MA3104

Linear Algebra II

Autumn 2021

Satvik Saha
19MS154

Indian Institute of Science Education and Research, Kolkata,
Mohanpur, West Bengal, 741246, India.

Contents

1 Linear operators on a vector space

1.1 Preliminaries

We discuss finite dimensional vector spaces V over some field $\mathbb{F}$, along with linear operators $T\colon V \to V$. We also assume that V has the inner product $\langle \cdot, \cdot \rangle$.

Theorem 1.1. *Let $\mathscr{L}(V)$ be the set of all linear operators on the vector space V. Then, $\mathscr{L}(V)$ is a linear algebra over the field $\mathbb{F}$.*

1.2 Ideals in a ring

Definition 1.1. Let $(R, +, \cdot)$ be a ring, where $(R, +)$ is its additive subgroup. A set $I \subseteq R$ is a left ideal of R if $(I, +)$ is a subgroup of $(R, +)$, and $rx \in I$ for every $r \in R$, $x \in I$.

Example. Let $\mathbb{Z}$ be the ring of integers. For some $n \in \mathbb{N}$, the set $n\mathbb{Z}$ is an ideal. In fact, these are the only ideals (along with $\{0\}$).

Definition 1.2. The principal left ideal generated by $x \in R$ is the set

$$I_x = Rx = \{rx : r \in R\}.$$

Example. In the ring of integers $\mathbb{Z}$, every ideal is a principal ideal. This follows directly from the fact that $(\mathbb{Z}, +)$ is a cyclic group, thus any subgroup is cyclic and generated by a single element.

Let $I \subseteq \mathbb{Z}$ be an ideal. If $I = \{0\}$, we are done. Otherwise, let n be the smallest positive integer in I (note that if $a \in I$, then $-a \in I$ which means that I must contain positive integers). This immediately gives $I \supseteq n\mathbb{Z}$. Now for any $m \in I$, use Euclid's Division Lemma to write $m = nq + r$, where $q, r \in \mathbb{Z}$, $0 \leq r < n$. Since I is an ideal, $nq \in I$ hence $m - nq = r \in I$. The minimality of n in I forces $r = 0$, hence $m = nq$ and $I \subseteq n\mathbb{Z}$. This proves $I = n\mathbb{Z}$.

Theorem 1.2. *Let $\mathbb{F}$ be a field and let $\mathbb{F}[x]$ denote the ring of polynomials with coefficients from $\mathbb{F}$. Then, every ideal in $\mathbb{F}[x]$ is a principal ideal.*

Remark. This is analogous to the theorem which states that every subgroup of a cyclic group is cyclic. Both lead to a precise definition of the greatest common divisor.

Corollary 1.2.1. *Let I be a non-trivial ideal in $\mathbb{F}[x]$. Then, there exists a unique monic polynomial $p \in \mathbb{F}[x]$ (leading coefficient 1) such that I is precisely the principal ideal generated by p.*

1.3 Eigenvalues and eigenvectors

Definition 1.3. Let $T \in \mathscr{L}(V)$ and $c \in \mathbb{F}$. We say that c is an eigenvalue or characteristic value of T if $T\boldsymbol{v} = c\boldsymbol{v}$ for some non-zero $\boldsymbol{v} \in V$. The vector $\boldsymbol{v}$ is called an eigenvector of T.

Theorem 1.3. *Let $T \in \mathscr{L}(V)$ and $c \in \mathbb{F}$. The following are equivalent.*

1. *c is an eigenvalue of T.*
2. *$T - cI$ is singular.*
3. $\det(T - cI) = 0$.

Definition 1.4. The polynomial $\det(T - xI)$ is called the characteristic polynomial of T.

Definition 1.5. Two linear operators $S, T \in \mathscr{L}(V)$ are similar if there exists an invertible operator $X \in \mathscr{L}(V)$ such that $S = X^{-1}TX$.

Remark. Similarity is an equivalence relation on $\mathscr{L}(V)$, thus partitioning it into similarity classes.

Lemma 1.4. *Similar linear operators have the same characteristic polynomial.*

Proof. Let S, T be similar with $S = X^{-1}TX$. Then,

$$\begin{aligned} \det(S - xI) &= \det(X^{-1}TX - xX^{-1}X) \\ &= \det(X^{-1}) \det(T - xI) \det(X) \\ &= \det(T - xI). \end{aligned}$$

□

Definition 1.6. A linear operator $T \in \mathscr{L}(V)$ is diagonalizable if there is a basis of V consisting of eigenvectors of T.

Remark. The matrix of T with respect to such a basis is diagonal.

Theorem 1.5. *Let $T \in \mathscr{L}(V)$ where V is finite dimensional, let $c_1, \ldots, c_k$ be distinct eigenvalues of T, and let $W_i = \ker(T - c_iI)$ be the corresponding eigenspaces. The following are equivalent.*

1. *T is diagonalizable.*
2. *The characteristic polynomial of T is of the form*

$$f(x) = (x - c_1)^{d_1} \ldots (x - c_k)^{d_k}$$

   where each $d_i = \dim W_i$.
3. $\dim V = \dim W_1 + \cdots + \dim W_k$.

1.4 Annihilating polynomials

Definition 1.7. An polynomial p such that $p(T) = 0$ for a given linear operator $T \in \mathscr{L}(V)$ is called an annihilating polynomial of T.

Lemma 1.6. *Every linear operator $T \in \mathscr{L}(V)$, where V is finite dimensional, has a non-trivial annihilating polynomial.*

Proof. Note that the operators $I, T, T^2, \ldots, T^{n^2} \in \mathscr{L}(V)$, of which there are n^2+1, are linearly dependent, since $\dim \mathscr{L}(V) = n^2$. $\square$

Lemma 1.7. *The annihilating polynomials of T form an ideal in $\mathbb{F}[x]$.*

Definition 1.8. The minimal polynomial of T is the unique monic generator of the annihilating polynomials of T.

Remark. The minimal polynomial of T divides all its annihilating polynomials.

Theorem 1.8. *The minimal polynomial and characteristic polynomial of T share the same roots, except for multiplicities.*

Proof. Let p be the minimal polynomial of T and let f be its characteristic polynomial.

First, let $c \in \mathbb{F}$ be a root of the minimal polynomial, i.e. $p(c) = 0$. The Division Algorithm guarantees

$$p(x) = (x - c)\, q(x)$$

for some monic polynomial q. By the minimality of the degree of p, we have $q(T) \neq 0$, hence there exists non-zero $\boldsymbol{v} \in V$ such that $\boldsymbol{w} = q(T)\, \boldsymbol{v} \neq \mathbf{0}$. Thus, $p(T)\, \boldsymbol{v} = \mathbf{0}$ gives

$$(T - cI)\, q(T)\, \boldsymbol{v} = \mathbf{0}, \qquad T\boldsymbol{w} = c\boldsymbol{w},$$

which shows that c is an eigenvalue, i.e. a root of the characteristic polynomial f.

Next, suppose that c is a root of the characteristic polynomial, i.e. $f(c) = 0$. Thus, c is an eigenvalue of T, hence there exists non-zero $\boldsymbol{v} \in V$ such that $T\boldsymbol{v} = c\boldsymbol{v}$. This gives $p(T)\, \boldsymbol{v} = p(c)\, \boldsymbol{v}$, but $p(T) = 0$ identically, forcing $p(c) = 0$. $\square$

Theorem 1.9 (Cayley-Hamilton)**.** *The characteristic polynomial of T annihilates T.*

Proof. Set $S = \operatorname{adj}(T - xI)$. This is a matrix with polynomial entries, satisfying

$$(T - xI)S = \det(T - xI)I = f(x)I,$$

where f is the characteristic polynomial of T. Now, we can also collect the powers x^n from S and write

$$S = \sum_{k=0}^{n-1} x^k S_k$$

for matrices S_k. Now, calculate

$$\begin{aligned} f(x)I &= (T - xI)S \\ &= (T - xI) \sum_{k=0}^{n-1} x^k S_k \\ &= -x^k S_{k-1} + \sum_{k=1}^{n-1} x^k (TS_k - S_{k-1}) + TS_0. \end{aligned}$$

Compare coefficients with

$$f(x)I = x^n I + a_{n-1}x^{n-1} + \cdots + a_0 I$$

to get

$$S_{n-1} = -I, \qquad TS_0 = a_0 I, \qquad TS_k - S_{k-1} = a_k I \text{ for } 1 \leq k \leq n-1.$$

Thus,

$$\begin{aligned} f(T) &= \sum_{k=0}^{n} a_k T^k \\ &= -T^n S_{n-1} + \sum_{k=1}^{n-1} (TS_k - S_{k-1})T^k + TS_0 \\ &= 0. \end{aligned}$$

□

Corollary 1.9.1. *The minimal polynomial of T divides its characteristic polynomial.*

Corollary 1.9.2. *The minimal polynomial of T in a finite-dimensional vector space V is at most* $\dim V$*.*

Theorem 1.10. *The minimal polynomial for a diagonalizable linear operator T in a finite-dimensional vector space is*

$$p(x) = (x - c_1) \ldots (x - c_k),$$

where $c_1, \ldots, c_k$ are distinct eigenvalues of T.

