A note on the Jordan canonical form
(work in progress)

Shuto Noda

January 22, 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 $\boldsymbol{v} \in \mathbb{C}^n$ for which

$$A\boldsymbol{v} = \lambda \boldsymbol{v}$$

In this case, $\boldsymbol{v}$ is called an eigenvector for A associated with λ.

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

We say that the subspace W of V is invariant under $A \in M_n$ if $\boldsymbol{x} \in W$ implies $A\boldsymbol{x} \in W$.

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

$$\mathcal{E}(\lambda) = \{\, \boldsymbol{v} \in \mathbb{C}^n \mid (A - \lambda I_n)\boldsymbol{v} = 0 \,\}$$

where I_n is the identity matrix. The **generalized eigenspace** of τ with eigenvalue λ is defined by

$$\mathcal{E}_k(\lambda) = \left\{\, \boldsymbol{v} \in \mathbb{C}^n \mid (A - \lambda I_n)^k \boldsymbol{v} = 0 \text{ for some } k \,\right\}.$$

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

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

(ii) $\mathcal{E}(\lambda) = \mathcal{E}_1(\lambda)$.

(iii) $\mathcal{E}_k(\lambda) \subset \mathcal{E}_{k+1}(\lambda)$.

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

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

$$(A - \lambda I_n)^m (c\boldsymbol{u} + d\boldsymbol{v}) = c(A - \lambda I_n)^m \boldsymbol{u} + d(A - \lambda I_n)^m \boldsymbol{v} = 0.$$

Moreover, $\boldsymbol{u} \in \mathcal{E}(\lambda)$ if and only if

$$(A - \lambda I_n)^k \boldsymbol{u} = 0 \text{ for } k = 1.$$

Hence, we get $\mathcal{E}(\lambda) = \mathcal{E}_1(\lambda)$. If $(A - \lambda I_n)^k \boldsymbol{v} = 0$, then

$$(A - \lambda I_n)^{k+1} \boldsymbol{v} = (A - \lambda I_n)(A - \lambda I_n)^k \boldsymbol{v} = 0$$

We have $\mathcal{E}_k(\lambda) \subset \mathcal{E}_{k+1}(\lambda)$ as desired. $\square$

$$\mathcal{E}(\lambda) = \mathcal{E}_1(\lambda) \subset \mathcal{E}_2(\lambda) \subset \cdots \subset \mathcal{E}_k(\lambda) = \mathcal{E}_{k+1}(\lambda) = \cdots$$

(i) $\mathcal{E}_j(\lambda) = \ker(A - \lambda I)^j$ is invariant under A.

(ii) $\mathcal{E}_j(\lambda) = \ker(A - \lambda I)^j$ is invariant under $\ker(A - \lambda I)$.

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 matrix $A \in \mathcal{L}(V)$.

(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.

$$\gamma(\lambda) \leq \alpha(\lambda)$$

2 Minimal polynomials

Definition 2.1. Let $\tau \in \mathcal{L}(V)$ and let $\lambda_1, \ldots, \lambda_r$ be distinct eigenvalues of τ. We define the **charasteristic polynomial** of τ by

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

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

In particular, the characteristic polynomial of a matrix $A \in M_n$ can be written as

$$c_A(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$$

where $\lambda_i \neq \lambda_j$ for $i \neq j$ and $n = n_1 + n_2 + \cdots + n_r$.

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

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

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

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. $\square$

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 $\{e_1, \dots, e_n\}$ 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 Nilpotent matrix

Definition 5.1. A linear transfomation $\tau \in \mathcal{L}(V)$ 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

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

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 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 $N^k = 0$.

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

We construct a basis of $\mathbb{C}$ according to this sequence. We define $\dim \ker N^i = d_i, d_i - d_{i-1} = r_i$ $(1 \leq i \leq k), m_0 = 0$. Choose vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{r_k} \in \ker N^k$ such that

(i) $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{r_k}\}$ is linearly independent.

(ii) $\langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{r_k} \rangle \cap \ker N^{k-1} = \{\mathbf{0}\}$.

(iii) $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{r_k}\}$ and any basis of $\ker N^{k-1}$ consist a basis of N^k.

