A note on the Jordan canonical form

(work in progress)

Shuto Noda

January 14, 2024

1 Basic concepts

We begin by reviewing the concepts of the diagonalization of square matrices and the eigenvalue decomposition of a complex vector space.

Definition 1.1. A scalar $\lambda \in \mathbb{C}$ is an eigenvalue for a matrix A if there exists a nonzero column vector $v \in \mathbb{C}^n$ for which

$$Av = \lambda v$$

In this case, v is called an eigenvector for A associated with λ.

Let $\tau \colon V \to V$ be a linear transfomation of a finite-dimentional vector space V over $\mathbb{C}$. We denote $\mathcal{L}(V)$ the set of all linear transformations on V.

Given any $\lambda \in \mathbb{C}$ the eigenspace of τ with eigenvalue λ is

$$\mathcal{E}(\lambda) = \{\, v \in V \mid (\tau - \lambda \operatorname{id}_V)v = 0 \,\}$$

where $\operatorname{id}_V \colon V \to V$ is the identity transformation in V. The **generalized eigenspace** of τ with eigenvalue λ is defined by

$$\mathcal{E}_k(\lambda) = \left\{\, v \in V \mid (\tau - \lambda \operatorname{id}_V)^k v = 0 \text{ for some } k \,\right\}.$$

Proposition 1.2. The generalized eigenspace $\mathcal{E}_k(\lambda)$ of τ has the following properties.

(i) $\mathcal{E}_k(\lambda)$ is a subspace of V.

(ii) $\mathcal{E}(\lambda) \subset \mathcal{E}_k(\lambda)$.

In this sence, the generalized eigenspaces are generalization of the eigenspaces.

Proof. Let $u, v \in \mathcal{E}_k(\lambda)$ and $c, d \in \mathbb{C}$. There is $k, l \in \mathbb{N}$ such that $(\tau - \lambda \operatorname{id}_V)^k u = 0$, $(\tau - \lambda \operatorname{id}_V)^l v = 0$. Then, by letting $m = \max\{k, l\}$, we have

$$(\tau - \lambda \operatorname{id}_V)^m (cu + dv) = c(\tau - \lambda \operatorname{id}_V)^m u + d(\tau - \lambda \operatorname{id}_V)^m v = 0.$$

Moreover, for any $u \in \mathcal{E}(\lambda)$, we have

$$(\tau - \lambda \operatorname{id}_V)^k v = 0 \text{ for } k = 1.$$

Hence, we get $u \in \mathcal{E}_k(\lambda)$. □

$$\mathcal{E}(\lambda) = \mathcal{E}_1(\lambda) \subset \mathcal{E}_2(\lambda) \subset \cdots$$

We also express block diagonal matrices as follows.

$$\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_l \end{bmatrix} = A_1 \oplus A_2 \oplus \cdots \oplus A_l$$

Note that the zeros in the off-diagonal elements are omitted.

2 Minimal polynomials

Definition 2.1. Let $\tau \in \mathcal{L}(V)$ and let $\alpha_1, \ldots, \alpha_r$ be distinct eigenvalues of τ. We define the **charasteristic polynomial** of τ by

$$c_\tau(t) = \prod_{i=1}^{r} (t - \alpha_i)^{n_i}$$

where n_i is called the **algebraic multiplicities** of the eigenvalue α_i.

$$\text{For } A = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}, \text{ we have } \varphi_A(t) = (t-1)^2(t-2),\ \mu_A(t) = (t-1)(t-2).$$

The minimal polynomial and the charasteristic polynomial of a matrix may coincide;

$$\text{For } B = \begin{bmatrix} 1 & 3 & \\ & 1 & 3 \\ & & 1 \end{bmatrix}, \text{ we have } \varphi_B(t) = \mu_B(t) = (t-1)^3.$$

3 Jordan block and Jordan matrix

An $n \times n$ matrix

$$J_n(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

is called a **Jordan block** associated with the scalar λ. Note that a Jordan block has λ's on the main diagonal, 1's on the superdiagonal and 0's elsewhere. For example,

$$J_3(2) = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix},\ J_2(-5) = \begin{bmatrix} -5 & 1 \\ & -5 \end{bmatrix},\ J_4(0) = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}.$$

A **Jordan matrix** is a block-diagonal matrix where each block along the diagonal is a Jordan block.

$$J_3(2) \oplus J_2(-5) \oplus J_1(-1) = \begin{bmatrix} 2 & 1 & & & & \\ & 2 & 1 & & & \\ & & 2 & & & \\ & & & 5 & 1 & \\ & & & & 5 & \\ & & & & & -1 \end{bmatrix}$$

4 Simultaneous diagonalization

Simultaneously diagonalizable matrices are important for a proof of the uniqueness of the Jordan decomposition we shall see later.

Proposition 4.1. A nilpotent matrix is diagonalizable if and only if it is a zero matrix.

Definition 4.2. Two matrices $A, B \in M_n$ are said to be **simultaneously diagonalizable** if there is a nonsingular $S \in M_n$ such that $S^{-1}AS$ and $S^{-1}BS$ are both diagonal.

Suppose that A and B are square matrices that commute; $AB = BA$. Let λ be an eigenvalue for A and let $\mathcal{E}(\lambda)$ be the eigenspace of A corresponding to λ. Let $e_1, \cdots, e_k$ be a basis for $\mathcal{E}(\lambda)$.

