\{e_i\}_1^n$ is an orthonormal basis, then for any $v \in V$, we have*

$$v = \sum_{i=1}^{n} \langle v, e_i \rangle e_i.$$

This lemma tells us that the decomposition of a vector within an orthonormal basis is very easy to find. The following is a consequence:

Proposition 34.7. *Let $\{e_i\}_1^n$ be an orthonormal basis of V, and write $v = \sum a_i e_i$. Then*

$$\|v\|^2 = |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2.$$

Proof. By Lemma 34.6, we have $a_i = \langle v, e_i \rangle$. Then, expand $\|v\|^2 = \langle v, v \rangle$ and simplify: all of the cross-terms disappear. □

We view one example.

Example 34.8. Let $\{e_i\}_1^m \subset V$ be an orthonormal system, and fix $v \in V$. Show that $\|v\|^2 = \sum_i |\langle v, e_i \rangle|^2$ if and only if $v \in \operatorname{span}(e_1, \ldots, e_m)$.

Proof. ($\Longleftarrow$): Suppose $v \in \operatorname{span}(e_1, \ldots, e_m) =: W$, so that $\{e_i\}_1^m$ is an orthonormal *basis* of W. Then we immediately have

$$v = \langle v, e_1 \rangle e_1 + \langle v, e_2 \rangle e_2 + \cdots + \langle v, e_m \rangle e_m.$$

Hence $\|v\|^2 = \sum_i |\langle v, e_i \rangle|^2$ by Proposition 34.7. Note that these steps are reversible, so the ($\Longrightarrow$) direction is completed as well. □

38 Orthonormal Bases

In this section, we will show that orthonormal bases exist in any inner product space V via induction. Take a linearly independent set $\{v_i\}_1^m$. We *normalize* v_1: define $e_1 := \alpha v_1$, where $\alpha = 1/\|v_1\|$. Hence $\|e_1\| = 1$, so the set $\{e_1\}$ is orthonormal. Now, $\{v_1, v_2\}$ is linearly independent, and thus so is $\{e_1, v_2\}$, but now we will find some e_2 which is orthonormal to e_1.

Letting $e_2 = \alpha e_1 + \beta v_2$, we must have

$$0 = \langle e_2, e_1 \rangle = \langle \alpha e_1 + \beta v_2, e_1 \rangle = \alpha \langle e_1, e_1 \rangle + \beta \langle v_2 e_1 \rangle = \alpha + \beta \langle v_2, e_1 \rangle.$$

Furthermore, we have $1 = \langle e_2, e_2 \rangle = \langle \alpha e_1 + \beta v_2, \alpha e_1 + \beta v_2 \rangle$. Setting $\beta = 1$ to simplify, we have

$$e_2 = \frac{v_2 - \langle v_2, e_1 \rangle e_1}{\|v_2 - \langle v_2, e_1 \rangle e_1\|}.$$

We can, of course, continue this inductively: this is the idea of the Gram-Schmidt Procedure.

Theorem 38.1 (Gram-Schmidt Procedure)**.** *Let $\{v_j\}_1^m$ be a linearly independent set, and let $e_1 := v_1/\|v_1\|$. For $j \geq 2$, recursively define*

$$u_j := v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1} \text{ and } e_j := \frac{u_j}{\|u_j\|}.$$

Then $\{e_j\}_1^m$ is an orthonormal list of vectors with $\operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(e_1, \ldots, e_m)$.

Example 38.2. Find an orthonormal basis of $\mathcal{P}_2(\mathbb{R})$ with the inner product $\langle p, q\rangle = \int_{-1}^{1} pq\, dx$.

Solution. We apply Gram-Schmidt (GS) to the standard basis $\{1, x, x^2\}$. We see $\|1\|^2 = 2$, so $e_1 = 1/\sqrt{2}$. Now,

$$u_2 = x - \langle x, e_1\rangle e_1 = x - \frac{1}{\sqrt{2}} \int_{-1}^{1} x \frac{1}{\sqrt{2}}\, dx = x,$$

and we have $\|x\|^2 = 2/3$, so $e_2 = \sqrt{\frac{3}{2}}x$. Finally, we have

$$u_3 = x^2 - \langle x^2, e_1\rangle e_1 - \langle x^2, e_2\rangle e_2$$

$$= x^2 - \left(\int_{-1}^{1} x^2 \sqrt{\frac{1}{2}}\, dx\right)\sqrt{\frac{1}{2}} - \cancel{\left(\int_{-1}^{1} x^2 \sqrt{\frac{3}{2}} x\, dx\right)} x\sqrt{\frac{3}{2}} = x^2 - \frac{1}{3}.$$

Now $\|u_3\|^2 = 8/45$, so $e_3 = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)$. Hence, an orthonormal basis of $\mathcal{P}_2(\mathbb{R})$ is

$$\boxed{\left\{\sqrt{\frac{1}{2}}, x\sqrt{\frac{3}{2}}, \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)\right\}}.$$

•

We have the following corollaries of the GS procedure.

Corollary 38.3. *If V is a finite-dimensional inner product space, then V has an orthonormal basis.*

Proof. Take a basis of V and apply GS. □

Corollary 38.4. *Let $\dim V = n$ and let $\{e_i\}_1^n$ be an orthonormal basis of V. If $v = \sum_i \alpha_i e_i$ and $u = \sum_i \beta_i e_i$, then $\langle v, u\rangle = \sum_i \alpha_i \overline{\beta_i}$.*

These lemmas are also important.

Lemma 38.5. *Let $\dim V =: n$, and let $\varepsilon := \{e_i\}_1^k \subset V$ be an orthonormal system. Then ε can be extended to an orthonormal basis $\{e_i\}_1^n$.*

Proof. Extend ε to a basis $\varepsilon \cup \{f_j\}_{k+1}^n$. Applying GS to this basis, we see that the vectors in ε remain unchanged, while the vectors f_j, $j \geq k+1$, are converted to orthonormal e_j. Hence $\{e_i\}_1^n$ is an orthonormal basis of V. □

Lemma 38.6. *Let V be a complex finite-dimensional inner product space, and let $T \in \mathcal{L}(V)$. Then there exists an orthonormal basis in which T has an upper triangular matrix.*

Proof. Choose a basis $\{v_i\}_1^n$ such that the matrix of T is upper triangular, and apply GS to obtain a corresponding $\{e_i\}_1^n$. Now $Te_1 = (Tv_1)/\|v_1\|$, but because the matrix of T is upper-triangular, we conclude that $Te_1 = \lambda_1 e_1$ for some $\lambda_1 \in \mathbb{C}$. This continues inductively: each e_i is not dependent on the e_j's with $j > i$. Hence, the matrix of T with respect to $\{e_i\}_1^n$ is upper triangular. □

40 Linear Functionals and Orthogonal Complements

As usual, let V denote an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

[6B] Linear Functionals

Here is one important application of inner products.

Definition 40.1. A *linear functional* on V is a linear map $\varphi : V \to \mathbb{F}$.

The next theorem tells us that linear functionals are completely determined by the inner product on V.

Theorem 40.2 (Riesz Representation Theorem)**.** *Any linear functional $\varphi : V \to \mathbb{F}$ is of the form $\varphi(v) = \langle v, u\rangle$ for a unique vector $u \in V$.*

Proof. Choose an orthonormal basis $\{e_i\}_1^n \subset V$, and let $v \in V$. Then $v = \sum_i a_i e_i$, so if $\varphi : V \to \mathbb{F}$ is a linear functional,

$$\varphi(v) = \varphi\left(\sum_{i=1}^n a_i e_i\right) = \sum_{i=1}^n \langle v, e_i\rangle\, \varphi(e_i),$$

where we know $a_i = \langle v, e_i\rangle$ from Lemma 34.6. We can continue to expand by properties of the inner product:

$$\varphi(v) = \sum_{i=1}^n \langle v, e_i\rangle\, \varphi(e_i) = \sum_{i=1}^n \left\langle v, \overline{\varphi(e_i)} e_i \right\rangle = \left\langle v, \sum_{i=1}^n \overline{\varphi(e_i)} e_i \right\rangle.$$

Now, let $u = \sum_i \overline{\varphi(e_i)} e_i$, which proves existence. Note that u is unique as φ is completely determined by its values on the basis $\{e_i\}_1^n$: if for all $v \in V$, we had $\langle v, u\rangle = \langle v, u'\rangle = \varphi(v)$ for some $u' \in V$, then $\langle v, u - u'\rangle = 0$. In particular, setting $v = u - u'$ gives $\|u - u'\| = 0 \iff u = u'$, so u is indeed unique. □

Example 40.3. Let $\varphi : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ be any linear functional. Then every φ has the form

$$\varphi(f(x)) = \int_{-1}^{1} f(x) g(x)\, dx$$

for a fixed $g(x) \in \mathcal{P}_2(\mathbb{R})$. This is because the integral is a well-defined inner product.

[6C] Orthogonal Complements

Just as we can discuss orthogonal lists of vectors, we can discuss orthogonal *spaces*.

Definition 40.4. Let $U \subseteq V$. The *orthogonal complement* of U is the set

$$U^\perp := \{v \in V : \langle v, u\rangle = 0 \text{ for all } u \in U\}.$$

Although U was just a subset, $U^\perp$ turns out to be a *subspace*.

Proposition 40.5. *For any $U \subseteq V$, we have $U^\perp \leq V$.*

Proof. Use the properties of the inner product. $\square$

The orthogonal complement satisfies these basic properties.

Proposition 40.6. *Let $U, W \subseteq V$. Then the following hold:*

1. $\{0\}^\perp = V$,
2. $V^\perp = \{0\}$,
3. $U \cap U^\perp = \{0\}$ *or* $\varnothing$, *and*
4. $U \subseteq V$ *implies* $W^\perp \subseteq U^\perp$.

We remark that it is **not true** that $U \cup U^\perp = V$! But we have a useful alternative to this that *is* true.

Theorem 40.7. *Let $U \leq V$ be a subspace. Then $V = U \oplus U^\perp$.*

Proof. Since $U \leq V$, we have $0 \in U$ so $U \cap U^\perp = \{0\}$ by Proposition 40.6(3). Thus, it suffices to show $V = U + U^\perp$. Choose an orthonormal basis $\{e_i\}_1^k$ of U, and fix $v \in V$. Then

$$v = w + \sum_{i=1}^{k} \langle v, e_i \rangle e_i$$

for some $w \in V$. Letting $u = \sum_i \langle v, e_i \rangle e_i$, we see $u \in U$, so we show $w \in U^\perp$. Now, for each e_i, $i \leq k$,

$$\langle w, e_i \rangle = \langle v - u, e_i \rangle = \langle v, e_i \rangle - \langle u, e_i \rangle = \langle v, e_i \rangle - \langle v, e_i \rangle = 0$$

by orthogonality. Hence $w \perp e_i$ for every $i \leq k$, so $w \in U^\perp$. $\square$

44 Complements and Adjoints

As usual, let V be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

[6C] Orthogonal Complements

Definition 44.1. Let $U \leq V$, so that $V = U \oplus U^\perp$ by Theorem 40.7. We define the *orthogonal projection* P_U by $P_U(v) := u$, where $v = u + w$, $u \in U$, $w \in U^\perp$.

Proposition 44.2 (Properties of the Orthogonal Projection)**.** *Let $U \leq V$, and let P_U be the orthogonal projection for U. Then for all $v \in V$:*

1. $P_U : V \to V$ *is linear,*
2. $\mathrm{im}(P_U) = U$,
3. $\ker(P_U) = U^\perp$,
4. $v - P_U(v) \in U^\perp$,
5. $P_U^2 = P_U$,
6. $\|P_U(v)\| \leq \|v\|$,
7. *If U has an orthonormal basis $\{e_i\}_1^m$, then $P_U(v) = \sum_i \langle v, e_i \rangle e_i$.*

Proof. (1) through (5) are fairly obvious. For (6) and (7), choose orthonormal bases $\langle e_i \rangle_1^m$ for U and $\langle e_i \rangle_{m+1}^n$ for $U^\perp$, so that $\langle e_i \rangle_1^n$ is an orthonormal basis[1] for V. Hence

$$v = \sum_{i=1}^n \langle v, e_i \rangle e_i = \sum_{i=1}^m \langle v, e_i \rangle e_i + \sum_{j=m+1}^n \langle v, e_j \rangle e_j =: u + w.$$

Hence $P_U(v) = u = \sum_{i=1}^m \langle v, e_i \rangle e_i$, proving (7). Now

$$\|v\|^2 = \sum_{i=1}^n |\langle v, e_i \rangle|^2 \geq \sum_{i=1}^m |\langle v, e_i \rangle|^2 ,$$

so we have proven (6). □

[7A] Adjoint Operators

Now, we define a type of operator closely related to the orthogonal complement.

Definition 44.3. Let V, W be inner product spaces, and let $T : V \to W$ be linear. Then the *adjoint* of T is a function $T^* : W \to V$ satisfying $\langle Tv, w \rangle = \langle v, T^* w \rangle$.

Example 44.4. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be given by $(x_1, x_2, x_3) \mapsto (x_2 + 3x_3, 2x_1)$. We compute T^*. Under the standard dot product (in $\mathbb{R}^n$), we write (for $x \in \mathbb{R}^3, y \in \mathbb{R}^2$)

$$Tx \cdot y = x \cdot T^* y,$$

so that we have

$$(x_1, x_2, x_3) \cdot T^*(y_1, y_2) = (x_2 + 3x_3, 2x_1) \cdot (y_1, y_2) = x_2 y_1 + 3x_3 y_1 + 2x_1 y_2.$$

But $x_2 y_1 + 3x_3 y_1 + 2x_1 y_2 = (x_1, x_2, x_3) \cdot (2x_2, y_1, 3y_1)$, so we have $T^*(y_1, y_2) = (2y_2, y_1, 3y_1)$.

Lemma 44.5. *If $T : V \to W$ is a linear map, T^* exists uniquely and is also a linear map.*

Proof. Fix a linear map $T : V \to W$, and define a linear functional $f : V \to \mathbb{F}$ by $v \mapsto \langle Tv, w \rangle \in \mathbb{F}$. By the Riesz Representation Theorem, there exists a *unique* $u \in V$ such that $f(v) = \langle v, u \rangle$. We claim $T^* w = u$. We leave it for the reader to check uniqueness and that T^* is linear. □

47 Basic Properties of the Adjoint

Example 47.1. Let $\mathcal{P}_2(\mathbb{R})$ have the inner product $\langle p, q \rangle := \int_0^1 pq \, dx$. Define $T \in \mathcal{L}(\mathcal{P}_2(\mathbb{R}))$ by $a_0 = a_1 x + a_2 x^2 \mapsto a_1 x$. Is T *self-adjoint*, i.e., does $T = T^*$?

Solution. If T is self-adjoint, then we have $\langle Tv, w \rangle = \langle v, Tw \rangle$. Taking $v = 1$ and $w = x$, we verify that

$$\langle Tv, w \rangle = \langle 0, x \rangle = 0, \text{ but}$$

$$\langle v, Tw \rangle = \langle 1, x \rangle = \int_0^1 x \, dx = \frac{1}{2} \neq 0.$$

Hence, T is *not* self-adjoint. •

[1] Here, we used the fact that $U \perp U^\perp$.

We also introduce the following definition.

Definition 47.2. Let $T \in \mathcal{L}(V)$. Then T is *normal* if $TT^* = T^*T$.

Proposition 47.3. *Let $S, T \in \mathcal{L}(V, W)$, and $\lambda \in \mathbb{F}$. Then the following hold:*

1. $(S+T)^* = S^* + T^*$,
2. $(ST)^* = T^*S^*$,
3. $(T^*)^* = T$,
4. $(\lambda T)^* = \bar{\lambda} T^*$,
5. $I^* = I$,
6. *If β and γ are orthonormal bases of V resp. W and T has matrix A with respect those bases, then T^* has matrix $\overline{A^t}$ with respect to the same bases.*

We will prove (1), and leave the rest to the reader.

Proof. (1): We proceed by definition of the adjoint: $(S+T)^*$ must satisfy for all $v \in V$, $w \in W$

$$\langle Sv, w\rangle + \langle Tv, w\rangle = \langle (S+T)v, w\rangle = \langle v, (S+T)^* w\rangle .$$

But notice $\langle v, S^*w\rangle + \langle v, T^*w\rangle = \langle v, S^*w + T^*w\rangle$, and

$$\langle v, S^*w\rangle + \langle v, T^*w\rangle = \langle Sv, w\rangle + \langle Tv, w\rangle = \langle (S+T)v, w\rangle .$$

By uniqueness of the adjoint, we are done. □

58 Normal Operators

As usual, let V be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and let $T \in \mathcal{L}(V)$. We will study 3 special types of operators.

Definition 58.1. T is *self-adjoint* if $T^* = T$.

Definition 58.2. T is *skew-adjoint* if $T^* = -T$.

Definition 58.3. T is an *isometry* if $T^* = T^{-1}$. In particular, if $\mathbb{F} = \mathbb{C}$, we say that T is *unitary*.

In all three cases, we see that by Definition 47.2, T is normal. For a demonstration of the importance of these operators, let T be self-adjoint, and let $v \neq 0$ be a λ-eigenvector of T. Then

$$\langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda \|v\|^2 , \text{ but}$$
$$\langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \bar{\lambda} \|v\|^2 .$$

Hence $\lambda = \bar{\lambda}$, i.e., λ is real. We have just established that self-adjoint operators only have real eigenvalues.

Here is an important lemma that will come up later.

Lemma 58.4. *Let V be an inner product space over $\mathbb{C}$. Then if $\langle Tv, v\rangle = 0$ for all $v \in V$, then $T = 0$.*

The proof of this lemma is not enlightening, so we will skip it.

60 Spectral Theorems

In this section, our goal is to show that certain types of normal operators are guaranteed to be diagonalizable. As usual, let V be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and let $T \in \mathcal{L}(V)$. We start with a theorem.

Theorem 60.1. *T is normal if and only if $\|Tv\| = \|T^*v\|$.*

Proof. Let T be normal. Then $\langle (T^*T - TT^*)v, v\rangle = 0$ as $T^*T - TT^* = 0$, but this means $\langle T^*Tv, v\rangle = \langle TT^*v, v\rangle$, so by flipping the operators around, we get $\langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle$. □

Theorem 60.2. *Let T be normal. If $v \neq 0$ is an λ-eigenvector of T, then v is an $\bar{\lambda}$-eigenvector of T^*.*

Proof. Fix v as in the statement. Then $(T - \lambda I)v = 0$, so $\|(T - \lambda I)v\| = 0$, so we have

$$0 = \langle (T - \lambda I)v, (T - \lambda I)v\rangle = \left\langle v, (T^* - \bar{\lambda} I)(T - \lambda I)v\right\rangle .$$

By normality of T, write

$$0 = \left\langle v, (T^*T - \bar{\lambda}T + \lambda T^* + \lambda\bar{\lambda} I)v\right\rangle = \left\langle (T^* - \bar{\lambda} I)v, (T^* - \bar{\lambda} I)v\right\rangle .$$

But this implies $\left\|T^* - \bar{\lambda} I\right\| v = 0$, so that v is an $\bar{\lambda}$-eigenvector of T. □

Our next theorem relates normality to orthogonality.

Theorem 60.3. *Let T be normal with distinct eigenvalues $\alpha \neq \beta$. Then α-eigenvectors are orthogonal to β-eigenvectors.*

Proof. Let u resp. v be an α- resp. β-eigenvector. Then $Tu = \alpha u$ ad $Tv = \beta v$. Hence

$$(\alpha - \beta)\langle u, v\rangle = \langle (\alpha - \beta)u, v\rangle = \langle \alpha u, v\rangle - \left\langle u, \bar{\beta} v\right\rangle = \langle Tu, v\rangle - \langle u, T^*v\rangle ,$$

where the last equality follows from Theorem 60.2. Now $\langle u, T^*v\rangle = \langle Tu, v\rangle$ by definition of the adjoint, so $(\alpha - \beta)\langle u, v\rangle = 0 \implies \langle u, v\rangle = 0$. □

Example 60.4. Let $T \in \mathcal{L}(\mathbb{C}^2)$ be given by the matrix $\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$. We can verify $TT^* = 13I$, so that $T^* = 13T^{-1}$. Hence T is normal, and the eigenvalues of T are $\lambda_{1,2} = 2 \pm 3i$. In the eigenbasis $\left\{\frac{1}{\sqrt{2}}(i, 1), \frac{1}{\sqrt{2}}(-i, 1)\right\}$, which is orthogonal by the previous theorem, T has the matrix $\operatorname{diag}(2 + 3i, 2 - 3i)$.

In fact, the basis above is orthonormal. This suggests to us this fundamental theorem in linear algebra: normal operators over a complex vector space are diagonalizable.

Theorem 60.5 (Complex Spectral Theorem)**.** *Let V be a finite-dimensional inner product space over $\mathbb{C}$, and let $T \in \mathcal{L}(V)$. Then the following are equivalent:*

1. *T is normal,*
2. *V has an orthonormal basis of eigenvectors of T,*
3. *T is diagonal in some orthonormal basis of V.*

Proof. Clearly, $(2) \iff (3)$ holds. We show $(1) \iff (3)$.

