

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
1.1	Preliminaries	1
1.2	Ideals in a ring	1
1.3	Eigenvalues and eigenvectors	2
1.4	Annihilating polynomials	3
1.5	Invariant subspaces	5
1.6	Triangulability and diagonalizability	7
1.7	Simultaneous triangulation and diagonalization	8
1.8	Direct sum decompositions	9
1.9	Projections maps	10
1.10	Cyclic subspaces and the Rational form	15
1.11	Jordan form	18

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 $T: V \rightarrow V$. We also assume that V has the inner product $\langle \cdot, \cdot \rangle$.

Theorem 1.1. *Let $\mathcal{L}(V)$ be the set of all linear operators on the vector space V . Then, $\mathcal{L}(V)$ 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 $T \in \mathcal{L}(V)$ and $c \in \mathbb{F}$. We say that c is an eigenvalue or characteristic value of T if $T\mathbf{v} = c\mathbf{v}$ for some non-zero $\mathbf{v} \in V$. The vector \mathbf{v} is called an eigenvector of T .

Theorem 1.3. Let $T \in \mathcal{L}(V)$ 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 $S, T \in \mathcal{L}(V)$ are similar if there exists an invertible operator $X \in \mathcal{L}(V)$ such that $S = X^{-1}TX$.

Remark. Similarity is an equivalence relation on $\mathcal{L}(V)$, 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 $T \in \mathcal{L}(V)$ 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 $T \in \mathcal{L}(V)$ where V is finite dimensional, let c_1, \dots, 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

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

where each $d_i = \dim W_i$.

3. $\dim V = \dim W_1 + \dots + \dim W_k$.

1.4 Annihilating polynomials

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

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

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

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 $\mathbf{v} \in V$ such that $\mathbf{w} = q(T)\mathbf{v} \neq \mathbf{0}$. Thus, $p(T)\mathbf{v} = \mathbf{0}$ gives

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

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 $\mathbf{v} \in V$ such that $T\mathbf{v} = c\mathbf{v}$. This gives $p(T)\mathbf{v} = p(c)\mathbf{v}$, but $p(T) = 0$ identically, forcing $p(c) = 0$. \square

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

Proof. Set $S = \text{adj}(T - xI)$. 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

$$S_{n-1} = -I, \quad TS_0 = a_0I, \quad TS_k - S_{k-1} = a_kI \text{ for } 1 \leq k \leq n-1.$$

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*

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

where c_1, \dots, 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 \mathbf{v}_i , the operator $T - c_i I$ kills it where c_i is the corresponding eigenvalue. Thus, $p(T)\mathbf{v}_i$ vanishes for every basis vector \mathbf{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 $T \in \mathcal{L}(V)$ 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 $T_W \in \mathcal{L}(W)$ as the restriction of T to W in the natural way, by setting $T_W(\mathbf{w}) = T(\mathbf{w})$ for all $\mathbf{w} \in W$.

Lemma 1.11. *If W is an invariant subspace under $T \in \mathcal{L}(V)$, 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 $\beta_W = \{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ be an ordered basis of W , and extend it to an ordered basis $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of V . Thus, the matrix $[T]_\beta$ has coefficients a_{ij} such that

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

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

Lemma 1.12. *If W is an invariant subspace under $T \in \mathcal{L}(V)$, 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, $T^k\mathbf{v} = \mathbf{0}$ implies $T_W^k\mathbf{v} = \mathbf{0}$, hence any polynomial which annihilates T also annihilates T_W . This gives the second result. \square

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

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

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

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

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

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

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

Lemma 1.14. *Let $T \in \mathcal{L}(V)$ 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 $\mathbf{v} \in V$ such that $\mathbf{v} \notin W$, and $(T - cI)\mathbf{v} \in W$ for some eigenvalue c .

Proof. What we must show is that the T -conductor of \mathbf{v} into W is a linear polynomial. Choose arbitrary $\mathbf{w} \in V \setminus W$, and let g be the T -conductor of \mathbf{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 $\mathbf{v} = h(T)\mathbf{w} \notin W$. Finally, note that

$$(T - c_i I)\mathbf{v} = (T - c_i I)h(T)\mathbf{w} = g(T)\mathbf{w} \in W. \quad \square$$

1.6 Triangulability and diagonalizability

Theorem 1.15. *Let $T \in \mathcal{L}(V)$ 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 $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 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

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

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

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

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

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. \square

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

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

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 $\mathbf{v} \notin W$ and an eigenvalue c_j such that $\mathbf{w} = (T - c_j I)\mathbf{v} \in W$. Now, \mathbf{w} can be written as the sum of eigenvectors

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

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

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

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,

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

Since

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

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

1.7 Simultaneous triangulation and diagonalization

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

An analogous definition holds for simultaneous diagonalizability.

