LINEAR ALGEBRA, PART 4. JORDAN CANONICAL FORM THEOREM FOR COMPLEX LINEAR OPERATORS

1. Characteristic polynomial

Definition 1.1. Let V be a finite-dimensional vector space. For $A \in \mathrm{End}(V)$ we define its characteristic polynomial by

$$\chi_A(\lambda) := \det(\lambda \cdot \mathrm{Id} - A). \tag{1.1}$$

Exercise 1.2. Suppose that $A \in \mathrm{End}(\mathbb{F}^n)$ is given by the following matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ 0 & 0 & 1 & \dots & 0 & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \end{pmatrix}. \tag{1.2}$$

Find $\chi_A(\lambda)$.

Definition 1.3. Let $A \in \mathrm{End}(V)$. A subspace $U \subseteq V$ is called an invariant subspace of A, or said to be A-invariant if $\forall u \in U$ we have that $Au \in U$. Note that if U is A-invariant, then the restriction of the map A to the subspace U is an endomorphism of U.

Proposition 1.4. *Suppose U is an invariant subspace of $A \in \mathrm{End}(V)$, $(u_1, \dots, u_m)$ is a basis of U, and $\alpha = (u_1, \dots, u_m, w_1, \dots, w_k)$ is a basis of V. Then the matrix of A in the basis α has the form*

$$\begin{pmatrix} B & C \\ 0 & D \end{pmatrix}, \quad \text{for some } B \in \mathrm{Mat}(m \times m, \mathbb{F}), C \in \mathrm{Mat}(m \times k, \mathbb{F}), D \in \mathrm{Mat}(k \times k, \mathbb{F}). \tag{1.3}$$

Exercise 1.5. Show that

$$\det \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} = \det(B) \cdot \det(D), \tag{1.4}$$

where B and D are square matrices.

From Proposition 1.4 we may deduce the following result about characteristic polynomials.

Proposition 1.6. *Suppose U is invariant with respect to $A \in \mathrm{End}(V)$. Denote by B the restriction of A to the subspace U. Then the characteristic polynomial χ_B divides the characteristic polynomial χ_A.*

Date: August 30, 2025.

2. Cayley-Hamilton Theorem

In this section, we begin by considering polynomials evaluated at linear operators.

Definition 2.1. For a polynomial $p(x) = \sum_{j=0}^{k} a_j x^j \in \mathbb{F}[x]$ and a linear operator $B \in \mathrm{End}(V)$ we define another linear operator $p(B) \in \mathrm{End}(V)$ by the following equality:

$$p(B) := a_0 \cdot \mathrm{Id} + \sum_{j=1}^{k} a_j B^j. \tag{2.1}$$

It is also convenient to adopt the notation $B^0 := \mathrm{Id}$.

The following proposition will help us to prove the Cayley-Hamilton Theorem.

Proposition 2.2. *Suppose that $A \in \mathrm{End}(V)$, $z \in V, z \neq 0$, $\{z, Az, A^2z, \ldots, A^{k-1}z\}$ is linearly independent, and*

$$A^k z = \sum_{j=0}^{k-1} b_j A^j z. \tag{2.2}$$

Then a) $U = \langle z, Az, A^2z, \ldots, A^{k-1}z\rangle$ is invariant subspace of A.

b) Denote by B the restriction of A to the subspace U. Then

$$\chi_B(\lambda) = \lambda^k - \sum_{j=0}^{k-1} b_j \lambda^j. \tag{2.3}$$

c) $(\chi_B(B))(z) = 0$.

Proposition 2.3 (Cayley-Hamilton)**.** *Suppose that $A \in \mathrm{End}(V)$ then*

$$\chi_A(A) = 0. \tag{2.4}$$

3. Eigenvalues and Eigenvectors.

In this section, we will consider the case when an invariant subspace is one-dimensional.

Definition 3.1. Suppose that A is a linear operator in a vector space V. An element $v \in V$ is called an eigenvector of the endomorphism A if:

(1) $v \neq 0$;
(2) $\exists \lambda \in \mathbb{F}$ such that $Av = \lambda v$.

The number λ is called an eigenvalue of the linear operator A. The set of all eigenvalues of A is called the spectrum of A and is denoted by $\mathrm{Spec}(A)$.

Now, we will formulate a possible algorithm explaining how to find all eigenvalues of a given linear operator.