$(1) \implies (3)$: Let T be normal. Then T has an upper triangular matrix $A = (a_{ij})$ in some orthonormal basis $\{e_i\}_{i=1}^n$ of V. We verify that

$$\|Te_1\|^2 = \|a_{1,1}e_1\|^2 = a_{1,1}\overline{a_{1,1}}, \text{ and}$$

$$\|T^*e_1\|^2 = \|\overline{a_{11}}e_1 + \cdots + \overline{a_{1n}}e_n\|^2 = \sum_{j=1}^{n} |a_{1j}|^2$$

by taking the conjugate transpose of A. But $\|Te_1\|^2 = \|T^*e_1\|^2$ by normality and Theorem 60.1, so $a_{1,2} = \cdots = a_{1n} = 0$. This forces e_2 to be an $a_{2,2}$-eigenvector of T, but by a similar argument, $a_{2,3} = \cdots = a_{2n} = 0$. Continuing inductively, the off-diagonal entries are all 0, so A is in fact diagonal.

$(3) \implies (1)$: Let T have the matrix $A = \operatorname{diag}(a_1, \dots, a_n)$ in some orthonormal basis. Taking the conjugate transpose gives a diagonal matrix, and diagonal matrices commute, implying $TT^* = T^*T$. □

In the real case, being normal is not sufficient, but T needs to be *self-adjoint* in order for orthonormal diagonalization to be possible.

Theorem 60.6 (Real Spectral Theorem)**.** *Let V be a finite-dimensional inner product space over $\mathbb{R}$, and let $T \in \mathcal{L}(V)$. Then the following are equivalent:*

1. *T is* **self-adjoint**,
2. *V has an orthonormal basis of eigenvectors of T,*
3. *T is diagonal in some orthonormal basis of V.*

64 The Real Spectral Theorem, Positive Operators

[7B] Proof of the Real Spectral Theorem

In this subsection, we prove Theorem 60.6, so we will assume that V is a finite-dimensional inner product space over $\mathbb{R}$, and $T \in \mathcal{L}(V)$. The proof is far more complicated, so we need two lemmas.

Lemma 64.1. *Let $f(x) = x^2 + bx + c \in \mathbb{R}[x]$. Then f is irreducible over $\mathbb{R}$ if and only if $b^2 < 4c$. When this happens, if $T \in \mathcal{L}(V)$ is self-adjoint, then $T^2 + bT + cI$ is invertible.*

Proof. The first part is just high school algebra: check the discriminant of the quadratic. We will show that $\langle (T^2 + bT + cI)v, v \rangle > 0$ whenever $v \neq 0$ to show that $T^2 + bT + cI$ has nontrivial kernel. Fix $v \neq 0$, and let $T = T^*$. We compute

$$\begin{aligned}
\langle (T^2 + bT + cI)v, v \rangle &= \langle T^2 v, v \rangle + b \langle Tv, v \rangle + c \langle v, v \rangle \\
&= \langle Tv, Tv \rangle + b \langle Tv, v \rangle + c \langle v, v \rangle \\
&\geq \|Tv\|^2 + c \|v\|^2 - |b| \, \|Tv\| \, \|v\| \text{ (Cauchy-Schwarz)} \\
&= \left(\|Tv\| - \frac{|b| \, \|v\|}{2} \right)^2 + \left(c - \frac{b^2}{4} \right) \|v\|^2 > 0 \text{ as } b^2 < 4c.
\end{aligned}$$

Hence $\ker(T^2 + bT + cI) = \{0\}$, so $T^2 + bT + cI$ is invertible. □

We now state the second lemma we will use.

Lemma 64.2. *Suppose $T = T^*$ and $U \leq V$ is T-invariant. Then the following hold:*

1. *$U^\perp$ is T-invariant;*

2. *$T|_U$ is self-adjoint;*

3. *$T|_{U^\perp}$ is self-adjoint.*

Proof. Of course, (2) and (3) hold immediately by definition: $T = T^*$ on all of V, so the restrictions on the subspaces are equal. For (1), we fix $v \in U$ and $u \in U^\perp$. If suffices to show $\langle v, Tu \rangle = 0$, i.e., $Tu \in U^\perp$. But $\langle v, Tu \rangle = \langle T^* v, u \rangle = \langle Tv, u \rangle = 0$, as $Tv \in U$ by assumption. $\square$

We are now ready to prove the Real Spectral Theorem.

Proof. We prove $(1) \implies (2)$ in Theorem 60.6. The rest is easy.

We proceed by induction on the dimension of V. The one-dimensional case is obvious, and suppose the theorem holds for all vector spaces with dimension at most $n-1$, for some $n \in \mathbb{N}$. Now, let $\dim V := n$. Start with some $v \neq 0$, and consider the set $\{v, Tv, T^2 v, \ldots, T^n v\}$, which is linearly dependent by counting. Hence, there exist nontrivial $a_i \in \mathbb{R}$, $0 \leq i \leq n$, with $a_0 v + a_1 T v + \cdots + a_n T^n v = 0$. Let $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{R}[x]$. By the Fundamental Theorem of Algebra, $\mathbb{C}$ is a splitting field for p, so write

$$p(x) = \left[\prod_{\lambda_i \in \mathbb{R}} (x - \lambda_i)\right] \left[\prod_{\mu_j \in \mathbb{C} \setminus \mathbb{R}} (x - \mu_j)\right].$$

Since $p(x) \in \mathbb{R}[x]$, complex conjugation is an automorphism of $\mathbb{R}[\mu_j]$, so $\overline{\mu_j}$ is also a root[2] for p for any possible choice of j. It follows that $(x - \mu_j)(x - \overline{\mu_j}) = x^2 + |\mu_i|^2 \in \mathbb{R}[x]$ is an irreducible real factor of p, by Lemma 64.1. Hence, we can rewrite our factorization as

$$p(x) = \left[\prod_{\lambda_i \in \mathbb{R}} (x - \lambda_i)\right] \prod_k q_k(x),$$

where the q_k are irreducible quadratics in $\mathbb{R}[x]$. It follows by Lemma 64.1 that $q_k(T)$ is invertible, so that $\prod q_k(T)$ is also invertible. But $p(T)v = 0$, so at least one of the $(T - \lambda_i I)$'s must be uninvertible, so we have found a **real** eigenvalue λ for T.

Now, let $u_1 \in V$ be an λ-eigenvector of norm 1 of T. Set $U := \operatorname{span}(u_1) < V$; clearly, U is T-invariant. By Lemma 64.2, $U^\perp$ is T-invariant, and by construction $V = U \oplus U^\perp$. Applying the inductive hypothesis on $U^\perp$, we write $U^\perp = \operatorname{span}(u_2, \ldots, u_n)$, where the u_i are orthonormal eigenvectors of $T|_{U^\perp}$, and hence are eigenvectors of T. Hence $\{u_1, \ldots, u_n\}$ is an orthonormal basis of eigenvectors of T. $\square$

Examples

Example 64.3. Let $V = \mathbb{R}^2$ and let T be given by the matrix $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$, for $a, b, c \in \mathbb{R}$. Now $\det(A - \lambda I) = \lambda^2 - \lambda(a+b) + (ab - c^2)$. The discriminant of this polynomial is $\Delta = (a-b)^2 + 4c^2 \geq 0$. If $\Delta > 0$, then we two distinct eigenvalues, so we can diagonalize. If $\Delta = 0$, then our matrix is already diagonalized.

[2]This is just a fancy way of expressing the Conjugate Roots Theorem from high school algebra.

Example 64.4. Let $A = \begin{pmatrix} 14 & -13 & 8 \\ -13 & 14 & 8 \\ 8 & 8 & -7 \end{pmatrix}$. By inspection, A is self-adjoint, and hence normal, so the Real Spectral Theorem states that this matrix is diagonalizable. Indeed, we can verify that

$$\left\{ \frac{(1,-1,0)}{\sqrt{2}}, \frac{(1,1,1)}{\sqrt{3}}, \frac{(1,1,-2)}{\sqrt{6}} \right\}$$

is a basis of $\mathbb{R}^3$ under which A is a diagonal matrix.

[7C] Positive Operators

We will assume that V is a finite-dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $T \in \mathcal{L}(V)$.

Definition 64.5. An operator $T \in \mathcal{L}(V)$ is *positive* if T is self-adjoint and $\langle Tv, v \rangle \geq 0$ for all $v \in V$.

Positive operators are useful, as it turns out that all eigenvalues of a positive T are non-negative. Hence, "positive" operators should really be called "non-negative" operators.

Lemma 64.6. *If $T \in \mathcal{L}(V)$ is positive, then the following hold:*

1. *T only has non-negative eigenvalues;*
2. *There exists a self-adjoint operator S satisfying $S^2 = T$.*

If this is the case, we say that S is a square root of T.

Proof. (1): By the Spectral Theorem, T has a diagonal matrix $A := \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, where the λ_i are *real*, with respect to some orthonormal eigenbasis $\{u_i\}_1^n$. Hence $0 \leq \langle Tu_i, u_i \rangle = \lambda_i$.

(2): Fix the same diagonal matrix A as above. Let S have the matrix $\operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$ with respect to the same orthonormal eigenbasis. Clearly $S^2 = T$ and $S^* = S$. $\square$

67 Spectral Theorems: Worked Examples

As usual, assume V is a finite-dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and let $T \in \mathcal{L}(V)$.

Example 67.1. Let $\mathbb{F} = \mathbb{C}$, and let $T \in \mathcal{L}(V)$ be normal. Prove that T is self-adjoint if and only if its eigenvalues are all real.

Proof. ($\Longrightarrow$): Suppose that T is self-adjoint. By the Complex Spectral Theorem (CST), T has a diagonal matrix in an orthonormal eigenbasis of V, say $M_T = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Now, the matrix of T^* is the conjugate transpose, but this must coincide with M_T. Hence $\lambda_i = \overline{\lambda_i}$ for all $i \leq n$, so $\lambda_i \in \mathbb{R}$.

($\Longleftarrow$): Suppose T is normal, and all its eigenvalues are real. By CST, T has a diagonal matrix $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, where the λ_i are eigenvalues of T. Taking the conjugate transpose of A does nothing, so $T = T^*$. $\square$

Example 67.2. Give an alternate proof to the ($\Longrightarrow$) direction of Ex. 67.1 without the CST.

Proof. Suppose T is self-adjoint. If $v \neq 0$ is a λ-eigenvector of T, we have $\langle Tv, v \rangle = \lambda \|v\|^2$. But $\langle Tv, v \rangle = \langle v, Tv \rangle = \bar{\lambda} \|v\|^2$, so $\lambda = \bar{\lambda} \implies \lambda \in \mathbb{R}$. $\square$

Example 67.3. Let $\mathbb{F} = \mathbb{C}$, and let $T \in \mathcal{L}(V)$ be normal, satisfying $T^8 = T^9$. Show that T is self-adjoint and idempotent.

Proof. By the CST, find an λ-eigenvector $e \neq 0$ of T. Now $T^8 = T^9$ implies $T^8 e = T^9 e \iff \lambda^8 e = \lambda^9 e$, so that $\lambda^8(\lambda - 1)e = 0$. This implies $\lambda = 0, 1$, so T can only have 0 or 1 as eigenvalues. Hence, the matrix of T is $B = \operatorname{diag}(b_1, \ldots, b_n)$, where each b_i is binary (either 0 or 1). Clearly, $B^2 = B$ and B is self-adjoint. $\square$

Example 67.4. Let $T \in \mathcal{L}(V)$ be normal on $\mathbb{C}$. Prove that there exists an $S \in \mathcal{L}(V)$ such that $S^2 = T$.

Proof. By the CST, find a basis of T such that T has the matrix $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, where we have $\lambda_i \in \mathbb{C}$. Square roots always exist over complex numbers, so simply set S to have the matrix $\operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$. $\square$

Example 67.5. Give an example of a real vector space V, some $T \in \mathcal{L}(V)$, some $b, c \in \mathbb{R}$, such that $b^2 < 4c$ yet $T^2 + bT + cI$ is uninvertible.

Proof. To keep things simple, set $b = 0$ and $c = 1$, and $V = \mathbb{R}^2$. Then $0 < 4$, so $b^2 < 4c$, and we have $T^2 + bT + cI = T^2 + I$. It would be nice if we had $T^2 + I = 0$. Luckily, this is possible: set T to be the rotation by $\pi/2$, which has matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. $\square$

We remark that the above example shows that the *self-adjoint* condition in Lemma 64.1 and hence in the Real Spectral Theorem are crucial for those statements to be true.

70 Positive Operators and Isometries

In the proof of Lemma 64.6, the square root S we chose was in fact a positive operator. This leads us to the following statement.

Of course, we will continue to assume that V is a finite-dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and that $T \in \mathcal{L}(V)$.

Lemma 70.1. *If T is positive, then it has a unique positive square root.*

The proof is not too enlightening, so we will skip it. However, this allows us to introduce the notation $\sqrt{T}$ for the unique positive square root of T, assuming that T is also positive.

[7C] Isometries

Recall from Definition 58.3 that an operator is an *isometry* if $T^* = T^{-1}$. Alternatively, we can define an isometry as follows.

Definition 70.2. $T \in \mathcal{L}(V)$ is an *isometry* if it preserves norms; i.e., $\|Tv\| = \|v\|$ for all $v \in V$.

The proof of the next lemma is also not enlightening, so we will skip it as well.

Lemma 70.3. *Let $T \in \mathcal{L}(V)$. Then T is an isometry if and only if $\langle Tv, Tw \rangle = \langle v, w \rangle$ for all $v, w \in V$.*

Now, we show that Definitions 58.3 and 70.2 are equivalent.

Proposition 70.4. *Let $T \in \mathcal{L}(V)$. Then $\|Tv\| = \|v\|$ if and only if $T^* = T^{-1}$.*

Proof. Pick any $v \in V$. Then $\langle v, v\rangle = \langle Tv, Tv\rangle = \langle T^*Tv, v\rangle$ by Lemma 70.3. Define the operator S by $S := I - T^*T$. Now, $\langle v, v\rangle - \langle T^*Tv, v\rangle = 0$, so $\langle Sv, v\rangle = 0$. But $S^* = I^* - (T^*T)^* = I - T^*T = S$, so S is self-adjoint. By the CST, $\langle Sv, v\rangle = 0$ implies that 0 is an eigenvalue, so $S = 0$, so $TT^* = I$. The other direction is similar. $\square$

Corollary 70.5. *If T is an isometry, then T is normal.*

The next proposition allows us to classify isometries.

Proposition 70.6. *If T is an isometry, then if $v \neq 0$ is a λ-eigenvector of T, then $|\lambda| = 1$.*

Proof. Write $\langle v, v\rangle = \langle Tv, Tv\rangle = \langle \lambda v, \lambda v\rangle = |\lambda|^2 \langle v, v\rangle$, so $|\lambda| = 1$. $\square$

Theorem 70.7. *Let $\{e_i\}_1^n \subset V$ be an orthonormal basis, and let $\{\lambda_i\}_1^n \subset \mathbb{C}$ with $|\lambda_i| = 1$. Define the operator S by $Se_i = \lambda_i e_i$ for each basis vector. Then S is an isometry.*

Proof. Take $v = \sum_i a_i e_i$. Then $\langle v, v\rangle = \sum |a_i|^2$ by orthonormality. Now

$$\langle Sv, Sv\rangle = \sum_{i=1}^n |a_i\lambda_i|^2 = \sum_{i=1}^n |a_i|^2|\lambda_i|^2 = \sum_{i=1}^n |a_i|^2.$$

Hence $\langle v, v\rangle = \langle Sv, Sv\rangle$, so S is an isometry. $\square$

Corollary 70.8. *If S is an isometry and $\{e_i\}_1^n$ is an orthonormal basis of V, then $\{Se_i\}_1^n$ is also an orthonormal basis of V.*

73 Positive Operators: Worked Examples

[The numbers in parentheses next to each Example number below refers to the textbook exercise number.]

Example 73.1 (7C.5)**.** Prove that the sum of two positive operators is positive.

Proof. Let S, T be positive operators over an inner product space V. Then S, T are self-adjoint and $\langle Sv, v\rangle, \langle Tv, v\rangle \geq 0$ for every $v \in V$. Now

$$\langle (S+T)v, v\rangle = \langle Sv + Tv, v\rangle = \langle Sv, v\rangle + \langle Tv, v\rangle \geq 0 + 0 = 0,$$

so this completes the proof after noting that $(S+T)^* = S^* + T^* = S + T$. $\square$

Example 73.2 (7C.6)**.** Suppose that $T \in \mathcal{L}(V)$ is positive. Prove that T^k is positive for every positive integer k.

Proof. Let T be positive, and fix $k \in \mathbb{Z}^+$. Then T is self-adjoint, so we immediately see that T^k is self-adjoint. Since T is positive, all of its eigenvalues are non-negative, so fix an eigenbasis of V. Then T has a diagonal matrix $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Clearly, T^k has the matrix $\operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)$, and since $\lambda_i \geq 0$, $\lambda_i^k \geq 0$. Then by Theorem 7.35 in the text, T^k is positive. $\square$

Example 73.3 (7C.7)**.** Suppose that $T \in \mathcal{L}(V)$ is positive. Prove that T is invertible if and only if $\langle Tv, v\rangle > 0$ for every $v \in V \setminus 0$.

Proof. Suppose T is positive, and suppose $\langle Tv, v\rangle > 0$ for every $v \in V \setminus 0$. Certainly, this implies $Tv \neq 0$, so T has trivial kernel and is thus invertible.

Now, suppose T is invertible and positive. Then there exists some S, also positive, such that $S^2 = T$. Now for any $v \in V \setminus 0$, $\langle Tv, v\rangle = \left\langle S^2 v, v\right\rangle = \langle Sv, Sv\rangle = \|Sv\|^2 > 0$. □

Example 73.4 (7C.9)**.** Prove or disprove: the identity operator on $\mathbb{F}^2$ has infinitely many square roots.

Proof. This is true: fix $\theta \in [0, \pi)$, and define T_θ to be the reflection about the line with angle θ to the x-axis. It is easy to check that T_θ is self-adjoint, and that $T_\theta^2 = I$. □

77 Singular Values: Worked Examples

[The material here is covered in Section 78, so that should be read first. This discussion section was ahead of the lecture.]

Example 77.1. Suppose $T \in \mathcal{L}(V)$, $S \in \mathcal{L}(V)$ is an isometry, and $R \in \mathcal{L}(V)$ is a positive operator such that $T = SR$. Prove that $R = \sqrt{T^*T}$.

Proof. We want to introduce a T^* into our equation, so we take adjoints and find that $T^* = R^*S^* = R^*S^{-1}$. Hence

$$T^*T = R^*S^{-1}SR = R^*R = R^2,$$

as R is positive, hence self-adjoint. As T^*T is positive, this implies $\sqrt{T^*T} = R$. □

Example 77.2. Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that the singular values of T equal the absolute values of the eigenvalues of T, repeated appropriately.

Proof. Pick an orthonormal eigenbasis $\{e_i\}$ of T. Then $Te_i = \lambda_i e_i$, where the λ_i are the eigenvalues, possibly repeated, of T. Because T is self-adjoint, we see that $T^*Te_i = T^2 e_i = \lambda_i^2 e_i$, so by selecting the positive square root, we see that $\sqrt{T^*T}e_i = |\lambda_i| e_i$, so we are done. □

Example 77.3. Suppose $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if 0 is not a singular value of T.

Proof. ($\implies$): Suppose T is invertible. Then T^* is also invertible, so T^*T is invertible, which forces $\sqrt{T^*T}$ to be invertible. This means that 0 is not an eigenvalue of $\sqrt{T^*T}$, so we are done.

($\impliedby$): Reverse everything from the preceding argument. □

78 Polar and Singular-Value Decompositions

Our goal in this section is to develop a decomposition of a matrix that is diagonal, but we allow ourselves to use two different bases of the same vector space. Assume our usual assumptions: that V is a finite-dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and let $T \in \mathcal{L}(V)$.

Theorem 78.1 (Polar Decomposition)**.** *Suppose $T \in \mathcal{L}(V)$. Then there exists an isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.*

Proof. See [7D] in the text. □

We now view an example of finding a polar decomposition of a matrix.