Proposition 4.3. Let A and B commute. Any eigenspace $\mathcal{E}(\lambda)$ of A is invariant under B, i.e. ,

$$v \in \mathcal{E}(\lambda) \Rightarrow Bv \in \mathcal{E}(\lambda)$$

Proof. It is enough to show that if $e_i \in \mathcal{E}(\lambda)$ then $Be_i \in \mathcal{E}(\lambda)$. We have

$$A(Be_i) = (AB)e_i = (BA)e_i = B(\lambda e_i) = \lambda(Be_i)$$

since $AB = BA$. Therefore $Be_i \in \mathcal{E}(\lambda)$ as desired. □

We extend $\{ e_1, \cdots, e_k \}$ to a basis $\{ e_1, \cdots, e_k, e_{k+1}, \cdots, e_n \}$ for V.
We express $Be_1, \ldots, Be_k \in \mathcal{E}(\lambda)$ as a linear combination of the basis for V.

$$\begin{aligned}
Be_1 &= b_{11}e_1 + \cdots + b_{k1}e_k + 0e_{k+1} + \cdots + 0e_n \\
Be_2 &= b_{12}e_1 + \cdots + b_{k2}e_k + 0e_{k+1} + \cdots + 0e_n \\
&\vdots \\
Be_k &= b_{1k}e_1 + \cdots + b_{kk}e_k + 0e_{k+1} + \cdots + 0e_n
\end{aligned}$$

where the coefficients $b_{ij} \in \mathbb{C}$. The matrix B corresponding to $\{e_1, \ldots, e_n\}$ would be something like

$$B = \begin{bmatrix} b_{11} & \cdots & b_{1k} & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{k1} & \cdots & b_{kk} & * & \cdots & * \\ & & & * & \cdots & * \\ & 0 & & \vdots & & \vdots \\ & & & * & \cdots & * \end{bmatrix}$$

Proposition 4.4. Suppose A is diagonalizable and matrices A, B commute, i.e. , $AB = BA$. Then there exists an invertible matrix P such that $P^{-1}AP$ and $P^{-1}BP$ are both diagonal.

Let the characteristic polynomial of A to be

$$\varphi_A(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k}$$

where the m_i is the multiplicity of an eigenvalue λ_i.

5 Jordan decomposition

Proposition 5.1 (Jordan decomposition). Let X be a square matrix of size n. Then, X can be uniquely written as the sum of two matrices: a diagonalizable matrix S and a nilpotent matrix N.

$$X = S + N$$

Furthermore, S and N commute, i.e., $SN = NS$ and the decomposition is unique.

Here is an example of Jordan decomposition. Let a matrix X be in the Jordan canonical form.

$$X = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} J_3(\lambda_1) & \\ & J_2(\lambda_2) \end{bmatrix} = J_3(\lambda_1) \oplus J_2(\lambda_2)$$

Then X can be written as a sum

$$X = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where the first matrix is diagonalizable and the second matrix is nilpotent. In particular, we have the Jordan decomposition of a Jordan block as a sum

$$J_n(\lambda) = \lambda I_n + J_n(0).$$

Proposition 5.2 (Generalized eigenspace decomposition). Let V be a finite dimensional vector space over $\mathbb{C}$. Then, V can be decomposed into the sum of generalized eigenspaces:

$$V = \bigoplus_i \mathcal{E}_{k_i}(\lambda_i)\ .$$

If the multiplicity of λ_i is m_i, we have

$$\dim \mathcal{E}_{k_i}(\lambda_i) = m_i\ .$$

Note that the minimal polynomial of X in the example above is

$$p(t) = (t-\lambda_1)^3(t-\lambda_2)^2$$

This corresponds to a decomposition of V as a direct sum

$$V = \mathcal{E}_3(\lambda_1) \oplus \mathcal{E}_2(\lambda_2)$$

where $\mathcal{E}_{k_i}(\lambda_i)$ are generalized eigenspace of X defined by

$$\mathcal{E}_k(\lambda) = \left\{\, v \in V \mid (X-\lambda 1)^k v = 0 \text{ for some } k \,\right\}.$$

The exponents $3, 2$ in the minimal polynomial correspond to the sizes of the Jordan blocks in this case.

Jordan form of nilpotent matrix of size 3 is one of the following.

$$J_1(0) \oplus J_1(0) \oplus J_1(0) = \begin{bmatrix} J_1(0) & & \\ & J_1(0) & \\ & & J_1(0) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$J_2(0) \oplus J_1(0) = \begin{bmatrix} J_2(0) & \\ & J_1(0) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$J_3(0) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Theorem 5.3. Every square matrix A is similar to a Jordan matrix. In other words, there is an invertible matrix P such that

$$P^{-1}AP = J_{n_{11}}(\lambda_1) \oplus J_{n_{1m_1}}(\lambda_1) \oplus \cdots \oplus J_{n_{r1}}(\lambda_r) \oplus J_{n_{rm_r}}(\lambda_r)$$

Moreover, the Jordan blocks are unique up to order.

Theorem 5.4 (Jordan canonical form of nilpotent matrix). Every nilpotent matrix N is similar to a diagonal matrix which has Jordan blocks in the main diagonal.

$$P^{-1}NP = \begin{bmatrix} J_{n_1}(0) & & & \\ & J_{n_2}(0) & & \\ & & \ddots & \\ & & & J_{n_l}(0) \end{bmatrix} = J_{n_1}(0) \oplus \cdots \oplus J_{n_l}(0)$$

The matrix $P^{-1}NP$ is unique by considering the order of numbers $n_1 \geq \cdots \geq n_l \geq 1$.

References

[1] S. Roman. *Advanced Linear Algebra*. Graduate Texts in Mathematics. Springer New York, 2007.

[2] Wikipedia. Row and column spaces — Wikipedia, the free encyclopedia. `http://en.wikipedia.org/w/index.php?title=Rowandcolumnspaces`, 2023. [Online; accessed 31-October-2023].