Proof. The diagonalizability of T implies that V admits a basis of eigenvectors of T. Thus, for any such eigenvector $\boldsymbol{v}_i$, the operator $T - c_i I$ kills it where c_i is the corresponding eigenvalue. Thus, $p(T)\boldsymbol{v}_i$ vanishes for every basis vector $\boldsymbol{v}_i$ □

Remark. The converse is also true, i.e. T is diagonalizable if and only if the minimal polynomial is the product of distinct linear factors.

1.5 Invariant subspaces

Definition 1.9. Let $T \in \mathscr{L}(V)$ where V is finite-dimensional, and let $W \subseteq V$ be a subspace. We say that W is invariant under T if $T(W) \subseteq W$.

If a subspace W is invariant under T, we define the linear map $T_W \in \mathscr{L}(W)$ as the restriction of T to W in the natural way, by setting $T_W(\boldsymbol{w}) = T(\boldsymbol{w})$ for all $\boldsymbol{w} \in W$.

Lemma 1.11. *If W is an invariant subspace under $T \in \mathscr{L}(V)$, then there is a basis of V in which T has the block triangular form*

$$[T]_\beta = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix},$$

where A is an $r \times r$ matrix, $r = \dim W$.

Proof. Let $\beta_W = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$ be an ordered basis of W, and extend it to an ordered basis $\beta = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ of V. Thus, the matrix $[T]_\beta$ has coefficients a_{ij} such that

$$T\boldsymbol{v}_j = a_{1j}\boldsymbol{v}_1 + \cdots + a_{rj}\boldsymbol{v}_r + \cdots + a_{nj}\boldsymbol{v}_n.$$

However for all $j \leq r$, $T\boldsymbol{v}_j \in W$ by the invariance of W, so the coefficients of $\boldsymbol{v}_{i>r}$ in the expansion of $T\boldsymbol{v}_j$ must vanish. Thus, all $a_{ij} = 0$ where $i > r$, $j \leq r$. □

Lemma 1.12. *If W is an invariant subspace under $T \in \mathscr{L}(V)$, the characteristic polynomial of T_W divides the characteristic polynomial of T, and the minimal polynomial of T_W divides the minimal polynomial of T.*

Proof. Choose an ordered basis β of V such that

$$[T]_\beta = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = D.$$

Note that the matrix of T_W in the restricted basis β_W is just A. It can be shown that

$$\det(xI - D) = \det(xI - A)\,\det(xI - C),$$

which immediately gives the first result.

Now, it can also be shown that the powers of D are of the form

$$[T^k]_\beta = \begin{bmatrix} A^k & B_k \\ 0 & C^k \end{bmatrix} = D^k.$$

Now, $T^k\boldsymbol{v} = \boldsymbol{0}$ implies $T_W^k\boldsymbol{v} = \boldsymbol{0}$, hence any polynomial which annihilates T also annihilates T_W. This gives the second result. □

Definition 1.10. Let W be an invariant subspace under $T \in \mathscr{L}(V)$, and let $\boldsymbol{v} \in V$. We define the T-conductor of $\boldsymbol{v}$ into W as the set $S_T(\boldsymbol{v}; W)$ of all polynomials g such that $g(T)\boldsymbol{v} \in W$.

When $W = \{0\}$, $S_T(\boldsymbol{v}, \{0\})$ is called the T-annihilator of $\boldsymbol{v}$.

Lemma 1.13. *If W is invariant under T, then it is invariant under all polynomials of T. Thus, the conductor $S_T(\boldsymbol{v}, W)$ is an ideal in the ring of polynomials $\mathbb{F}[x]$.*

Definition 1.11. If W is an invariant subspace under $T \in \mathscr{L}(V)$, and $\boldsymbol{v} \in V$, then the unique monic generator of $S_T(\boldsymbol{v}, W)$ is also called the T-conductor of $\boldsymbol{v}$ into W.

The unique monic generator of $S_T(\boldsymbol{v}, \{0\})$ is also called the T-annihilator of $\boldsymbol{v}$.

Remark. The T-annihilator of $\boldsymbol{v}$ is the unique monic polynomial g of least degree such that $g(T)\boldsymbol{v} = \mathbf{0}$.

Remark. The minimal polynomial is a T-conductor for every $\boldsymbol{v} \in V$, thus every T-conductor divides the minimal polynomial of T.

Lemma 1.14. *Let $T \in \mathscr{L}(V)$ for finite-dimensional V, where the minimal polynomial of T is a product of linear operators*

$$p(x) = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}.$$

Let W be a proper subspace of V which is invariant under T. Then, there exists a vector $\boldsymbol{v} \in V$ such that $\boldsymbol{v} \notin W$, and $(T - cI)\boldsymbol{v} \in W$ for some eigenvalue c.

Proof. What we must show is that the T-conductor of $\boldsymbol{v}$ into W is a linear polynomial. Choose arbitrary $\boldsymbol{w} \in V \setminus W$, and let g be the T-conductor of $\boldsymbol{w}$ into W. Thus, g divides the minimal polynomial of T, and hence is a product of linear factors of the form $x - c_i$ for eigenvalues c_i. Thus write

$$g = (x - c_i)h.$$

The minimality of g ensures that $\boldsymbol{v} = h(T)\boldsymbol{w} \notin W$. Finally, note that

$$(T - c_iI)\boldsymbol{v} = (T - c_iI)h(T)\boldsymbol{w} = g(T)\boldsymbol{w} \in W.$$ □

1.6 Triangulability and diagonalizability

Theorem 1.15. *Let $T \in \mathscr{L}(V)$ for finite-dimensional V. Then, T is triangulable if and only if the minimal polynomial is a product of linear polynomials.*

Proof. First suppose that the minimal polynomial is of the form

$$p(x) = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}.$$

We want to find an ordered basis $\beta = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ in which

$$[T]_\beta = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Thus, we demand

$$T\boldsymbol{v}_j = a_{1j}\boldsymbol{v}_1 + \cdots + a_{jj}\boldsymbol{v}_j,$$

i.e. each $T\boldsymbol{v}_j$ is in the span of $\boldsymbol{v}_1, \dots, \boldsymbol{v}_j$.

Apply the previous lemma on $W = \{\mathbf{0}\}$ to obtain $\boldsymbol{v}_1$. Next, let W_1 be the subspace spanned by $\boldsymbol{v}_1$ and use the lemma to obtain $\boldsymbol{v}_2$. Then let W_2 be the subspace spanned by $\boldsymbol{v}_1, \boldsymbol{v}_2$ and use the lemma to obtain $\boldsymbol{v}_3$, and so on. Note that at each step, the newly generated vector $\boldsymbol{v}_j$ satisfies $\boldsymbol{v}_j \notin W_{j-1}$ and $(T - c_i I)\boldsymbol{v}_j \in W_{j-1}$, hence

$$T\boldsymbol{v}_j = a_{ij}\boldsymbol{v}_1 + \cdots + a_{(j-1)j}\boldsymbol{v}_{j-1} + c_i\boldsymbol{v}_j$$

as desired.