Example 78.2. Let $T = \begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix} \in \mathcal{L}(\mathbb{R}^2)$. Then $T^* = \begin{pmatrix} 1 & 0 \\ a & 0 \end{pmatrix}$, so $T^*T = \begin{pmatrix} 1 & a \\ a & a^2 \end{pmatrix}$. We check that T^*T has eigenvalues $(1+a^2)$ and 0, and an orthonormal eigenbasis of $\mathbb{R}^2$ is

$$e_1 = \sqrt{a^2+1}(1,a), e_2 = \frac{(-a,1)}{\sqrt{a^2+1}}.$$

Thus, T^*T maps $e_1 \mapsto (a^2+1)e_1$ and $e_2 \mapsto 0$. This means $\sqrt{T^*T}$ maps $e_1 \mapsto \sqrt{a^2+1}e_1$ and $e_2 \mapsto 0$. Using this, we can find a matrix for $\sqrt{T^*T}$. If $\sqrt{T^*T} = \begin{pmatrix} x & z \\ z & y \end{pmatrix}$,[3] we have

$$\begin{pmatrix} x & z \\ z & y \end{pmatrix}\begin{pmatrix} 1 \\ a \end{pmatrix} = \sqrt{a^2+1}\begin{pmatrix} 1 \\ a \end{pmatrix} \text{ and } \begin{pmatrix} x & z \\ z & y \end{pmatrix}\begin{pmatrix} -a \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This yields $z = ax$ and $y = az$, so $\sqrt{T^*T} = \begin{pmatrix} x & ax \\ ax & a^2x \end{pmatrix}$. This gives $(1+a^2)x = \sqrt{1+a^2} \implies x = (1+a^2)^{-1/2}$, so we have found the matrix for $\sqrt{T^*T}$. Now, $\operatorname{im} T = \mathbb{R}\begin{pmatrix} 1 \\ a \end{pmatrix}$, so we define a transformation $S_1 : \operatorname{im}\sqrt{T^*T} \to \operatorname{im} T$ by $S(\sqrt{T^*T}v) \mapsto Tv$. We have

$$S_1\begin{pmatrix} 1 \\ a \end{pmatrix} = T\left[\frac{(1,a)}{\sqrt{a^2+1}}\right] = \frac{1}{\sqrt{a^2+1}}\begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ a \end{pmatrix} = \sqrt{a^2+1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Now, define $S_2 : (\operatorname{im}\sqrt{T^*T})^\perp \to (\operatorname{im} T)^\perp$, which is a map $S_2 : \mathbb{R}\cdot e_2 \to \mathbb{R}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ by $S_2 e_2 = (0,1)$.

Referring to page 235 in the text, we see that $S = \boxed{\frac{1}{\sqrt{a^2+1}}\begin{pmatrix} 1 & a \\ -a & 1 \end{pmatrix}}$.

The Polar Decomposition allows us to arrive at a nicer diagonalization, which uses two different bases for one space.

Definition 78.3. Let $T \in \mathcal{L}(V)$. The *singular values* of T are the eigenvalues of $\sqrt{T^*T}$.

Theorem 78.4 (Singular Value Decomposition)**.** *Let T have the singular values $s_1, \dots, s_n$. Then, there exist orthonormal bases $\{e_1, \dots, e_n\}$, $\{f_1, \dots, f_n\}$ of V such that*

$$Tv = \sum_{i=1}^n s_i \langle v, e_i \rangle f_i \text{ for every } v \in V.$$

Alternatively, T has a diagonal matrix with respect to these two different bases.

Proof. By the Polar Decomposition, write $T = S\sqrt{T^*T}$, where S is an isometry. By the Spectral Theorem, find an orthonormal basis $\{e_i\}_1^n$ of V such that $\sqrt{T^*T}(e_i) = s_i e_i$. Define $f_i := Se_i$. Now, S is an isometry, so by Corollary 70.8, $\{f_i\}_1^n$ is also an orthonormal basis of V. Now, pick $v \in V$, so that $v = \sum_i \langle v, e_i \rangle e_i$. Since

$$\sqrt{T^*T}v = s_1 \langle v, e_1 \rangle e_1 + \cdots + s_n \langle v, e_n \rangle e_n, \text{ we have}$$

$$Tv = S\sqrt{T^*T}v = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n,$$

completing the proof. □

[3]...and we are allowed to assume this as $\sqrt{T^*T}$ is by definition positive, and thus self-adjoint, and thus its matrix is symmetric.

80 Generalized Eigenvectors

In this section, we have no need of inner products, so we will just assume that V is a finite-dimensional vector space over $\mathbb{C}$.

Definition 80.1. An operator $N \in \mathcal{L}(V)$ is *nilpotent* if there exists some $k \in \mathbb{N}$ such that $N^k = 0$. The minimum of all such k is called the *nilpotency degree* of N.

Before we proceed further with this definition, let us verify some easy facts about the kernels of powers of transformations.

Lemma 80.2. *If $T \in \mathcal{L}(V)$, and if there exists an $m \in \mathbb{N}$ such that $\ker T^m = \ker T^{m+1}$, then*

$$\ker T^m = \ker T^{m+1} = \ker T^{m+2} = \cdots .$$

Proof. Certainly, for any $j \in \mathbb{N}$ we have $\ker T^{m+j} \leq \ker T^{m+j+1}$. Now, fix $v \in \ker T^{m+j+1}$. The $T^{m+1}(T^j v) = 0$, so $T^j v \in \ker T^{m+1} = \ker T^m$ by assumption, so that $T^m(T^j v) = 0 \iff T^{m+j}v = 0$, so $v \in \ker T^{m+j}$. Hence, we are done. □

Proposition 80.3. *Let $T \in \mathcal{L}(V)$. Then the chain of subspaces*

$$\{0\} = \ker T^0 \leq \ker T \leq \ker T^2 \leq \ker T^3 \leq \cdots$$

satisfies the ascending chain condition: i.e., there exists an $n \in \mathbb{N}$ such that

$$\ker T^{n-1} \leq \ker T^n = \ker T^{n+1} = \ker T^{n-2} = \cdots .$$

In fact, such n satisfies $n \leq \dim V$.

Proof. Suppose otherwise. By Lemma 80.2, we see that we must have strict inclusions

$$\ker T^0 < \ker T^1 < \ker T^2 < \cdots ,$$

so the dimensions must keep increasing:

$$0 = \dim \ker T^0 < \dim \ker T^1 < \dim \ker T^2 < \cdots .$$

But this is impossible as $\ker T^k \leq V$, and V is finite-dimensional by our assumption at the beginning of this section. □

We continue exploring nilpotency.

Lemma 80.4. *Let $N \in \mathcal{L}(V)$ be nilpotent. Then N has an upper triangular matrix with all zeros on the diagonal, for some basis of V.*

Proof. Build a basis of $\ker N$, say $\{e_i\}_1^k$, and extend this to a basis $\{e_i\}_1^\ell$ of $\ker N^2$, then to a basis $\{e_i\}_1^m$ of $\ker N^3$, and so on until we hit the nilpotency degree of N. Then $Ne_1 = Ne_2 = \cdots = Ne_k = 0$, and $Ne_{k+1}, \ldots, N_\ell \in \ker N$, and $Ne_{\ell+1}, \ldots, Ne_m \in \ker N^2$, and so on. This constructs an upper-triangular matrix with 0 on the diagonal. Of course, this process terminates as per Proposition 80.3. □

This allows us to consider an important class of kernels.

Definition 80.5. Let λ be an eigenvalue for some $T \in \mathcal{L}(V)$. then the *generalized eigenspace corresponding to λ* is

$$G(\lambda, T) := \ker(T - \lambda I)^n.$$

The nonzero vectors in $G(\lambda, T)$ are *generalized λ-eigenvectors* of T.

Immediately, we have an important fact.

Theorem 80.6. *Let T have distinct eigenvalues $\lambda_1, \dots, \lambda_m$, and let v_i be generalized λ_i-eigenvectors of T. Then $\{v_1, \dots, v_m\}$ is linearly independent.*

Proof. Suppose we had a linear combination summing to zero: $a_1 v_1 + \cdots + a_n v_n = 0$, for $a_i \in \mathbb{C}$. Set $k \in \mathbb{N}$ to be the largest natural number with $(T - \lambda_1 I)^k v_1 \neq 0$; such a number must exist as v_1 is a generalized eigenvector, so at least $(T - \lambda_1 I)^n v_1 = 0$. If $w := (T - \lambda_1 I)^k v$, then $(T - \lambda_1)w = 0 \iff Tw = \lambda_1 w$.

Now, consider some $\lambda \neq \lambda_1$. We see that $(T - \lambda I)w = (\lambda_1 - \lambda)w$, so that $(T - \lambda I)^n w = (\lambda_1 - \lambda)^n w$. Applying the product $(T - \lambda_1)^k \prod_{j \geq 2} (T - \lambda_j I)^n$ to our linear combination, we see

$$0 = a_1 (T - \lambda_1)^k (T - \lambda_2 I)^n \cdots (T - \lambda_m I)^n v_1$$

as the other terms are annihilated by the $(T - \lambda_j I)^n$. Now, by commutativity and the above discussion

$$0 = a_1 (T - \lambda_2 I)^n \cdots (T - \lambda_m I)^n = a_1 (\lambda_1 - \lambda_2)^n \cdots (\lambda 1 - \lambda_m)^n w.$$

Since $\lambda_1 \neq \lambda_j$ for any $j \geq 2$, we see that $a_1 = 0$. This argument is not unique for λ_1, so replacing λ_1 with λ_ℓ show that $a_\ell = 0$, so our list is linearly independent. $\square$

From this, we can see that the $G(\lambda_i, T)$, when summed together, form a direct sum. The next proposition is extremely useful in the proofs we will see later.

Proposition 80.7. *Let $T \in \mathcal{L}(V)$ and let λ be an eigenvalue for T. Then $G(\lambda, T)$ is a T-invariant subspace.*

Proof. By definition, we have $G(\lambda, T) = \ker(T - \lambda I)^n$. Set $f(x) = (x - 1)^n \in \mathbb{C}[x]$, and apply Lemma 18.7. $\square$

84 The Generalized Eigenspace Decomposition

Assume that V is a finite-dimensional vector space over $\mathbb{C}$. We remarked above that the generalized eigenspaces for an operator T form a direct sum, and that each generalized eigenspace is T-invariant. In fact, this direct sum covers the whole space, as we shall see. We prove a short proposition first.

Proposition 84.1. *Let $\dim V =: n$. Then for any operator $T \in \mathcal{L}(V)$, we have $\ker T^n \oplus \operatorname{im} T^n$.*

Proof. Let $v \in \ker T^n \cap \operatorname{im} T^n$. Then $T^n v = 0$ and there exists some $w \in V$ such that $v = T^n w$. Hence $T^{2n} w = 0$, so $w \in \ker T^{2n} = \ker T^n$ by way of Proposition 80.3. This means $v = 0$. Now, the equality $V = \ker T^n + \operatorname{im} T^n$ is given by the Rank-Nullity Theorem. $\square$

What follows is arguably one of the most important theorems in the course.

Theorem 84.2 (Generalized Eigenspace Decomposition)**.** *Let V be a finite-dimensional vector space over $\mathbb{C}$, $T \in \mathcal{L}(V)$, and let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T. Then*

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_k, T),$$

and the operators $(T - \lambda_j I)|_{G(\lambda_j, T)}$, $1 \leq i \leq k$, are all nilpotent.

Proof. That the operators $(T - \lambda_j I)|_{G(\lambda_j, T)}$ are nilpotent follows from the definition of $G(\lambda_j, T)$, that they are the kernel of some operator power.

For the decomposition part of the theorem, we induct on the dimension of V. If V is one-dimensional, we have $V = G(\lambda_1, T)$. Now, assume the inductive hypothesis and assume the theorem holds for all vector spaces with dimension at most $n-1$. Now, let $\dim V = n$, and write $V = W \oplus U$, where $W := G(\lambda_1, T) = \ker(T - \lambda_1 I)^n$. By Proposition 84.1, we immediately have $U = \operatorname{im}(T - \lambda_1 I)^n$. We know $\dim W \leq 1$, so $\dim U =: m \leq n-1$. By Lemma 18.7 and Proposition 80.7, both U and W are T-invariant, so by the inductive hypothesis, write

$$U = G(\lambda_2, T|_U) \oplus \cdots \oplus G(\lambda_k, T|_U).$$

All of the generalized λ_1-eigenvectors lie in W, so we see $\lambda_1 \notin \{\lambda_2, \ldots, \lambda_k\}$ (so what we are doing is legal). Now, we show $G(\lambda_i, T|_U) \subseteq G(\lambda_i, T)$, for $i \geq 2$. The forward inclusion is obvious, so take $v \in G(\lambda_i, T)$. By our decomposition, we have

$$v = v_1 + v_2 + \cdots + v_k,$$

where $v_1 \in W$ and $v_i \in G(\lambda_i, T|_U)$ for $i \geq 2$. Then $0 = v_1 + v_2 + \cdots + (v_i - v) + v_{i+1} + v_k$ for some $i \geq 2$, and each term is in its own $G(\lambda_j, T)$. But generalized eigenvectors for distinct eigenvalues are linearly independent, so we must have $v_j = 0$, $j \neq i$ and $v = v_i \in G(\lambda_i, T|_U)$. Hence $V = G(\lambda_1, T) \oplus G(\lambda_2, T) \oplus \cdots \oplus G(\lambda_k, T)$. $\square$

The following is a restatement of our main theorem.

Corollary 84.3. *If $T \in \mathcal{L}(V)$, then V has a basis of generalized eigenvectors of T.*

The following may also be of use.

Lemma 84.4. *If $T \in \mathcal{L}(V)$ has eigenvectors $\lambda_1, \ldots, \lambda_k \neq 0$, then T has a square root.*

Proof. Since each $G(\lambda_i, T)$ is T-invariant, write $G(\lambda_i, T)$ as a direct sum of the $G(\lambda_j, T|_{G(\lambda_i, T)})$ by the Generalized Eigenspace Decomposition (GED). Now, set $N := T - \lambda I$. Then $N^{\dim V = 0}$, thus $T = \lambda I + N = \lambda(I + N/\lambda)$. Without loss of generality, if $\lambda = 1$, we have $T = I + N$ so use the power series representation

$$\sqrt{I + N} = I + a_1 T + a_2 T^2 + a_3 T^3 + \cdots,$$

which terminates by nilpotency. $\square$

Example 84.5. Let $T = \begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}$. We can read off the eigenvalues: $\lambda_1 = 6$ and $\lambda_2 = 7$. Thus $T - 6I = \begin{pmatrix} 0 & 3 & 4 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, so we see $(T - 6I)^2 = \begin{pmatrix} 0 & 0 & 10 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Hence

$$G(6, T) = \operatorname{span}((1,0,0), (0,1,0)) \text{ and } G(7, T) = \operatorname{span}((10, 2, 1)).$$

90 Characteristic Polynomials

In this section, V denotes a finite-dimensional vector space over $\mathbb{C}$, and $T \in \mathcal{L}(V)$. By the GED, write

$$V = G(\lambda_1, T) \oplus \cdots G(\lambda_m, T),$$

where the λ_i are distinct, and define $d_i := \dim G(\lambda_i, T)$.

Definition 90.1. Let $T \in \mathcal{L}(V)$ and fix notation as in the beginning of this section. The *characteristic polynomial* of T is $q_T(x) := \prod_{i=1}^{m}(x-\lambda_i)^{d_i} \in \mathbb{C}[x]$.

Theorem 90.2 (Cayley-Hamilton Theorem)**.** *Let $T \in \mathcal{L}(V)$, and let $q_T(x)$ be the characteristic polynomial of T. Then $q_T(T) = 0$.*

Proof. Write $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$ by the GED. Each subspace is T-invariant, so by choosing bases on these subspaces, we put T into *block-diagonal form*, with matrix $M_T := \operatorname{diag}(D_1, \ldots, D_m)$, where D_i has size $d_i \times d_i$ and is the matrix of $T|_{G(\lambda_i,T)}$. Now $M_T^k = \operatorname{diag}(D_1^k, \ldots, D_m^k)$, so $q(M_T) = \operatorname{diag}(q(D_1), \ldots, q(D_m))$, and $q(D_i) = (D_i - \lambda_i I)^{d_i}$. Hence, it suffices to show that $(D_i - \lambda_i I)^{d_i} = 0$. But $\ker(D_i - \lambda_i I)^{d_i} = G(\lambda_i, T)$ by definitions, so by the GED, we are done. $\square$

Example 90.3. Let $T = \begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}$. Then $q_T(x) = (x-6)^2(x-7)$, and we can check that

$$q_T(T) = \begin{pmatrix} 0 & 3 & 4 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}^2 \begin{pmatrix} -1 & 3 & 4 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 10 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 3 & 4 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{0}_{3\times 3},$$

as predicted by the Cayley-Hamilton Theorem.

Example 90.4. Let $T := \operatorname{diag}(6, 6, 7)$. Then $q_T(x) = (x-6)^2(x-7)$, but if $m(x) := (x-6)(x-7)$, we see $m(T) = 0$. This highlights a possible issue with the characteristic polynomial: it is not necessarily of minimal degree.

To fix that issue, we make the following definition.

Definition 90.5. Let $T \in \mathcal{L}(V)$. The *minimal polynomial* of T is the monic polynomial $m_T(x)$ of smallest degree satisfying $m_T(T) = 0$.

The following theorems are not unlike propositions in ring theory (Math 120B).

Theorem 90.6. *Let $T \in \mathcal{L}(V)$. Then the minimal polynomial m_T is unique, and m_T and q_T have the exact same roots.*

Proof. In view of Theorem 90.2, a minimal polynomial exists. Suppose m_1, m_2 are minimal polynomials of T. Write $m_1(x) = a_0 + a_1 x + \cdots + x^k$ and $m_2(x) = b_0 + b_1 x + \cdots + x^k$. Then $m_1(T) = m_2(T) = 0$, but this implies $(m_1 - m_2)(T) = 0$. Now $m_1 - m_2$ has strictly lower degree than both m_1 and m_2, which contradicts minimality. Hence, $m_T(x)$ is unique.

The second part of this theorem amounts to showing that the roots of m_T are precisely the eigenvalues of T.[4] Let $m_T(x) = a_0 + a_1 x + \cdots + x^k$. If $r \in \mathbb{C}$ is a zero of m_T, then we have $m_T(x) = (x - r)g(x)$, where g is monic. Since $m_T(T) = 0$, we see

$$0 = (T - rI)(g(T)v)$$

for any $v \in V$. But this is only legal if there exists at least a nonzero vector $v_0 \in V$ with $g(T)v_0 \neq 0$, otherwise we have a contradiction with minimality. Hence $(T - rI)v_0 = 0$, so that v_0 is an r-eigenvector of T.

[4] This is the textbook's proof, not the one presented in lecture; this one is far more clear. See [8C] in the text.

Conversely, suppose λ is an eigenvalue for T. Pick an λ-eigenvector $v \neq 0$; i.e., $Tv = \lambda v$. Hence

$$\begin{aligned} 0 &= m_T(T)v = a_0 v + a_1 Tv + a_2 T^2 v + \cdots + T^k v \\ \implies 0 &= a_0 + a_1 \lambda v + a_2 \lambda^2 v + \cdots + \lambda^k v = m_T(\lambda)v. \end{aligned}$$

Hence $m_T(\lambda) = 0$. □

93 Generalized Eigenspaces: Worked Examples

Example 93.1 (8A.1)**.** Define $T \in \mathcal{L}(\mathbb{C}^2)$ by $(w, z) \mapsto (z, 0)$. Find all generalized eigenvectors of T.

Solution. Let λ be an eigenvalue for T. Then if $(a, b) \neq \mathbf{0}$ is an eigenvector, we have $T(a, b) = (b, 0) = (\lambda a, \lambda b)$. Clearly, $\lambda = 0$ is an eigenvalue. If $\lambda \neq 0$ is an eigenvalue, we must have $b = 0$, but this implies $a = 0$, which is impossible. Hence, the only eigenvalue of T is $\lambda = 0$, and an eigenvector is $(1, 0)$. We also notice that $(0, 1)$ (as well as literally any other vector in $\mathbb{C}^2$) is a generalized eigenvector, as $(T - 0I)^2 = T^2 = 0$. Hence $G(0, T) = \boxed{\mathbb{C}^2}$. •

Example 93.2 (8A.5)**.** Suppose $T \in \mathcal{L}(V)$, m is a positive integer, and $v \in V$ is such that $T^{m-1}v \neq 0$ but $T^m v = 0$. Prove that $\{v, Tv, \ldots, T^{m-1}v\}$ is linearly independent.

Proof. Suppose there exist a_i, $0 \leq i \leq m-1$, such that

$$a_0 v + a_1 Tv + \cdots + a_{m-1} T^{m-1} v = 0.$$

This implies that

$$T^{m-1}(a_0 v + \cdots + a_{m-1} T^{m-1} v) = 0 \implies T^{m-1} a_0 v = 0 \implies a_0 = 0.$$

Similarly, we see

$$T^{m-2}(a_1 Tv + \cdots + a_{m-1} T^{m-1} v) = T^{m-1} a_1 v = 0 \implies a_1 = 0,$$

and in this fashion, we eliminate all of the a_i. Hence the set is linearly independent. □

Example 93.3 (8A.6)**.** Suppose $T \in \mathcal{L}(\mathbb{C}^3)$ is defined by $(z_1, z_2, z_3) \mapsto (z_2, z_3, 0)$. Prove that T has no square root.

