A note on the Jordan canonical form

(work in progress)

Shuto Noda

January 17, 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 transformation of a finite-dimensional 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 $id_V: V \to V$ 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.

Definition 1.3. Let λ be an eigenvalue of a linear operator $\tau \in \mathcal{L}(V)$.

- (i) The algebraic multiplicity $\alpha(\lambda)$ of λ is the multiplicity of λ as a root of the characteristic polynomial $c_{\tau}(t)$.
- (ii) The geometric multiplicity $\gamma(\lambda)$ of λ is the dimension of the eigenspace $\mathcal{E}(\lambda)$, i.e. , $\gamma(\lambda) = \dim \mathcal{E}(\lambda)$.

Proposition 1.4. The geometric multiplicity of an eigenvalue λ is less than or equal to its algebraic multiplicity.

$$\alpha(\lambda) \ge \gamma(\lambda)$$

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 characteristic 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 .

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

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

For
$$B = \begin{bmatrix} 1 & 3 \\ & 1 & 3 \\ & & 1 \end{bmatrix}$$
, 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, \dots, 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 $\{e_1, \dots, e_k\}$ to a basis $\{e_1, \dots, e_k, e_{k+1}, \dots, e_n\}$ for V. We express $Be_1, \dots, Be_k \in \mathcal{E}(\lambda)$ as a linear combination of the basis for V.

$$Be_{1} = b_{11}e_{1} + \dots + b_{k1}e_{k} + 0e_{k+1} + \dots + 0e_{n}$$

$$Be_{2} = b_{12}e_{1} + \dots + b_{k2}e_{k} + 0e_{k+1} + \dots + 0e_{n}$$

$$\vdots$$

$$Be_{k} = b_{1k}e_{1} + \dots + b_{kk}e_{k} + 0e_{k+1} + \dots + 0e_{n}$$

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

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

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 Nilpotent matrix

Definition 5.1. A linear transformation $\tau \in \mathcal{L}(V)$ is **nilpotent** if $\tau^m = 0$ for some positive $m \in \mathbb{N}$. A matrix $A \in M_n$ is **nilpotent** if $A^m = 0$ for some positive $m \in \mathbb{N}$.

The characteristic polynomial and the minimal polynomial of an nilpotent transfomation (or matrix) is t^m .

6 Jordan decomposition

Proposition 6.1 (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

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 6.2 (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 6.3 (Jordan canonical form of a nilpotent matrix). Every nilpotent matrix N is similar to a diagonal matrix which has Jordan blocks in the main diagonal.

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

The matrix J is unique by considering the order of numbers $n_1 \geq \cdots \geq n_l \geq 1$.

Proof. Let N be a nilpotent matrix with $N^{k-1} \neq 0$ and $N^k = 0$. We have

$$\{\mathbf{0}\} \subset \ker N \subset \ker N^2 \subset \cdots \subset \ker N^{k-1} \subset \ker N^k = \mathbb{C}^n$$

We define dim $\ker N^i = m_i, m_i - m_{i-1} = r_i (1 \le u \le k), m_0 = 0$. Choose a set of linearly independent vectors $\{\mathbf{a}_1, \dots, \mathbf{a}_{r_k}\}$ from $\ker N^k \setminus \ker N^{k-1}$ so that

$$\ker N^k = \langle \mathbf{a}_1, \cdots, \mathbf{a}_{r_k} \rangle \oplus \ker N^{k-1}$$

where the first r_k vectors and any basis of ker N^{k-1} consists of a basis of ker N^k . Now $N\mathbf{a}_1, \dots, N\mathbf{a}_{r_k} \in \ker N^{k-1}$ are linearly independent vectors such that

$$\langle N\mathbf{a}_1, \cdots, N\mathbf{a}_{r_k} \rangle \cap \ker N^{k-2} = \{\mathbf{0}\}\$$

In fact, suppose that $c_1N\mathbf{a}_1 + \cdots + c_{r_k}N\mathbf{a}_{r_k} \in \ker N^{k-2}$ then

$$N^{k-2}(c_1N\mathbf{a}_1 + \dots + c_{r_k}N\mathbf{a}_{r_k}) = N^{k-1}(c_1\mathbf{a}_1 + \dots + c_{r_k}\mathbf{a}_{r_k}) = \mathbf{0}.$$

Thus $c_1 \mathbf{a}_1 + \dots + c_{r_k} \mathbf{a}_{r_k} \in \ker N^{k-1}$.

$$c_1\mathbf{a}_1 + \dots + c_{r_k}\mathbf{a}_{r_k} \in \cap \langle N\mathbf{a}_1, \dots, N\mathbf{a}_{r_k} \rangle \cap \ker N^{k-2} = \{\mathbf{0}\}$$

implies that $c_1 = \cdots = c_{r_k} = 0$. We also have $r_k \leq r_{k-1}$ (since N is monomorphism). There exists $r_{k-1} - r_k$ many vectors $\mathbf{a}_{r_k+1}, \cdots, \mathbf{a}_{r_{k-1}} \in \ker N^{k-1} \setminus \ker N^{k-2}$ so that

$$\ker N^{k-1} = \langle N\mathbf{a}_1, \cdots, N\mathbf{a}_{r_k}, \mathbf{a}_{r_k+1}, \cdots, \mathbf{a}_{r_{k-1}} \rangle \oplus \ker N^{k-2}$$

where the first r_{k-1} vectors and any basis of ker N^{k-2} consists of a basis of ker N^{k-1} .

