LECTURE 9. JORDAN CANONICAL FORM THEOREM

In this lecture, the notation $\mathbb{F}$ stands for a field.

1. The Existence Part

Definition 1.1. A linear operator B is called nilpotent if there exists $k \geq 0$ such that $B^k = 0$.

Definition 1.2. For an arbitrary positive integer k and $\lambda \in \mathbb{F}$ the following $k \times k$ matrix

$$J_k(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \tag{1.1}$$

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

Exercise 1.3. Show that the matrix $J_k(\lambda) - \lambda \mathrm{Id}$ defines a nilpotent linear operator.

Exercise 1.4. Suppose that $J_k(0)$ is a matrix of $A \in \mathrm{End}(V)$ with respect to the basis $e_1, \dots, e_k$. Find bases of $\ker(A)$ and $\mathrm{im}(A)$.

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

Proposition 1.5 (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. 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)*

$$[A]_\alpha = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_m}(\lambda_m) \end{pmatrix} \tag{1.2}$$

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.

Exercise 1.6. Is a Jordan basis uniquely determined?

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

$$V = \bigoplus_{j=1}^{k} V^{\lambda_j}, \tag{1.3}$$

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,

$$\exists k \in \mathbb{Z}_+ \text{ such that } (A_j - \lambda_j \cdot \mathrm{Id})^k = 0. \tag{1.4}$$

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 convenient matrix form of A to the case of a nilpotent linear operator.

Proposition 1.7. *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*

$$[B]_\beta = \begin{pmatrix} J_{k_1}(0) & 0 & \dots & 0 \\ 0 & J_{k_2}(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_s}(0) \end{pmatrix} \tag{1.5}$$

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

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

Lemma 1.8. *Let B be a nilpotent linear operator in a finite-dimensional vector space V. Let q and $l_1, \dots, l_q$ be arbitrary positive integers. Suppose that a set*

$$S := \{v_{(i|j)} \in V \mid 1 \le j \le q,\, 1 \le i \le l_j\} \tag{1.6}$$

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*

$$\{v_{(1|j)} \mid 1 \le j \le q\} \tag{1.7}$$

is linearly independent.

Then the set S is also linearly independent.

To prove Lemma 1.8 we may assume the contrary. Next, one can introduce a function of elements of the set S, say, the height h that is defined as

$$h(v_{(i|j)}) := i. \tag{1.8}$$

There exists at least one element of maximum height in the equality describing the linear dependence. Suppose this height equals t. We apply the linear operator B to this equality $(t-1)$ times.

Exercise 1.9. Finish the proof of Lemma 1.8.

Proof of Proposition 1.7. We know that $\dim\mathrm{im}(B) < \dim V$ 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

$$M = \{v_{(1|1)}, v_{(2|1)}, \dots, v_{(l_1|1)}, v_{(1|2)}, \dots, v_{(l_2|2)}, \dots, v_{(1|q)}, \dots, v_{(l_q|q)}\} \tag{1.9}$$

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

$$\begin{pmatrix} J_{l_1}(0) & 0 & \dots & 0 \\ 0 & J_{l_2}(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{l_q}(0) \end{pmatrix}. \tag{1.10}$$

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

$$v_{(l_j+1|j)}. \tag{1.11}$$

Also, let us define

$$m_j := l_j + 1 \tag{1.12}$$

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

$$\{v_{(1|j)} \mid 1 \leq j \leq q\} \tag{1.13}$$

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

$$v_{(1|q+1)}, v_{(1|q+2)}, \ldots, v_{(1|q+r)}. \tag{1.14}$$

Also, let us set

$$m_j := 1 \tag{1.15}$$

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

$$\widetilde{M} := \{v_{(i|j)} \mid 1 \leq j \leq q+r, 1 \leq i \leq m_j\}. \tag{1.16}$$

Exercise 1.10. Show that $\widetilde{M}$ is linearly independent.

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

Exercise 1.11. Show that

$$M \subset \widetilde{M} \text{ and } |\widetilde{M} \setminus M| = \dim(\ker(B)). \tag{1.17}$$

Exercise 1.12. Show that

$$\left\langle \widetilde{M} \right\rangle = V. \tag{1.18}$$

□

There is another way to explain this construction. Suppose that a basis of the subspace $\ker(B)$ consists of s vectors. Denote those vectors by

$$v_{(1|1)}, v_{(1|2)}, \ldots, v_{(1|s)}. \tag{1.19}$$

Next, we consider the preimage $B^{-1}(v_{(1|j)})$ for every $1 \leq j \leq s$. If this preimage is non-empty, we pick its arbitrary element and denote it by $v_{(2|j)}$. If the preimage $B^{-1}(v_{(1|j)})$ is an empty set, we do nothing.

In the third step, we consider vectors of the form $v_{(2|j)}$ and their preimages $B^{-1}(v_{(2|j)})$ for those j, for which the vectors $v_{(2|j)}$ exist.

We continue this process as long as there exists at least one non-empty preimage. Lemma 1.8 implies that the set of vectors we have obtained is linearly independent. But it is not trivial to show that the span equals the whole space V.

As we explained above, Proposition 1.5 follows from Proposition 1.7.

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.

2. The Uniqueness Part

Exercise 2.1. Suppose that $J_k(0)$ is a matrix of $A \in \mathrm{End}(V)$ with respect to the basis $e_1, \dots, e_k$. Find a basis of $\ker(A^q)$.

Proposition 2.2 (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*

$$[A]_\alpha = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{k_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{k_m}(\lambda_m) \end{pmatrix} \tag{2.1}$$

for some, not necessarily distinct, $k_1, k_2, \dots, k_m \in \mathbb{Z}_+, \lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{F}$ then every other Jordan canonical form of A has 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.

Assume that the matrix $[A]_\alpha$ has the form (2.1). 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 2.3. Show that for every $\lambda \in \mathbb{F}, q \geq 1$

$$\dim\left(\ker(A - \lambda\mathrm{Id})^q\right) = \sum_{j=1}^{q} j m_j(\lambda) + q \sum_{k \geq q+1} m_k(\lambda). \tag{2.2}$$

Exercise 2.4. Show that for every $\lambda \in \mathbb{F}, q \geq 1$

$$\dim\left(\ker(A - \lambda\mathrm{Id})^q\right) - \dim\left(\ker(A - \lambda\mathrm{Id})^{q-1}\right) = \sum_{k \geq q} m_k(\lambda). \tag{2.3}$$

Exercise 2.5. Show that for every $\lambda \in \mathbb{F}, q \geq 1$

$$2\dim\left(\ker(A - \lambda\mathrm{Id})^q\right) - \left(\dim\left(\ker(A - \lambda\mathrm{Id})^{q+1}\right) + \dim\left(\ker(A - \lambda\mathrm{Id})^{q-1}\right)\right) = m_q(\lambda). \tag{2.4}$$

Therefore, for every root λ of the characteristic polynomial of A the sequence $(m_q(\lambda))$ is determined uniquely. Hence, any two Jordan canonical forms of a given linear operator differ only by a permutation of the blocks. □

Sections from the textbooks relevant to this lecture:

- Friedberg, Insel, Spence, Sections 7.1-7.3;
- Axler, Ch.8;
- Roman, pp. 185-192.