We denote $\mathcal{B}_1$ the set of kr_k many vectors as follows.

$$\mathcal{B}_1 = \left\{ \begin{matrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_{r_k} \\ N\boldsymbol{a}_1 & N\boldsymbol{a}_2 & \cdots & N\boldsymbol{a}_{r_k} \\ \vdots & \vdots & & \vdots \\ N^{k-1}\boldsymbol{a}_1 & N^{k-1}\boldsymbol{a}_2 & \cdots & N^{k-1}\boldsymbol{a}_{r_k} \end{matrix} \right\}$$

The set of vectors $\mathcal{B}_1$ is linearly independent. Suppose that

$$\sum_{\nu=1}^{r_k} c_\nu^{(0)} \boldsymbol{a}_\nu + \sum_{\nu=1}^{r_k} c_\nu^{(1)} N\boldsymbol{a}_\nu + \cdots + \sum_{\nu=1}^{r_k} c_\nu^{(k-1)} N^{k-1}\boldsymbol{a}_\nu = \mathbf{0} \tag{1}$$

Multiplying both sides (1) by the matrix N^{k-1} from the left, we obtain

$$\sum_{\nu=1}^{r_k} c_\nu^{(0)} N^{k-1}\boldsymbol{a}_\nu = N^{k-1}\left(\sum_{\nu=1}^{r_k} c_\nu^{(0)} \boldsymbol{a}_\nu\right) = \mathbf{0}$$

Thus, $\sum_{\nu=1}^{r_k} c_\nu^{(0)} \boldsymbol{a}_\nu \in \ker N^{k-1}$. However, from the assumption $\langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{r_k} \rangle \cap \ker N^{k-1} = \{\mathbf{0}\}$, we obtain $c_\nu^{(0)} = 0$ for all $\nu = 1, \ldots, r_k$. Multiplying both sides (1) by the matrix N^{k-2} instead, we obtain $c_\nu^{(1)} = 0$. By continuing this argument, we obtain $c_\nu^{(0)} = \cdots = c_\nu^{(k-1)} = 0$. Hence $(*)$ is linearly independent.

Choose vectors $\boldsymbol{a}_{r_k+1}, \cdots, \boldsymbol{a}_{r_{k-1}} \in \ker N^{k-1}$ such that

(i) $\left\{N\boldsymbol{a}_1, \cdots, N\boldsymbol{a}_{r_k}, \boldsymbol{a}_{r_k+1}, \cdots, \boldsymbol{a}_{r_{k-1}}\right\}$ is linearly independent.

(ii) $\left\langle N\boldsymbol{a}_1, \cdots, N\boldsymbol{a}_{r_k}, \boldsymbol{a}_{r_k+1}, \cdots, \boldsymbol{a}_{r_{k-1}}\right\rangle \cap \ker N^{k-2} = \{\mathbf{0}\}$.

(iii) $\left\{N\boldsymbol{a}_1, \cdots, N\boldsymbol{a}_{r_k}, \boldsymbol{a}_{r_k+1}, \cdots, \boldsymbol{a}_{r_{k-1}}\right\}$ and any basis of $\ker N^{k-2}$ consist a basis of N^{k-1}.

We denote $\mathcal{B}_2$ the set of vectors as follows.

$$\mathcal{B}_2 = \left\{ \begin{matrix} \boldsymbol{a}_{r_k+1} & \boldsymbol{a}_{r_k+2} & \cdots & \boldsymbol{a}_{r_{k-1}} \\ N\boldsymbol{a}_{r_k+1} & N\boldsymbol{a}_{r_k+2} & \cdots & N\boldsymbol{a}_{r_{k-1}} \\ \vdots & \vdots & & \vdots \\ N^{k-1}\boldsymbol{a}_{r_k+1} & N^{k-1}\boldsymbol{a}_{r_k+2} & \cdots & N^{k-1}\boldsymbol{a}_{r_{k-1}} \end{matrix} \right\}$$

In fact, the set of vectors $\mathcal{B}_1 \cup \mathcal{B}_2$ is linearly independent.