Proposition 3.2. *Suppose that $A \in \mathrm{End}(V)$. Then*

$$(\lambda \text{ is an eigenvalue of } A) \iff (\lambda \text{ is a root of the characteristic polynomial } \chi_A) \tag{3.1}$$

Definition 3.3. Suppose that A is a linear operator in a vector space V and λ is an eigenvalue of A. Put

$$V_\lambda := \{x \in V \mid Ax = \lambda x\}. \tag{3.2}$$

V_λ is called an eigenspace of A. We see that V_λ is a subspace of the vector space V. Its dimension $\dim V_\lambda$ is called the geometric multiplicity of λ.

Note that V_λ consists of zero vector and all eigenvectors with the eigenvalue λ.

If we know all roots of the characteristic polynomial of A, we can find all eigenvectors of the linear operator A by solving equations $(A-\lambda\mathrm{Id})(x)=0$ for every root λ. If $\dim V<\infty$ and we fix a basis of V, then for every fixed λ, the equation $(A-\lambda\mathrm{Id})(x)=0$ is a homogeneous system of linear equations, therefore,

$$\dim V_\lambda = \dim V - \mathrm{rk}(A-\lambda\mathrm{Id}). \tag{3.3}$$

Definition 3.4. Suppose that A is a linear operator in a finite-dimensional vector space V. The multiplicity of the factor $(x-\lambda)$ in the factorization of the characteristic polynomial χ_A into irreducible polynomials is called the algebraic multiplicity of λ.

Proposition 3.5. *Suppose that A is a linear operator in a finite-dimensional vector space V, λ is an eigenvalue of A. It holds that the geometric multiplicity of λ is not greater than its algebraic multiplicity.*

4. Diagonalization.

Proposition 4.1. *Suppose that $v_1,\dots,v_k$ are eigenvectors of a linear operator A with distinct eigenvalues, then the set $\{v_1,\dots,v_k\}$ is linearly independent.*

From Proposition 4.1, we may deduce a useful result about the corresponding eigenspaces.

Proposition 4.2. *Suppose that A is a linear operator in a finite-dimensional vector space V, $\lambda_1,\dots,\lambda_k$ are distinct eigenvalues of A, then the sum of the corresponding eigenspaces*

$$V_{\lambda_1}+\dots+V_{\lambda_k} \tag{4.1}$$

is a direct sum.

Definition 4.3. A square matrix

$$B=(b_i^j)_{1\le i,j\le n}\in \mathrm{Mat}(n\times n,\mathbb{F}) \tag{4.2}$$

is called diagonal if $b_i^j=0$ for all $i\neq j$:

$$\begin{pmatrix} a_1 & 0 & \dots & 0\\ 0 & a_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & a_n \end{pmatrix}. \tag{4.3}$$

A linear operator A in a finite-dimensional vector space V is called diagonalizable if and only if there exists a basis of V in which the matrix of A is diagonal.

Exercise 4.4. Show that the linear operator defined by the matrix

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

is not diagonalizable.

Definition 4.5. We say that a polynomial $p\in\mathbb{F}[x]$ splits (into linear factors) if it can be represented as

$$p(x)=c\cdot\prod_{j=1}^{k}(x-a_j)^{m_j} \tag{4.4}$$

for some $a_1,\dots,a_k,c\in\mathbb{F}$, $m_1,\dots,m_k\in\mathbb{Z}_+$.

Now we will state an obvious property of diagonalizable linear operators.

Proposition 4.6. *If a linear operator A in a finite-dimensional vector space V is diagonalizable then its characteristic polynomial splits into linear factors.*

Using Proposition 4.1, we can obtain a sufficient condition for diagonalizability.

Proposition 4.7. *If the characteristic polynomial of a linear operator A in a finite-dimensional vector space V has $\dim V$ distinct roots then A is diagonalizable.*

Also, we may easily formulate a statement about triangulation.

Definition 4.8. A square matrix

$$B = (b_i^j)_{1\leq i,j\leq n} \in \mathrm{Mat}(n\times n, \mathbb{F}) \tag{4.5}$$

is called upper triangular if $b_i^j = 0$ for all $i < j$.

Proposition 4.9 (Schur's Theorem)**.** *Suppose that A is a linear operator in a finite-dimensional vector space V and its characteristic polynomial splits into linear factors, then there exists a basis of V in which the matrix of A is upper triangular.*

In the proof of Schur's Theorem, we may recall how we connected the characteristic polynomials of a linear operator and its restriction to an invariant subspace in the previous lecture.

5. Primary Decomposition Theorem

Definition 5.1. Suppose that A is a linear operator in a vector space V, λ is an eigenvalue of A. Put

$$V^\lambda := \{x \in V \mid \exists k \in \mathbb{Z}_+ \text{ such that } (A - \lambda\cdot \mathrm{Id})^k(x) = 0\}. \tag{5.1}$$

V^λ is called a generalized eigenspace of A. We immediately get that V^λ is a subspace of the vector space V.

