Jordan Form

Dr. Yimao Chen

December 17, 2023

1 The Jordan Canonical Form I

1.1 Theoretical Preparation

Suppose V is a finite-dimensional complex vector space.

Generalized Eigenvectors

Definition 1.1.1. Let T be a linear operator on a vector space V, and let λ be a scalar. A nonzero vector v in V is called a generalized eigenvector of T corresponding to λ if $(T - \lambda I)^p(x) = 0$ for some positive integer p.

Generalized Eigenspace

Definition 1.1.2. Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. The generalized eigenspace of T corresponding to λ, denoted $G(\lambda, T)$, is the subset of V denoted by

$$G(\lambda, T) = \{x \in V : (T - \lambda I)^p(x) = 0 \text{ for some positive integer } p\}.$$

The relation between eigenspaces and generalized eigenspaces is given as follows:

Theorem 1.1.3. *Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. Then*

(a) $G(\lambda, T)$ is subspace which is invariant under T, and $E(\lambda, T) \subset G(\lambda, T)$.

(b) For any $\mu \neq \lambda$, $(T - \mu I)|_{G(\lambda,T)}$ is injective.

The following theorem is a characterization of generalized eigenspace.

Theorem 1.1.4. *Let T be a linear operator on a finite-dimensional complex vector space V. Suppose that λ is an eigenvalue of T with multiplicity m. Then*

(a) $\dim G(\lambda, T) = m$.

(b) $G(\lambda, T) = null\ ((T - \lambda I)^m)$.

Theorem 1.1.5. *Let T be a linear operator on a finite dimensional complex vector space V , and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be the distinct eigenvalues of T. Then, for every $v \in V$, there exist vectors $v_i \in G(\lambda_i, T), 1 \leq i \leq k$, such that*

$$v = v_1 + v_2 + \cdots + v_k.$$

Theorem 1.1.6. *Let T be a linear operator on a finite dimensional vector space V, and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be the distinct eigenvalues of T with corresponding multiplicities $m_1, m_2, \cdots, m_k$. For $1 \leq i \leq k$, let β_i be an ordered basis for $G(\lambda_i, T)$. Then the following statements are true.*

(a) $\beta_i \cap \beta_j = \emptyset$, *for* $i \neq j$.

(b) $\beta = \beta_1 \cup \cdots \cup \beta_k$ *is an ordered basis for* V.

(c) $\dim(G(\lambda, T)) = m_i$ *for all* i.

Definition of a cycle of generalized eigenvectors.

Definition 1.1.7. Let T be a linear operator on a vector space V, and let x be a generalized eigenvector of T corresponding to the eigenvalue λ. Suppose that p is the smallest positive integer for which $(T - \lambda I)^p(x) = 0$. Then the ordered set

$$\{(T - \lambda I)^{p-1}(x), (T - \lambda I)^{p-2}(x), \cdots, (T - \lambda I)(x), x\}$$

is called a **cycle of generalized eigenvectors** of T corresponding to λ. The vectors $(T - \lambda I)^{p-1}(x)$ and x are called the **initial vector** and the **end vector** of the cycle, respectively. We say that **length** of the cycle is p.

Notice that the initial vector of a cycle of generalized eigenvectors of a linear operator T is the only eigenvector of T in the cycle.

Now we try to find a Jordan basis.

Theorem 1.1.8. *Let T be a linear operator on a finite-dimensional complex vector space V, and suppose that β is a basis for V such that β is a disjoint union of cycles of generalized eigenvectors of T. Then the following statements are true.*

(a) For each cycle γ of generalized eigenvectors contained in β, $U = span(\gamma)$ is invariant under T, and the matrix with respect to which is a Jordan block.

(b) β is a Jordan canonical basis for V.