Next, suppose that T is triangulable. Thus, there is a basis in which the matrix of T is diagonal, which immediately means that the characteristic polynomial is the product of linear factors $x - a_{ii}$. Furthermore, the diagonal elements are precisely the eigenvalues of T. Since the minimal polynomial divides the characteristic polynomial, it too is a product of linear polynomials. □

Corollary 1.15.1. *In an algebraically closed field $\mathbb{F}$, any $n \times n$ matrix over $\mathbb{F}$ is triangulable.*

Theorem 1.16. *Let $T \in \mathscr{L}(V)$ for finite-dimensional V. Then, T is diagonalizable if and only if the minimal polynomial is a product of distinct linear factors, i.e.*

$$p(x) = (x - c_1) \ldots (x - c_k)$$

where c_i are distinct eigenvalues of T.

Proof. We have already shown that if T is diagonalizable, then its minimal polynomial must have the given form.

Next, let the minimal polynomial of T have the given form. Let W be the subspace spanned by all eigenvectors of V. Suppose that $W \neq V$. Using the fact that W is an invariant subspace under T and the previous lemma, we find $\boldsymbol{v} \notin W$ and an eigenvalue c_j such that $\boldsymbol{w} = (T - c_j I)\boldsymbol{v} \in W$. Now, $\boldsymbol{w}$ can be written as the sum of eigenvectors

$$\boldsymbol{w} = \boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k$$

where each $T\boldsymbol{w}_i = c_i\boldsymbol{w}_i$. Thus for every polynomial h, we have

$$h(T)\boldsymbol{w} = h(c_1)\boldsymbol{w}_1 + \cdots + h(c_k)\boldsymbol{w}_k \in W.$$

Since c_j is an eigenvalue of T, write $p = (x - c_j)q$ for some polynomial q. Further write $q - q(c_j) = (x - c_j)h$ using the Remainder Theorem. Thus,

$$q(T)\boldsymbol{v} - q(c_j)\boldsymbol{v} = h(T)(T - c_j I)\boldsymbol{v} = h(T)\boldsymbol{w} \in W.$$

Since

$$\boldsymbol{0} = p(T)\boldsymbol{v} = (T - c_j I)q(T)\boldsymbol{v},$$

the vector $q(T)\boldsymbol{v}$ is an eigenvector and hence in W. However, $\boldsymbol{v} \notin W$, forcing $q(c_j) = 0$. This contradicts the fact that the factor $x - c_j$ appears only once in the minimal polynomial. □

1.7 Simultaneous triangulation and diagonalization

Definition 1.12. Let V be a finite-dimensional vector space, and let $\mathscr{F}$ be a family of linear operators on V. The family $\mathscr{F}$ is said to be simultaneously triangulable if there exists a basis of V in which every operator in $\mathscr{F}$ is represented by an upper triangular matrix.

An analogous definition holds for simultaneous diagonalizability.

Lemma 1.17. *Let $\mathscr{F}$ be a simultaneously diagonalizable family of linear operators. Then, every pair of operators from $\mathscr{F}$ commute.*

Proof. This follows trivially from the fact that diagonal matrices commute. □

Definition 1.13. A subspace W is invariant under a family of linear operators $\mathscr{F}$ if it is invariant under every operator $T \in \mathscr{F}$.

Lemma 1.18. *Let $\mathscr{F}$ be a commuting family of triangulable linear operators on V, and let $W \subset V$ be a proper subspace invariant under $\mathscr{F}$. Then, there exists a vector $\boldsymbol{v} \in V$ such that $\boldsymbol{v} \notin W$ and $T\boldsymbol{v} \in \operatorname{span}\{\boldsymbol{v}, W\}$ for each $T \in \mathscr{F}$.*

Proof. We observe that we can assume that $\mathscr{F}$ contains only finitely many operators, without loss of generality. This is because of the finite dimensionality of V, which enables us to pick a finite basis of $\mathscr{L}(V)$.

Using Lemma 1.14, we can find vectors $\boldsymbol{v}_1 \notin W$ and c_1 such that $(T_1 - c_1 I)\boldsymbol{v}_1 \in W$, for $T_1 \in \mathscr{F}$. Define

$$V_1 = \{\boldsymbol{v} \in V : (T_1 - c_1 I)\boldsymbol{v} \in W\}.$$

Note that V_1 is a subspace which properly contains W. Furthermore, V_1 is invariant under $\mathscr{F}$ – this uses the fact that the operators from $\mathscr{F}$ commute. Now, let U_2 be the restriction of T_2 to V_1. Apply the lemma the find to U_2, W, V_1 to obtain $\boldsymbol{v}_2 \in V_1$, $\boldsymbol{v}_2 \notin W$ such that $(U_2 - c_2 I)\boldsymbol{v}_2 \in W$. Note that $(T_i - c_i I)\boldsymbol{v}_2 \in W$ for $i = 1, 2$. Construct V_2 as before, and repeat this process until we have exhausted all linear operators in $\mathscr{F}$. The final vector $\boldsymbol{v}_j$ satisfies the desired properties. □

Theorem 1.19. *Let $\mathscr{F}$ be a commuting family of triangulable linear operators on V. There exists an ordered basis of V which simultaneously triangulates $\mathscr{F}$.*

Proof. The proof is identical to that of Theorem 1.15. □

Theorem 1.20. *Let $\mathscr{F}$ be a commuting family of diagonalizable linear operators on V. There exists an ordered basis of V which simultaneously diagonalizes $\mathscr{F}$.*

Proof. We perform induction on the dimension of V. The theorem is trivial when $\dim V = 1$; suppose that it holds for vector spaces of dimension less than n, and let $\dim V = n$. Pick $T \in \mathscr{F}$ such that T is not a scalar multiple of I_n. Let $c_1, \dots, c_k$ be distinct eigenvalues of T, and let W_i be the corresponding eigenspaces. Each W_i is invariant under all operators which commute with T. Now let $\mathscr{F}_i$ be the family of operators from $\mathscr{F}$, restricted to the invariant subspace W_i. Note that each operator in $\mathscr{F}_i$ is diagonalizable. Furthermore, $\dim W_i < \dim V$, so the induction hypothesis says that $\mathscr{F}_i$ is simultaneously diagonalizable; let β_i be the corresponding basis. Each vector in β_i is an eigenvector for every operator in $\mathscr{F}_i$. Let β consist of the such vectors from all β_i generated in this way. Since T is diagonal, this is indeed an basis of V, as desired. □

1.8 Direct sum decompositions

Definition 1.14. Let $W_1, \ldots, W_k$ be subspaces of V. We say that these W_i are independent if

$$\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k = \mathbf{0}$$

where $\boldsymbol{w}_i \in W_i$ implies that each $\boldsymbol{w}_i = \mathbf{0}$.

Lemma 1.21. *If $W_1, \ldots, W_k$ are independent, then each vector $\boldsymbol{w} \in W_1 + \ldots + W_k$ has a unique representation*

$$\boldsymbol{w} = \boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k$$

where each $\boldsymbol{w}_i \in W_i$.

Definition 1.15. The sum of independent subspaces $W_1 + \cdots + W_k$ is called a direct sum, denoted

$$W_1 \oplus \cdots \oplus W_k.$$

Lemma 1.22. *Let V be a finite-dimensional vector space, let $W_1, \ldots, W_k$ be subspaces of V, and let $W = W_1 + \cdots + W_k$. Then, the following are equivalent.*

1. *$W_1, \ldots, W_k$ are independent.*
2. *For each $2 \le j \le k$,*
$$W_j \cap (W_1 + \cdots + W_{j-1}) = \{\mathbf{0}\}.$$
3. *If β_i are bases of W_i, then the set β consisting of all these vectors is a basis of W.*

1.9 Projections maps

Definition 1.16. A projection map on a vector space V is a linear operator E such that $E^2 = E$. In other words, E is idempotent.

Lemma 1.23. *Let E be a projection map on V, and let $R = \operatorname{im} E$, $N = \ker E$.*

1. *A vector $\boldsymbol{v} \in R$ if and only if $E\boldsymbol{v} = \boldsymbol{v}$.*
2. *Any vector $\boldsymbol{v} \in V$ has the unique representation $\boldsymbol{v} = E\boldsymbol{v} + (\boldsymbol{v} - E\boldsymbol{v})$, with $E\boldsymbol{v} \in R$ and $\boldsymbol{v} - E\boldsymbol{v} \in N$.*
3. *$V = R \oplus N$.*

Remark. If R and N are two subspaces of V such that $V = R \oplus N$, then there is exactly one projection map E such that $R = \operatorname{im} E$ and $N = \ker E$. Namely, send $\boldsymbol{v} \mapsto \boldsymbol{v}_R$ where $\boldsymbol{v} = \boldsymbol{v}_R + \boldsymbol{v}_N$ is the unique decomposition of $\boldsymbol{v}$.

