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

*Updated on September 8, 2021*

**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 $\boldsymbol{w} = q(T)\,\boldsymbol{v} \neq \boldsymbol{0}$. Thus, $p(T)\,\boldsymbol{v} = \boldsymbol{0}$ gives
$$(T - cI)\, q(T)\,\boldsymbol{v} = \boldsymbol{0}, \qquad T\boldsymbol{w} = c\boldsymbol{w},$$
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 $p(c) = 0$. $\qquad\square$

**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
$$
\begin{aligned}
f(x)I &= (T - xI)S \\
&= (T - xI)\sum_{k=0}^{n-1} x^k S_k \\
&= -x^k S_{k-1} + \sum_{k=1}^{n-1} x^k (T S_k - S_{k-1}) + T S_0.
\end{aligned}
$$
Compare coefficients with
$$f(x)I = x^n I + a_{n-1}x^{n-1} + \cdots + a_0 I$$

*Updated on September 8, 2021*

to get

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

Thus,

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

**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 $\boldsymbol{v}_i$ $\qquad \square$

*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 $\beta_W = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ be an ordered basis of $W$, and extend it to an ordered basis $\beta = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ 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 $\beta$ 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 $\beta_W$ 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 $W = \{\boldsymbol{0}\}$, $S_T(\boldsymbol{v}, \{\boldsymbol{0}\})$ 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 $S_T(\boldsymbol{v}, \{\boldsymbol{0}\})$ 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 $T \in \mathcal{L}(V)$ 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 $\boldsymbol{v} \in V$ 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

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

## 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 $\beta = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ 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 $W = \{\boldsymbol{0}\}$ 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. $\qquad \square$

**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 $T \in \mathscr{L}(V)$ 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

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

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

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

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.  $\square$

## 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.  $\square$

*Updated on September 8, 2021*

**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. $\square$

**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. $\square$

**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 $\beta_i$ be the corresponding basis. Each vector in $\beta_i$ is an eigenvector for every operator in $\mathscr{F}_i$. Let $\beta$ consist of the such vectors from all $\beta_i$ 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 $W_1, \ldots, W_k$ be subspaces of $V$. We say that these $W_i$ are independent if
$$\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k = \boldsymbol{0}$$
where $\boldsymbol{w}_i \in W_i$ implies that each $\boldsymbol{w}_i = \boldsymbol{0}$.

*Updated on September 8, 2021*

**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 $2 \leq j \leq k$,*
$$W_j \cap (W_1 + \cdots + W_{j-1}) = \{\boldsymbol{0}\}.$$

3. *If $\beta_i$ are bases of $W_i$, then the set $\beta$ 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, 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 $E_i E_j = 0$ 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. *$E_i E_j = 0$ 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,

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

**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 $T_1, \ldots, T_k$.

*Updated on September 8, 2021*

**Lemma 1.27.** *Let $V = W_1 \oplus \cdots \oplus W_k$, let $\beta_i$ be ordered basses of $W_i$, and let $\beta$ 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

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

*Updated on September 8, 2021*

giving property 4. Also,

$$TE_i = (c_1 E_1 + \cdots + c_k E_k)E_i = c_i E_i^2 = c_i E_i,$$

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

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

Pick non-zero $\mathbf{v} \in V$ such that $(T - cI)\mathbf{v} = 0$. Then, some $E_i \mathbf{v} \neq \mathbf{0}$ (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 $(c_i - c)E_i \mathbf{v} = \mathbf{0}$, 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 $\operatorname{im} E_i = \ker(T - c_i I)$. Pick $\mathbf{v} \in \ker(T - c_i I)$, 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 $\operatorname{im} E_i$, each $(c_j - c_i)E_j \mathbf{v} = \mathbf{0}$, or $E_j \mathbf{v} = \mathbf{0}$ for $j \neq i$. Thus,

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

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

**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 $T = c_1 E_1 + \cdots + c_k E_k$ 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 $T \in \mathcal{L}(V)$ 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 $p_i^{r_i}$.*

*Updated on September 8, 2021*

*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

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

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

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 (ii).

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 (iii). □

**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

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

*Updated on September 8, 2021*

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. $\square$