Proof. We give a dimensional argument. Notice that T has *nilpotency degree* 3; i.e., $T, T^2 \neq 0$, but $T^3 = 0$. Hence $\ker T^3 = \mathbb{C}^3$, but we know that $0 = \ker T^0 \leq \ker T^1 \leq \ker T^2 \leq \ker T^3 = \mathbb{C}^3$. Moreover, $\ker T^2, \ker T^3 \neq \mathbb{C}^3$, so by counting, we must have

$$0 = \dim\ker T^0 < \dim\ker T^1 < \dim\ker T^2 < \dim\ker T^3 = 3,$$

where each inequality is strict. Now, assume for contradiction that S is a square root of T. Then we know that $\dim\ker S^2 = 1$, $\dim\ker S^4 = 2$, and $\dim\ker S^6 = 3$, but this is a contradiction: we have $1 \leq \dim\ker S^3 \leq 2$, but if $\dim\ker S^3 = 1$, this implies $\ker S^2 = \ker S^3 = \cdots = \ker S^6$, which is impossible, and similarly if $\dim\ker S^3 = 2$. Hence, T cannot have a square root. □

Example 93.4 (8B.1)**.** Suppose V is a finite-dimensional complex vector space, $N \in \mathcal{L}(V)$, and 0 is the only eigenvalue of N. Prove that N is nilpotent.

Proof. By the GED, $V = G(0, N)$, which means that $(N - 0I)|_V = N$ is nilpotent. $\square$

Example 93.5 (8B.2)**.** Give an example of an operator T on a finite-dimensional real vector space such that 0 is the only real eigenvalue of T, but T is not nilpotent.

Solution. Define T on $\mathbb{R}^3$ by $(x, y, z) \mapsto (0, -z, y)$. If $\lambda \in \mathbb{R}$ is an eigenvalue, we must have $(0, -z, y) = (\lambda x, \lambda y, \lambda z)$, so clearly $\lambda = 0$ is an eigenvalue. If $\lambda \neq 0$, then we have $x = 0$, and we have $-z = \lambda y$, $y = \lambda z$, which implies $-z = \lambda^2 z \implies 0 = z\lambda^2 + z = z(\lambda^2 + 1)$, so $\lambda = \pm i \notin \mathbb{R}$. However, we show that T is not nilpotent: notice that $T^k(0, 1, 1) \neq \mathbf{0}$ for any $k \in \mathbb{Z}$. •

Example 93.6 (8B.4)**.** Suppose V is an n-dimensional complex vector space and T is an operator on V such that $\ker T^{n-2} \neq \ker T^{n-1}$. Prove that T has at most 2 distinct eigenvalues.

Proof. We must have

$$\{0\} = \ker T^0 < \ker T^1 < \cdots < \ker T^{n-2} < \ker T^{n-1},$$

where each inclusion is strict. Hence $\dim \ker T^{n-1} \geq n - 1$, but $\ker T^{n-1} \leq \ker T^n = G(0, T)$, so this leaves two cases to consider.

If $\dim G(0, T) = n$, then this implies $G(0, T) = V$. Hence 0 is the only eigenvalue of V by the GED. Otherwise, if $\dim G(0, T) = n - 1$, then $V = G(0, T) \oplus G(\lambda, T)$, where $\lambda \neq 0$ is the other eigenvalue of T. $\square$

94 Characteristic Polynomials, Jordan Form

As usual, assume that V is a finite-dimensional vector space over $\mathbb{C}$, and let $T \in \mathcal{L}(V)$. The GED tells us this fact, which we state separately for clarity.

Proposition 94.1. *Let $T \in \mathcal{L}(V)$, and let $\{\lambda_i\}_1^k$ be the distinct eigenvalues of T. Then $(T - \lambda_i I)|_{G(\lambda_i, T)}$ is nilpotent for all $1 \leq i \leq k$.*

This proposition allows us to define the characteristic polynomial q_T, as well as state the Cayley-Hamilton Theorem. We also have this final lemma about the relationship between m_T and q_T.

Lemma 94.2. *If $b(x) \in \mathbb{C}[x]$ is such that $b(T) = 0$, then $m_T \mid b$. In particular, the minimal polynomial divides the characteristic polynomial of T.*

Proof. By the Division Algorithm, write $b(x) = m(x)g(x) + r(x)$, where $\deg r < \deg m$. Evaluating for $T \in \mathcal{L}(V)$, we have $0 = b(T) = m(T)a(T) + r(T) = 0 + r(T) = r(T)$, so $m \mid b$. $\square$

[8D] Jordan Form

Example 94.3. Let $N : \mathbb{C}^4 \to \mathbb{C}^4$ be defined by $(z_1, z_2, z_3, z_4) \mapsto (0, z_1, z_2, z_3)$. Let $v = (1, 0, 0, 0)$. We check that $\{v, Nv, N^2v, N^3v\}$ is a basis of $\mathbb{C}^4$. Ordering this basis backwards, we see that N has the matrix $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ with 1's on the superdiagonal and 0's on the main diagonal.

Example 94.4. Define $N : \mathbb{C}^6 \to \mathbb{C}^6$ by $(z_1, \dots, z_6) \mapsto (0, z_1, z_2, 0, z_4, 0)$. Let $v_1 = (1, 0, 0, 0, 0, 0)$, $v_2 = (0, 0, 0, 1, 0, 0)$ and $v_3 = (0, 0, 0, 0, 0, 1)$. Then $N^3 v_1 = N^2 v_2 = N v_3 = 0$, and we can check that

$$\{N^2 v_1, N v_1, v_1, N v_2, v_2, v_3\}$$

is a basis. Now, the matrix of N under this basis is

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

which matches with our red, blue, and green groupings of our basis, but in blocks instead.

This is one case of a general phenomenon.

Theorem 94.5. *Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then there exist vectors $v_1, \dots, v_s$ and integers $m_1, \dots, m_s \geq 0$ such that $N^{m_i+1} v_i = 0$, and for each v_i and m_i pair, we have $N^j v_i \neq 0$ for $j \leq m_i$,* **and**

$$\beta := \{N^{m_1} v_1, \dots, v_1, N^{m_2} v_2, \dots, v_2, \dots, N^{m_s} v_s, \dots, v_s\}$$

is a basis for V. Additionally, this implies that N has the matrix $\operatorname{diag}(A_1, A_2, \dots, A_s)$ with respect to β, and each A_i is an $(m_i + 1) \times (m_i + 1)$ matrix with 0 on the diagonal and 1 on the superdiagonal, and 0 everywhere else.

Proof. We prove by induction on the dimension of V. When V is one-dimensional, the only nilpotent operator is 0, so in this case, just take $v_1 = (1)$. Now, suppose the theorem holds for all vector spaces with dimension at most $n - 1$, and let $n := \dim V$. Now, let $U := \operatorname{im} N \leq V$. We have $\dim U < \dim V$ by nilpotence, and U is N-invariant. Hence, we apply the inductive assumption on $N|_U$, so there exists a basis

$$\gamma := \{N^{m_1} v_1, \dots, v_1, \dots, N^{m_s} v_s, \dots, v_s\}$$

of U. Each v_i lies in the range $U := \operatorname{im} N$, so $v_i = N u_i$ for some $u_i \in V$. Hence basis γ can be rewritten as

$$\gamma = \{N^{m_1+1} u_1, \dots, N u_1, \dots, N^{m_s+1} u_s, \dots, N u_s\}$$

Add the vectors $\{u_i\}_1^s$ to this basis: we claim that the set $\beta' = \{u_i\}_1^s \cup \gamma$ is linearly independent. Indeed, if there were a linear combination (Σ) of β' summing to 0, then applying N to (Σ) shifts the combination over, so all of the coefficients drop off by linear independence of the smaller set except possibly the ones in front of the $N^{m_j+1} u_j$. But these are independent anyway, so no problem actually exists. Now, extend β' to a basis β of V by adding $\{w_i\}_1^p$, where w_i are just arbitrary vectors added to compensate for the difference. Now $N w_i \in U \iff N w_i \in \operatorname{span} \gamma$, so there exist x_i in $V = \operatorname{span} \beta$ such that $N w_i = N x_i \iff N(w_i - x_i) = 0$. Setting $u_{n+k} = w_k - x_k$, we replace the w_k's with u_{n+k}'s and we can check that this still forms a basis, completing the proof. $\square$

97 Characteristic Polynomials: Worked Examples

Example 97.1 (8C.2)**.** Suppose V is a complex vector space. Suppose $T \in \mathcal{L}(V)$ is such that 5 and 6 are eigenvalues of T, and T has no other eigenvalues. Prove that if $\dim V =: n$, we have that $(T-5I)^{n-1}(T-6I)^{n-1} = 0$.

Proof. The characteristic polynomial $q(z)$ of T has the form $q(z) = (x-5)^i(x-5)^j$, where $i, j \leq n-1$ by counting. Clearly, q divides $(x-5)^{n-1}(x-6)^{n-1}$, so Cayley-Hamilton finishes the proof. $\square$

Example 97.2 (8C.3)**.** Give an example of an operator $T \in \mathcal{L}(\mathbb{C}^4)$ whose characteristic polynomial equals $(z-7)^2(z-8)^2$.

Solution. Take $T = \operatorname{diag}(7,7,8,8)$. $\bullet$

Example 97.3. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that there exists a polynomial $p \in \mathbb{F}[x]$ such that $T^{-1} = p(T)$.

Proof. Let $q(z) = a_0 + a_1 z + \cdots + z^n$ be the minimal polynomial of T. Then $q(T) = 0$, and in particular, we know $a_0 \neq 0$. Hence

$$0 = a_0 I + a_1 T + \cdots + a_{n-1}T^{n-1} + T^n \implies -a_0 I = a_1 T + \cdots + a_{n-1}T^{n-1} + T^n$$

$$\implies I = -\frac{1}{a_0}(a_1 T + \cdots + a_{n-1}T^{n-1} + T^n) \implies T^{-1} = -\frac{1}{a_0}(a_1 + \cdots + a_{n-1}T^{n-2} + T^{n-1}),$$

so we are done. $\square$

98 Jordan Form, Complexification

[8D] Jordan Form

As usual, let V denote a finite-dimensional vector space over $\mathbb{C}$. From Theorem 94.5, we deduce the following.

Theorem 98.1 (Jordan Canonical Form)**.** *Let $T \in \mathcal{L}(V)$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. Then there exists a basis of V in which T has the matrix $M_T = \operatorname{diag}(A_1, \ldots, A_n)$, where each A_j has λ_j's on the diagonal, 1's on the superdiagonal, and 0's everywhere else.*

Proof. By the GED, write $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_k, T)$. Then in appropriate bases of each $V_i := G(\lambda_i, T)$ under Theorem 94.5, we have that $(T - \lambda_i I)|_{V_i}$ has the matrix $\operatorname{diag}(B_{i1}, \ldots, B_{ij})$ where each $B_{i\ell}$ has 0 on the diagonal and 1 above. Now, $T = T - \lambda_i I + \lambda_i I$, and put all of these bases together. $\square$

Definition 98.2. The basis chosen in Theorem 98.1 is called a *Jordan basis* for T, and the matrix in Theorem 98.1 is the *Jordan (canonical) form* of T.

Example 98.3. Recall the matrix $T = \begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}$, with $G(6,T) = \operatorname{span}\{(1,0,0),(0,1,0)\}$ and $G(7,T) = \operatorname{span}\{(10,2,1)\}$. There are two possible options for the Jordan form of T:

$$\begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{pmatrix} \text{ or } \begin{pmatrix} 6 & 1 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

The first of these choices implies that $\dim E(6,T) = 2$, but we check that $\dim E(6,T) = 1$:

$$E(6,T) = \ker(T - 6I) = \ker \begin{pmatrix} 0 & 3 & 4 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \operatorname{span}\{(1,0,0)\}.$$

Hence, the second option is the correct Jordan form for T.

We comment that the number of Jordan blocks with the same eigenvalue is precisely $\dim E(\lambda, T)$.

Example 98.4. Let T have a matrix with λ along its diagonal, and 1 above. This means that T is already in Jordan form, which only has 1 block that is the entire matrix.

[9A] Complexification

Over real vector spaces, we cannot use the Jordan form; however, we can extend our real vector space to take advantage of it.

Definition 98.5. Let V be an $\mathbb{R}$-vector space. The *complexification of* V, denoted $V_{\mathbb{C}}$, is the space $V_{\mathbb{C}} := V \times iV$, where addition is done component wise and the scalar field is $\mathbb{C}$. Scalar multiplication is defined as follows for any $z = a + bi \in \mathbb{C}$ and $u + vi \in V_{\mathbb{C}}$:

$$(a + bi)(u + vi) = (au - bv) + i(bu + av).$$

Similarly, the complexification of an operator $T \in \mathcal{L}(V)$ is the operator $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ by $T_{\mathbb{C}}(u + vi) = Tu + iTv$.

It is readily checked that $V_{\mathbb{C}}$ is a $\mathbb{C}$-vector space, and $T_{\mathbb{C}}$ is actually $\mathbb{C}$-linear. Similarly, if $\{e_j\}$ is an $\mathbb{R}$-basis for V, $\{e_j + 0i\} = \{e_j\}$ is also a $\mathbb{C}$-basis for $V_{\mathbb{C}}$. Of course, the matrix of T matches that of $T_{\mathbb{C}}$ by the obvious embedding $V \hookrightarrow V_{\mathbb{C}}$.

Proposition 98.6. *Let V be an $\mathbb{R}$-vector space, and let $v \in V_{\mathbb{C}}$ be a (generalized) λ-eigenvector of some $T_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}})$. Then $\bar{v}$ is a (generalized) $\bar{\lambda}$-eigenvector of $T_{\mathbb{C}}$.*

103 Minimal Polynomials and Nilpotency

In this section, we view one proposition, which is occasionally useful.

Proposition 103.1. *Let V be a finite-dimensional complex vector space, and let $N \in \mathcal{L}(V)$ be nilpotent. Then the minimal polynomial of N is x^{m+1}, where m is the largest consecutive string of 1's on the superdiagonal of the Jordan form of N.*

Proof. Write $M_N = \operatorname{diag}(A_1, \ldots, A_k)$ in Jordan form, so each A_i has 0's on the diagonal[5], 1's on the superdiagonal, and 0's everywhere else. The start of a block coincides with an interruption of the string of 1's, so the number m (the longest consecutive string of 1's) is simply the number of 1's in the largest block. That is, m is one less than the maximum of the sizes of each A_i. Take n to be this maximum, so it is easy to see that $N, \ldots, N^{n-1} \neq 0$, but $N^n = 0$. Hence $m_T(x) = x^n = x^{m+1}$ as desired. ☐

[5]The only eigenvalue of a nilpotent operator is 0.

107 Nilpotency: Worked Examples

Example 107.1. Let $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if its minimal polynomial has nonzero constant term.

Proof. We have these equivalences: T is not invertible $\iff$ 0 is an eigenvalue for T $\iff$ $x \mid m_T \iff m_T$ has 0 as a constant term. Now take the contrapositive. $\square$

Example 107.2. Let $N \in \mathcal{L}(V)$, $\dim V = n$, and $\ker N^{n-1} \neq \ker N^n$. Prove that N is nilpotent and for all $0 \le m \le n$, we have $\dim \ker N^m = m$.

Proof. We know $\ker N^0 \le \ker N^1 \le \cdots \le \ker N^{n-1} \le \ker N^n$. If $\ker N^j = \ker N^{j+1}$ for $j < n$, we get a contradiction as

$$\ker N^j = \ker N^{j+1} = \cdots = \ker N^{n-1} = \ker N^n.$$

Hence, each inclusion is strict:

$$\ker N^0 < \ker N^1 < \cdots < \ker N^{n-1} < \ker N^n.$$

This implies $\dim \ker N^n \ge n$, but $\dim V = n$ so $\dim \ker N^n = n \iff \ker N^n = V \iff N^n = 0$, so N is nilpotent. By counting, we also have $\dim \ker N^m = m$ as desired. $\square$

Example 107.3. Give an operator $T \in \mathcal{L}(\mathbb{C}^7)$ such that $T^2 + T + I$ is nilpotent.

Solution. Take $\zeta_3 := \exp(2\pi i/3)$. Then the map $T : \mathbb{C}^7 \to \mathbb{C}^7$ by $v \mapsto \zeta_3 v$ satisfies $T^2+T+I = 0$, which is indeed nilpotent. •

Example 107.4. In contrast to Ex. 107.3, show that there cannot exist a $T \in \mathcal{L}(\mathbb{R}^7)$ such that $T^2 + T + I$ is nilpotent.

Proof. Assume for contradiction that such a T exists. Define $S := T^2+T+I$, and say that S has nilpotency degree n, so the minimal polynomial of S is $m_S(x) = x^n$. Hence $(T^2+T+I)^n = 0$, but this implies the eigenvalues of T are roots of x^2+x+1; i.e., ζ_3 and ζ_3^2. But also, we see that

$$m_T(x) \mid (x^2+x+1)^n \implies m_T \mid x^2+x+1$$

(by irreducibility)[6], so we have $m_T \in (\mathbb{C} \setminus \mathbb{R})[x]$, a contradiction. $\square$

108 Adjoints and Canonical Forms

Finally, we end these notes with a connection between the two main topics of this course: inner product spaces and generalized eigenspaces.

Theorem 108.1. *Let V be a finite-dimensional inner product space over $\mathbb{R}$ or $\mathbb{C}$. If $T \in \mathcal{L}(V)$ is self-adjoint, then $E(\lambda, T) = G(\lambda, T)$ for every eigenvalue λ of T.*

Proof. Apply the Spectral Theorem on T and diagonalize T — all Jordan blocks have size 1. $\square$

[6] Prime ideals are radical, and the set of all polynomials that vanish for T is the ideal generated by m_T.

Homework Exercises

What follows are my attempted solutions to every homework exercise that has been assigned this quarter. With the exception of HW #8, these homework assignments received full points, **but I cannot guarentee the correctness of the work here,** as most of the grade is completion-based. I have revised a few of these in these notes for clarity.

Homework 1

[5A] Invariant Subspaces

5A.12 Eigenvalues and Eigenvectors of a Polynomial Map

Define $T : \mathcal{P}_4(\mathbb{R}) \to \mathcal{P}_4(\mathbb{R})$ by $(Tp)(x) = xp'(x)$. Find all eigenvalues and eigenvectors of T.

Solution. Let $p \in \mathcal{P}_4(\mathbb{R})$. Then $p(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$, so that

$$p'(x) = 4a_4x^3 + 3a_3x^2 + 2a_2x + a_1, \text{ and}$$

$$(Tp)(x) = xp'(x) = 4a_4x^4 + 3a_3x^3 + 2a_2x^2 + a_1x$$

Let us suppose $p \neq 0$ is an eigenvector of T. Then $Tp = \lambda p$, i.e., we must have $xp'(x) = \lambda p(x)$, or

$$\begin{aligned} 4a_4x^4 + 3a_3x^3 + 2a_2x^2 + a_1x &= \lambda(a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0) \\ \implies 4a_4x^4 + 3a_3x^3 + 2a_2x^2 + a_1x &= \lambda a_4x^4 + \lambda a_3x^3 + \lambda a_2x^2 + \lambda a_1x + \lambda a_0. \end{aligned}$$

If $a_4 \neq 0$, then comparing coefficients tells us that $\lambda_1 = 4$ is an eigenvalue as long as $a_0 = a_1 = a_2 = a_3 = 0$. Similarly, if $a_4 = 0$ but $a_3 \neq 0$, then another eigenvalue is $\lambda_2 = 3$ by setting $a_0 = a_1 = a_2 = 0$. Continuing in this fashion, we have the following eigenvalues with their corresponding eigenvectors:

Eigenvalue	Eigenvectors
4	$p(x) = ax^4$, $a \in \mathbb{R}$
3	$p(x) = ax^3$, $a \in \mathbb{R}$
2	$p(x) = ax^2$, $a \in \mathbb{R}$
1	$p(x) = ax$, $a \in \mathbb{R}$
0	$p(x) = a$, $a \in \mathbb{R}$

•

5A.14 Eigenvalues and Eigenvectors of a Projection

Suppose $V = U \oplus W$, where $U, W \leq V$ are nontrivial. Define $P : V \to V$ by $P(u + w) = u$ for $u \in U$ and $w \in W$. Find all eigenvalues and eigenvectors of P.

Solution. Since $V = U \oplus W$, every $v \in V$ can be written uniquely in the form $v = u + w$, for $u \in U$ and $w \in W$. Now, suppose $0 \neq v = u + w \in V$ is an eigenvector for P, with eigenvalue λ. Then $P(u+w) = u = \lambda(u+w) = \lambda u + \lambda w$, so that $0 = (\lambda - 1)u + \lambda w \implies (1-\lambda)u = \lambda w$. Now, $U \cap V = \{0\}$, so $0 = (1-\lambda)u = \lambda w$. If u and w are both zero, then $v = u+w = 0+0 = 0$, which contradicts our assumption, so we consider two cases.