Lemma 1.24. *A projection map is trivially diagonalizable.*

Proof. Note that $x^2 - x = x(x-1)$ annihilates any projection map. Also note that any projection map restricted to its range is the identity map. Thus, $\operatorname{trace} E = \operatorname{rank} E$. □

Lemma 1.25. *Let $V = W_1 \oplus \cdots \oplus W_k$, and let $\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k$ with $\boldsymbol{v}_i \in W_i$. Define the maps E_i such that $E_i\boldsymbol{v} = \boldsymbol{v}_i$. Then, each E_i is the projection map along W_i.*

Remark. Observe that

$$I = E_1 + \cdots + E_k.$$

Furthermore, we have $E_iE_j = 0$ for all $i \neq j$, which means that $\operatorname{im} E_j \subseteq \ker E_i$.

Theorem 1.26. *If $V = W_1 + \cdots + W_k$, then there exist k linear operators $E_1, \ldots, E_k$ on V such that*

1. $E_i^2 = E_i$.
2. $E_iE_j = 0$ *for all* $i \neq j$.
3. $I = E_1 + \cdots + E_k$.
4. $\operatorname{im} E_i = W_i$.

Conversely, if there exist linear k linear operators which satisfy properties 1, 2, 3 and label $\operatorname{im} E_i = W_i$*, then $V = W_1 \oplus \cdots \oplus W_k$.*

Proof. We only need to prove the converse. Let $E_i, \ldots, E_k$ satisfy the properties 1, 2, 3 and let $\operatorname{im} E_i = W_i$. Pick $\boldsymbol{v} \in V$, hence

$$\boldsymbol{v} = I_k\boldsymbol{v} = E_1\boldsymbol{v} + \cdots + E_k\boldsymbol{v} \in W_1 + \cdots + W_k,$$

which shows that $V = W_1 + \cdots + W_k$. We claim that this representation of $\boldsymbol{v}$ is unique. In other words, suppose that

$$\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k$$

where each $\boldsymbol{v}_i \in W_i$; we claim that $\boldsymbol{v}_i = E_i\boldsymbol{v}$ is the only choice. Since $\boldsymbol{v}_i \in W_i$, write $\boldsymbol{v}_i = E_i\boldsymbol{w}_i$. Then,

$$E_j\boldsymbol{v} = \sum_{i=1}^{k} E_j\boldsymbol{v}_i = \sum_{i=1}^{k} E_jE_i\boldsymbol{w}_i = E_j^2\boldsymbol{w}_j = E_j\boldsymbol{w}_j = \boldsymbol{v}_j.$$

□

Definition 1.17. Let $V = W_1 \oplus \cdots \oplus W_k$, and let $T \in \mathscr{L}(V)$. Additionally, let each W_i be invariant under T, hence $T\boldsymbol{v}_i \in W_i$. Define the linear operators $T_i \in \mathscr{L}(W_i)$, which are the restrictions of T to W_i. Then, given any $\boldsymbol{v} \in V$, there is a unique representation $\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k$ where $\boldsymbol{v}_i \in W_i$, so

$$T\boldsymbol{v} = T\boldsymbol{v}_1 + \cdots + T\boldsymbol{v}_k = T_1\boldsymbol{v}_1 + \cdots + T_k\boldsymbol{v}_k = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k.$$

This representation must be unique. We say that T is the direct sum of the linear operators $T_1, \ldots, T_k$.

Lemma 1.27. *Let $V = W_1 \oplus \cdots \oplus W_k$, let β_i be ordered basses of W_i, and let β be the basis formed by combining all these vectors. Let $T \in \mathscr{L}(V)$ and suppose that each W_i is invariant under T. Then, by setting $[T_i]_{\beta_i} = A_i$, we have the block diagonal form*

$$[T]_\beta = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{bmatrix}.$$

Theorem 1.28. *Let $V = W_1 \oplus \cdots \oplus W_k$, let E_i be the projections along W_i, and $T \in \mathscr{L}(V)$. Then, each W_i is invariant under T if and only if T commutes with each of the projections E_i.*

Proof. Suppose that T commutes with each E_i, i.e. $TE_i = E_iT$. We want to show that each $W_i = \operatorname{im} E_i$ is invariant under T. Let $\boldsymbol{v} \in W_i$, hence $\boldsymbol{v} = E_i\boldsymbol{v}$ and

$$T\boldsymbol{v} = TE_i\boldsymbol{v} = E_iT\boldsymbol{v}.$$

Thus, $T\boldsymbol{v} \in W_i$ as desired.

Conversely, suppose that each W_i is invariant under T. Pick $\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k \in V$ where $\boldsymbol{v}_i \in W_i$. Set $\boldsymbol{w}_i = T\boldsymbol{v}_i \in W_i$, and compute

$$E_iT\boldsymbol{v} = E_iT(\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k) = E_i(\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k) = \boldsymbol{w}_i = T\boldsymbol{v}_i = TE_i\boldsymbol{v}. \qquad \square$$

Theorem 1.29. *Let $T \in \mathscr{L}(V)$ where V is a finite-dimensional vector space. If T is diagonalizable and $c_1, \ldots, c_k$ are the distinct eigenvalues of T, then there are non-zero linear operators $E_1, \ldots, E_k$ on V which satisfy the following.*

1. $T = c_1E_1 + \cdots + c_kE_k$.
2. $I = E_1 + \cdots + E_k$.
3. $E_iE_j = 0$ *for all* $i \neq j$.
4. $E_i^2 = E_i$.
5. $\operatorname{im} E_i = \ker(T - c_iI)$.

Conversely, if there exist k distinct scalars $c_1, \ldots, c_k$ and k non-zero linear operators which satisfy properties 1, 2, 3, then T is diagonalizable, $c_1, \ldots, c_k$ are the eigenvalues of T, and properties $4, 5$ are also satisfied.

Proof. Suppose that T is diagonalizable, with distinct eigenvalues $c_1, \ldots, c_k$. Let $W_i = \ker(T - c_iI)$, and note that $V = W_1 \oplus \cdots \oplus W_k$. Let $E_1, \ldots, E_k$ be the projections associated with this decomposition. This immediately gives us the properties 2, 3, 4, 5. To show that property 1 holds, pick arbitrary $\boldsymbol{v} \in V$ and write $\boldsymbol{v} = E_1\boldsymbol{v} + \cdots + E_k\boldsymbol{v}$. Then, note that $E_i\boldsymbol{v}$ are eigenvectors, hence

$$T\boldsymbol{v} = TE_1\boldsymbol{v} + \cdots + TE_k\boldsymbol{v} = c_1E_1\boldsymbol{v} + \cdots + c_kE_k\boldsymbol{v}.$$

Conversely, let $T \in \mathscr{L}(V)$ and suppose that $c_1, \ldots, c_k$ and non-zero $E_1, \ldots, E_k$ satisfy properties 1, 2, 3. Then, note that

$$E_i = E_iI = E_i(E_1 + \cdots + E_k) = E_i^2,$$

giving property 4. Also,

$$TE_i = (c_1E_1 + \cdots + c_kE_k)E_i = c_iE_i^2 = c_iE_i,$$

hence $\operatorname{im} E_i \neq \{\mathbf{0}\}$ is an eigenspace of T corresponding to the eigenvalue c_i, i.e. $\operatorname{im} E_i \subseteq \ker(T - c_iI)$. We claim that there are no other eigenvalues; suppose that $\ker(T - cI)$ is non-zero. Write

$$T - cI = c_1E_1 + \cdots + c_kE_k - cI = (c_1 - c)E_1 + \cdots + (c_k - c)E_k.$$

Pick non-zero $\boldsymbol{v} \in V$ such that $(T - cI)\boldsymbol{v} = 0$. Then, some $E_i\boldsymbol{v} \neq \mathbf{0}$ (this is because the images of the projection operators are independent, and $I = E_1 + \cdots + E_k$). On the other hand, we must have each $(c_i - c)E_i\boldsymbol{v} = \mathbf{0}$, forcing $c = c_i$. Finally, $I = E_1 + \cdots + E_k$ says that V is the direct sum of the $\operatorname{im} E_i$, which are are contained within the eigenspaces of T. This means that T is diagonalizable.