Choose vectors $\boldsymbol{a}_{r_{k-1}+1}, \cdots, \boldsymbol{a}_{r_{k-2}} \in \ker N^{k-2}$ such that

(i) $\left\{N^2\boldsymbol{a}_1, \cdots, N^2\boldsymbol{a}_{r_k}, N\boldsymbol{a}_{r_k+1}, \cdots, N\boldsymbol{a}_{r_{k-1}}, \boldsymbol{a}_{r_{k-1}+1}, \cdots, \boldsymbol{a}_{r_{k-2}}\right\}$ is linearly independent.

(ii) $\left\langle N^2\boldsymbol{a}_1, \cdots, N^2\boldsymbol{a}_{r_k}, N\boldsymbol{a}_{r_k+1}, \cdots, N\boldsymbol{a}_{r_{k-1}}, \boldsymbol{a}_{r_{k-1}+1}, \cdots, \boldsymbol{a}_{r_{k-2}}\right\rangle \cap \ker N^{k-3} = \{\boldsymbol{0}\}$.

(iii) $\left\{N^2\boldsymbol{a}_1, \cdots, N^2\boldsymbol{a}_{r_k}, N\boldsymbol{a}_{r_k+1}, \cdots, N\boldsymbol{a}_{r_{k-1}}, \boldsymbol{a}_{r_{k-1}+1}, \cdots, \boldsymbol{a}_{r_{k-2}}\right\}$ and any basis of $\ker N^{k-3}$ consist a basis of N^{k-3}.

By continuing this construction, we have a basis $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ of $\mathbb{C}$.

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

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

Let

$$P = \begin{bmatrix} | & & | & & | & & | & & | & & | \\ N^{k-1}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_1 & \cdots & N^{k-1}\boldsymbol{a}_{r_k} & \cdots & \boldsymbol{a}_{r_k} & \cdots & \boldsymbol{a}_{r_2+1} & \cdots & \boldsymbol{a}_{r_1} \\ | & & | & & | & & | & & | & & | \end{bmatrix}$$

The matrix J that represents N is $J = J_{n_1}(0) \oplus \cdots \oplus J_{n_l}(0)$ as desired. $\square$

7 Jordan canonical form for any square matrix

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

$$A = S + N$$

$$P^{-1}AP = P^{-1}(S+N)P = P^{-1}SP + P^{-1}NP$$

The Jordan canonical form of each matrix would be

$$P^{-1}SP = \begin{bmatrix} \lambda_1 I_{p_1} & & \\ & \ddots & \\ & & \lambda_m I_{p_m} \end{bmatrix}, \quad P^{-1}NP = \begin{bmatrix} J_{p_1}(0) & & \\ & \ddots & \\ & & J_{p_m}(0) \end{bmatrix}$$

Hence, the Jordan canonical form of A must be

$$P^{-1}AP = \begin{bmatrix} \lambda_1 I_{p_1} + J_{p_1}(0) & & \\ & \ddots & \\ & & \lambda_m I_{p_m} + J_{p_m}(0) \end{bmatrix} = \begin{bmatrix} J_{p_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{p_m}(\lambda_m) \end{bmatrix}$$

where

$$J_{p_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}.$$

Theorem 7.1. Every square matrix $A \in M_n$ is similar to some Jordan matrix. In other words, there is an invertible matrix P such that

$$J = P^{-1}AP = \begin{bmatrix} J_{p_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{p_m}(\lambda_m) \end{bmatrix}$$

The matrix J is called the Jordan canonical form of A. Moreover, the Jordan canonical form is unique up to order of the Jordan blocks.

8 Jordan basis

$J = P^{-1}AP = J_{p_1}(\lambda_1) \oplus \cdots \oplus J_{p_m}(\lambda_m)$

We choose a basis of $\mathbb{C}^n$ according to the Jordan canonical form. The columns of P are the basis vectors

$$\boldsymbol{a}_{11}, \ldots, \boldsymbol{a}_{1p_1}, \boldsymbol{a}_{21}, \ldots, \boldsymbol{a}_{2p_2}, \ldots, \boldsymbol{a}_{mo_1}, \ldots, \boldsymbol{a}_{mp_m}$$