Case I: $u \neq 0$. In this case, we have $0 = (1-\lambda)u$, so $\boxed{\lambda = 1}$ is an eigenvalue. But $0 = \lambda w = 1w$, so $w = 0$, hence the eigenvectors corresponding to $\lambda = 1$ are exactly the nonzero $u \in U$.

Case II: $w \neq 0$. In this case, we have $0 = \lambda w \implies \boxed{\lambda = 0}$ is an eigenvalue. But $0 = (1-\lambda)u = 1u \implies u = 0$, so the eigenvectors corresponding to $\lambda = 0$ are exactly the nonzero $w \in W$. •

5A.15a Similar Transformations Have Same Eigenvalues

Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible. Show that T and $S^{-1}TS$ have the same eigenvalues.

Proof. Let λ be an eigenvalue of T. Then there exists some $v \in V$ with $Tv = \lambda v$. Notice that $S^{-1}T(v) = S^{-1}(\lambda v) = \lambda S^{-1}(v)$, but we can write $v = SS^{-1}v$ to obtain (using linearity of T and S, and thus linearity of S^{-1}):

$$\lambda S^{-1}v = S^{-1}Tv = S^{-1}T(SS^{-1}v) = S^{-1}TS(S^{-1}v).$$

Hence λ is an eigenvalue of T, with eigenvector $S^{-1}v$.

Conversely, let μ be an eigenvalue of $S^{-1}TS$. Then there exists some $u \in V$ with $S^{-1}TSu = \mu u$. A left-application of S yields $TSu = S(\mu u) = \mu Su$, so μ is an eigenvalue of T with eigenvector Su. Thus, T and $S^{-1}TS$ have the same eigenvalues. □

5A.16 Conjugate of an Eigenvalue

Suppose V is a complex vector space, $T \in \mathcal{L}(V)$, and the matrix of T with respect to some basis of V contains only real entries. Show that if λ is an eigenvalue of T, then so is $\bar{\lambda}$.

Proof. Let $E \subset V$ be a basis for V, and let $0 \neq v \in V$ be an λ-eigenvector of $T \in \mathcal{L}(V)$. Since E is a basis, we can write $v = \sum\limits_{e \in E} a_e e$, for appropriate scalars a_e. Since T is determined by its values on the basis E, for each $u \in E$ write $Tu = \sum\limits_{e \in E} A_{u,e} e$. Hence with some work we see

$$\lambda \sum_{e \in E} a_e e = \lambda v = Tv = T\left(\sum_{e \in E} a_e e\right) = \sum_{e \in E} a_e Te = \sum_{u \in E} a_u Tu = \sum_{u \in E}\sum_{e \in E} a_u A_{u,e} e.$$

Now, taking a complex conjugate yields (noting $\overline{\lambda v} = \overline{Tv} = \bar{T}\bar{v}$):

$$\overline{\lambda v} = \overline{\lambda \sum_{e \in E} a_e e} = \overline{T}\left(\overline{\sum_{e \in E} a_e e}\right) = \sum_{e \in E} \overline{a_e}\,\overline{T}e.$$

But since the matrix of T has only real entries, we have $\overline{T} = T$, so the above can be rewritten as

$$\overline{\lambda v} = \sum_{e \in E} \overline{a}_e \overline{T}e = \sum_{u \in E} \overline{a_u} Tu = \sum_{u \in E}\sum_{e \in E} \overline{a_u} A_{u,e} e = T\overline{v},$$

by comparing our first computation. Hence $\bar{\lambda}\bar{v} = T\bar{v}$, so $\bar{v}$ is an $\bar{\lambda}$-eigenvector of T. □

5A.19 Eigenvalues and Eigenvectors of a Matrix of all 1's

Suppose n is a positive integer and $T \in \mathcal{L}(\mathbb{F}^n)$ is defined by

$$T(x_1, \ldots, x_n) = \left(\sum_{i=1}^n x_i, \ldots, \sum_{i=1}^n x_i \right),$$

in other words, T is the operator whose matrix with respect to the standard basis consists of all 1's. Find all eigenvalues and eigenvectors of T.

Solution. Let $x = (x_1, \ldots, x_n) \neq (0, 0, \ldots, 0)$ be an eigenvector of T with eigenvalue λ. Then $Tx = \lambda x$ for some $\lambda \in \mathbb{F}$, or we have that

$$\sum_{i=1}^n x_i = \lambda x_j$$

for every $1 \leq j \leq n$. Hence $\lambda x_1 = \lambda x_2 = \ldots = \lambda x_n = \sum x_i$. Let us consider two cases, where $\lambda = 0$ and $\lambda \neq 0$.

Case I: $\lambda \neq 0$. In this case, we multiply by λ^{-1} to obtain $x_1 = x_2 = \ldots = x_n$; i.e., the eigenvectors corresponding to a nonzero λ is exactly the subspace $\boxed{\langle (1, 1, \ldots, 1) \rangle}$. Now $T(1, 1, \ldots, 1) = n \cdot 1 = n$, so $\boxed{\lambda = n}$.

Case II: $\lambda = 0$. In this case, we have $\lambda x_j = 0 = \sum x_i$ for any $1 \leq j \leq n$, so the eigenvectors in this case are exactly the vectors $(x_1, \ldots, x_n) \in \mathbb{F}^n$ such that $x_n = -(x_1 + \ldots + x_{n-1})$. In particular, these vectors are allowed to be nonzero, so $\boxed{\lambda = 0}$ is indeed an eigenvalue. •

5A.26 When All Non-trivial Vectors are Eigenvectors

Suppose $T \in \mathcal{L}(V)$ is such that every nonzero vector $v \in V$ is an eigenvector of T. Prove that T is a scalar multiple of the identity operator.

Proof. Since every nonzero vector $v \in V$ is an eigenvector, it suffices to show that T has exactly one eigenvalue. Let λ, μ be eigenvalues of T with eigenvectors $v, w \in V$ respectively. Then $Tv = \lambda v$ and $Tw = \mu w$. We consider two cases to show that $\lambda = \mu$.

Case I: v and w are linearly dependent. Then $w = cv$ for some scalar c, so that $Tw = T(cv) = cTv = c\lambda v = \lambda(cv) = \lambda w$. Hence $\lambda w = \mu w \iff (\lambda - \mu)w = 0$, and since $w \neq 0$ (as it is an eigenvector), we have $\lambda = \mu$.

Case II: v and w are linearly independent. Here, notice that $T(u+v) = Tu + Tv = \lambda u + \mu v$, but by linear independence, notice $u + v \neq 0$, so $u + v$ is an eigenvector. Hence, there exists some ξ such that $T(u+v) = \xi(u+v) = \xi u + \xi v$, so that

$$\lambda u + \mu v = \xi u + \xi v \implies (\xi - \lambda)u + (\xi - \mu)v = 0.$$

Since v and w are linearly independent, we must have $\xi - \lambda = 0$ and $\xi - \mu = 0$, hence $\xi = \lambda = \mu$.

It thus follows that $Tv = \lambda v$ for all $v \in V$, so that $T = \lambda I$, where I is the identity operator. □

Homework 2

[5B] Eigenvalues and Upper-Triangular Matrices

5B.2 Product of "Eigentransformations"

Suppose that $T \in \mathcal{L}(V)$ and $(T-2I)(T-3I)(T-4I) = 0$. Suppose λ is an eigenvalue of T. Prove that $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$.

Proof. Let λ be an eigenvalue of T, and let $v \neq 0$ be an eigenvector with eigenvalue λ. Then $Tv = \lambda v$, so that $(T-4I)(v) = Tv - 4I(v) = \lambda v - 4v = (\lambda - 4)v$. Similarly, we see that

$$(T-3I)(T-4I)v = (T-3I)(\lambda-4)v = (\lambda-4)(T-3I)v = (\lambda-4)(\lambda-3)v,$$

so continuing this inductively we see that

$$(T-2I)(T-3I)(T-4I)v = (\lambda-4)(\lambda-3)(\lambda-2)v.$$

But $(T-2I)(T-3I)(T-4I)v = 0$, and since v is nonzero, we must have $\lambda_{1,2,3} = 2, 3, 4$, which completes the proof. □

5B.4 Kernel and Image of Idempotent Transformation

Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$. Prove that $V = \ker P \oplus \operatorname{im} P$.

Proof. It suffices to show that $V \subseteq \ker P + \operatorname{im} P$, and that $\ker P \cap \operatorname{im} P = \{0\}$. Take $v \in V$. If $v \in \operatorname{im} P$, then clearly $v = 0 + v \in \ker P + \operatorname{im} P$. If $v \notin \operatorname{im} P$, we still have $Pv \in \operatorname{im} P$, so that $v = (v - Pv) + Pv$. We claim that $v - Pv \in \ker P$; to see this, note that

$$P(v - Pv) = Pv - PPv = Pv - P^2 v = Pv - Pv = 0.$$

Hence $v = (v - Pv) + Pv \in \ker P + \operatorname{im} P$.

Now, take $u \in \ker P \cap \operatorname{im} P$. Then $Pu = 0$, and there exists some $w \in V$ such that $Pw = v$. But now $P^2 w = Pv = 0$, but $P^2 = P$, so $v = Pw = P^2 w = 0$, hence $\ker P \cap \operatorname{im} P = \{0\}$. This shows that $V = \ker P \oplus \operatorname{im} P$. □

5B.7 Square Root of an Eigenvalue

Suppose $T \in \mathcal{L}(V)$. Prove that 9 is an eigenvalue of T^2 if and only if 3 or -3 is an eigenvalue of T.

Proof. ($\Longleftarrow$): Without loss of generality, suppose 3 is an eigenvalue of T. Then there exists $v \neq 0$ such that $Tv = 3v$. Now $T^2 v = TTv = T(3v) = 3Tv = 3 \cdot 3v = 9v$, so we immediately see that v is an eigenvector for T^2, with eigenvalue 9. The case where -3 is an eigenvalue is similar.

($\Longrightarrow$): Suppose that 9 is an eigenvalue of T^2. Then there exists $v \neq 0$ such that $T^2 v = 9v \iff (T^2 - 9I)v = 0$. Noting that $(T-3I)(T+3I) = T^2 - 9I$, we observe that $(T-3I)(T+3I)v = 0$. From here, we note that it is not possible that both $(T-3I)$ and $(T+3I)$ are injective, otherwise the composition $(T-3I)(T+3I) = (T^2 - 9I)$ is injective, implying that $v = 0$, which contradicts our assumption. Thus, this implies that there exists $u \neq 0$ such that either $(T-3I)u = 0 \iff Tu = 3u$ or $(T+3I)u = 0 \iff Tu = -3u$; i.e., at least one of 3 and -3 is an eigenvalue of T. □

[5C] Eigenvalues and Diagonal Matrices

5C.3 Conditions for a Direct Sum of Kernel and Image

Suppose that V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent:

(a) $V = \ker T \oplus \operatorname{im} T$;

(b) $V = \ker T + \operatorname{im} T$;

(c) $\ker T \cap \operatorname{im} T = \{0\}$.

We first state a theorem from Math 121A that makes this easier, which was proved by tediously choosing a basis for $W_1 \cap W_2$, extending them to bases of W_1 and W_2, and doing a linear independence argument:

Theorem. *If $W_1, W_2 \leq V$ are subspaces and $\dim V < \infty$, then*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Now, we prove the result.

Proof. Clearly, $(a) \implies (b)$ and $(a) \implies (c)$ hold, so it suffices to show that $(b) \implies (c)$ and $(c) \implies (a)$ hold.

$(b) \implies (c)$: Let $n := \dim V$, and suppose $\dim \ker T = k$ for some $k \leq n$. By the First Isomorphism Theorem, we know $\operatorname{im} T \cong V/\ker T$, so that $\dim \operatorname{im} T = \dim(V/\ker T) = n - k$. But by Theorem 1, we must have $\dim(\ker T \cap \operatorname{im} T) = 0$, which means that $\ker T \cap \operatorname{im} T = \{0\}$.

$(c) \implies (a)$: Suppose that $\ker T \cap \operatorname{im} T = \{0\}$. It suffices to show that $V = \ker T + \operatorname{im} T$. But this is really a restatement of the above, as if we set $\dim \ker T = k$ (and thus $\dim \operatorname{im} T = n - k$ by the Isomorphism Theorem), we must have by Theorem 1

$$\dim(\ker T + \operatorname{im} T) = \dim \ker T + \dim \operatorname{im} T - \dim(\ker T \cap \operatorname{im} T)$$

$$= k + (n - k) + \dim\{0\} = k + n - k + 0 = n.$$

But $\ker T + \operatorname{im} T \leq V$, so we must have $V = \ker T + \operatorname{im} T$. That $V = \ker T \oplus \operatorname{im} T$ follows from the fact that the kernel and image have trivial intersection, which we assumed. □

5C.5 Diagonalizable Iff Eigentransformations Split into Direct Sum

Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if $V = \ker(T - \lambda I) \oplus \operatorname{im}(T - \lambda I)$ for every $\lambda \in \mathbb{C}$.

Proof. $(\implies)$: Suppose $\dim V =: n < \infty$, and let T be diagonalizable. Then T has a diagonal matrix representation $M_T = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ in some basis. Then $T - \lambda I$ is also diagonalizable, as its matrix representation is $M_T - \lambda I = \operatorname{diag}(\lambda_1 - \lambda, \ldots, \lambda_n - \lambda)$. Denote $S := T - \lambda I$. Then $V = E(\mu_1, S) \oplus E(\mu_2, S) \oplus \cdots \oplus E(\mu_m, S)$, where the μ_i are eigenvalues of S.

If $\mu_i = 0$, then if $v \in E(\mu_i, S)$, then $Sv = \mu_i v = 0v = 0$, so $v \in \ker S$. Thus, $E(0, S) \leq \ker S$. Else, if $\mu_i \neq 0$, then if $v \in E(\mu_i, S)$, we notice that $S(\mu_i^{-1} v) = \mu_i \mu_i^{-1} v = v$, so $v \in \operatorname{im}(S)$, so $E(\mu_i, S) \leq \operatorname{im}(S)$ if $\mu_i \neq 0$. Now, if 0 is an eigenvalue of S, without loss of generality set $\mu_1 = 0$, so that if $v \in V$, we have $v = v_1 + (v_2 + \cdots + v_n)$ for appropriate v_i, so that $v \in \ker S + \operatorname{im} S$, so that $V \leq \ker S + \operatorname{im} S \implies V = \ker S + \operatorname{im} S$. By Exercise 3, we have that $V = \ker S \oplus \operatorname{im} S = \ker(T - \lambda I) \oplus \operatorname{im}(T - \lambda I)$, as desired.

($\Longleftarrow$): We prove by induction on $\dim V =: n$. If $n = 1$, every operator is diagonalizable, so this is obvious.

Now, suppose that the statement holds for all complex vector spaces with dimension less than n, for $n \geq 2$. Let V be a complex vector space $\dim V = n$, and let $T \in \mathcal{L}(V)$. Now, T has an eigenvalue λ_1 (as we are working over $\mathbb{C}$), so by assumption we know that $V = \ker(T - \lambda_1 I) \oplus \operatorname{im}(T - \lambda_1 I)$, but $\ker(T - \lambda_1 I) = E(\lambda_1, T)$, so $V = E(\lambda_1, T) \oplus \operatorname{im}(T - \lambda_1 I)$.

If we let $U = \operatorname{im}(T - \lambda_1 I) \leq V$, notice that U is T-invariant, so it suffices to show that the restriction $S := T|_U$ is diagonalizable. But $\dim U \leq \dim V$, so we can just show $U = \ker(S - \lambda I) \oplus \operatorname{im}(S - \lambda I)$ for each $\lambda \in \mathbb{C}$. But by Exercise 3, it is enough to show $\ker(S - \lambda I) \cap \operatorname{im}(S - \lambda I) = \{0\}$, but we know that S is just a restriction of T. Hence $\ker(S - \lambda I) \leq \ker(T - \lambda I)$ and $\operatorname{im}(S - \lambda I) \leq \operatorname{im}(T - \lambda I)$, and we already have $\ker(T - \lambda I) \cap \operatorname{im}(T - \lambda I) = \{0\}$ by assumption. Hence $\ker(S - \lambda I) \cap \operatorname{im}(S - \lambda I) = \{0\}$, so $T|_U$ is diagonalizable and $\operatorname{im}(T - \lambda_1 I)$ is a direct sum of eigenspaces, which implies V is as well; i.e., T is diagonalizable. □

5C.16 The Fibonacci Sequence

The Fibonacci sequence (F_n) is defined by

$$F_1 = F_2 = 1 \text{ and } F_n = F_{n-2} + F_{n-1} \text{ for } n \geq 3.$$

Define $T \in \mathcal{L}(\mathbb{R}^2)$ by $T(x, y) = (y, x + y)$.

(a) *Show that $T^n(0, 1) = (F_n, F_{n+1})$ for each positive integer n.*

Proof. We prove by induction on n. For $n = 1$, we have $T^1(0, 1) = T(0, 1) = (1, 0 + 1) = (1, 1) = (F_1, F_2)$. Now, suppose $T^k(0, 1) = (F_k, F_{k+1})$ for all $k < n$, for some $n \geq 2$. Now by the inductive hypothesis,

$$T^n(0, 1) = T(T^{n-1}(0, 1)) = T(F_{n-1}, F_n) = (F_n, F_{n-1} + F_n).$$

Now we see that $F_{n+1} = F_n + F_{n-1}$, so $T^n(0, 1) = (F_n, F_{n+1})$, thus completing the proof by induction. □

(b) *Find the eigenvalues of T.*

Solution. Let $v = (x, y) \neq (0, 0)$ be an eigenvector of T, and let λ be its associated eigenvalue. Then we have $Tv = \lambda v$, or $(y, x + y) = (\lambda x, \lambda y)$, so that $y = \lambda x$ and $x + y = \lambda y$. This is a system of equations, so we substitute the first into the second and write $x + y = \lambda y \implies x + \lambda x = \lambda^2 x \implies x(\lambda^2 - \lambda - 1) = 0$. If $x = 0$, then we see $y = 0$, which contradicts our assumption that v is nonzero, so we have $\lambda^2 - \lambda - 1 = 0 \implies \boxed{\lambda_{1,2} = \dfrac{1 \pm \sqrt{5}}{2}}$. •

(c) *Find a basis of $\mathbb{R}^2$ consisting of eigenvectors of T.*

Solution. For brevity, denote the two eigenvalues of T as $\varphi := \frac{1+\sqrt{5}}{2}$ and $\psi := \frac{1-\sqrt{5}}{2}$.

For the eigenvalue φ, we see that if $v = (x, y)$ is a φ-eigenvector, then $Tv = \varphi v \implies (y, x + y) = (\varphi x, \varphi y)$. Letting $x = 1$ (this is legal as eigenvectors are unique up to rescaling), we see that $(y, 1 + y) = (\varphi, \varphi y)$, so $y = \varphi$, so $v = (1, \varphi)$ is a φ-eigenvector. [Indeed, we see that $\varphi^2 = \varphi + 1$, so the calculation does check out.]

For the eigenvalue ψ, we see that if $w = (x, y)$ is a ψ-eigenvector, then $Tw = \psi w \implies (y, x+y) = (\psi x, \psi y)$. Again, picking $x = 1$, we see that $(y, 1+y) = (\psi, \psi y)$ so that $w = (1, \psi)$ is a ψ-eigenvector.