We finally show that $\operatorname{im} E_i = \ker(T - c_iI)$. Pick $\boldsymbol{v} \in \ker(T - c_iI)$, which means that

$$(c_1 - c_i)E_1\boldsymbol{v} + \cdots + (c_k - c_i)E_k\boldsymbol{v} = 0.$$

By the independence of each $\operatorname{im} E_i$, each $(c_j - c_i)E_j\boldsymbol{v} = \mathbf{0}$, or $E_j\boldsymbol{v} = \mathbf{0}$ for $j \neq i$. Thus,

$$\boldsymbol{v} = E_1\boldsymbol{v} + \cdots + E_k\boldsymbol{v} = E_i\boldsymbol{v},$$

so $\boldsymbol{v} \in \operatorname{im} E_i$. This proves that $\operatorname{im} E_i = \ker(T - c_iI)$. □

Lemma 1.30. *The Lagrange polynomials p_i of degree n form a basis of the vector space of polynomials of degree at most n. If we have $p_i(t_j) = \delta_{ij}$, then for any polynomial f of degree n, we have*

$$f = \sum f(t_i)p_i.$$

Lemma 1.31. *If T is diagonalizable with $T = c_1E_1 + \cdots + c_kE_k$ where E_i are projections as discussed earlier, Then, for any polynomial g, we have*

$$g(T) = g(c_1)E_1 + \cdots + g(c_k)E_k.$$

Thus, if $p_1, \ldots, p_k$ are the Lagrange polynomials corresponding to the points $c_1, \ldots, c_k$ and we put $g = c_i$, then each $p_i(T) = E_i$. Thus, each E_i is a polynomial in T.

Theorem 1.32 (Primary Decomposition Theorem)**.** *Let $T \in \mathscr{L}(V)$ where V is finite-dimensional, and let p be the minimal polynomial of T, where*

$$p = p_1^{r_1} \ldots p_k^{r_k}$$

where p_i are distinct, irreducible polynomials. Let $W_i = \ker p_i(T)^{r_i}$, then

1. *$V = W_1 \oplus \cdots \oplus W_k$.*
2. *Each W_i is invariant under T.*
3. *If T_i is the restriction of T to W_i, then the minimal polynomial of T_i is $p_i^{r_i}$.*

Proof. Set

$$f_i = \frac{p}{p_i^{r_i}} = \prod_{j \neq i} p_j^{r_j}.$$

Since the polynomials f_i are relatively prime, we can pick polynomials g_i such that

$$f_1 g_1 + \cdots + f_k g_k = 1.$$

Note that when $i \neq j$, we have $p | f_i f_j$. Set $h_i = f_i g_i$, and let $E_i = h_i(T)$. We have $E_1 + \cdots + E_k = I$, and $E_i E_j = 0$ for $i \neq j$ (the $f_i f_j(T)$ term contains $p(T) = 0$). This shows that E_i are projections corresponding to some direct sum decomposition of V. We claim that $\operatorname{im} E_i = W_i$. To see this, first let $\boldsymbol{v} \in \operatorname{im} E_i$, whence $\boldsymbol{v} = E_i \boldsymbol{v}$ so

$$p_i(T)^{r_i} \boldsymbol{v} = p_i(T)^{r_i} E_i \boldsymbol{v} = p_i(T)^{r_i} f_i(T) g_i(T) \boldsymbol{v} = \mathbf{0}.$$

Conversely, if $\boldsymbol{v} \in W_i$, when $p_i(T)^{r_i} \boldsymbol{v} = \mathbf{0}$. Now, for $i \neq j$, we have $p_i^{r_i} | f_j g_j$ hence $E_j \boldsymbol{v} = f_j g_j(T) \boldsymbol{v} = \mathbf{0}$ for $i \neq j$. This leaves

$$\boldsymbol{v} = I\boldsymbol{v} = (E_1 + \cdots + E_k)\boldsymbol{v} = E_i \boldsymbol{v},$$

hence $\boldsymbol{v} \in \operatorname{im} E_i$. This proves 1.

It is clear that W_i is invariant under T. Pick arbitrary $\boldsymbol{v} \in W_i$, whence $\boldsymbol{v} = E_i \boldsymbol{v}$ so $T\boldsymbol{v} = TE_i \boldsymbol{v} = E_i T \boldsymbol{v} \in W_i$. This proves 2.

Since $p_i(T)^{r_i} = 0$ on W_i, we have $p_i(T_i)^{r_i} = 0$, hence the minimal polynomial of T_i divides $p_i^{r_i}$. Conversely, if $g(T_i) = 0$ for some polynomial g, then $g(T) f_i(T) = 0$ (g kills everything in W_i, while f_i kills everything in the other $W_j \neq W_i$). Thus, $p = p_i^{r_i} f_i$ divides $g f_i$, or $p_i^{r_i}$ divides g. Hence, the minimal polynomial of T_i is precisely $p_i^{r_i}$. This proves 3. □

Corollary 1.32.1. *Let $E_1, \ldots, E_k$ be the projections associated with the primary decomposition of T. Then, each E_i is a polynomial in T, so any operator which commutes with T must also commute with each E_i. The subspaces W_i are thus invariant under any any operator which commutes with T.*

Theorem 1.33. *Let $T \in \mathscr{L}(V)$ where V is finite-dimensional, and let the minimal polynomial of p be of the form*

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}.$$

Then, there is a unique diagonalizable operator D and a unique nilpotent operator N such that $T = D + N$, $DN = ND$, and both are polynomials in T.

Proof. Set $D = c_1 E_1 + \cdots + c_k E_k$, $N = T - D$. Note that D is diagonalizable, and

$$N = (T - c_1 I)E_i + \cdots + (T - c_k I)E_k.$$

It can be shown that

$$N^r = (T - c_1 I)^r E_i + \cdots + (T - c_k I)^r E_k,$$

hence $N^r = 0$ when r is equal to the maximum of the r_i.

We now claim that this choice of D and N is unique. Let D' and N' also satisfy the above properties; since D' and N' commute and $T = D' + N'$, all the operators T, D, N, D', N' commute. Write $D + N = D' + N'$, hence

$$D - D' = N' - N.$$

Since D and D' commute, they are simultaneously diagonalizable, hence $D-D'$ is diagonalizable. Now, note that
$$(N'-N)^r = \sum_{j=0}^{r} \binom{r}{j} (N')^{r-j}(-N)^j.$$
Since N' and N are both nilpotent, the right hand side is zero for sufficiently high r. In other words, $N'-N$ is nilpotent, hence so is $D-D'$. This forces $D=D'$, since the only nilpotent diagonalizable operator is the zero operator. □

1.10 Cyclic subspaces and the Rational form

Lemma 1.34. *Let $T \in \mathscr{L}(V)$ where V is finite-dimensional, and let $\boldsymbol{v} \in V$. There is a smallest invariant subspace W containing $\boldsymbol{v}$, namely the intersection of all invariant subspaces containing $\boldsymbol{v}$. Then, W is the collection of $g(T)\boldsymbol{v}$, for all polynomials g.*

Proof. It is clear that the collection $\{g(T)\boldsymbol{v}\}$ is a T-invariant subspace containing $\boldsymbol{v}$. We now show that this is contained within every T-invariant subspace containing $\boldsymbol{v}$. Let W' be a T-invariant subspace containing $\boldsymbol{v}$. Then, $T\boldsymbol{v} \in W'$, hence all $T^k\boldsymbol{v} \in W'$. This means that all polynomials $g(T)\boldsymbol{v} \in W'$, as desired. □

Definition 1.18. Let $T \in \mathscr{L}(V)$, and $\boldsymbol{v} \in V$. We define the T-cyclic subspace generated by $\boldsymbol{v}$ as
$$Z(\boldsymbol{v}, T) = \{g(T)\boldsymbol{v} : g \in \mathbb{F}[x]\}.$$
If $V = Z(\boldsymbol{v}, T)$, then $\boldsymbol{v}$ is called a cyclic vector for T.