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

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

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

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

Proof. We observe that we can assume that \mathcal{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 $\mathcal{L}(V)$.

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

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

Note that V_1 is a subspace which properly contains W . Furthermore, V_1 is invariant under \mathcal{F} – this uses the fact that the operators from \mathcal{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 $\mathbf{v}_2 \in V_1, \mathbf{v}_2 \notin W$ such that $(U_2 - c_2 I)\mathbf{v}_2 \in W$. Note that $(T_i - c_i I)\mathbf{v}_2 \in W$ for $i = 1, 2$. Construct V_2 as before, and repeat this process until we have exhausted all linear operators in \mathcal{F} . The final vector \mathbf{v}_j satisfies the desired properties. \square

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

Proof. The proof is identical to that of Theorem 1.15. \square

Theorem 1.20. Let \mathcal{F} be a commuting family of diagonalizable linear operators on V . There exists an ordered basis of V which simultaneously diagonalizes \mathcal{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 $T \in \mathcal{F}$ 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 \mathcal{F}_i be the family of operators from \mathcal{F} , restricted to the invariant subspace W_i . Note that each operator in \mathcal{F}_i is diagonalizable. Furthermore, $\dim W_i < \dim V$, so the induction hypothesis says that \mathcal{F}_i is simultaneously diagonalizable; let β_i be the corresponding basis. Each vector in β_i is an eigenvector for every operator in \mathcal{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. \square

1.8 Direct sum decompositions

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

$$\mathbf{w}_1 + \dots + \mathbf{w}_k = \mathbf{0}$$

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

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

$$\mathbf{w} = \mathbf{w}_1 + \dots + \mathbf{w}_k$$

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

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

$$W_1 \oplus \dots \oplus W_k.$$

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

1. W_1, \dots, W_k are independent.

2. For each $2 \leq j \leq k$,

$$W_j \cap (W_1 + \dots + W_{j-1}) = \{\mathbf{0}\}.$$

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 $R = \text{im } E$, $N = \ker E$.*

1. A vector $\mathbf{v} \in R$ if and only if $E\mathbf{v} = \mathbf{v}$.

2. Any vector $\mathbf{v} \in V$ has the unique representation $\mathbf{v} = E\mathbf{v} + (\mathbf{v} - E\mathbf{v})$, with $E\mathbf{v} \in R$ and $\mathbf{v} - E\mathbf{v} \in N$.

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 $R = \text{im } E$ and $N = \ker E$. Namely, send $\mathbf{v} \mapsto \mathbf{v}_R$ where $\mathbf{v} = \mathbf{v}_R + \mathbf{v}_N$ is the unique decomposition of \mathbf{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, $\text{trace } E = \text{rank } E$. \square

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

Remark. Observe that

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

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

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

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

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

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

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

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

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

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

$$E_j \mathbf{v} = \sum_{i=1}^k E_j E_i \mathbf{v}_i = \sum_{i=1}^k E_j E_i \mathbf{w}_i = E_j^2 \mathbf{w}_j = E_j \mathbf{w}_j = \mathbf{v}_j. \quad \square$$

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

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

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

Lemma 1.27. Let $V = W_1 \oplus \cdots \oplus W_k$, let β_i be ordered bases of W_i , and let β be the basis formed by combining all these vectors. Let $T \in \mathcal{L}(V)$ 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 $T \in \mathcal{L}(V)$. 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 $W_i = \text{im } E_i$ is invariant under T . Let $\mathbf{v} \in W_i$, hence $\mathbf{v} = E_i\mathbf{v}$ and

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

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

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

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

Theorem 1.29. Let $T \in \mathcal{L}(V)$ where V is a finite-dimensional vector space. If T is diagonalizable and c_1, \dots, c_k are the distinct eigenvalues of T , then there are non-zero linear operators E_1, \dots, 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. $\text{im } E_i = \ker(T - c_iI)$.