Since $\varphi \neq \psi$, it is easy to see that $\{v, w\}$ is linearly independent, and thus a basis for $\mathbb{R}^2$ is given by

$$\{v, w\} = \{(1, \varphi), (1, \psi)\} = \boxed{\left\{\left(1, \frac{1+\sqrt{5}}{2}\right), \left(1, \frac{1-\sqrt{5}}{2}\right)\right\}}.$$

•

(d) *Use the solution to part (c) to compute $T^n(0,1)$. Conclude that*

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right] = \frac{1}{\sqrt{5}}(\varphi^n - \psi^n).$$

Proof. Fix the same notation as in part (c). Notice that

$$v - w = (1, \varphi) - (1, \psi) = \left(0, \frac{1+\sqrt{5}-(1-\sqrt{5})}{2}\right) = \left(0, \sqrt{5}\right),$$

so that $(0,1) = \frac{1}{\sqrt{5}}(v - w)$. Now note that $Tv = \varphi v$ and $Tw = \psi w$, so we have

$$T^n(0,1) = T^n\left(\frac{1}{\sqrt{5}}(v-w)\right) = \frac{1}{\sqrt{5}}T^n(v-w)$$

$$= \frac{1}{\sqrt{5}}(T^n v - T^n w) = \boxed{\frac{1}{\sqrt{5}}(\varphi^n v - \psi^n w)}.$$

Now, by expanding the above, we see that

$$T^n(0,1) = \frac{1}{\sqrt{5}}(\varphi^n v - \psi^n w) = \frac{1}{\sqrt{5}}[\varphi^n(1,\varphi) - \psi^n(1,\psi)]$$

$$= \frac{1}{\sqrt{5}}[(\varphi^n, \varphi^{n+1}) - (\psi^n, \psi^{n+1})] = \left(\frac{1}{\sqrt{5}}(\varphi^n - \psi^n), \frac{1}{\sqrt{5}}(\varphi^{n+1} - \psi^{n+1})\right).$$

But by part (a), we see that $T^n(0,1) = (F_n, F_{n+1})$, so that by matching components, we have $F_n = \frac{1}{\sqrt{5}}(\varphi^n - \psi^n)$, as desired. □

(e) *Use part (d) to conclude that for each positive integer n, the Fibonacci number F_n is the integer that is closest to $\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n = \frac{\varphi^n}{\sqrt{5}}$.*

Proof. Write $F_n = \frac{1}{\sqrt{5}}\varphi^n - \frac{1}{\sqrt{5}}\psi^n$. Clearly from part (a), F_n is an integer, so it suffices to show that

$$\left|F_n - \frac{\varphi^n}{\sqrt{5}}\right| < \frac{1}{2}.$$

Now

$$\left|F_n - \frac{\varphi^n}{\sqrt{5}}\right| = \left|-\frac{1}{\sqrt{5}}\psi^n\right| = \left|\frac{1}{\sqrt{5}}\right| \cdot \left|\frac{1-\sqrt{5}}{2}\right|^n.$$

Noting that $2 < \sqrt{5} < 3$, we see that $1/\sqrt{5} < 1/2$ and

$$-3 < -\sqrt{5} < -2 \implies -2 < 1 - \sqrt{5} < -1 < 0 \implies -1 < \frac{1-\sqrt{5}}{2} < 0,$$

so that $\left|\frac{1-\sqrt{5}}{2}\right| < 1 \implies \left|\frac{1-\sqrt{5}}{2}\right|^n < 1$. Hence

$$\left|F_n - \frac{\varphi^n}{\sqrt{5}}\right| = \left|-\frac{1}{\sqrt{5}}\psi^n\right| = \left|\frac{1}{\sqrt{5}}\right| \cdot \left|\frac{1-\sqrt{5}}{2}\right|^n < \frac{1}{2} \cdot 1 = \frac{1}{2},$$

which shows that F_n is the integer closest to $\varphi^n/\sqrt{5}$. □

Homework 3

[6A] Inner Products and Norms

6A.11 A Basic Cauchy-Schwarz Inequality Proof

Prove that

$$16 \leq (a+b+c+d)\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right)$$

for all positive integers a, b, c, d.

We first derive a special case of the Cauchy-Schwarz inequality in $\mathbb{R}^4$. Take two vectors $x = (x_1, x_2, x_3, x_4), y = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4$, where the inner product is the usual dot product. The Cauchy-Schwarz inequality thus tells us that

$$\langle x, y\rangle \leq \|x\| \, \|y\| \implies \langle x, y\rangle^2 \leq \|x\|^2 \, \|y\|^2 = \langle x, x\rangle \langle y, y\rangle .$$

For the regular dot product in $\mathbb{R}^4$, we thus have

$$(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 \leq (x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2). \tag{1}$$

Now, we apply equation (1) to the exercise.

Proof. Fix $a, b, c, d \in \mathbb{R}^+$, and let $x = \left(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}\right)$ and $y = \left(\frac{1}{\sqrt{a}}, \frac{1}{\sqrt{b}}, \frac{1}{\sqrt{c}}, \frac{1}{\sqrt{d}}\right)$. By (1) we immediately see that

$$\begin{aligned} 16 = 4^2 = (1+1+1+1)^2 &= \left(\frac{\sqrt{a}}{\sqrt{a}} + \frac{\sqrt{b}}{\sqrt{b}} + \frac{\sqrt{c}}{\sqrt{c}} + \frac{\sqrt{d}}{\sqrt{d}}\right)^2 \\ &\leq \left((\sqrt{a})^2 + (\sqrt{b})^2 + (\sqrt{c})^2 + (\sqrt{d})^2\right)\left(\left(\frac{1}{\sqrt{a}}\right)^2 + \left(\frac{1}{\sqrt{b}}\right)^2 + \left(\frac{1}{\sqrt{c}}\right)^2 + \left(\frac{1}{\sqrt{d}}\right)^2\right) \\ &= (a+b+c+d)\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right), \end{aligned}$$

so we are done. □

[Note that this is in more detail than I would have preferred.]

6A.19 Norm Identity I

Suppose V is a real inner product space. Prove that for all $u, v \in V$,

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}.$$

Proof. Let $u, v \in V$, where V is a **real** vector space. Then $\langle u, v\rangle = \overline{\langle v, u\rangle} = \langle v, u\rangle$, so that $\langle u, -v\rangle = -\langle u, v\rangle = \langle -u, v\rangle$, so we use these to simplify as follows:

$$\frac{\|u+v\|^2 - \|u-v\|^2}{4} = \frac{\langle u+v, u+v\rangle - \langle u-v, u-v\rangle}{4}$$
$$= \frac{(\langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle) - (\langle u,u\rangle + \langle u,-v\rangle + \langle -v,u\rangle + \langle -v,-v\rangle)}{4}$$
$$= \frac{\langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle - \langle u,u\rangle + 2\langle u,v\rangle - \langle v,v\rangle}{4}$$
$$= \frac{2\langle u,v\rangle + 2\langle u,v\rangle}{4} = \frac{4\langle u,v\rangle}{4} = \langle u,v\rangle,$$

which completes the proof. □

6A.27 Norm Identity II

Suppose $u, v, w \in V$. Prove that

$$\left\|w - \frac{1}{2}(u+v)\right\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}. \tag{2}$$

Discussion: This equality roughly resembles the parallelogram equality, so we try to rearrange it to see if anything is useful. Note that for any scalar $\alpha \in F$ and any $v \in V$, we have $\|\alpha v\| = |\alpha|\,\|v\|$. We write

$$\left\|w - \frac{1}{2}(u+v)\right\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}$$
$$\iff \left\|w - \frac{1}{2}(u+v)\right\|^2 + \frac{\|u-v\|^2}{4} = \frac{\|w-u\|^2}{2} + \frac{\|w-v\|^2}{2}$$
$$\iff \left\|w - \frac{1}{2}(u+v)\right\|^2 + \left(\frac{\|u-v\|}{2}\right)^2 = \frac{\|w-u\|^2}{2} + \frac{\|w-v\|^2}{2}$$
$$\iff \left\|w - \frac{1}{2}(u+v)\right\|^2 + \left\|\frac{u-v}{2}\right\|^2 = \frac{\|w-u\|^2}{2} + \frac{\|w-v\|^2}{2}.$$

This suggests to us based on the parallelogram equality that we should find vectors u', v' such that

$$u' + v' = w - \frac{1}{2}(u+v) \text{ and } u' - v' = \frac{1}{2}(u-v).$$

Combining the two equations, we see

$$2u' = w - \frac{1}{2}(u+v) + \frac{1}{2}(u-v) = w - \frac{1}{2}u - \frac{1}{2}v + \frac{1}{2}u - \frac{1}{2}v = w - v,$$

so that $u' = (w-v)/2$. Hence,

$$v' = u' - \frac{1}{2}(u-v) = \frac{1}{2}(w-v) - \frac{1}{2}(u-v) = \frac{1}{2}(w-v-u+v) = \frac{1}{2}(w-u).$$

We present our formal proof below.

Proof 1. Fix $u, v, w \in V$, and define $u' := \frac{1}{2}(w-v)$ and $v' = \frac{1}{2}(w-u)$. Then by the parallelogram equality, we have

$$\|u'+v'\|^2 + \|u'-v'\|^2 = 2\left(\|u'\|^2 + \|v'\|^2\right).$$

Now, observe that $u'+v' = \frac{1}{2}(w-v+w-u) = w - \frac{1}{2}(u+v)$ and $u'-v' = \frac{1}{2}(w-v) - \frac{1}{2}(w-u) = \frac{1}{2}(w-v-w+u) = \frac{1}{2}(u-v)$. Observing that for any scalar $\alpha \in F$ and any $x \in V$, we have $\|\alpha x\| = |\alpha|\,\|x\|$, we write

$$\begin{aligned}
&\left\|w - \frac{1}{2}(u+v)\right\|^2 + \left\|\frac{u-v}{2}\right\|^2 = 2\|u'\|^2 + 2\|v'\|^2 \\
\iff &\left\|w - \frac{1}{2}(u+v)\right\|^2 + \frac{\|u-v\|^2}{2^2} = 2\left\|\frac{w-v}{2}\right\|^2 + 2\left\|\frac{w-u}{2}\right\|^2 \\
\iff &\left\|w - \frac{1}{2}(u+v)\right\|^2 + \frac{\|u-v\|^2}{4} = \frac{2\|w-v\|^2}{4} + \frac{2\|w-u\|^2}{4} \\
\iff &\left\|w - \frac{1}{2}(u+v)\right\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4},
\end{aligned}$$

which completes the proof. □

Alternatively, we can brute-force our way through. In the proof below, we write inner products via juxtaposition for clarity purposes. This means that we have $uv = \overline{vu}$ and $u(\lambda v) = \bar{\lambda} uv$. We also shorten $\langle v, v\rangle$ as just v^2. This is *not* the same as $\|v\|^2$ below.

Proof 2. By direct computation, we expand out both sides. On one side,

$$\begin{aligned}
&\frac{1}{2}(\|w-u\|^2 + \|w-v\|^2) - \frac{1}{4}\|u-v\|^2 \\
&= \frac{1}{2}[(w-u)(w-u) + (w-v)(w-v)] - \frac{1}{4}(u-v)(u-v) \\
&= \frac{1}{2}(ww + uu - uw - wu) + \frac{1}{2}(ww + vv - vw - wv) - \frac{1}{4}(u-v)(u-v) \\
&= ww + \frac{1}{2}uu + \frac{1}{2}vv - \frac{1}{2}(uw + wu + vw + wv) - \frac{1}{4}(uu + vv - vu - uv) \\
&= w^2 + \frac{1}{2}(u^2 + v^2) - \frac{1}{2}(uw + wu + vw + wv) - \frac{1}{4}(u^2 + v^2) + \frac{1}{4}(vu + uv) \\
&= w^2 + \frac{1}{4}(u^2 + v^2) - \frac{1}{2}(\mathrm{Re}(uw) + \mathrm{Re}(vw)) + \frac{1}{4}(\mathrm{Re}(vu))
\end{aligned}$$

On the other side,

$$\begin{aligned}\left\|w-\frac{1}{2}(u+v)\right\|^2 &= \left(w-\frac{1}{2}(u+v)\right)\left(w-\frac{1}{2}(u+v)\right)\\ &= ww+\frac{1}{4}(u+v)^2-\frac{1}{2}(u+v)w-(w)\frac{1}{2}(u+v)\\ &= w^2+\frac{1}{4}(u+v)(u+v)-\frac{1}{2}((u+v)w+w(u+v))\\ &= w^2+\frac{1}{4}(uu+vv+vu+uv)-\frac{1}{2}(uw+vw+wu+wv)\\ &= w^2+\frac{1}{4}(u^2+v^2)-\frac{1}{2}(\mathrm{Re}(uw)+\mathrm{Re}(vw))+\frac{1}{4}(\mathrm{Re}(vu)),\end{aligned}$$

so both sides must be equal. $\square$

6A.28 Closest Points

Suppose $C \subseteq V$ with the property that $u, v \in C$ implies $\frac{1}{2}(u+v) \in C$. Let $w \in V$. Show that there is at most one point in C that is closest to w. In other words, show that there is at most one point $u \in C$ such that $\|w-u\| \le \|w-v\|$ for all $v \in C$. **Hint:** *Use the previous exercise.*

Proof. Fix $w \in V$. For contradiction, let $u, v \in C$ be such that $u \neq v$ and that for all $x \in C$, we have $\|w-u\|, \|w-v\| \le \|w-x\|$. That is, assume that there are two distinct "closest" vectors u, v of w. Then by applying equation (2) from Exercise 27, we see that (by appealing to minimality of u)

$$\left\|w-\frac{1}{2}(u+v)\right\|^2 = \frac{\|w-u\|^2+\|w-v\|^2}{2}-\frac{\|u-v\|^2}{4}$$

$$\implies \left\|w-\frac{1}{2}(u+v)\right\|^2 \le \frac{\|w-u\|^2}{2}+\frac{\|w-u\|^2}{2}-\frac{\|u-v\|^2}{4} \le \|w-u\|^2,$$

so that $\left\|w-\frac{1}{2}(u+v)\right\| \le \|w-u\|$, and thus $\frac{1}{2}(u+v) \in C$ is closer to w than u, contradicting the fact that we assumed that u was closest to w. Thus, there can only be at most one vector in C that is closest to w. $\square$

[6B] Orthonormal Bases

6B.2 Condition for Vector Lying in Span of Orthonormal System

Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V, and let $v \in V$. Prove that $\|v\|^2 = \sum_{i=1}^{m} |\langle v, e_i\rangle|^2$ if and only if $v \in \langle e_1, \ldots, e_m\rangle$.

Proof. This is Example 34.8 in the main text. We print a more explicit version here.

($\implies$) : Suppose $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$. Now, since $\{e_i\}_{i=1}^m$ is linearly independent, extend $\{e_i\}_{i=1}^m$ to be a basis $\{e_i\}_{i=1}^m \cup \{v_j\}_{j>m}$ of V. Then $V = \sum a_i e_i + \sum b_j v_j$, so

$$\|v\|^2 = \langle v, v\rangle = \left\langle \sum_i a_i e_i + \sum_j b_j v_j, v \right\rangle = \sum_i \langle a_i e_i, v\rangle + \sum_j \langle b_j v_j, v\rangle = \sum_{i=1}^{m} |\langle v, e_i\rangle|^2.$$

Equating like terms, we see that we must have $b_j = 0$, so that we must have $v \in \langle e_1, \dots, e_m \rangle$.

$(\Longleftarrow)$: Suppose $v \in \langle e_1, \dots, e_m \rangle$. Then it follows from Lemma 34.6 that

$$v = \sum_{i=1}^{m} \langle v, e_i \rangle e_i, \text{ so that } \|v\|^2 = \langle v, v \rangle = \sum_{i=1}^{m} \langle v, e_i \rangle,$$

from Proposition 34.7. This completes the proof. □

6B.14 Small Distance to Orthonormal Basis

Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $v_1, \dots, v_n$ are vectors in V such that $\|e_j - v_j\| < 1/\sqrt{n}$ for each j. Prove that $v_1, \dots, v_n$ is a basis for V.

Proof. Suppose for contradiction that $\{v_i\}_{i=1}^n$ is **not** a basis for V. Because we have n vectors v_i, this means that the v_i are linearly dependent, so there exists some $a_i \in F$, not all zero, such that $\sum_{i=1}^{n} a_i v_i = 0$. Observe that $\sum_{i=1}^{n} a_i v_i = \sum_{i=1}^{n} a_i(v_i - e_i + e_i) = \sum_{i=1}^{n} a_i(v_i - e_i) + \sum_{i=1}^{n} a_i e_i$, hence

$$\left\| \sum_{i=1}^{n} a_i(v_i - e_i) \right\| = \left\| \sum_{i=1}^{n} a_i e_i \right\|. \tag{3}$$

Now, notice that by the Triangle Inequality we have

$$\left\| \sum_{i=1}^{n} a_i(v_i - e_i) \right\| \leq \sum_{i=1}^{n} \|a_i(v_i - e_i)\| = \sum_{i=1}^{n} |a_i| \|v_i - e_i\|,$$

so by our assumption that $\|e_i - v_i\| < 1/\sqrt{n}$ for all $i \leq n$, we have

$$\sum_{i=1}^{n} |a_i| \|v_i - e_i\| < \sum_{i=1}^{n} |a_i| \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |a_i|.$$

Now, by Cauchy-Schwarz, we see that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} |a_i| = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (1 \cdot |a_i|) \leq \frac{1}{\sqrt{n}} \cdot \sqrt{\sum_{i=1}^{n} 1 \cdot \sum_{i=1}^{n} |a_i|^2} = \frac{\sqrt{n}}{\sqrt{n}} \cdot \sqrt{\sum_{i=1}^{n} |a_i|^2}$$

$$\implies \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |a_i| \leq \sqrt{\sum_{i=1}^{n} |a_i|^2},$$

but by the fact that the e_i are orthonormal, we see that $\left\| \sum_{i=1}^{n} a_i e_i \right\| = \sqrt{\sum_{i=1}^{n} |a_i|^2}$. Hence, we have just shown that from equation (3) that

$$\left\| \sum_{i=1}^{n} a_i e_i \right\| \leq \sum_{i=1}^{n} |a_i| \|v_i - e_i\| < \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |a_i| \leq \sqrt{\sum_{i=1}^{n} |a_i|^2} = \left\| \sum_{i=1}^{n} a_i e_i \right\|,$$

a contradiction. Hence, the $\{v_i\}_{i=1}^n$ must be linearly independent, and since $\dim V = n$, we have that $\{v_i\}_{i=1}^n$ is a basis for V. □

Homework 4

Notation: In the Gram-Schmidt Procedure, we define the vector u_k, for $k \geq 2$, inductively by $u_k = v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i$, as we did in Theorem 38.1.

[6B] Orthonormal Bases

6B.7 Example of the Riesz Representation Theorem

Find a polynomial $q \in \mathcal{P}_2(\mathbb{R})$ such that

$$p\left(\frac{1}{2}\right) = \int_0^1 p(x)q(x)\,dx$$

for every $p \in \mathcal{P}_2(\mathbb{R})$.

Solution. We claim (without proof) that $\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx$ is an inner product and that the mapping $\varphi : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ given by $p \mapsto p(1/2)$ is a linear functional.[7] Hence, the Riesz Representation Theorem applies and there exists a unique $q \in \mathbb{R}$ such that

$$p\left(\frac{1}{2}\right) = \langle p, q \rangle = \int_0^1 p(x)q(x)\,dx.$$

We know that $\{1, x, x^2\}$ is a basis for $\mathcal{P}_2(\mathbb{R})$, so we apply the Gram-Schmidt Procedure to this basis to obtain an orthonormal basis. If $v_1 = 1$, $v_2 = x$, and $v_3 = x^2$, we see that

$$\|v_1\|^2 = \int_0^1 1 \cdot 1\,dx = x\Big|_0^1 = 1 - 0 = 1,$$

so $e_1 = v_1 / \|v_1\| = v_1 = 1$. Now

$$u_2 = v_2 - \langle v_2, e_1 \rangle e_1 = x - 1\int_0^1 x \cdot 1\,dx = x - \frac{1}{2} \text{ and}$$

$$\|u_2\|^2 = \int_0^1 x^2 - x + \frac{1}{4}\,dx = \frac{1}{12},$$

so $e_2 = u_2 / \|u_2\| = \sqrt{12}\left(x - \frac{1}{2}\right) = 2\sqrt{3}\left(x - \frac{1}{2}\right)$. Finally, we have

$$\begin{aligned} u_3 &= v_3 - \langle v_3, e_1 \rangle e_1 - \langle v_3, e_2 \rangle e_2 \\ &= x^2 - 1\int_0^1 x^2\,dx - 2\sqrt{3}\left(x - \frac{1}{2}\right)\int_0^1 x^2 \cdot 2\sqrt{3}\left(x - \frac{1}{2}\right)\,dx \\ &= x^2 - \frac{1}{3} - 12\left(x - \frac{1}{2}\right)\int_0^1 x^3 - \frac{1}{2}x^2\,dx = x^2 - \frac{1}{3} - \left(x - \frac{1}{2}\right) = x^2 - x + \frac{1}{6}. \end{aligned}$$

Now

$$\|u_3\|^2 = \int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 dx = \frac{1}{180},$$

[7] For the reader familiar with Math 120B, this should be fairly evident.

so $e_3 = u_3/\|u_3\| = \sqrt{180}\left(x^2 - x + \frac{1}{6}\right) = 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right)$. Hence, an orthonormal basis of $\mathcal{P}_2(\mathbb{R})$ is $\{1, 2\sqrt{3}\left(x - \frac{1}{2}\right), 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right)\}$.

Now, in the proof of the Riesz Representation Theorem, we let $q = \overline{\varphi(e_1)}e_1 + \overline{\varphi(e_2)}e_2 + \overline{\varphi(e_3)}e_3$. Since we are working over the reals, we just have $q = \sum_{i=1}^{3} \varphi(e_i)e_i$, so we have

$$\varphi(e_1) = 1, \varphi(e_2) = 2\sqrt{3}\left(\frac{1}{2} - \frac{1}{2}\right) = 0, \varphi(e_3) = 6\sqrt{5}\left(\frac{1}{4} - \frac{1}{2} + \frac{1}{6}\right) = -\frac{\sqrt{5}}{2}.$$

Hence

$$q(x) = 1 - \frac{\sqrt{5}}{2} \cdot 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right) = 1 - 15\left(x^2 - x + \frac{1}{6}\right) = \boxed{-15x^2 + 15x - \frac{3}{2}},$$

so we are done. •

[6C] Orthogonal Complements

6C.4 Finding an Orthonormal Basis of $U^\perp$

Suppose $U \leq \mathbb{R}^4$ is given by $U = \operatorname{span}((1, 2, 3, -4), (-5, 4, 3, 2))$. Find and orthonormal basis of U and an orthonormal basis for $U^\perp$.