By continuing this construction of the basis of \mathbb{C} , we have the basis

$$\bigcup_{1 \le i \le k} \{ N^j(\mathbf{a}_l) \mid 0 \le j \le k - i, r_k + 1 \le l \le r_{k-1} \}.$$

For $r_i + 1 \le j \le r_{i-1}$, $\langle N^{i-1}\mathbf{a}_j, N^{i-2}\mathbf{a}_j, \cdots, N\mathbf{a}_j, \mathbf{a}_j \rangle$ is N-invariant subspace. For this basis (called the Jordan chain for j), the matrix N is something like

$$N(N^{i-1}\mathbf{a}_j, N^{i-2}\mathbf{a}_j, \cdots, \mathbf{a}_j) = (N^{i-1}\mathbf{a}_j, N^{i-2}\mathbf{a}_j, \cdots, \mathbf{a}_j) \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$$

Order the Jordan chains in descending length. Under this basis, the matrix J that represents N would be $J = J_{n_1}(0) \oplus \cdots \oplus J_{n_l}(0)$ as desired.

7 Jordan canonical form for any matrix

Any matrix X can be written as the sum of the semisimple matrix S and the nilpotent matrix N, i.e. ,

$$X = S + N$$

The Jordan canonical form of each matrix would be

$$S = \begin{bmatrix} \lambda_1 E_{n_1} & & & \\ & \ddots & & \\ & & \lambda_l E_{n_l} \end{bmatrix}, \quad N = \begin{bmatrix} J_1(0) & & & \\ & \ddots & & \\ & & J_{n_l}(0) \end{bmatrix}$$
$$\begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

Theorem 7.1. 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.

8 Examples

$$A = \begin{bmatrix} 0 & 1 & 0 & 2 \\ -1 & 2 & 0 & 1 \\ 1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

The characteristic polynomial of A is

$$c_A(t) = \det(tI_4 - A) = \begin{bmatrix} t & -1 & 0 & -2 \\ 1 & t - 2 & 0 & -1 \\ -1 & 1 & t - 2 & 0 \\ 0 & 0 & 0 & t - 2 \end{bmatrix} = (t - 1)^2 (t - 2)^2.$$

The eigenvalues for A are $\lambda_1 = 1, \lambda_2 = 2$.

$$A - I_4 = \begin{bmatrix} -1 & 1 & 0 & 2 \\ -1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (A - I_4)^2 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$A - 2I_4 = \begin{bmatrix} -2 & 1 & 0 & 2 \\ -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (A - 2I_4)^2 = \begin{bmatrix} 3 & -2 & 0 & -3 \\ 2 & -1 & 0 & -2 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, the generalized eigenspaces $\mathcal{E}_{k_i}(\lambda_i)$ for λ_i are given as follows.

$$\mathcal{E}_2(1) = \left\{ \begin{bmatrix} x \\ y \\ y - x \\ 0 \end{bmatrix} \mid x, y \in \mathbb{C} \right\}, \quad \mathcal{E}_2(2) = \left\{ \begin{bmatrix} x \\ 0 \\ y \\ x \end{bmatrix} \mid x, y \in \mathbb{C} \right\}$$

Let $\tau_1(v)=(A-I4)^2(v)$ be a linear transformation on $\mathcal{E}_2(1)$. The fact that $\tau_1\neq 0$ and $\tau_1^2=0$ implies that τ_1 is nilpotent with its nilpotence index 2 and the Jordan canonical form of τ_1 is $J_2(0)$. Similarly, let $\tau_2(v)=(A-2I4)^2(v)$ be a linear transformation on $\mathcal{E}_2(2)$. Then the Jordan canonical form of τ_2 is $J_2(0)$. Then the Jordan canonical form of $\tau=A$ on $V=\mathcal{E}_2(1)\oplus\mathcal{E}_2(2)$ is

$$J_2(1) \oplus J_2(2) = \begin{bmatrix} J_2(1) & & & \\ & J_2(2) \end{bmatrix} = \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}$$

In fact, let

$$P = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, P^{-1} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then, we have

$$P^{-1}AP = \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}$$

as desired.

9 How do I find Jordan basis

$$\begin{bmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{bmatrix}$$

$$c_A(t) = (t-2)^3$$

Consider the sequence of kernels

$$\{\mathbf{0}\} \subsetneq \ker(A - 2I) \subsetneq \ker(A - 2I)^2 \subsetneq \cdots$$

The sequence stops after 2 steps since

$$A - 2I = \begin{bmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ -2 & 1 & 0 \end{bmatrix}, \quad (A - 2I)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A-2I has rank 1, hence its kernel (the eigenspace) has codimension 1, i.e., has dimension 2. $(A-2I)^2$ is the null matrix hence its kernel has dimension 3.

Take any vector in $\ker(A-2I)^2 \setminus \ker(A-2I)$ i.e. any vector of \mathbb{R}^3 which is not an eigenvector. Say

$$e_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Note that

$$e_2 = (A - 2I)e_3 = \begin{bmatrix} 1\\2\\1 \end{bmatrix}$$

is an eigenvector by construction.

We complete this set of two vectors to a basis, by choosing another eigenvector which is linearly independent from e_2 , say

$$e_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

The definition of e_2 from e_3 can be written as $Ae_3 = 2e_3 + e_2$.

$$A \sim J = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

10 Finding the Jordan canonical form

- (i) Find eigenvalues of A.
- (ii) Find algebraic multiplicity and geometric multiplicity.
- (iii) Calculate the dimension of the generalized eigenspace.
- (iv) Create the Jordan basis chain and the transformation matrix P.

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].
- [3] "How do I find Jordan basis?" Stack Exchange. https://math.stackexchange.com/questions/2433644/how-do-i-find-jordan-basis. [Online; accessed 16-January-2024].