Theorem 1.1.9. *Let T be a linear operator on a vector space, and let λ be an eigenvalue of T. Suppose that $\gamma_1, \gamma_2, \cdots, \gamma_q$ are cycles of generalized eigenvectors of T corresponding to λ such that the initial vectors of the γ_i's are distinct and form a linearly independent set. Then the γ_i's are disjoint, and their union $\gamma = \bigcup_{i=1}^{q} \gamma_i$ is linearly independent.*

Corollary 1.1.10. *Every cycle of generalized eigenvectors of a linear operator is linearly independent.*

Theorem 1.1.11. *Let T be a linear operator on a finite-dimensional vector space V, and let λ be an eigenvalue of T. Then $G(\lambda, T)$ has an ordered basis consisting of a union of disjoint cycles of generalized eigenvectors corresponding to λ.*

Corollary 1.1.12. *Let T be a linear operator on a finite-dimensional complex vector space V. Then T has a Jordan canonical form.*

2 The Jordan Canonical Form II

2.1 The Dot Diagram

To help visualize each of the matrices A_i and ordered basis β_i, we use an array of dots called a **dot diagram** of $T|_{G(\lambda_i,T)}$. Suppose that β_i is a disjoint union of cycles of generalized eigenvectors $\gamma_1, \gamma_2, \cdots, \gamma_{n_i}$, with lengths $p_1 \geq p_2 \geq \cdots \geq p_{n_i}$, respectively. The dot diagram of $T|_{G(\lambda_i,T)}$ contains one dot for each vector in β_i, and the dots are configured according to the following rules.

1. The array consists of n_i columns (one column for each cycle).

2. Counting from left to right, the jth column consists of the p_j dots that correspond to the vectors of γ_j starting with the initial vector at the top and continuing down to the end vector.

Denote the **end vectors** of the cycles by $v_1, v_2, \cdots, v_{n_i}$. In the following **dot diagram** of $T|_{G(\lambda_i,T)}$, each dot is labeled with the name of the vector in β_i to which it corresponds.

$$
\begin{array}{llllllll}
\bullet & (T-\lambda_i I)^{p_1-1}(v_1) & \bullet & (T-\lambda_i I)^{p_2-1}(v_2) & \cdots\cdots & \bullet & (T-\lambda_i I)^{p_{n_i}-1}(v_{n_i}) \\
\bullet & (T-\lambda_i I)^{p_1-2}(v_1) & \bullet & (T-\lambda_i I)^{p_2-2}(v_2) & \cdots\cdots & \bullet & (T-\lambda_i I)^{p_{n_i}-2}(v_{n_i}) \\
 & \qquad\vdots & & \qquad\vdots & \ddots & & \qquad\vdots \\
 & \qquad\vdots & & \qquad\vdots & \ddots & \bullet & (T-\lambda_i I)(v_{n_i}) \\
 & \qquad\vdots & & \qquad\vdots & \ddots & \bullet & v_{n_i} \\
 & \qquad\vdots & \bullet & (T-\lambda_i I)(v_2) & & & \\
 & \qquad\vdots & \bullet & v_2 & & & \\
\bullet & (T-\lambda_i I)(v_1) & & & & & \\
\bullet & v_1 & & & & &
\end{array}
$$

Notice that the dot diagram of $T|_{G(\lambda_i,T)}$ has n_i columns (one for each cycle) and p_1 rows. Since $p_1 \geq p_2 \geq \cdots \geq p_{n_i}$, the columns of the dot diagram become shorter (or at least not longer) as we move from left to right.

Now let r_j denote the number of dots in the jth row of the dot diagram. Observe that $r_1 \geq r_2 \geq \cdots \geq r_{p_1}$. Furthermore, the diagram can be constructed from the values of the r_i's.

2.2 Computing the Jordan Canonical form

Theorem 2.2.1. *For any positive integer r, the vectors in β_i that are associated with the dots in the first r rows of the dot diagram of T_i constitute a basis for null $((T-\lambda_i I)^r)$. Hence the number of dots in the first r rows of the dot diagram equals nullity $((T-\lambda_i I)^r)$.*

Corollary 2.2.2. *The dimension of $E(\lambda_i, T)$ is n_i. Hence in a Jordan canonical form of T, the number of Jordan blocks corresponding to λ_i equals the dimension of $E(\lambda_i, T)$.*

Theorem 2.2.3. *Let r_j denote the number of dots in the jth row of the dot diagram of $T_{G(\lambda_i,T)}$. Then the following statements are true.*

(a) $r_1 = \dim\,(V) - \dim \textit{range}\,(T - \lambda_i I)$.