This problem is very tedious to solve (as well as a waste of space), so we describe a solution here. We can apply the Gram-Schmidt Procedure to the two basis vectors given to obtain an orthonormal basis $\{e_1, e_2\}$ of U. Since $V = U \oplus U^\perp$, we extend to an orthonormal basis $\{e_1, e_2, e_3, e_4\}$. Now, $\{e_3, e_4\}$ is an orthonormal basis for $U^\perp$.

6C.5 Orthogonal Projection of Orthogonal Complement

Suppose V is finite-dimensional and $U \leq V$. Show that $P_{U^\perp} = I - P_U$, where I is the identity operator on V.

Proof. Fix an arbitrary $v \in V$, and let $U \leq V$. Denote $W := U^\perp$. Then $V = U \oplus W$, so there exist unique $u \in U$, $w \in W$ such that $v = u + w$. However, $W^\perp = U$, so that $v = w + u \in W \oplus W^\perp$ (where we are justified in using the $\oplus$ symbol as $U \cap W = W^\perp \cap W = \varnothing$). Hence $P_{U^\perp}(v) = P_W(v) = P_W(w + u) = w$, but we also note that

$$(I - P_U)(v) = (I - P_U)(u + w) = I(u + w) - P_U(u + w) = u + w - u = w = P_W(v).$$

Hence $I - P_U = P_W = P_{U^\perp}$. □

6C.8 When are Idempotent Operators Orthogonal Projections?

Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ satisfies $P^2 = P$ and $\|Pv\| \leq \|v\|$ for every $v \in V$. Prove that there exists a subspace $U \leq V$ such that $P = P_U$.

Proof. It would be nice if we had $U = \operatorname{im} P$, so we claim that $U = \operatorname{im} P$, and naturally, that $U^\perp = \ker P$. Fix $u \in \operatorname{im} P$ and $w \in \ker P$. We first show $\langle u, w \rangle = 0$. Since $u \in \operatorname{im} P$, there exists $x \in V$ such that $u = Px$. Hence $Pu = PPx = P^2x = Px = u$, so $Pu = u$. Hence for some $w \in \ker P$, $P(u + w) = Pu + Pw = u + Pw = u$. Similarly, for any $a \in F$, we have

$P(u+aw) = Pu + P(aw) = u + aPw = u$, so that $\|u\|^2 = \|P(u+aw)\|^2 \leq \|u+aw\|^2$ by assumption. This implies $\|u\| \leq \|u+aw\|$, so by Example 33.3, we have $\langle u, w\rangle = 0$. Hence $u \perp w$, so $\ker P \perp \operatorname{im} P$, so that $\ker P \subseteq (\operatorname{im} P)^\perp$.

It remains to show that $V = \ker P \oplus \operatorname{im} P$, so that $(\operatorname{im} P)^\perp = \operatorname{im} P$. Fix $v \in V$, so that $Pv \in \operatorname{im} P$. Now $P^2 v = Pv$, so that $P^2 v - Pv = 0 \iff P(v - Pv) = 0$, so that $v - Pv \in \ker P$. Hence $v = Pv + (v - Pv) \in \operatorname{im} P + \ker P$, hence $V = \operatorname{im} P + \ker P$. Now, suppose $v \in \operatorname{im} P \cap \ker P$. Then there exists $u \in V$ such that $v = Pu$, and we know $Pv = 0$. But $0 = Pv = P^2 u = Pu = v$, so $v = 0$; i.e., $\operatorname{im} P \cap \ker P = 0$. Hence $V = \ker P \oplus \operatorname{im} P$.

It follows that for any $v \in V$, we can uniquely write $v = u + w$, where $u \in \operatorname{im} P$ and $w \in \ker P$, so that we have $P(v) = P(u+w) = Pu + Pw = Pu + 0 = Pu = u$. Since $(\operatorname{im} P)^\perp = \ker P$, we have that $P = P_{\operatorname{im} P}$, and we are done. ☐

[7A] Self-Adjoint and Normal Operators

7A.1 Calculating an Adjoint

Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbb{F}^n)$ by $T(z_1, \ldots, z_n) = (0, z_1, \ldots, z_{n-1})$. Find a formula for $T^*(z_1, \ldots, z_n)$.

Solution. Let $z = (z_1, \ldots, z_n)$ and $w = (w_1, \ldots, w_n)$ in $\mathbb{F}^n$. Then the adjoint must satisfy $\langle Tz, w\rangle = \langle z, T^* w\rangle$. We have

$$\begin{aligned}\langle Tz, w\rangle = \langle (0, z_1, \ldots, z_{n-1}), (w_1, \ldots, w_n)\rangle &= \sum_{i=2}^{n} z_{i-1}\overline{w_i} \\ &= z_1\overline{w_2} + z_2\overline{w_3} + \cdots + z_{n-1}\overline{w_n} = \langle z, T^* w\rangle .\end{aligned}$$

From here, it is easy to guess that $T^*(w) = T^*(w_1, \ldots, w_n) = (w_2, w_3, \ldots, w_n, 0)$, and we verify:

$$\langle z, T^* w\rangle = \langle (z_1, \ldots, z_n), (w_2, \ldots, w_n, 0)\rangle = \sum_{j=1}^{n-1} z_j\overline{w_{j+1}} = \sum_{i=2}^{n} z_{i-1}\overline{w_i} = \langle Tz, w\rangle .$$

Hence $T^*(z_1, \ldots, z_n) = \boxed{(z_2, \ldots, z_n, 0)}$. •

7A.3 Invariance of Orthogonal Complement

Suppose $T \in \mathcal{L}(V)$ and $U \leq V$. Prove that U is invariant under T if and only if $U^\perp$ is invariant under T^*.

Proof. Fix $T \in \mathcal{L}(V)$, and $U \leq V$, so that $V = U \oplus U^\perp$.

($\Longrightarrow$): Suppose U is invariant under T. Fix $u \in U$, so that $Tu \in U$. Hence for any $w \in U^\perp$, we have $Tu \perp w$ (as $Tu \in U$), so that $\langle Tu, w\rangle = 0$. But $\langle Tu, w\rangle = \langle u, T^* w\rangle = 0$, so that $u \perp T^* w$. By definition of $U^\perp$, we have $T^* w \in U^\perp$, so that $U^\perp$ is invariant under T^*.

($\Longleftarrow$): Suppose $U^\perp$ is invariant under T. Fix $w \in U^\perp$, so that $T^* w \in U^\perp$. Hence for any $u \in U = (U^\perp)^\perp$, we have $u \perp Tw$, so that $\langle u, T^* w\rangle = 0$. But $\langle Tu, w\rangle = \langle u, T^* w\rangle = 0$, so that $Tu \perp w$. But by definition of $U^\perp$, we see $Tu \in U$, so that U is invariant under T. ☐

Homework 5

[7A] Self-Adjoint and Normal Operators

7A.11 Projection Iff Self-Adjoint

Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that there is a subspace $U \leq V$ such that $P = P_U$ if and only if P is self-adjoint.

Proof. ($\implies$) : Suppose $P = P_U$ for some $U \leq V$. Then we write $V = U \oplus U'$. Pick $v, v' \in V$. Then $v = u + w$ and $v' = u' + w'$ for $u, u' \in U$ and $w, w' \in U^{\perp}$. Now

$$\langle Pv, v' \rangle = \langle u, v' \rangle = \langle u, u' + w' \rangle = \langle u, u' \rangle + \langle u, w' \rangle = \langle u, u' \rangle .$$

However, since $w \perp u'$, we write

$$\langle Pv, v' \rangle = \langle u, u' \rangle + 0 = \langle u, u' \rangle + \langle w, u' \rangle = \langle u + w, u' \rangle = \langle v, u' \rangle = \langle v, Pv' \rangle ,$$

so that P is self-adjoint.

($\impliedby$): Suppose P is self-adjoint and satisfies $P^2 = P$. We claim $U := \operatorname{im} P$. Since P is self-adjoint, we see that $\operatorname{im} P = (\ker P)^{\perp} = U^{\perp}$, so $P = \operatorname{im} P \oplus \ker P$. Now, fix $v \in V$, and write $v = Pv + (v - Pv)$. Now $P(v - Pv) = Pv - P^2 v = Pv - Pv = 0$, so $v - Pv \in \ker P$. Hence $Pv = u$, where $u = Pv \in \operatorname{im} P = U$, so that $P = P_U$. □

[7B] The Spectral Theorem

7B.2 Polynomial of Operator is Zero

Suppose that T is a self-adjoint operator on a finite-dimensional inner product space and that 2 and 3 are the only eigenvalues of T. Prove that $T^2 - 5T + 6I = 0$.

Proof. Since T is self-adjoint, by the Spectral Theorem there exists an orthonormal eigenbasis $\{e_i\}_{i=1}^{n}$ of V, where $n := \dim V$. Since the only eigenvalues are 2 and 3, we either have $(T - 2I)e_i = 0$ or $(T - 3I)e_i = 0$ for each e_i. Hence $(T - 2I)(T - 3I)e_i = (T^2 - 5T + 6I)e_i = 0$ for all e_i, so that $T^2 - 5T + 6I = 0$. □

7B.5 Normality and Eigenvectors

Suppose V is a inner product space over $\mathbb{R}$, and let $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where the λ_i are distinct eigenvalues of T.

Proof. ($\implies$): Suppose T is self-adjoint. Then by the Real Spectral Theorem, V has an orthonormal eigenbasis, so we immediately have $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$. That all pairs of eigenvectors corresponding to distinct eigenvalues are orthogonal comes from the fact that we chose an orthonormal eigenbasis by the Complex Spectral Theorem.

($\impliedby$): Suppose all pairs of eigenvectors of T corresponding to distinct eigenvectors of T are orthogonal, and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$. Now, take orthonormal bases β_i of each $E(\lambda_i, T)$. Then a basis for V is the union $\beta := \beta_1 \cup \cdots \cup \beta_m$. If $v \in \beta$, then v is normalized, and for if $v \in \beta_i$ and $w \in \beta_j$, $j \neq i$, we have $v \perp w$ by assumption. Hence β is an orthonormal basis of V consisting of eigenvectors of T, so by the Real Spectral Theorem, T is self-adjoint. □

7B.7 $T^9 = T^8$ Implies $T^2 = T$ If T is Self-Adjoint

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is a normal operator such that $T^9 = T^8$. Prove that T is self-adjoint and $T^2 = T$.

Proof. This is Example 67.3 in the main text. For convenience, we print a more explicit proof here.

Suppose $T^9 = T^8$. Since V is a complex vector space, T has an eigenvector v with some eigenvalue λ. Hence $T^8 v = \lambda^8 v = \lambda^9 v = T^9 v$, so that $\lambda^8(\lambda - 1)v = 0$. Since $v \neq 0$, we have $\lambda = 0, 1$, so 0 and 1 are the only possible eigenvalues of T. By the Complex Spectral Theorem, fix an orthonormal basis of V of eigenvectors of T such that T has the matrix $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Now, the λ_i are either 0 or 1, so that $\overline{A^T} = A$ (hence T is normal) and $A^2 = \operatorname{diag}(\lambda_1^2, \ldots, \lambda_n^2) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = A$, so that $T^2 = T$. □

7B.11 Cube Root of an Operator

Prove or give a counterexample: every self-adjoint operator on V has a cube root, i.e., if $T \in \mathcal{L}(V)$ is self-adjoint, then there exists $S \in \mathcal{L}(V)$ such that $S^3 = T$.

Solution. **This is true.** [Compare this with Example 67.4 in the main text.] •

Proof. Suppose T is self-adjoint. Then by the Spectral Theorem, there exists an orthonormal eigenbasis $\{e_i\}_{i=1}^n$ of V. Let λ_i be the eigenvalue associated with each e_i, so that T has matrix $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ in this basis. Now, take S to be the linear transformation given by the matrix $B := \operatorname{diag}\left(\sqrt[3]{\lambda_1}, \ldots, \sqrt[3]{\lambda_n}\right)$ in the basis $\{e_i\}$. Clearly, $B^3 = A$, so that $S^3 = T$, so we are done. □

Homework 6

[7C] Positive Operators and Isometries

7C.7 Positive Operator is Positive Iff it is Actually Positive

Suppose $T \in \mathcal{L}(V)$ is positive. Prove that T is invertible if and only if $\langle Tv, v\rangle > 0$ for every $v \in V$ with $v \neq 0$.

Proof. ($\Longrightarrow$) : Suppose T is invertible, so $\ker T = \{0\}$. Since T is positive, there exists S, self-adjoint, such that $S^2 = T$. Since T is invertible, so is S, so for any $v \neq 0$, we see that $\langle Tv, v\rangle = \left\langle S^2 v, v\right\rangle = \langle Sv, Sv\rangle = \|Sv\|^2 > 0$.

($\Longleftarrow$): Suppose $\langle Tv, v\rangle > 0$ for every $v \in V$ with $v \neq 0$. Certainly, this implies $Tv \neq 0$. Hence, T has trivial kernel, and is thus invertible. □

7C.11 Similar Normal Operators

Suppose T_1, T_2 are normal operators on $\mathcal{L}(\mathbb{F}^3)$ and both operators have $2, 5, 7$ as eigenvalues. Prove that there exists an isometry $S \in \mathcal{L}(\mathbb{F}^3)$ such that $T_1 = S^* T_2 S$.

Proof. By the spectral theorem, there exists an orthonormal eigenbasis $\{v_2, v_5, v_7\}$ for T_1, and an orthonormal eigenbasis $\{w_2, w_5, w_7\}$ for T_2. Define the operator S by $Sv_i = w_i$, for $i = 2, 5, 7$. Since $\{v_2, v_5, v_7\}$ is an orthonormal basis and $\{Sv_2, Sv_5, Sv_7\} = \{w_2, w_5, w_7\}$ is orthonormal,

S is an isometry, so $S^* = S^{-1}$. Now $S^*T_2S(v) = S^*T_2S(a_2v_2 + a_5v_5 + a_7v_7) = S^*T_2(a_2w_2 + a_5w_5 + a_7w_7) = S^*(2a_2w_2 + 5a_5w_5 + 7a_7w_7) = 2a_2v_2 + 5a_5v_5 + 7a_7v_7 = T_1(v)$. □

7C.12 Similar Normal Operators?

Give an example of two self-adjoint operators $T_1, T_2 \in \mathcal{L}(\mathbb{F}^4)$ such that the eigenvalues of both operators are $2, 5, 7$ but there does not exist an isometry $S \in \mathcal{L}(\mathbb{F}^4)$ such that $T_1 = S^*T_2S$.

Solution. Let $\{e_1, \dots, e_4\}$ be the standard basis in $\mathbb{F}^4$ (which is orthonormal), and let T_1 and T_2 have the matrices $\operatorname{diag}(2,2,5,7)$ and $\operatorname{diag}(2,5,7,7)$ respectively. Clearly, T_1 and T_2 are both self-adjoint. But there is no invertible operator S such that $T_1 = S^{-1}T_2S$, let alone an isometry. If S were an invertible operator satisfying $T_1 = S^{-1}T_2S$, then consider the vector $v \in V$ such that $Sv = e_3$. Then

$$T_1v = S^{-1}T_2Sv = S^{-1}T_2(e_3) = S^{-1}(7e_3) = 7v.$$

Hence, v is an eigenvector of T_1 with eigenvalue 7, i.e., $v \in E(T_1, 7)$. Now, consider the vector $v' \in V$ such that $Sv' = e_4$. Then

$$T_1v' = S^{-1}T_2Sv' = S^{-1}T_2(e_4) = S^{-1}(7e_4) = 7v'.$$

Hence, v' is also an eigenvector of T_2 with eigenvalue 7, i.e, $v, v' \in E(T_1, 7)$. Now, $\dim E(T_1, 7) = 1$, so $v' = \alpha v$ for some $\alpha \in \mathbb{F}$. But this is problematic, as $S(v' - v) = S((\alpha - 1)v) = (\alpha - 1)e_3$ and $S(v' - v) = Sv' - Sv = e_4 - e_3 = (\alpha - 1)e_3 \implies e_4 = \alpha e_3$, a contradiction, as e_3 and e_4 are linearly independent. Hence, no such invertible S can exist, let alone an isometry (which is invertible by definition). •

[7D] Polar and Singular-Value Decompositions

7D.3 Reversed Polar Decomposition

Suppose $T \in \mathcal{L}(V)$. Prove that there exists an isometry $S \in \mathcal{L}(V)$ such that $T = \sqrt{TT^*}S$.

Proof. Fix $T \in \mathcal{L}(V)$, and consider $T^* \in \mathcal{L}(V)$. By the Polar Decomposition, write $T^* = R\sqrt{T^{**}T^*} = R\sqrt{TT^*}$ for some isometry R. Now, write

$$T = (T^*)^* = (R\sqrt{TT^*})^* = (\sqrt{TT^*})^*R^* = (\sqrt{TT^*})^*R^{-1}.$$

But $\sqrt{TT^*}$ is self-adjoint by definition, so we have $T = \sqrt{TT^*}R^{-1}$. Set $S := R^{-1}$, which is also an isometry. □

7D.7 Calculating a Polar Decomposition

Define $T \in \mathcal{L}(\mathbb{F}^3)$ by $T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2)$. Find an isometry $S \in \mathcal{L}(\mathbb{F}^3)$ such that $T = S\sqrt{T^*T}$.

Solution. The matrix of T with respect to the standard basis is $M_T = \begin{pmatrix} 0 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$, so the matrix of the adjoint is $M_{T^*} = \overline{M_T} = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix}$. Hence, we compute that the matrix of

T^*T is $M = \operatorname{diag}(4, 9, 1)$. From here, we read off the matrix of $\sqrt{T^*T}$ as $\operatorname{diag}(2, 3, 1)$, so we want a matrix S, isometric, such that

$$\begin{pmatrix} 0 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} = S \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, we let S be the matrix that corresponds to the permutation cycle $(1\ 2\ 3)$ in the symmetric group S_3, acting row-wise, i.e.,

$$\begin{pmatrix} 0 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, we define S by the transformation that sends $(z_1, z_2, z_3) \mapsto (z_3, z_1, z_2)$, and it is easy to check that S^* corresponds to S^{-1}, so S is the isometry we are looking for. •

7D.17 Facts About the Singular Value Decomposition

Suppose $T \in \mathcal{L}(V)$ has a singular-value decomposition given by

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

for every $v \in V$, where the s_i are the singular values of T and the e_i and f_i form orthonormal bases of V.

(a) *Prove that if $v \in V$, then $T^*v = s_1 \langle v, f_1 \rangle e_1 + \cdots + s_n \langle v, f_n \rangle e_n$.*

Proof. Observe that by construction of the decomposition $Te_i = s_i f_i$, which means that T can be viewed as a diagonal matrix $\operatorname{diag}(s_1, \ldots, s_n)$, where we write the matrix with respect to both bases $\{e_i\}$ and $\{f_i\}$. Because singular values are real, we see that the matrix of T^* is the same as that for T, but that the roles of $\{e_i\}$ and $\{f_i\}$ have been exchanged, so $T^* f_i = s_i e_i$. Extending by linearity, we see that $T^*v = s_1 \langle v, f_1 \rangle e_1 + \cdots + s_n \langle v, f_n \rangle e_n$, which completes the proof. □

(b) *Prove that if $v \in V$, then $T^*Tv = s_1^2 \langle v, e_1 \rangle e_1 + \cdots + s_n^2 \langle v, e_n \rangle e_n$.*

Proof. We compute:

$$\begin{aligned} T^*(Tv) &= T^*\big(s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n\big) = \sum_{i=1}^n T^*\big(s_i \langle v, e_i \rangle f_i\big) \\ &= \sum_{i=1}^n s_i \langle v, e_i \rangle T^*(f_i) = \sum_{i=1}^n s_i \langle v, e_i \rangle (s_i e_i) = \sum_{i=1}^n s_i^2 \langle v, e_i \rangle e_i, \end{aligned}$$

as expected. □

(c) *Prove that if $v \in V$, then $\sqrt{T^*T}v = s_1 \langle v, e_1 \rangle e_1 + \cdots + s_n \langle v, e_n \rangle e_n$.*

Proof. The s_i are by definition the eigenvalues of $\sqrt{T^*T}$, with corresponding eigenvectors e_i. Hence $\sqrt{T^*T}e_i = s_i e_i$, so extending by linearity, we get the formula we want. □

(d) *Suppose T is invertible. Prove that if $v \in V$, then*

$$T^{-1}v = \frac{\langle v, f_1 \rangle e_1}{s_1} + \cdots + \frac{\langle v, f_n \rangle e_n}{s_n}$$

for every $v \in V$.