Conversely, if there exist k distinct scalars c_1, \dots, c_k and k non-zero linear operators which satisfy properties 1, 2, 3, then T is diagonalizable, c_1, \dots, 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, \dots, c_k . Let $W_i = \ker(T - c_iI)$, and note that $V = W_1 \oplus \cdots \oplus W_k$. Let E_1, \dots, 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 $\mathbf{v} \in V$ and write $\mathbf{v} = E_1\mathbf{v} + \cdots + E_k\mathbf{v}$. Then, note that $E_i\mathbf{v}$ are eigenvectors, hence

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

Conversely, let $T \in \mathcal{L}(V)$ and suppose that c_1, \dots, c_k and non-zero E_1, \dots, 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 $\text{im } E_i \neq \{0\}$ is an eigenspace of T corresponding to the eigenvalue c_i , i.e. $\text{im } E_i \subseteq \ker(T - c_iI)$. 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 $\mathbf{v} \in V$ such that $(T - cI)\mathbf{v} = 0$. Then, some $E_i\mathbf{v} \neq 0$ (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 $(c_i - c)E_i\mathbf{v} = 0$, forcing $c = c_i$. Finally, $I = E_1 + \cdots + E_k$ says that V is the direct sum of the $\text{im } E_i$, which are contained within the eigenspaces of T . This means that T is diagonalizable.

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

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

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

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

so $\mathbf{v} \in \text{im } E_i$. This proves that $\text{im } E_i = \ker(T - c_iI)$. \square

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, \dots, p_k are the Lagrange polynomials corresponding to the points c_1, \dots, 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 $T \in \mathcal{L}(V)$ where V is finite-dimensional, and let p be the minimal polynomial of T , where*

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

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 $\text{im } E_i = W_i$. To see this, first let $\mathbf{v} \in \text{im } E_i$, whence $\mathbf{v} = E_i \mathbf{v}$ so

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

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

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

hence $\mathbf{v} \in \text{im } E_i$. This proves 1.

It is clear that W_i is invariant under T . Pick arbitrary $\mathbf{v} \in W_i$, whence $\mathbf{v} = E_i \mathbf{v}$ so $T \mathbf{v} = T E_i \mathbf{v} = E_i T \mathbf{v} \in W_i$. 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. \square

Corollary 1.32.1. *Let E_1, \dots, 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 operator which commutes with T .*

Theorem 1.33. *Let $T \in \mathcal{L}(V)$ 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

$$N = (T - c_1 I) E_1 + \cdots + (T - c_k I) E_k.$$

It can be shown that

$$N^r = (T - c_1 I)^r E_1 + \cdots + (T - c_k I)^r E_k,$$

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. \square

1.10 Cyclic subspaces and the Rational form

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

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

Definition 1.18. Let $T \in \mathcal{L}(V)$, and $\mathbf{v} \in V$. We define the T -cyclic subspace generated by \mathbf{v} as

$$Z(\mathbf{v}, T) = \{g(T)\mathbf{v} : g \in \mathbb{F}[x]\}.$$

If $V = Z(\mathbf{v}, T)$, then \mathbf{v} is called a cyclic vector for T .

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

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

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

Proof. First note that

$$\mathbf{0} = p_{\mathbf{v}}(T)\mathbf{v} = a_k T^k \mathbf{v} + a_{k-1} T^{k-1} \mathbf{v} + \dots + a_0 \mathbf{v}.$$

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

We now show that $\mathbf{v}, \dots, T^{k-1} \mathbf{v}$ are linearly independent. If not, then

$$a_0 \mathbf{v} + \dots + a_{k-1} T^{k-1} \mathbf{v} = \mathbf{0}$$

for at least one $a_i \neq 0$. This contradicts the minimality of the degree of the T -annihilator of \mathbf{v} . Thus, we have properties 1, 2.

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

Definition 1.19. Let p be the following monic polynomial.

$$p(x) = a_0 + a_1x + \dots + a_{k-1}x^{k-1} + x^k.$$

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 $T \in \mathcal{L}(V)$ such that $v \in V$ is a cyclic vector of T . Then, the matrix representation of T in the basis $v, T v, \dots, T^{n-1}v$ 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 $T \in \mathcal{L}(V)$. 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 v , we have already shown that the desired basis is $\{v, T v, \dots, T^{n-1}v\}$.

Conversely, suppose that in the basis $\{v_0, v_1, \dots, v_{k-1}\}$, the matrix representation of T is the companion matrix of its characteristic polynomial. Then we immediately have $T v_0 = v_1$, $T v_1 = v_2$, \dots , $T^{n-2} v_{n-2} = v_{n-1}$. This immediately shows that v_0 is a cyclic vector of T . \square

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

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

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

1. $V = W_0 \oplus Z(\mathbf{v}_1, T) \oplus \dots \oplus Z(\mathbf{v}_r, T)$.
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 $Z(\mathbf{v}_1, T)$. \square

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

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

Definition 1.21. Let $T \in \mathcal{L}(V)$, 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. \square

Theorem 1.40 (Generalized Cayley-Hamilton Theorem). *Let $T \in \mathcal{L}(V)$, 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*

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

then the prime factorization of f is of the form

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

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.*