(b) $r_j = \dim \textit{range}(T - \lambda_i I)^{j-1} - \dim \textit{range}(T - \lambda_i I)^j$.

Corollary 2.2.4. *For any eigenvalue λ_i of T, the dot diagram of T_i is unique. Thus, subject to the convention that the cycles of generalized eigenvectors for the bases of each generalized eigenspace are listed in order of decreasing length, the Jordan canonical form of a linear operator or a matrix is unique up to the ordering of the eigenvalues.*

3 Typical Applications

3.1 Examples

Example 3.1.1. *Let*

$$A = \begin{pmatrix} 2 & -4 & 2 & 2 \\ -2 & 0 & 1 & 3 \\ -2 & -2 & 3 & 3 \\ -2 & -6 & 3 & 7 \end{pmatrix}.$$

We find the Jordan canonical form J of A, a Jordan canonical basis for T ($Tx = Ax$), and a matrix Q such that $J = Q^{-1}AQ$.

Solution.
Jordan Basis:

$$\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix} \right\}.$$

Jordan Canonical Form:

$$J = \mathcal{M}(T, \alpha) = \begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

We define Q to be the matrix whose columns are the vectors of β listed in the same order, namely,

$$Q = \begin{pmatrix} 2 & 0 & 0 & 1 \\ 1 & 1 & 1 & -1 \\ 0 & 2 & 1 & -1 \\ 2 & 0 & 1 & 0 \end{pmatrix}.$$

Then $J = Q^{-1}AQ$.

Example 3.1.2. *Let V be the vector space of polynomial functions in two real variables x and y of degree at most 2. Then V is a vector space over $\mathbb{R}$ and $\alpha = \{1, x, y, x^2, y^2, xy\}$ is an ordered basis for V. Let T be the linear operator on V defined by*

$$T(f(x,y)) = \frac{\partial}{\partial x} f(x,y).$$

Find the Jordan canonical form and a Jordan canonical basis for T.

Solution.
The Jordan canonical form of T is:

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

A Jordan canonical basis for T is:

$$\beta = \{2, 2x, x^2, y, xy, y^2\}.$$

Theorem 3.1.3. *Let A and B be $n \times n$ matrices, each having Jordan canonical form computed according to the conventions of this section. Then A and B are similar if and only if they have (up to an ordering of their eigenvalues) the same Jordan canonical form.*

Example 3.1.4. *Which of the following matrices are similar?*

$$A = \begin{pmatrix} -3 & 3 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & -1 \\ -4 & 4 & -2 \\ -2 & 1 & 1 \end{pmatrix},$$
$$C = \begin{pmatrix} 0 & -1 & -1 \\ -3 & -1 & -2 \\ 7 & 5 & 6 \end{pmatrix}, D = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Solution. Observe that A, B, and C have the same characteristic polynomial $-(t-1)(t-2)^2$, whereas D has $-t(t-1)(t-2)$ as its characteristic polynomial. Because similar matrices have the same characteristic polynomials, D can not be similar to A, B, or C. Let J_A, J_B, and J_C be the Jordan canonical forms of A, B, and C, respectively, using the ordering 1,2 for their common eigenvalues. Then

$$J_A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, J_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \text{and } J_C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Since $J_A = J_C$, A is similar to C. Since J_B is different from J_A and J_C, B is similar to neither A nor C.

Remarks:

- The reader should observe that any diagonal matrix is a Jordan canonical form.
- Thus a linear operator T on a finite-dimensional vector space V is diagonalizable if and only if its Jordan canonical form is a diagonal matrix.
- Hence T is diagonalizable if and only if the Jordan canonical basis for T consists of eigenvectors of T.

3.2 Minimal Polynomial

Recall:

Definition 3.2.1. Let T be a linear operator on a finite-dimensional vector space. A polynomial $p(t)$ is called a minimal polynomial of T if $p(t)$ is a monic polynomial of least positive degree for which $p(T) = 0$.