Proposition 5.2. *Suppose that A is a linear operator in a vector space V, λ is an eigenvalue of A. We get the following properties:*
a) $V_\lambda \subseteq V^\lambda$;
b) The subspace V^λ is A-invariant.

Proposition 5.3. *Suppose that A is a linear operator in a finite-dimensional vector space V and $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of its characteristic polynomial. If for some vectors $x_1, \ldots, x_k$ such that $x_j \in V^{\lambda_j}$ we have an equality*

$$\sum_{j=1}^{k} x_j = 0 \tag{5.2}$$

then $x_j = 0$, $\forall 1 \leq j \leq k$.

The following two lemmas may help us prove Proposition 5.3.

Lemma 5.4. *Suppose $A \in \mathrm{End}(V)$ and λ is its eigenvalue, then $\forall x \in V^\lambda, x \neq 0$, $\exists k \geq 0$ such that $(A - \lambda\mathrm{Id})^k(x)$ is an eigenvector of the linear operator A with an eigenvalue λ.*

Lemma 5.5. *Let v be an eigenvector of a linear operator A with an eigenvalue λ. For every polynomial $p \in \mathbb{F}[x]$ the element v is an eigenvector of $p(A)$ with the eigenvalue $p(\lambda)$.*

Proposition 5.6 (Primary Decomposition Theorem)**.** *Suppose that A is a linear operator in a finite-dimensional vector space V and its characteristic polynomial splits into linear factors:*

$$\chi_A(x) = \prod_{j=1}^{k} (x - \lambda_j)^{m_j}, \tag{5.3}$$

where $\lambda_1, \ldots, \lambda_k$ are distinct, $m_j \geq 1$, $\forall 1 \leq j \leq k$, then we obtain the equality

$$\boxed{V = \bigoplus_{j=1}^{k} V^{\lambda_j};} \tag{5.4}$$

To prove the Primary Decomposition Theorem, one may consider the polynomials

$$q_j := \prod_{1\le i\le k,\, i\neq j} (x-\lambda_i)^{m_i} \tag{5.5}$$

and note that

$$\operatorname{im}(q_j(A)) \subseteq V^{\lambda_j}, \quad \forall 1 \le j \le k. \tag{5.6}$$

Bézout's Lemma for the $GCD(q_1,\dots,q_k)$ implies a part of the theorem statement.

Another part of this theorem follows from Proposition 5.3.

6. Other Useful Characteristics of a Linear Operator

Definition 6.1. Suppose that A is a linear operator in a finite-dimensional vector space V and

$$\chi_A(\lambda) = \lambda^n + \sum_{j=0}^{n-1} b_j \lambda^j. \tag{6.1}$$

The trace $\operatorname{tr}(A)$ of a linear operator A, by definition, is equal to $-b_{n-1}$.

Note that because the characteristic polynomial of a linear operator does not depend on the choice of the basis, neither does the trace.

Exercise 6.2. Let (a_i^j) be the matrix of a linear operator $A \in \operatorname{End}(V)$ in some basis. Find an expression for $\operatorname{tr}(A)$ in terms of the matrix entries.

Exercise 6.3. Show that

$$\operatorname{tr}(AB) = \operatorname{tr}(BA), \quad \forall A, B \in \operatorname{End}(V). \tag{6.2}$$

Definition 6.4. Suppose that A is a linear operator in a finite-dimensional vector space V. Consider the set

$$I(A) := \{p \in \mathbb{F}[x] \mid p \neq 0, p(A) = 0\}. \tag{6.3}$$

One can show that an element of minimal degree of $I(A)$ is determined uniquely up to multiplication by a non-zero scalar. As usual, we may require that its leading coefficient equals 1 to ensure uniqueness. This monic polynomial is called the minimal (annihilating) polynomial for the linear operator A.

Exercise 6.5. Show that the minimal polynomial for a linear operator is indeed well-defined.

7. The Existence of a Jordan Canonical Form

Definition 7.1. A linear operator B is called nilpotent if there exists $k \ge 0$ such that $B^k = 0$.

Definition 7.2. For an arbitrary positive integer k and $\lambda \in \mathbb{F}$ the following $k \times k$ matrix

$$J_k(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \tag{7.1}$$

is called a Jordan block, or a Jordan matrix, or a Jordan cell.

Exercise 7.3. Show that the matrix $J_k(\lambda) - \lambda \mathrm{Id}$ defines a nilpotent linear operator.