Proof. First, we make sure that the division by the s_i is legal. If T is invertible, then T^* must also be invertible as $\ker T = \{0\} \iff (\operatorname{im} T^*)^\perp = \{0\} \iff \operatorname{im} T^* = V$. Hence, neither T nor T^* have 0 as an eigenvalue, so neither can T^*T nor $\sqrt{T^*T}$. Hence, 0 is not a singular value, so the division is legal here. Now, we compute — it suffices to check that the given formula shows that $TT^{-1}e_i = T^{-1}Te_i = I$ for the basis vectors e_i and f_i. Denote $Sv := \frac{\langle v, f_1 \rangle e_1}{s_1} + \cdots + \frac{\langle v, f_n \rangle e_n}{s_n}$. Then by orthonormality of $\{e_i\}$ and $\{f_i\}$:

$$TSf_i = T\left(\frac{\langle f_i, f_i \rangle e_i}{s_i}\right) = \frac{1}{s_i}T(e_i) = \frac{1}{s_i}(s_i \langle e_i, e_i \rangle f_i) = \frac{s_i}{s_i} \cdot 1 f_i = f_i, \text{ and}$$

$$STe_i = S(s_i \langle e_i, e_i \rangle f_i) = s_i S(1f_i) = s_i \frac{\langle f_i, f_i \rangle e_i}{s_i} = 1e_i = e_i,$$

so $S = T^{-1}$. □

Homework 7

[8A] Generalized Eigenvectors and Nilpotent Operators

8A.4 Distinct Generalized Eigenspaces Have Trivial Intersection

Suppose $T \in \mathcal{L}(V)$ and $\alpha, \beta \in \mathbb{F}$ with $\alpha \neq \beta$. Prove that $G(\alpha, T) \cap G(\beta, T) = \{0\}$.

Proof. Let $v \in G(\alpha, T) \cap G(\beta, T)$. If $v \neq 0$, then v is both a generalized eigenvector corresponding to α, and a generalized eigenvector corresponding to β. By Theorem 80.6, $\{v, v\}$ must be linearly independent, which is a contradiction, so $v = 0$. □

8A.6 This Operator Has No Square Root

Suppose $T \in \mathcal{L}(\mathbb{C}^3)$ is defined by $(z_1, z_2, z_3) \mapsto (z_2, z_3, 0)$. Prove that T has no square root, i.e., there is no $S \in \mathcal{L}(\mathbb{C}^3)$ such that $S^2 = T$.

Proof. This is Example 93.3. □

8A.15 Counting

Suppose $\dim V =: n$ and $N \in \mathcal{L}(V)$ is such that $\ker N^{n-1} \neq \ker N^n$. Prove that N is nilpotent and $\dim \ker N^j = j$ for each integer j with $0 \leq j \leq n$.

Proof. This is Example 107.2. Here is a proof in more detail.

Since $\ker N^{n-1} \neq \ker N^n$, we see that $\ker N^j \neq \ker N^{j+1}$ for every $0 \leq j \leq n-1$, as if $\ker N^j = \ker N^{j+1}$, we would have $\ker N^j = \ker N^{j+1} = \cdots = \ker N^{n-1} = \ker N^n$, a contradiction. Thus, we must have

$$0 = \dim \ker N^0 < \dim \ker N^1 < \dim \ker N^2 < \cdots < \dim \ker N^{n-1} < \dim \ker N^n,$$

but we know that $\dim \ker N^n \leq n$. The inequality above is only possible if each $\ker N^j$, $0 \leq j \leq n$, has dimension j, i.e., that the dimension of the kernel increases by 1 for each application of N. This proves the second statement. Now, we know that $\dim \ker N^n = n$, but $n = \dim V$, so $\ker N^n = V$, which implies $N^n = 0$, so N is nilpotent. $\square$

[8B] Decomposition of an Operator

8B.3 Similar Transformations Have the Same Eigenvalues... the Same Number of Times

Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible. Prove that T and $S^{-1}TS$ have the same eigenvalues with the same multiplicities.

Proof. It suffices to show that T and $S^{-1}TS$ have the same characteristic polynomial. Let $p(\lambda) := \det(T - \lambda I)$ be the characteristic polynomial of T. However, observe that

$$S^{-1}TS - \lambda I = S^{-1}TS - S^{-1}(\lambda I)S = S^{-1}(T - \lambda I)S, \text{ so}$$

$$\det(S^{-1}TS - \lambda I) = \det(S^{-1}(T - \lambda I)S) = \det S^{-1} \cdot \det(T - \lambda I) \cdot \det S = \det(T - \lambda I),$$

so T and $S^{-1}TS$ have the same characteristic polynomial. Hence T and $S^{-1}TS$ share the same eigenvalues, and the eigenvalues have the same multiplicity. $\square$

8B.6 Computing a Square Root of $I + N$

Define $N \in \mathcal{L}(\mathbb{F}^5)$ by $(x_1, \ldots, x_5) \mapsto (2x_2, 3x_3, -x_4, 4x_5, 0)$. Find a square root of $I + N$.

Solution. We see that $N^5 = 0$. Now, apply the Taylor series of $\sqrt{1+x}$ as seen in lecture and write

$$\sqrt{I+N} = 1 + a_2 N + a_2 N^2 + a_3 N^3 + a_4 N^4,$$

where we are justified in leaving out the rest of the terms as $N^5 = 0$. From calculus, we know that if $f(x) = \sqrt{1+x}$, then the Taylor Series of x centered at $x = 0$ is

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots,$$

so we compute the relevant derivatives and get

$$f(0) = 1, f'(0) = \frac{1}{2}, f''(0) = -\frac{1}{4}, f'''(0) = \frac{3}{8}, \text{ and } f''''(0) = -\frac{15}{16}.$$

Hence

$$\begin{aligned}\sqrt{1+x} &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{3}{8 \cdot 3!}x^3 - \frac{15}{16 \cdot 4!}x^4 + \cdots \\ &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4,\end{aligned}$$

and replacing x's with N's, we see that

$$\sqrt{1+N} = \boxed{1 + \frac{1}{2}N - \frac{1}{8}N^2 + \frac{1}{16}N^3 - \frac{5}{128}N^4},$$

which is good enough. We can find an explicit formula via computation of N^i, $i \leq 4$. •

8B.10 The Jordan-Chevalley Decomposition

Suppose V is a vector space over $\mathbb{C}$ and $T \in \mathcal{L}(V)$. Prove that there exist $D, N \in \mathcal{L}(V)$ such that $T = D + N$, the operator D is diagonalizable, N is nilpotent, and $DN = ND$.

Proof. Let $v_1, \ldots, v_n$ be a generalized eigenbasis of V, consisting of generalized eigenvectors of T. We know that $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T) =: U_1 \oplus \cdots \oplus U_m$, where the λ_i are the distinct eigenvalues of T. As such, the projection operators P_i are well-defined, where $P_i(v)$ does the obvious thing of sending v into its projection in U_i. Thus, we write T as a sum of restrictions and projections as follows:

$$T = T|_{U_1} P_1 + \cdots + T|_{U_m} P_m.$$

Now, we know that $T|_{U_i} - \lambda_i I =: N_i \in \mathcal{L}(U_i)$ is nilpotent, so we substitute this into the above to get

$$T = (\lambda_1 I + N_1)P_1 + \cdots + (\lambda_m I + N_m)P_m = \sum_{i=1}^{m} \lambda_i I P_i + \sum_{i=1}^{m} N_i P_i = \sum_{i=1}^{m} \lambda_i P_i + \sum_{i=1}^{m} N_i P_i.$$

We claim that $D = \sum \lambda_i P_i$ and $N = \sum N_i P_i$ as above. First, we see that for each generalized eigenvector v_j in the eigenbasis ($j \leq n$), we see that $Dv_j = \lambda_i v_j$, where λ_i is the eigenvalue corresponding to v_j. Hence, D has the matrix $\operatorname{diag}(\lambda_1, \ldots, \lambda_m)$, where some of the λ_i are repeated appropriately, so D is diagonalizable as required. Similar to the computation for D, we observe that $Nv_j = N_i v_j$, where i is chosen such that λ_i is the corresponding eigenvalue for v_j. But each N_i was nilpotent, so the sum N is also nilpotent. But because $N_i \in \mathcal{L}(U_i)$ and is nilpotent, we see that certainly $N^n v_j = N_i^n v_j = 0$, so N is nilpotent. To see that D and N commute, we compute $DNv_j = D(N_i v_j) = \lambda_i N_i v_j \text{ and } NDv_j = N(\lambda_i v_j) = \lambda_i N_i v_j$, so we are done. □

Homework 8

8C Characteristic and Minimal Polynomials

8C.1 Product of Eigentransformations is Zero: Cayley-Hamilton Edition

Suppose $T \in \mathcal{L}(\mathbb{C}^4)$ is such that the eigenvalues of T are $3, 5, 8$. Prove that $(T - 3I)^2(T - 5I)^2(T - 8I)^2 = 0$.

Proof. We know that $T \in \mathcal{L}(\mathbb{C}^4)$, so one of the eigenvalues of T must have multiplicity 2. Hence, the characteristic polynomial of T has the form $q(x) = (x - \lambda_1)^2(x - \lambda_2)(x - \lambda_3) \in \mathbb{C}[x]$, where $\lambda_1, \lambda_2, \lambda_3$ are $3, 5, 8$ in some order. Clearly q divides $(x-3)^2(x-5)^2(x-8)^2$, so by Cayley-Hamilton, $q(T) = 0 \implies (T - 3I)^2(T - 5I)^2(T - 8I)^2 = 0$. □

8C.4 An Introduction to Jordan Form

Give an example of an operator on $\mathbb{C}^4$ whose characteristic polynomial equals $(z-1)(z-5)^3$ and whose minimal polynomial equals $(z-1)(z-5)^2$.

Solution. Let $T \in \mathcal{L}(\mathbb{C}^4)$ be the matrix (on the next page) given by

$$M_T := \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Clearly, the characteristic polynomial of T is $(z-1)(z-5)^3$. Also, we can check that

$$(M_T - I)(M_T - 5I)^2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} -4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}^2 = \mathbf{0}_{4\times 4},$$

so the minimal polynomial is $(z-1)(z-5)^2$. •

We should notice these two Jordan blocks of M_T above if we swap the second and third columns and adjust the basis accordingly:

$$\tilde{M_T} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

8C.18 The Companion Matrix

Suppose $a_0, \ldots, a_{n-1} \in \mathbb{C}$. Find the minimal and characteristic polynomials of the operator on $\mathbb{C}^n$ whose matrix with respect to the standard basis is

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

Solution. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{C}^n$. Reading off the matrix, we see $Te_i = e_{i+1}$ for $i \leq n-1$, and $Te_n = -a_0 e_1 - \cdots - a_{n-1}e_n$. This implies that $T^j e_1 = e_{j+1}$ for $j \leq n-1$, and $T^n e_1 = -a_0 e_1 - \cdots - a_{n-1}e_n$. Hence, the set $\{T^j e_1\}_{j=0}^{n-1}$ is just the standard basis rewritten, so it is linearly independent, so if

$$b_0 T^0 e_1 + b_1 T^1 e_1 + \cdots + b_{n-1} T^{n-1} e_1 = 0,$$

we must have $b_0 = \cdots = b_{n-1} = 0$. Rewriting this, if $q(z) = b_0 + b_1 z + \cdots + b_{n-1} z^{n-1}$, then $q(z) = 0$. We have just shown that the minimal polynomial of T must have degree at least n. Now, we know $T^n e_1 = -a_0 e_1 - \cdots - a_{n-1} e_n$. But we know that $e_i = T^{i-1} e_1$, so we have

$$T^n e_1 = -a_0 T^0 e_1 - \cdots - a_{n-1} T^{n-1} e_1 \iff (a_0 I + a_1 T + \cdots + a_{n-1} T^{n-1} + T^n)(e_1) = 0,$$

so setting $\boxed{p(z) = a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n}$, we have $p(T)(e_1) = 0$. Now $p(T)(e_i) = p(T)(T^{i-1}e_1) = T^{i-1}(p(T)e_1) = T^{i-1}(0) = 0$, and p is monic, so p is both the minimal and characteristic polynomial of T. •

[8D] Jordan Form

8D.5 Matrix of T^2 Given Jordan Form

Suppose $T \in \mathcal{L}(V)$ and $\{v_i\}_{i=1}^n$ is a basis for V that is a Jordan basis for T. Describe the matrix of T^2 with respect to this basis.

Solution. If the matrix of T is $M_T = \operatorname{diag}(A_1, A_2, \dots, A_m)$, then the matrix of T^2 can be found by simply squaring each block separately: $\boxed{M_{T^2} = \operatorname{diag}(A_1^2, A_2^2, \dots, A_m^2)}$. By checking with computation, each block looks like

$$A_j^2 = \begin{pmatrix} \lambda_j^2 & 2\lambda_j & 1 & \cdots & \cdots & \cdots & 0 \\ 0 & \lambda_j^2 & 2\lambda_j & 1 & \cdots & \cdots & 0 \\ 0 & 0 & \lambda_j^2 & 2\lambda_j & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \lambda_j^2 & 2\lambda_j & 1 \\ 0 & \cdots & \cdots & 0 & 0 & \lambda_j^2 & 2\lambda_j \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & \lambda_j^2 \end{pmatrix},$$

with λ_j^2 on the diagonal, $2\lambda_j$ on the super-diagonal, and 1 on the diagonal above that. •

8D.6 Basis for Kernel of Nilpotent Transformation

Suppose $N \in \mathcal{L}(V)$ is nilpotent and $\{v_1, \dots, v_n\}$ and $m_1, \dots, m_n$ are as in Proposition 8.55 in the text. Prove that $\{N^{m_j} v_j\}_{j=1}^n$ is a basis of $\ker N$.

Note: Proposition 8.55 in the text is Theorem 94.5 in these notes.

Proof. By Proposition 8.55 in the text, we know that

$$\{N^{m_1} v_1, \dots, v_1, \dots, N^{m_n} v_n, v_n\} \tag{4}$$

is a basis of V, and we have $N^{m_1+1} v_1 = \cdots = N^{m_n+1} v_n = 0$. By definition, the vectors $N^{m_j} v_j$, $1 \le j \le n$ lie in $\ker N$, and they are linearly independent by the list above, so we have shown that $\dim \ker N \ge n$. It suffices to just show that $\dim \ker N \le n$, so $\{N^{m_j} v_j\}_{j=1}^n$ is automatically a basis for $\ker N$. We know that $N(N^k v_j) = N^{k+1} v_j$. But we also know that all of the vectors in (4) are linearly independent, and hence nonzero, and all of these vectors lie in the range of N:

$$N^{m_1} v_1, \dots, N v_1, \dots, N^{m_n} v_n, N v_n.$$

This list is linearly independent, sits inside the range, and has length $m_1 + \cdots + m_n$, so $\dim \operatorname{im} N \ge m_1 + \cdots + m_n$. But (4) is a basis for V, and that list has length $m_1 + m_2 + \cdots + m_n + n$, so by Rank-Nullity, we see that

$$\dim V = \dim \operatorname{im} N + \dim \ker N \iff \dim V - \dim \operatorname{im} N + \dim \ker N$$

$$\implies (m_1 + m_2 + \cdots + m_n + n) - (m_1 + m_2 + \cdots + m_n) \ge \dim \ker N,$$

so $\dim \ker N \le n$, so that $\dim \ker N = n$ and the linearly independent list $\{N^{m_j} v_j\}_{j=1}^n \subseteq \ker N$ is a basis for $\ker N$. □

8D.8 When does V Not Split?

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Prove that there does not exist a direct sum decomposition of V into two proper subspaces invariant under T if and only if the minimal polynomial of T is of the form $(z-\lambda)^{\dim V}$ for some $\lambda \in \mathbb{C}$.

Proof. ($\Longrightarrow$): Suppose there does not exist a direct sum decomposition of V into two proper subspaces invariant under T. Then after fixing a Jordan basis for T, we see that the matrix of T consists of a single block, so it must contain one value λ along its entire diagonal. Hence λ is the only eigenvalue of T, which means that the matrix of T looks like

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \lambda & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & \lambda \end{pmatrix}.$$

By multiplying, we can check that the minimal polynomial of T is $(z-\lambda)^{\dim V}$.

($\Longleftarrow$): By contradiction, suppose that the minimal polynomial of T has the form $(z-\lambda)^{\dim V}$ but there exists a direct sum decomposition of V into two proper subspaces invariant under T, say $V = U \oplus W$ for $\dim U, \dim W \geq 1$, and both U, W are invariant under T. (Of course, this is equivalent to saying that the Jordan form of T splits into at least two blocks). Now, the characteristic polynomial of $T|_U$ is just $(z-\lambda)^{\dim U}$, as it must share the same roots as the minimal polynomial of T. Similarly, the characteristic polynomial of $T|_W$ is $(z-\lambda)^{\dim W}$. Without loss of generality, if $\dim U \geq \dim W$, we see that $p(z) := (z-\lambda)^{\dim U}$ satisfies $p(T) = 0$, but $\dim V > \dim U$ by assumption, so $(z-\lambda)^{\dim V}$ cannot possibly be minimal, a contradiction. ☐

Homework 9

[9A] Complexification

9A.4 Spanning Sets Are Preserved Under Complexification

Suppose V is a real vector space and $v_1, \ldots, v_m \in V$. Prove that $\{v_1, \ldots, v_m\}$ span $V_{\mathbb{C}}$ if and only if $\{v_1, \ldots, v_m\}$ span V.

Proof. ($\Longrightarrow$): Suppose $\{v_i\}$ spans $V_{\mathbb{C}}$. Pick $v \in V \leq V_{\mathbb{C}}$. Then because we assumed $\{v_i\}$ spans $V_{\mathbb{C}}$,

$$v = v + 0i = a_1 v_1 + \cdots + a_m v_m = (\operatorname{Re} a_1 v_1 + \cdots + \operatorname{Re} a_m v_m) + (\operatorname{Im} a_1 v_1 + \cdots + \operatorname{Im} a_m v_m).$$

But by comparing components, we have written v as an $\mathbb{R}$-linear combination of the v_i, so $v \in \operatorname{span}_{\mathbb{R}}(v_1, \ldots, v_m)$. Hence the v_i span V.

($\Longleftarrow$): Suppose $\{v_i\}$ spans V. Pick some $v + wi \in V_C$. Since $v, w \in V$, write

$$v + wi = (a_1 v_1 + \cdots + a_m v_m) + (b_1 v_1 + \cdots + b_m v_m)i$$

for appropriate $a_i, b_i \in \mathbb{R}$. Bringing this together gives

$$v + wi = (a_1 + b_1 i)v_1 + \cdots + (a_m + b_m i)v_m,$$

which writes $v + wi$ as a $\mathbb{C}$-linear combination of the v_i, so $v \in \operatorname{span}_{\mathbb{C}}(v_1, \ldots, v_m)$, so the v_i span $V_{\mathbb{C}}$. □

9A.8 The Unreal Eigenvalues of $T_{\mathbb{C}}$

Suppose $T \in \mathcal{L}(\mathbb{R}^3)$ and $5, 7$ are eigenvalues of T. Prove that $T_{\mathbb{C}}$ has no non-real eigenvalues.

Proof. We know that T and $T_{\mathbb{C}}$ have the same real eigenvalues. Additionally, we know that for $T_{\mathbb{C}}$, non-real eigenvalues must come in pairs, so if $\lambda \notin \mathbb{R}$ is an eigenvalue of $T_{\mathbb{C}}$, then $\bar{\lambda} \notin \mathbb{R}$ is also an eigenvalue of $T_{\mathbb{C}}$. But this would imply that $T_{\mathbb{C}}$ and hence T has 4 distinct eigenvalues: $5, 7, \lambda, \bar{\lambda}$, a contradiction. □

[10A] Trace

10A.16 Is the Trace Multiplicative?

Prove or give a counterexample: if $S, T \in \mathcal{L}(V)$, then $\operatorname{tr}(ST) = (\operatorname{tr} S)(\operatorname{tr} T)$.

Solution. This is false: take $S = T = I$. Then $\operatorname{tr}(I^2) = \operatorname{tr}(I) = 1 + 1 = 2$, but $(\operatorname{tr} I)^2 = 2^2 = 4 \neq 2$. •

[10B] Determinant

10B.5 Is the Determinant Additive?

Prove or give a counterexample: if $S, T \in \mathcal{L}(V)$, then $\det(S + T) = \det S + \det T$.

Solution. This is false: again take $S = T = I$. Then $\det(2I) = 4$, but $2 \det I = 2 \neq 4$. •

10B.6 Determinant of a Block Upper-Triangular Matrix

Suppose A is a block upper-triangular matrix $A = \operatorname{diag}(A_1, \ldots, A_m)$ where each A_j is square. Prove that $\det A = (\det A_1) \cdots (\det A_m)$.

Proof. Each block A_j represents some operator on some subset of the basis of V. For each A_j, find an appropriate basis such that A_j is upper triangular and A remains block triangular. This turns A into an upper-triangular matrix. Then the eigenvalues of the A_j lie on their diagonals, and the eigenvalues of A consist of all of the eigenvalues of the A_j, so the statement is proved. □