Theorem 3.2.2. *Let T be a linear operator on a finite-dimensional vector space V. Then T is diagonalizable if and only if the minimal polynomial of T is of the form*

$$p(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k),$$

where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the distinct eigenvalues of T.

Proof. Suppose that T is diagonalizable. Let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be the distinct eigenvalues of T, and define

$$p(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k),$$

$p(t)$ divides the minimal polynomial of T. Let $v_1, v_2, \cdots, v_n$ be a basis for V consisting of eigenvectors of T, and consider any v_i in the list, we have $(T - \lambda_j I)(v_i) = 0$ for some eigenvalue λ_j. Since $(t - \lambda_j)$ divides $p(t)$, there is a polynomial $q_j(t)$ such that $p(t) = q_j(t)(t - \lambda_j)$. Hence

$$p(T)(v_i) = q_j(T)(T - \lambda_j I)(v_i) = 0.$$

It follows that $p(T) = 0$, since $p(T)$ takes each vector in a basis for V into 0. Therefore $p(t)$ is the minimal polynomial of T.

Conversely, suppose that there are distinct scalars $\lambda_1, \cdots, \lambda_k$ such that the minimal polynomial $p(t)$ of T factors as

$$p(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k),$$

the λ_i's are eigenvalues of T. We apply mathematical induction on $n = \dim(V)$. Clearly T is diagonalizable for $n = 1$. Now assume that T is diagonalizable whenever $\dim(V) < n$ for some $n > 1$, and let $\dim(V) = n$ and $W = \text{range}(T - \lambda_k I)$. Obviously $W \neq V$, because λ_k is an eigenvalue of T. If $W = \{0\}$, then $T = \lambda_k I$, which is clearly diagonalizable. So suppose that $0 < \dim(W) < n$. Then W is invariant under T, and for any $x \in W$,

$$(T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_{k-1} I)(x) = 0.$$

It follows that the minimal polynomial of $T|_W$ divides the polynomial $(t - \lambda_1) \cdots (t - \lambda_{k-1})$. Hence by the induction hypothesis, $T|_W$ is diagonalizable. Furthermore, λ_k is not an eigenvalue of $T|_W$. Therefore

$$W \cap \text{null}(T - \lambda_k I) = \{0\}.$$

Now let $v_1, \cdots, v_m$ be a basis for W consisting of eigenvectors of $T|_W$ (and hence of T), and let $w_1, \cdots, w_p$ be a basis for $\text{null}(T - \lambda_k I)$, the eigenspace of T corresponding to λ_k. $m + p = n$ by the fundamental theorem of linear maps applied to $T - \lambda_k I$. We show that $v_1, \cdots, v_m, w_1, \cdots, w_p$ is linear independent. Consider scalars $a_1, \cdots, a_m$ and $b_1, \cdots, b_p$ such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m + b_1 w_1 + b_2 w_2 + \cdots + b_p w_p = 0.$$

Let

$$x = \sum_{i=1}^{m} a_i v_i \text{ and } y = \sum_{i=1}^{p} b_i w_i.$$

Then $x \in W, y \in \text{null}(T - \lambda_k I)$, and $x + y = 0$. It follows that

$$x = -y \in W \cap \text{null}(T - \lambda_k I) = \{0\},$$

and therefore $x = 0$. Since $v_1, \cdots, v_m$ is linearly independent, we have that $a_1 = a_2 = \cdots = a_m = 0$. Similarly, $b_1 = b_2 = \cdots = b_p = 0$, we conclude that $v_1, \cdots, v_m, w_1, \cdots, w_p$ is linear independent subset of V consisting of n eigenvectors. It follows that $v_1, \cdots, v_m, w_1, \cdots, w_p$ is a basis for V consisting of eigenvectors of T, and consequently T is diagonalizable. □

3.3 Further Remarks

Remarks:

- In addition to diagonalizable operators, there are methods for determining the minimal polynomial of any linear operator on a finite-dimensional vector space.
- In the case that the characteristic polynomial of the operator splits, the minimal polynomial can be described using the Jordan canonical form of the operator.
- In the case that the characteristic polynomial does not split, the minimal polynomial can be described using the rational canonical form.