In this case, we have

$$\begin{aligned}
(A - \lambda_i I_n)\boldsymbol{a}_{i1} &= \boldsymbol{0} \\
(A - \lambda_i I_n)\boldsymbol{a}_{i2} &= \boldsymbol{a}_{i1} \\
&\vdots \\
(A - \lambda_i I_n)\boldsymbol{a}_{ip_i} &= \boldsymbol{a}_{ip_{i-1}}
\end{aligned}$$

Let A be a square matrix which has only one Jordan block. Then we have $P^{-1}AP = J_n(\lambda)$ and $P = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]$. Multiplication by $A - \lambda I_n$ yields the following sequence.

$$\begin{aligned}
(A - \lambda I_n)\boldsymbol{a}_1 &= \boldsymbol{0} \\
(A - \lambda I_n)\boldsymbol{a}_2 &= \boldsymbol{a}_1 \\
&\vdots \\
(A - \lambda I_n)\boldsymbol{a}_n &= \boldsymbol{a}_{n-1}
\end{aligned}$$

We denote this relation as follows.

$$\begin{array}{l}
\boldsymbol{a}_n \\
\downarrow \\
\boldsymbol{a}_{n-1} \\
\downarrow \\
\vdots \\
\downarrow \\
\boldsymbol{a}_1 \\
\downarrow \\
\boldsymbol{0}
\end{array}$$

We call this diagram the **Jordan chain** for λ. In this case, the vector $\boldsymbol{a}_n \in \ker(A - \lambda I_n)^n$ is called the **generating element**. The length of the Jordan chain corresponds the size of the Jordan blocks.

$$\begin{array}{c}
* \\
\downarrow \\
* \\
\downarrow \\
\vdots \\
\downarrow \\
* \\
\downarrow \\
\boldsymbol{0}
\end{array}$$

9 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.

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

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.

Find the Jordan canonical form of A.

$$A = \begin{bmatrix} 4 & 2 & -1 \\ -3 & -1 & 2 \\ -2 & -2 & 4 \end{bmatrix}$$

The characteristic polynomial of A is $c_A(t) = (t-2)^2(t-3)$. For eigenvalue 3, the Jordan block is $J_1(3)$.

$$A - 3I_3 = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -4 & 2 \\ -2 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

We have the eigenspace for $\lambda = 3$.

$$\mathcal{E}_1(3) = \ker(A - 3I_3) = \langle \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \rangle$$

$$A - 2I_3 = \begin{bmatrix} 2 & 2 & -1 \\ -3 & -3 & 2 \\ -2 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$
$$(A - 2I_3)^2 = \begin{bmatrix} 0 & 0 & 0 \\ -1 & -1 & 1 \\ -2 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{E}_1(2) = \ker(A - 2I_3) = \langle \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \rangle, \ \mathcal{E}_2(2) = \ker(A - 2I_3)^2 = \langle \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \rangle$$

The fact that $\dim\ker(A - 2I_3)^2 - \dim\ker(A - 2I_3)^1 = 2 - 1 = 1$ and $\dim\ker(A - 2I_3)^1 = 1$ imply that the Jordan chain would be something like;

$$\begin{array}{ll} \ker(A - 2I_3)^2 \setminus \ker(A - 2I_3)^1 \ni & * \\ & \downarrow \\ \ker(A - 2I_3)^1 \ni & * \\ & \downarrow \\ & \mathbf{0} \end{array}$$

Choose a vector $\boldsymbol{a}$ from $\ker(A - 2I_3)^2 \setminus \ker(A - 2I_3)^1$, say

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \notin \ker(A - 2I_3)^1$$

Then the basis of $\mathbb{C}^3$ corresponding to the Jordan canonical form consists of the vectors

$$(A - 2I_3)\boldsymbol{a} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

Then

$$P^{-1}AP = J_2(2) \oplus J_1(3) = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \text{ where } P = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

10 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 $Ae_3 = 2e_3 + e_2$.

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

11 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.

12 How do I find Jordan basis

When the size of the matrix becomes large, it is known that finding the Jordan normal form is difficult, even for computers. However, the Jordan canonical form of a matrix plays an important role when analyzing the matrix exponential functions.

For example, the Jordan canonical form can be used to prove the equation

$$\exp\left(P^{-1}\begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix} P\right) = P^{-1}\begin{bmatrix} e^a & e^a \\ 0 & e^a \end{bmatrix} P$$

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].