Theorem 1.35. *Let $T \in \mathscr{L}(V)$, let $\boldsymbol{v} \in V$ be non-zero, and let $p_{\boldsymbol{v}}$ be the T-annihilator of $\boldsymbol{v}$. Then,*

1. $\dim Z(\boldsymbol{v}, T) = \deg p_{\boldsymbol{v}}$.
2. *If $\deg p_{\boldsymbol{v}} = k$, then $\boldsymbol{v}, T\boldsymbol{v}, \dots, T^{k-1}\boldsymbol{v}$ forms a basis of $Z(\boldsymbol{v}, T)$.*
3. *If U is the restriction of T to $Z(\boldsymbol{v}, T)$, then $p_{\boldsymbol{v}}$ is the minimal polynomial of U.*

Remark. If V contains a T-cyclic vector $\boldsymbol{v}$, then $Z(\boldsymbol{v}, T)$, then the minimal polynomial of T is precisely its characteristic polynomial. The converse of this is also true.

Proof. First note that
$$\mathbf{0} = p_{\boldsymbol{v}}(T)\boldsymbol{v} = a_k T^k \boldsymbol{v} + a_{k-1}T^{k-1}\boldsymbol{v} + \cdots + a_0\boldsymbol{v}.$$
Since $a_k \neq 0$, this immediately gives $T^k\boldsymbol{v}$ as a linear combination of $\boldsymbol{v}, \dots, T^{k-1}\boldsymbol{v}$. Thus, $Z(\boldsymbol{v}, T)$ is spanned by $\boldsymbol{v}, \dots, T^{k-1}\boldsymbol{v}$. The same thing can be shown by using the Division Lemma to write $g = p_{\boldsymbol{v}}q + r$ where $0 \leq \deg r < k$.

We now show that $\boldsymbol{v}, \dots, T^{k-1}\boldsymbol{v}$ are linearly independent. If not, then
$$a_0\boldsymbol{v} + \cdots + a_{k-1}T^{k-1}\boldsymbol{v} = \mathbf{0}$$
for at least one $a_i \neq 0$. This contradicts the minimality of the degree of the T-annihilator of $\boldsymbol{v}$. Thus, we have properties 1, 2.

Note that $p_{\boldsymbol{v}}(U) = 0$. Any polynomial of lower degree such that $p(U)\boldsymbol{v} = 0$ must be the zero polynomial by the linear independence of $\boldsymbol{v}, \dots, T^{k-1}\boldsymbol{v}$. This means that $p_{\boldsymbol{v}}$ must be the minimal polynomial of $Z(\boldsymbol{v}, T)$, proving 3. □

Definition 1.19. Let p be the following monic polynomial.

$$p(x) = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + x^k.$$

The following matrix is called its companion matrix.

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{bmatrix}.$$

Lemma 1.36. *Let $T \in \mathscr{L}(V)$ such that $\boldsymbol{v} \in V$ is a cyclic vector of T. Then, the matrix representation of T in the basis $\boldsymbol{v}, T\boldsymbol{v}, \dots, T^{n-1}\boldsymbol{v}$ is the companion matrix of the characteristic/minimal polynomial of T.*

Remark. If T is also nilpotent, then $T^n = 0$ hence the last column in our matrix vanishes.

Theorem 1.37. *Let $T \in \mathscr{L}(V)$. Then, T admits a cyclic vector if and only if there is an ordered basis of V in which the matrix representation of T is the companion matrix of its characteristic polynomial.*

Proof. If T admits a cyclic vector $\boldsymbol{v}$, we have already shown that the desired basis is $\{\boldsymbol{v}, T\boldsymbol{v}, \dots, T^{n-1}\boldsymbol{v}\}$.

Conversely, suppose that in the basis $\{\boldsymbol{v}_0, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{k-1}\}$, the matrix representation of T is the companion matrix of its characteristic polynomial. Then we immediately have $T\boldsymbol{v}_0 = \boldsymbol{v}_1$, $T\boldsymbol{v}_1 = \boldsymbol{v}_2$, ..., $T^{n-2}\boldsymbol{v}_{n-2} = \boldsymbol{v}_{n-1}$. This immediately shows that $\boldsymbol{v}_0$ is a cyclic vector of T. □

Corollary 1.37.1. *If A is companion matrix of a monic polynomial p, then p is both the minimal and characteristic polynomial of A.*

Corollary 1.37.2. *If $S, T \in \mathscr{L}(V)$ both have cyclic vectors in V, then they are similar if and only if they have the same characteristic polynomial.*

Definition 1.20. Let $T \in \mathscr{L}(V)$ and let $W \subseteq V$ be a T-invariant subspace. We say that W is T-admissible if the following condition holds: if $f(T)\boldsymbol{v} \in W$ for some polynomial f, then there exists $\boldsymbol{w} \in W$ such that $f(T)\boldsymbol{v} = f(T)\boldsymbol{w}$.

Theorem 1.38 (Cyclic Decomposition Theorem)**.** *Let $T \in \mathscr{L}(V)$, and let $W_0 \subset V$ be a proper T-admissible subspace. Then, there exist non-zero vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_1$, with respective T-annihilators $p_1, \dots, p_r$ such that*

1. *$V = W_0 \oplus Z(\boldsymbol{v}_1, T) \oplus \cdots \oplus Z(\boldsymbol{v}_r, T)$.*
2. *p_k divides p_{k-1}.*

The integer r and the annihilators $p_1, \dots, p_r$ are uniquely determined by 1 and 2.

Corollary 1.38.1. *Every T-admissible subspace of V has a complementary T-invariant subspace.*

Corollary 1.38.2. *The annihilator p_1 is the minimal polynomial of T.*

Proof. Choose $W_0 = \{\mathbf{0}\}$, hence V is the direct sum of T-cyclic subspaces. Since each p_k divides p_{k-1}, we see that p_1 annihilates every vector in V. Its minimality is guaranteed by the fact that it is the minimal polynomial of $Z(\boldsymbol{v}_1, T)$. □

Corollary 1.38.3. *Given any $T \in \mathscr{L}(V)$, there exists $\boldsymbol{v} \in V$ such that its T-annihilator is the minimal polynomial of T.*

Corollary 1.38.4. *Given, $T \in \mathscr{L}(V)$, T has a cyclic vector if and only if its minimal and characteristic polynomials are identical.*

Definition 1.21. Let $T \in \mathscr{L}(V)$, and let V be written as the direct sum of T-cyclic subspaces as described by the Cyclic Decomposition Theorem. Then, there is a basis of V in which T is represented in a block diagonal form, with each block being a companion matrix, with the sizes of the blocks being weakly decreasing. This matrix is called the rational form of T.

Theorem 1.39. *Each matrix is similar to exactly one matrix in the rational form.*

Proof. This is guaranteed by the uniqueness of the polynomials $p_1, \dots, p_r$ generated by the Cyclic Decomposition Theorem. Note that if two blocks happen to be of equal size, the divisibility property forces $p_i = p_j$ for the corresponding blocks, so these blocks are exactly equal. □

Theorem 1.40 (Generalized Cayley-Hamilton Theorem)**.** *Let $T \in \mathscr{L}(V)$, let p be its minimal polynomial, and let f be its characteristic polynomial. Then p divides f, p and f have the same prime factors except for multiplicities, and if the prime factorization of p is*

$$p = f_1^{r_1} \dots f_k^{r_k},$$

then the prime factorization of f is of the form

$$f = f_1^{d_1} \dots f_k^{d_k}$$

with $d_i = \dim \ker (f_i^{r_i}) / \deg f_i$.

1.11 Jordan form

Lemma 1.41. *The rational form of a nilpotent matrix contains only 1's and 0's on the lower off-diagonal. Each choice of $k_1 \geq k_2 \geq \cdots \geq k_r \geq 1$ with $k_1 + \cdots + k_r = n$, i.e. each partition of n completely determines a similarity class of nilpotent $n \times n$ matrices.*

Remark. Note that $r = \dim \ker N$.