Exercise 7.4. Suppose that $J_k(0)$ is a matrix of $A \in \operatorname{End}(V)$ in the basis $e_1,\dots,e_k$. Find bases of $\ker(A)$ and $\operatorname{im}(A)$.

The purpose of this section is to show the following theorem.

Proposition 7.5 (Jordan Canonical Form Theorem(The Existence Part)). *Suppose that A is a linear operator in a finite-dimensional vector space V and the characteristic polynomial of A splits into linear factors. Then there exists a basis α of the vector space V in which the matrix of A has the following block-diagonal form (Jordan canonical form)*

$$\boxed{[A]_\alpha = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_m}(\lambda_m) \end{pmatrix}} \tag{7.2}$$

for some, not necessarily distinct, $k_1, k_2, \dots, k_m \in \mathbb{Z}_+, \lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{F}$. The basis α is called a Jordan basis.

Suppose that A is a linear operator in a finite-dimensional vector space V and the characteristic polynomial of A splits into linear factors. In the previous lecture, we have established that

$$V = \bigoplus_{j=1}^{k} V^{\lambda_j}, \tag{7.3}$$

where V^{λ_j} is a generalized eigenspace of A corresponding to λ_j. Also, we have shown that each V^{λ_j} is A-invariant. Denote by A_j the restriction of the linear operator A to the subspace V^{λ_j}. The subspace V^{λ_j} is finite-dimensional. Therefore, by its definition,

$$\exists k \in \mathbb{Z}_+ \text{ such that } (A_j - \lambda_j \cdot \mathrm{Id})^k = 0. \tag{7.4}$$

Consider $B := A_j - \lambda_j \mathrm{Id} \in \mathrm{End}(V^{\lambda_j})$. We see that B is a nilpotent linear operator. Hence, we have reduced the problem of finding a "nice" matrix form of A to the case of a nilpotent linear operator.

Proposition 7.6. *Suppose that B is a nilpotent linear operator in a finite-dimensional vector space V. Then there exists a basis β of the vector space V in which the matrix of B has the following block-diagonal form*

$$[B]_\beta = \begin{pmatrix} J_{k_1}(0) & 0 & \dots & 0 \\ 0 & J_{k_2}(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_s}(0) \end{pmatrix} \tag{7.5}$$

for some $k_1, k_2, \dots, k_s$.

Before proving this key proposition, we will state the following lemma.

Lemma 7.7. *Let B be a nilpotent linear operator in a finite-dimensional vector space V. We will consider a set of vectors from V that satisfies the following properties. First, we pick an arbitrary positive integer q and q arbitrary positive integers $l_1, \dots, l_q$. Now, suppose that a set*

$$S := \{v_{(i|j)} \in V \mid 1 \le j \le q,\ 1 \le i \le l_j\} \tag{7.6}$$

satisfies the following conditions

(1) $Bv_{(i|j)} = v_{(i-1|j)}, \quad \forall 1 \le j \le q,\ 2 \le i \le l_j;$
(2) $Bv_{(1|j)} = 0, \quad \forall 1 \le j \le q;$
(3) *The set of vectors*

$$\{v_{(1|j)} \mid 1 \le j \le q\} \tag{7.7}$$

is linearly independent.

Then the set S is also linearly independent.

Proof of Proposition 7.6. We know that $\mathrm{im}(B) < n$ because B is nilpotent. Also $\mathrm{im}(B)$ is invariant under B. Denote the restriction of B to $\mathrm{im}(B)$ by D. Applying the induction hypothesis to D we get that there exists a basis

$$M = \{v_{(1|1)}, v_{(2|1)}, \dots, v_{(l_1|1)}, v_{(1|2)}, \dots, v_{(l_2|2)}, \dots, v_{(1|q)}, \dots, v_{(l_q|q)}\} \tag{7.8}$$

of $\mathrm{im}(B)$ such that the matrix of D in that basis has the form

$$\begin{pmatrix} J_{l_1}(0) & 0 & \dots & 0 \\ 0 & J_{l_2}(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{l_q}(0) \end{pmatrix}. \tag{7.9}$$

