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 $\mathcal{L}(V)$ 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 $T \in \mathscr{L}(V)$ 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 $W_i = \ker(T - c_i I)$ 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 $\dim \mathscr{L}(V) = n^2$. □

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 $p(c) = 0$. ☐

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*

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

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 $\boldsymbol{v}_i$, 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 $i > r$, $j \leq r$. $\square$

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. $\square$

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 $g(T)\boldsymbol{v} = \boldsymbol{0}$.

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

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

□

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

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

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

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

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.*

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

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

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

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 $T\boldsymbol{v} \in \{\boldsymbol{v}, W\}$ 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 $c_1, \ldots, c_k$ 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. □