Definition 1.22. Let $T \in \mathscr{L}(V)$ such that its minimal polynomial is a product of linear factors,

$$p = (x - c_1)^{r_1} \dots (x - x_k)^{r_k}.$$

The Primary Decomposition Theorem guarantees that by defining $W_i = \ker (T - c_i I)^{r_i}$, we have $V = W_1 \oplus \cdots \oplus W_k$. Furthermore, if T_i are the restrictions of T to W_i, the minimal polynomials for T_i are $(x - c_i)^{r_i}$, hence $T_i = N_i + c_i I$ for nilpotent operators N_i. In a cyclic basis, each T_i is the direct sum of matrices

$$\begin{bmatrix} c_i & 0 & \dots & 0 & 0 \\ 1 & c_i & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & c_i & 0 \\ 0 & 0 & \dots & 1 & c_i \end{bmatrix},$$

descending in size. These are called elementary Jordan matrices with characteristic value c_i. Since T is a direct sum of each W_i, the matrix representation of T in a appropriate basis is in a block diagonal form with eigenvalues along the diagonal, and 1's and 0's along the off-diagonal. This is called the Jordan form of T.

Theorem 1.42. *The Jordan form of a linear operator is unique, up to permutation of the blocks.*

2 Inner product spaces

2.1 Preliminaries

Definition 2.1. Let $\mathbb{F}$ either the field of real or complex numbers, and let V be a vector space over $\mathbb{F}$. An inner product on V is a function $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{F}$ satisfying the following conditions.

1. $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
2. $\langle \alpha \boldsymbol{v}, \boldsymbol{w} \rangle = \alpha \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
3. $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \overline{\langle \boldsymbol{w}, \boldsymbol{v} \rangle}$.
4. $\langle \boldsymbol{v}, \boldsymbol{v} \rangle > 0$ for all $\boldsymbol{v} \neq \boldsymbol{0}$.

Remark. An inner product is completely determined by its real part.

Definition 2.2. The standard inner product on the vector space $\mathbb{F}^n$ is defined as

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \sum_{i=1}^{n} v_i \overline{w_i}.$$

Definition 2.3. A norm is a function $\| \cdot \| \colon V \to \mathbb{R}$ if

1. $\|\boldsymbol{v}\| \geq 0$, and $\|\boldsymbol{v}\| = 0$ implies $\boldsymbol{v} = \boldsymbol{0}$.
2. $\|\alpha \boldsymbol{v}\| = |\alpha| \|\boldsymbol{v}\|$.
3. $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$.

Remark. Any inner product induces a norm, via $\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$.

Lemma 2.1 (Polarization identity)**.**

$$4\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \sum_{k=1}^{4} i^k \|\boldsymbol{v} + i^k \boldsymbol{w}\|^2.$$

Lemma 2.2. *A norm arises from an inner product if and only if it satisfies the parallelogram identity,*

$$\|\boldsymbol{v} + \boldsymbol{w}\|^2 + \|\boldsymbol{v} - \boldsymbol{w}\|^2 = 2(\|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2).$$

Lemma 2.3. *Let V be finite dimensional, with an ordered basis $\beta = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$. Then for any $\boldsymbol{u}, \boldsymbol{w} \in V$, we have*

$$\langle \boldsymbol{u}, \boldsymbol{w} \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} u_i \overline{w_j} \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle.$$

By setting $a_{ij} = \langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle$ and letting $A = [a_{ij}]$, we can write

$$\langle \boldsymbol{u}, \boldsymbol{w} \rangle = \boldsymbol{w}^* A \boldsymbol{u}.$$

Note that the $\boldsymbol{u}, \boldsymbol{w}$ on the right hand side denote the coordinate column vectors in the basis β. We see that A is a Hermitian matrix, satisfying $A^* = A$. Furthermore, A is invertible because $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = \boldsymbol{u}^* A \boldsymbol{u} > 0$ for all $\boldsymbol{u} \neq \boldsymbol{0}$. Conversely, any such matrix defined an inner product.

Definition 2.4. A positive linear operator satisfies $\langle T\boldsymbol{v}, \boldsymbol{v} \rangle > 0$ for all $\boldsymbol{v} \in V$.

Remark. We will see that this conditions on T implies that it is Hermitian.

Lemma 2.4. *If $\langle T\boldsymbol{v}, \boldsymbol{v} \rangle = 0$ for all $\boldsymbol{v} \in V$ where V is a complex inner product space, then $T = 0$.*

Proof. Expand $\langle T(\boldsymbol{v}+\boldsymbol{w}), \boldsymbol{v}+\boldsymbol{w} \rangle = 0$ and $\langle T(\boldsymbol{v}+i\boldsymbol{w}), \boldsymbol{v}+i\boldsymbol{w} \rangle = 0$ to conclude that $\langle T\boldsymbol{v}, \boldsymbol{w} \rangle = 0$ for all $\boldsymbol{v}, \boldsymbol{w} \in V$. Setting $\boldsymbol{w} = T\boldsymbol{v}$ immediately gives the result. □

2.2 Orthogonality

Definition 2.5. Two vectors $\boldsymbol{v}, \boldsymbol{w} \in V$ are orthogonal if $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$.

Definition 2.6. If $W \subset V$, then $W^\perp$ is the set of vectors which are orthogonal to all vectors in W.

Lemma 2.5. *Let $W \subset V$ be a subspace. Then, $V = W \oplus W^\perp$.*

Theorem 2.6. *An orthogonal set of non-zero vectors is linearly independent.*

Theorem 2.7 (Gram-Schmidt)**.** *Every finite-dimensional inner product space admits an orthonormal basis.*

2.3 Dual spaces and adjoints

Lemma 2.8. *Every member of the dual space V^* is of the form $\boldsymbol{v} \mapsto \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ for unique $\boldsymbol{w} \in V$. This gives a canonical identification between V and V^*.*

Definition 2.7. The adjoint of a linear operator T on a finite-dimensional inner product space V is the unique linear operator T^* which satisfies

$$\langle T\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, T^*\boldsymbol{w} \rangle$$

for all $\boldsymbol{v}, \boldsymbol{w} \in V$.

Remark. The matrix of T^* is the conjugate transpose of the matrix of T, given an orthonormal basis.

Remark. An operator such that $T = T^*$ is called self-adjoint, or Hermitian.

Example. Given any operator T, we can write

$$T = \frac{1}{2}(T + T^*) + i \cdot \frac{1}{2i}(T - T^*)$$

where both $(T + T^2)/2$ and $(T - T^*)/2i$ are Hermitian.

Lemma 2.9. *Let $T \in \mathscr{L}(V)$. Then, the subspace W is invariant under T if and only if $W^\perp$ is invariant under T^*.*

Definition 2.8. A reducing subspace of $T \in \mathscr{L}$ is invariant under both T and T^*.

Remark. If W is a reducing subspace of T, then W and $W^\perp$ are both invariant under T,

Lemma 2.10. *Let $V = R \oplus N$; then, there is a unique projection E with range R and kernel N. When $N = R^\perp$, we call E the orthogonal projection into R. Here, we have $E = E^*$.*

Proof. We prove the latter, i.e for all $\boldsymbol{v}, \boldsymbol{w} \in V$, we have

$$\langle E\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, E\boldsymbol{w} \rangle.$$

Note that we can expand

$$\boldsymbol{v} = E\boldsymbol{v} + (I - E)\boldsymbol{v}, \qquad \boldsymbol{w} = E\boldsymbol{w} + (I - E)\boldsymbol{w},$$

hence

$$\begin{aligned} \langle E\boldsymbol{v}, \boldsymbol{w} \rangle &= \langle E\boldsymbol{v}, E\boldsymbol{w} \rangle + \langle E\boldsymbol{v}, (I - E)\boldsymbol{w} \rangle = \langle E\boldsymbol{v}, E\boldsymbol{w} \rangle, \\ \langle \boldsymbol{v}, E\boldsymbol{w} \rangle &= \langle E\boldsymbol{v}, E\boldsymbol{w} \rangle + \langle (I - E)\boldsymbol{v}, E\boldsymbol{w} \rangle = \langle E\boldsymbol{v}, E\boldsymbol{w} \rangle. \end{aligned}$$

□

Lemma 2.11. *A linear operator T is Hermitian if and only if $\langle T\boldsymbol{v}, \boldsymbol{v}\rangle$ is real for all $\boldsymbol{v} \in V$.*

Corollary 2.11.1. *Every eigenvalue of a Hermitian operator is real.*

Theorem 2.12. *Hermitian operators are diagonalizable, with an orthonormal basis of eigenvectors.*

Definition 2.9. A normal operator is one which commutes with its adjoint.