We will construct a new set of vectors $\widetilde{M}$ by adding some vectors to the set M. As the first step, for every $1 \leq j \leq q$ we pick an arbitrary vector from the preimage $B^{-1}(v_{(l_j|j)}$ and denote this vector by

$$v_{(l_j+1|j)}. \tag{7.10}$$

Also, let us put

$$m_j := l_j + 1 \tag{7.11}$$

for $1 \leq j \leq q$. In the second step, we notice that the set of vectors

$$\{v_{(1|j)} \mid 1 \leq j \leq q\} \tag{7.12}$$

is linearly independent. Suppose we need to add r vectors to this set to make it a basis of $\ker(B)$. Let us denote those vectors by

$$v_{(1|q+1)}, v_{(1|q+2)}, \dots, v_{(1|q+r)}. \tag{7.13}$$

Also, let us put

$$m_j := 1 \tag{7.14}$$

for $q+1 \leq j \leq q+r$. We define our new set of vectors $\widetilde{M}$ as

$$\widetilde{M} := \{v_{(i|j)} \mid 1 \leq j \leq q+r, 1 \leq i \leq m_j\}. \tag{7.15}$$

Exercise 7.8. Show that $\widetilde{M}$ is linearly independent.

Now let us consider $\left\langle \widetilde{M} \right\rangle$.

Exercise 7.9. Show that

$$M \subset \widetilde{M} \text{ and } |\widetilde{M} \setminus M| = \dim(\ker(B)). \tag{7.16}$$

Exercise 7.10. Show that

$$\left\langle \widetilde{M} \right\rangle = V. \tag{7.17}$$

□

As we explained above, Proposition 7.5 follows from Proposition 7.6.

The field $\mathbb{F}$ is called algebraically closed if for every polynomial $p \in \mathbb{F}[x]$ there exists $a \in \mathbb{F}$ such that $p(a) = 0$. We have shown that for every linear operator in a finite-dimensional vector space over an algebraically closed field there exists a Jordan basis. In particular, it is true for every complex linear operator.

8. The Uniqueness of the Jordan Canonical Form

Exercise 8.1. Suppose that $J_k(0)$ is the matrix representing $A \in \mathrm{End}(V)$ in the basis $e_1, \ldots, e_k$. Find a basis of $\ker(A^q)$.

Proposition 8.2 (Jordan Canonical Form Theorem(The Uniqueness Part))**.** *Suppose that A is a linear operator in a finite-dimensional vector space V. If there exists a basis α of the vector space V in which the matrix of A has the following block-diagonal form*

$$[A]_\alpha = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \ldots & 0 \\ 0 & J_{k_2}(\lambda_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & J_{k_m}(\lambda_m) \end{pmatrix} \tag{8.1}$$

for some, not necessarily distinct, $k_1, k_2, \ldots, k_m \in \mathbb{Z}_+, \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F}$ then every other Jordan canonical form of A have the same multiset of Jordan blocks. In other words, the sequence of pairs (k_i, λ_i) is determined uniquely, up to a permutation of terms.

Proof. We will show how to determine this multiset from some characteristics of the linear operator A.

Note that we obtain the numbers λ_i as roots of the characteristic polynomial of A.

Suppose that the matrix $[A]_\alpha$ has the form (8.1). Denote the number of Jordan blocks that have the form $J_k(\lambda)$ by $m_k(\lambda)$ for every $k \geq 1, \lambda \in \mathbb{F}$.

Exercise 8.3. Show that for every $\lambda \in \mathbb{F}, q \geq 1$

$$\dim\left(\ker(A - \lambda\mathrm{Id})^q\right) = \sum_{j=1}^{q} j m_j(\lambda) + q \sum_{k \geq q+1} m_k(\lambda). \tag{8.2}$$

Exercise 8.4. Show that for every $\lambda \in \mathbb{F}, q \geq 1$

$$\dim\left(\ker(A - \lambda\mathrm{Id})^q\right) - \dim\left(\ker(A - \lambda\mathrm{Id})^{q-1}\right) = \sum_{k \geq q} m_k(\lambda). \tag{8.3}$$

Exercise 8.5. Show that for every $\lambda \in \mathbb{F}, q \geq 1$

$$2\dim\left(\ker(A - \lambda\mathrm{Id})^q\right) - \left(\dim\left(\ker(A - \lambda\mathrm{Id})^{q+1}\right) + \dim\left(\ker(A - \lambda\mathrm{Id})^{q-1}\right)\right) = m_q(\lambda). \tag{8.4}$$

Therefore, for every root λ of the characteristic polynomial of A the sequence $(m_q(\lambda))$ is determined uniquely. Hence, one JCF of a given linear operator can be always obtained from its other JCF using a permutation of blocks. □

Chapters from the textbooks relevant to the lecture:

- Friedberg, Insel, Spence, Sections 5.1, 5.2, 7.1-7.3;
- Axler, Sections 5A-5D, 8A-8D;
- Roman, pp. 185-197, 363-378, 393-395.