Lemma 2.13. *A linear operator T is normal if and only if $\|T\boldsymbol{v}\| = \|T^*\boldsymbol{v}\|$ for all $\boldsymbol{v} \in V$.*

Corollary 2.13.1. *Given two distinct eigenvalues of a normal operator, the corresponding eigenspaces are orthogonal.*

Theorem 2.14. *Normal operators are diagonalizable, with an orthonormal basis of eigenvectors. Conversely, any diagonalizable operator with an orthonormal basis is normal.*

2.4 Inner product space isomorphisms

Definition 2.10. A linear map $T: V \to W$ is an inner product space isomorphism if it is bijective and preserves the inner products of V and W.

Theorem 2.15. *Let $T: V \to W$, where V and W are finite dimensional inner product spaces of the same dimension. The following are equivalent.*

1. *T preserves the inner product.*
2. *T is an inner product space isomorphism.*
3. *T maps every orthonormal basis of V to an orthonormal basis of W.*
4. *T maps some orthonormal basis of V to an orthonormal basis of W.*

Proof. The implications $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4$ are trivial. To show $4 \Rightarrow 1$, suppose that T maps the orthonormal basis $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ to $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$. Then given any $\boldsymbol{v} = a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n$, $\boldsymbol{u} = b_1\boldsymbol{v}_1 + \cdots + b_n\boldsymbol{v}_n$, we can calculate

$$\langle \boldsymbol{v}, \boldsymbol{u}\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\overline{b_j}\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \sum_{i=1}^{n} a_i\overline{b_i},$$

$$\langle T\boldsymbol{v}, T\boldsymbol{u}\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\overline{b_j}\langle T\boldsymbol{v}_i, T\boldsymbol{v}_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\overline{b_j}\langle \boldsymbol{w}_i, \boldsymbol{w}_j\rangle = \sum_{i=1}^{n} a_i\overline{b_i}. \qquad \square$$

Corollary 2.15.1. *Two finite dimensional inner product spaces over the same field are isomorphic if and only if they have the same dimension.*

Lemma 2.16. *A linear isomorphism $T\colon V \to W$ is inner product preserving if and only if it is norm preserving.*

Definition 2.11. A unitary operator on an inner product space is an isomorphism of the space to itself.

Remark. The products and inverses of unitary operators are also unitary. Thus, the unitary operators on an inner product space form a group.

Theorem 2.17. *A linear operator U is unitary if and only if its adjoint U^* exists and $U^*U = UU^* = I$.*

Proof. We have $U^* = U^{-1}$, since

$$\langle U\boldsymbol{v}, \boldsymbol{w}\rangle = \langle U\boldsymbol{v}, UU^{-1}\boldsymbol{w}\rangle = \langle \boldsymbol{v}, U^{-1}\boldsymbol{w}\rangle.$$

Conversely if we know that $U^*U = UU^* = I$, then we immediately get $U^{-1} = U^*$. To show that it is inner product preserving, write

$$\langle U\boldsymbol{v}, U\boldsymbol{w}\rangle = \langle \boldsymbol{v}, U^*U\boldsymbol{w}\rangle = \langle \boldsymbol{v}, \boldsymbol{w}\rangle. \qquad \square$$

Theorem 2.18. *A linear operator is unitary if and only if its matrix representation in some orthonormal basis is unitary, i.e. satisfies $A^*A = I$.*

Definition 2.12. A square matrix A is called orthogonal if it satisfies $A^\top A = I$.

Remark. A unitary matrix is orthogonal if and only if it is real.

Theorem 2.19. *For every complex, invertible $n \times n$ matrix B, there exists a unique lower triangular M with positive real entries on the main diagonal such that MB is unitary.*

Proof. The rows of B are linearly independent, and hence form a basis of $\mathbb{C}^n$. Perform Gram-Schmidt orthonormalization on B, and simply let M be the matrix of this transformation. $\square$

Definition 2.13. Two complex matrices are said to be unitarily equivalent if they are conjugate via a unitary matrix.

Remark. The analogous definition for orthogonally equivalent matrices applies for real matrices.

Theorem 2.20. *Let $T \in \mathscr{L}(V)$, and let β be an orthonormal basis of V. Suppose that $A = [T]_\beta$ is upper triangular. Then, T is normal if and only if A is diagonal.*

Proof. It is clear that if A is diagonal, then T is normal since A and A^* commute. Now, suppose that T is normal. Enumerate $\beta = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$. Since $T\boldsymbol{v}_1 = a_{11}\boldsymbol{v}_1$, we have $T^*\boldsymbol{v}_1 = \overline{a}_{11}\boldsymbol{v}_1$. We can also use $A^* = [T^*]_\beta$ to compute

$$T^*\boldsymbol{v}_1 = \overline{a}_{11}\boldsymbol{v}_1 + \overline{a}_{12}\boldsymbol{v}_2 + \cdots + \overline{a}_{1n}\boldsymbol{v}_n,$$

hence we must have all $a_{12}, \ldots, a_{1n} = 0$. Now, $T\boldsymbol{v}_2 = a_{22}\boldsymbol{v}_2$, hence $T^*\boldsymbol{v}_2 = \overline{a}_{22}\boldsymbol{v}_2$: repeat the same procedure as above to conclude that all off-diagonal entries of A must vanish, completing the proof. $\square$

Theorem 2.21. *Let $T \in \mathscr{L}(V)$ where V is a finite-dimensional complex inner product space. Then, there exists an orthonormal basis of V such that $[T]_\beta$ is upper triangular.*

Corollary 2.21.1. *Every complex $n \times n$ matrix is unitarily similar to an upper triangular matrix.*

Corollary 2.21.2. *Let $T \in algL(V)$ be a normal operator on a finite-dimensional complex inner product space. Then, there exists an orthonormal basis of V in which the matrix representation of T is diagonal.*

Theorem 2.22 (Spectral theorem). *Let T be a normal operator on a finite-dimensional complex inner product space, and let $c_1, \ldots, c_k$ be distinct eigenvalues. Let $W_i = \ker(T - c_i I)$, and E_i be the orthogonal projections of V into W_i. Then, V is the orthogonal direct sum of $W_1, \ldots, W_k$ and $T = c_1E_1 + \cdots + c_kE_k$.*

Definition 2.14. Let T be a normal operator on a finite-dimensional complex inner product space, and let

$$T = c_1E_1 + \cdots + c_kE_k$$

be its spectral resolution. Suppose that f is a function whose domain contains the spectrum of T. Then, the linear operator $f(T)$ is defined by

$$f(T) = f(c_1)E_1 + \cdots + f(c_k)E_k.$$

Definition 2.15. A linear operator $T \in \mathscr{L}(V)$ is called non-negative if $\langle T\boldsymbol{v}, \boldsymbol{v}\rangle \geq 0$ for all $\boldsymbol{v} \in V$.

Lemma 2.23. *A non-negative operator is self-adjoint, and its spectrum is non-negative.*

Example. For any operator T, the operators T^*T and TT^* are non-negative. To see this,

$$\langle T^*T\boldsymbol{v}, \boldsymbol{v}\rangle = \langle T\boldsymbol{v}, T\boldsymbol{v}\rangle \geq 0.$$

Lemma 2.24. *An linear operator T is positive if and only if we can write $T = S^*S$ for some linear operator S.*

Theorem 2.25. *Let T be a normal operator on a finite-dimensional complex inner product space. Then, T is self-adjoint, non-negative, or unitary according to whether the spectrum of T is contained in $\mathbb{R}$, $[0, \infty)$, or the unit circle respectively.*

Theorem 2.26. *Let T be a non-negative operator on a finite-dimensional complex inner product space. Then, T has a unique square root, i.e. there is precisely one non-negative operator N such that $N^2 = T$.*

Theorem 2.27. *For any $T \in \mathscr{L}(V)$, there exists a unitary operator U and a non-negative operator N such that $T = UN$. Furthermore, N is unique; if T is invertible, then U is also unique. This is called a polar decomposition of T.*

Remark. The operators U and N commute if and only if T is normal.